Generalized Smoothing Splines and the Optimal Discretization of the Wiener Filter

Michael Unser, *Fellow, IEEE,* and Thierry Blu, *Member, IEEE*

***Abstract*—We introduce an extended class of cardinal L*L-splines, where L is a pseudo-differential operator satisfying some admissibility conditions. We show that the L*L-spline signal interpolation problem is well posed and that its solution is the unique minimizer of the spline energy functional $\|Ls\|_{L_2}^2$, subject to the interpolation constraint. Next, we consider the corresponding regularized least squares estimation problem, which is more appropriate for dealing with noisy data. The criterion to be minimized is the sum of a quadratic data term, which forces the solution to be close to the input samples, and a "smoothness" term that privileges solutions with small spline energies. Here, too, we find that the optimal solution, among all possible functions, is a cardinal L*L-spline. We show that this smoothing spline estimator has a stable representation in a B-spline-like basis and that its coefficients can be computed by digital filtering of the input signal. We describe an efficient recursive filtering algorithm that is applicable whenever the transfer function of L is rational (which corresponds to the case of exponential splines).**

We justify these algorithms statistically by establishing an equivalence between L*L smoothing splines and the minimum mean square error (MMSE) estimation of a stationary signal corrupted by white Gaussian noise. In this model-based formulation, the optimum operator L is the whitening filter of the process, and the regularization parameter is proportional to the noise variance. Thus, the proposed formalism yields the optimal discretization of the classical Wiener filter, together with a fast recursive algorithm. It extends the standard Wiener solution by providing the optimal interpolation space. We also present a Bayesian interpretation of the algorithm.

***Index Terms*—Nonparametric estimation, recursive filtering, smoothing splines, splines (polynomial and exponential), stationary processes, variational principle, Wiener filter.**

I. Introduction

In a series of companion papers, we have introduced a general continuous/discrete approach to signal processing that uses an exponential B-spline representation of signals [1], [2]. A key feature of this framework is that it allows one to implement continuous-time signal processing operations exactly, and quite efficiently, by simple discrete processing of the B-spline coefficients of the signal. Its main advantage over the traditional bandlimited approach is that the underlying basis functions are compactly supported. The E-spline framework offers the choice of a large variety of basis functions, which are specified in terms of their poles and zeros, and which constitute a considerable extension of the standard, piecewise polynomial ones. While the availability of such a rich collection of signal models provides many new design opportunities, it also raises the issue of the selection of the most appropriate one for a given application. One possible criterion is to favor simplicity and to choose the input signal representation (e.g., piecewise constant or linear) that minimizes algorithmic complexity. Another option, which is the one developed in this paper, is to optimize the choice of the signal space based on the characteristics (e.g., general smoothness properties or prior statistical distribution) of the class of continuous-time signals from which the measurements (discrete samples) are derived. This can be accomplished in essentially two ways: a) within a deterministic framework by searching for a solution that minimizes some suitable smoothness energy functional and b) within a stochastic framework by deriving a Bayesian or a minimum mean square error (MMSE) estimator of the signal. Similar to what has been noted in the related area of image restoration [3]–[5], we will see that these various points of views (which may be referred to as Lagrange/Tikhonov, Bayesian, and Wiener) are mathematically equivalent and that they lead to generalized spline solutions that can be determined by digital filtering techniques. We will also take advantage of the proposed formalism to specify a rather general class of spline estimators that can handle noisy input data, similar to the classical smoothing splines proposed by Schoenberg and Reinsh [6], [7].

The model-based approach to splines that we are suggesting has its roots in approximation theory: There is an extensive literature that deals with the variational aspects of splines and even a whole sub-branch of spline theory that uses the energy minimization property as starting point for the definition of a generalized notion of spline [8]–[10]. This kind of formulation is also the one that carries over best to multiple dimensions; for instance, it has led to the methods of thin-plate splines and radial basis functions [11]–[13], which have become quite popular for the interpolation of scattered data in multiple dimensions.

The equivalence between variational splines and statistical estimation techniques—in particular, Bayesian ones—has been recognized early on [14], [15]. Smoothing splines and their multidimensional extensions have been studied in depth by Wahba, who used a powerful Reproducing-Kernel-Hilbert-Space formulation [16]; these efforts led to a variety of nonparametric linear estimation techniques, which are now widely used in statistics [17]–[19]. Somewhat lesser known is the work of Weinert, who uncovered the connection between spline interpolation and MMSE estimation and who used this formal link to derive Kalman-type recursive smoothing algorithms [20],

Manuscript received February 19, 2004; revised June 22, 2004. This work was supported in part by the Swiss National Science Foundation under Grant 200020-101821. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Karim Drouiche.

The authors are with the Biomedical Imaging Group, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland (e-mail: michael.unser@epfl.ch; Thierry.Blu@epfl.ch).

Digital Object Identifier 10.1109/TSP.2005.847821

1053-587X/$20.00 © 2005 IEEE

[21]. The theoretical connection that exists between spline estimation techniques and the kriging methods that were developed in geostatistics for the interpolation of scattered data is also relevant to the issue [22]–[24].

Unfortunately, the adaptation of these mathematical results to the cardinal framework is not straightforward. The major difficulty is that the above-mentioned formulations all consider a finite number of samples—these cannot be simply transposed to infinite dimensions without resolving some delicate convergence and stability issues. In addition, the generality of some of the methods (they consider nonuniform samples as well as space-varying differential models) is such that it does not facilitate the identification of fast computational schemes. This is a typical situation where it is more efficient to formulate the constrained version of the problem and to derive the solution we are seeking. This justifies our present effort, which is to revisit the variational aspects of splines within the cardinal framework. As in our previous work, we apply a signal processing formulation and take advantage of the shift-invariant structure of the problem, with the following benefits:

- *Simplification of the theory:* Even though there are a few technicalities associated with the infinite dimensionality of the data sequences, we can simplify the mathematics by applying Fourier techniques. In particular, there is no more need for a special treatment of the homogeneous component of the solution (in our view, the most delicate aspect of the classical formulation) because there are no boundary conditions.
- *Generality:* The Fourier domain formulation allows us to consider a larger class of spline-defining (pseudo-differential) operators than what is done usually—with the restriction that they need to be shift-invariant.
- *Self-contained formulation:* The present derivations and algorithms can all be explained in standard signal processing terms, making them accessible and more appealing to this community.
- *New computational solutions and link with standard signal processing methods:* As in our previous work, we can solve the spline fitting problems using digital filters. The formulation also yields an optimal discretization of the classical Wiener filter.

The paper is organized as follows. In Section II, we start with a brief exposition of the key mathematical notions that are required for this paper. In Section III, we enlarge our previous family of cardinal E-splines by introducing the notion of *spline-admissible* convolution operator. We then consider the class of positive-definite operators $\mathrm{L}^*\mathrm{L}$ that yield generalized cardinal splines with interesting mathematical properties. In Section IV, we show that these $\mathrm{L}^*\mathrm{L}$-splines can also be specified via the minimization of a pseudo-differential energy functional. We use this functional to define a smoothing spline estimator for noisy signals. Taking advantage of the existence of a B-spline-like representation of these splines, we derive efficient smoothing spline algorithms, extending our earlier results for polynomial splines [25]. In Section V, we propose a statistical interpretation of the proposed spline fitting techniques. In particular, we show that the smoothing spline algorithm solves the problem of the MMSE estimation of a stationary signal corrupted by additive white noise. This result is important because it allows us to optimally tune the parameters of the spline algorithm based on some *a priori* knowledge of second-order statistics of the signal and noise. It also yields a fast recursive implementation of the classical Wiener filter (cf. (18) and Appendix II), as an interesting by-product. Finally, we also show how to recast the spline estimation problem into a Bayesian framework, which is more in line with the interpretation of Kimeldorf and Wahba [14].

II. Preliminaries

We start with a presentation of the notations and mathematical tools that are used throughout the paper.

A. Continuous-Time Operators and Function Spaces

In our formulation, we consider an extended family of linear shift-invariant operators L, which are characterized by a convolution kernel $L(t)$, i.e., $\mathrm{L}f = L(t) * f(t)$, where the argument $f = f(t)$ is a continuous-time function with $t \in \mathbb{R}$. We assume that $L(t)$ is a tempered distribution (i.e., $L \in S'$) and that its Fourier transform $\mathcal{F}\{L\}(\omega) = \hat{L}(\omega)$ is a *true function* of ω. The adjoint operator is denoted by L^*, and its impulse response is $L^T(t) := L(-t)$.

We classify the operators L according to an equivalent differentiation order r.

Definition 1: The convolution operator L is of order r if and only if, for all positive real $\rho < r - 1/2$, we have that

$$\sum_{n\in\mathbb{Z}} \frac{|\omega + 2n\pi|^{2\rho}}{1 + |\hat{L}(\omega + 2n\pi)|^2} \leq C_\rho < \infty. \tag{1}$$

A prototypical example of an rth order operator is the rth fractional derivative D^r, whose frequency response is $(j\omega)^r$. An alternative notation is $\mathrm{D}^r f = f^{(r)}$.

To each operator L, we associate a corresponding generalized Sobolev space

$$W_2^L = \left\{ f(t), t \in \mathbb{R} \right. \\ \left. : \int_{-\infty}^{+\infty} |\hat{f}(\omega)|^2 (1 + |\hat{L}(\omega)|^2) d\omega < +\infty \right\}$$

where $\hat{f}(\omega) = \mathcal{F}\{f\} = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt$ is the Fourier transform of $f(t)$. For the particular choice $\mathrm{L} = \mathrm{D}^r$, we recover the usual Sobolev spaces of order r, i.e., $W_2^{\mathrm{D}^r} = W_2^r$.

Our first result clarifies the link between the order of the operator L and the classical notion of Sobolev smoothness (for a proof, see Appendix I).

Theorem 1: Let L be a linear shift-invariant operator of order $r > 1/2$; then, for all $0 \leq \rho < r - 1/2$, we have the following properties.

i) $W_2^{\mathrm{L}} \subset W_2^\rho$.

ii) If $f \in W_2^{\mathrm{L}}$, then the Poisson summation formula holds

$$\sum_{n\in\mathbb{Z}} \mathcal{F}\{f^{(\rho)}\}(\omega + 2n\pi) = \sum_{n\in\mathbb{Z}} f^{(\rho)}(n) e^{-jn\omega}$$

almost everywhere.

iii) If $f \in W_2^{\mathrm{L}}$, then the integer samples $f^{(\rho)}(n)$ are square summable: $\{f^{(\rho)}(n)\}_{n\in\mathbb{Z}} \in \ell_2$.

In particular, when L is of smoothness order $r > 1/2$, then the samples $f(n)$ of $f \in W_2^{\mathrm{L}}$ are in ℓ_2.

B. Digital Filters and Sequence Spaces

Discrete signals are indexed by the discrete variable $k \in \mathbb{Z}$ using square brackets to distinguish them from their continuous-time counterparts. Such signals will be distinguished based on their appartenance to the classical sequence spaces: $\ell_p = \{a[k] : \|a\|_{\ell_p} < +\infty\}$ with $\|a\|_{\ell_p} := \left(\sum_{k\in\mathbb{Z}} |a[k]|^p\right)^{1/p}$ for $1 \le p < \infty$, and $\|a\|_{\ell_\infty} := \sup_{k\in\mathbb{Z}} |a[k]|$.

The digital filtering of a sequence $a[k]$ yields the signal $(h * a)[k] = \sum_{l\in\mathbb{Z}} a[l]h[k-l]$, where h is the impulse response response of the filter; when the discrete context is obvious, we will sometimes drop the time indices and simply write $h * a$. The digital filter can also be represented by its equivalent continuous-time impulse response $h(t) = \sum_{k\in\mathbb{Z}} h[k]\delta(t-k)$.

Here, we will only consider "stable" filters whose impulse responses are in ℓ_1. The frequency response of such a filter, which is obtained by evaluating the z-transform of h, $H(z) = \sum_{k\in\mathbb{Z}} h[k]z^{-k}$, for $z = e^{j\omega}$, is guaranteed to be bounded and continuous [26].

An important related mathematical result is Young's inequality

$$\|h * a\|_{\ell_p} \le \|h\|_{\ell_1} \cdot \|a\|_{\ell_p}, \text{ for } 1 \le p \le +\infty \tag{2}$$

which implies that the application of such a filter to an ℓ_p-sequence produces an output signal that is in ℓ_p as well. For instance, by taking $p = \infty$, we can deduce that the digital filtering of a bounded input (BI) signal produces a bounded output (BO) signal. In fact, there is a well-known equivalence between $h \in \ell_1$ and BIBO-stability, where the latter property is crucial for signal processing applications [27]. Another direct consequence of Young's inequality for $p = 1$ is that the cascade of two BIBO-stable filters $h_1 \in \ell_1$ and $h_2 \in \ell_1$ yields a filter $h_1 * h_2 \in \ell_1$ that is BIBO-stable as well. In addition, note that $h \in \ell_1$ implies that h has finite energy (i.e., $h \in \ell_2$) because of the inclusion property of the discrete ℓ_p spaces:

$$\ell_p \subset \ell_{p'}, \text{ for } 1 \le p \le p' \le +\infty.$$

The convolution inverse of h will be denoted by $(h)^{-1}$. A sufficient condition for the existence of this inverse is provided by Wiener's Lemma [29], which states that $(h)^{-1} \in \ell_1$ if $h \in \ell_1$ and $H(e^{j\omega}) \neq 0$ for $\forall\, \omega \in [0, 2\pi]$. Hence, the inverse filter is stable provided that the frequency response of h is nonvanishing.

The adjoint operator of h (real-valued) is the time-reversed filter whose impulse response is $h^T[k] = h[-k]$. Indeed, we have the ℓ_2-inner product relation

$$\forall\, a, c \in \ell_2, \quad \langle a, h * c\rangle_{\ell_2} = \langle h^T * a, c\rangle_{\ell_2}$$

which is established by simple change of summation variable.

C. Riesz Bases

Cardinal splines provide a one-to-one mapping between discrete sequences and continuous-time functions. A convenient mathematical way of describing this mapping is through the specification of a Riesz basis for a given spline family. In the present context, the Riesz bases have a convenient integer-shift-invariant structure $\{\varphi(t-k)\}_{k\in\mathbb{Z}}$, where the generator $\varphi(t)$ is a generalized B-spline, which is typically the shortest possible (or most localized) spline within the given family. A standard result in sampling theory is that the integer shifts of a function $\varphi(t) \in L_1 \cap L_2$ form a Riesz basis[1] if and only if there exist two positive constants A_φ and B_φ such that

$$\begin{cases} 0 < A_\varphi^2 := \inf_{\omega\in[0,2\pi]} \sum_{k\in\mathbb{Z}} |\hat{\varphi}(\omega + 2k\pi)|^2 \\ B_\varphi^2 := \sup_{\omega\in[0,2\pi]} \sum_{k\in\mathbb{Z}} |\hat{\varphi}(\omega + 2k\pi)|^2 < +\infty \end{cases} \tag{3}$$

where $\hat{\varphi}(\omega)$ denotes the Fourier transform of $\varphi(t)$. The corresponding function space, which is a subspace of L_2, is

$$V(\varphi) = \left\{ s(t) = \sum_{k\in\mathbb{Z}} c[k]\varphi(t-k) : c \in \ell_2 \right\}. \tag{4}$$

The Riesz basis conditions implies that we have an equivalence between the L_2-norm of a signal $s(t) \in V(\varphi)$ and the ℓ_2-norm of its coefficients $c[k]$:

$$A_\varphi \cdot \|c\|_{\ell_2} \le \left\| \sum_{k\in\mathbb{Z}} c[k]\varphi(t-k) \right\|_{L_2} \le B_\varphi \cdot \|c\|_{\ell_2}$$

with equality if and only if the basis is orthonormal, that is, when $A_\varphi = B_\varphi = 1$.

In some instances—for example, when one is processing samples of a stationary process as in Section V —one is interested in enlarging the space of admissible splines $s(t)$ by allowing for B-spline coefficients $c[k] \in \ell_\infty$. This is possible, provided that the Riesz basis is L_p-stable, that is, when there exist two positive constants $0 < A_{\varphi,p}$ and $B_{\varphi,p} < +\infty$ such that

$$A_{\varphi,p} \cdot \|c\|_{\ell_p} \le \| \sum_{k\in\mathbb{Z}} c[k]\varphi(t-k)\|_{L_p} \le B_{\varphi,p} \cdot \|c\|_{\ell_p}$$

for any $1 \le p \le \infty$. It can be shown (cf. [30]) that a sufficient condition for an integer-shift-invariant Riesz basis $\{\varphi(t-k)\}_{k\in\mathbb{Z}}$ to be L_p-stable (for all $1 \le p \le +\infty$) is

$$B_{\varphi,p} \le \|\varphi\|_{L_{1,\infty}} := \sup_{t\in[0,1]} \sum_{k\in\mathbb{Z}} |\varphi(t+k)| < +\infty \tag{5}$$

which adds a stronger constraint to (3). In particular, our L_p-stability condition $\|\varphi\|_{L_{1,\infty}} < +\infty$ implies that $\varphi(t)$ must be bounded and included in L_p for all $p \in [1, +\infty]$. Conversely, we note that (5) is satisfied whenever $\varphi(t)$ is bounded and decays faster than $O(1/|t|^{1+\epsilon})$ with $\epsilon > 0$. This is obviously the

[1]The condition $\varphi(t) \in L_1$ is not absolutely necessary, but it simplifies the formulation by ensuring that the sum in (3) is continuous in ω when B_φ exists.

case for the exponential B-splines described in [2], which are all compactly supported.

We now establish a Young-type convolution inequality for the L_p-stability condition, which will be required later on.

Proposition 1: Let $\varphi_1(t)$ and $\varphi_2(t)$ be two L_p-stable functions. Then

$$\|\varphi_1 * \varphi_2\|_{L_{1,\infty}} \leq \|\varphi_1\|_{L_{1,\infty}} \cdot \|\varphi_2\|_{L_{1,\infty}} < +\infty.$$

Proof: We consider the following 1-periodic function in t, which we bound from above:

$$\begin{aligned}
&\sum_{k\in\mathbb{Z}} |(\varphi_1 * \varphi_2)(t+k)| \\
&\quad \leq \sum_{k\in\mathbb{Z}} \int_{-\infty}^{+\infty} |\varphi_1(\tau)||\varphi_2(t+k-\tau)|d\tau \\
&\quad = \sum_{k\in\mathbb{Z}} \sum_{l\in\mathbb{Z}} \int_0^1 |\varphi_1(\tau+l)||\varphi_2(t+k-\tau-l)|d\tau \\
&\quad = \int_0^1 \sum_{l\in\mathbb{Z}} |\varphi_1(\tau+l)| \\
&\qquad \cdot \sum_{k'\in\mathbb{Z}} |\varphi_2(t+k'-\tau)|\, d\tau \text{ (Fubini's Theorem)} \\
&\quad \leq \sup_{\tau\in[0,1]} \sum_{l\in\mathbb{Z}} |\varphi_1(\tau+l)| \cdot \int_0^1 \sum_{k'\in\mathbb{Z}} |\varphi_2(t+k'-\tau)|\, d\tau \\
&\quad \leq \sup_{\tau\in[0,1]} \sum_{l\in\mathbb{Z}} |\varphi_1(\tau+l)| \cdot \sup_{t'\in[0,1]} \sum_{k'\in\mathbb{Z}} |\varphi_2(t'+k')|
\end{aligned}$$

where we have used the fact that the functions $\sum_{l\in\mathbb{Z}} |\varphi_1(\tau + l)|$ and $\sum_{k'\in\mathbb{Z}} |\varphi_2(t + k' - \tau)|$ are 1-periodic in τ and $t' = t - \tau$, respectively, and where we have bounded them by their maximum within this period. Since this inequality is true for all t, it also holds for the value of t at which the left-hand side function reaches its maximum, which yields the desired result. ■

III. L*L-Splines and Cardinal Interpolation

A. Spline-Admissible Operators

In our previous series of papers, we have investigated the cardinal exponential splines and have shown how these can be specified via a differential operator L, whose transfer function can be expressed as a polynomial in ω or, more generally, the ratio of two such polynomials. Here, we extend our class of linear, shift-invariant operators L further to encompass an even wider family of splines. The operators that we will be considering are referred to as "spline-admissible"; they must satisfy the following properties.

Definition 2: L is a *spline-admissible* operator of order r if and only if we have the following.

1) L is a linear shift-invariant operator of smoothness order $r > 1/2$.
2) L has a well-defined inverse L^{-1} whose impulse response $\rho(t)$ is a function of slow growth (i.e., $\rho(t) \in S'$). Thus, L admits $\rho(t)$ as Green's function: $\mathrm{L}\{\rho\} = \delta(t)$.
3) There exists a corresponding spline-generating function $\beta(t) = \Delta_{\mathrm{L}}\{\rho(t)\} := \sum_{k\in\mathbb{Z}} d[k]\rho(t-k)$ (generalized B-spline) that satisfies the Riesz basis condition (3).
4) The localization operator Δ_{L} in 4) is a stable digital filter in the sense that $d \in \ell_1$.
5) The spline generator $\beta(t)$ satisfies the L_p-stability condition (5).

Properties 1) to 3) are quite explicit and easy to check. Properties 4) and 5) are less direct statements; the most delicate issue is to establish that there is indeed a function $\beta(t)$ in $L_2 \cap \operatorname{span}\left\{L^{-1}\{\delta(t-k)\}\right\}_{k\in\mathbb{Z}}$ that satisfies the Riesz basis condition. This typically needs to be done on a case-by-case basis for a given family of operators.

We will now elaborate some more on the crucial role played by the localization filter and justify the use of this terminology. The Fourier transform of $\beta(t)$ is given by

$$\hat{\beta}(\omega) = \frac{\Delta_{\mathrm{L}}(e^{j\omega})}{\hat{L}(\omega)}$$

so that a necessary requirement for it to be in L_2 is that the localization filter $\Delta_{\mathrm{L}}(e^{j\omega}) = \sum_{k\in\mathbb{Z}} d[k]e^{-j\omega k}$ be such that is compensates for the potential singularities of the inverse filter $1/\hat{L}(\omega)$. In particular, this means that $\Delta_{\mathrm{L}}(e^{j\omega})$ must locally have the same behavior as $\hat{L}(\omega)$, especially around the frequencies where $\hat{L}(\omega)$ is vanishing. An important practical requirement is that $\hat{\beta}(\omega)$ should have the largest possible degree of continuity (ideally, $\hat{\beta}(\omega) \in C^\infty$) to guarantee that $\beta(t)$ has very fast decay (ideally, compact support). It also makes good sense to select a digital filter $\Delta_{\mathrm{L}}(e^{j\omega})$ such that $\hat{\beta}(\omega)$ is close to one over a reasonable frequency range. This will have the following desirable effects:

1) It will ensure that the discrete operator $\Delta_{\mathrm{L}}\{\cdot\}$ is a good approximation of the continuous-time operator $\mathrm{L}\{\cdot\}$.
2) It will yield an impulse-like spline generator $\beta(t)$ that is well localized in time; this function can be thought of as a regularized approximation of $\delta(t)$ in the space generated by the Green function of L, which is typically nonlocal.

Prominent examples of spline-admissible operators are the nth-order derivatives $\mathrm{D}^n = (d^n/dt^n)$ and, more generally, the whole class of operators with rational transfer functions that generate the generalized exponential splines [2]; the only constraint is that the degree of the numerator (N) must be greater than that of the denominator (M) so that the order $r = M - N$ is positive. The generalized splines considered in [31] are also included. We should note, however, that the present class is considerably larger; for instance, it includes fractional derivatives that admit B-spline like generators, albeit not compactly supported when the order is noninteger [32]. It also contains some more exotic self-similar operators that can be associated with refinable basis functions and wavelets [33].

This naturally leads to the following definition of a generalized cardinal L-spline.

Definition 3: The continuous-time function $s(t)$ is a cardinal L-spline if and only if $\mathrm{L}\{s(t)\} = \sum_{k\in\mathbb{Z}} a[k]\delta(t-k)$ with $a[k] \in \ell_\infty$.

Since L is a generalized rth-order differentiation operator, the intuitive meaning of this definition is that $s(t)$ is piecewise "smooth" with rth-order discontinuities at the integers.

In general, when L is spline-admissible, it is possible to represent such a spline in terms of its generalized B-spline expansion $s(t) = \sum_{k\in\mathbb{Z}} c[k]\beta(t-k)$ with $c[k] \in \ell_\infty$. Note that the $a[k]$'s in Definition 3 are related to the $c[k]$'s through the discrete convolution relation $a[k] = (d * c)[k]$, where $d \in \ell_1$ is the localization filter.

B. Positive Definite Operators and Symmetric B-Spline Interpolation

One potential problem when considering general L-splines is that the corresponding cardinal interpolation problem is not necessarily well-posed. In this section, we will show that this problem can be avoided by considering the cardinal splines associated with the class of positive definite operators L*L. To this end, we will first specify an L_p-stable Riesz basis for these splines. We will then show that the corresponding interpolation problem is always well-defined and that it can be solved efficiently by digital filtering.

Proposition 2: Let L be a spline-admissible operator of order r with spline generator $\beta(t)$ such that $\mathrm{L}\{\beta(t)\} = \sum_{k\in\mathbb{Z}} d[k]\delta(t-k)$. Then, the positive definite operator L*L is spline-admissible of order $2r - 1/2$ with symmetric spline generator $\varphi(t) = \beta(t) * \beta(-t)$ such that $\mathrm{L}^*\mathrm{L}\{\varphi(t)\} = \sum_{k\in\mathbb{Z}}(d * d^T)[k]\delta(t-k)$.

Proof: We need to show that all conditions in Definition 2 are satisfied for the operator L*L. Using the inequality

$$\left(\sum_{n\in\mathbb{Z}} \frac{|\omega+2n\pi|^{2\rho}}{1+|\hat{L}(\omega+2n\pi)|^2}\right)^2 \geq \sum_{n\in\mathbb{Z}} \frac{|\omega+2n\pi|^{4\rho}}{(1+|\hat{L}(\omega+2n\pi)|^2)^2} \geq \frac{1}{2}\sum_{n\in\mathbb{Z}} \frac{|\omega+2n\pi|^{4\rho}}{1+|\hat{L}(\omega+2n\pi)|^4}$$

we see that if the left-hand side converges for any $\rho < r - 1/2$, then the right-hand side converges for any $\rho' = 2\rho < r' - 1/2$, where $r' = 2r - 1/2$; i.e., L*L is of order $2r - 1/2$ if L is of order r.

Because $\hat{L}(\omega)$ is a true function and $1/\hat{L}(\omega)$ is a tempered distribution, then $1/|\hat{L}(\omega)|^2$ is tempered as well. Its inverse Fourier transform $\rho_2(t)$ thus satisfies $\mathrm{L}^*\mathrm{L}\{\rho_2(t)\} = \delta(t)$. Next, we evaluate $\mathrm{L}^*\mathrm{L}\{\varphi(t)\} = \mathrm{L}\{\beta(t)\} * \mathrm{L}^*\{\beta(-t)\} = \left(\sum_{k\in\mathbb{Z}} d[k]\delta(t-k)\right) * \left(\sum_{k'\in\mathbb{Z}} d[k']\delta(t+k')\right) = \sum_{k\in\mathbb{Z}}(d * d^T)[k]\delta(t-k)$, which proves that $\varphi(t)$ is an L*L-spline. The corresponding localization operator $(d^T * d)$ is symmetric and guaranteed to be in ℓ_1 as long as $d \in \ell_1$.

The spline generator $\varphi(t)$ is L_p-stable, as a direct consequence of Proposition 2. L_p-stability, in particular, ensures that the upper Riesz bound in (3) is well-defined. The existence of the lower Riesz for $\beta(t)$ implies that $\sum_{k\in\mathbb{Z}} |\hat{\beta}(\omega+2k\pi)|^2$ is nonvanishing for all $\omega \in [0, 2\pi]$. By writing this sum for a given ω_0 as $\sum_{k\in\mathbb{Z}} |a_k|^2 > 0$, we deduce that $\sum_{k\in\mathbb{Z}} |a_k|^4 > 0$ because there must always be at least one a_k that is nonzero. This implies that the lower Riesz bound for $\varphi(t)$ is strictly greater than zero, which completes the proof. ■

Now that we have identified an L_p stable B-spline-like basis for the cardinal L*L-splines, we want to use this representation to determine the spline function $s_{\mathrm{int}}(t) \in V(\varphi)$ that interpolates a given discrete input signal $\{s[k]\}_{k\in\mathbb{Z}}$. It turns out that such a spline interpolant always exists and that its B-spline coefficients can be computed by simple digital filtering of the input data.

Proposition 3: Let $V(\varphi) = \mathrm{span}\{\varphi(t-k)\}_{k\in\mathbb{Z}}$ be the cardinal spline space associated with the symmetric operator L*L, where L is spline-admissible, as in Proposition 2. Then, the problem of interpolating a bounded sequence $s[k]$ in $V(\varphi)$ has a unique and well-defined solution. The interpolating spline is given by

$$s_{\mathrm{int}}(t) = \sum_{k\in\mathbb{Z}} (h_{\mathrm{int}} * s)[k]\varphi(t-k)$$

where $h_{\mathrm{int}} \in \ell_1$ is the impulse response of the digital filter whose transfer function is

$$H_{\mathrm{int}}(z) = \frac{1}{\sum_{k\in\mathbb{Z}} \varphi(k)z^{-k}}.$$

Moreover, we have the following equivalence:

$$s[k] \in \ell_2 \Longleftrightarrow s_{\mathrm{int}} \in W_2^{\mathrm{L}}. \tag{6}$$

Proof: To establish this result, it suffices to note that there is an exact correspondence between the integer samples of $\varphi(t)$ and the Gram sequence of the Riesz basis generated by $\beta(t)$:

$$b[k] = \varphi(k) = \langle \beta(t), \beta(t+k)\rangle_{L_2}. \tag{7}$$

This relation allows us to write the discrete-time Fourier transform of b as follows:

$$B_{\mathrm{L}}(e^{j\omega}) = \sum_{k\in\mathbb{Z}} b[k]e^{-j\omega k} = \sum_{k\in\mathbb{Z}} |\hat{\beta}(\omega+2k\pi)|^2 \tag{8}$$

which is positive and greater than the lower Riesz bound $A_\beta > 0$. Since $b \in \ell_1$ (as a consequence of L_p-stability) and $B_{\mathrm{L}}(e^{j\omega})$ is nonvanishing, we can apply Wiener's Lemma to show that its convolution inverse h_{int} is in ℓ_1. Finally, we verify that the solution is interpolating by resampling $s_{\mathrm{int}}(t)$ at the integers, which amounts to evaluating the discrete convolution relation: $s_{\mathrm{int}}(k) = (h_{\mathrm{int}} * s * b)[k] = s[k]$ (because h_{int} is the convolution inverse of b).

To prove (6), we evaluate the quantity $\|s_{\mathrm{int}}\|^2_{W_2^{\mathrm{L}}} := \|s_{\mathrm{int}}\|^2_{L_2} + \|\mathrm{L}s_{\mathrm{int}}\|^2_{L_2}$:

$$\begin{aligned}\|s_{\mathrm{int}}\|^2_{W_2^{\mathrm{L}}} &= \frac{1}{2\pi}\int |\hat{s}_{\mathrm{int}}(\omega)|^2(1+|\hat{L}(\omega)|^2)d\omega \\ &= \frac{1}{2\pi}\int \frac{|S(e^{j\omega})|^2}{|B_{\mathrm{L}}(e^{j\omega})|^2} \cdot \left(|\hat{\varphi}(\omega)|^2 + |\Delta_{\mathrm{L}}(e^{j\omega})|^2\hat{\varphi}(\omega)\right) d\omega \\ &= \frac{1}{2\pi}\int_0^{2\pi} \frac{|S(e^{j\omega})|^2}{|B_{\mathrm{L}}(e^{j\omega})|^2} \cdot \left(C_{\mathrm{L}}(e^{j\omega}) + |\Delta_{\mathrm{L}}(e^{j\omega})|^2 B_{\mathrm{L}}(e^{j\omega})\right) d\omega\end{aligned}$$

where $S(e^{j\omega}) = \sum_{k\in\mathbb{Z}} s[k]e^{-j\omega k}$, and $C_{\mathrm{L}}(e^{j\omega}) = \sum_{k\in\mathbb{Z}} |\hat{\varphi}(\omega + 2k\pi)|^2$. Because φ and β satisfy the Riesz basis condition and because $|\Delta_{\mathrm{L}}(e^{j\omega})|$ is bounded (since $d \in \ell_1$), there exist two constants $0 < U \leq V < \infty$ such that

$$U\frac{1}{2\pi}\int_0^{2\pi} |S(e^{j\omega})|^2 d\omega \leq \|s_{\mathrm{int}}\|^2_{W_2^{\mathrm{L}}} \leq V\frac{1}{2\pi}\int_0^{2\pi} |S(e^{j\omega})|^2 d\omega.$$

This means that $\|s_{\mathrm{int}}\|_{W_2^{\mathrm{L}}}$ is bounded from below and from above by Const $\times \|s[k]\|_{\ell_2}$. ∎

Proposition 3 implies that there is a one-to-one mapping between the spline interpolant $s_{\mathrm{int}}(t)$ and the discrete samples of the signal $s[k] = s_{\mathrm{int}}(k)$. This mapping is not restricted to ℓ_2-sequences; it is also valid when $s[k] \in \ell_\infty$, which is the most general setting for doing discrete signal processing. The spline coefficients $c[k]$ are obtained by filtering the discrete input signal $f[k]$ with the interpolation filter h_{int}, which is guaranteed to be BIBO-stable. The representation is reversible since $f[k] = (b * c)[k]$, where b is the sampled version of the generalized B-spline.

When $\varphi(t)$ is compactly supported, the interpolation can be implemented quite efficiently by using the recursive filtering procedure described in Appendix II; see also [34]. To take full advantage of this type of algorithm, we need to have an explicit time-domain expression for the spline-generating function $\varphi(t)$. This is possible when the spline-defining operator has a rational transfer function $L_{\vec{\alpha}}(s) = \left(\prod_{n=1}^{N}(s-\alpha_n)/\prod_{m=1}^{M}(s-\gamma_m)\right)$ with $r = N - M \geq 1$. In this case, $\varphi(t)$ is a rescaled and recentered exponential B-spline with parameters $(\vec{\alpha} : -\vec{\alpha}^*) = (\alpha_1, \ldots, \alpha_N, -\alpha_1^*, \ldots, -\alpha_N^*; \gamma_1, \ldots, \gamma_M, -\gamma_1^*, \ldots, -\gamma_N^*)$:

$$\varphi_{\vec{\alpha}}(t) = \beta_{\vec{\alpha}}(t) * \beta^*_{\vec{\alpha}}(-t) \tag{9}$$

$$= (-1)^M \left(\prod_{n=1}^{N} e^{\alpha_n^*}\right) \beta_{(\vec{\alpha}:-\vec{\alpha}^*)}(t+N) \tag{10}$$

which can be determined explicitly using the generalized spline formulas given in [2]. This is a symmetric function of class C^{2r-2} that is supported in $[-N, N]$. The corresponding Fourier domain formula is

$$\hat{\varphi}_{\vec{\alpha}}(\omega) = |\hat{\beta}_{\vec{\alpha}}(\omega)|^2 = \prod_{n=1}^{N} \frac{|1 - e^{\alpha_n} e^{-j\omega}|^2}{|j\omega - \alpha_n|^2} \prod_{m=1}^{M} |j\omega - \gamma_m|^2 \tag{11}$$

in which we can also identify the transfer function of the localization filter $\Delta_{\mathrm{L}}(z)\Delta^*_{\mathrm{L}}(z^{-1}) = \prod_{n=1}^{N}(1-e^{\alpha_n}z^{-1})(1-e^{\alpha_n^*}z)$.

When the underlying L-spline basis $\{\beta(t-k)\}_{k\in\mathbb{Z}}$ is orthonormal, the interpolation algorithm becomes trivial since the filter h_{int} reduces to the identity (because the evaluation of (7) yields $b[k] = \delta_k$). This happens, for example, when $\beta(t)$ is a Haar function (B-spline of degree 0) or a Daubechies scaling function of order N [35]. The corresponding L*L interpolators are the second-order piecewise linear interpolator, and the $2N$th-order Dubuc–Deslaurier interpolators whose basis functions are the autocorrelation of a Daubechies scaling function [36].

IV. Variational Splines and Algorithms

We have just seen that it is always possible to interpolate a data sequence with an L*L-spline and that this process establishes a one-to-one mapping between the discrete and continuous-time domain representations of a signal. In this section, we show that this kind of interpolation can also be justified theoretically based on an energy minimization principle (variational formulation). We also consider a variation of the data fitting problem that relaxes the interpolation constraint and is therefore better suited for dealing with noisy signals. This is achieved by introducing a regularization constraint, which leads to a class of nonparametric estimation algorithms, referred to as "smoothing splines."

A. Energy Minimization Properties

Theorem 2: Let L be a spline-admissible operator of order $r > 1/2$. Then, we have the remarkable norm identity

$$\forall\, f \in W_2^{\mathrm{L}}, \quad \|\mathrm{L}f\|^2_{L_2} = \|\mathrm{L}s_{\mathrm{int}}\|^2_{L_2} + \|\mathrm{L}\{f - s_{\mathrm{int}}\}\|^2_{L_2}$$

where $s_{\mathrm{int}}(t)$ is the unique L*L-spline that interpolates $f(t)$, i.e., $f(k) = s_{\mathrm{int}}(k), \forall\, k \in \mathbb{Z}$, as specified in Proposition 3.

Proof: We will prove that if $g \in W_2^{\mathrm{L}}$ such that $g(k) = 0$ for $k \in \mathbb{Z}$, then the inner product between $\mathrm{L}s_{\mathrm{int}}$ and $\mathrm{L}g$ vanishes. Note that $s_{\mathrm{int}} \in W_2^{\mathrm{L}}$ because $s_{\mathrm{int}}[k]$—the sampled version of $f(t)$—is in ℓ_2 (as a consequence of (6) and *iii* of Theorem 1). Moreover, because $g \in W_2^{\mathrm{L}}$, $\|\sum_{n\in\mathbb{Z}} \hat{g}(\omega + 2n\pi)\|_{L_2([0,2\pi])} = 0$ (see *ii* of Theorem 1). With the same notations as for the proof of Proposition 3, we have

$$\begin{aligned}
\langle \mathrm{L}s_{\mathrm{int}}, \mathrm{L}g\rangle &= \frac{1}{2\pi}\int \frac{S(e^{j\omega})}{B_{\mathrm{L}}(e^{j\omega})} |\hat{L}(\omega)|^2 \varphi(\omega)\hat{g}(\omega) d\omega \\
&= \frac{1}{2\pi}\int \frac{S(e^{j\omega})|\Delta_L(e^{j\omega})|^2}{B_L(e^{j\omega})}\hat{g}(\omega) d\omega \\
&= \lim_{N\to\infty} \frac{1}{2\pi}\int_0^{2\pi} \frac{S(e^{j\omega})|\Delta_{\mathrm{L}}(e^{j\omega})|^2}{B_{\mathrm{L}}(e^{j\omega})} \cdot \sum_{|n|\leq N} \hat{g}(\omega + 2n\pi) d\omega \\
&\leq \underbrace{\frac{1}{2\pi}\left\| \frac{S(e^{j\omega})|\Delta_{\mathrm{L}}(e^{j\omega})|^2}{B_{\mathrm{L}}(e^{j\omega})} \right\|_{L_2([0,2\pi])}}_{\text{bounded}} \cdot \underbrace{\lim_{N\to\infty} \left\| \sum_{|n|\leq N} \hat{g}(\omega + 2n\pi) \right\|_{L_2([0,2\pi])}}_{=0}
\end{aligned}$$

where the first term on the right-hand side is bounded because $S(e^{j\omega}) = \sum_{n\in\mathbb{Z}} f[k]e^{-j\omega k}$ is in $L_2([0, 2\pi])$; $|\Delta_L(e^{j\omega})|$ is upper bounded (because $d \in \ell_1$); and $B_{\mathrm{L}}(e^{j\omega})$ is lower bounded (Riesz basis condition).

Thus, we have $\|\mathrm{L}g + \mathrm{L}s_{\mathrm{int}}\|^2_{L_2} = \|\mathrm{L}s_{\mathrm{int}}\|^2_{L_2} + 2\langle \mathrm{L}s_{\mathrm{int}}, \mathrm{L}g\rangle + \|\mathrm{L}g\|^2_{L_2} = \|\mathrm{L}s_{\mathrm{int}}\|^2_{L_2} + \|\mathrm{L}g\|^2_{L_2}$, which provides the result of the theorem when we make $g = f - s_{\mathrm{int}}$. ∎

Corollary 1: Let $f[k] \in \ell_2$ be a discrete input signal and L be a spline-admissible operator of order $r > 1/2$. Among all possible interpolating functions $f(t) \in W_2^{\mathrm{L}}$, the optimal one

that minimizes $\|\mathrm{L}f\|_{L_2}$ subject to the interpolation constraint $f[k] = f(t)$ is the L*L-spline interpolant $s_{\mathrm{int}}(t)$ specified in Proposition 3.

Indeed, for any interpolating function $f(t)$, the quadratic norm $\|\mathrm{L}f\|_{L_2}^2$ can be written as the sum of two positive terms where the first one is fixed and uniquely determined by the input samples $f[k]$. Therefore, the optimal solution among all possible interpolators is $f(t) = s_{\mathrm{int}}(t)$, which makes the second term vanish.

Corollary 1 expresses a well-known property of splines that goes back to the pioneering work of Schoenberg [37]. For L= D^2, the optimal interpolant is the cubic spline with $\mathrm{L}^*\mathrm{L} = \mathrm{D}^4$. Since $\mathrm{D}^2 f$ is a good approximation of the curvature of $f(t)$, the cubic spline interpolant is often said to have the "minimum curvature" property; in fact, this is the argument[2] that was used by Schoenberg for calling these functions "splines".

The energy minimization property of splines in Theorem 2 constitutes a variation on a well-known theme in spline theory [8]. It is a complementary result because the traditional formulation usually considers a finite number of data points on a nonuniform grid. Since our formulation is restricted to the cardinal framework, we are obviously giving up on some level of generality. However, there are some positive compensations: First, the result in Theorem 2 is also applicable to certain nonlocal differential operators, such as those with rational transfer functions, which are not covered by the standard theory. Second, the cardinal framework is ideally suited for signal processing because it allows for very efficient digital-filter-based solutions that are not available otherwise. Last, but not least, the present formulation leads to important theoretical simplifications because the derivation are entirely Fourier-based; in particular, this keeps us from having to worry about the technicalities associated with the part of the solution that is in the null space of the operator L*L.

B. Smoothing Splines

When the input data $\{f[k]\}_{k\in\mathbb{Z}}$ is corrupted by noise, it can be counterproductive to determine its exact spline fit. Instead, one would rather seek a solution that is close to the data but has some inherent smoothness to counterbalance the effect of the noise. There are two possible ways to look at this problem within our variational framework: The first is to say that we want an approximate fit where the magnitude of the approximation error is fixed (i.e., $\xi_{\mathrm{data}} := \sum_{k\in\mathbb{Z}} [f[k] - s(k)]^2 = C_1$, where the constant C_1 is typically set to a fraction of the noise variance) and where all remaining degrees of freedom are taken care of by searching for the continuous-time solution $s(t)$ that minimizes $\xi_{\mathrm{L}} := \|\mathrm{L}s\|_{L_2}^2$. Alternatively, we may assume that we have some *a priori* knowledge on the solution, which is given in the form of an upper bound on the spline energy functional (i.e., $\xi_{\mathrm{L}} \leq C_2$, where C_2 needs to be smaller than the energy of the spline interpolant), and search for the solution that minimizes the quadratic fitting error ξ_{data}, subject to this constraint. In general, when the data is noisy and L is of order r, any attempt to decrease the fitting error ξ_{data} will increase the spline energy ξ_{L}, and vice versa. This means that both formulations are qualitatively equivalent because they attempt to achieve the right balance between having a solution that is close to the data (i.e., ξ_{data} small) and one that is smooth in the sense that ξ_{L} is reasonably small. It turns out that the first problem is mathematically equivalent to the second one. The solution can be determined by using the method of Lagrange multipliers; in both cases, this boils down to the minimization of the functional $\xi_{\mathrm{data}} + \lambda\xi_{\mathrm{L}}$, where λ is a free scalar parameter that is selected such as to fulfill the constraint. A remarkable result is that the solution of this approximation problem, among all possible continuous-time functions $s(t)$, is a cardinal L*L-spline and that it can be determined by digital filtering.

Theorem 3: Let L be a spline-admissible operator of order $r > 1/2$ with spline generator $\beta(t)$ such that $\mathrm{L}\{\beta(t)\} = \Delta_L\{\delta(t)\} = \sum_{k\in\mathbb{Z}} d[k]\delta(t-k)$. Then, the continuous-time solution of the variational problem with discrete input data $f[k] \in \ell_2$ and regularization parameter $\lambda \geq 0$

$$\min_{s(t)\in W_2^L} \sum_{k\in\mathbb{Z}} (f[k] - s(k))^2 + \lambda\|\mathrm{L}\{s(t)\}\|_{L_2}^2 \tag{12}$$

is a unique and well-defined cardinal L*L-spline. It is given by

$$s_\lambda(t) = \sum_{k\in\mathbb{Z}} (h_\lambda * f)[k]\varphi(t-k) \tag{13}$$

where $\varphi(t) = \beta(t) * \beta(-t)$ and where $h_\lambda \in \ell_1$ is the impulse response of the digital smoothing spline filter whose transfer function is

$$H_\lambda(z) = \frac{1}{\sum_{k\in\mathbb{Z}} \varphi(k)z^{-k} + \lambda\Delta_{\mathrm{L}}(z)\Delta_{\mathrm{L}}(z^{-1})}. \tag{14}$$

Proof: Since the cost function is quadratic in s, the problem is guaranteed to have a solution. Using Theorem 2, we write the criterion to minimize as

$$\xi = \sum_{k\in\mathbb{Z}} [f[k] - s(k)]^2 + \lambda\|\mathrm{L}\{s_{\mathrm{int}}(t)\}\|_{L_2}^2 + \lambda\|\mathrm{L}\{s(t) - s_{\mathrm{int}}(t)\}\|_{L_2}^2$$

where $s_{\mathrm{int}}(t)$ is the L*L-spline interpolator of the sequence $s(k)$. The left-most data term is entirely specified by the integer samples $s(k)$. If we assume that the $s(k)$'s are the samples of the optimal solution, we can further minimize the regularization part of the criterion by selecting the solution $s(t) = s_{\mathrm{int}}(t)$ that sets the third positive term to zero. This proves the first part of the theorem; namely, that the optimal solution is a cardinal L*L-spline.

To determine this spline solution, we substitute its B-spline expansion (4) into the criterion. Using the fact that $\|\mathrm{L}\{s\}\|_{L_2}^2 = \langle \mathrm{L}^*\mathrm{L}\{s\}, s\rangle_{L_2}$ together with the property $\mathrm{L}^*\mathrm{L}\{\varphi\} = \sum_{k\in\mathbb{Z}} (d^T * d)[k]\delta(t-k)$, we manipulate ξ to express its explicit dependency on $c[k]$

$$\begin{aligned}\xi &= \|f - b*c\|_{\ell_2} + \lambda\|\mathrm{L}\{s(t)\}\|_{L_2}^2 \\ &= \langle f - b*c, f - b*c\rangle_{\ell_2} + \lambda\langle d^T * d * c, b * c\rangle_{\ell_2} \\ &= \langle f, f\rangle_{\ell_2} - 2\langle b^T * f, c\rangle_{\ell_2} + \langle b^T * b * c, c\rangle_{\ell_2} \\ &\quad + \lambda\langle b^T * d^T * d * c, c\rangle_{\ell_2}\end{aligned}$$

[2]According to the American Heritage dictionary, a spline is a thin rod that was used for drawing smooth curves.

which now involves discrete convolution operators and ℓ_2-inner products only; b is the digital convolution kernel defined by (7). The optimal solution is determined by setting the partial derivative of ξ with respect to $c[k]$ to zero, which yields

$$\frac{\partial \xi}{\partial c[k]} = (-2b^T * f + 2b^T * b * c + 2\lambda b^T * d^T * d * c)[k] = 0.$$

Using the fact that the symmetric convolution operator $b^T = b$ is invertible, we get the discrete convolution relation

$$f[k] = (b * c)[k] + \lambda(d^T * d * c)[k] = \left((b + \lambda d^T * d) * c\right)[k].$$

This equation can be solved by going into the discrete-time Fourier domain

$$F(e^{j\omega}) = \left(B_{\mathrm{L}}(e^{j\omega}) + \lambda|\Delta_{\mathrm{L}}(e^{j\omega})|^2\right) \cdot C(e^{j\omega})$$

where $B_{\mathrm{L}}(e^{j\omega})$ is given by (8) and is guaranteed to be strictly positive. Hence, the filter $B_{\mathrm{L}}(e^{j\omega}) + \lambda|\Delta_{\mathrm{L}}(e^{j\omega})|^2$, which is made up of positive terms only, cannot vanish and can be safely inverted, irrespective of the value of $\lambda \geq 0$, to yield the smoothing spline filter described by (14). ■

The optimal solution $s_\lambda(t)$ in Theorem 3 is called a smoothing spline. By adjusting the regularization parameter λ, we can control the amount of smoothing. When $\lambda = 0$, there is no smoothing at all, and the solution coincides with the spline interpolant specified in Proposition 3. For larger values of λ, the smoothing kicks in and tends to attenuate high-frequency components because the localization filter has qualitatively the same rth-order behavior as the operator: $\Delta_{\mathrm{L}}(e^{j\omega}) \approx \omega^r$ (cf. argument in [2]). In the limit, when $\lambda \to +\infty$, the only signal components left are the ones in the null space of the operator L, for instance, the best fitting polynomial of degree n when $\mathrm{L} = \mathrm{D}^n$ (polynomial spline case). Note that when $\hat{L}(\omega)$ is rational, the smoothing spline filter h_λ can be implemented quite efficiently using the recursive algorithm that is given in Appendix II.

The key practical question is how to select the most suitable operator L and the optimal value of λ for the problem at hand. While this can be done empirically along the lines of what has been exposed at the beginning of this section, it can also be approached in a rigorous statistical way by introducing a stochastic model for the signal. In Section V, we will establish a formal equivalence between the present smoothing spline algorithm and MMSE-estimation for stationary processes, which will allow us to select these parameters in an optimal fashion.

To illustrate the smoothing spline concept, we now consider the case of the first-order differential operator $\mathrm{L} = \mathrm{D} - \alpha\mathrm{I}$ with $\alpha \in R$. Its causal Green function is $\rho_\alpha(t) = 1_+(t) \cdot e^{\alpha t}$. We localize this function by applying to it the digital filter $\Delta_\alpha(z) = 1 - e^\alpha z^{-1}$, which yields the first order exponential B-spline: $\beta_\alpha(t) = \rho_\alpha(t) - e^\alpha \rho_\alpha(t-1)$, as described in [1]. The generating function for the optimal smoothing spline in Theorem 3

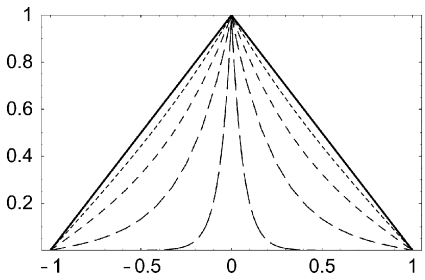

Fig. 1. Optimal interpolation kernels associated with the first-order differential operators $LL = \mathrm{D}-\alpha\mathrm{I}$. The corresponding parameter values from top to bottom are $\alpha = 0$ (piecewise linear interpolator) and $\alpha = \pm 1, \pm 2, \pm 4, \pm 16$. The magnitude of α controls the degree of concavity of the curve and its peakedness at the origin, while the sign has no influence.

corresponds to the autocorrelation function of $\beta_\alpha(t)$, which is given by

$$\varphi(t) = \beta_\alpha(t) * \beta_\alpha(-t) = \begin{cases} \varphi(0) \cdot \frac{e^{\alpha|t|} - e^{2\alpha}e^{-\alpha|t|}}{1-e^{2\alpha}}, & \text{for } t \in [-1,1] \text{ and } \alpha \neq 0 \\ 1 - |t|, & \text{for } t \in [-1,1] \text{ and } \alpha = 0 \\ 0, & \text{for } t \notin [-1,1] \end{cases} \tag{15}$$

with

$$\varphi(0) = \begin{cases} \frac{e^{2\alpha}-1}{2\alpha}, & \text{for } \alpha \neq 0 \\ 1, & \text{for } \alpha = 0. \end{cases} \tag{16}$$

These functions are continuous, symmetric, and supported in $[-1, 1]$. The corresponding interpolation kernels are obtained by dividing these B-splines by $\varphi(0)$. Some examples are shown in Fig. 1.

The B-spline coefficients of the smoothing spline are obtained by filtering the data with the digital filter whose transfer function is

$$H_{\alpha,\lambda}(z) = \frac{1}{\varphi(0) + \lambda(1 - e^\alpha z)(1 - e^\alpha z^{-1})}. \tag{17}$$

At the end of the process, the integer samples of the smoothing spline are recovered by convolving the computed B-spline coefficients with b (the sampled B-spline kernel). In the first-order case that has just been considered, this amounts to a simple scalar multiplication by $\varphi(0)$. In practice, it makes good sense to incorporate this rescaling into the smoothing spline filter, which is equivalent to considering an expansion in terms of renormalized basis functions $\varphi(t)/\varphi(0)$, such as the ones displayed in Fig. 1.

This smoothing spline filter can be implemented quite efficiently from a cascade of first-order causal and anticausal operators. It is a filter that is commonly used for image processing and goes under the name of the Shen–Castan detector. As far as we know, this type of exponential filtering has not yet been cataloged as a smoothing spline estimator, expect for the piecewise linear case ($\alpha = 0$), which is investigated in [25].

C. Splines Under Tension and Generalizations

In some applications, it can be desirable to have more free parameters for the specification of the regularization functional. For instance, one may consider a linear combination of energies

associated with derivatives of increasing order. It turns out that this type of problem can be solved in essentially the same way as the previous one.

Corollary 2: Let L_i, $i = 1, \ldots, n$ be a series of spline-admissible operators of order $r_i \geq 1$ with stable spline generators $\beta_i(t)$ such that $\mathrm{L}_i\{\beta_i(t)\} = \Delta_{\mathrm{L}_i}\{\delta(t)\} = \sum_{k\in\mathbb{Z}} d_i[k]\delta(t-k)$ with $d_i \in \ell_1$. Then, the continuous-time solution of the variational problem with discrete input data $f[k] \in \ell_2$ and regularization parameters $\lambda_i \geq 0$

$$\min_{s(t)\in\cap_{i=1}^n W_2^{\mathrm{L}_i}} \sum_{k\in\mathbb{Z}} (f[k]-s(k))^2 + \sum_{i=1}^n \lambda_i \|\mathrm{L}_i\{s(t)\}\|_{L_2}^2$$

is a $(\sum \lambda_i \mathrm{L}_i^*\mathrm{L}_i)$-spline, which is unique and well-defined. It is given by

$$s(t) = \sum_{k\in\mathbb{Z}} c[k]\varphi(t-k),$$

where $\varphi(t) = \sum_{i=1}^n \lambda_i\,(\beta_i(t) * \beta_i(-t))$; the B-spline coefficients $c[k]$ are determined by digitally filtering the input data $f[k]$ with the filter $h_{(\lambda_1,\ldots,\lambda_n)}$, whose transfer function is

$$H_{(\lambda_1,\ldots,\lambda_n)}(z) = \frac{1}{\sum\limits_{k\in\mathbb{Z}} \varphi(k)z^{-k} + \sum\limits_{i=1}^n \lambda_i \Delta_{\mathrm{L}_i}(z)\Delta_{\mathrm{L}_i}(z^{-1})}.$$

Proof: We bring this problem back to the previous one by using the linearity of the L_2-inner product and noticing that

$$\sum_{i=1}^n \lambda_i \|\mathrm{L}_i\{s\}\|_{L_2}^2 = \sum_{i=1}^n \lambda_i \langle \mathrm{L}_i^*\mathrm{L}_i\{s\}, s\rangle = \left\langle \sum_{i=1}^n \lambda_i \mathrm{L}_i^*\mathrm{L}_i\{s\}, s \right\rangle.$$

Thus, we may consider an equivalent operator L such that $\mathrm{L}^*\mathrm{L} = \sum_{i=1}^n \lambda_i \mathrm{L}_i^*\mathrm{L}_i$. The simplest way to get such an L is by taking a square-root in the Fourier domain. When the $\hat{L}_i(\omega)$'s are rational transfer functions that satisfy the Paley–Wiener condition, we can apply a standard spectral factorization technique to determine a solution that corresponds to a causal operator. The operator L will be associated with a spline-generating function $\beta(t)$ whose Gram sequence is the inverse Fourier transform of $A(e^{j\omega}) = \sum_{n\in\mathbb{Z}} |\hat{\beta}(\omega + 2\pi n)|^2 = \sum_{n\in\mathbb{Z}} \sum_{i=1}^n \lambda_i |\hat{\beta}_i(\omega + 2\pi n)|^2$. Because of the stability assumption on each individual components, we have that $0 < \min\{A_i\} \leq A(e^{j\omega}) \leq \sum_{i=1}^n B_i < +\infty$, where A_i and B_i are the lower and upper Riesz bounds of the individual spline-generators $\beta_i(t)$. This guarantees that L—irrespective of whether it is causal or not—is a spline-admissible operator; the result then follows directly from the application of Theorem 3. ∎

The well-known splines-under-tension, initially introduced by Schweikert, correspond to the choice $\mathrm{L}_1 = \mathrm{D}$ and $\mathrm{L}_2 = \mathrm{D}^2$ [38]. The present formalism provides a most efficient way to implement this smoothing method on a uniform grid.

V. Application to Stochastic Signal Processing

In this section, we propose two equivalent stochastic interpretations of the deterministic spline-fitting methods that have been considered so far. The main difficulty here is that the integer samples of a stationary process $x(t)$ defined over the entire real line are not in ℓ_2, which has the unfortunate consequence of making the variational quantities considered in Section IV ill-defined. First, we will bypass this difficulty by concentrating on the task of determining the minimum mean square estimate (MMSE) of the process at a given time t_0. We will then propose an alternative Bayesian formulation that is restricted to the periodic case where the various energy terms can be renormalized. In either case, we will see that the general spline-fitting algorithms that have been presented so far are optimal for the estimation of a wide family of stationary signals. The relevant class of stationary processes are the *regular* ones (cf. [39]) whose whitening operator is spline-admissible. The advantage of the present formulation is that it yields a proper discretization and a recursive implementation of the corresponding Wiener filter for stationary signals corrupted by additive white noise.

A. Stationary Processes: Innovations and Whitening Filters

Here, we will consider the case of a continuous-time signal $x(t)$ that is a realization of a wide sense stationary process. The process is assumed to have zero mean, $E\{x(t)\} = 0$, $\forall\, t \in \mathbb{R}$, and is characterized by its autocorrelation function $c_{xx}(\tau) = E\{x(t)x(t+\tau)\} \in L_2$. Its spectral power density is given by $C_{xx}(\omega) = \mathcal{F}\{c_{xx}(\tau)\} = \int_{-\infty}^{+\infty} c_{xx}(\tau)e^{-j\omega\tau}d\tau$ [39]. Furthermore, we assume that the process is *regular* in the sense that there exists a whitening filter L—which is generally not BIBO stable but whose impulse response is a tempered distribution—that transforms $x(t)$ into an innovations signal $i(t) = L\{x(t)\}$, which is completely uncorrelated. The whitening filter is the inverse of the innovations filter H that transforms $w(t)$ back into $x(t) = H\{w(t)\}$.

The whitening property implies that $E\{w(t)w(t+\tau)\} = \mathrm{L}^*\mathrm{L}\{c_{xx}(\tau)\} = \sigma_0^2 \cdot \delta(\tau)$. If, in addition, L is spline-admissible, this is equivalent to saying that $c_{xx}(\tau)$ can be represented in terms of the B-spline basis functions associated with the operator $\mathrm{L}^*\mathrm{L}$. Thus, we can write that

$$c_{xx}(\tau) = \sigma_0^2 \sum_{k\in\mathbb{Z}} p[k]\varphi(\tau - k)$$

where p is the discrete convolution inverse of the localization operator $d * d^T$.

B. MMSE Estimation and Spline Interpolation

The next result provides a strong statistical justification for the use of $\mathrm{L}^*\mathrm{L}$-spline interpolation.

Theorem 4: Let $x(k)$ be the samples of a realization of a continuous-time stationary process whose whitening filter L is spline-admissible. Then, the linear MMSE estimator of $x(t)$ at time $t = t_0$, given the samples $\{x(k)\}_{k\in\mathbb{Z}}$, is $s_{\mathrm{int}}(t_0)$, where $s_{\mathrm{int}}(t)$ is the $\mathrm{L}^*\mathrm{L}$-spline interpolator of $\{x[k]\}_{k\in\mathbb{Z}}$, as specified in Proposition 3.

Proof: The goal is to determine the linear estimator $\tilde{x}(t_0) = \sum_{k\in\mathbb{Z}} a_{t_0}[k]x(k)$ such that the mean square error $E\left\{|x(t_0) - \tilde{x}(t_0)|^2\right\}$ is minimized. The estimator is specified by the bi-infinite sequence of regression coefficients $a_{t_0}[k]$, which are collected in the vector $\mathbf{a}$. By applying the orthogonality principle $E\{x(l)\,(x(t_0) - \tilde{x}(t_0))\} = 0, \forall\, l \in \mathbb{Z}$, it is not difficult to show that the optimal regression coefficients are the solution of the corresponding Yule–Walker equations:

$$\mathbf{R}\cdot\mathbf{a} = \mathbf{r}$$

where $\mathbf{R}$ is an infinite Toeplitz matrix whole entries are $[\mathbf{R}]_{k,l} = E\{x(k)x(l)\} = c_{xx}(k-l)$ and where $\mathbf{r}$ is the infinite correlation vector whose lth entry is $[\mathbf{r}]_l = c_{xx}(t_0 - l)$. This system of equations can also be written as a discrete convolution equation

$$(p * b * a_{t_0})[k] = p[k] * \varphi(t_0 - k).$$

By convolving this equation with $d * d^T$ (the BIBO-stable convolution inverse of p), we find that the optimal coefficients are given by

$$a_{t_0}[k] = h_{\text{int}}[k] * \varphi(t_0 - k)$$

where $h_{\text{int}}[k]$ is the convolution inverse of $b[k]$. Thus, the final form of the MMSE estimator is

$$\tilde{x}(t_0) = \sum_{k\in\mathbb{Z}} c[k]\varphi(t_0 - k)$$

where the B-spline coefficients are obtained as $c[k] = (h_{\text{int}} * x)[k]$. ∎

Thus, we have a simple, efficient spline interpolation algorithm that provides the linear MMSE estimator of a stationary signal at noninteger locations. If, in addition, we assume that the process is Gaussian, then the spline interpolation corresponds to the conditional mean that is the optimum solution among all estimators (including the nonlinear ones). Interestingly, we note that the result remains valid when the autocorrelation function $c_{xx}(\tau)$ is a cardinal L*L-spline, which is less restrictive than the requirement that L is the whitening filter of the process.

C. MMSE Estimation and Smoothing Splines

We consider a similar estimation task but in a noisy situation. The goal is now to estimate a realization $x(t)$ of the process given some noisy measurements $y[k] = x(k) + n[k]$, where $n[k]$ is some additive, signal-independent white noise. Here, too, there is a direct connection with our previous variational formulation: The optimal solution is provided by the smoothing spline algorithm described in Section IV-B.

Theorem 5: Let $x(t)$ be a realization of a continuous-time wide sense stationary process whose autocorrelation function $c_{xx}(\tau) = E\{x(t)x(t+\tau)\}$ is such that $\text{L*L}\{c_{xx}(\tau)\} = \sigma_0^2 \cdot \delta(\tau)$, where L is spline-admissible. Then, the linear MMSE estimator of $x(t)$ at time $t = t_0$, given the measurements $\{y[k] = x(k) + n[k]\}_{k\in\mathbb{Z}}$, where $n[k]$ is white noise with variance σ^2, is $s_\lambda(t_0)$ with $\lambda = (\sigma^2/\sigma_0^2)$, where $s_\lambda(t)$ is the L*L-smoothing spline fit of $\{y[k]\}_{k\in\mathbb{Z}}$, as specified in the second part of Theorem 3.

Proof: The argument is essentially the same as before, with the linear estimator now being $\tilde{x}(t_0) = \sum_{k\in\mathbb{Z}} a_{t_0}[k]y(k)$. The matrix entries for the corresponding Yule–Walker equations are $[\mathbf{R}]_{k,l} = E\{y[k]y[l]\} = c_{xx}(k-l) + \sigma^2\delta_{k-l}$, which now takes into account the effect of independent additive noise, and $[\mathbf{r}]_l = c_{yx}(t_0 - l) = c_{xx}(t_0 - l)$, which is unchanged. Here, too, the system can be rewritten as a discrete convolution equation

$$\left(\sigma_0^2 \cdot p * b + \sigma^2 \cdot \delta\right)[k] * a_{t_0}[k] = \sigma_0^2 \cdot [k] * \varphi(t_0 - k)$$

where $\delta[k] = \delta_k$ denotes the discrete Kronecker impulse. By convolving each side of this identity with $(d * d^T)$—the convolution inverse of p—and dividing by σ_0^2, we get

$$\left(b + \left(\frac{\sigma^2}{\sigma_0^2}\right) d * d^T\right)[k] * a_{t_0}[k] = \varphi(t_0 - k).$$

Finally, we solve this equation by applying the convolution inverse of the filter on the left-hand side, which yields

$$a_{t_0}[k] = h_\lambda[k] * \varphi(t_0 - k)$$

where the relevant inverse filter is $h_\lambda[k]$, as defined by (14), with $\lambda = (\sigma^2/\sigma_0^2)$. ∎

This result is interesting in two respects. First, it provides us with the optimum regularization parameter $\lambda = (\sigma^2/\sigma_0^2)$ for the smoothing spline algorithm, which is quite valuable. Second, it yields an estimation algorithm that is an optimal discretization of the classical continuous-time Wiener filter. The discrete version of the output signal is recovered by resampling the smoothing spline at the integers, which is equivalent to a digital post-filtering with the sampled version of the basis function (B-spline). Both filters can be combined to yield an equivalent discrete filter Wiener, whose frequency response is

$$H_{\text{Wiener}}(e^{j\omega}) = \frac{B_{\text{L}}(e^{j\omega})}{B_{\text{L}}(e^{j\omega}) + \lambda|\Delta_{\text{L}}(e^{j\omega})|^2} \tag{18}$$

where $B_L(e^{j\omega})$ is defined by (8).

While smoothing splines and conventional Wiener filtering should lead to similar results, the former has a conceptual advantage over the latter because it provides a solution that is valid for any $t_0 \in \mathbb{R}$ and not just the integers. If we were designing a classical Wiener filter, we would have the choice between two options: i) Restrict ourselves to the integer samples and use a purely discrete formulation, or ii) derive the continuous-time solution and discretize the corresponding filter assuming, as one usually does, that the signal is bandlimited. With the present formulation, the answer that we get is more satisfying because we are addressing both issues simultaneously: a) the specification of the optimal signal space and b) the search for the best solution within that space, which leads to a digital filtering solution. In addition, when the spline generator $\varphi(t)$ is compactly supported, the smoothing spline formulation yields a recursive filtering algorithm (cf. Appendix II) that is typically faster than the traditional Wiener filter, which is specified and implemented in the Fourier domain.

As an illustration, we consider the estimation of a first-order Markov signal corrupted by additive white noise. This measurement model is characterized by two parameters: the normalized

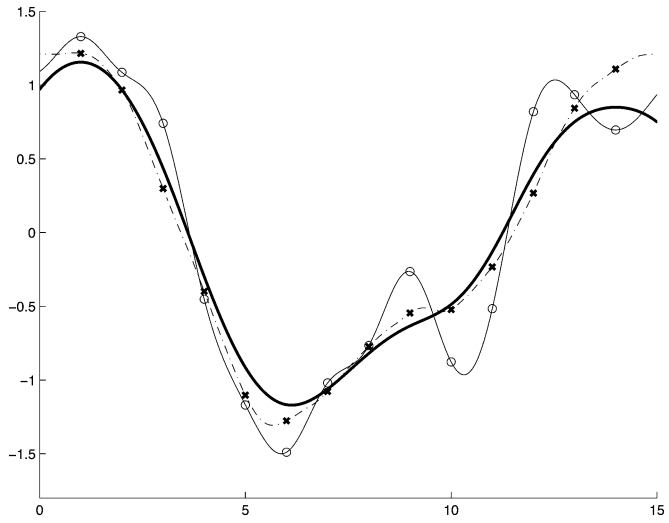

Fig. 2. Example of continuous-time signal estimation from discrete noisy data. The underlying third-order continuous-time process $x(t)$ is displayed in dotted lines, and its noise-free samples are marked by crosses. The noisy input samples are shown as circles. The thin continuous curve is the corresponding interpolating L*L-spline. The optimal smoothing spline (thicker line) is comparatively much closer to the noise-free reference signal.

correlation coefficient $0 < \rho := E\{x(t)x(t+1)\}/E\{x^2\} < 1$ and the quadratic signal-to-noise ratio: SNR $= E\{x^2\}/\sigma^2$. The autocorrelation function of the normalized first-order Markov process is $c_{xx}(\tau) = \rho^{|\tau|}$, and its spectral power density is $C_{xx}(\omega) = ((-2\log\rho)/(\omega^2 + \log^2\rho))$ [39]. We whiten the process by applying the filter $\hat{L}_\alpha(\omega) = j\omega - \alpha$ with $\alpha = \log\rho < 0$, which corresponds to the first-order differential operator D $-$ αI. The variance of the normalized whitened signal is given by $C_{xx}(\omega)\cdot|\hat{L}_\alpha(\omega)|^2 = -2\log\rho > 0$. Thus, the optimal estimator is an exponential spline with parameter $(\alpha, -\alpha)$, and the smoothing spline algorithm can be implemented efficiently via the recursive filtering procedure described at the end of Section IV-B. The optimal smoothing parameters is $\lambda = (-2\log\rho/\text{SNR})$. It can also be checked that the smoothing spline filter $\varphi(0)\cdot H_{\alpha,\lambda}(e^{j\omega})$ defined by (16) and (17) is equivalent to the classical discrete-time Wiener filter that can be specified for the corresponding discrete first-order Markov model. The essential difference here is that the smoothing spline fit also fills in for the noninteger values of the signal.

Some concrete results of another similar signal estimation problem are shown in Fig. 2. The corresponding spectral shaping filter is a third-order allpole system with $(\alpha_1 = -0.001, \alpha_2 = -1, \alpha_3 = -2)$. The noise-free continuous-time process $x(t)$ is represented by a dotted line. The noisy input samples $y[k] = x(k) + n[k]$ are displayed as circles. The corresponding L*L-spline estimators were computed in Matlab using the recursive filtering procedure described in Appendix II. The spline interpolant $s_{\text{int}}(t) = s_0(t)$, which fits the noisy input data exactly, is displayed using a thin continuous line, while the optimal smoothing spline estimator with $\lambda = \sigma^2/\sigma_0^2 = 0.09$ is superimposed using a thicker line. Clearly, $s_\lambda(t)$ is less oscillating than $s_0(t)$ and closer to the noise-free signal.

D. Bayesian Formulation

Following the lead of Wahba and others, it is also tempting to interpret smoothing spline estimation in Bayesian terms. The difficulty, when $x(t)$ is a realization of a stationary process, is that the cost function in (12) is not defined because the input sequence $y[k] = x(k) + n[k]$ is no longer in ℓ_2. We can circumvent the problem by concentrating on the time interval $t \in [0, T]$ and assuming that the process $x(t)$ is T-periodic with T integer. The corresponding normalized cost function with input data $Y = \{y[k], k = 0, \ldots, T-1\}$ is written as

$$\xi(X|Y) = \frac{1}{2\sigma^2}\sum_{k=0}^{T-1}|y[k] - x(k)|^2 + \frac{1}{2\sigma_0^2}\int_0^T |\mathrm{L}\{x(t)\}|^2 dt \tag{19}$$

and its minimization yields the periodic maximum a posteriori (MAP) estimate of the signal $X = \{x(t), t \in [0,T]\}$. This is because the criterion corresponds to the *a posteriori* log-likelihood function of a Bayesian measurement model: $Y|X = X + N$, where the unknown noise (N) and signal random variables are Gaussian-distributed. In the Bayesian framework, the data term of the criterion is $-\log p(Y|X) = -\log p(N)$, which corresponds to the additive white Gaussian noise component, while the regularization term is $-\log p(X)$, which is the log-likelihood prior derived from the signal model. In the present case, the log-likelihood prior corresponds to a periodic stationary signal whose autocorrelation function is the T-periodized version of the one considered in Sections V-A–C. This can be shown by applying a standard argument from time-series analysis: First, we expand the periodic signal $x(t)$ into a Fourier series, noticing that the Fourier coefficients $c_n = (1/T)\int_0^T x(t)e^{-j2\pi n/T}dt$ are Gaussian distributed with zero mean and variance $E\{|c_n|^2\} = \left(\sigma_0^2/|\hat{L}(2\pi n/T)|^2\right)$. Next, we use the property that the Fourier coefficients of a stationnary signal are independent and write the corresponding Gaussian log-likelihood function $\log p(X) = \lim_{N\to\infty}\sum_{n=-N}^{N}|c_n|^2\left(|\hat{L}(2\pi n/T)|^2/2\sigma_0^2\right) + C_N$, where C_N is a signal-independent constant. Finally, we map the result back into the time domain using Parseval's formula, which yields the right-hand side integral in (19) up to a constant offset.

From a mathematical point of view, the minimization of the MAP criterion (19), subject to the periodicity constraint $x(t) = x(t+T)$, is the periodized version of the smoothing spline problem in Theorem 3. This means that the solution can be determined by applying the exact same algorithm (digital filtering with $h_\lambda[k]$) to an augmented T-periodic input signal $\{y[k]\}_{k\in\mathbb{Z}}$. The solution is a T-periodic cardinal L*L-spline, which can be represented as $s(t) = \sum_{k\in\mathbb{Z}} c[k]\varphi(t-k)$ with $c[k] = c[k+T]$ or, equivalently, as $s(t) = \sum_{k=0}^{T-1} c[k]\varphi_T(t-k)$ with $\varphi_T(t) = \sum_{n\in\mathbb{Z}}\varphi(t+nT)$.

VI. Conclusion

We have presented a series of mathematical arguments to justify the use of splines in signal processing applications. They come into two flavors:

1) *Deterministic results:* The L*L-spline interpolator is the optimal solution among all possible interpolators $f(t)$ of a discrete signal in the sense that it minimizes the energy functional $\|\mathrm{L}f\|_{L_2}^2$. This is a result that holds for a relatively general class of generalized differential operators L, including those with rational transfer functions. The same type of property also carries over to the regularized version of the problem that involves a quadratic data term. The corresponding solution defines the L*L smoothing spline estimator, which is better suited for fitting noisy signals.
2) *Statistical results:* Here, the premise is that we are observing the integer samples of a realization of a continuous-time stationary process $x(t)$, whose power spectrum is $C_{xx}(\omega) = \sigma_0^2/|\hat{L}(\omega)|^2$. Then, the L*L-spline interpolator (resp., smoothing spline estimator) evaluated at $t = t_0$ yields the MMSE estimator of $x(t_0)$ when the measurements are noise-free (resp., corrupted by additive white Gaussian noise). This is a strong result that implies that the E-spline framework is optimal for the estimation of stationary signals whose power spectrum is rational.

The solutions of these problems are defined in the continuous-time domain, and a key contribution of this work has been to show how to compute them efficiently using digital filters. The smoothing spline and interpolation algorithms that have been described yield the expansion coefficients of the continuous-time solution in the corresponding B-spline basis; once these coefficients are known, the spline function is entirely specified and can be easily evaluated at any location t_0 with an $O(1)$ cost. When the transfer function of the operator L is rational, then the corresponding exponential B-spline is compactly supported, and the whole process is implemented efficiently using recursive filtering techniques.

Thanks to the above-mentioned MMSE property, our generalized smoothing spline algorithm, when evaluated at the integers (cf. (18)), is in fact equivalent to a classical discrete Wiener filter. In this respect, the present formulation brings in two advantages. First, it yields a direct characterization of the restoration filter in the z-transform domain together with a fast recursive algorithm—this has to be contrasted with the standard frequency-domain specification and implementation of the Wiener filter. Second, the underlying spline also gives the optimal solution at noninteger locations, which is an aspect that is not addressed in the traditional formulation. Thus, we may think of the smoothing spline as an optimum discretization of the Wiener filter: In addition to yielding the optimally filtered sample values, it also specifies the function space that allows us to map the solution back into the continuous-time domain. Note that the optimum function space is generally not bandlimited, which constitutes a conceptual departure from the standard signal processing paradigm.

In this paper, we have limited ourselves to the class of variational problems that have explicit spline solutions that can be computed using simple linear algorithms (i.e., digital filters). We believe that the proposed spline framework can also be extended to yield some interesting nonlinear algorithms for continuous/discrete signal processing. For instance, we note that the variational argument that is used in the beginning of the proof of Theorem 3 will also hold in the more general case where the data term is an arbitrary nonlinear function of the the signal samples $s(k)$. This implies that there is a much more general class of variational problems that admit a cardinal L*L-spline solutions and may require the development of specific computational techniques. This may constitute an interesting direction for future research.

Appendix I
Proof of Proposition 1

A. Proof of Property i

Because L is of order r, (1) implies that

$$\frac{|\omega|^{2\rho}}{1+|\hat{L}(\omega)|^2} \leq C_\rho < \infty$$

for every $\rho < r - 1/2$. Moreover, if f belongs to W_2^L, then

$$\begin{aligned}\int |\omega|^{2\rho}|\hat{f}(\omega)|^2 d\omega &= \int \frac{|\omega|^{2\rho}}{1+|\hat{L}(\omega)|^2} \\ &\quad \cdot (1+|\hat{L}(\omega)|^2)|\hat{f}(\omega)|^2 d\omega \\ &\leq C_\rho \int (1+|\hat{L}(\omega)|^2)|\hat{f}(\omega)|^2 d\omega < \infty\end{aligned}$$

which proves that $f \in W_2^{\mathrm{L}}$ implies $f \in W_2^\rho$.

B. Proof of Property ii

Again, let f belong to W_2^L and $\rho < r - 1/2$. We build the function $u_N(\omega) = \sum_{|n|\leq N}(j(\omega+2n\pi))^\rho \hat{f}(\omega+2n\pi)$. Then, using Cauchy–Schwarz inequality, we have

$$\begin{aligned}|u_N(\omega)|^2 &= \left| \sum_{|n|\leq N} \frac{j(\omega+2n\pi))^\rho}{\sqrt{1+|\hat{L}(\omega+2n\pi)|^2}} \right. \\ &\quad \left. \cdot \sqrt{1+|\hat{L}(\omega+2n\pi)|^2}\hat{f}(\omega+2n\pi) \right|^2 \\ &\leq \sum_{|n|\leq N} \frac{|\omega+2n\pi|^{2\rho}}{1+|\hat{L}(\omega+2n\pi)|^2} \\ &\quad \cdot \sum_{|n|\leq N} (1+|\hat{L}(\omega+2n\pi)|^2)|\hat{f}(\omega+2n\pi)|^2 \\ &\leq C_\rho \sum_{n\in\mathbb{Z}} (1+|\hat{L}(\omega+2n\pi)|^2) \\ &\quad \cdot |\hat{f}(\omega+2n\pi)|^2.\end{aligned}$$

The left-hand side is integrable in $[0, 2\pi]$ since $f \in W_2^L$, which implies that $u_N(\omega)$ is square integrable over $[0, 2\pi]$. This also implies that $u = \lim_{N\to\infty} u_N$ is square integrable as well, while the convergence from u_N to u is dominated by the above right-hand side. Thanks to Fourier's theorem, we know that the 2π-periodic function $u(\omega)$ can be expressed as $\sum_{n\in\mathbb{Z}} c_n e^{-jn\omega}$ almost everywhere. The coefficients c_n are given by

$$c_n = \frac{1}{2\pi}\int_0^{2\pi} u(\omega)e^{jn\omega}d\omega = \frac{1}{2\pi}\int_0^{2\pi} \lim_{N\to\infty} u_N(\omega)e^{jn\omega}d\omega$$

$$c_n = \lim_{N\to\infty} \frac{1}{2\pi}\int_0^{2\pi} u_N(\omega)e^{jn\omega}d\omega$$

(Lebesgue's dominated convergence Theorem)

$$c_n = \frac{1}{2\pi}\int_{-\infty}^{\infty}(j\omega)^{\rho}\hat{f}(\omega)e^{jn\omega}d\omega = f^{(\rho)}(n)$$

which proves Poisson's summation formula for f and all its derivatives up to order $r - 1/2$.

C. Proof of Property iii

As a corollary of *ii.*, the coefficients c_n are known to belong to ℓ_2 because $u \in L_2([0, 2\pi])$.

Appendix II
Recursive Implementation of the Smoothing Spline/Wiener Filter

When the underlying B-spline and localization filter are compactly supported, which is necessarily the case when $\hat{L}(\omega)$ is rational, the smoothing spline filter $H_\lambda(z)$ specified by (14) is an allpole symmetrical system. This implies that its roots (poles) come in reciprocal pairs $\{(z_i, 1/z_i)\}_{(i=1,\ldots,n)}$ with $|z_i| < 1$ because the filter is stable (cf. Theorem 3). To obtain a stable recursive implementation, we need to separate the filter into causal and anticausal components.

A first solution is based on the product decomposition

$$H_\lambda(z) = \frac{1}{P_n(z)} \cdot \frac{1}{P_n(z^{-1})}$$

where $P_n(z)$ is an nth-order polynomial in z^{-1} of the form

$$P_n(z) = z^{-n}p_n\prod_{i=1}^{n}(z - z_i).$$

This suggests a cascade structure where the allpole causal filter $1/P_n(z)$ is implemented recursively from left to right and is followed by the anticausal filter $1/P_n(z^{-1})$, whose implementation is essentially the same, except that the recursion is applied from right to left. The only delicate aspect is the handling of the boundary conditions, which should be chosen to be mirror symmetric to minimize artifacts. When all roots are real, this can be done via the algorithm described in [34], which is based on a finer decomposition into a cascade of firstorder systems, with an appropriate treatment of boundary conditions.

When some of the roots are complex, the decomposition into a cascade of first-order systems is less attractive computationally. The grouping into second-order systems is not a good solution either because of the complications introduced by the boundary conditions. We therefore propose an alternative approach that is based on the sum decomposition

$$H_\lambda(z) = S_\lambda(z) + S_\lambda(z^{-1}) \tag{20}$$

where $S_\lambda(z) = Q_n(z)/P_n(z)$ is a stable causal rational filter. The coefficients of the polynomial $Q_n(z)$ in the denominator of $S_\lambda(z)$ are determined by solving a linear system of equation or by regrouping the causal terms of a decomposition of $H_\lambda(z)$ in simple partial fractions.

We now propose a simple practical solution for implementing this filter using a standard causal recursive filtering module (e.g., a Matlab routine) while enforcing the constraint of mirror symmetric boundary conditions. Let $\{x[k]\}_{(k=0,\ldots,N-1)}$ denote our input signal of size N. This signal is extended to infinity, at least conceptually, by imposing the mirror symmetric conditions

$$\begin{cases} x[k] = x[-k], & \forall\, k \in \mathbb{Z} \\ x[N-1+k] = x[N-1-k], & \forall\, k \in \mathbb{Z} \end{cases} \tag{21}$$

which produces an extended signal that has centers of symmetry at the boundaries ($k = 0$ and $k = N - 1$) and is $(2N - 1)$-periodic. The reason for using these particular boundary conditions is that they are preserved at the output when the filter symmetric, which is precisely the case here. This ensures that the spline model is consistent with the same boundary conditions on $c[k]$. For instance, in the case of an exact fit ($\lambda = 0$), we recover $x[k]$ exactly by filtering the B-spline coefficients with $b[k] = \varphi(k)$, which is center-symmetric as well.

For a signal that satisfies boundary conditions (21), it is not difficult to show that

$$c[k] = (h_\lambda * x)[k] = y[k] + y[2N - 1 - k]$$

with $y[k] = (s_\lambda * x)[k]$, where s_λ is the causal filter in the sum decomposition (20). Thus, a quick and simple way of implementing the smoothing spline filter is by constructing an augmented input vector $\{x[k]\}_{(k=-k_0,\ldots,2N-1)}$, feeding it into a causal recursive filter to obtain the values $\{y[k]\}_{(k=0,\ldots,2N-1)}$, and finally computing $\{c[k]\}_{(k=0,\ldots,N-1)}$ as indicated above. The integer $k_0 = \log(\epsilon)/\log(\max_i |z_i|)$ is selected sufficiently large for the transient effect of the general purpose causal filtering routine to be ϵ-negligible at $k = 0$.

Acknowledgment

The authors would like to thank I. Khalidov for implementing the recursive filtering algorithm described in Appendix II and for his help with the production of Fig. 2.

References

[1] 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.

[2] M. Unser, "Cardinal exponential splines: Part II—Think analog, act digital," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1439–1449, Apr. 2005.

[3] G. Demoment, "Image reconstruction and restoration—Overview of common estimation structures and problems," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 2024–2036, 1989.

[4] N. B. Karayiannis and A. N. Venetsanopoulos, "Regularization theory in image restoration—The stabilizing functional approach," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 7, pp. 1155–1179, Jul. 1990.

[5] M. R. Banham and A. K. Katsaggelos, "Digital image restoration," *IEEE Signal Process. Mag.*, vol. 14, no. 2, pp. 24–41, Feb. 1997.

[6] I. J. Schoenberg, "Spline functions and the problem of graduation," *Proc. Nat. Acad. Sci.*, vol. 52, pp. 947–950, 1964.

[7] C. H. Reinsh, "Smoothing by spline functions," *Numer. Math.*, vol. 10, pp. 177–183, 1967.

[8] C. De Boor and R. E. Lynch, "On splines and their minimum properties," *J. Math. Mech.*, vol. 15, no. 6, pp. 953–969, 1966.

[9] J. H. Ahlberg, E. N. Nilson, and J. L. Walsh, *The Theory of Splines and Their Applications*. New York: Academic, 1967.

[10] P. M. Prenter, *Splines and Variational Methods*. New York: Wiley, 1975.

[11] J. Duchon, "Splines minimizing rotation-invariant semi-norms in sobolev spaces," in *Constructive Theory of Functions of Several Variables*, W. Schempp and K. Zeller, Eds. Berlin, Germany: Springer-Verlag, 1977, pp. 85–100.

[12] J. Meinguet, "Multivariate interpolation at arbitrary points made simple," *Zeitschrift fur Angewandte Mathematik und Physik*, vol. 30, pp. 292–304, 1979.
[13] C. Micchelli, "Interpolation of scattered data: Distance matrices and conditionally positive definite functions," *Constructive Approx.*, vol. 2, pp. 11–22, 1986.
[14] G. Kimeldorf and G. Wahba, "A correspondence between bayesian estimation on stochastic processes and smoothing by splines," *Ann. Math. Statist.*, vol. 41, pp. 495–502, 1970.
[15] G. K. Wahba, "Spline functions and stochastic processes," *Sankhya Series A*, vol. 32, pp. 173–180, 1970.
[16] G. Wahba, *Spline Models for Observational Data*. Philadelphia, PA: SIAM, 1990.
[17] D. D. Cox, "An analysis of bayesian-inference for nonparametric regression," *Ann. Statist.*, vol. 21, no. 2, pp. 903–923, 1993.
[18] R. Eubank, *Nonparametric Regression and Spline Smoothing*. New York: Marcel Dekker, 1999.
[19] A. Van der Linde, "Smoothing errors," *Statist.*, vol. 31, no. 2, pp. 91–114, 1998.
[20] H. L. Weinert and G. S. Sidhu, "Stochastic framework for recursive computation of spline functions: Part I, Interpolating splines," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 1, pp. 45–50, Jan. 1978.
[21] H. L. Weinert, R. H. Byrd, and G. S. Sidhu, "Stochastic framework for recursive computation of spline functions: Part II, Smoothing splines," *J. Optim. Theory Applicat.*, vol. 30, no. 2, pp. 255–268, 1980.
[22] G. Matheron, "The intrinsic random functions and their applications," *Appl. Probab.*, vol. 5, no. 12, pp. 439–468, 1973.
[23] D. E. Myers, "Kriging, cokriging, radial basis functions and the role of positive definiteness," *Comput. Math. Applicat.*, vol. 24, no. 12, pp. 139–148, 1992.
[24] N. Cressie, "Geostatistics," *Amer. Statist.*, vol. 43, no. 4, pp. 197–202, 1989.
[25] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part II—Efficient design and applications," *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 834–848, Feb. 1993.
[26] E. M. Stein and G. Weiss, *Fourier Analysis on Euclidean Spaces*. Princeton, NJ: Princeton Univ. Press, 1971.
[27] A. V. Oppenheim and A. S. Willsky, *Signal and Systems*. Upper Saddle River, NJ: Prentice-Hall, 1996.
[28] B. P. Lathy, *Signal Processing and Linear Systems*. Carmichael, CA: Berkeley-Cambridge Press, 1998.
[29] Y. Katznelson, *An Introduction to Harmonic Analysis*. New York: Dover, 1976.
[30] A. Aldroubi and K. Grochenig, "Nonuniform sampling and reconstruction in shift invariant spaces," *SIAM Rev.*, vol. 43, pp. 585–620, 2001.
[31] R. Panda, G. S. Rath, and B. N. Chatterji, "Generalized B-spline signal processing," *Signal Process.*, vol. 55, pp. 1–14, 1996.
[32] M. Unser and T. Blu, "Fractional splines and wavelets," *SIAM Rev.*, vol. 42, no. 1, pp. 43–67, 2000.
[33] T. Blu and M. Unser, "Wavelets, fractals, and radial basis functions," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 543–553, Mar. 2002.
[34] P. Thévenaz, T. Blu, and M. Unser, "Image interpolation and resampling," in *Handbook of Medical Imaging, Processing and Analysis*, I. N. Bankman, Ed. San Diego, CA: Academic, 2000, ch. 25, pp. 393–420.
[35] I. Daubechies, "Orthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, pp. 909–996, 1988.
[36] G. Deslauriers and D. Dubuc, "Symmetric iterative interpolation processes," *Constr. Approx.*, vol. 5, pp. 49–68, 1989.
[37] I. J. Schoenberg, "Contribution to the problem of approximation of equidistant data by analytic functions," *Quart. Appl. Math.*, vol. 4, pp. 45–99, 112, 1946.
[38] G. Schweikert, "An interpolating curve using a spline in tension," *J. Math. Phys.*, vol. 45, pp. 312–317, 1966.
[39] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1991.

Michael Unser (M'89–SM'94–F'99) received the M.S. (summa cum laude) and Ph.D. degrees in electrical engineering in 1981 and 1984, respectively, from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland.

From 1985 to 1997, he was with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, MD. He is now Professor and Head of the Biomedical Imaging Group at EPFL. His main research area is biomedical image processing. He has a strong interest in sampling theories, multiresolution algorithms, wavelets, and the use of splines for image processing. He is the author of 100 published journal papers in these areas.

Dr. Unser is Associate Editor-in-Chief for the IEEE TRANSACTIONS ON MEDICAL IMAGING. He is on the editorial boards of several other journals, including IEEE SIGNAL PROCESSING MAGAZINE, IEEE TRANSACTIONS ON IMAGE PROCESSING (from 1992 to 1995), and IEEE SIGNAL PROCESSING LETTERS (from 1994 to 1998). He serves as regular chair for the SPIE Conference on Wavelets, which has been held annually since 1993. He was general co-chair of the first IEEE International Symposium on Biomedical Imaging, Washington, DC, 2002. He received the 1995 and 2003 Best Paper Awards and the 2000 Magazine Award from the IEEE Signal Processing Society.

Thierry Blu (M'96) was born in Orléans, France, in 1964. He received the "Diplôme d'ingénieur" from École Polytechnique, Paris, France, in 1986 and from Télécom Paris (ENST), France, in 1988. In 1996, he received the Ph.D. degree in electrical engineering from ENST for a study on iterated rational filterbanks applied to wideband audio coding.

He is with the Biomedical Imaging Group, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, on leave from the France Télécom National Center for Telecommunications Studies (CNET), Issy-les-Moulineaux, France. His research interests include (multi)wavelets, multiresolution analysis, multirate filterbanks, approximation and sampling theory, psychoacoustics, optics, wave propagation, etc.

Dr. Blu received the 2003 best paper award (SP Theory and Methods) from the IEEE Signal Processing Society. He is currently serving as an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING.