

Support and Approximation Properties of Hermite Splines

Julien Fageot*, Shayan Aziznejad*, Michael Unser*, Virginie Uhlmann*^{†‡}

June 18, 2019

Abstract: In this paper, we formally investigate two mathematical aspects of Hermite splines which translate to features that are relevant to their practical applications. We first demonstrate that Hermite splines are maximally localized in the sense that their support sizes are minimal among pairs of functions with identical reproduction properties. Then, we precisely quantify the approximation power of Hermite splines for reconstructing functions and their derivatives, and show that they are asymptotically identical to cubic B-splines for these tasks. Hermite splines therefore combine optimal localization and excellent approximation power, while retaining interpolation properties and closed-form expression, in contrast to existing similar approaches. These findings shed a new light on the convenience of Hermite splines for use in computer graphics and geometrical design.

1 Introduction

In his seminal 1973 monograph on cardinal interpolation and spline functions [1], I.J. Schoenberg explains and characterizes B-spline interpolation, which still inspire researchers and yield exciting applications nowadays. In the same work, he also sets the basis of Hermite interpolation [2, 3]. In the classical B-spline framework, a continuous-domain function is constructed from a discrete sequence of samples. By contrast, the Hermite interpolation problem involves two sequences of discrete samples, which impose constraints not only on the resulting interpolated function but also on its derivatives up to a given order.

Curves in the plane or surfaces in the volume can be constructed from one-dimensional interpolation schemes $\mathbb{R} \rightarrow \mathbb{R}$ by interpolating along each spatial coordinate. The practical value of Hermite splines in this context is to offer tangential control on the interpolated curve. This can be easily understood through their link with Bézier curves [7]. The latter lie at the heart of vector graphics and are popular tools for computer-aided geometrical

*Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne (EPFL), Station 17, 1015 Lausanne, Switzerland.

[†]European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Cambridge CB10 1SD, UK.

[‡]Corresponding author.

Contact: julien.fageot@epfl.ch; shayan.aziznejad@epfl.ch; michael.unser@epfl.ch; uhlmann@ebi.ac.uk.

design and modelling [8, 9, 10]. Hermite splines are also an interesting option for the design of multiwavelets, that is, wavelets with multiple generators [12, 13] because of their small support. In practice, Hermite splines thus provides a suitable solution to a number of problems, whether with respect to simplicity of construction, efficiency, or convenience. This hands-on intuition can be translated to formal properties of Hermite splines and mathematically characterized. As examples, the joint interpolation properties of Hermite splines (see Section 1.2), which ensure that, at integer values, the interpolated function exactly matches the sequences of samples and derivatives samples that were used to build it; their smoothness properties [11], which guarantee low curvature of the interpolated curve under some mild conditions; and their statistical optimality (in terms of MMSE) for the reconstruction of second-order Brownian motion from direct and first derivative samples [14]. In that spirit, we investigate in this work the theoretical counterpart of two additional features that are observed to grant Hermite splines their practical usefulness.

1.1 Contributions

Our contributions state the minimal support property of Hermite splines and investigate their approximation power. In the following, we describe the practical observations motivating them, the results themselves, and related works.

Minimal support property. The short support of Hermite splines is an important feature making them attractive in practice. The size of the support relates to the local extent of modifications on the continuously-defined spline curve. We formally demonstrate that Hermite splines have the minimal support property among basis functions generating cubic and quadratic splines. When dealing with B-splines, there is a trade-off between the ability to reproduce smooth functions, which only depends on the B-spline order and thus maximum polynomial degree, and the possibility to allow for sharp transitions, as the regularity of the spline increases with its order. Typically, cubic splines can efficiently reproduce smooth functions, but lack the power to capture irregular transitions. On the other hand, quadratic splines have lesser approximation power, but are preferred when dealing with sharp transitions. Hermite splines combine these two strengths in one scheme and are, in terms of support size, better than any two-functions scheme, including the one composed of the classical cubic and quadratic B-splines. In addition, we also show that one necessarily requires two generators to achieve this optimality. This result relates to similar ones involving a single generator [15, 16].

Rate of decay of the approximation error. Hermite splines can provide faithful approximation reasonably fast as the number of parameters increases. This feature relates to the rate of decay of the approximation error. Numerous works approach these questions by restricting themselves to a specific interpolation framework, such as [17, 18] for the specific case of Hermite approximation. They provide precise estimations of the optimal

bound on the approximation error relying on L_∞ norms, but do not allow comparison with other schemes. In contrast, approaches have been developed for single [19] and multiple generators [20] to provide a unifying setting for comparison. Their rigorous mathematical framework offer generalized measures that can be applied to a wide variety of basis functions for estimating approximation constants. Hermite interpolation however violates some of the core assumptions of the analysis, preventing it from being studied with these tools. Taking strong inspiration from this previous work, we provide a novel study of Hermite approximation with a analogous analysis strategy, relieving the boundedness assumptions and allowing to consider additional spline approximation schemes. We precisely quantify the rate of decay of the approximation error of the Hermite scheme and quantitatively estimate the corresponding approximation constants. Hermite splines have excellent approximation properties when it comes to reconstructing a function and its first derivative, actually being close to optimal. The investigation of approximation error on the derivative relates to [21, 22], although following a completely different line: while they focused on the reconstruction of the derivative from signal samples, the Hermite scheme grants direct access on the function and derivative samples, allowing to reconstruct the derivative in a multi-function setting.

Pioneering works cover the study of spline schemes to an impressive degree of generality [23, 24, 25, 26]. They include results on minimum support and approximation errors in a large variety of cases. Regarding minimum support, these previous results are however restricted to single generator schemes. To the best of our knowledge, we are not aware of previous work considering multi-generators and therefore covering the Hermite case. These studies also do not focus on providing a way to quantify the approximation error constants, thus not enabling comparison between schemes of the same order. As mentioned beforehand, frameworks have been proposed to address this, but the Hermite scheme falls outside of their core hypotheses, therefore requiring a formal adaptation.

1.2 Hermite Splines

Schoenberg defines the cardinal cubic Hermite interpolation problem as follows [2, 3]. Knowing the discrete sequences of numbers $c[k]$ and $d[k]$, $k \in \mathbb{Z}$, we look for a continuously-defined function $f_{\text{Her}}(t)$, $t \in \mathbb{R}$, satisfying $f_{\text{Her}}(k) = c[k]$, $f'_{\text{Her}}(k) = d[k]$ for all $k \in \mathbb{Z}$, such that f_{Her} is piecewise polynomial of degree at most 3 and once differentiable with continuous derivative at the integers. The existence and uniqueness of the solution is guaranteed [2, Theorem 1] for any sequences $c = (c[k])$ and $d = (d[k])$ bounded by a polynomial, but we shall restrict to sequences in $\ell_2(\mathbb{Z})$ thereafter. In [3], it is shown that the Hermite spline f_{Her} associated to the sequences c and d can be expressed as

$$f_{\text{Her}}(t) = \sum_{k \in \mathbb{Z}} c[k] \phi_1(\cdot - k) + d[k] \phi_2(\cdot - k) \quad (1)$$

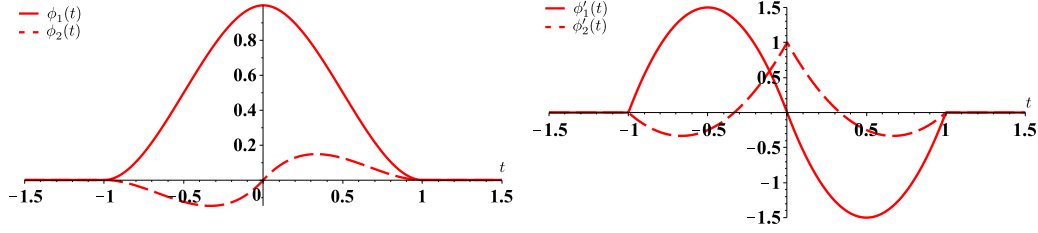


Figure 1 – Cubic Hermite splines ϕ_1 and ϕ_2 . The two functions and their derivatives are vanishing at the integers with the exception of $\phi_1(0) = 1$ and $\phi'_2(0) = 1$ (interpolation properties). They are supported in $[-1, 1]$.

where the functions ϕ_1 and ϕ_2 are given by

$$\phi_1(t) = (2|t| + 1)(|t| - 1)^2 \mathbb{1}_{0 \leq |t| \leq 1}, \quad (2)$$

$$\phi_2(t) = t(|t| - 1)^2 \mathbb{1}_{0 \leq |t| \leq 1}. \quad (3)$$

In addition to their fairly simple analytical expression, the cubic Hermite splines have other important properties. First, they are of finite support in $[-1, 1]$. Moreover, the generating functions ϕ_1 , ϕ_2 and their derivatives ϕ'_1 , ϕ'_2 satisfy the joint interpolation conditions

$$\phi_1(k) = \delta[k], \quad \phi'_2(k) = \delta[k], \quad \phi'_1(k) = 0, \quad \phi_2(k) = 0, \quad (4)$$

for all $k \in \mathbb{Z}$, where $\delta[k]$ is the discrete unit impulse. The functions and their first derivative are depicted in Figure 1, where the interpolation properties can easily be observed. The functions ϕ_1 and ϕ_2 are deeply intertwined as $c[k] = f_{\text{Her}}(k)$ and $d[k] = f'_{\text{Her}}(k)$ in (1). The cubic Hermite splines are differentiable with continuous derivatives at the integer knots points $t = k$. As a result, functions generated by cubic Hermite splines are C^1 -continuous piecewise-cubic polynomials with knots at integer locations.

2 Minimum Support Properties of Hermite Splines

B-splines are known to be maximally localized, meaning that they are compactly supported with minimal support properties among functions with the same approximation properties [27, 28]. Hermite splines possess a similar fundamental minimal support property: they are of minimal support among the pair of functions being able to generate both quadratic and cubic splines (Theorem 2). In addition, a single generating function is not sufficient for this purpose (Proposition 4). This demonstrates that Hermite splines are maximally localized for the purpose of representing piecewise-cubic functions that are continuously differentiable, as exploited for instance in image processing [11].

2.1 Shift invariant spaces and support properties

Consider a set of $N \geq 1$ functions $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_N) \in (L_2(\mathbb{R}))^N$ and define the space of functions

$$V(\boldsymbol{\varphi}) = \left\{ \sum_{i=1}^N \sum_{k \in \mathbb{Z}} c_i[k] \varphi_i(\cdot - k) : \mathbf{c} = (c_1, \dots, c_N) \in (\ell_2(\mathbb{Z}))^N \right\}. \quad (5)$$

We say that φ_i are *basis functions* of the set $V(\boldsymbol{\varphi})$. The space $V(\boldsymbol{\varphi})$ is *shift-invariant* in the sense that, by definition, $f(\cdot - k)$ is in $V(\boldsymbol{\varphi})$ for $f \in V(\boldsymbol{\varphi})$ and $k \in \mathbb{Z}$ [23, 24, 25, 29].

We consider shift-invariant spaces generated by a single ($N = 1$) or two functions ($N = 2$), corresponding to B-splines and Hermite splines, respectively. We only consider spaces (5) for which the family $(\varphi_i(\cdot - k))_{i=1 \dots N, k \in \mathbb{Z}}$ is a Riesz basis; that is,

$$A \sum_{i=1}^N \sum_{k \in \mathbb{Z}} c_i[k]^2 \leq \left\| \sum_{i=1}^N \sum_{k \in \mathbb{Z}} c_i[k] \varphi_i(\cdot - k) \right\|_{L_2}^2 \leq B \sum_{i=1}^N \sum_{k \in \mathbb{Z}} c_i[k]^2 \quad (6)$$

for any $c_i \in \ell_2(\mathbb{Z})$ with $i = 1, \dots, N$ and for some constants $0 < A \leq B < \infty$. This ensures that any $f \in V(\boldsymbol{\varphi})$ has a unique and stable representation.

Support properties of B-splines. Under the Riesz condition, a natural question is the ability of the basis functions $\boldsymbol{\varphi}$ to exactly reproduce classes of functions. The possibility to perfectly reproduce polynomials is of crucial importance. The constant function 1 can be reproduced if and only if the basis functions satisfy the partition of unity, which is the minimal requirement for a practical approximation scheme [30].

The polynomial B-spline of order $L \geq 1$ is classically known to be able to reproduce polynomials up to degree $(L - 1)$; that is, t^ℓ for every $\ell = 0, \dots, (L - 1)$. Following [31], we denote it by $\beta^{(L-1)}$. In particular, the cubic B-spline (of order $L = 4$), can perfectly reproduce any polynomial of degree at most 3. The ability of the basis functions $\boldsymbol{\varphi}$ to perfectly reproduce polynomials is intimately linked to their approximation power, as will be developed in Section 3. B-splines are actually the most localized functions satisfying this property, as formalised in Proposition 1.

Proposition 1. *Let $\varphi \in L_2(\mathbb{R})$ be a function such that $(\varphi(\cdot - k))_{k \in \mathbb{Z}}$ is a Riesz basis and can reproduce polynomials up to degree $(L - 1) \geq 0$. Then the support of φ is at least of size L .*

In particular, the B-spline of order L , whose support is of size L , is optimal in terms of support localization among basis functions that are able to reproduce polynomials of degree at most $(L - 1)$.

This result is classical in approximation theory: Schoenberg showed that B-splines effectively have the adequate approximation order [1]. A complete characterization of the

functions of minimal support with a given approximation order can be found in [16, Theorem 1]. To the best of our knowledge, very little is known about the localization of basis functions when $N > 1$, which is what we propose to investigate.

2.2 Minimal Support Properties for Two Basis Functions

Hermite splines ϕ_1 and ϕ_2 , given by (2) and (3), are able to reproduce both β^2 and β^3 , the quadratic and cubic B-splines of order 3 and 4, respectively [11]. In particular, this means that $V(\phi_1, \phi_2)$ contains polynomials of degree at most 3. Many other pairs of basis functions, starting with β^2 and β^3 themselves, can also reproduce quadratic and cubic splines. In line with Proposition 1, investigating the support localization of a pair of functions having the same reproduction properties as the Hermite splines follows naturally. This boils down to studying basis functions for which $\beta^2, \beta^3 \in V(\varphi_1, \varphi_2)$.

We characterize the support size of such pairs of functions in Theorem 2. This result then allows us to deduce the minimum support property of Hermite splines in Corollary 3.

Theorem 2. *Let $\varphi_1, \varphi_2 \in L_2(\mathbb{R}^d)$ be two basis functions. We assume that*

$$\beta^2(t) = \sum_{k \in \mathbb{Z}} a_k \varphi_1(t - k) + b_k \varphi_2(t - k), \quad (7)$$

$$\beta^3(t) = \sum_{k \in \mathbb{Z}} c_k \varphi_1(t - k) + d_k \varphi_2(t - k), \quad (8)$$

with reproduction sequences a, b, c, d satisfying

$$\sum_{k \in \mathbb{Z}} k^3 (|a_k| + |b_k| + |c_k| + |d_k|) < \infty. \quad (9)$$

In particular, the quadratic and cubic B-splines β^2, β^3 are in $V(\varphi_1, \varphi_2)$. Then, we have that

$$|\text{Supp} \varphi_1| + |\text{Supp} \varphi_2| \geq 4 \quad (10)$$

Proof. First of all, one can restrict to compactly supported basis functions φ_1 and φ_2 (otherwise, $|\text{Supp} \varphi_1| + |\text{Supp} \varphi_2| = \infty$). If one of the basis function, for instance φ_2 , is identically zero, then the cubic spline $\beta^3 \in V(\varphi_1)$. This means in particular that the basis function φ_1 reproduces polynomials up to degree 3, and its support is therefore at least of size four [16, Theorem 1]. Hence, we again have that $|\text{Supp} \varphi_1| + |\text{Supp} \varphi_2| = |\text{Supp} \varphi_1| \geq 4$. We now assume that φ_1 and φ_2 are not identically 0.

Step 1. We show that the extreme points of the supports of φ_1 and φ_2 are integers. For $x = a, b, c, d$, we set $X(\omega) = \sum_{k \in \mathbb{Z}} x_k e^{-j\omega k}$, the 2π -periodic Fourier transform of x .

Condition (9) ensures that X has a continuous third derivative. In Fourier domain, (7) and (8) become

$$\widehat{\beta^2}(\omega) = \frac{(1 - e^{-j\omega})^3}{(j\omega)^3} = A(\omega)\widehat{\varphi_1}(\omega) + B(\omega)\widehat{\varphi_2}(\omega), \quad (11)$$

$$\widehat{\beta^3}(\omega) = \frac{(1 - e^{-j\omega})^4}{(j\omega)^4} = C(\omega)\widehat{\varphi_1}(\omega) + D(\omega)\widehat{\varphi_2}(\omega). \quad (12)$$

We set $\det(\omega) = A(\omega)D(\omega) - B(\omega)C(\omega)$, which is itself a function with continuous third derivative. From (11) and (12), we obtain

$$\det(\omega)\widehat{\varphi_1}(\omega) = D(\omega)\frac{(1 - e^{-j\omega})^3}{(j\omega)^3} - B(\omega)\frac{(1 - e^{-j\omega})^4}{(j\omega)^4}, \quad (13)$$

$$\det(\omega)\widehat{\varphi_2}(\omega) = -C(\omega)\frac{(1 - e^{-j\omega})^3}{(j\omega)^3} + A(\omega)\frac{(1 - e^{-j\omega})^4}{(j\omega)^4}. \quad (14)$$

We show that $\det(\omega) \neq 0$ for $\omega \neq 0$ by contradiction. Let us fix $\omega_0 \in (0, 2\pi)$ and assume that $\det(\omega_0) = 0$. We set $\alpha = \frac{1 - e^{-j\omega_0}}{j\omega_0}$ and $\beta = \frac{1 - e^{-j\omega_0}}{j\omega_0 + 2\pi}$. Then, $\alpha \neq 0$, $\beta \neq 0$, and $\alpha \neq \beta$. Moreover, (13) for $\omega = \omega_0$ and $(\omega_0 + 2\pi)$ implies that

$$\begin{pmatrix} \alpha^3 & -\alpha^4 \\ \beta^3 & -\beta^4 \end{pmatrix} \begin{pmatrix} D(\omega_0) \\ B(\omega_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (15)$$

The matrix being invertible (with determinant $\alpha^3\beta^3(\alpha - \beta) \neq 0$), we deduce that $D(\omega_0) = B(\omega_0) = 0$. Similarly, (14) with $\omega = \omega_0$ and $(\omega_0 + 2\pi)$ implies that $A(\omega_0) = C(\omega_0) = 0$. Injecting this in (11) with $\omega = \omega_0$, we deduce that $\widehat{\beta_2}(\omega_0) = \alpha^3 = 0$, which contradicts our initial assumption.

We now study $\det(\omega)$ around the origin. Assume that $\det^{(p)}(0) = 0$ for $p = 0, 1, 2, 3$. We will show that this is impossible. Under our assumption, Taylor's expansion gives that $\det(\omega) = o(\omega^3)$ around 0. Hence, the function $\frac{\det(\omega)}{(1 - e^{-j\omega})^3}$ can be continuously extended at $\omega = 0$, for which it is 0. By periodicity, this is also valid for $\omega = 2\pi$. In particular, $\frac{\det(\omega)}{(1 - e^{-j\omega})^3}\widehat{\varphi_1}(\omega)$ is continuous with limit 0 at $\omega = 0$ and 2π . From (13), we moreover remark that

$$\frac{\det(\omega)}{(1 - e^{-j\omega})^3}\widehat{\varphi_1}(\omega) = D(\omega) - B(\omega)\frac{1 - e^{-j\omega}}{j\omega}. \quad (16)$$

Knowing the limit of the left term at $\omega = 0$ and 2π , and using the 2π -periodicity of $D(\omega)$, one deduces that $D(0) - B(0) = 0 = D(2\pi) - 0 = D(0)$. Therefore, $D(0) = B(0) = 0$. An identical reasoning based on (14) implies that $A(0) = C(0) = 0$. From (11) with $\omega = 0$, we obtain $\widehat{\beta_3}(0) = 0$, which is false. As a consequence, $\det^{(p)}(0) \neq 0$ for some $0 \leq p \leq 3$, imposing that $\frac{(1 - e^{j\omega})^3}{\det(\omega)}$ has a limit (possibly 0) at 0. Combined with the fact that \det

does not vanish over $(0, 2\pi)$, we deduce that $F(\omega) = \frac{(1-e^{j\omega})^3}{\det(\omega)}$ is a continuous 2π -periodic function. We can then rewrite (13) as

$$(j\omega)^4 \widehat{\varphi}_1(\omega) = (j\omega)F(\omega)D(\omega) - (1 - e^{-j\omega})F(\omega)B(\omega). \quad (17)$$

The functions $F(\omega)D(\omega)$ and $(1 - e^{-j\omega})F(\omega)B(\omega)$ are 2π -periodic, hence their inverse Fourier transforms are sum of Dirac impulses located at the integers. It means in particular that we have, in time domain,

$$\varphi_1^{(4)}(t) = \sum_{k \in \mathbb{Z}} (y_k \delta(t - k) + z_k \delta'(t - k)). \quad (18)$$

Since $\varphi_1^{(4)}$ is compactly supported, like φ_1 , only finitely many y_k and z_k are non-zero. Then, φ_1 is a compactly supported function whose fourth derivative has a support with integer extreme points (due to (18)), and therefore has a support with integer extreme points. The same reasoning also applies for φ_2 , concluding this part of the proof.

Step 2. We know that the supports of φ_1 and φ_2 are of the form $[a, b]$ with $a < b$, $a, b \in \mathbb{Z}$. We assume that $|\text{Supp}\varphi_1| + |\text{Supp}\varphi_2| < 4$. Then, one of the two basis functions has a support of size one, for instance φ_1 . We also assume without loss of generality that $\text{Supp}\varphi_1 = [0, 1]$, implying that only y_0, y_1, z_0, z_1 are possibly non-zero in (18). Going back to the Fourier domain, one obtains

$$(j\omega)^4 \widehat{\varphi}_1(\omega) = y_0 + y_1 e^{-j\omega} + j\omega(z_0 + z_1 e^{-j\omega}). \quad (19)$$

The function φ_1 is compactly supported. Its Fourier transform is hence infinitely smooth, and we can do the Taylor expansion of both sides in (19). In particular, we obtain the relations

$$y_0 + y_1 = z_0 + z_1 + y_1 = \frac{y_1}{2} + z_1 = \frac{y_1}{6} + \frac{z_1}{2} = 0. \quad (20)$$

This imposes that $y_0 = y_1 = z_0 = z_1 = 0$, which is absurd. Finally, it shows that $|\text{Supp}\varphi_1| + |\text{Supp}\varphi_2| \geq 4$, as expected. \square

Condition (9) plays an important role in our proof by imposing some regularity in Fourier domain. In practice, one even expects that compactly supported basis functions can generate the B-splines β^2 and β^3 with finitely many coefficients, in which case (9) is automatically satisfied. However, we believe that the condition (9) can be relaxed up to some extent. From Theorem 2, we easily deduce that Hermite splines have the minimum support property.

Corollary 3. *The Hermite splines (ϕ_1, ϕ_2) are of minimal support among the pair of functions that are able to reproduce both quadratic and cubic B-splines with reproduction sequences satisfying (9).*

Proof. From [11, Appendix A], we know that Hermite splines can reproduce both quadratic and cubic B-splines, hence β^2 and β^3 are in $V(\phi_1, \phi_2)$ with compactly supported reproduction sequences obviously satisfying (9). The supports of ϕ_1 and ϕ_2 are of size two, implying that $|\text{Supp}\phi_1| + |\text{Supp}\phi_2| = 4$. Finally, the pair (ϕ_1, ϕ_2) is maximally localized due to (10). \square

It is worth noting that the supports of the pair of Hermite splines jointly has the same size as the B-spline β^3 . However, β^3 alone has less reproduction property. Being of class C^2 , it is in particular unable to reproduce the quadratic spline β^2 , which only has C^1 transitions at the integers. The simplest way of reproducing β^2, β^3 is to consider the basis pair (β^2, β^3) itself, which is not maximally localized since the sum of the supports is 7. An important additional remark is that two functions are needed to reproduce both cubic and quadratic spline, as formalised in Proposition 4.

Proposition 4. *There exists no single function $\varphi \in L_2(\mathbb{R})$ that is able to reproduce β^2 and β^3 with summable reproduction sequences.*

Proof. By contradiction, let us assume that there exists φ such that $\beta^2 = \sum_{k \in \mathbb{Z}} a_k \varphi(\cdot - k)$ and $\beta^3 = \sum_{k \in \mathbb{Z}} b_k \varphi(\cdot - k)$ with $a, b \in \ell_1(\mathbb{Z})$. Then, the Fourier transforms $A(e^{j\omega})$ and $B(e^{j\omega})$ are continuous 2π -periodic functions. In Fourier domain, we have that

$$\left(\frac{1 - e^{-j\omega}}{j\omega} \right)^3 = A(e^{j\omega})\hat{\varphi}(\omega), \quad \left(\frac{1 - e^{-j\omega}}{j\omega} \right)^4 = B(e^{j\omega})\hat{\varphi}(\omega). \quad (21)$$

Set $\omega_0 \in (0, 2\pi)$ and $\omega_1 = \omega_0 + 2\pi$. The relation (21) imposes that $A(e^{j\omega_i})$, $B(e^{j\omega_i})$, and $\hat{\varphi}(\omega_i)$ are non-zero for $i = 1, 2$, and

$$\frac{1 - e^{-j\omega_i}}{j\omega_i} A(e^{j\omega_i})\hat{\varphi}(\omega_i) = B(e^{j\omega_i})\hat{\varphi}(\omega_i). \quad (22)$$

After simplifications, we deduce that

$$j\omega_i = \frac{(1 - e^{-j\omega_i})A(e^{j\omega_i})}{B(e^{j\omega_i})}. \quad (23)$$

The right term in (23) is equal for ω_0 and ω_1 by periodicity, while the left term is different. This contradicts our initial assumption, and implies Proposition 4. \square

3 Approximation Properties of Hermite Splines

Existing approaches have been proposed to characterize the behaviour of the approximation error for single [19] and multi-generators [20]. They however assume technical conditions that Hermite interpolation does not satisfy. We therefore formulate new theoretical tools to

quantify the asymptotic constant of the approximation error of this scheme. Since functions are estimated with derivative samples in the Hermite case, we also propose to investigate the approximation error on the first derivative. In our setting, other existing approximation schemes can be considered as well, allowing us to put the excellent approximation properties of Hermite splines in perspective with other related schemes, such as cubic B-splines and interlaced derivative sampling.

3.1 Generalized Sampling and Approximation Error

The approximation of a continuously-defined signal from a collection of its samples in a generalized sampling scheme relies on two ingredients: some basis functions $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_N) \in (L_2(\mathbb{R}))^N$, and some *sampling functions* $\tilde{\boldsymbol{\varphi}} = (\tilde{\varphi}_1, \dots, \tilde{\varphi}_N)$ that are rapidly decaying generalized functions, including rapidly decaying L_2 -integrable functions¹ together with the Dirac impulse δ , its derivative, and their shifts. The term *generalized* is motivated by the fact that sampling functions allow accessing more than the values of the signal at the sampling points [33]. The set of pairs of basis and sampling functions fully characterize an approximation scheme. Moreover, because we aim at a fair comparison, the quantity of information per unit of time should be equal among the considered schemes. This requires to slightly adapt the definition of $W(\boldsymbol{\varphi})$ given in (5). From now, we shall consider

$$W(\boldsymbol{\varphi}) = \left\{ \sum_{i=1}^N \sum_{k \in \mathbb{Z}} c_i[k] \varphi_i(\cdot - Nk) : \mathbf{c} = (c_1, \dots, c_N) \in (\ell_2(\mathbb{Z}))^N \right\}, \quad (24)$$

so that, for any $N \geq 1$, there is on average a single degree of freedom on each interval of size one. The sampling and reconstruction problem is then formally defined as follows. The function f is reconstructed by its approximation \tilde{f} associated to the basis functions $\boldsymbol{\varphi}$ and for the sampling functions $\tilde{\boldsymbol{\varphi}}$, defined as

$$\tilde{f} = \sum_{i=1}^N \sum_{k \in \mathbb{Z}} \langle f, \tilde{\varphi}_i(\cdot - Nk) \rangle \varphi_i(\cdot - Nk) \in W(\boldsymbol{\varphi}). \quad (25)$$

We denote by $\mathcal{Q}_{\tilde{\boldsymbol{\varphi}}}^{\boldsymbol{\varphi}}$ the operator such that $\mathcal{Q}_{\tilde{\boldsymbol{\varphi}}}^{\boldsymbol{\varphi}} f = \tilde{f}$. Hermite spline approximation thus corresponds to $N = 2$ with basis functions $\varphi_1(t) = \phi_1(\frac{t}{2})$ and $\varphi_2(t) = 2\phi_2(\frac{t}{2})$, and sampling functions $\tilde{\varphi}_1(t) = \delta(t)$ and $\tilde{\varphi}_2(t) = -\delta'(t)$.

The best approximation of a given scheme is obtained when the pairs of sampling and basis functions are properly chosen such that \tilde{f} is the orthogonal projection of f onto $W(\boldsymbol{\varphi})$.

¹We say that an L_2 function is *rapidly decaying* if it decays faster than any polynomial at infinity. A generalized function in $\mathcal{S}'(\mathbb{R})$ is rapidly decaying if its convolution with any infinitely differentiable and rapidly decaying function is a rapidly decaying function [32]. In particular, a rapidly decaying (generalized) function has an infinitely differentiable Fourier transform, which we shall rely on thereafter.

This implies imposing a particular condition [20] on the sampling functions $\tilde{\varphi}$, namely that

$$\tilde{\varphi} = \begin{pmatrix} \varphi_{1,d} \\ \vdots \\ \varphi_{N,d} \end{pmatrix} = \mathcal{F}^{-1} \{ \mathbf{G}_{\varphi}(\cdot)^{-1} \hat{\varphi} \}, \quad (26)$$

where \mathbf{G}_{φ} is the Gram matrix of size $N \times N$ associated to φ , given for $\omega \in \mathbb{R}$ by

$$\mathbf{G}_{\varphi}(\omega) = \sum_{k \in \mathbb{Z}} \hat{\varphi}(\omega + 2k\pi) \hat{\varphi}^{*T}(\omega + 2k\pi). \quad (27)$$

This particular collection of sampling functions are called the *dual functions* associated to φ and are denoted by φ_d . Note that $\mathbf{G}_{\varphi}(\omega)$ is invertible for every $\omega \in \mathbb{R}$ — and therefore (27) is meaningful — because (6) is equivalent to $0 < A \leq \lambda_{\min}(\omega) \leq \lambda_{\max}(\omega) \leq B < \infty$, where $\lambda_{\min}(\omega)$ ($\lambda_{\max}(\omega)$, respectively) is the minimum (maximum, respectively) eigenvalue of $\mathbf{G}_{\varphi}(\omega)$ [34]. When $\tilde{\varphi}$ is defined as (26), the operator $\mathcal{Q}_{\tilde{\varphi}}$ is the orthogonal projector over $W(\varphi)$, and is denoted as $\mathcal{P}_{\varphi} = \mathcal{Q}_{\tilde{\varphi}}^d$. In this situation, (25) is reformulated as

$$\tilde{f}(t) = \mathcal{P}_{\varphi} f = \sum_{i=1}^N \sum_{k \in \mathbb{Z}} \langle f, \varphi_{i,d}(\cdot - Nk) \rangle \varphi_i(\cdot - Nk). \quad (28)$$

To simplify the notation, we shall write $\mathcal{P} = \mathcal{P}_{\varphi}$ and $\mathcal{Q} = \mathcal{Q}_{\tilde{\varphi}}$ thereafter. The quality of the approximation is then evaluated in terms of approximation error which is expressed as

$$E_{\tilde{\varphi}}^{\tilde{\varphi}}(f) = \|f - \tilde{f}\|_{L_2} = \|f - \mathcal{Q}f\|_{L_2}. \quad (29)$$

When $\tilde{f} = \mathcal{P}f$, the error is denoted as $E_{\varphi}(f)$. A direct implication is that $E_{\varphi}(f) = E_{\tilde{\varphi}}^d(f) \leq E_{\tilde{\varphi}}^{\tilde{\varphi}}(f)$, reaching the equality if and only if $\tilde{\varphi} = \varphi_d$.

Up to now, we considered approximation schemes with (generalized) samples taken on the integer grid, corresponding to a sampling step size $T = 1$. This parameter affects the coarseness of the approximation: when $T \rightarrow 0$, the error is expected to vanish. For $T > 0$, the approximation space (24) becomes

$$W_T(\varphi) = \left\{ \sum_{i=1}^N \sum_{k \in \mathbb{Z}} c_i[k] \varphi_i\left(\frac{\cdot}{T} - Nk\right) : \mathbf{c} \in (\ell_2(\mathbb{Z}))^N \right\} \quad (30)$$

and the approximation of f is given by

$$\tilde{f}_T = \mathcal{Q}_T f = \sum_{i=1}^N \sum_{k \in \mathbb{Z}} \left\langle f, \frac{1}{T} \tilde{\varphi}_i\left(\frac{\cdot}{T} - Nk\right) \right\rangle \varphi_i\left(\frac{\cdot}{T} - Nk\right), \quad (31)$$

with resulting error

$$E_{\tilde{\varphi}}^{\tilde{\varphi}}(f, T) = \|f - \mathcal{Q}_T f\|_{L_2}. \quad (32)$$

The orthogonal projector (28) and its associated error are easily reformulated accordingly.

A number of hypotheses on the basis functions φ and the sampling functions $\tilde{\varphi}$ have to be met in order to study the errors $E_{\varphi}(f, T)$ and $E_{\tilde{\varphi}}^{\tilde{\varphi}}(f, T)$ in terms of rate of decay and asymptotic constant. The first one is the Riesz basis condition (6), which ensures a unique and stable representation. The second one is the basis functions order.

Definition 5. *A family of $N \leq 1$ basis functions $\varphi = (\varphi_1, \dots, \varphi_N)$ is of order L if they can reproduce polynomial functions up to degree $(L - 1)$, meaning that, for $\ell = 0, \dots, (L - 1)$, there exists sequences $c_{i,\ell}$ such that*

$$t^\ell = \sum_{i=1}^N \sum_{k \in \mathbb{Z}} c_{i,\ell}[k] \varphi_i(t - Nk). \quad (33)$$

This condition is known to be equivalent to the Strang and Fix conditions [35]. When it is met, the decrease of the optimal error $E_{\varphi}(f, T)$ is bounded from above by T^L [20]. The last important condition is that the sampling and basis functions are quasi-biorthonormal of order L .

Definition 6. *Two families of basis φ and sampling functions $\tilde{\varphi}$ are quasi-biorthonormal of order L if the basis functions are of order L and, for the dual function φ_d given by (26) and all $\ell = 0, \dots, (L - 1)$, we have*

$$\int_{\mathbb{R}} t^\ell \tilde{\varphi}(t) dt = \int_{\mathbb{R}} t^\ell \varphi_d(t) dt. \quad (34)$$

It is worth noting that (34) is a slight abuse of notation, since the $\tilde{\varphi}_i$ are not necessarily defined pointwise. However, they are assumed to be rapidly decaying generalized functions and can therefore be taken against a slowly growing smooth functions such as $t \mapsto t^\ell$. The quasi-biorthonormality ensures that the decrease of the error $E_{\tilde{\varphi}}^{\tilde{\varphi}}(f, T)$ is also bounded by T^L [20]. Finally, the rate of decay of the approximation error being under control in all generality, the asymptotic constant can be obtained as

$$C_{\tilde{\varphi}}^{\tilde{\varphi}}(f) = \lim_{T \rightarrow 0} T^{-L} E_{\tilde{\varphi}}^{\tilde{\varphi}}(f, T), \quad (35)$$

which, in practice can be computed relying on a Fourier-domain approximation error kernel. For more details, we refer the interested reader to [36, 37, 38] and references therein.

3.2 Approximation Constants of Irregular Sampling Schemes

Our goal in this section is to extend the range of applicability of the main results of [20] so as to include Hermite spline approximation. The original framework is indeed restricted to sampling functions $\tilde{\varphi}$ with bounded Fourier transforms. While this allows considering the case of interpolation, it excludes Hermite spline approximation since the Fourier transform $j\omega$ of δ' is unbounded.

Our development follows the key contributions of [20]. We therefore only detail the adaptations that are required in our case. We start with a brief summary of the general approach, which is common to many works of approximation theory in shift-invariant spaces. We first introduce the kernels associated to the functions $\varphi, \tilde{\varphi}$ as

$$\mathcal{E}_{\min}(\omega) = 1 + \tilde{\varphi}^{*T}(\omega) \mathbf{G}_{\varphi}^{-1}(\omega) \tilde{\varphi}(\omega), \quad (36)$$

$$\mathcal{E}_{\text{res}}(\omega) = (\hat{\tilde{\varphi}} - \hat{\varphi}_d)^{*T}(\omega) \mathbf{G}_{\varphi}(\omega) (\hat{\tilde{\varphi}} - \hat{\varphi}_d)(\omega), \quad (37)$$

$$\mathcal{E}(\omega) = \mathcal{E}_{\min}(\omega) + \mathcal{E}_{\text{res}}(\omega), \quad (38)$$

where \mathbf{G}_{φ} is the Gram matrix (27). The kernel \mathcal{E}_{\min} relates to the *minimum error* case achieved using the orthogonal projector (*i.e.*, $\tilde{\varphi} = \varphi_d$), and \mathcal{E}_{res} to the *residual error* arising when using sampling functions that differ from the dual functions. We furthermore note that $\mathcal{E}_{\text{res}}(\omega) = 0$ when $\tilde{\varphi} = \varphi_d$, as expected. The key ideas of relying on \mathcal{E} are as follows.

- The kernel \mathcal{E} measures the approximation power of a reasonable approximation scheme $(\varphi, \tilde{\varphi})$ in the sense that $\|f - \mathcal{Q}_T f\|_{L_2} \approx \left(\int_{\mathbb{R}} |\hat{f}(\omega)|^2 \mathcal{E}(T\omega) d\omega \right)^{1/2}$ for small $T > 0$.
- The precise behaviour of $\|f - \mathcal{Q}_T f\|_{L_2}$ is then deduced from that of \mathcal{E} around the origin. It depends on the approximation order L of the scheme and, typically, behaves like CT^L . The constant C depends on the function to approximate f and on the scheme $(\varphi, \tilde{\varphi})$ via the Taylor expansion of \mathcal{E} .

We give a precise meaning to these two points in Propositions 8 and 9 below. Before that, we recall an important lemma taken from [20] that will play a fundamental role in our proofs.

Lemma 7. *For $k \geq 0$, we set $\hat{f}_k(\omega) = \hat{f}(\omega) 1_{k/T \leq |\omega| < (k+1)/T}$. Then, the following relations hold:*

$$\hat{f} = \sum_{k \geq 0} \hat{f}_k; \quad (39)$$

$$\|f_k - \mathcal{Q}_T f_k\|_{L_2}^2 = \int_{\mathbb{R}} |\hat{f}_k(\omega)|^2 \mathcal{E}(T\omega) d\omega \text{ for } k \geq 0; \text{ and} \quad (40)$$

$$\left| \|f - \mathcal{Q}_T f\|_{L_2} - \|f_0 - \mathcal{Q}_T f_0\|_{L_2} \right| \leq \sum_{k > 0} \|f_k - \mathcal{Q}_T f_k\|_{L_2}. \quad (41)$$

The equality (39) is obvious. The two next relations comes from [20, Theorem 1] and we have simply reformulated the results with our notation. First, it is proven (see (27) in [20]) that $\|f - \mathcal{Q}_T f\|_{L_2} = \sum_{k \in \mathbb{Z}} \int_{\mathbb{R}} |\widehat{f}_k(\omega)|^2 \mathcal{E}(T\omega) d\omega$ as soon as $\widehat{f}(\omega)\widehat{f}(\omega - n/T) = 0$ for any $\omega \in \mathbb{R}$ and $n \in \mathbb{Z}$, a condition that is satisfied by the f_k by construction, giving (40). Moreover, the inequality (41) appears in the proof of [20, Theorem 1] (see (63)).

Proposition 8. *Let $(\varphi, \tilde{\varphi})$ be a set of N basis and sampling functions that are biorthonormal and provide an approximation scheme of order L . We assume moreover that the kernel \mathcal{E} given by (38) satisfies*

$$|\mathcal{E}(\omega)| \leq C^2 \max(1, |\omega|^{2p}) \quad (42)$$

for some $C > 0$, some integer $0 \leq p \leq L$, and every $\omega \in \mathbb{R}$. Then, for every $f \in W_2^{L+1}(\mathbb{R})$, we have

$$\|f - \mathcal{Q}_T f\|_{L_2} = \left(\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \mathcal{E}(T\omega) d\omega \right)^{1/2} + O(T^{L+1}). \quad (43)$$

The case $p = 0$ corresponds to \mathcal{E} being bounded and can be found in [20, Theorem 1].

Proof. We set $\widehat{f}_k(\omega) = \widehat{f}(\omega) 1_{k/T \leq |\omega| < (k+1)/T}$. Then, we have

$$\begin{aligned} & \left| \|f - \mathcal{Q}_T f\|_{L_2} - \left(\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \mathcal{E}(T\omega) d\omega \right)^{1/2} \right| \\ & \leq \left| \|f - \mathcal{Q}_T f\|_{L_2} - \|f_0 - \mathcal{Q}_T f_0\|_{L_2} \right| + \left| \|f_0 - \mathcal{Q}_T f_0\|_{L_2} - \left(\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \mathcal{E}(T\omega) d\omega \right)^{1/2} \right| \\ & \stackrel{(i)}{\leq} \sum_{k>0} \|f_k - \mathcal{Q}_T f_k\|_{L_2} + \left| \|f_0 - \mathcal{Q}_T f_0\|_{L_2} - \left(\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \mathcal{E}(T\omega) d\omega \right)^{1/2} \right| \\ & = (I) + (II), \end{aligned} \quad (44)$$

where we used (41) in (i). As a consequence, (43) follows if one shows that the two terms (I) and (II) in (44) are $O(T^{L+1})$. Using (40), we have

$$\begin{aligned} \|f_k - \mathcal{Q}_T f_k\|_{L_2}^2 &= \int_{k/T < |\omega| \leq (k+1)/T} |\widehat{f}_k(\omega)|^2 \mathcal{E}(T\omega) d\omega \\ &= \int_{k/T < |\omega| \leq (k+1)/T} |\widehat{f}_k(\omega)|^2 |\omega|^{2(L+1)} \frac{T^{2p}}{|\omega|^{2(L-p+1)}} \frac{\mathcal{E}(T\omega)}{T^{2p} |\omega|^{2p}} d\omega \\ &\leq C^2 \frac{T^{2p}}{(k/T)^{2(L-p+1)}} \int_{\mathbb{R}} |\widehat{f}_k(\omega)|^2 |\omega|^{2(L+1)} d\omega \\ &= C^2 \frac{T^{2(L+1)}}{k^{2(L-p+1)}} \|f_k^{(L+1)}\|_{L_2}^2, \end{aligned} \quad (45)$$

where the inequality is due to $|\omega| \geq k/T$ over the domain and to (42), which implies that $C^2 \geq \sup_{\omega \in \mathbb{R}} \mathcal{E}(\omega) / \max(1, |\omega|^p) \geq \sup_{|\omega| \geq 1} \mathcal{E}(\omega) / |\omega|^p$. Summing over $k > 0$, we deduce that

$$\begin{aligned}
(I) &\leq CT^{L+1} \sum_{k>0} \frac{1}{k^{L-p+1}} \|f_k^{(L+1)}\|_{L_2} \\
&\leq CT^{L+1} \left(\sum_{k>0} \frac{1}{k^{2(L-p+1)}} \right)^{1/2} \left(\sum_{k>0} \|f_k^{(L+1)}\|_{L_2}^2 \right)^{1/2} \\
&= CT^{L+1} \sqrt{\zeta(2(L-p+1))} \|(f - f_0)^{(L+1)}\|_{L_2} = O(T^{L+1}), \tag{46}
\end{aligned}$$

where the second inequality is derived from Cauchy-Schwarz, and ζ is the Riemann zeta function. For the second term, we remark that, again due to [20, (27)], $\|f_0 - \mathcal{Q}_T f_0\|_{L_2}^2 = \int_{|\omega| \leq 1/T} |\widehat{f}(\omega)|^2 \mathcal{E}(T\omega) d\omega$. Therefore, using the relation $\|g\|_{L_2} - \|h\|_{L_2} \leq \|g - h\|_{L_2}$, which is a consequence of the Minkowski inequality, we deduce that

$$\begin{aligned}
(II) &\leq \left(\int_{|\omega| > 1/T} |\widehat{f}(\omega)|^2 \mathcal{E}(T\omega) d\omega \right)^{1/2} \\
&= \left(\int_{|\omega| > 1/T} |\widehat{f}(\omega)|^2 |\omega|^{2(L+1)} \frac{T^{2p}}{|\omega|^{2(L-p+1)}} \frac{\mathcal{E}(T\omega)}{T^{2p} |\omega|^{2p}} d\omega \right)^{1/2} \\
&\leq CT^p T^{L-p+1} \left(\int_{|\omega| > 1/T} |\widehat{f}(\omega)|^2 |\omega|^{2(L+1)} d\omega \right)^{1/2} \\
&= CT^{L+1} \|(f - f_0)^{(L+1)}\|_{L_2} = O(T^{L+1}), \tag{47}
\end{aligned}$$

where the inequality once more follows from $|\omega| \geq 1/T$ over the domain and from our assumption (42). Combining (46) and (47) in (44) completes the proof. \square

Our next result connects the expression $\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \mathcal{E}(T\omega) d\omega$ to the Taylor expansion of \mathcal{E} .

Proposition 9. *Let $(\varphi, \tilde{\varphi})$ be a set of N basis and sampling functions that are biorthonormal and provide an approximation scheme of order L . We assume moreover that the kernel \mathcal{E} given by (38) is $(2L+1)$ -times continuously differentiable and satisfies*

$$|\mathcal{E}^{(2L+1)}(\omega)| \leq C^2 \max(1, |\omega|^{2p}) \tag{48}$$

for some $C > 0$, some integer $0 \leq p \leq L$, and every $\omega \in \mathbb{R}$. Then, for every $f \in W_2^{L+p+1/2}(\mathbb{R})$, we have

$$\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \mathcal{E}(T\omega) d\omega = \frac{\mathcal{E}^{(2L)}(0)}{(2L)!} \|f^{(L)}\|_{L_2}^2 T^{2L} + O(T^{2L+1}). \tag{49}$$

Proof. The kernel \mathcal{E} being symmetric, we deduce that $\mathcal{E}^{(2k+1)}(0) = 0$ for $k = 0, \dots, (L-1)$. Moreover, the condition of $(\varphi, \tilde{\varphi})$ ensures that $\mathcal{E}^{(2k)}(0) = 0$ for $k = 0, \dots, (L-1)$, together with $\mathcal{E}^{(2L)}(0) \neq 0$. Therefore, \mathcal{E} being $(2L+1)$ -times continuously differentiable, the Taylor expansion around the origin is given by

$$\mathcal{E}(\omega) = \frac{\mathcal{E}^{(2L)}(0)}{(2L)!} \omega^{2L} + \frac{\mathcal{E}^{(2L+1)}(\omega\theta)}{(2L+1)!} \omega^{2L+1}, \quad (50)$$

with $\theta = \theta(\omega) \in (0, 1)$. Using (50), we deduce that

$$\begin{aligned} \int_{\mathbb{R}} |\hat{f}(\omega)|^2 \mathcal{E}(T\omega) d\omega &= \int_{\mathbb{R}} |\hat{f}(\omega)|^2 \frac{\mathcal{E}^{(2L)}(0)}{(2L)!} T^{2L} \omega^{2L} d\omega \\ &\quad + \frac{T^{2L+1}}{(2L+1)!} \int_{\mathbb{R}} |\hat{f}(\omega)|^2 \omega^{2L+1} \mathcal{E}^{(2L+1)}(\omega T \theta(\omega T)) d\omega \\ &= \frac{\mathcal{E}^{(2L)}(0)}{(2L)!} \|f^{(L)}\|_{L_2}^2 T^{2L} \\ &\quad + \frac{T^{2L+1}}{(2L+1)!} \int_{\mathbb{R}} |\hat{f}(\omega)|^2 \omega^{2L+1} \mathcal{E}^{(2L+1)}(\omega T \theta(\omega T)) d\omega. \end{aligned} \quad (51)$$

Due to (48) and $0 < \theta(\omega T) < 1$,

$$\mathcal{E}^{(2L+1)}(\omega T \theta(\omega T)) \leq C^2 \max(1, |\omega|^{2p} T^{2p}) \leq C^2 (1 + |\omega|^{2p} T^{2p}), \quad (52)$$

from which we deduce that $\int_{\mathbb{R}} |\hat{f}(\omega)|^2 \omega^{2L+1} \mathcal{E}^{(2L+1)}(\omega T \theta(\omega T)) d\omega \leq C^2 (\|f^{(L+1/2)}\|_{L_2}^2 + T^{2p} \|f^{(L+p+1/2)}\|_{L_2}^2)$. Injecting this to (51) implies

$$\begin{aligned} \int_{\mathbb{R}} |\hat{f}(\omega)|^2 \mathcal{E}(T\omega) d\omega - \frac{\mathcal{E}^{(2L)}(0)}{(2L)!} \|f^{(L)}\|_{L_2}^2 T^{2L} &= C^2 \frac{T^{2L+1}}{(2L+1)!} (\|f^{(L+1/2)}\|_{L_2}^2 \\ &\quad + T^{2p} \|f^{(L+p+1/2)}\|_{L_2}^2) \\ &= O(T^{2L+1}), \end{aligned} \quad (53)$$

which concludes the proof. \square

Note that we use derivatives of fractional order $\gamma = L + 1/2$ and $\gamma = L + p + 1/2$ in the proof, which are defined in the Fourier domain as $\mathcal{F}\{f^{(\gamma)}\}(\omega) = (j\omega)^\gamma \hat{f}(\omega)$. Finally, we conclude with an extension of [20, Theorem 4] to sampling functions that are not necessarily bounded in Fourier domain.

Theorem 10. *We consider an approximation scheme $(\varphi, \tilde{\varphi})$ with N basis and sampling functions such that*

- *the basis functions are rapidly decaying L_2 functions such that the family $\{\varphi_i(\cdot - Nk)\}_{i=1 \dots N, k \in \mathbb{Z}}$ is a Riesz basis in the sense of (6), with approximation power of order $L \geq 1$ (see Definition 5);*

- The sampling functions are rapidly decaying generalized functions such that $(\varphi, \tilde{\varphi})$ is biorthonormal of order L (see Definition 6);
- There exists an integer $0 \leq p \leq L$ such that, for any $\omega \in \mathbb{R}$, any $0 \leq k \leq L$, and any $1 \leq i \leq N$,

$$|\hat{\tilde{\varphi}}_i^{(k)}(\omega)| \leq C \max(1, |\omega|^p). \quad (54)$$

Then, for any $f \in W^{L+\max(p+1/2, 1)}(\mathbb{R})$, we have that

$$\|f - \mathcal{Q}_T f\|_{L_2} \underset{T \rightarrow 0}{\sim} \sqrt{\frac{\mathcal{E}^{(2L)}(0)}{(2L)!}} \|f^{(L)}\|_{L_2} T^L. \quad (55)$$

The two first conditions in Theorem 10 are necessary to ensure that the approximation scheme has an approximation power of order L . The last condition allow us to show that such an approximation power is attained with restricted condition on the sampling functions having possibly unbounded Fourier transform.

Proof. We prove that the conditions of Theorem 10 imply that we fulfill the hypotheses of both Propositions 8 and 9.

Since $\varphi, \tilde{\varphi}$ are rapidly decaying (generalized) functions, their Fourier transforms $\hat{\varphi}$ and $\hat{\tilde{\varphi}}$ are infinitely differentiable. The same holds true for $\hat{\varphi}_d = \mathbf{G}_\varphi^{-1} \hat{\varphi}$ since \mathbf{G}_φ is a matrix-valued infinitely differentiable function with infinitely differentiable inverse \mathbf{G}_φ^{-1} due to the Riesz basis condition. This implies that \mathcal{E} is infinitely differentiable and therefore $(2L+1)$ -times continuously differentiable.

The basis functions are rapidly decaying, hence in $L_1(\mathbb{R})$, implying that $\hat{\varphi}$ is bounded. The Riesz basis condition then easily implies that both \mathbf{G}_φ and its inverse are bounded as functions of ω . It then follows that $\mathcal{E}_{\min}(\omega)$ is bounded, while $\mathcal{E}_{\text{res}}(\omega)$ is dominated by $\|(\hat{\varphi} - \hat{\varphi}_d)(\omega)\|$, the latter being dominated by $\max(1, |\omega|^{2p})$ due to (54). It finally implies (42) for some constant $C > 0$. Similarly, for $k \leq L$, the function $\hat{\varphi}^{(k)}$ is bounded, being the Fourier transform of $t \mapsto t^k \varphi(t)$, which is in $L_1(\mathbb{R})$ due to the rapid decay of φ . Again, this property is transferred to the derivative of \mathbf{G}_φ and its inverse. Then, exploiting Leibnitz rule, one shows that $\mathcal{E}_{\min}^{(2L+1)}$ is bounded, while $\mathcal{E}^{(2L+1)}$ is a sum of products constituted of bounded terms and of two terms of the form $\hat{\tilde{\varphi}}_i(\omega)$, which are controlled by (54). Putting things together, we easily deduce (48) for some constant $C > 0$.

Finally, the hypotheses of Propositions 8 and 9 are satisfied, implying (43) and (49) together with

$$\|f - \mathcal{Q}_T f\|_{L_2} = \sqrt{\frac{\mathcal{E}^{(2L)}(0)}{(2L)!}} \|f^{(L)}\|_{L_2} T^L + O(T^{L+1/2}). \quad (56)$$

This proves (55). □

Note that [20, Theorem 4] corresponds to Theorem 10 with $p = 1$, and with less restrictive assumptions on φ and $\tilde{\varphi}$ (essentially, the rapid decay is replaced by polynomial decay adapted to L). The condition $f \in W^{L+\max(p+1/2,1)}(\mathbb{R})$ reflects the cost of covering the scenario of possibly irregular sampling functions with unbounded Fourier transform.

3.3 Approximation Properties of Hermite Splines

We can now evaluate the approximation error on f in different frameworks, including the Hermite scheme. In our analysis, for $(\varphi, \tilde{\varphi})$ being fixed, we also quantify the error on the derivative f' when we approximate it by $(\mathcal{Q}_T f)'$. We therefore study the quantities $\|f - \mathcal{Q}_T f\|_{L_2}$ and $\|f' - (\mathcal{Q}_T f)'\|_{L_2}$. Knowing the order of approximation $L \geq 1$, the quality of the approximation is quantified by the two asymptotic constants

$$\mathbf{C}_{\tilde{\varphi}}^{\tilde{\varphi}} = \begin{pmatrix} C_{\tilde{\varphi},1}^{\tilde{\varphi}} \\ C_{\tilde{\varphi},2}^{\tilde{\varphi}} \end{pmatrix} = \begin{pmatrix} \lim_{T \rightarrow 0} T^{-L} \|f^{(L)}\|_{L_2}^{-1} \|f - \mathcal{Q}_T f\|_{L_2} \\ \lim_{T \rightarrow 0} T^{-(L-1)} \|f^{(L-1)}\|_{L_2}^{-1} \|f' - (\mathcal{Q}_T f)'\|_{L_2} \end{pmatrix}. \quad (57)$$

The asymptotic constant of $C_{\tilde{\varphi},2}^{\tilde{\varphi}}$ can be computed with the same tools as that of $C_{\tilde{\varphi},1}^{\tilde{\varphi}}$. Using integration by parts, we indeed have that

$$(\mathcal{Q}_T f)' = \frac{1}{T} \sum_{i=1}^N \sum_{k \in \mathbb{Z}} \left\langle f, \frac{1}{T} \tilde{\varphi}_i \left(\frac{\cdot}{T} - k \right) \right\rangle \varphi_i' \left(\frac{\cdot}{T} - k \right) \quad (58)$$

$$= \frac{1}{T} \sum_{i=1}^N \sum_{k \in \mathbb{Z}} \left\langle f', \frac{1}{T} \tilde{\varphi}_{i,\text{int}} \left(\frac{\cdot}{T} - k \right) \right\rangle \varphi_i' \left(\frac{\cdot}{T} - k \right), \quad (59)$$

where the new sampling functions are best defined in the Fourier domain as

$$\widehat{\tilde{\varphi}_{i,\text{int}}}(\omega) = -\frac{1}{j\omega} \widehat{\tilde{\varphi}_i}(\omega). \quad (60)$$

As one power of T gets lost in the differentiation process, the rate of decay of the error on f' is thus equal to $(L - 1)$. Note that the apparent singularity in (60) around 0 is counter-balanced in the analysis since one only approximates functions f' that are derivatives functions, hence for which $\widehat{f}'(\omega) = j\omega \widehat{f}(\omega)$.

In the Hermite framework, for which $N = 2$, sampling functions are taken as $\tilde{\varphi}_1 = \delta$ and $\tilde{\varphi}_2 = -\delta'$ and basis functions as (2) and (3). For comparison purpose, we also consider two relevant schemes that fit our analysis framework: classical cubic B-splines and interlaced derivative sampling. Cubic B-spline approximation corresponds to $N = 1$, with $\tilde{\varphi} = \sum_{k \in \mathbb{Z}} (b^3)^{-1} [k] \delta(\cdot - k)$, where $(b_1^3)^{-1}$ is the direct B-spline filter, and $\tilde{\varphi} = \beta^3$ the cubic B-spline [31]. Interlaced derivative sampling can be defined in the more general framework of generalized sampling without band-limited constraints [36]. In this setting, $N = 2$ and

the sampling functions correspond to $\tilde{\varphi}_1 = \delta$ and $\tilde{\varphi}_2 = -\delta'(\cdot - \frac{1}{2})$. The basis functions are constructed from the cubic B-spline to allow for a fair comparison, and are given by

$$\hat{\varphi}_1(\omega) = \frac{3e^{-2j\omega}(-1 + e^{j\omega})^4}{2\omega^4}, \quad (61)$$

$$\hat{\varphi}_2(\omega) = \frac{e^{-2j\omega}(-1 + e^{j\omega})^4(1 + e^{j\omega}(-4 + e^{j\omega}))}{(2 - 2e^{2j\omega})\omega^4}. \quad (62)$$

in the Fourier domain. A comprehensive description of their derivation is provided in [38]. It is worth noting that the sampling functions do not correspond to the dual functions in these frameworks. This is easily motivated by practical considerations: the sampling process must, in practice, be implemented with digital filtering, excluding dual functions due to their continuous nature. The dual functions can, however, still be constructed from the Gram matrix following (26) so as to estimate the optimal approximation error.

We now reveal the approximation power of these three different schemes. The results are known for the cubic B-splines [19] and are included for comparison purposes. For Hermite approximation and interlaced sampling, they are deduced from Theorem 10, and are not included in the multigenerator framework [20], whose hypotheses exclude the use of derivative samples.

Proposition 11. *Let $(\varphi, \tilde{\varphi})$ be one of the three approximation schemes considered above. Then, we have that*

$$\begin{aligned} \|f - \mathcal{Q}_T f\|_{L_2} &\underset{T \rightarrow 0}{\sim} \frac{1}{72\sqrt{70}} \|f^{(4)}\|_{L_2} T^4, \text{ and} \\ \|f' - (\mathcal{Q}_T f)'\|_{L_2} &\underset{T \rightarrow 0}{\sim} \frac{1}{12\sqrt{210}} \|f^{(3)}\|_{L_2} T^3, \end{aligned} \quad (63)$$

for every $f \in W_2^5(\mathbb{R})$ (cubic B-splines), or $f \in W_2^{11/2}(\mathbb{R})$ (Hermite splines or interlaced derivative sampling). Moreover, in the three cases, we have the following relation to the optimal approximation scheme associated to φ :

$$\begin{aligned} \|f - \mathcal{Q}_T f\|_{L_2} &\underset{T \rightarrow 0}{\sim} \sqrt{\frac{10}{3}} \|f - \mathcal{P}_T f\|_{L_2}, \text{ and} \\ \|f' - (\mathcal{Q}_T f)'\|_{L_2} &\underset{T \rightarrow 0}{\sim} \|f' - (\mathcal{P}_T f)'\|_{L_2}, \end{aligned} \quad (64)$$

for every $f \in W_2^5(\mathbb{R})$ (cubic B-splines), or $f \in W_2^{11/2}(\mathbb{R})$ (Hermite splines and interlaced derivative sampling).

Proof. The three frameworks are readily known to define approximation schemes of order $L = 4$ [31, 11, 36]. All the considered basis functions specify a Riesz basis, are rapidly decaying (Hermite and cubic B-splines are compactly supported, while the basis functions

for interlaced derivative sampling, despite being non-compactly supported, are exponentially decaying [36]), and reproduce polynomials up to degree 3. The sampling functions are rapidly decaying generalized functions. Indeed, they are compactly supported for the Hermite and interlaced derivative sampling schemes. For cubic B-splines, we have seen that $\tilde{\varphi} = \sum_{k \in \mathbb{Z}} (b^3)^{-1} [k] \delta(\cdot - k)$, where the sequence $(b^3)^{-1}$ is known to be exponentially decaying [31], implying the result. In addition, for the Hermite and interlaced derivative sampling schemes (for cubic splines, respectively), (54) is clearly satisfied with $p = 1$ ($p = 0$, respectively). The conditions of Theorem 10 are therefore satisfied for $L = 4$, implying (55).

The value of $\mathcal{E}^{(2L)}(0)$ is computed by computing the Taylor expansion of the kernel \mathcal{E} around 0^2 . The analysis of the approximation error on the derivative follows the same principle, the kernel being given for $(\varphi', \tilde{\varphi}_{\text{int}})$ according to (60), giving (63).

We obtain the asymptotic behaviour of the optimal approximation errors $\|f - \mathcal{P}_T f\|_{L_2}$ and $\|f' - (\mathcal{P}_T f)'\|_{L_2}$ associated to $\tilde{\varphi} = \varphi_d$ in the same way, leading to (64). \square

Our findings call for the following comments.

- The three compared schemes have the same approximation order, the same approximation constant, and the same optimal approximation constant (associated to (φ, φ_d)) for reconstructing both f and its derivative.
- In all cases, the sampling functions result in a near-to-optimal asymptotic constant $C_{\varphi,1}^{\tilde{\varphi}}$ for the reconstruction of f . This minor discrepancy of a factor of $\sqrt{10/3} \approx 1.83$ is expected from the fact that the sampling functions are not dual functions. It is also remarkable that the reconstruction of the derivative, while not being associated to the dual functions, is associated to an optimal approximation constant $C_{\varphi,2}^{\tilde{\varphi}}$.
- Dual functions, although offering the smallest approximation error, have strong practical disadvantages. First, the φ_d , are non-compactly supported splines. More importantly, they cannot be easily implemented as they do not have a digital filter equivalent (*i.e.*, computing the $\langle f, \varphi_d \rangle$ does not only depend on knowing f and its derivative on a fixed grid, in contrast to usual cubic B-splines and Hermite splines). Finally, they do not possess a closed-form expression in general. For these reasons, the sampling functions $\tilde{\varphi}$ classically used in the three considered approximation schemes, although non-optimal, are preferable in practice.

In Table 1, we sum up these findings, including the asymptotic constants for cubic B-spline, interlaced derivative sampling and cubic Hermite splines and their comparison with optimal constants. Hermite interpolation is thus not a unique way of approximating a function and its first derivative, even if one wishes the error to remain close to optimal. The notable difference lies in the fact that the Hermite scheme provides functions that are simultaneously of finite support, which is not the case for interlaced derivative sampling (see (61) and (62)), and interpolating, which is not the case for cubic B-splines.

²We relied on the technical computing software Mathematica 11 for this task.

Table 1 – Comparison of approximation methods.

Approximation Method	Cubic B-splines	Interlaced Derivative Sampling	Hermite Splines
Digital filter implementation	✓	✓	✓
Interpolating	✗	✓	✓
Finite Support	✓	✗	✓
Closed-form expression	✓	✗	✓
Rate of Decay (L)	4	4	4
Asymptotic Constant ($C_{\varphi,1}^{\tilde{\varphi}}$)	$\frac{1}{72\sqrt{70}}$	$\frac{1}{72\sqrt{70}}$	$\frac{1}{72\sqrt{70}}$
Ratio to optimal ($C_{\varphi,1}^{\tilde{\varphi}}$)	$\frac{\sqrt{3}}{\sqrt{10}}$	$\frac{\sqrt{3}}{\sqrt{10}}$	$\frac{\sqrt{3}}{\sqrt{10}}$
Rate of Decay ($L - 1$)	3	3	3
Asymptotic Constant ($C_{\varphi,2}^{\tilde{\varphi}}$)	$\frac{1}{12\sqrt{210}}$	$\frac{1}{12\sqrt{210}}$	$\frac{1}{12\sqrt{210}}$
Ratio to optimal ($C_{\varphi,2}^{\tilde{\varphi}}$)	1	1	1

4 Concluding Remarks

Our work focused on the formal investigation of two practical aspects of Hermite splines, namely their short support and their good approximation properties. We show that Hermite splines are of minimal support among pairs of functions with similar reproduction properties, and provide a framework to quantify their approximation power. The latter result not only allows us to prove that Hermite splines are asymptotically identical to cubic B-splines, but also offers a more general framework for the quantitative approximation of functions and their derivatives.

In summary, Hermite splines are found to offer an approximation scheme that (1) has the same approximation power than the notorious cubic B-splines, (2) is interpolating (possibly with the derivative), (3) is based on maximally localized compactly supported basis functions. The resulting cost is the need to use two basis functions instead of a single one, and have access to derivative samples.

5 Acknowledgements

This work was supported by core funding from EMBL, the Swiss National Science Foundation under Grant 200020_162343/1, and the European Research Council under Grant H2020-ERC (ERC grant agreement No 692726 - GlobalBioIm). The authors thank Yoann Pradat for useful remarks on the manuscript.

Competing interests: The authors have no competing interests.

References

- [1] I. Schoenberg, *Cardinal Spline Interpolation*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1973.
- [2] P. Lipow and I. Schoenberg, “Cardinal interpolation and spline functions. III. Cardinal Hermite interpolation,” *Linear Algebra and its Applications*, vol. 6, pp. 273–304, 1973.
- [3] I. Schoenberg and A. Sharma, “Cardinal interpolation and spline functions. V. The B-splines for cardinal Hermite interpolation,” *Linear Algebra and its Applications*, vol. 7, no. 1, pp. 1–42, January 1973.
- [4] M. Unser, A. Aldroubi, and M. Eden, “B-Spline signal processing: Part I—Theory,” *IEEE Transactions on Signal Processing*, vol. 41, no. 2, pp. 821–833, February 1993.
- [5] —, “B-Spline signal processing: Part II—Efficient design and applications,” *IEEE Transactions on Signal Processing*, vol. 41, no. 2, pp. 834–848, February 1993.
- [6] C. De Boor, *A practical guide to splines*. New York, NY, USA: Springer-Verlag, 1978.
- [7] R. Farouki, “The bernstein polynomial basis: A centennial retrospective,” *Computer Aided Geometric Design*, vol. 29, no. 6, pp. 379–419, August 2012.
- [8] H. Prautzsch, W. Boehm, and M. Paluszny, *Bézier and B-spline techniques*. Berlin, Germany: Springer-Verlag, 2013.
- [9] G. Farin, *Curves and surfaces for CAGD: a practical guide*. Burlington, MA, USA: Morgan Kaufmann Publishers, 2002.
- [10] W. Böhm, G. Farin, and J. Kahmann, “A survey of curve and surface methods in cagd,” *Computer Aided Geometric Design*, vol. 1, no. 1, pp. 1–60, July 1984.
- [11] V. Uhlmann, J. Fageot, and M. Unser, “Hermite snakes with control of tangents,” *IEEE Transactions on Image Processing*, vol. 25, no. 6, pp. 2803–2816, June 2016.
- [12] W. Dahmen, B. Han, R.-Q. Jia, and A. Kunoth, “Biorthogonal multiwavelets on the interval: Cubic Hermite splines,” *Constructive Approximation*, vol. 16, no. 2, pp. 221–259, February 2000.
- [13] R. Warming and R. Beam, “Discrete multiresolution analysis using hermite interpolation: Biorthogonal multiwavelets,” *SIAM Journal on Scientific Computing*, vol. 22, no. 4, pp. 1269–1317, July 2000.
- [14] V. Uhlmann, J. Fageot, H. Gupta, and M. Unser, “Statistical optimality of Hermite splines,” in *Proceedings of the Eleventh International Workshop on Sampling Theory and Applications (SampTA’15)*, Washington DC, USA, May 25–29, 2015, pp. 226–230.

- [15] R. Delgado-Gonzalo, P. Thévenaz, and M. Unser, “Exponential splines and minimal-support bases for curve representation,” *Computer Aided Geometric Design*, vol. 29, no. 2, pp. 109–128, February 2012.
- [16] T. Blu, P. Thévenaz, and M. Unser, “MOMS: Maximal-order interpolation of minimal support,” *IEEE Transactions on Image Processing*, vol. 10, no. 7, pp. 1069–1080, July 2001.
- [17] C. Hall, “On error bounds for spline interpolation,” *Journal of approximation theory*, vol. 1, no. 2, pp. 209–218, September 1968.
- [18] G. Birkhoff and A. Priver, “Hermite interpolation errors for derivatives,” *Studies in Applied Mathematics*, vol. 46, no. 1-4, pp. 440–447, April 1967.
- [19] T. Blu and M. Unser, “Quantitative Fourier analysis of approximation techniques: Part I—Interpolators and projectors,” *IEEE Transactions on Signal Processing*, vol. 47, no. 10, pp. 2783–2795, October 1999.
- [20] —, “Approximation error for quasi-interpolators and (multi-) wavelet expansions,” *Applied and Computational Harmonic Analysis*, vol. 6, no. 2, pp. 219–251, March 1999.
- [21] L. Condat, “Reconstruction of derivatives: Error analysis and design criteria,” in *Signal Processing Conference, 2011 19th European*. IEEE, 2011, pp. 839–843.
- [22] L. Condat and T. Moller, “Quantitative error analysis for the reconstruction of derivatives,” *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2965–2969, 2011.
- [23] C. de Boor, R. DeVore, and A. Ron, “Approximation from shift-invariant subspaces of $L_2(\mathbb{R}^d)$,” *Transactions of the American Mathematical Society*, vol. 341, no. 2, pp. 787–806, 1994.
- [24] —, “The structure of finitely generated shift-invariant spaces in $L_2(\mathbb{R}^d)$,” *Journal of Functional Analysis*, vol. 119, no. 1, pp. 37–78, 1994.
- [25] —, “Approximation orders of FSI spaces in $L_2(\mathbb{R}^d)$,” *Constructive approximation*, vol. 14, no. 4, pp. 631–652, 1998.
- [26] O. Holtz and A. Ron, “Approximation orders of shift-invariant subspaces of $W_2^s(\mathbb{R}^d)$,” *arXiv preprint math/0512609*, 2005.
- [27] A. Ron, “Factorization theorems for univariate splines on regular grids,” *Israel Journal of Mathematics*, vol. 70, no. 1, pp. 48–68, 1990.
- [28] T. Blu, P. Thévenaz, and M. Unser, “Moms: Maximal-order interpolation of minimal support,” *IEEE Transactions on Image Processing*, vol. 10, no. 7, pp. 1069–1080, 2001.

- [29] K. Jetter and G. Plonka, “A survey on l_2 -approximation order from shift-invariant spaces,” in *Multivariate approximation and applications*. Cambridge University Press, 2001, pp. 73–111.
- [30] C. De Boor and R. DeVore, “Partitions of unity and approximation,” *Proceedings of the American Mathematical Society*, vol. 93, no. 4, pp. 705–709, 1985.
- [31] M. Unser, “Splines: A perfect fit for signal and image processing,” *IEEE Signal Processing Magazine*, vol. 16, no. 6, pp. 22–38, November 1999.
- [32] L. Schwartz, *Théorie des distributions*. Hermann, 1966.
- [33] A. Papoulis, “Generalized sampling expansion,” *IEEE Transactions on Circuits and Systems*, vol. 24, no. 11, pp. 652–654, November 1977.
- [34] A. Aldroubi, “Oblique projections in atomic spaces,” *Proceedings of the American Mathematical Society*, vol. 124, no. 7, pp. 2051–2060, 1996.
- [35] G. Strang and G. Fix, “A Fourier analysis of the finite element variational method,” in *Constructive Aspect of Functional Analysis*. Rome, Italy: Cremonese, 1971, pp. 796–830.
- [36] M. Unser and J. Zerubia, “A generalized sampling theory without band-limiting constraints,” *IEEE Transactions on Circuits and Systems—II: Analog and Digital Signal Processing*, vol. 45, no. 8, pp. 959–969, August 1998.
- [37] —, “Generalized sampling: Stability and performance analysis,” *IEEE Transactions on Signal Processing*, vol. 45, no. 12, pp. 2941–2950, December 1997.
- [38] V. Uhlmann, “Landmark active contours for bioimage analysis: A tale of points and curves,” EPFL Thesis no. 7951 (2017), 263 p., Swiss Federal Institute of Technology Lausanne (EPFL), December 1, 2017.