IEEE Open Journal of Signal Processing

Received 22 March 2021; revised 19 September 2021; accepted 20 September 2021. Date of publication 29 September 2021; date of current version 9 November 2021. The review of this article was arranged by Associate Editor Mr. Vincenzo Matta.

Digital Object Identifier 10.1109/OJSP.2021.3116482

Continuous-Domain Formulation of Inverse Problems for Composite Sparse-Plus-Smooth Signals

THOMAS DEBARRE, SHAYAN AZIZNEJAD, AND MICHAEL UNSER (Fellow, IEEE)

Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

CORRESPONDING AUTHOR: THOMAS DEBARRE (e-mail: thomas.debarre@gmail.com)

This work was supported in part by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under Grant 692726-GlobalBioIm and in part by Swiss National Science Foundation under Grant 200020_184646/1.

ABSTRACT We present a novel framework for the reconstruction of 1D composite signals assumed to be a mixture of two additive components, one sparse and the other smooth, given a finite number of linear measurements. We formulate the reconstruction problem as a continuous-domain regularized inverse problem with multiple penalties. We prove that these penalties induce reconstructed signals that indeed take the desired form of the sum of a sparse and a smooth component. We then discretize this problem using Riesz bases, which yields a discrete problem that can be solved by standard algorithms. Our discretization is exact in the sense that we are solving the continuous-domain problem over the search space specified by our bases without any discretization error. We propose a complete algorithmic pipeline and demonstrate its feasibility on simulated data.

INDEX TERMS Composite signals, total variation, Tikhonov regularization, B-splines.

I. INTRODUCTION

In the traditional discrete formalism of linear inverse problems, the goal is to recover a signal $\mathbf{c}_0 \in \mathbb{R}^N$ based on some measurement vector $\mathbf{y} \in \mathbb{R}^M$. These measurements are typically acquired via a linear operator $\mathbf{H} \in \mathbb{R}^{M\times N}$ that models the physics of our acquisition system (forward model), so that $\mathbf{Hc}_0 \approx \mathbf{y}$. The recovery is often achieved by solving an optimization problem that aims at minimizing the discrepancy between the measurements $\mathbf{Hc}$ of the reconstructed signal $\mathbf{c}$ and the acquired data $\mathbf{y}$. This data fidelity is measured with a suitable convex loss function $E : \mathbb{R}^M \times \mathbb{R}^M \to \mathbb{R}$, the prototypical example being the quadratic error $E(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|_2^2$. A regularization term is often added to the cost functional, which yields the optimization problem

$$\arg\min_{\mathbf{c}\in\mathbb{R}^N} \left\{ \underbrace{E(\mathbf{Hc}, \mathbf{y})}_{\text{Data fidelity}} + \underbrace{\lambda\mathcal{R}(\mathbf{Lc})}_{\text{Regularization}} \right\}, \quad (1)$$

where $\mathcal{R}$ is the regularization functional, $\mathbf{L}$ specifies a suitable transform domain, and $\lambda > 0$ is a tuning parameter that determines the strength of the regularization. The use of regularization can have multiple motivations:

1) to handle the ill-posedness of the inverse problem, which occurs when different signals yield identical measurements;
2) to favor certain types of reconstructed signal (*e.g.*, sparse or smooth) based on our prior knowledge;
3) to improve the conditioning of the inverse problem and thus increase its numerical stability and robustness to noise.

Historically, the first instance of regularization dates back to Tikhonov [1] with a quadratic regularization functional $\mathcal{R} = \|\cdot\|_2^2$. Tikhonov regularization constrains the energy of $\mathbf{Lc}$ which, when $\mathbf{L}$ is a finite-difference matrix, leads to a smooth signal $\mathbf{c}$. Tikhonov regularization has the practical advantage of being mathematically tractable which leads to a closed-form solution. More recently, there has been growing interest in ℓ_1 regularization $\mathcal{R} = \|\cdot\|_1$, which has peaked in popularity for compressed sensing (CS) [2]–[5]. With ℓ_1 regularization, the prior assumption is that the transform signal $\mathbf{Lc}_0$ is sparse, meaning that it has few nonzero coefficients:

indeed, the ℓ_1 norm can be seen as a convex relaxation of the ℓ_0 "norm," which counts the number of nonzero entries of a vector. The sparsity-promoting effect of ℓ_1 regularization is well understood and documented [6]–[8]. It is now generally considered to be superior to Tikhonov regularization for most applications [9]. Moreover, despite its non-differentiability, numerous efficient proximal algorithms based on the proximity operator of the ℓ_1 norm have emerged to solve ℓ_1-regularized problems [10]–[13].

A. DISCRETE INVERSE PROBLEMS FOR COMPOSITE SIGNALS

Despite their success, ℓ_1 and ℓ_2 regularization methods are too simple to model many real-world signals. In this paper, we investigate composite models of the form $s = s_1 + s_2$ where the two components have different characteristics. More precisely, s_1 is assumed to be sparse in some given domain and is treated with ℓ_1 regularization, while s_2 is assumed to be smooth and is treated with ℓ_2 regularization. In discrete settings, a natural way of reconstructing such signals is to solve the optimization problem

$$\min_{\mathbf{c}_1,\mathbf{c}_2\in\mathbb{R}^N} \left\{E(\mathbf{H}(\mathbf{c}_1+\mathbf{c}_2),\mathbf{y}) + \lambda_1\|\mathbf{L}_1\mathbf{c}_1\|_1 + \lambda_2\|\mathbf{L}_2\mathbf{c}_2\|_2^2\right\}, \tag{2}$$

where $\mathbf{c}_1, \mathbf{c}_2$ are the two components of the signal $\mathbf{c} = \mathbf{c}_1 + \mathbf{c}_2$, $\lambda_1, \lambda_2 > 0$ are tuning parameters, $\mathbf{L}_1 \in \mathbb{R}^{N\times N}$ is a sparsifying transform for $\mathbf{c}_1$, and $\mathbf{L}_2 \in \mathbb{R}^{N\times N}$ is a low-energy-promoting transform for $\mathbf{c}_2$. Amongst others, this modeling is considered in [14]–[18].

B. CONTINUOUS-DOMAIN FORMULATION

Until now, we have focused on the discrete setting, as it constitutes the vast majority of the inverse-problem literature for the obvious reason of computational feasibility. However, most real-world signals are inherently continuous. Therefore, when feasible, to formulate the inverse problem in the continuous domain is a natural and desirable objective.

In this work, we adapt the discrete approach of (2) to 1D continuous-domain composite signals by solving an optimization problem of the form

$$\min_{s_1,s_2} \left\{E(\nu(s_1+s_2),\mathbf{y}) + \lambda_1\|\mathrm{L}_1\{s_1\}\|_{\mathcal{M}} + \lambda_2\|\mathrm{L}_2\{s_2\}\|_{L_2}^2\right\}, \tag{3}$$

where s_1, s_2 are the two components of the signal $s = s_1 + s_2 : \mathbb{R} \to \mathbb{R}$, $\nu = (\nu_1, \ldots, \nu_M) : s \mapsto \nu(s) \in \mathbb{R}^M$ is the continuous-domain linear forward model, and $\mathrm{L}_1, \mathrm{L}_2$ are suitable continuously defined regularization operators. Typical choices are $\mathrm{L}_i = \mathrm{D}^{N_{0,i}}$ for $i \in \{1, 2\}$, where D is the derivative operator and $N_{0,i}$ the order of the derivative. The regularization norm $\|\cdot\|_{\mathcal{M}}$ is the total-variation (TV) norm for measures, which is the continuous counterpart of the discrete ℓ_1 norm [19]–[21]. We refer to this term as the generalized TV (gTV) regularizer due to the presence of the operator L_1. Finally, $\|\cdot\|_{L_2}$ is the usual norm over the space $L_2(\mathbb{R})$ of signals with finite energy; we refer to the corresponding term as the *generalized Tikhonov* (gTikhonov) regularizer, which promotes smoothness in combination with the operator L_2.

C. REPRESENTER THEOREMS AND DISCRETIZATION

A classical way of discretizing a continuous-domain problem is to reformulate it as a finite-dimensional one by relying on a *representer theorem* that gives a parametric form of the solution. Prominent examples include representer theorems for problems formulated over reproducing-kernel Hilbert spaces (RKHS), which are foundational to the field of machine learning [22], [23]. As demonstrated in [24, Theorem 3], the minimization problem (3) over the component s_2 (with a fixed s_1) — *i.e.*, gTikhonov regularization — falls into this category: the representer theorem states that there is a unique solution of the form

$$s_2^*(x) = p_2(x) + \sum_{m=1}^{M} a_{m,2}h_m(x), \tag{4}$$

where the additional component p_2 lies in the null space of L_2 (*i.e.*, $\mathrm{L}_2\{p_2\} = 0$), h_m is a (typically quite smooth) kernel function that is fully determined by the choice of ν_m and L_2, and $a_{m,2} \in \mathbb{R}$ are expansion coefficients. Therefore, to solve the continuous-domain problem, one need only optimize over the $a_{m,2}$ coefficients and the null-space component p_2 which lives in a finite-dimensional space. This leads to a standard finite-dimensional problem with Tikhonov regularization.

Concerning the minimization over the component s_1 (gTV regularization), several representer theorems give a parametric form of a sparse solution in different settings [20], [25]–[27], the most general one that includes ours being [28]. The more specific case of our setting is tackled by [24, Theorem 4], which states that there is an L_1-spline solution of the form

$$s_1^*(x) = p_1(x) + \sum_{k=1}^{K} a_{k,1}\rho_{\mathrm{L}_1}(x - x_k), \tag{5}$$

where $a_{k,1}, x_k \in \mathbb{R}$, ρ_{L_1} is a Green's function of L_1 (*i.e.*, $\mathrm{L}_1\{\rho_{\mathrm{L}_1}\} = \delta$, where δ is the Dirac impulse), K is the number of atoms of s_1 which is bounded by $K \leq (M - N_{0,1})$, $N_{0,1}$ being the dimension of the null space of L_1, and p_1 lies in the null space of L_1. For example, when $\mathrm{L}_1 = \mathrm{D}^{N_{0,1}}$, s_1 is a piecewise polynomial of degree $(N_{0,1} - 1)$ with smooth junctions at the knots x_k. These representer theorems have paved the way for various exact discretization methods. In the gTikhonov case, one can optimize over the $a_{m,2}$ coefficients in (4) directly [24]. For the gTV case (5), grid-based techniques using a well-conditioned B-spline basis [29] as well as grid-free techniques [30] have been proposed.

D. OUR CONTRIBUTION

In this work, we show that the representer theorems presented in the previous section can be combined into a composite one when dealing with Problem (3). More specifically, we prove that there exists a solution to (3) of the form $s_1^* = s_1^* + s_2^*$ such that s_1^* is of the form (5) and s_2^* is of the form (4): a "sparse plus smooth" solution. Building on this representation, we propose an exact discretization scheme. Both components s_i for $i \in \{1, 2\}$ are expressed in a suitable Riesz basis as $s_i = \sum_k c_i[k]\varphi_{i,k}$, where $c_i[k]$ are the coefficients to be optimized. This leads to an infinite-dimensional optimization problem reminiscent of the infinite-dimensional compressed sensing framework of Adcock and Hansen [31] and the wavelet-based model of Daubechies *et al.* [32]. However, these frameworks differ from our original Problem (3) in that their native spaces admit a countable basis which leads to a more straighforward discretization process.

To solve this infinite-dimensional problem numerically, we cast it as a finite-dimensional problem under some mild assumptions. This requires a careful handling of the boundaries of our interval of interest. In our implementation, we choose basis functions $\varphi_{1,k} = \beta_{\mathrm{L}_1}(\cdot - k)$ and $\varphi_{2,k} = \beta_{\mathrm{L}_2^*\mathrm{L}_2}(\cdot - k)$, where β_{L} is the B-spline for the operator L. B-splines are popular choices of basis functions [33]–[35], in large part due to their minimal-support property. Indeed, β_{L_i} has finite support when $\mathrm{L}_i = \mathrm{D}^{N_0}$, and it is the shortest-support generating function of the space of uniform L_i splines [36]. We show that optimizing over the spline coefficients leads to a discrete problem similar to (2) of the form

$$\min_{(\mathbf{c}_1,\mathbf{c}_2)\in\mathbb{R}^{N_1}\times\mathbb{R}^{N_2}} \left\{E(\mathbf{H}_1\mathbf{c}_1 + \mathbf{H}_2\mathbf{c}_2, \mathbf{y}) + \lambda_1\|\mathbf{L}_1\mathbf{c}_1\|_1 + \lambda_2\|\mathbf{L}_2\mathbf{c}_2\|_2^2\right\}, \tag{6}$$

where $\mathbf{H}_i \in \mathbb{R}^{M\times N_i}$ and $\mathbf{L}_i \in \mathbb{R}^{P_i\times N_i}$ for $i \in \{1, 2\}$. This discretization is exact in the sense that it is equivalent to the continuous problem (3) when each component s_i lies in the space generated by the basis functions $\{\varphi_{i,k}\}_{k\in\mathbb{Z}}$. This is a consequence of our informed choice of these basis functions $\varphi_{i,k}$. Moreover, the short support of the B-splines leads to well-conditioned $\mathbf{H}_i$ matrices and, thus, to a computationally feasible problem.

E. RELATED WORKS

The use of multiple regularization penalties is quite common in the literature. However, in most cases, each penalty is applied to the full signal instead of a component-wise [37]–[42]. A prominent example of such an approach is the elastic net [43], which is widely used in statistics. The spirit of these approaches is however quite different from ours: the reconstructed signal is encouraged to satisfy different priors *simultaneously*. Conversely, in (2), each component satisfies different priors independently from the others, which will give very different results.

Closer to our framework is the popular low-rank plus sparse matrix decomposition model [44], which imposes some form of sparsity on both components.

The model of Meyer [45] and its generalization by Vese and Osher [46], [47] follow the same idea as Problem (3), with the important difference that they use calculus-of-variation techniques to solve it. There is a connection as well with the Mumford-Shah functional [48], which is commonly used to segment an image in piecewise-smooth regions. The main difference lies in the fact that the optimization is not performed over the different components of the signal, but over the region boundaries. Another difference is that these models assume that one has full access to the noisy signal over a continuum, whereas (3) assumes that we only have access to some discrete measurements specified by the forward model ν.

F. OUTLINE

In Section II, we give the necessary mathematical preliminaries to formulate our optimization problem. In Section III, we formulate the continuous-domain problem and present our representer theorem, which is our main theoretical result. In Section IV, we explain our discretization strategy, which relies on the selection of suitable Riesz bases. Finally, in Section VI, we present experiments on simulated data.

II. PRELIMINARIES

In this section, we provide the mathematical background needed for this manuscript, which includes tools from functional analysis, particularly distribution theory. We refer to the textbook [49] for more in-depth background on this topic.

A. OPERATORS AND SPLINES

The crucial elements of our formulation are the regularization operators L_1 and L_2. In this section, we specify which type of operators are suitable in our framework.

Let $\mathcal{S}'(\mathbb{R})$ denote the space of tempered distributions, defined as the dual of the Schwartz space $\mathcal{S}(\mathbb{R})$ of infinitely smooth functions on $\mathbb{R}$ whose successive derivatives are rapidly decaying. Let $\mathcal{F}$ be the generalized Fourier transform with $\widehat{f} \triangleq \mathcal{F}\{f\}$. Then, it is a standard result in distribution theory that the frequency response $\widehat{L} = \mathcal{F}\{\mathrm{L}\{\delta\}\}$ of an ordinary differential operator $\mathrm{L} : \mathcal{S}'(\mathbb{R}) \to \mathcal{S}'(\mathbb{R})$ is a slowly increasing smooth function $\widehat{L} : \mathbb{R} \to \mathbb{C}$ [49, Chapter 7, §5]. Moreover, for any $f \in \mathcal{S}'(\mathbb{R})$, we have that $\mathcal{F}\{\mathrm{L}\{f\}\} = \widehat{L}\widehat{f} \in \mathcal{S}'(\mathbb{R})$. Next, we require L to be spline-admissible in the sense of Definition 1.

Definition 1 (Spline-admissible operator): A continuous linear shift-invariant operator $\mathrm{L} : \mathcal{S}'(\mathbb{R}) \to \mathcal{S}'(\mathbb{R})$ is *spline-admissible* if it verifies the following properties:

- there exists a function of slow growth $\rho_{\mathrm{L}} : \mathbb{R} \to \mathbb{R}$ (the Green's function of L) that satisfies $\mathrm{L}\{\rho_{\mathrm{L}}\} = \delta$;
- its null space $\mathcal{N}_{\mathrm{L}} = \{f \in \mathcal{S}'(\mathbb{R}) : \mathrm{L}\{f\} = 0\}$ has finite dimension N_0.

The prototypical example of a spline-admissible operator is the multiple-order derivative $\mathrm{L} = \mathrm{D}^{N_0}$ for $N_0 \geq 1$. Its causal Green's function is the one-sided power function $\rho_{\mathrm{L}} = \frac{x_+^{N_0-1}}{(N_0-1)!}$, where $x_+ = \max(0, x)$. The null space of L is the space of polynomials of degree less than N_0.

A spline-admissible operator L specifies the family of L-splines provided in Definition 2.

Definition 2 (Nonuniform L*-spline):* Let L be a spline-admissible operator in the sense of Definition 1. A nonuniform L-spline with K knots $x_1 < \cdots < x_K$ is a function $s : \mathbb{R} \mapsto \mathbb{R}$

that verifies

$$\mathrm{L}\{s\}(x) = \sum_{k=1}^{K} a_k \delta(x - x_k), \tag{7}$$

where $a_k \in \mathbb{R}$ is the amplitude of the kth singularity. The weighted sum of Dirac impulses in (7) is known as the *innovation* of s. The spline s can equivalently be written as

$$s(x) = p(x) + \sum_{k=1}^{K} a_k \rho_{\mathrm{L}}(x - x_k), \tag{8}$$

where $p \in \mathcal{N}_{\mathrm{L}}$.

For example, the operator $\mathrm{L} = \mathrm{D}^{N_0}$ leads to the well-known polynomial splines, which are piecewise polyomials of degree $(N_0 - 1)$ and of differentiability class $\mathcal{C}^{N_0-2}$.

B. NATIVE SPACES

1) SPARSE COMPONENT

The other crucial elements of our framework are the native spaces for each component. Let $\mathcal{M}(\mathbb{R})$ be the space of bounded Radon measures, which is known by the Riesz-Markov theorem [50, Chapter 6] to be the continuous dual of $C_0(\mathbb{R})$. The latter is the space of continuous functions vanishing at infinity, which is a Banach space when equipped with the supremum norm $\|\cdot\|_\infty$. The sparsity-promoting regularization norm $\|\cdot\|_{\mathcal{M}}$ is defined for a tempered distribution $w \in \mathcal{S}'(\mathbb{R})$ as

$$\|w\|_{\mathcal{M}} \triangleq \sup_{\varphi \in \mathcal{S}(\mathbb{R}), \|\varphi\|_\infty = 1} \langle w, \varphi \rangle. \tag{9}$$

Practically, the two critical features of the $\|\cdot\|_{\mathcal{M}}$ norm are the following:

1) it generalizes the L_1 norm in the sense that $\|w\|_{\mathcal{M}} = \|w\|_{L_1}$ for any $w \in L_1(\mathbb{R})$;
2) the $\|\cdot\|_{\mathcal{M}}$ norm of a weighted sum of Dirac impulses is $\|\sum_k a_k \delta(\cdot - x_k)\|_{\mathcal{M}} = \sum_k |a_k|$.

Accordingly, the native space for s_1 in (3) is defined as

$$\mathcal{M}_{\mathrm{L}_1}(\mathbb{R}) = \{s \in \mathcal{S}'(\mathbb{R}) : \mathrm{L}_1\{s\} \in \mathcal{M}(\mathbb{R})\}, \tag{10}$$

which is a Banach space when equipped with the direct-sum topology. It is also the largest space for which the regularization is well-defined. We refer to [51] for technical details on the construction of $\mathcal{M}_{\mathrm{L}_1}(\mathbb{R})$.

2) SMOOTH COMPONENT

The regularization norm $\|\cdot\|_{L_2}$ for the smooth component s_2 in (3) is defined over the Hilbert space $L_2(\mathbb{R})$. The corresponding native space of the smooth component s_2 is the Hilbert space

$$\mathcal{H}_{\mathrm{L}_2}(\mathbb{R}) = \{f \in \mathcal{S}'(\mathbb{R}) : \mathrm{L}_2\{f\} \in L_2(\mathbb{R})\}. \tag{11}$$

III. CONTINUOUS-DOMAIN INVERSE PROBLEM

When the intersection $\mathcal{N}_0 = \mathcal{N}_{\mathrm{L}_1} \cap \mathcal{N}_{\mathrm{L}_2}$ of the null spaces of our two native spaces is non-trivial, we need to specify boundary conditions on one of the two spaces to ensure the well-posedness of our optimization problem. This is done by introducing a biorthogonal system $(\phi_0, \mathbf{p}_0)$ for $\mathcal{N}_0$ in the sense of [20, Definition 3]. An example of a valid choice is given in (56) in Appendix B. The search space with boundary conditions ϕ_0 is then given by

$$\mathcal{M}_{\mathrm{L}_1,\phi_0}(\mathbb{R}) = \{f \in \mathcal{M}_{\mathrm{L}_1}(\mathbb{R}) : \phi_0(f) = \mathbf{0}\}. \tag{12}$$

We now present in Theorem 1 our problem formulation to reconstruct sparse-plus-smooth composite signals. This representer Theorem gives a parametric form of a solution of our optimization problem; the proof is given in Appendix A.

Theorem 1 (Continuous-domain representer theorem): Let $\mathrm{E} : \mathbb{R}^M \times \mathbb{R}^M \to \mathbb{R}$ be a nonnegative, coercive, proper, convex, and lower-semicontinuous functional. Let $\mathrm{L}_1, \mathrm{L}_2$ be spline-admissible operators in the sense of Definition 1 and let $\nu = (\nu_1, \ldots, \nu_M)$ be a linear measurement operator composed of the M linear functionals $\nu_m : f \mapsto \nu_m(f) \in \mathbb{R}$ that are weak*-continuous over $\mathcal{M}_{\mathrm{L}_1}(\mathbb{R})$ and over $\mathcal{H}_{\mathrm{L}_2}(\mathbb{R})$. We assume that $\mathcal{N}_\nu \cap (\mathcal{N}_{\mathrm{L}_1} + \mathcal{N}_{\mathrm{L}_2}) = \{0\}$, where $\mathcal{N}_\nu$ is the null space of ν (well-posedness assumption). Then, for any $\lambda_1, \lambda_2 > 0$, the optimization problem

$$\mathcal{S} = \left\{ \arg\min_{\substack{s_1 \in \mathcal{M}_{\mathrm{L}_1,\phi_0}(\mathbb{R}) \\ s_2 \in \mathcal{H}_{\mathrm{L}_2}(\mathbb{R})}} \mathcal{J}(s_1, s_2) \right\} \quad \text{with}$$

$$\mathcal{J}(s_1, s_2) = \mathrm{E}(\nu(s_1 + s_2), \mathbf{y}) + \lambda_1 \|\mathrm{L}_1\{s_1\}\|_{\mathcal{M}} + \lambda_2 \|\mathrm{L}_2\{s_2\}\|_{L_2}^2 \tag{13}$$

has a solution $(s_1^*, s_2^*) \in \mathcal{S}$ with the following components:

- the component s_1^* is a nonuniform L_1-spline of the form
$$s_1^*(x) = p_1(x) + \sum_{k=1}^{K_1} a_{1,k} \rho_{\mathrm{L}_1}(x - x_k) \tag{14}$$
for some $K_1 \leq (M - N_{0,1})$, where $p_1 \in \mathcal{N}_{\mathrm{L}_1}$, and $a_{1,k}, x_k \in \mathbb{R}$;
- the component s_2^* is of the form
$$s_2^*(x) = p_2(x) + \sum_{m=1}^{M} a_{2,m} h_m(x), \tag{15}$$
where $h_m(x) = (\nu_m * \mathcal{F}^{-1}\{\frac{1}{|\hat{L}_2|^2}\})(x)$, $p_2 \in \mathcal{N}_{\mathrm{L}_2}$, $a_{2,k} \in \mathbb{R}$, and where $\sum_{m=1}^{M} a_{2,m} \langle q_2, \nu_m \rangle = 0$ for any $q_2 \in \mathcal{N}_{\mathrm{L}_2}$.

Moreover, for any pair of solutions $(s_1^*, s_2^*), (\tilde{s}_1^*, \tilde{s}_2^*) \in \mathcal{S}$, s_2^* and $\tilde{s}_2^*$ differ only up to an element of the null space $\mathcal{N}_{\mathrm{L}_2}$, so that $(s_2^* - \tilde{s}_2^*) \in \mathcal{N}_{\mathrm{L}_2}$.

A pleasing outcome of Theorem 1 is that it combines Theorems 3 and 4 of [24] into one. There is, however, an added technicality due to the boundary conditions ϕ_0. The latter are necessary to ensure the well-posedness of Problem (13). Otherwise, for any $(s_1^*, s_2^*) \in \mathcal{S}$ and $p \in \mathcal{N}_0$, we would have that $(s_1^* + p, s_2^* - p) \in \mathcal{S}$ which would imply that $\mathcal{S}$ is unbounded. Note, however, that these conditions do not restrict the search space, since $\mathcal{M}_{\mathrm{L}_1,\phi_0}(\mathbb{R}) + \mathcal{H}_{\mathrm{L}_2}(\mathbb{R}) = \mathcal{M}_{\mathrm{L}_1}(\mathbb{R}) + \mathcal{H}_{\mathrm{L}_2}(\mathbb{R})$.

IV. EXACT DISCRETIZATION

In order to discretize Problem (13), we restrict the search spaces $\mathcal{M}_{\mathrm{L}_1,\phi_0}(\mathbb{R})$ and $\mathcal{H}_{\mathrm{L}_2}(\mathbb{R})$. The standard approach to achieve this is to choose a sequence of appropriate basis functions $\{\varphi_{i,k}\}_{k\in\mathbb{Z}}$ that span the reconstruction spaces

$$V_i(\mathbb{R}) = \left\{ \sum_{k\in\mathbb{Z}} c_i[k]\varphi_{i,k} : c_i \in V_i(\mathbb{Z}) \right\} \tag{16}$$

for $i \in \{1, 2\}$ that are subject to the constraints $V_1(\mathbb{R}) \subset \mathcal{M}_{\mathrm{L}_1,\phi_0}(\mathbb{R})$ and $V_2(\mathbb{R}) \subset \mathcal{H}_{\mathrm{L}_2}(\mathbb{R})$. These continuous spaces are linked to discrete spaces $V_i(\mathbb{Z})$, the choices of which will be made explicit in (20) and (27). More precisely, there is a one-to-one mapping between them using the basis functions $\varphi_{i,k}$.

A. RIESZ BASES AND B-SPLINES

For numerical purposes, a desirable property is that our basis functions satisfy the Riesz property. Riesz bases are highly important concepts in that they generalize orthonormal bases, while leaving more flexibility for other desirable properties such as short support [52].

Definition 3 (Riesz basis): A sequence of functions $\{\varphi_k\}_{k\in\mathbb{Z}}$ with $\varphi_k \in L_2(\mathbb{R})$ is said to be a Riesz basis if there exist constants $0 < A \leq B$ such that, for any $c \in \ell_2(\mathbb{Z})$, we have that

$$A\|c\|_{\ell_2} \leq \left\| \sum_{k\in\mathbb{Z}} c[k]\varphi_k \right\|_{L_2} \leq B\|c\|_{\ell_2}. \tag{17}$$

Popular examples of Riesz bases are B-spline bases, which are introduced in Definition 4.

Definition 4 (B-spline): The B-spline for a spline-admissible operator L is characterized by a finite-difference-like filter $(d_{\mathrm{L}}[k])_{k\in\mathbb{Z}}$, and is defined as

$$\beta_{\mathrm{L}}(x) = \mathcal{F}^{-1}\left\{ \frac{\sum_{k\in\mathbb{Z}} d_{\mathrm{L}}[k]\mathrm{e}^{-\mathrm{j}k(\cdot)}}{\widehat{L}(\cdot)} \right\}(x). \tag{18}$$

The criteria for choosing a valid filter d_{L} for a general class of operators L are given in [53, Theorem 2.7].

The best-known example of a B-spline is the polynomial B-spline for the operator $\mathrm{L} = \mathrm{D}^{N_0}$, whose filter d_{L} is characterized by its z-transform $D_{\mathrm{L}}(z) = (1 - z^{-1})^{N_0}$. The corresponding B-spline β_{L} is supported in $[0, N_0]$.

A key feature of B-splines is that they are the L-splines with the shortest support or, when finite support is impossible, with the fastest decay. Moreover, by [53, Theorem 2.7], for a valid B-spline β_{L} as specified by Definition 4, the sequence of functions $\{\beta_{\mathrm{L}}(\cdot - k)\}_{k\in\mathbb{Z}}$ forms a Riesz basis in the sense of Definition 3.

It is clear from (18) that the innovation of the B-spline is a sum of Dirac impulses given by

$$\mathrm{L}\{\beta_{\mathrm{L}}\} = \sum_{k\in\mathbb{Z}} d_{\mathrm{L}}[k]\delta(\cdot - k). \tag{19}$$

B. CHOICE OF BASIS FUNCTIONS

We now present and discuss our choice for the basis functions $\varphi_{1,k}$ and $\varphi_{2,k}$.

1) SPARSE COMPONENT

For the sparse component, we choose basis functions $\varphi_{1,k} = \beta_{\mathrm{L}_1}(\cdot - k)$ (defined in (18)) for all $k \in \mathbb{Z}$. With this choice,

$$V_1(\mathbb{R}) = \left\{ f = \sum_{k\in\mathbb{Z}} c_1[k]\varphi_{1,k} : c_1 \in V_1(\mathbb{Z}) \right\} \subset \mathcal{M}_{\mathrm{L}_1,\phi_0}(\mathbb{R})$$

with the digital-filter space

$$V_1(\mathbb{Z}) = \left\{ (c_1[k])_{k\in\mathbb{Z}} : \ (d_{\mathrm{L}_1} * c_1) \in \ell_1(\mathbb{Z}) \text{ and } \sum_{k\in\mathbb{Z}} c_1[k]\phi_0(\varphi_{1,k}) = \mathbf{0} \right\}, \tag{20}$$

is the largest possible native reconstruction space [54, Equation (22)]. The choice of the basis functions $\varphi_{1,k}$ is guided by the following considerations:

- they generate the space of uniform L_1 splines. This conforms with Theorem 1, which states that the component s_1^* is an L_1-spline;
- they enable exact computations in the continuous domain. In particular, we have that $\|\mathrm{L}_1\{\sum_{k\in\mathbb{Z}} c_1[k]\varphi_{1,k}\}\|_{\mathcal{M}} = \|d_{\mathrm{L}_1} * c_1\|_{\ell_1}$;
- the Riesz-basis property of B-splines leads to a well-conditioned system matrix, which is paramount in numerical applications.

B-splines are the only functions that satisfy all these properties. Based on these criteria, B-splines are thus optimal.

2) SMOOTH COMPONENT

At first glance, the most natural choice for $\varphi_{2,k}$ is to select the basis functions suggested by (15) in Theorem 1: h_m for $1 \leq m \leq M$ and a basis of $\mathcal{N}_{\mathrm{L}_2}$, which yield a finite number $M + N_{0,2}$ of basis functions. However, this approach runs into the following hitches:

- the basis functions h_m are typically increasing at infinity, which contradicts the Riesz-basis requirement and leads to severely ill-conditioned optimization tasks [24];
- depending on the measurements operator ν, h_m may lack a closed-form expression.

We therefore focus on these criteria, in a spirit similar to [55]. The $\varphi_{2,k}$ are chosen to be regular shifts of a generating function φ_2, with $\varphi_{2,k} = \varphi_2(\cdot - k)$ such that $\{\mathrm{L}_2\{\varphi_{2,k}\}\}_{k\in\mathbb{Z}}$ forms a Riesz basis in the sense of Definition 3. Contrary to $\varphi_{1,k}$, these requirements allow for many choices of $\varphi_{2,k}$. In order to perform exact discretization, one then only needs to compute the following autocorrelation filter.

Proposition 1 (Autocorrelation filter for the smooth component): Let φ_2 be a generating function such that $\varphi_{2,k} = \varphi_2(\cdot - k)$ form a Riesz basis. Then, the following two items hold:

- the inner product $\langle \mathrm{L}_2\{\varphi_{2,k}\}, \mathrm{L}_2\{\varphi_{2,k'}\}\rangle_{L_2}$ only depends on the difference $(k-k')$. We can thus introduce the autocorrelation filter

$$\begin{aligned}\rho[k] &= \langle \mathrm{L}_2\{\varphi_{2,k}\}, \mathrm{L}_2\{\varphi_{2,0}\}\rangle_{L_2}\\ &= \langle \mathrm{L}_2\{\varphi_{2,k+k'}\}, \mathrm{L}_2\{\varphi_{2,k'}\}\rangle_{L_2}\end{aligned} \tag{21}$$

for any $k, k' \in \mathbb{Z}$;
- the filter ρ is positive semidefinite, with $\sum_{k,k'\in\mathbb{Z}} c[k]c[k']\rho[k-k'] \geq 0$ for any finitely supported real digital filter c.

Proof: The first item is proved with a simple change of variable in the integral that defines the inner product. The second item is derived by observing that, for any c_2, we have

$$\begin{aligned}\left\| \mathrm{L}_2\left\{\sum_{k\in\mathbb{Z}} c_2[k]\varphi_{2,k}\right\}\right\|_{L_2}^2 &= \sum_{k,k'\in\mathbb{Z}} \big(c_2[k]c_2[k'] \\ &\quad \langle \mathrm{L}_2\{\varphi_{2,k}\}, \mathrm{L}_2\{\varphi_{2,k'}\}\rangle\big)\\ &= \sum_{k,k'\in\mathbb{Z}} c_2[k]c_2[k']\rho[k-k'] \geq 0.\end{aligned} \tag{22}$$

■

For our implementation, we make a specific choice of basis function φ_2 among the many choices for which the autocorrelation filter (21) can be computed analytically. We choose the $\mathrm{L}_2^*\mathrm{L}_2$ B-spline basis $\varphi_2 = \beta_{\mathrm{L}_2^*\mathrm{L}_2}$ and $\varphi_{2,k} = \varphi_2(\cdot - k)$, where L_2^* denotes the adjoint operator of L_2. This choice has the following additional advantages:

- the generator φ_2 has a simple explicit expression that does not depend on the measurement operator ν;
- the autocorrelation filter ρ also has a simple expression, as will be shown in Proposition 4;
- in the special case of the sampling operator $\nu_m = \delta(\cdot - x_m)$, where the x_m are the sampling locations, this choice conforms with (15) in Theorem 1 since s_2^* is then an $\mathrm{L}_2^*\mathrm{L}_2$-spline. Note, however, that we do not exploit the knowledge that s_2^* has knots at the sampling locations x_m.

With this basis function $\varphi_2 = \beta_{\mathrm{L}_2^*\mathrm{L}_2}$, there is no straightforward choice of the digital-filter space $V_2(\mathbb{Z})$. Our practical choice is given in (27); it depends on our discretization method for reasons that are discussed in Remark 1.

C. FORMULATION OF THE DISCRETE PROBLEM

The autocorrelation filter introduced in Proposition 1 enables us to discretize Problem (13) in an exact way in the $V_i(\mathbb{R})$ spaces.

Proposition 2 (Riesz-basis Discretization): Let $\varphi_{i,k}$ be chosen as specified in Section IV-B for $i \in \{1,2\}$, $k \in \mathbb{Z}$, and

$$\mathcal{S}_\mathrm{d} = \left\{\arg\min_{(c_1,c_2)\in V_1(\mathbb{Z})\times V_2(\mathbb{Z})} \mathcal{J}_\mathrm{d}(c_1,c_2)\right\}. \tag{23}$$

The cost function is given by

$$\begin{aligned}\mathcal{J}_\mathrm{d}(c_1,c_2) &= E\left(\sum_{k_1\in\mathbb{Z}} c_1[k]\nu(\varphi_{1,k}) + \sum_{k\in\mathbb{Z}} c_2[k]\nu(\varphi_{2,k}), \mathbf{y}\right)\\ &\quad + \lambda_1 \|d_{\mathrm{L}_1} * c_1\|_{\ell_1} + \lambda_2\langle c_2, \rho * c_2\rangle_{\ell_2},\end{aligned} \tag{24}$$

where d_{L_1} is the finite-difference-like filter from Definition 4, ρ is defined in Proposition 1, and $\langle\cdot,\cdot\rangle_{\ell_2}$ is the inner product over $\ell_2(\mathbb{Z})$. Then, Problem (23) is equivalent to a restriction of the search spaces $\mathcal{M}_{\mathrm{L}_1,\phi_0}(\mathbb{R})$ and $\mathcal{H}_{\mathrm{L}_2}(\mathbb{R})$ to the spaces $V_1(\mathbb{R})$ and $V_2(\mathbb{R})$ defined in (16), respectively, so that

$$\mathcal{S}_\mathrm{res} = \left\{\arg\min_{(s_1,s_2)\in V_1(\mathbb{R})\times V_2(\mathbb{R})} \mathcal{J}(s_1,s_2)\right\}, \tag{25}$$

in the sense that there exists a bijective linear mapping $(c_1,c_2) \mapsto (\sum_{k\in\mathbb{Z}} c_1[k]\varphi_{1,k}, \sum_{k\in\mathbb{Z}} c_2[k]\varphi_{2,k})$ from $\mathcal{S}_\mathrm{d}$ to $\mathcal{S}_\mathrm{res}$.

Proof: By plugging the expansions $s_i = \sum_{k\in\mathbb{Z}} c_i[k]\varphi_{i,k}$ into the cost function $\mathcal{J}$, using the linearity of ν, we get the data-fidelity term of (24). Using (19), we readily deduce that $\|\mathrm{L}_1\{\sum_{k\in\mathbb{Z}} c_1[k]\varphi_{1,k}\}\|_\mathcal{M} = \|d_{\mathrm{L}_1} * c_1\|_{\ell_1}$ [29, Equation (25)]. As for the second regularization term, we observe that

$$\begin{aligned}\langle c_2, c_2 * \rho\rangle_{\ell_2} &= \sum_{k,k'\in\mathbb{Z}} c_2[k]c_2[k']\rho[k-k']\\ &= \left\|\mathrm{L}_2\left\{\sum_{k\in\mathbb{Z}} c_2[k]\varphi_{2,k}\right\}\right\|_{L_2}^2,\end{aligned} \tag{26}$$

using (22) for the last step. This proves the equivalence between Problems (25) and (23), up to the specified mapping which is indeed a bijective linear mapping due to the Riesz-basis properties of $\{\varphi_{1,k}\}_{k\in\mathbb{Z}}$ and $\{\varphi_{2,k}\}_{k\in\mathbb{Z}}$. ■

V. PRACTICAL IMPLEMENTATION

We now discuss how to solve our discretized problem (23) in practice, which involves recasting it as a finite-dimensional problem.

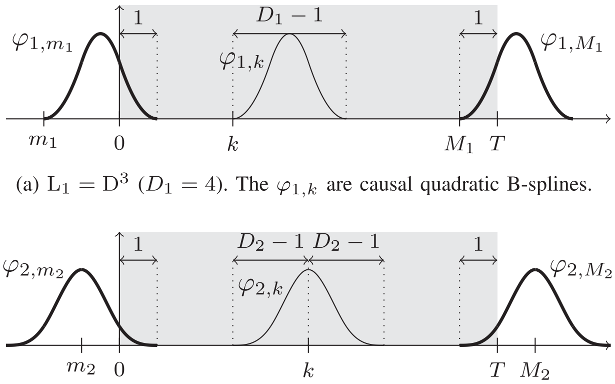
A. FINITE DOMAIN ASSUMPTIONS

To solve problem (23) numerically in an exact way, we must make assumptions that will enable us to restrict the problem to a finite interval of interest.

1) The operators L_i for $i \in \{1,2\}$ admit a B-spline with finite support, which implies that the filters d_{L_i} (introduced in Definition (18)) and ρ (Proposition 4) have finite support. Without loss of generality, the support of d_{L_i} is chosen to be $[0\ldots D_i-1]$ (of length $D_i > 0$), which leads to causal B-splines β_{L_i}.
2) The measurement functionals ν_m are supported in an interval $I_T = [0,T]$, where $T \in \mathbb{N}$.

The first item is fulfilled for common one-dimensional operators L_1 such as ordinary differential operators [56] or rational operators [57].

The second assumption is natural and is often fulfilled in practice, for instance in imaging with a finite field of view. The support length T then roughly corresponds to the number of



(a) $\mathrm{L}_1 = \mathrm{D}^3$ ($D_1 = 4$). The $\varphi_{1,k}$ are causal quadratic B-splines.

(b) $\mathrm{L}_2 = \mathrm{D}^2$ ($D_2 = 3$). The $\varphi_{2,k}$ are centered cubic B-splines.

FIGURE 1. Examples of boundary basis functions φ_{i,m_i} and φ_{i,M_i} for $i \in \{1, 2\}$.

grid points in the interval of interest. Note that, for simplicity, we only consider integer grids. However, the finesse of the grid can be tuned at will by adjusting T and rescaling the problem over the interval of interest.

B. FORMULATION OF THE FINITE-DIMENSIONAL PROBLEM

Our choice of basis functions together with the assumptions in Section V-A enable us to restrict Problem (23) to the interval of interest I_T. More precisely, we introduce the indices $m_i, M_i \in \mathbb{Z}$ for $i \in \{1, 2\}$; the range $[m_i \dots M_i]$ corresponds to the set of indices k for which $\mathrm{Supp}(\varphi_{i,k}) \cap I_T \neq \emptyset$, so that the basis function $\varphi_{i,k}$ affects the measurements. Hence, the number of active basis functions (*i.e.*, the number of spline coefficients to be optimized) is $N_i = (M_i - m_i - 1)$. It can easily be verified that we have $m_1 = (-D_1 + 2)$, $M_1 = (T - 1)$, $m_2 = (-D_2 + 2)$, and $M_2 = (T + D_2 - 2)$. See Fig. 1 for an illustrative example.

Finally, we introduce the native digital-filter space

$$V_2(\mathbb{Z}) = \left\{(c_2[k])_{k\in\mathbb{Z}} : \mathrm{Supp}(d_{\mathrm{L}_2} * c_2) \subset [1 \dots M_2]\right\}, \quad (27)$$

which is a valid choice because $V_2(\mathbb{R}) \subset \mathcal{H}_{\mathrm{L}_2}(\mathbb{R})$. Indeed, we can verify that, for any $c_2 \in V_2(\mathbb{Z})$, the function $s_2 = \sum_{k\in\mathbb{Z}} c_2[k]\varphi_{2,k}$ satisfies $\|\mathrm{L}_2\{s_2\}\|_{L_2}^2 = \|g * c_2\|_{\ell_2}^2 = \|b^{1/2} * (d_{\mathrm{L}_2} * c_2)\|_{\ell_2}^2 < +\infty$, which proves that $s_2 \in \mathcal{H}_{\mathrm{L}_2}(\mathbb{R})$. This is due to the finite support of both $(d_{\mathrm{L}_2} * c_2)$ and $b^{1/2}$, where the filter $b^{1/2}$ and the decomposition $g = b^{1/2} * d_{\mathrm{L}_2}$ are introduced in Proposition 4 in Appendix C.

Remark 1: Contrary to $V_1(\mathbb{Z})$ defined in (20), our choice of $V_2(\mathbb{Z})$ in (27) is not the largest valid space: there exist larger vector spaces such that $V_2(\mathbb{R}) \subset \mathcal{H}_{\mathrm{L}_2}(\mathbb{R})$. However, the support restriction implies that for any $s_2 = \sum_{k\in\mathbb{Z}} c_2[k]\beta_{\mathrm{L}_2^*\mathrm{L}_2}(\cdot - k) \in V_2(\mathbb{R})$, the function $\mathrm{L}_2\{s_2\} = \sum_{k\in\mathbb{Z}} (d_{\mathrm{L}_2} * c_2)[k]\beta_{\mathrm{L}_2}^{\vee}(\cdot - k)$ has a finite support. This is a desirable property both for simplicity of implementation and because it conforms with Theorem 1, since s_2^* in (15) also satisfies this property. Our specific choice of support for $(d_{\mathrm{L}_2} * c_2)$ is guided by boundary considerations and will be justified in the proof of Proposition 3.

The restriction to a finite number of active spline coefficients leads to finite-dimensional system and regularization matrices. The system matrices are of the form

$$\mathbf{H}_i = \left[\nu(\varphi_{i,m_i}) \quad \cdots \quad \nu(\varphi_{i,M_i})\right] \in \mathbb{R}^{M\times N_i}. \quad (28)$$

The regularization matrix for the sparse component, denoted by $\mathbf{L}_1 \in \mathbb{R}^{(N_1-D_1+1)\times N_1}$, is of the form

$$\mathbf{L}_1 = \begin{pmatrix} d_{\mathrm{L}_1}[D_1-1] & \cdots & d_{\mathrm{L}_1}[0] & 0 & \cdots & 0 \\ 0 & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ 0 & \cdots & 0 & d_{\mathrm{L}_1}[D_1-1] & \cdots & d_{\mathrm{L}_1}[0] \end{pmatrix}. \quad (29)$$

The second component requires a careful handling of the boundaries in order to achieve exact discretization. This leads to a more complicated expression for the associated regularization matrix, which is given in (61) in Appendix C.

Finally, we introduce the matrix $\mathbf{A} \in \mathbb{R}^{N_0\times N_1}$ associated to the boundary condition functionals ϕ_0. Our choice of boundary condition functionals ϕ_0 is presented in Appendix B. With this choice, the constraint $\phi_0(\sum_{k\in\mathbb{Z}} c_1[k]\varphi_{1,k}) = \mathbf{0}$ leads to N_0 linear constraints on the coefficients $\mathbf{c}_1 = (c_1[m_1], \dots, c_1[M_1])$, which can be written in matrix form as $\mathbf{A}\mathbf{c}_1 = \mathbf{0}$. In common cases, these constraints simply lead to the N_0 first coefficients of $\mathbf{c}_1$ to be set to zero, which thus reduces the dimension of the optimization problem.

These matrices enable an exact discretization of Problem (23), as shown in Proposition 3, the proof of which being given in Appendix D.

Proposition 3 (Recasting as a finite problem): Let $\varphi_{1,k} = \beta_{\mathrm{L}_1}(\cdot - k)$, $\varphi_{2,k} = \beta_{\mathrm{L}_2^*\mathrm{L}_2}(\cdot - k)$, and let the assumptions in Section V-A be satisfied. Then, Problem (23) is equivalent to the optimization problem

$$S = \left\{ \arg\min_{\substack{(\mathbf{c}_1,\mathbf{c}_2)\in\mathbb{R}^{N_1}\times\mathbb{R}^{N_2} \\ \mathbf{A}\mathbf{c}_1=\mathbf{0}}} J(\mathbf{c}_1, \mathbf{c}_2) \right\}, \quad (30)$$

where the cost function is given by

$$J(\mathbf{c}_1, \mathbf{c}_2) = E(\mathbf{H}_1\mathbf{c}_1 + \mathbf{H}_2\mathbf{c}_2, \mathbf{y}) \quad (31)$$

$$+ \lambda_1\|\mathbf{L}_1\mathbf{c}_1\|_1 + \lambda_2\|\mathbf{L}_2\mathbf{c}_2\|_2^2. \quad (32)$$

The matrices $\mathbf{H}_i$ and $\mathbf{L}_i$ for $i \in \{1, 2\}$ are defined in (28), (29), and (61). This equivalence holds in the sense that there exists a bijective linear mapping from $\mathcal{S}_{\mathrm{d}}$ to S

$$(c_1, c_2) \mapsto ((c_1[m_1], \dots, c_1[M_1]), (c_2[m_2], \dots, c_2[M_2])) \quad (33)$$

between their solution sets.

The combination of Propositions 2 and 3 allows us to solve the continuous-domain infinite-dimensional problem (25) by finding a solution $(\mathbf{c}_1^*, \mathbf{c}_2^*) \in S$ of the finite-dimensional problem (30). We obtain the corresponding solution of (25) by extending these vectors to digital filters $(c_1^*, c_2^*) \in \mathcal{S}_\mathrm{d}$ (this extension is unique as specified by Proposition 3), which yields the continuous-domain reconstruction $s^* = s_1^* + s_2^*$, where $s_i^* = \sum_{k\in\mathbb{Z}} c_d^*[k]\varphi_{i,k}$.

C. SPARSIFICATION STEP

Although problem (30) can be solved using standard solvers such as the alternating-direction method of multipliers (ADMM), there is no guarantee that such solvers will yield a solution of the desired form specified by Theorem 1, *i.e.*, s_1^* is an L_1-spline with fewer than $(M - N_{0,1})$ knots, and s_2^* is a sum of M kernel functions and a null space element. This is a particularly relevant observation for the first component since, at fixed second component s_2^*, only extreme-point solutions s_1^* of Problem (13) take the prescribed form [20]. This problem can be alleviated by computing a solution $(\mathbf{c}_1^*, \mathbf{c}_2^*)$ to Problem (30), and then finding an extreme point of the solution set $\mathbf{c}_1^{\mathrm{extr}} \in \arg\min_{\mathbf{c}_1\in\mathbb{R}^{N_1}} J(\mathbf{c}_1, \mathbf{c}_2^*)$, which leads to a solution $(\mathbf{c}_1^{\mathrm{extr}}, \mathbf{c}_2^*)$ of the prescribed form. This is achieved by recasting the problem as a linear program and using the simplex algorithm [58] to reach an extreme-point solution [24, Theorem 7].

VI. EXPERIMENTAL VALIDATION

We now validate our reconstruction algorithm in a simulated setting.

A. EXPERIMENTAL SETTING

1) GRID SIZE

We rescale the problem by a factor T so that the interval of interest I_T is mapped into [0, 1]. We tune the finesse of the grid (and the dimension of the optimization task) by varying T, which amounts to varying the grid size $h = 1/T$ in the rescaled problem.

2) GROUND TRUTH

We generate a ground-truth signal $s^{\mathrm{GT}} = s_1^{\mathrm{GT}} + s_2^{\mathrm{GT}}$. The sparse component s_1^{GT} is chosen to be an L_1-spline of the form (8) with few jumps, for which gTV is an adequate choice of regularization, as demonstrated by (14) in our representer theorem. For the smooth component s_2^{GT}, we generate a realization of a solution s_2 of the stochastic differential equation $\mathrm{L}_2 s_2 = w$, where w is a Gaussian white noise with standard deviation σ_2 by following the method of [59]. The operator L_2 then acts as a whitening operator for the stochastic process s_2. The reason for this choice is the connection between the minimum mean-square estimation of such stochastic processes and the solutions to variational problems with gTikhonov regularization $\|\mathrm{L}_2 s_2\|_{L_2}^2$ [22], [60], [61].

3) FORWARD OPERATOR

Our forward model is the Fourier-domain cosine sampling operator of the form $\nu_1(s) = \int_0^1 s(t)\mathrm{d}t$ (DC term) and

$$\nu_m(s) = \int_0^1 \cos(\omega_m t + \theta_m) s(t)\mathrm{d}t \tag{34}$$

for $2 \le m \le M$, where the sampling pulsations ω_m are chosen at random within the interval $(0, \omega_{\max}]$, and the phases θ_m are chosen at random within the interval $[0, 2\pi)$. Notice that ν_m is a Fourier-domain measurement of the restriction of s to the interval of interest [0,1], in conformity with the finite-domain assumption in Section V-A.

For the data-fidelity term, we use the standard quadratic error $E(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|_2^2$.

B. COMPARISON WITH NON-COMPOSITE MODELS

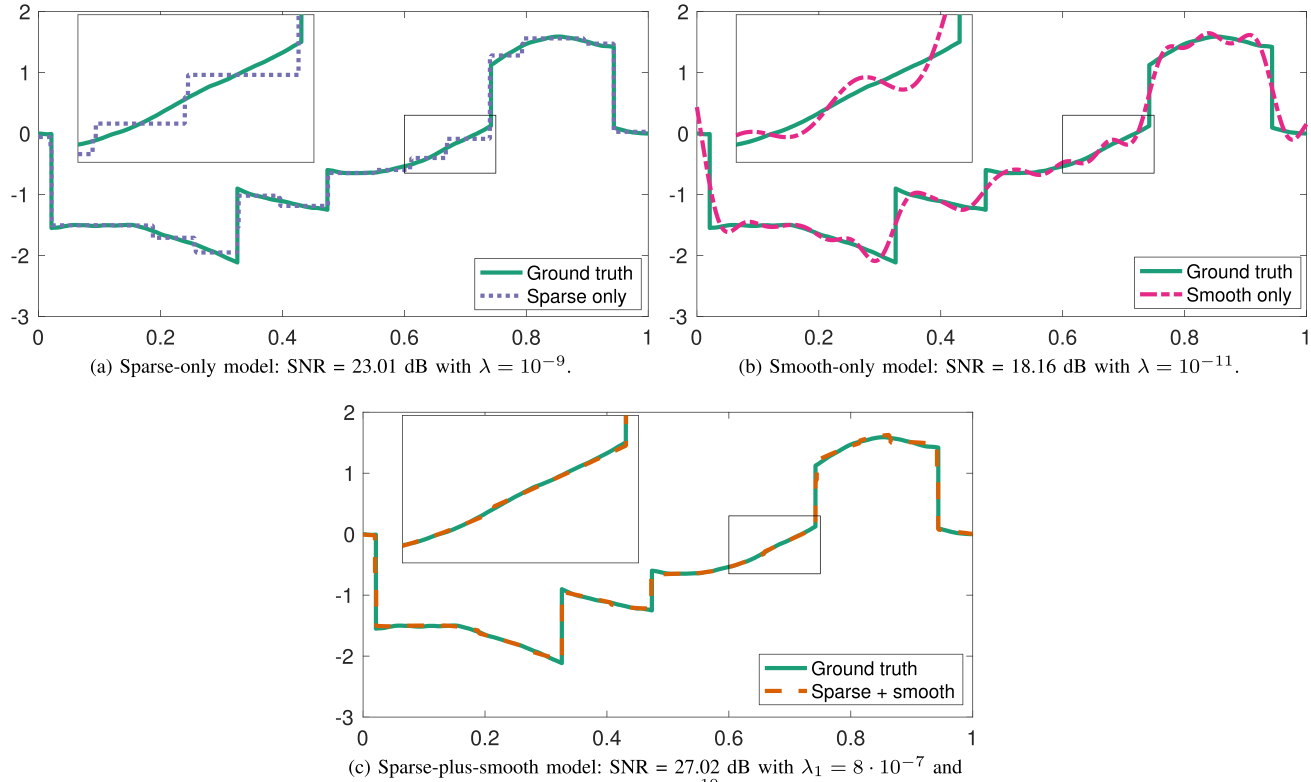
We now validate our new sparse-plus-smooth model against more standard non-composite models. More precisely, for $i \in \{1, 2\}$ we solve the regularized problems

$$\arg\min_{f\in\mathcal{X}_i} \{E(\nu(f), \mathbf{y}) + \lambda\mathcal{R}_i(f)\} \tag{35}$$

with regularizers $\mathcal{R}_1(f) = \|\mathrm{L}_1\{f\}\|_{\mathcal{M}}$ (sparse model with native space $\mathcal{X}_1 = \mathcal{M}_{\mathrm{L}_1}(\mathbb{R})$) and $\mathcal{R}_2(f) = \|\mathrm{L}_2\{f\}\|_{L_2}$ (smooth model with native space $\mathcal{X}_2 = \mathcal{H}_{\mathrm{L}_2}(\mathbb{R})$). We discretize these problems using the reconstruction spaces $V_i(\mathbb{R})$ described in this paper (without restricting $V_1(\mathbb{R})$ with the boundary conditions ϕ_0). The sparse model thus amounts to an ℓ_1-regularized discrete problem which we solve using ADMM, while the smooth model has a closed-form solution that can be obtained by inverting a matrix.

For this comparison, we choose regularization operators $\mathrm{L}_1 = \mathrm{D}$ and $\mathrm{L}_2 = \mathrm{D}^2$ with $M = 50$ Fourier-domain measurements (cosine sampling with $\omega_{\max} = 100$). We generate the ground-truth signal according to Section VI-A2, with $K_1 = 5$ jumps whose i.i.d. Gaussian amplitudes have the variance $\sigma_1^2 = 1$ for s_1^{GT}. For the smooth component s_2^{GT}, we generate a realization of a Gaussian white noise w with the variance $\sigma_2^2 = 100$, such that $\mathrm{L}_2\{s_2^{\mathrm{GT}}\} = w$. The measurements are corrupted by some i.i.d. Gaussian white noise $\mathbf{n} \in \mathbb{R}^M$ so that $\mathbf{y} = \nu(s_{\mathrm{GT}}) + \mathbf{n}$. We set the signal-to-noise ratio (SNR) between $\nu(s_{\mathrm{GT}})$ and $\mathbf{n}$ to be 50 dB. The regularization parameters are selected through a grid search with $h = 1/2^9$ to maximize the SNR of the reconstructed signal s with respect to the ground truth (defined as $\mathrm{SNR} = 10\log_{10}(\frac{\int_0^T (s^{\mathrm{GT}}(t))^2\mathrm{d}t}{\int_0^T (s(t)-s^{\mathrm{GT}}(t))^2\mathrm{d}t})$).

The results of this comparison in terms of SNR are shown in Table 1 for varying grid sizes h. For all methods, the SNR values increase when the grid size decreases, which is to be expected since the grids are embedded. The only exception is the sparse-plus-smooth reconstruction for the finest grid size, which is likely due to numerical issues arising from the increased dimension ($N \approx 2^{11}$) of the optimization problem. The effect of the grid size on the quality of the reconstruction varies between the models: it is almost non-existent for the smooth-only model, whereas it is most significant for our



(a) Sparse-only model: SNR = 23.01 dB with $\lambda = 10^{-9}$.

(b) Smooth-only model: SNR = 18.16 dB with $\lambda = 10^{-11}$.

(c) Sparse-plus-smooth model: SNR = 27.02 dB with $\lambda_1 = 8 \cdot 10^{-7}$ and $\lambda_2 = 5 \cdot 10^{-10}$.

FIGURE 2. Comparison between our sparse-plus-smooth model and non-composite models with regularization operators $\mathrm{L}_1 = \mathrm{D}$, $\mathrm{L}_2 = \mathrm{D}^2$, and $M = 50$ Fourier-domain measurements.

TABLE 1 SNR Values (In Db) of the Reconstructed Signal With Respect to the Ground Truth With Varying Grid Size *h*

h	Sparse plus smooth	Sparse only	Smooth only
2^6	18.02	17.72	18.16
2^7	21.87	21.08	18.16
2^8	23.46	22.07	18.16
2^9	27.02	23.01	18.16
2^{10}	25.70	23.04	18.16

sparse-plus-smooth model. Over all grid sizes, due to the fact that our sparse-plus-smooth signal model matches the ground truth, our reconstructed signal yields a higher SNR (27.02 dB) than the sparse-only (23.04 dB) and smooth-only (18.16 dB) models.

The reconstruction results for the grid size $h = 1/2^9$ are shown in Fig. 2. Our sparse-plus-smooth reconstruction is qualitatively much more satisfactory. As can be observed in the zoomed-in section, the sparse-only model is subject to a staircasing phenomenon in the smooth regions of the ground-truth signal, a well-known shortcoming of total-variation regularization. Conversely, our reconstruction is remarkably accurate in the smooth regions. Finally, the smooth-only model fails both visually and in terms of SNR, due to its inability to represent sharp jumps.

VII. CONCLUSION

We have introduced a continuous-domain framework for the reconstruction of multicomponent signals. It assumes two additive components, the first one being sparse and the other being smooth. The reconstruction is performed by solving a regularized inverse problem, using a finite number of measurements of the signal. The form of a solution to this problem is given by our representer theorem. This form justifies the choice of the search space in which we discretize the problem. Our discretization is exact, in the sense that it amounts to solving a continuous-domain optimization problem restricted to our search space. The discretized problem is then solved using our ADMM-based algorithm, which we validate on simulated data.

APPENDIX

A. PROOF OF THEOREM 1

The main technical part of our proof is to show the existence of a minimizer; once this is ensured, the optimization problem can be decoupled into two separate problems. Then, we can invoke representer theorems proven in [24] for these problems to obtain a parametric form of the solution of the original problem. Finally, the uniqueness of the smooth component follows from the strict convexity of the associated regularization penalty. We now recall some relevant notions that will be used throughout the proof.

Preliminaries

We extend the biorthogonal system $(\boldsymbol{\phi}_0, \mathbf{p}_0)$ for $\mathcal{N}_0$ to the 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, where $\tilde{\boldsymbol{\phi}}_i = \begin{bmatrix}\boldsymbol{\phi}_0 & \boldsymbol{\phi}_i\end{bmatrix}$ and $\tilde{\mathbf{p}}_i = \begin{bmatrix}\mathbf{p}_0 & \mathbf{p}_i\end{bmatrix}$ for $i \in \{1, 2\}$. It is known from [51, Theorem 4] that any function $s_1 \in \mathcal{M}_{\mathrm{L}_1}(\mathbb{R})$ has the unique decomposition

$$s_1 = \mathrm{L}_{1,\tilde{\boldsymbol{\phi}}_1}^{-1}\{w\} + \tilde{\mathbf{c}}_0^T \mathbf{p}_0 + \mathbf{c}_1^T \mathbf{p}_1, \tag{36}$$

where $w \in \mathcal{M}(\mathbb{R})$, $\tilde{\mathbf{c}}_0 = \boldsymbol{\phi}_0(s_1) \in \mathbb{R}^{N_0}$, $\mathbf{c}_1 = \boldsymbol{\phi}_1(s_1) \in \mathbb{R}^{N_{0,1}-N_0}$, and $\mathrm{L}_{1,\tilde{\boldsymbol{\phi}}_1}^{-1}$ is the pseudo-inverse operator of L_1 for the biorthogonal system $(\tilde{\boldsymbol{\phi}}_1, \tilde{\mathbf{p}}_1)$ [51, Section 3.2]. Using this decomposition, we can equip the space $\mathcal{M}_{\mathrm{L}_1}(\mathbb{R})$ with the norm

$$\|s_1\|_{\mathcal{M}_{\mathrm{L}_1}} \triangleq \|w\|_{\mathcal{M}} + \|\tilde{\mathbf{c}}_0\|_2 + \|\mathbf{c}_1\|_2. \tag{37}$$

Finally, an element $s_1 \in \mathcal{M}_{\mathrm{L}_1}(\mathbb{R})$ is in the restricted search space $\mathcal{M}_{\mathrm{L}_1,\phi_0}(\mathbb{R})$ if and only if $\tilde{\mathbf{c}}_0 = \mathbf{0}$.

Similarly, for any $s_2 \in \mathcal{H}_{\mathrm{L}_2}(\mathbb{R})$, there is a unique decomposition

$$s_2 = \mathrm{L}_{2,\tilde{\boldsymbol{\phi}}_2}^{-1}\{h\} + \mathbf{c}_0^T \mathbf{p}_0 + \mathbf{c}_2^T \mathbf{p}_2, \tag{38}$$

where $\mathbf{c}_0 = \boldsymbol{\phi}_0(s_2) \in \mathbb{R}^{N_0}$, $\mathbf{c}_2 = \boldsymbol{\phi}_2(s_2) \in \mathbb{R}^{N_{0,2}-N_0}$, and $h \in L_2(\mathbb{R})$. Consequently, the associated norm for the space $\mathcal{H}_{\mathrm{L}_2}(\mathbb{R})$ is defined as

$$\|s_2\|_{\mathcal{H}_{\mathrm{L}_2}} \triangleq \|h\|_{L_2} + \|\mathbf{c}_0\|_2 + \|\mathbf{c}_1\|_2. \tag{39}$$

Existence of a Solution

The first step is to prove that (13) has a minimizer. We do so by reformulating the problem as the minimization of a weak*-lower semicontinuous functional over a weak*-compact domain. We then prove the existence by relying on the generalized Weierstrass theorem.

We denote the cost at the trivial point (0,0) as $\mathcal{J}_0 = \mathcal{J}(0, 0) = E(\mathbf{0}, \mathbf{y})$. Adding the constraint $\mathcal{J}(s_1, s_2) \le \mathcal{J}_0$ does not change the solution set of the original problem, as it must hold for any minimizer of (13). So, from now on, we assume that the cost functional is upper-bounded by $\mathcal{J}_0$. This readily implies that

$$E(\boldsymbol{\nu}(s_1 + s_2), \mathbf{y}) \le \mathcal{J}_0, \tag{40}$$

$$\|\mathrm{L}_1\{s_1\}\|_{\mathcal{M}} \le \frac{\mathcal{J}_0}{\lambda_1}, \tag{41}$$

$$\|\mathrm{L}_2\{s_2\}\|_{L_2} \le \sqrt{\frac{\mathcal{J}_0}{\lambda_2}}. \tag{42}$$

The coercivity of $E(\cdot, \mathbf{y})$ implies the existence of a constant $C_1 > 0$ such that $E(\mathbf{z}, \mathbf{y}) \le \mathcal{J}_0 \Rightarrow \|\mathbf{z}\|_2 \le C_1$. Together with (40), this yields

$$\|\boldsymbol{\nu}(s_1 + s_2)\|_2 \le C_1. \tag{43}$$

Moreover, since $\boldsymbol{\nu}$ is weak*-continuous over $\mathcal{M}_{\mathrm{L}_1}(\mathbb{R})$, it is also continuous. This is due to the fact that a Banach space (in this case, the predual of $\mathcal{M}_{\mathrm{L}_1}(\mathbb{R})$) is isometrically embedded in its bidual [50]. Moreover, by assumption, $\boldsymbol{\nu}$ is continuous over $\mathcal{H}_{\mathrm{L}_2}(\mathbb{R})$. Hence, there exists a second constant $C_2 > 0$ such that

$$\|f_1\|_{\mathcal{M}_{\mathrm{L}_1}} + \|f_2\|_{\mathcal{H}_{L_2}} \le \frac{\mathcal{J}_0}{\lambda_1} + \sqrt{\frac{\mathcal{J}_0}{\lambda_2}} \Rightarrow \|\boldsymbol{\nu}(f_1 + f_2)\|_2 \le C_2. \tag{44}$$

Now, by taking

$$\begin{aligned} f_1 &= s_1 - \boldsymbol{\phi}_1(s_1)^T \mathbf{p}_1, \\ f_2 &= s_2 - \boldsymbol{\phi}_0(s_2)^T \mathbf{p}_0 - \boldsymbol{\phi}_2(s_2)^T \mathbf{p}_2, \end{aligned} \tag{45}$$

and, together with (41) and (42), we deduce that

$$\begin{aligned} &\left\| \boldsymbol{\nu}\left(s_1 - \boldsymbol{\phi}_1(s_1)^T \mathbf{p}_1 + s_2 - \boldsymbol{\phi}_0(s_2)^T \mathbf{p}_0 - \boldsymbol{\phi}_2(s_2)^T \mathbf{p}_2\right) \right\|_2 \\ &\le C_2. \end{aligned} \tag{46}$$

By using the triangle inequality and the two bounds (43) and (46), we have

$$\left\| \boldsymbol{\nu}\left(\boldsymbol{\phi}_1(s_1)^T \mathbf{p}_1 + \boldsymbol{\phi}_0(s_2)^T \mathbf{p}_0 + \boldsymbol{\phi}_2(s_2)^T \mathbf{p}_2\right) \right\|_2 \le C_1 + C_2. \tag{47}$$

Finally, the well-posedness assumption in Theorem 1 ensures the existence of a constant $B > 0$ such that

$$\forall q \in \mathcal{N}_{\mathrm{L}_1} + \mathcal{N}_{\mathrm{L}_2} : B\|\boldsymbol{\phi}_i(q)\|_2 \le \|\boldsymbol{\nu}(q)\|_2, \quad i \in \{0, 1, 2\}. \tag{48}$$

Hence, by taking

$$q = \boldsymbol{\phi}_1(s_1)^T \mathbf{p}_1 + \boldsymbol{\phi}_0(s_2)^T \mathbf{p}_0 + \boldsymbol{\phi}_2(s_2)^T \mathbf{p}_2 \tag{49}$$

and by applying the Inequality (48), we have that

$$\|\boldsymbol{\phi}_1(s_1)\|_2, \|\boldsymbol{\phi}_0(s_2)\|_2, \|\boldsymbol{\phi}_2(s_2)\| \le \frac{C_1 + C_2}{B}. \tag{50}$$

Therefore, the original problem (13) is equivalent to the constrained minimization problem

$$\min_{\substack{s_1 \in \mathcal{M}_{\mathrm{L}_1,\phi_0}(\mathbb{R}) \\ s_2 \in \mathcal{H}_{\mathrm{L}_2}(\mathbb{R})}} \mathcal{J}(s_1, s_2) \quad \text{s.t.} \quad \|s_1\|_{\mathcal{M}_{\mathrm{L}_1}} \le A_1, \|s_2\|_{\mathcal{H}_{\mathrm{L}_2}} \le A_2, \tag{51}$$

where $A_1 = \frac{\mathcal{J}_0}{\lambda_1} + \frac{C_1+C_2}{B}$ and $A_2 = \sqrt{\frac{\mathcal{J}_0}{\lambda_2}} + \frac{C_1+C_2}{B}$.

The cost functional in (51), which is the same as in (13), is weak* lower-semicontinuous. Moreover, the constraint cube is weak*-compact in the product topology due to the Banach-Alaoglu theorem [62, Theorem 3.15]. Hence, (51) reaches its infimum, and so does (13).

Form of the Solution Let $(\tilde{s}_1, \tilde{s}_2)$ be a solution of (13) and consider the minimization problem

$$\min_{s_1 \in \mathcal{M}_{\mathrm{L}_1,\boldsymbol{\phi}_0}(\mathbb{R})} \|\mathrm{L}_1\{s_1\}\|_{\mathcal{M}} \quad \text{s.t.} \quad \boldsymbol{\nu}(s_1) = \boldsymbol{\nu}(\tilde{s}_1). \tag{52}$$

Unser *et al.* have shown in [20] that (52) has a minimizer s_1^* of the form (14). One can also readily verify that $(s_1^*, \tilde{s}_2)$ is a minimizer of the original problem. Similarly, one can consider the minimization problem

$$\min_{s_2\in\mathcal{H}_{\mathrm{L}_2}(\mathbb{R})} \|\mathrm{L}_2\{s_2\}\|_{L_2} \quad \text{s.t.} \quad \boldsymbol{\nu}(s_2)=\boldsymbol{\nu}(\tilde{s}_2). \tag{53}$$

It is known from [24, Theorem 3] that (53) has a minimizer s_2^* of the form (15). Again, (s_1^*, s_2^*) is a solution of the original problem, which matches the form specified by Theorem 1.

Uniqueness of the Second Component To prove the final statement of Theorem 1, let us consider two arbitrary pairs of solutions $(\bar{f}_1,\bar{f}_2)$ and $(\tilde{f}_1,\tilde{f}_2)$ of Problem (13) and let us denote by $\mathcal{J}_{\min}$ their minimal cost value. The convexity of the cost functional yields that, for any $\alpha\in(0,1)$ and $(f_{\alpha,1},f_{\alpha,2})=\alpha(\bar{f}_1,\bar{f}_2)+(1-\alpha)(\tilde{f}_1,\tilde{f}_2)$, we have

$$\mathcal{J}(f_{\alpha,1},f_{\alpha,2})\le\alpha\mathcal{J}(\bar{f}_1,\bar{f}_2)+(1-\alpha)\mathcal{J}(\tilde{f}_1,\tilde{f}_2)=\mathcal{J}_{\min}. \tag{54}$$

The optimality of $(\bar{f}_1,\bar{f}_2)$ and $(\tilde{f}_1,\tilde{f}_2)$ implies that (54) must be an equality. In particular, we must have that

$$\|\mathrm{L}_2\{f_{\alpha,2}\}\|_2^2=\alpha\|\mathrm{L}_2\{\bar{f}_2\}\|_2^2+(1-\alpha)\|\mathrm{L}_2\{\tilde{f}_2\}\|_2^2. \tag{55}$$

Now, due to the strict convexity of $\|\mathrm{L}_2\{\cdot\}\|_{L_2}^2$, we deduce that $\mathrm{L}_2\{\bar{f}_2-\tilde{f}_2\}=0$, and hence that $(\bar{f}_2-\tilde{f}_2)\in\mathcal{N}_{\mathrm{L}_2}$. This implies that all solutions have the same second component up to a term in the null space of L_2.

B. CHOICE OF BOUNDARY CONDITION FUNCTIONALS ϕ_0

We discuss here our choice of the boundary-condition functionals ϕ_0 for certain common choices of operators L_i. We focus on multiple-order derivative operators $\mathrm{L}_i=\mathrm{D}^{N_{0,i}}$, although the discussion remains valid for the more general class of rational operators [57], which, to the best of our knowledge, is the largest class of spline-admissible operators that satisfy the first assumption in Section V-A. The null spaces $\mathcal{N}_{\mathrm{L}_i}$ are thus the spaces of polynomials of degree smaller than $N_{0,i}$. We assume for now that we have $N_{0,1}\le N_{0,2}$, in which case we have $\mathcal{N}_{\mathrm{L}_1}\subset\mathcal{N}_{\mathrm{L}_2}$ and thus $\mathcal{N}_0=\mathcal{N}_{\mathrm{L}_1}$ and $N_0=(D_1-1)$. Then, for any $\epsilon>0$, the functionals $\phi_0=\frac{1}{\epsilon}\mathrm{rect}(\frac{\cdot}{\epsilon})$ for $N_0=1$ and

$$\phi_0=\left(\delta,\dots,\delta^{(N_0-2)},\delta^{(N_0-1)}*\frac{1}{\epsilon}\mathrm{rect}\left(\frac{\cdot}{\epsilon}\right)\right)*\delta(\cdot-\frac{\epsilon}{2}), \tag{56}$$

for $N_0>1$, where $\mathrm{rect}(t)=1$ for $-1/2\le t<1/2$ and 0 elsewhere, are valid choices of a biorthogonal system matched to the basis $\mathbf{p}_0=(1,(\cdot),\dots,\frac{(\cdot)^{N_0-1}}{(N_0-1)!})*\delta(\cdot-\frac{\epsilon}{2})$ of $\mathcal{N}_0$. Indeed, one can easily verify that this choice satisfies the biorthonormality relation $\langle\phi_{0,i},p_{0,j}\rangle=\delta_{i-j}$ (Kronecker delta). Moreover, we have $\phi_{0,i}\in\mathcal{X}_{\mathrm{L}_1}$ (the predual of $\mathcal{M}_{\mathrm{L}_1}(\mathbb{R})$), which implies that $(\mathbf{p}_0,\phi_0)$ is indeed a valid biorthogonal system of $\mathcal{N}_0$ [51, Proposition 5]. The fact that $\phi_{0,i}\in\mathcal{X}_{\mathrm{L}_1}$ is proved in [63] for the case $N_0=2$; this proof can readily be extended to higher orders.

The boundary conditions (56), along with a choice of ϵ such that $\frac{\epsilon}{h}$ is arbitrarily small, is numerically equivalent to $\phi_0(f)=(f(0),\dots,f^{(N_0-1)}(0^+))$, where $f^{(N_0-1)}(0^+)$ is the right limit of $f^{(N_0-1)}$ at 0. It can easily be shown in this case that, for $s_1=\sum_{k\in\mathbb{Z}}c_1[k]\varphi_{1,k}\in V_1(\mathbb{R})$ with $\varphi_{1,k}=\beta_{\mathrm{L}_1}(\cdot-kh)$, we have $\phi_0(f)=\mathbf{0}\Leftrightarrow c_1[-N_0+1]=\dots=c_1[0]=0$, which leads to the constraints $\mathbf{A}\mathbf{c}_1=(c_{1,1},\dots,c_{1,N_0})=\mathbf{0}$ in Problem (30). This choice simplifies the optimization task by reducing the dimension of the problem, whereas other boundary conditions could lead to more complicated linear constraints and would make the optimization task more difficult.

So far, we have assumed that $N_{0,1}\le N_{0,2}$ since this condition is always satisfied for the most common case $\mathrm{L}_1=\mathrm{D}$. However, if $N_{0,1}>N_{0,2}$, then it is more convenient to apply the boundary conditions ϕ_0 to the second component, which leads to the same simple boundary conditions $\mathbf{A}\mathbf{c}_2=(c_{2,1},\dots,c_{2,N_0})=\mathbf{0}$. By default, we implicitly consider the more common case $N_{0,1}\le N_{0,2}$ throughout the paper and thus impose the boundary conditions on the first component.

C. EXPRESSION OF THE REGULARIZATION MATRIX $\mathbf{L}_2$

Factorization of the Autocorrelation Filter

To specify the regularization matrix for the second component, we must first express in a convenient form the autocorrelation filter ρ defined in Proposition 1. This is done in Proposition 4, which gives the expression of ρ and its "square root" g for the choice of basis function $\varphi_2=\beta_{\mathrm{L}_2^*\mathrm{L}_2}$ made in Section IV-B2.

Proposition 4 (Factorization of the autocorrelation filter): Let the assumptions in Section V-A be satisfied, and let $\varphi_2=\beta_{\mathrm{L}_2^*\mathrm{L}_2}=\beta_{\mathrm{L}_2}*\beta_{\mathrm{L}_2}^\vee$. Then, the basis $\{\varphi_{2,k}\}_{k\in\mathbb{Z}}$ forms a Riesz basis as required in Section IV-B, and the autocorrelation filter ρ defined in Proposition 1 is of the form

$$\rho=d_{\mathrm{L}_2}*d_{\mathrm{L}_2}^\vee*b, \tag{57}$$

where $b[k]=\beta_{\mathrm{L}_2^*\mathrm{L}_2}(k)$ is the B-spline kernel of the operator $\mathrm{L}_2^*\mathrm{L}_2$, which is a positive-semidefinite filter supported in $[-(D_2-2)\dots D_2-2]$. The filter ρ can thus be factorized as $\rho=g*g^\vee$ with

$$g=d_{\mathrm{L}_2}*b^{1/2}, \tag{58}$$

where the filter $b^{1/2}$ satisfies $b=b^{1/2}*(b^{1/2})^\vee$ and is of length $B=(D_2-1)$, and g is thus of length $G=2B=2(D_2-1)$.

Proof: We have that

$$\begin{aligned}\rho[k]&=\langle\mathrm{L}_2\{\varphi_{2,k}\},\mathrm{L}_2\{\varphi_{2,0}\}\rangle_{L_2}\\&=\langle\mathrm{L}_2^*\mathrm{L}_2\{\varphi_{2,k}\},\varphi_{2,0}\rangle_{\mathcal{H}'_{\mathrm{L}_2}\times\mathcal{H}_{\mathrm{L}_2}}\\&=\langle\sum_{k'\in\mathbb{Z}}d_{\mathrm{L}_2^*\mathrm{L}_2}[k]\delta(\cdot-(k+k')),\varphi_{2,0}\rangle_{\mathcal{H}'_{\mathrm{L}_2}\times\mathcal{H}_{\mathrm{L}_2}}\\&=\sum_{k'\in\mathbb{Z}}d_{\mathrm{L}_2^*\mathrm{L}_2}[k]b[k+k']\\&=(d_{\mathrm{L}_2^*\mathrm{L}_2}*b^\vee)[-k]\\&=(d_{\mathrm{L}_2^*\mathrm{L}_2}*b)[k],\end{aligned}$$

TABLE 2 **Relevant Filters and Their Supports ($C = \sqrt{\frac{2-\sqrt{3}}{6}}$)**

	d_{L_1}	d_{L_2}	$\rho = b * d_{\mathrm{L}_2} * d_{\mathrm{L}_2}^{\vee}$	$g = b^{1/2} * d_{\mathrm{L}_2}$	$b = \left(\beta_{\mathrm{L}_2^*\mathrm{L}_2}(k)\right)_{k\in\mathbb{Z}}$	$b^{1/2}$
Description	Finite-difference filter for L_1	Finite-difference filter for L_2	Autocorrelation filter for L_2	"Square root" of ρ ($\rho = g * g^{\vee}$)	Samples of basis function ($\mathrm{L}_2^*\mathrm{L}_2$ B-spline)	"Square root" of b ($b = b^{1/2} * (b^{1/2})^{\vee}$)
Introduced in	Definition 4	Definition 4	Proposition 1	Proposition 4	Proposition 4	Proposition 4
Support length	D_1	D_2	$2G-1 = 4D_2-5$	$G = 2D_2 - 2$	$2B-1 = 2D_2-3$	$B = D_2 - 1$
Support	$[0 \dots D_1 - 1]$	$[0 \dots D_2 - 1]$	$[-(G-1) \dots G-1]$	$[0 \dots G-1]$	$[-(B-1) \dots B-1]$	$[0 \dots B-1]$
Example $\mathrm{L}_2 = \mathrm{D}$		[1, -1]	[-1, 2, -1]	[1, -1]	[1]	[1]
Example $\mathrm{L}_2 = \mathrm{D}^2$		[1, -2, 1]	$\frac{1}{6}[1, 0, -9, 16, -9, 0, 1]$	$C[1, \sqrt{3}, -(3+2\sqrt{3}), 2+\sqrt{3}]$	$\frac{1}{6}[1, 4, 1]$	$C[1, 2+\sqrt{3}]$

where $\langle \cdot, \cdot \rangle_{\mathcal{H}'_{\mathrm{L}_2} \times \mathcal{H}_{\mathrm{L}_2}}$ denotes the duality product between $\mathcal{H}_{\mathrm{L}_2}(\mathbb{R})$ and its dual $\mathcal{H}'_{\mathrm{L}_2}(\mathbb{R})$, and the last line results from the symmetry of ρ and b.

Next, we prove that b is positive-semidefinite. Indeed, for any finitely supported filter c, we have that

$$\sum_{k,k'\in\mathbb{Z}} c[k]c[k']b[k-k'] = \left\| \sum_{k\in\mathbb{Z}} c[k]\beta_{\mathrm{L}_2}(\cdot - k) \right\|_{L_2}^2 \geq 0, \tag{59}$$

where we have used the property

$$b[k] = (\beta_{\mathrm{L}_2} * \beta_{\mathrm{L}_2}^{\vee})(k) = \langle \beta_{\mathrm{L}_2}, \beta_{\mathrm{L}_2}(\cdot - k) \rangle_{L_2}. \tag{60}$$

Finally, to prove the existence of $b^{1/2}$, we notice that b has the finite support $[-(B-1) \dots B-1]$ due to the finite support $(-D_2, D_2)$ of $\beta_{\mathrm{L}_2^*\mathrm{L}_2}$, and we have $B = (D_2 - 1)$. Since b is also symmetric, its z-transform satisfies $B(z) = B(z^{-1})$; therefore, for any zero z_k of $B(z)$, z_k^{-1} is also a zero. Moreover, it is well known that $B(\pm 1) \neq 0$, so that zeros must come in pairs $z_k \neq z_k^{-1}$. Hence, $B(z)$ can be written as $B(z) = \prod_{k=1}^{B}(1 - z_k z)(1 - z_k z^{-1})$. Hence, to take $b^{1/2}$ to be the inverse z-transform of $B^{1/2}(z) = \prod_{k=1}^{B}(1 - z_k z^{-1})$ is a valid choice (we clearly have $b = b^{1/2} * (b^{1/2})^{\vee}$), and (58) is readily obtained. ■

We summarize in Table 2 the different filters and their mutual relations. Without loss of generality, we take the filters d_{L_i} for $i \in \{1, 2\}$ to be causal, which leads to causal B-splines. These filters will be useful for the definition of the regularization matrix $\mathbf{L}_2$.

Expression of $\mathbf{L}_2$

The regularization matrix $\mathbf{L}_2 \in \mathbb{R}^{(N_2-1)\times N_2}$ for the smooth component is given by

$$\mathbf{L}_2 = \begin{pmatrix} \mathbf{M}^- & \mathbf{0} \\ \hline \multicolumn{2}{c}{\mathbf{M}} \\ \hline \mathbf{0} & \mathbf{M}^+ \end{pmatrix}. \tag{61}$$

The central matrix $\mathbf{M} \in \mathbb{R}^{(N_2-G+1)\times N_2}$ is given by

$$\mathbf{M} = \begin{pmatrix} g[G-1] & \cdots & g[0] & 0 & \cdots & 0 \\ 0 & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ 0 & \cdots & 0 & g[G-1] & \cdots & g[0] \end{pmatrix}, \tag{62}$$

where g is defined in Proposition 4. The matrices $\mathbf{M}^{\pm} \in \mathbb{R}^{(B-1)\times(G-1)}$ are defined as $[\mathbf{M}^-]_{i,j} = g^{-(B_2-i)}[G - B + (i-1) - (j-1)]$ and $[\mathbf{M}^+]_{i,j} = g^{+i}[G + (i-1) - j)]$ for $1 \leq i \leq (B_2 - 1)$ and $1 \leq j \leq (G-1)$, where the filter $g^{\pm k}$ are given by $g^{-k} = b^{1/2}|_{\{0,\dots,B-1-k\}} * d_{\mathrm{L}_2}$ (supported in $[0 \dots G-1-k]$) and $g^{+k} = b^{1/2}|_{\{k,\dots,B-1\}} * d_{\mathrm{L}_2}$ (supported in $\{k, \dots G-1\}$). Here, the notation $a|_J$ refers to the filter a restricted to the set J of indices, with $a|_J[k] = a[k]$ if $k \in J$, and $a|_J[k] = 0$ otherwise.

As an illustration, for $\mathrm{L}_2 = \mathrm{D}$, we have that $B = 1$ and, hence, simply that $\mathbf{L}_2 = \mathbf{M}$. For $\mathrm{L}_2 = \mathrm{D}^2$, we have $B = 2$ and

$$\mathbf{L}_2 = \begin{pmatrix} C & -2C & C & 0 & \cdots & \cdots & 0 \\ \hline \multicolumn{7}{c}{\mathbf{M}} \\ \hline 0 & \cdots & \cdots & 0 & C' & -2C' & C' \end{pmatrix}, \tag{63}$$

where $C = \sqrt{\frac{2-\sqrt{3}}{6}}$ and $C' = C(2+\sqrt{3})$.

D. PROOF OF PROPOSITION 3

Let $s_i = \sum_{k\in\mathbb{Z}} c_i[k]\varphi_{i,k}$ with $c_i \in V_i(\mathbb{Z})$ for $i \in \{1, 2\}$. The filters c_i are assumed to have values determined by the vector $\mathbf{c}_i = (c_i[m_i], \dots, c_i[M_i])$ at certain points. By definition of m_i and M_i, the values of c_i outside these intervals do not affect the measurements $\nu(s_i)$, and we clearly have that $\nu(s_i) = \mathbf{H}_i\mathbf{c}_i$. Therefore, these coefficients solely affect the regularization terms. We now show that, for a solution $(c_1, c_2) \in \mathcal{S}_{\mathrm{d}}$ to problem (23), the coefficients are uniquely determined by the vectors $\mathbf{c}_i$, and that the regularization terms $\|d_{\mathrm{L}_1} * c_1\|_{\ell_1}$ and $\langle c_2, \rho * c_2\rangle_{\ell_2}$ can thus be expressed exclusively in terms of these vectors.

Concerning the first component, this is proved in [29, Proposition 2], which shows that $\|\mathrm{L}_1\{s_1\}\|_{\mathcal{M}} = \|\mathbf{L}_1\mathbf{c}_1\|_1$. The additional constraint $\mathbf{A}\mathbf{c}_1$ comes from $\phi_0(s_1) = \mathbf{0}$ imposed on the search space $V_1(\mathbb{Z})$ of Problem (23).

We now consider the regularization term for the component $\langle c_2, \rho * c_2\rangle_{\ell_2}$. By Proposition 4, we have that $\rho = g * g^\vee$, and, hence, that $\langle c_2, \rho * c_2\rangle_{\ell_2} = \langle g * c_2, g * c_2\rangle_{\ell_2} = \|g * c_2\|_{\ell_2}^2$, where $g = b^{1/2} * d_{\mathrm{L}_2}$. We also have $g * c_2 = b^{1/2} * a$, where $a = d_{\mathrm{L}_2} * c_2$ is supported in $[1 \ldots M_2]$ by definition of the native space $V_2(\mathbb{Z})$ given in (27). Since $a[n] = \sum_{k=0}^{D_2-1} d_{\mathrm{L}_2}[k]c_2[n-k]$, $a[n]$ is entirely determined by the vector $\mathbf{c}_2$ for $1 \leq n \leq M_2$, which justifies our choice of the space $V_2(\mathbb{Z})$. For values of n outside this interval, there is a unique way of setting the coefficients $c_2[k]$ in order to nullify $a[n]$ and thus obtain that $c_2 \in V_2(\mathbb{Z})$. For example, $c_2[M_2+1]$ can be set to nullify $a[M_2+1]$ based on the (D_2-1) previous coefficients of c_2, and, similarly, all the $c_2[n]$ for $n > M_2+1$ can be set recursively to nullify all the $a[k]$ for all $k > M_2+1$. The same argument can be made to show that there is a unique choice $c_2[n]$ for $n < m_2$ that nullifies $a[k]$ for all $k < 1$.

We now compute the values of $(g * c_2)[n]$ in different regimes for n. We have that $(g * c_2)[n] = (b^{1/2} * a)[n] = \sum_{k=0}^{B_2-1} b^{1/2}[k]a[n-k]$ where a is supported in $[1 \ldots M_2]$. For $B_2 \leq n \leq M_2$, this sum is solely affected by the coefficients $\mathbf{c}_2 = (c_2[m_2], \ldots, c_2[M_2])$, so that the corresponding terms can be written in matrix form as $\mathbf{M}\mathbf{c}$ (the central part of the $\mathbf{L}_2$ matrix defined in (61)). Outside this interval, for example for $n = M_2+1$, we have that $(g * c_2)[n] = \sum_{k\in\mathbb{Z}} b^{1/2}|_{\{1,\ldots,B_2-1\}}[k]a[n-k]$, since the $k=0$ term is anyway nullified by the fact that $a[n] = 0$. An analogous reformulation allows us to have $(g * c_2)[n]$ only depend on the $\mathbf{c}_2$ coefficients. The same reformulation for all the coefficients $M_2+1 \leq n \leq (M_2+B_2-1)$ leads to the matrix $\mathbf{M}^+$ in (61), while a similar argument for coefficients $(g * c_2)[n]$ with $1 \leq n \leq (B_2-1)$ leads to the matrix $\mathbf{M}^-$.

We have thus proved that the solutions $(c_1, c_2) \in \mathcal{S}_\mathrm{d}$ to Problem (23) are uniquely determined by their coefficients $\mathbf{c}_i = (c_i[m_i], \ldots, c_i[M_i])$ for $i \in \{1, 2\}$, and that the regularization terms can be written $\|d_{\mathrm{L}_1} * c_1\|_{\ell_1} = \|\mathbf{L}_1\mathbf{c}_1\|_1$ and $\|g * c_2\|_{\ell_2}^2 = \|\mathbf{L}_2\mathbf{c}_2\|_2^2$. This, together with the fact that $\nu(\sum_{k\in\mathbb{Z}} c_i[k]\varphi_{i,k}) = \mathbf{H}_i\mathbf{c}_i$, proves that $\mathcal{J}_\mathrm{d}(c_1, c_2) = J(\mathbf{c}_1, \mathbf{c}_2)$. Conversely, for any $(\mathbf{c}_1, \mathbf{c}_2) \in \mathbb{R}^{N_1} \times \mathbb{R}^{N_2}$, there is a unique extension of these vectors to filters $c_i \in V_i(\mathbb{R})$ such that $\mathbf{c}_i = (c_i[m_i], \ldots, c_i[M_i])$ and $\mathcal{J}_\mathrm{d}(c_1, c_2) = J(\mathbf{c}_1, \mathbf{c}_2)$. These extensions are explicited in [29, Proposition 2] for c_1 and earlier in this proof for c_2. This proves the existence of the bijective linear mapping between the solution sets $\mathcal{S}_\mathrm{d}$ and $\mathcal{S}$ specified in Proposition 3.

REFERENCES

[1] A. Tikhonov, "Solution of incorrectly formulated problems and the regularization method," *Sov. Math.*, vol. 4, pp. 1035–1038, 1963.

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

[3] E. Candès, "Compressive Sampling," in *Proc. Int. Congr. Mathematicians*, Madrid, Spain, 2006, pp. 1433–1452.

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

[5] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Cambridge, MA, USA: Birkhäuser Basel, 2013.

[6] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *J. Roy. Stat. Soc.: Ser. B. (Methodol.)*, vol. 58, no. 1, pp. 267–288, 1996.

[7] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.

[8] M. Unser, J. Fageot, and H. Gupta, "Representer theorems for sparsity-promoting ℓ_1 regularization," *IEEE Trans. Inf. Theory*, vol. 62, no. 9, pp. 5167–5180, Sep. 2016.

[9] T. Hastie, R. Tibshirani, and M. Wainwright, *Statistical Learning With Sparsity*. London, U.K.: Chapman and Hall/CRC, 2015.

[10] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009.

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

[12] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imag. Vis.*, vol. 40, no. 1, pp. 120–145, 2010.

[13] 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.

[14] C. De Mol and M. Defrise, "Inverse imaging with mixed penalties," in *Proc. URSI EMTS*, Pisa, Italy, 2004, pp. 798–800.

[15] A. Gholami and S. Hosseini, "A balanced combination of tikhonov and total variation regularizations for reconstruction of piecewise-smooth signals," *Signal Process.*, vol. 93, no. 7, pp. 1945–1960, 2013.

[16] V. Naumova and S. Peter, "Minimization of multi-penalty functionals by alternating iterative thresholding and optimal parameter choices," *Inverse Problems*, vol. 30, no. 12, 2014, Art. no. 125003.

[17] M. Grasmair, T. Klock, and V. Naumova, "Adaptive multi-penalty regularization based on a generalized lasso path," *Appl. Comput. Harmon. Anal.*, vol. 49, no. 1, pp. 30–55, 2018.

[18] V. Debarnot, P. Escande, T. Mangeat, and P. Weiss, "Learning low-dimensional models of microscopes," *IEEE Trans. Comput. Imag.*, vol. 7, pp. 178–190, Dec. 2021, doi: 10.1109/TCI.2020.3048295.

[19] 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.

[20] 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, 2017.

[21] S. Aziznejad and M. Unser, "Multi-kernel regression with sparsity constraint," *J. Math. Data Sci.*, vol. 3, no. 1, pp. 201–224, 2021.

[22] G. Wahba, *Spline Models for Observational Data*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1990.

[23] B. Schölkopf, R. Herbrich, and A. Smola, "A generalized representer theorem," in *Lecture Notes Computer Science*, D. Helmbold and R. Williamson, Eds. Berlin, Germany: Springer, 2001, pp. 416–426.

[24] 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.

[25] S. Fisher and J. Jerome, "Spline solutions to L^1 extremal problems in one and several variables," *J. Approximation Theory*, vol. 13, no. 1, pp. 73–83, 1975.

[26] K. Bredies and M. Carioni, "Sparsity of solutions for variational inverse problems with finite-dimensional data," *Calculus Variations Partial Differ. Equ.*, vol. 59, no. 1, pp. 1–26, 2019.

[27] J. Fageot and M. Simeoni, "TV-based reconstruction of periodic functions," *Inverse Problems*, vol. 36, no. 11, 2020, Art. no. 115015.

[28] C. Boyer, A. Chambolle, Y. De Castro, V. Duval, F. de Gournay, and P. Weiss, "On representer theorems and convex regularization," *SIAM J. Optim.*, vol. 29, no. 2, pp. 1260–1281, 2019.

[29] 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, Jul. 2019.

[30] A. Flinth and P. Weiss, "Exact solutions of infinite dimensional total-variation regularized problems," *Inf. Inference: J. IMA*, vol. 8, no. 3, pp. 407–443, 2019.
[31] B. Adcock and A. Hansen, "Generalized sampling and infinite-dimensional compressed sensing," *Found. Comput. Math.*, vol. 16, no. 5, pp. 1263–1323, 2016.
[32] I. Daubechies, M. Defrise, and C. De Mol, "Sparsity-enforcing regularisation and ISTA revisited," *Inverse Problems*, vol. 32, no. 10, 2016, Art. no. 104001.
[33] C. de Boor, *A Practical Guide to Splines*. Berlin, Germany: Springer-Verlag GmbH, 2001.
[34] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part I-theory," *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 821–833, Feb. 1993.
[35] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Process. Mag.*, vol. 16, no. 6, pp. 22–38, Nov. 1999.
[36] I. Schoenberg, *Cardinal Spline Interpolation*. Philadelphia, PA, USA: SIAM, 1973.
[37] M. Belge, M. Kilmer, and E. Miller, "Efficient determination of multiple regularization parameters in a generalized l-curve framework," *Inverse Problems*, vol. 18, no. 4, pp. 1161–1183, 2002.
[38] S. Roth and M. Black, "Fields of experts: A framework for learning image priors," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, San Diego, CA, USA, Jun. 2005, pp. 860–867.
[39] Z. Chen, Y. Lu, Y. Xu, and H. Yang, "Multi-parameter tikhonov regularization for linear ill-posed operator equations," *J. Comput. Math.*, vol. 26, no. 1, pp. 37–55, 2008.
[40] S. Lu and S. Pereverzev, "Multi-parameter regularization and its numerical realization," *Numerische Mathematik*, vol. 118, no. 1, pp. 1–31, 2010.
[41] Z. Wang, "Multi-parameter tikhonov regularization and model function approach to the damped morozov principle for choosing regularization parameters," *J. Comput. Appl. Math.*, vol. 236, no. 7, pp. 1815–1832, 2012.
[42] R. Abhishake and S. Sivananthan, "Multi-penalty regularization in learning theory," *J. Complexity*, vol. 36, pp. 141–165, 2016.
[43] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *J. Roy. Stat. Soc.: Ser. B. (Stat. Methodol.)*, vol. 67, no. 2, pp. 301–320, 2005.
[44] 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. Med.*, vol. 73, no. 3, pp. 1125–1136, 2015.
[45] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*. Providence RI, USA: Amer. Math. Soc., Sep. 2001, doi: 10.1090/ulect/022.
[46] L. Vese and S. Osher, "Modeling textures with total variation minimization and oscillating patterns in image processing," *J. Sci. Comput.*, vol. 19, no. 1–3, pp. 553–572, 2003.
[47] L. Vese and S. Osher, "Image denoising and decomposition with total variation minimization and oscillatory functions," *J. Math. Imag. Vis.*, vol. 20, no. 1/2, pp. 7–18, 2004.
[48] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math.*, vol. 42, no. 5, pp. 577–685, 1989.
[49] L. Schwartz, *Théorie Des Distributions*. Paris, France: Hermann Paris, 1951.
[50] W. Rudin, *Real and Complex Analysis*. New York, NY, USA: McGraw-Hill, 1986.
[51] M. Unser and J. Fageot, "Native Banach spaces for splines and variational inverse problems," 2019, *arXiv:1904.10818*.
[52] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia PA, USA: Soc. Ind. Appl. Math., Jan. 1992, doi: 10.1137/1.9781611970104.
[53] A. Amini, R. Madani, and M. Unser, "A universal formula for generalized cardinal B-splines," *Appl. Comput. Harmon. Anal.*, vol. 45, no. 2, pp. 341–358, 2018.
[54] T. Debarre, S. Aziznejad, and M. Unser, "Hybrid-spline dictionaries for continuous-domain inverse problems," *IEEE Trans. Signal Process.*, vol. 67, no. 22, pp. 5824–5836, Nov. 2019.
[55] P. Bohra and M. Unser, "Computation of "best" interpolants in the l_p sense," in *Proc. 45th IEEE Int. Conf. Acoust., Speech, Signal Process.*, Barcelona, Spain, May 2020, pp. 5505–5509.
[56] 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.
[57] M. Unser, "Cardinal exponential splines: Part II-Think analog, act digital," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1439–1449, Apr. 2005.
[58] 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.
[59] L. Dadi, S. Aziznejad, and M. Unser, "Generating sparse stochastic processes using matched splines," *IEEE Trans. Signal Process.*, vol. 68, pp. 4397–4406, Jul. 2020, doi: 10.1109/TSP.2020.3011632.
[60] M. Unser and T. Blu, "Generalized smoothing splines and the optimal discretization of the Wiener filter," *IEEE Trans. Signal Process.*, vol. 53, no. 6, pp. 2146–2159, Jun. 2005.
[61] A. Badoual, J. Fageot, and M. Unser, "Periodic splines and Gaussian processes for the resolution of linear inverse problems," *IEEE Trans. Signal Process.*, vol. 66, no. 22, pp. 6047–6061, Nov. 2018.
[62] W. Rudin, *Functional Analysis* (International Series in Pure and Applied Mathematics.) New York, NY, USA: McGraw-Hill, Inc., 1991.
[63] M. Unser, "A representer theorem for deep neural networks," *J. Mach. Learn. Res.*, vol. 20, no. 110, pp. 1–30, 2019.

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 under the direction of Prof. Unser and Dr. Fageot, EPFL. His research interests include splines, inverse problems, and variational methods for reconstructing sparse signals.

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

MICHAEL UNSER (Fellow, IEEE) is a Professor and the Director of EPFL's Biomedical Imaging Group, 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 coauthored more than 350 journal papers on those topics. He is the 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 was on the Editorial Board of most of the primary journals in his field including the IEEE TRANSACTIONS ON MEDICAL IMAGING (Associate Editor-in-Chief 2003–2005), IEEE TRANSACTIONS IMAGE PROCESSING, Proc. of IEEE, and *SIAM Journal of Imaging Sciences*. He is the Founding Chair of the technical committee on Bio Imaging and Signal Processing (BISP) of the IEEE Signal Processing Society.

He is an EURASIP fellow (2009), and a member of the Swiss Academy of Engineering Sciences. He was the recipient of several international prizes including five IEEE-SPS best paper awards, two Technical Achievement Awards from the IEEE (2008 SPS and EMBS 2010), the Technical Achievement Award from EURASIP (2018), and a recent Career Achievement Award (IEEE EMBS 2020).