Hybrid-Spline Dictionaries for Continuous-Domain Inverse Problems

Thomas Debarre, Shayan Aziznejad, and Michael Unser, *Fellow, IEEE*

***Abstract*—We study one-dimensional continuous-domain inverse problems with multiple generalized total-variation regularization, which involves the joint use of several regularization operators. Our starting point is a new representer theorem that states that such inverse problems have *hybrid-spline* solutions with a total sparsity bounded by the number of measurements. We show that such continuous-domain problems can be discretized in an exact way by using a union of B-spline dictionary bases matched to the regularization operators. We then propose a multiresolution algorithm that selects an appropriate grid size that depends on the problem. Finally, we demonstrate the computational feasibility of our algorithm for multiple-order derivative regularization operators.**

***Index Terms*—Inverse problems, total variation, sparsity, compressed sensing, B-splines.**

I. Introduction

THE sparse representation of signals is a major topic of interest in the field of signal processing. In particular, it is the driving principle of compressed sensing (CS) [1], [2], a field which has gained considerable popularity in the past decade. The CS approach to find a sparse representation of a signal given its measurements is to solve an inverse problem with sparsity-promoting regularization. The aim is to reconstruct a signal $\mathbf{c} \in \mathbb{R}^N$ given measurements $\mathbf{y} = \mathbf{Hc} + \mathbf{n} \in \mathbb{R}^M$ (with $M \ll N$), where $\mathbf{H} : \mathbb{R}^N \to \mathbb{R}^M$ is the system (or measurement) matrix and $\mathbf{n}$ is some additive noise which models measurement errors. The a priori assumption on the signal $\mathbf{c}$ is that it is sparse in a certain dictionary frame $\mathbf{D} : \mathbb{R}^P \to \mathbb{R}^N$, so that $\mathbf{c} = \mathbf{Da}$, where the vector of coefficients $\mathbf{a} \in \mathbb{R}^P$ is sparse. When the dictionary basis is orthonormal (i.e., $P = N$ and $\mathbf{D}$ is orthogonal), and under certain conditions on the system matrix $\mathbf{H}$ and the noise $\mathbf{n}$, standard CS theory guarantees stable recovery of the signal $\mathbf{c}$ [3]. This is achieved by solving the optimization problem

$$\hat{\mathbf{a}} = \underset{\mathbf{a}\in\mathbb{R}^P}{\arg\min} \left(\|\mathbf{HDa} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{a}\|_1 \right), \tag{1}$$

where $\lambda > 0$ is a regularization parameter, and by computing $\mathbf{c} = \mathbf{D}\hat{\mathbf{a}}$.

In many cases, however, the sparsity of a signal is expressed in a dictionary basis that is redundant and thus not orthonormal. More precisely, in this work, we are interested in multicomponent signals $\mathbf{c} = \sum_{i=1}^{Q} \mathbf{c}_i$ such that each of the Q components $\mathbf{c}_i$ is sparse in a different dictionary $\mathbf{D}_i$ ($i \in \{1, \ldots, Q\}$). For simplicity of exposition, we set $Q = 2$. A natural way of formulating the recovery problem is to solve (1) with a concatenated dictionary $\mathbf{D} = (\mathbf{D}_1, \mathbf{D}_2)$, where we can assume for now that $\mathbf{D}_1$ and $\mathbf{D}_2$ are orthonormal bases. In this case, $P = 2N$ and the dictionary basis is clearly highly redundant. Although this setting introduces practical and theoretical difficulties, it is extremely useful in many applications when a single dictionary is insufficient to represent the richness of a signal.

The problem of accurately reconstructing both components $\mathbf{c}_1$ and $\mathbf{c}_2$ is known as data separation [2, Chapter 11], and has been studied extensively both theoretically and practically. In fact, some of the first theoretical works concerning sparse vector recovery using ℓ_1-norm minimization involved a concatenated dictionary consisting of a mixture of sinusoids and spikes [4], [5]. The goal was to provide a condition under which ℓ_0 and ℓ_1 minimization yield the same solution. This sparked an abundance of research, which extended and improved these results for more general (non-orthonormal) dictionaries [6]–[9]. An overview is given in [10]. Later, these results were extended to images to separate point-like and curve-like structures [11]. These works mostly tackle denoising problems characterized by $\mathbf{H} = \mathbf{I}_N$ and $M = N$. In the CS field, in which we have $M \ll N$, [12] considers redundant dictionaries in general. On the practical side, data separation is intimately related to morphological component analysis (MCA), a method popularized by Starck *et al.* [13]–[17] with applications in inpainting removal or the separation of texture and natural parts of an image.

A. Analysis Formulation With Unions of Dictionaries

For overcomplete dictionaries $\mathbf{D}$ in general, optimizing over the synthesis coefficients $\mathbf{a}$ as in (1) seems natural, since the sparsity of the signal is precisely enforced on $\mathbf{a}$. This leads to a separable (if not differentiable) regularization term and, thus, to standard soft-thresholding-based algorithms. Conversely, optimizing over the analysis coefficients $\mathbf{c}$ is less straightforward since the representation of the signal $\mathbf{c}$ by the synthesis coefficients $\mathbf{a}$ is not unique. Therefore, the enforcement of the sparsity

Manuscript received February 14, 2019; revised August 5, 2019; accepted September 24, 2019. Date of publication October 4, 2019; date of current version October 22, 2019. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Mark A. Davenport. This research was supported in part by the European Research Council (ERC), under Grant 692726-GlobalBioIm and in part by the Swiss National Science Foundation, under Grant 200020_184646/1. *(Corresponding author: Thomas Debarre.)*

The authors are with the Biomedical Imaging Group, École polytechnique fédérale de Lausanne, 1015 Lausanne, Switzerland (e-mail: thomas.debarre@gmail.com; shayan.aziznejad@epfl.ch; michael.unser@epfl.ch).

Digital Object Identifier 10.1109/TSP.2019.2944754

1053-587X © 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

of the signal would require one to find its sparsest representation in the dictionary basis, which seems to add unnecessary complications. This explains why most of the published works choose to optimize over the synthesis coefficients.

However, in the case of a concatenated dictionary $\mathbf{D} = (\mathbf{D}_1, \mathbf{D}_2)$, one can also think of optimizing over both analysis coefficients $\mathbf{c}_1$ and $\mathbf{c}_2$. This leads to an analysis problem of the form

$$(\hat{\mathbf{c}}_1, \hat{\mathbf{c}}_2) = \underset{\mathbf{c}_1, \mathbf{c}_2 \in \mathbb{R}^N}{\arg\min} \left(\|\mathbf{H}(\mathbf{c}_1 + \mathbf{c}_2) - \mathbf{y}\|_2^2 + \lambda \left(\|\mathbf{L}_1 \mathbf{c}_1\|_1 + \|\mathbf{L}_2 \mathbf{c}_2\|_1 \right) \right), \quad (2)$$

which yields the reconstructed signal $\mathbf{c} = \mathbf{c}_1 + \mathbf{c}_2$. When $\mathbf{D}_1$ and $\mathbf{D}_2$ are orthonormal bases, taking $\mathbf{L}_1 = \mathbf{D}_1^T$ and $\mathbf{L}_2 = \mathbf{D}_2^T$ makes Problems (1) and (2) exactly equivalent. In other cases, the Moore-Penrose pseudoinverses of $\mathbf{D}_1$ and $\mathbf{D}_2$ can be used for $\mathbf{L}_1$ and $\mathbf{L}_2$, in which case (1) and (2) are not equivalent. Although the analysis formulation (2) may seem more tedious, there are several reasons to believe that it might be appealing:

- in recent years, efficient splitting algorithms that overcome the non-separability of the regularization terms in Problem (2) have been designed (e.g., [18] or the well-known alternating-direction method of multipliers (ADMM) [19]);
- many successful practical CS applications favor the analysis formulation, e.g., [20];
- there is an intimate connexion between regularization operators (e.g., $\mathbf{L}_1$ and $\mathbf{L}_2$ in (2)) and splines. Indeed, it has recently been proved that continuous-domain inverse problems with generalized total-variation (gTV) regularization have sparse spline solutions [21]. The analysis formulation (2) therefore seems like the natural discrete counterpart to this continuous framework, as we clarify in Section II.

The difference between the two formulations has been studied in [22] where the authors conclude that, in the general case, both problems are inherently different. Several practicioners have used the analysis formulation for data-separation problems, most notably Starck *et al.* in the context of MCA [13]–[16]. More recently, Problem (2) was applied to the task of separating cartoon and texture parts of an image in [23]. There, the authors favor the analysis formulation for general redundant dictionaries. A similar approach is used in low-rank plus sparse decomposition methods [24].

Despite these empirical works, virtually no theoretical study of Problem (2) has been carried out. In [25], Candès *et al.* have named it the "split-analysis" problem. They recommend solving it precisely in the case of concatenated dictionaries (without further investigation). A theoretical study was later done by Lin *et al.* in [26], where they show that the data-separation problem (i.e., the recovery of both components of the original signal) can be solved via Problem (2). This result requires that $\mathbf{H}$ satisfies the restricted-isometry property adapted to a dictionary (D-RIP) and that $\mathbf{L}_1$ and $\mathbf{L}_2$ satisfy a mutual coherence condition.

While the literature on the topic is scarce in the discrete setting, to the best of our knowledge, it is nonexistent in the continuous domain. Since most real-world signals are continuously defined, the reconstruction of continuous-domain solutions is a desirable objective. Moreover, although handling discrete signals is obviously appealing from a computational perspective, it introduces discretization errors in the measurements. For instance, the discrete Fourier transform (DFT) is often used as surrogate for the continuous Fourier transform to model MRI measurements, which is by no means an exact discretization.

B. Continuous-Domain Problems for Hybrid Splines

In this work, we propose to use unions of dictionaries in a continuous-domain framework. Our goal is to reconstruct a multicomponent continuous-domain 1D signal $s = s_1 + s_2$, where s_1 and s_2 have different characteristics. Similarly to the discrete setting, we are given measurements $\mathbf{y} = \boldsymbol{\nu}(s) + \mathbf{n} \in \mathbb{R}^M$, where $\boldsymbol{\nu} : s \to \boldsymbol{\nu}(s) \in \mathbb{R}^M$ is a (continuous-domain) linear measurement operator and $\mathbf{n} \in \mathbb{R}^M$ is some additive noise. We focus on continuous-domain inverse problems of the form

$$s^* = \underset{f}{\arg\min} \left(\|\boldsymbol{\nu}(f) - \mathbf{y}\|_2^2 + \lambda \mathcal{R}_{\text{hyb}}(f) \right), \quad (3)$$

where $\lambda > 0$ is the regularization parameter. The hybrid regularization term is given by

$$\mathcal{R}_{\text{hyb}}(f) = \min_{\substack{f_1, f_2 \\ f_1 + f_2 = f}} \left((1-\alpha) \|\mathrm{L}_1\{f_1\}\|_{\mathcal{M}} + \alpha \|\mathrm{L}_2\{f_2\}\|_{\mathcal{M}} \right), \quad (4)$$

where $\alpha \in (0, 1)$ controls the weighing of the two regularization terms. The regularization norm $\|\cdot\|_{\mathcal{M}}$ norm generalizes the L_1 norm [21] and is the continuous counterpart of the ℓ_1 norm used in discrete problems.

Although they are formulated in slightly different forms, there is an obvious connection between Problem (3) and the discrete analysis Problem (2). Similarly to the discrete case, the sparsity of each component s_i is promoted in a different dictionary, which is determined by the choice of the regularization operator L_i for $i \in \{1, 2\}$. Our main theoretical result is a representer theorem that states that, for differential operators L_i, Problem (3) leads to spline solutions $s^* = s_1^* + s_2^*$, where each component s_i^* is an L_i-spline. The reconstructed signal s^* is therefore a sum of different splines, which we coin as a *hybrid spline*. Moreover, the total sparsity of s^* in this union of spline dictionaries is no larger than the number M of measurements. This representer theorem generalizes the main result of [21].

An important characteristic of our framework is its compatibility with the class of piecewise-polynomial functions, which can accurately model a large variety of real-world signal. For example, by taking $\mathrm{L}_1 = \mathrm{D}$ (piecewise-constant splines) and $\mathrm{L}_2 = \mathrm{D}^4$ (cubic splines), the dictionary consists of piecewise-cubic polynomials that admit discontinuities.

A key feature of our continuous-domain formulation is that Problem (3) can be discretized in an exact way using B-splines, based on the methodology of [27]. B-splines are basis functions that are popular in the field of signal processing [28]–[30] because they have compact support. Continuous-domain inverse problems with gTV-type norms are an active field of research [31]–[38]. However, to the best of our knowledge, ours is the first instance of a continuous-domain inverse problem for multicomponent signals.

C. Outline

This paper is organized as follows: in Section II, we present the continuous-domain framework of Problem (3), which is a multicomponent extension of that of [21]. We then prove our representer theorem; it states that Problem (3) admits a global minimizer that is a hybrid spline. We then provide some background information on splines in Section III. In Section IV, we show how we discretize Problem (3) in an exact way using B-splines and propose a practical solution to the corresponding discrete problem. In Section V, we present our proposed multiresolution algorithm. Finally, we illustrate some applications of our algorithm in Section VI, including curve fitting and CS-type problems.

II. Continuous-Domain Inverse Problem

In this section, we present a class of continuous-domain inverse problems that involve hybrid gTV regularization. These problems extend the framework of [21], which is the basis of this work. For more detailed information, we refer to that original article.

A. Definitions and Notations

Our aim is to recover a continuous-domain signal $s : \mathbb{R} \to \mathbb{R}$ given M noisy measurements modeled as $\mathbf{y} = \boldsymbol{\nu}(s) + \mathbf{n}$, where $\mathbf{n} \in \mathbb{R}^M$ is some additive noise. The noiseless measurements $\boldsymbol{\nu}(s)$ are acquired through M linear measurement functionals $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_M)$, with $\boldsymbol{\nu}(s) = (\langle \nu_1, s\rangle, \ldots, \langle \nu_M, s\rangle)$. There, $\langle \nu_m, s\rangle$ stands for the duality product, which is given by $\int_{\mathbb{R}} \nu_m(x)s(x)\mathrm{d}x$ when ν_m and s are ordinary functions. The ν_m functionals constitute the (assumed) forward model.

The regularization norm $\|\cdot\|_{\mathcal{M}}$ is defined for a tempered distribution $w \in \mathcal{S}'(\mathbb{R})$ as

$$\|w\|_{\mathcal{M}} = \sup_{\phi \in \mathcal{S}(\mathbb{R}), \|\phi\|_\infty = 1} \langle w, \phi \rangle, \tag{5}$$

where $\mathcal{S}(\mathbb{R})$ is the Schwartz space of smooth and rapidly decaying functions. Leaving aside the abstractness of this definition, let us focus on two critical features of the $\|\cdot\|_{\mathcal{M}}$ norm:

1) it generalizes the L_1 norm in the sense that $\|w\|_{\mathcal{M}} = \|w\|_1$ for any $w \in L_1(\mathbb{R})$;
2) the $\|\cdot\|_{\mathcal{M}}$ norm of a weighted sum of Dirac impulses is $\|\sum_{k\in\mathbb{Z}} a[k]\delta(\cdot - x_k)\|_{\mathcal{M}} = \|a\|_1$, where $(a[k])_{k\in\mathbb{Z}}$ is a real sequence and the $x_k \in \mathbb{R}$ are pairwise distinct.

Next, we sum up all the relevant information and notations that concern the regularization operators L_i ($i \in \{1, 2\}$).

1) The operator L_i is assumed to be spline-admissible in the sense of [21, Definition 1]. It therefore has a Green's function denoted by ρ_{L_i} which verifies $\mathrm{L}_i\{\rho_{\mathrm{L}_i}\} = \delta$.
2) The native space of L_i is denoted by $\mathcal{M}_{\mathrm{L}_i}(\mathbb{R})$ and verifies $\|\mathrm{L}_i\{f_i\}\|_{\mathcal{M}} < \infty$ for any $f_i \in \mathcal{M}_{\mathrm{L}_i}(\mathbb{R})$.
3) The null space of L_i is denoted by $\mathcal{N}_{\mathrm{L}_i}$ and has finite dimension $N_{0,i}$.
4) The intersection of the null spaces is denoted by $\mathcal{N}_0 = \mathcal{N}_{\mathrm{L}_1} \cap \mathcal{N}_{\mathrm{L}_2}$ and has dimension N_0. If $N_0 > 0$, then we introduce the biorthogonal system $(\boldsymbol{\phi}_0, \mathbf{p}_0) = (\phi_n, p_n)_{n=1}^{N_0}$ for $\mathcal{N}_0$ in the sense of Definition 3 in [21] that satisfies $\phi_n(p_m) = \delta_{nm}$ (Kronecker delta).
5) The restricted search space for L_1 is defined as

$$\mathcal{M}_{\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{R}) = \{f \in \mathcal{M}_{\mathrm{L}_1}(\mathbb{R}) : \boldsymbol{\phi}_0(f) = \mathbf{0}\}, \tag{6}$$

where $\boldsymbol{\phi}_0 : \mathcal{M}_{\mathrm{L}_1}(\mathbb{R}) \to \mathbb{R}^{N_0}$.

We then define splines matched to these operators.

Definition 1 (Nonuniform L*-splines):* Let L be a spline-admissible operator in the sense of [21, Definition 1]. A nonuniform L-spline is a function $s \in \mathcal{M}_{\mathrm{L}}(\mathbb{R})$ such that $\mathrm{L}\{s\} = \sum_k a_k\delta(\cdot - x_k)$ where $a_k, x_k \in \mathbb{R}$.

In the case of multiple-order derivative operators $\mathrm{L} = \mathrm{D}^{N_0}$, nonuniform L-splines are piecewise polynomials of degree $(N_0 - 1)$ with differentiability class $\mathcal{C}^{N_0-2}$. The locations x_k and amplitudes a_k of the singularities (knots) are characterized by the innovation $\mathrm{L}\{s\} = \sum_k a_k\delta(\cdot - x_k)$ of the spline. An example of a D-spline with its innovation is given in Fig. 1.

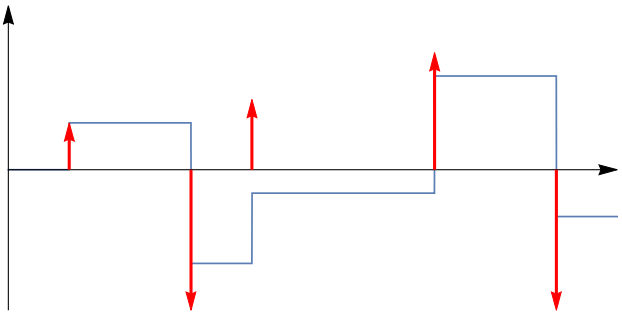


Fig. 1. Nonuniform D-spline (thin curve). The vertical arrows represent the innovation $\mathrm{D}\{s\} = \sum_k a_k\delta(\cdot - x_k)$.

Finally, we introduce the *hybrid regularization functional* for functions $f \in \mathcal{M}_{\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{R}) + \mathcal{M}_{\mathrm{L}_2}(\mathbb{R})$ as

$$\mathcal{R}_{\mathrm{hyb}}(f) = \min_{\substack{f_1 \in \mathcal{M}_{\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{R}) \\ f_2 \in \mathcal{M}_{\mathrm{L}_2}(\mathbb{R}) \\ f_1 + f_2 = f}} ((1-\alpha)\|\mathrm{L}_1\{f_1\}\|_{\mathcal{M}} + \alpha\|\mathrm{L}_2\{f_2\}\|_{\mathcal{M}}). \tag{7}$$

Proposition 1: The hybrid regularization functional $\mathcal{R}_{\mathrm{hyb}}$ defined in (7) has the following properties:

1) it is well defined in the sense that, for any $f \in \mathcal{M}_{\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{R}) + \mathcal{M}_{\mathrm{L}_2}(\mathbb{R})$, there exist functions $(f_1, f_2) \in \mathcal{M}_{\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{R}) \times \mathcal{M}_{\mathrm{L}_2}(\mathbb{R})$ such that $f = f_1 + f_2$ and

$$\mathcal{R}_{\mathrm{hyb}}(f) = (1-\alpha)\|\mathrm{L}_1\{f_1\}\|_{\mathcal{M}} + \alpha\|\mathrm{L}_2\{f_2\}\|_{\mathcal{M}}; \tag{8}$$

2) it is convex.

The proof of Proposition 1 is given in Appendix A.

B. Representer Theorem

We now have the necessary tools to present the main theoretical result of this paper, on which our implementation is based.

Theorem 1 (Continuous-domain representer theorem): Let $\mathrm{L}_1, \mathrm{L}_2$ be spline-admissible operators and let $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_M)$ be a linear measurement operator composed of the M linear functionals $\nu_m : f \mapsto \nu_m(f) \in \mathbb{R}$ which are weak*-continuous on both $\mathcal{M}_{\mathrm{L}_1}(\mathbb{R})$ and $\mathcal{M}_{\mathrm{L}_2}(\mathbb{R})$. Assume that

$\mathcal{N}_{\boldsymbol{\nu}} \cap (\mathcal{N}_{\mathrm{L}_1} + \mathcal{N}_{\mathrm{L}_2}) = \{0\}$, where $\mathcal{N}_{\boldsymbol{\nu}}$ is the null space of $\boldsymbol{\nu}$ (well-posedness assumption). Then, the linear inverse problem

$$\mathcal{S} = \left\{ \underset{f \in \mathcal{M}_{\mathrm{L}_1,\phi_0}(\mathbb{R}) + \mathcal{M}_{\mathrm{L}_2}(\mathbb{R})}{\arg\min} \underbrace{\left(\|\boldsymbol{\nu}(f) - \mathbf{y}\|_2^2 + \lambda \mathcal{R}_{\mathrm{hyb}}(f) \right)}_{\mathcal{J}(f)} \right\} \tag{9}$$

has a solution s of the form $s = s_1 + s_2$, where s_i are nonuniform L_i-splines (in the sense of Definition 1) of the form

$$s_i(x) = \sum_{k=1}^{K_i} a_{i,k} \rho_{\mathrm{L}_i}(x - x_{i,k}) + q_i(x) \tag{10}$$

for $q_i \in \mathcal{N}_{\mathrm{L}_i}$, $a_{i,k}, x_{i,k} \in \mathbb{R}$. Moreover, the sparsity indices K_i verify $K_1 + K_2 \leq M$.

The proof is given in Appendix B. Theorem 1 is a powerful generalization of Theorem 1 in [21] for hybrid regularization terms. It states that the continuous-domain Problem (9) has a hybrid spline solution that consists of the sum of an L_1-spline and an L_2-spline. The unknowns are the locations $x_{i,k}$ and amplitudes $a_{i,k}$ of the knots (which specify the innovation of the spline corresponding to the vertical arrows in Fig. 1), as well as the null-space components q_i. The total sparsity of this solution is bounded by the number of measurement M, meaning that it is characterized by very few parameters. The following observations can be made concerning Theorem 1:

- We use the restricted search space $\mathcal{M}_{\mathrm{L}_1,\phi_0}(\mathbb{R})$ defined in (6) instead of the complete space $\mathcal{M}_{\mathrm{L}_1}(\mathbb{R})$ in order to ensure that Problem (9) is well-posed. [1] This does not restrict the native space of the reconstructed signal s since $\mathcal{M}_{\mathrm{L}_1,\phi_0}(\mathbb{R}) + \mathcal{M}_{\mathrm{L}_2}(\mathbb{R}) = \mathcal{M}_{\mathrm{L}_1}(\mathbb{R}) + \mathcal{M}_{\mathrm{L}_2}(\mathbb{R})$.
- Theorem 1 can readily be extended to Q operators $\mathrm{L}_1, \ldots, \mathrm{L}_Q$: however, for $Q > 2$, the handling of the pairwise null space intersections would make the general formulation more tedious. For the sake of clarity, we therefore only consider the case $Q = 2$.
- A remarkable feature of Theorem 1 is that the bound on the sparsity of the solutions does not increase with the number Q of operators. This is particularly appealing from a theoretical point of view since, compared to the single-operator framework of [21], we essentially enrich our dictionary at no cost in terms of sparsity.

III. Derivative Operators and Splines

For the sake of clarity, we focus on N_0th-order derivative operators $\mathrm{L} = \mathrm{D}^{N_0}$. They are the most standard choices of regularization operators; for instance, $\mathrm{L} = \mathrm{D}$ leads to TV regularization. Higher orders $N_0 \geq 2$ lead to total-variation (TV) regularization of the $(N_0 - 1)$th derivative of the argument function f. Examples of more general operators that are compatible with the framework of [21] (and, *a fortiori*, ours) are generic differential operators, which lead to exponential splines [30], [39].

[1] An unbounded solution set $\mathcal{S}$ would arise if we allowed ourselves to add contributions $(p, -p)$ to a solution, with arbitrary $p \in \mathcal{N}_0$. Indeed, the extended solution would still be an element of $\mathcal{S}$.

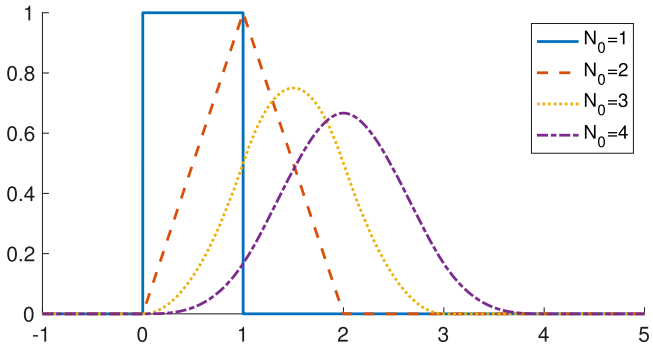


Fig. 2. Cardinal B-splines of operators D^{N_0} for $N_0 = 1, \ldots, 4$.

The null space of D^{N_0} is the set of polynomials of degree smaller than N_0, so that

$$\mathcal{N}_{\mathrm{D}^{N_0}} = \mathrm{span}\{p_n = (\cdot)^{n-1}\}_{n=1}^{N_0}. \tag{11}$$

The causal Green's function of D^{N_0} is

$$\rho_{\mathrm{D}^{N_0}}(x) = \frac{x_+^{N_0-1}}{(N_0-1)!}. \tag{12}$$

A. Cardinal B-Splines

The cardinal B-spline of D^{N_0} has a compact expression in the Fourier domain. It is defined by

$$\beta_{\mathrm{D}^{N_0}}(x) = \mathcal{F}^{-1}\left\{ \left(\frac{1 - \mathrm{e}^{-\mathrm{j}\omega}}{\mathrm{j}\omega} \right)^{N_0} \right\}(x). \tag{13}$$

It is indeed a D^{N_0}-spline in the sense of Definition 1 since

$$\mathrm{D}^{N_0}\{\beta_{\mathrm{D}^{N_0}}\}(x) = \sum_{k \in \mathbb{Z}} d_{\mathrm{D}^{N_0}}[k] \delta(x - k), \tag{14}$$

where $(d_{\mathrm{D}^{N_0}}[k])_{k \in \mathbb{Z}}$ is a discrete finite-impulse-response filter defined by its z transform

$$d_{\mathrm{D}^{N_0}}(z) = (1 - z^{-1})^{N_0}, \tag{15}$$

which is supported in $\{0, \ldots, N_0\}$. We provide in Table I a summary of the relevant characteristics for small values of N_0. Notice that the B-spline of D^{N_0} is supported in $[0, N_0]$. As shown in Fig. 2, higher-order derivatives lead to smoother B-splines.

B. Scaling

An important property of the B-spline is its scalability. We define the scaled B-spline as

$$\beta_{\mathrm{L},h}(x) = \beta_{\mathrm{L}}\left(\frac{x}{h}\right), \tag{16}$$

where $h > 0$ is the grid size. This means that the knots of $\beta_{\mathrm{L},h}$, which are integers for cardinal B-splines in Table I, lie in $h\mathbb{Z}$. The major asset of scaled B-splines is that they can represent any L-spline that has knots in $h\mathbb{Z}$. Hence, we introduce the space, which we use as our search space to discretize Problem (9).

Definition 2: The discretized search space $\mathcal{M}_{\mathrm{L},h}(\mathbb{R})$ associated to L is specified by

TABLE I
CHARACTERISTICS OF N_0TH-ORDER DERIVATIVE OPERATORS

L	$\rho_{\mathrm{L}}(x)$	$\beta_{\mathrm{L}}(x)$	$(d_{\mathrm{L}}[0],\dots,d_{\mathrm{L}}[N_0])$
D	$\mathbb{1}_+(x)$	$\beta_{\mathrm{D}}(x)=\begin{cases}1, & 0\le x<1\\ 0, & \text{otherwise}\end{cases}$	$(1,-1)$
D^2	x_+	$\beta_{\mathrm{D}^2}(x)=\begin{cases}x, & 0\le x<1\\ 2-x, & 1\le x<2\\ 0, & \text{otherwise}\end{cases}$	$(1,-2,1)$
D^3	$x_+^2/2$	$\beta_{\mathrm{D}^3}(x)=\begin{cases}x^2/2, & 0\le x<1\\ -x^2+3x-3/2, & 1\le x<2\\ (3-x^2)/2, & 2\le x<3\\ 0, & \text{otherwise}\end{cases}$	$(1,-3,3,-1)$
D^4	$x_+^3/6$	$\beta_{\mathrm{D}^4}(x)=\begin{cases}x^3/6, & 0\le x<1\\ -x^3/2+2x^2-2x+2/3, & 1\le x<2\\ x^3/2-4x^2+10x-22/3, & 2\le x<3\\ (4-x)^3/6, & 3\le x<4\\ 0, & \text{otherwise}\end{cases}$	$(1,-4,6,-4,1)$

$$\mathcal{M}_{\mathrm{L},h}(\mathbb{R}) = \left\{ p+\sum_{k\in\mathbb{Z}} a[k]\rho_{\mathrm{L}}(\cdot-kh) : a\in\ell_1(\mathbb{Z}), p\in\mathcal{N}_{\mathrm{L}} \right\} \quad (17)$$

$$= \left\{ \sum_{k\in\mathbb{Z}} c[k]\beta_{\mathrm{L},h}(\cdot-kh) : c\in\ell_{1,\mathrm{L}}(\mathbb{Z}) \right\} \quad (18)$$

where

$$\ell_{1,\mathrm{L}}(\mathbb{Z}) = \left\{ (c[k])_{k\in\mathbb{Z}} : (d_{\mathrm{L}} * c)\in\ell_1(\mathbb{Z}) \right\}. \quad (19)$$

The proof of the equivalence of the Green's function and B-spline representations is given in Proposition 1 in [27]. Given the form of the solutions (10) of our continuous-domain inverse Problem (9), this choice of search space is a natural one. By picking h sufficiently small, it contains functions that are arbitrarily close to (10). Moreover, the use of the underlying B-spline representation of $\mathcal{M}_{\mathrm{L},h}(\mathbb{R})$ leads to well-conditioned problems and, thus, to computationally effective algorithms. This is due to the advantageous properties of B-splines, namely, their finite support and the fact that they produce a Riesz basis [39, Theorem 1].

IV. EXACT DISCRETIZATION

We now detail our discretization method for Problem (9), which is based on the methodology of [27]. For all that follows, the conditions of Assumptions 1 apply.

Assumptions 1:

- The measurement functionals ν_m are supported in an interval $I_T=[0,T]$ where $T\in h\mathbb{Z}$.
- The regularization operators are $\mathrm{L}_1=\mathrm{D}^{N_{0,1}}$ and $\mathrm{L}_2=\mathrm{D}^{N_{0,2}}$ where $N_{0,1}<N_{0,2}$.
- For the space $\mathcal{M}_{\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{R})$ defined in (6), we choose the canonical boundary conditions $\boldsymbol{\phi}_0:\mathcal{M}_{\mathrm{L}_1}(\mathbb{R})\to\mathbb{R}^{N_{0,1}}$, where $\boldsymbol{\phi}_0(f)=(f(0),\dots,f^{(N_{0,1}-1)}(0))$.

The first assumption is natural and is often fulfilled in practice, for instance in imaging with a finite field of view. Since L_1 and L_2 have almost symmetrical roles, the second assumption is only significant in relation to the boundary conditions $\boldsymbol{\phi}_0$. It implies that $\mathcal{N}_1\subset\mathcal{N}_2$ and $\mathcal{N}_0=\mathcal{N}_{\mathrm{L}_1}$. Finally, we make a canonical choice for $\boldsymbol{\phi}_0$ for the sake of simplicity. As specified in [21], $(\boldsymbol{\phi}_0,\mathbf{p}_0)=(\phi_{0,n},p_{0,n})_{n=1}^{N_0}$ forms a biorthogonal system for $\mathcal{N}_0$, where $\mathbf{p}_0$ is defined as in (11).

A. Specification of $\mathcal{M}_{\mathrm{L}_1,\boldsymbol{\phi}_0,h}(\mathbb{R})$

The discretized search space for L_1 with boundary conditions $\boldsymbol{\phi}_0$ is given by

$$\mathcal{M}_{\mathrm{L}_1,\boldsymbol{\phi}_0,h}(\mathbb{R}) = \left\{ s\in\mathcal{M}_{\mathrm{L}_1,h}(\mathbb{R}) : \boldsymbol{\phi}_0(s)=\mathbf{0} \right\} \quad (20)$$

$$= \left\{ \sum_{k\in\mathbb{Z}} c[k]\beta_{\mathrm{L}_1,h}(\cdot-kh) : c\in\ell_{1,\mathrm{L}_1}(\mathbb{Z}), \boldsymbol{\phi}_{0,h}(c)=\mathbf{0} \right\}, \quad (21)$$

where $\boldsymbol{\phi}_{0,h}:\ell_{1,\mathrm{L}}(\mathbb{Z})\to\mathbb{R}^{N_0}$ satisfies $\boldsymbol{\phi}_{0,h}(c)=\boldsymbol{\phi}_0(s)$. We thus define the matching search space for B-spline coefficients as

$$\ell_{1,\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{Z}) = \left\{ c\in\ell_{1,\mathrm{L}_1}(\mathbb{Z}) : \boldsymbol{\phi}_{0,h}(c)=\mathbf{0} \right\}. \quad (22)$$

Since $\beta_{\mathrm{D}^{N_{0,1}}}$ is supported in $[0,N_{0,1}]$, the boundary conditions $\boldsymbol{\phi}_{0,h}$ impose linear constraints on only a few coefficients of c. For example, for $\mathrm{L}_1=\mathrm{D}$ (i.e., $N_{0,1}=1$), we have that $\phi_0(s)=s(0)=0$, which simply amounts to $\phi_{0,h}(c)=c[0]=0$. In general, it can easily be shown that the canonical boundary conditions are

$$\boldsymbol{\phi}_{0,h}(c)=(c[0],\dots,c[-N_{0,1}+1])=\mathbf{0}. \quad (23)$$

B. Discrete Problem Formulation

Next, we introduce the sets of indices $I_i=\{m_i,\dots,M_i\}$ of cardinality $N_i=\#I_i=M_i-m_i+1$ such that $k\in I_i$ if and only if the support of the B-spline atom $\beta_{\mathrm{L}_i}(\cdot-kh)$ and I_T have a nontrivial intersection. The indices $\{-N_{0,1}+1,\dots,0\}$ are excluded from I_1 due to the boundary conditions (23). We thus have that $m_1=1$, $m_2=(-N_{0,1}+1)$, and $M_1=M_2=(T/h-1)$, which implies that $N_1=(T/h-1)$ and $N_2=(T/h+N_{0,2}-1)$. By optimizing over the B-spline coefficients in I_i, we get the

following finite optimization problem

$$S_h = \left\{ \underset{(\mathbf{c}_1,\mathbf{c}_2)\in\mathbb{R}^{N_1}\times\mathbb{R}^{N_2}}{\arg\min} J_h(\mathbf{c}_1,\mathbf{c}_2) \right\} \tag{24}$$

$$J_h(\mathbf{c}_1,\mathbf{c}_2) = \|\mathbf{H}_1\mathbf{c}_1 + \mathbf{H}_2\mathbf{c}_2 - \mathbf{y}\|_2^2 + \lambda\left((1-\alpha)\|\mathbf{L}_1\mathbf{c}_1\|_1 + \alpha\|\mathbf{L}_2\mathbf{c}_2\|_1\right). \tag{25}$$

The system matrices $\mathbf{H}_i \in \mathbb{R}^{M\times N_i}$ are defined as

$$\mathbf{H}_i = \begin{bmatrix} \mathbf{h}_{m_i} & \cdots & \mathbf{h}_{M_i} \end{bmatrix} : \quad \mathbf{h}_k = \boldsymbol{\nu}(\beta_{\mathrm{L}_i}(\cdot - kh)), \tag{26}$$

and the Toeplitz-like regularization matrices are given by $\mathbf{L}_1 \in \mathbb{R}^{N_1\times N_1}$, with

$$\mathbf{L}_1 = \frac{1}{h^{N_{0,1}-1}} \times \begin{pmatrix} d_{\mathrm{L}_1}[0] & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & & & \vdots \\ d_{\mathrm{L}_1}[N_{0,1}] & & \ddots & \ddots & & \vdots \\ 0 & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ 0 & \cdots & 0 & d_{\mathrm{L}_1}[N_{0,1}] & \cdots & d_{\mathrm{L}_1}[0] \end{pmatrix}, \tag{27}$$

and $\mathbf{L}_2 \in \mathbb{R}^{(N_2-N_{0,2})\times N_2}$, with

$$\mathbf{L}_2 = \frac{1}{h^{N_{0,2}-1}} \times \begin{pmatrix} d_{\mathrm{L}_2}[N_{0,2}] & \cdots & d_{\mathrm{L}_2}[0] & 0 & \cdots & 0 \\ 0 & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ 0 & \cdots & 0 & d_{\mathrm{L}_2}[N_{0,2}] & \cdots & d_{\mathrm{L}_2}[0] \end{pmatrix}. \tag{28}$$

The cost function can thus be rewritten as

$$J_h(\mathbf{c}_1,\mathbf{c}_2) = \left\| \mathbf{H}\begin{pmatrix}\mathbf{c}_1\\ \mathbf{c}_2\end{pmatrix} - \mathbf{y}\right\|_2^2 + \lambda\left\|\mathbf{L}\begin{pmatrix}\mathbf{c}_1\\ \mathbf{c}_2\end{pmatrix}\right\|_1, \tag{29}$$

where the concatenated system and regularization matrices are

$$\mathbf{H} = \begin{bmatrix}\mathbf{H}_1 & \mathbf{H}_2\end{bmatrix} \in \mathbb{R}^{M\times(N_1+N_2)} \quad \text{and} \tag{30}$$

$$\mathbf{L} = \begin{pmatrix}(1-\alpha)\mathbf{L}_1 & \mathbf{0} \\ \mathbf{0} & \alpha\mathbf{L}_2\end{pmatrix} \in \mathbb{R}^{(N_1+N_2-N_{0,2})\times(N_1+N_2)}, \tag{31}$$

respectively. Therefore, Problem (24) is a standard penalized basis pursuit (PBP) problem which can be solved using off-the-shelf algorithms such as ADMM. Furthermore, it satisfies the property of Theorem 2.

Theorem 2: Under the assumptions of Theorem 1 and Assumptions 1, the following conditions hold true.

- We have that $\ker\mathbf{H} \cap \ker\mathbf{L} = \{\mathbf{0}\}$, and the solution set S_h of Problem (24) is a nonempty compact convex set.
- Problem (24) is truly equivalent to the continuous Problem (9) in the discretized search spaces defined in (17) and (6), with
$$\mathcal{S}_h = \underset{f\in\mathcal{M}_{\mathrm{L}_1,h,\boldsymbol{\phi}_0}(\mathbb{R})+\mathcal{M}_{\mathrm{L}_2,h}(\mathbb{R})}{\arg\min} \mathcal{J}(f), \tag{32}$$
in the sense that there exists a bijective linear mapping between the solution sets.
- The reconstructed signal associated to a solution $(\mathbf{c}_1,\mathbf{c}_2) \in S_h$ of Problem (24) is a hybrid spline
$$s = s_1 + s_2 = \sum_{k\in\mathbb{Z}} \Big(c_1[k]\beta_{\mathrm{L}_1}(\cdot - kh) + c_2[k]\beta_{\mathrm{L}_2}(\cdot - kh)\Big), \tag{33}$$
where the sequences $(c_1,c_2) \in \ell_{1,\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{Z}) \times \ell_{1,\mathrm{L}_2}(\mathbb{Z})$ satisfy $c_i|_{I_i} = (c_i[m_i],\ldots,c_i[M_i]) = \mathbf{c}_i$;
- The sparsity of s_i in the Green's function basis is given by $\|\mathbf{L}_i\mathbf{c}_i\|_0$.

The proof of Theorem 2, which is given in Appendix C, is a multicomponent extension of Proposition 2 in [27]. The main difference is to be found in the boundary conditions. Note that, by definition of I_i, the B-spline coefficients outside of I_i do not affect the reconstructed signal outside the interval of interest I_T. Therefore, by Theorem 2, the expression of s over I_T and its sparsity is entirely determined by the coefficients $\mathbf{c}_i$. Computing the complete sequences of B-spline coefficients c_i is therefore unnecessary.

C. Reaching a Sparse Solution

Since Problem (24) can be rewritten as a standard PBP problem of the form

$$S_h = \left\{ \underset{\mathbf{c}\in\mathbb{R}^N}{\arg\min} \left(\|\mathbf{H}\mathbf{c} - \mathbf{y}\|_2^2 + \lambda\|\mathbf{L}\mathbf{c}\|_1\right)\right\}, \tag{34}$$

where $N = N_1 + N_2$, all the results of Section V. C in [27] are applicable. Since $\mathbf{L}$ is clearly of full rank, its null space has dimension $N_{0,2}$, which therefore plays the role of N_0 in [27]. We now sum up the properties of Problem (24).

- Its solution set S_h is a compact convex set whose extreme points $\mathbf{c}^* = (\mathbf{c}_1^*,\mathbf{c}_2^*)$ are sparse, in the sense that they verify $\|\mathbf{L}\mathbf{c}^*\|_0 \le (M - N_{0,2})$.
- All its solutions yield the same measurements $\mathbf{y}_\lambda$, so that $\forall\mathbf{c} \in S_h$, $\mathbf{H}\mathbf{c} = \mathbf{y}_\lambda$.
- It is equivalent to the linear program
$$S_h^{\mathrm{LP}} = \left\{ \underset{(\mathbf{c},\mathbf{u})\in\mathbb{R}^N\times\mathbb{R}^{(N-N_{0,2})}}{\arg\min} \sum_{n=1}^{N-N_{0,2}} u_n : \mathbf{u} + \mathbf{L}\mathbf{c} \ge \mathbf{0}, \quad \mathbf{u} - \mathbf{L}\mathbf{c} \ge \mathbf{0}, \quad \mathbf{H}\mathbf{c} = \mathbf{y}_\lambda \right\}, \tag{35}$$
where for any $\mathbf{x},\mathbf{y} \in \mathbb{R}^P$, $\mathbf{x} \le \mathbf{y}$ implies that $x_p \le y_p$ for all $p \in \{1,\ldots,P\}$. Moreover, an extreme point $(\mathbf{c}^*,\mathbf{u}^*)$ of S_h^{LP} yields an extreme point $\mathbf{c}^*$ of S_h.

The same pipeline can also be used to compute a sparse extreme point of S_h. The procedure is as follows:

Algorithm 1: Pseudocode of our Algorithm.

Input: $\boldsymbol{\nu}$, $N_{0,1}$, $N_{0,2}$, T, $\mathbf{y}$, λ, α, n_0, ϵ
$n = n_0$; $\mathbf{c}_1 = \mathbf{0}$; $\mathbf{c}_2 = \mathbf{0}$; $cost_p = +\infty$; $error = \epsilon + 1$
while $error > \epsilon$ **do**
  $h = T/2^n$
  update $\mathbf{H}$, $\mathbf{L}$; // Depend on $h, \boldsymbol{\nu}, N_{0,1}, N_{0,2}, T$
  $(\mathbf{c}_1, \mathbf{c}_2) \leftarrow \mathrm{ADMM}((\mathbf{c}_{1,\uparrow 2}, \mathbf{c}_{2,\uparrow 2}); \mathbf{H}, \mathbf{L}, \mathbf{y}, \lambda)$
  $error = |J_h(\mathbf{c}_1, \mathbf{c}_2) - cost_p|/cost_p$
  $cost_p = J_h(\mathbf{c}_1, \mathbf{c}_2)$
  $n \leftarrow n + 1$
end
$\mathbf{y}_\lambda = \mathbf{H}\begin{pmatrix}\mathbf{c}_1\\ \mathbf{c}_2\end{pmatrix}$
return $(\mathbf{c}_1^*, \mathbf{c}_2^*) = \mathrm{Simplex}(\mathbf{H}, \mathbf{L}, \mathbf{y}, \lambda, \mathbf{y}_\lambda)$

1) Run any iterative solver (e.g., ADMM) to find a solution $\mathbf{c}_{\mathrm{ADMM}} \in S_h$ of Problem (24).
2) Compute $\mathbf{y}_\lambda = \mathbf{H}\mathbf{c}_{\mathrm{ADMM}}$.
3) Solve the linear program (35) using the simplex or dual-simplex algorithm [40], [41]. These algorithms are known to converge to an extreme point $(\mathbf{c}^*, \mathbf{u}^*)$ of S_h^{LP}, where $\mathbf{c}^*$ is an extreme point of S_h.

V. Multiresolution Strategy

Although the algorithm introduced in Section IV-C yields a sparse solution of Problem (24), it does so for a fixed grid size h. Yet, choosing a suitable grid size clearly depends on the problem at hand. We therefore propose a multiresolution strategy that selects the grid size automatically. We introduce the optimal costs of the continuous and discrete problems, which are defined as $\mathcal{J}^0 = \min_{f \in \mathcal{M}_{\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{R}) + \mathcal{M}_{\mathrm{L}_2}(\mathbb{R})} \mathcal{J}(f)$ and $J_h^0 = \min_{(\mathbf{c}_1,\mathbf{c}_2) \in \mathbb{R}^{N_1} \times \mathbb{R}^{N_2}} J_h(\mathbf{c}_1, \mathbf{c}_2)$, respectively. Our approach is based on Theorem 3.

Theorem 3 (Convergence of the cost function of the discrete problem): Assume that the hypotheses of Theorem 1 and Assumptions 1 are met. Then,

$$\lim_{h \to 0} J_h^0 = \mathcal{J}^0. \tag{36}$$

The proof of Theorem 3, which is based on Theorem 3 in [27], is given in Appendix D. This result justifies our choice of the search space (17) since it contains functions which yield costs arbitrarily close to the optimal cost $\mathcal{J}^0$ of the continuous-domain Problem (9). It also justifies the use of a multiresolution algorithm, the pseudocode of which being given in Algorithm 1.

The principle of this algorithm is to split the grid in half by taking grid sizes $(h_n = T/2^n)_{n \geq n_0}$, where h_{n_0} is an initial coarse grid size. We stop the refinement as soon as the optimal cost $J_{h_n}^0$ ceases to decrease within a certain tolerance ϵ chosen by the user. Further refinement of the grid is then useless, since the reconstructed signal does not vary significantly anymore. The variable $cost_p$ denotes the cost $J_{h_{n-1}}^0$ from the previous grid size. Because the search spaces are embedded like $\mathcal{M}_{\mathrm{L}_1,h_n,\boldsymbol{\phi}_0}(\mathbb{R}) \subset \mathcal{M}_{\mathrm{L}_1,h_{n+1},\boldsymbol{\phi}_0}(\mathbb{R})$ and $\mathcal{M}_{\mathrm{L}_2,h_n}(\mathbb{R}) \subset \mathcal{M}_{\mathrm{L}_2,h_{n+1}}(\mathbb{R})$, a decrease in the grid size can only improve the reconstruction in terms of cost, i.e., $(J_{h_n}^0)_{n \geq n_0}$ is decreasing. Moreover, Theorem 3 states that this quantity converges to $\mathcal{J}^0$, which guarantees the convergence of our algorithm. This embedding also allows us to use the reconstruction from the previous grid size as a warm start for ADMM. This is done by converting the B-spline coefficients to $\mathbf{c}_{i,\uparrow 2}$ for $i \in \{1, 2\}$, which represent the same continuous-domain signal on the finer grid. In practice, for a tolerance $\epsilon = 10^{-3}$, Algorithm 1 typically converges for a grid size of $h = 1/2^9$ or coarser in our experiments, which leads to reasonable computation times of a few seconds in the conditions of Section VI.

VI. Applications

Our algorithms are implemented using GlobalBioIm [42], a Matlab inverse-problem library developed in our group, as well as the Gurobi optimizer[2] for the simplex algorithm. In order for our hybrid regularization (7) to be an adequate prior, the ground truth signal in our experiments is a sparse hybrid spline $s = s_1 + s_2$. Each component s_i is a sparse L_i-spline of the form

$$s_i(x) = \sum_{k=1}^{K_{s_i}} a_{k,i} \rho_{\mathrm{L}_i}(x - x_{k,i}) + \sum_{n=1}^{N_{0,i}} b_{n,i} p_{n,i}(x), \tag{37}$$

where $\{p_{n,i}\}_{n=1}^{N_{0,i}}$ form a basis of $\mathcal{N}_{\mathrm{L}_i}$. The sparsity K_{s_i} of s_i is chosen by the user, the knot locations $x_{i,k} \in I_T$ are chosen at random and the coefficients $a_{k,i}$ and $b_{n,i}$ are i.i.d. Gaussian random variables. The tuning parameters λ and α are selected using grid search, and we pick a tolerance $\epsilon = 10^{-3}$.

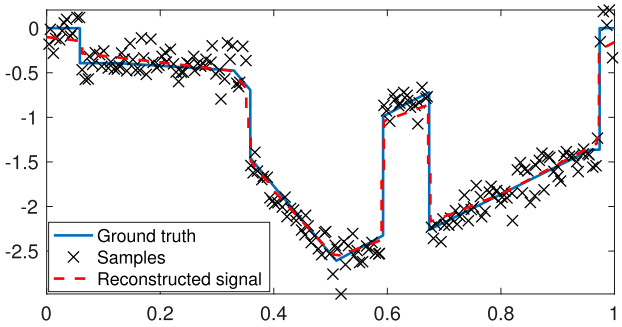
A. Curve Fitting

Curve fitting is particularly well-suited for smoothing problems, which consist in fitting a continuous-domain function which is sparse in a certain dictionary basis from many noisy data points. The measurement functionals are then given by $\nu_m(s) = s(x_m)$ where $x_m \in [0, T]$ are the sampling locations. Such reconstruction problems are commonly tackled in a single-operator framework, where the dictionary is associated with a single brand of splines. A limit case is polynomial regression, which is achieved by taking $\mathrm{L} = \mathrm{D}^{N+1}$ and picking a very high value of the regularization parameter λ.

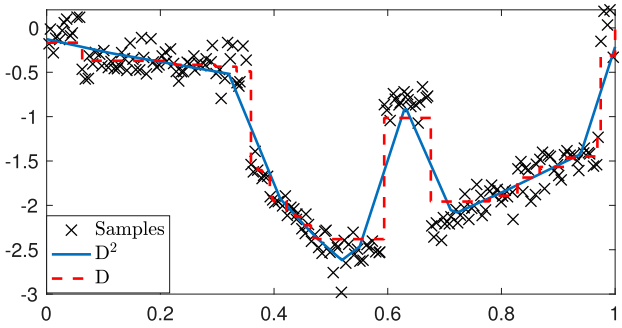
In contrast with standard single-operator frameworks, our approach allows for the joint use of several families of basis functions, and can therefore represent a richer class of signals. An example of a curve-fitting reconstruction using Algorithm 1 is given in Fig. 3(a). The chosen regularization operators are $\mathrm{L}_1 = \mathrm{D}$ and $\mathrm{L}_2 = \mathrm{D}^2$; our dictionary thus consists of both piecewise-constant and piecewise-linear splines. We compare our results to single-operator reconstructions with regularization operators $\mathrm{L} = \mathrm{D}$ and $\mathrm{L} = \mathrm{D}^2$ in Fig. 3(b).

Observe that the reconstructed signal is quite sparse ($K_1 = 14$ and $K_2 = 6$), and is satisfactory in that it is close to what a human would reconstruct, except for some shearing effects around $x = 0$ and $x = 0.6$ that are typical of regularization methods. It was not obvious *a priori* that such results could be

[2]LLC Gurobi Optimization, Gurobi optimizer reference manual, 2018.

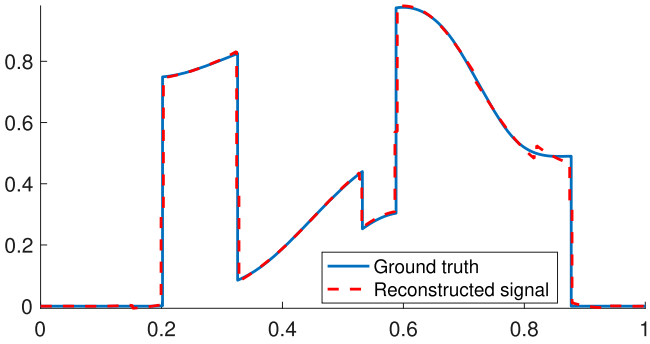


(a) Hybrid spline reconstruction with $\mathrm{L}_1 = \mathrm{D}$, $\mathrm{L}_2 = \mathrm{D}^2$, $\lambda = 2$, $\alpha = 0.04$ and grid size $h = 1/2^9$. Sparsity of the reconstruction: $14 + 6 = 20$.

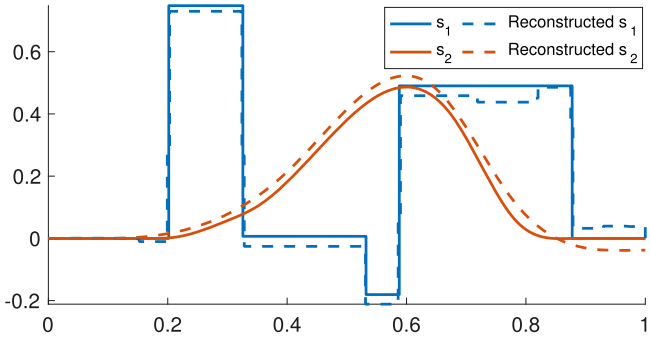


(b) Single-operator reconstructions $\mathrm{L} = \mathrm{D}$ ($\lambda = 3$, sparsity 26) and $\mathrm{L} = \mathrm{D}^2$ ($\lambda = 0.1$, sparsity 15).

Fig. 3. Curve fitting comparison with $M = 200$ data points.



(a) Complete reconstructed signal (SNR = 22.2 dB, sparsity 17 + 7 = 24)



(b) Separate components

Fig. 4. Reconstruction result with noiseless Fourier measurements for $\mathrm{L}_1 = \mathrm{D}, \mathrm{L}_2 = \mathrm{D}^4$, $M = 30$, $\lambda = 10^{-15}$, $\alpha = 5 \times 10^{-5}$. Final grid size: $h = 1/2^8$.

achieved. Indeed, although we are using the sparsity-promoting ℓ_1 norm, the large number of measurements M leads us far from a CS-type framework. Theorem 2 in [27] states that the sparsity of the reconstructed signal is bounded by $(M - N_{0,2})$. We observe experimentally that increasing λ tends to produce sparser solutions, and that the level of sparsity can be adjusted at will by tuning λ. This is not altogether surprising, since $\lambda = +\infty$ yields a reconstructed signal in the null space of L_2 with sparsity zero.

Another promising feature of this experiment is that our algorithm is able to strike the right balance between both dictionaries (i.e., D-splines and D^2-splines), in that the selected dictionary elements explain the ground truth signal well. Note that this requires careful tuning of the weight parameter α. Indeed, a badly tuned α leads to a lopsided use of one of the dictionaries, the other being too strongly penalized. We observe that a suitable balance can consistently be found when the ground truth fits the signal model as in (37).

By contrast, the single-operator reconstructions in Fig. 3(b) do not do well in regions that call for different dictionary elements. More precisely, linear regimes in the ground truth signal lead to a staircasing effect in the piecewise-constant reconstruction (D), which results in a loss of sparsity. As for the sharp jumps in the ground truth, they lead to gradual increases in the piecewise-linear reconstruction (D^2). This is because sharp increases are heavily penalized by the regularization, which is undesirable given the form of the ground truth.

B. Compressed Sensing

A second application of our framework is CS-type problems, which attempt to the recover a sparse multicomponent signals given a small number of measurements. [3] We use the same type of test signals as in (37), namely hybrid splines with low sparsity $K_{s_1} + K_{s_2}$. The measurement functionals are assumed to take the form

$$\nu_m(s) = \int_0^T \cos(\omega_m x + \phi_m) s(x) \mathrm{d}x, \tag{38}$$

where $\omega_m \in \mathbb{R}$ are the sampling frequencies and $\phi_m \in [0, 2\pi)$ some given phase offsets. This amounts to sampling in the Fourier space — a rectangular window is applied in order to satisfy Assumptions 1. We choose Fourier measurements because they are known to have good recovery properties according to the theory of CS [43]. However, the absence of a D-RIP-type assumption prevents us from making any theoretical claims on the quality of the recovery.

In this setting, perfect recovery is illusory since the $x_{k,i}$ are not on a grid and the s_i signals are thus not included in the native spaces $\mathcal{M}_{\mathrm{L}_1,\phi_0,h}(\mathbb{R})$ and $\mathcal{M}_{\mathrm{L}_2,h}(\mathbb{R})$. Even in purely discrete settings, the recovery of a multicomponent discrete signal is a difficult problem, which requires conditions on the measurement matrix $\mathbf{H}$ and some form of incoherence between

[3] Note that we are not interested in the data separation problem as in [26], but only in recovering the complete multicomponent signal.

dictionaries [11], [26]. Here, no such assumptions are made, which makes the recovery problem even more challenging. To avoid adding to the difficulty, we only consider noiseless measurements.

An example run is shown in Fig. 4 for a test signal with sparsity 15 ($K_{s_1} = 5$ and $K_{s_2} = 10$) and $M = 30$. Since the measurements are noiseless and we are interested in recovering the test signal as faithfully as possible, the data-fidelity term should be penalized much more than the regularization term. We therefore pick the regularization parameter $\lambda = 10^{-15} \ll 1$. The reconstructed signal in Fig. 4(a) is remarkably close to the test signal, considering the difficulties of the problem. Notice that the final grid size selected by Algorithm 1 ($h = 1/2^8$) is still computationally tractable. The separate components of the reconstructed signal compared to those of the test signal are provided in Fig. 4(b). We observe that the separation is not perfect: there is a small compensation effect between the two reconstructed components.

VII. Conclusion

We have established a representer theorem that states that hybrid splines are solutions of continuous-domain inverse problems with multiple generalized total-variation regularization. The regularization operators L_1 and L_2 are taken to be multiple-order derivatives, which lead to piecewise-polynomial splines. This result implies that such problems can be solved exactly using a concatenated dictionary that consists of L_1 and L_2-splines. We propose an exact B-spline-based discretization scheme and a multiresolution algorithm to solve the continuous-domain problem in a suitable search space. We then apply our algorithm to curve fitting and CS-type problems, and show that it is both computationally feasible and very successful experimentally.

To the best of our knowledge, this is the first instance of a continuous-domain inverse problem for multicomponent signals. Our algorithm can be viewed as the continuous-domain counterpart of discrete data-separation problems, such as morphological component analysis.

Appendix

A. Proof of Proposition 1

Recall that $(\boldsymbol{\phi}_0, \mathbf{p}_0)$ is a biorthogonal system in the sense of Definition 3 in [21] for $\mathcal{N}_0 = \mathcal{N}_{\mathrm{L}_1} \cap \mathcal{N}_{\mathrm{L}_2}$. The latter can be extended to two biorthogonal systems $(\tilde{\boldsymbol{\phi}}_1, \tilde{\mathbf{p}}_1)$ and $(\tilde{\boldsymbol{\phi}}_2, \tilde{\mathbf{p}}_2)$ for $\mathcal{N}_{\mathrm{L}_1}$ and $\mathcal{N}_{\mathrm{L}_2}$, respectively. There, $\tilde{\boldsymbol{\phi}}_i = [\boldsymbol{\phi}_0 \ \boldsymbol{\phi}_i]$ and $\tilde{\mathbf{p}}_i = [\mathbf{p}_0 \ \mathbf{p}_i]$ for $i \in \{1, 2\}$.

Based on Theorem 4 in [21], we know that there exist stable right-inverse operators $\mathrm{L}_{\tilde{\boldsymbol{\phi}}_i}^{-1}$ such that any $f \in \mathcal{M}_{\mathrm{L}_i}(\mathbb{R})$ has a unique representation as

$$f(x) = \mathrm{L}_{\tilde{\boldsymbol{\phi}}_i}^{-1}\{w_i\}(x) + \mathbf{c}_0^T \mathbf{p}_0(x) + \mathbf{c}_i^T \mathbf{p}_i(x), \tag{39}$$

where $w_i \in \mathcal{M}(\mathbb{R})$, $\mathbf{c}_0 \in \mathbb{R}^{N_0}$, and $\mathbf{c}_i \in \mathbb{R}^{N_{0,i} - N_0}$ with $i \in \{1, 2\}$. We then proceed in two steps.

1) The Minimum is Reached in (7):

Proof: Let us denote

$$\mathcal{C}_f = \{(f_1, f_2) \in \mathcal{M}_{\mathrm{L}_1, \boldsymbol{\phi}_0}(\mathbb{R}) \times \mathcal{M}_{\mathrm{L}_2}(\mathbb{R}) : f = f_1 + f_2\} \tag{40}$$

the set of feasible pairs (f_1, f_2) and

$$\mathcal{R}_{1,2}(f_1, f_2) = (1 - \alpha)\|\mathrm{L}_1\{f_1\}\|_{\mathcal{M}} + \alpha\|\mathrm{L}_2\{f_2\}\|_{\mathcal{M}}. \tag{41}$$

The hybrid regularization can be rewritten as

$$\mathcal{R}_{\mathrm{hyb}}(f) = \min_{(f_1, f_2) \in \mathcal{C}_f} \mathcal{R}_{1,2}(f_1, f_2). \tag{42}$$

The feasible set $\mathcal{C}_f$ is nonempty since $f \in \mathcal{M}_{\mathrm{L}_1, \boldsymbol{\phi}_0}(\mathbb{R}) + \mathcal{M}_{\mathrm{L}_2}(\mathbb{R})$. Therefore, there exists a sequence $(f_{1,k}, f_{2,k})_{k \in \mathbb{N}} \in \mathcal{C}_f$ that monotonically decreases to the infimum of $\mathcal{R}_{1,2}$ over $\mathcal{C}_f$. Using (39), both components can be represented as $f_{1,k} = \mathrm{L}_{\tilde{\boldsymbol{\phi}}_1}^{-1}\{w_{1,k}\} + \mathbf{c}_{1,k}^T \mathbf{p}_1$ and $f_{2,k} = \mathrm{L}_{\tilde{\boldsymbol{\phi}}_2}^{-1}\{w_{2,k}\} + \mathbf{c}_{0,k}^T \mathbf{p}_0 + \mathbf{c}_{2,k}^T \mathbf{p}_2$. The assumption of monotonic decrease implies that

$$\forall k \in \mathbb{N}, i \in \{1, 2\} : \quad 0 \leq \|w_{i,k}\|_{\mathcal{M}} \leq A, \tag{43}$$

where $A = \mathcal{R}_{1,2}(f_{1,0}, f_{2,0})$. Next, since $(f_{1,k}, f_{2,k})_{k \in \mathbb{N}} \in \mathcal{C}_f$, we have that

$$\forall k \in \mathbb{N} : f = \mathrm{L}_{\tilde{\boldsymbol{\phi}}_1}^{-1}\{w_{1,k}\} + \mathrm{L}_{\tilde{\boldsymbol{\phi}}_2}^{-1}\{w_{2,k}\} + \sum_{i=0}^{2} \mathbf{c}_{i,k}^T \mathbf{p}_i. \tag{44}$$

Therefore, for all test functions $\varphi \in \mathcal{S}(\mathbb{R})$, we have that

$$\begin{aligned} \langle f, \varphi \rangle = &\left\langle w_{1,k}, \mathrm{L}_{\tilde{\boldsymbol{\phi}}_1}^{-1*}\{\varphi\} \right\rangle + \left\langle w_{2,k}, \mathrm{L}_{\tilde{\boldsymbol{\phi}}_2}^{-1*}\{\varphi\} \right\rangle \\ &+ \left\langle \sum_{i=0}^{2} \mathbf{c}_{i,k}^T \mathbf{p}_i, \varphi \right\rangle, \end{aligned} \tag{45}$$

where $\mathrm{L}_{\tilde{\boldsymbol{\phi}}_1}^{-1*}$ is the adjoint operator of $\mathrm{L}_{\tilde{\boldsymbol{\phi}}_1}^{-1}$. Let $B_k = \max_i(\|\mathbf{c}_{i,k}\|_\infty)$, and

$$\mathbf{b}_{i,k} = \begin{cases} \dfrac{\mathbf{c}_{i,k}}{B_k}, & B_k \neq 0, \\ \mathbf{1}, & \text{otherwise.} \end{cases} \tag{46}$$

Clearly, the vectors $\mathbf{b}_{i,k}$ are in the unit ball. Moreover, as stated in (43), $w_{i,k}$ is in the ball of radius A, which is weak*-compact according to the Banach-Alaoglu theorem. Therefore, there exists a subsequence $(f_{1,k_n}, f_{2,k_n})_{n \in \mathbb{N}}$ with the following properties

- the sequences $(w_{1,k_n})_{n \in \mathbb{N}}$ and $(w_{2,k_n})_{n \in \mathbb{N}}$ are converging for the weak*-topology. Their limits are denoted by $w_{1,\lim}, w_{2,\lim} \in \mathcal{M}(\mathbb{R})$, respectively;
- there exist vectors $\mathbf{b}_{i,\lim}$ such that $\lim_{n \to +\infty} \mathbf{b}_{i,k_n} = \mathbf{b}_{i,\lim}$ for $i \in \{0, 1, 2\}$, at least one of which is nonzero since, for any $k \in \mathbb{N}$, there exists an $i \in \{0, 1, 2\}$ such that $\mathbf{b}_{i,k} = \mathbf{1}$;
- we have the convergence $\lim_{n \to +\infty} B_{k_n} = B_{\lim}$ where $B_{\lim} \in \mathbb{R} \cup \{+\infty\}$.

Rewriting (45) for $k = k_n$ yields

$$\langle f, \varphi \rangle - \left\langle w_{1,k_n}, \mathrm{L}_{\tilde{\boldsymbol{\phi}}_1}^{-1*}\{\varphi\} \right\rangle - \left\langle w_{2,k_n}, \mathrm{L}_{\tilde{\boldsymbol{\phi}}_2}^{-1*}\{\varphi\} \right\rangle = B_{k_n} \left\langle \sum_{i=0}^{2} \mathbf{b}_{i,k_n}^T \mathbf{p}_i, \varphi \right\rangle. \quad (47)$$

Assume by contradiction that $B_{\lim} = +\infty$. As the left-hand side of the equality in (47) converges to a finite limit, so must the right-hand side, which implies that $\lim_{n\to+\infty} \langle \sum_{i=0}^{2} \mathbf{b}_{i,k_n}^T \mathbf{p}_i, \varphi \rangle = 0$. Since $\lim_{n\to+\infty} \langle \sum_{i=0}^{2} \mathbf{b}_{i,k_n}^T \mathbf{p}_i, \varphi \rangle = \langle q_{\lim}, \varphi \rangle$ for all $\varphi \in \mathcal{S}(\mathbb{R})$, where $q_{\lim} = \sum_{i=0}^{2} \mathbf{b}_{i,\lim}^T \mathbf{p}_i$, it follows that $q_{\lim} = 0$. Yet, $\{\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2\}$ are linearly independent while not all $\mathbf{b}_{i,\lim}$ are zero, which yields a contradiction. Therefore, we must have that $B_{\lim} < +\infty$.

Let $f_1 = \mathrm{L}_{\tilde{\boldsymbol{\phi}}_1}^{-1}\{w_{1,\lim}\} + \mathbf{c}_{1,\lim}^T \mathbf{p}_1$ and $f_{2,\lim} = \mathrm{L}_{\tilde{\boldsymbol{\phi}}_2}^{-1}\{w_{2,\lim}\} + \mathbf{c}_{0,\lim}^T \mathbf{p}_0 + \mathbf{c}_{2,\lim}^T \mathbf{p}_2$, where $\mathbf{c}_{i,\lim} = B_{\lim} \mathbf{b}_{i,\lim}$ for $i \in \{0,1,2\}$. By taking the limit in (47), we get that

$$\langle f, \varphi \rangle - \left\langle w_{1,\lim}, \mathrm{L}_{\tilde{\boldsymbol{\phi}}_1}^{-1*}\{\varphi\} \right\rangle - \left\langle w_{2,\lim}, \mathrm{L}_{\tilde{\boldsymbol{\phi}}_2}^{-1*}\{\varphi\} \right\rangle = \left\langle \sum_{i=0}^{2} \mathbf{c}_{i,\lim}^T \mathbf{p}_i, \varphi \right\rangle. \quad (48)$$

Since (48) is valid for all $\varphi \in \mathcal{S}(\mathbb{R})$, we have that $f = f_{1,\lim} + f_{2,\lim}$ and $\mathcal{R}_{\mathrm{hyb}}(f) = \mathcal{R}_{1,2}(f_{1,\lim}, f_{2,\lim})$. ■

2) The Regularizer $\mathcal{R}_{\mathrm{hyb}}$ is Convex:

Proof: Consider two functions $f, g \in \mathcal{M}_{\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{R}) + \mathcal{M}_{\mathrm{L}_2}(\mathbb{R})$ and let $f = f_1 + f_2$ and $g = g_1 + g_2$ be their decomposition, as specified by (8). Denote $h_\beta = \beta f + (1-\beta) g$ as a convex combination of f and g, where $\beta \in [0,1]$. Since $h_\beta = (\beta f_1 + (1-\beta) g_1) + (\beta f_2 + (1-\beta) g_2)$, we have that

$$\begin{aligned} \mathcal{R}_{\mathrm{hyb}}(h_\beta) &\le (1-\alpha)\|\beta f_1 + (1-\beta) g_1\|_{\mathcal{M}} \\ &\quad + \alpha \|\beta f_2 + (1-\beta) g_2\|_{\mathcal{M}} \\ &\le (1-\alpha)(\beta \|f_1\|_{\mathcal{M}} + (1-\beta)\|g_1\|_{\mathcal{M}}) \\ &\quad + \alpha(\beta \|f_2\|_{\mathcal{M}} + (1-\beta)\|g_2\|_{\mathcal{M}}) \\ &= \beta \mathcal{R}_{\mathrm{hyb}}(f) + (1-\beta)\mathcal{R}_{\mathrm{hyb}}(g). \end{aligned} \quad (49)$$

■

B. Proof of Theorem 1

The proof is divided into two parts. We first prove the existence of a solution, and then that some solutions are of the form (10). Note that we use some notations and results introduced as preliminaries in Appendix A.

1) Existence of a Solution:

Proof: Consider a sequence $(f_k)_{k\in\mathbb{N}} \in \mathcal{M}_{\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{R}) + \mathcal{M}_{\mathrm{L}_2}(\mathbb{R})$ that monotonically decreases to the infimum value $\mathcal{J}_0$ of the cost functional. For every k, consider the decomposition $f_k = f_{1,k} + f_{2,k}$ given by (8), such that $\mathcal{R}_{\mathrm{hyb}}(f_k) = (1-\alpha)\|\mathrm{L}_1\{f_{1,k}\}\|_{\mathcal{M}} + \alpha \|\mathrm{L}_2\{f_{2,k}\}\|_{\mathcal{M}}$.

Expanding $f_{1,k}$ and $f_{2,k}$ using (39) yields

$$f_{1,k} = \mathrm{L}_{\tilde{\boldsymbol{\phi}}_1}^{-1}\{w_{1,k}\} + \mathbf{c}_{1,k}^T \mathbf{p}_1, \quad (50)$$

$$f_{2,k} = \mathrm{L}_{\tilde{\boldsymbol{\phi}}_2}^{-1}\{w_{2,k}\} + \mathbf{c}_{0,k}^T \mathbf{p}_0 + \mathbf{c}_{2,k}^T \mathbf{p}_2. \quad (51)$$

The proof will now consist in extracting a converging subsequence of $(f_k)_{k\in\mathbb{N}}$. To achieve this, the first step is to prove that the sequences $(w_{i,k})_{k\in\mathbb{N}}$ $(i \in \{1,2\})$ and $(\mathbf{c}_{i,k})_{k\in\mathbb{N}}$ $(i \in \{0,1,2\})$ are bounded. ■

Due to the assumption of monotonic decrease, we have that

$$\forall k \in \mathbb{N}: \quad \|w_{1,k}\|_{\mathcal{M}}, \|w_{2,k}\|_{\mathcal{M}} \le C_1, \quad (52)$$

where $C_1 = \frac{\mathcal{J}(f_1)}{\lambda \min(\alpha, 1-\alpha)}$. Next, we bound the $(\mathbf{c}_{i,k})_{k\in\mathbb{N}}$ sequences. Using the triangular inequality, we get that

$$\begin{aligned} \|\boldsymbol{\nu}(f_k)\|_2 \ge & \left\| \boldsymbol{\nu} \left(\sum_{i=0}^{2} \mathbf{c}_{i,k}^T \mathbf{p}_i \right) \right\|_2 \\ & - \left\| \boldsymbol{\nu} \left(\mathrm{L}_{\tilde{\boldsymbol{\phi}}_1}^{-1}\{w_{1,k}\} + \mathrm{L}_{\tilde{\boldsymbol{\phi}}_2}^{-1}\{w_{2,k}\} \right) \right\|_2. \end{aligned} \quad (53)$$

Using Proposition 8 in [21], the well-posedness assumption in Theorem 1 is equivalent to the existence of a constant $B > 0$ such that

$$\left\| \boldsymbol{\nu} \left(\sum_{i=0}^{2} \mathbf{c}_{i,k}^T \mathbf{p}_i \right) \right\|_2 \ge B \sum_{i=0}^{2} \|\mathbf{c}_{i,k}\|_2. \quad (54)$$

Next, to handle the second term in (53), we prove Lemma 1.

Lemma 1: The operators $\boldsymbol{\nu}(\mathrm{L}_{\tilde{\boldsymbol{\phi}}_i}^{-1}\{\cdot\}) : \mathcal{M}(\mathbb{R}) \to \mathbb{R}^M$ with $i \in \{1,2\}$ are weak*-continuous.

Proof: Let $(w_n)_{n\in\mathbb{N}} \in \mathcal{M}(\mathbb{R})$ be a sequence that converges to $w \in \mathcal{M}(\mathbb{R})$ for the weak*-topology. Since $\boldsymbol{\nu}$ is weak*-continuous, it is sufficient to prove that $\mathrm{L}_{\tilde{\boldsymbol{\phi}}_i}^{-1}\{w_n - w\} \to 0$ for the weak*-topology in $\mathcal{M}_{\mathrm{L}_i}(\mathbb{R})$. According to Theorem 6 in [21], we have that $(C_{\mathrm{L}_i}(\mathbb{R}))' = \mathcal{M}_{\mathrm{L}_i}(\mathbb{R})$, where $C_{\mathrm{L}_i}(\mathbb{R}) = \mathrm{L}_i^*\{C_0(\mathbb{R})\} \oplus \mathcal{N}'_{\mathrm{L}_i}$. Let $f \in C_{\mathrm{L}_i}(\mathbb{R})$; there exists functions $(f_1, f_2) \in C_0(\mathbb{R}) \times \mathcal{N}'_{\mathrm{L}_i}$ such that $f = \mathrm{L}_i^*\{f_1\} + f_2$. Using (18) in [21], we have that $\tilde{\boldsymbol{\phi}}_i(\mathrm{L}_{\tilde{\boldsymbol{\phi}}_i}^{-1}\{w_n - w\}) = \mathbf{0}$ which, by definition of $\mathcal{N}'_{\mathrm{L}_i}$, implies that $\langle \mathrm{L}_{\tilde{\boldsymbol{\phi}}_i}^{-1}\{w_n - w\}), f_2 \rangle = 0$. We thus have that

$$\begin{aligned} \left\langle \mathrm{L}_{\tilde{\boldsymbol{\phi}}_i}^{-1}\{w_n - w\}), f \right\rangle &= \left\langle \mathrm{L}_{\tilde{\boldsymbol{\phi}}_i}^{-1}\{w_n - w\}), \mathrm{L}_i^*\{f_1\} \right\rangle \\ &= \langle w_n - w, f_1 \rangle \to 0 \end{aligned} \quad (55)$$

since $w_n \to w$ for the weak*-topology. This proves that $\mathrm{L}_{\tilde{\boldsymbol{\phi}}_i}^{-1}\{w_n - w\} \to 0$ for the weak*-topology and, thus, the desired result. ■

Next, Lemma 1 implies that the operators $\boldsymbol{\nu}(\mathrm{L}_{\tilde{\boldsymbol{\phi}}_i}^{-1}\{\cdot\})$ are continuous for the norm topology $\|\cdot\|_{\mathcal{M}}$ since the latter is stronger than the weak*-topology (i.e., it has more open sets). This implies the existence of a constant $A > 0$ such that

$$\left\| \boldsymbol{\nu} \left(\mathrm{L}_{\tilde{\boldsymbol{\phi}}_1}^{-1}\{w_{1,k}\} + \mathrm{L}_{\tilde{\boldsymbol{\phi}}_2}^{-1}\{w_{2,k}\} \right) \right\|_2 \le A \left(\|w_{1,k}\|_{\mathcal{M}} + \|w_{2,k}\|_{\mathcal{M}} \right). \quad (56)$$

Combining (52), (54), and (56) yields that

$$\|\boldsymbol{\nu}(f_k)\|_2 \geq B\left(\sum_{i=0}^{2}\|\mathbf{c}_{i,k}\|_2\right) - A\left(\|w_{1,k}\|_{\mathcal{M}} + \|w_{2,k}\|_{\mathcal{M}}\right)$$
$$\geq B\left(\sum_{i=0}^{2}\|\mathbf{c}_{i,k}\|_2\right) - 2AC_1. \tag{57}$$

Using the assumption of monotonic decrease and the triangular inequality, we get that

$$\forall k \in \mathbb{N} : \sqrt{\mathcal{J}(f_0)} + \|\mathbf{y}\|_2 \geq \|\boldsymbol{\nu}(f_k)\|_2. \tag{58}$$

This shows that the norms of $\mathbf{c}_{0,k}$, $\mathbf{c}_{1,k}$, and $\mathbf{c}_{2,k}$ are bounded by a constant, which was our initial goal. Together with (52), this implies the existence of a subsequence $(f_{k_n})_{n\in\mathbb{N}}$ such that

- the sequences w_{i,k_n} are converging to a limit $w_{i,\lim} \in \mathcal{M}(\mathbb{R})$ for the weak*-topology for $i \in \{1,2\}$ (Banach-Alaoglu theorem);
- for $i \in \{0,1,2\}$, the sequences $\mathbf{c}_{i,k_n}$ converge to a limit $\mathbf{c}_{i,\lim}$.

Using Lemma 1, we thus have that $\lim_{n\to+\infty}\boldsymbol{\nu}(f_{k_n}) = \boldsymbol{\nu}(f_{\lim})$ where

$$f_{\lim} = \mathrm{L}_{\tilde{\boldsymbol{\phi}}_1}^{-1}\{w_{1,\lim}\} + \mathrm{L}_{\tilde{\boldsymbol{\phi}}_2}^{-1}\{w_{2,\lim}\} + \sum_{i=0}^{2}\mathbf{c}_{i,\lim}^T\mathbf{p}_i. \tag{59}$$

This yields that

$$\mathcal{J}_0 = \lim_{n\to+\infty}\mathcal{J}(f_{k_n})$$
$$= \|\boldsymbol{\nu}(f_{\lim}) - \mathbf{y}\|_2^2 + \lambda((1-\alpha)\|w_{1,\lim}\|_{\mathcal{M}} + \alpha\|w_{2,\lim}\|_{\mathcal{M}}) \tag{60}$$
$$= \mathcal{J}(f_{\lim}), \tag{61}$$

using the weak*-continuity of the $\|\cdot\|_{\mathcal{M}}$ norm in (60). This proves the desired result $f_{\lim} \in \mathcal{S}$.

2) Form of the Solutions:

Lemma 2: All solutions of Problem (9) yield the same measurement vector $\mathbf{z}_\lambda \in \mathbb{R}^M$.

Proof: Let f, g be two solutions of Problem (9) and define $h = \frac{f+g}{2}$. Since $\mathcal{R}_{\mathrm{hyb}}$ (Proposition 1) and $\|\cdot\|_2^2$ are convex functionals, we have that

$$\mathcal{R}_{\mathrm{hyb}}(h) \leq \frac{\mathcal{R}_{\mathrm{hyb}}(f) + \mathcal{R}_{\mathrm{hyb}}(g)}{2} \text{ and} \tag{62}$$
$$\|\boldsymbol{\nu}(h) - \mathbf{y}\|_2^2 \leq \frac{\|\boldsymbol{\nu}(f) - \mathbf{y}\|_2^2 + \|\boldsymbol{\nu}(g) - \mathbf{y}\|_2^2}{2}. \tag{63}$$

Summing these inequalities yields $\mathcal{J}(h) \leq \mathcal{J}_0$. Yet $\mathcal{J}(h) = \mathcal{J}_0$, which implies that the cases of equality are met. Since $\|\cdot\|_2^2$ is strictly convex, we necessarily have that $\boldsymbol{\nu}(f) = \boldsymbol{\nu}(g) = \mathbf{z}_\lambda$. ■

Using Lemma 2, Problem (9) can be reformulated as

$$\mathcal{S} = \left\{ \underset{f\in\mathcal{M}_{\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{R})+\mathcal{M}_{\mathrm{L}_2}(\mathbb{R})}{\arg\min} \mathcal{R}_{\mathrm{hyb}}(f) \quad \text{s.t.} \quad \boldsymbol{\nu}(f) = \mathbf{z}_\lambda \right\}. \tag{64}$$

Now, consider the problem

$$\min_{\substack{w_1,w_2\in\mathcal{M}(\mathbb{R}) \\ q\in\mathcal{N}_{\mathrm{L}_1}+\mathcal{N}_{\mathrm{L}_2} \\ F(w_1,w_2,q)=\mathbf{z}_\lambda}} \Big((1-\alpha)\|w_1\|_{\mathcal{M}} + \alpha\|w_2\|_{\mathcal{M}}\Big), \quad \text{where} \tag{65}$$

$$F(w_1,w_2,q) = \boldsymbol{\nu}\left(\mathrm{L}_{\tilde{\boldsymbol{\phi}}_1}^{-1}\{w_1\} + \mathrm{L}_{\tilde{\boldsymbol{\phi}}_2}^{-1}\{w_2\} + q\right) \tag{66}$$

is a continuous linear functional. Due to the well-posedness assumption,

$$\forall q \in \mathcal{N}_{\mathrm{L}_1} + \mathcal{N}_{\mathrm{L}_2} : \quad F(0,0,q) \geq B\|q\|_{\mathcal{N}_{\mathrm{L}_1}+\mathcal{N}_{\mathrm{L}_2}}. \tag{67}$$

Using the vector-valued Fisher-Jerome theorem [44], Problem (65) has a solution (w_1, w_2, q) such that

$$w_i = \sum_{k=1}^{K_i} a_{i,k}\delta(\cdot - x_{i,k}), \quad i \in \{1,2\}, \tag{68}$$

where $a_{i,n}, x_{i,k} \in \mathbb{R}$ and $K_1 + K_2 \leq M$. Next, let

$$s = \mathrm{L}_{\tilde{\boldsymbol{\phi}}_1}^{-1}\{w_1\} + \mathrm{L}_{\tilde{\boldsymbol{\phi}}_2}^{-1}\{w_2\} + q. \tag{69}$$

Clearly, $\boldsymbol{\nu}(s) = \mathbf{z}_\lambda$ and $s = s_1 + s_2$ where the components s_i are of the form (10). Now, assume by contradiction that s is not a solution of Problem (9). According to Lemma 2, this implies the existence of a function $g \in \mathcal{M}_{\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{R}) + \mathcal{M}_{\mathrm{L}_2}(\mathbb{R})$ such that $\boldsymbol{\nu}(g) = \mathbf{z}_\lambda$ and $\mathcal{R}_{\mathrm{hyb}}(g) < \mathcal{R}_{\mathrm{hyb}}(s)$. As stated in (39), g can rewritten as

$$g = \mathrm{L}_{\tilde{\boldsymbol{\phi}}_1}^{-1}\{u_1\} + \mathrm{L}_{\tilde{\boldsymbol{\phi}}_2}^{-1}\{u_2\} + r, \tag{70}$$

where $u_1, u_2 \in \mathcal{M}(\mathbb{R})$ and $r \in \mathcal{N}_{\mathrm{L}_1} + \mathcal{N}_{\mathrm{L}_2}$. Thus, $F(u_1, u_2, r) = \mathbf{z}_0$ and

$$(1-\alpha)\|u_1\|_{\mathcal{M}} + \alpha\|u_2\|_{\mathcal{M}} = \mathcal{R}_{\mathrm{hyb}}(g)$$
$$< \mathcal{R}_{\mathrm{hyb}}(s) = (1-\alpha)\|w_1\|_{\mathcal{M}} + \alpha\|w_2\|_{\mathcal{M}}, \tag{71}$$

which is in contradiction with (w_1, w_2, q) being a solution of (65). This proves that s is indeed the solution of (9) and thus the desired result. ■

C. Proof of Theorem 2

We first prove the Lemma 3.

Lemma 3: Let $s^* = s_1^* + s_2^* \in \mathcal{S}_h$ be a solution of Problem (32) and let $(c_1^*, c_2^*) \in \ell_{1,\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{Z}) \times \ell_{1,\mathrm{L}_2}(\mathbb{Z})$ be the corresponding sequences of B-spline coefficients. Under Assumptions 1, c_i^* is uniquely determined by its N_i coefficients $c^*|_{i_i}$.

Proof: Let us first observe that Problem (32) can be rewritten by using the B-spline representations of $\mathcal{M}_{\mathrm{L}_1,h,\boldsymbol{\phi}_0}(\mathbb{R})$ in (6) and $\mathcal{M}_{\mathrm{L}_2,h}(\mathbb{R})$ in (17). By optimizing over the spline coefficients, we get the optimization problem

$$\mathcal{S}_h = \left\{ \underset{(c_1,c_2)\in\ell_{1,\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{Z})\times\ell_{1,\mathrm{L}_2}(\mathbb{Z})}{\arg\min} \mathcal{J}_h(c_1,c_2) \right\}, \tag{72}$$

where the reconstructed signal for $(c_1^*, c_2^*) \in \mathcal{S}_h$ is given by $s^* = s_1^* + s_2^*$, whose components s_i^* are L_i-splines with B-spline coefficients c_1^*. From now on, in this proof, let $\mathcal{J}_h$ denote the cost

function and $\mathcal{S}_h$ the solution set for sequences of B-spline coefficients $(c_1, c_2) \in \ell_{1,\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{Z}) \times \ell_{1,\mathrm{L}_2}(\mathbb{Z})$. As demonstrated by (25) in [27], the regularization term associated to a signal $s_i \in \mathcal{M}_{\mathrm{L}_i}(\mathbb{R})$ with B-spline coefficients $c_i \in \ell_{1,\mathrm{L}_i}(\mathbb{Z})$ can be expressed as $\|\mathrm{L}_i\{s_i\}\|_{\mathcal{M}} = 1/h^{N_{0,i}-1}\|d_{\mathrm{L}_i} * c_i\|_1$. Next, consider a sequence c_1 such that $c_1|_{i_1} = c_1^*|_{i_1}$, $c_1[-N_{0,1}+1] = \cdots = c_1[0] = 0$ (boundary conditions in Assumptions 1), and whose remaining coefficients are free. This freedom does not affect the data fidelity term due to the finite-support assumption on ν_m. The coefficient $c_1[M_1+1]$ is uniquely chosen such that $(d_{\mathrm{L}_1} * c_1)[M_1+1] = \sum_{k=0}^{N_{0,1}} d_{\mathrm{L}_1}[k]c_1[M_1-k+1] = 0$. Similarly, all $c[k]$ coefficients for $k > M_1$ can be uniquely determined recursively to nullify $(d_{\mathrm{L}_1} * c_1)[k]$ as a linear combination of the $(N_{0,1}-1)$ previous coefficients.

Next, due to the boundary conditions, we have that $(d_{\mathrm{L}_1} * c_1)[0] = \sum_{k=0}^{N_{0,1}} d_{\mathrm{L}_1}[k]c_1[-k] = 0 \Rightarrow c[-N_{0,1}] = 0$. Analogously, the unique way of canceling all coefficients $(d_{\mathrm{L}_1} * c_1)[k]$ for $k \leq 0$ is to set $c_1[k] = 0$ for all $k \leq (-N_{0,1})$. By construction, this sequence c_1 yields a regularization cost no greater than that of c_1^*. Indeed, we have that $(d_{\mathrm{L}_1} * c_1^*)[k] = (d_{\mathrm{L}_1} * c_1)[k]$ for $1 \leq k \leq M_1$ and that $(d_{\mathrm{L}_1} * c_1)[k] = 0$ otherwise. Since the two sequences c_1 and c_1^* yield the same measurements, we have that $\mathcal{J}_h(c_1, c_2^*) \leq \mathcal{J}_h(c_1^*, c_2^*)$. Yet, since (s_1^*, s_2^*) is a solution of (32) and the construction of c_1 is unique, we necessarily have that $c_1 = c_1^*$.

Finally, we construct a sequence c_2 in a similar fashion. The only difference is the absence of boundary conditions, which means that the coefficients $c_2[k]$ for $k < m_2$ are determined as a linear combination of the $(N_{0,2}-1)$ next coefficients to cancel out $(d_{\mathrm{L}_2} * c_2)[k]$ for all $k \leq 0$. The same argument as for c_1 can then be used to prove that $c_2 = c_2^*$. ∎

The proof of Lemma 3 details the construction of two injective linear maps $\tilde{\theta}_1 : \mathbb{R}^{N_1} \to \ell_{1,\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{Z})$ and $\tilde{\theta}_2 : \mathbb{R}^{N_2} \to \ell_{1,\mathrm{L}_2}(\mathbb{Z})$ such that, for any $\mathbf{c}_i \in \mathbb{R}^{N_i}$, $\tilde{\theta}_i(\mathbf{c}_i)|_{i_i} = \mathbf{c}_i$ and $\phi_{0,h}(\tilde{\theta}_1(\mathbf{c}_1)) = \mathbf{0}$. We then define the mapping $\tilde{\theta} : \mathbb{R}^{N_1} \times \mathbb{R}^{N_2} : (\mathbf{c}_1, \mathbf{c}_2) \mapsto (\tilde{\theta}_1(\mathbf{c}_1), \tilde{\theta}_2(\mathbf{c}_2))$. Let $(\mathbf{c}_1, \mathbf{c}_2) \in \mathbb{R}^{N_1 \times N_2}$ and consider $\tilde{\theta}(\mathbf{c}_1, \mathbf{c}_2) \in \ell_{1,\mathrm{L}_1}(\mathbb{Z}) \times \ell_{1,\mathrm{L}_2}(\mathbb{Z})$. Following the proof of Lemma 3, $\mathcal{J}_h(\tilde{\theta}(\mathbf{c}_1, \mathbf{c}_2))$ can be computed using only the $N_1 + N_2$ coefficients $\mathbf{c}_1$ and $\mathbf{c}_2$. Indeed, all other coefficients $(\tilde{\theta}_i(\mathbf{c}_i)[k])_{k \notin I_i}$ do not affect the data fidelity term and cancel out all the regularization terms which they affect. This implies that $\mathcal{J}_h(\tilde{\theta}(\mathbf{c}_1, \mathbf{c}_2)) = J_h(\mathbf{c}_1, \mathbf{c}_2) = \|\mathbf{H}\binom{\mathbf{c}_1}{\mathbf{c}_2} - \mathbf{y}\|_2^2 + \lambda\|\mathbf{L}\binom{\mathbf{c}_1}{\mathbf{c}_2}\|_1$, where $\mathbf{H}$ and $\mathbf{L}$ are defined as in (30) and (31), respectively. Since, by Lemma 3, $\mathcal{S}_h \subset \tilde{\theta}(\mathbb{R}^{N_1} \times \mathbb{R}^{N_2})$, Problems (72) and (24) are equivalent in the sense that $\tilde{\theta}(S_h) = \mathcal{S}_h$, and the restriction $\theta = \tilde{\theta}|_{S_h} : S_h \to \mathcal{S}_h$ is a bijective linear map.

Concerning the first item of Theorem 2, let $\binom{\mathbf{c}_1}{\mathbf{c}_2} \in \ker \mathbf{H} \cap \ker \mathbf{L}$. Since $\mathbf{L}_1$ is invertible (27), we have that $\mathbf{c}_1 = \mathbf{0}$. Hence, the continuous-domain signal $s = \sum_{k \in \mathbb{Z}} \tilde{\theta}_2(\mathbf{c}_2)[k]\beta_{\mathrm{L}_2}(\cdot - kh)$ verifies $s \in \mathcal{N}_{\boldsymbol{\nu}} \cap \mathcal{N}_{\mathrm{L}_2} = \{0\}$ (well-posedness assumption in Theorem 1). We thus have that $\tilde{\theta}_2(\mathbf{c}_2) = 0$, which yields $\mathbf{c}_2 = 0$. Therefore, $\ker \mathbf{H} \cap \ker \mathbf{L} = \{\mathbf{0}\}$, which implies that Problem (24) is well-posed and that its solution set S_h is a nonempty compact set. It is also convex due to the convexity of the cost function J_h.

D. Proof of Theorem 3

Let $s = s_1 + s_2 \in \mathcal{S}$, where s_1 and s_2 are of the form (10). Assume by contradiction that there exists a knot $x_{1,k_0} \notin I_T$. We can then construct a spline $\tilde{s}_1$ which has the same expression as s_1 in I_T and the same knot locations, except for the knot at x_{1,k_0} which is removed. Since the ν_m functionals are supported in I_T, we have that $\boldsymbol{\nu}(s_1) = \boldsymbol{\nu}(\tilde{s}_1)$ and that $\|\mathrm{L}_1\{s_1\}\|_{\mathcal{M}} = \|\mathrm{L}_1\{\tilde{s}_1\}\|_{\mathcal{M}} + |a_{1,k_0}| > \|\mathrm{L}_1\{\tilde{s}_1\}\|_{\mathcal{M}}$. This implies that $\mathcal{J}(s) > \mathcal{J}(\tilde{s}_1 + s_2)$, which contradicts $s \in \mathcal{S}$. Therefore, we have that $x_{1,k} \geq 0$ for all k; we can thus construct the spline

$$s_{1,h}(x) = p_1(x) + \sum_{k=1}^{K_1} a_{1,k}\rho_{\mathrm{L}_1}(x - x_{1,k}^h), \tag{73}$$

where $x_{1,k}^h \in h\mathbb{Z}$ and such that $x_{1,k}^h \geq 0$ converges to $x_{1,k}$ when h goes to zero — we choose $x_{1,k}^h = 0$ if $x_{1,k} = 0$. Since this spline has the same expression as s_1 in the neighborhood of zero, we have that $\boldsymbol{\phi}_0(s_1) = \mathbf{0} \Rightarrow \boldsymbol{\phi}_0(s_{1,h}) = \mathbf{0}$ and, thus, that $s_{1,h} \in \mathcal{M}_{\mathrm{L}_1,h,\boldsymbol{\phi}_0}(\mathbb{R})$. Similarly, we construct $s_{2,h} \in \mathcal{M}_{\mathrm{L}_2,h}(\mathbb{R})$. The desired result then follows from the proof Theorem 3 in [27].

REFERENCES

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[2] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[3] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathmatiques*, vol. 346, no. 9/10, pp. 589–592, 2008.

[4] D. L. Donoho and P. B. Stark, "Uncertainty principles and signal recovery," *SIAM J. Appl. Math.*, vol. 49, no. 3, pp. 906–931, 1989.

[5] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.

[6] M. Elad and A. M. Bruckstein, "A generalized uncertainty principle and sparse representation in pairs of bases," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2558–2567, Sep. 2002.

[7] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3320–3325, Dec. 2003.

[8] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," *Proc. Nat. Acad. Sci.*, vol. 100, no. 5, pp. 2197–2202, 2003.

[9] J.-J. Fuchs, "On sparse representations in arbitrary redundant bases," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1341–1344, Jun. 2004.

[10] M. Elad, M. A. Figueiredo, and Y. Ma, "On the role of sparse and redundant representations in image processing," *Proc. IEEE*, vol. 98, no. 6, pp. 972–982, 2010.

[11] D. L. Donoho and G. Kutyniok, "Microlocal analysis of the geometric separation problem," *Commun. Pure Appl. Math.*, vol. 66, no. 1, pp. 1–47, 2013.

[12] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 2210–2219, May 2008.

[13] J.-L. Starck, D. L. Donoho, and M. Elad, "Redundant multiscale transforms and their application for morphological component analysis," in *Advances in Imaging and Electron Physics*, vol. 132. Amsterdam, The Netherlands: Elsevier, 2004.

[14] J.-L. Starck, M. Elad, and D. L. Donoho, "Image decomposition via the combination of sparse representations and a variational approach," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1570–1582, Oct. 2005.

[15] J.-L. Starck, Y. Moudden, J. Bobin, M. Elad, and D. L. Donoho, "Morphological component analysis," in *Proc. SPIE*, vol. 5914, 2005, Art. no. 59140Q.

[16] M. Elad, J.-L. Starck, P. Querre, and D. L. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *Appl. Comput. Harmon. Anal.*, vol. 19, no. 3, pp. 340–358, 2005.

[17] J. Bobin, J.-L. Starck, J. M. Fadili, Y. Moudden, and D. L. Donoho, "Morphological component analysis: An adaptive thresholding strategy," *IEEE Trans. Image Process.*, vol. 16, no. 11, pp. 2675–2681, Nov. 2007.
[18] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM J. Imag. Sci.*, vol. 2, no. 2, pp. 323–343, 2009.
[19] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
[20] M. Lustig, D. L. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.
[21] M. Unser, J. Fageot, and J. Ward, "Splines are universal solutions of linear inverse problems with generalized TV regularization," *SIAM Rev.*, vol. 59, no. 4, pp. 769–793, Dec. 2017.
[22] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis *versus* synthesis in signal priors," *Inverse Problems*, vol. 23, no. 3, pp. 947–968, 2007.
[23] J.-F. Cai, S. Osher, and Z. Shen, "Split Bregman methods and frame based image restoration," *Multiscale Model. Simul.*, vol. 8, no. 2, pp. 337–369, 2009.
[24] R. Otazo, E. J. Candès, and D. K. Sodickson, "Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components," *Magn. Reson. Medicine*, vol. 73, no. 3, pp. 1125–1136, 2015.
[25] E. J. Candès, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Appl. Comput. Harmon. Anal.*, vol. 31, no. 1, pp. 59–73, 2011.
[26] J. Lin, S. Li, and Y. Shen, "Compressed data separation with redundant dictionaries," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4309–4315, Jul. 2013.
[27] T. Debarre, J. Fageot, H. Gupta, and M. Unser, "B-spline-based exact discretization of continuous-domain inverse problems with generalized TV regularization," *IEEE Trans. Inf. Theory*, vol. 65, no. 7, pp. 4457–4470, Mar. 5, 2019.
[28] I. Schoenberg, *Cardinal Spline Interpolation*. Philadelphia, PA, USA: SIAM, 1973.
[29] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Process. Mag.*, vol. 16, no. 6, pp. 22–38, Nov. 1999.
[30] L. Schumaker, *Spline Functions: Basic Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2007.
[31] K. Bredies and H. K. Pikkarainen, "Inverse problems in spaces of measures," *ESAIM: Control, Optim. Calculus Variations*, vol. 19, no. 1, pp. 190–218, 2013.
[32] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Commun. Pure Appl. Math.*, vol. 67, no. 6, pp. 906–956, 2014.
[33] Q. Denoyelle, V. Duval, and G. Peyré, "Support recovery for sparse super-resolution of positive measures," *J. Fourier Anal. Appl.*, vol. 23, no. 5, pp. 1153–1194, 2017.
[34] V. Duval and G. Peyré, "Exact support recovery for sparse spikes deconvolution," *Founda. Comput. Math.*, vol. 15, no. 5, pp. 1315–1355, 2015.
[35] N. Boyd, T. Hastie, S. Boyd, B. Recht, and M. I. Jordan, "Saturating splines and feature selection," *J. Mach. Learn. Res.*, vol. 18, no. 1, pp. 7172–7203, 2017.
[36] C. Dossal, V. Duval, and C. Poon, "Sampling the Fourier transform along radial lines," *SIAM J. Numer. Anal.*, vol. 55, no. 6, pp. 2540–2564, 2017.
[37] C. Aubel, D. Stotz, and H. Bölcskei, "A theory of super-resolution from short-time Fourier transform measurements," *J. Fourier Anal. Appl.*, vol. 24, no. 1, pp. 45–107, 2018.
[38] H. Gupta, J. Fageot, and M. Unser, "Continuous-domain solutions of linear inverse problems with Tikhonov *versus* generalized TV regularization," *IEEE Trans. Signal Process.*, vol. 66, no. 17, pp. 4670–4684, Sep. 2018.
[39] M. Unser and T. Blu, "Cardinal exponential splines: Part I—Theory and filtering algorithms," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1425–1438, Apr. 2005.
[40] G. Dantzig, A. Orden, and P. Wolfe, "The generalized simplex method for minimizing a linear form under linear inequality restraints," *Pacific J. Math.*, vol. 5, no. 2, pp. 183–195, 1955.
[41] D. G. Luenberger, *Introduction to Linear and Nonlinear Programming*, vol. 28. Reading, MA, USA: Addison-Wesley, 1973.
[42] M. Unser, E. Soubies, F. Soulez, M. McCann, and L. Donati, "GlobalBioIm: A unifying computational framework for solving inverse problems," in *Proc. OSA Imag. Appl. Opt. Congr. Comput. Opt. Sens. Imag.*, San Francisco CA, USA, Jun. 2017, Paper no. CTu1B.
[43] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
[44] S. Aziznejad and M. Unser, "An L1 representer theorem for multiple-kernel regression," 2019, *arXiv:1811.00836*.

Thomas Debarre graduated from Mines ParisTech, Paris, France in 2016 and received the M.Sc. degree in applied mathematics from the École Normale Supérieure, Cachan, France, in 2017. He is currently working toward the Ph.D. degree with the Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, under the direction of M. Unser and M. Fageot. His research focuses on splines, inverse problems, and variational methods for reconstructing sparse signals.

Shayan Aziznejad received the B.Sc. degrees in electrical engineering and mathematics in 2017, from Sharif University of Technology, Tehran, Iran. He is currently working toward the Ph.D. degree with the Biomedical Imaging Group, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, under the direction of M. Unser. His research focuses on variational formulation of inverse problems and machine learning, spline theory, and stochastic processes. He is the recipient of the Best Student Paper Award at ICASSP 2019.

Michael Unser (M'89–SM'94–F'99) is currently a Professor and Director of Biomedical Imaging Group, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Lausanne, Switzerland. His primary area of investigation is biomedical image processing. He is internationally recognized for his research contributions to sampling theory, wavelets, the use of splines for image processing, stochastic processes, and computational bioimaging. He has authored or co-authored more that 250 journal papers on those topics. He is the co-author with P. Tafti of the book "*An Introduction to Sparse Stochastic Processes* (Cambridge University Press, 2014).

From 1985 to 1997, he was with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, USA, conducting research on bioimaging.

Dr. Unser has held the position of Associate Editor-in-Chief (2003–2005) for the IEEE TRANSACTIONS ON MEDICAL IMAGING. He is currently a member of the editorial boards of *SIAM J. Imaging Sciences and Foundations* and *Trends in Signal Processing*. He is the Founding chair of the Technical Committee on Bio Imaging and Signal Processing of the IEEE Signal Processing Society. He is an EURASIP Fellow (2009) and a member of the Swiss Academy of Engineering Sciences. He is the recipient of several international prizes including three IEEE-SPS Best Paper Awards and two Technical Achievement Awards from the IEEE (2008 SPS and EMBS 2010).