

Full Length Article

Optimal nonlinear approximation with exponential splines

Bassam El Rawas^{a,*}, Michael Unser^a, Rahul Parhi^b^a Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland^b Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA, USA

Received 5 August 2025; received in revised form 10 September 2026; accepted 12 September 2026

Communicated by Ding-Xuan Zhou

Abstract

The present paper is concerned with the L^2 -approximation properties of exponential splines with n knots for functions defined on an interval $\Omega = (a, b)$. We first specify the Banach native space for these splines and show that it continuously embeds in $L^2(\Omega)$. We derive sharp bounds on the approximation error rate for functions in this native space by exponential splines with n knots. To prove the optimality of these rates, we show that these native spaces are equivalent (as Banach spaces) to spaces of higher-order bounded variation on Ω .

© 2026 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Keywords: Approximation rate; Bounded variation; Dictionary approximation; Exponential splines; Nonlinear approximation

1. Introduction

Polynomial splines are ubiquitous in science and engineering due to their remarkable approximation properties [8]. Polynomial splines of order N are tightly linked to the N th order differential operator D^N . Indeed, a function g is a polynomial spline of order N if and only if

$$D^N\{g\}(x) = \sum_k a_k \delta(x - x_k), \quad (1)$$

where the $(x_k)_k$ are the knots of the spline. This shows that g is a polynomial of degree $(N - 1)$ on each interval (x_k, x_{k+1}) , and is $(N - 2)$ times continuously differentiable everywhere. This

* Corresponding author.

E-mail address: bassam.elrawas@epfl.ch (B. El Rawas).

motivates the definition of $BV^N(\Omega)$, the space of functions of N th-order bounded variation on an interval $\Omega = (a, b) \subset \mathbb{R}$. This space is defined as

$$BV^N(\Omega) = \{f \in \mathcal{D}'(\Omega) : \|D^N f\|_{\mathcal{M}(\Omega)} < +\infty\}, \quad (2)$$

where $\mathcal{D}'(\Omega)$ is the space of distributions on Ω and $\|\cdot\|_{\mathcal{M}(\Omega)}$ (defined in (12)) is the total-variation norm in the sense of measures, which can be viewed as a generalization of the L^1 -norm to “absolutely integrable” distributions. Observe that, if g is a polynomial spline of order N , then $\|D^N g\|_{\mathcal{M}} = \|\alpha\|_{\ell^1}$. Therefore, polynomial splines of order N with coefficient sequences in ℓ^1 are examples of function in $BV^N(\Omega)$. These function spaces are particularly well-suited to model real-world objects, as its members have a piecewise-smooth structure. Furthermore, the approximation properties of polynomial splines for functions that lie in $BV^N(\Omega)$ are well-studied [9]. We also remark that wavelet-based approximations for BV^N have also been studied [2].

While polynomial splines are widely used in practice, *exponential splines* [23] offer even more flexibility. A function g is an exponential spline associated to the differential operator $L_\alpha = \prod_{n=1}^N (D - \alpha_n \text{Id})$, $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{C}^N$, if and only if

$$L_\alpha\{g\}(x) = \sum_k a_k \delta(x - x_k). \quad (3)$$

This definition is similar to that of polynomial splines, with the main difference being that g is now a piecewise-exponential polynomial, whose structure depends on α . Note that, when $\alpha = \mathbf{0}$, we get back to the case where g is a polynomial spline.

The use of exponential splines is partly motivated by the fact that differential operators often appear with physical models [12,13] whose differential equations are not simply dictated by a single derivative operator. In the context of inverse problems, the use of differential operators for generalized total-variation (gTV) regularization promotes solutions that are exponential splines [24], which changes the geometry of the solution space and allows for more freedom in the choice of the shape of the solutions. Exponential splines also offer more flexibility in the design of the approximants, an important consideration when it comes to the choice of basis functions appropriate for the representation of certain types of signals.

Exponential splines are also popular in the context of curve reconstruction. For example, in [7], it is shown that, unlike polynomial splines, exponential splines can reproduce trigonometric functions, which allows for the reconstruction of circles and ellipses. Other classes of exponential splines also enjoy a Hermite interpolation property [5], which allows for the reproduction of various shapes at arbitrary scales due to their multiresolution properties. These curve-reconstruction techniques have been applied to biomedical imaging, for instance in [22] where exponential B-splines are used to segment cells in microscopic images, or in [6] where ellipse-like targets are detected in different types of medical images.

Furthermore, with the goal of flexibility and ease of implementation, there are examples of non-stationary subdivision schemes that give rise to exponential splines with tunable parameters [3,4,19]. These allow for the control of properties such as their smoothness or the curves they can reconstruct. Motivated by the flexibility of exponential splines and the applications listed above, it is of interest to investigate their approximation properties. While there are works that explore their linear approximation capacities (such as [13]), none is specifically devoted to the characterization of their nonlinear approximation properties. This is a gap that is being filled with this paper. In particular, we are interested in nonlinear n -term approximations.

1.1. Contributions

To each differential operator L_α , we associate the *native space* $\mathcal{M}^\alpha(\Omega)$ defined as

$$\mathcal{M}^\alpha(\Omega) = \{f \in \mathcal{D}'(\Omega) : \|L_\alpha f\|_{\mathcal{M}(\Omega)} < +\infty\}, \quad (4)$$

where the relevant distributions can be identified as locally integrable functions. This space plays the same role as $BV^N(\Omega)$, which is precisely the native space of the differential operator $D^N = L_{(0,\dots,0)}$. Our contributions in this paper are as follows.

1. We consider the native space $\mathcal{M}^\alpha(\Omega)$ of a linear differential operator

$$L_\alpha = \prod_{n=1}^N (D - \alpha_n \text{Id}), \quad \alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{C}^N, \quad (5)$$

defined on the bounded open interval $\Omega = (a, b)$ and equipped with a Banach direct-sum topology. We show that the continuous embedding $\mathcal{M}^\alpha(\Omega) \xhookrightarrow{c} L^2(\Omega)$ holds with respect to the aforementioned topology.

2. We derive an upper bound on the exponential-spline n -term nonlinear approximation in $L^2(\Omega)$ for functions in $\mathcal{M}^\alpha(\Omega)$. In particular, we show that the approximation rate with respect to the $L^2(\Omega)$ norm is n^{-N} , where N is the highest-order derivative in L_α .
3. We prove that the native space $\mathcal{M}^\alpha(\Omega)$ is topologically equivalent to $BV^N(\Omega)$. The well-known continuous embeddings (cf., [17])

$$B_{1,1}^N(\Omega) \xhookrightarrow{c} BV^N(\Omega) \xhookrightarrow{c} B_{1,\infty}^N(\Omega) \quad (6)$$

then allow us to prove the optimality of the rate n^{-N} . The topological equivalence also gives access to the suboptimality of linear approximation via the known Kolmogorov n -widths.

1.2. Organization of the paper

We begin by recalling necessary mathematical notions in Section 2 and briefly present preliminary results on the Green's function of differential operators on a bounded domain. We also discuss our choice of norm on $\mathcal{M}^\alpha(\Omega)$. In Section 3, we prove that the native space $\mathcal{M}^\alpha(\Omega)$ is continuously embedded into $L^2(\Omega)$. We also define a dictionary of exponential-spline components on Ω and show that it is uniformly bounded and compact in $L^2(\Omega)$. We show that it admits a smooth parameterization that depends on the order N of the differential operator L_α . We then derive in Section 4 an upper approximation bound that strictly improves on the universal rates of Maurey. Our result generalizes known rates for splines with uniformly spaced knots that depend on the spacing of the knots rather than on their quantity. It also allows us to retrieve known results for n -term polynomial-spline approximation rates, which constitutes a special case of our framework. Finally, in Section 5, we establish a topological equivalence between $\mathcal{M}^\alpha(\Omega)$ and $BV^N(\Omega)$. It allows us to discuss the optimality of our rates.

2. Preliminaries and notations

Let $\Omega \subset \mathbb{R}$ be an open interval and let $\mathcal{D}'(\Omega)$ be the space of distributions on Ω . Let $\alpha \in \mathbb{C}^N$. The *linear differential operator* $L_\alpha : \mathcal{D}'(\Omega) \rightarrow \mathcal{D}'(\Omega)$ is defined as

$$L_\alpha = (D - \alpha_1 \text{Id}) \circ \dots \circ (D - \alpha_N \text{Id}) = \prod_{n=1}^N (D - \alpha_n \text{Id}), \quad (7)$$

where the α_n will be referred to as its *poles*. Its characteristic polynomial can be factorized as

$$p_\alpha(z) = \prod_{n=1}^N (z - \alpha_n) = \prod_{m=1}^{N_d} (z - \alpha_{(m)})^{n_{(m)}}, \quad (8)$$

where N_d is the number of distinct roots, and $\alpha_{(m)}$ denotes its m th distinct root of multiplicity $n_{(m)}$. Furthermore, we denote by $\mathcal{N}_\alpha = \{f \in \mathcal{D}'(\Omega) : L_\alpha f = 0\}$ the *null space* of L_α . The latter is known [23] to be a finite-dimensional vector space with $\dim(\mathcal{N}_\alpha) = N = \sum_{m=1}^{N_d} n_{(m)}$. It consists of exponentials and/or exponential polynomials. Specifically,

$$\mathcal{N}_\alpha = \text{span}\{x \mapsto x^{n-1} e^{\alpha_{(m)} x}\}_{m=1, \dots, N_d; n=1, \dots, n_{(m)}}. \quad (9)$$

We also define the *causal Green's function* of L_α as the unique causal function $\rho_\alpha : \Omega \rightarrow \mathbb{R}$ that satisfies $L_\alpha \rho_\alpha = \delta$ in $\mathcal{D}'(\Omega)$.

We now introduce the exponential spline associated to the linear differential operator L_α on the bounded open interval $\Omega = (a, b)$.

Definition 2.1. A function s is an *exponential spline* (or an L_α -spline) on $\Omega = (a, b)$ if and only if

$$L_\alpha \{s\}(x) = \sum_{k=1}^K a_k \delta(x - x_k) \quad (10)$$

for a finite series of knots $a < x_1 < \dots < x_K < b$ and some linear coefficients $a_1, \dots, a_K \in \mathbb{R}$.

Starting from (10), we retrieve the explicit form of s by “integration”, which yields that

$$s(x) = \sum_{k=1}^K a_k \rho_\alpha(x - x_k) + q(x), \quad (11)$$

where ρ_α is the causal Green's function of L_α and $q \in \mathcal{N}_\alpha$ is some null-space component whose determination requires additional boundary conditions. When $\alpha = \mathbf{0}$, $L_\alpha = D^N$ and Definition 2.1 reduces to that of classical polynomial splines of degree $(N - 1)$.

The *total-variation norm* of a distribution $f \in \mathcal{D}'(\Omega)$ can be defined as

$$\|f\|_{\mathcal{M}(\Omega)} := \sup_{\varphi \in \mathcal{D}(\Omega), \|\varphi\|_\infty = 1} \langle f, \varphi \rangle, \quad (12)$$

where $\|\varphi\|_\infty := \sup_{x \in \Omega} |\varphi(x)|$. It is used to define the Banach space $(\mathcal{M}(\Omega), \|\cdot\|_{\mathcal{M}(\Omega)})$ with

$$\mathcal{M}(\Omega) = \{f \in \mathcal{D}'(\Omega) : \|f\|_{\mathcal{M}(\Omega)} < +\infty\}, \quad (13)$$

which coincides with the space of Radon measures on Ω . Since $\|f\|_{L^1(\Omega)} = \|f\|_{\mathcal{M}(\Omega)}$ for any $f \in L^1(\Omega)$, this norm is a generalization of the L^1 norm. Note that $L^1(\Omega)$ is isometrically included in $\mathcal{M}(\Omega)$, and the inclusion is strict.

This total-variation norm allows us to identify the *native space* of the differential operator L_α as

$$\mathcal{M}^\alpha(\Omega) = \{f \in \mathcal{D}'(\Omega) : \|L_\alpha f\|_{\mathcal{M}(\Omega)} < +\infty\}, \quad (14)$$

where we observe that $\|L_\alpha \cdot\|_{\mathcal{M}(\Omega)}$ is a semi-norm on $\mathcal{M}^\alpha(\Omega)$.

Let $\mathcal{X}$ be a topological vector space whose continuous dual is denoted $\mathcal{X}'$. Given a linear functional $\phi \in \mathcal{X}'$, we sometimes write $\phi(x) = \langle \phi, x \rangle$ for all $x \in \mathcal{X}$. Then, given a measure

μ , we write that $w(t) dt = d\mu(t)$, with the implicit understanding that

$$\langle \varphi, w \rangle = \int_{\Omega} \varphi(t) d\mu(t) = \int_{\Omega} \varphi(t) w(t) dt \quad (15)$$

for any $\varphi \in \mathcal{D}(\Omega)$, where $w \in \mathcal{D}'(\Omega)$ is the representer of μ in $\mathcal{D}'(\Omega)$.

Given a Banach space $(\mathcal{X}, \|\cdot\|)$, the *diameter* of a subset $A \subseteq \mathcal{X}$ is defined as

$$\text{diam}(A) = \sup\{\|u - v\| : u, v \in A\} \in [0, +\infty]. \quad (16)$$

2.1. Green's function of differential operators on a bounded domain

A linear operator $T : \mathcal{D}'(\Omega) \rightarrow \mathcal{D}'(\Omega)$ is said to be *shift-invariant* if $T\{f(\cdot - x_0)\}(x) = T\{f\}(x - x_0)$ for any $f \in \mathcal{D}'(\Omega)$ and any $x_0 \in \Omega$. Linear differential operators of the same form as L_{α} fall under this category, and the Malgrange–Ehrenpreis theorem [25] ensures that they admit a *Green's function*: a function $g(\cdot)$ of a single variable such that $Tg = \delta$. Note that $\tilde{g} = g + q$ where g is a Green's function of T and $q \in \mathcal{N}_T$ is also a Green's function of T , which means that they generally are not unique. However, linear differential operators with constant coefficients always admit a unique *causal* Green's function.

In the sequel, all Green's functions of L_{α} will be assumed to be causal and will be denoted ρ_{α} . When L_{α} is defined on $\mathbb{R}$, its Green's function satisfies the convolution property

$$\rho_{\alpha} = \rho_{\alpha_1} * \cdots * \rho_{\alpha_N}, \quad (17)$$

where $\rho_{\alpha_i}(x) = \mathbb{1}_{+}(x)e^{\alpha_i x}$ is the Green's function of the first-order operator $(D - \alpha_i \text{Id})$. In our case, L_{α} is defined on $\Omega = (a, b)$. However, it turns out that this does not change the expression of its Green's function. Indeed, a test function $\varphi \in \mathcal{D}(\Omega)$ can be extended to $\varphi_{\text{ext}} \in \mathcal{D}(\mathbb{R})$ as

$$\varphi_{\text{ext}}(x) = \begin{cases} \varphi(x), & \text{if } x \in \Omega \\ 0, & \text{otherwise.} \end{cases} \quad (18)$$

Then, as $L_{\alpha} \rho_{\alpha} = \delta$ on $\mathcal{D}'(\mathbb{R})$, we have that, for any $x_0 \in \Omega$,

$$\varphi(x_0) = \varphi_{\text{ext}}(x_0) = \langle L_{\alpha} \rho_{\alpha}(\cdot - x_0), \varphi_{\text{ext}} \rangle_{\mathcal{D}'(\mathbb{R})} = \langle L_{\alpha} \rho_{\alpha}(\cdot - x_0), \varphi \rangle_{\mathcal{D}'(\Omega)}, \quad (19)$$

which implies that $L_{\alpha} \rho_{\alpha} = \delta$ on $\mathcal{D}'(\Omega)$ as $\varphi \in \mathcal{D}(\Omega)$ is arbitrary.

2.2. Choice of norms on the native space

The space $\mathcal{M}^{\alpha}(\Omega)$ defined in (14) as a set can be equipped with a variety of norms to turn it into a Banach space. To do so, we need to factor out $\mathcal{N}_{\alpha}$ from $\mathcal{M}^{\alpha}(\Omega)$, which can be achieved with the help of the *biorthogonal system* proposed in [24] and stated in Definition 2.2.

Definition 2.2. Let $\mathcal{N}$ be a finite-dimensional vector subspace of some Banach space $\mathcal{X}$, with $N_0 = \dim(\mathcal{N})$. The pair $(\phi_n, p_n)_{n=1}^{N_0}$ is called a *biorthogonal system* for $\mathcal{N}$ if

- The set $\{p_n\}_{n=1}^{N_0}$ is a basis for $\mathcal{N}$,
- For all $n = 1, \dots, N_0$, the functions $\phi_n \in \mathcal{X}'$ and $\phi_n(p_m) = \delta_{nm}$ (biorthogonality).

We also denote by $\boldsymbol{\phi}(f) \in \mathbb{R}^{N_0}$ the vector $\boldsymbol{\phi}(f) = (\phi_1(f), \dots, \phi_{N_0}(f))$.

Remark 2.3. The biorthogonality condition can also be rewritten as $\boldsymbol{\phi}(p_n) = \mathbf{e}_n$, where $\mathbf{e}_n$ is the n th element of the canonical basis of $\mathbb{R}^{N_0}$.

In what follows, we consider a biorthogonal system $(\phi_n, p_n)_{n=1}^{N_0}$ for the finite-dimensional vector space $\mathcal{N}_\alpha$ and we define a composite norm for any $f \in \mathcal{M}^\alpha(\Omega)$ as

$$\|f\|_\phi := \|L_\alpha f\|_{\mathcal{M}(\Omega)} + \|\phi(f)\|_2, \quad (20)$$

which turns $\mathcal{M}^\alpha(\Omega)$ into a Banach space. The definition of this norm relies on the direct-sum decomposition of $\mathcal{M}^\alpha(\Omega)$ as

$$\mathcal{M}^\alpha(\Omega) = \mathcal{M}_\rho^\alpha(\Omega) \oplus \mathcal{N}_\alpha, \quad (21)$$

where

$$\mathcal{M}_\rho^\alpha(\Omega) = \left\{ f \in \mathcal{D}'(\Omega) : f = \int_\Omega \rho_\alpha(\cdot - y) d\mu(y) \text{ for some measure } \mu \in \mathcal{M}(\Omega) \right\}. \quad (22)$$

The proof of this decomposition is detailed in [Appendix A](#). We discuss the choice of norms on $\mathcal{M}^\alpha(\Omega)$ and their equivalence in [Section 5](#).

3. Exponential-spline dictionary approximation

3.1. Continuous embedding of the native space into $L^2(\Omega)$

We derive the approximation rate with respect to the L^2 norm. Hence, we need to establish a continuous embedding of $\mathcal{M}^\alpha(\Omega)$ into $L^2(\Omega)$ with respect to the norm $\|\cdot\|_\phi$, in order to give meaning to the expression $\|f\|_{L^2(\Omega)}$ for $f \in \mathcal{M}^\alpha(\Omega)$. This is the object of [Proposition 3.1](#).

Proposition 3.1. *There exists a constant $C > 0$ such that, for all $f \in \mathcal{M}^\alpha(\Omega)$, $\|f\|_{L^2(\Omega)} \leq C\|f\|_\phi$.*

Proof. We start from the representation $f = \int_\Omega \rho_\alpha(\cdot - y) d\mu(y) + q = f_1 + q$, where $q \in \mathcal{N}_\alpha$ and μ is a measure specified by $L_\alpha f$, which is proven in [Appendix A \(Theorem A.3\)](#). By the triangle inequality,

$$\|f\|_{L^2(\Omega)} \leq \left\| \int_\Omega \rho_\alpha(\cdot - y) d\mu(y) \right\|_{L^2(\Omega)} + \|q\|_{L^2(\Omega)}. \quad (23)$$

We shall bound each term separately.

We first have that $\phi(q) = \phi(f - f_1) = \phi(f)$, which allows us to write

$$q = \sum_{n=1}^{N_0} \langle \phi_n, q \rangle p_n = \sum_{n=1}^{N_0} \langle \phi_n, f \rangle p_n. \quad (24)$$

This leads to

$$\begin{aligned} \|q\|_{L^2(\Omega)}^2 &= \int_\Omega \left| \sum_{n=1}^{N_0} \langle \phi_n, f \rangle p_n(x) \right|^2 dx \leq \int_\Omega N_0 \sum_{n=1}^{N_0} |\langle \phi_n, f \rangle|^2 |p_n(x)|^2 dx \\ &= \sum_{n=1}^{N_0} |\langle \phi_n, f \rangle|^2 N_0 \|p_n\|_{L^2(\Omega)}^2 \\ &\leq C_1^2 \sum_{n=1}^{N_0} |\langle \phi_n, f \rangle|^2 = C_1^2 \|\phi(f)\|_2^2, \end{aligned} \quad (25)$$

where $C_1^2 = N_0 \max_{n=1, \dots, N_0} \|p_n\|_{L^2(\Omega)}^2$, the maximum running over all N_0 elements of the basis of $\mathcal{N}_\alpha$. This maximum exists because each p_n is continuous and Ω is bounded.

To bound the integral of ρ_α , we start by bounding ρ_α itself. Let $x, y \in \Omega$. Then,

$$|\rho_\alpha(x - y)| \leq \sup_{x \in \Omega} |\rho_\alpha(x)| = \tilde{C}. \quad (26)$$

Hence for any $y \in \Omega$, $\|\rho_\alpha(\cdot - y)\|_\infty \leq \tilde{C}$. Since Ω has finite Lebesgue measure, if $1 \leq p < q \leq +\infty$, then there exists a constant $K > 0$ which depends on p, q , and the measure of Ω such that $\|f\|_{L^p(\Omega)} \leq K \|f\|_{L^q(\Omega)}$ for any $f \in L^q(\Omega) \subseteq L^p(\Omega)$. In our case, this yields

$$\|\rho_\alpha(\cdot - y)\|_{L^2(\Omega)} \leq K \|\rho_\alpha(\cdot - y)\|_{L^\infty(\Omega)}. \quad (27)$$

From the uniform bound and Inequality (27), we obtain

$$\begin{aligned} \left\| \int_\Omega \rho_\alpha(\cdot - y) d\mu(y) \right\|_{L^2(\Omega)} &\leq \int_\Omega \|\rho_\alpha(\cdot - y)\|_{L^2(\Omega)} d|\mu|(y) \\ &\leq \int_\Omega K \|\rho_\alpha(\cdot - y)\|_{L^\infty(\Omega)} d|\mu|(y) \\ &\leq K \tilde{C} \int_\Omega d|\mu|(y) = C_2 \|\mu\|_{\mathcal{M}(\Omega)} = C_2 \|L_\alpha f\|_{\mathcal{M}(\Omega)}, \end{aligned} \quad (28)$$

with $C_2 = K \tilde{C}$, and μ being the measure corresponding to $L_\alpha f$, as shown in Theorem A.3. We conclude that

$$\|f\|_{L^2} \leq \left\| \int_\Omega \rho_\alpha(\cdot - y) d\mu(y) \right\|_{L^2(\Omega)} + \|q\|_{L^2(\Omega)} \leq C_2 \|L_\alpha f\|_{\mathcal{M}(\Omega)} + C_1 \|\phi(f)\|_2 \leq C \|f\|_\phi. \quad \square \quad (29)$$

3.2. Choice of dictionary

A uniformly bounded dictionary in a Banach space $\mathcal{X}$ is a subset $\mathcal{A} \subseteq \mathcal{X}$ such that $\sup_{z \in \mathcal{A}} \|z\|_{\mathcal{X}} < +\infty$. This definition does not impose any conditions on the size of $\mathcal{A}$, which can be uncountable. A dictionary approximation is a type of nonlinear approximation [8]. Given a uniformly bounded dictionary $\mathcal{A} \subseteq \mathcal{X}$, we define the approximation sets

$$\Sigma_n(\mathcal{A}) = \left\{ \sum_{k=1}^n c_k g_k : g_k \in \mathcal{A}, c_k \in \mathbb{R} \right\} \subseteq \mathcal{X} \quad (30)$$

and

$$\Sigma(\mathcal{A}) = \left\{ \sum_{k=1}^n c_k g_k : g_k \in \mathcal{A}, n \in \mathbb{N}_{\geq 1}, \|c\|_1 < +\infty \right\} \subseteq \mathcal{X}, \quad (31)$$

and we seek to determine an n -term approximation of an element $f \in \overline{\Sigma(\mathcal{A})}$ (i.e., an approximation of f that uses an element of $\Sigma_n(\mathcal{A})$). This approximation is nonlinear in the sense that, given two elements in $\Sigma_n(\mathcal{A})$ each described by n terms, their sum will in general be described by $2n$ terms, which means that $\Sigma_n(\mathcal{A})$ is not a subspace of $\mathcal{X}$.

In our case, we are working with the Hilbert space $\mathcal{X} = L^2(\Omega)$. We know from Theorem A.3 that any $f \in \mathcal{M}(\Omega)$ can be written as

$$f(x) = \int_\Omega \rho_\alpha(x - y) d\mu(y) + q(x) = f_1(x) + q(x) \quad (32)$$

with $q \in \mathcal{N}_\alpha$ and $\mu \in \mathcal{M}(\Omega)$, which means that it suffices to approximate functions that belong to the space $\mathcal{M}_\rho^\alpha(\Omega)$ defined in (22). Indeed, to approximate $f \in \mathcal{M}^\alpha(\Omega)$, we make use of the fact that $f = f_1 + q$ with $f_1 \in \mathcal{M}_\rho^\alpha(\Omega)$, approximate f_1 with an exponential spline with n knots, and add $q \in \mathcal{N}_\alpha$ to it. Consequently, our dictionary is given by

$$\mathcal{A}_\alpha = \{\rho_\alpha(\cdot - y)\}_{y \in \overline{\Omega}} \subseteq L^2(\Omega). \quad (33)$$

The corresponding approximation set $\Sigma_n(\mathcal{A}_\alpha)$ is a set of exponential splines with n knots, which we write explicitly as

$$\Sigma_n(\mathcal{A}_\alpha) = \left\{ \sum_{k=1}^n c_k \rho_\alpha(\cdot - y_k) : y_k \in \overline{\Omega}, c_k \in \mathbb{R} \right\}. \quad (34)$$

To show that we can use this dictionary to approximate elements of $\mathcal{M}_\rho^\alpha(\Omega)$, we need to characterize the closure of our approximation sets in $L^2(\Omega)$, which we do in Proposition 3.2.

Proposition 3.2. *The dictionary $\mathcal{A}_\alpha = \{\rho_\alpha(\cdot - y)\}_{y \in \overline{\Omega}} \subseteq L^2(\Omega)$ is compact in $L^2(\Omega)$.*

Proof. As $\mathcal{A}_\alpha$ is a subset of a Banach space, it suffices to show that every sequence in $\mathcal{A}_\alpha$ has a converging subsequence. First, because Ω is bounded, $\mathcal{A}_\alpha$ is uniformly bounded, which means that there exists a constant $M > 0$ such that $\|\rho_\alpha(\cdot - y)\|_{L^2(\Omega)} < M$ for all $y \in \Omega$.

Next, we consider a sequence $\{g_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A}_\alpha$. We can write each g_n as $g_n(\cdot) = \rho_\alpha(\cdot - y_n)$, which defines a sequence $\{y_n\}_{n \in \mathbb{N}} \subseteq \overline{\Omega} = [a, b]$. This new sequence of real numbers is bounded and, by the Bolzano–Weierstrass theorem, admits a converging subsequence $\{y_{n_k}\}_{k \in \mathbb{N}}$ that converges to some $\tilde{y}$. This subsequence corresponds to the subsequence $\{g_{n_k}\}_{k \in \mathbb{N}} = \{\rho_\alpha(\cdot - y_{n_k})\}_{k \in \mathbb{N}}$.

Next, let $x \in \Omega$. The function ρ_α is continuous almost everywhere (everywhere when $N > 1$), which means that

$$\lim_{k \rightarrow +\infty} \rho_\alpha(x - y_{n_k}) = \rho_\alpha(x - \tilde{y}) \quad (35)$$

for almost all $x \in \Omega$. Using this fact along with the dominated convergence theorem, we have

$$\lim_{k \rightarrow +\infty} \|\rho_\alpha(x - y_{n_k}) - \rho_\alpha(x - \tilde{y})\|_{L^2(\Omega)}^2 = \lim_{k \rightarrow +\infty} \int_{\Omega} |\rho_\alpha(x - y_{n_k}) - \rho_\alpha(x - \tilde{y})|^2 dx = 0. \quad (36)$$

This shows that any sequence $\{g_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A}_\alpha$ admits a converging subsequence, which shows that $\mathcal{A}_\alpha$ is sequentially compact and completes the proof. $\square$

We then show that $\overline{\Sigma(\mathcal{A}_\alpha)}$, the space this dictionary of exponential splines can approximate, is exactly $\mathcal{M}_\rho^\alpha(\Omega)$.

Proposition 3.3. *The closure of $\Sigma(\mathcal{A}_\alpha) = \{\sum_{k=1}^n c_k \rho_\alpha(\cdot - y_k) : n \in \mathbb{N}_{\geq 1}, \|c\|_{\ell^1} < +\infty\} \subseteq L^2(\Omega)$ in $L^2(\Omega)$ is equal to $\{\int_{\Omega} \rho_\alpha(\cdot - y) d\mu(y)\} = \mathcal{M}_\rho^\alpha(\Omega)$.*

Proof. Let $f \in \overline{\Sigma(\mathcal{A}_\alpha)}$. We invoke [20, Lemma 3] and claim that $f \in \overline{\Sigma(\mathcal{A}_\alpha)}$ if and only if f can be written as a Bochner integral of the form

$$f = \int_{\mathcal{A}_\alpha} i_{\mathcal{A}_\alpha \rightarrow L^2(\Omega)} d\mu \quad (37)$$

with μ a Borel measure on $\mathcal{A}_\alpha$ and $i_{\mathcal{A}_\alpha \rightarrow L^2(\Omega)}$ the inclusion map. By parameterizing the dictionary $\mathcal{A}_\alpha$ with the help of the operator $P : \Omega \rightarrow \mathcal{A}_\alpha$, $y \mapsto \rho_\alpha(\cdot - y)$, we make a “change of variables” by means of the pushforward of μ by P , to obtain

$$f(\cdot) = \int_{\Omega} \rho_\alpha(\cdot - y) dv(y), \quad (38)$$

where v is the pushforward of μ . Since we know that $\mathcal{M}_\rho^\alpha(\Omega)$ is equal to functions of the same form as (38), and by making use of the equivalence given by [20, Lemma 3], we conclude that

$$\overline{\Sigma(\mathcal{A}_\alpha)} = \mathcal{M}_\rho^\alpha(\Omega), \quad (39)$$

which is the desired result. $\square$

3.3. A first approximation rate

The following result of Theorem 3.4 was proved by Maurey [18]. It enables us to get a first approximation rate for our dictionary.

Theorem 3.4. *Let $\mathcal{H}$ be a Hilbert space, $\mathcal{A} \subseteq \mathcal{H}$ a uniformly bounded dictionary, and let $n \in \mathbb{N}_{\geq 1}$. Then there exists $C > 0$ such that for all $f \in \overline{\Sigma(\mathcal{A})}$, there exists an n -term approximant $f_n \in \Sigma_n(\mathcal{A})$ such that*

$$\|f - f_n\|_{\mathcal{H}} \leq C \|f\|_{\overline{\Sigma(\mathcal{A})}} n^{-\frac{1}{2}}. \quad (40)$$

Remark 3.5. The constant C depends on the dictionary. In particular, $C \leq \sup_{z \in \mathcal{A}} \|z\|_{\mathcal{H}}$.

The application of this result to $\mathcal{H} = L^2(\Omega)$ and $\mathcal{A} = \mathcal{A}_\alpha = \{\rho_\alpha(\cdot - y)\}_{y \in \overline{\Omega}}$ already shows that the approximation rate for our exponential spline dictionary is at least $-\frac{1}{2}$.

4. Improved dictionary approximation rate

Our goal in this section is to find an approximation rate of order $O(n^{-\frac{1}{2}-s})$ with $s > 0$. We achieve this by taking advantage of the smoothness of our dictionary $\mathcal{A}_\alpha$.

4.1. Smoothness constant of the dictionary

The appropriate notion of smoothness for our purpose is in the sense of Lipschitz, as formalized in Definition 4.1.

Definition 4.1. Let $\Omega \subseteq \mathbb{R}^d$. A function $f : \Omega \rightarrow \mathbb{R}$ is said to belong to $\text{Lip}(s, L^\infty(\Omega))$ if the seminorm defined for $k \in \mathbb{N}$ and $0 < \alpha \leq 1$

$$\|f\|_{\text{Lip}(s, L^\infty(\Omega))} := \sup_{x, y \in \Omega} \frac{|D^k f(x) - D^k f(y)|}{|x - y|^\alpha} \quad (41)$$

satisfies $\|f\|_{\text{Lip}(s, L^\infty(\Omega))} < \infty$, where $s = k + \alpha$.

Definition 4.2. Let $\mathcal{X}$ be a Banach space, $\Omega \subseteq \mathbb{R}^d$ an open set, and $s > 0$. A function $G : \Omega \rightarrow \mathcal{X}$ is of smoothness class s , which we write $G \in \text{Lip}_\infty(s, \mathcal{X})$, if for all $\xi \in \mathcal{X}'$ the function $f_\xi : \Omega \rightarrow \mathbb{R}$ defined by

$$f_\xi(x) = \langle \xi, G(x) \rangle \quad (42)$$

satisfies

$$\|f_\xi\|_{\text{Lip}(s, L^\infty(\Omega))} \leq K \|\xi\|_{\mathcal{X}'} \quad (43)$$

for a constant $K < \infty$.

Our next goal is to parameterize our spline dictionary $\mathcal{A}_\alpha$ with a map that is in $\text{Lip}_\infty(s, L^2(\Omega))$, where $s = k + \alpha$ is a constant to determine. The parameterization map we use is

$$\begin{aligned} P : \overline{\Omega} &\rightarrow \mathcal{A}_\alpha \\ y &\mapsto \rho_\alpha(\cdot - y). \end{aligned} \quad (44)$$

We calculate the smoothness constant of P in [Theorem 4.3](#), whose proof is in [Appendix B](#).

Theorem 4.3. Let $L_\alpha = \prod_{n=1}^N (D - \alpha_n \text{Id})$ be a differential operator of order $N \in \mathbb{N}_{\geq 1}$. Let $\mathcal{A}_\alpha$ be its associated dictionary and let P be its parameterization map. Then $P \in \text{Lip}_\infty(s, L^2(\Omega))$, where $s = k + \alpha$ with $k = N - 1$ and $\alpha = \frac{1}{2}$.

4.2. Derivation of the improved approximation rate

We shall now refine the initial approximation rate of $n^{-\frac{1}{2}}$ by using the smoothness constant $s = k + \alpha = N - \frac{1}{2}$ in [Theorem 4.3](#) and a result of Siegel and Xu [[21](#), Theorem 2]. We state a simplified version of it in [Theorem 4.4](#), more suitable for our purposes.

Theorem 4.4 (Siegel, Xu). Let $s > 0$ and $\mathcal{H}$ be a Hilbert space. Suppose that Ω is a compact subset of $\mathbb{R}^d$, $P \in \text{Lip}_\infty(s, \mathcal{H})$, and the dictionary $\mathcal{A} \subseteq P(\Omega)$. Then for $f \in \overline{\Sigma(\mathcal{A})}$ we have that

$$\inf_{f_n \in \Sigma_n(\mathcal{A})} \|f - f_n\|_{\mathcal{H}} \lesssim \|f\|_{\overline{\Sigma(\mathcal{A})}} n^{-\frac{1}{2} - \frac{s}{d}}. \quad (45)$$

The implied constant depends only on s , d , and the parameterization map P .

We now state our main result, which provides an approximation rate for an arbitrary element of $\mathcal{M}^\alpha(\Omega)$ with exponential splines.

Theorem 4.5. Let $L_\alpha = \prod_{n=1}^N (D - \alpha_n \text{Id})$. Then, for all $f \in \mathcal{M}^\alpha(\Omega)$, there exists an exponential spline f_n with n knots such that

$$\|f - f_n\|_{L^2(\Omega)} \leq C \|L_\alpha f\|_{\mathcal{M}(\Omega)} n^{-N}, \quad (46)$$

where $C > 0$ is a constant that depends on N , α and Ω .

Proof. Let $f \in \mathcal{M}^\alpha(\Omega)$. We established in [Theorem A.3](#) that $f = f_1 + q$, where $f_1 \in \mathcal{M}_\rho^\alpha(\Omega)$ and $q \in \mathcal{N}_\alpha$. We apply [Theorem 4.4](#) to f_1 , with the parameterization map P defined in [\(44\)](#) and with the smoothness constant $s = N - \frac{1}{2}$ computed in [Theorem 4.3](#). This yields the existence of an n -term approximant $\tilde{f}_n$ (i.e., an exponential spline with n knots) such that

$$\|f_1 - \tilde{f}_n\|_{L^2(\Omega)} \leq C \|L_\alpha f_1\|_{\mathcal{M}(\Omega)} n^{-N} = C \|L_\alpha f\|_{\mathcal{M}(\Omega)} n^{-N}. \quad (47)$$

Defining

$$f_n = \tilde{f}_n + q, \quad (48)$$

which is still an exponential spline with n knots because $q \in \mathcal{N}_\alpha$, we obtain the inequality

$$\|f - f_n\|_{L^2(\Omega)} = \|f_1 + q - \tilde{f}_n - q\|_{L^2(\Omega)} = \|f_1 - \tilde{f}_n\|_{L^2(\Omega)} \leq C \|L_\alpha f\|_{\mathcal{M}(\Omega)} n^{-N}, \quad (49)$$

which is the desired result. $\square$

This new approximation rate n^{-N} depends on the order N of L_α , which constitutes a great improvement upon the rate $n^{-\frac{1}{2}}$. As N grows, so does the regularity of L_α -splines and, intuitively, so does their approximation capabilities as these smoother functions have more degrees of freedom when it comes to their design. This intuition is confirmed by the result of [Theorem 4.5](#). We note that this rate was achieved by means of *linear* approximation [13] for functions f such that $D^N f \in L^2(\Omega)$, through projections on a shift-invariant space spanned by exponential B-splines. This smoothness condition is much stronger than that of our [Theorem 4.5](#) as it only requires that $D^N f \in \mathcal{M}(\Omega)$. Furthermore, the Kolmogorov n -width (which captures the fundamental limits of linear approximation) of the unit ball in $\mathcal{M}^\alpha(\Omega)$ scales as $n^{-N-\frac{1}{2}}$ (cf., [21, Equation (1.33)]) combined with the equivalence of norms in [Proposition 5.3](#). Thus, we see that nonlinear methods of approximation are actually *necessary* to achieve the rate n^{-N} for functions in $\mathcal{M}^\alpha(\Omega)$.

5. Topological equivalence with $BV^N(\Omega)$ and optimality

Our final goal is to establish a topological equivalence between $\mathcal{M}^\alpha(\Omega)$ and $BV^N(\Omega)$, where

$$BV^N(\Omega) = \{f \in \mathcal{D}'(\Omega) : \|D^N f\|_{\mathcal{M}(\Omega)} < +\infty\} \quad (50)$$

is the space of functions with bounded variations of order N . Equipped with the norm

$$\|f\|_{BV^N(\Omega)} := \|D^N f\|_{\mathcal{M}(\Omega)} + \|f\|_{L^1(\Omega)}, \quad (51)$$

$BV^N(\Omega)$ is a Banach space. Note that a distribution $f \in \mathcal{D}'(\Omega)$ such that Df is a Radon measure is actually an $L^1(\Omega)$ function [11], which shows that, for $f \in BV^N(\Omega)$, $\|f\|_{L^1(\Omega)}$ is well-defined.

The space $\mathcal{M}^\alpha(\Omega)$ can be equipped with a variety of norms that turn it into a Banach space, for instance the norm induced by the biorthogonal system $(\phi_n, p_n)_{n=1}^{N_0}$ for $\mathcal{N}_\alpha$ given by $\|f\|_\phi = \|L_\alpha f\|_{\mathcal{M}(\Omega)} + \|\phi(f)\|_2$, as described in [Section 2.2](#). A norm that can establish a more meaningful link between $\mathcal{M}^\alpha(\Omega)$ and $BV^N(\Omega)$ is the following Sobolev-style norm

$$\|f\|_{L_\alpha,1} := \|L_\alpha f\|_{\mathcal{M}(\Omega)} + \|f\|_{L^1(\Omega)}. \quad (52)$$

We show in [Proposition 5.1](#) that $\|\cdot\|_\phi$ is equivalent to $\|\cdot\|_{L_\alpha,1}$.

Proposition 5.1. *The norms $\|\cdot\|_\phi$ and $\|\cdot\|_{L_\alpha,1}$ defined on $\mathcal{M}^\alpha(\Omega)$ are equivalent.*

Proof. We consider a biorthogonal system $(\phi_n, p_n)_{n=1}^{N_0}$ in which each ϕ_n is well-defined on $L^1(\Omega)$, so that the continuous operator $\phi = (\phi_1, \dots, \phi_{N_0})$ can be defined from $L^1(\Omega)$ to $\mathbb{R}^{N_0}$. Because Ω is bounded, $\mathcal{N}_\alpha \subseteq L^1(\Omega)$, which implies that $\phi_n \in (L^1(\Omega))' \subseteq (\mathcal{N}_\alpha)'$. Therefore, $(\phi_n, p_n)_{n=1}^{N_0}$ is a valid biorthogonal system for $\mathcal{N}_\alpha$. Next, continuity of ϕ corresponds to the

existence of a constant $A > 0$ such that $\|\phi(f)\|_2 \leq A\|f\|_{L^1(\Omega)}$, which implies the desired result, as

$$\|f\|_{\phi} = \|L_{\alpha} f\|_{\mathcal{M}(\Omega)} + \|\phi(f)\|_2 \leq \max(1, A)(\|L_{\alpha} f\|_{\mathcal{M}(\Omega)} + \|f\|_{L^1(\Omega)}) = B\|f\|_{L_{\alpha,1}}, \quad (53)$$

where $B = \max(1, A) > 0$. The lower bound for the norm equivalence then follows from the bounded inverse theorem. $\square$

We now establish the topological equivalence between $\mathcal{M}^{\alpha}(\Omega)$ and $\text{BV}^N(\Omega)$. We use [Lemma 5.2](#), whose proof is in [Appendix C](#).

Lemma 5.2. *For $f \in \text{BV}^N(\Omega)$, we have that $\|D^k f\|_{\mathcal{M}(\Omega)} \lesssim \|D^N f\|_{\mathcal{M}(\Omega)} + \|f\|_{\mathcal{M}(\Omega)}$ for all $k = 1, \dots, N-1$.*

Proposition 5.3. *The norms $\|\cdot\|_{L_{\alpha,1}}$ and $\|\cdot\|_{\text{BV}^N(\Omega)}$ are equivalent.*

Proof. From [Lemma 5.2](#),

$$\begin{aligned} \|f\|_{L_{\alpha,1}} &= \|L_{\alpha} f\|_{\mathcal{M}(\Omega)} + \|f\|_{L^1(\Omega)} \leq \sum_{n=0}^N |\beta_n| \|D^n f\|_{\mathcal{M}(\Omega)} + \|f\|_{L^1(\Omega)} \\ &\lesssim \|D^N f\|_{\mathcal{M}(\Omega)} + \|f\|_{L^1(\Omega)} = \|f\|_{\text{BV}^N(\Omega)}, \end{aligned} \quad (54)$$

where we wrote $L_{\alpha} = \prod_{n=1}^N (D - \alpha_n \text{Id}) = \sum_{n=0}^N \beta_n D^n$. The lower bound for the norm equivalence then follows from the bounded inverse theorem. $\square$

Remark 5.4. [Proposition 5.3](#) combined with the embeddings [\(6\)](#) readily proves the optimality (in the sense of entropy) of the rate n^{-N} in [Theorem 4.5](#). Recall that the L^2 -entropy numbers of a compact set $K \subset L^2(\Omega)$ is defined as

$$\varepsilon_n(K)_{L^2} = \inf\{\varepsilon > 0 : K \text{ is covered by } 2^n \text{ balls of radius } \varepsilon\}. \quad (55)$$

This quantity captures how precisely the elements of K can be specified with n bits. The notion of entropy of a compact set was introduced by Kolmogorov [\[14\]](#) as a way to quantify its compactness. The Birman–Solomyak theorem [\[1\]](#) establishes that the L^2 -entropy numbers of the unit balls in $U(B_{1,1}^N(\Omega))$ and $U(B_{1,\infty}^N(\Omega))$ both scale as

$$\varepsilon_n(U(B_{1,1}^N(\Omega)))_{L^2} \asymp \varepsilon_n(U(B_{1,\infty}^N(\Omega)))_{L^2} \asymp n^{-N}. \quad (56)$$

Therefore, by [\(6\)](#), the entropy numbers of the unit ball of $\text{BV}^N(\Omega)$ and hence of $\mathcal{M}^{\alpha}(\Omega)$ also scale as n^{-N} . Thus, we see that the rate in [Theorem 4.5](#) is sharp in the sense of entropy.

6. Conclusion

We have investigated the nonlinear approximation properties of exponential splines associated with linear differential operators. By specifying the native space $\mathcal{M}^{\alpha}(\Omega)$ on a finite interval Ω , we have extended the classic theory of polynomial-spline approximation to a broader and more flexible setting. We have endowed this space with a direct-sum Banach topology that accounts for the nontrivial null space of the differential operator and we have proved its continuous embedding into $L^2(\Omega)$. We constructed a well-behaved dictionary of exponential-spline components and showed that it yields an n -term approximation rate of n^{-N} in $L^2(\Omega)$

for functions in $\mathcal{M}^\alpha(\Omega)$. This rate strictly improves upon universal approximation bounds and generalizes known results for both exponential and polynomial splines. Moreover, we have established the topological equivalence between $\mathcal{M}^\alpha(\Omega)$ and the space $\text{BV}^N(\Omega)$, which allows us to leverage classic results from approximation theory to prove the optimality of our rate.

Acknowledgments

This work was supported by the European Research Council (ERC Project FunLearn) under Grant 101020573.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Direct-sum decomposition of the native space

The function $f \mapsto \|L_\alpha f\|_{\mathcal{M}(\Omega)}$ defined on $\mathcal{M}^\alpha(\Omega)$ is only a seminorm, due to the fact that $\mathcal{N}_\alpha$ is not empty. One way to handle this difficulty is to formally consider the quotient space $\mathcal{M}^\alpha(\Omega)/\mathcal{N}_\alpha$ and to define a composite norm on it. Concretely, this procedure allows us to write $\mathcal{M}^\alpha(\Omega)$ as the direct sum of a space on which $f \mapsto \|L_\alpha f\|_{\mathcal{M}(\Omega)}$ is a norm with $\mathcal{N}_\alpha$ equipped with its own norm as well. This direct-sum decomposition then allows us to define a composite norm on $\mathcal{M}^\alpha(\Omega)$. We do this using biorthogonal systems, defined in Section 2.2, and we compute the direct-sum decomposition in Theorem A.3. This technique has been used in [24] for the native space $\mathcal{M}^\alpha(\mathbb{R})$, as well as in [15,16] with seminorms stemming from Radon transforms in the context of ridge splines. We adapt here their results to our bounded interval Ω ,

We give an explicit example of a biorthogonal system for $\mathcal{N}_\alpha \subseteq \mathcal{M}^\alpha(\Omega)$. We know that $\mathcal{N}_\alpha$ is finite-dimensional with basis $\{p_1, \dots, p_{N_0}\}$. Therefore, by applying the Gram–Schmidt orthonormalization procedure as Ω is bounded, we obtain an orthonormal basis $\{u_1, \dots, u_{N_0}\}$ of $\mathcal{N}_\alpha$. Next, we define, for $n = 1, \dots, N_0$ and $f \in \mathcal{N}_\alpha$,

$$\phi_n(f) = \langle u_n, f \rangle_{L^2(\Omega)}. \quad (\text{A.1})$$

Then $(\phi_n, u_n)_{n=1}^{N_0}$ is a biorthogonal system for $\mathcal{N}_\alpha$, as the ϕ_i are linear, continuous and satisfy $\phi_n(u_m) = \delta_{nm}$ by orthonormality of $\{u_1, \dots, u_{N_0}\}$.

In what follows, we consider a biorthogonal system $(\phi_n, p_n)_{n=1}^{N_0}$ for $\mathcal{N}_\alpha$, for example the one defined above. We first state an adaptation of Theorem 4 in [24] to bounded sets, which helps us to invert L_α under certain conditions. Its proof is the same as in [24], up to consideration of restrictions of measures in $\mathcal{M}(\mathbb{R})$ to $\mathcal{M}(\Omega)$.

Theorem A.1. *Let $N_0 = \dim(\mathcal{N}_\alpha)$ and let $(\phi_n, p_n)_{n=1}^{N_0}$ be a biorthogonal system for $\mathcal{N}_\alpha$. Then L_α admits a stable right inverse $L_{\alpha,\phi}^{-1} : \mathcal{M}(\Omega) \rightarrow \mathcal{M}^\alpha(\Omega)$, $w \mapsto L_{\alpha,\phi}^{-1} w = \int_\Omega g_\phi(\cdot, y)w(y)dy$ with kernel*

$$g_\phi(x, y) = \rho_\alpha(x - y) - \sum_{n=1}^{N_0} p_n(x)q_n(y), \quad (\text{A.2})$$

where $q_n(y) = \langle \phi_n, \rho_\alpha(\cdot - y) \rangle$ for $y \in \Omega$. It satisfies:

- *Right inverse property:* $L_\alpha L_{\alpha,\phi}^{-1} \varphi = \varphi$,

- *Boundary conditions:* $\phi(L_{\alpha,\phi}^{-1}\varphi) = \mathbf{0}$.

We also define

$$\mathcal{M}_\phi^\alpha(\Omega) = \{f \in \mathcal{M}^\alpha(\Omega) : \phi(f) = \mathbf{0}\}, \quad (\text{A.3})$$

which is a subspace of $\mathcal{M}^\alpha(\Omega)$. Lemma A.2 characterizes the surjectivity of $L_{\alpha,\phi}^{-1}$ on this space.

Lemma A.2. *Let $f \in \mathcal{M}_\phi^\alpha(\Omega)$. Then, there exists a unique measure $w \in \mathcal{M}(\Omega)$ such that $f = L_{\alpha,\phi}^{-1}w$.*

Proof. Let $f \in \mathcal{M}_\phi^\alpha(\Omega)$ and consider the linear functional

$$\begin{aligned} T_f : \mathcal{D}(\Omega) &\rightarrow \mathbb{R} \\ \varphi &\mapsto \langle L_\alpha f, \varphi \rangle. \end{aligned} \quad (\text{A.4})$$

We show that it is continuous by showing that it is bounded, since

$$\sup_{\varphi \in \mathcal{D}(\Omega)} \frac{T_f(\varphi)}{\|\varphi\|_\infty} = \sup_{\varphi \in \mathcal{D}(\Omega)} \frac{\langle L_\alpha f, \varphi \rangle}{\|\varphi\|_\infty} = \|L_\alpha f\|_{\mathcal{M}(\Omega)} < +\infty. \quad (\text{A.5})$$

The left-hand side of this inequality is finite because $f \in \mathcal{M}_\phi^\alpha(\Omega) \subseteq \mathcal{M}^\alpha(\Omega)$, which shows that T_f is a continuous linear functional on $\mathcal{D}(\Omega)$. By the Riesz–Markov–Kakutani representation theorem, there exists a unique Radon measure μ on Ω such that, for any $\varphi \in \mathcal{D}(\Omega)$,

$$T_f(\varphi) = \int_\Omega \varphi(y) d\mu(y) = \int_\Omega \varphi(y)w(y) dy, \quad (\text{A.6})$$

where $w \in \mathcal{M}(\Omega)$ uniquely defines μ .

We now define the function $g_\mu(x) = \int_\Omega g_\phi(x, y) d\mu(y) = \int_\Omega g_\phi(x, y)w(y) dy = L_{\alpha,\phi}^{-1}\{w\}(x)$. Then, for any $\varphi \in \mathcal{D}(\Omega)$,

$$\begin{aligned} \langle L_\alpha g_\mu, \varphi \rangle &= \langle L_\alpha L_{\alpha,\phi}^{-1}w, \varphi \rangle = \langle w, \varphi \rangle = \int_\Omega \varphi(y)w(y) dy = \int_\Omega \varphi(y) d\mu(y) \\ &= T_f(\varphi) = \langle L_\alpha f, \varphi \rangle. \end{aligned} \quad (\text{A.7})$$

The second equality is a consequence of the definition of $L_{\alpha,\phi}^{-1}$. This shows that $L_\alpha g_\mu = w = L_\alpha f$ in $\mathcal{D}'(\Omega)$, which leads to $g_\mu = L_{\alpha,\phi}^{-1}w = f + c$, where $c \in \mathcal{N}_\alpha$. However, since $f \in \mathcal{M}_\phi^\alpha(\Omega)$, we know that $\phi(f) = \mathbf{0}$. Coupled with the fact that $\phi(L_{\alpha,\phi}^{-1}w) = \mathbf{0}$, we conclude that

$$\phi(L_{\alpha,\phi}^{-1}w) = \phi(f) + \phi(c) \implies \mathbf{0} = \mathbf{0} + \phi(c) \implies c = 0. \quad (\text{A.8})$$

The last implication in (A.8) comes from the fact that $\mathcal{M}_\phi^\alpha(\Omega) \cap \mathcal{N}_\alpha = \{0\}$. Indeed, let $h \in \mathcal{M}_\phi^\alpha(\Omega) \cap \mathcal{N}_\alpha$. Because $h \in \mathcal{N}_\alpha$, we can write that $h = \sum_{n=1}^{N_0} \langle \phi_n, h \rangle p_n$. But, because $h \in \mathcal{M}_\phi^\alpha(\Omega)$, we have that $\phi(h) = \mathbf{0}$ and, hence, that $h = 0$, which shows that $\mathcal{M}_\phi^\alpha(\Omega) \cap \mathcal{N}_\alpha = \{0\}$. This shows that, for any $f \in \mathcal{M}_\phi^\alpha(\Omega)$, there exists a unique measure $w \in \mathcal{M}(\Omega)$ such that $f = L_{\alpha,\phi}^{-1}w$. $\square$

We now have the tools to provide a direct-sum decomposition of $\mathcal{M}^\alpha(\Omega)$. The proof draws inspiration from [24, Theorem 5].

Theorem A.3. We have that $\mathcal{M}^\alpha(\Omega) = \mathcal{M}_\phi^\alpha(\Omega) \oplus \mathcal{N}_\alpha$. Consequently, for every $f \in \mathcal{M}^\alpha(\Omega)$, there exists a Radon measure $\mu \in \mathcal{M}(\Omega)$ and a function $c \in \mathcal{N}_\alpha$ such that

$$f(x) = \int_{\Omega} \rho_\alpha(x - y) d\mu(y) + c(x). \quad (\text{A.9})$$

Proof. Let $N_0 = \dim(\mathcal{N}_\alpha)$ and let $(\phi_n, p_n)_{n=1}^{N_0}$ be a biorthogonal system for $\mathcal{N}_\alpha$. Let $f \in \mathcal{M}^\alpha(\Omega)$. Since $\mathcal{N}_\alpha$ is a closed subspace of $\mathcal{M}^\alpha(\Omega)$, by continuity of the projection operator $\text{Proj}_{\mathcal{N}_\alpha} : f \mapsto \sum_{n=1}^{N_0} \langle \phi_n, f \rangle p_n$, we can write that $f = f_1 + q$, where $q = \text{Proj}_{\mathcal{N}_\alpha}(f)$. The function $f_1 = f - q$ is such that

$$\begin{aligned} \phi(f_1) &= \phi(f) - \phi(q) = \phi(f) - \phi\left(\sum_{n=1}^{N_0} \langle \phi_n, f \rangle p_n\right) \\ &= \phi(f) - \sum_{n=1}^{N_0} \langle \phi_n, f \rangle \underbrace{\phi(p_n)}_{=e_n} = \phi(f) - \phi(f) = 0. \end{aligned} \quad (\text{A.10})$$

This means that $f_1 \in \mathcal{M}_\phi^\alpha(\Omega)$, which implies that $\mathcal{M}^\alpha(\Omega) = \mathcal{M}_\phi^\alpha(\Omega) + \mathcal{N}_\alpha$. Since we showed that $\mathcal{M}_\phi^\alpha(\Omega) \cap \mathcal{N}_\alpha = \{0\}$, we can conclude that

$$\mathcal{M}^\alpha(\Omega) = \mathcal{M}_\phi^\alpha(\Omega) \oplus \mathcal{N}_\alpha. \quad (\text{A.11})$$

Next, by [Lemma A.2](#), there exists a unique measure $w \in \mathcal{M}(\Omega)$ such that $f_1 = L_{\alpha, \phi}^{-1} w$. By definition of $L_{\alpha, \phi}^{-1}$, we have that $w = L_\alpha f_1 = L_\alpha(f - q) = L_\alpha f$. This finally leads to

$$\begin{aligned} f &= L_{\alpha, \phi}^{-1} w + q = \int_{\Omega} g_\phi(\cdot, y) w(y) dy + q \\ &= \int_{\Omega} \rho_\alpha(\cdot - y) \underbrace{w(y) dy}_{=d\mu(y)} - \underbrace{\sum_{n=1}^{N_0} p_n(\cdot) \int_{\Omega} q_n(y) dy}_{=c \in \mathcal{N}_\alpha} + q, \end{aligned} \quad (\text{A.12})$$

which shows that, for any $f \in \mathcal{M}^\alpha(\Omega)$, $f(x) = \int_{\Omega} \rho_\alpha(x - y) d\mu(y) + c(x)$ where $c \in \mathcal{N}_\alpha$. $\square$

The direct-sum decomposition $\mathcal{M}^\alpha(\Omega) = \mathcal{M}_\phi^\alpha(\Omega) \oplus \mathcal{N}_\alpha$ of [Theorem A.3](#) allows us to define the composite norm

$$\|f\|_\phi := \|L_\alpha f\|_{\mathcal{M}(\Omega)} + \|\phi(f)\|_2 \quad (\text{A.13})$$

for any $f \in \mathcal{M}^\alpha(\Omega)$, which turns $\mathcal{M}^\alpha(\Omega)$ into a Banach space, as the spaces $(\mathcal{M}_\phi^\alpha(\Omega), \|L_\alpha \cdot\|_{\mathcal{M}(\Omega)})$ and $(\mathcal{N}_\alpha, \|\phi(\cdot)\|_2)$ are themselves Banach spaces [\[24\]](#).

Appendix B. Proof of [Theorem 4.3](#)

We first need a technical lemma.

Lemma B.1. Let L_α be a differential operator of order $N \in \mathbb{N}_{\geq 1}$, ρ_α its causal Green's function, $k = N - 1$ and $\alpha = \frac{1}{2}$. Then, for all $y, y' \in \Omega$,

$$\|D^k \rho_\alpha(\cdot - y) - D^k \rho_\alpha(\cdot - y')\|_{L^2(\Omega)} \leq K |y - y'|^\alpha \quad (\text{B.1})$$

for some $K > 0$.

Proof. We begin by using the property that the causal Green's function of $L_\alpha = \prod_{n=1}^N (D - \alpha_n \text{Id})$ is the convolution of the causal Green's functions ρ_{α_n} of each first-order term, with the latter being locally bounded and discontinuous at $x = 0$. More generally, if $N \geq 2$, $\rho_\alpha \in C^{N-2}(\mathbb{R})$. This implies that $D^k \rho_\alpha = D^{N-1} \rho_\alpha$ is discontinuous at $x = 0$ as well. Next, supposing without loss of generality that $y' > y$, we have that

$$\|D^k \rho_\alpha(\cdot - y) - D^k \rho_\alpha(\cdot - y')\|_{L^2(\Omega)}^2 = \int_{\Omega} |D^k \rho_\alpha(x - y) - D^k \rho_\alpha(x - y')|^2 dx. \quad (\text{B.2})$$

Using the causality of ρ_α to separate the integral, and using the fact that $\Omega = (a, b)$, we write the integral in (B.2) as

$$\int_y^{y'} |D^k \rho_\alpha(x - y)|^2 dx + \int_{y'}^b |D^k \rho_\alpha(x - y) - D^k \rho_\alpha(x - y')|^2 dx = I_1 + I_2. \quad (\text{B.3})$$

We proceed to find an upper bound for I_1 and I_2 separately. The case of I_1 is straightforward since, for $y, y' \in \Omega$ fixed, the function $h(x) = D^k \rho_\alpha(x - y)$ is continuous on $[y, y']$. Then, we have that

$$I_1 \leq |y' - y| \cdot \sup_{y, y' \in \Omega} \left\{ \max_{x \in [y, y']} |D^k \rho_\alpha(x - y)|^2 \right\} = K_1 |y - y'|. \quad (\text{B.4})$$

As for I_2 , we fix $x \in [y', b)$ and let $g_x(y) = D^k \rho_\alpha(x - y)$. This allows us to write

$$D^k \rho_\alpha(x - y) - D^k \rho_\alpha(x - y') = g_x(y) - g_x(y'). \quad (\text{B.5})$$

As g_x is continuous on $[y, y']$ and differentiable on (y, y') , we then apply the mean-value theorem to obtain the upper bound

$$|g_x(y) - g_x(y')| \leq K_{x, y, y'}^2 |y - y'| \leq \left(\sup_{x, y, y' \in \Omega} K_{x, y, y'}^2 \right) \text{diam}(\Omega) |y - y'|, \quad (\text{B.6})$$

where $K_{x, y, y'} = \sup_{t \in (y, y')} |g'_x(t)|$ depends on x, y , and y' . We use this to upper-bound I_2 as

$$\begin{aligned} I_2 &\leq (b - y') \left(\sup_{x, y, y' \in \Omega} K_{x, y, y'}^2 \right) \text{diam}(\Omega) |y - y'| \\ &\leq \left(\sup_{x, y, y' \in \Omega} K_{x, y, y'}^2 \right) \text{diam}(\Omega)^2 |y - y'| = K_2 |y - y'|. \end{aligned} \quad (\text{B.7})$$

Putting (B.4) and (B.7) together, we get

$$\|D^k \rho_\alpha(\cdot - y) - D^k \rho_\alpha(\cdot - y')\|_{L^2(\Omega)}^2 = I_1 + I_2 \leq (K_1 + K_2) |y - y'| = \tilde{K} |y - y'|, \quad (\text{B.8})$$

which finally leads to

$$\|D^k \rho_\alpha(\cdot - y) - D^k \rho_\alpha(\cdot - y')\|_{L^2(\Omega)} \leq (\tilde{K})^{\frac{1}{2}} |y - y'|^{\frac{1}{2}} = K |y - y'|^{\frac{1}{2}}, \quad (\text{B.9})$$

as desired. $\square$

Proof of Theorem 4.3. We begin by recalling that $L^2(\Omega)$ is its own dual. We also notice that, for $y, y' \in \Omega$, $y' \neq y$, the statement we want to show is, for all $\xi \in L^2(\Omega)$,

$$\|f_\xi\|_{\text{Lip}(s, L^\infty(U))} \leq K \|\xi\|_{L^2(\Omega)} \iff |D^k f_\xi(y) - D^k f_\xi(y')| \leq K \|\xi\|_{L^2(\Omega)} |y - y'|^\alpha. \quad (\text{B.10})$$

Recalling that $f_\xi(y) = \langle \xi, P(y) \rangle = \int_\Omega \xi(x) \rho_\alpha(x - y) dx$, we have the bound

$$\begin{aligned} |D^k f_\xi(y) - D^k f_\xi(y')| &\leq \int_\Omega |\xi(x)| \cdot |D^k \rho_\alpha(x - y) - D^k \rho_\alpha(x - y')| dx \\ &\leq \|\xi\|_{L^2(\Omega)} \|D^k \rho_\alpha(\cdot - y) - D^k \rho_\alpha(\cdot - y')\|_{L^2(\Omega)}, \end{aligned} \quad (\text{B.11})$$

where we used the Cauchy–Schwarz inequality. It is hence sufficient to show that

$$\|D^k \rho_\alpha(\cdot - y) - D^k \rho_\alpha(\cdot - y')\|_{L^2(\Omega)} \leq K |y - y'|^\alpha \quad (\text{B.12})$$

for some $K > 0$, where $k = N - 1$ and $\alpha = \frac{1}{2}$. Using [Lemma B.1](#) above, the result is immediate. $\square$

Appendix C. Proof of [Lemma 5.2](#)

When f is a function that belongs to the Sobolev space $W^{N,1}(\Omega)$, where N is the highest order derivative in L_α , the Gagliardo–Nirenberg interpolation inequality on a bounded domain [\[10\]](#) implies that

$$\|D^k f\|_{L^1(\Omega)} \lesssim \|D^N f\|_{L^1(\Omega)} + \|f\|_{L^1(\Omega)} \quad (\text{C.1})$$

for all $k = 1, \dots, N - 1$. Before extending this inequality to the $\mathcal{M}(\Omega)$ norm for $f \in \text{BV}^N(\Omega)$, we need a result on the approximation of BV functions using smooth functions, which we state in [Theorem C.1](#).

Theorem C.1 ([\[11, Theorem 1.17\]](#)). *Let $f \in \text{BV}^1(\Omega)$. Then, there exists a sequence $(f_n)_{n \in \mathbb{N}}$ in $C^\infty(\Omega)$ such that $\|f_n\|_{L^1(\Omega)} \rightarrow \|f\|_{L^1(\Omega)}$ and $\|D f_n\|_{\mathcal{M}(\Omega)} \rightarrow \|D f\|_{\mathcal{M}(\Omega)}$ as $n \rightarrow +\infty$.*

This theorem generalizes by induction to $f \in \text{BV}^N(\Omega)$ for any $N \in \mathbb{N}_{\geq 1}$, and allows us to prove [Lemma 5.2](#).

Proof of [Lemma 5.2](#). We prove the case $N = 2$, as the result for $N \geq 3$ can be deduced by induction. Let $f \in \text{BV}^2(\Omega)$, and consider a sequence $(f_n)_{n \in \mathbb{N}}$ in $C^\infty(\Omega)$ given by [Theorem C.1](#). As Ω is bounded, $f_n \in W^{2,1}(\Omega)$ for all $n \in \mathbb{N}$. Applying the Gagliardo–Nirenberg inequality to f_n , we obtain

$$\|D f_n\|_{\mathcal{M}(\Omega)} = \|D f_n\|_{L^1(\Omega)} \lesssim \|f_n\|_{L^1(\Omega)} + \|D^2 f_n\|_{L^1(\Omega)} = \|f_n\|_{\mathcal{M}(\Omega)} + \|D^2 f_n\|_{\mathcal{M}(\Omega)}. \quad (\text{C.2})$$

Taking the limit $n \rightarrow +\infty$ and using [Theorem C.1](#) results in the inequality

$$\|D f\|_{\mathcal{M}(\Omega)} \lesssim \|f\|_{\mathcal{M}(\Omega)} + \|D^2 f\|_{\mathcal{M}(\Omega)}, \quad (\text{C.3})$$

which is the desired result. $\square$

Data availability

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

References

- [1] M.S. Birman, M.Z. Solomyak, Piecewise-polynomial approximations of functions of the classes W_p^α , *Matematicheskii Sbornik* 115 (3) (1967) 331–355.
- [2] A. Cohen, W. Dahmen, I. Daubechies, R.A. DeVore, Harmonic analysis of the space BV, *Rev. Mat. Iberoam.* 19 (1) (2003) 235–263, <http://dx.doi.org/10.4171/rmi/345>.
- [3] C. Conti, L. Gemignani, L. Romani, From approximating to interpolatory non-stationary subdivision schemes with the same generation properties, *Adv. Comput. Math.* 35 (2–4) (2011) 217–241, <http://dx.doi.org/10.1007/s10444-011-9175-6>.
- [4] C. Conti, L. Romani, Affine combination of B-spline subdivision masks and its non-stationary counterparts, *BIT Numer. Math.* 50 (2) (2010) 269–299, <http://dx.doi.org/10.1007/s10543-010-0263-7>.
- [5] C. Conti, L. Romani, M. Unser, Ellipse-preserving Hermite interpolation and subdivision, *J. Math. Anal. Appl.* 426 (1) (2015) 211–227, <http://dx.doi.org/10.1016/j.jmaa.2015.01.017>.
- [6] R. Delgado-Gonzalo, P. Thévenaz, C.S. Seelamantula, M. Unser, Snakes with an ellipse-reproducing property, *IEEE Trans. Image Process.* 21 (3) (2012) 1258–1271, <http://dx.doi.org/10.1109/tip.2011.2169975>.
- [7] R. Delgado-Gonzalo, P. Thévenaz, M. Unser, Exponential splines and minimal-support bases for curve representation, *Comput. Aided Geom. Design* 29 (2) (2012) 109–128, <http://dx.doi.org/10.1016/j.cagd.2011.10.005>.
- [8] R.A. DeVore, Nonlinear approximation, *Acta Numerica* 7 (1998) 51–150, <http://dx.doi.org/10.1017/s0962492900002816>.
- [9] R.A. DeVore, G.G. Lorentz, *Constructive approximation*, in: *Grundlehren der mathematischen Wissenschaften*, (303) Springer Berlin, Heidelberg, 1993.
- [10] A. Fiorenza, M.R. Formica, T.G. Roskovec, F. Soudský, Detailed proof of classical gagliardo–nirenberg interpolation inequality with historical remarks, *Z. Anal. Ihre Anwend.* 40 (2) (2021) 217–236, <http://dx.doi.org/10.4171/zaa/1681>.
- [11] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, Birkhäuser Boston, 1984, <http://dx.doi.org/10.1007/978-1-4684-9486-0>.
- [12] I. Khalidov, J. Fadili, F. Lazeyras, D. Van De Ville, M. Unser, Activelets: Wavelets for sparse representation of hemodynamic responses, *Signal Process.* 91 (12) (2011) 2810–2821, <http://dx.doi.org/10.1016/j.sigpro.2011.03.008>.
- [13] I. Khalidov, M. Unser, From differential equations to the construction of new wavelet-like bases, *IEEE Trans. Signal Process.* 54 (4) (2006) 1256–1267, <http://dx.doi.org/10.1109/tsp.2006.870544>.
- [14] A.N. Kolmogorov, On linear dimensionality of topological vector spaces, in: *Doklady Akademii Nauk*, vol. 120, (2) Russian Academy of Sciences, 1958, pp. 239–241.
- [15] R. Parhi, R.D. Nowak, Banach space representer theorems for neural networks and ridge splines, *J. Mach. Learn. Res.* 22 (43) (2021) 1–40.
- [16] R. Parhi, R.D. Nowak, Near-minimax optimal estimation with shallow ReLU neural networks, *IEEE Trans. Inform. Theory* 69 (2) (2023) 1125–1140, <http://dx.doi.org/10.1109/tit.2022.3208653>.
- [17] J. Peetre, *New thoughts on besov spaces*, in: *Duke University Mathematics Series*, Mathematics Department, Duke University, 1976.
- [18] G. Pisier, Remarques sur un résultat non publié de B. Maurey, *Séminaire Analyse Fonctionnelle (Dit "Maurey-Schwartz")* (1981) 1–12.
- [19] L. Romani, From approximating subdivision schemes for exponential splines to high-performance interpolating algorithms, *J. Comput. Appl. Math.* 224 (1) (2009) 383–396, <http://dx.doi.org/10.1016/j.cam.2008.05.013>.
- [20] J.W. Siegel, J. Xu, Characterization of the variation spaces corresponding to shallow neural networks, *Constr. Approx.* 57 (3) (2023) 1109–1132, <http://dx.doi.org/10.1007/s00365-023-09626-4>.
- [21] J.W. Siegel, J. Xu, Sharp bounds on the approximation rates, metric entropy, and n-widths of shallow neural networks, *Found. Comput. Math.* 24 (4) (2024) 481–537, <http://dx.doi.org/10.1007/s10208-022-09595-3>.
- [22] V. Uhlmann, R. Delgado-Gonzalo, C. Conti, L. Romani, M. Unser, Exponential Hermite splines for the analysis of biomedical images, in: *2014 IEEE International Conference on Acoustics, Speech and Signal Processing*, ICASSP, IEEE, 2014, pp. 1631–1634, <http://dx.doi.org/10.1109/icassp.2014.6853874>.
- [23] M. Unser, T. Blu, Cardinal exponential splines: Part I-Theory and filtering algorithms, *IEEE Trans. Signal Process.* 53 (4) (2005) 1425–1438, <http://dx.doi.org/10.1109/TSP.2005.843700>.
- [24] M. Unser, J. Fageot, J.P. Ward, Splines are universal solutions of linear inverse problems with generalized TV regularization, *SIAM Rev.* 59 (4) (2017) 769–793, <http://dx.doi.org/10.1137/16M1061199>, (Accessed 2024-07-28).
- [25] V.S. Vladimirov, *Methods of the Theory of Generalized Functions*, CRC Press, 2002, <http://dx.doi.org/10.1201/9781482288162>.