

Monte Carlo wavelets: a randomized approach to frame discretization

Zeljko Kereta
SSRI - MIND

Simula Research Laboratory
1364 Fornebu - Norway
Email: zeljko@simula.no

Stefano Vigogna
Dibris

University of Genova
16146 Genova - Italy
Email: vigogna@dibris.unige.it

Valeriya Naumova
MIND - SimulaMet

Simula Research Laboratory
0167 Oslo - Norway
Email: valeriya@simula.no

Lorenzo Rosasco
Dibris - LCSL

University of Genova, MIT, IIT
16146 Genova - Italy
Email: lorenzo.rosasco@unige.it

Ernesto De Vito

Department of Mathematics
University of Genova
16146 Genova - Italy
Email: devito@dima.unige.it

Abstract—In this paper we propose and study a family of continuous wavelets on general domains, and a corresponding stochastic discretization that we call Monte Carlo wavelets. First, using tools from the theory of reproducing kernel Hilbert spaces and associated integral operators, we define a family of continuous wavelets by spectral calculus. Then, we propose a stochastic discretization based on Monte Carlo estimates of integral operators. Using concentration of measure results, we establish the convergence of such a discretization and derive convergence rates under natural regularity assumptions.

I. INTRODUCTION

To construct a discrete frame it is common practice to first construct a frame on a continuous parameter (measure) space, where the mathematical properties are nice and the structure is rich, and then obtain a discrete frame by carefully selecting a discrete subset of parameters. The discretization problem has been studied extensively and in several settings [5], [2], [7], [6]. Although theoretically relevant in establishing density theorems and complete characterizations, some general discretization procedures require assumptions which might be difficult to confirm, and they are often not constructive. On the other hand, discrete parameter selection in constructive discretization methods is usually sensitive, and it does not generalize trivially from the Euclidean space to more general geometries [1]. In the following, we propose a different approach based on random sampling.

First, we consider a general domain with an associated positive definite kernel, and exploit the theory of reproducing kernel Hilbert spaces (RKHS) and corresponding integral operators to define a family of continuous wavelets. Here, we borrow ideas from [8], [9], [2], and develop them using the properties of a RKHS. Then, we propose a stochastic discretization approach which replaces constructive discrete parameter selection with random sampling. We note that the corresponding discrete frame is actually only finite-dimensional. In particular, it does not solve the discretization

problem in the classical sense. Indeed, each discrete frame is a frame only up to a certain sampling resolution. The frame dimensions increase with the number of samples, and our main technical contribution is proving convergence to the continuous frame as the number of samples goes to infinity. Further, we derive convergence rates under natural regularity assumptions [4]. Our analysis combines classical tools from approximation theory and spectral calculus with results of concentration of measure to deal with random sampling.

The paper is organized as follows. In Section II we introduce the general framework in which we will conduct our analysis. This will allow us to state the continuous and discrete settings within a unified formalism. The key ingredient is a reproducing kernel and the associated integral operator. Section III contains the general frame construction based on spectral filters and eigenfunctions of the integral operator. In Section IV we consider a specific class of filters, stemming from the theory of inverse problems, for which we will derive explicit approximation rates. In Section V we specialize the construction of Section III in a continuous and a discrete frame, regarding the latter as a Monte Carlo estimate of the former. Our main result, Theorem 4, establishes quantitative consistency of the Monte Carlo estimate on a class of Sobolev-regular signals. In Section VI we provide an implementable formula to compute our Monte Carlo wavelets from the eigendecomposition of a sample kernel matrix.

II. SETTING

Let \mathcal{X} be a locally compact, second countable topological space endowed with a Borel probability measure ρ . We fix a continuous positive semi-definite kernel

$$K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}, \quad (1)$$

and let \mathcal{H} be the corresponding reproducing kernel Hilbert space

$$\mathcal{H} = \{f : \mathcal{X} \rightarrow \mathbb{C} : f(x) = \langle f, K_x \rangle\} = \overline{\text{span}}\{K_x\}_{x \in \mathcal{X}}, \quad (2)$$

symbol	definition	Notation
$\langle \cdot, \cdot \rangle$	inner product in a Hilbert space	
$\ \cdot \ $	norm of a Hilbert space, or operator norm	
$\ \cdot \ _{\text{HS}}$	Hilbert-Schmidt norm	
$\text{span}\{S\}$	linear span of the set S	
\overline{S}	topological closure of the set S	
P_S	orthogonal projection onto the closed subspace S	
$\text{supp}(\rho)$	support of the measure ρ	
δ_x	Dirac measure at x	
$X \lesssim Y$	$X \leq CY$ for some constant $C > 0$	
$\mathbf{u}[i]$	i -th component of the vector $\mathbf{u} \in \mathbb{R}^N$	
$\mathbf{M}[i, j]$	(i, j) -th component of the matrix $\mathbf{M} \in \mathbb{R}^{N \times N}$	

with $K_x = K(\cdot, x)$ and inner product given by $\langle K_x, K_y \rangle = K(x, y)$. Assuming $\sup_{x \in X} K(x, x) \leq \kappa^2$, we define the integral operator $T : \mathcal{H} \rightarrow \mathcal{H}$

$$Tf(x) = \int_{\mathcal{X}} K(x, y) f(y) d\rho(y). \quad (3)$$

Under these assumptions, T is a positive self-adjoint trace class operator with spectrum $\sigma(T) \subset [0, \kappa^2]$. Let

$$Tv_i = \begin{cases} \lambda_i v_i & i \in I_\rho \\ 0 & i \in I_0 \end{cases}$$

be the spectral decomposition of T over the orthonormal basis $\{v_i\}_{i \in I_\rho \cup I_0}$. The subspace

$$\mathcal{H}_\rho = \overline{\text{span}}\{K_x\}_{x \in \text{supp}(\rho)} = \overline{\text{span}}\{v_i\}_{i \in I_\rho} \quad (4)$$

splits \mathcal{H} as $\mathcal{H} = \mathcal{H}_\rho \oplus \ker T$.

As examples of this setting, we can think of \mathcal{X} as \mathbb{R}^d , or non-Euclidean domains such as a compact connected Riemannian manifold or a weighted graph. In all these cases, we can take K as the heat kernel associated with the proper notion of Laplacian, be it the Laplace-Beltrami operator or the graph Laplacian.

III. FRAME CONSTRUCTION

To construct our frame, we take inspiration from [8] and references therein. We start by defining filters on the spectrum of T . Let $(G_j)_{j \in \mathbb{N}}$ be a family of bounded measurable functions $G_j : [0, \infty) \rightarrow [0, \infty)$ satisfying

$$\sum_{j \in \mathbb{N}} \lambda G_j(\lambda)^2 = 1 \quad \lambda \in (0, \kappa^2]. \quad (5)$$

$G_j(T)$ is then a positive, self-adjoint operator on \mathcal{H} with $\sigma(G_j(T)) = G_j(\sigma(T)) \subset [0, G_j(\kappa^2)]$, for every j . We can thus define

$$\psi_{j,x} = G_j(T)K_x \quad j \in \mathbb{N}, x \in \mathcal{X}. \quad (6)$$

Note that, if $x \in \text{supp}(\rho)$, then $\psi_{j,x} \in \mathcal{H}_\rho$ since $G_j(T)\mathcal{H}_\rho \subset \mathcal{H}_\rho$. Moreover, using self-adjointness of $G_j(T)$, the reproducing property (2), and expression (4), we can rewrite $\psi_{j,x}$ as

$$\psi_{j,x} = \sum_{i \in I_\rho} G_j(\lambda_i) \overline{v_i(x)} v_i. \quad (7)$$

This reveals (6) as a generalization of classical wavelets, where v_i 's correspond to eigenfunctions of the Laplacian. We may therefore interpret x and j as location and scale parameters, respectively.

The following proposition shows that the family defined through (6) is a Parseval frame on \mathcal{H}_ρ .

Proposition 1. For every $f \in \mathcal{H}$,

$$\|P_{\mathcal{H}_\rho} f\|^2 = \sum_{j \in \mathbb{N}} \int_{\mathcal{X}} |\langle f, \psi_{j,x} \rangle|^2 d\rho(x).$$

Proof. Let $f \in \text{span}\{v_i\}_{i \in I_\rho \cup I_0}$. Since $G_j(T)$ is self-adjoint, we have $\langle f, \psi_{j,x} \rangle = \langle G_j(T)f, K_x \rangle$. Hence, integrating over \mathcal{X} we get

$$\begin{aligned} \int_{\mathcal{X}} |\langle f, \psi_{j,x} \rangle|^2 d\rho(x) &= \langle TG_j(T)f, G_j(T)f \rangle \\ &= \langle TG_j(T)^2 f, f \rangle \\ &= \sum_{i \in I_\rho} \lambda_i G_j(\lambda_i)^2 |\langle f, v_i \rangle|^2. \end{aligned}$$

Summation over \mathbb{N} along with (5) thus gives

$$\begin{aligned} \sum_{j \in \mathbb{N}} \langle G_j(T)^2 Tf, f \rangle &= \sum_{i \in I_\rho} |\langle f, v_i \rangle|^2 \sum_{j \in \mathbb{N}} \lambda_i G_j(\lambda_i)^2 \\ &= \sum_{i \in I_\rho} |\langle f, v_i \rangle|^2 = \|P_{\mathcal{H}_\rho} f\|^2. \end{aligned}$$

The assertion follows since $\text{span}\{v_i\}_{i \in I_\rho \cup I_0}$ is dense in \mathcal{H} . \square

IV. EXAMPLES OF SPECTRAL FILTERS

Let $(g_j)_{j \in \mathbb{N}}$ be a family of functions $g_j : [0, \infty) \rightarrow [0, \infty)$ such that

$$0 \leq g_j \leq g_{j+1}, \quad \lim_{j \rightarrow \infty} \lambda g_j(\lambda) = 1. \quad (8)$$

We can define filters obeying (5) by means of (8), as

$$G_0(\lambda) = \sqrt{g_0(\lambda)}, \quad G_{j+1}(\lambda) = \sqrt{g_{j+1}(\lambda) - g_j(\lambda)}.$$

The *qualification* of a spectral function $g_j(\lambda)$ is a constant $\bar{\nu} \in (0, \infty]$ such that, for all $\nu > 0$,

$$\sup_{\lambda \in (0, \kappa^2]} \lambda^\nu (1 - \lambda g_j(\lambda)) \lesssim j^{-\min\{\nu, \bar{\nu}\}}.$$

As instances of $g_j(\lambda)$, we can take spectral functions in Table I (see also [3]).

TABLE I

method	$g_j(\lambda)$	qualification
Tikhonov regularization	$\frac{1}{\lambda+1/j}$	1
Iterated Tikhonov	$\frac{(\lambda+1/j)^m - (1/j)^m}{\lambda(\lambda+1/j)^m}$	m
Landweber iteration	$\frac{1}{\lambda}(1 - (1 - \gamma\lambda)^j)$	∞
Asymptotic regularization	$\frac{1}{\lambda}(1 - \exp(-j\lambda))$	∞

These functions play a fundamental role in ill-posed inverse problems regularization, where the error of a given regularizer decays with a rate governed by its qualification.

V. CONSISTENCY

From now on, assume $\text{supp } \rho = \mathcal{X}$, whence $\mathcal{H}_\rho = \mathcal{H}$ by (4). Therefore, (6) is a frame on the whole space \mathcal{H} , thanks to Proposition 1. Suppose we are given N independent and identically distributed samples

$$x_1, \dots, x_N \sim \rho. \quad (9)$$

Then we can consider the empirical distribution

$$\widehat{\rho}_N = \frac{1}{N} \sum_{i=1}^N \delta_{x_i},$$

define the empirical integral operator

$$\widehat{T}f(x) = \frac{1}{N} \sum_{i=1}^N K(x, x_i) f(x_i) \quad (10)$$

and repeat the construction of Section III, to get

$$\widehat{\psi}_{j,x_i} = G_j(\widehat{T})K_{x_i} \quad j \in \mathbb{N}, i = 1, \dots, N. \quad (11)$$

In view of Proposition 1, (11) defines a discrete Parseval frame on

$$\widehat{\mathcal{H}}_N = \mathcal{H}_{\widehat{\rho}_N} = \text{span}\{K_{x_i}\}_{i=1}^N \simeq \mathbb{R}^N.$$

Since (10) is a Monte Carlo estimate of (3), we call (11) a family of *Monte Carlo wavelets*. Note that, as we take more and more samples, we obtain a sequence of discrete Parseval frames on a sequence of nested subspaces of increasing dimension:

$$\widehat{\mathcal{H}}_N \subset \widehat{\mathcal{H}}_{N+1} \subset \dots \subset \mathcal{H}.$$

Our goal is to establish consistency of the Monte Carlo estimate (11), that is, to see whether, and in what sense, (11) tends to (6) as $N \rightarrow \infty$.

Now, let

$$T_j f = \int_{\mathcal{X}} \langle f, \psi_{j,x} \rangle \psi_{j,x} d\rho(x), \quad \widehat{T}_j f = \frac{1}{N} \sum_{i=1}^N \langle f, \widehat{\psi}_{j,x_i} \rangle \widehat{\psi}_{j,x_i} \quad (12)$$

be the frame operator at scale j and its empirical counterpart. Proposition 1 gives a resolution of the identity, which can be approximated by a truncated empirical version:

$$I_{\mathcal{H}} = \sum_{j \in \mathbb{N}} T_j \approx \sum_{j \leq \tau} \widehat{T}_j. \quad (13)$$

We split the resolution error of (13) on a signal $f \in \mathcal{H}$ into an approximation and a sampling error,

$$\|f - \sum_{j \leq \tau} \widehat{T}_j f\| \leq \left\| \sum_{j > \tau} T_j f \right\| + \left\| \sum_{j \leq \tau} (T_j - \widehat{T}_j) f \right\|, \quad (14)$$

and derive quantitative bounds for each term (Propositions 2 and 3). Tuning the resolution τ in terms of the sample size N , will then yield our consistency result in Theorem 4.

For what concerns the approximation error, note that Proposition 1 already implies

$$\left\| \sum_{j > \tau} T_j f \right\| \xrightarrow{\tau \rightarrow \infty} \rightarrow 0,$$

by dominated convergence. However, in order to obtain a convergence rate, we need to assume some regularity of the signal f . Specifically, we will assume the following Sobolev smoothness condition [4]:

$$f = T^\alpha h \quad \text{for some } h \in \mathcal{H}, \alpha > 0.$$

Proposition 2 (Approximation error). *Let $g_j(\lambda)$ as in (8). Suppose $g_j(\lambda)$ has qualification $\bar{\nu}$. Then, for every $f \in T^\alpha \mathcal{H}$,*

$$\left\| \sum_{j > \tau} T_j f \right\| \lesssim \|T^{-\alpha} f\| \kappa^{2(\alpha-\beta)} \tau^{-\beta}$$

where $\beta = \min\{\alpha, \bar{\nu}\}$.

Proof. By Lemma 5 we have $\sum_{j \leq \tau} T_j = T g_\tau(T)$, hence

$$\begin{aligned} \left\| \sum_{j > \tau} T_j f \right\|^2 &= \sum_i |1 - \lambda_i g_\tau(\lambda_i)|^2 |\langle f, v_i \rangle|^2 \\ &= \sum_i \lambda_i^{2\beta} |1 - \lambda_i g_\tau(\lambda_i)|^2 |\langle T^{-\beta} f, v_i \rangle|^2. \end{aligned}$$

Thus,

$$\begin{aligned} \left\| \sum_{j > \tau} T_j f \right\| &\leq \|T^{-\beta} f\| \sup_i \lambda_i^\beta (1 - \lambda_i g_\tau(\lambda_i)) \\ &\lesssim \|T^{-\beta} f\| \tau^{-\beta}. \quad \square \end{aligned}$$

Bound on the sampling error follows by concentration of the empirical integral operator \widehat{T} with respect to the continuous operator T .

Proposition 3 (Sampling error). *Let $g_j(\lambda)$ be as in (8). Suppose $\lambda \mapsto \lambda g_j(\lambda)$ is Lipschitz continuous on $[0, \kappa^2]$ with Lipschitz constant bounded by j . Then, for every $f \in \mathcal{H}$, with probability higher than $1 - 2e^{-t}$,*

$$\left\| \sum_{j \leq \tau} (T_j - \widehat{T}_j) f \right\| \lesssim \|f\| \kappa^2 \sqrt{t} \tau N^{-1/2}.$$

Proof. Using Lemma 6 we have

$$\begin{aligned} \left\| \sum_{j \leq \tau} (T_j - \widehat{T}_j) f \right\| &\leq \|T g_\tau(T) - \widehat{T} g_\tau(\widehat{T})\| \|f\| \\ &\leq \|T g_\tau(T) - \widehat{T} g_\tau(\widehat{T})\|_{\text{HS}} \|f\| \\ &\leq \tau \|T - \widehat{T}\|_{\text{HS}} \|f\|. \end{aligned}$$

Thanks to the reproducing property (2), we can write

$$T = \int_{\mathcal{X}} K_x \otimes K_x d\rho(x), \quad \widehat{T} = \frac{1}{N} \sum_{i=1}^N K_{x_i} \otimes K_{x_i}.$$

This allows for the concentration in Lemma 7, which yields

$$\|T - \widehat{T}\|_{\text{HS}} \lesssim \kappa^2 \sqrt{t} N^{-1/2}$$

with probability higher than $1 - 2e^{-t}$. \square

We are finally ready to state the consistency of our Monte Carlo discretization.

Theorem 4. *Let $g_j(\lambda)$ as in (8). Suppose $g_j(\lambda)$ has qualification $\bar{\nu}$, and $\lambda \mapsto \lambda g_j(\lambda)$ is Lipschitz continuous on $[0, \kappa^2]$ with Lipschitz constant bounded by j . Then, for every $f \in T^\alpha \mathcal{H}$, with probability higher than $1 - 2e^{-t}$,*

$$\begin{aligned} \|f - \sum_{j \leq N^{\frac{1}{2\beta+2}}} \sum_{i=1}^N \langle f, \widehat{\psi}_{j,x_i} \rangle \widehat{\psi}_{j,x_i}\| \\ \lesssim \|T^{-\alpha} f\| (\kappa^{2(\alpha-\beta)} + \kappa^{2\alpha+2} \sqrt{t}) N^{-\frac{\beta}{2\beta+2}} \end{aligned}$$

where $\beta = \min\{\alpha, \bar{\nu}\}$.

Proof. We split $\|f - \sum_{j \leq \tau} \widehat{T}_j f\|$ in (14), and combine the bounds in Propositions 2 and 3. In order to balance the two terms, we choose $\tau = N^{\frac{1}{2\beta+2}}$, which concludes the proof. \square

In particular, in view of Lemma 8, we can use the spectral filters defined in Table I, for which we obtain the following rates:

TABLE II

method	rate
Tikhonov regularization	$N^{-\frac{\min\{\alpha, 1\}}{2 \min\{\alpha, 1\} + 2}}$
Iterated Tikhonov	$N^{-\frac{\min\{\alpha, m\}}{2 \min\{\alpha, m\} + 2}}$
Landweber iteration	$N^{-\frac{\alpha}{2\alpha+2}}$
Asymptotic regularization	$N^{-\frac{\alpha}{2\alpha+2}}$

VI. NUMERICAL IMPLEMENTATION

To implement the system of Monte Carlo wavelets in (11) we exploit equation (7), that is,

$$\widehat{\psi}_{j,x_k} = \frac{1}{N} \sum_{i=1}^N G_j(\widehat{\lambda}_i) \widehat{v}_i(x_k) \widehat{v}_i,$$

where $\widehat{T} \widehat{v}_i = \widehat{\lambda}_i \widehat{v}_i$ is the spectral decomposition of \widehat{T} . In particular, we need to compute eigenvalues and eigenvectors of \widehat{T} , translating the calculations from $\widehat{\mathcal{H}}$ to \mathbb{R}^N and back [10]. To this end, we introduce the restriction map $\widehat{S} : \widehat{\mathcal{H}} \rightarrow \mathbb{R}^N$

$$(\widehat{S}f)[i] = f(x_i) \quad i = 1, \dots, N, \quad (15)$$

whose adjoint $\widehat{S}^* : \mathbb{R}^N \rightarrow \widehat{\mathcal{H}}$ defines the out-of-sample extension

$$\widehat{S}^* \mathbf{u} = \frac{1}{N} \sum_{\ell=1}^N \mathbf{u}[\ell] K_{x_\ell}.$$

Let

$$\widehat{\mathbf{K}}[i, k] = \frac{1}{N} K(x_i, x_k), \quad i, k = 1, \dots, N \quad (16)$$

be the (normalized) sample kernel matrix. Now $\widehat{T} = \widehat{S}^* \widehat{S}$ and $\widehat{\mathbf{K}} = \widehat{S} \widehat{S}^*$, hence \widehat{T} and $\widehat{\mathbf{K}}$ share the same eigenvalues. Moreover, the eigenvectors of \widehat{T} are related to the ones of $\widehat{\mathbf{K}}$ by the singular value decomposition of \widehat{S} :

$$\widehat{S} \widehat{v}_i = \sqrt{\widehat{\lambda}_i} \widehat{\mathbf{u}}_i, \quad i = 1, \dots, N.$$

Thus, we have

$$\widehat{v}_i = \widehat{\lambda}_i^{-\frac{1}{2}} \widehat{S}^* \widehat{\mathbf{u}}_i = \widehat{\lambda}_i^{-\frac{1}{2}} \frac{1}{N} \sum_{\ell=1}^N \widehat{\mathbf{u}}_i[\ell] K_{x_\ell},$$

which evaluated at x_k gives

$$\widehat{v}_i(x_k) = \frac{\widehat{\lambda}_i^{-\frac{1}{2}}}{N} \sum_{\ell=1}^N K(x_k, x_\ell) \widehat{\mathbf{u}}_i[\ell] = \widehat{\lambda}_i^{-\frac{1}{2}} (\widehat{\mathbf{K}} \widehat{\mathbf{u}}_i)[k] = \sqrt{\widehat{\lambda}_i} \widehat{\mathbf{u}}_i[k].$$

We therefore obtain

$$\widehat{\psi}_{j,x_k}(x) = \frac{1}{N} \sum_{i,\ell=1}^N G_j(\widehat{\lambda}_i) \widehat{\mathbf{u}}_i[k] \widehat{\mathbf{u}}_i[\ell] K(x, x_\ell). \quad (17)$$

In summary, given a kernel (1), filters (5) and samples (9), a numerical implementation of the Monte Carlo wavelets (11) can be performed from the sample kernel matrix (16) alone following the steps below:

1. compute eigenvalues and eigenvectors $(\widehat{\lambda}_i, \widehat{\mathbf{u}}_i)_{i=1}^N$ of $\widehat{\mathbf{K}}$;
2. apply equation (17).

LEMmata

Lemma 5. *Let T , G_j and T_j be defined as in (3), (5) and (12). Then $T_j = T G_j(T)^2$.*

Lemma 6. *Let A, B be self-adjoint operators on a Hilbert space \mathcal{H} , and let $F : \mathbb{R} \rightarrow \mathbb{C}$ be a Lipschitz continuous function with Lipschitz constant L . Then*

$$\|F(A) - F(B)\|_{\text{HS}} \leq L \|A - B\|_{\text{HS}}.$$

Lemma 7 ([10, Theorem 7]). *The operators T and \widehat{T} defined in (3) and (10) are Hilbert-Schmidt and*

$$\|T - \widehat{T}\|_{\text{HS}} \leq 2\sqrt{2} \kappa^2 \sqrt{t} N^{-1/2}$$

with probability higher than $1 - 2e^{-t}$.

Lemma 8. *Let $g_j(\lambda)$ be as in Table I. The function $\lambda \mapsto \lambda g_j(\lambda)$ is Lipschitz continuous on $[0, \kappa^2]$, with Lipschitz constant bounded by j .*

ACKNOWLEDGMENTS

Z. K. and V. N. acknowledge the support from RCN-funded FunDaHD project No 251149/O70.

L. R. acknowledges the financial support of the AFOSR projects FA9550-17-1-0390 and BAA-AFRL-AFOSR-2016-0007 (European Office of Aerospace Research and Development), and the EU H2020-MSCA-RISE project NoMADS - DLV-777826 and Axpo Italia SpA. L.R. acknowledge the support of NVIDIA Corporation.

E. De Vito is a member of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM).

REFERENCES

- [1] M. Christ. A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral. *Colloq. Math.*, 60/61(2):601–628, 1990.
- [2] T. Coulhon, G. Kerkycharian, and P. P. Petrushev. Heat kernel generated frames in the setting of Dirichlet spaces. *J Fourier Anal Appl*, 18(5):995–1066, 2012.
- [3] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.
- [4] H. G. Feichtinger, H. Führ, and I. Z. Pesenson. Geometric space–frequency analysis on manifolds. *J Fourier Anal Appl*, 22(6):1294–1355, 2016.
- [5] M. Fornasier and H. Rauhut. Continuous frames, function spaces, and the discretization problem. *J Fourier Anal Appl*, 11(3):245–287, 2005.
- [6] D. Freeman and D. Speegle. The discretization problem for continuous frames. *Advances in Mathematics*, 345:784 – 813, 2019.
- [7] H. Führ, K. Gröchenig, A. Haimi, A. Klotz, and J. L. Romero. Density of sampling and interpolation in reproducing kernel Hilbert spaces. *J. Lond. Math. Soc. (2)*, 96(3):663–686, 2017.
- [8] F. Göbel, G. Blanchard, and U. von Luxburg. Construction of tight frames on graphs and application to denoising. In W. K. Härdle and X. Shen H.-S. Lu, editors, *Handbook of Big Data Analytics*, chapter 20, pages 503–522. Springer, 2018.
- [9] D. K. Hammond, P. Vandergheynst, and R. Gribonval. Wavelets on graphs via spectral graph theory. *Appl. Comput. Harmon. Anal.*, 30(2):129 – 150, 2011.
- [10] L. Rosasco, M. Belkin, and E. De Vito. On learning with integral operators. *J. Mach. Learn. Res.*, 11:905–934, 2010.