



**HAL**  
open science

## A convex approach to super-resolution and regularization of lines in images

Kévin Polisano, Laurent Condat, Marianne Clausel, Valérie Perrier

► **To cite this version:**

Kévin Polisano, Laurent Condat, Marianne Clausel, Valérie Perrier. A convex approach to super-resolution and regularization of lines in images. [Technical Report] Laboratoire Jean Kuntzmann (LJK). 2017. hal-01599010v2

**HAL Id: hal-01599010**

**<https://hal.science/hal-01599010v2>**

Submitted on 27 Apr 2018 (v2), last revised 19 Nov 2018 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A convex approach to super-resolution and regularization of lines in images\*

Kévin Polisano<sup>†</sup>    Laurent Condat<sup>‡</sup>    Marianne Clausel<sup>†</sup>  
Valérie Perrier<sup>†</sup>

## Abstract

We present a new convex formulation for the problem of recovering lines in degraded images. Following the recent paradigm of super-resolution, we formulate a dedicated atomic norm penalty and we solve this optimization problem by means of a primal-dual algorithm. This parsimonious model enables the reconstruction of lines from lowpass measurements, even in presence of a large amount of noise or blur. Furthermore, a Prony method performed on rows and columns of the restored image, provides a spectral estimation of the line parameters, with subpixel accuracy.

**Keywords.** Super-resolution, sparse recovery, convex optimization, line detection, splitting method, spectral estimation

**AMS subject classifications.** 94A08, 94A12, 94A20, 47N10, 90C22, 90C25

## 1 Introduction

Many restoration or reconstruction imaging problems are ill-posed and must be regularized. So, they can be formulated as convex optimization problems formed by the combination of a data fidelity term with a norm-based regularizer. Typically, given the data  $\mathbf{y} = \mathbf{A}\mathbf{x}^\dagger + \boldsymbol{\epsilon}$ , for some unknown image  $\mathbf{x}^\dagger$  to estimate, some known observation operator  $\mathbf{A}$  and some noise  $\boldsymbol{\epsilon}$ , one aims at solving a problem like

$$\text{Find } \tilde{\mathbf{x}} \in \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|^2 + \lambda R(\mathbf{x}),$$

where  $\lambda$  controls the tradeoff between data fidelity and regularization and  $R$  is a convex regularization functional.  $R$  can be chosen to promote some kind of smoothness. The classical Tikhonov regularizer  $R(\mathbf{x}) = \|\nabla \mathbf{x}\|_2^2$  generally makes the problem easy

---

\*Part of this work has appeared in preliminary form in the conference proceedings of the *24th European Signal Processing Conference (EUSIPCO)*, 2016 [41].

<sup>†</sup>Univ. Grenoble Alpes, CNRS, Grenoble INP, LJK, F-38000, Grenoble, France (kevin.polisano@univ-grenoble-alpes.fr).

<sup>‡</sup>Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, F-38000, Grenoble, France (laurent.condat@gipsa-lab.grenoble-inp.fr).

to solve, but yields over-smoothing of the textures and edges in the recovered image  $\tilde{\mathbf{x}}$ . A popular and better regularizer is the total variation  $R(\mathbf{x}) = \|\nabla \mathbf{x}\|_1$ , see e.g [11, 20]; it yields images with sharp edges, but the textures are still over-smoothed, there are staircasing effects and the pixel values tend to be clustered in piecewise constant areas. To overcome these drawbacks, one can penalize higher order derivatives [32, 6] or make use of non-local penalties [39, 18, 15]. Another approach, which is at the heart of the recent paradigm of sparse recovery [49, 27] and compressed sensing [25, 53], is to choose  $R$  to favor some notion of low complexity. Indeed, many phenomena, when observed by instruments, yield data living in high dimensional spaces, but inherently governed by a small number of degrees of freedom. One early choice was to set  $R$  as the  $\ell_1$  norm of wavelet coefficients of the image. But the signals encountered in applications like radar, array processing, communication, seismology, or remote sensing, are usually specified by parameters in a *continuous* domain, from which they depend nonlinearly. So, modern sampling theory has widened its scope to a broader class of signals, with so-called finite rate of innovation, i.e. ruled by parcimonious models [34, 26, 4, 57]. This encompasses reconstruction of pulses from lowpass measurements [23] and spectral estimation, which is the reconstruction of sinusoids from point samples [50, 51], with many applications [13, 31, 56, 7, 52, 33, 43, 48, 21]. The knowledge of the kind of elements we want to promote in the image makes it possible to estimate them from coarse-scale measurements, even with infinite precision if there is no noise. Methods achieving this goal are qualified as *super-resolution* methods, because they uncover fine scale information, which was lost in the data, beyond the Rayleigh or Nyquist resolution limit of the acquisition system [28, 8]. However, in this context, maximum likelihood estimation amounts to structured low rank approximation, which forms nonconvex and very difficult, even NP-hard in general, problems [35, 22]. An elegant and unifying formulation, which yields convex problems, is based on the *atomic norm* [2, 19]. We place ourselves in this general framework of atomic norm minimization: the sought-after image  $\mathbf{x}^\sharp$  is supposed to be a sparse positive combination of the elements of an infinite dictionary  $\mathcal{A}$ , indexed by continuously varying parameters. Then, one can choose  $R$  as the atomic norm  $\|\mathbf{x}\|_{\mathcal{A}}$  of the image  $\mathbf{x}$ , which can be viewed as the  $\ell_1$  norm of the coefficients, when the image is expressed in terms of the unit-norm elements of  $\mathcal{A}$ , called *atoms*:

$$\|\mathbf{x}\|_{\mathcal{A}} = \inf \{t > 0 : \mathbf{x} \in t \text{conv}(\mathcal{A})\} , \quad (1)$$

where  $\text{conv}(\mathcal{A})$  is the convex hull of the atoms. In this paper, we consider the setting, which is new to our knowledge, where the atoms are *lines*. Expressed in the Fourier domain, these atoms can be characterized with respect to their rows and columns, and the problem can be reduced to a dictionary of 1-D complex exponential samples, indexed by their frequency and phase, and the atomic norm can be computed via semidefinite programming [58]. This formulation enables us to derive a convex optimization problem under constraints, solved by mean of a splitting primal-dual algorithm [17]. Then, performing a Prony-like method [45] onto the solution of the algorithm allows us to extract the parameters of the lines. This approach provides a very high accuracy for the line estimation, where the Hough [29, 30, 36] and the Radon [44, 24, 37] transforms fail, due to their discrete nature. Our motivation stems from the frequent presence in biomedical images, e.g. in microscopy, of elongated structures like filaments, neurons, veins, which are deteriorated when reconstructed with classical penalties.

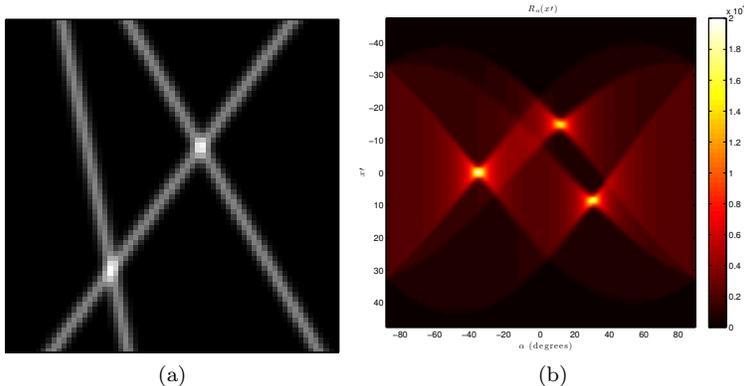


Figure 1: (a) The image  $\mathbf{b}^\sharp$  of three blurred lines with  $\kappa = 1$  and (b) the Radon transform of  $\mathbf{b}^\sharp$ .

The paper is organized as follows. The model is exposed in [section 2](#), the framework of atomic norm minimization underlying the super-resolution principle is introduced in [section 3](#), the algorithms we derive are in [section 4](#). Then a Prony-like method is developed in [section 5](#) as a way to perform spectral estimation of the line parameters. Finally some experimental results are shown in [section 6](#), and appendices give a comparison with the classic Hough and Radon procedure of lines detection. Part of this work has been published in the preliminary paper [\[41\]](#). In this paper, we present the mathematical developments that did not appear in the previous one, but also another algorithm, a new procedure estimation of the line parameters, an extension for the whole range of line angles with no more restriction, an application to inpainting problems and many other numerical experiments.

## 2 An image model of blurred lines

Our aim is to restore a blurred image  $\mathbf{b}^\sharp$  containing lines, and to estimate the parameters—angle, offset, amplitude—of the lines, given degraded data  $\mathbf{y}$ . In this section, we formulate what we precisely mean by an image containing lines. In short,  $\mathbf{b}^\sharp$  is a sum of perfect lines which have been blurred and then sampled. Both processes are detailed in the following.

### 2.1 The Ideal Continuous Model and the Objectives

We place ourselves in the quotient space  $\mathbb{P} = \mathbb{R}/(W\mathbb{Z}) \times \mathbb{R}$ , corresponding to the 2-D plane with horizontal  $W$ -periodicity, for some integer  $W \geq 1$ . To simplify the notations, we suppose that  $W$  is odd and we set  $M = (W - 1)/2$ .

A line of infinite length, with angle  $\theta \in (-\pi/2, \pi/2]$  with respect to verticality, amplitude  $\alpha > 0$ , and offset  $\eta \in \mathbb{R}$  from the origin on  $x$ -axis, is defined as the distribution

$$(t_1, t_2) \in \mathbb{P} \mapsto \alpha \delta(\cos \theta(t_1 - \eta) + \sin \theta t_2),$$

where  $\delta$  is the Dirac distribution. We define the distribution  $x^\sharp$ , which is a sum of

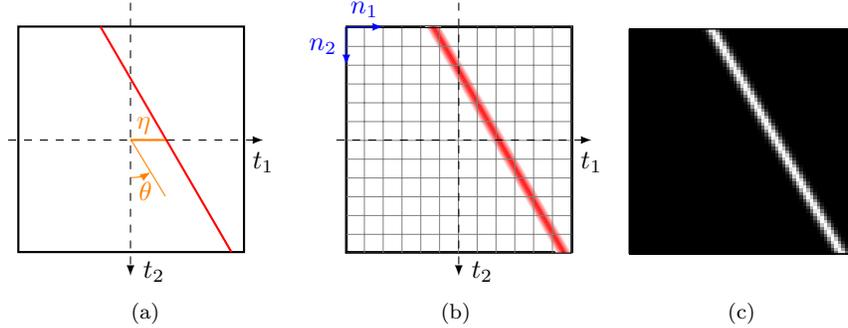


Figure 2: (a) Parameters  $(\theta, \eta)$  characterizing the position of a line in the 2-D plane, (b) the matrix convention we use to display the image obtained by applying the sampling operator with unit step  $\Delta$  onto the blurred line  $x^\sharp * \phi$ , and (c) the resulting discrete image  $\mathbf{b}^\sharp[n_1, n_2] = (x^\sharp * \phi)(n_1, n_2)$ .

$K$  different such perfect lines, for some integer  $K \geq 1$ , as

$$x^\sharp : (t_1, t_2) \in \mathbb{P} \mapsto \sum_{k=1}^K \alpha_k \delta(\cos \theta_k (t_1 - \eta_k) + \sin \theta_k t_2). \quad (2)$$

**Remark.** *Figure 2a* illustrates the line parameters and *Figure 2b* the convention we use for representing images. At this time, we suppose that the lines are rather vertical; that is, for every  $k = 1, \dots, K$ ,  $\theta_k \in (-\pi/4, \pi/4]$ . We will treat the general case in [subsection 4.3](#).

Since the ideal model  $x^\sharp$  is made up of Dirac distributions, the horizontal Fourier transform  $\hat{x}^\sharp = \mathcal{F}_1 x^\sharp$  is composed of a sum of exponentials. Our goal will be to reconstruct  $\hat{x}^\sharp$  by a super-resolution method, from its observations through a known degradation operator  $\mathbf{A}$  and some noise, which is an ill-posed problem. Then, spectral estimation of these exponentials will allow us to recover the line parameters. Let us first characterize the blur operator  $\mathbf{A}$ .

## 2.2 A Blur Model for an Exact Sampling Process

The observed image  $\mathbf{b}^\sharp$  of size  $W \times H$  in [Figure 2c](#) is obtained by convolution of the distribution  $x^\sharp$  with a blur function  $\phi$ , followed by a sampling with unit step denoted by the operator  $\Delta$ :

$$\mathbf{b}^\sharp[n_1, n_2] = (x^\sharp * \phi)(n_1, n_2), \quad \forall n_1 = 0, \dots, W-1, \quad n_2 = 0, \dots, H-1. \quad (3)$$

We also consider that the Point Spread Function (PSF)  $\phi$  is separable; that is the function  $x^\sharp * \phi$  can be obtained by a first horizontal convolution with  $\varphi_1$  and then a second vertical convolution with  $\varphi_2$ . Formally,  $x^\sharp * \phi = (x^\sharp * \phi_1) * \phi_2$  with  $\phi_1(t_1, t_2) = \varphi_1(t_1)\delta(t_2)$  and  $\phi_2(t_1, t_2) = \delta(t_1)\varphi_2(t_2)$ , where  $\varphi_1$  and  $\varphi_2$  are  $L^1$  functions.

So, after the first horizontal convolution, using the fact that  $\delta(at) = \delta(t)/|a|$  for any  $a \neq 0$ , we obtain the function:

$$\Psi = x^\sharp * \phi_1 : (t_1, t_2) \mapsto \sum_{k=1}^K \frac{\alpha_k}{\cos \theta_k} \varphi_1(t_1 + \tan \theta_k t_2 - \eta_k). \quad (4)$$

We can show that, after the second vertical convolution, we get the function

$$x^\sharp * \phi : (t_1, t_2) \in \mathbb{P} \mapsto \sum_{k=1}^K \alpha_k \psi_k (\cos \theta_k (t_1 - \eta_k) + \sin \theta_k t_2) , \quad (5)$$

where

$$\psi_k = \left( \frac{1}{\cos \theta_k} \varphi_1 \left( \frac{\cdot}{\cos \theta_k} \right) \right) * \left( \frac{1}{\sin \theta_k} \varphi_2 \left( \frac{\cdot}{\sin \theta_k} \right) \right) , \quad (6)$$

if  $\theta_k \neq 0$  and  $\psi_k = \varphi_1$  else.

**Remark.** We can notice that (5) can also be interpreted as follows: every line has undergone a 1-D convolution with  $\psi_k$  in the direction transverse to it. We can also notice that if  $\varphi_1$  and  $\varphi_2$  are Gaussian functions and have same variance  $\kappa^2$ , it follows from (6) that  $\psi_k$  has variance  $\kappa^2 (\cos^2 \theta + \sin^2 \theta) = \kappa^2$  as well.

**Assumptions.** In order to avoid any approximation when passing from the continuous to the discrete formulation, we assume that  $\varphi_1$  and  $\varphi_2$  have the following properties:

- (i)  $\varphi_1 \in L^1(0, W)$  is  $W$ -periodic, bounded, such that  $\frac{1}{W} \int_0^W \varphi_1 = 1$ , and band-limited; that is, its Fourier coefficients

$$c_m(\varphi_1) = \frac{1}{W} \int_0^W \varphi_1(t_1) e^{-j2\pi m t_1 / W} dt_1 ,$$

are zero for every  $m \in \mathbb{Z}$  with  $|m| \geq (W + 1)/2 = M + 1$ . Then, the discrete filter

$$(\mathbf{g}[n] = \varphi_1(n))_{n \in \mathbb{Z}} , \quad (7)$$

has for discrete Fourier coefficients:

$$\widehat{\mathbf{g}}[m] = \frac{1}{W} \sum_{n=0}^{W-1} \mathbf{g}[n] e^{-j \frac{2\pi m n}{W}} = c_m(\varphi_1) .$$

- (ii)  $\varphi_2 \in L^1(\mathbb{R})$  is such that  $\int_{\mathbb{R}} \varphi_2 = 1$ , and denoting  $\text{sinc}(t_2) = \sin(\pi t_2)/(\pi t_2)$ , the discrete filter

$$(\mathbf{h}[n] = (\varphi_2 * \text{sinc})(n))_{n \in \mathbb{Z}} \quad (8)$$

has compact support of length  $2S + 1$ , for some  $S \in \mathbb{N}$ ; i.e.,

$$\mathbf{h}[n] = 0, \quad \text{if } |n| \geq S + 1 .$$

Let us deduce from these assumptions some other properties satisfied by  $\varphi_1$  and  $\varphi_2$ , and their associated discrete filters  $\mathbf{g}$  and  $\mathbf{h}$ . First, we have

$$\frac{1}{W} \int_0^W \varphi_1(t) dt = \frac{1}{W} \sum_{n=-M}^M \mathbf{g}[n] = 1 ,$$

that is, the filter  $\mathbf{g}$  is normalized as well. Moreover

$$\int_0^W \varphi_1(t)^2 dt = \sum_{n=-M}^M \mathbf{g}[n]^2 = \sum_{n=-M}^M \varphi_1(n)^2 ,$$

and by Parseval relation:

$$\sum_{m \in \mathbb{Z}} |c_m(\varphi_1)|^2 = \sum_{m=-M}^M |\widehat{\mathbf{g}}[m]|^2 = \frac{1}{W} \int_0^W \varphi_1(t)^2 dt.$$

Now, we describe the sampling process leading from continuous to discrete formulation, based on the following proposition:

**Proposition 1.** *It is equivalent to perform the vertical convolution of  $\Psi = x^\# * \phi_1$  with  $\varphi_2$ , with  $\varphi_2 * \text{sinc}$ , or with the Dirac comb  $\gamma : t_2 \mapsto \sum_{n=-S}^S \mathbf{h}[n] \delta(t_2 - n)$ , where  $\mathbf{h}[n] = (\varphi_2 * \text{sinc})(n)$ .*

*Proof.* For every  $k = 1, \dots, K$ , the assumption  $\theta_k \in (-\pi/4, \pi/4]$  yields  $|\tan \theta_k| \leq 1$ . So, the function  $\Psi = x^\# * \phi_1$  given in (4), as a function of  $t_2$  at fixed  $t_1$ , is bandlimited: for every  $t_1 \in [0, W)$ , the Fourier transform  $\mathcal{F}_2 \Psi : \omega_2 \mapsto \int_{\mathbb{R}} (x^\# * \phi_1)(t_1, t_2) e^{-j2\pi\omega_2 t_2} dt_2$ , which is a distribution (sum of  $K$  Dirac combs), is zero for every  $|\omega_2| \geq 1/2$ . Indeed, we have:

$$[\mathcal{F}_2 \Psi](\omega_2) = \sum_{k=1}^K \frac{\alpha_k}{\sin \theta_k} \widehat{\varphi}_1 \left( \frac{\omega_2}{\tan \theta_k} \right) \exp \left( j2\pi\omega_2 \frac{t_1 - \eta_k}{\tan \theta_k} \right).$$

Since  $|\tan \theta_k| \leq 1$ , we have  $|\omega_2 / \tan \theta_k| \geq |\omega_2|$ . The support of  $\widehat{\varphi}_1$  is included in  $[-1/2, 1/2]$  ( $c_m(\varphi_1) = 0$  for  $|m| \geq M + 1$  and  $M/W < 1/2$ ), as well as the support of  $\mathcal{F}_2 \Psi$  which is necessarily included in the support of  $\widehat{\varphi}_1$ . Then, we have the equivalence  $\mathcal{F}_2 \Psi = \mathcal{F}_2 \Psi \cdot \mathbf{1}_{[-1/2, 1/2]} \Leftrightarrow \Psi = \Psi * \text{sinc}$ , and furthermore  $\Psi * \varphi_2 = \Psi * (\varphi_2 * \text{sinc})$ . In the Fourier domain, the function  $h = \varphi_2 * \text{sinc}$  is bandlimited, so  $[\mathcal{F}_2 \Psi] \widehat{h} = [\mathcal{F}_2 \Psi] \widehat{h}_{\text{per}}$  where  $\widehat{h}_{\text{per}}$  corresponds to the periodization of the spectrum of  $\widehat{h}$  with period 1, which amounts to saying that  $\Psi * h = \Psi * (\sum_n h[n] \delta(\cdot - n))$ .  $\square$

**Remark.** *Assumption (ii) implies that the filter  $(\mathbf{h}[n])_n$  should have compact support, but we can notice that the function  $h = \varphi_2 * \text{sinc}$  does not have compact support, since it is bandlimited. This means that the continuous function  $h$  has to vanish at integer points  $t = n$  for  $|n| > S$ . Given such a compact filter  $(\mathbf{h}[n])_{n=-S}^S$ , the unique bandlimited function  $h$  satisfying these conditions is obtained by the Shannon interpolation formula:*

$$h(t) = \sum_{n=-S}^S \mathbf{h}[n] \text{sinc}(t - n).$$

*By unicity, we necessarily have  $\varphi_2 * \text{sinc} = h$ , and we can notice that there always exists a bandlimited solution  $\varphi_2$  of this equation, which is simply  $\varphi_2 = h$ . In practice, we can always approximate a PSF by a bandlimited function  $h$ , with  $2S + 1$  samples  $\mathbf{h}[n]$  of this PSF; that is why we argued that the compact support assumption is not restrictive.*

Now, to obtain the discrete image  $\mathbf{b}^\#$  of (3), let us first define  $\mathbf{u}^\#$  by sampling  $x^\# * \phi_1$  with unit step:

$$\mathbf{u}^\#[n_1, n_2] = (x^\# * \phi_1)(n_1, n_2), \quad \forall n_1 = 0, \dots, W - 1, \quad n_2 = -S, \dots, H - 1 + S. \quad (9)$$

With the above assumptions and [Proposition 1](#), we can express  $\mathbf{b}^\sharp$  from  $\mathbf{u}^\sharp$  using a discrete vertical convolution with the filter  $\mathbf{h}$ :

$$\mathbf{b}^\sharp[n_1, n_2] = \sum_{p=-S}^S \mathbf{u}^\sharp[n_1, n_2 - p] \mathbf{h}[p], \quad \forall n_1 = 0, \dots, W-1, \quad n_2 = 0, \dots, H-1. \quad (10)$$

Altogether, we completely and exactly characterized the sampling process, which involves a continuous blur  $\phi$ , using the two discrete and finite filters  $(\mathbf{g}[n])_{n=0}^{W-1}$  and  $(\mathbf{h}[n])_{n=-S}^S$ . We insist on the fact that no discrete approximation is made during this sampling process, due to the assumptions (i) and (ii). An example of three blurred lines is depicted in [Figure 1](#), with the normalized filter  $\mathbf{h}$  approximating a Gaussian function of standard deviation  $\kappa$ ; that is,  $\varphi_2 : t \mapsto (2\pi\kappa^2)^{-1/2} \exp(-t^2/(2\kappa^2))$ , on the compact set  $[-S, S]$  with  $S = \lceil 4\kappa \rceil - 1$ ; and the normalized filter  $\mathbf{g} = [\mathbf{0}_{M-S}, \mathbf{h}, \mathbf{0}_{M-S}]$  whose Discrete Fourier Transform (DFT) is an interpolation of  $\widehat{\mathbf{h}}$ , which approaches the continuous Fourier transform  $\widehat{\varphi}_2 : \nu \mapsto \exp(-2\pi^2\kappa^2\nu^2)$ . So,  $\|\widehat{\mathbf{g}}\|_\infty = \|\widehat{\mathbf{h}}\|_\infty = 1$ .

### 2.3 Toward an Inverse Problem in Fourier Domain

Let us further characterize the image  $\mathbf{b}^\sharp$  in Fourier domain. [Figure 3](#) explains our notations in more details and illustrates the relation between all continuous and discrete variables.

First, we consider the image  $\widehat{\mathbf{u}}^\sharp$  obtained by applying the 1-D DFT on every row of  $\mathbf{u}^\sharp$  (9):

$$\widehat{\mathbf{u}}^\sharp[m, n_2] = \frac{1}{W} \sum_{n_1=0}^{W-1} \mathbf{u}^\sharp[n_1, n_2] e^{-j\frac{2\pi m}{W} n_1}, \quad (11)$$

$$\forall m = -M, \dots, M, \quad n_2 = -S, \dots, H-1+S,$$

which are the exact Fourier coefficients of the function  $t \mapsto (x^\sharp * \phi_1)(t, n_2)$  following assumption (ii). Hence, from (4) and  $\widehat{\mathbf{u}}^\sharp[m, n_2] = \frac{1}{W} \int_0^W (x^\sharp * \phi_1)(t, n_2) e^{-j\frac{2\pi m}{W} t} dt$ , we obtain:

$$\widehat{\mathbf{u}}^\sharp[m, n_2] = \widehat{\mathbf{g}}[m] \widehat{\mathbf{x}}^\sharp[m, n_2], \quad \forall m = -M, \dots, M, \quad n_2 = -S, \dots, H-1+S, \quad (12)$$

with

$$\widehat{\mathbf{x}}^\sharp[m, n_2] = \sum_{k=1}^K \frac{\alpha_k}{\cos \theta_k} e^{j2\pi(\tan \theta_k n_2 - \eta_k) \frac{m}{W}}. \quad (13)$$

**Remark.** The image  $\widehat{\mathbf{x}}^\sharp$  is the sampled version of the continuous function  $\widehat{x}^\sharp = \mathcal{F}_1 x^\sharp$ , which is the horizontal Fourier transform of  $x^\sharp$  (2).

Now we apply a 1-D DFT on the first component of the discrete image  $\mathbf{b}^\sharp$  (10), leading to the elements

$$\widehat{\mathbf{b}}^\sharp[m, n_2] = (\widehat{\mathbf{u}}^\sharp[m, :] * \mathbf{h}) [n_2], \quad \forall m = -M, \dots, M, \quad n_2 = 0, \dots, H-1. \quad (14)$$

Since the image  $\mathbf{u}^\sharp$  and the filter  $\mathbf{g}$  are real, then  $\widehat{\mathbf{x}}^\sharp$  is Hermitian, so we can only deal with the right part  $\widehat{\mathbf{x}}^\sharp[0 : M, :]$  and notice that the column corresponding to  $m = 0$  is real and equal to  $\sum_{k=1}^K \frac{\alpha_k}{\cos \theta_k}$ . We consider in the following the Fourier

image  $\widehat{\mathbf{x}}^\sharp[m, n_2]$  of size  $(M+1) \times H_S$ , with  $H_S = H + 2S$ , due to the addition of  $S$  pixels beyond the borders for the convolution by the filter  $\mathbf{h}$ . More precisely,

$$\widehat{\mathbf{x}}^\sharp \in \mathcal{X} = \{\widehat{\mathbf{x}} \in \mathcal{M}_{M+1, H_S}(\mathbb{C}) : \text{Im}(\widehat{\mathbf{x}}[0, :]) = 0\} , \quad (15)$$

endowed with the following inner product, where  $\cdot^*$  denotes the complex conjugation:

$$\langle \widehat{\mathbf{x}}_1, \widehat{\mathbf{x}}_2 \rangle_{\mathcal{X}} = \sum_{n_2=0}^{H_S-1} \widehat{\mathbf{x}}_1[0, n_2] \widehat{\mathbf{x}}_2[0, n_2] + 2\text{Re} \left( \sum_{m=1}^M \sum_{n_2=0}^{H_S-1} \widehat{\mathbf{x}}_1[m, n_2] \widehat{\mathbf{x}}_2[m, n_2]^* \right) . \quad (16)$$

We also define the Hilbert space  $\mathcal{Y}$ , which is equivalent to (15) for  $S = 0$ :

$$\mathcal{Y} = \{\widehat{\mathbf{y}} \in \mathcal{M}_{M+1, H}(\mathbb{C}) : \text{Im}(\widehat{\mathbf{y}}[0, :]) = 0\} , \quad (17)$$

endowed with the inner product (16) for  $S = 0$ ; that is:

$$\langle \widehat{\mathbf{y}}_1, \widehat{\mathbf{y}}_2 \rangle_{\mathcal{Y}} = \sum_{n_2=0}^{H-1} \widehat{\mathbf{y}}_1[0, n_2] \widehat{\mathbf{y}}_2[0, n_2] + 2\text{Re} \left( \sum_{m=1}^M \sum_{n_2=0}^{H-1} \widehat{\mathbf{y}}_1[m, n_2] \widehat{\mathbf{y}}_2[m, n_2]^* \right) . \quad (18)$$

Thereafter, we will have to deal with rows and columns of  $\widehat{\mathbf{x}} \in \mathcal{X}$ , which respectively belong to the Hilbert space denoted by  $\mathcal{X}_l$  and  $\mathcal{X}_t$ , endowed with the following inner products:

$$\langle \mathbf{l}_1, \mathbf{l}_2 \rangle_{\mathcal{X}_l} = \mathbf{l}_1[0] \mathbf{l}_2[0] + 2\text{Re} \left( \sum_{m=1}^M \mathbf{l}_1[m] \mathbf{l}_2[m]^* \right) , \quad (19)$$

$$\langle \mathbf{t}_1, \mathbf{t}_2 \rangle_{\mathcal{X}_t} = 2\text{Re} \left( \sum_{n_2=1}^{H_S} \mathbf{t}_1[n_2] \mathbf{t}_2[n_2]^* \right) . \quad (20)$$

Let  $\mathbf{A} : \mathcal{X} \rightarrow \mathcal{Y}$  be the operator which multiplies each row vector  $\widehat{\mathbf{x}}[m, :]$  of  $\widehat{\mathbf{x}} \in \mathcal{X}$  by the corresponding Fourier coefficient  $\widehat{\mathbf{g}}[m]$  for  $m = 0, \dots, M$  and convolves it with the filter  $\mathbf{h} = [h_{-S}, \dots, h_0, \dots, h_S]$ . From a matricial point of view, it corresponds to a right and left matrix multiplication with the matrices  $\widehat{\mathbf{G}}$  of size  $(M+1) \times (M+1)$  and  $\check{\mathbf{H}}$  of size  $H \times H_S$  defined by:

$$\widehat{\mathbf{G}} = \text{diag}(\widehat{g}_0, \dots, \widehat{g}_M), \quad \check{\mathbf{H}} = \begin{pmatrix} h_{-S} & \cdots & h_S & 0 & 0 & \cdots & 0 \\ 0 & h_{-S} & \cdots & h_S & 0 & \cdots & 0 \\ \vdots & \cdots & \ddots & \ddots & \ddots & \cdots & \vdots \\ 0 & \cdots & 0 & h_{-S} & \cdots & h_S & 0 \\ 0 & \cdots & 0 & 0 & h_{-S} & \cdots & h_S \end{pmatrix} , \quad (21)$$

that is,

$$\mathbf{A} \widehat{\mathbf{x}} = \check{\mathbf{H}} \widehat{\mathbf{x}} \widehat{\mathbf{G}} . \quad (22)$$

Following (12) and (14), the inverse problem writes

$$\mathbf{A} \widehat{\mathbf{x}}^\sharp = \widehat{\mathbf{b}}^\sharp . \quad (23)$$

Finally, the image  $\mathbf{b}^\sharp$  of the blurred lines is corrupted by noise, so that we observe the degraded image

$$\mathbf{y} = \mathbf{b}^\sharp + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(0, \zeta^2) , \quad (24)$$

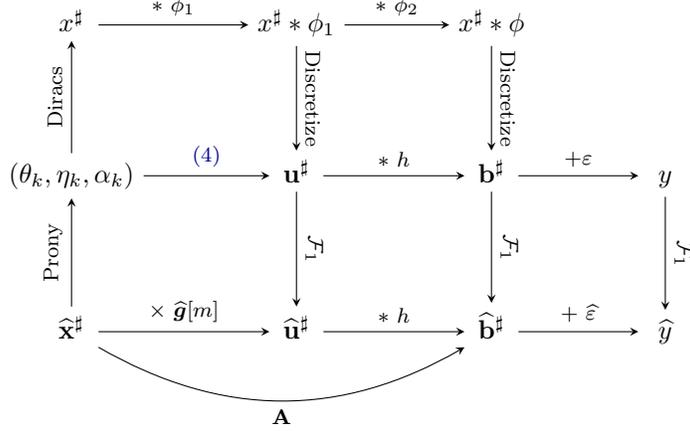


Figure 3: Relation between variables.

with  $\zeta$  the noise level. The problem reverts to an inverse problem we need to solve. In absence of noise, one can find a theoretical solution of this ill-posed problem, by solving on each column  $m$  the following reduced problems:  $\hat{\mathbf{g}}[m]\hat{\mathbf{H}}\hat{\mathbf{x}}[m, :] = \hat{\mathbf{b}}[m, :]$  where  $\hat{\mathbf{H}}$  is the convolution matrix (21) corresponding to the vertical convolution by filter  $\mathbf{h}$ .

In order to address the noisy case, we will need to derive an optimization problem of this inverse problem, under constraints exploiting the sparse structure of the signal we are looking for, namely a combination of lines. The super-resolution consists to recover the high frequencies (lost because of the blur operator) from the degraded image  $\mathbf{y}$ , so it can be viewed as a spectral extrapolation process, and then to recover the parameters  $(\theta_k, \eta_k, \alpha_k)$  of these lines. This procedure can be decomposed as:

1. First solve the optimization problem

$$\text{Minimize } \|\hat{\mathbf{y}} - \mathbf{A}\hat{\mathbf{x}}\|_{\mathcal{Y}}, \text{ under the constraint that } \mathbf{x} \text{ is made of lines ; } \quad (25)$$

that is, to go to the bottom line of the diagram in Figure 3 from  $\hat{\mathbf{y}}$  to  $\hat{\mathbf{x}}^{\#}$ .

2. Second perform a Prony-like method onto  $\hat{\mathbf{x}}^{\#}$  in order to estimate the  $K$  parameters  $(\theta_k, \eta_k, \alpha_k)$ .

These two steps are summarized in Figure 4. Note that this work also covers the case where a mask is applied; that is, it can encompass inpainting problems. In the next section, we present the framework of atomic norm from which the optimization problem will be derived.

## 3 Super-Resolution Detection of Lines

### 3.1 Atomic Norm and Semidefinite Characterizations

Consider a complex signal  $\mathbf{z} \in \mathbb{C}^N$  represented as a  $K$ -sparse mixture of atoms from the set

$$\mathcal{A} = \{\mathbf{a}(\omega) \in \mathbb{C}^N : \omega \in \Omega\} ,$$

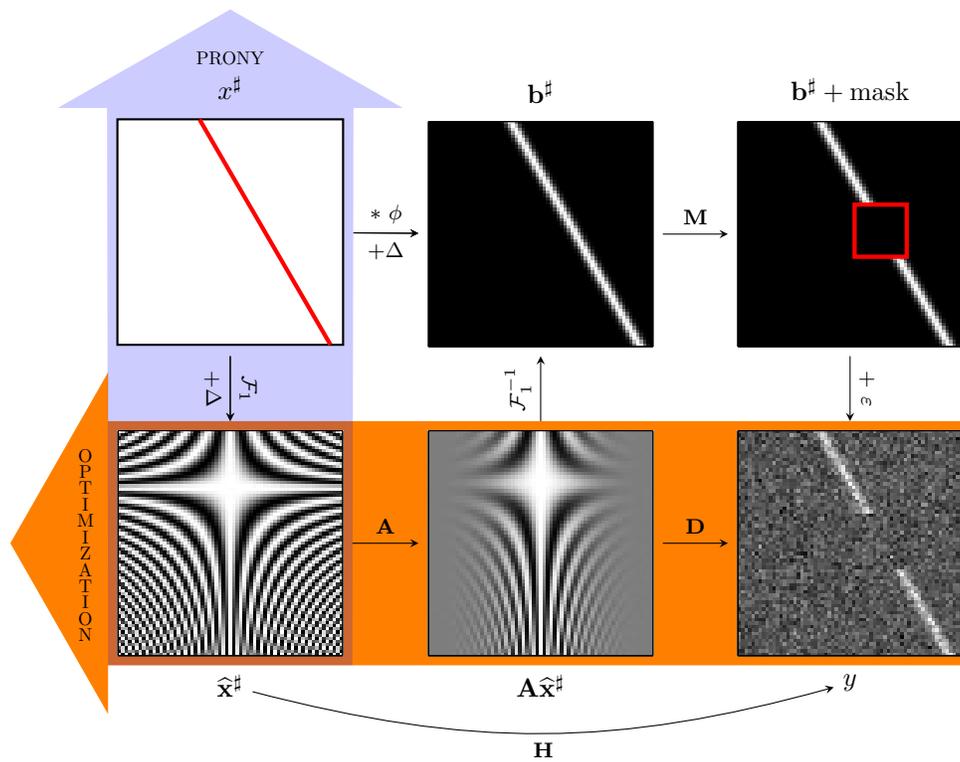


Figure 4: The two steps of the procedure : a convex optimization for the reconstruction of lines (in orange) and a Prony-like method for the estimation of their parameters (in purple).

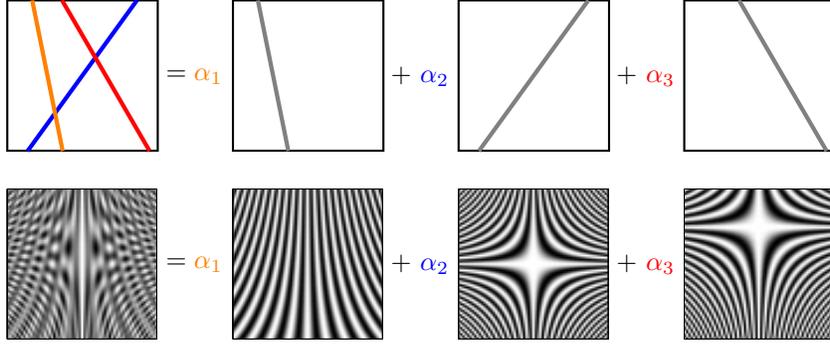


Figure 5: Illustration with a signal made of a weighted combination of three lines atoms (in gray)  $\mathbf{a}(\theta_k, \eta_k)$ . In the Fourier domain, we have the same kind of combination but with 2-D exponentials atoms  $\hat{\mathbf{a}}(\theta_k, \eta_k)$ . In both cases the sum of the weights  $\alpha_1 + \alpha_2 + \alpha_3$ , where  $\alpha_i$  are the amplitudes of the lines, corresponds to the atomic norm on the appropriate dictionary.

that is,

$$\mathbf{z} = \sum_{k=1}^K c_k \mathbf{a}(\omega_k), \quad c_k \in \mathbb{C}, \quad \omega_k \in \Omega.$$

We consider atoms  $\mathbf{a}(\omega) \in \mathbb{C}^N$  that are continuously indexed in the dictionary  $\mathcal{A}$  by the parameter  $\omega$  in a compact set  $\Omega$ . The atomic norm, first introduced in [14], is defined as

$$\|\mathbf{z}\|_{\mathcal{A}} = \inf \{t > 0 : \mathbf{z} \in t \text{conv}(\mathcal{A})\},$$

where  $\text{conv}(\mathcal{A})$  denotes the convex hull of a general atomic set  $\mathcal{A}$ , enforcing sparsity. Chandrasekaran et al. [14] argue that the atomic norm is the best convex heuristic for underdetermined, structured linear inverse problems, which generalizes the  $\ell_1$  norm for sparse recovery and the nuclear norm for low-rank matrix completion.

In our problem (2) the atoms are *lines*, so one can consider the dictionary  $\mathcal{A}_{2D}$  indexed by the angle and the offset; that is, composed by the line atoms

$$\mathbf{a}(\theta_k, \eta_k) = \delta(\cos \theta_k (t_1 - \eta_k) + \sin \theta_k t_2).$$

Or alternatively in the Fourier domain, the dictionary  $\hat{\mathcal{A}}_{2D}$ , composed by the 2-D exponentials atoms

$$\hat{\mathbf{a}}(\theta_k, \eta_k) = \frac{1}{\cos \theta_k} e^{j2\pi(\tan \theta_k n_2 - \eta_k)m/W},$$

as illustrated in Figure 5. The problem is that there is no closed-form expression for the atomic norm in these 2-D dictionaries, to our knowledge. However, in the case of 1-D complex exponentials, there is a way to compute the atomic norm via semidefinite programming. So, we reformulate the problem using the simplified 1-D case. From now on, we consider the dictionaries

$$\mathcal{A} = \left\{ \mathbf{a}(f, \phi) \in \mathbb{C}^{|I|}, \quad f \in [0, 1], \quad \phi \in [0, 2\pi) \right\}, \quad (26)$$

$$\mathcal{A}_0 = \left\{ \mathbf{a}(f) \in \mathbb{C}^{|I|}, \quad f \in [0, 1] \right\}, \quad (27)$$

in which the atoms are the vectors of components  $[\mathbf{a}(f, \phi)]_i = e^{j(2\pi f i + \phi)}$ ,  $i \in I$ , and  $[\mathbf{a}(f)]_i = [\mathbf{a}(f, 0)]_i = e^{j2\pi f i}$ ,  $i \in I$ . The atomic norm writes:

$$\|\mathbf{z}\|_{\mathcal{A}} = \inf_{\substack{c'_k \geq 0 \\ f'_k \in [0, 1[ \\ \phi'_k \in [0, 2\pi)}} \left\{ \sum_k c'_k : \mathbf{z} = \sum_k c'_k \mathbf{a}(f'_k, \phi'_k) \right\}. \quad (28)$$

**Theorem 1** (Caratheodory). *A vector  $\mathbf{z} = (z_{N-1}^*, \dots, z_1^*, z_0, z_1, \dots, z_{N-1})$  of length  $2N - 1$ , with  $z_0 \in \mathbb{R}$ , is a positive combination of  $K \leq N$  atoms  $\mathbf{a}(f_k)$  if and only if  $\mathbf{T}_N(\mathbf{z}_+) \succcurlyeq 0$  and is of rank  $K$ , where  $\mathbf{z}_+ = (z_0, \dots, z_{N-1})$  is of length  $N$  and  $\mathbf{T}_N : \mathbb{C}^N \rightarrow \mathbb{C}^N \times \mathbb{C}^N$  is the Toeplitz operator*

$$\mathbf{T}_N : \mathbf{z}_+ = (z_0, \dots, z_{N-1}) \mapsto \begin{pmatrix} z_0 & z_1^* & \cdots & z_{N-1}^* \\ z_1 & z_0 & \cdots & z_{N-2}^* \\ \vdots & \vdots & \ddots & \vdots \\ z_{N-1} & z_{N-2} & \cdots & z_0 \end{pmatrix}, \quad (29)$$

and  $\succcurlyeq 0$  denotes positive semidefiniteness. Moreover, this decomposition is unique, if  $K < N$ .

*Proof.* See references [9, 10, 55, 40].  $\square$

We also derived the above characterization, which gets rid of a useless variable.

**Proposition 2.** *The atomic norm  $\|\mathbf{z}\|_{\mathcal{A}}$  can be characterized by the following semidefinite program:*

$$\|\mathbf{z}\|_{\mathcal{A}} = \min_{\mathbf{q} \in \mathbb{C}^N, q_0 \geq 0} \left\{ q_0 : \mathbf{T}'_N(\mathbf{z}, \mathbf{q}) = \begin{pmatrix} \mathbf{T}_N(\mathbf{q}) & \mathbf{z} \\ \mathbf{z}^* & q_0 \end{pmatrix} \succcurlyeq 0 \right\}. \quad (30)$$

where  $q_0$  is the first component of vector  $\mathbf{q} = (q_0, \dots, q_{N-1}) \in \mathbb{R}^+ \times \mathbb{C}^{N-1}$  and  $\mathbf{z}^* = \bar{\mathbf{z}}^T$ .

*Proof.* This result is an improvement of [54, Proposition II.1] and the proof is given in [Appendix A](#).  $\square$

Since the matrix  $\mathbf{T}'_N(\mathbf{z}, \mathbf{q})$  in (30) is Hermitian and positive semidefinite, its eigenvalues  $(\lambda_i)_{0 \leq i \leq N}$  are positive reals. So,  $q_0 = \frac{1}{N+1} \text{Trace}(\mathbf{T}'_N(\mathbf{z}, \mathbf{q})) = \frac{1}{N+1} \sum_{i=0}^N \lambda_i$  is real and positive. We also define, for the upcoming convex optimization problem, the following set of complex matrices

$$\mathcal{Q} = \{ \mathbf{q} \in \mathbb{C}^{M+1} \times \mathbb{C}^{H_S} : \text{Im}(\mathbf{q}[:, 0]) = 0 \}, \quad (31)$$

endowed with the following inner product:

$$\langle \mathbf{q}_1, \mathbf{q}_2 \rangle_{\mathcal{Q}} = \sum_{m=0}^M \mathbf{q}_1[m, 0] \mathbf{q}_2[m, 0] + 2\text{Re} \left( \sum_{n_2=1}^{H_S-1} \sum_{m=0}^M \mathbf{q}_1[m, n_2] \mathbf{q}_2[m, n_2]^* \right). \quad (32)$$

**Remark.** *The columns of  $\mathbf{q} \in \mathcal{Q}$  belong to the Hilbert space  $\mathcal{Q}_t$  which is  $\mathbb{R} \times \mathcal{C}^{H_S-1}$  endowed with the inner product:*

$$\langle \mathbf{q}_1, \mathbf{q}_2 \rangle_{\mathcal{Q}_t} = \mathbf{q}_1[0] \mathbf{q}_2[0] + 2\text{Re} \left( \sum_{n_2=1}^{H_S-1} \mathbf{q}_1[n_2] \mathbf{q}_2[n_2]^* \right). \quad (33)$$

### 3.2 Properties of the Model $\widehat{\mathbf{x}}^\sharp$ with respect to the Atomic Norm

In Fourier domain, the discrete image  $\widehat{\mathbf{x}}^\sharp$  given by (13) can be viewed as a sum of atoms: first considering the rows  $\mathbf{l}_{n_2}$  of the matrix  $\widehat{\mathbf{x}}^\sharp$ , with  $I = \{0, \dots, H_S - 1\}$ :

$$\mathbf{l}_{n_2}^\sharp = \widehat{\mathbf{x}}^\sharp[:, n_2] = \sum_{k=1}^K c_k \mathbf{a}(f_{n_2, k}), \quad (34)$$

second considering the columns  $\mathbf{t}_m$ , with  $I = \{-M, \dots, M\}$ :

$$\mathbf{t}_m^\sharp = \widehat{\mathbf{x}}^\sharp[m, :] = \sum_{k=1}^K c_k \mathbf{a}(f_{m, k}, \phi_{m, k})^\top, \quad (35)$$

where

$$\begin{aligned} c_k &= \frac{\alpha_k}{\cos \theta_k}, & f_{n_2, k} &= \frac{\tan \theta_k n_2 - \eta_k}{W}, \\ \phi_{m, k} &= -\frac{2\pi \eta_k m}{W}, & f_{m, k} &= \frac{\tan \theta_k m}{W}. \end{aligned} \quad (36)$$

We define for later use, the frequency  $\nu_k = \eta_k/W$  and the coefficients  $d_{m, k} = c_k e^{j\phi_{m, k}}$ ,  $e_{m, k} = e^{j\phi_{m, k}}$ . The vectors  $\mathbf{l}_{n_2}^\sharp$  of size  $W = 2M + 1$  are positive combinations of  $K$  atoms  $\mathbf{a}(f_{n_2, k})$ , with  $K \leq M$  since we can reasonably assume that the number of lines  $K$  is smaller than half the number of pixels  $M$ . Thus, [Theorem 1](#) ensures that the decomposition (34) is unique, hence, following (28):

$$\|\mathbf{l}_{n_2}^\sharp\|_{\mathcal{A}} = \sum_{k=1}^K c_k = \widehat{\mathbf{x}}^\sharp[0, n_2], \quad \forall n_2 = 0, \dots, H_S - 1. \quad (37)$$

By contrast, since the coefficients  $d_{m, k}$  are complex, [Theorem 1](#) no longer holds and we simply have from [Proposition 2](#):

$$\|\mathbf{t}_m^\sharp\|_{\mathcal{A}} \leq \sum_{k=1}^K c_k, \quad \forall m = -M, \dots, M. \quad (38)$$

Let us take a closer look at the case of one line; that is,  $K = 1$ , characterized by parameters  $(\theta, \eta, \alpha)$ . We recall by (13) that  $\widehat{\mathbf{x}}^\sharp$  can be written as:

$$\widehat{\mathbf{x}}^\sharp[m, n_2] = c_1 e^{j2\pi((f_1 - f_0)n_2 + f_0)m}, \quad c_1 = \frac{\alpha}{\cos \theta}, \quad f_0 = -\frac{\eta}{W}, \quad f_1 = \frac{\tan \theta - \eta}{W}.$$

Let  $\mathbf{z} = (z_0, \dots, z_{|I|-1})$  be a complex vector, whose elements  $z_i$  are rearranged in a Toeplitz matrix  $\mathbf{P}_K(\mathbf{z})$  of size  $(|I| - K) \times (K + 1)$  and rank  $K$  as follows:

$$\mathbf{P}_K(\mathbf{z}) = \begin{pmatrix} z_K & \cdots & z_0 \\ \vdots & \ddots & \vdots \\ z_{|I|-1} & \cdots & z_{|I|-K-1} \end{pmatrix}.$$

We get the following characterization of one line in Fourier domain:

**Proposition 3.** An image  $\hat{\mathbf{x}}$  is of the form  $\hat{\mathbf{x}}[m, n] = c_1 e^{j2\pi((f_1 - f_0)n + f_0)m}$  if and only if the rows  $\mathbf{l}_n$  and columns  $\mathbf{t}_m$  of  $\hat{\mathbf{x}}$  satisfy  $\mathbf{T}_M(\mathbf{l}_n)$  is positive semidefinite and of rank one,  $\mathbf{P}_1(\mathbf{t}_m)$  is of rank one, and  $\hat{\mathbf{x}}[0, n] = \hat{\mathbf{x}}[0, 0]$  for all  $m$  and  $n$ .

*Proof.* See [Appendix B](#).  $\square$

Besides, with  $\mathbf{D} = \text{diag}(c_1, \dots, c_K)$  and  $\mathbf{V}_{n_2} = [\mathbf{a}(f_{n_2,1}) \ \cdots \ \mathbf{a}(f_{n_2,K})]$ , we can remark that

$$\mathbf{T}_M(\mathbf{l}_{n_2}^\#) = \sum_{k=1}^K c_k \mathbf{T}_M(\mathbf{a}(f_{n_2,k})) = \sum_{k=1}^K c_k \mathbf{a}(f_{n_2,k}) \mathbf{a}(f_{n_2,k})^* = \mathbf{V}_{n_2} \mathbf{D} \mathbf{V}_{n_2}^*,$$

where  $\cdot^*$  denotes the hermitian conjugate. Hence the nuclear norm of  $\mathbf{T}_M(\mathbf{l}_{n_2}^\#)$  is

$$\|\mathbf{T}_M(\mathbf{l}_{n_2}^\#)\|_* = \sum_{k=1}^K c_k = \|\mathbf{l}_{n_2}^\#\|_{\mathcal{A}}.$$

The first equality is explained by the singular value decomposition (SVD) of a matrix  $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}$ , since the nuclear norm of a matrix corresponds to the sum of its singular values. It is often used as a convex approximation of the rank of this matrix [47, 46]. Consequently, in the following, we consider a convex relaxation of the line characterization given in [Proposition 3](#), in which the rank constraint on  $\mathbf{T}_M(\mathbf{l}_{n_2}^\#)$  is replaced by an atomic norm constraint on  $\|\mathbf{l}_{n_2}^\#\|_{\mathcal{A}}$ . Since the minimum value to achieve is  $c^\# = \sum_{k=1}^K c_k$ , and since the atomic norm lies on the first column  $\hat{\mathbf{x}}[0, n_2] = \hat{\mathbf{x}}[0, 0]$ , we impose the constraint  $\hat{\mathbf{x}}[0, n_2] = \hat{\mathbf{x}}[0, 0] \leq c^\#$ . We do the same for the columns.

**Remark.** We did not establish a similar characterization for  $K \geq 2$  lines, but the philosophy remains: the aim is to minimize atomic norms of rows and columns simultaneously, so that the solution will be composed of sparse sums of exponentials in both directions. This leads to the following optimization problem presented in the next section.

## 4 Minimization Problem with the Atomic Norm Regularization

Given the operator  $\mathbf{A}$  defined in (22) by means of the filters (7)-(8) and  $\hat{\mathbf{y}}$  the Fourier version of the degraded image observed (24), we are looking for an image  $\hat{\mathbf{x}} \in \mathcal{X}$  which minimizes  $\|\mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{y}}\|_{\mathcal{Y}}$ , for the norm derived from the inner product (18), and whose rows and columns satisfy properties (37) and (38). We fix a constant  $c$ . Consequently, the following optimization problem provides an estimator of  $\hat{\mathbf{x}}^\#$  defined in (13):

$$\hat{\mathbf{x}} \in \arg \min_{(\hat{\mathbf{x}}, \mathbf{q}) \in \mathcal{X} \times \mathcal{Q}} \frac{1}{2} \|\mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{y}}\|_{\mathcal{Y}}^2, \quad (39)$$

$$s.t. \begin{cases} \hat{\mathbf{x}}[0, n_2] = \hat{\mathbf{x}}[0, 0] \leq c, & (40a) \\ \mathbf{q}[m, 0] \leq c, & (40b) \\ \mathbf{T}_{M+1}(\hat{\mathbf{x}}[:, n_2]) \succcurlyeq 0, & (40c) \\ \mathbf{T}'_{H_S}(\hat{\mathbf{x}}[m, :], \mathbf{q}[m, :]) \succcurlyeq 0, & (40d) \\ \forall n_2 = 0, \dots, H_S - 1, \forall m = 1, \dots, M, \end{cases}$$

where the Hilbert spaces  $(\mathcal{X}, \langle \cdot, \cdot \rangle_{\mathcal{X}})$ ,  $(\mathcal{Y}, \langle \cdot, \cdot \rangle_{\mathcal{Y}})$  and  $(\mathcal{Q}, \langle \cdot, \cdot \rangle_{\mathcal{Q}})$  are respectively defined in (15)-(16), (17)-(18) and (31)-(32); and the operators  $\mathbf{T}_{M+1} : \mathcal{X}_t \rightarrow \mathcal{T}_{M+1}$  and  $\mathbf{T}'_{H_S} : \mathcal{X}_t \times \mathcal{Q}_t \rightarrow \mathcal{T}_{H_S+1}$  are defined respectively on rows and columns of  $\widehat{\mathbf{x}} \in \mathcal{X}$  and  $\mathbf{q} \in \mathcal{Q}$ , endowed respectively with the inner products (19) and (20)-(33), to Hermitian-Toeplitz matrices of dimension  $M+1$  and  $H_S+1$  respectively, whose spaces are denoted by  $\mathcal{T}_{M+1}$  and  $\mathcal{T}_{H_S+1}$  endowed with the classical inner product on complex matrices:

$$\langle \mathbf{M}, \mathbf{N} \rangle_{\mathcal{M}} = \sum_{i,j} \mathbf{M}_{ij}^* \mathbf{N}_{ij}, \quad (41)$$

and corresponding Frobenius norm

$$\|\mathbf{M}\|_{\text{F}} = \left( \sum_{i,j} |\mathbf{M}_{ij}|^2 \right)^{1/2}. \quad (42)$$

The expressions of the operators  $\mathbf{T}_N$  and  $\mathbf{T}'_N$  are given respectively in (29) and (30).

**Remark.** *This optimization problem could be rewritten in a regularized form involving a parameter  $\lambda$  to tune, what is not any better than the tuning parameter  $c$ , which has the advantage of having a physical meaning, related to the line intensities.*

We keep this constrained formulation and write it in a more suitable way as follows. Let  $\mathcal{H} = \mathcal{X} \times \mathcal{Q}$  be the Hilbert space in which the variable optimization  $\mathbf{X} = (\widehat{\mathbf{x}}, \mathbf{q})$  lies, endowed with the following inner product:

$$\langle (\widehat{\mathbf{x}}_1, \mathbf{q}_1), (\widehat{\mathbf{x}}_2, \mathbf{q}_2) \rangle_{\mathcal{H}} = \langle \widehat{\mathbf{x}}_1, \widehat{\mathbf{x}}_2 \rangle_{\mathcal{X}} + \langle \mathbf{q}_1, \mathbf{q}_2 \rangle_{\mathcal{Q}}. \quad (43)$$

Let us define  $L_m^{(1)} : \mathcal{H} \rightarrow \mathcal{T}_{H_S+1}$  and  $L_{n_2}^{(2)} : \mathcal{H} \rightarrow \mathcal{T}_{M+1}$  by

$$L_m^{(1)}(\mathbf{X}) = \mathbf{T}'_{H_S}(\widehat{\mathbf{x}}[m, :], \mathbf{q}[m, :]), \quad (44)$$

$$L_{n_2}^{(2)}(\mathbf{X}) = \mathbf{T}_{M+1}(\widehat{\mathbf{x}}[:, n_2]). \quad (45)$$

If  $\iota_C$  is the indicator function of a convex set  $C$  defined by

$$\iota_C : x \mapsto \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{if } x \notin C \end{cases},$$

where  $C$  will be either  $\mathcal{B} \subset \mathcal{H}$  the set corresponding to the boundary constraints

$$\mathcal{B} = \left\{ (\widehat{\mathbf{x}}, \mathbf{q}) \in \mathcal{H} : \widehat{\mathbf{x}}[0, n_2] = \widehat{\mathbf{x}}[0, 0] \leq c, \mathbf{q}[m, 0] \leq c \right\}, \quad (46)$$

or the cone of positive semidefinite matrices  $\mathcal{C}$ . Then, the optimization problem (39) under constraints (40) can be rewritten as follows:

$$\tilde{\mathbf{X}} = \arg \min_{\mathbf{X} = (\widehat{\mathbf{x}}, \mathbf{q}) \in \mathcal{H}} \left\{ \frac{1}{2} \|\mathbf{A}\widehat{\mathbf{x}} - \widehat{\mathbf{y}}\|_{\mathcal{Y}}^2 + \iota_{\mathcal{B}}(\mathbf{X}) + \sum_{m=1}^M \iota_{\mathcal{C}}(L_m^{(1)}(\mathbf{X})) + \sum_{n_2=0}^{H_S-1} \iota_{\mathcal{C}}(L_{n_2}^{(2)}(\mathbf{X})) \right\}. \quad (47)$$

## 4.1 First Algorithm Design

The optimization problem (47) can be cast as a minimization problem, involving smooth, proximable and linear composite terms [17]:

$$\tilde{\mathbf{X}} = \arg \min_{\mathbf{X} \in \mathcal{H}} \left\{ F(\mathbf{X}) + G(\mathbf{X}) + \sum_{i=0}^{Q-1} H_i(\mathbf{L}_i(\mathbf{X})) \right\}, \quad (48)$$

with  $F(\mathbf{X}) = \frac{1}{2} \|\mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{y}}\|_{\mathcal{Y}}^2$ ,  $\mathbf{X} = (\hat{\mathbf{x}}, \mathbf{q})$ ,  $G = \iota_{\mathcal{B}}$ , which is proximable,  $Q = M + H_S$  linear composite terms where  $H_i = \iota_{\mathcal{C}}$ ,  $\mathbf{L}_i = \mathbf{L}_i^{(2)}$  when  $0 \leq i \leq H_S - 1$  and  $\mathbf{L}_i = \mathbf{L}_{i-H_S+1}^{(1)}$  when  $H_S \leq i \leq H_S + M - 1$ . We define  $\mathbf{H}\mathbf{x} = \sum_{i=0}^{Q-1} H_i x_i$ ,  $\mathbf{L}^{(1)}(\mathbf{X}) = (\mathbf{L}_1^{(1)}(\mathbf{X}), \dots, \mathbf{L}_M^{(1)}(\mathbf{X}))$  and  $\mathbf{L}^{(2)}(\mathbf{X}) = (\mathbf{L}_0^{(2)}(\mathbf{X}), \dots, \mathbf{L}_{H_S-1}^{(2)}(\mathbf{X}))$ .  $\mathbf{L} = (\mathbf{L}^{(1)}, \mathbf{L}^{(2)})$  is the linear operator such that the composite terms rewrite  $\mathbf{H} \circ \mathbf{L}$ . We define onto the range of  $\mathbf{L}^{(1)}$ ,  $\mathbf{L}^{(2)}$  and  $\mathbf{L}$ , which are cartesian product spaces, an inner product as the sum of the inner products defined on these spaces (similarly to (43)), and corresponding norms denoted by  $\|\cdot\|_{(1)}$ ,  $\|\cdot\|_{(2)}$  and  $\|\cdot\|_{(1,2)}$ . We define the operator norms

$$\|\mathbf{A}\| = \sup_{\hat{\mathbf{x}} \in \mathcal{X}} \frac{\|\mathbf{A}\hat{\mathbf{x}}\|_{\mathcal{Y}}}{\|\hat{\mathbf{x}}\|_{\mathcal{X}}}, \quad (49)$$

$$\|\mathbf{L}_i\| = \sup_{\mathbf{X} \in \mathcal{H}} \frac{\|\mathbf{L}_i(\mathbf{X})\|_{\mathcal{F}}}{\|\mathbf{X}\|_{\mathcal{H}}}, \quad (50)$$

$$\|\mathbf{L}^{(j)}\| = \sup_{\mathbf{X} \in \mathcal{H}} \frac{\|\mathbf{L}^{(j)}(\mathbf{X})\|_{(j)}}{\|\mathbf{X}\|_{\mathcal{H}}}, \quad j \in \{1, 2\}, \quad (51)$$

$$\|\mathbf{L}\| = \sup_{\mathbf{X} \in \mathcal{H}} \frac{\|\mathbf{L}(\mathbf{X})\|_{(1,2)}}{\|\mathbf{X}\|_{\mathcal{H}}}, \quad (52)$$

We now establish some properties of the functions, adjoint operators and norms previously defined.

**Lemma 1.** *The norm (49) of the operator  $\mathbf{A}$  defined in (22) is given by*

$$\|\mathbf{A}\| = \|\hat{\mathbf{g}}\|_{\infty} \|\hat{\mathbf{h}}\|_{\infty}. \quad (53)$$

*Proof.* By definition  $\mathbf{A}\hat{\mathbf{x}} = (\hat{\mathbf{x}}\hat{\mathbf{G}}) * \mathbf{h}$ , where  $\hat{\mathbf{G}} = \text{diag}(\hat{g}_{M+1}, \dots, \hat{g}_W)$ . Let compute the norm of this operator. If we denote by  $\hat{\mathbf{x}}_k$  the  $k$ -th column of  $\hat{\mathbf{x}}$ , then by considering in Fourier the norm operator  $f \mapsto f * h$  we have the inequality

$$\|\hat{\mathbf{x}}_k * \mathbf{h}\|_2 \leq \|\hat{\mathbf{h}}\|_{\infty} \|\hat{\mathbf{x}}_k\|_2.$$

Thus, with the norm derived from (16) we get

$$\begin{aligned} \|\mathbf{A}\hat{\mathbf{x}}\|_{\mathcal{Y}}^2 &= |\hat{g}_{M+1}|^2 \|\hat{\mathbf{x}}_0 * \mathbf{h}\|_2^2 + 2|\hat{g}_{M+2}|^2 \|\hat{\mathbf{x}}_1 * \mathbf{h}\|_2^2 + \dots + 2|\hat{g}_W|^2 \|\hat{\mathbf{x}}_M * \mathbf{h}\|_2^2, \\ &\leq \|\hat{\mathbf{g}}\|_{\infty}^2 \|\hat{\mathbf{h}}\|_{\infty}^2 (\|\hat{\mathbf{x}}_0\|_2^2 + 2\|\hat{\mathbf{x}}_1\|_2^2 + \dots + 2\|\hat{\mathbf{x}}_M\|_2^2), \\ &\leq \|\hat{\mathbf{g}}\|_{\infty}^2 \|\hat{\mathbf{h}}\|_{\infty}^2 \|\hat{\mathbf{x}}\|_{\mathcal{X}}^2. \end{aligned} \quad (54)$$

Since filter  $h$  is low-pass, then the equality is achieved for an image  $\hat{\mathbf{x}}$  whose all columns are null except one which is constant (non zero), of index  $m_0$  where  $\hat{g}_{m_0}$  corresponds to the maximum  $\|\hat{\mathbf{g}}\|_{\infty}$ , which proves the result (53).  $\square$

**Lemma 2.** *The adjoint operator of the operator  $\mathbf{A}$  is denoted by  $\mathbf{A}^*$  and defined such that  $\langle \mathbf{A}\hat{\mathbf{x}}, \hat{\mathbf{z}} \rangle_{\mathcal{X}} = \langle \hat{\mathbf{x}}, \mathbf{A}^*\hat{\mathbf{z}} \rangle_{\mathcal{Y}}$ . Its matricial expression is given by*

$$\mathbf{A}^*\hat{\mathbf{z}} = \check{\mathbf{H}}^*\hat{\mathbf{z}}\hat{\mathbf{G}}^*, \quad (55)$$

or alternatively by  $\mathbf{A}^*\hat{\mathbf{z}} = (\hat{\mathbf{z}}\hat{\mathbf{G}}^*) * \bar{\mathbf{h}}'$  where  $\mathbf{h}'$  is the flipped vector from  $\mathbf{h}$ .

*Proof.* From the definition (22) we have  $\mathbf{A}\hat{\mathbf{x}} = \check{\mathbf{H}}\hat{\mathbf{x}}\hat{\mathbf{G}}$ , which is a matrix product. Then, for a matrix  $\mathbf{M}$  we keep the usual property  $\langle \mathbf{M}\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2 \rangle_{\mathcal{X}} = \langle \hat{\mathbf{x}}_1, \mathbf{M}^*\hat{\mathbf{x}}_2 \rangle_{\mathcal{X}}$  noticing that

$$\langle \hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2 \rangle_{\mathcal{X}} = \langle \hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2 \rangle_{\mathcal{M}} + \langle \hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2 \rangle_{\mathcal{M}}^* - \langle \hat{\mathbf{x}}_1[:, 0], \hat{\mathbf{x}}_2[:, 0] \rangle_{\mathbb{C}^{H_S}},$$

and then the proof is straightforward.  $\square$

Then, we have the following proposition:

**Proposition 4.** *For  $\mathbf{X} = (\hat{\mathbf{x}}, \mathbf{q})$ , the gradient of  $F(\mathbf{X}) = \frac{1}{2}\|\mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{y}}\|_{\mathcal{Y}}^2$  is*

$$\nabla F(\mathbf{X}) = (\mathbf{A}^*(\mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{y}}), \mathbf{0})^{\top},$$

which is Lipschitz-continuous with Lipschitz constant  $\beta = \|\hat{\mathbf{g}}\|_{\infty}^2 \|\hat{\mathbf{h}}\|_{\infty}^2$ .

*Proof.* Let  $F_1$  and  $F_2$  be the applications:

$$\begin{aligned} F_1 : \hat{\mathbf{x}} \in \mathcal{X} &\mapsto \frac{1}{2}\|\mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{y}}\|_{\mathcal{Y}}^2 \in \mathbb{R}, \\ F_2 : \mathbf{X} = (\hat{\mathbf{x}}, \mathbf{q}) \in \mathcal{H} &\mapsto \hat{\mathbf{x}} \in \mathcal{X}. \end{aligned}$$

Then,  $F : \mathcal{H} \mapsto \mathbb{R}$  writes  $F = F_1 \circ F_2$  and its differential at  $\mathbf{X}_0$  is:

$$(\mathrm{d}F)_{\mathbf{X}_0}(\mathbf{X}) = (\mathrm{d}F_1)_{F_2(\mathbf{X}_0)} \circ (\mathrm{d}F_2)_{\mathbf{X}_0}(\mathbf{X}).$$

First, we have

$$\begin{aligned} F_1(\hat{\mathbf{x}} + \mathbf{h}) &= \frac{1}{2}\|\mathbf{A}(\hat{\mathbf{x}} + \mathbf{h}) - \hat{\mathbf{y}}\|_{\mathcal{Y}}^2, \\ &= \frac{1}{2}\|\mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{y}}\|_{\mathcal{Y}}^2 + \frac{1}{2}\langle \mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{y}}, \mathbf{A}\mathbf{h} \rangle_{\mathcal{Y}} + \frac{1}{2}\langle \mathbf{A}\mathbf{h}, \mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{y}} \rangle_{\mathcal{Y}} + \frac{1}{2}\|\mathbf{A}\mathbf{h}\|_{\mathcal{Y}}^2, \\ &= F_1(\hat{\mathbf{x}}) + \langle \mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{y}}, \mathbf{A}\mathbf{h} \rangle_{\mathcal{Y}} + o(\|\mathbf{h}\|_{\mathcal{X}}), \\ &= F_1(\hat{\mathbf{x}}) + \langle \mathbf{A}^*(\mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{y}}), \mathbf{h} \rangle_{\mathcal{X}} + o(\|\mathbf{h}\|_{\mathcal{X}}), \end{aligned}$$

that is

$$(\mathrm{d}F_1)_{\hat{\mathbf{x}}}(\mathbf{h}) = \langle \mathbf{A}^*(\mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{y}}), \mathbf{h} \rangle_{\mathcal{X}}.$$

Moreover,  $F_2$  is linear so  $(\mathrm{d}F_2)_{\mathbf{X}_0}(\mathbf{X}) = F_2(\mathbf{X})$ , hence

$$(\mathrm{d}F)_{\mathbf{X}_0}(\mathbf{X}) = \langle \mathbf{A}^*(\mathbf{A}\hat{\mathbf{x}}_0 - \hat{\mathbf{y}}), \hat{\mathbf{x}} \rangle_{\mathcal{X}} = \left\langle \begin{pmatrix} \mathbf{A}^*(\mathbf{A}\hat{\mathbf{x}}_0 - \hat{\mathbf{y}}) \\ \mathbf{0} \end{pmatrix}, \mathbf{X} \right\rangle_{\mathcal{H}},$$

that is

$$\nabla F(\mathbf{X}_0) = \begin{pmatrix} \mathbf{A}^*(\mathbf{A}\hat{\mathbf{x}}_0 - \hat{\mathbf{y}}) \\ \mathbf{0} \end{pmatrix}.$$

Consequently,

$$\|\nabla F(\mathbf{X}) - \nabla F(\mathbf{X}')\|_{\mathcal{H}} \leq \|\mathbf{A}^*\mathbf{A}(\hat{\mathbf{x}} - \hat{\mathbf{x}}')\|_{\mathcal{X}} \leq \|\mathbf{A}^*\mathbf{A}\| \|\hat{\mathbf{x}} - \hat{\mathbf{x}}'\|_{\mathcal{X}}.$$

We get  $\beta = \|\mathbf{A}^*\mathbf{A}\| = \|\mathbf{A}\|^2$  and Lemma 1 concludes the proof.  $\square$

We now give in the next proposition the adjoints of these operators:

$$\mathbf{T}_{M+1} : (\mathcal{X}_t, \langle \cdot, \cdot \rangle_{\mathcal{X}_t}) \rightarrow (\mathcal{T}_{M+1}, \langle \cdot, \cdot \rangle_{\mathcal{M}}), \quad (56)$$

$$\mathbf{T}'_{H_S} : (\mathcal{X}_t \times \mathcal{Q}_t, \langle \cdot, \cdot \rangle_{\mathcal{X}_t} + \langle \cdot, \cdot \rangle_{\mathcal{Q}_t}) \rightarrow (\mathcal{T}_{H_S+1}, \langle \cdot, \cdot \rangle_{\mathcal{M}}), \quad (57)$$

where the inner products are defined in (19), (20), (33) and (41).

**Proposition 5.** For  $\mathbf{M}^{(1)} \in \mathcal{T}_{H_S+1}$  and  $\mathbf{M}^{(2)} \in \mathcal{T}_{M+1}$  the adjoint operators of (56) and (57) applied to  $\mathbf{M}^{(1)}$  and  $\mathbf{M}^{(2)}$  give the following vectors:

$$\begin{aligned} \mathbf{z}_2 &= \mathbf{T}_{M+1}^* \mathbf{M}^{(2)} \in \mathbb{R} \times \mathbb{C}^M, \\ (\mathbf{z}_1, \mathbf{q}_1) &= \mathbf{T}'_{H_S}{}^* \mathbf{M}^{(1)} \in \mathbb{C}^{H_S} \times (\mathbb{R} \times \mathbb{C}^{H_S-1}), \end{aligned}$$

whose components are:

$$\begin{aligned} \mathbf{z}_2[k] &= \sum_{l=0}^{M-k} \mathbf{M}_{l+k,l}^{(2)}, \quad \forall k = 0, \dots, M, \\ \mathbf{z}_1[k] &= \mathbf{M}_{H_S+1,k}^{(1)}, \quad \mathbf{q}_1[k] = \sum_{l=0}^{H_S-1-k} \mathbf{M}_{l+k,l}^{(1)} + \delta_k \mathbf{M}_{H_S,H_S}^{(1)}, \quad \forall k = 0, \dots, H_S - 1. \end{aligned}$$

*Proof.* See in Appendix C. □

Now we will explicit a bound of the operator norm  $\|\mathbf{L}\|$ .

**Proposition 6.** The norm of the operator  $\mathbf{L} = (\mathbf{L}^{(1)}, \mathbf{L}^{(2)})$ , where

$$\begin{aligned} \mathbf{L}^{(1)}(\mathbf{X}) &= (\mathbf{L}_1^{(1)}(\mathbf{X}), \dots, \mathbf{L}_M^{(1)}(\mathbf{X})), \\ \mathbf{L}^{(2)}(\mathbf{X}) &= (\mathbf{L}_0^{(2)}(\mathbf{X}), \dots, \mathbf{L}_{H_S-1}^{(2)}(\mathbf{X})), \end{aligned}$$

and with  $\mathbf{L}_m^{(1)}$  and  $\mathbf{L}_{n_2}^{(2)}$  defined in (44)-(45), is given by

$$\|\mathbf{L}\|^2 \leq \|\mathbf{L}^{(1)}\|^2 + \|\mathbf{L}^{(2)}\|^2 = \|\mathbf{T}'_{H_S}\|^2 + \|\mathbf{T}_{M+1}\|^2 = (H_S + 1) + (M + 1).$$

*Proof.* See in Appendix D. □

To solve the problem (47), we first propose Algorithm 1, which uses the primal-dual method introduced in [41]. Following [17, Theorem 5.1], we know that the method converges to a solution  $(\tilde{\mathbf{X}}, \tilde{\boldsymbol{\xi}}_0, \dots, \tilde{\boldsymbol{\xi}}_{Q-1})$  of the problem (48), provided the parameters  $\tau > 0$  and  $\sigma > 0$  in Algorithm 1 are such that

$$\frac{1}{\tau} - \sigma \|\mathbf{L}\|^2 > \frac{\beta}{2}. \quad (58)$$

We then choose  $0 < \tau < 2$ ,  $\sigma = (H_S + M + 2)^{-2}(1/\tau - \beta/1.9)$  and  $\rho_n \equiv \rho = 1$ .

We detail below the other terms in lines 3 and 6 of Algorithm 1, involving proximal and adjoint operators computation. For more details on convex analysis, monotone operator theory and proximal splitting methods see [5, 1, 16, 38].

Set  $\bar{x}_0 = \frac{1}{H_S} \sum_{n_2=0}^{H_S-1} \hat{\mathbf{x}}[0, n_2]$ ,  $G = \iota_{\mathcal{B}}$  with  $\mathcal{B}$  defined in (46), we get  $\forall m, n_2$ :

$$\text{prox}_{\tau G}(\hat{\mathbf{x}}, \mathbf{q}) = \begin{cases} \hat{\mathbf{x}}[0, n_2] = \bar{x}_0 & \text{if } \bar{x}_0 \leq c \\ \hat{\mathbf{x}}[0, n_2] = c & \text{otherwise} \\ \mathbf{q}[m, 0] = c & \text{if } \mathbf{q}[m, 0] > c \end{cases}.$$

---

**Algorithm 1** Primal-dual splitting algorithm for (48)

---

**Input:**  $\hat{\mathbf{y}}$  1-D FFT of the blurred and noisy data image  $y$

**Output:**  $\tilde{x}$  solution of the optimization problem (39) under constraints (40)

- 1: Initialize primal and dual variables to zero  $\mathbf{X}_0 = 0, \boldsymbol{\xi}_{i,0} = 0, \forall i \in \llbracket 1, Q \rrbracket$
  - 2: **for**  $n = 1$  **to** Number of iterations **do**
  - 3:  $\mathbf{X}_{n+1} = \text{prox}_{\tau G}(\mathbf{X}_n - \tau \nabla F(\mathbf{X}_n) - \tau \sum_{i=0}^{Q-1} \mathbf{L}_i^* \boldsymbol{\xi}_{i,n}),$
  - 4:  $\mathbf{X}_{n+1} = \rho_n \tilde{\mathbf{X}}_{n+1} - (1 - \rho_n) \mathbf{X}_n,$
  - 5: **for**  $i = 0$  **to**  $Q - 1$  **do**
  - 6:  $\tilde{\boldsymbol{\xi}}_{i,n+1} = \text{prox}_{\sigma H_i^*}(\boldsymbol{\xi}_{i,n} + \sigma \mathbf{L}_i(2\mathbf{X}_{n+1} - \mathbf{X}_n)),$
  - 7:  $\boldsymbol{\xi}_{i,n+1} = \rho_n \tilde{\boldsymbol{\xi}}_{i,n+1} + (1 - \rho_n) \boldsymbol{\xi}_{i,n},$
  - 8: **end for**
  - 9: **end for**
- 

Let  $P_{\mathcal{C}}$  be the projection operator onto the cone of positive matrix  $\mathcal{C}$ , by Moreau identity [1]:

$$\text{prox}_{\sigma H_i^*}(\mathbf{M}) = \mathbf{M} - \sigma \text{prox}_{\frac{H_i}{\sigma}}\left(\frac{1}{\sigma} \mathbf{M}\right) = \mathbf{M} - P_{\mathcal{C}}(\mathbf{M}).$$

Finally, we need to compute on line 3 the adjoint operators  $\mathbf{L}_i^*$  where the  $\mathbf{L}_i$  are defined in (44)-(45). The dual variables  $(\boldsymbol{\xi}_{i,n})_i$  in Algorithm 1 refer to Hermitian-Toeplitz matrices  $\mathbf{M}_m^{(1)} \in \mathcal{T}_{H_S+1}$  or  $\mathbf{M}_{n_2}^{(2)} \in \mathcal{T}_{M+1}$ . By definition, the adjoint operator are the images  $(\mathbf{z}_m^{(1)}, \mathbf{q}_m^{(1)}) = \mathbf{L}_m^{(1)*} \mathbf{M}_m^{(1)}$  and  $(\mathbf{z}_{n_2}^{(2)}, \mathbf{q}_{n_2}^{(2)}) = \mathbf{L}_{n_2}^{(2)*} \mathbf{M}_{n_2}^{(2)}$ . According to the definitions (44)-(45), for a primal variable  $\mathbf{X} = (\hat{\mathbf{x}}, \mathbf{q})$  the operators  $\mathbf{L}_m^{(1)}$  and  $\mathbf{L}_{n_2}^{(2)}$  respectively act on the  $m$ -th columns of the images  $(\hat{\mathbf{x}}, \mathbf{q})$  and on the  $n_2$ -th row of the image  $\hat{\mathbf{x}}$  only, so we can easily see concerning the adjoints that  $\mathbf{q}_{n_2}^{(2)} = \mathbf{0}$  and  $\mathbf{z}_{n_2}^{(2)}$  (resp.  $\mathbf{z}_m^{(1)}, \mathbf{q}_m^{(1)}$ ) is null except at the corresponding row  $n_2$  (resp. column  $m$ ) where

$$\mathbf{z}_{n_2}^{(2)}[:, n_2] = \mathbf{T}_{M+1}^* \mathbf{M}_{n_2}^{(2)}, \quad (59)$$

$$(\mathbf{z}_m^{(1)}[m, :], \mathbf{q}_m^{(1)}[m, :]) = \mathbf{T}_{H_S}^{\prime*} \mathbf{M}_m^{(1)}, \quad (60)$$

with the expression of the adjoint operators  $\mathbf{T}_{M+1}^*$  and  $\mathbf{T}_{H_S}^{\prime*}$  given in Proposition 5. Thus, the operations onto  $\mathbf{X}_n = (\hat{\mathbf{x}}_n, \mathbf{q}_n)$  before applying  $\text{prox}_{\tau G}$  on line 3, consists in a gradient descent step  $\mathbf{X}_n - \tau \nabla F(\mathbf{X}_n)$ , following by an update of all its rows and columns caused by the terms  $-\tau \sum_{i=0}^{Q-1} \mathbf{L}_i^* \boldsymbol{\xi}_{i,n}$ , whose expressions are provided by (59) and (60).

## 4.2 Second Algorithm Design

Notice that in Algorithm 1,  $\tau$  must be smaller than  $2/\beta$ , which is a limitation in terms of convergence speed. To overcome this issue, we subsequently developed a second algorithm, similar to Algorithm 1, but with the data fidelity term  $\|\mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{y}}\|_{\mathcal{Y}}$  activated through its proximity operator, instead of its gradient. We consider the optimization problem as a over-relaxed version of the Chambolle–Pock algorithm [12]:

$$\mathbf{X}^* = \arg \min_{\mathbf{X} \in \mathcal{H}} \{G(\mathbf{X}) + \mathbf{H}(\mathbf{L}(\mathbf{X}))\}, \quad (61)$$

with now  $G = \frac{1}{2} \|\mathbf{A} \cdot -\hat{\mathbf{y}}\|_y^2$  which is proximable,  $\mathbf{H}\mathbf{x} = \sum_{i=0}^Q H_i x_i$  with  $H_i = \iota_C$  for  $i < Q$ , where  $L_i = L_i^{(2)}$  when  $0 \leq i \leq H_S - 1$ ,  $L_i = L_{i-H_S+1}^{(1)}$  when  $H_S \leq i \leq Q - 1$ , and  $H_Q = \iota_B$  with  $L_Q = \text{Id}$ . So now,  $\|\mathbf{L}\|^2 \leq H_S + M + 3$ .

Let  $\tau > 0$  and  $\sigma > 0$  such that  $\tau\sigma\|\mathbf{L}\|^2 = 1$ , then the primal-dual [Algorithm 2](#), with  $F = 0$  and weights  $\rho_n \equiv \rho = 1.9$ , which is an over-relaxed version of the Chambolle–Pock algorithm, converges to a solution  $(\tilde{\mathbf{X}}, \tilde{\boldsymbol{\xi}}_0, \dots, \tilde{\boldsymbol{\xi}}_{Q-1})$  of the problem (48) [17, Theorem 5.1].

---

**Algorithm 2** Primal-dual splitting algorithm for (61)

---

**Input:**  $\hat{\mathbf{y}}$  1-D FFT of the blurred and noisy data image  $y$

**Output:**  $\hat{\mathbf{x}}$  solution of the optimization problem (39) under constraints (40)

- 1: Initialize all primal and dual variables to zero
  - 2: **for**  $n = 1$  **to** Number of iterations **do**
  - 3:    $\tilde{\mathbf{X}}_{n+1} = \text{prox}_{\tau G}(\mathbf{X}_n - \tau \sum_{i=0}^{Q-1} L_i^* \boldsymbol{\xi}_{i,n})$ ,
  - 4:    $\mathbf{X}_{n+1} = \rho_n \tilde{\mathbf{X}}_{n+1} - (1 - \rho_n) \mathbf{X}_n$
  - 5:   **for**  $i = 0$  **to**  $Q - 1$  **do**
  - 6:      $\tilde{\boldsymbol{\xi}}_{i,n+1} = \text{prox}_{\sigma H_i^*}(\boldsymbol{\xi}_{i,n} + \sigma L_i(2\mathbf{X}_{n+1} - \mathbf{X}_n))$ ,
  - 7:      $\boldsymbol{\xi}_{i,n+1} = \rho_n \tilde{\boldsymbol{\xi}}_{i,n+1} + (1 - \rho_n) \boldsymbol{\xi}_{i,n}$
  - 8:   **end for**
  - 9: **end for**
- 

The [Algorithm 2](#) requires to compute  $\text{prox}_{\tau G}$ . Since we have

$$\begin{aligned} \mathbf{p} = \text{prox}_{\tau G}(\hat{\mathbf{x}}) &\Leftrightarrow \hat{\mathbf{x}} - \mathbf{p} = \nabla(\tau G)(\mathbf{p}), \\ &\Leftrightarrow \hat{\mathbf{x}} - \mathbf{p} = \tau \mathbf{A}^*(\mathbf{A}\mathbf{p} - \hat{\mathbf{y}}), \\ &\Leftrightarrow \hat{\mathbf{x}} + \tau \mathbf{A}^* \hat{\mathbf{y}} = (\mathbf{I} + \tau \mathbf{A}^* \mathbf{A}) \mathbf{p}, \end{aligned} \quad (62)$$

then the proximal operator has the following expression:

$$\text{prox}_{\tau G}(\hat{\mathbf{x}}) = (\mathbf{I} + \tau \mathbf{A}^* \mathbf{A})^{-1}(\hat{\mathbf{x}} + \tau \mathbf{A}^* \hat{\mathbf{y}}),$$

for which we propose below two ways of computing the inverse.

We proved in [Lemma 2](#) that  $\mathbf{A}^* \hat{\mathbf{y}} = \tilde{\mathbf{H}}^* \hat{\mathbf{y}} \hat{\mathbf{G}}^*$  and then

$$(\mathbf{I} + \tau \mathbf{A}^* \mathbf{A}) \hat{\mathbf{x}} = \hat{\mathbf{x}} + \mathbf{P} \hat{\mathbf{x}} \mathbf{Q}, \quad \mathbf{P} = \tau \tilde{\mathbf{H}}^* \tilde{\mathbf{H}}, \quad \mathbf{Q} = \hat{\mathbf{G}} \hat{\mathbf{G}}^*.$$

The square matrices  $\mathbf{P}$  and  $\mathbf{Q}$  are of size  $p = H_S$  and  $q = M + 1$ . We have to solve  $(\mathbf{I} + \tau \mathbf{A}^* \mathbf{A}) \hat{\mathbf{x}} = \mathbf{z}$ ; that is,  $\hat{\mathbf{x}} + \mathbf{P} \hat{\mathbf{x}} \mathbf{Q} = \mathbf{z}$ . This kind of system can be solved by the mean of the Kronecker Product as:

$$\hat{\mathbf{x}} + \mathbf{P} \hat{\mathbf{x}} \mathbf{Q} = \mathbf{z} \quad \Leftrightarrow \quad (\mathbf{I}_{pq,pq} + \mathbf{Q} \otimes \mathbf{P}^T) \text{Vec}(\hat{\mathbf{x}}) = \text{Vec}(\mathbf{z}).$$

where  $\text{Vec}(\hat{\mathbf{x}})$  denotes the vectorization of the matrix  $\hat{\mathbf{x}}$  formed by stacking the columns of  $\hat{\mathbf{x}}$  into a single column vector, and  $\mathbf{I}_{pq,pq} + \mathbf{Q} \otimes \mathbf{P}^T$  is a matrix of size  $pq \times pq$  which can be inverted, giving access to  $\text{Vec}(\hat{\mathbf{x}})$  and then to  $\hat{\mathbf{x}}$ . Finally, the operator  $\text{prox}_{\tau G}$  is nothing more than a large matrix-vector product.

Another option consists in working on the columns  $\hat{\mathbf{x}}_m$  of  $\hat{\mathbf{x}}$ , since the operator  $\mathbf{A}$  acts on them.

$$(\mathbf{I} + \tau \mathbf{A}^* \mathbf{A}) \hat{\mathbf{x}} = \mathbf{z} \quad \Leftrightarrow \quad (\mathbf{I} + |\hat{\mathbf{g}}_m|^2 \mathbf{P}) \hat{\mathbf{x}}_m = \mathbf{z}_m, \quad \forall m = 0, \dots, M.$$

This time, the operator  $\text{prox}_{\tau G}$  involves performing  $M + 1$  matrix-vector products of size  $p \times p$ , which appears to be more efficient in practice.

### 4.3 Extended Problem Formulation

We now consider a data image  $\mathbf{b}^\sharp$  containing lines with no angle restriction, which extends the previous case by relaxing the assumption made in [subsection 2.1](#). We can decompose this image into the sum of two images  $\mathbf{b}^\sharp = \mathbf{b}_1^\sharp + \mathbf{b}_2^\sharp$ , with  $\mathbf{b}_1^\sharp$  (resp.  $\mathbf{b}_2^\sharp$ ) containing vertical (resp. horizontal) lines. We can also define  $\widehat{\mathbf{x}}_1^\sharp$  of size  $(M+1) \times H_S$  and  $\widehat{\mathbf{x}}_2^\sharp$  of size  $W_S \times (P+1)$  with  $W_S = W + 2S$  and  $P = (H-1)/2$  such as  $\mathbf{A}\widehat{\mathbf{x}}_1^\sharp = \widehat{\mathbf{b}}_1^\sharp$  and  $\tilde{\mathbf{A}}\widehat{\mathbf{x}}_2^\sharp = \widehat{\mathbf{b}}_2^\sharp$ , where  $\mathbf{g}_2 = [\mathbf{0}_{P-S}, \mathbf{h}, \mathbf{0}_{P-S}]$  and  $\tilde{\mathbf{A}}$  denotes the operator which multiplies each row vector  $\widehat{\mathbf{x}}_2^\sharp[:, n_2]$  by the corresponding Fourier coefficient  $\widehat{\mathbf{g}}_2[n_2]$  and convolves it with the filter  $\mathbf{h}$ ; that is,  $\tilde{\mathbf{A}}\widehat{\mathbf{x}}_2 = (\widehat{\mathbf{G}}_2\widehat{\mathbf{x}}_2) * \mathbf{h}$  with  $\widehat{\mathbf{G}}_2 = \text{diag}(\widehat{\mathbf{g}}_2[P+1], \dots, \widehat{\mathbf{g}}_2[H])$ . We finally define the Hermitian symmetry operator  $\mathbf{S}_1$  (resp.  $\mathbf{S}_2$ ) which from each row  $\mathbf{v} = [v_0, v_1, \dots, v_M]$  (resp. column  $\mathbf{v} = [v_P, \dots, v_0, \dots, v_P]$ ) associates the symmetric extension  $[v_M^*, \dots, v_0, \dots, v_M]$  (resp.  $[v_P^*, \dots, v_0, \dots, v_P]$ ). Let  $\mathbf{X}_1 = (\widehat{\mathbf{x}}_1, \mathbf{q}_1)$  and  $\mathbf{X}_2 = (\widehat{\mathbf{x}}_2, \mathbf{q}_2)$  be the optimization variables, lying of spaces  $\mathcal{H}_1 = \mathcal{X}_1 \times \mathcal{Q}_1$  and  $\mathcal{H}_2 = \mathcal{X}_2 \times \mathcal{Q}_2$ . Let  $\mathcal{H} = \mathcal{H}_1 \times \mathcal{H}_2$ ,  $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$  and  $\mathcal{Q} = \mathcal{Q}_1 \times \mathcal{Q}_2$ . Then, the data fidelity term is now:

$$F(\mathbf{X}_1, \mathbf{X}_2) = \frac{1}{2} \|\mathcal{F}_1^{-1} \mathbf{S}_1 \mathbf{A} \widehat{\mathbf{x}}_1 + \mathcal{F}_2^{-1} \mathbf{S}_2 \tilde{\mathbf{A}} \widehat{\mathbf{x}}_2 - \mathbf{y}\|_{\mathbb{F}}^2 = \frac{1}{2} \|\mathbf{A}_1 \widehat{\mathbf{x}}_1 + \mathbf{A}_2 \widehat{\mathbf{x}}_2 - \mathbf{y}\|_{\mathbb{F}}^2,$$

with  $\mathbf{A}_1 = \mathcal{F}_1^{-1} \mathbf{S}_1 \mathbf{A}$ ,  $\mathbf{A}_2 = \mathcal{F}_2^{-1} \mathbf{S}_2 \tilde{\mathbf{A}}$ , where  $\mathcal{F}_1$  (resp.  $\mathcal{F}_2$ ) is the Fourier transform with respect to the rows (resp. columns) and  $\|\cdot\|_{\mathbb{F}}$  is the Frobenius norm.

**Proposition 7.** *The gradient of  $F$  is*

$$\nabla F(\mathbf{X}_1, \mathbf{X}_2) = \frac{1}{2} \begin{pmatrix} \mathbf{A}_1^* (\mathbf{A}_1 \widehat{\mathbf{x}}_1 + \mathbf{A}_2 \widehat{\mathbf{x}}_2 - \mathbf{y}) \\ \mathbf{A}_2^* (\mathbf{A}_1 \widehat{\mathbf{x}}_1 + \mathbf{A}_2 \widehat{\mathbf{x}}_2 - \mathbf{y}) \end{pmatrix}, \quad (63)$$

which is Lipschitz-continuous of Lipschitz constant  $\beta = \frac{1}{\min(W, H)}$ .

*Proof.* Let us compute the differential function of  $F$ :

$$\begin{aligned} & F(\mathbf{X}_1 + \mathbf{h}_1, \mathbf{X}_2 + \mathbf{h}_2) \\ &= \frac{1}{2} \langle \mathbf{A}_1 \widehat{\mathbf{x}}_1 + \mathbf{A}_2 \widehat{\mathbf{x}}_2 + \mathbf{A}_1 \mathbf{h}_1 + \mathbf{A}_2 \mathbf{h}_2 - \mathbf{y}, \mathbf{A}_1 \widehat{\mathbf{x}}_1 + \mathbf{A}_2 \widehat{\mathbf{x}}_2 + \mathbf{A}_1 \mathbf{h}_1 + \mathbf{A}_2 \mathbf{h}_2 - \mathbf{y} \rangle_{\mathcal{M}}, \\ &= F(\mathbf{X}_1, \mathbf{X}_2) + \frac{1}{2} \langle \mathbf{A}_1 \mathbf{h}_1, \mathbf{A}_1 \widehat{\mathbf{x}}_1 + \mathbf{A}_2 \widehat{\mathbf{x}}_2 - \mathbf{y} \rangle_{\mathcal{M}} + \frac{1}{2} \langle \mathbf{A}_1 \mathbf{h}_1, \mathbf{A}_2 \mathbf{h}_2 \rangle_{\mathcal{M}} + \frac{1}{2} \langle \mathbf{A}_1 \mathbf{h}_1, \mathbf{A}_1 \mathbf{h}_1 \rangle_{\mathcal{M}} \\ &\quad + \frac{1}{2} \langle \mathbf{A}_2 \mathbf{h}_2, \mathbf{A}_1 \widehat{\mathbf{x}}_1 + \mathbf{A}_2 \widehat{\mathbf{x}}_2 - \mathbf{y} \rangle_{\mathcal{M}} + \frac{1}{2} \langle \mathbf{A}_2 \mathbf{h}_2, \mathbf{A}_1 \mathbf{h}_1 \rangle_{\mathcal{M}} + \frac{1}{2} \langle \mathbf{A}_2 \mathbf{h}_2, \mathbf{A}_2 \mathbf{h}_2 \rangle_{\mathcal{M}}, \end{aligned}$$

and

$$|\langle \mathbf{A}_1 \mathbf{h}_1, \mathbf{A}_2 \mathbf{h}_2 \rangle_{\mathcal{M}}| \leq \|\mathbf{A}_1\| \|\mathbf{A}_2\| \|\mathbf{h}_1\|_{\mathcal{H}_1} \|\mathbf{h}_2\|_{\mathcal{H}_2} = o(\|(\mathbf{h}_1, \mathbf{h}_2)\|_{\mathcal{H}}),$$

and we deduce that

$$\nabla F(\mathbf{X}_1, \mathbf{X}_2) = \frac{1}{2} \begin{pmatrix} \mathbf{A}_1^* (\mathbf{A}_1 \widehat{\mathbf{x}}_1 + \mathbf{A}_2 \widehat{\mathbf{x}}_2 - \mathbf{y}) \\ \mathbf{A}_2^* (\mathbf{A}_1 \widehat{\mathbf{x}}_1 + \mathbf{A}_2 \widehat{\mathbf{x}}_2 - \mathbf{y}) \end{pmatrix}.$$

The adjoints are  $\tilde{\mathbf{A}}^* \mathbf{z} = (\widehat{\mathbf{G}}_2^* \mathbf{z}) * \bar{\mathbf{h}}'$ ,  $(\mathcal{F}_1^{-1})^* = \frac{1}{W} \mathcal{F}_1$  and  $(\mathcal{F}_2^{-1})^* = \frac{1}{H} \mathcal{F}_2$ , and  $\mathbf{S}_1^*(v_{-M}, \dots, v_0, \dots, v_M) = (v_0, \dots, v_M)$  and  $\mathbf{S}_2^*(v_{-P}, \dots, v_0, \dots, v_P) = (v_0, \dots, v_P)$ .

Let us determine the Lipschitz constant of the gradient  $\nabla F$ :

$$\begin{aligned} \|\nabla F(\mathbf{X}_1, \mathbf{X}_2) - \nabla F(\mathbf{X}'_1, \mathbf{X}'_2)\|_{\mathcal{X}}^2 &= \frac{1}{4} \|\mathbf{A}_1^* (\mathbf{A}_1 (\widehat{\mathbf{x}}_1 - \widehat{\mathbf{x}}'_1) + \mathbf{A}_2 (\widehat{\mathbf{x}}_2 - \widehat{\mathbf{x}}'_2))\|_{\mathcal{X}_1}^2 \\ &\quad + \frac{1}{4} \|\mathbf{A}_2^* (\mathbf{A}_1 (\widehat{\mathbf{x}}_1 - \widehat{\mathbf{x}}'_1) + \mathbf{A}_2 (\widehat{\mathbf{x}}_2 - \widehat{\mathbf{x}}'_2))\|_{\mathcal{X}_2}^2. \end{aligned}$$

We are looking for a majoration of each term. We treat the first one  $C_1$ , the second  $C_2$  is obtained in the same manner. Using the inequality  $(a + b)^2 \leq 2a^2 + 2b^2$ :

$$\begin{aligned} C_1 &\leq \frac{1}{4} (\|\mathbf{A}_1^* \mathbf{A}_1\| \|\widehat{\mathbf{x}}_1 - \widehat{\mathbf{x}}_1'\|_{\mathcal{X}_1} + \|\mathbf{A}_1^* \mathbf{A}_2\| \|\widehat{\mathbf{x}}_2 - \widehat{\mathbf{x}}_2'\|_{\mathcal{X}_2})^2, \\ &\leq \frac{1}{2} \|\mathbf{A}_1^* \mathbf{A}_1\|^2 \|\widehat{\mathbf{x}}_1 - \widehat{\mathbf{x}}_1'\|_{\mathcal{X}_1}^2 + \frac{1}{2} \|\mathbf{A}_1^* \mathbf{A}_2\|^2 \|\widehat{\mathbf{x}}_2 - \widehat{\mathbf{x}}_2'\|_{\mathcal{X}_2}^2. \end{aligned}$$

We have  $\|\mathbf{A}\| = \|\mathbf{A}^*\| = 1$ ,  $\|\mathbf{S}_i \widehat{\mathbf{x}}_i\|_{\mathbb{F}} = \|\widehat{\mathbf{x}}_i\|_{\mathcal{X}}$ , for  $i \in \{1, 2\}$ ; that is,  $\|\mathbf{S}_i\| = 1$ , and  $\|\mathcal{F}_i^{-1} \mathbf{v}\|_2^2 = \frac{1}{N^2} \|\mathbf{v}\|_2^2$ ; that is,  $\|\mathcal{F}_i^{-1}\| = \frac{1}{N}$ . Hence,  $\|\mathbf{A}_1\| \leq \frac{1}{W}$ ,  $\|\mathbf{A}_1^*\| \leq 1$ ,  $\|\mathbf{A}_2\| \leq \frac{1}{H}$  and  $\|\mathbf{A}_2^*\| \leq 1$ . Consequently, we get

$$C_1 \leq \frac{1}{2W^2} \|\widehat{\mathbf{x}}_1 - \widehat{\mathbf{x}}_1'\|_{\mathcal{X}_1}^2 + \frac{1}{2H^2} \|\widehat{\mathbf{x}}_2 - \widehat{\mathbf{x}}_2'\|_{\mathcal{X}_2}^2,$$

and exactly the same majoration for  $C_2$ . Thus, we have

$$\|\nabla F(\mathbf{X}_1, \mathbf{X}_2) - \nabla F(\mathbf{X}'_1, \mathbf{X}'_2)\|_{\mathcal{X}}^2 \leq \beta^2 (\|\mathbf{X}_1 - \mathbf{X}'_1\|_{\mathcal{H}_1}^2 + \|\mathbf{X}_2 - \mathbf{X}'_2\|_{\mathcal{H}_2}^2),$$

with

$$\beta = \frac{1}{\min(W, H)}.$$

□

The image  $\widehat{\mathbf{x}}_1^\sharp$  keeps the same kind of constraints as in the [Algorithm 1](#), which act similarly on the image  $\widehat{\mathbf{x}}_2^\sharp$  in a rotated way; that is, we define

$$\mathbf{L}_m^{(3)}(\mathbf{X}_2) = \mathbf{T}_{P+1}(\widehat{\mathbf{x}}_2[m, :]), \quad (64)$$

$$\mathbf{L}_{n_2}^{(4)}(\mathbf{X}_2) = \mathbf{T}'_{W_S}(\text{fliplr}(\widehat{\mathbf{x}}_2[:, n_2]), \text{fliplr}(\mathbf{q}_2[:, n_2])), \quad (65)$$

where `fliplr` performs a flip from left to right on each row of the matrix.

The boundary constraints on  $\widehat{\mathbf{x}}_1$  and  $\widehat{\mathbf{x}}_2$  are respectively given by:

$$\mathcal{B}_1 = \left\{ (\widehat{\mathbf{x}}_1, \mathbf{q}_1) \in \mathcal{H}_1 : \widehat{\mathbf{x}}_1[0, n_2] = \widehat{\mathbf{x}}_1[0, 0] \leq c_1, \mathbf{q}_1[m, 0] \leq c_1 \right\}, \quad (66)$$

$$\mathcal{B}_2 = \left\{ (\widehat{\mathbf{x}}_2, \mathbf{q}_2) \in \mathcal{H}_2 : \widehat{\mathbf{x}}_2[m, 0] = \widehat{\mathbf{x}}_2[0, 0] \leq c_2, \mathbf{q}_2[P, n_2] \leq c_2 \right\}. \quad (67)$$

Likewise, the inner product on spaces  $\mathcal{X}_2$  and  $\mathcal{Q}_2$  are:

$$\langle \mathbf{z}_1, \mathbf{z}_2 \rangle_{\mathcal{X}_2} = \sum_{m=0}^{W_S-1} \mathbf{z}_1[m, 0] \mathbf{z}_2[m, 0] + 2\text{Re} \left( \sum_{n_2=1}^P \sum_{m=0}^{W_S-1} \mathbf{z}_1[m, n_2] \mathbf{z}_2[m, n_2]^* \right), \quad (68)$$

$$\langle \mathbf{z}_1, \mathbf{z}_2 \rangle_{\mathcal{Q}_2} = 2\text{Re} \left( \sum_{n_2=0}^P \sum_{m=0}^{W_S-1} \mathbf{z}_1[m, n_2] \mathbf{z}_2[m, n_2]^* \right). \quad (69)$$

and so the adjoint of the operators remain the same.

As previously, we define  $\mathbf{L}^{(3)}(\mathbf{X}_2) = (\mathbf{L}_0^{(3)}(\mathbf{X}_2), \dots, \mathbf{L}_{W_S-1}^{(3)}(\mathbf{X}_2))$ , also  $\mathbf{L}^{(4)}(\mathbf{X}_2) = (\mathbf{L}_0^{(4)}(\mathbf{X}_2), \dots, \mathbf{L}_P^{(4)}(\mathbf{X}_2))$  and  $\mathbf{L} = (\mathbf{L}^{(1)}, \mathbf{L}^{(2)}, \mathbf{L}^{(3)}, \mathbf{L}^{(4)})$ . Again, an easy computation leads to

$$\begin{aligned} \|\mathbf{L}\|^2 &\leq \|\mathbf{L}^{(1)}\|^2 + \|\mathbf{L}^{(2)}\|^2 + \|\mathbf{L}^{(3)}\|^2 + \|\mathbf{L}^{(4)}\|^2, \\ &\leq (H_S - 1) + (M + 1) + (P + 1) + (W_S - 1). \end{aligned}$$

Finally, we have

$$\begin{aligned}
(\tilde{\mathbf{X}}_1, \tilde{\mathbf{X}}_2) = \arg \min_{(\mathbf{X}_1, \mathbf{X}_2) \in \mathcal{H}} & \left\{ \frac{1}{2} \|\mathbf{A}_1 \hat{\mathbf{x}}_1 + \mathbf{A}_2 \hat{\mathbf{x}}_2 - \mathbf{y}\|_{\mathbb{F}}^2 \right. \\
& + \iota_{\mathcal{B}_1}(\mathbf{X}_1) + \sum_{m=0}^M \iota_{\mathcal{L}}(\mathbf{L}_m^{(1)}(\mathbf{X}_1)) + \sum_{n_2=0}^{H_S-1} \iota_{\mathcal{L}}(\mathbf{L}_{n_2}^{(2)}(\mathbf{X}_1)) \\
& \left. + \iota_{\mathcal{B}_2}(\mathbf{X}_2) + \sum_{m=0}^{W_S-1} \iota_{\mathcal{L}}(\mathbf{L}_m^{(3)}(\mathbf{X}_2)) + \sum_{n_2=0}^P \iota_{\mathcal{L}}(\mathbf{L}_{n_2}^{(4)}(\mathbf{X}_2)) \right\}. \quad (70)
\end{aligned}$$

#### 4.4 Inpainting problems

We now consider the case in which a binary mask  $\mathbf{M}$  is applied on the data image, as in Figure 4. More precisely, the result  $\mathbf{M} \cdot \mathbf{x}$  is an element-wise multiplication of the matrix  $\mathbf{x}$  with the binary matrix  $\mathbf{M}$ , whose zero coefficients are the indices of the pixels unavailable to observation. We have  $\mathbf{M}^* = \mathbf{M}$ . The data fidelity term becomes  $F(\mathbf{X}) = \frac{1}{2} \|\mathbf{M} \mathcal{F}_1^{-1} \mathbf{S}_1 \mathbf{A} \hat{\mathbf{x}} - \mathbf{y}\|_{\mathbb{F}}^2$ , whose gradient can be expressed as previously, with  $\beta = 1/W$  (since  $\|\mathcal{F}_1^{-1}\| = 1/W$  and  $\|\mathbf{M}\| = 1$ ). The constraints remain the same as in (40), and the method is also easily transposable to the extended setting of subsection 4.3.

At this stage, we completed the first step; that is, we are able to restore the image  $\hat{\mathbf{x}}^\sharp$  from the degraded image  $\mathbf{y}$ . From now, we can for instance reduce the blur by taking other filters  $\mathbf{g}_r$  and  $\mathbf{h}_r$  with a smaller spread, and visualize the resulting image  $\mathbf{b}_r$  passing the solution  $\hat{\mathbf{x}}$  through this new blur operator  $\mathbf{A}_r$ ; that is,  $\hat{\mathbf{b}}_r = \mathbf{A}_r \hat{\mathbf{x}}$ .

## 5 Recovering Line Parameters by a Prony Method

In this section, we present the method that underlies the second step of this work (see Figure 4), namely the estimation of the line parameters, which is related to the spectral estimation field. We now focus on estimating the parameters  $(\theta_k, \alpha_k, \eta_k)$ , which characterize the  $K$  lines, from the solution of the minimization problem  $\hat{\mathbf{x}}$ , symmetrized to  $m = -M, \dots, -1$  beforehand. This requires the use of a classical spectral estimation method [50, 51]. The recovering procedure hereafter, based on [45], is an extended method of the famous Prony method [42].

Let us sketch this Prony-like method [45], which is based on an annihilating property [3]. Let  $\mathbf{z} = (z_0, \dots, z_{|I|-1})$  be a complex vector, whose components are:

$$z_i = \sum_{k=1}^K d_k (e^{j2\pi f_k})^i, \quad \forall i = 0, \dots, |I| - 1, \quad (71)$$

with  $d_k \in \mathbb{C}$ ,  $f_k \in [-1/2, 1/2)$ . Let  $\zeta_k = e^{j2\pi f_k}$ , we introduce the annihilating polynomial  $H(\zeta) = \prod_{l=1}^K (\zeta - \zeta_l) = \sum_{l=0}^K h_l \zeta^{K-l}$  with  $h_0 = 1$ . Then, we can notice

that for all  $r = K, \dots, |I| - 1$ :

$$\sum_{l=0}^K h_l z_{r-l} = \sum_{l=0}^K h_l \left( \sum_{k=1}^K d_k \zeta_k^{r-l} \right) = \sum_{k=1}^K d_k \zeta_k^{r-K} \underbrace{\left( \sum_{l=0}^K h_l \zeta_k^{K-l} \right)}_{H(\zeta_k)=0} = 0. \quad (72)$$

Rearranging the elements  $z_i$  in a Toeplitz matrix  $\mathbf{P}_K(\mathbf{z})$  of size  $(|I| - K) \times (K + 1)$  and rank  $K$  as follows:

$$\mathbf{P}_K(\mathbf{z}) = \begin{pmatrix} z_K & \cdots & z_0 \\ \vdots & \ddots & \vdots \\ z_{|I|-1} & \cdots & z_{|I|-K-1} \end{pmatrix},$$

(72) can be written with  $\mathbf{h} = (h_0, \dots, h_K)$  as:

$$(\mathbf{z} * \mathbf{h})(r) = 0, \quad \forall r = K, \dots, |I| - 1 \iff \mathbf{P}_K \mathbf{h} = \mathbf{0}.$$

Consequently, the method consists in finding a right singular vector of the matrix  $\mathbf{P}_K$  associated to the singular value zero, which leads to vector  $\mathbf{h} = (h_0, \dots, h_K)$  and then to the polynomial  $H(\zeta) = \sum_{l=0}^K h_l \zeta^{K-l}$ , whose roots are the searched complex  $\zeta_k = e^{j2\pi f_k}$  and then  $f_k = \arg(\zeta_k)/(2\pi)$ . The complex amplitudes can be retrieved as well by written (71) matricially  $\mathbf{z} = \mathbf{U}\mathbf{d}$  where  $\mathbf{d} = (d_1, \dots, d_K)$  and the following matrix  $\mathbf{U}$  of size  $|I| \times K$ :

$$\mathbf{U} = (\mathbf{a}(f_1) \quad \cdots \quad \mathbf{a}(f_K)) = \begin{pmatrix} 1 & \cdots & 1 \\ e^{-j2\pi f_1} & \cdots & e^{-j2\pi f_K} \\ e^{-j2\pi f_1 \times 2} & \cdots & e^{-j2\pi f_K \times 2} \\ \vdots & \vdots & \vdots \\ e^{-j2\pi f_1(|I|-1)} & \cdots & e^{-j2\pi f_K(|I|-1)} \end{pmatrix}.$$

Finally, we recover the amplitudes by a least square approximation:

$$\mathbf{d} = (\mathbf{U}^* \mathbf{U})^{-1} \mathbf{U}^* \mathbf{z}.$$

Applied on the columns of  $\hat{\mathbf{x}}^\sharp$ , the procedure becomes the following:

### Extraction of line parameters (Procedure 1)

– From  $m = 1, \dots, M$ ,

1. Compute  $\tilde{f}_{m,k} = \arg(\zeta_{m,k})/(2\pi)$ , where  $(\zeta_{m,k})_k$  are roots of the polynomial  $\sum_{k=0}^K h_{m,k} \zeta^k$  with  $\mathbf{h}_m = [h_{m,0}, \dots, h_{m,K}]^\top$  being the right singular vector of  $\mathbf{P}_K(\hat{\mathbf{x}}[m, :])$  with  $I = \{0, \dots, H_S - 1\}$ . It corresponds to the singular value zero (the smallest value in practice).
2. Form the matrix  $\tilde{\mathbf{U}}_m = [\mathbf{a}(\tilde{f}_{m,1}) \cdots \mathbf{a}(\tilde{f}_{m,K})]$ , and compute the complex amplitudes  $\tilde{\mathbf{d}}_m = [\tilde{d}_{m,1}, \dots, \tilde{d}_{m,K}]^\top$  by solving the least-squares linear system  $\tilde{\mathbf{U}}_m^* \tilde{\mathbf{U}}_m \tilde{\mathbf{d}}_m = \tilde{\mathbf{U}}_m^* \hat{\mathbf{x}}[m, :]$ .
3. Compute  $\tilde{\theta}_{m,k} = \arctan(W \tilde{f}_{m,k}/m)$  from (36).
4. Compute  $\tilde{\alpha}_{m,k} = |\tilde{d}_{m,k}| \cos \tilde{\theta}_{m,k}$ .

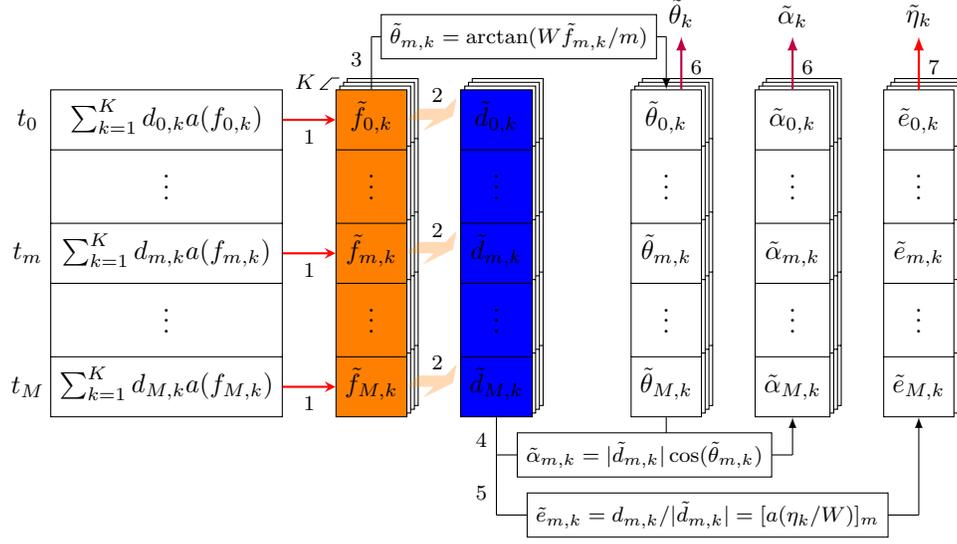


Figure 6: Parameter extraction procedure

5. Compute  $\tilde{e}_{m,k} = \tilde{d}_{m,k}/|\tilde{d}_{m,k}|$ .
- For  $k = 1, \dots, K$ 
  6. Sort the  $\tilde{f}_{m,k}$  with respect to  $k$ , and apply this permutation on the other array. Compute the mean of all estimated angles  $\tilde{\theta}_k = \frac{1}{M} \sum_{m=1}^M \tilde{\theta}_{m,k}$  and amplitudes  $\tilde{\alpha}_k = \frac{1}{M} \sum_{m=1}^M \tilde{\alpha}_{m,k}$ .
  7. Compute the frequency  $\tilde{\nu}_k$  as previously from  $\mathbf{P}_K(\tilde{\mathbf{e}}_k)$  with  $\tilde{\mathbf{e}}_k = (\tilde{e}_{m,k})_m$  and  $I = \{-M, \dots, M\}$ , and then the horizontal offset  $\tilde{\eta}_k = W\tilde{\nu}_k/(2\pi)$ .

**Remark.** The point 6. and 7. are possible to the extent that the sorting frequencies  $\tilde{f}_{m,k}$  are always related to the same angles  $\theta_1 \leq \dots \leq \theta_k \leq \dots \leq \theta_K$  for all  $m$ , which allows us to compute the mean according to  $m$ . It would not be possible to do the same with  $\tilde{f}_{n_2,k} = (\tan(\tilde{\theta}_k)n_2 - \tilde{\eta}_k)/W$  because the affine relation does not preserve the order, for instance one can find  $n$  and  $n'$  such that  $\tilde{f}_{n,k_1} \leq \tilde{f}_{n,k_2}$  and  $\tilde{f}_{n',k_1} \geq \tilde{f}_{n',k_2}$ .

Thus, the trick is to perform the Prony method on the vectors  $\tilde{\mathbf{e}}_k$  again, with this aim to preserve the order. Indeed, if we would chosen instead of step 6. to perform the Prony method on the rows, in order to extract the frequencies  $\tilde{f}_{n_2,k} = (\tan(\tilde{\theta}_k)n_2 - \tilde{\eta}_k)/W$ , then on one hand they are not uniquely determined since they belong to an interval of length greater than one  $\tilde{f}_{n_2,k} \in [-(H_S + M)/W, (H_S + M)/W]$ , and on the other hand we would lose the correspondence between the  $\tilde{f}_{n_2,k}$  and the previous estimated angles  $\tilde{\theta}_k$ , what compromises the estimation of the  $\eta_k$ . The procedure is summarized in Figure 6.

In the preliminary version of this work [41], we propose this simplistic method which is not sufficiently robust to noise remaining in the output solution of the

algorithm, whose stopping criteria is met when this one is sufficiently close to the exact solution. In the noiseless case, the Prony method is able to recover the frequencies with an infinite precision if the number of samples  $|I|$  is greater than  $2K$ . But in our case the estimate  $\tilde{f}_{m,k}$  is affected by some uncertainty  $\epsilon_{m,k}$ ; that is,  $\tilde{f}_{m,k} = f_{m,k} + \epsilon_{m,k}$ , due to the instability of roots finding in presence of noise, and then the error of  $\tilde{\theta}_{m,k} = \arctan(W\tilde{f}_{m,k}/m) = \arctan(Wf_{m,k}/m + W\epsilon_{m,k}/m) \approx \theta_{m,k} + W\epsilon_{m,k}/m$ , is amplified by a factor  $W/m$  and gives a bad result, in particularly for a small  $m$ . Consequently, the mean  $\tilde{\theta}_k = \frac{1}{M} \sum_{m=1}^M \tilde{\theta}_{m,k}$  does not lead to a robust estimation of the angles  $\tilde{\theta}_k$ .

We improve the robustness of the method by applying a linear regression to the data  $\{f_{m,k}\}_{1 \leq m \leq M}$  since

$$\tilde{f}_{m,k} = \frac{\tan \theta_k}{W} m + \epsilon_{m,k} ,$$

in order to estimate the slope  $\tan \theta_k$ . The errors  $\epsilon_{m,k}$  committed by evaluating the frequencies  $f_{m,k}$  are not of the same order of magnitude according to  $m$ . Indeed, for a small  $m$ , the frequencies  $f_{m,k} = \frac{\tan \theta_k}{W} m$ , are close to each other, and the Prony method fails to accurately determine the frequencies when the minimal separation is too small, especially when  $K$  and the amount of residual noise are not too. Consequently, it is preferable to start the linear regression from greater values of  $m$ , in order to space the frequencies on the unit disk.

**Remark.** *Make sure that for high values of  $m$  the two extremal frequencies, say  $f_{m,1} \leq 0$  and  $f_{m,K} \geq 0$ , are not both close respectively to  $-\pi$  and  $\pi$ , what would violate the separation criteria as well.*

The angle parameters  $\tilde{\theta}_k$  are now better estimated. We keep the same procedure of estimation for the offsets  $\tilde{\eta}_k$ , because it has the advantage to conserve the correspondance with the estimated angles  $\tilde{\theta}_k$ .

**Remark.** *Notice that we could compare the results by applying the Prony method on the middle line ( $n_2 = 0$ ) of the image, since  $\mathbf{l}_0[m] = \sum_{k=1}^K c_k e^{j2\pi\eta_k m/W}$  and that the argument of these exponentials are uniquely determined, giving a better estimation of the offsets  $\tilde{\eta}_k$  provided that the latter are well separated, but losing the correspondance with the angles  $\tilde{\theta}_k$ . To reconnect with the angles, one can test all the possible combinations and keep the one which minimizes the data fidelity term; else one can also find the peaks on the Radon transform at column  $\theta = \tilde{\theta}_k$ .*

Finally, as far as the amplitudes  $c_k$  are concerned, taking the modulus of  $\tilde{d}_{m,k}$ 's also gives bad results, since they are computed from the estimate  $\tilde{\mathbf{x}}$  whose amplitudes have been cut down, due to the choice of a parameter  $c < c^*$  for removing noise. It is more appropriate, given the estimated 2-D atoms  $\frac{1}{\cos \tilde{\theta}_k} \hat{\mathbf{a}}(\tilde{\theta}_k, \tilde{\eta}_k)$ , to evaluate the amplitudes  $\tilde{\alpha}_k$  by performing a least square method with respect to the noisy data  $\hat{\mathbf{y}}$ :

$$(\tilde{\alpha}_1, \dots, \tilde{\alpha}_K) = \arg \min_{\alpha_1, \dots, \alpha_K} \left\| \sum_{k=1}^K \alpha_k \hat{\mathbf{I}}_k - \hat{\mathbf{y}} \right\|^2, \quad \hat{\mathbf{I}}_k = \mathbf{A} \cdot \frac{1}{\cos \tilde{\theta}_k} \hat{\mathbf{a}}(\tilde{\theta}_k, \tilde{\eta}_k). \quad (73)$$

The new procedure is the following:

### Extraction of line parameters (Procedure 2)

- From  $m = 1, \dots, M$ ,
  - Compute  $\tilde{f}_{m,k} = \arg(\zeta_{m,k})/(2\pi)$ , where  $(\zeta_{m,k})_k$  are roots of the polynomial  $\sum_{k=0}^K h_{m,k} \zeta^k$  with  $\mathbf{h}_m = [h_{m,0}, \dots, h_{m,K}]^\top$  being the right singular vector of the matrix  $\mathbf{P}_K(\tilde{\mathbf{x}}[m, :])$  with  $I = \{0, \dots, H_S - 1\}$ . It corresponds to the singular value zero (the smallest value in practice).
- For  $k = 1, \dots, K$ 
  - Perform a linear regression on  $\{\tilde{f}_{m,k}\}_m$  to estimate  $\tan \tilde{\theta}_k$  and then  $\tilde{\theta}_k$ .
- From  $m = 1, \dots, M$ ,
  1. Form the matrix  $\tilde{\mathbf{U}}_m = [\mathbf{a}(\tan \tilde{\theta}_1 m/W) \cdots \mathbf{a}(\tan \tilde{\theta}_K m/W)]$ , and compute the complex amplitudes  $\tilde{\mathbf{d}}_m = [\tilde{d}_{m,1}, \dots, \tilde{d}_{m,K}]^\top$  by solving the least-squares linear system  $\tilde{\mathbf{U}}_m^* \tilde{\mathbf{U}}_m \tilde{\mathbf{d}}_m = \tilde{\mathbf{U}}_m^* \tilde{\mathbf{x}}[m, :]$ .
  2. Compute  $\tilde{e}_{m,k} = \tilde{d}_{m,k}/|\tilde{d}_{m,k}|$ .
- For  $k = 1, \dots, K$ 
  - Compute the frequency  $\tilde{\nu}_k$  as previously from  $\mathbf{P}_K(\tilde{\mathbf{e}}_k)$  with  $\tilde{\mathbf{e}}_k = (\tilde{e}_{m,k})_m$  and  $I = \{-M, \dots, M\}$ , and then the horizontal offset  $\tilde{\eta}_k = W\tilde{\nu}_k/(2\pi)$ .
- Find the amplitudes by debiasing as explained in (73).

## 6 Experimental results

The reconstruction procedure was implemented in MATLAB<sup>TM</sup> code, available on the webpage of the first author. We consider an image of size  $W = H = 65$ , containing three lines of parameters:  $(\theta_1, \eta_1, \alpha_1) = (-\pi/5, 0, 255)$ ,  $(\theta_2, \eta_2, \alpha_2) = (\pi/16, -15, 255)$  and  $(\theta_3, \eta_3, \alpha_3) = (\pi/6, 10, 255)$ . The first experiment consists in the reconstruction of the lines from  $\tilde{\mathbf{x}}$  in absence of noise, (1) by applying the operator  $\mathbf{A}$  on this solution, possibly with others kernels  $\mathbf{g}$  and  $\mathbf{h}$ , and then taking the 1-D inverse Fourier transform ; and (2) by applying the Prony method to recover the parameters of the lines, in the aim to display them as vectorial lines. We run the algorithm for  $10^6$  iterations. Results of relative errors for the solution  $\tilde{\mathbf{x}}$  and the estimated parameters are given Figure 7 and Table 1, where  $\Delta_{\theta_i}/\theta_i = |\theta_i - \tilde{\theta}_i|/|\theta_i|$ ,  $\Delta_{\alpha_i}/\alpha_i = |\alpha_i - \tilde{\alpha}_i|/|\alpha_i|$  and  $\Delta_{\eta_i} = |\eta_i - \tilde{\eta}_i|$ . Although the algorithm is quite slow to achieve high accuracy, convergence is guaranteed and we observe empirically perfect reconstruction of  $\mathbf{x}^\dagger$  when the lines are not too close to each other.

The purpose of the second experiment is to highlight the robustness of the method in presence of a strong noise level (Figure 8a). With  $c = 700$  and only  $2.10^3$  iterations, we are able to completely remove the noise and to estimate the line parameters with an error of  $10^{-2}$ . Finally, the last experiment for  $10^5$  iterations, illustrates the efficiency of the method even in presence of a large blur (Figure 8b), yielding an error of  $10^{-4}$ . For both experiments, the estimated images corresponding to step (1) and (2) appear identical, and are displayed in Figure 8c and Figure 8d.

**Remark.** *One way to get rid off the periodicity is to work with an image four times bigger, but it is computationally prohibitive.*

We emphasize that our algorithm has an accuracy which could not be achieved by detecting peaks of the Hough or Radon transform. These methods are relevant for giving a coarse estimation of line parameters. They are robust to strong noise, but completely fail with a strong blur, which prevents peak detection (see Figure 9).

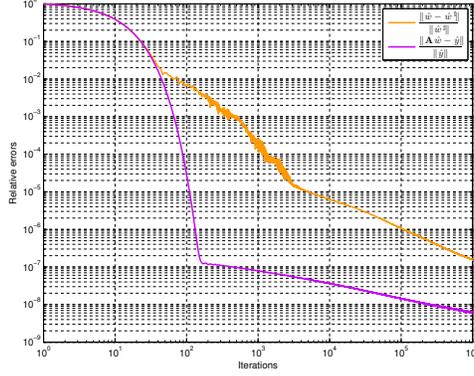


Figure 7: Decrease of the relative errors  $\frac{\|\hat{\mathbf{x}} - \hat{\mathbf{x}}^\# \|_{\mathcal{X}}}{\|\hat{\mathbf{x}}^\# \|_{\mathcal{X}}}$  and  $\frac{\|\mathbf{A}\hat{\mathbf{x}} - \hat{\mathbf{y}} \|_{\mathcal{Y}}}{\|\hat{\mathbf{y}} \|_{\mathcal{Y}}}$  for the first experiment.

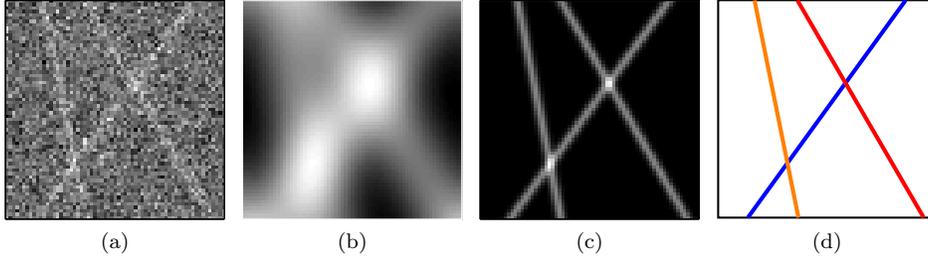


Figure 8: (a) Lines corrupted by a strong noise level ( $\zeta = 200$ ) for the second experiment, (b) Lines degraded by a strong blur ( $\kappa = 8$ ) for the third experiment, (c) denoised image and (d) reconstruction of exact lines.

Table 1: Errors on line parameters recovered by the proposed method.

	Experiment 1	Experiment 2	Experiment 3
$\Delta_\theta/\theta$	$(10^{-7}, 3.10^{-6}, 7.10^{-7})$	$(10^{-2}, 6.10^{-2}, 9.10^{-2})$	$(6.10^{-7}, 9.10^{-5}, 8.10^{-6})$
$\Delta_\alpha/\alpha$	$(10^{-7}, 10^{-7}, 10^{-7})$	$(10^{-2}, 9.10^{-2}, 2.10^{-1})$	$(4.10^{-5}, 2.10^{-5}, 2.10^{-5})$
$\Delta_\eta$	$(4.10^{-6}, 7.10^{-6}, 7.10^{-6})$	$(5.10^{-2}, 4.10^{-2}, 3.10^{-2})$	$(5.10^{-5}, 10^{-4}, 3.10^{-4})$

Notice that even by decreasing the discretization steps of the process, we rapidly reach a plateau, as illustrated by Figure 10. This method is limited in accuracy by the pixel grid. In contrary, our super-resolution method enables to achieve an infinite precision for the line parameters.

### 6.1 Close lines

For a reasonable amount of noise ( $\zeta = 20$ ), the algorithm succeed in separating two close lines  $(\theta_1, \eta_1, \alpha_1) = (-0.73, -1, 255)$  and  $(\theta_2, \eta_2, \alpha_2) = (-0.75, 1, 255)$  as illustrated in Figure 11. The estimation of the parameters gives  $(\tilde{\theta}_1, \tilde{\eta}_1, \tilde{\alpha}_1) = (-0.725, -0.7, 237)$  and  $(\tilde{\theta}_2, \tilde{\eta}_2, \tilde{\alpha}_2) = (-0.753, -0.6, 251)$ .

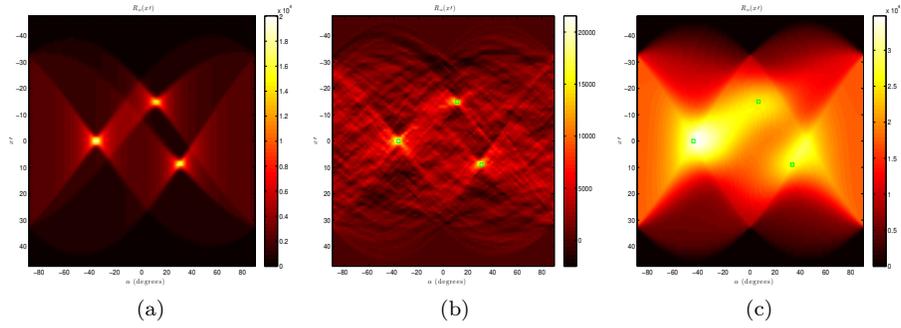


Figure 9: The Radon transform of the image  $\mathbf{y}$  for experiment 1, 2 and 3. The theoretical parameters of the lines are in green.

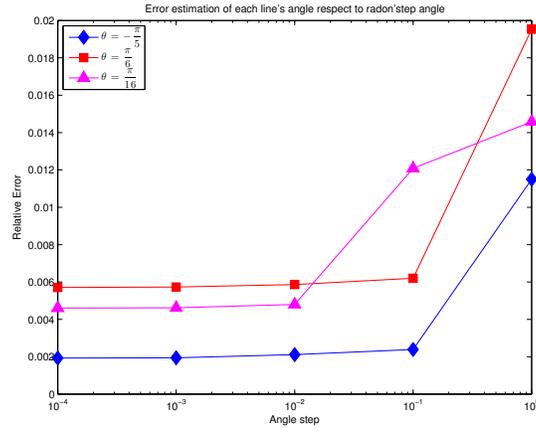


Figure 10: Variation of the accuracy with respect to the angle step of the Radon transform.

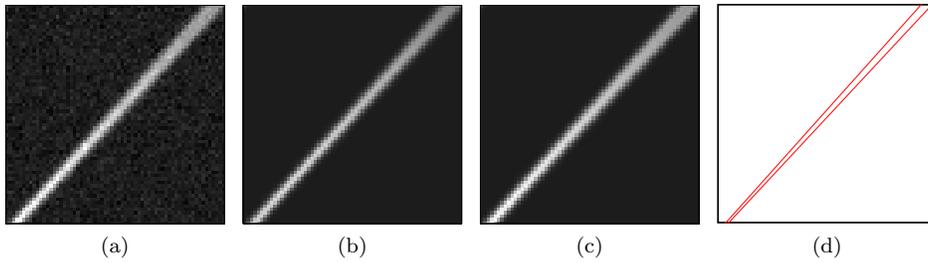


Figure 11: (a) Image  $\mathbf{y}$  of seven lines corrupted by noise ( $\zeta = 20$ ), (b) Result  $\tilde{\mathbf{x}}$  of the denoising step from the optimization, (c) Theoretical solution  $\hat{\mathbf{x}}^{\dagger}$  to compare with, (d) Extraction of the line parameters from the Prony step.

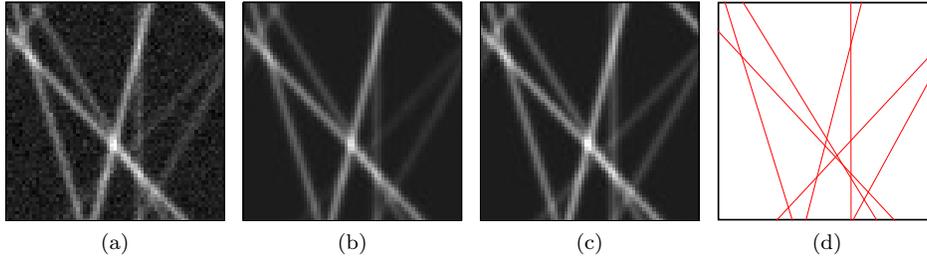


Figure 12: (a) Image  $\mathbf{y}$  of seven lines corrupted by noise ( $\zeta = 20$ ), (b) Result  $\tilde{\mathbf{x}}$  of the denoising step from the optimization, (c) Theoretical solution  $\hat{\mathbf{x}}^\sharp$  to compare with, (d) Extraction of the line parameters from the Prony step.

Table 2: Angles, offsets and amplitudes of the seven lines.

$\theta_k$	-0.75	-0.5	-0.25	$10^{-3}$	0.3	0.55	0.75
$\eta_k$	15	25	2	7	-20	-5	-10
$\alpha_k$	60	80	255	100	180	120	240

## 6.2 More lines and different amplitudes

A more complicated example is depicted in Figure 12 (a), containing seven lines whose parameters are enumerated in Table 2, corrupted by some noise with variance  $\zeta = 20$ . We run the algorithm with  $c = 0.8c^*$ ,  $\tau = 1$ ,  $\sigma = (\tau(M + H_S + 2))^{-1}$ , and after only  $2.10^3$  iterations, we are able to denoise the image as illustrated in Figure 12 (b), and to estimate, thanks to the new Prony procedure, the line parameters as illustrated in Figure 12 (c), with an error of  $10^{-2}$  as given in Table 3.

## 7 Conclusion

We provided a new formulation for the problem of recovering lines in degraded images using the framework of atomic norm minimization. A primal-dual splitting algorithm has been used to solve the convex optimization problem. We applied it successfully to several image recovery problems, recovering lines parameters by the Prony method, and we showed the robustness of the method for strong blur and strong noise level. We insist on the novelty of our approach, which is to estimate lines with parameters (angle, offset, amplitude) living in a continuum, with perfect reconstruction in absence of noise, without being limited by the discrete nature of the image, nor its finite size. This work can be viewed as a proof of concept for super-resolution line detection, and invite us to revisit the Hough transform in a continuous way. Many theoretical questions remain open, like the study of

Table 3: Errors on line parameters recovered by the proposed method.

$\Delta\theta$	$1.10^{-2}$	$2.10^{-2}$	$1.10^{-3}$	$2.10^{-3}$	$5.10^{-3}$	$5.10^{-3}$	$1.10^{-3}$
$\Delta\eta$	$5.10^{-1}$	$7.10^{-2}$	$4.10^{-2}$	$1.10^{-1}$	$1.10^{-2}$	$2.10^{-2}$	$1.10^{-2}$
$\Delta\alpha/\alpha$	$4.10^{-2}$	$5.10^{-2}$	$5.10^{-3}$	$4.10^{-2}$	$6.10^{-3}$	$1.10^{-2}$	$4.10^{-3}$

the separation conditions under which perfect reconstruction can be guaranteed. We should also investigate the possibility to relax the periodicity assumption, and from a practical point of view, parallel computing would be welcome to speed up the proposed algorithm, and to pretend to an industrial use. In future work, the super-resolution detection of lines can be applied on an diffracted image of tubulins, which are curved structures, blurred due to the diffraction through the microscope, and behaving locally like straight lines. Thus one could apply the method set out in this paper on patches, to beat the diffraction and find their exact positions.

## A Proof of Proposition 2

Let us denote

$$\text{SDP}(\mathbf{z}) = \inf_{\mathbf{q} \in \mathbb{C}^N, q_0 \geq 0} \left\{ q_0 : \mathbf{T}'_N(\mathbf{z}, \mathbf{q}) = \begin{pmatrix} \mathbf{T}_N(\mathbf{q}) & \mathbf{z} \\ \mathbf{z}^* & q_0 \end{pmatrix} \succcurlyeq 0 \right\}. \quad (74)$$

We want to prove that  $\text{SDP}(\mathbf{z}) = \|\mathbf{z}\|_{\mathcal{A}}$  and that the minimum in (74) is achieved.

- Suppose  $\mathbf{z} = \sum_{k=1}^K c_k \mathbf{a}(f_k, \phi_k)$  with  $c_k > 0$ .

Defining  $\mathbf{q} = \sum_{k=1}^K c_k \mathbf{a}(f_k)$  with  $\mathbf{q} = (q_0, q_1, \dots, q_{N-1})$ , then  $q_0 = \sum_{k=1}^K c_k$ . For  $i = 0, \dots, N-1$ , the atoms  $\mathbf{a}(f_k)$  have for components  $[\mathbf{a}(f_k)]_i = e^{j2\pi f_k i}$ , hence

$$\begin{aligned} \mathbf{T}_N(\mathbf{a}(f_k)) &= \begin{pmatrix} 1 & e^{-j2\pi f_k} & \dots & e^{-j2\pi f_k(N-1)} \\ e^{j2\pi f_k} & 1 & \dots & e^{-j2\pi f_k(N-2)} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j2\pi f_k(N-1)} & e^{j2\pi f_k(N-2)} & \dots & 1 \end{pmatrix}, \\ &= \begin{pmatrix} 1 \\ e^{j2\pi f_k} \\ \vdots \\ e^{j2\pi f_k(N-1)} \end{pmatrix} \begin{pmatrix} 1 & e^{-j2\pi f_k} & \dots & e^{-j2\pi f_k(N-1)} \end{pmatrix}, \\ &= \mathbf{a}(f_k) \mathbf{a}(f_k)^*. \end{aligned}$$

We deduce that

$$\begin{aligned} \mathbf{T}_N(\mathbf{q}) &= \sum_{k=1}^K c_k \mathbf{T}(\mathbf{a}(f_k)), \\ &= \sum_{k=1}^K c_k \mathbf{a}(f_k) \mathbf{a}(f_k)^*, \\ &= \sum_{k=1}^K c_k \mathbf{a}(f_k, \phi_k) \mathbf{a}(f_k, \phi_k)^*. \end{aligned}$$

Therefore, the matrix

$$\begin{pmatrix} \mathbf{T}_N(\mathbf{q}) & \mathbf{z} \\ \mathbf{z}^* & q_0 \end{pmatrix} = \sum_{k=1}^K c_k \begin{pmatrix} \mathbf{a}(f_k, \phi_k) \\ 1 \end{pmatrix} \begin{pmatrix} \mathbf{a}(f_k, \phi_k) \\ 1 \end{pmatrix}^*,$$

is positive semidefinite. Given  $q_0 = \sum_{k=1}^K c_k$ , we get  $\text{SDP}(\mathbf{z}) \leq \sum_{k=1}^K c_k$ . Since this holds for any decomposition of  $\mathbf{z}$ , we conclude that  $\text{SDP}(\mathbf{z}) \leq \|\mathbf{z}\|_{\mathcal{A}}$ .

- Conversely, let  $\mathbf{q} \in \mathbb{C}^N$  be a vector such that  $q_0 \geq 0$  and  $\begin{pmatrix} \mathbf{T}_N(\mathbf{q}) & \mathbf{z} \\ \mathbf{z}^* & q_0 \end{pmatrix} \succcurlyeq 0$ .

In particular we have  $\mathbf{T}_N(\mathbf{q}) \succcurlyeq 0$ . We denote by  $r$  the rank of  $\mathbf{T}_N(\mathbf{q})$ . [Theorem 1](#) ensures that  $\mathbf{T}_N(\mathbf{q}) \succcurlyeq 0$  and of rank  $r \leq N$  if and only if there exists  $d_k > 0$  and

distinct  $f_k$ , such that

$$\mathbf{q} = \sum_{k=1}^r d_k \mathbf{a}(f_k), \quad (75)$$

$$q_0 = \sum_{k=1}^r d_k. \quad (76)$$

Let us set  $\mathbf{D} = \text{diag}(d_1, \dots, d_r)$  and

$$\mathbf{V} = (\mathbf{a}(f_1) \quad \dots \quad \mathbf{a}(f_r)) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ e^{j2\pi f_1} & e^{j2\pi f_2} & \dots & e^{j2\pi f_r} \\ e^{j2\pi f_1 2} & e^{j2\pi f_2 2} & \dots & e^{j2\pi f_r 2} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j2\pi f_1(N-1)} & e^{j2\pi f_2(N-1)} & \dots & e^{j2\pi f_r(N-1)} \end{pmatrix}.$$

By linearity of the operator  $\mathbf{T}_N$ :

$$\begin{aligned} \mathbf{T}_N(\mathbf{q}) &= \sum_{k=1}^r d_k \mathbf{T}_N(\mathbf{a}(f_k)), \\ &= \sum_{k=1}^r d_k \mathbf{a}(f_k) \mathbf{a}(f_k)^*, \\ &= \mathbf{V} \mathbf{D} \mathbf{V}^*. \end{aligned}$$

Since  $\mathbf{T}_N(\mathbf{a}(f_k))$  contains only ones on the diagonal, we have

$$\frac{1}{N} \text{Tr}(\mathbf{T}_N(\mathbf{q})) = \sum_{k=1}^r d_k > 0.$$

Besides,  $\frac{1}{N} \text{Tr}(\mathbf{T}_N(\mathbf{q})) = q_0$ , therefore  $q_0 > 0$ .

Let be  $\mathbf{M}$  a general block matrix  $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{C} \end{pmatrix}$ , the Schur complement gives

$$[\mathbf{C} \succ 0 \Rightarrow \mathbf{M} \succ 0] \implies [\mathbf{A} - \mathbf{B} \mathbf{C}^{-1} \mathbf{B}^* \succ 0].$$

We apply this lemma to  $\mathbf{M} = \begin{pmatrix} \mathbf{T}_N(\mathbf{q}) & \mathbf{z} \\ \mathbf{z}^* & q_0 \end{pmatrix}$ , with  $\mathbf{A} = \mathbf{T}_N(\mathbf{q})$ ,  $\mathbf{B} = \mathbf{z}$  and  $\mathbf{C} = q_0$ . The left term is satisfied by hypothesis, hence

$$\mathbf{T}_N(\mathbf{q}) - q_0^{-1} \mathbf{z} \mathbf{z}^* \succ 0 \iff \mathbf{V} \mathbf{D} \mathbf{V}^* - q_0^{-1} \mathbf{z} \mathbf{z}^* \succ 0.$$

We define the square matrix  $\mathbf{V}_r$  by extracting the  $r$  first rows and columns of  $\mathbf{V}$ , which is a Vandermonde matrix, whose determinant is

$$\det(\mathbf{V}_r) = \prod_{1 \leq k < l \leq r} (\mathbf{a}(f_l) - \mathbf{a}(f_k)).$$

Since we assumed  $f_k \neq f_l, \forall k \neq l$ ,  $\mathbf{V}_r$  is invertible, and  $\text{rank}(\mathbf{V}) = r$ . Let us define  $\mathbf{v} : \mathbb{C}^r \rightarrow \mathbb{C}^N$  and  $\mathbf{v}^* : \mathbb{C}^N \rightarrow \mathbb{C}^r$  the endomorphisms corresponding to matrices  $\mathbf{V}$  and  $\mathbf{V}^*$ . We have  $\text{rank}(\mathbf{v}^*) = \text{rank}(\mathbf{v}) = r$ . By the rank-nullity theorem:

$$\dim(\ker \mathbf{v}^*) = N - r.$$

Thus, there exists a vector  $\mathbf{p} \in \mathbb{C}^N$  such that  $\mathbf{p} \neq \mathbf{0}$  and  $\mathbf{V}^* \mathbf{p} = \mathbf{0} \Leftrightarrow \mathbf{p}^* \mathbf{V} = \mathbf{0}$ . Consequently,

$$\begin{aligned} \mathbf{p}^* (\mathbf{V} \mathbf{D} \mathbf{V}^* - q_0^{-1} \mathbf{z} \mathbf{z}^*) \mathbf{p} &\geq 0 \Leftrightarrow (\mathbf{p}^* \mathbf{V}) \mathbf{D} (\mathbf{V}^* \mathbf{p}) - q_0^{-1} \mathbf{p}^* \mathbf{z} \mathbf{z}^* \mathbf{p} \geq 0, \\ &\Leftrightarrow q_0^{-1} \|\mathbf{p}^* \mathbf{z}\|_2^2 \leq 0, \\ &\Leftrightarrow \|\mathbf{p}^* \mathbf{z}\|_2^2 = 0, \\ &\Leftrightarrow \mathbf{p}^* \mathbf{z} = 0, \\ &\Leftrightarrow \mathbf{p} \perp \mathbf{z}. \end{aligned}$$

Since  $\mathbf{p} \in \ker \mathbf{v}^*$ , then  $\mathbf{z} \in (\ker \mathbf{v}^*)^\perp = \text{Im } \mathbf{v}$ , so there exists a vector  $\mathbf{w} \in \mathbb{C}^r$  such that  $\mathbf{z} = \mathbf{V} \mathbf{w} = \sum_{k=1}^r w_k \mathbf{a}(f_k)$ , hence

$$\mathbf{V} \mathbf{D} \mathbf{V}^* - q_0^{-1} \mathbf{V} \mathbf{w} \mathbf{w}^* \mathbf{V}^* \succcurlyeq 0.$$

Besides,  $\text{Im } \mathbf{v}^* \subset \mathbb{C}^r$  and  $\dim(\text{Im } \mathbf{v}^*) = \text{rank}(\mathbf{v}^*) = r = \dim(\mathbb{C}^r)$ , thus  $\text{Im } \mathbf{v}^* = \mathbb{C}^r$  and  $\mathbf{v}^*$  is surjective. Consequently, there exists a vector  $\mathbf{u} \in \mathbb{C}^N$  such that  $\mathbf{V}^* \mathbf{u} = \text{sgn}(\mathbf{w}) = (w_1/|w_1|, \dots, w_r/|w_r|)^\top$ , and

$$\begin{aligned} \mathbf{u}^* (\mathbf{V} \mathbf{D} \mathbf{V}^* - q_0^{-1} \mathbf{V} \mathbf{w} \mathbf{w}^* \mathbf{V}^*) \mathbf{u} &\geq 0 \Leftrightarrow (\mathbf{u}^* \mathbf{V}) \mathbf{D} (\mathbf{V}^* \mathbf{u}) - q_0^{-1} (\mathbf{u}^* \mathbf{V}) \mathbf{w} \mathbf{w}^* (\mathbf{V}^* \mathbf{u}) \geq 0, \\ &\Leftrightarrow \text{sgn}(\mathbf{w})^* \mathbf{D} \text{sgn}(\mathbf{w}) - \frac{1}{q_0} \text{sgn}(\mathbf{w})^* \mathbf{w} \mathbf{w}^* \text{sgn}(\mathbf{w}) \geq 0, \\ &\Leftrightarrow \sum_{k=1}^r d_k \left| \frac{w_k}{|w_k|} \right|^2 - q_0^{-1} \left( \sum_{k=1}^r \frac{w_k^*}{|w_k|} w_k \right)^2 \geq 0, \\ &\Leftrightarrow q_0^2 \geq \left( \sum_{k=1}^r |w_k| \right)^2, \text{ (since } q_0 = \sum_{k=1}^r d_k \text{)}, \\ &\Leftrightarrow q_0 \geq \sum_{k=1}^r |w_k| \geq \|\mathbf{z}\|_{\mathcal{A}}, \end{aligned}$$

by definition of the atomic norm (28). Taking the infimum leads to  $\text{SDP}(\mathbf{z}) \geq \|\mathbf{z}\|_{\mathcal{A}}$ .

• Finally, let us show that the infimum of the linear form  $\ell : \mathbf{q} \mapsto q_0$  is achieved on the set

$$A(\mathbf{z}) = \left\{ \mathbf{q} \in \mathbb{R}^+ \times \mathbb{C}^{N-1} : \mathbf{T}'_N(\mathbf{z}, \mathbf{q}) = \begin{pmatrix} \mathbf{T}_N(\mathbf{q}) & \mathbf{z} \\ \mathbf{z}^* & q_0 \end{pmatrix} \succcurlyeq 0 \right\};$$

that is,

$$\text{SDP}(\mathbf{z}) = \inf_{\mathbf{q} \in A(\mathbf{z})} \ell(\mathbf{q}) = \min_{\mathbf{q} \in A(\mathbf{z})} \ell(\mathbf{q}). \quad (77)$$

(i) Let us notice that since  $\mathbf{q} \in A(\mathbf{z})$  implies  $\mathbf{T}_N(\mathbf{q}) \succcurlyeq 0$  then

$$\ell(\mathbf{q}) = q_0 = \frac{1}{N} \text{Tr}(\mathbf{T}_N(\mathbf{q})) = \sum_{i=0}^{N-1} \lambda_i \geq 0,$$

with  $\lambda_i$  the eigenvalues of  $\mathbf{T}_N(\mathbf{q})$  which are positive reals.

(ii) First we show that  $A(\mathbf{z})$  is nonempty, since  $\mathbf{q} = (\|\mathbf{z}\|_2, 0, \dots, 0)^\top \in A(\mathbf{z})$ . Indeed for a fixed vector  $\mathbf{z} = (z_0, \dots, z_{N-1}) \in \mathbb{C}^N$ ,  $\mathbf{v} = (v_0, \dots, v_N) \in \mathbb{C}^{N+1}$  and  $\mathbf{v}' = (v_0, \dots, v_{N-1}) \in \mathbb{C}^N$  we have for this  $\mathbf{q}$  :

$$\begin{aligned} \mathbf{v}^* \begin{pmatrix} \mathbf{T}_N(\mathbf{q}) & \mathbf{z} \\ \mathbf{z}^* & q_0 \end{pmatrix} \mathbf{v} &= \|\mathbf{z}\|_2 \|\mathbf{v}\|_2^2 + 2 \operatorname{Re} \left( v_N \sum_{i=0}^{N-1} z_i v_i^* \right), \\ &\geq \|\mathbf{z}\|_2 \|\mathbf{v}\|_2^2 - 2|v_N| |\langle \mathbf{z}, \mathbf{v}' \rangle|, \\ &\geq \|\mathbf{z}\|_2 \|\mathbf{v}\|_2^2 - 2|v_N| \|\mathbf{z}\|_2 \|\mathbf{v}'\|_2, \\ &\geq \|\mathbf{z}\|_2 (\|\mathbf{v}'\|_2^2 - 2|v_N| \|\mathbf{v}'\|_2 + |v_N|^2), \\ &\geq \|\mathbf{z}\|_2 (\|\mathbf{v}'\|_2 - |v_N|)^2, \\ &\geq 0. \end{aligned}$$

Then  $\mathbf{q} = (\|\mathbf{z}\|_2, 0, \dots, 0)^\top \in A(\mathbf{z})$  and  $q_0 = \|\mathbf{z}\|_2$ , which means that  $A(\mathbf{z})$  is non-empty and the set  $\{\ell(\mathbf{q}) : \mathbf{q} \in A(\mathbf{z})\} \subset \mathbb{R}^+$  is non-empty so admits a lower bound lesser than  $\|\mathbf{z}\|_2$ , hence

$$0 \leq \operatorname{SDP}(\mathbf{z}) = \inf_{\mathbf{q} \in A(\mathbf{z})} \ell(\mathbf{q}) \leq \|\mathbf{z}\|_2.$$

(iii) From (ii), we have

$$\operatorname{SDP}(\mathbf{z}) = \inf_{\mathbf{q} \in A(\mathbf{z})} \ell(\mathbf{q}) = \inf_{B(\mathbf{z})} \ell(\mathbf{q}),$$

where

$$B(\mathbf{z}) = \{\mathbf{q} \in A(\mathbf{z}), q_0 \leq \|\mathbf{z}\|_2\} \subset A(\mathbf{z}).$$

Now, from (75) and (76) we can show that  $B(\mathbf{z})$  is bounded since:

$$\forall \mathbf{q} \in A(\mathbf{z}), \quad \|\mathbf{q}\|_2 \leq \sum_{k=1}^r d_k \|\mathbf{a}(f_k)\|_2 \leq \sqrt{N} \sum_{k=1}^r d_k = \sqrt{N} q_0,$$

hence

$$\forall \mathbf{q} \in B(\mathbf{z}), \quad \|\mathbf{q}\|_2 \leq \sqrt{N} q_0 \leq \sqrt{N} \|\mathbf{z}\|_2.$$

Consequently,

$$B(\mathbf{z}) \subset B_{\|\cdot\|_2}(0, \|\mathbf{z}\|_2).$$

(iv) Moreover,  $A(\mathbf{z}) = \mathbf{T}'_N(\mathbf{z}, \cdot)^{-1}(\mathcal{C})$  is a closed set since the cone of positive matrix  $\mathcal{C}$  is closed and the application  $\mathbf{T}'_N(\mathbf{z}, \cdot)$  is linear so continuous in finite dimension. Thus,

$$B(\mathbf{z}) = A(\mathbf{z}) \cap \{\mathbf{q} \in \mathbb{R}^+ \times \mathbb{C}^{N-1} : q_0 \leq \|\mathbf{z}\|_2\}$$

is a closed set as intersection of two closed sets.

(v) From (iii) and (iv), we conclude that  $B(\mathbf{z})$  is a compact set and that the application  $\ell$ , which is linear and then continuous, achieved its minimum on  $B(\mathbf{z})$  so on  $A(\mathbf{z})$ , which proves the result (77).

## B Proof of Proposition 3

The proof of the direct implication is straightforward. Let us consider the converse one.

By [Theorem 1](#), since  $\forall n, \mathbf{T}_M(\mathbf{l}_n) \succcurlyeq 0$  and is of rank one, then there exists  $\gamma_n \geq 0$  and  $f_n \in [0, 1[$  such that

$$\mathbf{l}_n[m] = \gamma_n \exp(j2\pi f_n m) .$$

Since we assume  $\forall n, \hat{\mathbf{x}}[0, n] = \hat{\mathbf{x}}[0, 0] = c_1$ , then  $\mathbf{l}_n[0] = \mathbf{l}_0[0] = c_1$ , we have

$$\mathbf{l}_n[m] = c_1 \exp(j2\pi f_n m) . \quad (78)$$

Let  $m$  be fixed. The Prony matrix  $\mathbf{P}_1(\mathbf{t}_m)$  of size  $2 \times (N - 1)$

$$\mathbf{P}_1(\mathbf{t}_m) = \begin{pmatrix} \mathbf{t}_m[1] & \mathbf{t}_m[0] \\ \vdots & \vdots \\ \mathbf{t}_m[N-1] & \mathbf{t}_m[N-2] \end{pmatrix} ,$$

is of rank one, consequently there exists  $\lambda_m \in \mathbb{C}$  such that

$$\mathbf{t}_m[n+1] = \lambda_m \mathbf{t}_m[n], \quad \forall 0 \leq n \leq N-2 .$$

Thus,

$$\mathbf{t}_m[n] = \lambda_m^n \mathbf{t}_m[0], \quad \forall 0 \leq n \leq N-1 .$$

From (78)  $\mathbf{t}_m[0] = \mathbf{l}_0[m] = c_1 \exp(j2\pi f_0 m)$ ,  $\mathbf{t}_m[1] = \mathbf{l}_1[m] = c_1 \exp(j2\pi f_1 m)$  and then

$$\lambda_m = \frac{\mathbf{t}_m[1]}{\mathbf{t}_m[0]} = \frac{\ell_1[m]}{\ell_0[m]} = \exp(j2\pi(f_1 - f_0)m) .$$

Therefore we have

$$\begin{aligned} \mathbf{t}_m[n] &= \lambda_m^n \mathbf{t}_m[0] , \\ &= \exp(j2\pi(f_1 - f_0)m)^n c_1 \exp(j2\pi f_0 m) , \\ &= c_1 \exp[j2\pi((f_1 - f_0)n + f_0)m] . \end{aligned}$$

## C Proof of Proposition 5

For  $\mathbf{z} = (z_0, \dots, z_{N-1}) \in \mathbb{R} \times \mathbb{C}^{N-1}$  and  $\mathbf{M} \in \mathcal{T}_N \subset \mathcal{M}_N(\mathbb{C})$  a Hermitian-Toeplitz matrix of dimension  $N$ . We have:

$$\begin{aligned} \langle \mathbf{T}_N(\mathbf{z}), \mathbf{M} \rangle_{\mathcal{M}} &= \sum_{0 \leq i, j \leq N-1} [\mathbf{T}_N(\mathbf{z})]_{ij}^* \mathbf{M}_{ij} , \\ &= \sum_{0 \leq i \leq j \leq N-1} z_{j-i} \mathbf{M}_{ij} + \sum_{0 \leq j < i \leq N-1} z_{i-j}^* \mathbf{M}_{ij} , \\ &\stackrel{(*)}{=} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1-k} z_k \mathbf{M}_{l, l+k} + \sum_{k=1}^{N-1} \sum_{l=0}^{N-1-k} z_k^* \mathbf{M}_{l+k, l} , \\ &\stackrel{(**)}{=} z_0 \left( \sum_{l=0}^{N-1} \mathbf{M}_{l, l} \right) + 2\text{Re} \left\{ \sum_{k=1}^{N-1} z_k^* \left( \sum_{l=0}^{N-1-k} \mathbf{M}_{l+k, l} \right) \right\} , \end{aligned}$$

with (\*) a change of variable  $k \leftarrow j - i$  and (\*\*) using that  $\mathbf{M}_{l,l+k} = \mathbf{M}_{l+k,l}^*$ .

Then, by writing

$$\mathbf{T}'_N(\mathbf{z}, \mathbf{q}) = \left( \begin{array}{c|c} \mathbf{0} & \begin{matrix} z_0 \\ \vdots \\ z_{N-1} \end{matrix} \\ \hline z_0^* & \cdots & z_{N-1}^* \\ \hline & & 0 \end{array} \right) + \left( \begin{array}{c|c} \mathbf{T}_N(\mathbf{q}) & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0 & \cdots & 0 \\ \hline & & q_0 \end{array} \right)$$

we obtain as well for  $\mathbf{M} \in \mathcal{T}_{N+1}$ :

$$\begin{aligned} \langle \mathbf{T}'_N(\mathbf{z}, \mathbf{q}), \mathbf{M} \rangle_{\mathcal{M}} &= 2\text{Re} \left\{ \sum_{k=0}^{N-1} z_k^* \mathbf{M}_{N+1,k} \right\} \\ &\quad + q_0 \left( \sum_{l=0}^N \mathbf{M}_{l,l} \right) + 2\text{Re} \left\{ \sum_{k=1}^{N-1} q_k^* \left( \sum_{l=0}^{N-1-k} \mathbf{M}_{l+k,l} \right) \right\}. \end{aligned}$$

Consequently, the adjoint of the operator

$$\mathbf{T}_{M+1} : (\mathcal{X}_l, \langle \cdot, \cdot \rangle_{\mathcal{X}_l}) \rightarrow (\mathcal{T}_{M+1}, \langle \cdot, \cdot \rangle_{\mathcal{M}}),$$

with the inner product  $\langle \cdot, \cdot \rangle_{\mathcal{X}_l}$  defined in (19), and applied to a matrix  $\mathbf{M}^{(2)} \in \mathcal{T}_{M+1}$ , is given by the vector

$$\mathbf{z}_2 = \mathbf{T}_{M+1}^* \mathbf{M}^{(2)} \in \mathbb{R} \times \mathbb{C}^M,$$

whose components are:

$$\mathbf{z}_2[k] = \sum_{l=0}^{M-k} \mathbf{M}_{l+k,l}^{(2)}, \quad \forall k = 0, \dots, M.$$

As well, the adjoint of the operator

$$\mathbf{T}'_{H_S} : (\mathcal{X}_t \times \mathcal{Q}_t, \langle \cdot, \cdot \rangle_{\mathcal{X}_t} + \langle \cdot, \cdot \rangle_{\mathcal{Q}_t}) \rightarrow (\mathcal{T}_{H_S+1}, \langle \cdot, \cdot \rangle_{\mathcal{M}}),$$

with the inner products  $\langle \cdot, \cdot \rangle_{\mathcal{X}_t}$  and  $\langle \cdot, \cdot \rangle_{\mathcal{Q}_t}$  defined in (20)-(33), and applied to  $\mathbf{M}^{(2)} \in \mathcal{T}_{H_S+1}$ , is given by the following pair of vectors

$$(\mathbf{z}_1, \mathbf{q}_1) = \mathbf{T}'_{H_S}{}^* \mathbf{M}^{(2)} \in \mathbb{C}^{H_S} \times (\mathbb{R} \times \mathbb{C}^{H_S-1}),$$

whose components are:

$$\mathbf{z}_1[k] = \mathbf{M}_{H_S+1,k}^{(1)}, \quad \mathbf{q}_1[k] = \sum_{l=0}^{H_S-1-k} \mathbf{M}_{l+k,l}^{(1)} + \delta_k \mathbf{M}_{H_S,H_S}^{(1)}, \quad \forall k = 0, \dots, H_S - 1.$$

## D Proof of Proposition 6

First, let us compute the operator norm  $\|\mathbf{T}_{M+1}\|^2 = \sup_{\mathbf{z} \in \mathcal{X}_l} \frac{\|\mathbf{T}_{M+1}(\mathbf{z})\|_{\mathbb{F}}^2}{\|\mathbf{z}\|_{\mathcal{X}_l}^2}$ .

By definition, we have  $\|\mathbf{z}\|_{\mathcal{X}_t}^2 = z_0^2 + 2|z_1|^2 + \dots + 2|z_M|^2$ . Moreover, we get:

$$\|\mathbf{T}_{M+1}(\mathbf{z})\|_{\mathbb{F}}^2 = (M+1)z_0^2 + 2M|z_1|^2 + 2(M-1)|z_2|^2 + \dots + 2|z_M|^2 \leq (M+1)\|\mathbf{z}\|_{\mathcal{X}_t}^2,$$

with equality when  $\mathbf{z} = (1, 0, \dots, 0)$ , hence

$$\|\mathbf{T}_{M+1}\|^2 = M + 1. \quad (79)$$

Let us now decompose the operator  $\mathbf{T}'_{H_S}$  as follows:

$$\mathbf{T}'_{H_S}(\mathbf{z}, \mathbf{q}) = \underbrace{\left( \begin{array}{c|c} \mathbf{0} & \begin{matrix} z_0 \\ \vdots \\ z_{H_S-1} \end{matrix} \\ \hline \begin{matrix} z_0^* & \dots & z_{H_S-1}^* \end{matrix} & 0 \end{array} \right)}_{\mathbf{T}_1(\mathbf{z})} + \underbrace{\left( \begin{array}{c|c} \mathbf{T}_{H_S}(\mathbf{q}) & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0 & \dots & 0 & q_0 \end{array} \right)}_{\mathbf{T}_2(\mathbf{q})}.$$

We directly have  $\|\mathbf{T}_1(\mathbf{z})\|_{\mathbb{F}}^2 = \|\mathbf{z}\|_{\mathcal{X}_t}^2$ , that is  $\|\mathbf{T}_1\| = 1$ . Besides, we have

$$\|\mathbf{T}_2(\mathbf{q})\|_{\mathbb{F}}^2 = (H_S + 1)q_0^2 + 2(H_S - 1)|q_1|^2 + \dots + 2|q_{H_S-1}|^2 \leq (H_S + 1)\|\mathbf{q}\|_{\mathcal{Q}_t}^2,$$

with equality when  $\mathbf{q} = (1, 0, \dots, 0)$ , hence  $\|\mathbf{T}_2\|^2 = H_S + 1$ .

Now, we compute

$$\begin{aligned} \frac{\|\mathbf{T}'_{H_S}(\mathbf{z}, \mathbf{q})\|_{\mathbb{F}}^2}{\|(\mathbf{z}, \mathbf{q})\|_{\mathcal{X}_t \times \mathcal{Q}_t}^2} &= \frac{\|\mathbf{T}_1(\mathbf{z})\|_{\mathbb{F}}^2 + \|\mathbf{T}_2(\mathbf{q})\|_{\mathbb{F}}^2}{\|\mathbf{z}\|_{\mathcal{X}_t}^2 + \|\mathbf{q}\|_{\mathcal{Q}_t}^2}, \\ &\leq \frac{\|\mathbf{T}_1\|^2 \|\mathbf{z}\|_{\mathcal{X}_t}^2 + \|\mathbf{T}_2\|^2 \|\mathbf{q}\|_{\mathcal{Q}_t}^2}{\|\mathbf{z}\|_{\mathcal{X}_t}^2 + \|\mathbf{q}\|_{\mathcal{Q}_t}^2}, \\ &\leq \alpha \|\mathbf{T}_1\|^2 + (1 - \alpha) \|\mathbf{T}_2\|^2, \\ &\leq \max(\|\mathbf{T}_1\|^2, \|\mathbf{T}_2\|^2) = \|\mathbf{T}_2\|^2, \end{aligned}$$

with  $\alpha = \frac{\|\mathbf{q}\|_{\mathcal{Q}_t}^2}{\|\mathbf{z}\|_{\mathcal{X}_t}^2 + \|\mathbf{q}\|_{\mathcal{Q}_t}^2}$  and is achieved when  $\mathbf{z} = \mathbf{0}$  and  $\mathbf{q} = (1, 0, \dots, 0)$ , hence

$$\|\mathbf{T}'_{H_S}\|^2 = H_S + 1.$$

We are now able to derive the operator norms (51):

$$\begin{aligned} \left\| \mathbf{L}^{(1)}(\mathbf{X}) \right\|_{(1)}^2 &= \sum_{m=1}^M \left\| \mathbf{L}_m^{(1)}(\mathbf{X}) \right\|_{\mathbb{F}}^2, \\ &\leq \|\mathbf{T}'_{H_S}\|^2 \sum_{m=1}^M \left( \|\widehat{\mathbf{x}}[m, :]\|_{\mathcal{X}_t}^2 + \|\mathbf{q}[m, :]\|_{\mathcal{Q}_t}^2 \right), \\ &\leq \|\mathbf{T}'_{H_S}\|^2 \|\mathbf{X}\|_{\mathcal{H}}^2, \end{aligned}$$

with equality when  $\widehat{\mathbf{x}} = \mathbf{0}$ ,  $\mathbf{q}[0, :] = \mathbf{0}$  and  $\mathbf{q}[m, :] = (1, 0, \dots, 0)$  for all  $m \in \llbracket 1, M \rrbracket$ .

As well, we have

$$\begin{aligned} \left\| \mathbf{L}^{(2)}(\mathbf{X}) \right\|_{(2)}^2 &= \sum_{n_2=0}^{H_S-1} \left\| \mathbf{L}_{n_2}^{(2)}(\mathbf{X}) \right\|_{\mathbf{F}}^2, \\ &\leq \left\| \mathbf{T}_{M+1} \right\|^2 \sum_{n_2=0}^{H_S-1} \left\| \widehat{\mathbf{x}}[:, n_2] \right\|_{\mathcal{X}_l}^2, \\ &\leq \left\| \mathbf{T}_{M+1} \right\|^2 \left\| \mathbf{X} \right\|_{\mathcal{H}}^2, \end{aligned}$$

with equality when  $\mathbf{q} = \mathbf{0}$  and  $\widehat{\mathbf{x}}[:, n_2] = (1, 0, \dots, 0)$  for all  $n_2 \in \llbracket 0, H_S - 1 \rrbracket$ .

We conclude that

$$\begin{aligned} \left\| \mathbf{L}^{(1)} \right\|_{(1)}^2 &= \left\| \mathbf{T}'_{H_S} \right\|^2 = H_S + 1, \\ \left\| \mathbf{L}^{(2)} \right\|_{(2)}^2 &= \left\| \mathbf{T}_{M+1} \right\|^2 = M + 1. \end{aligned}$$

Finally,

$$\begin{aligned} \left\| \mathbf{L}(\mathbf{X}) \right\|_{(1,2)}^2 &= \left\| \mathbf{L}^{(1)}(\mathbf{X}) \right\|_{(1)}^2 + \left\| \mathbf{L}^{(2)}(\mathbf{X}) \right\|_{(2)}^2, \\ &\leq \left( \left\| \mathbf{L}^{(1)} \right\|_{(1)}^2 + \left\| \mathbf{L}^{(2)} \right\|_{(2)}^2 \right) \left\| \mathbf{X} \right\|_{\mathcal{H}}^2. \end{aligned}$$

## References

- [1] H. H. BAUSCHKE AND P. L. COMBETTES, *Convex analysis and monotone operator theory in Hilbert spaces*, Springer Science & Business Media, 2011.
- [2] B. N. BHASKAR, G. TANG, AND B. RECHT, *Atomic norm denoising with applications to line spectral estimation*, 61 (2013), pp. 5987–5999.
- [3] T. BLU, *The generalized annihilation property: A tool for solving finite rate of innovation problems*, in *Sampling Theory and Applications (SAMPTA)*, 2009, pp. Special-Session.
- [4] T. BLU, P.-L. DRAGOTTI, M. VETTERLI, P. MARZILIANO, AND L. COULOT, *Sparse sampling of signal innovations*, *IEEE Signal Processing Magazine*, 25 (2008), pp. 31–40.
- [5] S. BOYD AND L. VANDENBERGHE, *Convex optimization*, Cambridge University Press, 2004.
- [6] K. BREDIES, K. KUNISCH, AND T. POCK, *Total generalized variation*, *SIAM Journal on Imaging Sciences*, 3 (2010), pp. 492–526.
- [7] Y. BRESLER AND A. MACOVSKI, *Exact maximum likelihood parameter estimation of superimposed exponential signals in noise*, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34 (1986), pp. 1081–1089.

- [8] E. J. CANDÈS AND C. FERNANDEZ-GRANDA, *Towards a mathematical theory of super-resolution*, Communications on Pure and Applied Mathematics, 67 (2014), pp. 906–956.
- [9] C. CARATHÉODORY, *Über den Variabilitätsbereich der Fourier'schen Konstanten von positiven harmonischen Funktionen*, Rendiconti del Circolo Matematico di Palermo (1884-1940), 32 (1911), pp. 193–217.
- [10] C. CARATHÉODORY AND L. FEJÉR, *Über den Zusammenhang der Extremen von harmonischen Funktionen mit ihren Koeffizienten und über den Picard-Landau'schen Satz*, Rendiconti del Circolo Matematico di Palermo (1884-1940), 32 (1911), pp. 218–239.
- [11] A. CHAMBOLLE, V. CASELLES, D. CREMERS, M. NOVAGA, AND T. POCK, *An introduction to total variation for image analysis*, Theoretical foundations and numerical methods for sparse recovery, 9 (2010), p. 227.
- [12] A. CHAMBOLLE AND T. POCK, *A first-order primal-dual algorithm for convex problems with applications to imaging*, Journal of Mathematical Imaging and Vision, 40 (2011), pp. 120–145.
- [13] Y. CHAN, J. LAVOIE, AND J. PLANT, *A parameter estimation approach to estimation of frequencies of sinusoids*, IEEE Transactions on Acoustics, Speech, and Signal Processing, 29 (1981), pp. 214–219.
- [14] V. CHANDRASEKARAN, B. RECHT, P. A. PARRILO, AND A. S. WILLSKY, *The convex geometry of linear inverse problems*, Foundations of Computational mathematics, 12 (2012), pp. 805–849.
- [15] G. CHIERCHIA, N. PUSTELNIK, B. PESQUET-POPESCU, AND J.-C. PESQUET, *A nonlocal structure tensor-based approach for multicomponent image recovery problems*, IEEE Transactions on Image Processing, 23 (2014), pp. 5531–5544.
- [16] P. L. COMBETTES AND J.-C. PESQUET, *Proximal splitting methods in signal processing*, in Fixed-point algorithms for inverse problems in science and engineering, Springer, 2011, pp. 185–212.
- [17] L. CONDAT, *A primal–dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms*, Journal of Optimization Theory and Applications, 158 (2013), pp. 460–479.
- [18] L. CONDAT, *Semi-local total variation for regularization of inverse problems.*, in EUSIPCO, 2014, pp. 1806–1810.
- [19] L. CONDAT, *Atomic norm minimization for the decomposition of vectors into complex exponentials*, research report, GIPSA-Lab, Grenoble, France, (2017).
- [20] L. CONDAT, *Discrete total variation: New definition and minimization*, SIAM Journal on Imaging Sciences, 10 (2017), pp. 1258–1290.
- [21] L. CONDAT, J. BOULANGER, N. PUSTELNIK, S. SAHNOUN, AND L. SENG-MANIVONG, *A 2-d spectral analysis method to estimate the modulation parameters in structured illumination microscopy*, in 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), IEEE, 2014, pp. 604–607.

- [22] L. CONDAT AND A. HIRABAYASHI, *Cadzow denoising upgraded: A new projection method for the recovery of Dirac pulses from noisy linear measurements*, Sampling Theory in Signal and Image Processing, (2014). in press.
- [23] L. CONDAT AND A. HIRABAYASHI, *Cadzow denoising upgraded: A new projection method for the recovery of dirac pulses from noisy linear measurements*, Sampling Theory in Signal and Image Processing, 14 (2015), pp. p–17.
- [24] S. R. DEANS, *Hough transform from the Radon transform*, IEEE Transactions on Pattern Analysis and Machine Intelligence, (1981), pp. 185–188.
- [25] D. L. DONOHO, *Compressed sensing*, IEEE Transactions on information theory, 52 (2006), pp. 1289–1306.
- [26] P. L. DRAGOTTI, M. VETTERLI, AND T. BLU, *Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets strang-fix*, IEEE Transactions on Signal Processing, 55 (2007), pp. 1741–1757.
- [27] M. ELAD, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, Springer Publishing Company, Incorporated, 1st ed., 2010.
- [28] C. FERNANDEZ-GRANDA, *Super-resolution and compressed sensing*, SIAM News, 46 (2013).
- [29] P. HOUGH, *Method and means for recognizing complex patterns*, Dec. 18 1962. US Patent 3,069,654.
- [30] J. ILLINGWORTH AND J. KITTLER, *A survey of the Hough transform*, Computer vision, graphics, and image processing, 44 (1988), pp. 87–116.
- [31] D. H. JOHNSON, *The application of spectral estimation methods to bearing estimation problems*, Proceedings of the IEEE, 70 (1982), pp. 1018–1028.
- [32] S. LEFKIMMIATIS, J. P. WARD, AND M. UNSER, *Hessian Schatten-norm regularization for linear inverse problems*, IEEE transactions on image processing, 22 (2013), pp. 1873–1888.
- [33] M. D. MACLEOD, *Fast nearly ML estimation of the parameters of real or complex single tones or resolved multiple tones*, IEEE Transactions on Signal processing, 46 (1998), pp. 141–148.
- [34] I. MARAVIC AND M. VETTERLI, *Sampling and reconstruction of signals with finite rate of innovation in the presence of noise*, IEEE Transactions on Signal Processing, 53 (2005), pp. 2788–2805.
- [35] I. MARKOVSKY, *Low rank approximation: algorithms, implementation, applications*, Springer Science & Business Media, 2011.
- [36] P. MUKHOPADHYAY AND B. B. CHAUDHURI, *A survey of Hough transform*, Pattern Recognition, 48 (2015), pp. 993–1010.
- [37] L. M. MURPHY, *Linear feature detection and enhancement in noisy images via the Radon transform*, Pattern Recognition Letters, 4 (1986), pp. 279–284.

- [38] N. PARIKH, S. P. BOYD, ET AL., *Proximal algorithms*, Foundations and Trends in optimization, 1 (2014), pp. 127–239.
- [39] G. PEYRÉ, S. BOUGLEUX, AND L. COHEN, *Non-local regularization of inverse problems*, in European Conference on Computer Vision, Springer, 2008, pp. 57–68.
- [40] K. POLISANO, M. CLAUSEL, L. CONDAT, AND V. PERRIER, *Modélisation de textures anisotropes par la transformée en ondelettes monogènes et super-résolution de lignes 2-D*, PhD thesis, Université Grenoble Alpes, 2017.
- [41] K. POLISANO, L. CONDAT, M. CLAUSEL, AND V. PERRIER, *Convex super-resolution detection of lines in images*, 24th European Signal Processing Conference (EUSIPCO), Aug. 2016, pp. 336–340.
- [42] R. PRONY, *Essai experimental et analytique : sur les lois de la dilatabilité de fluides élastique et sur celles de la force expansive de la vapeur de l’alkool, à différentes températures*, Journal de l’Ecole Polytechnique Floréal et Plairial, an III, Volume 1, cahier 22 (1795), pp. 24–76.
- [43] B. G. QUINN AND E. J. HANNAN, *The estimation and tracking of frequency*, vol. 9, Cambridge University Press, 2001.
- [44] J. RADON, *1.1 über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeitenestimmung von funktionen durch ihre integralwerte längs gewisser mannigfaltigkeiten*, Classic papers in modern diagnostic radiology, 5 (2005).
- [45] M. RAHMAN AND K.-B. YU, *Total least squares approach for frequency estimation using linear prediction*, IEEE Transactions on Acoustics, Speech and Signal Processing, 35 (1987), pp. 1440–1454.
- [46] B. RECHT, M. FAZEL, AND P. A. PARRILO, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM review, 52 (2010), pp. 471–501.
- [47] B. RECHT, W. XU, AND B. HASSIBI, *Necessary and sufficient conditions for success of the nuclear norm heuristic for rank minimization*, in IEEE Conference on Decision and Control (CDC) 47th, IEEE, 2008, pp. 3065–3070.
- [48] H.-C. SO, K. W. CHAN, Y.-T. CHAN, AND K. HO, *Linear prediction approach for efficient frequency estimation of multiple real sinusoids: algorithms and analyses*, IEEE Transactions on Signal Processing, 53 (2005), pp. 2290–2305.
- [49] J.-L. STARCK, F. MURTAGH, AND J. M. FADILI, *Sparse image and signal processing: wavelets, curvelets, morphological diversity*, Cambridge university press, 2010.
- [50] P. STOICA, *List of references on spectral line analysis*, Signal Processing, 31 (1993), pp. 329–340.
- [51] P. STOICA AND R. MOSES, *Spectral Analysis of Signals*, Prentice Hall, NJ, 2005.

- [52] P. STOICA, R. L. MOSES, B. FRIEDLANDER, AND T. SODERSTROM, *Maximum likelihood estimation of the parameters of multiple sinusoids from noisy measurements*, IEEE Transactions on Acoustics, Speech, and Signal Processing, 37 (1989), pp. 378–392.
- [53] T. STROHMER, *Measure what should be measured: progress and challenges in compressive sensing*, IEEE Signal Processing Letters, 19 (2012), pp. 887–893.
- [54] G. TANG, B. N. BHASKAR, P. SHAH, AND B. RECHT, *Compressed sensing off the grid*, IEEE Transactions on Information Theory, 59 (2013), pp. 7465–7490.
- [55] O. TOEPLITZ, *Zur Theorie der quadratischen und bilinearen Formen von unendlichvielen Veränderlichen*, Mathematische Annalen, 70 (1911), pp. 351–376.
- [56] D. W. TUFTS AND R. KUMARESAN, *Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood*, Proceedings of the IEEE, 70 (1982), pp. 975–989.
- [57] J. A. URIGHIEN, Y. C. ELGAR, AND P. DRAGOTTI, *Sampling at the rate of innovation: Theory and applications*, Compressed Sensing: Theory and Applications, (2012), p. 148.
- [58] L. VANDENBERGHE AND S. BOYD, *Semidefinite programming*, SIAM review, 38 (1996), pp. 49–95.