

# Fast Single Image Super-Resolution

Ningning Zhao, Qi Wei, Adrian Basarab, Denis Kouamé and Jean-Yves  
Tourneret

## Abstract

This paper addresses the problem of single image super-resolution, which consists of recovering a high resolution image from its blurred, decimated and noisy version. Given the well-known ill-posedness of image super-resolution, prior information is used for regularization purpose in order to obtain a well-posed problem. Among the existing algorithms, the alternating direction method of multipliers (ADMM) has been used intensively because of its effective implementation due to the possibility of splitting the super-resolution problem into up-sampling and deconvolution problems, which all can be easily solved. Instead of following this splitting strategy, we propose to consider the decimation and blurring operators simultaneously by taking advantage of their particular properties, leading to a new fast super-resolution approach. Based on this new scheme, different types of priors or regularizations are considered following the Bayesian framework. For a Gaussian prior, an analytical solution is easily obtained, which can be implemented in a very efficient way. In the case of non-Gaussian priors, we show that this analytical solution derived from the Gaussian case can be embedded into the ADMM framework, which accelerates the existing algorithms significantly. Simulation results on several images show the effectiveness of our fast scheme compared with the traditional ADMM implementation.

## Index Terms

Single image super-resolution, deconvolution, block circulant matrix, ADMM, Bayesian framework.

Part of this work has been supported by the Chinese Scholarship Council and the thematic trimester on image processing of the CIMI Labex, Toulouse, France, under grant ANR-11-LABX-0040-CIMI within the program ANR-11-IDEX-0002-02.

Ningning Zhao and Jean-Yves Tourneret are with University of Toulouse, IRIT/INP-ENSEEIH, Toulouse, France (e-mail: {nzhao, jean-yves.tourneret}@enseeiht.fr).

Qi Wei is with Department of Engineering, University of Cambridge, U.K. (e-mail: {qw245}@cam.ac.uk).

Adrian Basarab and Denis Kouamé are with University of Toulouse, IRIT, CNRS UMR 5505, 118 Route de Narbonne, F-31062, Toulouse Cedex 9, France (e-mail: {adrian.basarab, denis.kouame}@irit.fr).

## I. INTRODUCTION

Single image super-resolution (SR), also known as image scaling up or image enhancement, aims at estimating a high-resolution (HR) image from a low-resolution (LR) observed image [1]. Resolution enhancement is still an ongoing research problem with applications in various fields, such as remote sensing [2], video surveillance [3], hyperspectral imaging [4] or medical imaging [5]. The methods dedicated to single image SR can be classified into three categories [6]–[8]. The first category includes the interpolation based algorithms, such as nearest neighbor interpolation, bicubic interpolation [9] or adaptive interpolation techniques [10], [11]. Despite their simplicity and easy implementation, it is well-known that these algorithms generally over-smooth the high frequency details. The second type of methods considers example-based (or learning-based) algorithms that learn the relations between LR and HR image patches from a given database [6], [12]–[15]. Note that the effectiveness of the learning-based algorithms highly depends on the training image database and generally have high computational complexity. Reconstruction-based approaches that will be considered in this paper belong to the third category of SR approaches [7], [8], [16], [17]. These approaches formulate the image super-resolution as an optimization problem, either by incorporating priors in the Bayesian framework or by introducing regularizations into the ill-posed reconstruction problem. Priors or regularizations considered in the literature include the Tikhonov regularization [18]–[20], the total variation (TV) [17], [21], [22], the sparsity in transformed domains [23]–[26] or the generic image prior [7], [8], [16], [27]. It is also interesting to mention that regularization terms dedicated to specific areas (e.g., the gradient profile prior [16] and Fattal’s edge statistics [28]) also provided promising reconstruction results. The traditional optimization techniques used to solve the single image SR reconstruction mainly include the first order gradient-based methods [7], [8], [16] and the alternating direction method of multipliers (ADMM) [2], [5], [17], [22]. In this work, we focus on the ADMM-based algorithms, due to i) their efficiency in handling non-differential target distributions (e.g.,  $\ell_1$ -norm and TV regularizations) against the traditional gradient based methods [29], [30] and ii) the possibility they offer, through variable splitting, to divide the original optimization problem into several easier sub-problems [30]–[32].

With the conventional ADMM [2], [5], [17] (denoted by “direct” ADMM hereafter), the decimation and blurring operators are split and solved separately. Due to this splitting, a cumbersome SR problem can be decomposed into an up-sampling problem and a deconvo-

lution problem, which both can be solved efficiently. Our main contribution is to propose a new approach which is able to handle the two operators simultaneously by exploring their intrinsic properties. More precisely, if a Gaussian prior is considered, an analytical solution to implement the MAP estimation of the HR image can be obtained, avoiding any iterative updates. When handling non Gaussian priors, the proposed analytical solution can be embedded within the ADMM framework, leading to a fast version of ADMM (denoted by “FSR-ADMM” hereafter).

### A. Model of Image Formation

We focus in this paper on the techniques scaling up the LR images. Thus, the observed LR image can be modeled as a noisy version of the blurred and decimated HR image to be estimated as follows

$$\mathbf{y} = \mathbf{S}\mathbf{H}\mathbf{x} + \mathbf{n} \quad (1)$$

where the vector  $\mathbf{y} \in \mathbb{R}^{N_l \times 1}$  ( $N_l = m_l \times n_l$ ) denotes the LR observed image and  $\mathbf{x} \in \mathbb{R}^{N_h \times 1}$  ( $N_h = m_h \times n_h$ ) is the vectorized HR image to be estimated, with  $N_h > N_l$ . The vectors  $\mathbf{y}$  and  $\mathbf{x}$  are obtained by stacking the corresponding images (LR image  $\in \mathbb{R}^{m_l \times n_l}$  and HR image  $\in \mathbb{R}^{m_h \times n_h}$ ) into column vectors in a lexicographic order. Note that the vector  $\mathbf{n} \in \mathbb{R}^{N_l \times 1}$  is the independent identically distributed (*i.i.d.*) additive white Gaussian noise (AWGN) and that the matrices  $\mathbf{S} \in \mathbb{R}^{N_l \times N_h}$  and  $\mathbf{H} \in \mathbb{R}^{N_h \times N_h}$  represent the decimation and the blurring/convolution operations, respectively. More specifically,  $\mathbf{H}$  is a block circulant matrix with circulant blocks, which corresponds to cyclic convolution boundaries and left multiplying by  $\mathbf{S}$  performs down-sampling with an integer factor  $d$  ( $d = d_r \times d_c$ ), *i.e.*,  $N_h = N_l \times d$ . The decimation factors  $d_r$  and  $d_c$  represent the number of discarding rows and columns from the input image, thus satisfying the following relationships  $m_h = m_l \times d_r$  and  $n_h = n_l \times d_c$ .

### B. Paper Organization

The remainder of the paper is organized as follows. In Section II, we recall the SR reconstruction problem and the “direct” ADMM for a Gaussian prior. In Section III, we study the properties of the down-sampling and blurring operators and introduce a fast SR scheme based on an analytical solution obtained for a Gaussian prior. Section IV generalizes the proposed fast SR scheme to non-Gaussian priors in direct or transformed domains. Our algorithm is compared to the conventional ADMM implementation in Section V. Conclusions and perspectives are reported in Section VI.

## II. IMAGE SUPER-RESOLUTION FORMULATION

Similar to traditional image reconstruction problems, the estimation of the HR image from the LR observation is not invertible, leading to an ill-posed problem. To overcome this ill-posedness (or to sidestep the over-fitting problem), incorporating some prior information or regularizations is necessary. Assuming that the noise  $\mathbf{n}$  in (1) is AWGN and incorporating a proper regularization to the target image  $\mathbf{x}$ , the maximum *a posteriori* (MAP) estimator, i.e., the mode of the posterior distribution for the SR problem of a single image can be obtained by solving the following optimization problem

$$\min_{\mathbf{x}} \frac{1}{2} \underbrace{\|\mathbf{y} - \mathbf{S}\mathbf{H}\mathbf{x}\|_2^2}_{\text{data fidelity}} + \tau \underbrace{\phi(\mathbf{x})}_{\text{regularizer}} \quad (2)$$

where  $\|\mathbf{y} - \mathbf{S}\mathbf{H}\mathbf{x}\|_2^2$  is associated with the likelihood, referred to as *data fidelity term*, and  $\phi(\mathbf{x})$  corresponds to the image prior information, referred to as *regularizer or penalty* [33]. The role of the regularization parameter  $\tau$  is to weight the importance of the regularization with respect to (w.r.t.) the data term.

The choice of the regularization usually depends on the specific tasks of interest, the information resulting from previous experiments or a subjective view on the constraints affecting the unknown model parameters [34], [35]. Various regularizations have been already advocated to regularize the image super-resolution problem, such as, e.g., Tikhonov [36],  $\ell_1$ -norm [11] or total variation (TV) [5].<sup>1</sup>

Before proceeding to more complicated regularizations, let us consider the basic Tikhonov regularization, i.e.,  $\phi(\mathbf{x}) = \|\mathbf{x} - \bar{\mathbf{x}}\|_2^2$ . With the Tikhonov regularization, the problem (2) reduces to

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{H}\mathbf{x}\|_2^2 + \tau \|\mathbf{x} - \bar{\mathbf{x}}\|_2^2 \quad (3)$$

where  $\bar{\mathbf{x}}$  is a rough estimation of the HR image, e.g., an interpolated version of the observed image or a cleaner image obtained from other sensors [19], [20], [37]. The Tikhonov regularization with respect to  $\mathbf{x}$  implies that the target image  $\mathbf{x}$  is *a priori* close to the rough estimation  $\bar{\mathbf{x}}$ , whereas the parameter  $\tau$  weights the confidence in  $\bar{\mathbf{x}}$  w.r.t. to the data fidelity term. The solution of (3) can be expressed as following

$$\hat{\mathbf{x}} = (\mathbf{H}^H \mathbf{S}^H \mathbf{S} \mathbf{H} + 2\tau \mathbf{I}_{N_h})^{-1} (\mathbf{H}^H \mathbf{S}^H \mathbf{y} + 2\tau \bar{\mathbf{x}}) \quad (4)$$

<sup>1</sup>Proposing new regularization terms or comparing the performance of existing ones is out of the scope of this paper.

Note that if we do not consider the decimation matrix  $\mathbf{S}$  in (3), i.e.,  $\mathbf{S} = \mathbf{I}_{N_h}$  ( $\mathbf{I}_{N_h} \in \mathbb{R}^{N_h \times N_h}$  is an identity matrix), an analytical solution (seen in (5)) can be efficiently computed since the blurring matrix  $\mathbf{H}$  is BCCB (block circulant with circulant blocks) and can be diagonalized by a 2D discrete Fourier transform  $\mathcal{F}$ , leading to the well-known Wiener filter [38].

$$\tilde{\mathbf{x}} = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}(\mathbf{H})^H \circ \mathcal{F}(\mathbf{y}) + 2\tau \mathcal{F}(\bar{\mathbf{x}})}{\mathcal{F}(\mathbf{H})^H \circ \mathcal{F}(\mathbf{H}) + 2\tau \mathbf{I}_{N_h}} \right\}, \quad (5)$$

where  $(\cdot)^H$  denotes complex conjugation and “ $\circ$ ” is the componentwise multiplication. However, the diagonalization is no longer possible when  $\mathbf{S}$  is different from  $\mathbf{I}_{N_h}$ , making the problem (3) much more challenging. More specifically, the direct implementation of this analytical solution (4) requires the inversion of a high dimensional matrix, whose the computational complexity is of order  $\mathcal{O}(N_h^3)$ . One can think of using optimization or simulation-based methods to overcome this computational difficulty. The optimization-based methods, such as the gradient-based methods [16] or more recently ADMM [17], [31], approximate the solution of (4) by iterative updates. The simulation-based methods, e.g., the Markov Chain Monte Carlo methods [39]–[41], look for a the solution  $\hat{\mathbf{x}}$  as the posterior mean of a multivariate Gaussian distribution and propose to draw samples from this distribution. The average of the generated samples yields the minimum mean square error (MMSE) estimator of  $\mathbf{x}$ . Despite this formal appeal, simulation-based methods have the major drawback of being computationally expensive, which prevents their effective use when processing large images.

In this work, we focus on ADMM, which is one of the state-of-art algorithms able to solve convex optimization problems by splitting them into smaller problems, that are easier to handle than the initial problem [42]. The standard ADMM introduces an additional variable  $\mathbf{z} \triangleq \mathbf{H}\mathbf{x}$  and transforms problem (3) into the following constrained optimization problem

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{z}} \quad & \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{z}\|_2^2 + \tau \|\mathbf{x} - \bar{\mathbf{x}}\|_2^2 \\ \text{subject to} \quad & \mathbf{z} = \mathbf{H}\mathbf{x}. \end{aligned} \quad (6)$$

The augmented Lagrangian function associated with the constrained optimization problem (6) is

$$L(\mathbf{x}, \mathbf{z}, \boldsymbol{\lambda}) = \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{z}\|_2^2 + \tau \|\mathbf{x} - \bar{\mathbf{x}}\|_2^2 + \boldsymbol{\lambda}^T (\mathbf{H}\mathbf{x} - \mathbf{z}) + \frac{\mu}{2} \|\mathbf{H}\mathbf{x} - \mathbf{z}\|_2^2 \quad (7)$$

where  $\boldsymbol{\lambda} \in \mathbb{R}^{N_h \times 1}$  is a vector of the Lagrangian multipliers. Using the following variable substitution

$$\mathbf{d} = (1/\mu)\boldsymbol{\lambda}$$

we obtain the scaled augmented Lagrangian function

$$L(\mathbf{x}, \mathbf{z}, \mathbf{d}) = \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{z}\|_2^2 + \tau \|\mathbf{x} - \bar{\mathbf{x}}\|_2^2 + \frac{\mu}{2} \|\mathbf{H}\mathbf{x} - \mathbf{z} + \mathbf{d}\|_2^2 \quad (8)$$

where  $\mathbf{d}$  is a vector of scaled dual variables. The ADMM algorithm designed to solve the optimization problems (3) or (6) can be summarized in the following three steps (each one corresponding to a much easier optimization problem)

$$\begin{aligned} &\text{For } k = 0, 1, \dots \\ &\left[ \begin{array}{l} \mathbf{z}^{k+1} \in \operatorname{argmin}_{\mathbf{z}} \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{z}\|_2^2 + \frac{\mu}{2} \|\mathbf{H}\mathbf{x}^k - \mathbf{z} + \mathbf{d}^k\|_2^2 \\ \mathbf{x}^{k+1} \in \operatorname{argmin}_{\mathbf{x}} \tau \|\mathbf{x} - \bar{\mathbf{x}}\|_2^2 + \frac{\mu}{2} \|\mathbf{H}\mathbf{x} - \mathbf{z}^{k+1} + \mathbf{d}^k\|_2^2 \\ \mathbf{d}^{k+1} = \mathbf{d}^k + (\mathbf{H}\mathbf{x}^{k+1} - \mathbf{z}^{k+1}), \end{array} \right. \quad (9) \end{aligned}$$

where the superscript  $k$  represents the  $k$ th iteration. The solutions of the minimization problems w.r.t.  $\mathbf{z}$  and  $\mathbf{x}$  are reminded hereafter.

- Update  $\mathbf{z}$ : The optimization w.r.t.  $\mathbf{z}$  can be solved as

$$\mathbf{z}^{k+1} = (\mathbf{S}^H \mathbf{S} + \mu \mathbf{I}_{N_h})^{-1} (\mathbf{S}^H \mathbf{y} + \mu (\mathbf{H}\mathbf{x}^k + \mathbf{d}^k)).$$

The update of  $\mathbf{z}$  mainly consists of interpolating zeros and has linear complexity w.r.t. the number of pixels.

- Update  $\mathbf{x}$ : The optimization w.r.t.  $\mathbf{x}$  can be solved as

$$\mathbf{x}^{k+1} = (\mathbf{H}^H \mathbf{H} + 2\tau/\mu \mathbf{I}_{N_h})^{-1} (2\tau/\mu \bar{\mathbf{x}} + \mathbf{H}^H \boldsymbol{\rho})$$

where  $\boldsymbol{\rho} = \mathbf{z}^{k+1} - \mathbf{d}^k$ . It represents a basic deconvolution step that can be efficiently implemented in the Fourier domain.

To summarize, the standard ADMM algorithm decomposes the SR problem into an alternating update between an up-sampling and a deconvolution step, both efficiently solvable [30], [31]. Despite the efficiency of ADMM, it is still appealing to get rid of the iterative update to obtain the solution of (3). Equivalently, is it possible to implement (4) in a non-iterative way? Our paper aims at giving a positive answer to this question by proposing an original strategy that will be illustrated in the following sections, representing the main contribution of this study.

### III. PROPOSED FAST SUPER-RESOLUTION USING AN ANALYTICAL SOLUTION

Before going into further details to solve the problem (4), we list hereafter the two basic assumptions on the blurring and decimation operators used in this work. Note that these

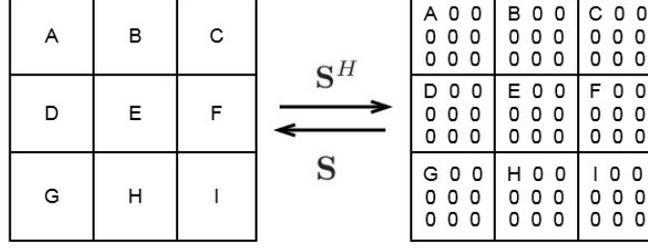


Fig. 1. Effect of up-sampling matrix  $\mathbf{S}^H$  on image of  $3 \times 3$  and down-sampling matrix  $\mathbf{S}$  on the corresponding  $9 \times 9$  image (scale up factor is 3).

assumptions are also commonly used in image deconvolution or image scaling up problems, see e.g., [6], [15], [19], [43], [44].

**Assumption 1.** *The blurring matrix  $\mathbf{H}$  is the matrix representation of the cyclic convolution operator, i.e.,  $\mathbf{H}$  is a block circulant matrix with circulant blocks (BCCB).*

Using the cyclic convolution assumption, the blurring matrix and its conjugate transpose can be decomposed as

$$\mathbf{H} = \mathbf{F}^H \mathbf{\Lambda} \mathbf{F} \quad (10)$$

$$\mathbf{H}^H = \mathbf{F}^H \mathbf{\Lambda}^H \mathbf{F} \quad (11)$$

where the matrices  $\mathbf{F}$  and  $\mathbf{F}^H$  are associated with the Fourier and inverse Fourier transforms (satisfying  $\mathbf{F}\mathbf{F}^H = \mathbf{F}^H\mathbf{F} = \mathbf{I}_{N_h}$ ) and  $\mathbf{\Lambda} = \text{diag}\{\mathbf{F}\mathbf{h}\} \in \mathbb{C}^{N_h \times N_h}$  is a diagonal matrix, whose diagonal elements are the Fourier coefficients of the first column of the blurring matrix  $\mathbf{H}$ , denoted as  $\mathbf{h}$ . Using the decompositions (10) and (11), the blurring operator  $\mathbf{H}\mathbf{x}$  and its conjugate  $\mathbf{H}^H\mathbf{x}$  can be efficiently implemented in the Fourier domain, see, e.g., [45]–[47].

**Assumption 2.** *The decimation matrix  $\mathbf{S} \in \mathbb{R}^{N_l \times N_h}$  is a down-sampling operator, while its conjugate transpose  $\mathbf{S}^H \in \mathbb{R}^{N_h \times N_l}$  interpolates the decimated image with zeros.*

Fig. 1 shows a toy example highlighting the roles of the decimation matrix  $\mathbf{S}$  and its conjugate transpose  $\mathbf{S}^H$ . The decimation matrix satisfies the relationship  $\mathbf{S}\mathbf{S}^H = \mathbf{I}_{N_l}$ . Denoting  $\underline{\mathbf{S}} \triangleq \mathbf{S}^H\mathbf{S}$ , multiplying an image by  $\underline{\mathbf{S}}$  can be achieved by making an entry-wise multiplication with an  $N_h \times N_h$  mask having ones at the sampled positions and zeros elsewhere.

Under the two assumptions mentioned above, the analytical solution of the estimated image

in (4) can be rewritten as

$$\begin{aligned}
\hat{\mathbf{x}} &= (\mathbf{H}^H \mathbf{S}^H \mathbf{S} \mathbf{H} + 2\tau \mathbf{I}_{N_h})^{-1} \mathbf{r} \\
&= (\mathbf{F}^H \mathbf{\Lambda}^H \mathbf{F} \mathbf{S} \mathbf{F}^H \mathbf{\Lambda} \mathbf{F} + 2\tau \mathbf{I}_{N_h})^{-1} \mathbf{r} \\
&= \mathbf{F}^H (\mathbf{\Lambda}^H \mathbf{F} \mathbf{S} \mathbf{F}^H \mathbf{\Lambda} + 2\tau \mathbf{I}_{N_h})^{-1} \mathbf{F} \mathbf{r}.
\end{aligned} \tag{12}$$

where

$$\mathbf{r} \triangleq (\mathbf{H}^H \mathbf{S}^H \mathbf{y} + 2\tau \bar{\mathbf{x}}).$$

In order to compute the analytical solution of (4) or (12) in practical situations, we use hereinafter an interesting property of the matrix  $\mathbf{F} \mathbf{S} \mathbf{F}^H$  summarized in Lemma 1 (see [19]).

**Lemma 1** (Wei *et al.*, [19]). *The following equality holds*

$$\mathbf{F} \mathbf{S} \mathbf{F}^H = \frac{1}{d} \mathbf{J}_d \otimes \mathbf{I}_{N_l} \tag{13}$$

where  $\mathbf{J}_d \in \mathbb{R}^{d \times d}$  is a matrix of ones,  $\mathbf{I}_{N_l} \in \mathbb{R}^{N_l \times N_l}$  is the identity matrix and  $\otimes$  represents the Kronecker product.

Inspired by this elegant property of matrix  $\mathbf{F} \mathbf{S} \mathbf{F}^H$ , the matrix  $\mathbf{\Lambda}^H \mathbf{F} \mathbf{S} \mathbf{F}^H \mathbf{\Lambda}$  in (12) can be rewritten as follows

$$\begin{aligned}
&\mathbf{\Lambda}^H \mathbf{F} \mathbf{S} \mathbf{F}^H \mathbf{\Lambda} \\
&= \frac{1}{d} \mathbf{\Lambda}^H (\mathbf{J}_d \otimes \mathbf{I}_{N_l}) \mathbf{\Lambda}
\end{aligned} \tag{14}$$

$$= \frac{1}{d} \mathbf{\Lambda}^H ((\mathbf{1}_d \mathbf{1}_d^T) \otimes (\mathbf{I}_{N_l} \mathbf{I}_{N_l})) \mathbf{\Lambda} \tag{15}$$

$$= \frac{1}{d} \mathbf{\Lambda}^H (\mathbf{1}_d \otimes \mathbf{I}_{N_l}) (\mathbf{1}_d^T \otimes \mathbf{I}_{N_l}) \mathbf{\Lambda} \tag{16}$$

$$= \frac{1}{d} \left( \mathbf{\Lambda}^H \underbrace{[\mathbf{I}_{N_l}, \dots, \mathbf{I}_{N_l}]^T}_d \right) \left( \underbrace{[\mathbf{I}_{N_l}, \dots, \mathbf{I}_{N_l}] \mathbf{\Lambda}}_d \right) \tag{17}$$

$$= \frac{1}{d} \underline{\mathbf{\Lambda}}^H \underline{\mathbf{\Lambda}}. \tag{18}$$

Note that (15) was obtained from (14) by replacing  $\mathbf{J}_d$  by  $\mathbf{1}_d \mathbf{1}_d^T$ , with  $\mathbf{1}_d \in \mathbb{R}^{d \times 1}$  a vector of ones. Obtaining (16) from (15) is straightforward using the following property of the Kronecker product

$$\mathcal{A} \mathcal{B} \otimes \mathcal{C} \mathcal{D} = (\mathcal{A} \otimes \mathcal{C})(\mathcal{B} \otimes \mathcal{D}).$$

In (17),  $\mathbf{\Lambda} \in \mathbb{R}^{N_h \times N_h}$ ,  $\underbrace{(\mathbf{I}_{N_l}, \dots, \mathbf{I}_{N_l})}_d \in \mathbb{R}^{N_l \times N_h}$  and  $\underbrace{(\mathbf{I}_{N_l}, \dots, \mathbf{I}_{N_l})^T}_d \in \mathbb{R}^{N_h \times N_l}$  are block matrices whose blocks are equal to the identity matrix  $\mathbf{I}_{N_l}$ .

By rewriting  $\mathbf{\Lambda}$  as  $\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{\Lambda}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{\Lambda}_d \end{bmatrix}$ , the matrix  $\underline{\mathbf{\Lambda}} \in \mathbb{R}^{N_l \times N_h}$  in (18) can be

obtained as follows

$$\begin{aligned} \underline{\mathbf{\Lambda}} &= [\mathbf{I}_{N_l}, \dots, \mathbf{I}_{N_l}] \mathbf{\Lambda} \\ &= [\mathbf{I}_{N_l}, \dots, \mathbf{I}_{N_l}] \begin{bmatrix} \mathbf{\Lambda}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{\Lambda}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{\Lambda}_d \end{bmatrix} \\ &= [\mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \dots, \mathbf{\Lambda}_d]. \end{aligned} \quad (19)$$

Note that each block  $\mathbf{\Lambda}_i \in \mathbb{R}^{N_l \times N_l}$  (for  $i = 1, \dots, d$ ) is a diagonal matrix. To further simplify the expression (12), the following well-know matrix inverse lemma is employed.

**Lemma 2** (Matrix inverse lemma/Woodbury formula [48]). *The following equality holds conditional on the existence of  $\mathbf{A}_1^{-1}$  and  $\mathbf{A}_3^{-1}$*

$$\begin{aligned} &(\mathbf{A}_1 + \mathbf{A}_2 \mathbf{A}_3 \mathbf{A}_4)^{-1} \\ &= \mathbf{A}_1^{-1} - \mathbf{A}_1^{-1} \mathbf{A}_2 (\mathbf{A}_3^{-1} + \mathbf{A}_4 \mathbf{A}_1^{-1} \mathbf{A}_2)^{-1} \mathbf{A}_4 \mathbf{A}_1^{-1}, \end{aligned} \quad (20)$$

where  $\mathbf{A}_1$ ,  $\mathbf{A}_2$ ,  $\mathbf{A}_3$  and  $\mathbf{A}_4$  are matrices of the correct sizes.

Combining the Woodbury formula in Lemma 2 with (18), the analytical solution (12) can be computed very efficiently as stated in the following theorem.

**Theorem 1.** *The solution  $\bar{\mathbf{x}}$  in (4) can be computed as*

$$\hat{\mathbf{x}} = \frac{1}{2\tau} \mathbf{r} - \frac{1}{2\tau} \mathbf{F}^H \underline{\mathbf{\Lambda}}^H (2\tau d \mathbf{I}_{N_l} + \underline{\mathbf{\Lambda}} \underline{\mathbf{\Lambda}}^H)^{-1} \underline{\mathbf{\Lambda}} \mathbf{F} \mathbf{r} \quad (21)$$

where  $\mathbf{r}$  and  $\underline{\mathbf{\Lambda}}$  are defined in (III) and (19) respectively.

*Proof:* Substituting (18) into (12) leads to

$$\hat{\mathbf{x}} = \mathbf{F}^H \left( \frac{1}{d} \underline{\mathbf{\Lambda}}^H \underline{\mathbf{\Lambda}} + 2\tau \mathbf{I}_{N_h} \right)^{-1} \mathbf{F} \mathbf{r}. \quad (22)$$

By using Lemma 2, the inverse matrix in (22) can be simplified leading to

$$\begin{aligned}\hat{\mathbf{x}} &= \mathbf{F}^H \left( \frac{1}{2\tau} \mathbf{I}_{N_h} - \frac{1}{2\tau} \underline{\mathbf{\Lambda}}^H \left( d\mathbf{I}_{N_l} + \underline{\mathbf{\Lambda}} \underline{\mathbf{\Lambda}}^H \frac{1}{2\tau} \right)^{-1} \underline{\mathbf{\Lambda}} \frac{1}{2\tau} \right) \mathbf{F} \mathbf{r} \\ &= \frac{1}{2\tau} \mathbf{r} - \frac{1}{2\tau} \mathbf{F}^H \underline{\mathbf{\Lambda}}^H \left( 2\tau d\mathbf{I}_{N_l} + \underline{\mathbf{\Lambda}} \underline{\mathbf{\Lambda}}^H \right)^{-1} \underline{\mathbf{\Lambda}} \mathbf{F} \mathbf{r}.\end{aligned}$$

■

Algorithm 1 summarizes the implementation of the proposed SR solution resulting from Theorem 1, which is referred to as *fast super-resolution (FSR)* approach. Note that the main contribution here is that the solution  $\hat{\mathbf{x}}$  can be calculated directly, requiring no iterative steps as well as avoiding any tuning of parameters, such as  $\mu$ . The proposed approach is “fast” in the sense that each step in Algorithm 1 is easy and light to implement. A more detailed complexity analysis is given hereinafter.

---

**Algorithm 1:** FSR implementing the analytical solution (4)

---

**Input:**  $\mathbf{y}$ ,  $\mathbf{H}$ ,  $\mathbf{S}$ ,  $\bar{\mathbf{x}}$ ,  $\tau$ ,  $d$

// Factorization of  $\mathbf{H}$  (FFT of the blurring kernel):  $\mathbf{H} = \mathbf{F}\mathbf{\Lambda}\mathbf{F}^H$

1  $\mathbf{\Lambda} \leftarrow \text{Dec}(\mathbf{H})$ ;

// Compute  $\underline{\mathbf{\Lambda}}$

2  $\underline{\mathbf{\Lambda}} = \mathbf{\Lambda}\mathbf{\Lambda}^H$ ;

// Calculate FFT of  $\mathbf{r}$ :  $\mathbf{F}\mathbf{r}$

3  $\mathbf{F}\mathbf{r} \leftarrow \mathbf{F}(\mathbf{H}^H \mathbf{S}^H \mathbf{y} + 2\tau \bar{\mathbf{x}})$ ;

// Hadamard (or entrywise) product in frequency domain

4  $\mathbf{x}_f \leftarrow \left( \underline{\mathbf{\Lambda}}^H \left( 2\tau d\mathbf{I}_{N_l} + \underline{\mathbf{\Lambda}} \underline{\mathbf{\Lambda}}^H \right)^{-1} \underline{\mathbf{\Lambda}} \right) \mathbf{F}\mathbf{r}$  ;

// inverse FFT and subtract from  $\mathbf{r}$

5  $\hat{\mathbf{x}} \leftarrow \frac{1}{2\tau} \mathbf{r} - \frac{1}{2\tau} \mathbf{F}^H \mathbf{x}_f$  ;

**Output:**  $\hat{\mathbf{x}}$

---

### Complexity Analysis

In Algorithm 1, Steps 2 and 4 have a low complexity of the order  $\mathcal{O}(N_h)$ . Note that in Step 4, though the implementation is in the frequency domain, the matrix  $2\tau d\mathbf{I}_{N_l} + \underline{\mathbf{\Lambda}} \underline{\mathbf{\Lambda}}^H$  to be inverted is a real matrix as  $\underline{\mathbf{\Lambda}} \underline{\mathbf{\Lambda}}^H$  is real, implying that only  $N_l$  divisions are required.

Steps 1, 3 and 5 have a complexity of the order  $\mathcal{O}(N_h \log N_h)$  because of the FFT or iFFT operations. More specifically, there is one FFT/iFFT in Step 1/ Step 5 while two FFT

operations are necessary in Step 3, where the matrix  $\mathbf{F}\mathbf{r}$  can be implemented in the Fourier domain as

$$\begin{aligned}\mathbf{F}\mathbf{r} &= \mathbf{F}(\mathbf{F}^H \boldsymbol{\Lambda}^H \mathbf{F}\mathbf{S}^H \mathbf{y} + 2\tau\bar{\mathbf{x}}) \\ &= \boldsymbol{\Lambda}^H \mathbf{F}\mathbf{S}^H \mathbf{y} + 2\tau\mathbf{F}\bar{\mathbf{x}}.\end{aligned}\tag{23}$$

Thus, the most computationally expensive part lies in the implementation of Step 3. In total, four FFT/iFFT computations are required in our implementation. Comparing with the original problem (4), the order of computation complexity has decreased significantly from  $\mathcal{O}(N_h^3)$  to  $\mathcal{O}(N_h \log N_h)$ , which allows the analytical solution (21) to be computed efficiently.

#### IV. GENERALIZATION OF THE FAST SUPER-RESOLUTION SCHEME

In addition to the Tikhonov regularizer (Gaussian prior) addressed in the previous section, there are other various regularizers (associated with non-Gaussian priors) in either image or transformed domains, commonly used for single image SR. In this section, we will show that the analytical solution studied in Section III can be generalized to solve the SR problem associated with more complicated regularizations. This generalization is accomplished by embedding the analytical solution (21) in the standard ADMM framework, leading to a fast version of ADMM, denoted by ‘‘FSR-ADMM’’ hereafter.

In the following, we focus on two regularizations widely used for the SR problem, *i.e.*, the TV regularization in the image domain [17], [49], [50] and the  $\ell_1$ -norm regularization (Laplace prior) in the wavelet domain [25]. In both cases, we start by recalling a standard way of solving the optimization problem with conventional ADMM and then present the proposed fast scheme.

##### A. TV regularization

TV regularization is well-known for its ability to remove unwanted details whilst preserving the important ones such as edges. Consequently, it has been widely used in reconstruction applications, such as image SR [5], [49], [50] or compressed sensing [2]. Considering the TV regularizer, *i.e.*,  $\phi(\mathbf{x}) = \|\mathbf{x}\|_{\text{TV}}$ , we obtain the following optimization problem

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{H}\mathbf{x}\|_2^2 + \tau \|\mathbf{x}\|_{\text{TV}}\tag{24}$$

where

$$\|\mathbf{x}\|_{\text{TV}} = \sqrt{\|\mathbf{D}_h \mathbf{x}\|_2^2 + \|\mathbf{D}_v \mathbf{x}\|_2^2}$$

is the isotropic TV [49],  $\mathbf{D}_h \mathbf{x}$  and  $\mathbf{D}_v \mathbf{x}$  represent the horizontal and vertical spatial derivatives of the image  $\mathbf{x}$ .

1) “direct” ADMM with TV regularization: In the classical ADMM scheme, problem (24) can be solved using the variable splitting strategy as recalled below [17]

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{u}_{\text{TV}}} \quad & \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{u}_{\text{TV1}}\|_2^2 + \tau \sqrt{\|\mathbf{u}_{\text{TV2}}\|_2^2 + \|\mathbf{u}_{\text{TV3}}\|_2^2} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{u}_{\text{TV}}, \end{aligned} \quad (25)$$

where  $\mathbf{A} \in \mathbb{R}^{3N_h \times N_h}$  and  $\mathbf{u}_{\text{TV}} \in \mathbb{R}^{3N_h \times 1}$  have the following definitions

$$\mathbf{A} = \begin{bmatrix} \mathbf{H} \\ \mathbf{D}_h \\ \mathbf{D}_v \end{bmatrix} \quad \text{and} \quad \mathbf{u}_{\text{TV}} = \begin{bmatrix} \mathbf{u}_{\text{TV1}} \\ \mathbf{u}_{\text{TV2}} \\ \mathbf{u}_{\text{TV3}} \end{bmatrix}$$

where the vectors  $\mathbf{u}_{\text{TV1}}$ ,  $\mathbf{u}_{\text{TV2}}$  and  $\mathbf{u}_{\text{TV3}}$  belong to  $\mathbb{R}^{N_h \times 1}$ . The augmented Lagrangian function of the constrained optimization problem (25) is

$$L(\mathbf{x}, \mathbf{u}_{\text{TV}}, \boldsymbol{\lambda}_{\text{TV}}) = \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{u}_{\text{TV1}}\|_2^2 + \tau \sqrt{\|\mathbf{u}_{\text{TV2}}\|_2^2 + \|\mathbf{u}_{\text{TV3}}\|_2^2} + \boldsymbol{\lambda}_{\text{TV}}^T (\mathbf{A}\mathbf{x} - \mathbf{u}_{\text{TV}}) + \frac{\mu}{2} \|\mathbf{A}\mathbf{x} - \mathbf{u}_{\text{TV}}\|_2^2$$

or equivalently

$$L(\mathbf{x}, \mathbf{u}_{\text{TV}}, \mathbf{d}_{\text{TV}}) = \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{u}_{\text{TV1}}\|_2^2 + \tau \sqrt{\|\mathbf{u}_{\text{TV2}}\|_2^2 + \|\mathbf{u}_{\text{TV3}}\|_2^2} + \frac{\mu}{2} \|\mathbf{A}\mathbf{x} - \mathbf{u}_{\text{TV}} + \mathbf{d}_{\text{TV}}\|_2^2.$$

The dual variables  $\mathbf{d}_{\text{TV}} \in \mathbb{R}^{3N_h \times 1}$  are summarized in the following vector

$$\mathbf{d}_{\text{TV}} = \left[ \mathbf{d}_{\text{TV1}}^T, \mathbf{d}_{\text{TV2}}^T, \mathbf{d}_{\text{TV3}}^T \right]^T$$

where the vectors  $\mathbf{d}_{\text{TV1}}$ ,  $\mathbf{d}_{\text{TV2}}$ , and  $\mathbf{d}_{\text{TV3}} \in \mathbb{R}^{N_h \times 1}$ . The “direct” ADMM scheme solving (24) can be summarized into the three steps provided below

For  $k = 0, 1, \dots$

$$\begin{cases} \mathbf{x}^{k+1} \in \operatorname{argmin}_{\mathbf{x}} \frac{\mu}{2} \|\mathbf{A}\mathbf{x} - \mathbf{u}_{\text{TV}}^k + \mathbf{d}_{\text{TV}}^k\|_2^2 \\ \mathbf{u}_{\text{TV}}^{k+1} \in \operatorname{argmin}_{\mathbf{u}_{\text{TV}}} \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{u}_{\text{TV1}}\|_2^2 + \\ \quad \tau \sqrt{\|\mathbf{u}_{\text{TV2}}\|_2^2 + \|\mathbf{u}_{\text{TV3}}\|_2^2} + \frac{\mu}{2} \|\mathbf{A}\mathbf{x}^{k+1} - \mathbf{u}_{\text{TV}} + \mathbf{d}_{\text{TV}}^k\|_2^2 \\ \mathbf{d}_{\text{TV}}^{k+1} = \mathbf{d}_{\text{TV}}^k + (\mathbf{A}\mathbf{x}^{k+1} - \mathbf{u}_{\text{TV}}^{k+1}) \end{cases} \quad (26)$$

where the superscript  $k$  denotes the  $k$ th iteration of the algorithm. The optimization problems in (26) w.r.t.  $\mathbf{x}$  and  $\mathbf{u}_{\text{TV}}$  are reminded hereinafter.

- Update  $\mathbf{x}$ :

$$\mathbf{x}^{k+1} = (\mathbf{H}^H \mathbf{H} + \mathbf{D}_h^H \mathbf{D}_h + \mathbf{D}_v^H \mathbf{D}_v)^{-1} \boldsymbol{\rho}_{\text{TV}}$$

where  $\boldsymbol{\rho}_{\text{TV}} = \mathbf{H}^H \boldsymbol{\rho}_{\text{TV1}} + \mathbf{D}_h^H \boldsymbol{\rho}_{\text{TV2}} + \mathbf{D}_v^H \boldsymbol{\rho}_{\text{TV3}}$ ,  $\boldsymbol{\rho}_{\text{TV1}} = \mathbf{u}_{\text{TV1}}^k - \mathbf{d}_{\text{TV1}}^k$ ,  $\boldsymbol{\rho}_{\text{TV2}} = \mathbf{u}_{\text{TV2}}^k - \mathbf{d}_{\text{TV2}}^k$  and  $\boldsymbol{\rho}_{\text{TV3}} = \mathbf{u}_{\text{TV3}}^k - \mathbf{d}_{\text{TV3}}^k$ .

- Update  $\mathbf{u}_{\text{TV}}$ : The optimization w.r.t.  $\mathbf{u}_{\text{TV}}$  can be divided into three sub-steps w.r.t.  $\mathbf{u}_{\text{TV1}}$ ,  $\mathbf{u}_{\text{TV2}}$  and  $\mathbf{u}_{\text{TV3}}$ :

$$\begin{aligned} \mathbf{u}_{\text{TV1}}^{k+1} &= \underset{\mathbf{u}_{\text{TV1}}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{u}_{\text{TV1}}\|_2^2 \\ &\quad + \frac{\mu}{2} \|\mathbf{H}\mathbf{x}^{k+1} - \mathbf{u}_{\text{TV1}} + \mathbf{d}_{\text{TV1}}^k\|_2^2 \\ &= (\mathbf{S}^H \mathbf{S} + \mu \mathbf{I}_{N_h})^{-1} (\mathbf{S}^H \mathbf{y} + \mathbf{H}\mathbf{x}^{k+1} + \mathbf{d}_{\text{TV1}}^k) \\ \mathbf{u}_{\text{TV2}}^{k+1} &= \underset{\mathbf{u}_{\text{TV2}}}{\operatorname{argmin}} \tau \sqrt{\|\mathbf{u}_{\text{TV2}}\|^2 + \|\mathbf{u}_{\text{TV3}}\|^2} + \frac{\mu}{2} \|\mathbf{u}_{\text{TV2}} - \boldsymbol{\nu}_1\|_2^2 \\ \mathbf{u}_{\text{TV3}}^{k+1} &= \underset{\mathbf{u}_{\text{TV3}}}{\operatorname{argmin}} \tau \sqrt{\|\mathbf{u}_{\text{TV2}}\|^2 + \|\mathbf{u}_{\text{TV3}}\|^2} + \frac{\mu}{2} \|\mathbf{u}_{\text{TV3}} - \boldsymbol{\nu}_2\|_2^2 \end{aligned}$$

where  $\boldsymbol{\nu}_1 = (\mathbf{D}_h \mathbf{x}^{k+1} + \mathbf{d}_{\text{TV2}}^k)$  and  $\boldsymbol{\nu}_2 = (\mathbf{D}_v \mathbf{x}^{k+1} + \mathbf{d}_{\text{TV3}}^k)$ . Both optimizations w.r.t.  $\mathbf{u}_{\text{TV2}}$  and  $\mathbf{u}_{\text{TV3}}$  can be obtained numerically using the vector-soft-thresholding, based on the fact that they are pixel-wise decoupled [49], [51], [52]. More specifically, by denoting  $\mathbf{u}_{2,3} = [\mathbf{u}_{\text{TV2}}, \mathbf{u}_{\text{TV3}}]$  and  $\boldsymbol{\nu} = [\boldsymbol{\nu}_1, \boldsymbol{\nu}_2]$  ( $\mathbf{u}_{2,3}, \boldsymbol{\nu} \in \mathbb{R}^{N_h \times 2}$ ), we have

$$\mathbf{u}_{2,3}[i] = \operatorname{prox}_{\|\cdot\|_2}(\boldsymbol{\nu}[i]) = \max\{\mathbf{0}, \|\boldsymbol{\nu}[i]\|_2 - \tau/\mu\} \frac{\boldsymbol{\nu}[i]}{\|\boldsymbol{\nu}[i]\|_2} \quad (27)$$

where the vectors  $\mathbf{u}_{2,3}[i], \boldsymbol{\nu}[i] \in \mathbb{R}^{1 \times 2}$  are the first lines of the corresponding matrices, for  $i \in \{1, \dots, N_h\}$ .

2) *Fast ADMM with TV regularization*: Compared to the classical ADMM implementation shown above, we propose a faster ADMM scheme by using the developed analytical solution in Section III. More specifically, we show in what follows that the use of this analytical solution allows a more compact variable splitting strategy, *i.e.*, two variables splitting instead of three in Section IV-A1 as follows

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{u}_{\text{TVF}}} \quad & \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{H}\mathbf{x}\|_2^2 + \tau \sqrt{\|\mathbf{u}_{\text{TVF1}}\|^2 + \|\mathbf{u}_{\text{TVF2}}\|^2} \\ \text{subject to} \quad & \mathbf{B}\mathbf{x} = \mathbf{u}_{\text{TVF}} \end{aligned} \quad (28)$$

where  $\mathbf{B} \in \mathbb{R}^{2N_h \times N_h}$  and  $\mathbf{u}_{\text{TVF}} \in \mathbb{R}^{2N_h \times 1}$  are defined by

$$\mathbf{B} = \begin{bmatrix} \mathbf{D}_h \\ \mathbf{D}_v \end{bmatrix} \quad \text{and} \quad \mathbf{u}_{\text{TVF}} = \begin{bmatrix} \mathbf{u}_{\text{TVF1}} \\ \mathbf{u}_{\text{TVF2}} \end{bmatrix}.$$

The corresponding augmented Lagrange function is

$$L(\mathbf{x}, \mathbf{u}_{\text{TVF}}, \boldsymbol{\lambda}_{\text{TVF}}) = \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{H}\mathbf{x}\|_2^2 + \tau \sqrt{\|\mathbf{u}_{\text{TVF1}}\|^2 + \|\mathbf{u}_{\text{TVF2}}\|^2} + \boldsymbol{\lambda}_{\text{TVF}}^T (\mathbf{B}\mathbf{x} - \mathbf{u}_{\text{TVF}}) + \frac{\mu}{2} \|\mathbf{B}\mathbf{x} - \mathbf{u}_{\text{TVF}}\|_2^2 \quad (29)$$

or equivalently

$$L(\mathbf{x}, \mathbf{u}_{\text{TVF}}, \mathbf{d}_{\text{TVF}}) = \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{H}\mathbf{x}\|_2^2 + \tau \sqrt{\|\mathbf{u}_{\text{TVF1}}\|^2 + \|\mathbf{u}_{\text{TVF2}}\|^2} + \frac{\mu}{2} \|\mathbf{B}\mathbf{x} - \mathbf{u}_{\text{TVF}} + \mathbf{d}_{\text{TVF}}\|_2^2 \quad (30)$$

where the dual variables are contained in a vector  $\in \mathbb{R}^{2N_h \times 1}$

$$\mathbf{d}_{\text{TVF}} = \begin{bmatrix} \mathbf{d}_{\text{TVF1}} \\ \mathbf{d}_{\text{TVF2}} \end{bmatrix}.$$

The fast SR algorithm proposed to solve the optimization problem (28) or (24) can be summarized into three steps detailed below:

For  $k = 0, 1, \dots$

$$\begin{cases} \mathbf{x}^{k+1} \in \operatorname{argmin}_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{H}\mathbf{x}\|_2^2 + \frac{\mu}{2} \|\mathbf{B}\mathbf{x} - \mathbf{u}_{\text{TVF}}^k + \mathbf{d}_{\text{TVF}}^k\|_2^2 \\ \mathbf{u}_{\text{TVF}}^{k+1} \in \operatorname{argmin}_{\mathbf{u}_{\text{TVF}}} \tau \sqrt{\|\mathbf{u}_{\text{TVF1}}\|^2 + \|\mathbf{u}_{\text{TVF2}}\|^2} \\ \quad + \frac{\mu}{2} \|\mathbf{B}\mathbf{x}^{k+1} - \mathbf{u}_{\text{TVF}} + \mathbf{d}_{\text{TVF}}^k\|_2^2 \\ \mathbf{d}_{\text{TVF}}^{k+1} = \mathbf{d}_{\text{TVF}}^k + (\mathbf{B}\mathbf{x}^{k+1} - \mathbf{u}_{\text{TVF}}^{k+1}). \end{cases} \quad (31)$$

The step updating  $\mathbf{x}$  in (31) can be treated as an SR problem with Gaussian regularization, and is thus efficiently solved using the fast implementation (21) derived in Section III

$$\begin{aligned} \mathbf{x}^{k+1} &= (\mathbf{H}^H \mathbf{S}^H \mathbf{S} \mathbf{H} + \mu \mathbf{D}_h^H \mathbf{D}_h + \mu \mathbf{D}_v^H \mathbf{D}_v)^{-1} \mathbf{r} \\ &= \mathbf{F}^H (\boldsymbol{\Lambda}^H \mathbf{F} \mathbf{S} \mathbf{F}^H \boldsymbol{\Lambda} + \mu \boldsymbol{\Sigma}_h^H \boldsymbol{\Sigma}_h + \mu \boldsymbol{\Sigma}_v^H \boldsymbol{\Sigma}_v)^{-1} \mathbf{F} \mathbf{r} \\ &= \mathbf{F}^H \left( \frac{1}{d} \boldsymbol{\Lambda}^H \boldsymbol{\Lambda} + \mu \boldsymbol{\Sigma}_h^H \boldsymbol{\Sigma}_h + \mu \boldsymbol{\Sigma}_v^H \boldsymbol{\Sigma}_v \right)^{-1} \mathbf{F} \mathbf{r} \\ &= \mathbf{F}^H \left( \frac{1}{\mu} \boldsymbol{\Sigma} - \frac{1}{\mu} \boldsymbol{\Sigma} \boldsymbol{\Lambda}^H \left( d \mathbf{I}_{N_i} + \frac{1}{\mu} \boldsymbol{\Lambda} \boldsymbol{\Sigma} \boldsymbol{\Lambda}^H \right)^{-1} \boldsymbol{\Lambda} \boldsymbol{\Sigma} \frac{1}{\mu} \right) \mathbf{F} \mathbf{r} \\ &= \frac{1}{\mu} \mathbf{F}^H \boldsymbol{\Sigma} \mathbf{F} \mathbf{r} - \frac{1}{\mu} \mathbf{F}^H \boldsymbol{\Sigma} \boldsymbol{\Lambda}^H (d \mu \mathbf{I}_{N_i} + \boldsymbol{\Lambda} \boldsymbol{\Sigma} \boldsymbol{\Lambda}^H)^{-1} \mathbf{F} \mathbf{r} \end{aligned} \quad (32)$$

where  $\mathbf{r} = (\mathbf{H}^H \mathbf{S}^H \mathbf{y} + \mu \mathbf{D}_h^H \boldsymbol{\rho}_h + \mu \mathbf{D}_v^H \boldsymbol{\rho}_v)$  with  $\boldsymbol{\rho}_h = \mathbf{u}_{\text{TVF1}}^k - \mathbf{d}_{\text{TVF1}}^k$  and  $\boldsymbol{\rho}_v = \mathbf{u}_{\text{TVF2}}^k - \mathbf{d}_{\text{TVF2}}^k$ , and  $\boldsymbol{\Sigma} = [\boldsymbol{\Sigma}_h^H \boldsymbol{\Sigma}_h + \boldsymbol{\Sigma}_v^H \boldsymbol{\Sigma}_v]^{-1}$ . Note that  $\boldsymbol{\Sigma}_h$  and  $\boldsymbol{\Sigma}_v$  are two diagonal matrices whose elements are the Fourier transforms of the first columns of matrices  $\mathbf{D}_h$  and respectively  $\mathbf{D}_v$ . The analytical solution (32) is obtained using the cyclic boundary assumption of the derivative

operator, implying that  $\mathbf{D}_h$  and  $\mathbf{D}_v$  are BCCB matrices. Thus, (32) can be implemented efficiently using the results of Section III.

The step updating  $\mathbf{u}_{\text{TVF}}$  can be efficiently solved by the vector-soft-thresholding operator in (27). The resulting fast SR method with TV regularization is summarized in Algo. 2.

---

**Algorithm 2:** FSR-ADMM with TV regularization

---

- 1 Set  $k = 0$ , choose  $\mu > 0$ ,  $\mathbf{x}^0$ ,  $\mathbf{d}^0$ , input ;
  - 2 Repeat
    - // Update  $\mathbf{x}$  using the analytic solution in (32)
    - 3  $\boldsymbol{\rho}_h = \mathbf{u}_{\text{TVF1}}^k - \mathbf{d}_{\text{TVF1}}^k$ ;
    - 4  $\boldsymbol{\rho}_v = \mathbf{u}_{\text{TVF2}}^k - \mathbf{d}_{\text{TVF2}}^k$  ;
    - 5  $\mathbf{Fr} = \mathbf{F}(\mathbf{H}^H \mathbf{S}^H \mathbf{y} + \mu \mathbf{D}_h \boldsymbol{\rho}_h + \mu \mathbf{D}_v \boldsymbol{\rho}_v)$ ;
    - 6  $\boldsymbol{\Lambda} \leftarrow \text{Dec}(\mathbf{H})$ ;
    - 7  $\underline{\boldsymbol{\Lambda}} = \boldsymbol{\Lambda} \boldsymbol{\Lambda}^H$ ;
    - 8  $\boldsymbol{\Sigma}_h \leftarrow \text{Dec}(\mathbf{D}_h)$ ;
    - 9  $\boldsymbol{\Sigma}_v \leftarrow \text{Dec}(\mathbf{D}_v)$ ;
    - 10  $\mathbf{P} \leftarrow (\boldsymbol{\Sigma}_h^H \boldsymbol{\Sigma}_h + \boldsymbol{\Sigma}_v^H \boldsymbol{\Sigma}_v)^{-1}$  ;
    - 11  $\mathbf{x}_f \leftarrow \left( \mathbf{P} \underline{\boldsymbol{\Lambda}}^H (\mu d \mathbf{I}_{N_t} + \underline{\boldsymbol{\Lambda}} \mathbf{P} \underline{\boldsymbol{\Lambda}}^H)^{-1} \underline{\boldsymbol{\Lambda}} \mathbf{P} \right) \mathbf{Fr}$  ;
    - 12  $\hat{\mathbf{x}} \leftarrow \frac{1}{\mu} \mathbf{F}^H \mathbf{P} \mathbf{Fr} - \frac{1}{\mu} \mathbf{F}^H \mathbf{x}_f$  ;
    - // Update  $\mathbf{u}$  using the vector-soft-thresholding operator
    - 13  $\boldsymbol{\nu} = [\mathbf{D}_h \mathbf{x}^{k+1} + \mathbf{d}_{\text{TVF1}}^k, \mathbf{D}_v \mathbf{x}^{k+1} + \mathbf{d}_{\text{TVF1}}^k]$  ;
    - 14  $\mathbf{u}_{\text{TVF}}^{k+1}[i] = \max\{\mathbf{0}, \|\boldsymbol{\nu}[i]\|_2 - \tau/\mu\} \frac{\boldsymbol{\nu}[i]}{\|\boldsymbol{\nu}[i]\|_2}$  ;
    - // Update the dual variables  $\mathbf{d}$
    - 15  $\mathbf{d}_{\text{TVF}}^{k+1} = \mathbf{d}_{\text{TVF}}^k + (\mathbf{B} \mathbf{x}^{k+1} - \mathbf{u}_{\text{TVF}}^{k+1})$ ;
    - 16  $k \leftarrow k + 1$
    - 17 until stopping criterion is satisfied.
- 

### B. $\ell_1$ -norm regularization in the wavelet domain

Single image enhancement in transformed domains such as the wavelet domain (based on the sparsity of wavelet coefficients of piecewise constant images) has been widely studied in the literature, see e.g., [11], [25]. In this section, we consider the SR problem with

Laplace/ $\ell_1$ -norm regularization in the wavelet domain. To our knowledge, most of the SR methods implemented in the wavelet domain only consider the blurring kernel, resulting into a deconvolution problem (e.g., see [23]). However, in this paper, we take into account both the decimation operator and the blurring kernel as considered in the previous sections. We assume that  $\mathbf{x}$  can be decomposed as a linear combination of wavelets, *i.e.*,  $\mathbf{x} = \mathbf{W}\boldsymbol{\theta}$ , where  $\boldsymbol{\theta} \in \mathbb{R}^{M \times 1}$  is the vector containing the wavelet coefficients,  $M$  is the number of wavelets considered in the decomposition,  $\mathbf{W} \in \mathbb{R}^{N_h \times M}$  and  $\mathbf{W}^H \in \mathbb{R}^{M \times N_h}$  are the wavelet and inverse wavelet transforms, thus satisfying the relationship  $\mathbf{W}\mathbf{W}^H = \mathbf{I}_{N_h}$ ,  $\mathbf{W}^H\mathbf{W} = \mathbf{I}_M$ , where  $\mathbf{I}_q$  is the identity matrix  $\in \mathbb{R}^{q \times q}$ ,  $q \in \{M, N_h\}$ . The single image SR problem can then be rewritten as

$$\min_{\boldsymbol{\theta}} \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{H}\mathbf{W}\boldsymbol{\theta}\|_2^2 + \tau \|\boldsymbol{\theta}\|_1. \quad (33)$$

Solving (33) provides the MAP estimator of the wavelet coefficients, which in turn, via an inverse wavelet transform, provide the super-resolved image.

1) “*direct*” ADMM with  $\ell_1$ -norm regularization: Considering the traditional variable splitting method, (33) can be transformed into the following optimization problem

$$\begin{aligned} \min_{\boldsymbol{\theta}, \mathbf{u}_L} \quad & \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{u}_{L1}\|_2^2 + \tau \|\mathbf{u}_{L2}\|_1 \\ \text{subject to} \quad & \mathbf{G}\boldsymbol{\theta} = \mathbf{u}_L \end{aligned} \quad (34)$$

where  $\mathbf{G} \in \mathbb{R}^{(N_h+M) \times M}$ ,  $\mathbf{u}_L \in \mathbb{R}^{(N_h+M) \times 1}$  are defined as

$$\mathbf{G} = \begin{bmatrix} \mathbf{H}\mathbf{W} \\ \mathbf{I} \end{bmatrix}, \quad \mathbf{u}_L = \begin{bmatrix} \mathbf{u}_{L1} \\ \mathbf{u}_{L2} \end{bmatrix}.$$

with  $\mathbf{u}_{L1} \in \mathbb{R}^{N_h \times 1}$  and  $\mathbf{u}_{L2} \in \mathbb{R}^{M \times 1}$ . The corresponding augmented Lagrange function associated with (34) is

$$L(\boldsymbol{\theta}, \mathbf{u}_L, \boldsymbol{\lambda}_L) = \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{u}_{L1}\|_2^2 + \tau \|\mathbf{u}_{L2}\|_1 + \boldsymbol{\lambda}_L^T (\mathbf{G}\boldsymbol{\theta} - \mathbf{u}_L) + \frac{\mu}{2} \|\mathbf{G}\boldsymbol{\theta} - \mathbf{u}_L\|_2^2 \quad (35)$$

or equivalently

$$L(\boldsymbol{\theta}, \mathbf{u}_L, \mathbf{d}_L) = \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{u}_{L1}\|_2^2 + \tau \|\mathbf{u}_{L2}\|_1 + \frac{\mu}{2} \|\mathbf{G}\boldsymbol{\theta} - \mathbf{u}_L + \mathbf{d}_L\|_2^2 \quad (36)$$

where the dual variables are contained in the vector

$$\mathbf{d}_L = \left[ \mathbf{d}_{L1}^T, \mathbf{d}_{L2}^T \right]^T \in \mathbb{R}^{(N_h+M) \times 1}$$

with  $\mathbf{d}_{L1} \in \mathbb{R}^{N_h \times 1}$  and  $\mathbf{d}_{L2} \in \mathbb{R}^{M \times 1}$ . The ‘‘direct’’ ADMM solving (33) can be summarized into three steps

For  $k = 0, \dots$

$$\left[ \begin{array}{l} \boldsymbol{\theta}^{k+1} \in \operatorname{argmin}_{\boldsymbol{\theta}} \frac{\mu}{2} \|\mathbf{G}\boldsymbol{\theta} - \mathbf{u}_L^k + \mathbf{d}_L^k\|_2^2 \\ \mathbf{u}_L^{k+1} \in \operatorname{argmin}_{\mathbf{u}_L} \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{u}_{L1}\|_2^2 + \tau \|\mathbf{u}_{L2}\|_1 \\ \quad + \frac{\mu}{2} \|\mathbf{G}\boldsymbol{\theta}^{k+1} - \mathbf{u}_L + \mathbf{d}_L^k\|_2^2 \\ \mathbf{d}_L^{k+1} = \mathbf{d}_L^k + (\mathbf{G}\boldsymbol{\theta}^{k+1} - \mathbf{u}_L^{k+1}) \end{array} \right. \quad (37)$$

The optimization problems w.r.t.  $\boldsymbol{\theta}$  and  $\mathbf{u}_L$  are detailed below

- Update  $\boldsymbol{\theta}$ :

$$\boldsymbol{\theta}^{k+1} = (\mathbf{W}^H \mathbf{H}^H \mathbf{H} \mathbf{W} + \mathbf{I}_M)^{-1} (\mathbf{W}^H \mathbf{H}^H \boldsymbol{\rho}_{L1} + \boldsymbol{\rho}_{L2}) \quad (38)$$

where  $\boldsymbol{\rho}_{L1} = \mathbf{u}_{L1}^k - \mathbf{d}_{L1}^k$  and  $\boldsymbol{\rho}_{L2} = \mathbf{u}_{L2}^k - \mathbf{d}_{L2}^k$ .

- Update  $\mathbf{u}_L$ : Denoting  $\boldsymbol{\nu}_{L1} = \mathbf{H}\mathbf{W}\boldsymbol{\theta}^{k+1} + \mathbf{d}_{L1}^k$  and  $\boldsymbol{\nu}_{L2} = \boldsymbol{\theta}^{k+1} + \mathbf{d}_{L2}^k$ , the variables  $\mathbf{u}_{L1}^{k+1}$  and  $\mathbf{u}_{L2}^{k+1}$  are updated as following

$$\begin{aligned} \mathbf{u}_{L1}^{k+1} &= \operatorname{argmin}_{\mathbf{u}_{L1}} \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{u}_{L1}\|_2^2 + \frac{\mu}{2} \|\boldsymbol{\nu}_{L1} - \mathbf{u}_{L1}\|_2^2 \\ &= (\mathbf{S}^H \mathbf{S} + \mu \mathbf{I}_{N_h})^{-1} (\mathbf{S}^H \mathbf{y} + \mu \boldsymbol{\nu}_{L1}) \end{aligned}$$

$$\mathbf{u}_{L2}^{k+1} = \operatorname{argmin}_{\mathbf{u}_{L2}} \tau \|\mathbf{u}_{L2}\|_1 + \frac{\mu}{2} \|\boldsymbol{\nu}_{L2} - \mathbf{u}_{L2}\|_2^2$$

where the MAP estimator of  $\mathbf{u}_{L2}$  can be calculated by the following soft-thresholding operator [51]:

$$\operatorname{prox}_{|\cdot|}(\nu) = \max\{0, |\nu| - \tau/\mu\} \quad (39)$$

where  $\nu$  is an element from the vector  $\boldsymbol{\nu}_{L2}$ .

2) *Fast ADMM  $\ell_1$ -norm regularization*: With the proposed fast SR scheme, we transform (33) into the following constrained optimization problem

$$\begin{aligned} \min_{\boldsymbol{\theta}, \mathbf{u}_{LF}} \quad & \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{H}\mathbf{W}\boldsymbol{\theta}\|_2^2 + \tau \|\mathbf{u}_{LF}\|_1 \\ \text{subject to} \quad & \boldsymbol{\theta} = \mathbf{u}_{LF}. \end{aligned} \quad (40)$$

The augmented Lagrange function corresponding to (40) is

$$L(\boldsymbol{\theta}, \mathbf{u}_{LF}, \boldsymbol{\lambda}_{LF}) = \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{H}\mathbf{W}\boldsymbol{\theta}\|_2^2 + \tau \|\mathbf{u}_{LF}\|_1 + \boldsymbol{\lambda}_{LF}^T (\boldsymbol{\theta} - \mathbf{u}_{LF}) + \frac{\mu}{2} \|\boldsymbol{\theta} - \mathbf{u}_{LF}\|_2^2 \quad (41)$$

or equivalently

$$L(\boldsymbol{\theta}, \mathbf{u}_{\text{LF}}, \mathbf{d}_{\text{LF}}) = \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{H}\mathbf{W}\boldsymbol{\theta}\|_2^2 + \tau \|\mathbf{u}_{\text{LF}}\|_1 + \frac{\mu}{2} \|\boldsymbol{\theta} - \mathbf{u}_{\text{LF}} + \mathbf{d}_{\text{LF}}\|_2^2. \quad (42)$$

Similar to the TV regularization, the proposed fast SR algorithm for  $\ell_1$ -norm regularization uses the analytical solution of Section III resulting into the following three steps

$$\begin{aligned} & \text{For } k = 0, \dots \\ & \left[ \begin{array}{l} \boldsymbol{\theta}^{k+1} \in \operatorname{argmin}_{\boldsymbol{\theta}} \frac{1}{2} \|\mathbf{y} - \mathbf{S}\mathbf{H}\mathbf{W}\boldsymbol{\theta}\|_2^2 + \frac{\mu}{2} \|\boldsymbol{\theta} - \mathbf{u}_{\text{LF}}^k + \mathbf{d}_{\text{LF}}^k\|_2^2 \\ \mathbf{u}_{\text{LF}}^{k+1} \in \operatorname{argmin}_{\mathbf{u}_{\text{LF}}} \tau \|\mathbf{u}_{\text{LF}}\|_1 + \frac{\mu}{2} \|\boldsymbol{\theta}^{k+1} - \mathbf{u}_{\text{LF}} + \mathbf{d}_{\text{LF}}^k\|_2^2 \\ \mathbf{d}_{\text{LF}}^{k+1} = \mathbf{d}_{\text{LF}}^k + (\boldsymbol{\theta}^{k+1} - \mathbf{u}_{\text{LF}}^{k+1}) \end{array} \right. \end{aligned} \quad (43)$$

The estimation of  $\boldsymbol{\theta}$  consists of an SR problem with Gaussian regularization and can thus be calculated analytically following Section III

$$\begin{aligned} \boldsymbol{\theta}^{k+1} &= (\mathbf{W}^H \mathbf{H}^H \mathbf{S}^H \mathbf{S} \mathbf{H} \mathbf{W} + \mu \mathbf{I}_M)^{-1} \mathbf{r} \\ &= \mathbf{W}^H \mathbf{F}^H (\boldsymbol{\Lambda}^H \mathbf{F} \mathbf{S} \mathbf{F}^H \boldsymbol{\Lambda} + \mu \mathbf{I}_{N_h})^{-1} \mathbf{F} \mathbf{W} \mathbf{r} \\ &= \mathbf{W}^H \mathbf{F}^H \left( \frac{1}{d} \boldsymbol{\Lambda}^H \boldsymbol{\Lambda} + \mu \mathbf{I}_{N_h} \right)^{-1} \mathbf{F} \mathbf{W} \mathbf{r} \\ &= \mathbf{W}^H \mathbf{F}^H \left( \frac{1}{\mu} \mathbf{I}_{N_h} - \frac{1}{\mu} \boldsymbol{\Lambda}^H \left( d \mathbf{I}_{N_l} + \boldsymbol{\Lambda} \boldsymbol{\Lambda}^H \frac{1}{\mu} \right)^{-1} \boldsymbol{\Lambda} \frac{1}{\mu} \right) \mathbf{F} \mathbf{W} \mathbf{r} \\ &= \frac{1}{\mu} \mathbf{r} - \frac{1}{\mu} \mathbf{W}^H \mathbf{F}^H \boldsymbol{\Lambda}^H (\mu d \mathbf{I}_{N_l} + \boldsymbol{\Lambda} \boldsymbol{\Lambda}^H)^{-1} \boldsymbol{\Lambda} \mathbf{F} \mathbf{W} \mathbf{r} \end{aligned} \quad (44)$$

where  $\mathbf{r} = \mathbf{W}^H \mathbf{H}^H \mathbf{S}^H \mathbf{y} + \mu (\mathbf{u}_{\text{LF}} - \mathbf{d}_{\text{LF}})$ . The estimation of  $\mathbf{u}_{\text{LF}}$  is conducted with the soft thresholding operator (39), since the corresponding optimization problem is pixelwise decoupled.

The resulting algorithm for Laplace regularization in the wavelet domain is summarized in Algo. 3.

## V. EXPERIMENTAL RESULTS

In this section, we compare the proposed fast SR algorithm based on ADMM and the traditional ADMM, on several images and using the three regularization terms considered in the previous sections. Our main objective is to analyze the reduced computational efficiency of the proposed technique. All the experiments were performed using MATLAB 2013A on a computer with Windows 7, Intel(R) Core(TM) i7-4770 CPU @3.40GHz and 8 GB RAM.

<sup>2</sup> $|\boldsymbol{\nu}| \triangleq [|\nu_1|, \dots, |\nu_M|]^T \in \mathbb{R}^{M \times 1}$ .

---

**Algorithm 3:** FSR-ADMM with  $\ell_1$ -norm regularization in wavelet domain
 

---

- 1 Set  $k = 0$ , choose  $\mu > 0$ ,  $\mathbf{x}^0$ ,  $\mathbf{d}^0$ , input ;
  - 2 Repeat
    - // Update  $\mathbf{x}$  using the analytic solution in (32)
  - 3  $\mathbf{r} = \mathbf{W}^H \mathbf{H}^H \mathbf{S}^H \mathbf{y} + \mu(\mathbf{u}_{\text{LF}} - \mathbf{d}_{\text{LF}})$ ;
  - 4  $\Lambda \leftarrow \text{Dec}(\mathbf{H})$ ;
  - 5  $\underline{\Lambda} = \Lambda \Lambda^H$ ;
  - 6  $\mathbf{x}_f \leftarrow \left( \underline{\Lambda}^H (\mu \mathbf{d} \mathbf{I}_{N_l} + \underline{\Lambda} \mathbf{P} \underline{\Lambda}^H)^{-1} \underline{\Lambda} \right) \mathbf{F} \mathbf{W} \mathbf{r}$  ;
  - 7  $\hat{\mathbf{x}} \leftarrow \frac{1}{\mu} \mathbf{r} - \frac{1}{\mu} \mathbf{W}^H \mathbf{F}^H \mathbf{x}_f$  ;
    - // Update  $\mathbf{u}$  using the soft-thresholding operator
  - 8  $\boldsymbol{\nu} = \boldsymbol{\theta}^{k+1} + \mathbf{d}_{\text{LF}}^k$  ;
  - 9  $\mathbf{u}_{\text{TVF}}^{k+1} = \max\{0, |\boldsymbol{\nu}| - \tau/\mu\}^2$ ;
  - // Update the dual variables  $\mathbf{d}$
  - 10  $\mathbf{d}_{\text{LF}}^{k+1} = \mathbf{d}_{\text{LF}}^k + (\boldsymbol{\theta}^{k+1} - \mathbf{u}_{\text{LF}}^{k+1})$ ;
  - 11  $k \leftarrow k + 1$
  - 12 until stopping criterion is satisfied.
- 

The MATLAB codes and all the simulation results are available in the first author's homepage

<sup>3</sup>.

#### A. Quality assessment metrics

The performances of the different SR algorithms are evaluated in terms of the following metrics: improved signal-to-noise ratio (ISNR), normalized root mean square error (NRMSE) and mean structural similarity (MSSIM). The definitions of these metrics are given below

<sup>3</sup><http://zhao.perso.enseeiht.fr/>

$$\text{ISNR} = 10 \log_{10} \frac{\|\mathbf{x} - \mathbf{y}\|^2}{\|\mathbf{x} - \hat{\mathbf{x}}\|^2} \quad (45)$$

$$\text{NRMSE} = \sqrt{\frac{\|\mathbf{x} - \hat{\mathbf{x}}\|^2}{\|\mathbf{x}\|^2}} \quad (46)$$

$$\text{MSSIM} = \frac{1}{M} \sum_{j=1}^M \text{SSIM}(\mathbf{x}_j, \hat{\mathbf{x}}_j) \quad (47)$$

where the vectors  $\mathbf{x}, \mathbf{y}, \hat{\mathbf{x}}$  are the ground truth (reference image/HR image), bicubic interpolation of the observation and the restored SR image respectively. Note that MSSIM is implemented blockwise, with  $M$  the number of local windows,  $\mathbf{x}_j$  and  $\hat{\mathbf{x}}_j$  local regions extracted from  $\mathbf{x}$  and  $\hat{\mathbf{x}}$  and SSIM the structural similarity measure of each window (defined in [53]).

### B. Example 1: Results with Gaussian regularization

We first consider the proposed SR approach with a Gaussian prior, leading to the  $\ell_2$ -norm/Tikhonov regularization (3) for the ‘‘pepper’’ image of size  $134 \times 198$ . The HR image (shown in Fig. 2(e)) was blurred by a Gaussian filter (whose variance was set to  $\sigma_h^2 = 3$ ) of size  $7 \times 7$  and downsampled by a factor of 2 in each spatial direction ( $d_r = d_c = 2$ ). The resulting blurred and downsampled image was contaminated by an additive white Gaussian noise (AWGN) corresponding to a blurred-signal-to-noise ratio (BSNR) = 40 dB defined as below

$$\text{BSNR} = 10 \log_{10} \left( \frac{\|\mathbf{SHx} - E(\mathbf{SHx})\|_2^2}{N\sigma_n^2} \right) \quad (48)$$

where  $N$  is the total number of pixels of the observed image, shown in Fig. 2(a) and  $E(\cdot)$  is the arithmetic mean operator.

With the formulated problem (3), the quality of the SR is highly influenced by the choice of  $\bar{\mathbf{x}}$ , which is the rough approximation of the HR image to reconstruct. In the experiments considered in this part, we have used two different ways to generate  $\bar{\mathbf{x}}$ .

In the first case,  $\bar{\mathbf{x}}$  (denoted as  $\bar{\mathbf{x}}_{\text{obs}}$ ) is obtained by bicubic interpolation of the observed image, shown in Fig. 2(b). In the second case,  $\bar{\mathbf{x}}$  is obtained by a more accurate approximation of the HR image computed by bicubic interpolation of the decimated reference image, without the influence of the blurring operator. We denote this approximated HR image by ‘‘ $\bar{\mathbf{x}}_{\text{ref}}$ ’’, as shown in Fig. 2(f). In each case, the regularization parameter  $\tau$  was manually fixed to its

optimal value by cross-validation:  $\tau = 2 \times 10^{-3}$  in the first case and  $\tau = 8 \times 10^{-3}$  in the second case. Note that the same values of  $\tau$  have been employed for the “direct” ADMM and for our implementation based on the analytical solution (23) detailed in Section III. With the “direct” ADMM, two other parameters have to be tuned: the stepsize that was set to  $\mu = 0.05$  and the tolerance of the stopping criterion that was set to  $10^{-4}$ . Again these two values were obtained by cross-validation. In this paper, the stopping criterion is chosen as the relative cost function error, i.e.,

$$\frac{|f(\mathbf{x}^{k+1}) - f(\mathbf{x}^k)|}{f(\mathbf{x}^k)} \quad (49)$$

where  $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{y} - \mathbf{S}\mathbf{H}\mathbf{x}\|_2^2 + \tau\|\mathbf{x} - \bar{\mathbf{x}}\|_2^2$ . Note that other ways of defining the stopping criterion such as those studied in [30] could be also investigated.

The restored images obtained with the proposed solution (referred to as analytical solution (AS)) are displayed in Figs. 2(c) (case 1) and 2(g) (case 2) corresponding to the two different values of  $\bar{\mathbf{x}}$ . The results obtained with the “direct” ADMM are shown in Figs. 2(d) (case 1) and 2(h) (case 2) for comparison. The visual impression, confirmed by the numerical results of Tab. I, show that the HR image reconstructions obtained with our method are similar to those obtained with the “direct” ADMM. However, the reconstruction process with the proposed analytical solution is much faster than the iterative “direct” ADMM. Indeed, the computational time with our method is divided by a factor of 60 for the first value of  $\bar{\mathbf{x}}$  and by a factor of 80 for its second value.

Note that the restored images obtained with  $\bar{\mathbf{x}}_{\text{ref}}$  (i.e., case 2) are visually much better than the ones obtained with  $\bar{\mathbf{x}}_{\text{obs}}$  (i.e., case 1), especially considering the artifacts near the object boundaries. These artifacts visible in case 1 are mainly due to the blurring operation, stemming from  $\bar{\mathbf{x}}_{\text{obs}}$ . This result was expected since “ $\bar{\mathbf{x}}_{\text{ref}}$ ” is a more accurate approximation of the HR image than “ $\bar{\mathbf{x}}_{\text{obs}}$ ”.

### C. Example 2: fast ADMM-based SR scheme

In this second group of experiments, we consider the non Gaussian priors discussed in Section IV, namely the TV regularization in the spatial domain and the Laplace/ $\ell_1$ -norm regularization in the wavelet domain. The results obtained with our fast ADMM implementation (referred to as FSR-ADMM) are compared with those obtained by a standard ADMM algorithm (referred to as “direct” ADMM).

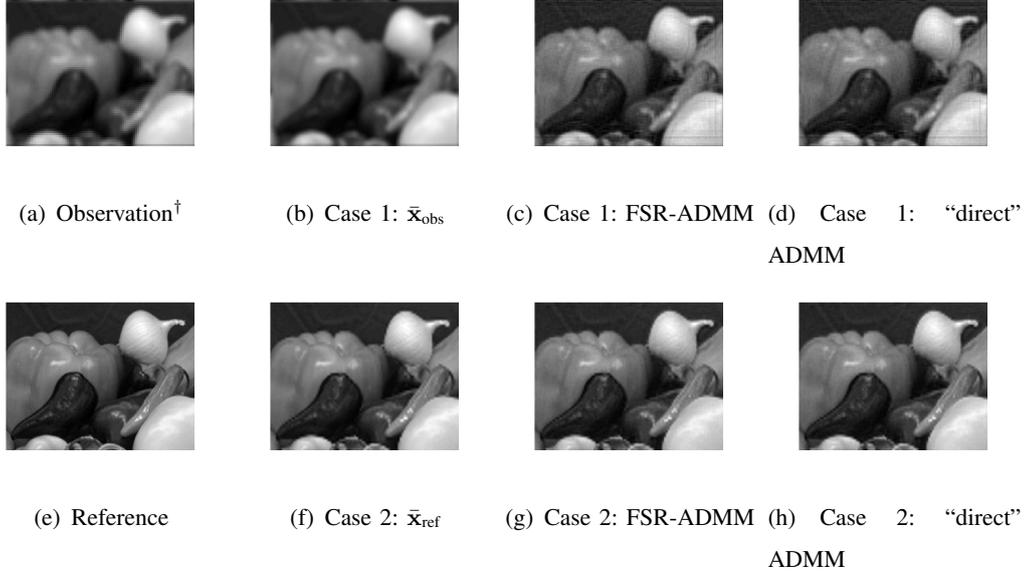


Fig. 2. Super-resolution results with up-sampling factor of 2 for the image “pepper”.

†Note that the LR images (observations) have been scaled for better visualization in this paper (i.e., the LR images contain  $d$  times fewer pixels than the corresponding HR images).

TABLE I  
EX. 1: QUALITY ASSESSMENT FOR THE IMAGE “PEPPER”

Method	ISNR(dB)	NRMSE	MSSIM	CPU(s)
Case 1				
AS	3.9311	0.0955	0.4986	<b>0.0041</b>
“direct”	3.9346	0.0954	0.4995	0.2647
Case 2				
AS	6.4798	0.05988	0.6998	<b>0.0039</b>
“direct”	6.4784	0.05989	0.7004	0.3085

The images “Lena”, “monarch” and “Barbara” of size  $512 \times 512$  pixels were considered in these experiments. These reference or HR images were blurred by a Gaussian filter of size  $9 \times 9$  (whose variance was set to  $\sigma_h^2 = 3$ ), downsampled by a decimation factor of  $4 \times 4$  (i.e.  $d_r = d_c = 4$ ) and contaminated by an AWGN corresponding to  $\text{BSNR} = 40$  dB. The observations (LR images) and the ground truths/original images are shown in Fig. 3 (first two columns) for illustration.

1) *TV regularization in the spatial domain:* The regularization parameter was manually fixed to  $2 \times 10^{-3}$  for the image “Lena”, to  $1.8 \times 10^{-3}$  for the image “monarch” and to  $2.5 \times 10^{-3}$  for the image “Barbara” in this Section.

Fig. 3 shows the SR results obtained using the bicubic interpolation technique (3rd column), the proposed FSR-ADMM (4th column) and the “direct” ADMM (last column). As expected, the ADMM reconstructions are more accurate than a simple interpolation of the LR image that is not able to solve the deblurring problem. Note that the results with FSR-ADMM and with “direct” ADMM are visually similar. This visual inspection is confirmed by the quantitative results provided in Tab. II. We emphasize that the proposed FSR-ADMM outperforms the “direct” ADMM implementation in terms of reconstruction times. Indeed, a speed-up of at least a factor of 2 is reported in Tab. II. Moreover, the plots in Fig. 4 illustrate the convergence of the two ADMM implementations. The proposed fast reconstruction method (FSR-ADMM) converges faster and with less fluctuations than the “direct” ADMM. The fluctuations of the convergence curves obtained with the “direct” ADMM are caused by the fact that there are more additional variables to solve the optimization problem when compared with the proposed FSR-ADMM. This fact is also responsible of the larger number of iterations acquired to reach convergence with the “direct” ADMM.

2)  *$\ell_1$ -norm regularization in the wavelet domain:* In this section, we evaluate the two ADMM implementations using an  $\ell_1$ -norm regularization in the wavelet domain (taking advantage of the sparsity of the wavelet coefficients). All experiments were conducted using the discrete Haar wavelet transform and the Rice wavelet toolbox [54]. The regularization parameter for both ADMM implementations was fixed to  $2 \times 10^{-4}$  for the “Lena” image, to  $1.8 \times 10^{-4}$  for the “monarch” image and to  $2.5 \times 10^{-4}$  for the image “Barbara”.

Fig. 5 shows the SR reconstruction results with an  $\ell_1$ -norm minimization in the wavelet domain. The HR images estimated with the FSR-ADMM and the “direct” ADMM are visually similar and better than a simple interpolation. The numerical results in Tab. III confirm that the two ADMM implementations present similar performance. As in the previous case, the FSR-ADMM outperforms the “direct” ADMM in terms of CPU time. The faster and smoother convergence obtained with our method is also illustrated by the plots in Fig. 6. Note that the fluctuations in the evolution of the objective function and the ISNR during the iterations obtained with the “direct” ADMM is due to the variable splitting, which requires more variables and constraints for the standard ADMM when compared to the proposed method.

TABLE II  
EX. 2: SR QUALITY ASSESSMENT WITH TV REGULARIZATION

Image	Method	ISNR (dB)	NRMSE ( $\times 10^{-2}$ )	MSSIM	Time (sec.)	Iter.
Lena	FSR-ADMM	5.2874	6.14	0.4069	<b>2.0593</b>	<b>19</b>
	“direct” ADMM	5.2929	6.14	0.4138	7.7142	125
Monarch	FSR-ADMM	6.4343	7.16	0.4165	<b>3.1090</b>	<b>28</b>
	“direct” ADMM	6.4323	7.17	0.4836	9.1289	150
Barbara	FSR-ADMM	2.0980	12.64	0.3599	<b>2.5550</b>	<b>23</b>
	“direct” ADMM	2.0979	12.65	0.3605	9.0444	150



Fig. 3. Super-resolution results with TV regularization. First column: observed LR images, second column: original HR images, third column: results using bicubic interpolation, fourth column: results using our fast ADMM-based SR scheme, fifth column: results using standard ADMM scheme.

## VI. CONCLUSION

This paper has studied a new fast single image super-resolution approach based on the widely used image formation model. The proposed super-resolution approach computes the image maximum a posterior estimator efficiently by exploiting the properties of the decimation and the blurring operators. For the Gaussian prior, computing the maximum *a posteriori* of the target image can be solved analytically, getting rid of any iterative steps. For the non-Gaussian priors, variable splittings has allowing this analytical solution to be embedded into

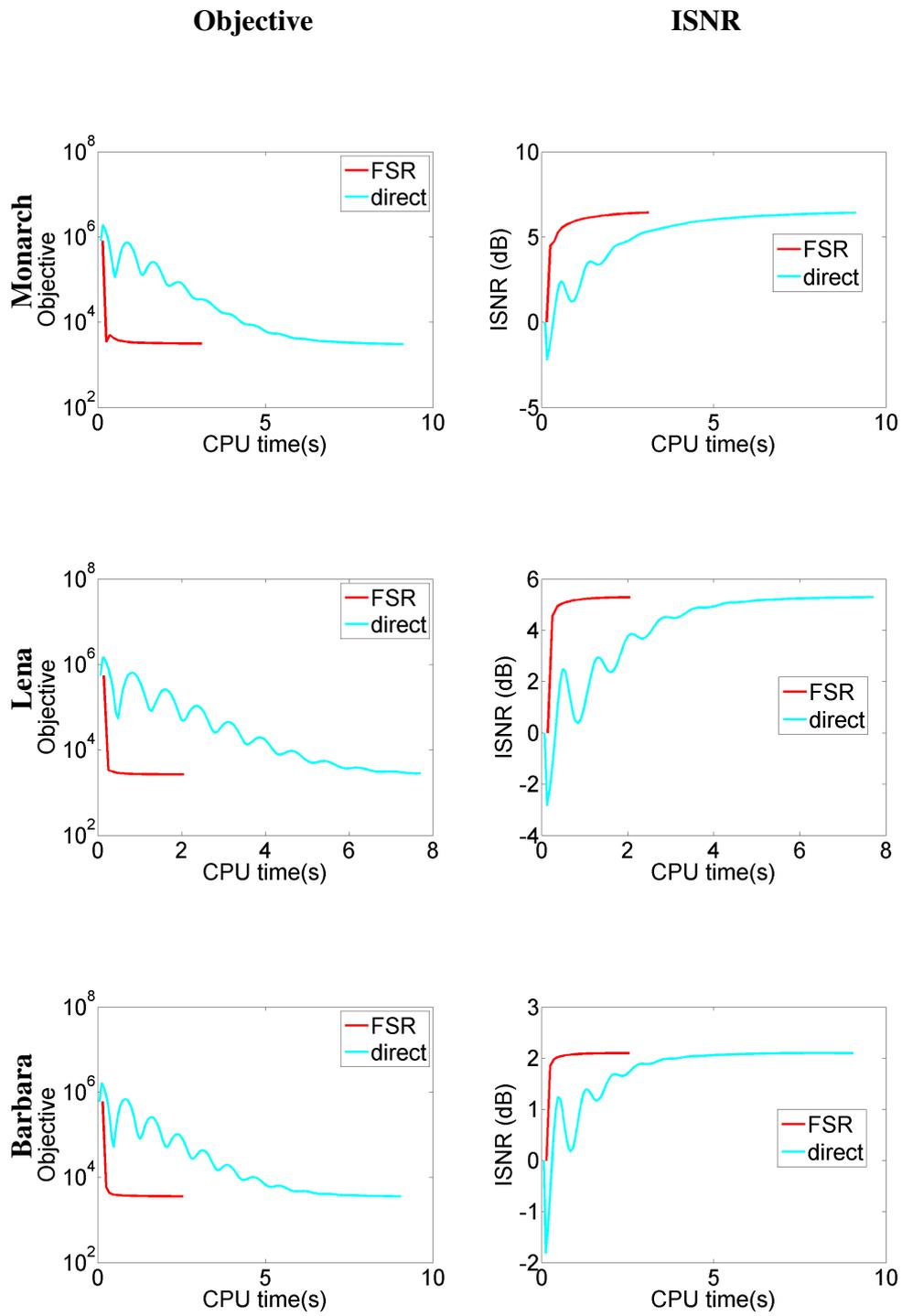


Fig. 4. TV regularization: evolution of the objective function (left) and the ISNR (right) over time.

TABLE III  
EX. 2: SR QUALITY ASSESSMENT WITH  $\ell_1$ -NORM REGULARIZATION

Image	Method	ISNR (dB)	NRMSE ( $\times 10^{-2}$ )	MSSIM	Time (sec.)	Iter.
Lena	FSR-ADMM	4.8977	6.42	0.3903	<b>5.3092</b>	<b>28</b>
	“direct” ADMM	4.8584	6.45	0.3912	32.7638	176
monarch	FSR-ADMM	5.2197	8.24	0.3859	<b>4.4268</b>	<b>23</b>
	“direct” ADMM	5.1727	8.28	0.4253	32.2093	200
Barbara	FSR-ADMM	2.0093	12.77	0.3401	<b>5.9850</b>	<b>32</b>
	“direct” ADMM	1.9904	12.80	0.3398	36.7896	200

the ADMM framework, thus accelerating the standard ADMM scheme. Results on several natural images have confirmed the computational efficiency of our approach and have shown fast and smooth convergence. As a perspective of this work, an interesting research track consists of extending the proposed method to some online applications such as video super-resolution and medical imaging.

## REFERENCES

- [1] S. C. Park, M. K. Park, and M. G. Kang, “Super-resolution image reconstruction: a technical overview,” *IEEE Signal Process. Mag.*, vol. 20, pp. 21–36, 2003.
- [2] G. Martin and J. M. Bioucas-Dias, “Hyperspectral compressive acquisition in the spatial domain via blind factorization,” in *Proc. IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, June 2015.
- [3] J. Yang and T. Huang, *Super-resolution imaging*. Boca Raton, FL, USA: CRC Press, 2010, ch. Image super-resolution: Historical overview and future challenges, pp. 20–34.
- [4] T. Akgun, Y. Altunbasak, and R. M. Mersereau, “Super-resolution reconstruction of hyperspectral images,” *IEEE Trans. Image Process.*, vol. 14, no. 11, pp. 1860–1875, 2005.
- [5] R. Morin, A. Basarab, and D. Kouame, “Alternating direction method of multipliers framework for super-resolution in ultrasound imaging,” in *Proc. IEEE International Symposium on Biomedical Imaging (ISBI)*, May 2012, pp. 1595–1598.
- [6] J. Yang, J. Wright, T. S. Huang, and Y. Ma, “Image super-resolution via sparse representation,” *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, 2010.
- [7] J. Sun, J. Sun, Z. Xu, and H.-Y. Shum, “Image super-resolution using gradient profile prior,” in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008, pp. 1–8.
- [8] Y.-W. Tai, S. Liu, M. S. Brown, and S. Lin, “Super resolution using edge prior and single image detail synthesis,” in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 2400 – 2407.
- [9] P. Thévenaz, T. Blu, and M. Unser, “Handbook of medical imaging,” I. N. Bankman, Ed. Orlando, FL, USA: Academic Press, Inc., 2000, ch. Image Interpolation and Resampling, pp. 393–420.



Fig. 5. Super-resolution results with  $\ell_1$ -norm regularization in the wavelet domain. First column: results using our fast ADMM-based SR scheme, second column: results using standard ADMM scheme.

- [10] X. Zhang and X. Wu, “Image interpolation by adaptive 2-D autoregressive modeling and soft-decision estimation,” *IEEE Trans. Image Process.*, vol. 17, no. 6, pp. 887–896, 2008.
- [11] S. Mallat and G. Yu, “Super-resolution with sparse mixing estimators,” *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2889–2900, 2010.
- [12] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, “Learning low-level vision,” *Int. J. Comput. Vision*, vol. 40, no. 1, pp. 25–47, Oct. 2000.
- [13] D. Glasner, S. Bagon, and M. Irani, “Super-resolution from a single image,” in *Proc. IEEE Int. Conf. Comp. Vision (ICCV)*, 2009, pp. 349 – 356.
- [14] J.-B. Huang, A. Singh, and N. Ahuja, “Single image super-resolution from transformed self-exemplars,” in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

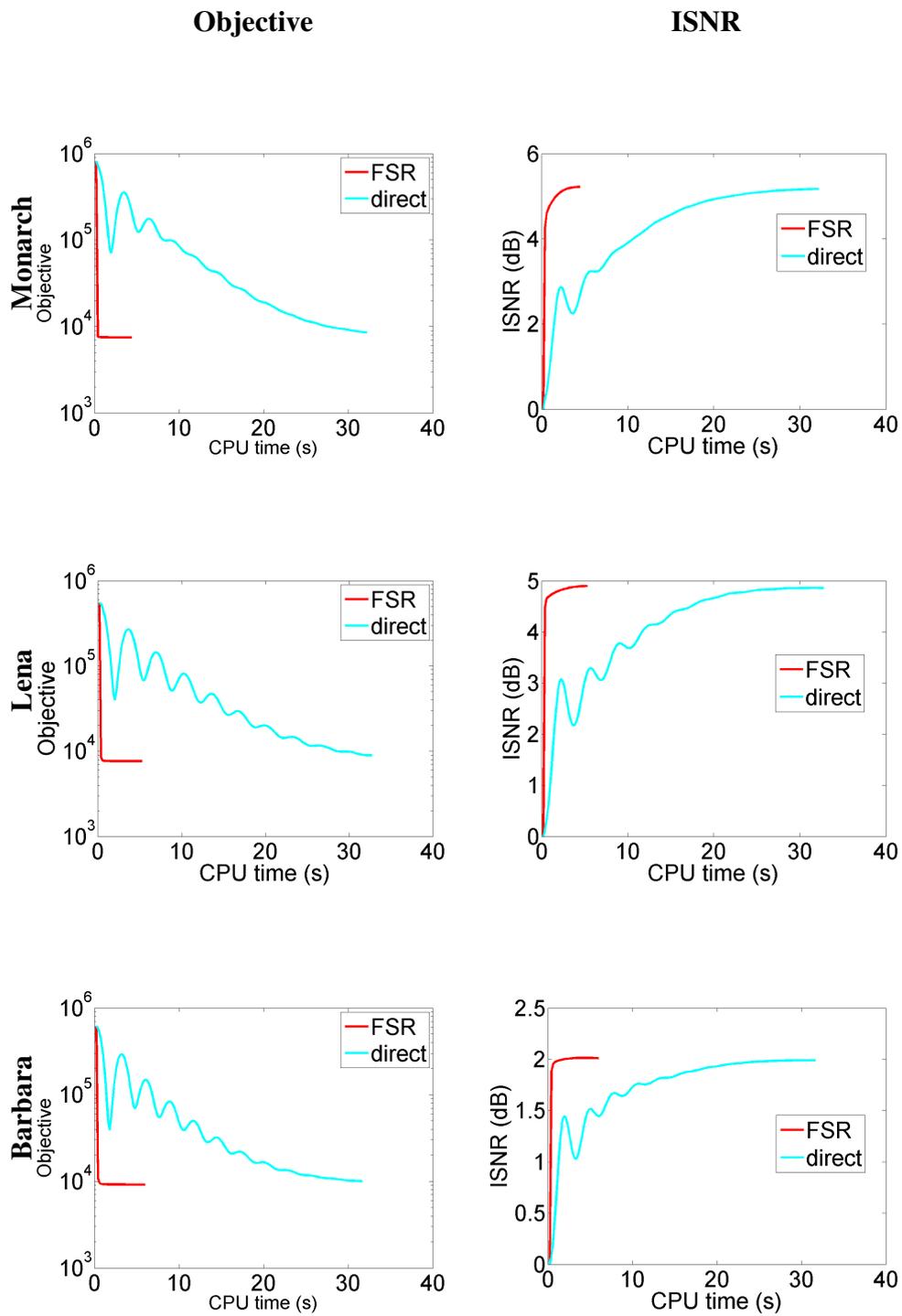


Fig. 6.  $\ell_1$ -norm regularization: evolution of the objective function (left) and the ISNR (right) over time.

- [15] R. Zeyde, M. Elad, and M. Protter, “On single image scale-up using sparse-representations,” in *Curves and Surfaces*, ser. Lecture Notes in Computer Science, J.-D. Boissonnat, P. Chenin, A. Cohen, C. Gout, T. Lyche, M.-L. Mazure, and L. Schumaker, Eds. Springer Berlin Heidelberg, 2012, vol. 6920, pp. 711–730.
- [16] J. Sun, J. Sun, Z. Xu, and H.-Y. Shum, “Gradient profile prior and its applications in image super-resolution and enhancement,” *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1529 – 1542, 2011.
- [17] M. K. Ng, P. Weiss, and X. Yuan, “Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods,” *SIAM J. Sci. Comput.*, vol. 32, pp. 2710–2736, 2010.
- [18] N. Nguyen, P. Milanfar, and G. Golub, “A computationally efficient superresolution image reconstruction algorithm,” *IEEE Trans. Image Process.*, vol. 10, no. 4, pp. 573–583, 2001.
- [19] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, “Fast fusion of multi-band images based on solving a Sylvester equation,” *IEEE Trans. Image Process.*, 2015, to appear.
- [20] M. Ebrahimi and E. R. Vrscay, “Regularization schemes involving self-similarity in imaging inverse problems,” in *Proc. 4 th AIP international Conference and the 1st Congress of the IPIA*, 2008.
- [21] H. A. Aly and E. Dubois, “Image up-sampling using total-variation regularization with a new observation model,” *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1647–1659, 2005.
- [22] A. Marquina and S. Osher, “Image super-resolution by TV-regularization and Bregman iteration,” *J. Sci. Comput.*, vol. 37, no. 3, pp. 367–382, 2008.
- [23] J. M. Bioucas-Dias, “Bayesian wavelet-based image deconvolution: A GEM algorithm exploiting a class o heavy-tailed priors,” *IEEE Trans. Image Process.*, vol. 15, no. 4, pp. 937–951, 2006.
- [24] J. Ng, R. Prager, N. Kingsbury, G. Treece, and A. Gee, “Wavelet restoration of medical pulse-echo ultrasound images in an EM framework,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 54, no. 3, pp. 550–568, 2007.
- [25] C. V. Jiji, M. V. Joshi, and S. Chaudhuri, “Single-frame image super-resolution using learned wavelet coefficients,” *Int. J. Imaging Syst. Technol.*, vol. 14, pp. 105–112, 2004.
- [26] M. A. T. Figueiredo and R. D. Nowak, “An EM algorithm for wavelet-based image restoration,” *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 906–916, 2003.
- [27] S. Roth and M. J. Black, “Fields of experts: a framework for learning image priors,” in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005, pp. 860–867.
- [28] R. Fattal, “Edge-avoiding wavelets and their applications,” *ACM Trans. Graph.*, vol. 28, no. 3, pp. 1–10, Aug. 2009.
- [29] J. R. Shewchuk, “An introduction to the conjugate gradient method without the agonizing pain,” School of Computer Science of Carnegie Mellon University, Tech. Rep., 1994.
- [30] 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.*, vol. 3, no. 1, pp. 1–122, 2011.
- [31] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, “Fast image recovery using variable splitting and constrained optimization,” *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2345–2356, 2010.
- [32] —, “An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems,” *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, 2011.
- [33] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*. Springer Science & Business Media, 1996, vol. 375.
- [34] C. P. Robert, *The Bayesian Choice: from Decision-Theoretic Motivations to Computational Implementation*, 2nd ed., ser. Springer Texts in Statistics. New York, NY, USA: Springer-Verlag, 2007.
- [35] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*, 3rd ed. Boca Raton, FL: CRC press, 2013.
- [36] A. Tikhonov and V. Arsenin, *Solutions of ill-posed problems*, ser. Scripta series in mathematics. Winston, 1977.

- [37] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, “Bayesian fusion of multi-band images,” *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 6, pp. 1–11, 2015.
- [38] A. K. Jain, *Fundamentals of Digital Image Processing*, ser. Information and System Sciences. Prentice Hall, 1989.
- [39] O. Féron, F. Orieux, and J.-F. Giovannelli, “Gradient scan Gibbs sampler: an efficient algorithm for high-dimensional Gaussian distributions,” 2015. [Online]. Available: <http://arxiv.org/abs/1509.03495>
- [40] F. Orieux, O. Féron, and J.-F. Giovannelli, “Sampling high-dimensional Gaussian distributions for general linear inverse problems,” *IEEE Signal Process. Lett.*, vol. 19, no. 5, pp. 251–254, 2012.
- [41] C. Gilavert, S. Moussaoui, and J. Idier, “Efficient Gaussian sampling for solving large-scale inverse problems using MCMC,” *IEEE Trans. Image Process.*, vol. 63, no. 1, pp. 70–80, 2015.
- [42] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [43] M. Elad and A. Feuer, “Restoration of a single superresolution image from several blurred, noisy and undersampled measured images,” *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1646–1658, 1997.
- [44] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, “Advances and challenges in super-resolution,” *Int. J. Imaging Syst. Technol.*, vol. 14, no. 2, pp. 47–57, 2004.
- [45] J. K. H. Ng, “Restoration of medical pulse-echo ultrasound images,” Ph.D. dissertation, Trinity College, University of Cambridge, 2006.
- [46] M. Elad and A. Feuer, “Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images,” *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1646–1658, 1997.
- [47] N. Zhao, A. Basarab, D. Kouamé, and J.-Y. Tourneret, “Joint segmentation and deconvolution of ultrasound images using a hierarchical Bayesian model based on generalized Gaussian priors,” 2015. [Online]. Available: <http://arxiv.org/abs/1412.2813>
- [48] W. W. Hager, “Updating the inverse of a matrix,” *SIAM Rev.*, pp. 221–239, 1989.
- [49] Y. Wang, J. Yang, W. Yin, and Y. Zhang, “A new alternating minimization algorithm for total variation image reconstruction,” *SIAM J. Imag. Sci.*, vol. 1, pp. 248–272, 2008.
- [50] H. Huang, A. G. Christodoulou, and W. Sun, “Super-resolution hyperspectral imaging with unknown blurring by low-rank and group-sparse modeling,” in *Proc. IEEE Int. Conf. on Image Process (ICIP)*, 2014.
- [51] P. L. Combettes and J.-C. Pesquet, “Proximal splitting methods in signal processing,” in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, ser. Springer Optimization and Its Applications, H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Eds. Springer New York, 2011, pp. 185–212.
- [52] J. Bioucas-Dias, F. Condessa, and J. Kovacevic, “Alternating direction optimization for image segmentation using hidden Markov measure field models,” in *Proc. IS&T/SPIE Electronic Imaging*, vol. 9019. International Society for Optics and Photonics, 2014.
- [53] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.
- [54] R. Baraniuk, H. Choi, R. Neelamani, V. Ribeiro, J. Romberg, H. Guo, F. Fernandes, B. Hendricks, R. Gopinath, M. Lang, J. E. Odegard, D. Wei, and J. Jackson, “Rice wavelet toolbox.” [Online]. Available: <http://dsp.rice.edu/software/rice-wavelet-toolbox>