Noname manuscript No.
(will be inserted by the editor)

Reproducing Kernels of Generalized Sobolev Spaces via a Green Function Approach with Distributional Operators

Gregory E. Fasshauer · Qi Ye

Abstract In this paper we extend the definition of generalized Sobolev space and subsequent theoretical results established recently for positive definite kernels and differential operators in the article [21]. In the present paper the semi-inner product of the generalized Sobolev space is set up by a vector distributional operator $\mathbf{P}$ consisting of finitely or countably many distributional operators P_n, which are defined on the dual space of the Schwartz space. The types of operators we now consider include not only differential operators, but also more general distributional operators such as pseudo-differential operators. We deduce that a certain appropriate full-space Green function G with respect to $L := \mathbf{P}^{*T}\mathbf{P}$ now becomes a conditionally positive function. In order to support this claim we ensure that the distributional adjoint operator $\mathbf{P}^*$ of $\mathbf{P}$ is well-defined in the distributional sense. Under sufficient conditions, the native space (reproducing-kernel Hilbert space) associated with the Green function G can be imbedded into or even be equivalent to a generalized Sobolev space. As an application, we take linear combinations of translates of the Green function with possibly added polynomial terms and construct a multivariate minimum-norm interpolant $s_{f,X}$ to data values sampled from an unknown generalized Sobolev function f at data sites located in some set $X \subset \mathbb{R}^d$. We will provide several examples, such as Matérn kernels or Gaussian kernels, that illustrate how many reproducing-kernel Hilbert spaces of well-known reproducing kernels are equivalent to a generalized Sobolev space.

Keywords kernel approximation · reproducing kernel Hilbert spaces · generalized Sobolev spaces · Green functions · conditionally positive definite functions

Mathematics Subject Classification (2000): Primary 41A30, 65D05; Secondary 34B27, 41A63, 46E22, 46E35

Gregory E. Fasshauer
Department of Applied Mathematics, Illinois Institute of Technology, Chicago IL 60616
E-mail: fasshauer@iit.edu

Qi Ye
Department of Applied Mathematics, Illinois Institute of Technology, Chicago IL 60616
Tel.: +1-312-451-7342
Fax: +1-312-567-3135
E-mail: qye3@iit.edu

1 Introduction

A large and increasing number of recent books and research papers apply radial basis functions or other kernel-based approximation methods to such fields as scattered data approximation, statistical or machine learning and the numerical solution of partial differential equations, e.g., [2,3,5,8,13,14,19,20]. Generally speaking, the fundamental underlying practical problem common to many of these applications can be represented in the following way. Given a set of data sites $X \subset \mathbb{R}^d$ and associated values $Y \subset \mathbb{R}$ sampled from an unknown function f, we use translates of a kernel function Φ and possible polynomial terms to set up an interpolant $s_{f,X}$ to approximate the function f. When f belongs to the related native space of Φ, we can obtain error bounds and optimality properties of this interpolation method. If Φ is only conditionally positive definite (instead of the more straightforward positive definite case), then it is known that the native space can also become a reproducing-kernel Hilbert space with a reproducing kernel computed from Φ along with additional polynomial terms (see Section 3 and [20]). Nevertheless, there still remains a couple of difficult and challenging questions to be answered for kernel methods: *What kind of functions belong to the related native space of a given kernel function, and which kernel function is the best for us to utilize for a particular application?* In particular, a better understanding of the native space in relation to traditional smoothness spaces (such as Sobolev spaces) is highly desirable. The latter question is partially addressed by the use of techniques such as cross-validation and maximum likelihood estimation to obtain optimally scaled kernels for any particular application (see e.g., [18,19]). However, at the function space level, the question of scale is still in need of a satisfactory answer.

By generalizing the ideas of [21], we will deal with these questions in a different way than the authors of the survey paper [13] did. In this paper, we want to show that the kernel functions and native spaces (reproducing kernels and reproducing-kernel Hilbert spaces) can be computed via Green functions and generalized Sobolev spaces induced by some vector distributional operators $\mathbf{P} := (P_1, \cdots, P_n, \cdots)^T$ consisting of finitely or countably many distributional operators P_n (see Definition 4.1). We can further check that the differential operators defined in [21] are special cases of these distributional operators. As a consequence, all theoretical results in [21] can also be verified by the conclusions of this article.

Some well known examples covered by our theory include the Duchon spaces and Beppo-Levi spaces associated with polyharmonic splines (see Example 5.1 and 5.6). Moreover, in [12] the author expressed a desire to choose the "best" scale parameter of a given kernel function for a particular interpolation problem by looking at scaled versions of the classical Sobolev space via different scale parameters. Example 5.3 and 5.7 tell us that we can balance the derivatives through selecting appropriate scale parameters to reconstruct the classical Sobolev space by the reproducing-kernel inner products of the Sobolev splines (Matérn functions). Finally, Example 5.8 shows that the native space of the ubiquitous Gaussian function (the reproducing-kernel Hilbert space of the Gaussian kernel) is equivalent to a generalized Sobolev space, which can be applied to support vector machines and in the study of motion coherence (see e.g., [16,22]).

In this article, we use the notation $\mathrm{Re}(\mathcal{E})$ to be the collection of all real-valued functions of the function space $\mathcal{E}$. For example, $\mathrm{Re}(\mathrm{C}(\mathbb{R}^d))$ denotes the collection of all real-valued continuous functions on $\mathbb{R}^d$. $\mathcal{SI}$ is defined as the collection of *slowly increasing functions* which grow at most like any particular fixed polynomial, i.e.,

$$\mathcal{SI} := \left\{ f : \mathbb{R}^d \to \mathbb{C} : f(\boldsymbol{x}) = O(\|\boldsymbol{x}\|_2^m) \text{ as } \|\boldsymbol{x}_2\| \to \infty \text{ for some } m \in \mathbb{N}_0 \right\}.$$

(The notation $f = O(g)$ means that there is a positive number M such that $|f| \leqslant M\,|g|$.) Roughly speaking, our generalized Soblev space is a generalization of the classical real-valued $\mathrm{L}_2(\mathbb{R}^d)$-based Sobolev space. Here the classical Sobolev space is usually given by

$$\mathcal{H}^n(\mathbb{R}^d) := \left\{ f \in \mathrm{Re}(\mathrm{L}_1^{loc}(\mathbb{R}^d)) \cap \mathcal{SI} : D^\alpha f \in \mathrm{L}_2(\mathbb{R}^d) \text{ for all } |\alpha| \le n, \alpha \in \mathbb{N}_0^d \right\}$$

with inner product

$$(f,g)_{\mathcal{H}^n(\mathbb{R}^d)} := \sum_{|\alpha|\le n} (D^\alpha f, \overline{D^\alpha g}), \quad f,g \in \mathcal{H}^n(\mathbb{R}^d),$$

where $(\cdot, \bar{\cdot})$ is the standard $\mathrm{L}_2(\mathbb{R}^d)$-inner product. Our concept of a *generalized* Sobolev space (to be defined in detail below, see Definition 4.4) will be of a very similar form, namely

$$\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d) := \left\{ f \in \mathrm{Re}(\mathrm{L}_1^{loc}(\mathbb{R}^d)) \cap \mathcal{SI} : \{P_n f\}_{n=1}^\infty \subseteq \mathrm{L}_2(\mathbb{R}^d) \text{ and } \sum_{n=1}^\infty \|P_n f\|_{\mathrm{L}_2(\mathbb{R}^d)}^2 < \infty \right\}$$

with the semi-inner product

$$(f,g)_{\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)} := \sum_{n=1}^\infty (P_n f, \overline{P_n g}), \quad f,g \in \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d).$$

Since the Dirac delta function δ_0 at the origin is just a tempered distribution belonging to the dual space of the Schwartz space, the Green function G we introduce in Definition 4.3 needs to be regarded as a tempered distribution as well. Thus we want to define a distributional operator L on the dual space of the Schwartz space so that $LG = \delta_0$. The distributional operator and its distributional adjoint operator are well-defined in Section 4.1. According to Theorem 4.1, we can prove that an even Green function $G \in \mathrm{Re}(\mathrm{C}(\mathbb{R}^d)) \cap \mathcal{SI}$ is a conditionally positive definite function of some order $m \in \mathbb{N}_0$. Therefore, we can construct the related native space $\mathcal{N}_G^m(\mathbb{R}^d)$ of G as a complete semi-inner product space (see Section 3 and [20]). Moreover, the distributional operator L can be computed by a vector distributional operator $\mathbf{P} := (P_1, \cdots, P_n)^T$ and its distributional adjoint $\mathbf{P}^*$, i.e., $L = \mathbf{P}^{*T}\mathbf{P} = \sum_{j=1}^n P_j^* P_j$. Under some sufficient conditions, we will further obtain a result in Theorem 4.2 that shows that the native space $\mathcal{N}_G^m(\mathbb{R}^d)$ is always a subspace of the generalized Sobolev space $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ and that their semi-inner products are the same on $\mathcal{N}_G^m(\mathbb{R}^d)$. This implies that the native spaces can be imbedded into the generalized Sobolev spaces. By Lemma 4.5, we know that $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d) \cap \mathrm{C}(\mathbb{R}^d) \cap \mathrm{L}_2(\mathbb{R}^d)$ is also a subspace of $\mathcal{N}_G^m(\mathbb{R}^d)$. Theorems 4.4 and 4.6 tell us that $\mathcal{N}_G^m(\mathbb{R}^d)$ may even be equivalent to $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$. However, we provide Example 5.5 to show that $\mathcal{N}_G^m(\mathbb{R}^d)$ is not always equivalent to $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$. In other words, $\mathcal{N}_G^m(\mathbb{R}^d)$ is sometimes just a proper subspace of $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$. We complete the proofs needed for the theoretical framework in this article by an application of the techniques of distributional Fourier transforms (see Section 4.2 and [17]) and generalized Fourier transforms (see Section 4.1 and [20]).

2 Background

Given data sites $X = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N\} \subset \mathbb{R}^d$ (which we also identify with the centers of our kernel functions below) and sampled values $Y = \{y_1, \cdots, y_N\} \subset \mathbb{R}$ of a real-valued continuous function f on X, we wish to approximate this function f by a linear combination of translates of a conditionally positive definite function Φ of order m (see Section 3.1) along with possible real-valued polynomial terms.

To this end we set up the interpolant in the form

$$s_{f,X}(\boldsymbol{x}) := \sum_{j=1}^{N} c_j \Phi(\boldsymbol{x} - \boldsymbol{x}_j) + \sum_{k=1}^{Q} \beta_k p_k(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^d, \tag{2.1}$$

and require it to satisfy the additional interpolation and moment conditions

$$\begin{aligned} s_{f,X}(\boldsymbol{x}_j) &= y_j, \quad j = 1, \ldots, N, \\ \sum_{j=1}^{N} c_j p_k(\boldsymbol{x}_j) &= 0, \quad k = 1, \ldots, Q, \end{aligned} \tag{2.2}$$

where Q is the dimension of $\pi_{m-1}(\mathbb{R}^d)$ and $\{p_1, \cdots, p_Q\}$ is a basis of $\pi_{m-1}(\mathbb{R}^d)$. Here $\pi_{m-1}(\mathbb{R}^d)$ denotes the *space of real-valued polynomials of degree less than m*, i.e.,

$$\pi_{m-1}(\mathbb{R}^d) := \Big\{ p(\boldsymbol{x}) := \sum_{|\alpha| \leqslant m-1} c_\alpha \boldsymbol{x}^\alpha : c_\alpha \in \mathbb{R},\ |\alpha| < m,\ \alpha \in \mathbb{N}_0^d \Big\}.$$

If X contains a $\pi_{m-1}(\mathbb{R}^d)$-unisolvent set $\{\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_Q\}$, i.e., $p \in \pi_{m-1}(\mathbb{R}^d)$ such that $p(\boldsymbol{\xi}_k) = 0$ for all $k = 1, \ldots, Q$ if and only if $p \equiv 0$, then the above system (2.2) is equivalent to a uniquely solvable linear system

$$\begin{pmatrix} \mathbf{A}_{\Phi,X} & \mathbf{P}_X \\ \mathbf{P}_X^T & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{c} \\ \boldsymbol{\beta} \end{pmatrix} = \begin{pmatrix} \mathbf{Y} \\ 0 \end{pmatrix}, \tag{2.3}$$

where $\mathbf{A}_{\Phi,X} := \big(\Phi(\boldsymbol{x}_j - \boldsymbol{x}_k)\big)_{j,k=1}^{N,N} \in \mathbb{R}^{N\times N}$, $\mathbf{P}_X := \big(p_k(\boldsymbol{x}_j)\big)_{j,k=1}^{N,Q} \in \mathbb{R}^{N\times Q}$, $\boldsymbol{c} := (c_1, \cdots, c_N)^T$, $\boldsymbol{\beta} := (\beta_1, \cdots, \beta_Q)^T$ and $\mathbf{Y} := (y_1, \cdots, y_N)^T$. All of the above is discussed in detail in [20, Chapter 8.5].

Example 2.1 One of the best known examples that fits into this framework is the *cubic spline* interpolant

$$s_{f,X}(x) := \sum_{j=1}^{N} c_j \left| x - x_j \right|^3 + \beta_2 x + \beta_1, \quad x \in \mathbb{R}.$$

If we let $\Phi(x) = |x|^3$, then Φ is known to be a conditionally positive definite function of order 2.

In addition, if we take $\mathbf{P} := d^2/dx^2$ and $\mathbf{P}^* := d^2/dx^2$ so that $L = \mathbf{P}^{*T}\mathbf{P} = d^4/dx^4$, then $\frac{1}{12}\Phi$ is a Green function with respect to L and the interpolant $s_{f,X}$ as set up in (2.1)-(2.2) is the minimum semi-norm interpolant for all real-valued functions $f \in \mathrm{C}^2(\mathbb{R})$ such that $f'' \in \mathrm{L}_2(\mathbb{R})$ (see [2, Chapter 6.1.5]).

Example 2.2 Another well-known example is the *tension spline* interpolant (see [11, 15])

$$s_{f,X}(x) := \sum_{j=1}^{N} c_j \Phi(x - x_j) + \beta_1, \text{ where } \Phi(x) := -\frac{\exp(-\sigma|x|) + \sigma|x|}{2\sigma^3}, \; x \in \mathbb{R}, \quad (2.4)$$

and $\sigma > 0$ is called the *tension parameter*.

We can check that Φ is a conditionally positive definite function of order 1. If we define $\mathbf{P} = (P_1, P_2)^T := (d^2/dx^2, \sigma d/dx)^T$ and $\mathbf{P}^* = (P_1^*, P_2^*)^T := (d^2/dx^2, -\sigma d/dx)^T$, then Φ is a Green function with respect to $L := \mathbf{P}^{*T}\mathbf{P} = d^4/dx^4 - \sigma^2 d^2/dx^2$.

Moreover, the interpolant $s_{f,X}$ from (2.1)-(2.2) minimizes the semi-norm $|\cdot|_{TS}$ of all real-valued functions $f \in \mathrm{C}^2(\mathbb{R})$ and

$$|f|_{TS}^2 := (\mathbf{P}f, \overline{\mathbf{P}f})_{TS} = \int_{\mathbb{R}} \left|f''(x)\right|^2 \mathrm{d}x + \int_{\mathbb{R}} \sigma^2 \left|f'(x)\right|^2 \mathrm{d}x < \infty$$

subject to the constraints $s_{f,X}(x_j) = y_j$, $j = 1, \cdots, N$.

As we will show in Section 3.2, we can in general construct a reproducing-kernel Hilbert space $\mathcal{N}_\Phi^m(\mathbb{R}^d)$ from Φ such that the interpolant $s_{f,X}$ is the best approximation of the function f in $\mathcal{N}_\Phi^m(\mathbb{R}^d)$ fitting the sample values Y on the data sites X. The construction consists of first introducing a complete $\mathcal{N}_\Phi^m(\mathbb{R}^d)$-semi-inner product for the native space $\mathcal{N}_\Phi^m(\mathbb{R}^d)$ and then using this semi-inner product to set up a new inner product such that the native space $\mathcal{N}_\Phi^m(\mathbb{R}^d)$ becomes a reproducing-kernel Hilbert space whose reproducing kernel is determined by Φ.

3 Conditionally Positive Definite Functions and Native Spaces

Most of the material presented in this section can be found in the excellent monograph [20]. For the reader's convenience we repeat here what is essential to our discussion later on.

3.1 Conditionally Positive Definite Functions

Definition 3.1 ([20, Definition 8.1]) A continuous even function $\Phi : \mathbb{R}^d \to \mathbb{R}$ is said to be a *conditionally positive definite function of order* $m \in \mathbb{N}_0$ if, for all $N \in \mathbb{N}$, all pairwise distinct centers $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N \in \mathbb{R}^d$, and all $\boldsymbol{c} = (c_1, \cdots, c_N)^T \in \mathbb{R}^N \setminus \{0\}$ satisfying

$$\sum_{j=1}^{N} c_j p(\boldsymbol{x}_j) = 0$$

for all $p \in \pi_{m-1}(\mathbb{R}^d)$, the quadratic form

$$\sum_{j=1}^{N}\sum_{k=1}^{N} c_j c_k \Phi(\boldsymbol{x}_j - \boldsymbol{x}_k) > 0.$$

In the case $m = 0$ with $\pi_{-1}(\mathbb{R}^d) := \{0\}$ the function Φ is called *positive definite*.

In general, we can not hope for $\Phi \in \mathrm{C}(\mathbb{R}^d)$ to be integrable. However, we can restrict Φ to be a slowly increasing function, i.e., $\Phi \in \mathcal{SI}$.

Next, we want to have a criterion to decide whether Φ is a conditionally positive definite function of order $m \in \mathbb{N}_0$. In Wendland's book [20], the generalized Fourier transform of order m is employed to determine the conditional positive definiteness of Φ. Let a special test function space $\mathcal{S}_{2m}$ be defined as

$$\mathcal{S}_{2m} := \left\{\gamma \in \mathcal{S} : \gamma(\boldsymbol{x}) = O\left(\|\boldsymbol{x}\|_2^{2m}\right) \text{ as } \|\boldsymbol{x}\|_2 \to 0\right\},$$

where the *Schwartz space* $\mathcal{S}$ consists of all functions $\gamma \in \mathrm{C}^\infty(\mathbb{R}^d)$ that satisfy

$$\sup_{\boldsymbol{x}\in\mathbb{R}^d} \left|\boldsymbol{x}^\beta D^\alpha \gamma(\boldsymbol{x})\right| \leqslant C_{\alpha,\beta,\gamma}$$

for all multi-indices $\alpha, \beta \in \mathbb{N}_0^d$ with a constant $C_{\alpha,\beta,\gamma}$.

Definition 3.2 ([20, Definition 8.9]) Suppose that $\Phi \in \mathrm{C}(\mathbb{R}^d) \cap \mathcal{SI}$. A measurable function $\hat{\phi} \in \mathrm{L}_2^{loc}(\mathbb{R}^d \backslash \{0\})$ is called a *generalized Fourier transform* of Φ if there exists an integer $m \in \mathbb{N}_0$ such that

$$\int_{\mathbb{R}^d} \Phi(\boldsymbol{x})\hat{\gamma}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^d} \hat{\phi}(\boldsymbol{x})\gamma(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \quad \text{for each } \gamma \in \mathcal{S}_{2m}.$$

The integer m is called the *order* of $\hat{\phi}$.

Remark 3.1 If Φ has a generalized Fourier transform of order m, then it has also order $l \geqslant m$. If $\Phi \in \mathrm{L}_2(\mathbb{R}^d) \cap \mathrm{C}(\mathbb{R}^d)$, then its $\mathrm{L}_2(\mathbb{R}^d)$-Fourier transform is a generalized Fourier transform of any order.

Theorem 3.1 ([20, Theorem 8.12]) *Suppose an even function $\Phi \in \mathrm{Re}(\mathrm{C}(\mathbb{R}^d)) \cap \mathcal{SI}$ possesses a generalized Fourier transform $\hat{\phi}$ of order m which is continuous on $\mathbb{R}^d \setminus \{0\}$. Then Φ is conditionally positive definite of order m if and only if $\hat{\phi}$ is nonnegative and nonvanishing.*

3.2 Native Space and Reproducing-Kernel Hilbert Space

If Φ is conditionally positive definite of order m, then the linear space

$$F_\Phi := \left\{ \sum_{j=1}^{N} c_j \Phi(\cdot - \boldsymbol{x}_j) : N \in \mathbb{N}, \boldsymbol{c} \in \mathbb{R}^N, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_N \in \mathbb{R}^d, \right.$$
$$\left. \text{with } \sum_{j=1}^{N} c_j p(\boldsymbol{x}_j) = 0 \text{ for all } p \in \pi_{m-1}(\mathbb{R}^d) \right\}$$

can become a pre-Hilbert space by introduction of the inner product

$$\left(\sum_{k=1}^{M} b_k \Phi(\cdot - \boldsymbol{x}_k), \sum_{j=1}^{N} c_j \Phi(\cdot - \boldsymbol{x}_j)\right)_{\Phi} := \sum_{k=1}^{M}\sum_{j=1}^{N} b_k c_j \Phi(\boldsymbol{x}_k - \boldsymbol{x}_j).$$

Hence, we can form the Hilbert-space completion $\mathcal{F}_\Phi$ of the pre-Hilbert space F_Φ with respect to the Φ-inner product.

Now we construct the so-called native space of Φ. Let $\Xi := \{\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_Q\} \subset \mathbb{R}^d$ be a $\pi_{m-1}(\mathbb{R}^d)$-unisolvent set and $\{q_1, \cdots, q_Q\}$ be a Lagrange basis of $\pi_{m-1}(\mathbb{R}^d)$ with respect to Ξ, where $Q = \dim \pi_{m-1}(\mathbb{R}^d)$. First we set up the injective linear mapping

$$R : \mathcal{F}_\Phi \to \mathrm{Re}(\mathrm{C}(\mathbb{R}^d)), \quad R(f)(\boldsymbol{x}) := (f, W(\cdot - \boldsymbol{x}))_\Phi,$$

where $W(\cdot - \boldsymbol{x}) := \Phi(\cdot - \boldsymbol{x}) - \sum_{k=1}^{Q} q_k(\boldsymbol{x})\Phi(\cdot - \boldsymbol{\xi}_k)$. We also need the projection operator

$$\Pi : \mathrm{Re}(\mathrm{C}(\mathbb{R}^d)) \to \pi_{m-1}(\mathbb{R}^d), \quad \Pi(f) := \sum_{k=1}^{Q} f(\boldsymbol{\xi}_k) q_k.$$

Definition 3.3 ([20, Definition 10.16]) Suppose that Φ is conditionally positive definite of order $m \in \mathbb{N}_0$. Then the *native space* corresponding to Φ with respect to $\pi_{m-1}(\mathbb{R}^d)$ is defined by

$$\mathcal{N}_\Phi^m(\mathbb{R}^d) := R(\mathcal{F}_\Phi) + \pi_{m-1}(\mathbb{R}^d)$$

and it is equipped with the semi-inner product

$$(f, g)_{\mathcal{N}_\Phi^m(\mathbb{R}^d)} := \left(R^{-1}\left(f - \Pi(f)\right), R^{-1}\left(g - \Pi(g)\right)\right)_\Phi, \quad f, g \in \mathcal{N}_\Phi^m(\mathbb{R}^d).$$

Since $\mathcal{F}_\Phi$ is a Hilbert space with respect to the Φ-inner product, the native space $\mathcal{N}_\Phi^m(\mathbb{R}^d)$ is complete meaning that every Cauchy sequence has a (not necessarily unique) limit. We can verify that $F_\Phi \subseteq \mathcal{N}_\Phi^m(\mathbb{R}^d)$ so that $|f|_{\mathcal{N}_\Phi^m(\mathbb{R}^d)} = \|f\|_\Phi$ whenever $f \in F_\Phi$. The null space of $\mathcal{N}_\Phi^m(\mathbb{R}^d)$ is given by $\pi_{m-1}(\mathbb{R}^d)$, i.e., $|p|_{\mathcal{N}_\Phi^m(\mathbb{R}^d)} = 0$ if and only if $p \in \pi_{m-1}(\mathbb{R}^d) \subseteq \mathcal{N}_\Phi^m(\mathbb{R}^d)$. Moreover, according to [20, Theorem 10.20], $\mathcal{N}_\Phi^m(\mathbb{R}^d)$ will become a *reproducing-kernel Hilbert space* with the new inner product

$$(f, g)_K := (f, g)_{\mathcal{N}_\Phi^m(\mathbb{R}^d)} + \sum_{k=1}^{Q} f(\boldsymbol{\xi}_k) g(\boldsymbol{\xi}_k), \quad f, g \in \mathcal{N}_\Phi^m(\mathbb{R}^d),$$

and its *reproducing kernel* is

$$\begin{aligned} K(\boldsymbol{x}, \boldsymbol{y}) :=& \Phi(\boldsymbol{x} - \boldsymbol{y}) - \sum_{k=1}^{Q} q_k(\boldsymbol{x})\Phi(\boldsymbol{\xi}_k - \boldsymbol{y}) - \sum_{l=1}^{Q} q_l(\boldsymbol{y})\Phi(\boldsymbol{x} - \boldsymbol{\xi}_l) \\ &+ \sum_{k=1}^{Q}\sum_{l=1}^{Q} q_k(\boldsymbol{x}) q_l(\boldsymbol{y}) \Phi(\boldsymbol{\xi}_k - \boldsymbol{\xi}_l) + \sum_{k=1}^{Q} q_k(\boldsymbol{x}) q_k(\boldsymbol{y}). \end{aligned}$$

This means that $K(\cdot, \boldsymbol{y}) \in \mathcal{N}_\Phi^m(\mathbb{R}^d)$ for each $\boldsymbol{y} \in \mathbb{R}^d$ and

$$(f, K(\cdot, \boldsymbol{y}))_K = f(\boldsymbol{y}), \quad \text{for each } f \in \mathcal{N}_\Phi^m(\mathbb{R}^d) \text{ and } \boldsymbol{y} \in \mathbb{R}^d.$$

Theorem 3.2 **([20, Theorem 10.21])** *Suppose that Φ is a conditionally positive definite function of order $m \in \mathbb{N}_0$. Further suppose that Φ has a generalized Fourier transform $\hat{\phi}$ of order m which is continuous on $\mathbb{R}^d \setminus \{0\}$. Then its native space is characterized by*

$$\mathcal{N}_\Phi^m(\mathbb{R}^d) = \Big\{ f \in \mathrm{Re}(\mathrm{C}(\mathbb{R}^d)) \cap \mathcal{SI} : f \textit{ has a generalized Fourier transform } \hat{f} \\ \textit{of order } m/2 \textit{ such that } \hat{\phi}^{-1/2}\hat{f} \in \mathrm{L}_2(\mathbb{R}^d) \Big\},$$

and its semi-inner product satisfies

$$(f, g)_{\mathcal{N}_\Phi^m(\mathbb{R}^d)} = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \frac{\hat{f}(\boldsymbol{x})\overline{\hat{g}(\boldsymbol{x})}}{\hat{\phi}(\boldsymbol{x})} \mathrm{d}\boldsymbol{x}, \quad f, g \in \mathcal{N}_\Phi^m(\mathbb{R}^d).$$

All further details of the above results are discussed in [20, Chapter 10.3].

4 Green Functions and Generalized Sobolev Space Connected to Conditionally Positive Definite Functions and Native Space

In this section, we will generalize the idea of using differential operators [21] to generate a generalized Sobolev space to distributional operators. After this is done, we will be able to extend all results for positive definite functions and their reproducing-kernel Hilbert space mentioned in [21] to conditionally positive definite functions of some order and their native space. For the benefit of the readers familiar with [21] we use the same notation here whenever possible. However, to keep the current paper self-contained, we will introduce some of that notation again in the following sections.

4.1 Distributional Operators and Distributional Adjoint Operators

First, we define a metric ρ on the Schwartz space $\mathcal{S}$ so that it becomes a Fréchet space, where the metric ρ is given by

$$\rho(\gamma_1, \gamma_2) := \sum_{\alpha,\beta \in \mathbb{N}_0^d} 2^{-|\alpha|-|\beta|} \frac{\rho_{\alpha\beta}(\gamma_1 - \gamma_2)}{1 + \rho_{\alpha\beta}(\gamma_1 - \gamma_2)}, \quad \rho_{\alpha\beta}(\gamma) := \sup_{\boldsymbol{x} \in \mathbb{R}^d} \left| \boldsymbol{x}^\beta D^\alpha \gamma(\boldsymbol{x}) \right|,$$

for each $\gamma_1, \gamma_2, \gamma \in \mathcal{S}$. This means that a sequence $\{\gamma_n\}_{n=1}^\infty$ of $\mathcal{S}$ converges to an element $\gamma \in \mathcal{S}$ if and only if $\boldsymbol{x}^\beta D^\alpha \gamma_n(\boldsymbol{x})$ converges uniformly to $\boldsymbol{x}^\beta D^\alpha \gamma(\boldsymbol{x})$ on $\mathbb{R}^d$ for each $\alpha, \beta \in \mathbb{N}_0^d$. Together with its metric ρ the Schwartz space $\mathcal{S}$ is regarded as the classical test function space.

Let $\mathcal{S}'$ be the space of tempered distributions associated with $\mathcal{S}$ (the dual space of $\mathcal{S}$, or space of continuous linear functionals on $\mathcal{S}$). We introduce the notation

$$\langle T, \gamma \rangle := T(\gamma), \quad \text{for each } T \in \mathcal{S}' \text{ and } \gamma \in \mathcal{S}.$$

For any $f, g \in \mathrm{L}_1^{loc}(\mathbb{R}^d)$ whose product fg is integrable on $\mathbb{R}^d$ we define a bilinear form by

$$(f, g) := \int_{\mathbb{R}^d} f(\boldsymbol{x})g(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.$$

If $f, g \in \mathrm{L}_2(\mathbb{R}^d)$ then $(f, \overline{g})$ is equal to the standard $\mathrm{L}_2(\mathbb{R}^d)$-inner product.

For each $f \in \mathrm{L}_1^{loc}(\mathbb{R}^d) \cap \mathcal{SI}$ there exists a unique tempered distribution $T_f \in \mathcal{S}'$ such that

$$\langle T_f, \gamma \rangle = (f, \gamma), \quad \text{for each } \gamma \in \mathcal{S}.$$

So $f \in \mathrm{L}_1^{loc}(\mathbb{R}^d) \cap \mathcal{SI}$ can be viewed as an element of $\mathcal{S}'$ and we rewrite $T_f := f$. This means that $\mathrm{L}_1^{loc}(\mathbb{R}^d) \cap \mathcal{SI}$ can be embedded into $\mathcal{S}'$, i.e., $\mathrm{L}_1^{loc}(\mathbb{R}^d) \cap \mathcal{SI} \subseteq \mathcal{S}'$. The Dirac delta function (Dirac distribution) δ_0 concentrated at the origin is also an element of $\mathcal{S}'$, i.e., $\langle \delta_0, \gamma \rangle = \gamma(0)$ for each $\gamma \in \mathcal{S}$ (see [17, Chapter 1] and [7, Chapter 11]).

Given a linear operator $P : \mathcal{S}' \to \mathcal{S}'$, is it always possible to define a linear (adjoint) operator $P^* : \mathcal{S}' \to \mathcal{S}'$ which also satisfies the usual adjoint properties? The answer to this question is that it may be not be possible for all P. However, adjoint operators are well-defined for certain special linear operators. We will refer to these special linear operators as *distributional operators* and to their adjoint operators as *distributional adjoint operators* in this article.

We first introduce these linear operators on $\mathcal{S}'$. Let $\mathcal{P}^* : \mathcal{S} \to \mathcal{S}$ be a continuous linear operator. Then a linear operator $P : \mathcal{S}' \to \mathcal{S}'$ induced by $\mathcal{P}^*$ can be denote via the form

$$\langle PT, \gamma \rangle := \langle T, \mathcal{P}^* \gamma \rangle, \quad \text{for each } T \in \mathcal{S}' \text{ and } \gamma \in \mathcal{S}.$$

Furthermore, if $P|_{\mathcal{S}}$ is a continuous operator from $\mathcal{S}$ into $\mathcal{S}$, i.e., $\{P\gamma : \gamma \in \mathcal{S}\} \subseteq \mathcal{S}$ and $\rho(P\gamma_n, P\gamma) \to 0$ when $\rho(\gamma_n, \gamma) \to 0$, then we call the linear operator P a *distributional operator*.

Next we will show that the adjoint operators of these distributional operators are well-defined in the following way. In the same manner as before, we can denote another linear operator $P^* : \mathcal{S}' \to \mathcal{S}'$ induced by $P|_{\mathcal{S}}$, i.e.,

$$\langle P^* T, \gamma \rangle := \langle T, P|_{\mathcal{S}} \gamma \rangle = \langle T, P\gamma \rangle, \quad \text{for each } T \in \mathcal{S}' \text{ and } \gamma \in \mathcal{S}.$$

Fixing any $\tilde{\gamma} \in \mathcal{S}$, we have

$$\langle P^* \tilde{\gamma}, \gamma \rangle = \langle \tilde{\gamma}, P\gamma \rangle = (\tilde{\gamma}, P\gamma) = \langle P\gamma, \tilde{\gamma} \rangle = \langle \gamma, \mathcal{P}^* \tilde{\gamma} \rangle = (\gamma, \mathcal{P}^* \tilde{\gamma}) = \langle \mathcal{P}^* \tilde{\gamma}, \gamma \rangle,$$

for each $\gamma \in \mathcal{S}$ which implies that $P^* \tilde{\gamma} = \mathcal{P}^* \tilde{\gamma}$. Hence $P^*|_{\mathcal{S}} = \mathcal{P}^*$ on $\mathcal{S}$ and $P^*|_{\mathcal{S}}$ is also a continuous operator from $\mathcal{S}$ into $\mathcal{S}$. Therefore P^* is also a distributional operator. This motivates us to call P^* the *distributional adjoint operator* of P. According to the above definition, P is also the distributional adjoint operator of the distributional operator P^*.

We can summarize the definitions of the distributional operator and its adjoint operator as below.

Definition 4.1 Let $P, P^* : \mathcal{S}' \to \mathcal{S}'$ be two linear operators. If $P|_{\mathcal{S}}$ and $P^*|_{\mathcal{S}}$ are continuous operators from $\mathcal{S}$ into $\mathcal{S}$ such that

$$\langle PT, \gamma\rangle := \langle T, P^*\gamma\rangle \;\text{ and }\; \langle P^*T, \gamma\rangle := \langle T, P\gamma\rangle, \quad \text{for each } T \in \mathcal{S}' \text{ and } \gamma \in \mathcal{S},$$

then P and P^* are said to be *distributional operators* and, moreover, P^* (or P) is called a *distributional adjoint operator* of P (or P^*).

We will simplify the term distributional adjoint operator to adjoint operator in this article.

Remark 4.1 Our distributional adjoint operator differs from the adjoint operator of a bounded operator defined in Hilbert space or Banach space. Our operator is defined in the dual space of the Schwartz space and it may not be a bounded operator. However, since the fundamental idea of our construction is similar to the classical ones we also call this an adjoint.

If $P = P^*$, then we call P *self-adjoint*. A distributional operator P is called *translation invariant* if

$$\tau_h P\gamma = P\tau_h\gamma, \quad \text{for each } h \in \mathbb{R}^d \text{ and } \gamma \in \mathcal{S},$$

where τ_h is defined by $\tau_h\gamma(\boldsymbol{x}) := \gamma(\boldsymbol{x} - h)$. A distributional operator is called *complex-adjoint invariant* if

$$\overline{P\gamma} = P\overline{\gamma}, \quad \text{for each } \gamma \in \mathcal{S}.$$

Now we will show that the *differential operator* used in [21, Definition 3] is a special kind of distributional operator. Roughly speaking, the differential operator is a linear combination of *distributional derivatives* [21] induced by the classical derivatives

$$D^\alpha := \prod_{k=1}^{d} \frac{\partial^{\alpha_k}}{\partial x_k^{\alpha_k}}, \quad |\alpha| := \sum_{k=1}^{d} \alpha_k, \quad \alpha := (\alpha_1, \cdots, \alpha_d)^T \in \mathbb{N}_0^d.$$

For convenience, we also denote the distributional derivative as $P := D^\alpha : \mathcal{S}' \to \mathcal{S}'$. According to [17, Chapter 1] and [21, Section 4.1], $P|_{\mathcal{S}} = D^\alpha|_{\mathcal{S}}$ is a continuous linear operator from $\mathcal{S}$ into $\mathcal{S}$ and the distributional derivative also has the property

$$\langle D^\alpha T, \gamma\rangle := (-1)^{|\alpha|}\langle T, D^\alpha\gamma\rangle, \quad \text{for each } T \in \mathcal{S}' \text{ and } \gamma \in \mathcal{S}.$$

Therefore the distributional derivative $P = D^\alpha$ is a distributional operator and its distributional adjoint operator is given by $P^* = (-1)^{|\alpha|}D^\alpha$. Moreover, if $f \in \mathrm{L}_1^{loc}(\mathbb{R}^d) \cap \mathcal{SI}$ and $D^\alpha f \in \mathrm{L}_1^{loc}(\mathbb{R}^d) \cap \mathcal{SI}$ then we can verify that $D^\alpha f$ is equivalent to the *weak derivative* of f defined in [1] and [21]. (This shows that the classical real-valued Sobolev space $\mathcal{H}^n(\mathbb{R}^d)$ defined in this article and [21] is the same as in [1].) Hence we can determine that the differential operator

$$P := \sum_{|\alpha| \leqslant N} c_\alpha D^\alpha, \quad \text{where } c_\alpha \in \mathbb{C} \text{ and } \alpha \in \mathbb{N}_0^d,\ N \in \mathbb{N}_0,$$

is also a distributional operator and its distributional adjoint operator has the form

$$P^* := \sum_{|\alpha|\leqslant N} (-1)^{|\alpha|} c_\alpha D^\alpha.$$

Much more detail is mentioned in [21, Section 4.1].

Finally, we introduce the second kind of distributional operator which is defined for any fixed function

$$\hat{p} \in \mathcal{FT} := \left\{ f \in \mathrm{C}^\infty(\mathbb{R}^d) : D^\alpha f \in \mathcal{SI} \text{ for each } \alpha \in \mathbb{N}_0^d \right\}.$$

It is obvious that all complex-valued polynomials belong to $\mathcal{FT}$. Since $\hat{p}\gamma \in \mathcal{S}$ for each $\gamma \in \mathcal{S}$, we can verify that the linear operator $\gamma \mapsto \hat{p}\gamma$ is a continuous operator from $\mathcal{S}$ into $\mathcal{S}$. Thus we can define the distributional operator P related to $\hat{p}$ by

$$\langle PT, \gamma \rangle := \langle T, \hat{p}\gamma \rangle, \quad \text{for each } T \in \mathcal{S}' \text{ and } \gamma \in \mathcal{S}.$$

Then we call this kind of distributional operator an $\mathcal{FT}$ *operator*. We can further check that this operator is self-adjoint and $Pg = \hat{p}g \in \mathrm{L}_1^{loc}(\mathbb{R}^d) \cap \mathcal{SI}$ if $g \in \mathrm{L}_1^{loc}(\mathbb{R}^d) \cap \mathcal{SI}$. Therefore we use the notation $P := \hat{p}$ for convenience. The $\mathcal{FT}$ space is also applied to the definition of distributional Fourier transforms of distributional operators in Section 4.2.

4.2 Distributional Fourier Transforms

Following [7, Chapter 11] and [1, Chapter 7], we will work with the following definitions of the *Fourier transform* and *inverse Fourier transform* of any $\gamma \in \mathcal{S}$:

$$\hat{\gamma}(\boldsymbol{x}) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} \gamma(\boldsymbol{y}) e^{-i\boldsymbol{x}^T\boldsymbol{y}} \mathrm{d}\boldsymbol{y}, \quad \check{\gamma}(\boldsymbol{x}) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} \gamma(\boldsymbol{y}) e^{i\boldsymbol{x}^T\boldsymbol{y}} \mathrm{d}\boldsymbol{y},$$

where $i := \sqrt{-1}$. Since $\hat{\gamma}$ belongs to $\mathcal{S}$ for each $\gamma \in \mathcal{S}$ and the Fourier transform map is a homeomorphism of $\mathcal{S}$ onto itself we can define the *distributional Fourier transform* $\hat{T} \in \mathcal{S}'$ of the tempered distribution $T \in \mathcal{S}'$ by

$$\langle \hat{T}, \gamma \rangle := \langle T, \hat{\gamma} \rangle, \quad \text{for each } \gamma \in \mathcal{S}.$$

Since $\overline{\hat{\gamma}} = \check{\overline{\gamma}}$ for each $\gamma \in \mathcal{S}$, we have

$$\langle T, \overline{\gamma} \rangle = \langle \hat{T}, \overline{\hat{\gamma}} \rangle, \quad \text{for each } T \in \mathcal{S}' \text{ and } \gamma \in \mathcal{S}.$$

This implies that the Fourier transform of $\gamma \in \mathcal{S}$ is the same as its distributional transform. If $f \in \mathrm{L}_2(\mathbb{R}^d)$, then its $\mathrm{L}_2(\mathbb{R}^d)$-Fourier transform is equal to its distributional Fourier transform. The distributional Fourier transform $\hat{\delta}_0$ of the Dirac delta function δ_0 is equal to $(2\pi)^{-d/2}$. Moreover, we can check that the distributional Fourier transform map is an isomorphism of the topological vector space $\mathcal{S}'$ onto itself. This shows that the distributional Fourier transform map is also a distributional operator.

If $\Phi \in \mathrm{C}(\mathbb{R}^d) \cap \mathcal{SI}$ has the generalized Fourier transform $\hat{\phi}$ of order m, then its generalized Fourier transform and its distributional Fourier transform coincide on the set $\mathcal{S}_{2m}$, i.e.,

$$\langle \hat{\Phi}, \gamma \rangle = \langle \Phi, \hat{\gamma} \rangle = (\Phi, \hat{\gamma}) = (\hat{\phi}, \gamma), \quad \text{for each } \gamma \in \mathcal{S}_{2m}.$$

Even if Φ does not have any generalized Fourier transform, it always has a distributional Fourier transform $\hat{\Phi}$ since Φ can be seen as a tempered distribution (see [17, Chapter 1] and [7, Chapter 11]).

Our main goal in this subsection is to define the distributional Fourier transform of a distributional operator induced by the $\mathcal{FT}$ space introduced in Section 4.1.

Definition 4.2 Let P be a distributional operator. If there is a function $\hat{p} \in \mathcal{FT}$ such that

$$\langle \widehat{PT}, \gamma \rangle = \langle \hat{p}\hat{T}, \gamma \rangle = \langle \hat{T}, \hat{p}\gamma \rangle, \quad \text{for each } T \in \mathcal{S}' \text{ and } \gamma \in \mathcal{S},$$

then $\hat{p}$ is said to be a *distributional Fourier transform* of P.

If P has the distributional Fourier transform $\hat{p}$, then P is translation-invariant because $\widehat{\tau_h P\gamma}(\boldsymbol{x}) = e^{-i\boldsymbol{x}^T h}\hat{p}(\boldsymbol{x})\hat{\gamma}(\boldsymbol{x}) = \widehat{P\tau_h\gamma}(\boldsymbol{x})$ for each $h \in \mathbb{R}^d$ and $\gamma \in \mathcal{S}$. Moreover, if P is complex-adjoint invariant and has the distributional Fourier transform $\hat{p}$, then

$$\langle \overline{\hat{p}}\hat{T}, \overline{\gamma} \rangle = \langle \hat{T}, \overline{\hat{p}\hat{\check{\overline{\gamma}}}} \rangle = \langle \hat{T}, \overline{\widehat{P\check{\overline{\gamma}}}} \rangle = \langle T, \overline{P\check{\gamma}} \rangle = \langle T, P\overline{\check{\gamma}} \rangle = \langle P^* T, \overline{\check{\gamma}} \rangle = \langle \widehat{P^* T}, \overline{\gamma} \rangle$$

for each $T \in \mathcal{S}'$ and $\gamma \in \mathcal{S}$. This shows that $\overline{\hat{p}}$ is the distributional Fourier transform of the adjoint operator P^* of P.

Now we show that any distributional derivative D^α has the distributional Fourier transform $\hat{p}(\boldsymbol{x}) := (i\boldsymbol{x})^\alpha$ where $i = \sqrt{-1}$. According to [17, Chapter 1], we know that $D^\alpha\hat{\gamma} = \widehat{\left(\overline{\hat{p}}\gamma\right)}$ for each $\gamma \in \mathcal{S}$. So

$$\langle \widehat{D^\alpha T}, \gamma \rangle = \langle D^\alpha T, \hat{\gamma} \rangle = (-1)^{|\alpha|}\langle T, D^\alpha\hat{\gamma} \rangle = (-1)^{|\alpha|}\langle T, \widehat{\overline{\hat{p}}\gamma} \rangle = (-1)^{|\alpha|}\langle \hat{T}, \overline{\hat{p}}\gamma \rangle = \langle \hat{p}\hat{T}, \gamma \rangle,$$

for each $\gamma \in \mathcal{S}$. This also implies that the distributional Fourier transform $\hat{p}^*$ of its adjoint operator $(-1)^{|\alpha|}D^\alpha$ is equal to $\hat{p}^*(\boldsymbol{x}) = (-i\boldsymbol{x})^\alpha = \overline{\hat{p}(\boldsymbol{x})}$. Furthermore, we can also obtain the distributional Fourier transform of a differential operator in the same way, e.g.,

$$\hat{p}(\boldsymbol{x}) = \sum_{|\alpha| \leqslant N} c_\alpha (i\boldsymbol{x})^\alpha, \quad \text{where } P = \sum_{|\alpha| \leqslant N} c_\alpha D^\alpha,\ c_\alpha \in \mathbb{C} \text{ and } \alpha \in \mathbb{N}_0^d,\ N \in \mathbb{N}_0.$$

Finally, we can check that the distributional Fourier transform of the differential operator is equivalent to [21, Definition 4].

4.3 Green Functions and Generalized Sobolev Space

Definition 4.3 ([21, Definition 4]) G is the (full-space) *Green function with respect to the distributional operator* L if $G \in \mathcal{S}'$ satisfies the equation

$$LG = \delta_0. \tag{4.1}$$

Equation (4.1) is to be interpreted in the distribution sense which means that $\langle G, L^*\gamma\rangle = \langle LG, \gamma\rangle = \langle \delta_0, \gamma\rangle = \gamma(0)$ for each $\gamma \in \mathcal{S}$.

According to Theorem 3.1 and [10] we can obtain the following theorem.

Theorem 4.1 *Let L be a distributional operator with distributional Fourier transform $\hat{l}$. Suppose that $\hat{l}$ is positive on $\mathbb{R}^d \setminus \{0\}$. Further suppose that $\hat{l}^{-1} \in \mathcal{SI}$ and that $\hat{l}(\boldsymbol{x}) = \Theta(\|\boldsymbol{x}\|_2^{2m})$ as $\|\boldsymbol{x}\|_2 \to 0$ for some $m \in \mathbb{N}_0$. If the Green function $G \in$ $\mathrm{Re}(\mathrm{C}(\mathbb{R}^d)) \cap \mathcal{SI}$ with respect to L is an even function, then G is a conditionally positive definite function of order m and*

$$\hat{\mathsf{g}}(\boldsymbol{x}) := (2\pi)^{-d/2}\hat{l}(\boldsymbol{x})^{-1}, \quad \boldsymbol{x} \in \mathbb{R}^d.$$

is the generalized Fourier transform of order m of G. (Here the notation $f = \Theta(g)$ means that there are two positive numbers M_1 and M_2 such that $M_1 |g| \leqslant |f| \leqslant M_2 |g|$.)

Proof First we want to prove that $\hat{\mathsf{g}}$ is the generalized Fourier transform of order m of G. Since $\hat{l}^{-1} \in \mathcal{SI}$ and $\hat{l}(\boldsymbol{x}) = \Theta(\|\boldsymbol{x}\|_2^{2m})$ as $\|\boldsymbol{x}\|_2 \to 0$ for some $m \in \mathbb{N}_0$, the product $\hat{\mathsf{g}}\gamma$ is integrable for each $\gamma \in \mathcal{S}_{2m}$. So, if we let $\widehat{G}$ be the distributional Fourier transform of G and we can verify that

$$\langle \widehat{G}, \gamma\rangle = (\hat{\mathsf{g}}, \gamma), \quad \text{for each } \gamma \in \mathcal{S}_{2m},$$

then we can conclude that $\hat{\mathsf{g}}$ is the generalized Fourier transform of G.

Since $\hat{l}$ is the distributional Fourier transform of the distributional operator L we know that $\hat{l} \in \mathcal{FT}$. Thus $D^\alpha\left(\hat{l}^{-1}\right) \in \mathcal{SI}$ for each $\alpha \in \mathbb{N}_0^d$ because of $D^\alpha\hat{l} \in \mathcal{SI}$ and $\hat{l}^{-1} \in \mathcal{SI}$. If $\hat{l}(0) > 0$, then $\hat{l}^{-1} \in \mathcal{FT}$, which implies that $\hat{l}^{-1}\gamma \in \mathcal{S}$ for each fixed $\gamma \in \mathcal{S}_{2m}$. Hence

$$\begin{aligned}\langle \widehat{G}, \gamma\rangle &= \langle \hat{l}\widehat{G}, \hat{l}^{-1}\gamma\rangle = \langle \widehat{LG}, \hat{l}^{-1}\gamma\rangle = \langle \hat{\delta}_0, \hat{l}^{-1}\gamma\rangle \\ &= \langle (2\pi)^{-d/2}, \hat{l}^{-1}\gamma\rangle = ((2\pi)^{-d/2}, \hat{l}^{-1}\gamma) = (\hat{\mathsf{g}}, \gamma).\end{aligned}$$

If $\hat{l}(0) = 0$, then $\hat{l}^{-1}$ does not belong to $\mathcal{FT}$. However, since $\hat{l} \in \mathcal{FT}$ is positive on $\mathbb{R}^d \setminus \{0\}$ we can find a positive-valued sequence $\{\hat{l}_n\}_{n=1}^\infty \subset \mathrm{C}^\infty(\mathbb{R}^d)$ such that

$$\hat{l}_n(\boldsymbol{x}) = \begin{cases}\hat{l}(\boldsymbol{x}), & \|\boldsymbol{x}\|_2 > n^{-1}, \\ \hat{l}(\boldsymbol{x}) + n^{-1}, & \|\boldsymbol{x}\|_2 < n^{-2}.\end{cases}$$

In particular $l_1 \equiv 1$. And then $\{\hat{l}_n\}_{n=1}^\infty \subset \mathcal{FT}$. It further follows that $D^\alpha\hat{l}_n$ converges uniformly to $D^\alpha\hat{l}$ on $\mathbb{R}^d$ for each $\alpha \in \mathbb{N}_0^d$.

We now fix a arbitrary $\gamma \in \mathcal{S}_{2m}$. Since $\hat{l}_n^{-1}\gamma$ and $\hat{l}^{-1}\gamma$ have absolutely finite integral, $\hat{l}_n^{-1}\gamma$ also converges to $\hat{l}^{-1}\gamma$ in the integral sense. Let $\gamma_k := \hat{l}_k^{-1}\gamma$ for each fixed $k \in \mathbb{N}$. Since $\gamma_k \in \mathcal{S}$,

$$\lim_{n\to\infty}\langle \widehat{G}, \hat{l}_n\gamma_k\rangle = \langle \widehat{G}, \hat{l}\gamma_k\rangle = \langle \widehat{\hat{l}G}, \gamma_k\rangle = \langle \widehat{LG}, \gamma_k\rangle = \langle \hat{\delta}_0, \gamma_k\rangle = \langle (2\pi)^{-d/2}, \gamma_k\rangle.$$

Moreover, we can also obtain that $\langle \widehat{G}, \hat{l}_k\gamma_k\rangle = \langle \widehat{G}, \gamma\rangle$ for each $k \in \mathbb{N}$ and that

$$\lim_{k\to\infty}\langle (2\pi)^{-d/2}, \gamma_k\rangle = \lim_{k\to\infty}((2\pi)^{-d/2}, l_k^{-1}\gamma) = ((2\pi)^{-d/2}, \hat{l}^{-1}\gamma) = (\hat{\mathrm{g}}, \gamma).$$

By taking the limit at the diagonal as $k = n$, we have $\langle \widehat{G}, \gamma\rangle = (\hat{\mathrm{g}}, \gamma)$ for each $\gamma \in \mathcal{S}_{2m}$.

Since $\hat{\mathrm{g}} \in \mathrm{C}(\mathbb{R}^d \setminus \{0\})$ is positive on $\mathbb{R}^d \setminus \{0\}$ and $G \in \mathrm{Re}(\mathrm{C}(\mathbb{R}^d)) \cap \mathcal{SI}$ is an even function, we can use Theorem 3.1 to conclude that G is a conditionally positive definite function of order m.

□

Remark 4.2 If L is a differential operator, then its distributional Fourier transform $\hat{l}$ satisfies the conditions of Theorem 4.1 if and only if $\hat{l}$ has a polynomial of the form $\hat{l}(\boldsymbol{x}) := q(\boldsymbol{x}) + a_{2m}\,\|\boldsymbol{x}\|_2^{2m}$, where $a_{2m} > 0$ and q is a polynomial of degree greater than $2m$ so that it is positive on $\mathbb{R}^d \setminus \{0\}$, or $q \equiv 0$.

Now we can define the generalized Sobolev space induced by a vector distributional operator $\mathbf{P} = (P_1, \cdots, P_n, \cdots)^T$ similar as in [21, Definition 6].

Definition 4.4 Consider the vector distributional operator $\mathbf{P} = (P_1, \cdots, P_n, \cdots)^T$ consisting of countably many distributional operators $\{P_n\}_{n=1}^{\infty}$. The *generalized Sobolev space* induced by $\mathbf{P}$ is defined by

$$\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d) := \left\{ f \in \mathrm{Re}(\mathrm{L}_1^{loc}(\mathbb{R}^d)) \cap \mathcal{SI} : \{P_n f\}_{n=1}^{\infty} \subseteq \mathrm{L}_2(\mathbb{R}^d) \text{ and } \sum_{n=1}^{\infty} \|P_n f\|_{\mathrm{L}_2(\mathbb{R}^d)}^2 < \infty \right\}$$

and it is equipped with the semi-inner product

$$(f, g)_{\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)} := \sum_{n=1}^{\infty}(P_n f, \overline{P_n g}), \quad f, g \in \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d).$$

What is the meaning of $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$? By the definition of the generalized Sobolev space we know that $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ is a real-valued subspace of $\mathrm{L}_1^{loc}(\mathbb{R}^d) \cap \mathcal{SI}$ and it is equipped with a semi-inner product induced by the vector distributional operator $\mathbf{P}$. Viewed in another way, $f \in \mathrm{Re}(\mathrm{L}_1^{loc}(\mathbb{R}^d)) \cap \mathcal{SI}$ belongs to $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ if and only if there is a sequence $\{g_n\}_{n=1}^{\infty} \subset \mathrm{L}_2(\mathbb{R}^d)$ such that $\sum_{n=1}^{\infty}\|g_n\|_{\mathrm{L}_2(\mathbb{R}^d)}^2 < \infty$ and

$$(g_n, \gamma) = \langle g_n, \gamma\rangle = \langle P_n f, \gamma\rangle = \langle f, P_n^*\gamma\rangle = (f, P_n^*\gamma), \quad \text{for each } \gamma \in \mathcal{S},\ n \in \mathbb{N}.$$

In the following theorems of this section we only consider $\mathbf{P}$ constructed by a finite number of distributional operators $P_1, \ldots, P_n$. If $\mathbf{P} := (P_1, \cdots, P_n)^T$, then the distributional operator

$$L := \mathbf{P}^{*T}\mathbf{P} = \sum_{j=1}^{n} P_j^* P_j$$

is well-defined, where $\mathbf{P}^* := (P_1^*, \cdots, P_n^*)^T$ is the adjoint operator of $\mathbf{P}$ as defined in Section 4.1. If we suppose that $\mathbf{P}$ is complex-adjoint invariant with distributional Fourier transform $\hat{\mathbf{p}} = (\hat{p}_1, \cdots, \hat{p}_n)^T$, then the distributional Fourier transform $\hat{\mathbf{p}}^* = (\hat{p}_1^* \cdots, \hat{p}_n^*)^T$ of its adjoint operator $\mathbf{P}^*$ is equal to $\overline{\hat{\mathbf{p}}} = (\overline{\hat{p}}_1, \cdots, \overline{\hat{p}}_n)^T$. Since

$$\langle \widehat{P_j^* P_j T}, \gamma \rangle = \langle \hat{p}_j^* \widehat{P_j T}, \gamma \rangle = \langle \hat{p}_j \hat{T}, \hat{p}_j^* \gamma \rangle = \langle \overline{\hat{p}}_j \hat{p}_j \hat{T}, \gamma \rangle = \langle \left|\hat{p}_j\right|^2 \hat{T}, \gamma \rangle$$

for each $T \in \mathcal{S}'$ and $\gamma \in \mathcal{S}$, the distributional Fourier transform $\hat{l}$ of L is given by

$$\hat{l}(\boldsymbol{x}) := \sum_{j=1}^{n} \left|\hat{p}_j(\boldsymbol{x})\right|^2 = \|\hat{\mathbf{p}}(\boldsymbol{x})\|_2^2, \quad \boldsymbol{x} \in \mathbb{R}^d.$$

Moreover, since $\mathbf{P}$ has a distributional Fourier transform, $\mathbf{P}$ is translation invariant (see Section 4.2).

We are now ready to state and prove our main theorem about the generalized Sobolev space $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ induced by a vector distributional operator $\mathbf{P} := (P_1, \cdots, P_n)^T$.

Theorem 4.2 *Let $\mathbf{P} := (P_1, \cdots, P_n)^T$ be a complex-adjoint invariant vector distributional operator with vector distributional Fourier transform $\hat{\mathbf{p}} := (\hat{p}_1, \cdots, \hat{p}_n)^T$ which is nonzero on $\mathbb{R}^d \setminus \{0\}$. Further suppose that $\boldsymbol{x} \mapsto \|\hat{\mathbf{p}}(\boldsymbol{x})\|_2^{-1} \in \mathcal{SI}$ and that $\|\hat{\mathbf{p}}(\boldsymbol{x})\|_2 = \Theta(\|\boldsymbol{x}\|_2^m)$ as $\|\boldsymbol{x}\|_2 \to 0$ for some $m \in \mathbb{N}_0$. If the Green function $G \in \mathrm{Re}(\mathrm{C}(\mathbb{R}^d)) \cap \mathcal{SI}$ with respect to $L = \mathbf{P}^{*T}\mathbf{P}$ is chosen so that it is an even function, then G is a conditionally positive definite function of order m and its native space $\mathcal{N}_G^m(\mathbb{R}^d)$ is a subspace of the generalized Sobolev space $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$. Moreover, their semi-inner products are the same on $\mathcal{N}_G^m(\mathbb{R}^d)$, i.e.,*

$$(f, g)_{\mathcal{N}_G^m(\mathbb{R}^d)} = (f, g)_{\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)}, \quad f, g \in \mathcal{N}_G^m(\mathbb{R}^d).$$

Proof By our earlier discussion the distributional Fourier transform $\hat{l}$ of L is equal to $\hat{l}(\boldsymbol{x}) = \|\hat{\mathbf{p}}(\boldsymbol{x})\|_2^2$. Thus $\hat{l}$ is positive on $\mathbb{R}^d \setminus \{0\}$, $\hat{l}^{-1} \in \mathcal{SI}$ and $\hat{l}(\boldsymbol{x}) = \Theta(\|\boldsymbol{x}\|_2^{2m})$ as $\|\boldsymbol{x}\|_2 \to 0$. According to Theorem 4.1, G is a conditionally positive definite function of order m and its generalized Fourier transform of order m is given by

$$\hat{\mathsf{g}}(\boldsymbol{x}) := (2\pi)^{-d/2}\hat{l}(\boldsymbol{x})^{-1} = (2\pi)^{-d/2} \|\hat{\mathbf{p}}(\boldsymbol{x})\|_2^{-2}, \quad \boldsymbol{x} \in \mathbb{R}^d.$$

With the material developed thus far we are able construct its native space $\mathcal{N}_G^m(\mathbb{R}^d)$ (see Section 3.2).

Next, we fix any $f \in \mathcal{N}_G^m(\mathbb{R}^d)$. According to Theorem 3.2, $f \in \mathrm{Re}(\mathrm{C}(\mathbb{R}^d)) \cap \mathcal{SI}$ possesses a generalized Fourier transform $\hat{f}$ of order $m/2$ and $\boldsymbol{x} \mapsto \hat{f}(\boldsymbol{x}) \|\hat{\mathbf{p}}(\boldsymbol{x})\|_2 \in \mathrm{L}_2(\mathbb{R}^d)$. This means that the functions $\hat{p}_j\hat{f}$ belong to $\mathrm{L}_2(\mathbb{R}^d)$, $j = 1, \ldots, n$. Hence we can define the function $f_{P_j} \in \mathrm{L}_2(\mathbb{R}^d)$ by

$$f_{P_j} := (\hat{p}_j\hat{f})^{\vee} \in \mathrm{L}_2(\mathbb{R}^d), \quad j = 1, \ldots, n$$

using the inverse $\mathrm{L}_2(\mathbb{R}^d)$-Fourier transform.

Since $\|\hat{\mathbf{p}}(\boldsymbol{x})\|_2 = \Theta(\|\boldsymbol{x}\|_2^m)$ as $\|\boldsymbol{x}\|_2 \to 0$ we have $\hat{p}_j(\boldsymbol{x}) = O(\|\boldsymbol{x}\|_2^m)$ as $\|\boldsymbol{x}\|_2 \to 0$ for each $j = 1, \ldots, n$. Thus $\hat{p}_j\check{\overline{\gamma}} \in \mathcal{S}_m$ for each $\gamma \in \mathcal{S}$. Moreover, since $\hat{p}_j\check{\overline{\gamma}} = \hat{p}_j\overline{\hat{\gamma}} =$

$\overline{\hat{p}_j^*\hat{\gamma}} = \overline{\widehat{P_j^*\gamma}}$ and the generalized and distributional Fourier transforms of f coincide on $\mathcal{S}_m$ we have

$$\begin{aligned}(f_{P_j}, \overline{\gamma}) &= ((\hat{p}_j\hat{f})\check{}, \overline{\gamma}) = (\hat{p}_j\hat{f}, \check{\overline{\gamma}}) = (\hat{f}, \hat{p}_j\check{\overline{\gamma}}) = \langle \hat{f}, \hat{p}_j\check{\overline{\gamma}}\rangle \\ &= \langle \hat{f}, \overline{\widehat{P_j^*\gamma}}\rangle = \langle f, \overline{P_j^*\gamma}\rangle = \langle f, P_j^*\overline{\gamma}\rangle = \langle P_j f, \overline{\gamma}\rangle, \quad \gamma \in \mathcal{S}.\end{aligned}$$

This shows that $P_j f = f_{P_j} \in \mathrm{L}_2(\mathbb{R}^d)$. Therefore we know that $f \in \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$.

To establish equality of the semi-inner products we let $f, g \in \mathcal{N}_G^m(\mathbb{R}^d)$. Then the Plancherel theorem [7] yields

$$\begin{aligned}(f,g)_{\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)} &= \sum_{j=1}^{n}(f_{P_j}, \overline{g_{P_j}}) = \sum_{j=1}^{n}(\hat{p}_j\hat{f}, \overline{\hat{p}_j\hat{g}}) = \sum_{j=1}^{n}\int_{\mathbb{R}^d}\hat{f}(\boldsymbol{x})\overline{\hat{g}(\boldsymbol{x})}\left|\hat{p}_j(\boldsymbol{x})\right|^2 \mathrm{d}\boldsymbol{x} \\ &= \int_{\mathbb{R}^d}\hat{f}(\boldsymbol{x})\overline{\hat{g}(\boldsymbol{x})}\,\|\hat{\mathbf{p}}(\boldsymbol{x})\|_2^2\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^d}\hat{f}(\boldsymbol{x})\overline{\hat{g}(\boldsymbol{x})}\hat{l}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ &= (2\pi)^{-d/2}\int_{\mathbb{R}^d}\frac{\hat{f}(\boldsymbol{x})\overline{\hat{g}(\boldsymbol{x})}}{\hat{\mathsf{g}}(\boldsymbol{x})}\mathrm{d}\boldsymbol{x} = (f,g)_{\mathcal{N}_G^m(\mathbb{R}^d)}.\end{aligned}$$

□

Remark 4.3 If each element of $\mathbf{P}$ is just a differential operator then each element of $\mathbf{P}$ can be written as $P_j := \sum_{|\alpha|\leqslant N} c_\alpha D^\alpha$ where $c_\alpha \in \mathbb{R}$ and $\alpha \in \mathbb{N}_0^d$, $N \in \mathbb{N}_0$ because $\mathbf{P}$ is complex-adjoint invariant.

The preceding theorem shows that $\mathcal{N}_G^m(\mathbb{R}^d)$ can be imbedded into $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$. Ideally, $\mathcal{N}_G^m(\mathbb{R}^d)$ would be equal to $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$, but this is not true in general. However, if we impose some additional conditions on $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$, then we can obtain equality.

Definition 4.5 Let $\mathbf{P} := (P_1, \cdots, P_n)^T$ be a vector distributional operator. We say that the generalized Sobolev space $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ possesses the $\mathcal{S}$*-dense* property if for every $f \in \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$, every compact subset $\Lambda \subset \mathbb{R}^d$ and every $\epsilon > 0$, there exists $\gamma \in \mathrm{Re}(\mathcal{S}) \subseteq \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ such that

$$|f-\gamma|_{\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)} < \epsilon \text{ and } \|f-\gamma\|_{\mathrm{L}_\infty(\Lambda)} < \epsilon, \tag{4.2}$$

i.e., there is a sequence $\{\gamma_n\}_{n=1}^\infty \subseteq \mathrm{Re}(\mathcal{S}) \subseteq \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ so that

$$|f-\gamma_n|_{\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)} \to 0 \text{ and } \|f-\gamma_n\|_{\mathrm{L}_\infty(\Lambda)} \to 0, \text{ when } n \to \infty.$$

Following the method of the proofs of [20, Theorems 10.41 and 10.43], we can complete the proofs of the following lemma and theorem.

Lemma 4.3 *Let* $\mathbf{P}$ *and* G *satisfy the conditions of Theorem 4.2 and suppose that* $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ *has the* $\mathcal{S}$*-dense property. Assume we are given the data sets* $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N\} \subset \mathbb{R}^d$ *and* $\{\lambda_1, \cdots, \lambda_N\} \subset \mathbb{R}$*. If we define* $f_\lambda := \sum_{k=1}^N \lambda_k G(\cdot - \boldsymbol{x}_k)$*, then for every* $f \in \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ *and every* $\boldsymbol{x} \in \mathbb{R}^d$ *we have the representation*

$$(f_\lambda(\boldsymbol{x}-\cdot), f)_{\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)} = \sum_{k=1}^{N}\lambda_k f(\boldsymbol{x}-\boldsymbol{x}_k). \tag{4.3}$$

Proof Let us first assume that $\gamma \in \mathrm{Re}(\mathcal{S})$. According to Theorem 4.2, $f_\lambda \in \mathcal{N}_G^m(\mathbb{R}^d) \subseteq \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$. Since $\mathbf{P}$ is translation invariant and complex-adjoint invariant we have

$$(f_\lambda(\boldsymbol{x}-\cdot),\gamma)_{\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)} = \sum_{j=1}^{n}(P_j f_\lambda(\boldsymbol{x}-\cdot),\overline{P_j\gamma}) = \sum_{j=1}^{n}(P_j f_\lambda(\boldsymbol{x}-\cdot),P_j\gamma)$$
$$=\sum_{j=1}^{n}\langle f_\lambda(\boldsymbol{x}-\cdot),P_j^*P_j\gamma\rangle = (f_\lambda,\mathbf{P}^{*T}\mathbf{P}\gamma(\boldsymbol{x}-\cdot)) = \sum_{k=1}^{N}\lambda_k(G(\cdot-\boldsymbol{x}_k),L\gamma(\boldsymbol{x}-\cdot))$$
$$=\sum_{k=1}^{N}\lambda_k\langle LG,\gamma(\boldsymbol{x}-\boldsymbol{x}_k-\cdot)\rangle = \sum_{k=1}^{N}\lambda_k\langle\delta_0,\gamma(\boldsymbol{x}-\boldsymbol{x}_k-\cdot)\rangle = \sum_{k=1}^{N}\lambda_k\gamma(\boldsymbol{x}-\boldsymbol{x}_k).$$

For a general $f \in \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ we fix $\boldsymbol{x} \in \mathbb{R}^d$ and choose a compact set $\Lambda \subset \mathbb{R}^d$ such that $\boldsymbol{x}-\boldsymbol{x}_k \in \Lambda$ for $k = 1,\ldots,N$. For any $\epsilon > 0$, there is a $\gamma \in \mathrm{Re}(\mathcal{S})$ which satisfies Equation (4.2). Then two applications of the triangle inequality show that the absolute value of the difference in the two sides of Equation (4.3) can be bounded by $\epsilon\left(\sum_{k=1}^{N}|\lambda_k| + |f_\lambda|_{\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)}\right)$, which tends to zero as $\epsilon \to 0$.
□

Theorem 4.4 *Let* $\mathbf{P}$ *and* G *satisfy the conditions of Theorem 4.2. If* $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ *possesses the* $\mathcal{S}$*-dense property, then*

$$\mathcal{N}_G^m(\mathbb{R}^d) \equiv \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d).$$

Proof By Theorem 4.2 we already know that $\mathcal{N}_G^m(\mathbb{R}^d)$ is contained in $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ and that their semi-inner products are the same in the subspace $\mathcal{N}_G^m(\mathbb{R}^d)$. Moreover, $\mathcal{N}_G^m(\mathbb{R}^d)$ is a complete subspace of $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$. So, if we assume that $\mathcal{N}_G^m(\mathbb{R}^d)$ were not the whole space $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$, then there would be an element $f \in \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ which is orthogonal to the native space $\mathcal{N}_G^m(\mathbb{R}^d)$.

Let $Q = \dim \pi_{m-1}(\mathbb{R}^d)$ and $\{q_1,\cdots,q_Q\}$ be a Lagrange basis of $\pi_{m-1}(\mathbb{R}^d)$ with respect to a $\pi_{m-1}(\mathbb{R}^d)$-unisolvent subset $\{\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_Q\} \subset \mathbb{R}^d$. We make the special choice of the data sets $\{-\boldsymbol{x},-\boldsymbol{\xi}_1,\cdots,-\boldsymbol{\xi}_Q\}$ and $\{1,-q_1(\boldsymbol{x}),\cdots,-q_Q(\boldsymbol{x})\}$ and correspondingly define

$$f_\lambda := G(\cdot+\boldsymbol{x}) - \sum_{k=1}^{Q} q_k(\boldsymbol{x})G(\cdot+\boldsymbol{\xi}_k).$$

Since $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ has the $\mathcal{S}$-dense property we can use Lemma 4.3 to represent any $f \in \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ in the form

$$f(\boldsymbol{w}+\boldsymbol{x}) = \sum_{k=1}^{Q} q_k(\boldsymbol{x})f(\boldsymbol{w}+\xi_k) + (f_\lambda(\boldsymbol{w}-\cdot),f)_{\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)}.$$

Since G is even, we have $\boldsymbol{x} \mapsto f_\lambda(-\boldsymbol{x}) \in \mathcal{N}_G^m(\mathbb{R}^d)$. We now set $\boldsymbol{w} = 0$. The fact that f is orthogonal to $\mathcal{N}_G^m(\mathbb{R}^d)$ gives us

$$f(\boldsymbol{x}) = \sum_{k=1}^{Q} q_k(\boldsymbol{x})f(\xi_k) + (f_\lambda(-\cdot),f)_{\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)} = \sum_{k=1}^{Q} f(\xi_k)q_k(\boldsymbol{x}).$$

This shows that $f \in \pi_{m-1}(\mathbb{R}^d) \subseteq \mathcal{N}_G^m(\mathbb{R}^d)$, and it contradicts our first assumption. It follows that $\mathcal{N}_G^m(\mathbb{R}^d) \equiv \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$.

□

Lemma 4.5 *Let* **P** *and G satisfy the conditions of Theorem 4.2. Then*

$$\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d) \cap \mathrm{C}(\mathbb{R}^d) \cap \mathrm{L}_2(\mathbb{R}^d) \subseteq \mathcal{N}_G^m(\mathbb{R}^d).$$

Proof We fix any $f \in \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d) \cap \mathrm{C}(\mathbb{R}^d) \cap \mathrm{L}_2(\mathbb{R}^d)$ and suppose that $\hat{f}$ and $\widehat{P_j f}$, respectively, are the $\mathrm{L}_2(\mathbb{R}^d)$-Fourier transforms of f and $P_j f$, $j = 1, \ldots, n$. Using the Plancherel theorem [7] we obtain

$$\int_{\mathbb{R}^d} \left|\hat{f}(\boldsymbol{x})\hat{p}_j(\boldsymbol{x})\right|^2 \mathrm{d}\boldsymbol{x} = (\hat{p}_j\hat{f}, \overline{\hat{p}_j\hat{f}}) = (\widehat{P_j f}, \overline{\widehat{P_j f}}) = (P_j f, \overline{P_j f}) < \infty.$$

And therefore, with the help of the proof of Theorem 4.2, we have

$$\begin{aligned}\int_{\mathbb{R}^d} \frac{\left|\hat{f}(\boldsymbol{x})\right|^2}{\hat{\mathsf{g}}(\boldsymbol{x})}\mathrm{d}\boldsymbol{x} &= (2\pi)^{d/2}\int_{\mathbb{R}^d} \left|\hat{f}(\boldsymbol{x})\right|^2 \hat{l}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = (2\pi)^{d/2}\int_{\mathbb{R}^d} \left|\hat{f}(\boldsymbol{x})\right|^2 \|\hat{\mathbf{p}}(\boldsymbol{x})\|_2^2\, \mathrm{d}\boldsymbol{x} \\ &= (2\pi)^{d/2}\sum_{j=1}^{n}\int_{\mathbb{R}^d} \left|\hat{f}(\boldsymbol{x})\hat{p}_j(\boldsymbol{x})\right|^2 \mathrm{d}\boldsymbol{x} < \infty\end{aligned}$$

showing that $\hat{\mathsf{g}}^{-1/2}\hat{f} \in \mathrm{L}_2(\mathbb{R}^d)$, where $\hat{\mathsf{g}}$ is the generalized Fourier transform of G. And now, according to Theorem 3.1, $f \in \mathcal{N}_G^m(\mathbb{R}^d)$.

□

This says that $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d) \cap \mathrm{C}(\mathbb{R}^d) \cap \mathrm{L}_2(\mathbb{R}^d)$ can be imbedded into $\mathcal{N}_G^m(\mathbb{R}^d)$. Moreover, according to Lemma 4.5 we can immediately obtain the following theorem.

Theorem 4.6 *Let* **P** *and G satisfy the conditions of Theorem 4.2. If* $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d) \subseteq \mathrm{C}(\mathbb{R}^d) \cap \mathrm{L}_2(\mathbb{R}^d)$*, then*

$$\mathcal{N}_G^m(\mathbb{R}^d) \equiv \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d).$$

Remark 4.4 As Example 5.5 in Section 5.2 shows, the native space $\mathcal{N}_G^m(\mathbb{R}^d)$ will not always be equal to the corresponding generalized Sobolev space $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$.

5 Examples of Conditionally Positive Definite Functions Generated By Green Functions

5.1 One-Dimensional Cases

With the theory we developed in the preceding section in mind we again discuss the cubic spline and the tension spline of Examples 2.1 and 2.2.

Example 5.1 (Piecewise Polynomial Splines) Consider the (scalar) distributional operator $\mathbf{P} := d^2/dx^2$ and $L := \mathbf{P}^{*T}\mathbf{P} = d^4/dx^4$. By integrating Equation (4.1) four times we can obtain a family of possible Green functions with respect to L, i.e.,

$$G(x) := \frac{|x|^3}{12} + a_3x^3 + a_2x^2 + a_1x + a_0, \quad x \in \mathbb{R},$$

where $a_j \in \mathbb{R}$, $j = 0, 1, 2, 3$. However, we want the Green function to be an even function. Hence, we choose

$$G(x) := \frac{1}{12}|x|^3, \quad x \in \mathbb{R}.$$

This ensures that $\mathbf{P}$ and G satisfy the conditions of Theorem 4.2 and $\|\hat{\mathbf{p}}(x)\|_2 = |x|^2$. As a result, the associated interpolant is given by

$$s_{f,X}(x) := \sum_{j=1}^{N} \frac{c_j}{12}\left|x - x_j\right|^3 + \beta_2 x + \beta_1, \quad x \in \mathbb{R}.$$

This is the same as the cubic spline interpolant of Example 2.1.

More generally, we let $\mathbf{P} := d^m/dx^m$ and $L := \mathbf{P}^{*T}\mathbf{P} = (-1)^m d^{2m}/dx^{2m}$ for some $m \in \mathbb{N}$. One of the Green functions with respect to L is

$$G(x) := \frac{(-1)^m}{2(2m-1)!}|x|^{2m-1}, \quad x \in \mathbb{R}.$$

We can verify that $\mathbf{P}$ and G satisfy the conditions of Theorem 4.2 and that $\|\hat{\mathbf{p}}(x)\|_2 = |x|^m$. Therefore G is a conditionally positive function of order m and we can obtain its associated interpolant in the form

$$s_{f,X}(x) := \sum_{j=1}^{N} \frac{(-1)^m c_j}{2(2m-1)!}\left|x - x_j\right|^{2m-1} + \sum_{k=0}^{m-1}\beta_k x^k, \quad x \in \mathbb{R}.$$

This is known as a *$(2m-1)$st-degree spline interpolant* (see [2, Chapter 6.1.5]).

In addition, according to [20, Theorem 10.40], we can check that

$$\mathrm{H}_{\mathbf{P}}(\mathbb{R}) \equiv \left\{f \in \mathrm{Re}(\mathrm{L}_1^{loc}(\mathbb{R})) \cap \mathcal{SI} : f^{(m)} \in \mathrm{L}_2(\mathbb{R})\right\}$$

has the $\mathcal{S}$-dense property. Therefore, Theorem 4.4 tells us that $\mathcal{N}_G^m(\mathbb{R}) \equiv \mathrm{H}_{\mathbf{P}}(\mathbb{R})$ and it follows that the $(2m-1)$st-degree spline is the optimal interpolant for all functions in the generalized Sobolev space $\mathrm{H}_{\mathbf{P}}(\mathbb{R})$.

Example 5.2 (Tension Splines) Let $\sigma > 0$ be a tension parameter and consider the vector distributional operator $\mathbf{P} := (d^2/dx^2, \sigma d/dx)^T$ and $L := \mathbf{P}^{*T}\mathbf{P} = d^4/dx^4 - \sigma^2 d^2/dx^2$. Then

$$G(x) := -\frac{1}{2\sigma^3}\left(\exp(-\sigma|x|) + \sigma|x|\right), \quad x \in \mathbb{R},$$

is a solution of Equation (4.1). We can verify that $\mathbf{P}$ and G satisfy the conditions of Theorem 4.2 and that $\|\hat{\mathbf{p}}(x)\|_2 = (|x|^4 + \sigma^2|x|^2)^{1/2} = \Theta(|x|)$ as $|x| \to 0$. So G is a conditionally positive definite function of order 1. This yields the same interpolant as the tension spline interpolant (2.4), i.e.,

$$s_{f,X}(x) := \sum_{j=1}^{N} c_j G(x - x_j) + \beta_1, \quad x \in \mathbb{R}.$$

It is easy to check that $|f|_{\mathrm{H}_{\mathbf{P}}(\mathbb{R})} = |f|_{TS}$ for each $f \in \mathrm{C}^2(\mathbb{R}) \cap \mathrm{H}_{\mathbf{P}}(\mathbb{R})$. According to the Sobolev inequality [1] and [20, Theorem 10.40], $\mathrm{H}_{\mathbf{P}}(\mathbb{R})$ has the $\mathcal{S}$-dense property which implies that $\mathcal{N}_G^1(\mathbb{R}) \equiv \mathrm{H}_{\mathbf{P}}(\mathbb{R})$. Theorem 4.4 and [20, Theorem 13.2] provide us with the same optimality property as stated in Example 2.2.

Example 5.3 (Univariate Sobolev Splines, [21, Example 2]) Let $\sigma > 0$ be a scaling parameter and consider $\mathbf{P} := (d/dx, \sigma I)^T$ and $L := \mathbf{P}^{*T}\mathbf{P} = \sigma^2 I - d^2/dx^2$. It is known that the Green function with respect to L is

$$G(x) := \frac{1}{2\sigma} \exp(-\sigma |x|), \quad x \in \mathbb{R}.$$

Since $\mathbf{P}$ and G satisfy the conditions of Theorem 4.2 and $\|\hat{\mathbf{p}}(x)\|_2 = \left(\sigma^2 + x^2\right)^{1/2} = \Theta(1)$ as $|x| \to 0$, we can determine that G is positive definite. By the same method as in [21, Example 2] we can verify that $\mathrm{H}_{\mathbf{P}}(\mathbb{R})$ is equivalent to the classical Sobolev space $\mathcal{H}^1(\mathbb{R})$ so that $\mathrm{H}_{\mathbf{P}}(\mathbb{R}) \subseteq \mathrm{C}(\mathbb{R}) \cap \mathrm{L}_2(\mathbb{R})$. According to Theorem 4.6,

$$\mathcal{N}_G^0(\mathbb{R}) \equiv \mathrm{H}_{\mathbf{P}}(\mathbb{R}) \equiv \mathcal{H}^1(\mathbb{R}).$$

5.2 Two and Three-Dimensional Cases

Example 5.4 (Thin Plate Splines) Let $\mathbf{P} := \left(\partial^2/\partial x_1^2, \sqrt{2}\partial^2/\partial x_1 \partial x_2, \partial^2/\partial x_2^2\right)^T$ and $L := \mathbf{P}^{*T}\mathbf{P} = \Delta^2$. It is well-known that the fundamental solution of the Poisson equation on $\mathbb{R}^2$ is given by $\boldsymbol{x} \mapsto \log \|\boldsymbol{x}\|_2$, i.e., $\Delta \log \|\boldsymbol{x}\|_2 = -2\pi\delta$. Therefore Equation (4.1) is solved by

$$G(\boldsymbol{x}) := \frac{1}{8\pi} \|\boldsymbol{x}\|_2^2 \log \|\boldsymbol{x}\|_2, \quad \boldsymbol{x} \in \mathbb{R}^2. \tag{5.1}$$

Since $\mathbf{P}$ and G satisfy the conditions of Theorem 4.2 and $\|\hat{\mathbf{p}}(\boldsymbol{x})\|_2 = \|\boldsymbol{x}\|_2^2$, G is a conditionally positive definite function of order 2 and its related interpolant has the form

$$s_{f,X}(\boldsymbol{x}) := \sum_{j=1}^{N} c_j G(\boldsymbol{x} - \boldsymbol{x}_j) + \beta_3 x_2 + \beta_2 x_1 + \beta_1, \quad \boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2. \tag{5.2}$$

Moreover, according to [20, Theorem 10.40], we can verify that $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^2)$ has the $\mathcal{S}$-dense property. Therefore, $\mathcal{N}_G^2(\mathbb{R}^2) \equiv \mathrm{H}_{\mathbf{P}}(\mathbb{R}^2)$ by Theorem 4.4. Equation (5.2) is known as the *thin plate spline* interpolant (see [4, 8]).

Finally, we consider the Duchon semi-norm mentioned in [4], i.e.,

$$|f|_{D_2}^2 := \int_{\mathbb{R}^2} \left|\frac{\partial^2 f(\boldsymbol{x})}{\partial x_1^2}\right|^2 + 2\left|\frac{\partial^2 f(\boldsymbol{x})}{\partial x_1 \partial x_2}\right|^2 + \left|\frac{\partial^2 f(\boldsymbol{x})}{\partial x_2^2}\right|^2 \mathrm{d}\boldsymbol{x}, \quad f \in \mathcal{H}_{D_2}(\mathbb{R}^2),$$

and the Duchon semi-norm space

$$\mathcal{H}_{D_2}(\mathbb{R}^2) := \left\{ f \in \mathrm{Re}(\mathrm{L}_1^{loc}(\mathbb{R}^2)) \cap \mathcal{SI} : |f|_{D_2} < \infty \right\}.$$

If we define $\mathbf{P}$ as above, then it is easy to check that $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^2) \equiv \mathcal{H}_{D_2}(\mathbb{R}^2)$. According to [20, Theorems 13.1 and 13.2] we can conclude that the Duchon semi-norm space possesses the same optimality properties as those listed in [4].

The following example shows that the same Green function G can generate different generalized Sobolev spaces $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$. Moreover, it illustrates the fact that the native space $\mathcal{N}_G^m(\mathbb{R}^d)$ may be a proper subspace of $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$.

Example 5.5 (Modified Thin Plate Splines) Let $\mathbf{P} := \Delta$ and $L := \mathbf{P}^{*T}\mathbf{P} = \Delta^2$. We find that the thin plate spline (5.1) is also the Green function with respect to the operator L defined here. The associated interpolant is again of the form (5.2).

We now consider the Laplacian semi-norm

$$|f|_\Delta^2 := \int_{\mathbb{R}^2} |\Delta f(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x}, \quad f \in \mathrm{C}(\mathbb{R}^2) \cap \mathcal{SI},$$

and the Laplacian semi-norm space

$$\mathcal{H}_\Delta(\mathbb{R}^2) := \left\{ f \in \mathrm{Re}(\mathrm{L}_1^{loc}(\mathbb{R}^2)) \cap \mathcal{SI} : |f|_\Delta < \infty \right\}.$$

It is easy to verify that $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^2) \equiv \mathcal{H}_\Delta(\mathbb{R}^2)$. However, it is known that $\mathcal{H}_{D_2}(\mathbb{R}^2)$ is a proper subspace of $\mathcal{H}_\Delta(\mathbb{R}^2)$ since $q \in \mathcal{H}_\Delta(\mathbb{R}^2)$ but $q \notin \mathcal{H}_{D_2}$ where $q(\boldsymbol{x}) := x_1 x_2$. Therefore, due to Example 5.4, we conclude that

$$\mathcal{N}_G^2(\mathbb{R}^2) \equiv \mathcal{H}_{D_2}(\mathbb{R}^2) \subsetneq \mathcal{H}_\Delta(\mathbb{R}^2) \equiv \mathrm{H}_{\mathbf{P}}(\mathbb{R}^2).$$

Instead of working with the polynomial space $\pi_1(\mathbb{R}^2)$ which is used to define $\mathcal{N}_G^2(\mathbb{R}^2)$, we can construct a new native space $\mathcal{N}_G^{\mathscr{P}}(\mathbb{R}^2)$ for G by using another finite-dimensional space $\mathscr{P}$ of $\mathrm{Re}(\mathrm{C}^2(\mathbb{R}^2)) \cap \mathcal{SI}$ such that $\mathcal{N}_G^{\mathscr{P}}(\mathbb{R}^2)$ may be equal to the other subspace of $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^2)$. First we can verify that the finite-dimensional space $\mathscr{P} := \mathrm{span}\left\{\pi_1(\mathbb{R}^2) \cup \{q\}\right\}$ is a subspace of the null space of $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^2)$. Since $\pi_1(\mathbb{R}^2) \subset \mathscr{P}$ and G is a conditionally positive definite function of order 2, we know that G is also conditionally positive definite with respect to $\mathscr{P}$. Hence, the new native space $\mathcal{N}_G^{\mathscr{P}}(\mathbb{R}^2)$ with respect to G and $\mathscr{P}$ is well-defined (see [20, Chapter 10.3]). We can further check that $\mathcal{N}_G^{\mathscr{P}}(\mathbb{R}^2)$ is a subspace of $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^2)$ but it is larger than $\mathcal{N}_G^2(\mathbb{R}^2)$, i.e., $\mathcal{N}_G^2(\mathbb{R}^2) \subsetneq \mathcal{N}_G^{\mathscr{P}}(\mathbb{R}^2) \subseteq \mathrm{H}_{\mathbf{P}}(\mathbb{R}^2)$.

So we can obtain a modification of the thin plate spline interpolant based on $\mathscr{P}$:

$$s_{f,X}^{\mathscr{P}}(\boldsymbol{x}) := \sum_{j=1}^N c_j G(\boldsymbol{x} - \boldsymbol{x}_j) + \beta_4 x_1 x_2 + \beta_3 x_2 + \beta_2 x_1 + \beta_1, \quad \boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2,$$

where the unknown coefficients are found by solving

$$\begin{pmatrix} \mathbf{A}_{G,X} & \mathbf{Q}_X \\ \mathbf{Q}_X^T & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{c} \\ \boldsymbol{\beta} \end{pmatrix} = \begin{pmatrix} \mathbf{Y} \\ 0 \end{pmatrix}, \quad \mathbf{Q}_X := \left(q_k(\boldsymbol{x}_j)\right)_{j,k=1}^{N,4} \in \mathbb{R}^{N \times 4},$$

with $\mathbf{A}_{G,X}$ as in (2.3) and $q_1(\boldsymbol{x}) := 1$, $q_2(\boldsymbol{x}) := x_1$, $q_3(\boldsymbol{x}) := x_2$, $q_4(\boldsymbol{x}) := x_1 x_2$.

Conjecture 5.1 *Motivated by Example 5.5 we audaciously* guess *the following extension of the theorems in Section 4.3: Let* $\mathbf{P}$ *and* G *satisfy the conditions of Theorem 4.2. If the subspace* $\mathscr{P}$ *of the null space of* $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ *is a finite-dimensional subspace and* $\pi_{m-1}(\mathbb{R}^d) \subseteq \mathscr{P}$*, then the new native space* $\mathcal{N}_G^{\mathscr{P}}(\mathbb{R}^2)$ *with respect to* G *and* $\mathscr{P}$ *is a subspace of* $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$.

5.3 d-Dimensional Cases

Example 5.6 (Polyharmonic Splines) This is a generalization of the earlier Examples 5.1 and 5.4. Let $\mathbf{P} := \left(\partial^m/\partial x_1^m, \cdots, \left(\frac{m!}{\alpha!}\right)^{1/2} D^\alpha, \cdots, \partial^m/\partial x_d^m\right)^T$ consisting of all $\left(\frac{m!}{\alpha!}\right)^{1/2} D^\alpha$ with $|\alpha| = m > d/2$. We further denote $L := \mathbf{P}^{*T}\mathbf{P} = (-1)^m \Delta^m$. Then the *polyharmonic spline on* $\mathbb{R}^d$ is the solution of Equation (4.1) (see [2, Chapter 6.1.5]), i.e.,

$$G(\boldsymbol{x}) := \begin{cases} \frac{\Gamma(d/2-m)}{2^{2m}\pi^{d/2}(m-1)!} \|\boldsymbol{x}\|_2^{2m-d} & \text{for } d \text{ odd},\\ \frac{(-1)^{m+d/2-1}}{2^{2m-1}\pi^{d/2}(m-1)!(m-d/2)!} \|\boldsymbol{x}\|_2^{2m-d} \log \|\boldsymbol{x}\|_2 & \text{for } d \text{ even}. \end{cases}$$

We can also check that $\mathbf{P}$ and G satisfy the conditions of Theorem 4.2 and that $\|\hat{\mathbf{p}}(\boldsymbol{x})\|_2 = \|\boldsymbol{x}\|_2^m$. Therefore G is a conditionally positive definite function of order m. Furthermore, according to [20, Theorem 10.40], we can verify that $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ has the $\mathcal{S}$-dense property. Therefore, $\mathcal{N}_G^m(\mathbb{R}^d) \equiv \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ by Theorem 4.4.

We now consider the Beppo-Levi space of order m on $\mathbb{R}^d$, i.e.,

$$BL_m(\mathbb{R}^d) := \left\{f \in \mathrm{Re}(\mathrm{L}_1^{loc}(\mathbb{R}^d)) : D^\alpha f \in \mathrm{L}_2(\mathbb{R}^d) \text{ for all } |\alpha| = m\right\}$$

equipped with the semi-inner product

$$(f, g)_{BL_m(\mathbb{R}^d)} := \sum_{|\alpha|=m} \frac{m!}{\alpha!}(D^\alpha f, D^\alpha g), \quad f, g \in BL_m(\mathbb{R}^d).$$

According to [9], we know that $BL_m(\mathbb{R}^d) \subseteq \mathrm{Re}(\mathrm{L}_1^{loc}(\mathbb{R}^d)) \cap \mathcal{SI}$ whenever $m > d/2$. Hence $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d) \equiv BL_m(\mathbb{R}^d)$.

By the way, it is well-known that G is also conditionally positive definite of order $l := m - \lceil d/2 \rceil + 1$ (see [20, Corollary 8.8]). However, the native space $\mathcal{N}_G^l(\mathbb{R}^d)$ induced by G and $\pi_{l-1}(\mathbb{R}^d)$ is a proper subspace of $\mathcal{N}_G^m(\mathbb{R}^d)$ when $d > 1$. Therefore

$$\mathcal{N}_G^l(\mathbb{R}^d) \subsetneq \mathcal{N}_G^m(\mathbb{R}^d) \equiv \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d) \equiv BL_m(\mathbb{R}^d), \quad d > 1.$$

Remark 5.1 If a vector distributional operator $\mathbf{P} := (P_1, \cdots, P_n)^T$ satisfies its distributional Fourier transform $\boldsymbol{x} \mapsto \|\hat{\mathbf{p}}(\boldsymbol{x})\|_2^2 \in \pi_{2m}(\mathbb{R}^d)$ and

$$\left\{a_\alpha D^\alpha : |\alpha| = m,\ \alpha \in \mathbb{N}_0^d\right\} \subseteq \left\{P_j : j = 1, \ldots, n\right\}, \quad \text{where } a_\alpha \neq 0 \text{ and } m > d/2,$$

then $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d) \subseteq BL_m(\mathbb{R}^d)$. According to the Sobolev inequality [1], there is a positive constant C such that $\|f\|_{\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)}^2 \leqslant C \|f\|_{BL_m(\mathbb{R}^d)}^2$ for each $f \in \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$. This implies that this generalized Sobolev space $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ also has the $\mathcal{S}$-dense property.

Example 5.7 (Sobolev Splines, [21, Example 3]) This is a generalization of Example 5.3. Let $\mathbf{P} := \left(\mathbf{Q}_0^T, \cdots, \mathbf{Q}_n^T\right)^T$, where

$$\mathbf{Q}_j := \begin{cases} \left(\frac{n!\sigma^{2n-2j}}{j!(n-j)!}\right)^{1/2} \Delta^k & \text{when } j = 2k, \\ \left(\frac{n!\sigma^{2n-2j}}{j!(n-j)!}\right)^{1/2} \Delta^k \nabla & \text{when } j = 2k+1, \end{cases} \qquad k \in \mathbb{N}_0,\ j = 0, 1, \ldots, n,\ n > d/2.$$

Here we use $\Delta^0 := I$. We further define $L := \mathbf{P}^{*T}\mathbf{P} = (\sigma^2 I - \Delta)^n$.

The *Sobolev spline* (or Matérn function) is known to be the Green function with respect to L (see [2, Chapter 6.1.6]), i.e.,

$$G(\boldsymbol{x}) := \frac{2^{1-n-d/2}}{\pi^{d/2}\Gamma(n)\sigma^{2n-d}} \left(\sigma \|\boldsymbol{x}\|_2\right)^{n-d/2} K_{d/2-n}\left(\sigma \|\boldsymbol{x}\|_2\right), \quad \boldsymbol{x} \in \mathbb{R}^d,$$

where $z \mapsto K_\nu(z)$ is the *modified Bessel function of the second kind of order* ν. Since $\mathbf{P}$ and G satisfy the conditions of Theorem 4.2 and $\|\hat{\mathbf{p}}(\boldsymbol{x})\|_2 = \Theta(1)$ as $\|\boldsymbol{x}\|_2 \to 0$, G is positive definite and the associated interpolant $s_{f,X}$ is the same as the Sobolev spline (or Matérn) interpolant.

Combining [21, Example 3] and Theorem 4.6, we can determine that

$$\mathcal{N}_G^0(\mathbb{R}^d) \equiv \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d) \equiv \mathcal{H}^n(\mathbb{R}^d).$$

Moreover, this shows that the classical Sobolev space $\mathcal{H}^n(\mathbb{R}^d)$ becomes a reproducing-kernel Hilbert space with $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$-inner product and its reproducing kernel is given by $K(\boldsymbol{x}, \boldsymbol{y}) := G(\boldsymbol{x} - \boldsymbol{y})$.

In the following example we are not able to establish that the operator $\mathbf{P}$ satisfies the conditions of Theorem 4.2 and so part of the connection to the theory developed in this paper is lost. We therefore use the symbol Φ to denote the kernel instead of G.

Example 5.8 (Gaussians, [21, Example 4]) The Gaussian kernel $K(\boldsymbol{x}, \boldsymbol{y}) := \Phi(\boldsymbol{x} - \boldsymbol{y})$ derived by the Gaussian function Φ is very important and popular in the current research fields of scattered data approximation and machine learning. A large group of people would like to know the native space of the Gaussian function or the reproducing-kernel Hilbert space of the Gaussian kernel. In this example we will show that the native space of the Gaussian function is equivalent to a generalized Sobolev space.

We firstly consider the *Gaussian* function

$$\Phi(\boldsymbol{x}) := \frac{\sigma^d}{\pi^{d/2}} \exp(-\sigma^2 \|\boldsymbol{x}\|_2^2), \quad \boldsymbol{x} \in \mathbb{R}^d, \quad \sigma > 0.$$

We know that Φ is a positive definite function and its native space $\mathcal{N}_\Phi^0(\mathbb{R}^d)$ is a reproducing kernel Hilbert space (see [5, Chapter 4]).

Let $\mathbf{P} := \left(\mathbf{Q}_0^T, \cdots, \mathbf{Q}_n^T, \cdots\right)^T$, where

$$\mathbf{Q}_n := \begin{cases} \left(\frac{1}{n!4^n\sigma^{2n}}\right)^{1/2} \Delta^k & \text{when } n = 2k, \\ \left(\frac{1}{n!4^n\sigma^{2n}}\right)^{1/2} \Delta^k \nabla & \text{when } n = 2k+1, \end{cases} \qquad k \in \mathbb{N}_0.$$

Here we again use $\Delta^0 := I$. Since the differential operators are just special cases of distributional operators, the generalized Sobolev space $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ defined by $\mathbf{P}$ is the same as that derived in [21, Example 4]. According to the proof of [21, Example 4], we have

$$\mathcal{N}_{\Phi}^{0}(\mathbb{R}^d) \equiv \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d).$$

Moreover, it is easy to verify that $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d) \subseteq \mathcal{H}^n(\mathbb{R}^d)$ for each $n \in \mathbb{N}$. According to the Sobolev embedding theorem [1], we also have $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d) \subseteq \mathrm{Re}(\mathrm{C}_b^{\infty}(\mathbb{R}^d))$. However, $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ does not contain polynomials. If $f \in \mathrm{Re}(\mathrm{C}_b^{\infty}(\mathbb{R}^d))$ and there is a positive constant C such that $\|D^{\alpha} f\|_{\mathrm{L}_{\infty}(\mathbb{R}^d)} \leqslant C^{|\alpha|}$ for each $\alpha \in \mathbb{N}_0^d$, then $f \in \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ which implies that $f \in \mathcal{N}_{\Phi}^{0}(\mathbb{R}^d)$.

6 Extensions and Future Works

In this paper we have presented a unified theory for the generation of conditionally positive definite functions of order m as (full-space) Green functions with respect to a distributional operator $L := \mathbf{P}^{*T}\mathbf{P}$ with an appropriate vector distributional operator $\mathbf{P}$. These even Green functions $G \in \mathrm{Re}(\mathrm{C}(\mathbb{R}^d)) \cap \mathcal{SI}$ can be used as basic functions of a translation invariant meshfree kernel-based approximation methods of the form (2.1)-(2.2). Our analysis is limited to this translation invariant setting which does not address the fully general situation with kernels of the form $K(\boldsymbol{x}, \boldsymbol{y})$, but is more general than the radial setting.

In Section 5 we were able to show that many different types of "splines" and radial basis functions can be treated with our Green function framework. Thus, reproducing kernel Hilbert space methods can be viewed as a natural generalization of univariate splines (including such variations as tension splines). Other forms of univariate splines such as smoothing splines or regression splines can be covered using a related least squares framework, and multivariate generalizations of these methods are widely used in statistics and machine learning.

We only consider real-valued functions as candidates for the generalized Sobolev spaces and Green functions in this paper, but all the conclusions and the theorems can be extended to complex-valued functions in a way similarly to [20]. $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ may not be complete even though we extend it to complex-valued functions. However, its completion is embedded into the tempered distribution space $\mathcal{S}'$ and has the explicit form

$$\overline{\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)} \equiv \left\{T \in \mathcal{S}' : P_j T \in \mathrm{L}_2(\mathbb{R}^d),\ j = 1, \ldots, n\right\}, \quad \text{if } \mathbf{P} = (P_1, \cdots, P_n)^T .$$

The vector distributional operator $\mathbf{P}$ can be further constructed by pseudo-differential operators with non constant coefficients. Therefore their generalized Sobolev spaces $\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)$ could be equivalent to the Beppo-Levi type spaces $X_{\tau}^{m}(\mathbb{R}^d)$. For example, if $\mathbf{P} := \left(\omega_{\tau}\partial^m/\partial x_1^m, \cdots, \omega_{\tau} D^{\alpha}, \cdots, \omega_{\tau}\partial^m/\partial x_d^m\right)^T$, then

$$\mathrm{H}_{\mathbf{P}}(\mathbb{R}^d) \equiv X_{\tau}^{m}(\mathbb{R}^d) := \left\{f \in \mathrm{Re}(\mathrm{L}_1^{loc}(\mathbb{R}^d)) \cap \mathcal{SI} : \omega_{\tau} D^{\alpha} f \in \mathrm{L}_2(\mathbb{R}^d), |\alpha| = m, \alpha \in \mathbb{N}_0^d\right\},$$

where $\omega_{\tau}(\boldsymbol{x}) := \|\boldsymbol{x}\|_2^{\tau}$ and $0 \leqslant \tau < 1$. However, $\mathbf{P}$ may not satisfy the condition of Theorem 4.2. We have reserved these situations for our future research.

Unfortunately, it is sometimes difficult for us to solve a Green function matching the conditions of Theorem 4.2 even if the vector distributional operator $\mathbf{P}$ satisfies the conditions of Theorem 4.2. However, there is usually an even Green function $G \in \mathrm{Re}(\mathrm{C}(\mathbb{R}^d \setminus \{0\})) \cap \mathcal{SI}$. This means that the Green function merely has a *singular point* at the origin. According to our numerical tests of some cases, we find that this kind of Green function can still play the role of a basic function for the construction of a multivariate interpolant $s_{f,X}$ via (2.1)-(2.2) after some techniques to remove the singularity. One of the numerical tests is a two-dimensional example as below. Let $\mathbf{P} := (\Delta, \sigma\nabla^T)^T$ with $\sigma > 0$ and the Green function with respect to $L := \mathbf{P}^{*T}\mathbf{P} = \Delta^2 - \sigma^2\Delta$ be given by

$$G(\boldsymbol{x}) := -\frac{1}{2\pi\sigma^2}\left(K_0\left(\sigma\|\boldsymbol{x}\|_2\right) + \log\left(\sigma\|\boldsymbol{x}\|_2\right)\right), \quad \boldsymbol{x} \in \mathbb{R}^2,$$

where $z \mapsto K_\nu(z)$ is the modified Bessel function of the second kind of order ν. We can use a transformations to remove the singularity of G as follows:

$$G_r(\boldsymbol{x}) := -\frac{1}{2\pi\sigma^2}\left(K_0\left(\sigma\|\boldsymbol{x}\|_2 + r\right) + \log\left(\sigma\|\boldsymbol{x}\|_2 + r\right)\right), \quad \boldsymbol{x} \in \mathbb{R}^2,\ r > 0.$$

We guess that the interpolant via this modified Green function may be used to approximate functions belonging to the related generalized Sobolev space.

We merely consider the Lebesgue measure here. However, we can further generalize our results to other measure spaces $(\Omega, \mathcal{B}_\Omega, \mu)$, where $\Omega \subseteq \mathbb{R}^d$ and $\mathcal{B}_\Omega$ is the Borel set of Ω. We suppose that the bijective map

$$\mathcal{A} : (\Omega, \mathcal{B}_\Omega, \mu) \to (\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$$

is differentiable at every point of Ω such that

$$\mathrm{d}\mu(\boldsymbol{x}) = \left|\det\left(J_{\mathcal{A}}(\boldsymbol{x})\right)\right| \mathrm{d}\boldsymbol{x}, \quad \text{where } J_{\mathcal{A}}(\boldsymbol{x}) \text{ is the Jacobian matrix of } \mathcal{A} \text{ at } \boldsymbol{x}.$$

According to the Radon-Nikodym Theorem [7] it is not difficult to gain similar conclusions when we transform the generalized Sobolev space to be

$$\mathrm{H}^\mu_{\mathbf{P}}(\Omega) := \left\{f_{\mathcal{A}} := f \circ \mathcal{A} : f \in \mathrm{H}_{\mathbf{P}}(\mathbb{R}^d)\right\},$$

with the semi-inner product

$$(f_{\mathcal{A}}, g_{\mathcal{A}})_{\mathrm{H}^\mu_{\mathbf{P}}(\Omega)} := \sum_{n=1}^{\infty}\int_\Omega P_n f(\mathcal{A}(\boldsymbol{x})) P_n g(\mathcal{A}(\boldsymbol{x}))\mathrm{d}\mu(\boldsymbol{x}), \quad f_{\mathcal{A}}, g_{\mathcal{A}} \in \mathrm{H}^\mu_{\mathbf{P}}(\Omega).$$

Finally, we do not specify any boundary conditions for the (full-space) Green functions. Thus we may have many choices of the Green functions with respect to the same distributional operator L. In our future work we will apply a vector distributional operator $\mathbf{P} := (P_1, \cdots, P_n)^T$ and a vector boundary operator $\mathbf{B} := (B_1, \cdots, B_s)^T$ on a bounded domain Ω to construct a reproducing kernel and its related reproducing-kernel Hilbert space (see [6]). We further hope to use the distributional operator L to approximate the eigenvalues and eigenfunctions of the kernel function with the goal of obtaining fast numerical methods to solve the interpolating systems (2.1)-(2.2) similar as fast multipole methods in [20, Chapter 15].

References

1. R. A. Adams and J. J. F. Fournier, *Sobolev Spaces* (2nd Ed.), Pure and Applied Mathematics, Vol. 140, Academic Press, 2003.
2. A. Berlinet and C. Thomas-Agnan, *Reproducing Kernel Hilbert Spaces in Probability and Statistics*, Kluwer Academic Publishers, 2004.
3. M. D. Buhmann, *Radial Basis Functions: Theory and Implementations*, Cambridge University Press (Cambridge), 2003.
4. J. Duchon, *Splines minimzing rotation-invariant semi-norms in Sobolev spaces*, in *Constructive Theory of Functions of Several Variables*, W. Schempp and K. Zeller (Eds.), Springer-Verlag (Berlin), 1977, 85–100.
5. G. E. Fasshauer. *Meshfree Approximation Methods with* MATLAB, Interdisciplinary Mathematical Sciences, Vol. 6, World Scientific Publishers (Singapore), 2007.
6. G. E. Fasshauer and Q. Ye, *Reproducing Kernels of Sobolev Spaces via a Green Function Approach with Differential Operators and Boundary Operators*, in preparation.
7. J. K. Hunter and B. Nachtergaele, *Applied Analysis*, World Scientific Publishers (Singapore), 2005.
8. J. Kybic, T. Blu and M. Unser, *Generalized sampling: A variational approach - Part I and II: Theory*, IEEE Trans. Signal Proc. 50/8 (2002), 1965–1985.
9. W. A. Light and H. Wayne, *Spaces of distributions, interpolation by translates of a basis function and error estimates*, Numer. Math. 81 (1999), 415–450.
10. W. R. Madych and S. A. Nelson, *Multivariate interpolation and conditionally positive definite functions II*, Math. Comp. 54 (1990), 211–230.
11. R. J. Renka, *Interpolatory tension splines with automatic selection of tension factors*, SIAM J. Sci. Stat. Comput. 8/3 (1987), 393–415.
12. R. Schaback, *Spectrally Optimized Derivative Formulae*, Data Page of R. Schaback's Research Group, 2008.
13. R. Schaback and H. Wendland, *Kernel techniques: From machine learning to meshless methods*, Acta Numerica 15 (2006), 543–639.
14. B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, MIT Press (Cambridge, MA), 2002.
15. D. G. Schweikert, *An interpolation curve using a spline in tension*, J. Math. Phys. 45 (1966), 312–317.
16. A. J. Smola, B. Schölkopf and K. R. Müller, *The connection between regularization operators and support vector kernels*, Neural Networks 11 (1998), 637–649.
17. E. M. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, 1975.
18. M. L. Stein, *Interpolation of Spatial Data. Some Theory for Kriging*, Springer Series in Statistics, Springer-Verlag (New York), 1999.
19. G. Wahba, *Spline Models for Observational Data*, CBMS-NSF Regional Conference Series in Applied Mathematics 59, SIAM (Philadelphia), 1990.
20. H. Wendland. *Scattered Data Approximation*, Cambridge University Press, 2005.
21. Q. Ye, *Reproducing Kernels of Generalized Sobolev Spaces via a Green Function Approach with Differential Operators*, submitted.
22. A. L. Yuille and N. M. Grzywacz, *The motion coherence theory*, in *Proceedings of the International Conference on Computer Vision*, IEEE Computer Society (1988), 344–353.