RESEARCH

Open Access

Random sampling and approximation of signals with bounded derivatives

Jianbin Yang[1]* and Xinzhu Tao[1]

*Correspondence:
jbyang@hhu.edu.cn
[1] Department of Mathematics,
College of Science, Hohai University,
Nanjing, China

Abstract

Approximation of analog signals from noisy samples is a fundamental, but nevertheless difficult problem. This paper addresses the problem of approximating functions in $H_{\gamma,\Omega}$ from randomly chosen samples, where

$$H_{\gamma,\Omega} = \left\{ f \mid f \text{ is continuous on } \overline{\Omega}, \text{and } \|Df\|_{L_\infty(\Omega)} \leq \gamma \|f\|_{L_\infty(\Omega)} \right\}.$$

We are concerned with the probability that functions in $H_{\gamma,\Omega}$ can be approximated from the noisy samples stably and how they can be approximated.

By calculating the upper bound of the covering number of a subset of $H_{\gamma,\Omega}$ and using the uniform law of large numbers, we conclude that functions in $H_{\gamma,\Omega}$ can be recovered stably with overwhelming probability provided that the sampling noise satisfies some mild conditions and the sampling size is sufficiently large. Furthermore, an ℓ_∞-regularized least squares model is proposed to approximate functions from noisy samples. The alternating direction method of multipliers (ADMM) algorithm is then applied to solve the model. In the end, numerical experiments are presented and discussed to illustrate the efficiency of our method.

MSC: 42C40; 60E15; 94A20; 62D05

Keywords: Sampling inequalities; Signals with bounded derivatives; Covering number; Regularized least squares

1 Introduction

In modern digital data processing of signals, such as sound, image or video processing, one always uses a discretized version of the original analog functions [1]. Then how the analog function can be recovered from its samples and whether the reconstruction is stable are fundamental problems in sampling theory [2]. For example, the Shannon–Whittaker sampling theorem says for each $f \in L_2(\mathbb{R})$ and $\text{supp}(\hat{f}) \subseteq [-\frac{1}{2}, \frac{1}{2}]$, it can be completely recovered from the sample points $\{f(j) : j \in \mathbb{Z}\}$ by the formula $f(x) = \sum_{j \in \mathbb{Z}} f(j) \frac{\sin \pi(x-j)}{\pi(x-j)}$, where the convergence is uniform on $\mathbb{R}$ and also in $L_2(\mathbb{R})$.

To characterize the conditions under which it is possible to recover particular classes of functions from the sampling points stably, we introduce the following sampling inequalities:

$$m_p \|f\|_{L_p(\mathbb{R}^d)}^p \leq \sum_{x_j \in X} |f(x_j)|^p \leq M_p \|f\|_{L_p(\mathbb{R}^d)}^p \quad \forall f \in V, \tag{1.1}$$

where m_p and M_p are positive constants independent of f and $1 \leq p < \infty$. The set $X = \{x_j : j \in J\}$ is said to be a stable set of sampling for function class $V \subseteq L_p(\mathbb{R}^d)$ if the inequalities (1.1) hold (see e.g. [2, 3]). The sampling inequalities (1.1) imply that a small perturbation of f causes only a small change of sampled values $\{f(x_j)\}$, and vice versa. This means that the sampling is a stable process and the reconstruction of functions in V from X is continuous. For the Shannon–Whittaker theorem, since the functions $\{\frac{\sin \pi (\cdot - j)}{\pi (\cdot - j)}\}_{j \in \mathbb{Z}}$ form an orthonormal basis for the space of bandlimited functions with highest cycle frequency $\frac{1}{2}$, we have $\|f\|^2_{L_2(\mathbb{R})} = \sum_{j \in \mathbb{Z}} |f(j)|^2$ and $m_2 = M_2 = 1$.

In the past years, there has been a considerable body of research on developing reconstruction algorithms and characterizing the conditions of stable set of sampling for various function classes. For instance, Aldroubi and Gröchenig [2] investigated the nonuniform sampling and reconstruction in shift-invariant spaces. Xian and Li [4, 5] studied the sampling set conditions and applications of weighted finitely generated shift-invariant spaces. Sun and Zhou [6], and Sun [7] characterized the local sampling in spline subspaces and shift-invariant spaces, respectively.

This paper addresses the problem of sampling and approximation of functions with bounded derivatives. We define the following class of functions on $\Omega \subset \mathbb{R}^d$:

$$H_{\gamma, \Omega} := \left\{ f \mid f \text{ is continuous on } \overline{\Omega}, \text{ and } \|Df\|_{L_\infty(\Omega)} \leq \gamma \|f\|_{L_\infty(\Omega)} \right\}, \tag{1.2}$$

where $|Df| = |D_{x_1} f| + |D_{x_2} f| + \cdots + |D_{x_d} f|$, and $D_{x_i} = \frac{\partial}{\partial x_i}$ is the weak derivative of order 1. The continuity of f at the boundary of Ω means that f can be continuously extended to $\overline{\Omega}$. The parameter γ characterizes the degree of oscillation of functions in $H_{\gamma, \Omega}$. For simplicity, throughout this paper, we always assume that $\Omega = (0, 1)^d$. Sampling in $H_{\gamma, \Omega}$ is an appropriate model for many applications, in particular in signal and image processing [1]. Moreover, by Bernstein's inequality [8], if $f \in L_2(\mathbb{R})$ and $\text{supp}(\hat{f}) \subseteq [-\frac{\gamma}{4\pi}, \frac{\gamma}{4\pi}]$, then $f \in H_{\gamma, \mathbb{R}}$. In other words, the space of bandlimited functions is a special case of $H_{\gamma, \Omega}$.

Random sampling approaches were used in a variety of fields, such as statistical learning theory [9], compressed sensing [10], image processing [11], and many others [12]. Recently, Bass and Gröchenig in [13] studied the stability problem of random sampling in the space of bandlimited functions. Yang et al. [14, 15] and Führ et al. [16] discussed the stability conditions and applications of random sampling in shift-invariant spaces.

The purpose of this paper is to investigate random sampling of functions in $H_{\gamma, \Omega}$. Let $\{(x_j, y_j)\}_{j=1}^n$ be the sampling of $f \in H_{\gamma, \Omega}$, where $\{x_j\}$ is uniformly drawn from Ω, and $y_j = f(x_j) + \epsilon_j$ with ϵ_j being a random noise. We consider the probability that $\{(x_j, y_j)\}_{j=1}^n$ is a stable set of sampling for $H_{\gamma, \Omega}$, and how f can be approximated from these samples. By estimating the upper bound of the capacity of $H_{\gamma, \Omega}$ and applying the uniform law of large numbers of the sampling values, we conclude that with overwhelming probability, the sampling inequalities (1.1) hold uniformly for all functions in $H_{\gamma, \Omega}$ when the sampling noise satisfies some mild conditions. Furthermore, an ℓ_∞-regularized least squares model is proposed, and the corresponding numerical algorithm is discussed.

The rest of this paper is structured as follows. Section 2 introduces some mathematical notation and the covering number of $H_{\gamma, \Omega}$. Section 3 characterizes the stability properties of random sampling in $H_{\gamma, \Omega}$. Section 4.1 presents an optimization model and the corresponding numerical algorithm to approximate $f \in H_{\gamma, \Omega}$ from its noisy samples. Section 4.2 discusses some numerical experiments. Finally, Sect. 5 concludes and discusses future directions.

2 Preliminaries

We first introduce some notation. Let $\mathbb{N}$ denote the set of positive integers. As usual, for $x \in \mathbb{R}$, let $\lfloor x \rfloor$ denote the largest integer smaller than or equal to x, and $\lceil x \rceil$ denote the smallest integer greater than or equal to x. For $x = (x_1, \ldots, x_d)$, $y = (y_1, \ldots, y_d) \in \mathbb{R}^d$, let $|x - y| := \max_{1 \le i \le d}\{|x_i - y_i|\}$.

For two sets A and B, the Cartesian product $A \times B$ is the set of all ordered pairs defined by $A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}$. Similarly, it can be generalized to an m-ary Cartesian product over m sets.

Let $\mathcal{M}$ be a Lebesgue measurable set. For $1 \le p < \infty$, we denote by $L_p(\mathcal{M})$ the Banach space of all functions such that

$$\|f\|_{L_p(\mathcal{M})} := \left(\int_{\mathcal{M}} |f(x)|^p \, dx \right)^{1/p} < \infty \quad \text{for } 1 \le p < \infty,$$

and $\|f\|_{L_\infty(\mathcal{M})}$ is the essential supremum of f on $\mathcal{M}$.

We can similarly define $\ell_p = \ell_p(\mathbb{Z}^d)$ the Banach space of all sequences $a = (a_k)_{k \in \mathbb{Z}^d}$ such that $\|a\|_{\ell_p} < \infty$, where

$$\|a\|_{\ell_p} := \left(\sum_{k \in \mathbb{Z}^d} |a_k|^p \right)^{1/p} \quad \text{for } 1 \le p < \infty,$$

and $\|a\|_{\ell_\infty}$ is the supremum of a on $\mathbb{Z}^d$.

The first main result of this paper is to estimate the probability such that the sampling inequalities (1.1) hold uniformly for all functions in $H_{\gamma,\Omega}$. Such similar problems are classical in statistical learning theory [9]. The most powerful tool used there is the capacity of the involved function set and the uniform law of large numbers [17–19]. Our analysis is following this strategy.

Since the covering number is a very convenient and powerful tool for metric space, we choose it to characterize the capacity of function sets. Let S be a metric space. For any $\eta > 0$, the covering number $\mathcal{N}(S, \eta)$ is the minimal number of balls with radius η that can cover S. When S is compact, $\mathcal{N}(S, \eta)$ is finite for any given $\eta > 0$.

It is not hard to verify that any $f \in H_{\gamma,\Omega}$ satisfies the inequalities (1.1) if and only if $f/\|f\|_{L_\infty(\mathbb{R}^d)}$ does. So we consider the subset

$$H_{\gamma,\Omega}^* := \left\{ f \in H_{\gamma,\Omega}, \text{ and } \|f\|_{L_\infty(\Omega)} = 1 \right\}. \tag{2.1}$$

The following proposition gives an upper bound for the covering number of $H_{\gamma,\Omega}^*$, and the proof follows the line of argument in [9, Proposition 5.4] where $d = 1$.

Proposition 2.1 *Let $H_{\gamma,\Omega}^*$ be defined by (2.1). For any $\eta > 0$, the covering number of $H_{\gamma,\Omega}^*$ with respect to $\| \cdot \|_{L_\infty}$*

$$\mathcal{N}\left(H_{\gamma,\Omega}^*, \eta \right) \le \exp\left(\ln 2 + \frac{2}{\eta} + \left(\frac{4\gamma}{\eta} \right)^d \ln 3 \right).$$

Proof We first define a set of regular grid points in Ω with equal distance $\delta = \frac{\eta}{4\gamma}$. Let $\mathrm{I} = \{1, 2, \ldots, \lfloor \frac{1}{\delta} \rfloor\}$ and $i := (i_1, i_2, \ldots, i_d) \in \mathrm{I}^d$. For each $i_\ell \in \mathrm{I}$, set $x_{i_\ell} = i_\ell \delta$. Then one can check

that the point set

$$X := \left\{ x_i = (x_{i_1}, x_{i_2}, \ldots, x_{i_d}) \mid x_{i_\ell} = i_\ell \delta, i \in I^d \right\}$$

is a regular grid in Ω with equal distance δ, which is also a δ-net of Ω.

For $f \in H^*_{\gamma,\Omega}$, we have $\|Df\|_{L_\infty} \leq \gamma \|f\|_{L_\infty} \leq \gamma$. Moreover, by the continuity of f, for any $x, y \in \Omega$, we have $-1 \leq f(x) \leq 1$ and $|f(x) - f(y)| \leq \gamma |x - y|$. It follows that

$$(v_i - 1)\frac{\eta}{2} \leq f(x_i) \leq v_i \frac{\eta}{2}, \quad \forall x_i \in X,$$

for some $v_i \in J := \{-\lceil \frac{2}{\eta} \rceil + 1, -\lceil \frac{2}{\eta} \rceil + 2, \ldots, \lceil \frac{2}{\eta} \rceil\}$. For each

$$v = \left(v_{(1,1,\ldots,1)}, v_{(2,1,\ldots,1)}, \ldots, v_i, \ldots, v_{(\lfloor \frac{1}{\delta} \rfloor, \lfloor \frac{1}{\delta} \rfloor, \ldots, \lfloor \frac{1}{\delta} \rfloor)} \right) \in J^{|I^d|},$$

define

$$H_v := \left\{ f \in H^*_{\gamma,\Omega} \mid (v_i - 1)\frac{\eta}{2} \leq f(x_i) \leq v_i \frac{\eta}{2}, \forall i \in I^d \right\}.$$

Then we obtain

$$H^*_{\gamma,\Omega} \subseteq \bigcup_{v \in J^{|I^d|}} H_v.$$

Besides, for every $f, g \in H_v$, there exist $\tilde{i} \in I^d$ and $x_{\tilde{i}} \in X$ such that

$$\max_{x \in \Omega} |f(x) - g(x)| = \max_{|x - x_{\tilde{i}}| \leq \delta} |f(x) - g(x)|$$

$$\leq \max_{|x - x_{\tilde{i}}| \leq \delta} |f(x) - f(x_{\tilde{i}})| + |g(x) - g(x_{\tilde{i}})| + |f(x_{\tilde{i}}) - g(x_{\tilde{i}})|$$

$$\leq 2\gamma\delta + \frac{\eta}{2} = \eta.$$

Hence, we conclude that the diameter of $H_v \subseteq H^*_{\gamma,\Omega}$ is at most η, and

$$\{H_v : v \in J^{|I^d|}\}$$

is an η-covering of $H^*_{\gamma,\Omega}$.

In the following, we count the number of sets $\{H_v\}$, $v \in J^{|I^d|}$. Let $f \in H_v$. For every fixed $i = (i_1, i_2, \ldots, i_\ell, \ldots, i_d) \in I^d$ and $x_i \in X$, we have

$$(v_i - 1)\frac{\eta}{2} \leq f(x_i) \leq v_i \frac{\eta}{2}.$$

Similarly, for $x_{\bar{i}} \in X$ with $\bar{i} = (i_1, i_2, \ldots, i_\ell + 1, \ldots, i_d) \in I^d$, a δ-neighboring point of x_i, we have

$$(v_{\bar{i}} - 1)\frac{\eta}{2} \leq f(x_{\bar{i}}) \leq v_{\bar{i}} \frac{\eta}{2}.$$

It follows that

$$(v_i - v_{\bar{i}} - 1)\frac{\eta}{2} \le f(x_i) - f(x_{\bar{i}}) \le (v_i - v_{\bar{i}} + 1)\frac{\eta}{2}. \tag{2.2}$$

Since $f \in H_v \subseteq H^*_{\gamma,\Omega}$ and $|x_i - x_{\bar{i}}| = \delta$, we have

$$\left| f(x_i) - f(x_{\bar{i}}) \right| \le \gamma\delta = \frac{\eta}{4}. \tag{2.3}$$

By (2.2) and (2.3), we obtain $|v_i - v_{\bar{i}}|\frac{\eta}{2} - \frac{\eta}{2} \le \frac{\eta}{4}$. Hence,

$$v_{\bar{i}} \in \{v_i, v_i + 1, v_i - 1\}.$$

In addition, since $v_{(1,1,\dots,1)}$ has at most $2\lceil\frac{2}{\eta}\rceil$ possible values, the number of nonempty sets $\{H_v\}$, $v \in \mathrm{J}^{|\mathrm{I}^d|}$ is at most

$$2\left\lceil\frac{2}{\eta}\right\rceil 3^{|\mathrm{I}^d|} \le 2\left(\frac{2}{\eta} + 1\right)3^{\lfloor\frac{1}{\delta}\rfloor^d}$$

$$\le \exp\left(\ln\left(2\left(\frac{2}{\eta} + 1\right)\right) + \left\lfloor\frac{1}{\delta}\right\rfloor^d \ln 3\right)$$

$$\le \exp\left(\ln\left(2\left(\frac{2}{\eta} + 1\right)\right) + \left\lfloor\frac{4\gamma}{\eta}\right\rfloor^d \ln 3\right).$$

Moreover, since $\ln(1 + t) \le t$ for $t \ge 0$, we have

$$\exp\left(\ln\left(2\left(\frac{2}{\eta} + 1\right)\right) + \left\lfloor\frac{4\gamma}{\eta}\right\rfloor^d \ln 3\right) \le \exp\left(\ln 2 + \frac{2}{\eta} + \left(\frac{4\gamma}{\eta}\right)^d \ln 3\right).$$

This completes the proof of Proposition 2.1. □

3 Stability of random sampling

Let $X = \{x_j : j \in \mathbb{N}\}$ be a sequence of independent and identically distributed (i.i.d.) random variables, each of which is uniformly drawn from Ω, and $y_j = f(x_j) + \epsilon_j$ with ϵ_j being a random noise. In this section, we investigate the probability that any $f \in H_{\gamma,\Omega}$ can be recovered from its samples $\{(x_j, y_j)\}_{j=1}^n$ stably. We will prove that if the random noise satisfies some mild conditions, then, with overwhelming probability, the sampling inequalities (1.1) hold uniformly for all functions in $H_{\gamma,\Omega}$ with (noisy) sampling values.

For every fixed $f \in H_{\gamma,\Omega}$, we define the random variable

$$X_j(f) := \left|f(x_j)\right|^p - \int_\Omega |f(x)|^p\,dx, \tag{3.1}$$

where $\{x_j\}$ is uniformly drawn from Ω. Then one can check that the sequence $\{X_j(f) : j \in \mathbb{N}\}$ is independent and the expectation $\mathbb{E}(X_j(f)) = 0$. Besides, for the variance of $X_j(f)$, we have

$$\mathrm{Var}\left(X_j(f)\right) = \mathbb{E}\left(X_j(f)^2\right) - \left(\mathbb{E}\left(X_j(f)\right)\right)^2$$

$$= \mathbb{E}\left(|f(x_j)|^{2p}\right) - \left(\int_\Omega |f(x)|^p\,dx\right)^2$$

$$\leq \mathbb{E}\big(\big|f(x_j)\big|^{2p}\big)$$

$$= \int_{\Omega} \big|f(x)\big|^{2p}\, dx$$

$$\leq \|f\|_{L_\infty(\Omega)}^{2p}.$$

The following Bernstein inequality plays an important role in probability theory, which gives bounds on the probability that the sum of independent random variables deviates from the expectation.

Lemma 3.1 ([20]) *Let $\xi_1, \xi_2, \ldots, \xi_n$ be independent random variables. Assume that $\mathbb{E}(\xi_j) = 0$, $\mathrm{Var}(\xi_j) \leq \sigma^2$ and $|\xi_j| \leq K$ almost surely for all j. Then, for any $\lambda \geq 0$,*

$$\mathrm{Prob}\left(\left|\frac{1}{n}\sum_{j=1}^{n}\xi_j\right| \geq \lambda\right) \leq 2\exp\left(-\frac{n\lambda^2}{2\sigma^2 + \frac{2}{3}K\lambda}\right).$$

Lemma 3.2 *Let $\{x_j : j = 1, 2, \ldots, n\}$ be a sequence of i.i.d. random variables that are uniformly drawn from $\Omega = (0,1)^d$. Let $H^*_{\gamma,\Omega}$ be defined by (2.1), and $X_j(f)$ be given by (3.1). Then, for any $\lambda \geq 0$ and $n \in \mathbb{N}$,*

$$\mathrm{Prob}\left(\sup_{f \in H^*_{\gamma,\Omega}}\left|\frac{1}{n}\sum_{j=1}^{n}X_j(f)\right| \geq \lambda\right) \leq 2\mathcal{N}\left(H^*_{\gamma,\Omega}, \frac{\lambda}{2p}\right)\exp\left(-\frac{3n\lambda^2}{24 + 4\lambda}\right).$$

Proof Let $\{f_\ell\}_{\ell=1,\ldots,L}$, where $L = \mathcal{N}(H^*_{\gamma,\Omega}, \frac{\lambda}{2p})$, be a sequence in $H^*_{\gamma,\Omega}$ such that $H^*_{\gamma,\Omega}$ can be covered by the L_∞ balls centered at f_ℓ with radius $\frac{\lambda}{2p}$. For each fixed $f_\ell \in H^*_{\gamma,\Omega}$, since $\|f_\ell\|_{L_\infty(\Omega)} = 1$, we have $\mathrm{Var}(X_j(f_\ell)) \leq 1$ and $|X_j(f_\ell)| \leq 1$. By Lemma 3.1, we obtain

$$\mathrm{Prob}\left(\left|\frac{1}{n}\sum_{j=1}^{n}X_j(f_\ell)\right| \geq \lambda\right) \leq 2\exp\left(-\frac{n\lambda^2}{2 + \frac{2}{3}\lambda}\right). \tag{3.2}$$

For any given $f \in H^*_{\gamma,\Omega}$, there exists some $\ell \in \{1, 2, \ldots, L\}$ such that $\|f - f_\ell\|_{L_\infty} \leq \frac{\lambda}{2p}$. Thus, by the mean value theorem,

$$\left|\frac{1}{n}\sum_{j=1}^{n}X_j(f) - \frac{1}{n}\sum_{j=1}^{n}X_j(f_\ell)\right|$$

$$= \left|\frac{1}{n}\sum_{j=1}^{n}\big(\big|f(x_j)\big|^p - \big|f_\ell(x_j)\big|^p\big)\right|$$

$$\leq p\big(\max\big\{\|f\|_{L_\infty(\Omega)}, \|f_\ell\|_{L_\infty(\Omega)}\big\}\big)^{p-1}\|f - f_\ell\|_{L_\infty(\Omega)}$$

$$\leq p\|f - f_\ell\|_{L_\infty(\Omega)}$$

$$\leq \frac{\lambda}{2}.$$

Combining this with (3.2), we conclude that, for each fixed ℓ, $1 \leq \ell \leq L$,

$$\text{Prob}\left\{ \sup_{\{f:\|f-f_\ell\|_{L_\infty} \leq \frac{\lambda}{2p}\}} \left| \frac{1}{n} \sum_{j=1}^{n} X_j(f) \right| \geq \lambda \right\}$$

$$\leq \text{Prob}\left\{ \left| \frac{1}{n} \sum_{j=1}^{n} X_j(f_\ell) \right| \geq \lambda - \frac{\lambda}{2} \right\}$$

$$\leq 2\exp\left(-\frac{3n\lambda^2}{24 + 4\lambda} \right).$$

Besides, since

$$H_{\gamma,\Omega}^* \subseteq \bigcup_{\ell=1}^{L} \left\{ f : \|f - f_\ell\|_{L_\infty} \leq \frac{\lambda}{2p} \right\},$$

we obtain

$$\text{Prob}\left(\sup_{f \in H_{\gamma,\Omega}^*} \left| \frac{1}{n} \sum_{j=1}^{n} X_j(f) \right| \geq \lambda \right) \leq \sum_{\ell=1}^{L} \text{Prob}\left\{ \sup_{\{f:\|f-f_\ell\|_{L_\infty} \leq \frac{\lambda}{2p}\}} \left| \frac{1}{n} \sum_{j=1}^{n} X_j(f) \right| \geq \lambda \right\}.$$

Therefore, noting that $L = \mathcal{N}(H_{\gamma,\Omega}^*, \frac{\lambda}{2p})$, we conclude that

$$\text{Prob}\left(\sup_{f \in H_{\gamma,\Omega}^*} \left| \frac{1}{n} \sum_{j=1}^{n} X_j(f) \right| \geq \lambda \right) \leq 2\mathcal{N}\left(H_{\gamma,\Omega}^*, \frac{\lambda}{2p} \right) \exp\left(-\frac{3n\lambda^2}{24 + 4\lambda} \right). \qquad \square$$

Theorem 3.3 *Let $H_{\gamma,\Omega}$ be defined by* (1.2). *Assume that $\{x_j : j = 1, 2, \ldots, n\}$ is a sequence of i.i.d. random variables that are uniformly drawn from Ω. Then, for any $0 < \lambda < \frac{1}{(p+1)^d \gamma^d}$, the following sampling inequalities:*

$$n\big(1 - (p+1)^d \gamma^d \lambda\big) \int_\Omega |f(x)|^p \, dx \leq \sum_{j=1}^{n} |f(x_j)|^p \leq n\big(1 + (p+1)^d \gamma^d \lambda\big) \int_\Omega |f(x)|^p \, dx \quad (3.3)$$

hold uniformly for all functions in $H_{\gamma,\Omega}$ with probability at least

$$1 - 2\exp\left(\ln 2 + \frac{4p}{\lambda} + \left(\frac{8p\gamma}{\lambda} \right)^d \ln 3 - \frac{3n\lambda^2}{24 + 4\lambda} \right).$$

Proof Obviously that every $f \in H_{\gamma,\Omega}$ satisfies (3.3) if and only if $f / \|f\|_{L_\infty}$ does. Thus, we assume that $\|f\|_{L_\infty} = 1$ and $f \in H_{\gamma,\Omega}^*$.

Let $X_j(f)$ be defined by (3.1), and one can check that the event

$$\mathcal{E} = \left\{ \sup_{f \in H_{\gamma,\Omega}^*} \left| \frac{1}{n} \sum_{j=1}^{n} X_j(f) \right| \geq \lambda \right\}$$

is the complement of

$$\tilde{\mathcal{E}} = \left\{ n \int_\Omega |f(x)|^p \, dx - \lambda n \leq \sum_{j=1}^{n} |f(x_j)|^p \leq n \int_\Omega |f(x)|^p \, dx + \lambda n, \forall f \in H_{\gamma,\Omega}^* \right\}.$$

For an arbitrary $f \in H^*_{\gamma,\Omega}$, there exists an $x^\star \in \overline{\Omega}$ such that $|f(x^\star)| = \|f\|_{L_\infty} = 1$. Without loss of generality, we assume that $f(x^\star) = 1$. Then, for $x \in \Omega^*$, where $\Omega^* := \{x \in \Omega, |x - x^\star| \leq \frac{1}{\gamma}\}$, we have

$$f(x^\star) - f(x) = |f(x) - f(x^\star)| \leq \gamma |x - x^\star| \leq 1.$$

It follows that

$$f(x) \geq 1 - \gamma |x - x^\star| \geq 0 \quad \text{for } x \in \Omega^*$$

and

$$\int_\Omega |f(x)|^p \, dx \geq \int_{\Omega^*} (1 - \gamma |x - x^\star|)^p \, dx \geq \frac{1}{(p+1)^d \gamma^d}.$$

Therefore, the event

$$\bar{\mathcal{E}} = \left\{ n\big(1 - (p+1)^d \gamma^d \lambda\big) \int_\Omega |f(x)|^p \, dx \leq \sum_{j=1}^n |f(x_j)|^p \leq n\big(1 + (p+1)^d \gamma^d \lambda\big) \int_\Omega |f(x)|^p \, dx, \right.$$

$$\left. \forall f \in H^*_{\gamma,\Omega} \right\}$$

contains the event $\tilde{\mathcal{E}}$.

Thus, by Lemma 3.2 and Proposition 2.1, the inequalities (3.3) hold uniformly for all functions in $H_{\gamma,\Omega}$ with probability

$$\text{Prob}(\bar{\mathcal{E}}) \geq \text{Prob}(\tilde{\mathcal{E}}) = 1 - \text{Prob}(\mathcal{E})$$

$$\geq 1 - 2\exp\left(\ln 2 + \frac{4p}{\lambda} + \left(\frac{8p\gamma}{\lambda}\right)^d \ln 3 - \frac{3n\lambda^2}{24 + 4\lambda} \right). \qquad \square$$

Corollary 3.4 *Under the same conditions of Theorem 3.3, let $y_j = f(x_j) + \epsilon_j$, $j = 1, 2, \ldots, n$ be the sampling of f. Suppose that the random noise $\{\epsilon_j\}$ are independent with $E(|\epsilon_j|^p) = \sigma^p$ and $\big||\epsilon_j|^p - \sigma^p\big| \leq M\|f\|^p_{L_p}$ for all ϵ_j. In addition, we assume that $\frac{\sigma^p}{\|f\|^p_{L_p}} \leq \rho \ll 1$. Then, for any $0 < \lambda < \frac{2^{1-p} - \rho}{2^{1-p}(p+1)^d \gamma^d + 1}$, the inequalities*

$$n\big(2^{1-p} - \rho - \big(2^{1-p}(p+1)^d \gamma^d + 1\big)\lambda\big) \int_\Omega |f(x)|^p \, dx$$

$$\leq \sum_{j=1}^n |f(x_j) + \epsilon_j|^p$$

$$\leq 2^{p-1} n\big(1 + \rho + \big((p+1)^d \gamma^d + 1\big)\lambda\big) \int_\Omega |f(x)|^p \, dx \qquad (3.4)$$

hold uniformly for all functions in $H_{\gamma,\Omega}$ with probability at least

$$\left(1 - 2\exp\left(-\frac{n\lambda^2}{2M^2}\right)\right)\left(1 - 2\exp\left(\ln 2 + \frac{4p}{\lambda} + \left(\frac{8p\gamma}{\lambda}\right)^d \ln 3 - \frac{3n\lambda^2}{24 + 4\lambda}\right)\right).$$

Proof One can check that every $f \in H_{\gamma,\Omega}$ satisfies the inequalities of (3.4) if and only if $f / \|f\|_{L_p}$ does. Thus, we assume that $\|f\|_{L_p} = 1$. By Hoeffding's inequality [21], we have

$$\text{Prob}\left\{ \left| \frac{1}{n} \sum_{j=1}^{n} |\epsilon_j|^p - \sigma^p \right| \geq \lambda \right\} \leq 2 \exp\left(-\frac{n\lambda^2}{2M^2} \right).$$

So, with probability $1 - 2\exp(-\frac{n\lambda^2}{2M^2})$,

$$\frac{1}{n} \sum_{j=1}^{n} |\epsilon_j|^p \leq \sigma^p + \lambda \|f\|_{L_p}^p.$$

For $1 \leq p < \infty$, since t^p is a convex function of t on $[0, +\infty)$, by Jensen's inequality, we have

$$\left| f(x_j) + \epsilon_j \right|^p \leq \left(|f(x_j)| + |\epsilon_j| \right)^p \leq 2^{p-1} \left(|f(x_j)|^p + |\epsilon_j|^p \right)$$

and

$$\left| f(x_j) + \epsilon_j \right|^p \geq \left(|f(x_j)| - |\epsilon_j| \right)^p \geq 2^{1-p} |f(x_j)|^p - |\epsilon_j|^p.$$

Hence, with the same probability, we have

$$\sum_{j=1}^{n} \left| f(x_j) + \epsilon_j \right|^p \leq 2^{p-1} \sum_{j=1}^{n} \left| f(x_j) \right|^p + 2^{p-1} n \left(\sigma^p + \lambda \|f\|_{L_p}^p \right)$$

and

$$\sum_{j=1}^{n} \left| f(x_j) + \epsilon_j \right|^p \geq 2^{1-p} \sum_{j=1}^{n} \left| f(x_j) \right|^p - n \left(\sigma^p + \lambda \|f\|_{L_p}^p \right).$$

Combining this with Theorem 3.3, we conclude that

$$n \left(2^{1-p} - \rho - \left(2^{1-p}(p+1)^d \gamma^d + 1 \right) \lambda \right) \int_{\Omega} |f(x)|^p \, dx$$

$$\leq \sum_{j=1}^{n} \left| f(x_j) + \epsilon_j \right|^p$$

$$\leq 2^{p-1} n \left(1 + \rho + \left((p+1)^d \gamma^d + 1 \right) \lambda \right) \int_{\Omega} |f(x)|^p \, dx$$

holds with probability at least

$$\left(1 - 2\exp\left(-\frac{n\lambda^2}{2M^2} \right) \right) \left(1 - 2\exp\left(\ln 2 + \frac{4p}{\lambda} + \left(\frac{8p\gamma}{\lambda} \right)^d \ln 3 - \frac{3n\lambda^2}{24 + 4\lambda} \right) \right). \qquad \square$$

We remark that ρ in Corollary 3.4 is connected with the signal-to-noise ratio (SNR) of f, and the signal can be recovered from its noisy samples stably only if the noise level is relatively small.

4 Approximation algorithm and numerical examples

4.1 Approximation model and algorithm

In this subsection, we consider how to approximate $f \in H_{\gamma,\Omega}$ from its noisy samples $\{(x_j, y_j)\}_{j=1}^n$, where $y_j = f(x_j) + \epsilon_j$ and ϵ_j is a random noise. The main idea is that we seek an approximant f^* by solving the following optimization problem:

$$\min_{g \in V} \sum_{i=1}^n \left(g(x_i) - y_i\right)^2 + \Gamma(g), \tag{4.1}$$

where the first term tries to fit $f^*(x_i)$ to y_i, and the second term is a regularization term.

For the function space V in (4.1), there are several choices, such as the Sobolev space, reproducing kernel Hilbert space, polynomial space or a principal shift-invariant space $S^h(\phi, \Omega)$, which is defined by

$$S^h(\phi, \Omega) = \left\{ \sum_{\alpha \in I} \mathbf{u}(\alpha)\phi\left(\frac{\cdot}{h} - \alpha\right) : \mathbf{u}(\alpha) \in \mathbb{R} \right\},$$

with

$$I = \left\{ \alpha \in \mathbb{Z}^d : \operatorname{supp} \phi\left(\frac{\cdot}{h} - \alpha\right) \cap \Omega \neq \emptyset \right\}.$$

Here, function ϕ is called the generator of $S^h(\phi, \Omega)$, and $h > 0$ is a scaling parameter that controls the refinement of the space. The compact support of ϕ is preferred since it generates sparse matrices. In addition, $S^h(\phi, \Omega)$ provides good approximation to smooth functions if ϕ satisfies certain conditions (see e.g. [22–24]).

The regularization term in (4.1) is chosen such that the derivative of f^* is controlled and the sampling noise can be reduced. For $g = \sum_{\alpha \in I} \mathbf{u}(\alpha)\phi(\frac{\cdot}{h} - \alpha)$, we take $\Gamma(g) = \|\operatorname{diag}(\boldsymbol{\lambda})\mathcal{W}\mathbf{u}\|_{\ell_\infty}$, where $\mathcal{W}$ is the discrete wavelet frame transform, and $\operatorname{diag}(\boldsymbol{\lambda})$ is a diagonal matrix based on the vector $\boldsymbol{\lambda}$ which scales different wavelet channels. The advantage of using wavelet frames here is that the model has fast algorithms, and wavelet frames can be regarded as certain discretizations to the general differential operator [25]. Moreover, with properly chosen parameters $\operatorname{diag}(\boldsymbol{\lambda})$, we have $\|Dg\|_{L_\infty(\Omega)} \sim \|\operatorname{diag}(\boldsymbol{\lambda})\mathcal{W}\mathbf{u}\|_{\ell_\infty}$ [26].

In summary, we determine the approximating function by minimizing

$$\sum_{j=1}^n \left(\sum_{\alpha \in I} \mathbf{u}(\alpha)\phi\left(\frac{x_j}{h} - \alpha\right) - y_j \right)^2 + \left\|\operatorname{diag}(\boldsymbol{\lambda})\mathcal{W}\mathbf{u}\right\|_{\ell_\infty}, \tag{4.2}$$

where $\mathbf{u}$ are the coefficients to solve. Finally, let

$$f^* = \sum_{\alpha \in I} \mathbf{u}^*(\alpha)\phi\left(\frac{\cdot}{h} - \alpha\right)$$

with $\mathbf{u}^*$ being the minimizer of (4.2).

Next, we show how to numerically solve (4.2), which can be written in the following matrix-vector form:

$$\min_{\mathbf{u} \in \mathbb{R}^m} \|\mathbf{A}\mathbf{u} - \mathbf{y}\|_{\ell_2}^2 + \left\|\operatorname{diag}(\boldsymbol{\lambda})\mathcal{W}\mathbf{u}\right\|_{\ell_\infty}, \tag{4.3}$$

where $\mathbf{y} = [y_1, y_2, \ldots, y_n]^T$, $\mathbf{A}_{jk} = \phi(x_j/h - \alpha_k)$, $X = \{x_1, \ldots, x_n\}$ and $I = \{\alpha_1, \ldots, \alpha_m\}$. This is equivalent to

$$\min_{\mathbf{u},\mathbf{d}} \|\mathbf{A}\mathbf{u} - \mathbf{f}\|_{\ell_2}^2 + \left\|\mathrm{diag}(\boldsymbol{\lambda})\mathbf{d}\right\|_{\ell_\infty} \quad \text{subject to } \mathbf{d} = \mathcal{W}\mathbf{u}. \tag{4.4}$$

Then the alternating direction method of multipliers (ADMM) method [12, 25] can be applied to solve (4.4) as follows:

$$\begin{cases} \mathbf{u}^{i+1} = \arg\min_{\mathbf{u}} \|\mathbf{A}\mathbf{u} - \mathbf{f}\|_{\ell_2}^2 + \frac{\mu}{2}\|\mathcal{W}\mathbf{u} - \mathbf{d}^i + \mathbf{b}^i\|_{\ell_2}^2, & (4.5) \\[2mm] \mathbf{d}^{i+1} = \arg\min_{\mathbf{d}} \|\mathrm{diag}(\boldsymbol{\lambda})\mathbf{d}\|_{\ell_\infty} + \frac{\mu}{2}\|\mathbf{d} - \mathcal{W}\mathbf{u}^{i+1} - \mathbf{b}^i\|_{\ell_2}^2, & (4.6) \\[2mm] \mathbf{b}^{i+1} = \mathbf{b}^i + \mathcal{W}\mathbf{u}^{i+1} - \mathbf{d}^{i+1}, & (4.7) \end{cases}$$

with initial guesses $\mathbf{u}^0$, $\mathbf{d}^0$, $\mathbf{b}^0$ and a parameter μ.

The quadratic problem (4.5) has the first order optimality condition

$$\left(2\mathbf{A}^T\mathbf{A} + \mu\mathbf{I}\right)\mathbf{u} = 2\mathbf{A}^T\mathbf{f} + \mu\mathcal{W}^T\left(\mathbf{d}^i - \mathbf{b}^i\right). \tag{4.8}$$

This is a sparse and positive definite linear system, and we use the conjugate gradient (CG) method to solve the problem. The solution to (4.6) is equivalent to the proximal operator of $\|\cdot\|_{\ell_\infty}$, i.e.,

$$\mathbf{prox}_{\frac{\lambda}{\mu}\|\cdot\|_{\ell_\infty}}\left(\mathcal{W}\mathbf{u}^{i+1} + \mathbf{b}^i\right) = \arg\min_{\mathbf{d}} \|\mathbf{d}\|_{\ell_\infty} + \frac{\mu}{2\lambda}\left\|\mathbf{d} - \left(\mathcal{W}\mathbf{u}^{i+1} + \mathbf{b}^i\right)\right\|_{\ell_2}^2.$$

By the property of the Moreau decomposition [12],

$$\mathbf{prox}_{\frac{\lambda}{\mu}\|\cdot\|_{\ell_\infty}}\left(\mathcal{W}\mathbf{u}^{i+1} + \mathbf{b}^i\right) = \mathcal{W}\mathbf{u}^{i+1} + \mathbf{b}^i - \mathbf{proj}_{B_1(0,\frac{\lambda}{\mu})}\left(\mathcal{W}\mathbf{u}^{i+1} + \mathbf{b}^i\right),$$

where $\mathbf{proj}_{B_1(0,\frac{\lambda}{\mu})}$ is the projection to the ℓ_1-norm ball given by

$$\mathbf{proj}_{B_1(0,\frac{\lambda}{\mu})}(\mathbf{t}) = \arg\min_{\|\mathbf{v}\|_{\ell_1} \leq \frac{\lambda}{\mu}} \|\mathbf{v} - \mathbf{t}\|_{\ell_2}^2. \tag{4.9}$$

The exact solution of (4.9) can be computed at most $O(N)$ time [27], where N is the dimension of the space.

Note that the numerical computation of $\mathcal{W}^T(\mathbf{d}^i - \mathbf{b}^i)$ in (4.8) is done by fast wavelet algorithm, similar to (4.6) and (4.7); see [25, 28] for detailed discussions. Finally, the proposed algorithm is summarized in Algorithm 1.

4.2 Numerical examples and discussions

In this subsection, we show the efficiency of the proposed approach (4.2) by approximating two functions in $H_{\gamma,\Omega}$ from some noisy samples.

In the first example, we test to approximate the well-known Franke function [29], which is defined as follows:

$$\mathrm{franke}(x_1, x_2) = \frac{3}{4}e^{-((9x_1-2)^2+(9x_2-2)^2)/4} + \frac{3}{4}e^{-((9x_1+1)^2)/49-(9x_2+1)/10}$$
$$+ \frac{1}{2}e^{-((9x_1-7)^2+(9x_2-3)^2)/4} - \frac{1}{5}e^{-(9x_1-4)^2-(9x_2-7)^2}.$$

Algorithm 1 ℓ_∞-regularized least squares algorithm

Initialize: Set initial guesses $\mathbf{d}^0$, $\mathbf{u}^0$ and $\mathbf{b}^0$. Choose an appropriate set of parameters μ and λ. Given a tolerance $\varepsilon > 0$. Set $i = 0$.

while $\|\mathbf{d}^i - \mathcal{W}\mathbf{u}^i\|_{\ell_2} \geq \epsilon$ **do**

$\quad \mathbf{u}^{i+1} = (2\mathbf{A}^T\mathbf{A} + \mu\mathbf{I})^{-1}(2\mathbf{A}^T\mathbf{f} + \mu\mathcal{W}^T(\mathbf{d}^i - \mathbf{b}^i))$

$\quad \mathbf{d}^{i+1} = \mathcal{W}\mathbf{u}^{i+1} + \mathbf{b}^i - \mathbf{proj}_{B_1(0,\frac{\lambda}{\mu})}(\mathcal{W}\mathbf{u}^{i+1} + \mathbf{b}^i)$

$\quad \mathbf{b}^{i+1} = \mathbf{b}^i + \mathcal{W}\mathbf{u}^{i+1} - \mathbf{d}^{i+1}$

$\quad i = i + 1$

end while

(a) Approximation by the proposed approach   (b) Original function $\mathrm{franke}(x_1, x_2)$

Figure 1 Approximation of 1000 samples from the Franke function
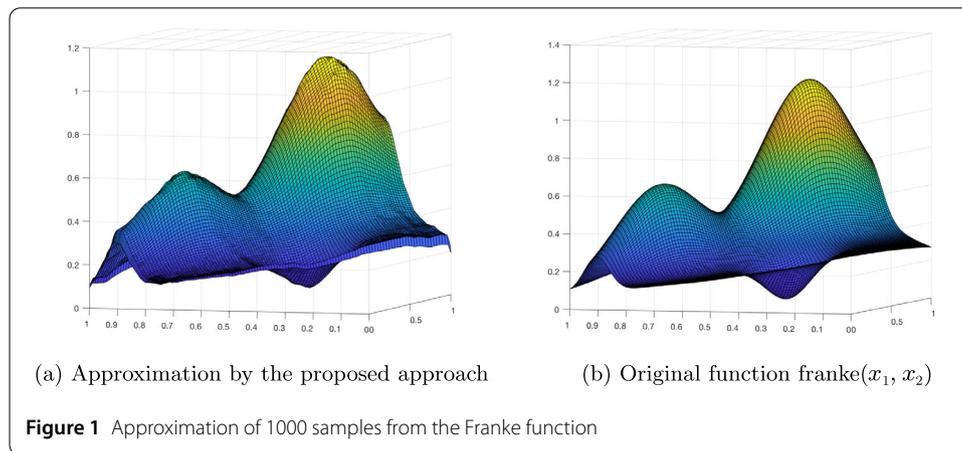
We randomly sample $\{\mathbf{x}_j\}_{j=1}^{1000}$ from $\Omega = (0,1)^2$, and take the sampling values $y_j = \mathrm{franke}(\mathbf{x}_j) + \epsilon_j$, where ϵ_j is a random variable drawn from the normal distribution $N(0, 0.01)$. Set the scaling parameter $h = 1/180$, and the tensor product of the cubic spline

$$
B_4(x) = \begin{cases}
x^3/6 & \text{if } 0 \leq x < 1, \\
(-3x^3 + 12x^2 - 12x + 4)/6 & \text{if } 1 \leq x < 2, \\
(3x^3 - 24x^2 + 60x - 44)/6 & \text{if } 2 \leq x < 3, \\
(4 - x)^3/6 & \text{if } 3 \leq x < 4, \\
0 & \text{else,}
\end{cases}
$$

as generator ϕ, with its associated wavelet tight frame masks $h_1 = [1/16, -1/4, 3/8, -1/4, 1/16]$, $h_2 = [-1/8, 1/4, 0, -1/4, 1/8]$, $h_3 = [\sqrt{6}/16, 0, -\sqrt{6}/8, 0, \sqrt{6}/16]$ and $h_4 = [-1/8, -1/4, 0, 1/4, 1/8]$.

Figure 1(a) illustrates the approximation result obtained by Algorithm 1, whereas Fig. 1(b) shows the original Franke function. It can be seen that the proposed model (4.2) is able to approximate function well.

In the second example, we test to approximate

$$
f(x_1, x_2) = \sin(2\pi x_1)^2 \cos(2\pi x_2). \tag{4.10}
$$

Using a similar method to the first example, we randomly sample $\{\mathbf{x}_j\}_{j=1}^{1000}$ from $\Omega = (0,1)^2$ and take the sampling values $y_j = f(\mathbf{x}_j) + \epsilon_j$ with ϵ_j drawn from $N(0, 0.01)$. We use the

(a) Approximation by the proposed approach   (b) Original function $f(x_1, x_2)$

Figure 2 Approximation of 1000 samples from $f(x_1, x_2)$ (4.10)
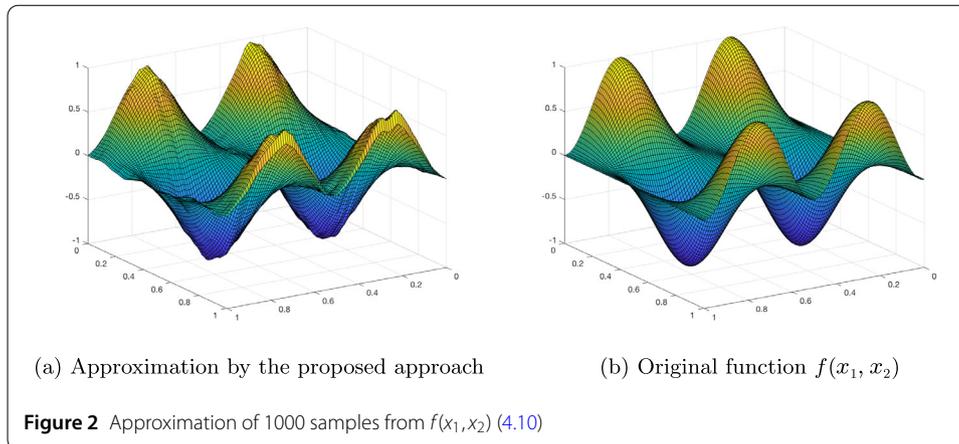
same generator ϕ and parameter h. In Fig. 2(a) the approximation is depicted, whereas in Fig. 2(b) the original function $f(x_1, x_2)$ is depicted.

5 Conclusion and future work

In this paper, we investigated random sampling of functions with bounded derivatives. We considered the probability that functions in $H_{\gamma,\Omega}$ can be recovered from noisy samples stably and how they can be approximated.

For the stability problem, we first estimated the capacity of $H_{\gamma,\Omega}$. By using the uniform law of large numbers, we concluded that functions in $H_{\gamma,\Omega}$ can be recovered stably with overwhelming probability when the sampling noise satisfies some mild conditions. Then we proposed an ℓ_∞-regularized least squares model in order to control the fluctuation of functions and suppress noise. The ADMM algorithm was applied to solve the optimization model. Finally, experiments on some function approximation tasks indicate the efficiency of the proposed approach.

In future work, it is of interest to present an error analysis for different kinds of denoising schemes when they are applied to reconstructing functions from random sampling.

Availability of data and materials
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors contributed equally to the manuscript and approved the final manuscript.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Lapidoth, A.: A Foundation in Digital Communication, 2nd edn. Cambridge University Press, Cambridge (2017)
2. Aldroubi, A., Gröchenig, K.: Nonuniform sampling and reconstruction in shift-invariant spaces. SIAM Rev. **43**(4), 585–620 (2001)
3. Xian, J.: Average sampling and reconstruction in a reproducing kernel subspace of homogeneous type space. Math. Nachr. **87**(8–9), 1042–1056 (2014)
4. Xian, J., Li, S.: Sampling set conditions in weighted finitely generated shift-invariant spaces and their applications. Appl. Comput. Harmon. Anal. **23**(2), 171–180 (2007)
5. Xian, J., Li, S.: Improved sampling and reconstruction in spline subspaces. Acta Math. Appl. Sin. Engl. Ser. **32**(2), 447–460 (2016)
6. Sun, W., Zhou, X.: Characterization of local sampling sequences for spline subspaces. Adv. Comput. Math. **30**(2), 153–175 (2009)
7. Sun, Q.: Local reconstruction for sampling in shift-invariant spaces. Adv. Comput. Math. **32**(3), 335–352 (2010)
8. Pinsky, M.A.: Introduction to Fourier Analysis and Wavelets, vol. 102. Am. Math. Soc., Providence (2008)
9. Cucker, F., Zhou, D.-X.: Learning Theory: An Approximation Theory Viewpoint. Cambridge University Press, Cambridge (2007)
10. Candès, E.J.: Compressive sampling. In: Proceedings of the International Congress of Mathematicians, vol. 3, pp. 1433–1452 (2006)
11. Piazzo, L.: Image estimation in the presence of irregular sampling, noise, and pointing jitter. IEEE Trans. Image Process. **28**(2), 713–722 (2019)
12. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends Mach. Learn. **3**(1), 1–122 (2011)
13. Bass, R.F., Gröchenig, K.: Random sampling of bandlimited functions. Isr. J. Math. **177**(1), 1–28 (2010)
14. Yang, J., Wei, W.: Random sampling in shift-invariant spaces. J. Math. Anal. Appl. **398**(1), 26–34 (2013)
15. Yang, J.: Random sampling and reconstruction in multiply generated shift-invariant spaces. Anal. Appl. **17**(2), 323–347 (2019)
16. Führ, H., Xian, J.: Relevant sampling in finitely generated shift-invariant spaces. J. Approx. Theory **240**, 1–15 (2019)
17. Wu, Q., Ying, Y., Zhou, D.-X.: Learning rates of least-square regularized regression. Found. Comput. Math. **6**(2), 171–192 (2006)
18. Cai, J.F., Shen, Z., Ye, G.B.: Approximation of frame based missing data recovery. Appl. Comput. Harmon. Anal. **31**(2), 185–204 (2011)
19. Lin, S., Guo, X., Zhou, D.-X.: Distributed learning with regularized least squares. J. Mach. Learn. Res. **18**(1), 3202–3232 (2017)
20. Bennett, G.: Probability inequalities for the sum of independent random variable. J. Am. Stat. Assoc. **57**(297), 33–45 (1962)
21. Hoeffding, W.: Probability inequalities for sums of bounded random variables. J. Am. Stat. Assoc. **58**(301), 13–30 (1963)
22. Jia, R.Q.: Approximation with scaled shift-invariant spaces by means of quasi-projection operators. J. Approx. Theory **131**(1), 30–46 (2004)
23. Dong, B., Shen, Z.: Pseudo-splines, wavelets and framelets. Appl. Comput. Harmon. Anal. **22**(1), 78–104 (2007)
24. Johnson, M.J., Shen, Z., Xu, Y.H.: Scattered data reconstruction by regularization in B-spline and associated wavelet spaces. J. Approx. Theory **159**(2), 197–223 (2009)
25. Cai, J.F., Osher, S., Shen, Z.: Split Bregman methods and frame based image restoration. Multiscale Model. Simul. **8**(2), 337–369 (2009)
26. Yang, J., Stahl, D., Shen, Z.: An analysis of wavelet frame based scattered data reconstruction. Appl. Comput. Harmon. Anal. **42**(3), 480–507 (2017)
27. Duchi, J., Shalev-Shwartz, S., Singer, Y., Chandra, T.: Efficient projections onto the ℓ_1-ball for learning in high dimensions. In: Proceedings of the 25th International Conference on Machine Learning, pp. 272–279. ACM, New York (2008)
28. Dong, B., Shen, Z.: MRA Based Wavelet Frames and Applications. IAS Lecture Notes Series. Summer Program on The Mathematics of Image Processing, Park City Mathematics Institute (2010)
29. Franke, R.: A critical comparison of some methods for interpolation of scattered data. Naval Postgraduate School Monterey Ca (1979)