Approximation of Non-Decaying Signals From Shift-Invariant Subspaces

Ha Q. Nguyen and Michael Unser

Abstract

In our recent work, the sampling and reconstruction of non-decaying signals, modeled as members of weighted-L_p spaces, were shown to be stable with an appropriate choice of the generating kernel for the shift-invariant reconstruction space. In this paper, we extend the Strang-Fix theory to show that, for d-dimensional signals whose derivatives up to order L are all in some weighted-L_p space, the weighted norm of the approximation error can be made to go down as $O(h^L)$ when the sampling step h tends to 0. The sufficient condition for this decay rate is that the generating kernel belongs to a particular hybrid-norm space and satisfies the Strang-Fix conditions of order L. We show that the $O(h^L)$ behavior of the error is attainable for both approximation schemes using projection (when the signal is prefiltered with the dual kernel) and interpolation (when a prefilter is unavailable). The requirement on the signal for the interpolation method, however, is slightly more stringent than that of the projection because we need to increase the smoothness of the signal by a margin of $d/p + \varepsilon$, for arbitrary $\varepsilon > 0$. This extra amount of derivatives is used to make sure that the direct sampling is stable.

Index Terms

Approximation theory, Strang-Fix conditions, shift-invariant spaces, spline interpolation, weighted L_p spaces, weighted Sobolev spaces, hybrid-norm spaces

I. Introduction

Sampling and reconstruction are important in signal processing because they provide an insightful connection between analog signals and their discrete representations. In the sampling procedure, oftentimes, a continuous-domain signal $f : \mathbb{R}^d \mapsto \mathbb{C}$ is uniformly sampled (with or without a *prefilter*) at multi-integer multiples of some sampling step h to produce a discrete-domain signal $c : \mathbb{Z}^d \mapsto \mathbb{C}$. The reconstruction, on the other hand, is commonly done by interpolating the samples $\{c[\boldsymbol{k}]\}_{\boldsymbol{k}\in\mathbb{Z}^d}$ with scaled and shifted copies of some kernel (generating function) φ positioned on the grid $h\mathbb{Z}^d$. Precisely, the reconstructed signal takes the (integer) shift-invariant form

$$\tilde{f}(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} c[\boldsymbol{k}]\varphi\left(\frac{\boldsymbol{x}}{h} - \boldsymbol{k}\right). \tag{1}$$

This interpolation model has been extensively used in the theory of splines [1]–[4]. It is general enough to include the celebrated reconstruction formula in Shannon's sampling theorem [5] in which the kernel φ is replaced with the sinc function. Although the sinc-based interpolation guarantees exact recovery of bandlimited signals (or signals prefiltered with an ideal lowpass filter) whenever $1/h$ exceeds Nyquist's rate, the slow decay of $\mathrm{sinc}(\boldsymbol{x})$ unfortunately prevents the application of this method in practice [6]. For other choices of φ with better localization properties, such as splines, exact reconstruction is no longer achievable but the quality of the approximation of a signal f by such $\tilde{f}$ given in (1) can be characterized as a power of the sampling step h via the Strang-Fix theory. Specifically, in early 1970's, Strang and Fix [7] extended Schoenberg's work [1] and introduced the concept of *controlled* approximation in which the ℓ_2-norm of the sampled coefficients is bounded by the L_2-norm of the original signal. They showed that, for compactly supported φ, the error of the controlled approximation is bound as

$$\forall f \in H_2^L(\mathbb{R}^d), \quad \min_c \left\| f - \tilde{f} \right\|_{L_2(\mathbb{R}^d)} \leq C_{\varphi,L} \cdot h^L \cdot \left\| f^{(L)} \right\|_{L_2(\mathbb{R}^d)}, \quad \text{as } h \to 0, \tag{2}$$

This research was funded by the Swiss National Science Foundation under Grant 200020-162343.

Ha Q. Nguyen is with the Viettel Research and Development Institute, Hoa Lac High-tech Park, Hanoi, Vietnam. The majority of this work was done when he was with the Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne (EPFL), Station 17, CH-1015, Lausanne, Switzerland (email: hanq8@viettel.com.vn).

Michael Unser is with the Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne (EPFL), Station 17, CH-1015, Lausanne, Switzerland (email: michael.unser@epfl.ch).

if and only if φ satisfies the Strang-Fix conditions of order L so that the representation (1) is able to reproduce all polynomials of degree less than L; this notion will be clarified later in Section II-D. Here, $f^{(L)}$ is the Lth derivative[1] of f and $H_2^L(\mathbb{R}^d)$ is the Sobolev space of L_2 functions whose first L derivatives are all in $L_2(\mathbb{R}^d)$. The order L in (2) is referred to in the literature as the order (power) of approximation.

The original result of Strang and Fix has been extended in various directions, including controlled L_p-approximation with globally supported (multi-) kernel [8]–[11], uncontrolled L_2-approximation [12], and finer estimations of the L_2-approximation error [13]–[17]; interested readers are also referred to the surveys [18]–[20]. More recently, the Strang-Fix theory was linked to the sampling of signals with finite rate of innovation [21]. Despite a rich literature on the Strang-Fix conditions, none of the existing results allows us to deal with the approximation of non-decaying (non-L_p) signals, such as sample paths of a Brownian motion, which can even grow at infinity. This is an important part that seems to be missing, and which, for instance, is relevant to the theory of *sparse stochastic processes* recently developed by Unser *et al.* [22]–[24].

In this paper, a follow-up of our recent works on the sampling theory for non-decaying signals [25]–[27], we provide an approximation theory for such objects. Recall that we showed in [25] that both the sampling and reconstruction of weighted-L_p signals, at a fixed sampling step, are stable, provided the generating kernel φ lies in an appropriate *hybrid-norm space*, a concept closely related to the Wiener amalgams that are frequently used in time-frequency analysis [28]–[30]. Note that, in the *direct* sampling scheme, where a prefilter is absent, not only the signal is required to live in a weighted-L_p space, but also its first $d/p+\varepsilon$ derivatives, for some $\varepsilon>0$. In the spirit of [25], we model non-decaying signals in this paper as members of the weighted space $L_{p,-\alpha}(\mathbb{R}^d)$ associated with the *Sobolev weight* $(1+\|\cdot\|^2)^{-\alpha/2}$, where $\alpha\geq 0$ specifies the order of growth of the signals. In particular, $f\in L_{p,-\alpha}(\mathbb{R}^d)$ if $(1+\|\cdot\|^2)^{-\alpha/2}f\in L_p(\mathbb{R}^d)$. We then extend the classical Strang-Fix theory to the approximation of such signals for the two common types of shift-invariant reconstructions: projection versus (direct) interpolation.

In the *projection* scheme, which provides the optimal L_2-approximation, the original signal is prefiltered with the *dual* kernel $h^{-d}\varphi_{\mathrm{d}}\left(-\frac{\cdot}{h}\right)$ [14] and the coefficients $\{c[\boldsymbol{k}]\}_{\boldsymbol{k}\in\mathbb{Z}^d}$ in (1) are obtained by sampling the resulting signal with step size h. It means that the reconstructed signal is given by

$$\tilde{f}_{\mathrm{proj}}(\boldsymbol{x})=\frac{1}{h^d}\sum_{\boldsymbol{k}\in\mathbb{Z}^d}\left\langle f,\varphi_{\mathrm{d}}\left(\frac{\cdot}{h}-\boldsymbol{k}\right)\right\rangle\varphi\left(\frac{\boldsymbol{x}}{h}-\boldsymbol{k}\right).$$

For this type of reconstruction, we show, in the first half of the paper, that if φ belongs to an appropriate hybrid-norm space and at the same time satisfies the Strang-Fix conditions of order L, then the weighted-L_p norm of the projection error is bounded as

$$\forall f\in H^L_{p,-\alpha}(\mathbb{R}^d),\quad \left\|f-\tilde{f}_{\mathrm{proj}}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)}\leq C_{\varphi,L,\alpha}\cdot h^L\cdot\left\|f^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)},\quad \text{as } h\to 0, \tag{3}$$

where the weighted Sobolev space $H^L_{p,-\alpha}(\mathbb{R}^d)$ is a collection of functions whose derivatives up to order L are all in $L_{p,-\alpha}(\mathbb{R}^d)$. We want to remark that this result is the weighted version of [11, Theorem 2.2].

In the *interpolation* scheme, the coefficients are sampled directly from the original signal; hence the reconstructed signal takes the form

$$\tilde{f}_{\mathrm{int}}(\boldsymbol{x})=\sum_{\boldsymbol{k}\in\mathbb{Z}^d}f(h\boldsymbol{k})\,\varphi_{\mathrm{int}}\left(\frac{\boldsymbol{x}}{h}-\boldsymbol{k}\right),$$

where φ_{int} is the *interpolant* generated from the kernel φ [6]. Similar to the projection case, we establish, in the second half of the paper, that if φ is an element of a particular hybrid-norm space that satisfies the Strang-Fix condition of order L, then, given $r>d/p$,

$$\forall f: D^r f\in H^L_{p,-\alpha}(\mathbb{R}^d),\quad \left\|f-\tilde{f}_{\mathrm{int}}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)}\leq C_{\varphi,L,\alpha}\cdot h^L\cdot\left\|(D^r f)^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)},\quad \text{as } h\to 0. \tag{4}$$

Here, $D^r f$ is a combination of all *fractional* derivatives up to order r of f defined in the frequency domain as $D^r f:=\mathcal{F}^{-1}\left\{(1+\|\cdot\|^2)^{r/2}\mathcal{F}f\right\}$ with $\mathcal{F}$ being the Fourier transform operator. Informally speaking, the interpolation error can also be made to decay like $O(h^L)$, when h tends to 0, for functions whose derivatives up to order $L+d/p+\varepsilon$ live in some weighted-L_p space, for arbitrary $\varepsilon>0$. This is not surprising because we need $d/p+\varepsilon$ derivatives to

[1]To be precise, when f is multivariate, $f^{(L)}$ is the summation of (the moduli of) all partial derivatives of order L of f.

take care of the sampling, as indicated in [25], and L derivatives more to reach the target approximation order. To the best of our knowledge, the bound (4) is new even in the unweighted L_p case (when all instances of the subscript α disappear), although similar results exist for the direct interpolation in L_2 [14] and L_∞ [9]. The (unweighted) L_p result presented in [10, Theorem 4.1], although similar to (4), does not fall into the realm of direct interpolation because the samples are taken from a smoothed version of the original signal. It is important to remark that the decay rate in (4) is integer while the order of smoothness of f is fractional. An extension of such bound for fractional decay rates, in a similar vein to the unweighted approximations in [31]–[35], might be considered for future research.

One of the challenges for the approximation in weighted spaces is that the beautiful Fourier-based methods commonly used in the Strang-Fix theory [7], [12]–[17] are no longer applicable, even in the weighted-L_2 case, due to the lack of a Parseval-type relation. In proving the bounds (3) and (4), we adapt the L_p-approximation techniques in [10], [11], which are carried entirely in the space domain, but our analysis is much more involved because of the handling of the weights. We also heavily rely on the preliminary results in [25]. Other works that are closely related to the present paper are [36], [37] in which similar bounds were derived in the weighted-L_p spaces associated with the so-called *Muckenhoupt weights* [38]. These weights, however, are strikingly different from the Sobolev weights used in this paper. They are characterized by the boundedness of the Hardy-Littlewood maximal operator [39]–[41] with respect to the weighted norm. Typical examples of the Muckenhoupt weights are $\|\cdot\|^\alpha$, for α being restricted in the interval $(-d/p, d-d/p)$ (cf. [42]). By contrast, the Sobolev weights $(1+\|\cdot\|^2)^{\alpha/2}$ can take arbitrary order $\alpha \in \mathbb{R}$ and therefore give us more freedom in quantifying the growth or decay of the signals. Moreover, the Muckenhoupt weights are not well-suited to time-frequency analysis because they are generally not *submultiplicative*, an important property that is satisfied by the Sobolev weights (cf. [43, Section 9]).

The remainder of the paper is organized as follows: preliminary notions are introduced in Section II; approximation error bounds for the projection and interpolation paradigms are derived in Sections III and IV, respectively; proofs of several auxiliary results are given in Section V.

II. Preliminaries

A. Notation

All functions in this paper are mappings from $\mathbb{R}^d$ to $\mathbb{C}$ for a fixed dimension $d \geq 1$. Vectors in $\mathbb{R}^d$ are denoted by bold letters and their Euclidean norms are denoted by $\|\cdot\|$. The constants throughout the paper are denoted by C with subscripts indicating the dependence of the constants on some parameters; we use the same notation for different constants that depend on the same set of parameters. The restriction of a function f on the multi-integer grid $\mathbb{Z}^d$ is denoted by $f[\cdot]$. $\mathbb{N}$ is the set of natural numbers and $\mathbb{Z}_+$ is the set of nonnegative integers, i.e., $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$. For brevity, we denote by $\langle\cdot\rangle$ the Sobolev weighting function $(1+\|\cdot\|^2)^{1/2}$. For $1 \leq p \leq \infty$, we use p' to denote the Hölder conjugate of p that satisfies $\frac{1}{p}+\frac{1}{p'}=1$.

$\mathcal{C}_{\mathrm{c}}^\infty(\mathbb{R}^d)$ is the space of smooth and compactly supported functions, $\mathcal{S}(\mathbb{R}^d)$ is Schwartz' class of smooth and rapidly decaying functions, and $\mathcal{S}'(\mathbb{R}^d)$ is the space of tempered distributions, which are continuous linear functionals on $\mathcal{S}(\mathbb{R}^d)$. As usual, the notation $\langle\cdot,\cdot\rangle$ is used interchangeably for the scalar product and for the action of a distribution on a test function. The (distributional) Fourier transform $\hat{f} = \mathcal{F}f$ of a tempered distribution $f \in \mathcal{S}'(\mathbb{R}^d)$ is also a tempered distribution defined as

$$\langle \mathcal{F}f, \varphi\rangle := \left\langle \hat{f}, \varphi \right\rangle := \langle f, \hat{\varphi}\rangle\,, \quad \text{for } \varphi \in \mathcal{S}(\mathbb{R}^d),$$

where

$$\hat{\varphi}(\boldsymbol{\omega}) := \int_{\mathbb{R}^d} \varphi(\boldsymbol{x}) \mathrm{e}^{-\mathrm{j}\langle \boldsymbol{\omega}, \boldsymbol{x}\rangle} \mathrm{d}\boldsymbol{x}.$$

We denote the inverse Fourier-transform operator by $\mathcal{F}^{-1}$. For a multi-index $\boldsymbol{\ell} \in \mathbb{Z}_+^d$, $|\boldsymbol{\ell}| := \sum_{i=1}^d \ell_i$ and $\partial^{\boldsymbol{\ell}}$ is a shorthand for $(\partial/\partial x_1)^{\ell_1} \cdots (\partial/\partial x_d)^{\ell_d}$. The (distributional) partial derivative with respect to $\boldsymbol{\ell}$ of a tempered distribution $f \in \mathcal{S}'(\mathbb{R}^d)$ is also a tempered distribution defined as

$$\left\langle \partial^{\boldsymbol{\ell}} f, \varphi \right\rangle := (-1)^{|\boldsymbol{\ell}|} \left\langle f, \partial^{\boldsymbol{\ell}} \varphi \right\rangle, \quad \text{for } \varphi \in \mathcal{S}(\mathbb{R}^d).$$

We also use the notation

$$f^{(n)} := \sum_{\boldsymbol{\ell}\in\mathbb{Z}_+^d:|\boldsymbol{\ell}|=n} |\partial^{\boldsymbol{\ell}} f|.$$

$\nabla := (\partial/\partial x_1, \ldots, \partial/\partial x_d)$ is the gradient operator and $D_{\boldsymbol{u}} := \langle \nabla, \boldsymbol{u}\rangle$ is the directional derivative operator with respect to $\boldsymbol{u} \in \mathbb{R}^d$. The shift and difference operators are defined as $S_{\boldsymbol{u}} f := f(\cdot - \boldsymbol{u})$ and $\Delta_{\boldsymbol{u}} f := f - S_{\boldsymbol{u}} f$, respectively. For $h > 0$, σ_h denotes the scaling operator given by $\sigma_h f := f(\cdot/h)$.

B. Weighted Normed Spaces

The spaces $L_p(\mathbb{R}^d)$ and $\ell_p(\mathbb{Z}^d)$ and their corresponding norms $\|\cdot\|_{L_p(\mathbb{R}^d)}$ and $\|\cdot\|_{\ell_p(\mathbb{Z}^d)}$ are defined as usual. We also need the hybrid-norm space $W_p(\mathbb{R}^d)$ which comprises all functions f whose hybrid (mixed) norm

$$\|f\|_{W_p(\mathbb{R}^d)} := \begin{cases} \left(\int_{[0,1]^d} \left(\sum_{\boldsymbol{k}\in\mathbb{Z}^d} |f(\boldsymbol{x}+\boldsymbol{k})|\right)^p \mathrm{d}\boldsymbol{x}\right)^{1/p}, & 1 \le p < \infty \\ \operatorname{ess\,sup}_{\boldsymbol{x}\in[0,1]^d} \sum_{\boldsymbol{k}\in\mathbb{Z}^d} |f(\boldsymbol{x}+\boldsymbol{k})|, & p = \infty \end{cases}$$

is finite. For any weighting function w, the weighted spaces $L_{p,w}(\mathbb{R}^d)$, $\ell_{p,w}(\mathbb{Z}^d)$ and $W_{p,w}(\mathbb{R}^d)$ are defined with respect to the following weighted norms:

$$\begin{aligned} \|f\|_{L_{p,w}(\mathbb{R}^d)} &:= \|f\cdot w\|_{L_p(\mathbb{R}^d)}, \\ \|c\|_{\ell_{p,w}(\mathbb{Z}^d)} &:= \|c\cdot w[\cdot]\|_{\ell_p(\mathbb{Z}^d)}, \\ \|f\|_{W_{p,w}(\mathbb{R}^d)} &:= \|f\cdot w\|_{W_p(\mathbb{R}^d)}. \end{aligned}$$

When $w = \langle\cdot\rangle^\alpha$, for some $\alpha \in \mathbb{R}$, we write $L_{p,\alpha}(\mathbb{R}^d)$ for $L_{p,w}(\mathbb{R}^d)$, $\ell_{p,\alpha}(\mathbb{Z}^d)$ for $\ell_{p,w}(\mathbb{Z}^d)$, and $W_{p,\alpha}(\mathbb{R}^d)$ for $W_{p,w}(\mathbb{R}^d)$. Note that, for $\alpha \ge 0$, the weight $w = \langle\cdot\rangle^\alpha$ is (weakly) submultiplicative, i.e.,

$$\langle \boldsymbol{x}+\boldsymbol{y}\rangle^\alpha \le C_\alpha \langle\boldsymbol{x}\rangle^\alpha \langle\boldsymbol{y}\rangle^\alpha, \quad \forall \boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d,$$

which is equivalent to

$$\langle \boldsymbol{x}+\boldsymbol{y}\rangle^{-\alpha} \le C_\alpha \langle\boldsymbol{x}\rangle^\alpha \langle\boldsymbol{y}\rangle^{-\alpha}, \quad \forall \boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d.$$

Furthermore, the weight $w = \langle\cdot\rangle^\alpha$ satisfies the Gelfand-Raikov-Shilov condition [44] that

$$\lim_{n\to\infty} w(n\boldsymbol{x})^{1/n} = 1, \quad \forall\, \boldsymbol{x}\in\mathbb{R}^d.$$

These two properties of $\langle\cdot\rangle^\alpha$, with $\alpha \ge 0$, will be crucial for us to manipulate weights.

Finally, let us define the weighted Sobolev spaces of integer and fractional orders. Given $1 \le p \le \infty$ and $\alpha \in \mathbb{R}$, the space $H^k_{p,\alpha}(\mathbb{R}^d)$ with $k \in \mathbb{Z}_+$ consists of all $f \in \mathcal{S}'(\mathbb{R}^d)$ such that

$$\|f\|_{H^k_{p,\alpha}(\mathbb{R}^d)} := \sum_{\boldsymbol{\ell}\in\mathbb{Z}_+^d:|\boldsymbol{\ell}|\le k} \left\|\partial^{\boldsymbol{\ell}} f\right\|_{L_{p,\alpha}(\mathbb{R}^d)} < \infty.$$

It is straightforward that if $f \in H^k_{p,\alpha}(\mathbb{R}^d)$ then $f^{(n)} \in L_{p,\alpha}(\mathbb{R}^d)$, for all $n \le k$. Meanwhile, the space $L^s_{p,\alpha}$ with $s \in \mathbb{R}$ consists of all $f \in \mathcal{S}'(\mathbb{R}^d)$ such that

$$\|f\|_{L^s_{p,\alpha}(\mathbb{R}^d)} := \left\|\mathcal{F}^{-1}\left\{\langle\cdot\rangle^s \hat{f}\right\}\right\|_{L_{p,\alpha}(\mathbb{R}^d)} < \infty.$$

From here on, the term $\mathcal{F}^{-1}\left\{\langle\cdot\rangle^s \hat{f}\right\}$ will be abbreviated as $D^s f$. When $s > 0$, D^{-s} is the Bessel potential of order s [45]. We also need the hybrid weighted Sobolev space $H^{k,s}_{p,\alpha}$ which encompasses all $f \in \mathcal{S}'(\mathbb{R}^d)$ such that $D^s f \in H^k_{p,\alpha}$. Note that, in the unweighted case ($\alpha = 0$), it is not difficult to show that $H^{k,s}_p(\mathbb{R}^d) = L^{k+s}_p(\mathbb{R}^d)$, for $1 < p < \infty$, using the Mikhlin-Hörmander theorem on Fourier multipliers (cf. [46, Chapter 5] and [45, Chapter 6]). For $\alpha \ne 0$, however, $H^{k,s}_{p,\alpha}(\mathbb{R}^d)$ is not necessarily the same as $L^{k+s}_{p,\alpha}(\mathbb{R}^d)$. This is due to the lack of a theory on weighted Fourier multipliers for the Sobolev weights; most of the existing literature are concerned with the Muckenhoupt weights, instead [42], [47], [48].

C. Shift-Invariant Spaces of Non-Decaying Functions

We are interested in the approximation of a non-decaying function living in the ambient space $L_{p,-\alpha}(\mathbb{R}^d)$, for some $\alpha \geq 0$, by an element in the (weighted) shift-invariant space $V_{p,-\alpha,h}(\varphi)$ generated by some kernel φ defined as

$$V_{p,-\alpha,h}(\varphi) := \left\{ f = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} c[\boldsymbol{k}] \varphi\left(\frac{\cdot}{h} - \boldsymbol{k}\right) : c \in \ell_{p,-\alpha}(\mathbb{Z}^d) \right\},$$

where $h > 0$ is a varying scale (sampling step). We write $V_{p,-\alpha}(\varphi)$ for $V_{p,-\alpha,1}(\varphi)$, write $V_{p,h}(\varphi)$ for $V_{p,0,h}(\varphi)$, and write $V_p(\varphi)$ for $V_{p,0,1}(\varphi)$. In addition to including many types of signal reconstruction models covered in the literature [6], this general formulation allows us to deal with (polynomially) growing signals. Similar to the unweighted case, we want to make sure that the (unscaled) space $V_{p,-\alpha}(\varphi)$ is a closed subspace of $L_{p,-\alpha}(\mathbb{R}^d)$ and each of its member $f \in V_{p,-\alpha}(\varphi)$ has an unambiguous representation in terms of the coefficients $c[\boldsymbol{k}]$. It turns out that, as shown in [25, Theorem 2], this wish list will be fulfilled if the generating kernel φ satisfies the following admissibility conditions:

(i) $\{\varphi(\cdot - \boldsymbol{k})\}_{\boldsymbol{k} \in \mathbb{Z}^d}$ is a Riesz basis for $V_2(\varphi)$ or, equivalently, the Fourier tranform of the autocorrelation sequence, $\hat{a}_\varphi(\boldsymbol{\omega}) := \sum_{\boldsymbol{k} \in \mathbb{Z}^d} |\hat{\varphi}(\boldsymbol{\omega} + 2\pi\boldsymbol{k})|^2$, is bounded from below and above for all $\boldsymbol{\omega} \in \mathbb{R}^d$;
(ii) φ belongs to the weighted hybrid-norm space $W_{q,\alpha}(\mathbb{R}^d)$ with $q := \max(p, p')$.

We want to emphasize that the above conditions, though mathematically cumbersome, are by no means restrictive since they are easily satisfied by all interpolation kernels used in practice, and in particular B-splines [6].

D. Strang-Fix Conditions

There are multiple forms of the Strang-Fix conditions; the equivalence between them was initially shown for compactly supported functions [7] but then extended to kernels with global supports [9], [10]. The most common form of the Strang-Fix conditions is characterized in the frequency domain: a kernel φ is said to satisfy the Strang-Fix conditions of order L if

(i) $\hat{\varphi}(\boldsymbol{0}) \neq 0$;
(ii) $\partial^{\boldsymbol{\ell}} \hat{\varphi}(2\pi\boldsymbol{k}) = 0, \quad \forall |\boldsymbol{\ell}| \leq L - 1, \forall \boldsymbol{k} \in \mathbb{Z}^d \setminus \{\boldsymbol{0}\}$.

To translate the Strang-Fix conditions into the space domain, we also need the Poisson summation formula (PSF) to hold. One of the most common conditions for the PSF [46] is that

$$\exists C, \varepsilon > 0 : |\varphi(\boldsymbol{x})| + |\hat{\varphi}(\boldsymbol{x})| \leq C(1 + \|\boldsymbol{x}\|)^{-d-\varepsilon}, \forall \boldsymbol{x} \in \mathbb{R}^d. \tag{5}$$

Although (5) is not satisfied by the rectangle kernel, it can be relaxed [10, Theorem 2.1] to cover that case. Assuming the PSF, the Strang-Fix conditions above are equivalent to the existence of a *quasi-interpolant* φ_{QI} of order L [49]–[51] in the shift-invariant subspace $V_2(\varphi)$. This quasi-interpolant exactly interpolates all polynomials of degree (strictly) less than L, i.e.

$$\sum_{\boldsymbol{k} \in \mathbb{Z}^d} \boldsymbol{k}^{\boldsymbol{\ell}} \varphi_{\mathrm{QI}}(\boldsymbol{x} - \boldsymbol{k}) = \boldsymbol{x}^{\boldsymbol{\ell}}, \quad \forall |\boldsymbol{\ell}| \leq L - 1, \forall \boldsymbol{x} \in \mathbb{R}^d, \tag{6}$$

where $\boldsymbol{x}^{\boldsymbol{\ell}}$ stands for $x_1^{\ell_1} \cdots x_d^{\ell_d}$. Therefore, the Strang-Fix conditions of order L can also be described as the ability of the space $V_2(\varphi)$ to reproduce polynomials of degree less than L. It is important to note that, for a particular φ, there are multiple choices for the quasi-interpolant within the subspace $V_2(\varphi)$, one of which is the interpolant φ_{int} that satisfies not only (6) but also the interpolating property

$$\varphi_{\mathrm{int}}(\boldsymbol{k}) = \delta[\boldsymbol{k}], \quad \forall \boldsymbol{k} \in \mathbb{Z}^d, \tag{7}$$

where $\delta[\cdot]$ denote the discrete unit impulse; the construction of this interpolant will be discussed in Section IV.

Most importantly, the Strang-Fix conditions of order L are necessary and sufficient for the controlled L_2-approximation of order L that for any $f \in H_2^L(\mathbb{R}^d)$, there exists $\tilde{f} = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} c[\boldsymbol{k}] \varphi(\cdot/h - \boldsymbol{k})$ in $V_2(\varphi)$ such that

(i) $\|c\|_{\ell_2(\mathbb{Z}^d)} \leq C \cdot \|f\|_{L_2(\mathbb{R}^d)}$ and
(ii) $\left\|f - \tilde{f}\right\|_{L_2(\mathbb{R}^d)} \leq C \cdot h^L \cdot \left\|f^{(L)}\right\|_{L_2(\mathbb{R}^d)}$, as $h \to 0$,

where the constants C are independent of f. Note that the controllability of the approximation is dictated by the first bound, whereas the order of the approximation is described by the second bound. This beautiful connection between the approximation of order L and the ability of the representation space to reproduce polynomials of degree less than L lies at the core of the Strang-Fix theory and its various extensions [10], [12], [15]. Finally, it is handy to keep in mind that the B-spline of order L [52], [53] satisfies the Strang-Fix conditions of order L.

III. Projection Error Bound

In this section, we derive the error bound for the approximation of a non-decaying function in the weighted Sobolev space $H^L_{p,-\alpha}(\mathbb{R}^d)$ by its projection onto the shift-invariant space $V_{p,-\alpha,h}(\varphi)$. Assume throughout this section that the kernel φ is such that $\{\varphi(\cdot-\boldsymbol{k})\}_{\boldsymbol{k}\in\mathbb{Z}^d}$ is a Riesz basis for $V_2(\varphi)$. This condition guarantees [6] that the dual kernel φ_{d} exists and is given in the Fourier domain by

$$\widehat{\varphi_{\mathrm{d}}}(\boldsymbol{\omega}) = \frac{\hat{\varphi}(\boldsymbol{\omega})}{\sum_{\boldsymbol{k}\in\mathbb{Z}^d}|\hat{\varphi}(\boldsymbol{\omega}+2\pi\boldsymbol{k})|^2}.$$

Let us define the operator

$$P_{\varphi,h} : f \mapsto \tilde{f}_{\mathrm{proj}} = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} c[\boldsymbol{k}]\varphi\left(\frac{\cdot}{h}-\boldsymbol{k}\right),$$

where, for each $\boldsymbol{k}\in\mathbb{Z}^d$, the coefficient $c[\boldsymbol{k}]$ is given by

$$c[\boldsymbol{k}] = \frac{1}{h^d}\int_{\mathbb{R}^d} f(\boldsymbol{y})\varphi_{\mathrm{d}}\left(\frac{\boldsymbol{y}}{h}-\boldsymbol{k}\right)\mathrm{d}\boldsymbol{y}.$$

In the language of signal processing, $c[\boldsymbol{k}]$ is the result of prefiltering the signal f with the filter $h^{-d}\varphi_{\mathrm{d}}\left(-\frac{\cdot}{h}\right)$ followed by a sampling at location $h\boldsymbol{k}$. We write P_φ for $P_{\varphi,1}$. It is well known in the (unweighted) L_2 case that $P_{\varphi,h}$ is an orthogonal projector from $L_2(\mathbb{R}^d)$ onto the subspace $V_{2,h}(\varphi)$ and therefore provides the best L_2-approximation. In the weighted-L_p setup, orthogonality no longer exists but the operator $P_{\varphi,h}$ still behaves properly. In particular, the following result shows that $P_{\varphi,h}$ is a bounded projector from $L_{p,-\alpha}(\mathbb{R}^d)$ onto $V_{p,-\alpha,h}(\varphi)$ whose norm is bounded as the scale h tends to 0. The essential condition for that to hold true is that the generating kernel φ is a member of an appropriate weighted hybrid-norm space.

Theorem 1. *Let $1\le p\le\infty$ and $\alpha\ge 0$. If $\varphi\in W_{q,\alpha}(\mathbb{R}^d)$ with $q:=\max(p,p')$ and $\{\varphi(\cdot-\boldsymbol{k})\}_{\boldsymbol{k}\in\mathbb{Z}^d}$ is a Riesz basis for $V_2(\varphi)$, then, for all $h>0$, $V_{p,-\alpha,h}(\varphi)$ is a closed subspace of $L_{p,-\alpha}(\mathbb{R}^d)$ and $P_{\varphi,h}$ is a projector from $L_{p,-\alpha}(\mathbb{R}^d)$ onto $V_{p,-\alpha,h}(\varphi)$. Furthermore, there exists a constant $C_{\varphi,\alpha}$ such that*

$$\|P_{\varphi,h}f\|_{L_{p,-\alpha}(\mathbb{R}^d)} \le C_{\varphi,\alpha}\|f\|_{L_{p,-\alpha}(\mathbb{R}^d)}, \quad \forall f\in L_{p,-\alpha}(\mathbb{R}^d), \forall h\in(0,1). \tag{8}$$

Proof. Since $\varphi\in W_{q,\alpha}(\mathbb{R}^d)$ and $\{\varphi(\cdot-\boldsymbol{k})\}_{\boldsymbol{k}\in\mathbb{Z}^d}$ is a Riesz basis for $V_2(\varphi)$, it is known from [25, Theorems 1 & 2] that $V_{p,-\alpha}(\varphi)$ is a closed subspace of $L_{p,-\alpha}(\mathbb{R}^d)$ and P_φ is a bounded projector from $L_{p,-\alpha}(\mathbb{R}^d)$ onto $V_{p,-\alpha}(\varphi)$. We now divide the rest of the proof into several steps.

First, we show that $V_{p,-\alpha,h}(\varphi)$ is a subspace of $L_{p,-\alpha}(\mathbb{R}^d)$, for all $h>0$. Given $f\in V_{p,-\alpha,h}(\varphi)$, it is clear that $\sigma_{1/h}f\in V_{p,-\alpha}(\varphi)\subset L_{p,-\alpha}(\mathbb{R}^d)$. On the other hand,

$$\begin{aligned}\|f\|^p_{L_{p,-\alpha}(\mathbb{R}^d)} &= h^d\int_{\mathbb{R}^d}\langle h\boldsymbol{x}\rangle^{-\alpha p}\left|(\sigma_{1/h}f)(\boldsymbol{x})\right|^p\mathrm{d}\boldsymbol{x}\\ &\le h^d\cdot\max(1,h^{-\alpha p})\int_{\mathbb{R}^d}\langle\boldsymbol{x}\rangle^{-\alpha p}\left|(\sigma_{1/h}f)(\boldsymbol{x})\right|^p\mathrm{d}\boldsymbol{x}\\ &= h^d\cdot\max(1,h^{-\alpha p})\cdot\left\|\sigma_{1/h}f\right\|^p_{L_{p,-\alpha}(\mathbb{R}^d)}.\end{aligned} \tag{9}$$

This implies that f also belongs to $L_{p,-\alpha}(\mathbb{R}^d)$, or $V_{p,-\alpha,h}(\varphi)$ is a subspace of $L_{p,-\alpha}(\mathbb{R}^d)$, for all $h>0$.

Second, we show that $V_{p,-\alpha,h}(\varphi)$ is closed under the norm of $L_{p,-\alpha}(\mathbb{R}^d)$, for all $h>0$. Let $\{f_n\}$ be a sequence in $V_{p,-\alpha,h}(\varphi)$ such that $f_n\to f$ in $L_{p,-\alpha}(\mathbb{R}^d)$ as $n\to\infty$. Similar to (9), we have that

$$\left\|\sigma_{1/h}f_n-\sigma_{1/h}f\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \le h^{-d/p}\cdot\max(1,h^\alpha)\cdot\|f_n-f\|_{L_{p,-\alpha}(\mathbb{R}^d)},$$

which implies that $\sigma_{1/h} f_n \to \sigma_{1/h} f$ in $L_{p,-\alpha}(\mathbb{R}^d)$ as $n \to \infty$. As $\{\sigma_{1/h} f_n\}$ is a sequence in $V_{p,-\alpha}(\varphi)$, it follows from the closedness of $V_{p,-\alpha}(\varphi)$ that $\sigma_{1/h} f \in V_{p,-\alpha}(\varphi)$, or $f \in V_{p,-\alpha,h}(\varphi)$. This shows the closedness of $V_{p,-\alpha,h}(\varphi)$.

Third, we show that $P_{\varphi,h}$ is a projector that maps $L_{p,-\alpha}(\mathbb{R}^d)$ to $V_{p,-\alpha,h}(\varphi)$, for all $h > 0$. Observe that $P_{\varphi,h} = \sigma_h P_\varphi \sigma_{1/h}$. From (9), $\sigma_{1/h}$ maps $L_{p,-\alpha}(\mathbb{R}^d)$ to itself. It is also known that P_φ maps $L_{p,-\alpha}(\mathbb{R}^d)$ to $V_{p,-\alpha}(\varphi)$ and σ_h maps $V_{p,-\alpha}(\varphi)$ to $V_{p,-\alpha,h}(\varphi)$. Therefore, $P_{\varphi,h}$ maps $L_{p,-\alpha}(\mathbb{R}^d)$ to $V_{p,-\alpha,h}(\varphi)$. The idempotence of $P_{\varphi,h}$ can be easily verified as

$$P_{\varphi,h}^2 = \sigma_h P_\varphi \sigma_{1/h} \sigma_h P_\varphi \sigma_{1/h} = \sigma_h P_\varphi^2 \sigma_{1/h} = \sigma_h P_\varphi \sigma_{1/h} = P_{\varphi,h},$$

where we have relied on the idempotence of the projector P_φ.

Finally, we show the bound (8). Let us consider the weighting function $w_h(\boldsymbol{x}) := \langle h\boldsymbol{x} \rangle^\alpha$. It is easy to see that w_h satisfies

$$w_h(\boldsymbol{x} + \boldsymbol{y}) \leq C_\alpha \, w_h(\boldsymbol{x}) w_h(\boldsymbol{y}), \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d, \forall h > 0. \tag{10}$$

By a change of variable and from the last bound in the proof of [25, Theorem 1], we have that, for all $h > 0$,

$$\begin{aligned} \|P_{\varphi,h} f\|_{L_{p,-\alpha}(\mathbb{R}^d)} &= \left\|\sigma_h P_\varphi \sigma_{1/h} f\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \\ &= h^{d/p} \cdot \left\|P_\varphi(\sigma_{1/h} f)\right\|_{L_{p,1/w_h}(\mathbb{R}^d)} \\ &\leq h^{d/p} \cdot C_\alpha^2 \cdot \|\varphi\|_{W_{p,w_h}(\mathbb{R}^d)} \, \|\varphi_{\mathrm{d}}\|_{W_{p',w_h}(\mathbb{R}^d)} \left\|\sigma_{1/h} f\right\|_{L_{p,1/w_h}(\mathbb{R}^d)} \\ &= C_\alpha^2 \cdot \|\varphi\|_{W_{p,w_h}(\mathbb{R}^d)} \, \|\varphi_{\mathrm{d}}\|_{W_{p',w_h}(\mathbb{R}^d)} \, \|f\|_{L_{p,-\alpha}(\mathbb{R}^d)}, \end{aligned} \tag{11}$$

where C_α is precisely the constant in (10) that does not depend on h. On the other hand, according to [25, Proposition 6], both φ and φ_{d} are elements of $W_{q,\alpha}(\mathbb{R}^d)$. Since $q = \max(p, p')$, it must be that $\varphi \in W_{p,\alpha}(\mathbb{R}^d)$ and $\varphi_{\mathrm{d}} \in W_{p',\alpha}(\mathbb{R}^d)$. Moreover, the assumption that $h \in (0, 1)$ gives

$$\|\varphi\|_{W_{p,w_h}(\mathbb{R}^d)} \leq \|\varphi\|_{W_{p,\alpha}(\mathbb{R}^d)} < \infty, \tag{12}$$

and

$$\|\varphi_{\mathrm{d}}\|_{W_{p',w_h}(\mathbb{R}^d)} \leq \|\varphi_{\mathrm{d}}\|_{W_{p',\alpha}(\mathbb{R}^d)} < \infty. \tag{13}$$

Putting together (11), (12), and (13) yields the desired bound (8). □

The main result of this section is as follows:

Theorem 2. *Let $1 \leq p \leq \infty$, $L \in \mathbb{N}$, and $\alpha \geq 0$. Assume that $\varphi \in W_{q,L+\alpha}(\mathbb{R}^d)$ with $q := \max(p, p')$ and that $\{\varphi(\cdot - \boldsymbol{k})\}_{\boldsymbol{k} \in \mathbb{Z}^d}$ is a Riesz basis for $V_2(\varphi)$. Assume also that φ satisfies the Strang-Fix conditions of order L. Then, there exists a constant $C_{\varphi,L,\alpha}$ such that, for all $f \in H_{p,-\alpha}^L(\mathbb{R}^d)$,*

$$\|f - P_{\varphi,h} f\|_{L_{p,-\alpha}(\mathbb{R}^d)} \leq C_{\varphi,L,\alpha} \cdot h^L \cdot \left\|f^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)}, \tag{14}$$

when $h \to 0$.

In what follows, we break the proof of Theorem 2 into several small results. Let us begin by defining the smoothing operator J_h as

$$J_h : f \mapsto \int_{\mathbb{R}^d} \left(f - \Delta_{h\boldsymbol{u}}^L f\right)(\cdot) \chi(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u}, \tag{15}$$

with some underlying function $\chi \in \mathcal{C}_{\mathrm{c}}^\infty(\mathbb{R}^d)$ such that $\mathrm{supp}(\chi) \subset [-1, 1]^d$ and $\int_{\mathbb{R}^d} \chi(\boldsymbol{u}) \mathrm{d}\boldsymbol{u} = 1$. This smoothing operator was also exploited in [10], [11].

Expanding $\Delta_{h\boldsymbol{u}}^L f$ as

$$\Delta_{h\boldsymbol{u}}^L f = \sum_{n=0}^{L} (-1)^n \binom{L}{n} f(\cdot - nh\boldsymbol{u}),$$

we obtain

$$f - \Delta_{h\boldsymbol{u}}^L f = \sum_{n=1}^{L} (-1)^{n-1} \binom{L}{n} f(\cdot - nh\boldsymbol{u}).$$

Therefore, J_h can also be expressed as

$$J_h f = \sum_{n=1}^{L} (-1)^{n-1} \binom{L}{n} \int_{\mathbb{R}^d} f(\cdot - nh\boldsymbol{u}) \chi(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u}.$$

This means that J_h is a convolution operator: $J_h f = f * \psi_h$, where

$$\psi_h := \sum_{n=1}^{L} (-1)^{n-1} \binom{L}{n} \frac{1}{(nh)^d} \sigma_{nh} \chi. \tag{16}$$

The following result shows that the weighted norm of the error between a function $f \in H_{p,-\alpha}^L(\mathbb{R}^d)$ and its smoothed version $J_h f$ is $O(h^L)$ as h tends to 0.

Proposition 1. *For $1 \leq p \leq \infty$, $L \in \mathbb{N}$, $\alpha \geq 0$, and J_h being the smoothing operator defined in* (15)*, there exists a constant $C_{L,\alpha}$ such that, for all $f \in H_{p,-\alpha}^L(\mathbb{R}^d)$ and for all $h \in (0,1)$,*

$$\|f - J_h f\|_{L_{p,-\alpha}(\mathbb{R}^d)} \leq C_{L,\alpha} \cdot h^L \cdot \left\| f^{(L)} \right\|_{L_{p,-\alpha}(\mathbb{R}^d)}. \tag{17}$$

Proof. We first need the following two lemmas whose proofs can be found in Section V.

Lemma 1. *Let $L \in \mathbb{N}$ and let β^{L-1} be the (1-D) B-spline of order $(L-1)$ given by the L-fold convolution*

$$\beta^{L-1} := \underbrace{\beta^0 * \beta^0 * \cdots * \beta^0}_{L \text{ times}},$$

where

$$\beta^0(x) := \begin{cases} 1, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}.$$

Then, for all $f \in \mathcal{S}'(\mathbb{R}^d)$, one has

$$\Delta_{\boldsymbol{u}}^L f = \int_{\mathbb{R}} D_{\boldsymbol{u}}^L f(\cdot - t\boldsymbol{u}) \beta^{L-1}(t) \mathrm{d}t. \tag{18}$$

Lemma 2. *Let $L \in \mathbb{N}$ and $\boldsymbol{u} \in \mathbb{R}^d$. If $f \in \mathcal{S}'(\mathbb{R}^d)$ such that its partial derivatives up to order L are locally integrable functions, then*

$$\left| D_{\boldsymbol{u}}^L f(\boldsymbol{x}) \right| \leq \|\boldsymbol{u}\|_\infty^L \cdot f^{(L)}(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in \mathbb{R}^d, \tag{19}$$

where $\|\boldsymbol{u}\|_\infty := \max\{|u_1|, \ldots, |u_d|\}$.

We remark that Lemma 1 is an extension of Peano's theorem [54, page 70] for smooth functions. It is needed to avoid the density argument in the proof of [10, Theorem 3.3] that is unavailable in the weighted case. Let us continue with the proof of Proposition 1. Observe that

$$\begin{aligned}(f - J_h f)(\boldsymbol{x}) &= f(\boldsymbol{x}) \int_{\mathbb{R}^d} \chi(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u} - \int_{\mathbb{R}^d} \left(f - \Delta_{h\boldsymbol{u}}^L f\right)(\boldsymbol{x}) \chi(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u} \\ &= \int_{\mathbb{R}^d} \Delta_{h\boldsymbol{u}}^L(\boldsymbol{x}) \chi(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u}.\end{aligned}$$

From Lemma 1 and by taking into account the fact that $\mathrm{supp}(\chi) \subset [-1,1]^d$ and $\mathrm{supp}(\beta^{L-1}) = [0, L]$, we write

$$(f - J_h f)(\boldsymbol{x}) = \int_{[-1,1]^d} \int_0^L D_{h\boldsymbol{u}}^L f(\boldsymbol{x} - th\boldsymbol{u}) \beta^{L-1}(t) \chi(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u} \, \mathrm{d}t.$$

It then follows from Minkowski's inequality and Lemma 2 that

$$\begin{aligned}\|f - J_h f\|_{L_{p,-\alpha}(\mathbb{R}^d)} &\leq \int_{[-1,1]^d} \int_{\mathbb{R}} \left\|D^L_{h\boldsymbol{u}} f(\cdot - th\boldsymbol{u})\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \beta^{L-1}(t)\chi(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}t \\ &\leq \int_{[-1,1]^d} \int_{\mathbb{R}} \|h\boldsymbol{u}\|_\infty^L \cdot \left\|f^{(L)}(\cdot - th\boldsymbol{u})\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \beta^{L-1}(t)\chi(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}t \\ &\leq h^L \cdot \int_{[-1,1]^d} \int_{\mathbb{R}} \left\|f^{(L)}(\cdot - th\boldsymbol{u})\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \beta^{L-1}(t)\chi(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}t. \end{aligned} \tag{20}$$

On the other hand,

$$\begin{aligned}\left\|f^{(L)}(\cdot - th\boldsymbol{u})\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} &= \left(\int_{\mathbb{R}^d} \left|\langle \boldsymbol{x}\rangle^{-\alpha} f^{(L)}(\boldsymbol{x} - th\boldsymbol{u})\right|^p \mathrm{d}\boldsymbol{x}\right)^{1/p} \\ &\leq C_\alpha \langle th\boldsymbol{u}\rangle^\alpha \left(\int_{\mathbb{R}^d} \left|\langle \boldsymbol{x} - th\boldsymbol{u}\rangle^{-\alpha} f^{(L)}(\boldsymbol{x} - th\boldsymbol{u})\right|^p \mathrm{d}\boldsymbol{x}\right)^{1/p} \\ &= C_\alpha \langle th\boldsymbol{u}\rangle^\alpha \left\|f^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)}. \end{aligned}$$

Thus, for $t \in [0, L]$ and $h \in (0, 1)$,

$$\left\|f^{(L)}(\cdot - th\boldsymbol{u})\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \leq C_\alpha L^\alpha \langle \boldsymbol{u}\rangle^\alpha \left\|f^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)}. \tag{21}$$

Combining (21) with (20) leads to

$$\begin{aligned}\|f - J_h f\|_{L_{p,-\alpha}(\mathbb{R}^d)} &\leq C_\alpha L^\alpha \cdot h^L \cdot \left\|f^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \int_0^L \beta^{L-1}(t)\mathrm{d}t \int_{[-1,1]^d} \langle \boldsymbol{u}\rangle^\alpha \chi(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} \\ &= C_{L,\alpha} \cdot h^L \cdot \left\|f^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)}, \end{aligned}$$

which completes the proof. □

Proposition 2. *Assume that $1 \leq p \leq \infty$, $L \in \mathbb{N}$, and $\alpha \geq 0$. Let $q := \max(p, p')$ and let J_h be the smoothing operator defined in* (15)*. If φ is an element of $W_{q,L+\alpha}(\mathbb{R}^d)$ that satisfies the Strang-Fix conditions of order L, then there exists a constant $C_{\varphi,L,\alpha}$ such that, for all $f \in H^L_{p,-\alpha}(\mathbb{R}^d)$ and for all $h \in (0, 1)$,*

$$\|J_h f - P_{\varphi,h} J_h f\|_{L_{p,-\alpha}(\mathbb{R}^d)} \leq C_{\varphi,L,\alpha} \cdot h^L \cdot \left\|f^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)}.$$

Proof. We begin the proof with a lemma; its proof is given in Section V.

Lemma 3. *Let $w_h(\boldsymbol{x}) := \langle h\boldsymbol{x}\rangle^\alpha$, $\alpha \geq 0$. Then, there exists a constant $C_{L,\alpha}$ such that, for all $f \in L_{p,-\alpha}(\mathbb{R}^d)$ and for all $h \in (0, 1)$,*

$$\left\|(\sigma_{1/h} J_h f)[\cdot]\right\|_{\ell_{p,1/w_h}(\mathbb{Z}^d)} \leq C_{L,\alpha} \cdot h^{-d/p} \cdot \|f\|_{L_{p,-\alpha}(\mathbb{R}^d)}. \tag{22}$$

Let us now put $g := J_h f$ and $e := g - P_{\varphi,h} g$. It is clear that g is infinitely differentiable. For $\boldsymbol{x} \in \mathbb{R}^d$, let $R_{\boldsymbol{x}}$ denote the remainder of the order-$(L-1)$ Taylor series of function g about $\boldsymbol{x}$. Since φ satisfies the Strang-Fix conditions of order L, it is known [14] that $P_{\varphi,h}$ maps every polynomial of degree less than L to itself. Therefore, it is possible to write

$$e(\boldsymbol{x}) = -\sum_{\boldsymbol{\ell} \in \mathbb{Z}^d} c_{\boldsymbol{x}}[\boldsymbol{\ell}] \varphi\left(\frac{\boldsymbol{x}}{h} - \boldsymbol{\ell}\right), \tag{23}$$

where the sequence $c_{\boldsymbol{x}}$ is given by

$$c_{\boldsymbol{x}}[\boldsymbol{\ell}] := \frac{1}{h^d} \int_{\mathbb{R}^d} R_{\boldsymbol{x}}(\boldsymbol{y}) \varphi_{\mathrm{d}}\left(\frac{\boldsymbol{y}}{h} - \boldsymbol{\ell}\right) \mathrm{d}\boldsymbol{y}. \tag{24}$$

The weighted-L_p norm of the projection error is then bounded as

$$\begin{aligned}\|e\|^p_{L_{p,-\alpha}(\mathbb{R}^d)} &= \sum_{\boldsymbol{k}\in\mathbb{Z}^d}\int_{[0,h]^d}\left|\langle \boldsymbol{x}+h\boldsymbol{k}\rangle^{-\alpha}\, e(\boldsymbol{x}+h\boldsymbol{k})\right|^p \mathrm{d}\boldsymbol{x}\\ &= h^d\cdot\int_{[0,1]^d}\sum_{\boldsymbol{k}\in\mathbb{Z}^d}\left|\langle h\boldsymbol{x}+h\boldsymbol{k}\rangle^{-\alpha}\, e(h\boldsymbol{x}+h\boldsymbol{k})\right|^p \mathrm{d}\boldsymbol{x}\\ &= h^d\cdot\int_{[0,1]^d}\sum_{\boldsymbol{k}\in\mathbb{Z}^d}\left|\langle h\boldsymbol{x}+h\boldsymbol{k}\rangle^{-\alpha}\sum_{\boldsymbol{\ell}\in\mathbb{Z}^d} c_{h\boldsymbol{x}+h\boldsymbol{k}}[\boldsymbol{\ell}]\cdot\varphi\left(\boldsymbol{x}+\boldsymbol{k}-\boldsymbol{\ell}\right)\right|^p \mathrm{d}\boldsymbol{x}\\ &\leq C_\alpha\cdot h^d\cdot\int_{[0,1]^d}\sum_{\boldsymbol{k}\in\mathbb{Z}^d}\left(\sum_{\boldsymbol{\ell}\in\mathbb{Z}^d}\langle h\boldsymbol{k}\rangle^{-\alpha}\left|c_{h\boldsymbol{x}+h\boldsymbol{k}}[\boldsymbol{k}-\boldsymbol{\ell}]\right|\cdot\left|\varphi\left(\boldsymbol{x}+\boldsymbol{\ell}\right)\right|\right)^p \mathrm{d}\boldsymbol{x}. \end{aligned} \tag{25}$$

The last estimate is due to a change of variable and to the fact that $\langle h\boldsymbol{x}+h\boldsymbol{k}\rangle^{-\alpha}\leq C_\alpha\langle h\boldsymbol{k}\rangle^{-\alpha}$, $\forall \boldsymbol{x}\in[0,1]^d$, $\forall h\in(0,1)$. Let us define the two sequences: $c_{\boldsymbol{x},\boldsymbol{\ell}}[\boldsymbol{k}]:=\langle h\boldsymbol{k}\rangle^{-\alpha}\left|c_{h\boldsymbol{x}+h\boldsymbol{k}}[\boldsymbol{k}-\boldsymbol{\ell}]\right|$ and $\varphi_{\boldsymbol{x}}[\cdot]=|\varphi(\boldsymbol{x}+\cdot)|$, for each $\boldsymbol{x}\in[0,1]^d$ and each $\boldsymbol{\ell}\in\mathbb{Z}^d$. Plugging these notations into (25) and applying Minkowski's inequality, we obtain

$$\begin{aligned}\|e\|^p_{L_{p,-\alpha}(\mathbb{R}^d)} &\leq C_\alpha\cdot h^d\cdot\int_{[0,1]^d}\sum_{\boldsymbol{k}\in\mathbb{Z}^d}\left(\sum_{\boldsymbol{\ell}\in\mathbb{Z}^d} c_{\boldsymbol{x},\boldsymbol{\ell}}[\boldsymbol{k}]\cdot\varphi_{\boldsymbol{x}}[\boldsymbol{\ell}]\right)^p \mathrm{d}\boldsymbol{x}\\ &\leq C_\alpha\cdot h^d\cdot\int_{[0,1]^d}\left(\sum_{\boldsymbol{\ell}\in\mathbb{Z}^d}\|c_{\boldsymbol{x},\boldsymbol{\ell}}\|_{\ell_p(\mathbb{Z}^d)}\cdot\varphi_{\boldsymbol{x}}[\boldsymbol{\ell}]\right)^p \mathrm{d}\boldsymbol{x}. \end{aligned} \tag{26}$$

We now proceed to bound the quantity $\|c_{\boldsymbol{x},\boldsymbol{\ell}}\|_{\ell_p(\mathbb{Z}^d)}$. By Taylor's theorem

$$\begin{aligned}R_{h\boldsymbol{x}+h\boldsymbol{k}}(h\boldsymbol{y}+h\boldsymbol{k}) &= \int_0^1\frac{(1-\tau)^{L-1}}{(L-1)!}S_{\tau h\boldsymbol{y}+(1-\tau)h\boldsymbol{x}}D^L_{h\boldsymbol{y}-h\boldsymbol{x}}(J_hf)(h\boldsymbol{k})\mathrm{d}\tau\\ &= \int_0^1 J_hT_{\boldsymbol{y},\tau}f(h\boldsymbol{k})\mathrm{d}\tau, \end{aligned} \tag{27}$$

where the operator $T_{\boldsymbol{y},\tau}$ is defined as

$$T_{\boldsymbol{y},\tau}:=\frac{(1-\tau)^{L-1}}{(L-1)!}S_{\tau h\boldsymbol{y}+(1-\tau)h\boldsymbol{x}}D^L_{h\boldsymbol{y}-h\boldsymbol{x}}. \tag{28}$$

Note that the swapping of $T_{\boldsymbol{y},\tau}$ and J_h in (27) is justified because J_h is a convolution operator and hence commutes with differential and shift operators. From (24) and the definition of $c_{\boldsymbol{x},\boldsymbol{\ell}}$, one has

$$\begin{aligned}c_{\boldsymbol{x},\boldsymbol{\ell}}[\boldsymbol{k}] &= \langle h\boldsymbol{k}\rangle^{-\alpha}\int_{\mathbb{R}^d}R_{h\boldsymbol{x}+h\boldsymbol{k}}(h\boldsymbol{y}+h\boldsymbol{k})\varphi_{\mathrm{d}}\left(\boldsymbol{y}+\boldsymbol{\ell}\right)\mathrm{d}\boldsymbol{y}\\ &= \int_{\mathbb{R}^d}\varphi_{\mathrm{d}}\left(\boldsymbol{y}+\boldsymbol{\ell}\right)\int_0^1\frac{1}{w_h(\boldsymbol{k})}\cdot J_hT_{\boldsymbol{y},\tau}f(h\boldsymbol{k})\mathrm{d}\tau\mathrm{d}\boldsymbol{y}, \end{aligned} \tag{29}$$

where $w_h:=\langle h\cdot\rangle^\alpha$. By Minkowski's inequality and by Lemma 3

$$\begin{aligned}\|c_{\boldsymbol{x},\boldsymbol{\ell}}\|_{\ell_p(\mathbb{Z}^d)} &\leq \int_{\mathbb{R}^d}\left|\varphi_{\mathrm{d}}\left(\boldsymbol{y}+\boldsymbol{\ell}\right)\right|\int_0^1\left\|(\sigma_{1/h}J_hT_{\boldsymbol{y},\tau}f)[\cdot]\right\|_{\ell_{p,1/w_h}(\mathbb{Z}^d)}\mathrm{d}\tau\mathrm{d}\boldsymbol{y}\\ &\leq C_{L,\alpha}\cdot h^{-d/p}\int_{\mathbb{R}^d}\left|\varphi_{\mathrm{d}}\left(\boldsymbol{y}+\boldsymbol{\ell}\right)\right|\int_0^1\|T_{\boldsymbol{y},\tau}f\|_{L_{p,-\alpha}(\mathbb{R}^d)}\mathrm{d}\tau\mathrm{d}\boldsymbol{y}. \end{aligned} \tag{30}$$

On the other hand

$$\begin{aligned}
\|T_{\boldsymbol{y},\tau} f\|_{L_{p,-\alpha}(\mathbb{R}^d)} &= \frac{(1-\tau)^{L-1}}{(L-1)!} \cdot \left\|S_{\tau h\boldsymbol{y}+(1-\tau)h\boldsymbol{x}} D^L_{h\boldsymbol{y}-h\boldsymbol{x}} f\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \\
&\leq C_L \cdot \left\|f^{(L)}(\cdot - \tau h\boldsymbol{y} - (1-\tau)h\boldsymbol{x})\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \|h\boldsymbol{y} - h\boldsymbol{x}\|^L \qquad (31) \\
&\leq C_{L,\alpha} \cdot \left\|f^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \cdot \langle \tau h\boldsymbol{y} + (1-\tau)h\boldsymbol{x}\rangle^{\alpha} \|h\boldsymbol{y} - h\boldsymbol{x}\|^L \qquad (32) \\
&\leq C_{L,\alpha} \cdot h^L \cdot \left\|f^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \cdot \langle \boldsymbol{y} - \boldsymbol{x}\rangle^{\alpha} \|\boldsymbol{y} - \boldsymbol{x}\|^L \qquad (33) \\
&\leq C_{L,\alpha} \cdot h^L \cdot \left\|f^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \cdot \langle \boldsymbol{y} - \boldsymbol{x}\rangle^{L+\alpha}, \qquad (34)
\end{aligned}$$

where (31) follows from Lemma 2; (32) is due to the submultiplicativity of the weight $\langle\cdot\rangle^{\alpha}$; and (33) is because $h, \tau \in (0,1)$ and $\boldsymbol{x} \in [0,1]^d$. Putting (30) and (34) together

$$\begin{aligned}
\|c_{\boldsymbol{x},\boldsymbol{\ell}}\|_{\ell_p(\mathbb{Z}^d)} &\leq C_{L,\alpha} \cdot h^{-d/p} \cdot h^L \left\|f^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \int_{\mathbb{R}^d} \langle \boldsymbol{y} - \boldsymbol{x}\rangle^{L+\alpha} |\varphi_{\mathrm{d}}(\boldsymbol{y} + \boldsymbol{\ell})| \mathrm{d}\boldsymbol{y} \\
&= C_{L,\alpha} \cdot h^{-d/p} \cdot h^L \left\|f^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \int_{\mathbb{R}^d} \langle \boldsymbol{y} - \boldsymbol{\ell} - \boldsymbol{x}\rangle^{L+\alpha} |\varphi_{\mathrm{d}}(\boldsymbol{y})| \mathrm{d}\boldsymbol{y} \\
&\leq C_{L,\alpha} \cdot h^{-d/p} \cdot h^L \left\|f^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \langle \boldsymbol{x} + \boldsymbol{\ell}\rangle^{L+\alpha} \|\varphi_{\mathrm{d}}\|_{L_{1,L+\alpha}(\mathbb{R}^d)}. \qquad (35)
\end{aligned}$$

The last estimate is again due to the submultiplicativity of the weight $\langle\cdot\rangle^{\alpha}$.

Since $\varphi \in W_{q,L+\alpha}(\mathbb{R}^d)$, it follows from [25, Proposition 6] that φ_{d} also belongs to $W_{q,L+\alpha}(\mathbb{R}^d)$. Since $W_{q,L+\alpha}(\mathbb{R}^d) \subset W_{1,L+\alpha}(\mathbb{R}^d) = L_{1,L+\alpha}(\mathbb{R}^d)$, it must be that $\varphi_{\mathrm{d}} \in L_{1,L+\alpha}(\mathbb{R}^d)$ and so the right-hand side of (35) is finite. Plugging (35) into (26) yields

$$\begin{aligned}
\|e\|_{L_{p,-\alpha}(\mathbb{R}^d)} &\leq C_{L,\alpha} \cdot h^L \cdot \left\|f^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \|\varphi_{\mathrm{d}}\|_{L_{1,L+\alpha}(\mathbb{R}^d)} \left(\int_{[0,1]^d} \left(\sum_{\boldsymbol{\ell} \in \mathbb{Z}^d} \langle \boldsymbol{x} + \boldsymbol{\ell}\rangle^{L+\alpha} |\varphi(\boldsymbol{x} + \boldsymbol{\ell})| \right)^p \mathrm{d}\boldsymbol{x} \right)^{1/p} \\
&= \underbrace{C_{L,\alpha} \cdot \|\varphi_{\mathrm{d}}\|_{L_{1,L+\alpha}(\mathbb{R}^d)} \cdot \|\varphi\|_{W_{p,L+\alpha}(\mathbb{R}^d)}}_{C_{\varphi,L,\alpha}} \cdot h^L \cdot \left\|f^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)},
\end{aligned}$$

which is the desired bound. □

With the above results in hands, we are now ready to prove Theorem 2.

Proof of Theorem 2. Without loss of generality, assume that $h \in (0,1)$. Put $g := J_h f$. By using the triangle inequality and by applying Theorem 1, we have that

$$\begin{aligned}
\|f - P_{\varphi,h} f\|_{L_{p,-\alpha}(\mathbb{R}^d)} &\leq \|f - g\|_{L_{p,-\alpha}(\mathbb{R}^d)} + \|P_{\varphi,h} f - P_{\varphi,h} g\|_{L_{p,-\alpha}(\mathbb{R}^d)} + \|g - P_{\varphi,h} g\|_{L_{p,-\alpha}(\mathbb{R}^d)} \\
&\leq (1 + C_{\varphi,\alpha}) \|f - g\|_{L_{p,-\alpha}(\mathbb{R}^d)} + \|g - P_{\varphi,h} g\|_{L_{p,-\alpha}(\mathbb{R}^d)}.
\end{aligned}$$

This bound together with Propositions 1 and 2 immediately implies (14), completing the proof. □

IV. Interpolation Error Bound

We consider in this section the approximation scheme in which a function is ideally sampled (without a prefilter) and reconstructed using an interpolating kernel. Consider throughout this section a kernel φ that satisfies condition (5). The interpolation operator associated with kernel φ and sampling step h is defined by

$$I_{\varphi,h} : f \mapsto \tilde{f}_{\mathrm{int}} = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} f(h\boldsymbol{k}) \varphi_{\mathrm{int}} \left(\frac{\cdot}{h} - \boldsymbol{k} \right), \qquad (36)$$

where the interpolant φ_{int} is related to the kernel φ by

$$\varphi_{\mathrm{int}} := \sum_{\boldsymbol{k} \in \mathbb{Z}^d} a[\boldsymbol{k}] \varphi(\cdot - \boldsymbol{k}), \qquad (37)$$

and where a is a discrete filter given by

$$a[\boldsymbol{n}] := \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \frac{\mathrm{e}^{\mathrm{j}\langle \boldsymbol{\omega}, \boldsymbol{n}\rangle}}{\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \varphi(\boldsymbol{k}) \mathrm{e}^{-\mathrm{j}\langle \boldsymbol{\omega}, \boldsymbol{k}\rangle}} \mathrm{d}\boldsymbol{\omega}, \quad \text{for } \boldsymbol{n} \in \mathbb{Z}^d. \tag{38}$$

This filter is to make sure that $f(h\boldsymbol{k}) = \tilde{f}_{\mathrm{int}}(h\boldsymbol{k})$, for all $\boldsymbol{k} \in \mathbb{Z}^d$. We have assumed implicitly in (38) that $\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \varphi(\boldsymbol{k}) \mathrm{e}^{-\mathrm{j}\langle \boldsymbol{\omega}, \boldsymbol{k}\rangle}$ is nonzero for all $\boldsymbol{\omega} \in \mathbb{R}^d$. It is noteworthy that, in the absence of a prefilter, the function f to be approximated has to be continuous everywhere for the sampling to make sense.

Another way to express (36) is

$$I_{\varphi,h} : f \mapsto \tilde{f}_{\mathrm{int}} = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} c[\boldsymbol{k}] \varphi\left(\frac{\cdot}{h} - \boldsymbol{k}\right), \tag{39}$$

where $c := (\sigma_{1/h} f)[\cdot] * a$ is the sampled sequence of f discretely filtered by a. This is the way interpolation is often implemented in practice since it is generally much easier to work with the kernel φ than with φ_{int}. To simplify the notation, we write I_φ for $I_{\varphi,1}$.

The following lemma says that the interpolant φ_{int} and the kernel φ can be made to lie in the same weighted hybrid-norm space by imposing on φ some mild conditions that are satisfied by, for example, B-splines of all orders.

Lemma 4. *Let $1 \le p \le \infty$ and $\alpha \ge 0$. Let $\varphi \in W_{p,\alpha}(\mathbb{R}^d)$ such that $\varphi[\cdot] \in \ell_{1,\alpha}(\mathbb{Z}^d)$ and $\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \varphi[\boldsymbol{k}] \mathrm{e}^{-\mathrm{j}\langle \boldsymbol{\omega}, \boldsymbol{k}\rangle}$ is nonzero for all $\boldsymbol{\omega} \in \mathbb{R}^d$. Then, the corresponding interpolant φ_{int} defined in* (37) *also belongs to $W_{p,\alpha}(\mathbb{R}^d)$.*

Proof. See Section V. □

The next result is the interpolation counterpart of Theorem 1 and can be thought of as the scaled version of [25, Proposition 9]. It asserts that $I_{\varphi,h}$ is a bounded operator from $L^{d/p+\varepsilon}_{p,-\alpha}(\mathbb{R}^d)$ to $V_{p,-\alpha,h}(\varphi)$ whose norm is bounded as $h \to 0$. The underlying condition is that the interpolant φ_{int} belongs to the weighted hybrid-norm space $W_{p,\alpha}(\mathbb{R}^d)$.

Theorem 3. *Assume that $1 \le p \le \infty$, $\alpha \ge 0$, and $r > d/p$. Let $\varphi \in W_{p,\alpha}(\mathbb{R}^d)$ such that $\varphi[\cdot] \in \ell_{1,\alpha}(\mathbb{Z}^d)$ and $\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \varphi[\boldsymbol{k}] \mathrm{e}^{-\mathrm{j}\langle \boldsymbol{\omega}, \boldsymbol{k}\rangle}$ is nonzero for all $\boldsymbol{\omega} \in \mathbb{R}^d$. Then, there exists a constant $C_{\varphi,r,\alpha}$ such that, for all continuous functions $f \in L^r_{p,-\alpha}(\mathbb{R}^d)$ and for all $h \in (0,1)$,*

$$\|I_{\varphi,h} f\|_{L_{p,-\alpha}(\mathbb{R}^d)} \le C_{\varphi,r,\alpha} \cdot \|f\|_{L^r_{p,-\alpha}(\mathbb{R}^d)}. \tag{40}$$

Proof. Let $B_r := \mathcal{F}^{-1}\{\langle\cdot\rangle^{-r}\}$ be the kernel associated with the Bessel potential of order r. Recall from [45, Proposition 6.1.5] that $B_r(\boldsymbol{x}) > 0$, for all $\boldsymbol{x} \in \mathbb{R}^d$, and that

$$B_r(\boldsymbol{x}) \le C_r \, \mathrm{e}^{-\frac{\|\boldsymbol{x}\|}{2}}, \quad \forall \|\boldsymbol{x}\| \ge 2. \tag{41}$$

Moreover, since $r > d/p$, it is also known [25, Proposition 7] that $B_r \in L_{p',\alpha}(\mathbb{R}^d)$.

Let us now define the weight $w_h(\boldsymbol{x}) := \langle h\boldsymbol{x}\rangle^\alpha$. Recall that w_h is submultiplicative with the same constant C_α for all $h > 0$. Observe from (36) that $I_{\varphi,h} = \sigma_h I_\varphi \sigma_{1/h}$. Therefore, by a change of variable, we have

$$\|I_{\varphi,h} f\|_{L_{p,-\alpha}(\mathbb{R}^d)} = \left\|\sigma_h I_\varphi \sigma_{1/h} f\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} = h^{d/p} \cdot \left\|I_\varphi\left(\sigma_{1/h} f\right)\right\|_{L_{p,1/w_h}(\mathbb{R}^d)}. \tag{42}$$

We now invoke [25, Proposition 4] to get

$$\left\|I_\varphi\left(\sigma_{1/h} f\right)\right\|_{L_{p,1/w_h}(\mathbb{R}^d)} \le C_\alpha \cdot \|\varphi_{\mathrm{int}}\|_{W_{p,w_h}(\mathbb{R}^d)} \cdot \left\|\left(\sigma_{1/h} f\right)[\cdot]\right\|_{\ell_{p,1/w_h}(\mathbb{Z}^d)}. \tag{43}$$

Note that, for all $h \in (0,1)$, $w_h(\boldsymbol{x}) \le \langle \boldsymbol{x}\rangle^\alpha$, and so, the quantity $\|\varphi_{\mathrm{int}}\|_{W_{p,w_h}(\mathbb{R}^d)}$ is bounded since

$$\|\varphi_{\mathrm{int}}\|_{W_{p,w_h}(\mathbb{R}^d)} \le \|\varphi_{\mathrm{int}}\|_{W_{p,\alpha}(\mathbb{R}^d)}, \tag{44}$$

which is finite due to Lemma 4. On the other hand, since $f = B_r * D^r f$, we can write

$$\sigma_{1/h} f = h^d \cdot \left(\sigma_{1/h} B_r\right) * \left(\sigma_{1/h} D^r f\right),$$

and apply [25, Proposition 5] to obtain

$$\begin{aligned}\left\|\left(\sigma_{1/h} f\right)[\cdot]\right\|_{\ell_{p,1/w_h}(\mathbb{Z}^d)} &\leq C_\alpha \cdot h^d \left\|\sigma_{1/h} B_r\right\|_{W_{p',w_h}(\mathbb{R}^d)} \cdot \left\|\sigma_{1/h} D^r f\right\|_{L_{p,1/w_h}(\mathbb{R}^d)} \\ &= C_\alpha \cdot h^d \left\|\sigma_{1/h} B_r\right\|_{W_{p',w_h}(\mathbb{R}^d)} \cdot h^{-d/p} \left\|f\right\|_{L^r_{p,-\alpha}(\mathbb{R}^d)}, \end{aligned} \tag{45}$$

where (45) is due to a change of variable and the definition of the Sobolev norm $\|\cdot\|_{L^r_{p,-\alpha}(\mathbb{R}^d)}$. Combining (42), (43), (44), and (45), we arrive at

$$\left\|I_{\varphi,h} f\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \leq C_\alpha^2 \cdot \left\|\varphi_{\text{int}}\right\|_{W_{p,\alpha}(\mathbb{R}^d)} \cdot h^d \left\|\sigma_{1/h} B_r\right\|_{W_{p',w_h}(\mathbb{R}^d)} \cdot \left\|f\right\|_{L^r_{p,-\alpha}(\mathbb{R}^d)}. \tag{46}$$

Hence, the desired bound (40) will be achieved if

$$\left\|\sigma_{1/h} B_r\right\|_{W_{p',w_h}(\mathbb{R}^d)} \leq C_{r,\alpha} \cdot h^{-d}, \quad \forall h \in (0,1), \tag{47}$$

for some constant $C_{r,\alpha}$. In the rest of the proof, we will show that this claim is true. Let us put $\mathbb{T} := [0,1]^d$, $\mathbb{T}_h := [0,h]^d$, and $B_{r,\alpha} := \langle\cdot\rangle^\alpha B_r$. From the positivity of B_r, it is clear that $B_{r,\alpha}(\boldsymbol{x}) > 0$, $\forall \boldsymbol{x} \in \mathbb{R}^d$. By the definition of the mixed norm, we express

$$\begin{aligned}\left\|\sigma_{1/h} B_r\right\|_{W_{p',w_h}(\mathbb{R}^d)} = \left\|\sigma_{1/h} B_{r,\alpha}\right\|_{W_{p'}(\mathbb{R}^d)} &= \left\| \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \left(\sigma_{1/h} B_{r,\alpha}\right)(\cdot + \boldsymbol{k}) \right\|_{L_{p'}(\mathbb{T})} \\ &= h^{-d/p'} \left\| \sum_{\boldsymbol{k}\in\mathbb{Z}^d} B_{r,\alpha}(\cdot + h\boldsymbol{k}) \right\|_{L_{p'}(\mathbb{T}_h)}. \end{aligned}$$

Applying Minkowski's inequality, we get

$$\begin{aligned}\left\|\sigma_{1/h} B_r\right\|_{W_{p',w_h}(\mathbb{R}^d)} &\leq h^{-d/p'} \sum_{\boldsymbol{k}\in\mathcal{S}_h} \left\|B_{r,\alpha}(\cdot + h\boldsymbol{k})\right\|_{L_{p'}(\mathbb{T}_h)} + h^{-d/p'} \left\| \sum_{\boldsymbol{k}\in\mathbb{Z}^d\setminus\mathcal{S}_h} B_{r,\alpha}(\cdot + h\boldsymbol{k}) \right\|_{L_{p'}(\mathbb{T}_h)} \\ &=: A + B, \end{aligned} \tag{48}$$

where $\mathcal{S}_h$ is a subset of $\mathbb{Z}^d$ defined by

$$\mathcal{S}_h := \left\{ \boldsymbol{k} \in \mathbb{Z}^d : \|\boldsymbol{k}\| \leq \frac{\sqrt{d}+2}{h} \right\}.$$

We complete the proof by showing that both terms A and B in (48) are bounded by $C_{r,\alpha}\, h^{-d}$. It is clear that $|\mathcal{S}_h| = Ch^{-d}$, for some constant C. Therefore, by Hölder's inequality

$$\begin{aligned} A &\leq h^{-d/p'} \cdot \left(\sum_{\boldsymbol{k}\in\mathcal{S}_h} 1^p \right)^{1/p} \cdot \left(\sum_{\boldsymbol{k}\in\mathcal{S}_h} \left\|B_{r,\alpha}(\cdot + h\boldsymbol{k})\right\|^{p'}_{L_{p'}(\mathbb{T}_h)} \right)^{1/p'} \\ &= h^{-d/p'} \cdot |\mathcal{S}_h|^{1/p} \cdot \left(\sum_{\boldsymbol{k}\in\mathcal{S}_h} \int_{\mathbb{T}_h} |B_{r,\alpha}(\boldsymbol{x} + h\boldsymbol{k})|^{p'} \,\mathrm{d}\boldsymbol{x} \right)^{1/p'} \\ &\leq h^{-d/p'} \cdot C \cdot h^{-d/p} \cdot \left(\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \int_{\mathbb{T}_h} |B_{r,\alpha}(\boldsymbol{x} + h\boldsymbol{k})|^{p'} \,\mathrm{d}\boldsymbol{x} \right)^{1/p'} \\ &\leq C \cdot h^{-d} \cdot \left(\int_{\mathbb{R}^d} |B_{r,\alpha}(\boldsymbol{x})|^{p'} \,\mathrm{d}\boldsymbol{x} \right)^{1/p'} \\ &\leq \underbrace{C \cdot \|B_r\|_{L_{p',\alpha}(\mathbb{R}^d)}}_{C_{r,\alpha}} \cdot h^{-d}. \end{aligned} \tag{49}$$

The constant $C_{r,\alpha}$ in (49) is finite because $B_r \in L_{p',\alpha}(\mathbb{R}^d)$. We now proceed to bound the term B in (48). As $h \in (0,1)$, we have that, for all $\boldsymbol{x} \in \mathbb{T}$ and for all $\boldsymbol{k} \notin \mathcal{S}_h$,

$$\|h\boldsymbol{x} + h\boldsymbol{k}\| \geq h\,\|\boldsymbol{k}\| - h\,\|\boldsymbol{x}\| > (\sqrt{d}+2) - h\sqrt{d} > 2,$$

which, according to (41), implies that

$$B_r(h\boldsymbol{x}+h\boldsymbol{k}) \leq C_r\,\mathrm{e}^{-\|h\boldsymbol{x}+h\boldsymbol{k}\|/2} \leq C_r\,\mathrm{e}^{\frac{\|h\boldsymbol{x}\|-\|h\boldsymbol{k}\|}{2}}.$$

Plugging this bound into the formula of B and using the submultiplicativity of the weight $\langle\cdot\rangle^\alpha$ and the fact that $h \in (0,1)$, we get

$$\begin{aligned}
B &= \left(\int_{\mathbb{T}} \left(\sum_{\boldsymbol{k}\in\mathbb{Z}^d\setminus\mathcal{S}_h} \langle h\boldsymbol{x}+h\boldsymbol{k}\rangle^\alpha\, B_r(h\boldsymbol{x}+h\boldsymbol{k})\right)^{p'} \mathrm{d}\boldsymbol{x}\right)^{1/p'} \\
&\leq C_{r,\alpha}\cdot\left(\int_{\mathbb{T}} \left(\sum_{\boldsymbol{k}\in\mathbb{Z}^d\setminus\mathcal{S}_h} \langle\boldsymbol{x}\rangle^\alpha\,\langle h\boldsymbol{k}\rangle^\alpha\,\mathrm{e}^{\frac{\|\boldsymbol{x}\|-\|h\boldsymbol{k}\|}{2}}\right)^{p'} \mathrm{d}\boldsymbol{x}\right)^{1/p'} \\
&\leq C_{r,\alpha}\cdot\left(\int_{\mathbb{T}} \langle\boldsymbol{x}\rangle^{p'\alpha}\,\mathrm{e}^{\frac{p'\|\boldsymbol{x}\|}{2}}\mathrm{d}\boldsymbol{x}\right)^{1/p'} \cdot \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \langle h\boldsymbol{k}\rangle^\alpha\,\mathrm{e}^{-\frac{\|h\boldsymbol{k}\|}{2}}\cdot
\end{aligned} \tag{50}$$

Since the integral in (50) is a constant independent of h, we only need to show that the sum is bounded by $C_\alpha\,h^{-d}$. Again, by the submultiplicativity of the weight $\langle\cdot\rangle^\alpha$ and by the assumption that $h \in (0,1)$, we have

$$\begin{aligned}
\int_{\mathbb{R}^d} \langle\boldsymbol{x}\rangle^\alpha\,\mathrm{e}^{-\frac{\|\boldsymbol{x}\|}{2}}\,\mathrm{d}\boldsymbol{x} &= \sum_{\boldsymbol{k}\in\mathbb{Z}^d}\int_{\mathbb{T}_h} \langle\boldsymbol{x}+h\boldsymbol{k}\rangle^\alpha\,\mathrm{e}^{-\frac{\|\boldsymbol{x}+h\boldsymbol{k}\|}{2}}\mathrm{d}\boldsymbol{x} \\
&\geq C_\alpha\int_{\mathbb{T}_h} \langle\boldsymbol{x}\rangle^{-\alpha}\,\mathrm{e}^{-\frac{\|\boldsymbol{x}\|}{2}}\mathrm{d}\boldsymbol{x}\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \langle h\boldsymbol{k}\rangle^\alpha\,\mathrm{e}^{-\frac{\|h\boldsymbol{k}\|}{2}} \\
&= C_\alpha\cdot h^d\int_{\mathbb{T}} \langle h\boldsymbol{x}\rangle^{-\alpha}\,\mathrm{e}^{-\frac{\|h\boldsymbol{x}\|}{2}}\mathrm{d}\boldsymbol{x}\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \langle h\boldsymbol{k}\rangle^\alpha\,\mathrm{e}^{-\frac{\|h\boldsymbol{k}\|}{2}} \\
&\geq C_\alpha\cdot h^d\int_{\mathbb{T}} \langle\boldsymbol{x}\rangle^{-\alpha}\,\mathrm{e}^{-\frac{\|\boldsymbol{x}\|}{2}}\mathrm{d}\boldsymbol{x}\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \langle h\boldsymbol{k}\rangle^\alpha\,\mathrm{e}^{-\frac{\|h\boldsymbol{k}\|}{2}},
\end{aligned}$$

which implies

$$\begin{aligned}
\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \langle h\boldsymbol{k}\rangle^\alpha\,\mathrm{e}^{-\frac{\|h\boldsymbol{k}\|}{2}} &\leq C_\alpha^{-1}\int_{\mathbb{R}^d} \langle\boldsymbol{x}\rangle^\alpha\,\mathrm{e}^{-\frac{\|\boldsymbol{x}\|}{2}}\mathrm{d}\boldsymbol{x}\left(\int_{\mathbb{T}} \langle\boldsymbol{x}\rangle^{-\alpha}\,\mathrm{e}^{-\frac{\|\boldsymbol{x}\|}{2}}\mathrm{d}\boldsymbol{x}\right)^{-1}\cdot h^{-d} \\
&= C_\alpha\cdot h^{-d}.
\end{aligned}$$

Combining this with (50) yields that $B \leq C_{r,\alpha}\,h^{-d}$ which, together with (49), establishes the claim (47) and therefore completes the proof. □

In the rest of this section, we state and prove the interpolation counterpart of Theorem 2.

Theorem 4. *Assume that $1 \leq p \leq \infty$, $L \in \mathbb{N}$, $\alpha \geq 0$, and $r > d/p$. Let φ be an element of $W_{p,L+\alpha}(\mathbb{R}^d)$ that satisfies the Strang-Fix conditions of order L. Assume also that $\varphi[\cdot] \in \ell_{1,L+\alpha}(\mathbb{Z}^d)$ and $\sum_{\boldsymbol{k}\in\mathbb{Z}^d}\varphi[\boldsymbol{k}]\mathrm{e}^{-\mathrm{j}\langle\boldsymbol{\omega},\boldsymbol{k}\rangle}$ is nonzero for all $\boldsymbol{\omega} \in \mathbb{R}^d$. Then, there exists a constant $C_{\varphi,L,\alpha}$ such that, for all continuous functions f in $H^{L,r}_{p,-\alpha}(\mathbb{R}^d)$,*

$$\|f - I_{\varphi,h}f\|_{L_{p,-\alpha}(\mathbb{R}^d)} \leq C_{\varphi,L,\alpha}\cdot h^L\cdot\left\|(D^r f)^{(L)}\right\|_{L_{p,-\alpha}}, \tag{51}$$

when $h \to 0$.

Similar to the proof of Theorem 2, we divide the proof of Theorem 4 into two propositions.

Proposition 3. *For $1 \leq p \leq \infty$, $L \in \mathbb{N}$, $\alpha \geq 0$, $r > 0$, and J_h being the smoothing operator defined in* (15)*, there exists a constant $C_{L,\alpha}$ such that, for all $f \in H^{L,r}_{p,-\alpha}(\mathbb{R}^d)$ and for all $h \in (0,1)$,*

$$\|f - J_h f\|_{L^r_{p,-\alpha}(\mathbb{R}^d)} \leq C_{L,\alpha} \cdot h^L \cdot \left\|(D^r f)^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)}.$$

Proof. Put $B_r := \mathcal{F}^{-1}\left\{\langle\cdot\rangle^{-r}\right\}$. Since J_h is a convolution operator, we have the expression

$$f - J_h f = B_r * D^r f - J_h(B_r * D^r f) = B_r * (D^r f - J_h D^r f).$$

Hence

$$\|f - J_h f\|_{L^r_{p,-\alpha}(\mathbb{R}^d)} = \|D^r f - J_h D^r f\|_{L_{p,-\alpha}(\mathbb{R}^d)}. \tag{52}$$

We now apply Proposition 1 to $D^r f \in H^L_{p,-\alpha}(\mathbb{R}^d)$ to obtain

$$\|D^r f - J_h D^r f\|_{L_{p,-\alpha}(\mathbb{R}^d)} \leq C_{L,\alpha} \cdot h^L \cdot \left\|(D^r f)^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)}. \tag{53}$$

Putting (52) and (53) together completes the proof. □

Proposition 4. *Assume that $1 \leq p \leq \infty$, $L \in \mathbb{N}$, $\alpha \geq 0$, and $r > 0$. Let J_h be the smoothing operator defined in* (15)*. If φ satisfies the conditions of Theorem 4, there exists a constant $C_{\varphi,r,L,\alpha}$ such that, for all $f \in H^{L,r}_{p,-\alpha}(\mathbb{R}^d)$ and for all $h \in (0,1)$,*

$$\|J_h f - I_{\varphi,h} J_h f\|_{L_{p,-\alpha}(\mathbb{R}^d)} \leq C_{\varphi,r,L,\alpha} \cdot h^L \cdot \left\|(D^r f)^{(L)}\right\|_{L_{p,-\alpha}(\mathbb{R}^d)}. \tag{54}$$

Proof. We first show that $f \in H^L_{p,-\alpha}(\mathbb{R}^d)$. Indeed, since $f = B_r * D^r f$, where $B_r := \mathcal{F}^{-1}\left\{\langle\cdot\rangle^{-r}\right\}$, we have the estimate

$$\begin{aligned}
\left\|\partial^{\boldsymbol{\ell}} f\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} &= \left\|\partial^{\boldsymbol{\ell}}(B_r * D^r f)\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} = \left\|B_r * \partial^{\boldsymbol{\ell}} D^r f\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \\
&\leq C_\alpha \cdot \|B_r\|_{L_{1,\alpha}(\mathbb{R}^d)} \cdot \left\|\partial^{\boldsymbol{\ell}} D^r f\right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \qquad (55) \\
&= C_{r,\alpha} \cdot \left\|\partial^{\boldsymbol{\ell}} D^r f\right\|_{L_{p,-\alpha}(\mathbb{R}^d)}, \quad \forall |\boldsymbol{\ell}| \leq L \qquad (56)
\end{aligned}$$

where (55) is a consequence of weighted Young's inequality. On the other hand, it was shown in [25, Proposition 7] that $B_r \in L_{1,\alpha}(\mathbb{R}^d)$, for $r > 0$. This means that the constant $C_{r,\alpha}$ in (56) is finite, which then implies that $f \in H^L_{p,-\alpha}(\mathbb{R}^d)$.

Let $R_{\boldsymbol{x}}$ be the remainder of the order-$(L-1)$ Taylor series of the infinitely differentiable function $g := J_h f$ about $\boldsymbol{x}$. Since φ_{int} is a quasi-interpolant of order L, $I_{\varphi,h}$ maps every polynomial of order less than L to itself. Following the path of the proof of Proposition 2, we write

$$e(\boldsymbol{x}) := g(\boldsymbol{x}) - (I_{\varphi,h} g)(\boldsymbol{x}) = -\sum_{\boldsymbol{\ell} \in \mathbb{Z}^d} c_{\boldsymbol{x}}[\boldsymbol{\ell}] \varphi_{\text{int}}\left(\frac{\boldsymbol{x}}{h} - \boldsymbol{\ell}\right),$$

where the sequence $c_{\boldsymbol{x}}$ is redefined as

$$c_{\boldsymbol{x}}[\boldsymbol{\ell}] := R_{\boldsymbol{x}}(h\boldsymbol{\ell}), \quad \text{for } \boldsymbol{\ell} \in \mathbb{Z}^d.$$

Therefore, (26) still holds and we only need to estimate $\|c_{\boldsymbol{x},\boldsymbol{\ell}}\|_{\ell_p(\mathbb{Z}^d)}$, where

$$c_{\boldsymbol{x},\boldsymbol{\ell}}[\boldsymbol{k}] := \langle h\boldsymbol{k}\rangle^{-\alpha} |c_{h\boldsymbol{x}+h\boldsymbol{k}}[\boldsymbol{k} - \boldsymbol{\ell}]| = \langle h\boldsymbol{k}\rangle^{-\alpha} |R_{h\boldsymbol{x}+h\boldsymbol{k}}(h\boldsymbol{k} - h\boldsymbol{\ell})|.$$

Similarly to (27), we express

$$R_{h\boldsymbol{x}+h\boldsymbol{k}}(h\boldsymbol{k} - h\boldsymbol{\ell}) = \int_0^1 J_h T_{-\boldsymbol{\ell},\tau} f(h\boldsymbol{k}) \mathrm{d}\tau,$$

where the operator $T_{\boldsymbol{y},\tau}$ is given in (28). Repeating the manipulations in the proof of Proposition 2, we obtain the counterpart of (35):

$$\|c_{\boldsymbol{x},\boldsymbol{\ell}}\|_{\ell_p(\mathbb{Z}^d)} \leq C_{L,\alpha} \cdot h^{-d/p} \left\| f^{(L)} \right\|_{L_{p,-\alpha}(\mathbb{R}^d)} \langle \boldsymbol{x} + \boldsymbol{\ell} \rangle^{L+\alpha}.$$

Substituting this bound into (26), we end up with

$$\|e\|_{L_{p,-\alpha}(\mathbb{R}^d)} \leq C_{L,\alpha} \cdot \|\varphi_{\text{int}}\|_{W_{p,L+\alpha}(\mathbb{R}^d)} \cdot h^L \cdot \left\| f^{(L)} \right\|_{L_{p,-\alpha}(\mathbb{R}^d)}, \tag{57}$$

where $\|\varphi_{\text{int}}\|_{W_{p,L+\alpha}(\mathbb{R}^d)}$ is a finite constant thanks to Lemma 4. Combining (57) and (56) gives us the desired bound (54). □

Proof of Theorem 4. Without loss of generality, assume that $h \in (0,1)$. Let $g := J_h f$. By the triangle inequality

$$\|f - I_{\varphi,h} f\|_{L_{p,-\alpha}(\mathbb{R}^d)} \leq \|f - g\|_{L_{p,-\alpha}(\mathbb{R}^d)} + \|I_{\varphi,h}(f-g)\|_{L_{p,-\alpha}(\mathbb{R}^d)} + \|g - I_{\varphi,h} g\|_{L_{p,-\alpha}(\mathbb{R}^d)}. \tag{58}$$

From Theorem 3 and Propositions 3, the first two terms in the right-hand side of (58) are bounded as

$$\begin{aligned} \|f - g\|_{L_{p,-\alpha}(\mathbb{R}^d)} + \|I_{\varphi,h}(f-g)\|_{L_{p,-\alpha}(\mathbb{R}^d)} &\leq \|f - g\|_{L^r_{p,-\alpha}(\mathbb{R}^d)} + C_{\varphi,r,\alpha} \cdot \|f - g\|_{L^r_{p,-\alpha}(\mathbb{R}^d)} \\ &\leq C_{\varphi,r,L,\alpha} \cdot h^L \cdot \left\| (D^r f)^{(L)} \right\|_{L_{p,-\alpha}(\mathbb{R}^d)}, \end{aligned} \tag{59}$$

whereas the third term is also bounded, according to Proposition 4, as

$$\|g - I_{\varphi,h} g\|_{L_{p,-\alpha}(\mathbb{R}^d)} \leq C_{\varphi,r,L,\alpha} \cdot h^L \cdot \left\| (D^r f)^{(L)} \right\|_{L_{p,-\alpha}(\mathbb{R}^d)}. \tag{60}$$

Finally, the desired bound (51) is obtained by combining (58), (59) and (60). □

V. Proofs of Auxiliary Results

A. Proof of Lemma 1

It is clear that

$$\mathcal{F}\left\{D^L_{\boldsymbol{u}} f\right\} = (\mathrm{j}\,\langle \boldsymbol{u}, \cdot \rangle)^L\, \hat{f}.$$

On the other hand, the Fourier transform of the B-spline β^{L-1} is given by [52]

$$\hat{\beta}^{L-1}(\omega) = \left(\frac{1 - \mathrm{e}^{-\mathrm{j}\omega}}{\mathrm{j}\omega} \right)^L.$$

Therefore, the Fourier transform of the right-hand side (RHS) of (18) is given by

$$\begin{aligned} \mathcal{F}\{\text{RHS}\} &= \int_{\mathbb{R}} \mathcal{F}\left\{D^L_{\boldsymbol{u}} f(\cdot - t\boldsymbol{u})\right\} \beta^{L-1}(t) \mathrm{d}t \\ &= \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{j}\langle t\boldsymbol{u}, \cdot \rangle} \mathcal{F}\left\{D^L_{\boldsymbol{u}} f\right\} \beta^{L-1}(t) \mathrm{d}t \\ &= (\mathrm{j}\,\langle \boldsymbol{u}, \cdot \rangle)^L\, \hat{f} \cdot \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{j}\langle \boldsymbol{u}, \cdot \rangle t} \beta^{L-1}(t) \mathrm{d}t \\ &= (\mathrm{j}\,\langle \boldsymbol{u}, \cdot \rangle)^L\, \hat{\beta}^{L-1}(\langle \boldsymbol{u}, \cdot \rangle) \hat{f} \\ &= \left(1 - \mathrm{e}^{-\mathrm{j}\langle \boldsymbol{u}, \cdot \rangle}\right)^L \hat{f}, \end{aligned}$$

which is exactly the Fourier transform of the left-hand side of (18), completing the proof.

B. Proof of Lemma 2

The claim is trivial for $L = 0$. We now show (19) based on the induction hypothesis that

$$\left|D_{\boldsymbol{u}}^{L-1} f(\boldsymbol{x})\right| \leq \|\boldsymbol{u}\|_{\infty}^{L-1} \cdot f^{(L-1)}(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in \mathbb{R}^d. \tag{61}$$

By definition of directional derivatives, we have that

$$\left|D_{\boldsymbol{u}}^{L} f(\boldsymbol{x})\right| = \left|\sum_{i=1}^{d} u_i \frac{\partial}{\partial x_i} D_{\boldsymbol{u}}^{L-1} f(\boldsymbol{x})\right| \leq \|\boldsymbol{u}\|_{\infty} \cdot \sum_{i=1}^{d} \left|D_{\boldsymbol{u}}^{L-1} \frac{\partial f}{\partial x_i}(\boldsymbol{x})\right|.$$

It then follows from (61) that

$$\begin{aligned}
\left|D_{\boldsymbol{u}}^{L} f(\boldsymbol{x})\right| &\leq \|\boldsymbol{u}\|_{\infty} \cdot \|\boldsymbol{u}\|_{\infty}^{L-1} \cdot \sum_{i=1}^{d} \left(\frac{\partial f}{\partial x_i}\right)^{(L-1)}(\boldsymbol{x}) \\
&\leq \|\boldsymbol{u}\|_{\infty}^{L} \cdot \sum_{i=1}^{d} \sum_{|\boldsymbol{k}|=L-1} \left|\partial^{\boldsymbol{k}} \left(\frac{\partial f}{\partial x_i}\right)(\boldsymbol{x})\right| \\
&= \|\boldsymbol{u}\|_{\infty}^{L} \cdot f^{(L)}(\boldsymbol{x}),
\end{aligned}$$

completing the proof.

C. Proof of Lemma 3

It is clear from the definition of J_h that $\sigma_{1/h} J_h = J \sigma_{1/h}$. Then, we write

$$\sigma_{1/h} J_h f = J \sigma_{1/h} f = (\sigma_{1/h} f) * \psi, \tag{62}$$

where the kernel ψ is given by

$$\psi := \sum_{n=1}^{L} (-1)^{n-1} \binom{L}{n} \frac{\sigma_n \chi}{n^d}.$$

Since χ is a compactly supported smooth function, it is easy to see that the kernel ψ given above is an element of the hybrid-norm space $W_{\infty,\alpha}(\mathbb{R}^d)$, which is clearly a subspace of $W_{p',\alpha}(\mathbb{R}^d)$. Then, the convolution expression in (62) allows us to invoke [25, Proposition 5] to obtain

$$\begin{aligned}
\left\|(\sigma_{1/h} J_h f)[\cdot]\right\|_{\ell_{p,1/w_h}(\mathbb{Z}^d)} &\leq C_{\alpha} \, \|\psi\|_{W_{p',w_h}} \cdot \left\|\sigma_{1/h} f\right\|_{L_{p,1/w_h}(\mathbb{R}^d)} \\
&\leq C_{\alpha} \, \|\psi\|_{W_{p',\alpha}} \cdot \left\|\sigma_{1/h} f\right\|_{L_{p,1/w_h}(\mathbb{R}^d)} \qquad (63) \\
&= C_{\alpha} \, \|\psi\|_{W_{p',\alpha}} \cdot h^{-d/p} \cdot \|f\|_{L_{p,-\alpha}(\mathbb{R}^d)}, \qquad (64)
\end{aligned}$$

where (63) is due to the assumption that $h \in (0, 1)$ and (64) is the result of a change of variable. Putting $C_{L,\alpha} := C_{\alpha} \|\psi\|_{W_{p',\alpha}}$ gives us the desired bound (22).

D. Proof of Lemma 4

Recall that, for $\alpha \geq 0$, the weight $\langle \cdot \rangle^{\alpha}$ is submultiplicative and satisfies the Gelfand-Raikov-Shilov condition. Since $\varphi[\cdot] \in \ell_{1,\alpha}(\mathbb{Z}^d)$ and since $\sum_{\boldsymbol{k} \in \mathbb{Z}^d} \varphi[\boldsymbol{k}] \mathrm{e}^{-\mathrm{j}\langle \boldsymbol{\omega}, \boldsymbol{k} \rangle}$ is nonzero for all $\boldsymbol{\omega} \in \mathbb{R}^d$, we are allowed to invoke the weighted version of Wiener's lemma [43, Theorem 6.2] to deduce that the sequence a defined in (38) also belongs to $\ell_{1,\alpha}(\mathbb{Z}^d)$. Now that φ_{int} has the representation (37) with $a \in \ell_{1,\alpha}(\mathbb{Z}^d)$ and $\varphi \in W_{p,\alpha}(\mathbb{R}^d)$, it must be that $\varphi_{\mathrm{int}} \in W_{p,\alpha}(\mathbb{R}^d)$ as a consequence of [25, Lemma 1].

REFERENCES

[1] I. J. Schoenberg, "Contributions to the problem of approximation of equidistant data by analytic functions," *Quart. Appl. Math.*, vol. 4, pp. 45–99, 112–141, 1946.

[2] ——, *Cardinal Spline Interpolation*. Philadelphia, PA: Society of Industrial and Applied Mathematics, 1973.

[3] C. de Boor, *A Practical Guide to Splines*. New York, NY: Springer-Verlag, 1978.

[4] L. L. Schumaker, *Spline Functions: Basic Theory*. New York, NY: Wiley, 1981.

[5] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949.

[6] M. Unser, "Sampling—50 years after Shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, Apr. 2000.

[7] G. Strang and G. Fix, "A Fourier analysis of the finite element variational method," in *Constructive Aspects of Functional Analysis*, G. Geymonat, Ed. Rome, Italy: Springer, 1971, pp. 796–830.

[8] C. de Boor and R.-Q. Jia, "Controlled approximation and a characterization of the local approximation order," *Proc. Amer. Math. Soc.*, vol. 95, no. 4, pp. 547–553, 1985.

[9] W. A. Light and E. W. Cheney, "Quasi-interpolation with translates of a function having noncompact support," *Constr. Approx*, vol. 8, no. 1, pp. 35–48, 1992.

[10] R.-Q. Jia and J. Lei, "Approximation by multiinteger translates of functions having global support," *J. Approx. Theory*, vol. 72, no. 1, pp. 2–23, 1993.

[11] J. Lei, "L_p-approximation by certain projection operators," *J. Math. Anal. Appl.*, vol. 185, no. 1, pp. 1–14, 1994.

[12] C. de Boor, R. A. DeVore, and A. Ron, "Approximation from shift-invariant subspaces of $L_2(\mathbb{R}^d)$," *Trans. Amer. Math. Soc.*, vol. 341, no. 2, pp. 787–806, 1994.

[13] M. Unser, "Quasi-orthogonality and quasi-projections," *Appl. Comput. Harmon. Anal.*, vol. 3, no. 3, p. 201214, Jul. 1996.

[14] M. Unser and I. Daubechies, "On the approximation power of convolution-based least squares versus interpolation," *IEEE Trans. Signal Process.*, vol. 45, no. 7, pp. 1697–1711, Jul. 1997.

[15] T. Blu and M. Unser, "Approximation error for quasi-interpolators and (multi-) wavelet expansions," *Appl. Comput. Harmon. Anal.*, vol. 6, no. 2, pp. 219–251, Mar. 1999.

[16] ——, "Quantitative Fourier analysis of approximation: Part I—Interpolators and projectors," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2783–2795, Oct. 1999.

[17] ——, "Quantitative Fourier analysis of approximation: Part II—Wavelets," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2796–2806, Oct. 1999.

[18] C. K. Chui, *Multivariate Splines*. Philadelphia, PA: Soc. Ind. Applied Math., 1988.

[19] C. de Boor, "Quasi-interpolants and approximation power of multivariate splines," in *Computation of Curves and Surfaces*, W. Dahmen, M. Gasca, and C. A. Micchelli, Eds. Dordrecht, The Netherlands: Kluwer, 1990, pp. 313–345.

[20] W. A. Light, "Recent developments in the Strang-Fix theory for approximation orders," in *Curves and Surfaces*, P. J. Laurent, A. L. Méhauté, and L. L. Schumaker, Eds. Boston, MA: Academic, 1991, pp. 285–292.

[21] P. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1741–1757, May 2007.

[22] M. Unser and P. D. Tafti, *An Introduction to Sparse Stochastic Processes*. Cambridge University Press, 2014.

[23] M. Unser, P. D. Tafti, and Q. Sun, "A unified formulation of Gaussian versus sparse stochastic processes—Part I: Continuous-domain theory," *IEEE Trans. Inform. Theory*, vol. 60, no. 3, pp. 1945–1962, Mar. 2014.

[24] M. Unser, P. D. Tafti, A. Amini, and H. Kirshner, "A unified formulation of Gaussian versus sparse stochastic processes—Part II: Discrete-domain theory," *IEEE Trans. Inform. Theory*, vol. 60, no. 5, pp. 3036–3051, May 2014.

[25] H. Q. Nguyen and M. Unser, "A sampling theory for non-decaying signals," *Appl. Comput. Harmon. Anal.*, vol. 43, no. 1, pp. 76–93, 2017.

[26] ——, "Generalized Poisson summation formula for tempered distributions," in *Proc. 11th Int. Conf. Sampling Theory and Applications (SampTA'15)*, May 25-29, 2015, pp. 1–5.

[27] H. Q. Nguyen, M. Unser, and J.-P. Ward, "Generalized Poisson summation formulas for continuous functions of polynomial growth," *J. Fourier Anal. Appl.*, vol. 23, no. 2, pp. 442–461, 2017.

[28] N. Wiener, *The Fourier Integral and Certain of its Applications*. Cambridge: MIT Press, 1933.

[29] H. G. Feichtinger, "New results on regular and irregular sampling based on Wiener amalgams," in *Proc. Conf. Function Spaces*, ser. Lecture Notes in Pure and Appl. Math., K. Jarosz, Ed. New York: Dekker, 1991, vol. 136, pp. 107–121.

[30] C. Heil, "An introduction to weighted Wiener amalgams," in *Wavelets and Their Applications*, M. Krishna, R. Radha, and S. Thangavelu, Eds. New Delhi: Allied Publishers, 2003, pp. 183–216.

[31] R. Q. Jia, "Approximation by quasi-projection operators in Besov spaces," *J. Approx. Theory*, vol. 162, no. 1, pp. 186–200, 2010.

[32] C. Bardaro, P. Butzer, R. Stens, and G. Vinti, "Approximation error of the Whittaker cardinal series in terms of an averaged modulus of smoothness covering discontinuous signals," *J. Math. Anal. Appl.*, vol. 316, no. 1, pp. 269–306, 2006.

[33] A. Krivoshein and M. Skopina, "Multivariate sampling-type approximation," *Anal. Appl.*, vol. 15, no. 4, pp. 521–542, 2017.

[34] Y. Kolomoitsev, A. Krivoshein, and M. Skopina, "Differential and falsified sampling expansions," *J. Fourier Anal. Appl.*, 2017, article in press. [Online]. Available: https://doi.org/10.1007/s00041-017-9559-1

[35] Y. Kolomoitsev and M. Skopina, "Approximation by multivariate Kantorovich-Kotelnikov operators," *J. Math. Anal. Appl.*, vol. 456, no. 1, pp. 195–213, 2017.

[36] N. Tomita, "Strang-Fix theory for approximation order in weighted L^p-spaces and Herz spaces," *J. Funct. Space Appl.*, vol. 4, no. 1, pp. 7–24, 2006.

[37] H. A. Aimar, A. L. Bernardis, and F. J. Martín-Reyes, "Multiresolution approximations and wavelet bases of weighted L^p spaces," *J. Fourier Anal. Appl.*, vol. 9, no. 5, pp. 497–510, 2003.

[38] B. Muckenhoupt, "Weighted norm inequalities for the Hardy maximal function," *Trans. Amer. Math. Soc.*, vol. 165, pp. 207–226, 1972.

[39] G. H. Hardy and J. E. Littlewood, "A maximal theorem with function-theoretic applications," *Acta Math.*, vol. 54, no. 1, pp. 81–116, 1930.
[40] E. M. Stein, "On certain operators on L_p spaces," Ph.D. dissertation, University of Chicago, Chicago, IL, 1955.
[41] C. Fefferman and E. M. Stein, "Some maximal inequalities," *Amer. J. Math.*, vol. 93, no. 1, pp. 107–115, Jan. 1971.
[42] D. S. Kurtz, "Littlewood-Paley and multiplier theorems on weighted L^p spaces," *Trans. Amer. Math. Soc.*, vol. 259, no. 1, pp. 235–254, May 1980.
[43] K. Gröchenig, "Weight functions in time-frequency analysis," arXiv:math/0611174 [math.FA], 2006.
[44] I. Gel'fand, D. Raikov, and G. Shilov, *Commutative normed rings*. Chelsea Publishing Co., 1964.
[45] L. Grafakos, *Modern Fourier Analysis*, 2nd ed. Springer, 2008.
[46] ——, *Classical Fourier Analysis*, 2nd ed. Springer, 2008.
[47] D. E. Edmunds, V. Kokilashvili, and A. Meskhi, "On Fourier multipliers in weighted Triebel–Lizorkin spaces," *J. Inequal. Appl.,* vol. 74, no. 4, pp. 555–591, 2002.
[48] S. Król, "Fourier multipliers on weighted L^p spaces," arXiv:1403.4477 [math.CA], 2014.
[49] C. de Boor and G. Fix, "Spline approximation by quasi-interpolants," *J. Approx. Theory*, vol. 8, pp. 19–45, 1973.
[50] C. de Boor, "The polynomials in the linear span of integer translates of a compactly supported function," *Constr. Approx.*, vol. 3, pp. 199–208, 1987.
[51] C. K. Chui and H. Diamond, "A characterization of multivariate quasi-interpolation formulas and applications," *Numer. Math.*, vol. 57, pp. 105–121, 1990.
[52] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part I—theory," *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 821–833, Feb. 1993.
[53] ——, "B-spline signal processing: Part II—efficiency design and applications," *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 834–848, Feb. 1993.
[54] P. J. Davis, *Interpolation and Approximation*. New York, NY: Dover, 1975.