Journal of Computational and Applied Mathematics 422 (2023) 114937

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

TV-based spline reconstruction with Fourier measurements: Uniqueness and convergence of grid-based methods

Thomas Debarre [a,*], Quentin Denoyelle [a], Julien Fageot [b]

[a] EPFL, Biomedical Imaging Group, Lausanne, Switzerland
[b] EPFL, AudioVisual Communications Laboratory, Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 3 March 2022
Received in revised form 22 August 2022

Keywords:
Generalized total-variation regularization
Inverse problems
Optimization
Splines
Fourier analysis
Grid-based algorithms

ABSTRACT

We study the problem of recovering piecewise-polynomial periodic functions from their low-frequency information. This means that we only have access to possibly corrupted versions of the Fourier samples of the ground truth up to a maximum cutoff frequency K_c. The reconstruction task is specified as an optimization problem with total-variation (TV) regularization (in the sense of measures) involving the Mth order derivative regularization operator $\mathrm{L} = \mathrm{D}^M$. The order $M \geq 1$ determines the degree of the reconstructed piecewise-polynomial spline, whereas the TV regularization norm, which is known to promote sparsity, guarantees a small number of pieces. We show that the solution of our optimization problem is always unique, which, to the best of our knowledge, is a first for TV-based problems. Moreover, we show that this solution is a periodic spline matched to the regularization operator L whose number of knots is upper-bounded by $2K_c$. We then consider the grid-based discretization of our optimization problem in the space of uniform L-splines. On the theoretical side, we show that any sequence of solutions of the discretized problem converges uniformly to the unique solution of the gridless problem as the grid size vanishes. Finally, on the algorithmic side, we propose a B-spline-based algorithm to solve the discretized problem, and we demonstrate its numerical feasibility experimentally. On both of these aspects, we leverage the uniqueness of the solution of the original problem.

© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

In recent years, TV[1]-based regularization techniques for continuous-domain inverse problems have shown to be very fruitful, with rapidly growing theoretical developments [2–5], algorithmic progress [6–8], and data science applications [9–11]. The goal of these techniques is to recover a continuous-domain signal of interest from a finite number of measurements. To tackle the obvious ill-posedness of this recovery problem, the prior assumption is that the signal is sparse in a certain basis. This prior is enforced by solving an optimization problem with a regularization term involving the TV norm.

In this work, we study the reconstruction of an unknown periodic real function $f_0 : \mathbb{T} \to \mathbb{R}$ from the knowledge of its possibly noise-corrupted low-frequency Fourier series coefficients, where $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z} = [0, 2\pi]$ is the torus whose

* Corresponding author.
E-mail addresses: thomas.debarre@gmail.com (T. Debarre), quentin.denoyelle@epfl.ch (Q. Denoyelle), julien.fageot@epfl.ch (J. Fageot).

[1] In this paper, the term "total variation" is understood in the sense of measure theory; note that it has a different meaning in other communities, *e.g.*, [1].

https://doi.org/10.1016/j.cam.2022.114937
0377-0427/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

two points 0 and 2π are identified. Fourier-domain measurements have historically been of particular interest in the TV-based community; some of the earliest works in this field focus specifically on this setting [2]. This interest is driven by the favorable theoretical recovery guarantees of Fourier-domain sampling as well as by potential applications such as MRI imaging, as demonstrated by existing discrete-domain sparsity-based methods [12].

Let $K_c \geq 0$ be the cutoff frequency; we therefore have access to

$$\boldsymbol{y} = (y_0, y_1, \dots, y_{K_c}) \in \mathbb{R} \times \mathbb{C}^{K_c}, \tag{1}$$

where $y_k \approx \widehat{f_0}[k]$, the kth Fourier-series coefficient of f_0. Note that since f_0 is a real function, $y_0 \in \mathbb{R}$ is approximately the mean $\widehat{f_0}[0] = \langle f_0, 1 \rangle$ of f_0, while $y_k \in \mathbb{C}$ for $k \neq 0$. Moreover, the Fourier series of f_0 is Hermitian symmetric, meaning that $\widehat{f_0}[-k] = \overline{\widehat{f_0}[k]} \in \mathbb{C}$ for every $k \in \mathbb{Z}$. The observation vector $\boldsymbol{y}$ in (1) has $2K_c + 1$ (real) degrees of freedom: one for the real mean y_0 and two for each of the other complex Fourier series coefficients.

1.1. Reconstruction via TV-based optimization

We formulate the reconstruction task as a regularized optimization problem with a sparsity prior, which is enforced via TV-based regularization. More precisely, the reconstruction f^* of f_0 is a solution of

$$f^* \in \arg\min_f E(\boldsymbol{y}, \boldsymbol{v}(f)) + \lambda \|\mathrm{L} f\|_{\mathcal{M}}, \tag{2}$$

where $\boldsymbol{y} \in \mathbb{R} \times \mathbb{C}^{K_c}$ is the observation vector; $\boldsymbol{v}(f)$ is the measurement vector

$$\boldsymbol{v}(f) = (\widehat{f}[0], \widehat{f}[1], \dots, \widehat{f}[K_c]) \in \mathbb{R} \times \mathbb{C}^{K_c}; \tag{3}$$

$E(\boldsymbol{y}, \cdot) : \mathbb{R} \times \mathbb{C}^{K_c} \to \mathbb{R}^+ \cup \{\infty\}$ is a data-fidelity functional which is a proper convex function, strictly convex over its effective domain,[2] lower semi-continuous (lsc), and coercive; $\|\cdot\|_{\mathcal{M}}$ is the total-variation (TV) norm on periodic Radon measures; and L is a regularization operator acting on periodic functions. The regularization operator specifies the transform domain in which the signal f_0 is assumed to be sparse. The composition of the TV norm with an operator L is known as *generalized TV regularization* [5]. For the sake of clarity, we focus on derivative operators of any order, *i.e.*, $\mathrm{L} = \mathrm{D}^M$ where $M \geq 1$, although our results can be extended to more general classes of operators such as fractional derivatives.

The data-fidelity term encourages the measurement vector $\boldsymbol{v}(f)$ to be close to the observations $\boldsymbol{y}$. A typical example is the quadratic functional

$$E(\boldsymbol{y}, \boldsymbol{v}(f)) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{v}(f)\|_2^2 = \frac{1}{2}\sum_{k=0}^{K_c} |y_k - \widehat{f}[k]|^2. \tag{4}$$

The data fidelity (4) is well-suited to an additive noise model where the measurements $\boldsymbol{y}$ are generated as $\boldsymbol{y} = \boldsymbol{v}(f_0) + \boldsymbol{n}$ with $\boldsymbol{n}$ a complex Gaussian vector (see [14, Section IV-B] for more details). Another case of interest is the indicator function $E(\boldsymbol{y}, \boldsymbol{v}(f)) = 0$ if $\boldsymbol{y} = \boldsymbol{v}(f)$ and ∞ otherwise, which leads to the constrained optimization problem[3] of the form

$$\arg\min_{f,\ \boldsymbol{v}(f)=\boldsymbol{y}} \|\mathrm{L} f\|_{\mathcal{M}}. \tag{5}$$

Other classical data-fidelity functionals can be found in [10, Section 7.5].

The choice of the TV norm promotes sparse and adaptive continuous-domain reconstruction, and has recently received a lot of attention (see Section 1.3). The operator L specifies the transform domain in which sparsity is enforced together with the regularity properties of the recovery. In the absence of a regularization operator L, Problem (2) leads to the recovery of periodic Dirac masses; this scenario is the subject of our previous paper [15]. In the latter, we thoroughly analyze the cases of uniqueness of the solution of the resulting optimization problem. In particular, we show that unlike in this manuscript, both cases (uniqueness and nonuniqueness) can occur, and we gave a necessary and sufficient condition on the data vector $\boldsymbol{y}$ that guarantees uniqueness.

1.2. Contributions

This paper is closely related to and relies on a result of our previous work [15], in which we also study problems of the form (2), but in the absence of a regularization operator L. Problems of the form (2), which deal with periodic functions, have previously been studied in [16] in a more general setting where the operators $\boldsymbol{v}$ and L can take various forms. In this paper, we focus on the specific case where $\boldsymbol{v}$ is a low-pass Fourier-domain operator and L is the Mth order derivative operator; this specificity allows us to prove additional theoretical results. The representer theorem from [16] guarantees

[2] The effective domain of a convex function $g : X \to \mathbb{R}^+ \cup \{\infty\}$ is the set $\{x \in X,\ g(x) < \infty\}$ [13].

[3] In this case, the value of the regularization parameter $\lambda > 0$ plays no role.

the existence of a solution to Problem (2), but without adjudicating on its uniqueness. Moreover, it gives the form of the extreme-point solution(s) as periodic L-splines, *i.e.*, functions f such that

$$\mathrm{L}f = \sum_{n=1}^{N} a_n \text{Ш}(\cdot - x_n) \tag{6}$$

is a finite sum of shifted Dirac combs, the distinct Dirac locations x_n being the knots of the spline (see Definition 1). Moreover, known proof techniques [5,16,17] allow us to show that the number of knots N is bounded by $N \leq 2K_c + 2$. Our contributions can be detailed as follows.

(i) Uniqueness of the Solution. Our main result is Theorem 1, in which we prove that the solution to Problem (2) is always unique. Moreover, we slightly improve the upper bound on the number of knots to $K \leq 2K_c$ in (6). To the best of our knowledge, Theorem 1 is the first systematic uniqueness result for the analysis of TV-based variational problems such as (2). This result has both theoretical and algorithmic implications, which we leverage in our other contributions. Our proof of Theorem 1 relies on a uniqueness result of our previous paper [15, Corollary 1] for Problem (2) in the absence of a regularization operator L. Interestingly, contrary to Problem (2), uniqueness is not systematic in that setting.

(ii) Uniform Convergence of Grid-Based Methods. We study the grid-based discretization of Problem (2). More precisely, we restrict its search space to the finite-dimensional space of uniform L-splines, *i.e.*, L-splines whose knots lie on a uniform grid. We show that as the grid gets finer, any sequence of solutions of the discretized problems converges in uniform norm towards the unique solution of Problem (2). This form of convergence is remarkably strong: in particular, it implies convergence for any L_p norm with $1 \leq p \leq \infty$. This result is a significant improvement over known convergence results for grid-based methods which exhibit weak*-type convergence such as [18, Proposition 4] and [19, Theorem 3].

(iii) Grid-Based Algorithm. We propose an adaptation of the B-spline-based algorithm developed in [19] to solve Problem (2). The main difference with the latter is the periodic setting, which does not add major difficulties but requires a different treatment of the boundaries. Thanks to our aforementioned uniform convergence result, the reconstructed signal is guaranteed to be uniformly close to the gridless solution when the grid is sufficiently fine. We provide some experimental results of our algorithm on some simulated data that demonstrate its numerical feasibility.

1.3. Related works

Optimization over radon measures. There exists a vast literature concerned with Dirac recovery using the TV norm as a regularizer, that is, problems of the form (2) in the absence of a regularization operator L. Some of the more recent works concerning this topic include [2,3,8,18,20–24]. However, our focus in this paper is on generalized TV regularization with a nontrivial operator L. We refer to the introduction of our previous paper [15] for a more detailed coverage of the Dirac recovery literature.

From sparse measures to splines and beyond. In recent years, several works have extended the TV-based Dirac recovery framework to smoother continuous-domain signals by considering generalized TV regularization, *i.e.*, problems such as (2) with a nontrivial regularization operator L. In [5], Unser et al. revealed the connection between the constrained Problem (2) (in a nonperiodic setting) and spline theory for general measurement functionals: the extreme-point solutions are necessarily L-splines. This result was revisited, extended, and refined by several authors [10,16,25–31]. This manuscript will strongly rely on the periodic theory of TV-based optimization problems recently developed in [16], of which Problem (2) is a special case.

The purpose of most of these works is to describe the solution sets of certain relevant optimization problems, which are typically nonunique. However, in our setting, as we show in Theorem 1, the solution to Problem (2) is always unique. The closest work in this direction is our recent paper [32], where we provide a full description of the solution set of nonperiodic TV-based optimization problems with a regularization operator $\mathrm{L} = \mathrm{D}^2$ (which leads to piecewise-linear reconstructions), and spatial sampling measurements $\boldsymbol{\nu}$. This study includes the characterization of the cases of uniqueness [32, Proposition 6 and Theorem 2], which, contrary to Problem (2), is not systematic.

Convergence results and algorithms for discretized problems. The convergence of discretized optimization schemes to the solutions of continuous-domain TV-regularized problems has been studied by several authors, such as [3,4,7,8,33]. Grid-based methods have specifically been considered in [4,18,23]. In these works, the authors prove convergence results in the weak* sense, which is suited to the space of Radon measures, in a setting where no systematic uniqueness results are known. To the best of our knowledge, our work is the first to prove the convergence of solutions of discretized generalized TV-based problems towards the solution of the original problem in a such a strong sense as the uniform norm. To achieve this, we leverage our uniqueness result of Theorem 1. On the algorithmic side, grid-based methods to solve optimization problems with TV-based regularization have been proposed in [19,25,34–36].

1.4. Outline

The paper is organized as follows. Section 2 introduces the mathematical material used in this paper. In Section 3, we present our optimization Problem (2) of interest and prove that it always has a unique solution. In Section 4, we present our proposed grid-based discretization of Problem (2), and prove that its solutions converge uniformly to that of the original problem when the grid size goes to zero. We present our method for solving this discretized problem using a B-spline basis in Section 5. Finally, we exemplify our results on simulations in Section 6.

2. Mathematical preliminaries

2.1. Periodic functions and periodic splines

We first introduce some notations and recall some basic facts concerning periodic functions and their Fourier series. More details can be found in [16, Section 2]. The Schwartz space of infinitely smooth periodic function is denoted by $\mathcal{S}(\mathbb{T})$, endowed with its usual Fréchet topology. Its topological dual is the space of periodic generalized functions $\mathcal{S}'(\mathbb{T})$.

For $k \in \mathbb{Z}$, let $e_k : \mathbb{T} \to \mathbb{C}$ be the complex exponential function $e_k(x) = \exp(\mathrm{i}kx)$, which is clearly in $\mathcal{S}(\mathbb{T})$. The Fourier series coefficients of $f \in \mathcal{S}'(\mathbb{T})$ are given by $\widehat{f}[k] = \langle f, e_k \rangle \in \mathbb{C}$. For a real function f, these coefficients are Hermitian symmetric, *i.e.*, $\widehat{f}[-k] = \overline{\widehat{f}[k]}$ for all $k \in \mathbb{Z}$, which implies in particular that $\widehat{f}[0] \in \mathbb{R}$. We then have that $f = \sum_{k\in\mathbb{Z}} \widehat{f}[k] e_k$ for any $f \in \mathcal{S}'(\mathbb{T})$, where the convergence is in $\mathcal{S}'(\mathbb{T})$. The Dirac stream is defined as $\mathrm{III} = \sum_{n\in\mathbb{Z}} \delta(\cdot - 2\pi n)$. Its Fourier coefficients are $\widehat{\mathrm{III}}[k] = 1$ for any $k \in \mathbb{Z}$. The derivative operator is denote by D. More generally, we consider the Mth-order derivative operator $\mathrm{L} = \mathrm{D}^M$ for a fixed integer $M \geq 1$. We then have that $\mathrm{L}f = \sum_{k\in\mathbb{Z}} (\mathrm{i}k)^M \widehat{f}[k] e_k$.

We define the *periodic Green function* g_{L} of $\mathrm{L} = \mathrm{D}^M$ as the function

$$g_{\mathrm{L}}(x) = \sum_{k\neq 0} \frac{e_k}{(\mathrm{i}k)^M}. \tag{7}$$

Then, g_{L} is the unique periodic and zero-mean function such that $\mathrm{D}^M g_{\mathrm{L}} = \mathrm{III} - 1$. It is worth noting that g_{L} is not a Green's function in the usual sense: there is no periodic function g such that $\mathrm{D}^M g = \mathrm{III}$, since $\mathrm{D}^M g$ necessarily has zero mean, whereas the Dirac stream does not. See [16, Section 2.2] for more details on this matter.

Definition 1. Let $M \geq 1$ and $\mathrm{L} = \mathrm{D}^M$. We say that f is a *periodic* L*-spline* (or simply a L-spline) if

$$\mathrm{L}f = w = \sum_{n=1}^{N} a_n \mathrm{III}(\cdot - x_n) \tag{8}$$

where $N \geq 0$, $a_n \in \mathbb{R}\backslash\{0\}$, and the knots $x_n \in \mathbb{T}$ are pairwise distinct. We call w the *innovation* of the L-spline f.

A function f satisfies (8) if and only if

$$f = a_0 + \sum_{n=1}^{N} a_n g_{\mathrm{L}}(\cdot - x_n) \tag{9}$$

for some $a_0 \in \mathbb{R}$. In this case, we necessarily have that $\sum_{n=1}^{N} a_n = 0$. This is a particular case of [16, Proposition 2.8] and simply follows from taking the mean (or 0th Fourier coefficient) in (8), giving $0 = \widehat{L}[0]\widehat{f}[0] = \sum_{n=1}^{N} a_n$. It is worth noting that the Green's function g_{L} is *not* a L-spline. However, for any $x_0 \in \mathbb{T} \setminus \{0\}$, $g_{\mathrm{L}} - g_{\mathrm{L}}(\cdot - x_0)$ is a periodic L-spline.

A L-spline is a periodic piecewise-polynomial function of degree at most $(M-1)$ and with $(M-2)$ continuous derivatives. The case $M = 1$ corresponds to piecewise-constant functions, while $M = 2$ leads to piecewise-linear continuous functions.

2.2. Periodic Radon measures and native spaces

Let $\mathcal{M}(\mathbb{T})$ be the space of periodic Radon measures. By the Riesz-Markov theorem [37], it is the continuous dual of the space $\mathcal{C}(\mathbb{T})$ of continuous periodic functions endowed with the supremum norm. The total-variation norm on $\mathcal{M}(\mathbb{T})$, for which it forms a Banach space, is given by

$$\|w\|_{\mathcal{M}} = \sup_{f\in\mathcal{C}(\mathbb{T}),\ \|f\|_\infty \leq 1} \langle w, f \rangle. \tag{10}$$

We denote by $\mathcal{M}_0(\mathbb{T})$ the set of Radon measures with zero mean, *i.e.* $\mathcal{M}_0(\mathbb{T}) = \{w \in \mathcal{M}(\mathbb{T}),\ \widehat{w}[0] = 0\}$. It is the continuous dual of the space $\mathcal{C}_0(\mathbb{T}) = \{f \in \mathcal{C}(\mathbb{T}),\ \widehat{f}[0] = 0\}$ of continuous functions with zero mean.

Let $\mathrm{L} = \mathrm{D}^M$ for some $M \geq 1$. We define the *native space* associated to L as

$$\mathcal{M}_{\mathrm{L}}(\mathbb{T}) = \{f \in \mathcal{S}'(\mathbb{T}) :\ \mathrm{L}f \in \mathcal{M}(\mathbb{T})\}. \tag{11}$$

Periodic native spaces have been studied for general spline-admissible operators (*i.e.*, periodic operators with finite-dimensional null space and which admit a pseudoinverse) in [16, Section 3]. Proposition 1 recalls some important properties of native spaces for the particular case of the Mth order derivative operator. For this purpose, we define the pseudo-inverse operator $\mathrm{L}^\dagger$ such that

$$\mathrm{L}^\dagger f = \sum_{k\neq 0} \frac{\widehat{f}[k]}{(\mathrm{i}k)^M} e_k \tag{12}$$

for any $f \in \mathcal{S}'(\mathbb{T})$. In particular, we have that $\mathrm{L}^\dagger \text{Ш} = g_{\mathrm{L}}$.

Proposition 1 (*Theorem 3.2 in [16]*)**.** *Let $\mathrm{L} = \mathrm{D}^M$ for some $M \geq 1$. We have the direct-sum relation*

$$\mathcal{M}_{\mathrm{L}}(\mathbb{T}) = \mathrm{L}^\dagger \mathcal{M}_0(\mathbb{T}) \oplus \mathrm{Span}\{1\}, \tag{13}$$

and any $f \in \mathcal{M}_{\mathrm{L}}(\mathbb{T})$ has a unique decomposition as

$$f = \mathrm{L}^\dagger w + a \tag{14}$$

where $w \in \mathcal{M}_0(\mathbb{T})$ and $a \in \mathbb{R}$ are given by $w = \mathrm{L}f$ and $a = \widehat{f}[0]$. Then, $\mathcal{M}_{\mathrm{L}}(\mathbb{T})$ is a Banach space for the norm

$$\|f\|_{\mathcal{M}_{\mathrm{L}}} = \|w\|_{\mathcal{M}} + |a|. \tag{15}$$

3. Uniqueness of TV-based penalized problems

It is well known that TV-based optimization problems with regularization operators lead to splines solutions [5]. This is both an existence result and a *representer theorem*, which provides the form of the (extreme-point) solutions of the optimization task. Here, we focus on problems of the form (2), whose main specificity compared to related works in the literature is the periodic setting and the Fourier-domain measurement operator $\boldsymbol{\nu}$.

We now state our main result, which guarantees the uniqueness of the solution to Problem (2). The proof relies on a result of our previous paper that focuses on the recovery of Dirac impulses from low-frequency Fourier measurements. We recall this result here.

Proposition 2. *[15, Corollary 1] If $|y_0| < |y_{k_0}|$ for some $k_0 \neq 0$, then the optimization problem*

$$\min_{w\in\mathcal{M}(\mathbb{T}),\ \boldsymbol{\nu}(w)=\boldsymbol{y}} \|w\|_{\mathcal{M}} \tag{16}$$

has a unique solution, which is the sum of at most $2K_c$ Dirac impulses.

Our proof technique consists in reformulating Problem (2) over the space of Radon measures and applying Proposition 2 to this equivalent problem. Interestingly, the regularization operator $\mathrm{L} = \mathrm{D}^M$ leads to systematic uniqueness, which is not true of Problem (16) (where there is no regularization operator) in the general case.

Theorem 1. *Let $\mathrm{L} = \mathrm{D}^M$ with $M \geq 1$, $K_c \geq 0$ be the cutoff frequency of the low-pass filter $\boldsymbol{\nu} : \mathcal{M}_{\mathrm{L}}(\mathbb{T}) \to \mathbb{R} \times \mathbb{C}^{K_c}$ defined in (3), $\boldsymbol{y} \in \mathbb{R} \times \mathbb{C}^{K_c}$, $E(\boldsymbol{y}, \cdot) : \mathbb{R} \times \mathbb{C}^{K_c} \to \mathbb{R}^+$ be a functional which is a proper convex function, strictly convex over its effective domain, lsc, and coercive, and $\lambda > 0$. Then, the optimization problem*

$$\mathcal{V}_\lambda(\boldsymbol{y}) = \operatorname*{arg\,min}_{f\in\mathcal{M}_{\mathrm{L}}(\mathbb{T})} E(\boldsymbol{y}, \boldsymbol{\nu}(f)) + \lambda \|\mathrm{L}f\|_{\mathcal{M}}, \tag{17}$$

admits a unique solution which is a L*-spline whose number of knots is bounded by $2K_c$.*

Proof. Using a classical argument based on the strict convexity of $E(\boldsymbol{y}, \cdot)$ (see for instance [32, Proposition 7]), we deduce that all solutions f^* of Problem (17) share an identical observation vector $\boldsymbol{y}_\lambda \in \mathbb{R}^M$, that is, $\forall f^* \in \mathcal{V}_\lambda(\boldsymbol{y})$, we have $\boldsymbol{\nu}(f^*) = \boldsymbol{y}_\lambda$. Hence, Problem (17) is equivalent to

$$\mathcal{V}_\lambda(\boldsymbol{y}) = \operatorname*{arg\,min}_{f\in\mathcal{M}_{\mathrm{L}}(\mathbb{T}),\ \boldsymbol{\nu}(f)=\boldsymbol{y}_\lambda} \|\mathrm{L}f\|_{\mathcal{M}}. \tag{18}$$

By Proposition 1, any $f \in \mathcal{M}_{\mathrm{L}}(\mathbb{T})$ admits a unique decomposition $f = \mathrm{L}^\dagger w + a \in \mathcal{M}_{\mathrm{L}}(\mathbb{T})$ with $(w, a) \in \mathcal{M}_0(\mathbb{T}) \times \mathbb{R}$. By plugging in this expansion into the cost functional of Problem (18), we get that the latter is equivalent to

$$\operatorname*{arg\,min}_{(w,a)\in\mathcal{M}_0(\mathbb{T})\times\mathbb{R},\ \boldsymbol{\nu}(\mathrm{L}^\dagger w+a)=\boldsymbol{y}_\lambda} \|w\|_{\mathcal{M}}. \tag{19}$$

We then observe that $\boldsymbol{\nu}(\mathrm{L}^\dagger w + a) = \left(a, \frac{\widehat{w}[1]}{\widehat{L}[1]}, \ldots, \frac{\widehat{w}[K_c]}{\widehat{L}[K_c]}\right)$ and thus that (w^*, a^*) is solution of (19) if and only if $a^* = (\boldsymbol{y}_\lambda)_0$ and w^* is solution of the new problem

$$\operatorname*{arg\,min}_{w\in\mathcal{M}(\mathbb{T}),\ \boldsymbol{\nu}(w)=\boldsymbol{z}} \|w\|_{\mathcal{M}} \tag{20}$$

where $\boldsymbol{z} \in \mathbb{R} \times \mathbb{C}^{K_c}$ is given by $z_0 = 0$ and $z_k = \widehat{L}[k](\boldsymbol{y}_\lambda)_k$ for $k \neq 0$. Note that the condition $w \in \mathcal{M}_0(\mathbb{T})$ in (19) (*i.e.*, w has zero mean) has been removed in (20) since it is automatically satisfied via the constraint $\boldsymbol{v}(w) = \boldsymbol{z}$ which gives $\widehat{w}[0] = z_0 = 0$.

We now prove that Problem (20) has a unique solution w^* which is a sum of at most $2K_c$ Dirac impulses. If $\boldsymbol{z} = \boldsymbol{0}$, then the result trivially holds, the unique solution being $w^* = 0$. We now assume that $\boldsymbol{z} \neq \boldsymbol{0}$. In this case, Problem (20) has a nonzero optimal value due to the fact that $w = 0$ cannot be a solution. Since $z_0 = 0$ and $\boldsymbol{z} \neq \boldsymbol{0}$, Problem (19) satisfies the assumptions of Proposition 2. This proves the uniqueness of w^* and the fact that it is a sum of at most $2K_c$ Dirac impulses. This in turn implies the uniqueness of the solution $f^* = \mathrm{L}^\dagger w^* + (\boldsymbol{y}_\lambda)_0$ as well as the fact that it is an L-spline with at most $2K_c$ knots. □

Remark. Theorem 1 remains valid for more general operators L, namely any spline-admissible operator in the sense of [16, Definition 2] whose null space includes constant functions, *i.e.*, $\mathrm{L}\{1\} = 0$.

Theorem 1 has three components: (i) it guarantees the uniqueness of the solution, it provides (ii) the form of the solution and (iii) an upper bound on the number of knots of the solution. The first item, arguably the most striking one, is completely new; existing results typically provide the form of extreme-point solutions of the problem. We are not aware of any other systematic uniqueness results concerning inverse problems with TV-based regularization in the literature. The second item is already known; it has been proved for our setting in [16, Theorem 4]. Finally, concerning the third item, known proof techniques [5,16,17] allow us to reach the bound $2K_c + 1$, which we improve to $2K_c$.

One can actually be slightly more precise and show that the mean of the solution is known under very mild conditions on the cost functional E. Under this assumption, we also provide a reformulation of Problem (17) over the space of Radon measures.

Proposition 3. *We assume that we are under the conditions of Theorem 1 and that the data-fidelity cost functional E is such that for any fixed $(z_1, \dots, z_{K_c}) \in \mathbb{C}^{K_c}$, we have*

$$y_0 = \operatorname*{arg\,min}_{z_0 \in \mathbb{R}} \; E(\boldsymbol{y}, \boldsymbol{z}) \tag{21}$$

where $\boldsymbol{y} = (y_0, y_1, \dots, y_{K_c}) \in \mathbb{R} \times \mathbb{C}^{K_c}$ and $\boldsymbol{z} = (z_0, z_1, \dots, z_{K_c}) \in \mathbb{R} \times \mathbb{C}^{K_c}$. Then, the unique solution f^* to (17) admits the decomposition $f^* = y_0 + \mathrm{L}^\dagger w^*$ where

$$w^* = \operatorname*{arg\,min}_{w \in \mathcal{M}_0(\mathbb{T})} E(\boldsymbol{y}, \boldsymbol{v}(\mathrm{L}^\dagger w + y_0)) + \lambda \|w\|_{\mathcal{M}}. \tag{22}$$

In particular, this implies that $\widehat{f^*}[0] = y_0$.

Proof. Similarly to our manipulation in (19), Problem (17) is equivalent to

$$(w^*, a^*) = \operatorname*{arg\,min}_{(w,a) \in \mathcal{M}_0(\mathbb{T}) \times \mathbb{R}} \; E(\boldsymbol{y}, \boldsymbol{v}(\mathrm{L}^\dagger w + a)) + \lambda \|w\|_{\mathcal{M}}, \tag{23}$$

with $f^* = \mathrm{L}^\dagger w^* + a^*$. Problem (23) has a unique solution due to that of Problem (17) (proved in Theorem 1), and to the uniqueness of the decomposition of f^* (Proposition 1). Then, we have that $\boldsymbol{v}(a^* + \mathrm{L}^\dagger w^*) = (a^*, \widehat{\mathrm{L}^\dagger w^*}[1], \dots, \widehat{\mathrm{L}^\dagger w^*}[K_c])$, which by (21) implies that $E(\boldsymbol{y}, \boldsymbol{v}(a^* + \mathrm{L}^\dagger w^*)) \geq E(\boldsymbol{y}, \boldsymbol{v}(y_0 + \mathrm{L}^\dagger w^*))$, with equality if and only if $a^* = y_0$. Hence, since the constant a does not impact the regularization in (22), we must have that $a^* = \widehat{f^*}[0] = y_0$. Problem (23) can thus be rewritten as (22). □

Remark. The relation (21) holds for virtually all classical cost functionals, including any ℓ_p norm-based cost such as the quadratic data fidelity (4), or any separable cost whose minimum over each component is reached when $y_m = z_m$, such as indicator functions. Proposition 3 ensures that the mean of the solution of Problem (17) is given by $\widehat{f^*}[0] = y_0$.

4. Uniform convergence of grid-based methods

A common way to solve infinite-dimensional continuous-domain problems such as (24) algorithmically is to discretize them using a uniform finite grid [18,23]. In this section, we propose such a discretization method of the problem

$$f^* = \operatorname*{arg\,min}_{f \in \mathcal{M}_{\mathrm{L}}(\mathbb{T})} \frac{1}{2} \|\boldsymbol{y} - \boldsymbol{v}(f)\|_2^2 + \lambda \|\mathrm{L}f\|_{\mathcal{M}}, \tag{24}$$

that is, Problem (17) with a quadratic data-fidelity cost $E(\boldsymbol{y}, \boldsymbol{z}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{z}\|_2^2$. Note that we no longer denote the solution of Problem (24) as a set but as a function f^*, since Theorem 1 guarantees that this solution is unique. We restrict to the case of the quadratic data fidelity for the sake of simplicity, although our results hereafter hold for more general choices of E. Our choice clearly satisfies the assumption of Proposition 3, hence the solution f^* of (24) satisfies $\widehat{f^*}[0] = y_0$.

Our discretization method, which was introduced for similar problems in [19], consists in restricting the search space of Problem (24) to the space of uniform L-splines $\mathcal{M}_{\mathrm{L}, \frac{2\pi}{P}}(\mathbb{T})$, *i.e.*, L-splines in the sense of Definition 1 with knots x_n on a

uniform grid. The space $\mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})$ is defined as follows for a grid size $h = \frac{2\pi}{P}$, where $P \in \mathbb{N}$, $P \geq 1$, is the number of grid points:

$$\mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T}) = \left\{ f \in \mathcal{S}'(\mathbb{T}),\ \mathrm{L}f = \sum_{p=0}^{P-1} a[p] \text{Ш}\left(\cdot - \frac{2\pi p}{P}\right) \right\}. \tag{25}$$

Our choice of restricting the search space of Problem (24) to $\mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})$ is guided by Theorem 1, which states that the unique solution to this problem is a L-spline. Hence, this choice of space is compatible with the sparsity-promoting regularization $\|\mathrm{L}\cdot\|_{\mathcal{M}}$. Although in general, the solution of our problem does not have knots on a uniform grid, it can be approximated arbitrary closely with an element of $\mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})$ when P is large. The other main feature of our method is that the computations are exact in the continuous domain, both those of the forward model and of the regularization term. Our discretized optimization problem then becomes

$$\mathcal{V}_{\lambda,P}(\boldsymbol{y}) = \underset{f \in \mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})}{\arg\min} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{v}(f)\|_2^2 + \lambda \|\mathrm{L}f\|_{\mathcal{M}}. \tag{26}$$

Note that contrary to the original Problem (24), the solution set $\mathcal{V}_{\lambda,P}(\boldsymbol{y})$ of the discretized problem (26) is not necessarily unique.

As we shall demonstrate in Section 5, Problem (24) can be solved algorithmically with standard finite-dimensional solvers. However, the important question of how well it approximates the original Problem (24) still remains. We answer this question in Theorem 2 by proving that any sequence of elements of $\mathcal{V}_{\lambda,P}(\boldsymbol{y})$ converge in a strong sense – namely, uniform convergence – towards f^* when $P \to \infty$.

Theorem 2. *Let* $\mathrm{L} = \mathrm{D}^M$ *with* $M \geq 2$, $\boldsymbol{y} \in \mathbb{R}^+ \times \mathbb{C}^{K_c}$, *and* $\lambda > 0$. *We denote by* f^* *the unique solution to* (24). *For any* $P \geq 1$, *we set* $f_P^* \in \mathcal{V}_{\lambda,P}(\boldsymbol{y})$. *Then, we have that*

$$\|f^* - f_P^*\|_\infty \underset{P\to\infty}{\longrightarrow} 0. \tag{27}$$

Remark 1. Despite the fact that the solutions to (26) may not be unique, Theorem 2 ensures that the convergence (27) holds for any choice of the f_P^*.

Remark 2. Uniform convergence implies convergence with respect to any L_p norm for $1 \leq p \leq \infty$, since we have $\|f\|_p \leq (2\pi)^{1/p}\|f\|_\infty$ for any $f \in \mathcal{M}_{\mathrm{L}}(\mathbb{T})$.

Remark 3. Theorem 2 holds for more general settings than Problem (24). More specifically, our proof seamlessly extends to the more general setting of Theorem 1 for any cost functional E that is continuous with respect to its second argument, such as ℓ_p losses of the form $E(\boldsymbol{y}, \boldsymbol{z}) = \|\boldsymbol{y} - \boldsymbol{z}\|_p^p$. Compared to the setting of Theorem 1, this notably excludes indicator functions, *i.e.*, the constrained optimization Problem (18). Concerning the regularization operator L, Theorem 2 readily extends to any operator L such that $\mathrm{L}\{1\} = 0$ and whose periodic Green's function is Lipschitz. This notably excludes the case $\mathrm{L} = \mathrm{D}$, *i.e.*, $M = 1$.

Proof. We first introduce $\mathcal{M}_{0,\frac{2\pi}{P}(\mathbb{T})} = \{w \in \mathcal{M}_0(\mathbb{T}),\ w = \sum_{p=0}^{P-1} a[p]\text{Ш}(\cdot - \frac{2\pi p}{P})\}$, the uniform discretization of $\mathcal{M}_0(\mathbb{T})$ using Dirac impulses. Then, using Proposition 3 (with a restriction of the search space which does not affect the proof), we have that $f_P^* = y_0 + \mathrm{L}^\dagger w_P^*$, where

$$w_P^* \in \underset{w \in \mathcal{M}_{0,\frac{2\pi}{P}}(\mathbb{T})}{\arg\min} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{v}(\mathrm{L}^\dagger w + y_0)\|_2^2 + \lambda\|w\|_{\mathcal{M}}. \tag{28}$$

We now prove that the Radon measures w_P^* converge towards the unique solution of

$$w^* = \underset{w \in \mathcal{M}_0(\mathbb{T})}{\arg\min} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{v}(\mathrm{L}^\dagger w + y_0)\|_2^2 + \lambda\|w\|_{\mathcal{M}} \tag{29}$$

for the weak* topology when $P \to \infty$, where the uniqueness of w^* follows from (22) in Proposition 3. This convergence is proved by following [18, Proposition 4]; the fact that the search space in (29) is $\mathcal{M}_0(\mathbb{T})$ rather than $\mathcal{M}(\mathbb{T})$ does not impact the proof. Then, the operator $\mathrm{L}^\dagger$ is linear and continuous between $\mathcal{M}_0(\mathbb{T})$ and $\mathcal{M}_{\mathrm{L}}(\mathbb{T})$ for their respective weak* topologies. This implies that f_P^* converges to f^* for the weak* topology over $\mathcal{M}_{\mathrm{L}}(\mathbb{T})$. According to [16, Proposition 9], $\mathrm{L} = \mathrm{D}^M$ is sampling-admissible for $M \geq 2$, which implies in particular that Ш is in the predual of $\mathcal{M}_{\mathrm{L}}(\mathbb{T})$. Equivalently, this implies that $f \mapsto f(x)$ is weak*-continuous over $\mathcal{M}_{\mathrm{L}}(\mathbb{T})$, which implies that $f_P^*(x) \to f^*(x)$ for any $x \in \mathbb{T}$ (pointwise convergence).

We now prove that the family $(f_P^*)_{P\in\mathbb{N}}$ is equicontinuous. Then, by [38, Theorem 15, Chapter 7], pointwise and uniform convergences are equivalent, which will conclude the proof. Since f_P^* is a L-spline, using the expansion (9), we have

$$f_P^* = y_0 + \mathrm{L}^\dagger w_P^* = y_0 + \sum_{n=1}^{N_p} a_P[n] g_\mathrm{L}(\cdot - x_{P,n}), \quad x_{P,n} \in \left\{ \frac{2\pi p}{P},\ 0 \le p \le P-1 \right\} \tag{30}$$

for some coefficients $a_P[n]$, $1 \le n \le N_p$, where g_L is the Green's function of L defined in (7). Moreover, g_L is a periodic Lipschitz function for $\mathrm{L} = \mathrm{D}^M$ and $M \ge 2$, hence $\|g_\mathrm{L}\|_{\mathrm{Lip}} := \sup_{x,y\in\mathbb{R},\ x\neq y} \frac{|g_\mathrm{L}(x)-g_\mathrm{L}(y)|}{|x-y|} < \infty$. For any $x, y \in \mathbb{R}$, we have that

$$\begin{aligned} |f_P^*(x) - f_P^*(y)| &\le \sum_{n=1}^{N_p} |a_P[n]| |g_\mathrm{L}(x - x_{P,n}) - g_\mathrm{L}(y - x_{P,n})| \le \left(\sum_{n=1}^{N_p} |a_P[n]| \right) \|g_\mathrm{L}\|_{\mathrm{Lip}} |x-y| \\ &= \left\| w_P^* \right\|_{\mathcal{M}} \|g_\mathrm{L}\|_{\mathrm{Lip}} |x-y|. \end{aligned} \tag{31}$$

We have seen that $w_P^* \to w^*$ when $P \to \infty$ for the weak* topology. It is therefore bounded for the total-variation norm, thanks to the uniform boundedness principle. We therefore deduce from (31) that the f_P^* are uniformly Lipschitz, and therefore equicontinuous, which proves the desired result. □

The first part of the proof of Theorem 2, dealing with the pointwise convergence, mostly relies on the generalization of the weak* convergence studied in [18, Proposition 4]. Duval and Peyré use tools from Γ-convergence (see [39] for an introduction) and are themselves inspired by [40].

Theorem 2 shows that our grid-based discretization yields spline solutions that are arbitrarily close to the unique solution f^* of (24) in the uniform sense when the discretization step $h = \frac{2\pi}{P}$ vanishes. It leverages the uniqueness of the spline reconstruction from Fourier measurements ensured by Theorem 1.

5. B-spline-based algorithm

We now introduce our proposed algorithm to solve the discretized Problem (26) in an *exact* way, *i.e.*, without any discretization error. The algorithm is based on [19] and uses the B-spline basis to represent the space of uniform splines $\mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})$. The main difference here with [19] is the periodic setting, which actually simplifies the treatment of the boundary conditions. Moreover, for the sake of conciseness, we focus here on discretizing for a fixed grid; we do not present the multiresolution aspect of the algorithm introduced in [19], although it can seamlessly be adapted to our setting.

5.1. Preliminaries on uniform periodic polynomial splines

A convenient feature of the space $\mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})$ is that it is generated by periodic B-splines, as will be proved in Proposition 4. To this end, we first provide some background information on B-splines and their periodized versions. B-splines are popular basis functions [41] that are widely used in signal processing applications [42,43], in part due to their short support which leads to well-conditioned optimization tasks. In the non-periodic setting, the scaled B-spline $\beta_{\mathrm{L},h}$ of the operator $\mathrm{L} = \mathrm{D}^M$ with grid size $h > 0$ is characterized by its Fourier transform

$$\widehat{\beta}_{\mathrm{L},h}(\omega) = \frac{1}{h^{M-1}} \left(\frac{1 - \mathrm{e}^{\mathrm{i}\omega h}}{\mathrm{i}\omega} \right)^M, \quad \forall \omega \in \mathbb{R}. \tag{32}$$

The scaled B-spline $\beta_{\mathrm{L},h}$ is a piecewise polynomial of order $(M-1)$ with continuous $(M-2)$-th derivative, and is supported over the interval $[0, hM]$.

For any integer $P \ge 1$, the periodized L B-spline with grid size $h = \frac{2\pi}{P}$ is then defined as

$$\beta^{\mathrm{per}}_{\mathrm{L},\frac{2\pi}{P}}(x) = \sum_{k\in\mathbb{Z}} \beta_{\mathrm{L},\frac{2\pi}{P}}(x - 2\pi k), \quad \forall x \in \mathbb{T}, \tag{33}$$

which is a converging sum due to the finite support of $\beta_{\mathrm{L},\frac{2\pi}{P}}$. In fact, for $P \ge M$, the periodic B-spline is not aliased, since we have $\mathrm{Supp}(\beta_{\mathrm{L},\frac{2\pi}{P}}) = [0, M\frac{2\pi}{P}] \subset [0, 2\pi] = \mathbb{T}$: we thus have $\beta^{\mathrm{per}}_{\mathrm{L},\frac{2\pi}{P}} = \beta_{\mathrm{L},\frac{2\pi}{P}}$ in the interval $\mathbb{T}$. Clearly, $\beta^{\mathrm{per}}_{\mathrm{L},\frac{2\pi}{P}}$ is 2π-periodic, and one readily shows from standard Fourier analysis that its Fourier series coefficients are given by

$$\widehat{\beta}^{\mathrm{per}}_{\mathrm{L},\frac{2\pi}{P}}[k] = \frac{1}{2\pi} \widehat{\beta}_{\mathrm{L},\frac{2\pi}{P}}(k) = P^{M-1} \left(\frac{1 - \mathrm{e}^{\mathrm{i}k\frac{2\pi}{P}}}{2\pi \mathrm{i}k} \right)^M. \tag{34}$$

Moreover, $\beta^{\mathrm{per}}_{\mathrm{L},\frac{2\pi}{P}}$ is a periodic L-spline in the sense of Definition 1, and its innovation is given by

$$\mathrm{L}\beta^{\mathrm{per}}_{\mathrm{L},\frac{2\pi}{P}} = \left(\frac{P}{2\pi} \right)^{M-1} \sum_{m=0}^{M} d_\mathrm{L}[m] \text{Ш}(\cdot - m\frac{2\pi}{P}), \tag{35}$$

where the P-periodic sequence d_{L} is characterized by its discrete Fourier transform (DFT) $D_{\mathrm{L}}[k] = (1 - \mathrm{e}^{-\mathrm{i}k\frac{2\pi}{P}})^M$. This relation is easily verified in the Fourier domain. As an example, for $\mathrm{L} = \mathrm{D}$, d_{L} is the P-periodized finite-difference sequence $d_{\mathrm{L}}[k] = \delta_P[k] - \delta_P[k-1]$ where $\delta_P[k]$ is the P-periodized Kronecker delta sequence. As stated earlier, periodic B-splines share the same celebrated property as regular B-splines: they are generators of the space of uniform (periodic) splines $\mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})$ introduced in (25).

Proposition 4. *The periodic B-spline $\beta^{\mathrm{per}}_{\mathrm{L},\frac{2\pi}{P}}$ is a generator of the space $\mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})$,* i.e., *we have*

$$\mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T}) = \left\{ f = \sum_{p=0}^{P-1} c[p]\beta^{\mathrm{per}}_{\mathrm{L},\frac{2\pi}{P}}\left(\cdot - \frac{2\pi p}{P}\right),\ \boldsymbol{c} = (c[0],\dots,c[P-1]) \in \mathbb{R}^P \right\}. \tag{36}$$

Proof. We first observe that the space $\mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})$ is a P-dimensional vector space: there are $(P-1)$ degrees of freedom for the $a[p]$ coefficients in (25) (P coefficients and one linear constraint $\sum_{p=0}^{P-1} a[p] = 0$), and one for the mean $\widehat{f}[0]$ (see Proposition 1). Next, we prove that for any $\boldsymbol{c} = (c[0],\dots,c[P-1]) \in \mathbb{R}^P$, we have $f = \sum_{p=0}^{P-1} c[p]\beta^{\mathrm{per}}_{\mathrm{L},\frac{2\pi}{P}}(\cdot - \frac{2\pi p}{P}) \in \mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})$. Indeed, we have

$$\begin{aligned}
\mathrm{L}f &= \left(\frac{P}{2\pi}\right)^{M-1} \sum_{p=0}^{P-1}\sum_{m=0}^{M} c[p]d_{\mathrm{L}}[m]\amalg\left(\cdot - (p+m)\frac{2\pi}{P}\right) \\
&= \left(\frac{P}{2\pi}\right)^{M-1} \sum_{p=0}^{P-1} (\boldsymbol{d}_{\mathrm{L}} * \boldsymbol{c})[p]\amalg\left(\cdot - \frac{2\pi p}{P}\right),
\end{aligned} \tag{37}$$

where (35) was used for the first line, and $(\boldsymbol{d}_{\mathrm{L}} * \boldsymbol{c})$ denotes here the cyclic convolution between the vectors $\boldsymbol{d}_{\mathrm{L}} = (d_{\mathrm{L}}[0],\dots,d_{\mathrm{L}}[P-1])$ and $\boldsymbol{c}$. This proves that $f \in \mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})$ with coefficients $(a[0],\dots,a[P-1]) = \left(\frac{P}{2\pi}\right)^{M-1}(\boldsymbol{d}_{\mathrm{L}} * \boldsymbol{c})$, and thus that the space generated by shifts of $\beta^{\mathrm{per}}_{\mathrm{L},\frac{2\pi}{P}}$ is included in $\mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})$. Yet both are P-dimensional vector spaces, which proves that they are in fact equal. □

5.2. Discrete problem formulation

In practice, to solve Problem (26), we use the B-spline representation (36) of $\mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})$. The choice of the B-spline representation is guided by numerical considerations: B-splines have the shortest support among any uniform L-spline, and thus lead to well-conditioned optimization tasks. The problem thus consists in optimizing over the $c[0],\dots,c[P-1]$ coefficients, which leads to a computationally feasible finite-dimensional problem, as demonstrated in the following proposition.

Proposition 5. *Problem (26) is exactly equivalent to solving the finite-dimensional problem*

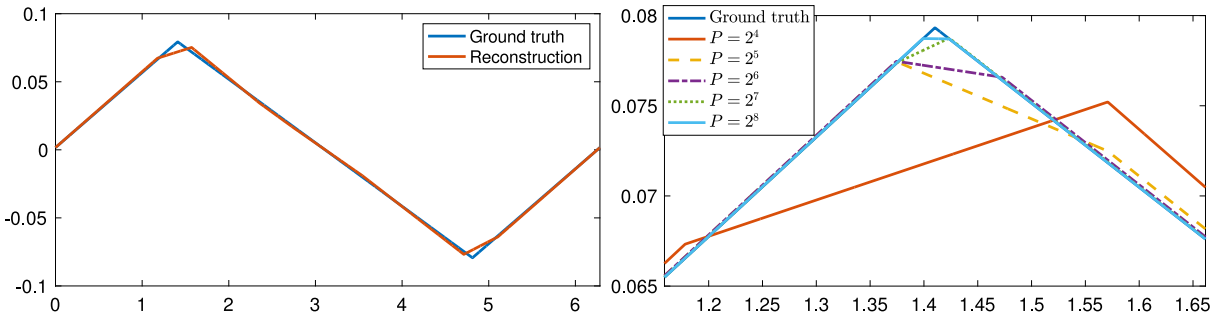
$$\mathcal{W}_{\lambda,P}(\boldsymbol{y}) = \operatorname*{arg\,min}_{\boldsymbol{c}\in\mathbb{R}^P} \frac{1}{2}\|\mathbf{H}\boldsymbol{c} - \boldsymbol{y}\|_2^2 + \lambda\left(\frac{P}{2\pi}\right)^{M-1}\|\boldsymbol{d}_{\mathrm{L}} * \boldsymbol{c}\|_1, \tag{38}$$

where the matrix $\mathbf{H} \in \mathbb{C}^{(K_c+1)\times P}$ is given by $\mathbf{H}_{k,\ell} = \nu_k\left(\beta^{\mathrm{per}}_{\mathrm{L},\frac{2\pi}{P}}(\cdot - \ell\frac{2\pi}{P})\right) = \mathrm{e}^{-\mathrm{i}\ell\frac{2\pi}{P}}\widehat{\beta}^{\mathrm{per}}_{\mathrm{L},\frac{2\pi}{P}}[k]$, $\boldsymbol{d}_{\mathrm{L}} = (d_{\mathrm{L}}[0],\dots,d_{\mathrm{L}}[P-1])$, and $\boldsymbol{c} = (c[0],\dots,c[P-1])$. The continuous-domain reconstructed signal is then $f = \sum_{p=0}^{P-1} c_p^*\beta^{\mathrm{per}}_{\mathrm{L},\frac{2\pi}{P}}(\cdot - \frac{2\pi p}{P})$, where $\boldsymbol{c}^* \in \mathcal{W}_{\lambda,P}(\boldsymbol{y})$.

Proof. This equivalence is obtained by plugging in $f = \sum_{p=1}^{P} c_p^*\beta^{\mathrm{per}}_{\mathrm{L},\frac{2\pi}{P}}(\cdot - \frac{2\pi p}{P}) \in \mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})$ into the cost function of problem (24). The expression of the system matrix $\mathbf{H}$ immediately follows. The expression of the regularization term follows from (37) and the fact that $\|\sum_{p=0}^{P-1} a_p\amalg(\cdot - x_p)\|_{\mathcal{M}} = \|\boldsymbol{a}\|_1$ for pairwise-distinct knot locations x_k. □

6. Experiments

In this section, we present some results of our discretization method presented in the previous section in various experimental settings. Our method amounts to solving Problem (38), which is a standard discrete problem with ℓ_1 regularization. This can be achieved using standard proximal solvers; in our experiments, we use the ADMM solver [44] of GlobalBioIm [45], a Matlab-based inverse-problem library developed by the Biomedical Imaging Group at EPFL. We solve the linear inversion step of ADMM with a direct matrix inversion, which is feasible due to the relatively small dimension of our problems.



(a) Reconstruction with $P = 16$ grid points (sparsity: $N = 7$ knots).

(b) Reconstructions with different grid sizes around the knot at 1.41

Fig. 1. Noiseless reconstruction of a piecewise-linear spline with $N = 2$ knots, $K_c = 3$, and $\lambda = 10^{-7}$.

6.1. Effect of gridding

6.1.1. Qualitative effect

We first present a toy experiment to illustrate the effect of gridding in our discretization method, *i.e.*, restricting the search space to $\mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})$. We therefore design an experiment in which the solution of the problem is known, in order to observe whether our algorithm is able to reconstruct it. To this end, we take $\mathrm{L} = \mathrm{D}^2$, and generate a ground-truth signal f_0 which is a periodic D^2-spline with 2 knots (the locations and amplitudes of the knots are picked at random). We then compute the noiseless data vector $\boldsymbol{y} = \boldsymbol{\nu}(f_0)$ for $K_c = 3$, and solve the corresponding problem (24) with a small regularization parameter $\lambda = 10^{-7}$ in order to enforce the constraints $\boldsymbol{\nu}(f) \approx \boldsymbol{y}$ with very low error. Since the form of f_0 is compatible with that of the solution given by Theorem 1, the hope is that f_0 will be very close to the solution f^* to problem (24), which is confirmed by our experiments.

In Fig. 1(a), we show the reconstruction result of our algorithm, using a voluntarily coarse grid with $P = 16$ points for visualization purposes. We observe that since the knots of f_0 are quite far from the grid, it is difficult to approximate f_0 with an element of $\mathcal{M}_{\mathrm{L},\frac{2\pi}{P}}(\mathbb{T})$. The reconstruction therefore requires several knots on the grid to mimic a single knot of f_0, and thus has a much higher sparsity ($N = 7$ knots versus $N = 2$ for f_0).

However, as we increase the number of grid points, the effect of gridding is greatly reduced, as illustrated in Fig. 1(b): with $P = 512$, the reconstruction using our algorithm is visually indistinguishable from f_0 (which is why we do not show it). However, the knot locations of f_0 still do not exactly lie on the grid, and thus our reconstruction still requires multiple knots to mimic a single knot of f_0, which leads to a sparsity of $N = 4$. Specifically, our reconstruction has two knots at consecutive grid points 1.3990 and 1.4113 mimicking the knot at 1.4103 of f_0, and two knots at 4.8106 and 4.8228 mimicking the knot at 4.8122 of f_0. This effect of knot multiplication due to gridding has already been observed and studied extensively in [18] in the absence of a regularization operator L.

The conclusion is thus that gridding leads to visually near-perfect reconstruction when the number of grid points is very large, which is in line with Theorem 2; however, the sparsity of the reconstruction is a poor indicator of the sparsity of the true solution of Problem (24), since gridding induces clusters of knots. Note that this effect is also present in the noisy scenario since, as shown in the proof of Theorem 1, Problem (17) can be reformulated as a noiseless problem of the form (18).

6.1.2. Quantitative effect

In Theorem 2, we have proved that any sequence of continuous-domain solutions f_P^* to the grid-restricted problem converges uniformly towards the unique solution f^* of problem (24) when P goes to infinity. In order to quantify the speed of this convergence, using the same experimental setting as in Fig. 1, we compute the error $\|f_P^* - f_0\|_\infty$ where f_P^* is the reconstructed signal using our grid-based algorithm, and the ground truth f_0 is a proxy for the solution f^* to problem (24). As explained earlier, this is a reasonable proxy due to the very small regularization parameter $\lambda = 10^{-7}$. In order to limit the effect of randomness in the choice of the knots of the ground truth, we apply a Monte Carlo-type method by generating 100 different ground-truth signals (following the methodology described in the previous section) and averaging the error over these 100 runs. These average errors for different grid sizes are shown in Fig. 2. The trend appears to be linear in log–log scale, which indicates an empirical speed of convergence of $\|f_P^* - f_0\|_\infty \approx (\frac{C}{P^s})$ for some constant $C > 0$ and where $-s < 0$ is the slope of the linear function. We observe here that $s \approx 1$ with $s < 1$. Note that there is no hope of having $s > 1$, since a nonuniform piecewise-linear spline cannot be approximated by a uniform piecewise-linear spline with an error smaller than $\mathcal{O}(h)$ in uniform norm, where $h = \frac{2\pi}{P}$ is the grid spacing.[4]

[4] This statement can easily be demonstrated geometrically.

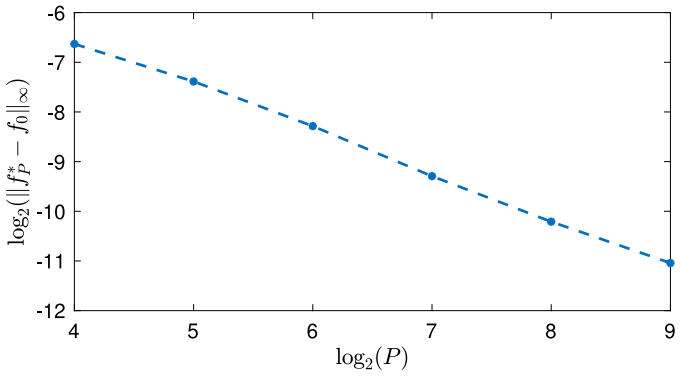


Fig. 2. Average error $\|f_P^* - f_0\|_\infty$ over 100 runs for different grid sizes P (in log–log scale).

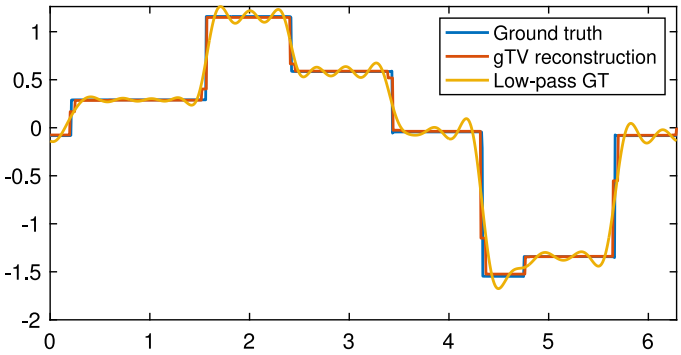


Fig. 3. Recovery of piecewise-constant spline with 7 knots with $K_c = 20$. For the our reconstruction (gTV), we use $\lambda = 10^{-2}$ and $P = 256$ grid points; the sparsity of the reconstruction is $N = 20$ knots. The data $\boldsymbol{y}$ is noisy in the gTV case, whereas the noiseless data $\boldsymbol{\nu}(f_0)$ is used for the low-pass reconstruction (partial Fourier series of the ground truth signal up to the cutoff frequency K_c).

6.2. Noisy recovery of sparse splines

In our next experiment, we attempt to recover a ground-truth signal f_0 based on noisy data $\boldsymbol{y} \in \mathbb{R} \times \mathbb{C}^{K_c}$ with $K_c = 20$ and a regularization operator $\mathrm{L} = \mathrm{D}$. Once again, the ground-truth signal fits the signal model of problem (24), *i.e.*, f_0 is a periodic D-spline (piecewise constant signal) with $N = 7$ knots. Each knot x_n is chosen at random within consecutive intervals of length $\frac{2\pi}{7}$, and the vector of amplitudes $\boldsymbol{a} = (a_1, \dots, a_n)$ is an i.i.d. Gaussian random vector projected on the space of zero-mean vectors. The measurements are corrupted by some additive i.i.d. Gaussian noise[5] $\boldsymbol{n} \in \mathbb{R} \times \mathbb{C}^{K_c}$ with standard deviation $\sigma = 10^{-3}$, *i.e.*, $\boldsymbol{y} = \boldsymbol{\nu}(f_0) + \boldsymbol{n}$.

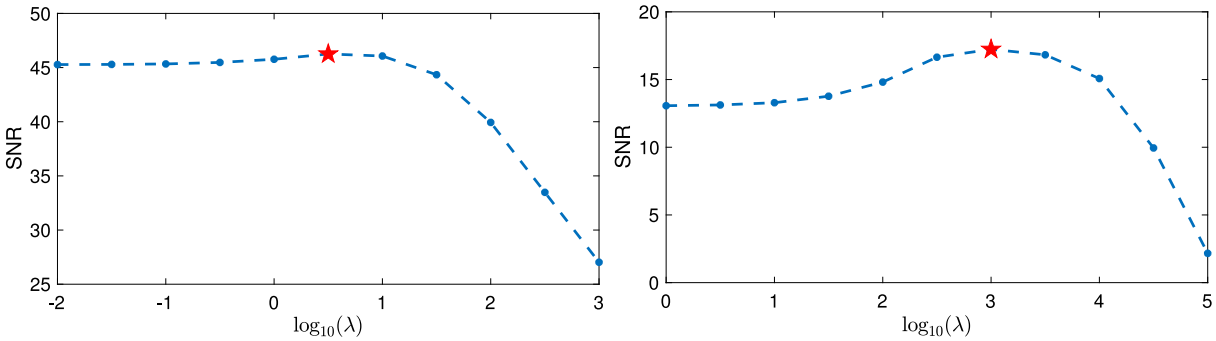
The reconstructed signal using our algorithm is shown in Fig. 3. Despite the presence of noise, the reconstruction of the ground truth f_0 is almost perfect. As observed in the previous experiment, the sparsity of the reconstruction ($N = 20$) is higher than that of the ground truth ($N = 7$) due to clusters of knots. We compare our reconstruction to the truncated Fourier series of f_0 up the K_c, *i.e.*, $f_{K_c} = \sum_{k=-K_c}^{K_c} \widehat{f_0}[k] e_k$, which solely depends on the noiseless data vector $\boldsymbol{\nu}(f_0)$. Without any prior knowledge, this is the simplest reconstruction one can think of based the available data $\boldsymbol{\nu}(f_0) = (\widehat{f_0}[0], \dots, \widehat{f_0}[K_c]$. As it turns out, f_{K_c} is also the unique solution to the following constrained L_2-regularized problem

$$f_{K_c} = \operatorname*{arg\,min}_{f : \boldsymbol{\nu}(f) = \boldsymbol{\nu}(f_0)} \|f\|_{L_2}, \tag{39}$$

as demonstrated in [25, Theorem 3]. In fact, adding any LSI regularization operator L in (39) still yields the same solution, since the basis functions φ_m in [25, Theorem 3] span the same space. This is due to the fact that the measurement functionals ν_m, *i.e.*, complex exponentials, are eigenfunctions of LSI operators.

As expected from the fact that f_{K_c} is a trigonometric polynomial whereas f_0 has sharps jumps, the reconstruction is quite poor and exhibits Gibbs-like oscillations, despite the absence of noise. This clearly demonstrates the superiority of gTV over L_2 regularization for sparse periodic splines reconstruction. Note however that the gap in performance decreases as the order of $\mathrm{L} = \mathrm{D}^M$ increases, since Gibbs-like phenomena are less significant for smoother functions.

[5] For complex entries, both the real and imaginary parts are i.i.d. Gaussian variables with the same σ.



(a) Low noise level $\left(\sigma = 0.01 \times \frac{\sum_{k=0}^{K_c} |\nu_k(f_0)|}{K_c+1}\right)$ (b) High noise level $\left(\sigma = 0.5 \times \frac{\sum_{k=0}^{K_c} |\nu_k(f_0)|}{K_c+1}\right)$.

Fig. 4. Evolution of SNR value between the ground-truth signal f_0 and the reconstructed signal f_λ^* for varying values of the regularization parameter λ averaged over 10 runs (in log–log scale). The star represents the highest SNR value.

6.3. Sensitivity to the regularization parameter

In our final experiment, we investigate the sensitivity of our reconstruction method to the choice of the regularization parameter λ. We choose the regularization operator $\mathrm{L} = \mathrm{D}^3$ ($M = 3$) and a cutoff frequency of $K_c = 5$. The ground-truth signal f_0 is generated in the same fashion as described in Section 6.2, only with $N = 4$ knots. The measurements $\boldsymbol{y}$ are corrupted by some additive i.i.d. Gaussian noise $\boldsymbol{n}$ with standard deviation σ. We monitor the evolution of the signal-to-noise ratio (SNR) between the ground truth f_0 and the reconstructed signal f_λ^* for varying values of the regularization parameter λ. In order to reduce the effect of randomness, in the same spirit as in Fig. 2, the results are averaged over 10 realizations of the ground truth f_0. All the experiments are performed with a fixed grid size of $P = 256$.

The results are shown in Fig. 4 for low ($\sigma = 0.01 \times \frac{\sum_{k=0}^{K_c} |\nu_k(f_0)|}{K_c+1}$, *i.e.*, one hundredth of the mean amplitude of the noiseless measurements $\boldsymbol{\nu}(f_0)$) and high ($\sigma = 0.5 \times \frac{\sum_{k=0}^{K_c} |\nu_k(f_0)|}{K_c+1}$) noise levels. We observe that in general, our method is more sensitive to the value of λ when the latter is too high than when it is too low. However, this effect is attenuated when the noise level increases: in Fig. 4(a), we observe that for low noise levels, any sufficiently small regularization parameter λ yields similar reconstruction results, whereas the SNR decreases dramatically past a certain value of λ. In comparison, as expected, for high noise levels (Fig. 4(b)), excessively low values of λ lead to a larger deterioration of the reconstruction quality.

7. Conclusion

This paper deals with continuous-domain inverse problems, where the goal is to recover a periodic function from its low-pass measurements. The reconstruction task is formalized as an optimization problem with a TV-based regularization involving a high-order derivative operator. It was known that spline solutions always exist (representer theorem). Our main result has proved that the solution is in fact always unique. We then studied the grid-based discretization of our optimization problem. We leveraged our uniqueness result to that any sequence of solutions of the discretized problems converge in uniform norm – a remarkably strong form of convergence – to the solution of the original problem when the grid size vanishes. Finally, we proposed a B-spline-based algorithm to solve the discretized problem, and we illustrated the relevance of our approach on simulations.

Data availability

No data was used for the research described in the article.

Acknowledgments

The authors thank Shayan Aziznejad, Adrian Jarret, Matthieu Simeoni, and Michael Unser for interesting discussions. Julien Fageot is supported by the Swiss National Science Foundation (SNSF) under Grant P400P2_194364. The work of Thomas Debarre is supported by the SNSF, Switzerland under Grant 200020_184646/1. Quentin Denoyelle is supported by the European Research Council (ERC) under Grant 692726-GlobalBioIm.

References

[1] L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1–4) (1992) 259–268.
[2] E. Candès, C. Fernandez-Granda, Super-resolution from noisy data, J. Fourier Anal. Appl. 19 (6) (2013) 1229–1254.
[3] K. Bredies, H. Pikkarainen, Inverse problems in spaces of measures, ESAIM Control Optim. Calc. Var. 19 (01) (2013) 190–218.
[4] V. Duval, G. Peyré, Exact support recovery for sparse spikes deconvolution, Found. Comput. Math. 15 (5) (2015) 1315–1355.
[5] M. Unser, J. Fageot, J.P. Ward, Splines are universal solutions of linear inverse problems with generalized TV regularization, SIAM Rev. 59 (4) (2017) 769–793.
[6] N. Boyd, G. Schiebinger, B. Recht, The alternating descent conditional gradient method for sparse inverse problems, SIAM J. Optim. 27 (2) (2017) 616–639.
[7] Q. Denoyelle, V. Duval, G. Peyré, E. Soubies, The sliding Frank–Wolfe algorithm and its application to super-resolution microscopy, Inverse Problems 36 (1) (2019) 014001.
[8] A. Flinth, F. de Gournay, P. Weiss, On the linear convergence rates of exchange and continuous methods for total variation minimization, Math. Program. 190 (1–2) (2020) 221–257.
[9] J.-B. Courbot, V. Duval, B. Legras, Sparse analysis for mesoscale convective systems tracking, Signal Process., Image Commun. 85 (2020) 115854.
[10] M. Simeoni, Functional Inverse Problems on Spheres: Theory, Algorithms and Applications (Ph.D. thesis), Swiss Federal Institute of Technology Lausanne (EPFL), 2020.
[11] J. Courbot, B. Colicchio, A fast homotopy algorithm for gridless sparse recovery, Inverse Problems 37 (2) (2021) 025002.
[12] M. Lustig, D. Donoho, J. Pauly, Sparse MRI: The application of compressed sensing for rapid MR imaging, Magn. Reson. Med. 58 (6) (2007) 1182–1195.
[13] R. Rockafellar, Convex Analysis, Princeton University Press, 1970.
[14] A. Badoual, J. Fageot, M. Unser, Periodic splines and Gaussian processes for the resolution of linear inverse problems, IEEE Trans. Signal Proc. 66 (22) (2018) 6047–6061.
[15] T. Debarre, Q. Denoyelle, J. Fageot, On the uniqueness of solutions for the basis pursuit in the continuum, Inverse Problems 38 (12) (2022) 125005.
[16] J. Fageot, M. Simeoni, TV-based reconstruction of periodic functions, Inverse Problems 36 (11) (2020) 115015.
[17] S. Fisher, J. Jerome, Spline solutions to L^1 extremal problems in one and several variables, J. Approx. Theory 13 (1) (1975) 73–83.
[18] V. Duval, G. Peyré, Sparse regularization on thin grids I: the LASSO, Inverse Problems 33 (5) (2017) 055008.
[19] T. Debarre, J. Fageot, H. Gupta, M. Unser, B-Spline-based exact discretization of continuous-domain inverse problems with generalized TV regularization, IEEE Trans. Inf. Theory 65 (7) (2019) 4457–4470.
[20] Y. de Castro, F. Gamboa, Exact reconstruction using beurling minimal extrapolation, J. Math. Anal. Appl. 395 (1) (2012) 336–354.
[21] E. Candès, C. Fernandez-Granda, Towards a mathematical theory of super-resolution, Comm. Pure Appl. Math. 67 (6) (2014) 906–956.
[22] J.-M. Azaïs, Y. de Castro, F. Gamboa, Spike detection from inaccurate samplings, Appl. Comput. Harmon. Anal. 38 (2) (2015) 177–195.
[23] V. Duval, G. Peyré, Sparse spikes super-resolution on thin grids II: the continuous basis pursuit, Inverse Problems 33 (9) (2017) 095008.
[24] C. Poon, N. Keriven, G. Peyré, Support localization and the Fisher metric for off-the-grid sparse regularization, in: K. Chaudhuri, M. Sugiyama (Eds.), Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics, in: Proceedings of Machine Learning Research, vol.89, PMLR, 2019, pp. 1341–1350.
[25] H. Gupta, J. Fageot, M. Unser, Continuous-domain solutions of linear inverse problems with Tikhonov vs. generalized TV regularization, IEEE Trans. Signal Process. 66 (17) (2018) 4670–4684.
[26] S. Aziznejad, M. Unser, Multikernel regression with sparsity constraint, SIAM J. Math. Data Sci. 3 (1) (2021) 201–224.
[27] C. Boyer, A. Chambolle, Y. de Castro, V. Duval, F.D. Gournay, P. Weiss, On representer theorems and convex regularization, SIAM J. Optim. 29 (2) (2019) 1260–1281.
[28] K. Bredies, M. Carioni, Sparsity of solutions for variational inverse problems with finite-dimensional data, Calc. Var. Partial Differential Equations 59 (1) (2019) 1–26.
[29] A. Flinth, P. Weiss, Exact solutions of infinite dimensional total-variation regularized problems, Inf. Inference: A J. IMA 8 (3) (2019) 407–443.
[30] M. Unser, J. Fageot, Native Banach spaces for splines and variational inverse problems, 2019, arXiv preprint arXiv:1904.10818.
[31] F. Filbir, K. Schröder, A. Veselovska, Super-resolution on the two-dimensional unit sphere, 2021, arXiv preprint arXiv:2107.10762.
[32] T. Debarre, Q. Denoyelle, M. Unser, J. Fageot, Sparsest piecewise-linear regression of one-dimensional data, J. Comput. Appl. Math. 406 (2022) 114044.
[33] C. Aubel, D. Stotz, H. Bölcskei, A theory of super-resolution from short-time Fourier transform measurements, J. Fourier Anal. Appl. 24 (1) (2018) 45–107.
[34] T. Debarre, S. Aziznejad, M. Unser, Hybrid-spline dictionaries for continuous-domain inverse problems, IEEE Trans. Signal Process. 67 (22) (2019) 5824–5836.
[35] T. Debarre, S. Aziznejad, M. Unser, Continuous-domain formulation of inverse problems for composite sparse-plus-smooth signals, IEEE Open J. Signal Proc. 2 (2021) 545–558.
[36] I. Lloréns Jover, T. Debarre, S. Aziznejad, M. Unser, Coupled splines for sparse curve fitting, IEEE Trans. Image Process. 31 (2022) 4707–4718.
[37] J. Gray, The shaping of the Riesz representation theorem: A chapter in the history of analysis, Arch. Hist. Exact Sci. 31 (2) (1984) 127–187.
[38] J.L. Kelley, General topology, Courier Dover Publications, 2017.
[39] G.D. Maso, An Introduction to Γ-Convergence, Vol. 8, Springer Science & Business Media, 2012.
[40] P. Heins, A Novel Regularization Technique and an Asymptotic Analysis of Spatial Sparsity Priors (Ph.D. thesis), Ph.D. thesis, Westfälische Wilhelms Universität Münster (WWU Münster), 2014. 32, 2017.
[41] I. Schoenberg, Cardinal Spline Interpolation, Society for Industrial and Applied Mathematics, 1973.
[42] C. de Boor, A Practical Guide to Splines, Vol. 27, Springer-Verlag New York, 1978.
[43] M. Unser, Splines: A perfect fit for signal and image processing, IEEE Signal Process. Mag. 16 (6) (1999) 22–38.
[44] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Found. Trends® Mach. Learn. 3 (1) (2010) 1–122.
[45] E. Soubies, F. Soulez, M. McCann, T.-a. Pham, L. Donati, T. Debarre, D. Sage, M. Unser, Pocket guide to solve inverse problems with GlobalBioIm, Inverse Problems 35 (10) (2019) 1–20, paper no. 104006.