

## DISTRIBUTIONAL EXTENSION AND INVERTIBILITY OF THE $k$ -PLANE TRANSFORM AND ITS DUAL\*

RAHUL PARHI<sup>†</sup> AND MICHAEL UNSER<sup>†</sup>

**Abstract.** We investigate the distributional extension of the  $k$ -plane transform in  $\mathbb{R}^d$  and of related operators. We parameterize the  $k$ -plane domain as the Cartesian product of the Stiefel manifold of orthonormal  $k$ -frames in  $\mathbb{R}^d$  with  $\mathbb{R}^{d-k}$ . This parameterization imposes an isotropy condition on the range of the  $k$ -plane transform which is analogous to the even condition on the range of the Radon transform. We use our distributional formalism to investigate the invertibility of the dual  $k$ -plane transform (the “backprojection” operator). We provide a systematic construction (via a completion process) to identify Banach spaces in which the backprojection operator is invertible and present some prototypical examples. These include the space of isotropic finite Radon measures and isotropic  $L^p$ -functions for  $1 < p < \infty$ . Finally, we apply our results to study a new form of regularization for inverse problems.

**Key words.**  $k$ -plane transform, Radon transform, X-ray transform, distributional extension, backprojection, invertibility

**MSC codes.** 46F12, 44A15, 44A12, 42B10

**DOI.** 10.1137/23M1556721

**1. Introduction.** The  $k$ -plane transform of a function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , denoted by  $\mathcal{R}_k\{f\}$ , is defined on the space of  $k$ -planes (affine subspaces of dimension  $k$ ) in  $\mathbb{R}^d$  for  $1 \leq k < d$ . The value of  $\mathcal{R}_k\{f\}$  at a given  $k$ -plane is the integral of  $f$  over that  $k$ -plane. The two extremes of this transform correspond to the Radon transform ( $k = (d - 1)$ ) and the X-ray transform ( $k = 1$ ). Intermediate dimensions are also of interest such as  $k = (d - 2)$ , which corresponds to the measurement model of nuclear magnetic resonance imaging [62, 64]. These transforms have had considerable interest in both pure and applied mathematics following the seminal work of Radon [47] and John [28, 29]. The Radon transform, in particular, has had profound impacts in the study of partial differential equations (PDEs) [30], computerized tomography [36], the theory of wavelets, ridgelets, and shearlets [3, 8, 9, 13, 33, 34, 52, 61, 66], and, more recently, neural networks [4, 35, 41, 42, 43, 44, 45]. Moreover, there is a large body of work that has investigated properties of the Radon transform. These include characterizations of its range on various spaces of functions [25, 26, 28, 37], distributional extensions [33, 37], the invertibility of the dual transform (the “backprojection” operator) [33, 49, 65], and several others [1, 31, 39].

The  $k$ -plane transform, sometimes referred to as the  $k$ -dimensional Radon transform or the Radon–John transform, is the most obvious generalization of the Radon transform. While there are works devoted to the  $k$ -plane transform that include inversion formulas [18, 32, 53, 56, 54] and range characterizations [22, 24, 38, 46, 51, 71, 72, 73], as well as its connections to wavelets [11, 53, 55], far less is known about the  $k$ -plane transform than the Radon transform.

---

\*Received by the editors March 2, 2023; accepted for publication (in revised form) April 16, 2024; published electronically July 2, 2024.

<https://doi.org/10.1137/23M1556721>

**Funding:** This work was supported by the Swiss National Science Foundation under grant 200020\_219356/1.

<sup>†</sup>Biomedical Imaging Group, École polytechnique fédérale de Lausanne, CH-1015 Lausanne, Switzerland (rahul.parhi@epfl.ch, michael.unser@epfl.ch).

This paper brings the following new results:

1. The extension of the  $k$ -plane transform and of related operators to distributions (section 6).
2. The formulation of an extended version of the Fourier slice theorem that is applicable to any tempered distribution (Theorem 7.7), a result that is new, to the best of our knowledge, for all  $1 \leq k < d$  (which includes the Radon transform). Theorem 7.7 also yields a general form of the intertwining properties for the  $k$ -plane transform (Corollary 7.8).
3. The systematic construction (via a completion process) of Banach spaces over which the dual  $k$ -plane transform (backprojection operator) is guaranteed to be invertible (Theorem 8.1).
4. The identification of prototypical examples of such  $k$ -plane-domain Banach spaces. These examples recover and extend many known results regarding the invertibility of the backprojection operator (see Remark 8.4 for a detailed discussion). They include the space of isotropic finite Radon measures and isotropic  $L^p$ -functions for  $1 < p < \infty$  (Corollary 8.3).

The primary technical contribution of this paper is the careful treatment of the extension of the  $k$ -plane transform and related operators to distributions. The main difficulty is that the range of the  $k$ -plane transform on Schwartz functions is not the full space of Schwartz functions on the  $k$ -plane domain. It is instead a subspace of Schwartz functions over which certain consistency conditions are met. As it turns out, when we extend the  $k$ -plane transform by duality, the resulting distribution spaces are actually *equivalence classes* of distributions. This introduces an inherent ambiguity and prevents one from identifying the  $k$ -plane transform of a distribution uniquely. To deal with this nonuniqueness issue, we construct  $k$ -plane-domain Banach spaces over which the backprojection operator is invertible. This then yields a setting where the distributional (filtered)  $k$ -plane transform can be *uniquely determined* (Theorem 8.1).

**2. Motivation.** The primary motivation for the proposed investigation is to study  *$k$ -plane ridges*. These are functions that map  $\mathbb{R}^d \rightarrow \mathbb{R}$  and that are constant along  $k$ -planes with a fixed orientation. They can be written as

$$(2.1) \quad \mathbf{x} \mapsto r(\mathbf{A}\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d,$$

where  $r: \mathbb{R}^{d-k} \rightarrow \mathbb{R}$  is the *profile* and  $\mathbf{A} \in \mathbb{R}^{(d-k) \times d}$  is the *orientation* of the  $k$ -plane ridge. Observe that (2.1) is constant along the  $k$ -planes  $\mathbf{A}\mathbf{x} = \mathbf{c}$ , where  $\mathbf{c} \in \mathbb{R}^{d-k}$ , whenever  $\mathbf{A}$  is full rank.

Early works that used the  $k$ -plane transform to study  $k$ -plane ridges go back to the 1990s and were spearheaded by Donoho [11] and Rubin [55]. There, the authors propose wavelet-like systems that interpolate between wavelets and ridgelets, as  $k$  varies. These systems, referred to as  $k$ -plane ridgelets, are built from  $k$ -plane ridges with  $(d-k)$ -variate wavelet profiles. Crucially, they rely on wavelet analysis in the  $k$ -plane domain. Donoho's motivation was to propose a system that could capture and exploit low-dimensional structures found in high-dimensional functions and data, with the purpose of understanding the curse of dimensionality in high-dimensional data-analysis problems [12].

More recently,  $k$ -plane ridges, sometimes referred to as *multi-index models* or  *$(d-k)$ -sparse functions*, have been studied in the context of dimensionality reduction [10, 19] and as a tool to understand why neural networks work well in high-dimensional problems—whereas kernel methods do not [2, 21]. In particular, the distributional filtered Radon transform ( $k = (d-1)$ ) and ridgelet analysis have been used to

characterize a quantitative gap between the performance of neural networks and kernel methods via nonlinear and linear minimax rates for nonparametric function estimation [45]. Notably, the authors of [45] prove that, for certain classes of problems, neural networks enjoy estimation error rates that do not grow with the input dimension, whereas the estimation error rates for kernel methods *necessarily* suffer the curse of dimensionality.

To summarize, we see that  $k$ -plane ridges and the  $k$ -plane transform are powerful tools to understand and quantify the curse of dimensionality. Therefore, we believe that the results of this paper provide the mathematical framework to further the investigation of this hallmark. Indeed, in section 9 we use our formalism to prove that the application of the distributional filtered  $k$ -plane transform to a  $k$ -plane ridge extracts the profile function and orientation matrix (see Theorem 9.4). This property allows for one to analyze combinations of  $k$ -plane ridges and to access the  $(d - k)$ -variate profiles directly, without knowledge of the underlying orientation matrices. Said differently, the filtered  $k$ -plane transform automatically captures intrinsic low-dimensional structures in high-dimensional functions. We exemplify the utility of this result in section 10, where we investigate the resolution of inverse problems with new forms of  $k$ -plane-domain regularization.

**3. Preliminaries and notation.** The Schwartz space of smooth and rapidly decreasing functions on  $\mathbb{R}^d$  is denoted by  $\mathcal{S}(\mathbb{R}^d)$ . Its continuous dual is the space  $\mathcal{S}'(\mathbb{R}^d)$  of tempered distributions. We let  $L^p(\mathbb{R}^d)$  denote the Lebesgue space for  $1 \leq p \leq \infty$  and  $L^1_{\text{temp}}(\mathbb{R}^d) = L^1_{\text{loc}}(\mathbb{R}^d) \cap \mathcal{S}'(\mathbb{R}^d)$  the space of tempered locally integrable functions. The Banach space of continuous functions vanishing at  $\pm\infty$  on  $\mathbb{R}^d$  equipped with the  $L^\infty$ -norm is denoted by  $C_0(\mathbb{R}^d)$ . By the Riesz–Markov–Kakutani representation theorem [17, Chapter 7], its continuous dual can be identified with the Banach space of finite Radon measures, denoted by  $\mathcal{M}(\mathbb{R}^d)$ . Since  $\mathcal{S}(\mathbb{R}^d) \xrightarrow{d} C_0(\mathbb{R}^d)$ , where  $\xrightarrow{d}$  denotes a dense and continuous embedding, we have by duality that  $\mathcal{M}(\mathbb{R}^d) \xrightarrow{c} \mathcal{S}'(\mathbb{R}^d)$ , where  $\xrightarrow{c}$  denotes a continuous embedding. In the case of a compact domain  $\Omega$ , we have  $\mathcal{M}(\Omega) = (C(\Omega))'$ , where  $C(\Omega)$  denotes the space of continuous functions on  $\Omega$ .

The Fourier transform of  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  is defined as

$$(3.1) \quad \widehat{\varphi}(\boldsymbol{\xi}) := \mathcal{F}_d\{\varphi\}(\boldsymbol{\xi}) = \int_{\mathbb{R}^d} \varphi(\mathbf{x}) e^{-i\boldsymbol{\xi}^\top \mathbf{x}} \, d\mathbf{x}, \quad \boldsymbol{\xi} \in \mathbb{R}^d,$$

where  $i^2 = -1$ . Consequently, the inverse Fourier transform of  $\widehat{\varphi} \in \mathcal{S}(\mathbb{R}^d)$  is given by

$$(3.2) \quad \mathcal{F}_d^{-1}\{\widehat{\varphi}\}(\mathbf{x}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \widehat{\varphi}(\boldsymbol{\xi}) e^{i\boldsymbol{\xi}^\top \mathbf{x}} \, d\boldsymbol{\xi}, \quad \mathbf{x} \in \mathbb{R}^d.$$

These operators are extended to act on  $\mathcal{S}'(\mathbb{R}^d)$  by duality. Any continuous linear shift-invariant (LSI) operator  $L : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}'(\mathbb{R}^d)$  is a convolution operator specified by a unique kernel  $h \in \mathcal{S}'(\mathbb{R}^d)$  such that  $L\varphi = h * \varphi$ . Such operators can also be specified in the Fourier domain by

$$(3.3) \quad L\varphi = \mathcal{F}_d^{-1}\{\widehat{L}\widehat{\varphi}\},$$

where  $\widehat{L} \in \mathcal{S}'(\mathbb{R}^d)$  is the Fourier transform of the kernel  $h \in \mathcal{S}'(\mathbb{R}^d)$ . The tempered distribution  $h$  is the *impulse response* of  $L$  and the tempered distribution  $\widehat{L}$  is the Fourier symbol or *frequency response* of  $L$ . We shall generally use upright, roman letters for LSI operators and use the italic variant with a hat to denote its frequency response. Finally, given a space  $\mathcal{X}$  and a norm  $\|\cdot\|$ , the completion of  $\mathcal{X}$  in  $\|\cdot\|$  is a Banach space, denoted by  $(\mathcal{X}, \|\cdot\|)$ .

**4. Parameterization of the  $k$ -plane transform.** The Radon transform is a special case of the  $k$ -plane transform. It corresponds to integrals over affine subspaces of dimension  $(d - 1)$  in  $\mathbb{R}^d$  (i.e., hyperplanes). The classical integral formula of the Radon transform of  $f \in L^1(\mathbb{R}^d)$  is

$$(4.1) \quad \mathcal{R}\{f\}(\boldsymbol{\alpha}, t) = \int_{\boldsymbol{\alpha}^\perp} f(\mathbf{x} + \boldsymbol{\alpha}t) \, d\mathbf{x}, \quad (\boldsymbol{\alpha}, t) \in \mathbb{S}^{d-1} \times \mathbb{R},$$

where  $\boldsymbol{\alpha}^\perp := \{\mathbf{x} \in \mathbb{R}^d : \boldsymbol{\alpha}^\top \mathbf{x} = 0\}$  and  $\mathbb{S}^{d-1} = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_2 = 1\}$  is the unit sphere in  $\mathbb{R}^d$ . This transform integrates  $f$  along the hyperplane  $\{\mathbf{x} \in \mathbb{R}^d : \boldsymbol{\alpha}^\top \mathbf{x} = t\}$ . In the case of  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ , this transform is equivalently specified by

$$(4.2) \quad \mathcal{R}\{\varphi\}(\boldsymbol{\alpha}, t) = \int_{\mathbb{R}^d} \varphi(\mathbf{x})\delta(\boldsymbol{\alpha}^\top \mathbf{x} - t) \, d\mathbf{x},$$

where  $\delta$  is the univariate Dirac impulse and the integral is understood as the action of  $\delta(\boldsymbol{\alpha}^\top(\cdot) - t) \in \mathcal{S}'(\mathbb{R}^d)$  on  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . Since  $(\boldsymbol{\alpha}, t)$  and  $(-\boldsymbol{\alpha}, -t)$  specify the same hyperplane, we see that the Radon transform imposes an even symmetry.

Many authors work with the  $k$ -plane transform in  $\mathbb{R}^d$  somewhat abstractly by considering integrals of a function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  over elements of the affine Grassmann manifold which is precisely the manifold of  $k$ -planes in  $\mathbb{R}^d$  (see [38] and the references therein). This abstract formulation does not recover the usual specifications of the Radon transform given by (4.1) and (4.2) in an immediate way. Following the work of Petrov [71, 72, 73], we can alternatively work with a concrete parameterization of the affine Grassman manifold with matrices and vectors. Here, we specify a  $k$ -plane in  $\mathbb{R}^d$  by the set of all  $\mathbf{x} \in \mathbb{R}^d$  that satisfy the system of equations

$$(4.3) \quad \mathbf{A}\mathbf{x} = \mathbf{t},$$

where  $\mathbf{A} \in \mathbb{R}^{(d-k) \times d}$ , and  $\mathbf{t} \in \mathbb{R}^{d-k}$ . Let us write  $\mathbf{A}$  rowwise as

$$(4.4) \quad \mathbf{A} = \begin{bmatrix} \boldsymbol{\alpha}_1^\top \\ \vdots \\ \boldsymbol{\alpha}_{d-k}^\top \end{bmatrix}.$$

While Petrov only requires that the rows of  $\mathbf{A}$  are linearly independent, we additionally impose, without loss of generality, that the rows are orthonormal, with  $\boldsymbol{\alpha}_n \in \mathbb{S}^{d-1}$ ,  $n = 1, \dots, (d - k)$ , and  $\boldsymbol{\alpha}_m^\top \boldsymbol{\alpha}_n = \delta[m - n]$ , where  $\delta[\cdot]$  denotes the Kronecker impulse. Accordingly,  $\mathbf{A}$  is an element of the Stiefel manifold of orthonormal  $k$ -frames in  $\mathbb{R}^d$  specified by

$$(4.5) \quad V_{d-k}(\mathbb{R}^d) = \{\mathbf{A} \in \mathbb{R}^{(d-k) \times d} : \mathbf{A}\mathbf{A}^\top = \mathbf{I}_{d-k}\}.$$

While we do not explicitly require the manifold structure of  $V_{d-k}(\mathbb{R}^d)$  in this paper, we do require the ability to integrate on  $V_{d-k}(\mathbb{R}^d)$  with respect to its Haar measure, which we denote by  $\sigma$ .

With this parameterization, we specify the  $k$ -plane transform of  $f \in L^1(\mathbb{R}^d)$  as

$$(4.6) \quad \mathcal{R}_k\{f\}(\mathbf{A}, \mathbf{t}) = \int_{\mathbf{A}^\perp} f(\mathbf{x} + \mathbf{A}^\top \mathbf{t}) \, d\mathbf{x}, \quad (\mathbf{A}, \mathbf{t}) \in (V_{d-k}(\mathbb{R}^d), \mathbb{R}^{d-k}),$$

where  $\mathbf{A}^\perp := \{\mathbf{x} \in \mathbb{R}^d : \mathbf{A}\mathbf{x} = \mathbf{0}\}$ . This transform integrates  $f$  along the  $k$ -plane  $\{\mathbf{x} \in \mathbb{R}^d : \mathbf{A}\mathbf{x} = \mathbf{t}\}$ . In the case of  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ , this transform is equivalently specified by

$$(4.7) \quad \mathcal{R}_k\{f\}(\mathbf{A}, \mathbf{t}) = \int_{\mathbb{R}^d} \varphi(\mathbf{x})\delta(\mathbf{A}\mathbf{x} - \mathbf{t}) \, d\mathbf{x},$$

Downloaded 07/02/24 to 128.178.48.127 . Redistribution subject to SIAM license or copyright; see https://pubs.siam.org/terms-privacy

where  $\delta$  is the  $(d - k)$ -variate Dirac impulse and the integral is understood as the action of  $\delta(\mathbf{A}(\cdot) - \mathbf{t}) \in \mathcal{S}'(\mathbb{R}^d)$  on  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . In Definition 7.3 and Remark 7.4, we expose the precise meaning of the tempered distribution  $\delta(\mathbf{A}(\cdot) - \mathbf{t})$ . We see that (4.6) and (4.7) immediately recover the specifications of the Radon transform in (4.1) and (4.2) when  $k = (d - 1)$ . Further, the dual transform (often called the “backprojection”) of  $g \in L^\infty(V_{d-k}(\mathbb{R}^d) \times \mathbb{R}^{d-k})$  is given by

$$(4.8) \quad \mathcal{R}_k^* \{g\}(\mathbf{x}) = \int_{V_{d-k}(\mathbb{R}^d)} g(\mathbf{A}, \mathbf{A}\mathbf{x}) d\sigma(\mathbf{A}), \quad \mathbf{x} \in \mathbb{R}^d.$$

Furthermore, since we impose that the rows of  $\mathbf{A}$  are orthonormal, we have that  $(\mathbf{A}, \mathbf{t})$  and  $(\mathbf{U}\mathbf{A}, \mathbf{U}\mathbf{t})$  define the same  $k$ -plane for any orthogonal transformation  $\mathbf{U} \in O_{d-k}(\mathbb{R})$  (the orthogonal group in dimension  $(d - k)$ ). The main advantage of the proposed parameterization is that it will allow us to identify the symmetries of the  $k$ -plane domain as “isotropic” symmetries.

**5. The  $k$ -plane transform of Schwartz functions.** Let  $\Xi_k = V_{d-k}(\mathbb{R}^d) \times \mathbb{R}^{d-k}$  denote the  $k$ -plane domain. We exclusively use the variables  $(\mathbf{A}, \mathbf{t}) \in \Xi_k$  to index the  $k$ -plane domain. The space of Schwartz functions on  $\Xi_k$ , denoted by  $\mathcal{S}(\Xi_k)$ , is the space of smooth functions that are rapidly decreasing in the  $\mathbf{t} \in \mathbb{R}^{d-k}$  variable [22]. More specifically, we have that  $\mathcal{S}(\Xi_k) = C^\infty(V_{d-k}(\mathbb{R}^d)) \widehat{\otimes} \mathcal{S}(\mathbb{R}^{d-k})$ , where  $\widehat{\otimes}$  denotes the topological tensor product, which is the completion of the algebraic tensor product with respect to the projective topology [67, Chapter 43]. We have the following classical result regarding the continuity and invertibility of the  $k$ -plane transform.

**THEOREM 5.1** (see [20, 22, 32, 53, 62, 63]). *The operator  $\mathcal{R}_k$  continuously maps  $\mathcal{S}(\mathbb{R}^d)$  into  $\mathcal{S}(\Xi_k)$ . Moreover,*

$$(5.1) \quad \mathcal{R}_k^* \mathbf{K}_{d-k} \mathcal{R}_k = c_{d,k} (-\Delta_d)^{\frac{k}{2}} \mathcal{R}_k^* \mathcal{R}_k = c_{d,k} \mathcal{R}_k^* \mathcal{R}_k (-\Delta_d)^{\frac{k}{2}} = \text{Id}$$

on  $\mathcal{S}(\mathbb{R}^d)$ , with

$$(5.2) \quad c_{d,k} = \frac{1}{(2\pi)^k} \frac{|\mathbb{S}^{k-1}|}{|\mathbb{S}^{d-k-1}|} \frac{1}{\sigma(V_{d-k-1}(\mathbb{R}^{d-1}))},$$

where  $|\cdot|$  denotes the surface area and  $\sigma(\cdot)$  is the Haar measure of the Stiefel manifold. The underlying operators are the  $d$ -variate Laplacian operator  $\Delta_d$  and the filtering operator<sup>1</sup>  $\mathbf{K}_{d-k} = c_{d,k} (-\Delta_{d-k})^{k/2}$ , where  $\Delta_{d-k}$  denotes the  $(d - k)$ -variate Laplacian applied to the  $\mathbf{t} \in \mathbb{R}^{d-k}$  variable. The filtering operator is equivalently specified by the frequency response  $\widehat{K}_{d-k}(\boldsymbol{\omega}) = c_{d,k} \|\boldsymbol{\omega}\|_2^k$ ,  $\boldsymbol{\omega} \in \mathbb{R}^{d-k}$ .

*Remark 5.2.* We shall often use subscripts on operators to remind the reader of the number of variables of the functions or distributions the operator applies to.

Using the expression for the volume of the Stiefel manifold from [15, p. 9], we simplify the constant in (5.2) to

$$(5.3) \quad c_{d,k} = \frac{1}{(2\pi)^k} \frac{|\mathbb{S}^{k-1}|}{|\mathbb{S}^{d-k-1}|} \frac{1}{\prod_{n=k}^{d-1} |\mathbb{S}^{n-1}|}.$$

We then strengthen the result of Theorem 5.1 by invoking an extended version of the open mapping theorem, where we use the key property that the range of the  $k$ -plane transform  $\mathcal{S}_k = \mathcal{R}_k(\mathcal{S}(\mathbb{R}^d))$  is a closed subspace of the Fréchet space  $\mathcal{S}(\Xi_k)$  [38, Chapter 4].

<sup>1</sup>In computed tomography (CT), this filter is referred to as the backprojection filter found in the filtered backprojection algorithm for CT image reconstruction.

COROLLARY 5.3. *The operator  $\mathcal{R}_k : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}_k$  is a homeomorphism with inverse  $\mathcal{R}_k^{-1} = \mathcal{R}_k^* \mathbf{K}_{d-k} : \mathcal{S}_k \rightarrow \mathcal{S}(\mathbb{R}^d)$ .*

*Proof.* From Theorem 5.1, we know that  $\mathcal{R}_k : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}_k$  is a continuous bijection with inverse given by  $\mathcal{R}_k^{-1} = \mathcal{R}_k^* \mathbf{K}_{d-k}$ . Since  $\mathcal{S}(\mathbb{R}^d)$  and  $\mathcal{S}_k$  are Fréchet spaces, where  $\mathcal{S}_k$  inherits its Fréchet topology from  $\mathcal{S}(\Xi_k)$ , by the open mapping theorem for Fréchet spaces [57, Theorem 2.11], we know that  $\mathcal{R}_k$  is an open map. Therefore, for  $\mathcal{R}_k^* \mathbf{K}_{d-k}$ , the inverse images of open sets are open, which implies the continuity of  $\mathcal{R}_k^* \mathbf{K}_{d-k}$ .  $\square$

Let  $\mathcal{S}_H$  denote the subspace of Schwartz functions on  $\Xi_k$  that satisfy the so-called Helgason consistency conditions (see [38, Chapter 4] for a precise definition). In the case of the Radon transform, these consistency conditions exactly characterize its range ( $\mathcal{S}_H = \mathcal{S}_{d-1}$ ). But, when  $k < (d - 1)$ , it turns out that  $\mathcal{S}_H \neq \mathcal{S}_k$  [22]. In fact,  $\mathcal{S}_k$  is a subspace of  $\mathcal{S}_H$  whose members satisfy a system of PDEs [22, 38]. By considering a different set of consistency conditions, Petrov could show in [71, 72] (see [46] for an English translation of [71]) that the range has an equivalent characterization which does not explicitly involve PDEs [71, Remark 6]. This is summarized in Theorem 5.4.

THEOREM 5.4 ([46, 71, 72]). *A function  $\psi$  is the  $k$ -plane transform of a function  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  (i.e.,  $\psi \in \mathcal{S}_k$ ) if and only if the following conditions hold:*

1.  $\psi \in \mathcal{S}(\Xi_k)$ ;
2.  $\psi(\mathbf{A}, \mathbf{t}) = \psi(\mathbf{U}\mathbf{A}, \mathbf{U}\mathbf{t})$  for any  $\mathbf{U} \in \mathbf{O}_{d-k}(\mathbb{R})$ , where  $\mathbf{O}_{d-k}(\mathbb{R})$  denotes the orthogonal group in dimension  $(d - k)$ ;
3. for  $n = 1, \dots, (d - k)$  and  $m \in \mathbb{N}_0$ ,

$$(5.4) \quad \int_{\mathbb{R}} \psi(\mathbf{A}, \mathbf{t}) t_n^m dt_n$$

*is a degree- $m$  homogeneous polynomial with respect to  $\alpha_n$  (the  $n$ th row of  $\mathbf{A}$ ).*

Item 2 is the  $k$ -plane analogue of the even symmetry imposed by the Radon transform. We refer to this as the *isotropic symmetry* of the  $k$ -plane transform since it says that, under the present parameterization, the  $k$ -plane transform of a function is isotropic in its variables.

**6. Distributional extension.** When  $k = (d - 1)$ , there are three different ways to extend the Radon transform to distributions. These are the formulations of (i) Gelfand, Graev, and Vilenkin [20]; (ii) Helgason [27]; and (iii) Ludwig [37]. It turns out that these three approaches are equivalent [48]. Here, for general  $k$ ,  $1 \leq k < d$ , we shall extend the  $k$ -plane transform to distributions by duality, taking inspiration from Ludwig [37]. In section 7, we shall also make explicit that our definitions are equivalent to Fourier-based definitions understood via the Fourier slice theorem, similar to those investigated in [48] for the Radon transform.

In the remainder of this paper,  $\langle \cdot, \cdot \rangle$  denotes the pairing of a space and its continuous dual defined on  $\mathbb{R}^d$  while  $\langle \cdot, \cdot \rangle_k$  denotes the pairing of a space and its continuous dual defined on  $\Xi_k$ . By invoking the property from Theorem 5.1 that  $\mathcal{R}_k^* \mathbf{K}_{d-k} \mathcal{R}_k = \text{Id}$  on  $\mathcal{S}(\mathbb{R}^d)$  and noting that  $\mathbf{K}_{d-k}$  is self-adjoint (since its frequency response is real-valued), we see that, for all  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  and  $f \in \mathcal{S}'(\mathbb{R}^d)$ ,

$$(6.1) \quad \langle f, \varphi \rangle = \langle f, \mathcal{R}_k^* \mathbf{K}_{d-k} \mathcal{R}_k \varphi \rangle = \langle \mathcal{R}_k f, \mathbf{K}_{d-k} \mathcal{R}_k \varphi \rangle_k = \langle \mathcal{R}_k f, \psi \rangle_k,$$

where  $\psi = \mathbf{K}_{d-k} \mathcal{R}_k \{\varphi\} \in \mathbf{K}_{d-k} \mathcal{R}_k(\mathcal{S}(\mathbb{R}^d))$  and so  $\varphi = \mathcal{R}_k \{\psi\}$ . Likewise, we have that

$$(6.2) \quad \langle f, \varphi \rangle = \langle f, \mathcal{R}_k^* \mathbf{K}_{d-k} \mathcal{R}_k \varphi \rangle = \langle \mathbf{K}_{d-k} \mathcal{R}_k f, \mathcal{R}_k \varphi \rangle_k = \langle \mathbf{K}_{d-k} \mathcal{R}_k f, \tilde{\psi} \rangle_k,$$

where  $\tilde{\psi} = \mathcal{R}_k\{\varphi\} \in \mathcal{S}_k$  and so  $\varphi = \mathcal{R}_k^* \mathbf{K}_{d-k}\{\tilde{\psi}\}$ . This manipulation reveals that, in order to have a well-defined distributional extension of these operators, some care has to be taken regarding their domain and range. This is summarized in Definition 6.1.

DEFINITION 6.1.

1. The distributional  $k$ -plane transform

$$(6.3) \quad \mathcal{R}_k : \mathcal{S}'(\mathbb{R}^d) \rightarrow (\mathbf{K}_{d-k} \mathcal{R}_k(\mathcal{S}(\mathbb{R}^d)))'$$

is defined to be the dual map of the homeomorphism  $\mathcal{R}_k^* : \mathbf{K}_{d-k} \mathcal{R}_k(\mathcal{S}(\mathbb{R}^d)) \rightarrow \mathcal{S}(\mathbb{R}^d)$ .

2. The distributional filtered  $k$ -plane transform

$$(6.4) \quad \mathbf{K}_{d-k} \mathcal{R}_k : \mathcal{S}'(\mathbb{R}^d) \rightarrow \mathcal{S}'_k$$

is defined to be the dual map of the homeomorphism  $\mathcal{R}_k^* \mathbf{K}_{d-k} : \mathcal{S}_k \rightarrow \mathcal{S}(\mathbb{R}^d)$ .

3. The distributional backprojection

$$(6.5) \quad \mathcal{R}_k^* : \mathcal{S}'_k \rightarrow \mathcal{S}'(\mathbb{R}^d)$$

is defined to be the dual map of the homeomorphism  $\mathcal{R}_k : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}_k$ .

4. The extended distributional backprojection

$$(6.6) \quad \mathcal{R}_k^* : \mathcal{S}'(\Xi_k) \rightarrow \mathcal{S}'(\mathbb{R}^d)$$

is defined to be the dual map of the continuous operator  $\mathcal{R}_k : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\Xi_k)$ , which is well-defined since  $\mathcal{S}_k$  is continuously embedded in  $\mathcal{S}(\Xi_k)$ .

Based on these definitions, we have the following result on the invertibility of the filtered  $k$ -plane transform on  $\mathcal{S}'(\mathbb{R}^d)$ , which is the dual of Theorem 5.1 and Corollary 5.3.

THEOREM 6.2. *It holds that  $\mathcal{R}_k^* \mathbf{K}_{d-k} \mathcal{R}_k = \text{Id}$  on  $\mathcal{S}'(\mathbb{R}^d)$ . Moreover, the filtered  $k$ -plane transform  $\mathbf{K}_{d-k} \mathcal{R}_k : \mathcal{S}'(\mathbb{R}^d) \rightarrow \mathcal{S}'_k$  is a homeomorphism with inverse given by the backprojection  $(\mathbf{K}_{d-k} \mathcal{R}_k)^{-1} = \mathcal{R}_k^* : \mathcal{S}'_k \rightarrow \mathcal{S}'(\mathbb{R}^d)$ .*

While Definition 6.1 properly specifies the distributional  $k$ -plane transform, filtered  $k$ -plane transform, and backprojection operators, the construction is inherently abstract in the sense that the underlying spaces  $\mathcal{S}'_k$  and  $(\mathbf{K}_{d-k} \mathcal{R}_k(\mathcal{S}(\mathbb{R}^d)))'$  are equivalence classes of distributions. This means that the distributional extensions of these operators lack unicity, except for the backprojection. In the case of the extended backprojection operator, we have a unique, concrete characterization of the backprojection of any distribution in  $\mathcal{S}'(\Xi_k)$ . Specifically,  $f \in \mathcal{S}'(\mathbb{R}^d)$  is defined to be the backprojection of  $g \in \mathcal{S}'(\Xi_k)$  (i.e.,  $f = \mathcal{R}_k^*\{g\}$ ) if

$$(6.7) \quad \langle f, \varphi \rangle = \langle g, \mathcal{R}_k\{\varphi\} \rangle_k$$

for all  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ .

Conversely, the range  $\mathcal{S}'_k$  of the distributional filtered  $k$ -plane transform can be identified with the abstract quotient space  $\mathcal{S}'(\Xi_k)/\mathcal{N}_{\mathcal{R}_k^*}$ , where

$$(6.8) \quad \mathcal{N}_{\mathcal{R}_k^*} = \{g \in \mathcal{S}'(\Xi_k) : \mathcal{R}_k^*\{g\} = 0 \Leftrightarrow \langle g, \tilde{\psi} \rangle_k = 0 \text{ for all } \tilde{\psi} \in \mathcal{S}_k\}$$

denotes the null space of the extended backprojection operator in Definition 6.1. It is important to note that this null space is huge and contains many exotic functions/distributions (see [37] for a characterization when  $k = (d - 1)$ ).

This nonuniqueness poses an issue when one needs to deal with practical applications. To this end, in section 8 we identify Banach subspaces of  $\mathcal{S}'_k$  from which we can systematically extract unique, concrete representers from the abstract equivalence classes via continuous projection operators. These  $k$ -plane-compatible Banach subspaces are such that the backprojection operator is guaranteed to be invertible.

For completeness, observe that when  $k$  is even, we analogously have that the range  $(\mathbb{K}_{d-k} \mathcal{R}_k(\mathcal{S}(\mathbb{R}^d)))' = (\mathbb{K}_{d-k}(\mathcal{S}_k))'$  of the distributional  $k$ -plane transform can be identified with the abstract quotient space  $\mathcal{S}'(\Xi_k)/\mathcal{N}_{\mathcal{R}_k^* \mathbb{K}_{d-k}}$ , where

$$(6.9) \quad \mathcal{N}_{\mathcal{R}_k^* \mathbb{K}_{d-k}} = \{g \in \mathcal{S}'(\Xi_k) : \mathcal{R}_k^* \mathbb{K}_{d-k} \{g\} = 0 \Leftrightarrow \langle g, \psi \rangle_k = 0 \text{ for all } \psi \in \mathbb{K}_{d-k}(\mathcal{S}_k)\}.$$

Here, we took advantage of the fact that when  $k$  is even,  $\mathbb{K}_{d-k} = c_{d,k}(-\Delta_{d-k})^{\frac{k}{2}}$  is a local operator. In this scenario,  $\mathbb{K}_{d-k}(\mathcal{S}_k)$  is continuously embedded in  $\mathcal{S}(\Xi_k)$ . Thus, we can define the extended distributional filtered backprojection

$$(6.10) \quad \mathcal{R}_k^* \mathbb{K}_{d-k} : \mathcal{S}'(\Xi_k) \rightarrow \mathcal{S}'(\mathbb{R}^d)$$

as the dual map of the continuous operator  $\mathbb{K}_{d-k} \mathcal{R}_k : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\Xi_k)$ . By contrast, when  $k$  is odd, such a quotient-space characterization of  $(\mathbb{K}_{d-k}(\mathcal{S}_k))'$  cannot exist. Indeed, in that scenario,  $\mathbb{K}_{d-k}$  is a global operator and so  $\mathbb{K}_{d-k}(\mathcal{S}_k) \not\subset \mathcal{S}(\Xi_k)$ .

**7. The Fourier slice theorem.** An important property of the  $k$ -plane transform is the Fourier slice theorem. It relates the Fourier transform of the  $k$ -plane transform of a function to its Fourier transform along *slices*. While this result has been reported many times in the literature (see [38] and references therein), to the best of our knowledge, it has not yet been established for the parameterization of the  $k$ -plane transform given in (4.6) and (4.7). We first rederive the theorem for Schwartz functions with our choice of parameterization and then extend the scope of the Fourier slice theorem to tempered distributions in full generality.

**7.1. The Fourier slice theorem for Schwartz functions.** In this section we begin by defining the notion of a  $k$ -plane ridge function. With this definition, we first prove a more general result (Theorem 7.2) and then retrieve the Fourier slice theorem as a special case.

**DEFINITION 7.1.** Given  $r \in L^1_{\text{temp}}(\mathbb{R}^{d-k})$  and  $\mathbf{A} \in V_{d-k}(\mathbb{R}^d)$ , the  $k$ -plane ridge function  $r_{\mathbf{A}} \in L^1_{\text{temp}}(\mathbb{R}^d)$  is defined by

$$(7.1) \quad r_{\mathbf{A}}(\mathbf{x}) := r(\mathbf{A}\mathbf{x}).$$

The function  $r$  is referred to as the profile and the matrix  $\mathbf{A}$  is referred to as the orientation of the  $k$ -plane ridge  $r_{\mathbf{A}}$ , which is constant along the  $k$ -planes  $\mathbf{A}\mathbf{x} = \mathbf{c}$ , where  $\mathbf{c} \in \mathbb{R}^{d-k}$ .

**THEOREM 7.2.** Given  $r \in L^1_{\text{temp}}(\mathbb{R}^{d-k})$  and  $\mathbf{A} \in V_{d-k}(\mathbb{R}^d)$ , we have the equality

$$(7.2) \quad \langle r_{\mathbf{A}}, \varphi \rangle = \langle r, \mathcal{R}_k \{\varphi\}(\mathbf{A}, \cdot) \rangle$$

for all  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ .

*Proof.* The result follows directly from the Fubini–Tonelli theorem. Indeed, we have that



$$\begin{aligned}
(7.3) \quad \langle r_{\mathbf{A}}, \varphi \rangle &= \int_{\mathbb{R}^d} r_{\mathbf{A}}(\mathbf{x}) \varphi(\mathbf{x}) \, d\mathbf{x} \\
&= \int_{\mathbb{R}^{d-k}} r(\mathbf{A}\mathbf{A}^\top \mathbf{t}) \int_{\mathbf{A}^\perp} \varphi(\mathbf{y} + \mathbf{A}^\top \mathbf{t}) \, d\mathbf{y} \, d\mathbf{t} \\
&= \int_{\mathbb{R}^{d-k}} r(\mathbf{t}) \mathcal{R}_k\{\varphi\}(\mathbf{A}, \mathbf{t}) \, d\mathbf{t} = \langle r, \mathcal{R}_k\{\varphi\}(\mathbf{A}, \cdot) \rangle
\end{aligned}$$

under the change of variables  $\mathbf{x} = \mathbf{y} + \mathbf{A}^\top \mathbf{t}$  with  $\mathbf{y} \in \mathbf{A}^\perp$  and  $\mathbf{t} \in \mathbb{R}^{d-k}$ .  $\square$

Given  $\mathbf{A} \in V_{d-k}(\mathbb{R}^d)$ , we see from Theorem 5.4 that  $\mathcal{R}_k\{\varphi\}(\mathbf{A}, \cdot) \in \mathcal{S}(\mathbb{R}^{d-k})$  for all  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . Therefore, Theorem 7.2 motivates the following *definition* of a  $k$ -plane ridge whose profile does not admit a pointwise interpretation.

**DEFINITION 7.3.** *Given the profile  $r \in \mathcal{S}'(\mathbb{R}^{d-k})$  and orientation  $\mathbf{A} \in V_{d-k}(\mathbb{R}^d)$ , the  $k$ -plane ridge  $r_{\mathbf{A}} \in \mathcal{S}'(\mathbb{R}^d)$  is the tempered distribution defined by*

$$(7.4) \quad \langle r_{\mathbf{A}}, \varphi \rangle := \langle r, \mathcal{R}_k\{\varphi\}(\mathbf{A}, \cdot) \rangle$$

for all  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ .

*Remark 7.4.* The formulation of the  $k$ -plane transform in (4.7) can be understood via this definition.

Under the present parameterization of the  $k$ -plane transform, the following corollary of Theorem 7.2 is our reformulation of the Fourier slice theorem on  $\mathcal{S}(\mathbb{R}^d)$ .

**COROLLARY 7.5.** *Given  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ , we have, for any  $\mathbf{A} \in V_{d-k}(\mathbb{R}^d)$ , that*

$$(7.5) \quad \mathcal{F}_{d-k}\{(\mathcal{R}_k\varphi)(\mathbf{A}, \cdot)\}(\boldsymbol{\omega}) = \mathcal{F}_d\{\varphi\}(\mathbf{A}^\top \boldsymbol{\omega}), \quad \boldsymbol{\omega} \in \mathbb{R}^{d-k}.$$

*Proof.* Choose  $r(\mathbf{t}) = e^{-i\boldsymbol{\omega}^\top \mathbf{t}}$  so that  $r_{\mathbf{A}}(\mathbf{x}) = e^{-i\boldsymbol{\omega}^\top \mathbf{A}\mathbf{x}} = e^{-i(\mathbf{A}^\top \boldsymbol{\omega})^\top \mathbf{x}}$  for any  $\mathbf{A} \in V_{d-k}(\mathbb{R}^d)$ . The result then follows for any  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  by the application of Theorem 7.2.  $\square$

*Remark 7.6.* This formulation of the Fourier slice theorem sets the stage for the investigation of the bijective isometry properties of the  $k$ -plane transform between  $L^2$ -Sobolev spaces. This has recently been established in full generality for the Radon transform ( $k = (d-1)$ ) [60]. The proofs of [60] crucially rely on the Fourier slice theorem. This suggests that their formulation can also be extended to the  $k$ -plane transform.

**7.2. The Fourier slice theorem for tempered distributions.** The general result for tempered distributions will be deduced by duality. To that end, we first define the operator  $F_{d,k}$  that acts on  $\mathcal{S}(\mathbb{R}^d)$  by

$$(7.6) \quad F_{d,k}\{\varphi\}(\mathbf{A}, \boldsymbol{\omega}) := \widehat{\varphi}(\mathbf{A}^\top \boldsymbol{\omega}) = \mathcal{F}_{d-k}\{(\mathcal{R}_k\varphi)(\mathbf{A}, \cdot)\}(\boldsymbol{\omega}) = \mathcal{F}_{d-k} \mathcal{R}_k\{\varphi\},$$

where the second equality is from Corollary 7.5. The range of  $F_{d,k}$  is the space  $\mathcal{F}_{d-k}(\mathcal{S}_k)$ . The operator is invertible with its inverse being

$$(7.7) \quad F_{d,k}^{-1} = \mathcal{R}_k^* K_{d-k} \mathcal{F}_{d-k}^{-1}.$$

Its adjoint is given by

$$(7.8) \quad F_{d,k}^* = \mathcal{R}_k^* \mathcal{F}_{d-k}^* = \frac{1}{(2\pi)^{d-k}} \mathcal{R}_k^* \mathcal{F}_{d-k}^{-1}.$$

Since  $\mathcal{R}_k^*$  is a homeomorphism from  $K_{d-k} \mathcal{R}_k(\mathcal{S}(\mathbb{R}^d)) \rightarrow \mathcal{S}(\mathbb{R}^d)$ , we have that

$$(7.9) \quad F_{d,k}^* : \mathcal{F}_{d-k} K_{d-k} \mathcal{R}_k(\mathcal{S}(\mathbb{R}^d)) \rightarrow \mathcal{S}(\mathbb{R}^d)$$

is a homeomorphism. This justifies the extension of  $F_{d,k}$  by duality,

$$(7.10) \quad F_{d,k} : \mathcal{S}'(\mathbb{R}^d) \rightarrow (\mathcal{F}_{d-k} K_{d-k} \mathcal{R}_k(\mathcal{S}(\mathbb{R}^d)))',$$

which is itself a homeomorphism. Next, from (7.6), we have the identity

$$(7.11) \quad \mathcal{R}_k = \mathcal{F}_{d-k}^{-1} F_{d,k} \quad \text{on } \mathcal{S}(\mathbb{R}^d).$$

The combination of (7.11) with Theorem 5.1 yields the identity

$$(7.12) \quad K_{d-k} \mathcal{R}_k = \mathcal{F}_{d-k}^{-1} \widehat{K}_{d-k} F_{d,k} \quad \text{on } \mathcal{S}(\mathbb{R}^d),$$

where  $\widehat{K}_{d-k}$  denotes multiplication in the  $\omega$  variable by the frequency response  $\widehat{K}_{d-k}(\omega) = c_{d,k} \|\omega\|_2^k$  of the filtering operator  $K_{d-k}$ . From Theorem 5.1,  $(\mathcal{R}_k^*)^{-1} = K_{d-k} \mathcal{R}_k$  on  $\mathcal{S}(\mathbb{R}^d)$ . Therefore, this also establishes the identity

$$(7.13) \quad \mathcal{R}^* = F_{d,k}^{-1} \widehat{K}_{d-k}^{-1} \mathcal{F}_{d-k} \quad \text{on } K_{d-k} \mathcal{R}_k(\mathcal{S}(\mathbb{R}^d)),$$

where  $\widehat{K}_{d-k}^{-1}$  denotes multiplication in the  $\omega$  variable by the frequency response  $\widehat{K}_{d-k}^{-1}(\omega) = c_{d,k}^{-1} \|\omega\|_2^{-k}$  of the inverse filtering operator  $K_{d-k}^{-1}$ .

Equation (7.13) is a new formula for the dual transform. In particular, if we fix  $f \in \mathcal{S}'(\mathbb{R}^d)$ , for any  $\psi \in K_{d-k} \mathcal{R}_k(\mathcal{S}(\mathbb{R}^d))$  we have

$$(7.14) \quad \langle \mathcal{R}_k f, \psi \rangle_k = \langle f, F_{d,k}^{-1} \widehat{K}_{d-k}^{-1} \mathcal{F}_{d-k} \psi \rangle.$$

By noticing that  $F_{d,k}^{-1} \widehat{K}_{d-k}^{-1} = \mathcal{R}_k^* K_{d-k} \mathcal{F}_{d-k}^{-1} \widehat{K}_{d-k}^{-1} = \mathcal{R}_k^* \mathcal{F}_{d-k}^{-1} = (2\pi)^{d-k} F_{d,k}^*$ , we have that

$$(7.15) \quad \langle \mathcal{R}_k f, \psi \rangle_k = (2\pi)^{d-k} \langle F_{d,k} f, \mathcal{F}_{d-k} \psi \rangle_k$$

for any  $\psi \in K_{d-k} \mathcal{R}_k(\mathcal{S}(\mathbb{R}^d))$ , where the right-hand side is well-defined from (7.10). An immediate consequence to (7.15) is the Fourier slice theorem for any  $f \in \mathcal{S}'(\mathbb{R}^d)$ .

**THEOREM 7.7.** *Let  $f \in \mathcal{S}'(\mathbb{R}^d)$ . Then,*

$$(7.16) \quad \mathcal{F}_{d-k} \mathcal{R}_k f = F_{d,k} f,$$

where the equality holds in  $(\mathcal{F}_{d-k} K_{d-k} \mathcal{R}_k(\mathcal{S}(\mathbb{R}^d)))'$ .

When  $f \in \mathcal{S}(\mathbb{R}^d)$ , the equality in (7.16) is precisely the equality in (7.5) and so Theorem 7.7 recovers Corollary 7.5. For general tempered distributions, the identity in Theorem 7.7 has to be interpreted with care since  $(K_{d-k} \mathcal{R}_k(\mathcal{S}(\mathbb{R}^d)))'$  is actually a space of equivalence classes of distributions. This is to say that, given  $f \in \mathcal{S}'(\mathbb{R}^d)$ ,  $\mathcal{R}_k\{f\}$  is, in general, nonunique. This is illustrated concretely in [50, Example 10.3.2], which provides an explicit calculation via the Fourier slice theorem to show that the Radon transform ( $k = (d - 1)$ ) of the function  $f \equiv 1 \in \mathcal{S}'(\mathbb{R}^d)$  is not uniquely determined.

The key takeaway message from the Fourier slice theorem (Corollary 7.5 and Theorem 7.7) is that the restriction of  $\mathcal{F}_d\{f\}$  to the slice  $\{\mathbf{A}^\top \omega \in \mathbb{R}^d : \omega \in \mathbb{R}^{d-k}\}$  is the  $(d - k)$ -variate Fourier transform of  $\mathcal{R}_k\{f\}(\mathbf{A}, \cdot)$ . Moreover, a further consequence of this theorem is a general form of the *intertwining properties* of the  $k$ -plane transform. Before we state our result, recall that the necessary and sufficient condition for an LSI operator  $L_{\text{rad}}$  to continuously map  $\mathcal{S}(\mathbb{R}) \rightarrow \mathcal{S}(\mathbb{R})$  (or, by duality, continuously map  $\mathcal{S}'(\mathbb{R}) \rightarrow \mathcal{S}'(\mathbb{R})$ ) is that  $\widehat{L}_{\text{rad}} \in \mathcal{O}_M(\mathbb{R})$ , where  $\mathcal{O}_M(\mathbb{R})$  is the space of smooth functions of slow growth on  $\mathbb{R}$  [59, p. 243].

COROLLARY 7.8. Let  $L_{d-k} : \mathcal{S}'(\mathbb{R}^{d-k}) \rightarrow \mathcal{S}'(\mathbb{R}^{d-k})$  and  $L_d : \mathcal{S}'(\mathbb{R}^d) \rightarrow \mathcal{S}'(\mathbb{R}^d)$  be isotropic LSI operators that share the same radial frequency profile. Their frequency responses satisfy

$$(7.17) \quad \widehat{L}_{d-k}(\boldsymbol{\omega}) = \widehat{L}_{\text{rad}}(\|\boldsymbol{\omega}\|_2), \quad \boldsymbol{\omega} \in \mathbb{R}^{d-k},$$

and

$$(7.18) \quad \widehat{L}_d(\boldsymbol{\xi}) = \widehat{L}_{\text{rad}}(\|\boldsymbol{\xi}\|_2), \quad \boldsymbol{\xi} \in \mathbb{R}^d,$$

where  $\widehat{L}_{\text{rad}} \in \mathcal{O}_M(\mathbb{R})$  is their shared radial frequency profile. Then, we have the intertwining properties

$$(7.19) \quad L_{d-k} \mathcal{R}_k = \mathcal{R}_k L_d \quad \text{on } \mathcal{S}'(\mathbb{R}^d)$$

and

$$(7.20) \quad \mathcal{R}_k^* L_{d-k}^* = L_d^* \mathcal{R}_k^* \quad \text{on } \mathcal{S}'_k.$$

*Proof.* With a slight abuse of notation, by Theorem 7.7, one has that

$$(7.21) \quad \begin{aligned} \mathcal{F}_{d-k} L_{d-k} \mathcal{R}_k &= \widehat{L}_{d-k}(\boldsymbol{\omega}) \mathcal{F}_{d-k} \mathcal{R}_k \\ &= \widehat{L}_{\text{rad}}(\|\boldsymbol{\omega}\|_2) F_{d,k} \\ &= \widehat{L}_{\text{rad}}(\|\mathbf{A}^\top \boldsymbol{\omega}\|_2) F_{d,k} \\ &= \widehat{L}_d(\mathbf{A}^\top \boldsymbol{\omega}) F_{d,k} \\ &= F_{d,k} L_d \\ &= \mathcal{F}_{d-k} \mathcal{R}_k L_d. \end{aligned}$$

Furthermore, since (7.19) holds on  $\mathcal{S}(\mathbb{R}^d)$ , by duality, we immediately have an intertwining property for the dual transform on  $\mathcal{S}'_k$  as

$$(7.22) \quad \mathcal{R}_k^* L_{d-k}^* = L_d^* \mathcal{R}_k^* \quad \text{on } \mathcal{S}'_k. \quad \square$$

**8. Systematic construction of  $k$ -plane-compatible spaces.** The distributional extension of the  $k$ -plane transform and related operators in section 6 provides a definition of the filtered  $k$ -plane transform of any  $f \in \mathcal{S}'(\mathbb{R}^d)$ . It also provides a way to invert it via the backprojection operator. So far, this is not very practical since the range of the distributional filtered  $k$ -plane transform  $\mathcal{S}'_k$  is a space of equivalence classes of distributions which has been identified as the quotient  $\mathcal{S}'(\Xi_k)/\mathcal{N}_{\mathcal{R}_k^*}$ . To make the framework more applicable, we now provide a systematic construction (via a completion process) of Banach subspaces of  $\mathcal{S}'_k$  for which the backprojection is invertible and whose elements are identified as ordinary functions or measures. We shall refer to these as  $k$ -plane-compatible (Banach) spaces. We then present some examples where, in particular, we can identify unique, concrete representers from the equivalence classes in  $\mathcal{S}'_k \cong \mathcal{S}'(\Xi_k)/\mathcal{N}_{\mathcal{R}_k^*}$  via continuous projection operators.

Let  $\mathcal{X}$  be a Banach space on  $\Xi_k$  such that

$$(8.1) \quad \mathcal{S}(\Xi_k) \xrightarrow{d} \mathcal{X} \xrightarrow{d} \mathcal{S}'(\Xi_k),$$

the corresponding dual pair of  $k$ -plane-compatible Banach spaces being

$$(8.2) \quad (\mathcal{X}_k, \mathcal{X}'_k),$$

with  $\mathcal{X}_k := \overline{(\mathcal{S}_k, \|\cdot\|_{\mathcal{X}})}$ , and for which we have the following powerful result.

THEOREM 8.1. *Let  $(\mathcal{X}_k, \mathcal{X}'_k)$  be as above. Then, the following holds.*

1. *The backprojection  $\mathcal{R}_k^* : \mathcal{X}'_k \rightarrow \mathcal{S}'(\mathbb{R}^d)$  is injective and  $K_{d-k} \mathcal{R}_k \mathcal{R}_k^* = \text{Id}$  on  $\mathcal{X}'_k$ .*
2. *Let  $\mathcal{Y} = (\mathcal{S}(\mathbb{R}^d), \|\cdot\|_{\mathcal{Y}})$  where, for  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ ,  $\|\varphi\|_{\mathcal{Y}} := \|\mathcal{R}_k \varphi\|_{\mathcal{X}}$ . Then, there exists a unique extension  $\mathcal{R}_k^* K_{d-k} : \mathcal{X}_k \rightarrow \mathcal{Y}$  that is an isometric isomorphism such that  $\mathcal{R}_k \mathcal{R}_k^* K_{d-k} = \text{Id}$  on  $\mathcal{X}_k$ .*
3. *The range spaces  $\mathcal{Y} = \mathcal{R}_k^* K_{d-k}(\mathcal{X}_k)$  and  $\mathcal{V} = \mathcal{R}_k^*(\mathcal{X}'_k)$  form a dual pair of Banach spaces (i.e., we can identify  $\mathcal{Y}' = \mathcal{V}$ ) such that  $(\mathcal{Y}, \mathcal{V})$  is isometrically isomorphic to  $(\mathcal{X}_k, \mathcal{X}'_k)$ .*

Moreover, if there exists a complementary Banach space  $\mathcal{X}_k^c$  such that  $\mathcal{X} = \mathcal{X}_k \oplus \mathcal{X}_k^c$ , then the following holds.

4. *The dual space is decomposable as  $\mathcal{X}' = \mathcal{X}'_k \oplus (\mathcal{X}_k^c)'$ .*
5. *The dual complement space  $(\mathcal{X}_k^c)'$  is the null space of  $\mathcal{R}_k^* : \mathcal{X}' \rightarrow \mathcal{Y}'$ .*
6. *The complement space  $\mathcal{X}_k^c$  is the null space of  $\mathcal{R}_k^* K_{d-k} : \mathcal{X} \rightarrow \mathcal{Y}$ .*
7. *The operators  $P_k = \mathcal{R}_k \mathcal{R}_k^* K_{d-k} : \mathcal{X} \rightarrow \mathcal{X}_k$  and  $P_k^* = K_{d-k} \mathcal{R}_k \mathcal{R}_k^* : \mathcal{X}' \rightarrow \mathcal{X}'_k$  form an adjoint pair of continuous projectors with  $P_k(\mathcal{X}) = \mathcal{X}_k$  and  $P_k^*(\mathcal{X}') = \mathcal{X}'_k$ .*

*Proof.*

1. Since  $\mathcal{S}_k \xrightarrow{\text{d.}} \mathcal{X}_k$ , by duality we have that  $\mathcal{X}'_k \xrightarrow{\text{c.}} \mathcal{S}'_k$ . Let  $\mathcal{V} = \mathcal{R}_k^*(\mathcal{X}'_k) \subset \mathcal{S}'(\mathbb{R}^d)$  denote the range of the restricted operator  $\mathcal{R}_k^*|_{\mathcal{X}'_k}$ . Clearly this restriction is a homeomorphism from  $\mathcal{X}'_k \rightarrow \mathcal{V}$ , where  $\mathcal{V}$  is a Banach space equipped with the norm

$$(8.3) \quad \|f\|_{\mathcal{V}} := \|\mathcal{R}_k^{*-1}\{f\}\|_{\mathcal{X}'_k}, \quad f \in \mathcal{V}.$$

In other words,  $\mathcal{R}_k^*$  is an isometric isomorphism from  $\mathcal{X}'_k$  to  $\mathcal{V}$ . In particular, this shows that  $\mathcal{R}_k^* : \mathcal{X}'_k \rightarrow \mathcal{S}'(\mathbb{R}^d)$  is an injection. Moreover, by the inversion formula in Theorem 6.2, the inverse of  $\mathcal{R}_k^*|_{\mathcal{X}'_k}$  on  $\mathcal{V}$  is given by the restriction of  $K_{d-k} \mathcal{R}_k$  to  $\mathcal{V}$ , denoted  $K_{d-k} \mathcal{R}_k|_{\mathcal{V}}$ . Therefore,  $K_{d-k} \mathcal{R}_k \mathcal{R}_k^* = \text{Id}$  on  $\mathcal{X}'_k$ .

2. From the proof of item 1, we have that  $K_{d-k} \mathcal{R}_k : \mathcal{V} \rightarrow \mathcal{X}'_k$  is an isometric isomorphism and, hence, it is continuous by design. By duality combined with the isometric embedding of a Banach space in its bidual, we have that  $\mathcal{R}_k^* K_{d-k} : \mathcal{X}_k \rightarrow \mathcal{V}'$  is continuous as well. We now prove that it is injective by showing that it has a trivial null space. For any  $\psi \in \mathcal{S}_k$ , the statement  $\mathcal{R}_k^* K_{d-k}\{\psi\} = 0$  is equivalent to

$$(8.4) \quad 0 = \langle v, \mathcal{R}_k^* K_{d-k}\{\psi\} \rangle = \langle K_{d-k} \mathcal{R}_k\{v\}, \psi \rangle_k \quad \text{for all } v \in \mathcal{V}.$$

Since  $K_{d-k} \mathcal{R}_k(\mathcal{V}) = \mathcal{X}'_k$ , (8.4) is equivalent to  $\langle g, \psi \rangle_k = 0$  for all  $g \in \mathcal{X}'_k$ , which implies  $\psi = 0$ . The result then follows from the dense embeddings  $\mathcal{S}_k \xrightarrow{\text{d.}} \mathcal{X}_k$  and  $\mathcal{S}(\mathbb{R}^d) \xrightarrow{\text{d.}} \mathcal{Y}$  and the continuity of the inversion formula for the  $k$ -plane transform. Indeed, these dense embeddings imply that the definition of  $\mathcal{Y}$  is equivalent to

$$(8.5) \quad \mathcal{Y} = \{\mathcal{R}_k^{-1}\{g\} : g \in \mathcal{X}_k\},$$

where  $\mathcal{R}_k^{-1} = \mathcal{R}_k^* K_{d-k}$  is continuous on  $\mathcal{S}_k$  and can be continuously extended, by density, to act on  $\mathcal{X}_k$ . Therefore,  $\mathcal{Y} = \mathcal{R}_k^* K_{d-k}(\mathcal{X}_k)$ . Since  $\mathcal{R}_k^* K_{d-k}$  is continuous and injective on  $\mathcal{X}_k$ , by the bounded inverse theorem [57], it has a continuous inverse on its range that maps  $\mathcal{Y} \rightarrow \mathcal{X}_k$ . Putting everything together, from the inversion formula for the  $k$ -plane transform and the definition

of the norm of  $\mathcal{Y}$ , we have that  $\mathcal{R}_k^* K_{d-k} : \mathcal{X}_k \rightarrow \mathcal{Y}$  is an isometric isomorphism with inverse given by  $(\mathcal{R}_k^* K_{d-k})^{-1} = \mathcal{R}_k : \mathcal{Y} \rightarrow \mathcal{X}_k$  and, so,  $\mathcal{R}_k \mathcal{R}_k^* K_{d-k} = \text{Id}$  on  $\mathcal{X}_k$ .

3. From item 2, we have that

$$(8.6) \quad \begin{aligned} \mathcal{R}_k^* K_{d-k} : \mathcal{X}_k &\rightarrow \mathcal{Y}, \\ \mathcal{R}_k : \mathcal{Y} &\rightarrow \mathcal{X}_k \end{aligned}$$

are bijective isometries. So, by duality, we have that

$$(8.7) \quad \begin{aligned} K_{d-k} \mathcal{R}_k : \mathcal{Y}' &\rightarrow \mathcal{X}'_k, \\ \mathcal{R}_k^* : \mathcal{X}'_k &\rightarrow \mathcal{Y}' \end{aligned}$$

are bijective isometries. From the proof of item 1, we know that  $\mathcal{X}'_k$  is isometrically isomorphic to  $\mathcal{V} = \mathcal{R}_k^*(\mathcal{X}'_k)$ . This means that  $\mathcal{V}$  is isometrically isomorphic to  $\mathcal{Y}'$ . Therefore, we can identify  $\mathcal{Y}' = \mathcal{V}$ .

4. This follows from the generic property that the dual of a direct sum is isometrically isomorphic to the direct sum of the duals.

5. Item 4 implies that  $(\mathcal{X}_k^c)'$  is the annihilator of  $\mathcal{X}_k$  in  $\mathcal{X}'$  and so

$$(8.8) \quad \begin{aligned} (\mathcal{X}_k^c)' &= \{g \in \mathcal{X}' : \langle g, \psi \rangle_k = 0 \text{ for all } \psi \in \mathcal{X}_k\} \\ &= \{g \in \mathcal{X}' : \langle g, \psi \rangle_k = 0 \text{ for all } \psi \in \mathcal{S}_k\}, \end{aligned}$$

where the second line follows from the denseness of  $\mathcal{S}_k$  in  $\mathcal{X}_k$ . From the definition of the null space in (6.8), we see that  $(\mathcal{X}_k^c)' \subset \mathcal{N}_{\mathcal{R}_k^*}$  is the null space of  $\mathcal{R}_k^* : \mathcal{X}' \rightarrow \mathcal{Y}'$ .

6. This result follows from item 5, by duality.

7. We first recall that existence (uniquity) and continuity of the projectors are guaranteed for any pair of complemented Banach spaces. Next, Items 1 and 2 imply that  $P_k^*(\mathcal{X}'_k) = \mathcal{X}'_k$  and  $P_k(\mathcal{X}_k) = \mathcal{X}_k$ . Similarly, items 5 and 6 imply that  $P_k^*((\mathcal{X}_k^c)') = \{0\}$  and  $P_k(\mathcal{X}_k^c) = \{0\}$ . Therefore,  $(P_k, P_k^*)$  form an adjoint pair of continuous projectors with  $P_k(\mathcal{X}) = \mathcal{X}_k$  and  $P_k^*(\mathcal{X}') = \mathcal{X}'_k$ .  $\square$

When there exists a complementary Banach space  $\mathcal{X}_k^c$  such that  $\mathcal{X} = \mathcal{X}_k \oplus \mathcal{X}_k^c$ , we depict the relevant Banach spaces from Theorem 8.1 and their relations in Figure 1. Our next theorem and corollary characterize some useful examples of  $k$ -plane-compatible Banach spaces whose elements are identifiable as concrete “isotropic” functions or measures (from the abstract equivalence classes in  $\mathcal{S}'_k \cong \mathcal{S}'(\Xi_k)/\mathcal{N}_{\mathcal{R}_k}$ ).

**THEOREM 8.2.** *Let  $\alpha_p$  denote the natural tensor norm for tensor products of  $L^p$  spaces and let  $\otimes_{\alpha_p}$  denote the completion of the algebraic tensor product with respect to the tensor norm  $\alpha_p$ . If  $\mathcal{X}$  is any of the Banach spaces*

- $L^p(\Xi_k) = L^p(V_{d-k}(\mathbb{R}^d) \otimes_{\alpha_p} L^p(\mathbb{R}^{d-k}))$  for  $1 < p < \infty$ ;
- $C_0(\Xi_k) = C(V_{d-k}(\mathbb{R}^d) \otimes_{\alpha_\infty} C_0(\mathbb{R}^{d-k}))$ ,

then  $\mathcal{X} = \mathcal{X}_k \oplus \mathcal{X}_k^c$ , where  $\mathcal{X}_k = \mathcal{X}_{\text{iso}}$ , where

$$(8.9) \quad \mathcal{X}_{\text{iso}} := \{g \in \mathcal{X} : g(\mathbf{A}, \mathbf{t}) = g(\mathbf{U}\mathbf{A}, \mathbf{U}\mathbf{t}) \text{ for any } (\mathbf{A}, \mathbf{t}) \in \Xi_k \text{ and } \mathbf{U} \in O_{d-k}(\mathbb{R})\}$$

is the subspace of isotropic functions in  $\mathcal{X}$ .

*Proof.* The space  $\mathcal{X}_{\text{iso}}$  is a closed subspace of  $\mathcal{X}$  by construction. To show that it is complemented in  $\mathcal{X}$ , we consider the self-adjoint operator  $P_{\text{iso}} : \mathcal{X} \rightarrow \mathcal{X}$  (see Appendix A) which extracts the isotropic part of a function as

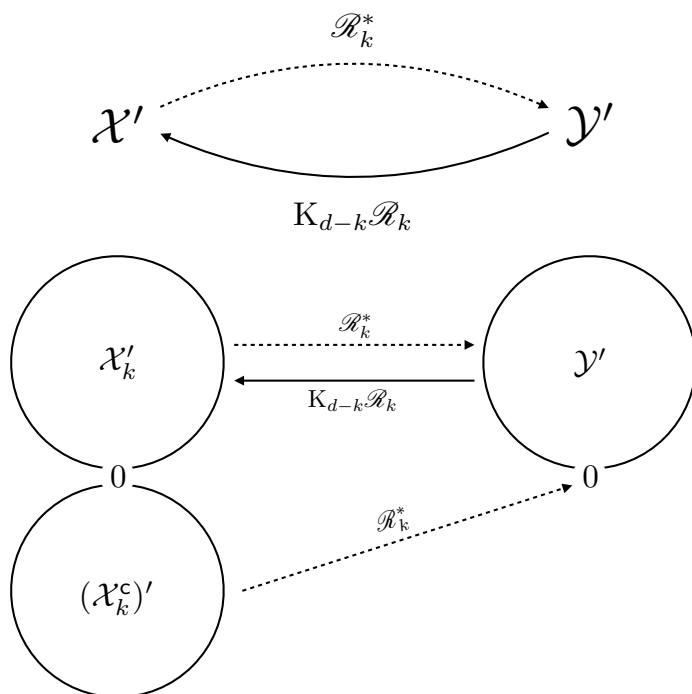


FIG. 1. Relevant Banach spaces when  $\mathcal{X} = \mathcal{X}_k \oplus \mathcal{X}_k^c$ .

$$(8.10) \quad P_{\text{iso}}\{g\}(\mathbf{A}, \mathbf{t}) = \int_{O_{d-k}(\mathbb{R})} g(\mathbf{U}\mathbf{A}, \mathbf{U}\mathbf{t}) \, d\sigma(\mathbf{U}),$$

where  $\int$  denotes the average integral and  $\sigma$  is the Haar measure on  $O_{d-k}(\mathbb{R})$ . Clearly,  $P_{\text{iso}}(\mathcal{X}) \subset \mathcal{X}_{\text{iso}}$  and, given  $g \in \mathcal{X}_{\text{iso}}$ ,  $P_{\text{iso}}P_{\text{iso}}g = g$ . This establishes that  $P_{\text{iso}}$  is a projector and that  $P_{\text{iso}}(\mathcal{X}) = \mathcal{X}_{\text{iso}}$ . This projector is bounded since

$$(8.11) \quad \begin{aligned} \|P_{\text{iso}}g\|_{\mathcal{X}} &= \left\| (\mathbf{A}, \mathbf{t}) \mapsto \int_{O_{d-k}(\mathbb{R})} g(\mathbf{U}\mathbf{A}, \mathbf{U}\mathbf{t}) \, d\sigma(\mathbf{U}) \right\|_{\mathcal{X}} \\ &\leq \int_{O_{d-k}(\mathbb{R})} \|(\mathbf{A}, \mathbf{t}) \mapsto g(\mathbf{U}\mathbf{A}, \mathbf{U}\mathbf{t})\|_{\mathcal{X}} \, d\sigma(\mathbf{U}) \\ &= \int_{O_{d-k}(\mathbb{R})} \|g\|_{\mathcal{X}} \, d\sigma(\mathbf{U}) \\ &= \|g\|_{\mathcal{X}}, \end{aligned}$$

where the second line follows from Jensen's inequality and the third line follows from the rotational invariance of the  $L^p$ -norms. The boundedness of the projector  $P_{\text{iso}}$  guarantees the existence of a complementary projector  $(\text{Id} - P_{\text{iso}}) : \mathcal{X} \rightarrow \mathcal{X}$  whose (closed) range  $(\text{Id} - P_{\text{iso}})(\mathcal{X}) =: \mathcal{X}_{\text{iso}}^c$  is the complementary subspace of  $\mathcal{X}_{\text{iso}}$  with the properties that  $\mathcal{X} = \mathcal{X}_{\text{iso}} \oplus \mathcal{X}_{\text{iso}}^c$  and so  $\mathcal{X}_{\text{iso}} \cap \mathcal{X}_{\text{iso}}^c = \{0\}$ .

The final part of the proof is to show that

$$(8.12) \quad \mathcal{X}_k = \overline{(\mathcal{S}_k, \|\cdot\|_{\mathcal{X}})} = \mathcal{X}_{\text{iso}}.$$

To that end, we consider the Lizorkin space

$$(8.13) \quad \mathcal{S}_{\text{Liz}}(\mathbb{R}^n) = \left\{ \varphi \in \mathcal{S}(\mathbb{R}^n) : \int_{\mathbb{R}^n} \mathbf{x}^{\mathbf{m}} \varphi(\mathbf{x}) \, d\mathbf{x} = 0 \text{ for every multi-index } \mathbf{m} \right\},$$

which is the closed subspace of Schwartz functions with all moments vanishing. We can define the Lizorkin space on  $\Xi_k$  as  $\mathcal{S}_{\text{Liz}}(\Xi_k) := C^\infty(V_{d-k}(\mathbb{R}^d)) \widehat{\otimes} \mathcal{S}_{\text{Liz}}(\mathbb{R}^{d-k})$ . Then, let

$$(8.14) \quad \mathcal{S}_{\text{Liz,iso}}(\Xi_k) = \left\{ \psi \in \mathcal{S}_{\text{Liz}}(\Xi_k) : \begin{array}{l} \psi(\mathbf{A}, \mathbf{t}) = \psi(\mathbf{U}\mathbf{A}, \mathbf{U}\mathbf{t}) \\ \text{for any } (\mathbf{A}, \mathbf{t}) \in \Xi_k \text{ and } \mathbf{U} \in O_{d-k}(\mathbb{R}) \end{array} \right\}$$

denote its isotropic subspace. From Theorem 5.4, we see that  $\mathcal{S}_{\text{Liz,iso}}(\Xi_k) \subset \mathcal{S}_k$ .

Since

$$(8.15) \quad \overline{(C^\infty(V_{d-k}(\mathbb{R}^d)), \|\cdot\|_{L^p})} = \begin{cases} L^p(V_{d-k}(\mathbb{R}^d)), & 1 < p < \infty, \\ C(V_{d-k}(\mathbb{R}^d)), & p = \infty, \end{cases}$$

and

$$(8.16) \quad \overline{(\mathcal{S}_{\text{Liz}}(\mathbb{R}^{d-k}), \|\cdot\|_{L^p})} = \begin{cases} L^p(\mathbb{R}^{d-k}), & 1 < p < \infty \text{ (from [58])}, \\ C_0(\mathbb{R}^{d-k}), & p = \infty \text{ (from [40])}, \end{cases}$$

and since the isotropic subspace is complete, then (8.15) and (8.16) imply that  $\mathcal{X}_{\text{iso}} = \overline{(\mathcal{S}_{\text{Liz,iso}}(\Xi_k), \|\cdot\|_{\mathcal{X}})}$ . Since  $\mathcal{S}_k \subset \mathcal{X}_{\text{iso}}$ , this proves (8.12), since the completion is unique.  $\square$

For these  $k$ -plane-compatible Banach spaces, the proof of Theorem 8.2 actually implies that  $P_k = P_k^* = P_{\text{iso}}$ . In other words, it is the projection operator that extracts the isotropic part of a function (or distribution). From Theorem 8.1 we also know that the backprojection operator is invertible on  $\mathcal{X}'_{\text{iso}}$ , where  $\mathcal{X}$  is any of the Banach spaces in Theorem 8.2. This is summarized in Corollary 8.3.

**COROLLARY 8.3.** *The backprojection operator  $\mathcal{R}_k^*$  is invertible on*

1.  $\mathcal{M}_{\text{iso}}(\Xi_k)$ ;
2.  $L^p_{\text{iso}}(\Xi_k)$  for  $1 < p < \infty$ .

*Proof.* This follows by duality directly from Theorem 8.2.  $\square$

**Remark 8.4.** Theorem 8.1 recovers the result of [69, Theorem 8] in the special case of  $k = (d - 1)$ . To the best of our knowledge, not much was previously known about the invertibility of  $\mathcal{R}_k^*$  beyond the fact that it is invertible on  $\mathcal{S}'_k$ . However, in the special case  $k = (d - 1)$ , a multitude of classical works have addressed the invertibility of the dual Radon transform  $\mathcal{R}_{d-1}^*$ . In particular, it is known that it is invertible on

1. the subspace  $\mathcal{S}_{\text{even}}(\mathbb{S}^{d-1} \times \mathbb{R})$  of even functions of the Schwartz space of smooth and rapidly decreasing functions on  $\mathbb{S}^{d-1} \times \mathbb{R}$  [65];
2. the space of essentially bounded, even, and compactly supported functions on  $\mathbb{S}^{d-1} \times \mathbb{R}$ , denoted as  $L^\infty_{\text{even,cpct}}(\mathbb{S}^{d-1} \times \mathbb{R})$  [49];
3. the subspace  $\mathcal{S}'_{\text{Liz,even}}(\mathbb{S}^{d-1} \times \mathbb{R})$  of even Lizorkin distributions on  $\mathbb{S}^{d-1} \times \mathbb{R}$  [33].

Therefore, Corollary 8.3 is among the first to provide results on the invertibility of  $\mathcal{R}_k^*$ , some of which are even new for the dual Radon transform, notably, the invertibility of  $\mathcal{R}_{d-1}^*$  on  $L^p_{\text{even}}(\mathbb{S}^{d-1} \times \mathbb{R})$  for  $1 < p < \infty$ . In particular, since both  $\mathcal{S}_{\text{iso}}(\Xi_k)$  and  $L^\infty_{\text{iso,cpct}}(\Xi_k)$  continuously embed into  $L^p_{\text{iso}}(\Xi_k)$  for any  $1 < p < \infty$ , we have that Corollary 8.3 provides a direct generalization of the results of [49, 65] for the  $k$ -plane transform.

**9. Specific  $k$ -plane transforms.** In this section, we provide explicit formulas for the  $k$ -plane transform of isotropic functions and  $k$ -plane ridge distributions.

**9.1. Isotropic functions.** An isotropic function  $\rho_{\text{iso}} : \mathbb{R}^d \rightarrow \mathbb{R}$  is characterized by its radial profile  $\rho : \mathbb{R} \rightarrow \mathbb{R}$  via the equality  $\rho_{\text{iso}}(\mathbf{x}) = \rho(t)$ , where  $t = \|\mathbf{x}\|_2$ . The frequency-domain counterpart of this characterization is  $\widehat{\rho}_{\text{iso}}(\boldsymbol{\xi}) = \widehat{\rho}_{\text{rad}}(\omega)$ , where  $\omega = \|\boldsymbol{\xi}\|_2$ . The radial frequency profile  $\widehat{\rho}_{\text{rad}}$  can be computed via the Hankel transform [70, Theorem 5.26]

$$(9.1) \quad \widehat{\rho}_{\text{rad}}(\omega) = \frac{(2\pi)^{d/2}}{|\omega|^{d/2-1}} \int_0^\infty J_{d/2-1}(\omega t) t^{d/2-1} \rho(t) t dt,$$

where  $J_\nu$  is the Bessel function of the first kind of order  $\nu$ .

**THEOREM 9.1.** *Let  $\rho_{\text{iso}} \in \mathcal{S}(\mathbb{R}^d)$  be an isotropic function with radial frequency profile  $\widehat{\rho}_{\text{rad}} \in \mathcal{S}(\mathbb{R})$ . Then, for any  $\mathbf{x}_0 \in \mathbb{R}^d$ ,*

$$(9.2) \quad \mathcal{R}_k\{\rho_{\text{iso}}(\cdot - \mathbf{x}_0)\}(\mathbf{A}, \mathbf{t}) = \widetilde{\rho}_{\text{rad}}(\mathbf{t} - \mathbf{A}\mathbf{x}_0),$$

where  $\widetilde{\rho}_{\text{rad}} = \mathcal{F}_{d-k}^{-1}\{\widehat{\rho}_{\text{rad}}(\|\cdot\|_2)\}$  and  $\mathcal{F}_{d-k}^{-1}$  denotes the  $(d-k)$ -variate inverse Fourier transform.

*Proof.* By the Fourier slice theorem,

$$(9.3) \quad \mathcal{F}_{d-k}\{\mathcal{R}_k\{\rho_{\text{iso}}\}(\mathbf{A}, \cdot)\}(\boldsymbol{\omega}) = \widehat{\rho}_{\text{iso}}(\mathbf{A}^\top \boldsymbol{\omega}) = \widehat{\rho}_{\text{rad}}(\|\mathbf{A}^\top \boldsymbol{\omega}\|_2) = \widehat{\rho}_{\text{rad}}(\|\boldsymbol{\omega}\|_2).$$

The  $(d-k)$ -variate inverse Fourier transform of (9.3) yields that

$$(9.4) \quad \mathcal{R}_k\{\rho_{\text{iso}}\}(\mathbf{A}, \mathbf{t}) = \widetilde{\rho}_{\text{rad}}(\mathbf{t}).$$

The result then follows by the shifting property of the  $k$ -plane transform. Indeed, for any  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  we have that

$$(9.5) \quad \begin{aligned} \mathcal{R}_k\{\varphi(\cdot - \mathbf{x}_0)\}(\mathbf{A}, \mathbf{t}) &= \int_{\mathbb{R}^d} \varphi(\mathbf{x} - \mathbf{x}_0) \delta(\mathbf{A}\mathbf{x} - \mathbf{t}) d\mathbf{x} \\ &= \int_{\mathbb{R}^d} \varphi(\mathbf{y}) \delta(\mathbf{A}\mathbf{y} - (\mathbf{t} - \mathbf{A}\mathbf{x}_0)) d\mathbf{y} \\ &= \mathcal{R}_k\{\varphi\}(\mathbf{A}, \mathbf{t} - \mathbf{A}\mathbf{x}_0). \end{aligned} \quad \square$$

**Remark 9.2.** Theorem 9.1 can be extended to general isotropic tempered distributions so long as the equality is understood in the distribution space  $(\mathbb{K}_{d-k} \mathcal{R}_k(\mathcal{S}(\mathbb{R}^d)))'$ .

**Example 9.3.** Consider the isotropic Gaussian density with zero mean and unit variance

$$(9.6) \quad g_{\text{iso}}(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}} \exp\left(-\frac{\|\mathbf{x}\|_2^2}{2}\right).$$

Its Fourier transform is given by  $\widehat{g}_{\text{iso}}(\boldsymbol{\xi}) = \exp(-\|\boldsymbol{\xi}\|_2^2/2)$ , which corresponds to the radial profile  $\widehat{g}_{\text{rad}}(\omega) = \exp(-|\omega|^2/2)$ . The determination of the  $(d-k)$ -variate inverse Fourier transform then gives

$$(9.7) \quad \mathcal{F}_{d-k}^{-1}\{\widehat{g}_{\text{rad}}(\|\cdot\|_2)\}(\mathbf{t}) = \frac{1}{(2\pi)^{\frac{d-k}{2}}} \exp\left(-\frac{\|\mathbf{t}\|_2^2}{2}\right).$$

Finally, by applying Theorem 9.1, we find for any  $\mathbf{x}_0 \in \mathbb{R}^d$  that

$$(9.8) \quad \mathcal{R}_k\{g_{\text{iso}}(\cdot - \mathbf{x}_0)\}(\mathbf{A}, \mathbf{t}) = \frac{1}{(2\pi)^{\frac{d-k}{2}}} \exp\left(-\frac{\|\mathbf{t} - \mathbf{A}\mathbf{x}_0\|_2^2}{2}\right).$$



**9.2.  $k$ -plane ridges.** From Definition 7.3, we have the definition of a  $k$ -plane ridge with any profile  $r \in \mathcal{S}'(\mathbb{R}^{d-k})$ . In this section, we write  $r(\mathbf{A}_0(\cdot) - \mathbf{t}_0) \in \mathcal{S}'(\mathbb{R}^d)$  to denote the  $k$ -plane ridge with profile  $r(\cdot - \mathbf{t}_0) \in \mathcal{S}'(\mathbb{R}^{d-k})$ , where  $(\mathbf{A}_0, \mathbf{t}_0) \in \Xi_k$ .

**THEOREM 9.4.** *Let  $(\mathbf{A}_0, \mathbf{t}_0) \in \Xi_k$  and  $r \in \mathcal{S}'(\mathbb{R}^{d-k})$ . Then,*

$$(9.9) \quad \mathbf{K}_{d-k} \mathcal{R}_k \{r(\mathbf{A}_0(\cdot) - \mathbf{t}_0)\} = [\delta(\cdot - \mathbf{A}_0) r(\cdot - \mathbf{t}_0)] \in \mathcal{S}'_k,$$

$$(9.10) \quad \mathcal{R}_k \{r(\mathbf{A}_0(\cdot) - \mathbf{t}_0)\} = [\delta(\cdot - \mathbf{A}_0) (q * r)(\cdot - \mathbf{t}_0)] \in (\mathbf{K}_{d-k} \mathcal{R}_k (\mathcal{S}(\mathbb{R}^d)))',$$

where  $q(\mathbf{t}) = \mathcal{F}_{d-k}^{-1} \{ \widehat{K}_k^{-1} \}(\mathbf{t})$  is the impulse response of the inverse filtering operator  $\mathbf{K}_{d-k}^{-1}$  and where the distributions on the right-hand sides must be understood as equivalence classes in their respective distribution spaces.

Furthermore, if

$$(9.11) \quad \delta(\cdot - \mathbf{A}_0) r(\cdot - \mathbf{t}_0) \in \mathcal{X}' = (\mathcal{X}_k \oplus \mathcal{X}_k^c)' \subset \mathcal{S}'(\Xi_k),$$

where  $\mathcal{X}_k$  and  $\mathcal{X}_k^c$  are complementary Banach spaces and  $(\mathcal{X}_k, \mathcal{X}_k')$  form a dual pair of  $k$ -plane-compatible Banach spaces as in section 8, then the filtered  $k$ -plane transform has the concrete interpretation

$$(9.12) \quad \mathbf{K}_{d-k} \mathcal{R}_k \{r(\mathbf{A}_0(\cdot) - \mathbf{t}_0)\} = \mathbf{P}_k^* \{ \delta(\cdot - \mathbf{A}_0) r(\cdot - \mathbf{t}_0) \} \in \mathcal{X}'_k.$$

*Proof.* It suffices to prove the results when  $\mathbf{t}_0 = \mathbf{0}$  since the general result follows by replacing  $r$  with  $r(\cdot - \mathbf{t}_0)$ . For all  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  we have that

$$(9.13) \quad \begin{aligned} \langle \delta(\cdot - \mathbf{A}_0) r(\cdot), \mathcal{R}_k \{ \varphi \} \rangle_k &= \langle r, \mathcal{R}_k \{ \varphi \}(\mathbf{A}_0, \cdot) \rangle \\ &= \langle r(\mathbf{A}_0(\cdot)), \varphi \rangle \\ &= \langle r(\mathbf{A}_0(\cdot)), \mathcal{R}_k^* \mathbf{K}_{d-k} \mathcal{R}_k \{ \varphi \} \rangle \\ &= \langle \mathbf{K}_{d-k} \mathcal{R}_k r(\mathbf{A}_0(\cdot)), \mathcal{R}_k \{ \varphi \} \rangle, \end{aligned}$$

where, in the second line, we used Definition 7.3 and, in the third line, we used Theorem 5.1. Since  $\mathcal{R}_k \{ \varphi \} \in \mathcal{S}_k$ , we see from item 2 in Definition 6.1 that  $\delta(\cdot - \mathbf{A}_0) r(\cdot) = \mathbf{K}_{d-k} \mathcal{R}_k r(\mathbf{A}_0(\cdot))$ , where the equality is understood in  $\mathcal{S}'_k$ , proving (9.9). This then yields (9.10) by substituting  $(q * r)$  for  $r$ . The identity in (9.11) follows from the application of Theorem 8.1.  $\square$

*Example 9.5.* If  $r \in \mathcal{M}(\mathbb{R}^{d-k})$ , then  $\delta(\cdot - \mathbf{A}_0) r(\cdot - \mathbf{t}_0) \in \mathcal{M}(\Xi_k)$ . By Theorem 8.2, we know that  $\mathcal{M}(\Xi_k) = \mathcal{M}_{\text{iso}}(\Xi_k) \oplus (\mathcal{M}_{\text{iso}}(\Xi_k))^c$ , where  $(C_{0,\text{iso}}(\Xi_k), \mathcal{M}_{\text{iso}}(\Xi_k))$  forms a dual pair of  $k$ -plane-compatible Banach spaces. Therefore, by Theorem 9.4,

$$(9.14) \quad \mathbf{K}_{d-k} \mathcal{R}_k \{r(\mathbf{A}_0(\cdot) - \mathbf{t}_0)\} = \mathbf{P}_{\text{iso}} \{ \delta(\cdot - \mathbf{A}_0) r(\cdot - \mathbf{t}_0) \} \in \mathcal{M}_{\text{iso}}(\Xi_k).$$

*Remark 9.6.* Let

$$(9.15) \quad \begin{aligned} \delta_{\text{iso}}((\mathbf{A}, \mathbf{t}) - (\mathbf{A}_0, \mathbf{t}_0)) &:= \mathbf{P}_{\text{iso}} \{ \delta(\cdot - (\mathbf{A}_0, \mathbf{t}_0)) \}(\mathbf{A}, \mathbf{t}) \\ &= \int_{\mathbf{O}_{d-k}(\mathbb{R})} \delta((\mathbf{U}\mathbf{A}, \mathbf{U}\mathbf{t}) - (\mathbf{A}_0, \mathbf{t}_0)) d\sigma(\mathbf{U}) \end{aligned}$$

denote the isotropic projection of the Dirac impulse on  $\Xi_k$ . Example 9.5 establishes the identity

$$(9.16) \quad \mathbf{K}_{d-k} \mathcal{R}_k \{ \delta(\mathbf{A}_0(\cdot) - \mathbf{t}_0) \} = \delta_{\text{iso}}(\cdot - (\mathbf{A}_0, \mathbf{t}_0)),$$

which is equivalent to

$$(9.17) \quad \mathcal{R}_k^* \{ \delta_{\text{iso}}(\cdot - (\mathbf{A}_0, \mathbf{t}_0)) \} = \delta(\mathbf{A}_0(\cdot) - \mathbf{t}_0),$$

where we use the consequence from Theorem 6.2 that  $\mathcal{R}_k^* \mathbf{K}_{d-k} \mathcal{R}_k = \text{Id}$  on  $\mathcal{S}'(\mathbb{R}^d)$ . Furthermore, if  $L_d$  is an isotropic LSI operator, then, by the intertwining properties (Corollary 7.8), we have that

$$(9.18) \quad L_d \mathcal{R}_k^* \{ \delta_{\text{iso}}(\cdot - (\mathbf{A}_0, \mathbf{t}_0)) \} = \mathcal{R}_k^* \{ L_{d-k} \delta_{\text{iso}}(\cdot - (\mathbf{A}_0, \mathbf{t}_0)) \} = r(\mathbf{A}_0(\cdot) - \mathbf{t}_0),$$

where  $r = \mathcal{F}_{d-k}^{-1} \{ \widehat{L}_{d-k} \}$  is the impulse response of  $L_{d-k}$ .

**10. Application to the regularization of inverse problems.** In this section, we demonstrate the use of our formalism to specify a new regularization for linear inverse problems and investigate its properties. The resolution of a linear inverse problem involves the estimation of an unknown object  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  based on a finite number of measurements of the form

$$(10.1) \quad \mathbf{y} = \mathbf{H}\{f\} + \boldsymbol{\varepsilon} \in \mathbb{R}^M.$$

Here,  $\mathbf{H}$  is a *linear* operator that models the response of the acquisition device (e.g., samples of the Fourier transform in the case of magnetic resonance imaging) and  $\boldsymbol{\varepsilon} \in \mathbb{R}^M$  is an unknown perturbation or noise term, often assumed to be a vector of independent and identically distributed Gaussian random variables. One typically solves the inverse problem by minimizing a functional of the form

$$(10.2) \quad \min_{f \in \mathcal{X}} \|\mathbf{y} - \mathbf{H}\{f\}\|_2^2 + \lambda \Phi(f),$$

where the first term is the *data-fidelity* term and the second term is the *regularization* term that injects prior knowledge on the solution [16, 5]. The parameter  $\lambda > 0$  controls the trade-off between data fidelity and regularization. The *native space*  $\mathcal{X}$  can be thought of as the largest space of functions mapping  $\mathbb{R}^d \rightarrow \mathbb{R}$  such that the regularizer  $\Phi(\cdot)$  is well-defined.

Here, we shall introduce a new sparsity-promoting regularizer in the  $k$ -plane domain and show that the solutions favor sparse superpositions of  $k$ -plane ridges. This is a generalization of recently studied regularizers in the Radon domain ( $k = (d - 1)$ ) which have been shown to favor neural-network-like solutions [42]. The sparse superpositions of  $k$ -plane ridges are, in essence, a hybridization of radial basis functions (RBFs) and neural networks.

Given  $s \geq 0$ , let

$$(10.3) \quad \mathbf{B}_d^s := (\text{Id} - \Delta)^{s/2}$$

denote the Bessel potential of order  $(-s)$  on  $\mathbb{R}^d$ , where  $\text{Id}$  denotes the identity operator and  $\Delta$  denotes the Laplacian. The operator  $\mathbf{B}_d^s$  can be understood in the Fourier domain via the frequency response

$$(10.4) \quad \widehat{B}_d^s(\boldsymbol{\xi}) = (1 + \|\boldsymbol{\xi}\|_2^2)^{s/2}, \quad \boldsymbol{\xi} \in \mathbb{R}^d.$$

These operators play an important role in the theory of Sobolev spaces. Indeed, when  $s$  is an integer and  $1 < p < \infty$ ,  $f \mapsto \|\mathbf{B}_d^s f\|_{L^p}$  is a norm for the Sobolev space  $W^{s,p}(\mathbb{R}^d)$ . Consider the space

$$(10.5) \quad \mathcal{M}_{\mathcal{R}_k}^s(\mathbb{R}^d) = \{f \in \mathcal{S}'(\mathbb{R}^d) : \|\mathbf{K}_{d-k} \mathcal{R}_k \mathbf{B}_d^s f\|_{\mathcal{M}(\Xi_k)} < \infty\}.$$

When equipped with the norm  $\|f\|_{\mathcal{M}_{\mathcal{R}_k}^s} := \|\mathbf{K}_{d-k} \mathcal{R}_k \mathbf{B}_d^s f\|_{\mathcal{M}(\Xi_k)}$ , this is a Banach space. Indeed, the operator

$$(10.6) \quad \mathbf{K}_{d-k} \mathcal{R}_k \mathbf{B}_d^s : \mathcal{M}_{\mathcal{R}_k}^s(\mathbb{R}^d) \rightarrow \mathcal{M}_{\text{iso}}(\Xi_k)$$

is an isometric isomorphism with inverse given by

$$(10.7) \quad \mathbf{B}_d^{-s} \mathcal{R}_k^* : \mathcal{M}_{\text{iso}}(\Xi_k) \rightarrow \mathcal{M}_{\mathcal{R}_k}^s(\mathbb{R}^d).$$

Since  $\mathcal{M}_{\text{iso}}(\Xi_k)$  is a Banach space,  $\mathcal{M}_{\mathcal{R}_k}^s(\mathbb{R}^d)$  is a Banach space too, when equipped with the previously specified norm. Moreover, it has a “canonical” predual which is the space

$$(10.8) \quad C_{0, \mathcal{R}_k}^s(\mathbb{R}^d) = (\mathbf{K}_{d-k} \mathcal{R}_k \mathbf{B}_d^s)^* (C_{0, \text{iso}}(\Xi_k)) = \mathbf{B}_d^s \mathcal{R}_k^* \mathbf{K}_{d-k} (C_{0, \text{iso}}(\Xi_k)),$$

i.e.,  $(C_{0, \mathcal{R}_k}^s(\mathbb{R}^d))' = \mathcal{M}_{\mathcal{R}_k}^s(\mathbb{R}^d)$ . This allows us to equip  $\mathcal{M}_{\mathcal{R}_k}^s(\mathbb{R}^d)$  with a weak\*-topology. With this formalism in hand, we have the following *representer theorem* for  $k$ -plane-domain regularization.

**THEOREM 10.1.** *Let  $\mathbf{H}\{f\} = (\langle h_1, f \rangle, \dots, \langle h_M, f \rangle) \in \mathbb{R}^M$  denote the linear measurement functionals where  $h_m \in C_{0, \mathcal{R}_k}^s(\mathbb{R}^d)$ ,  $m = 1, \dots, M$ , so that the measurement operator is weak\*-continuous on  $\mathcal{M}_{\mathcal{R}_k}^s(\mathbb{R}^d)$ . Then, for any fixed  $\mathbf{y} \in \mathbb{R}^M$ , the solution set to the variational problem*

$$(10.9) \quad \mathcal{V} := \arg \min_{f \in \mathcal{M}_{\mathcal{R}_k}^s(\mathbb{R}^d)} \|\mathbf{y} - \mathbf{H}\{f\}\|_2^2 + \lambda \|\mathbf{K}_{d-k} \mathcal{R}_k \mathbf{B}_d^s f\|_{\mathcal{M}(\Xi_k)}, \quad \lambda > 0,$$

is nonempty, convex, and weak\*-compact. The full solution set  $\mathcal{V}$  is the weak\* closure of the convex hull of its extreme points, which can all be expressed as

$$(10.10) \quad f_{\text{extreme}}(\mathbf{x}) = \sum_{n=1}^N a_n \rho_s(\mathbf{A}_n \mathbf{x} - \mathbf{t}_n),$$

with  $N$  parameters (number of atoms with  $N \leq M$ ),  $\{a_n\}_{n=1}^N \subset \mathbb{R} \setminus \{0\}$ , and  $\{(\mathbf{A}_n, \mathbf{t}_n)\}_{n=1}^N \subset \Xi_k$  which are data-dependent and not known a priori. The profile function  $\rho_s : \mathbb{R}^{d-k} \rightarrow \mathbb{R}$  is an RBF that corresponds to the Green's function of  $\mathbf{B}_{d-k}^s$ . Finally, the regularization cost, which is common to all solutions, is  $\|\mathbf{K}_{d-k} \mathcal{R}_k \mathbf{B}_d^s f_{\text{extreme}}\|_{\mathcal{M}} = \sum_{n=1}^N |a_n| = \|\mathbf{a}\|_1$ .

The key takeaway from Theorem 10.1 is that the solution set to the variational problem is completely characterized by sparse ( $N \leq M$ ) superpositions of *data-adaptive*  $k$ -plane ridges where the profiles are RBFs with exponential decay (see [23, p. 7] for a formula and decay properties for  $\rho_s$ ). By invoking the intertwining properties (Corollary 7.8), we may rewrite the regularizer as

$$(10.11) \quad \|\mathbf{B}_{d-k}^s \mathbf{K}_{d-k} \mathcal{R}_k f\|_{\mathcal{M}(\Xi_k)},$$

where we recall that  $\mathbf{B}_{d-k}^s$  is the  $(d-k)$ -dimensional Bessel potential of order  $(-s)$ .

We prove Theorem 10.1 using recent results characterizing the solution sets to optimization problems over Banach spaces with sparsity-promoting norms [6, 7, 68]. These results hinge on the characterization of the extreme points of the unit ball of the regularizer. Since  $\mathcal{M}_{\mathcal{R}_k}^s(\mathbb{R}^d)$  is isometrically isomorphic to  $\mathcal{M}_{\text{iso}}(\Xi_k)$ , one can find the extreme points of the unit ball in  $\mathcal{M}_{\mathcal{R}_k}^s(\mathbb{R}^d)$  by applying (10.7) to the extreme points of the unit ball in  $\mathcal{M}_{\text{iso}}(\Xi_k)$ . We characterize these extreme points in subsection 10.1

**10.1. Extreme points of the unit ball in  $\mathcal{M}_{\text{iso}}(\Xi_k)$ .** The canonical evaluation functionals that sample a continuous function defined on  $\Xi_k$  at  $(\mathbf{A}_0, \mathbf{t}_0)$  are the shifted Dirac impulses  $\delta(\cdot - (\mathbf{A}_0, \mathbf{t}_0)) = \delta(\cdot - \mathbf{A}_0)\delta(\cdot - \mathbf{t}_0) \in \mathcal{M}(\Xi_k)$ . It is well-known that these correspond to the extreme points of the unit ball in  $\mathcal{M}(\Xi_k)$ .

We now show that the  $k$ -plane-compatible sampling functionals are precisely the isotropic projections of the shifted Dirac impulses specified in (9.15) and that these are the extreme points of the unit ball in  $\mathcal{M}_{\text{iso}}(\Xi_k)$ . We summarize their properties in Lemma 10.2.

LEMMA 10.2. *The isotropic shifted Dirac impulse  $\delta_{\text{iso}}(\cdot - (\mathbf{A}_0, \mathbf{t}_0)) \in \mathcal{M}_{\text{iso}}(\Xi_k)$  specified by (9.15) satisfies the following properties.*

1. *Sampling: For any  $\psi \in C_{0,\text{iso}}(\Xi_k)$ ,*

$$(10.12) \quad \langle \delta_{\text{iso}}(\cdot - (\mathbf{A}_0, \mathbf{t}_0)), \psi \rangle_k = \psi(\mathbf{A}_0, \mathbf{t}_0).$$

2. *Rotation invariance: For any  $\mathbf{U} \in \text{O}_{d-k}(\mathbb{R})$ ,*

$$(10.13) \quad \delta_{\text{iso}}(\cdot - (\mathbf{A}_0, \mathbf{t}_0)) = \delta_{\text{iso}}(\cdot - (\mathbf{U}\mathbf{A}_0, \mathbf{U}\mathbf{t}_0)).$$

3. *Unit norm:  $\|\delta_{\text{iso}}(\cdot - (\mathbf{A}_0, \mathbf{t}_0))\|_{\mathcal{M}} = 1$ .*

4. *Linear combination: For any set  $\{(\mathbf{A}_n, \mathbf{t}_n)\}_{n=1}^N \subset \Xi_k$  of distinct points,*

$$(10.14) \quad \left\| \sum_{n=1}^N a_n \delta_{\text{iso}}(\cdot - (\mathbf{A}_n, \mathbf{t}_n)) \right\|_{\mathcal{M}} = \sum_{n=1}^N |a_n| = \|\mathbf{a}\|_1.$$

5. *Extreme points of  $B_{\mathcal{M}_{\text{iso}}} = \{e \in \mathcal{M}_{\text{iso}}(\Xi_k) : \|e\|_{\mathcal{M}} \leq 1\}$ : If  $e \in \text{Ext}B_{\mathcal{M}_{\text{iso}}}$ , then  $e = \pm \delta_{\text{iso}}(\cdot - (\mathbf{A}_0, \mathbf{t}_0))$  for some  $(\mathbf{A}_0, \mathbf{t}_0) \in \Xi_k$ .*

*Proof.* Items 1 to 3 are direct consequences of the definition of the isotropic Dirac impulse. Item 4 follows from the observation that the distributions  $\{\delta_{\text{iso}}(\cdot - (\mathbf{A}_n, \mathbf{t}_n))\}_{n=1}^N$  have disjoint support whenever  $\{(\mathbf{A}_n, \mathbf{t}_n)\}_{n=1}^N \subset \Xi_k$  are distinct since this implies equality in the triangle inequality, which proves (10.14).

As for item 5, the result follows from a classical result [14, p. 441] which characterizes the extreme points of the dual of any closed subspace of  $\mathcal{X} \subset C_0(Q)$  as the set of all evaluation functionals in  $\mathcal{X}'$  multiplied by  $\pm 1$ , where  $Q$  is any (locally compact) Hausdorff space. In the present scenario, we first note that  $C_{\text{iso},0}(\Xi_k) \subset C_0(\Xi_k)$  is a closed subspace with the domain  $\Xi_k$  being a locally compact Hausdorff space. From item 1, we know that the evaluation functionals in  $\mathcal{M}_{\text{iso}}(\Xi_k) = (C_{\text{iso},0}(\Xi_k))'$  are exactly the shifted isotropic Dirac impulses. We refer the reader to [14, p. 441] for a proof of this classical result when the domain is a Hausdorff space and to [7, Proposition 4.1] for a proof that applies to any locally compact Hausdorff space.  $\square$

**10.2. Proof of Theorem 10.1.**

*Proof.* From item 5 in Lemma 10.2, we see that the extreme points of the unit ball

$$(10.15) \quad B_{\mathcal{M}_{\mathcal{R}_k}^s} := \{f \in \mathcal{M}_{\mathcal{R}_k}^s(\mathbb{R}^d) : \|\mathbf{K}_{d-k} \mathcal{R}_k B_d^s f\|_{\mathcal{M}} \leq 1\}$$

take the form  $\pm B_d^{-s} \mathcal{R}_k^* \{\delta_{\text{iso}}(\cdot - (\mathbf{A}_0, \mathbf{t}_0))\} = \pm B_k^* B_{d-k}^{-s} \{\delta_{\text{iso}}(\cdot - (\mathbf{A}_0, \mathbf{t}_0))\}$ , where  $(\mathbf{A}_0, \mathbf{t}_0) \in \Xi_k$ . Applying (9.18) we find that  $e \in \text{Ext}B_{\mathcal{M}_{\mathcal{R}_k}^s}$  if and only if

$$(10.16) \quad e(\mathbf{x}) = \rho_s(\mathbf{A}_0 \mathbf{x} - \mathbf{t}_0), \quad (\mathbf{A}_0, \mathbf{t}_0) \in \Xi_k,$$

Downloaded 07/02/24 to 128.178.48.127 . Redistribution subject to SIAM license or copyright; see https://pubs.siam.org/terms-privacy

where  $\rho_s : \mathbb{R}^{d-k} \rightarrow \mathbb{R}$  is the impulse response of  $B_{d-k}^{-s}$  which is also the Green's function of  $B_{d-k}^s$ . The abstract representer theorem [68, Theorems 2 and 3] then tells us that the solution set  $\mathcal{V}$  is nonempty, convex, and weak\*-compact and that the full solution set is the weak\* closure of the convex hull of its extreme points. In general, the latter are linear combinations of  $N \leq M$  extreme points of the unit regularization ball. In the present scenario, this translates into

$$(10.17) \quad f_{\text{extreme}}(\mathbf{x}) = \sum_{n=1}^N a_n \rho_s(\mathbf{A}_n \mathbf{x} - \mathbf{t}_n),$$

where  $\{a_n\}_{n=1}^N \subset \mathbb{R} \setminus \{0\}$ ,  $\{\mathbf{A}_n\}_{n=1}^N \subset V_{d-k}(\mathbb{R}^d)$ ,  $\{\mathbf{t}_n\}_{n=1}^N \subset \mathbb{R}^{d-k}$ , and  $N \leq M$ . Moreover, every solution in  $\mathcal{V}$  shares the same regularization cost. A calculation in the Fourier domain shows that  $B_d^s \rho_s(\mathbf{A}(\cdot) - \mathbf{t}) = \delta(\mathbf{A}(\cdot) - \mathbf{t})$ . Thus, applying the identity (9.16), we find that

$$(10.18) \quad K_{d-k} \mathcal{R}_k B_d^s f_{\text{extreme}} = \sum_{n=1}^N a_n \delta_{\text{iso}}(\cdot - (\mathbf{A}_n, \mathbf{t}_n)),$$

which yields  $\|K_{d-k} \mathcal{R}_k B_d^s f_{\text{extreme}}\|_{\mathcal{M}} = \sum_{n=1}^N |a_n| = \|\mathbf{a}\|_1$  (see item 4, Lemma 10.2).  $\square$

**Appendix A. The isotropic projector.** In order to derive the explicit formula of the operator that projects a function onto its isotropic part, we first consider projecting functions in  $L^2(\mathbb{R}^n)$ . This allows us to define the projector as the solution to an optimization problem. Let  $L_{\text{iso}}^2(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$  denote the closed subspace of isotropic functions in  $L^2(\mathbb{R}^n)$ , as in

$$(A.1) \quad L_{\text{iso}}^2(\mathbb{R}^n) = \{f \in L^2(\mathbb{R}^n) : f(\mathbf{U}\mathbf{x}) = f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n, \mathbf{U} \in O_n(\mathbb{R})\}.$$

The isotropic projection of  $f \in L^2(\mathbb{R}^d)$  is then given by its orthoprojection onto  $L_{\text{iso}}^2(\mathbb{R}^n)$  as

$$(A.2) \quad P_{\text{iso}} f = \arg \min_{g \in L_{\text{iso}}^2(\mathbb{R}^n)} \|f - g\|_{L^2}^2,$$

which exists and is unique by the Hilbert projection theorem. We claim that for any  $f \in L^2(\mathbb{R}^d)$ ,

$$(A.3) \quad P_{\text{iso}}\{f\}(\mathbf{x}) = \int_{O_n(\mathbb{R})} f(\mathbf{U}\mathbf{x}) d\sigma(\mathbf{U}) \in L_{\text{iso}}^2(\mathbb{R}^n),$$

where  $\int$  denotes the average integral, i.e.,

$$(A.4) \quad \int_{O_n(\mathbb{R})} := \frac{1}{\sigma(O_n(\mathbb{R}))} \int_{O_n(\mathbb{R})}.$$

Clearly,  $P_{\text{iso}} P_{\text{iso}} = P_{\text{iso}}$  on  $L^2(\mathbb{R}^n)$ . Next, for any  $f, g \in L^2(\mathbb{R}^n)$ ,

$$(A.5) \quad \begin{aligned} \langle P_{\text{iso}} f, g \rangle &= \int_{\mathbb{R}^n} \left( \int_{O_n(\mathbb{R})} f(\mathbf{U}\mathbf{x}) d\sigma(\mathbf{U}) \right) g(\mathbf{x}) d\mathbf{x} \\ &= \int_{\mathbb{R}^n} \int_{O_n(\mathbb{R})} f(\mathbf{U}\mathbf{x}) g(\mathbf{x}) d\sigma(\mathbf{U}) d\mathbf{x} \end{aligned}$$

$$\begin{aligned}
&= \int_{O_n(\mathbb{R})} \int_{\mathbb{R}^n} f(\mathbf{U}\mathbf{x})g(\mathbf{x}) \, d\mathbf{x} \, d\sigma(\mathbf{U}) \\
&= \int_{O_n(\mathbb{R})} \int_{\mathbb{R}^n} f(\mathbf{y})g(\mathbf{U}^\top\mathbf{y}) \, d\mathbf{y} \, d\sigma(\mathbf{U}) \\
&= \int_{\mathbb{R}^n} \int_{O_n(\mathbb{R})} f(\mathbf{y})g(\mathbf{U}^\top\mathbf{y}) \, d\sigma(\mathbf{U}) \, d\mathbf{y} \\
&= \int_{\mathbb{R}^n} f(\mathbf{y}) \left( \int_{O_n(\mathbb{R})} g(\mathbf{U}^\top\mathbf{y}) \, d\sigma(\mathbf{U}) \right) \, d\mathbf{y} = \langle f, P_{\text{iso}}g \rangle,
\end{aligned}$$

where the change of the order of the integrals is justified by the Fubini–Tonelli theorem since

$$\begin{aligned}
\text{(A.6)} \quad \int_{O_n(\mathbb{R})} \int_{\mathbb{R}^n} |f(\mathbf{U}\mathbf{x})g(\mathbf{x})| \, d\mathbf{x} \, d\sigma(\mathbf{U}) &\leq \int_{O_n(\mathbb{R})} \|f(\mathbf{U}(\cdot))\|_{L^2} \|g\|_{L^2} \, d\sigma(\mathbf{U}) \\
&= \|f\|_{L^2} \|g\|_{L^2} < \infty.
\end{aligned}$$

Thus, for all  $f \in L^2(\mathbb{R}^n)$ , we have that  $P_{\text{iso}}f$  satisfies the orthogonality principle

$$\begin{aligned}
\text{(A.7)} \quad \langle f - P_{\text{iso}}f, P_{\text{iso}}f \rangle &= \langle P_{\text{iso}}f - P_{\text{iso}}P_{\text{iso}}f, f \rangle \\
&= \langle P_{\text{iso}}f - P_{\text{iso}}f, f \rangle \\
&= \langle 0, f \rangle = 0.
\end{aligned}$$

Therefore,  $P_{\text{iso}}f$  is indeed the solution to (A.2).

**Acknowledgments.** The authors would like to thank the anonymous referee for their careful reading of the manuscript and for providing concrete suggestions to improve the clarity of the exposition. Their feedback led to substantial improvement in the content and presentation of this work.

#### REFERENCES

- [1] G. S. ALBERTI, F. BARTOLUCCI, F. DE MARI, AND E. DE VITO, *Unitarization and inversion formulae for the Radon transform between dual pairs*, SIAM J. Math. Anal., 51 (2019), pp. 4356–4381, <https://doi.org/10.1137/18M1225628>.
- [2] F. BACH, *Breaking the curse of dimensionality with convex neural networks*, J. Mach. Learn. Res., 18 (2017), pp. 629–681.
- [3] F. BARTOLUCCI, F. DE MARI, E. DE VITO, AND F. Odone, *The Radon transform intertwines wavelets and shearlets*, Appl. Comput. Harmon. Anal., 47 (2019), pp. 822–847, <https://doi.org/10.1016/j.acha.2017.12.005>.
- [4] F. BARTOLUCCI, E. DE VITO, L. ROSASCO, AND S. VIGOGNA, *Understanding neural networks with reproducing kernel Banach spaces*, Appl. Comput. Harmon. Anal., 62 (2023), pp. 194–236, <https://doi.org/10.1016/j.acha.2022.08.006>.
- [5] M. BENNING AND M. BURGER, *Modern regularization methods for inverse problems*, Acta Numer., 27 (2018), pp. 1–111, <https://doi.org/10.1017/s0962492918000016>.
- [6] C. BOYER, A. CHAMBOLLE, Y. DE CASTRO, V. DUVAL, F. DE GOURNAY, AND P. WEISS, *On representer theorems and convex regularization*, SIAM J. Optim., 29 (2019), pp. 1260–1281, <https://doi.org/10.1137/18M1200750>.
- [7] K. BREDIES AND M. CARIONI, *Sparsity of solutions for variational inverse problems with finite-dimensional data*, Calc. Var. Partial Differential Equations, 59 (2020), 14, <https://doi.org/10.1007/s00526-019-1658-1>.
- [8] E. J. CANDÈS, *Ridgelets: Theory and Applications*, Ph.D. thesis, Stanford University, 1998.
- [9] E. J. CANDÈS, *Harmonic analysis of neural networks*, Appl. Comput. Harmon. Anal., 6 (1999), pp. 197–218, <https://doi.org/10.1006/acha.1998.0248>.
- [10] A. S. DALALYAN, A. JUDITSKY, AND V. SPOKOINY, *A new algorithm for estimating the effective dimension-reduction subspace*, J. Mach. Learn. Res., 9 (2008), pp. 1647–1678.

- [11] D. L. DONOHO, *Tight frames of  $k$ -plane ridgelets and the problem of representing objects that are smooth away from  $d$ -dimensional singularities in  $\mathbb{R}^n$* , Proc. Natl. Acad. Sci. USA, 96 (1999), pp. 1828–1833, <https://doi.org/10.1073/pnas.96.5.1828>.
- [12] D. L. DONOHO, *High-Dimensional Data Analysis: The Curses and Blessings of Dimensionality*, AMS Lectures, 2000.
- [13] D. L. DONOHO, *Orthonormal ridgelets and linear singularities*, SIAM J. Math. Anal., 31 (2000), pp. 1062–1099, <https://doi.org/10.1137/S0036141098344403>.
- [14] N. DUNFORD AND J. T. SCHWARTZ, *Linear Operators, Part I: General Theory*, Wiley Classics Library, John Wiley & Sons, New York, 1988.
- [15] A. EDELMAN, 18.325: *Finite Random Matrix Theory Volumes and Integration*, 2005, <https://web.mit.edu/18.325/www/handouts/handout4.pdf>.
- [16] H. W. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Math. Appl. 375, Kluwer, Dordrecht, 1996.
- [17] G. B. FOLLAND, *Real Analysis: Modern Techniques and Their Applications*, 2nd ed., Pure Appl. Math. (New York), John Wiley & Sons, New York, 1999.
- [18] B. FUGLEDE, *An integral formula*, Math. Scand., 6 (1958), pp. 207–212, <https://doi.org/10.7146/math.scand.a-10545>.
- [19] K. FUKUMIZU, F. R. BACH, AND M. I. JORDAN, *Dimensionality reduction for supervised learning with reproducing kernel Hilbert spaces*, J. Mach. Learn. Res., 5 (2004), pp. 73–99.
- [20] I. M. GEL'FAND, M. I. GRAEV, AND N. Y. VILENKIN, *Generalized Functions, Integral Geometry and Representation Theory 5*, Academic Press, New York, 1966.
- [21] B. GHORBANI, S. MEI, T. MISIAKIEWICZ, AND A. MONTANARI, *When do neural networks outperform kernel methods?*, Adv. Neural Inf. Process. Syst., 33 (2020), pp. 14820–14830.
- [22] F. B. GONZALEZ, *On the range of the Radon  $d$ -plane transform and its dual*, Trans. Amer. Math. Soc., 327 (1991), pp. 601–619, <https://doi.org/10.2307/2001816>.
- [23] L. GRAFAKOS, *Modern Fourier Analysis*, 2nd ed., Grad. Texts in Math. 250, Springer, New York, 2009, <https://doi.org/10.1007/978-0-387-09434-2>.
- [24] E. L. GRINBERG, *On images of Radon transforms*, Duke Math. J., 52 (1985), pp. 939–972, <https://doi.org/10.1215/S0012-7094-85-05251-2>.
- [25] S. HELGASON, *A duality in integral geometry; some generalizations of the Radon transform*, Bull. Amer. Math. Soc., 70 (1964), pp. 435–446, <https://doi.org/10.1090/S0002-9904-1964-11147-2>.
- [26] S. HELGASON, *The Radon transform on Euclidean spaces, compact two-point homogeneous spaces and Grassmann manifolds*, Acta Math., 113 (1965), pp. 153–180, <https://doi.org/10.1007/BF02391776>.
- [27] S. HELGASON, *Integral Geometry and Radon Transforms*, Springer, New York, 2011.
- [28] F. JOHN, *Bestimmung einer Funktion aus ihren Integralen über gewisse Mannigfaltigkeiten*, Math. Ann., 109 (1934), pp. 488–520, <https://doi.org/10.1007/BF01449151>.
- [29] F. JOHN, *Abhängigkeiten zwischen den Flächenintegralen einer stetigen Funktion*, Math. Ann., 111 (1935), pp. 541–559, <https://doi.org/10.1007/BF01472237>.
- [30] F. JOHN, *Plane Waves and Spherical Means Applied to Partial Differential Equations*, Interscience, New York, 1955.
- [31] A. KATSEVICH, *Resolution analysis of inverting the generalized Radon transform from discrete data in  $\mathbb{R}^3$* , SIAM J. Math. Anal., 52 (2020), pp. 3990–4021, <https://doi.org/10.1137/19M1295039>.
- [32] F. KEINERT, *Inversion of  $k$ -plane transforms and applications in computer tomography*, SIAM Rev., 31 (1989), pp. 273–298, <https://doi.org/10.1137/1031051>.
- [33] S. KOSTADINOVA, S. PILIPOVIĆ, K. SANEVA, AND J. VINDAS, *The ridgelet transform of distributions*, Integral Trans. Special Funct. Int. J., 25 (2014), pp. 344–358, <https://doi.org/10.1080/10652469.2013.853057>.
- [34] S. KOSTADINOVA, S. PILIPOVIĆ, K. SANEVA, AND J. VINDAS, *The ridgelet transform and quasiasymptotic behavior of distributions*, in Pseudo-Differential Operators and Generalized Functions, Springer, New York, 2015, pp. 185–197.
- [35] V. KÛARKOVÁA, P. C. KAINEN, AND V. KREINOVICH, *Estimates of the number of hidden units and variation with respect to half-spaces*, Neural Netw., 10 (1997), pp. 1061–1068.
- [36] B. F. LOGAN AND L. A. SHEPP, *Optimal reconstruction of a function from its projections*, Duke Math. J., 42 (1975), pp. 645–659.
- [37] D. LUDWIG, *The Radon transform on Euclidean space*, Comm. Pure Appl. Math., 19 (1966), pp. 49–81, <https://doi.org/10.1002/cpa.3160190207>.
- [38] A. MARKOE, *Analytic Tomography*, Encyclopedia Math. Appl. 106, Cambridge University Press, Cambridge, UK, 2006, <https://doi.org/10.1017/CBO9780511530012>.

- [39] A. MARKOE AND E. T. QUINTO, *An elementary proof of local invertibility for generalized and attenuated Radon transforms*, SIAM J. Math. Anal., 16 (1985), pp. 1114–1119, <https://doi.org/10.1137/0516082>.
- [40] S. NEUMAYER AND M. UNSER, *Explicit representations for Banach subspaces of Lizorkin distributions*, Anal. Appl., 21 (2023), pp. 1223–1250.
- [41] G. ONGIE, R. WILLET, D. SOUDRY, AND N. SREBRO, *A function space view of bounded norm infinite width ReLU nets: The multivariate case*, in Proceedings of the International Conference on Learning Representations, 2020, pp. 1–24.
- [42] R. PARHI AND R. D. NOWAK, *Banach space representer theorems for neural networks and ridge splines*, J. Mach. Learn. Res., 22 (2021), 43.
- [43] R. PARHI AND R. D. NOWAK, *What kinds of functions do deep neural networks learn? Insights from variational spline theory*, SIAM J. Math. Data Sci., 4 (2022), pp. 464–489, <https://doi.org/10.1137/21M1418642>.
- [44] R. PARHI AND R. D. NOWAK, *Deep learning meets sparse regularization: A signal processing perspective*, IEEE Signal Process. Mag., 40 (2023), pp. 63–74, <https://doi.org/10.1109/MSP.2023.3286988>.
- [45] R. PARHI AND R. D. NOWAK, *Near-minimax optimal estimation with shallow ReLU neural networks*, IEEE Trans. Inform. Theory, 69 (2023), pp. 1125–1140, <https://doi.org/10.1109/TIT.2022.3208653>.
- [46] E. E. PETROV, *Cavalieri conditions for a  $k$ -dimensional Radon transformation*, Math. Notes Acad. Sci. USSR, 50 (1991), pp. 1135–1141, <https://doi.org/10.1007/BF01157700>.
- [47] J. RADON, *Über die bestimmung von funktionen durch ihre integralwerte längs gewisser mannigfaltigkeiten*, Berichte Sächsischer Akademie der Wissenschaften, 69 (1917), pp. 262–277.
- [48] A. G. RAMM, *Radon transform on distributions*, Proc. Japan Acad. Ser. A. Math. Sci., 71 (1995), pp. 202–206.
- [49] A. G. RAMM, *Inversion formula and singularities of the solution for the backprojection operator in tomography*, Proc. Amer. Math. Soc., 124 (1996), pp. 567–577, <https://doi.org/10.1090/S0002-9939-96-03155-3>.
- [50] A. G. RAMM AND A. I. KATSEVICH, *The Radon Transform and Local Tomography*, CRC Press, Boca Raton, FL, 1996.
- [51] F. RICHTER, *On the  $k$ -dimensional Radon-transform of rapidly decreasing functions*, in Differential Geometry, Lecture Notes in Math. 1209, Springer, Berlin, 1986, pp. 243–258, <https://doi.org/10.1007/BFb0076636>.
- [52] B. RUBIN, *The Calderón reproducing formula, windowed  $X$ -ray transforms, and Radon transforms in  $L^p$ -spaces*, J. Fourier Anal. Appl., 4 (1998), pp. 175–197, <https://doi.org/10.1007/BF02475988>.
- [53] B. RUBIN, *Inversion of  $k$ -plane transforms via continuous wavelet transforms*, J. Math. Anal. Appl., 220 (1998), pp. 187–203, <https://doi.org/10.1006/jmaa.1997.5852>.
- [54] B. RUBIN, *Helgason-Marchaud inversion formulas for Radon transforms*, Proc. Amer. Math. Soc., 130 (2002), pp. 3017–3023, <https://doi.org/10.1090/S0002-9939-02-06554-1>.
- [55] B. RUBIN, *Convolution-backprojection method for the  $k$ -plane transform, and Calderón’s identity for ridgelet transforms*, Appl. Comput. Harmon. Anal., 16 (2004), pp. 231–242, <https://doi.org/10.1016/j.acha.2004.03.003>.
- [56] B. RUBIN, *Reconstruction of functions from their integrals over  $k$ -planes*, Israel J. Math., 141 (2004), pp. 93–117, <https://doi.org/10.1007/BF02772213>.
- [57] W. RUDIN, *Functional Analysis*, 2nd ed., Internat. Ser. Pure Appl. Math., McGraw-Hill, New York, 1991.
- [58] S. SAMKO, *Denseness of the spaces  $\Phi_V$  of Lizorkin type in the mixed  $L^{\vec{p}}(\mathbb{R}^n)$ -spaces*, Stud. Math., 113 (1995), pp. 199–210, <https://doi.org/10.4064/sm-113-3-199-210>.
- [59] L. SCHWARTZ, *Théorie des distributions*, Publications de l’Institut de Mathématique de l’Université de Strasbourg, IX-X, Hermann, Paris, 1966.
- [60] V. A. SHARAFUTDINOV, *Radon transform on Sobolev spaces*, Sib. Math. J., 62 (2021), pp. 560–580 (in English); Sibirsk. Mat. Zh. 62 (2021), pp. 690–714 (in Russian).
- [61] J. SHENOUDA, R. PARHI, AND R. D. NOWAK, *A continuous transform for localized ridgelets*, in Proceedings of the International Conference on Sampling Theory and Applications (SampTA), IEEE, 2023, pp. 1–5.
- [62] K. T. SMITH, D. C. SOLMON, AND S. L. WAGNER, *Practical and mathematical aspects of the problem of reconstructing objects from radiographs*, Bull. Amer. Math. Soc., 83 (1977), pp. 1227–1270, <https://doi.org/10.1090/S0002-9904-1977-14406-6>.
- [63] D. C. SOLMON, *The  $X$ -ray transform*, J. Math. Anal. Appl., 56 (1976), pp. 61–83, [https://doi.org/10.1016/0022-247X\(76\)90008-1](https://doi.org/10.1016/0022-247X(76)90008-1).



- [64] D. C. SOLMON, *A note on  $k$ -plane integral transforms*, J. Math. Anal. Appl., 71 (1979), pp. 351–358.
- [65] D. C. SOLMON, *Asymptotic formulas for the dual Radon transform and applications*, Math. Z., 195 (1987), pp. 321–343, <https://doi.org/10.1007/BF01161760>.
- [66] S. SONODA AND N. MURATA, *Neural network with unbounded activation functions is universal approximator*, Appl. Comput. Harmon. Anal., 43 (2017), pp. 233–268, <https://doi.org/10.1016/j.acha.2015.12.005>.
- [67] F. TRÈVES, *Topological Vector Spaces, Distributions and Kernels*, Academic Press, New York, 1967.
- [68] M. UNSER, *A unifying representer theorem for inverse problems and machine learning*, Found. Comput. Math., 21 (2021), pp. 941–960, <https://doi.org/10.1007/s10208-020-09472-x>.
- [69] M. UNSER, *Ridges, neural networks, and the Radon transform*, J. Mach. Learn. Res., 24 (2023), 37.
- [70] H. WENDLAND, *Scattered Data Approximation*, Cambridge Monogr. Appl. Comput. Math. 17, Cambridge University Press, Cambridge, UK, 2005.
- [71] Е. Е. ПЕТРОВ, *Условия Кавальери для  $k$ -мерного преобразования Радона*, Математические заметки, 50 (1991), pp. 61–68.
- [72] Е. Е. ПЕТРОВ, *Условия моментов для  $k$ -мерного преобразования Радона*, Известия высших учебных заведений, Математика, (1992), pp. 49–53.
- [73] Е. Е. ПЕТРОВ, *Условие Кавальери для интегралов по  $k$ -мерным плоскостям,  $k > n - 1$* , Алгебра и анализ, 5 (1993), pp. 191–205.