



Non-uniform Random Sampling and Reconstruction in Signal Spaces with Finite Rate of Innovation

Yancheng Lu¹ · Jun Xian^{1,2}

Received: 2 April 2019 / Accepted: 26 October 2019 / Published online: 6 November 2019
© Springer Nature B.V. 2019

Abstract We consider non-uniform random sampling in a signal space with finite rate of innovation $V^2(\Lambda, \Phi) \subset L^2(\mathbb{R}^d)$ generated by a series of functions $\Phi = (\phi_\lambda)_{\lambda \in \Lambda}$. A subset $V_{R,\delta}^2(\Lambda, \Phi)$ of $V^2(\Lambda, \Phi)$ is consisting of functions concentrates at least $1 - \delta$ of the whole energy in a cube with side lengths R . Under mild assumptions on the generators and the probability distribution, we show that for R sufficiently large, taking $O(R^d \log(R^d))$ many samples with such the non-uniform distribution yields a sampling set for $V_{R,\delta}^2(\Lambda, \Phi)$ with high probability. We impose compact support on the generators as an additional constraint for obtaining a reconstruction algorithm from non-uniform random sampling with high probability.

Keywords Random sampling · Non-uniform sampling · Spaces with finite rate of innovation · Non-uniform distribution · Reconstruction algorithm

Mathematics Subject Classification (2000) 94A20 · 42C15 · 60E15 · 62M30

1 Introduction

The space of signals with finite degree of freedom per unit of time is called the space with *finite rate of innovation (FRI)* [2, 5, 12, 15, 17, 19, 23]. The concept is introduced by Vetterli et al. [23] at first. The FRI model is ubiquitous and has wide scientific applications such as radar imaging [15], compression of electrocardiogram signals [2], curve fitting [12]. The model of signals with finite rate of innovation can cover following cases: (i) band-limited signals [3], (ii) signals in some shift-invariant spaces [1, 7–9, 24, 25], (iii) non-uniform

✉ J. Xian
xianjun@mail.sysu.edu.cn

Y. Lu
luych5@mail2.sysu.edu.cn

¹ Department of Mathematics, Sun Yat-sen University, 510275 Guangzhou, China

² Guangdong Province Key Laboratory of Computational Science, Sun Yat-sen University, 510275 Guangzhou, China

splines [16, 20, 21], (iv) stream of pulses $\sum_k a_k p(t - t_k)$ where p is some pulse signal shape, applied in GPS applications and cellular radio, (v) sum of some of the signals above [19], and so on.

In this paper, we consider the signal spaces of finite rate of innovation $V^2(\Lambda, \Phi) \subset L^2(\mathbb{R}^d)$, defined as the closed linear span of a tuple of generators $\{\phi_\lambda\}_{\lambda \in \Lambda} \subset L^2(\mathbb{R}^d)$, which is introduced in [17, 19]:

$$V^2(\Lambda, \Phi) = \left\{ \sum_{\lambda \in \Lambda} c(\lambda) \phi_\lambda : (c_\lambda)_{\lambda \in \Lambda} \in \ell^2(\Lambda) \right\},$$

where Λ is a relatively-separated subset of \mathbb{R}^d .

Realistically we can learn about $f \in V^2(\Lambda, \Phi)$ only if the samples are taken in the set where most of the L^2 -norm is localized. So motivated by [3, 8], we focus on the sampling problem on the following subset of $V^2(\Lambda, \Phi)$:

$$V_{R,\delta}^2(\Lambda, \Phi) := \left\{ f \in V^2(\Lambda, \Phi) : \int_{C_R} |f(x)|^2 dx \geq (1 - \delta) \int_{\mathbb{R}^n} |f(x)|^2 dx \right\},$$

where $C_R = [-R/2, R/2]^n$ and $0 < \delta < 1$. Thus $V_{R,\delta}^2(\Lambda, \Phi)$ represents the subset consisting of those functions whose energy is largely concentrated on C_R .

Usually, in order to guarantee the success of reconstruction, we demand the *sampling inequality* as below fulfilled:

$$A \|f\|_2^2 \leq \sum_{j=1}^s |f(x_j)|^2 \leq B \|f\|_2^2,$$

where A and B are some positive constants. However, owing to the randomness in choosing the sampling points, we can not guarantee the sampling inequality fulfilled surely. Instead, our goal is to chase the following probability estimate:

$$\mathbb{P} \left(A \|f\|_2^2 \leq \sum_{j=1}^s |f(x_j)|^2 \leq B \|f\|_2^2 \right) \geq 1 - \epsilon, \tag{1}$$

where $\epsilon > 0$ can be taken as arbitrarily small. This problem is studied for the cases of band-limited signal spaces, shift-invariant signal spaces and finitely generated shift-invariant signal spaces, respectively [3, 8, 10, 25].

In former research [3, 4, 6, 8, 13], the probability distribution of random sampling points is taken as the uniform distribution in C_R . It is often beneficial in statistical sampling and function integration to sample *uniformly* over the applicable parameter space. However, non-uniform distributions also has application in real world [11, 13, 14]. A sequence of points which is uniformly distributed can be mapped into another sequence that reflects a desired non-uniform joint probability distribution function [13] using different techniques like the rejection method or weighted sampling [11, 14]. In this paper, we will consider a more generalized probability distribution: we only demand that the probability density function exists, and has positive lower bound in C_R . We will show that the uniform random sampling in C_R is a special case in our results about non-uniform random sampling.

In this paper, we will show the probability estimate (1) for non-uniform random sampling in signal spaces of finite rate of innovation. This is one of the main results in this paper.

In [9], a reconstruction method from nonuniform sampling is presented in shift-invariant spaces. But the sampling is NOT random sampling in [9]. In [4], a linear algorithm is obtained for the random sampling from regular languages. But how to give a reconstruction algorithm from uniform random sampling is still an open question in bandlimited functions spaces [3]. In this paper, we try to show a reconstruction algorithm from non-uniform random sampling in spaces with finite rate of innovation. This is one direct application of our main results.

Many problems arising in industry, statistics, physics, and finance require to compute multivariate integration $\int_{C_R} f(x)dx$. The sample mean method has been recommended to give an approximation to $\int_{C_R} f(x)dx$. The Koksma–Hlawka inequality gives an upper bound for the approximation error

$$\left| \int_{C_R} f(x)dx - \frac{1}{s} \sum_{i=1}^s f(x_i) \right| \leq C(\Gamma)P(f),$$

where $\Gamma = (x_i)_{i=1}^s \subset C_R$ is the set of sampling points with uniform distribution, $C(\Gamma)$ is the discrepancy of Γ and $P(f)$ is a measure of the variation of f [6, 13]. A spontaneous question is that, can we use the mean of the squares of those to estimate $\int_{\mathbb{R}^d} |f(x)|^2 dx$? Or what is the maximum error between $\int_{\mathbb{R}^d} |f(x)|^2 dx$ and $\frac{1}{s} \sum_{i=1}^s |f(x_i)|^2$? In this paper, we provide that the estimate error is direct proportional to the L^2 norm of the function. If $\Gamma = \{x_i\}_{i=1}^s \subset \mathbb{R}^d$ with non-uniform distribution fulfills some conditions with probability at least $1 - \epsilon$ for sufficiently small $\epsilon > 0$, then we can get that the estimate error is a functional only dependent on f , as shown below:

$$\mathbb{P} \left(\left| \int_{\mathbb{R}^d} |f(x)|^2 dx - \frac{1}{s} \sum_{i=1}^s |f(x_i)|^2 \right| \leq D(f) \right) \geq 1 - \epsilon.$$

This is one direct application of our main results too.

The paper is organized as follows. In Sect. 2, we state the theorem of random sampling in a space of finite rate of innovation, and give notations and preliminaries. In Sect. 3, some examples are shown about the generators and probability density function. In Sect. 4, spectrum estimation based on Hilbert–Schmidt operator is described. In Sect. 5, we discuss non-uniform random sampling in finite sums of eigenspaces. The proof of main results are provided in Sect. 6. And in the last section, we give a reconstruction algorithm for a special case, and give some numerical testes to verify them. Some properties of signal spaces $V^2(\Lambda, \Phi)$ are presented in appendix.

2 Statement of the Main Result

2.1 Notations and Preliminaries

We first define relatively-separated point sets, the relative covering index and the absolute covering index.

We say a point set $X = x_j \subset \mathbb{R}^d$ is *relatively-separated* [17], if the relative covering index

$$D(X) := \sup_{x \in \mathbb{R}^d} \sum_{x_j \in X} \chi_{x_j + [0,1]^d}(x) \tag{2}$$

is finite. And we can define the absolute covering index as below:

$$N_0(X) := \sup_{k \in \mathbb{Z}^d} \#(X \cap (k + [0, 1]^d)) \tag{3}$$

We can easily see that $N_0(X) \leq D(X)$ and $\sup_{x_0 \in [0,1]^d} N_0(X + x_0) = D(X)$ for any relatively-separated set X . The two definitions are used in different usages.

Throughout the paper, we consider a space of finite rate of innovation $V^2(\Lambda, \Phi) \subset L^2(\mathbb{R}^d)$, defined as the closed linear span of a tuple of generators $\{\phi_\lambda\}_{\lambda \in \Lambda} \subset L^2(\mathbb{R}^d)$:

$$V^2(\Lambda, \Phi) = \left\{ \sum_{\lambda \in \Lambda} c(\lambda) \phi_\lambda : (c_\lambda)_{\lambda \in \Lambda} \in \ell^2(\Lambda) \right\},$$

where Λ is a relatively-separated subset of \mathbb{R}^d with relative covering number $D(\Lambda)$.

In this paper, we use the polynomial weight function u below [17]:

$$u(x) = u_\alpha(x) = (1 + |x|)^\alpha, \tag{4}$$

where $\alpha \geq 0$. Usually we demand that the below norm of Φ is finite for some $q, p \in [1, \infty]$:

$$\begin{aligned} & \|\Phi\|_{q,p,u} \\ & := \max \left\{ \sup_{k \in \mathbb{Z}^d} \left(\sum_{\lambda \in \Lambda} \|\phi_\lambda(\cdot) u(\cdot - \lambda)\|_{L^q(k+[0,1]^d)}^p \right)^{\frac{1}{p}}, \sup_{\lambda \in \Lambda} \left(\sum_{k \in \mathbb{Z}^d} \|\phi_\lambda(\cdot) u(\cdot - \lambda)\|_{L^q(k+[0,1]^d)}^p \right)^{\frac{1}{p}} \right\}. \end{aligned} \tag{5}$$

The definition is a bit different with that in [17] and [19]. However, the definition in [17] and that in this paper are equivalent. Some properties of the signal space $V^2(\Lambda, \Phi)$ will be shown in [Appendix](#).

We pay attention to the subset $V_{R,\delta}^2(\Phi, \Lambda)$ of $V^2(\Lambda, \Phi)$:

$$V_{R,\delta}^2(\Phi, \Lambda) := \left\{ f \in V^2(\Phi, \Lambda) : \int_{C_R} |f(x)|^2 dx \geq (1 - \delta) \int_{\mathbb{R}^d} |f(x)|^2 dx \right\}.$$

What’s more, we will adopt a measurable typical or classical probability density function (P.D.F.) $w_1 : \mathbb{R}^d \rightarrow \mathbb{R}$ which satisfies $\|w_1\|_{L^1(\mathbb{R}^d)} = 1$ and $w_1(x) \geq 0, \forall x \in \mathbb{R}^d$. And the density function of the probability distribution we use to sample is defined as:

$$w_R(x) = \frac{1}{R^d} w_1\left(\frac{x}{R}\right). \tag{6}$$

We denote $v_R = \sqrt{w_R}$, $\|x\|_\infty = \max(|x_1|, \dots, |x_d|)$, and $C_R = [-\frac{R}{2}, \frac{R}{2}]^d$. Throughout this paper, we will frequently refer to the constants as in [Table 1](#).

2.2 Assumptions and Main Results

We collect our assumptions on the generators in the following list:

- (A.0) Φ is a Riesz basis of $V^2(\Lambda, \Phi)$.
- (A.1) $\|\Phi\|_{\infty,1,u_0} < \infty$.
- (A.2) $\|\Phi\|_{2,2,u_\alpha} < \infty$ for some $\alpha > 0$.

Table 1 Some constants

Symbols and definitions	Remarks
$D_1 = \ \tilde{\Phi}\ _{2,1,u_0}^4 \ \Phi\ _{2,1,u_0}^2 \ \Phi\ _{2,2,u_\alpha}^2$	$\tilde{\Phi}$ is dual of Φ
$D_2 = \ \Phi\ _{\infty,2,u_0} \ \tilde{\Phi}\ _{\infty,2,u_0}$	$\tilde{\Phi}$ is dual of Φ
$D_3 = \ \Phi\ _{\infty,1,u_0}^2 \ \tilde{\Phi}\ _{2,1,u_0}^2$	$\tilde{\Phi}$ is dual of Φ
$\lambda^* = \arg \min_{\lambda \in \Lambda} \{\ \lambda\ _\infty\}$	the point closest to the origin in Λ with Chebyshev distance
$\alpha' = \min(d, 2\alpha)$	
$\beta = \frac{5}{2} + 2 \frac{2\alpha}{\sqrt{2^{d+2} D_1}}$	
$R_0 = 2((\frac{\sqrt{2}\ \Phi\ _{2,2,u_\alpha}}{\ \phi_{\lambda^*}\ _2})^{1/\alpha} + \ \lambda^*\ _\infty)$	$R_0 \geq 2$
$M = \max\{\beta^d, \frac{1}{K_w}\}$	$K_w = \ w_1\ _{L^\infty(\mathbb{R}^d)}$

By (A.1), we have $\forall 1 \leq p, q \leq \infty, \|\Phi\|_{q,p,u_0} \leq \|\Phi\|_{\infty,1,u_0} < \infty$.

By (A.0) and [19, Theorem 4.1], we have that $\forall 1 \leq p, q \leq \infty, \|\tilde{\Phi}\|_{q,p,u_0} < \infty$.

The assumptions on the probability density function are shown in the following:

(B.0) w_1 is essentially bounded, or $K_w := \|w_1\|_{L^\infty(\mathbb{R}^d)} < \infty$.

(B.1) The essential infimum of w_1 in C_1 is positive, that is, there exists $\rho > 0$ such that $\text{ess inf}_{x \in C_1} w_1(x) \geq \rho K_w$, where K_w is defined in (B.0).

According to (6) and the assumptions of w_1 , we get that $\|w_R\|_{L^\infty(\mathbb{R}^d)} = R^{-d} K_w$ by (B.0), and $\text{ess sup}_{x \in C_R} w_R(x) > \rho R^{-d} K_w$ by (B.1).

Theorem 1 Assume that Φ satisfies assumptions (A.0)–(A.2), and w_1 satisfies assumptions (B.0)–(B.1). Let $(x_j)_{j \in \mathbb{N}} \subset \mathbb{R}^d$ be a sequence of i.i.d random variables with the probability density function w_R . Assume that $R \geq R_0$ and $\delta, \nu \in (0, 1)$ are sufficiently small to guarantee that

$$\frac{\nu^2}{K_w + \frac{\nu}{3}} \leq 6K_w D_2^2 \log\left(\frac{3}{\epsilon}\right) \quad \text{and} \quad A = K_w \rho \left(\frac{1}{2} - \delta\right) - \nu - 12K_w D_3 \delta > 0. \tag{7}$$

Let $0 < \epsilon < 1$. If the number s of samples satisfies

$$s \geq \max\left\{R^d \frac{2D_2(K_w + \frac{\nu}{3})}{\nu^2} \log\left(\frac{2D(\Lambda)M}{\epsilon} R^{d^2/\alpha'}\right), \frac{R^d}{3K_w}\right\}, \tag{8}$$

then the sampling inequality

$$\forall f \in V_{R,\delta}^2(\Lambda, \Phi) : sAR^{-d} \|f\|_2^2 \leq \sum_{j=1}^s |f(x_j)|^2 \leq sD_2 \|f\|_2^2 \tag{9}$$

holds with probability at least $1 - \epsilon$.

Remark 2 The result above shows the sampling rate of $O(R^d \log R)$ holds in the space of finite rate of innovation when fulfills the assumptions above and d is fixed, even if the distribution of random sampling is not uniform. If we have one more condition that u_α is 2-admissible, or $\alpha > d/2$ (see [17]), then $\alpha' = d$ and the sampling rate could be actually $O(R^d \log R^d)$ when d is not fixed.

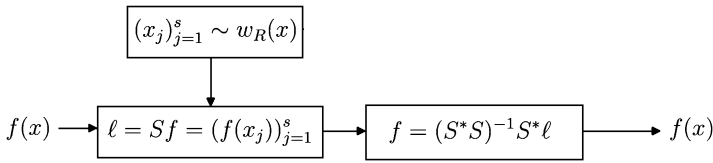


Fig. 1 A schematic showing random sampling and reconstruction

Remark 3 Theorem 1 is the main result in [3, 8, 25] in bandlimited signal spaces and shift-invariant spaces for uniform random sampling, respectively.

Corollary 4 Suppose that $\text{supp } \phi_\lambda \subset C_R$ for any $\lambda \in \Lambda$, thus $V_{R,\delta}^2(\Lambda, \Phi) = V^2(\Lambda, \Phi)$ for any small $\delta > 0$. For a sampling set $\Gamma = (x_j)_{j=1}^s \subset \mathbb{R}^d$, define the sampling operator $S : V^2(\Lambda, \Phi) \rightarrow \mathbb{C}^s$ by $f \mapsto (f(x_j))_{j=1}^s$. Denoting ℓ as the sampling result $(f(x_j))_{j=1}^s$, we have that S^*S is invertible under the conditions in Theorem 1 with probability at least $1 - \epsilon$, and thus $f = (S^*S)^{-1}S^*\ell$ in this case. A schematic of random sampling and reconstruction is shown in Fig. 1.

Corollary 5 Denote the conditional number of the operator S^*S in Corollary 4 by $\kappa(S^*S)$. Then the conditional number satisfies

$$\mathbb{P}\left(\kappa(S^*S) \leq \frac{2D_2R^d}{K_w\rho - 2v}\right) \geq 1 - \epsilon$$

under the conditions in Theorem 1. The upper bound above conditional number is the quotient of the upper bound and lower bound of the sampling inequality.

Corollary 6 For any $f \in V_{R,\delta}^2(\Lambda, \Phi)$, we have that

$$\mathbb{P}\left(\left|\int |f(x)|^2 dx - \frac{1}{s} \sum_{i=1}^s |f(x_i)|^2\right| \leq \max\{|1 - AR^{-d}|, |D_2 - 1|\} \|f\|_2^2\right) \geq 1 - \epsilon$$

when the conditions in Theorem 1 are fulfilled.

Remark 7 We expect that $D(s)$ in the below inequality tends to zero when $s \rightarrow \infty$:

$$\mathbb{P}\left\{\left|\int_{\mathbb{R}^d} |f(x)|^2 dx - \frac{1}{s} \sum_{i=1}^s |f(x_i)|^2\right| \leq D(s) \|f\|_2^2\right\} > 1 - \epsilon.$$

However, this can not reach in the FRI space. Thus the estimation in Corollary 6 is very coarse.

3 Some Examples of Generators and Probability Density Function

Two examples of generators of a space of finite rate of innovation fulfilling the assumptions of (A.0)–(A.2).

Example 8 The space S_m of all L^2 polynomial splines of order m fulfilling C^{m-2} continuity conditions at each knot t_i , where $m \in \mathbb{N}^+$ and $T = (t_i)_{i=-\infty}^{+\infty}$ is a bi-infinite increasing sequence satisfying

$$0 < T_0 \leq t_{i+m} - t_i \leq T_1 < \infty. \tag{10}$$

The space can be written as

$$S_m = \left\{ \sum_{i=-\infty}^{+\infty} c(i)B_i(x) : (c(i))_{i=-\infty}^{+\infty} \in \ell^2(\mathbb{Z}) \right\} \tag{11}$$

where B_i is the normalized B-spline associated with the knots t_i, \dots, t_{i+m} [16]. We have $B_i(x) = 0$ when $x \leq t_i$ or $x \geq t_{i+m}$. Let $\Lambda = T$ and $\phi_{t_i} = B_i$, then the support of B_i is a subset of $[t_i, t_{i+m}]$. Obviously, $D(T) \leq \frac{m}{T_0} + m$, so T is relatively separated.

Let $q = \infty, p = 1$, and take any weight function $u(x)$ that monotonically increases with $|x|$. Then we have that

$$\begin{aligned} \|\Phi\|_{\infty,1,u} &= \max \left\{ \sup_{i \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \|B_i(\cdot)u(\cdot - t_i)\|_{L^\infty(k+[0,1])}, \sup_{k \in \mathbb{Z}} \sum_{i \in \mathbb{Z}} \|B_i(\cdot)u(\cdot - t_i)\|_{L^\infty(k+[0,1])} \right\} \\ &\leq \sup_{i \in \mathbb{Z}} \sum_{k=\lfloor t_i \rfloor}^{\lceil t_{i+m} \rceil - 1} \|B_i(\cdot)u(\cdot - t_i)\|_{L^\infty([t_i, t_{i+m}])} \\ &\quad + \sup_{k \in \mathbb{Z}} \sum_{i: t_{i+m} \geq k \text{ and } t_i \leq k+1} \|B_i(\cdot)u(\cdot - t_i)\|_{L^\infty([t_i, t_{i+m}])} \\ &\leq \sup_{i \in \mathbb{Z}} \sum_{k=\lfloor t_i \rfloor}^{\lceil t_{i+m} \rceil - 1} \|B_i\|_{L^\infty([t_i, t_{i+m}])} u(t_{i+m} - t_i) \\ &\quad + \sup_{k \in \mathbb{Z}} \sum_{i: t_{i+m} \geq k \text{ and } t_i \leq k+1} \|B_i\|_{L^\infty([t_i, t_{i+m}])} u(t_{i+m} - t_i) \\ &\leq \sup_{i \in \mathbb{Z}} (\lceil t_{i+m} \rceil - \lfloor t_i \rfloor) u(t_{i+m} - t_i) \\ &\quad + \sup_{k \in \mathbb{Z}} \#\{i : t_{i+m} \geq k \text{ and } t_i \leq k + 1\} u(t_{i+m} - t_i) \\ &\leq \sup_{i \in \mathbb{Z}} (T_1 + 2) u(T_1) + \sup_{k \in \mathbb{Z}} \left(\frac{1}{T_0} + 2 \right) m u(T_1) \\ &= \left(T_1 + 2 + \left(\frac{1}{T_0} + 2 \right) m \right) u(T_1) < \infty \end{aligned}$$

by using that $\|B_i\|_\infty \leq 1$ according to [16]. So the assumption (A.1) holds.

At the same time, this implies that $\|\Phi\|_{q,p,u} \leq \|\Phi\|_{\infty,1,u} \leq (T_1 + 2 + (\frac{1}{T_0} + 2)m)u(T_1) < \infty$ for any $1 \leq q, p \leq \infty$. So the assumption (A.2) holds.

Moreover, by [16, Theorem 4.41], we have that for any $f = \sum_{i=-\infty}^{+\infty} c(i) B_i$,

$$\begin{aligned} \sum_{i=-\infty}^{+\infty} |c(i)|^2 &\leq \sum_{i=-\infty}^{+\infty} (2m + 1)^2 9^{2(m-1)} (t_{i+m} - t_i)^{-1} \|f\|_{L^2[t_i, t_{i+m}]}^2 \\ &\leq (2m + 1)^2 81^{m-1} T_0^{-1} \sum_{i=-\infty}^{+\infty} \|f\|_{L^2[t_i, t_{i+m}]}^2 \\ &= m(2m + 1)^2 81^{m-1} T_0^{-1} \|f\|_{L^2(\mathbb{R})}^2. \end{aligned}$$

In the other hand, we have

$$\begin{aligned} \int_{\mathbb{R}} |f(x)|^2 &= \int_{\mathbb{R}} \left| \sum_{i=-\infty}^{+\infty} c(i) B_i(x) \right|^2 dx \\ &= \sum_{j=-\infty}^{+\infty} \int_{[t_j, t_{j+1}]} \left| \sum_{i=j-m+1}^j c(i) B_i(x) \right|^2 dx \\ &\leq \sum_{j=-\infty}^{+\infty} \int_{[t_j, t_{j+1}]} \sum_{i=j-m+1}^j |c(i)|^2 \sum_{i=j-m+1}^j |B_i(x)|^2 dx \\ &\leq \sum_{j=-\infty}^{+\infty} m(t_{j+1} - t_j) \sum_{i=j-m+1}^j |c(i)|^2 \\ &\leq m^2 T_1 \sum_{i=-\infty}^{+\infty} |c(i)|^2. \end{aligned}$$

Thus $\Phi = \{B_i\}_{i=-\infty}^{+\infty}$ is a Riesz basis of S_m . Therefore the assumption (A.0) is fulfilled.

Example 9 The space of trigonometric polynomials segmented in C_R . This is a very trivial example. We will consider the following signal spaces:

$$T(k_0, t) = \left\{ f : f(x) = \sum_{k \in [-k_0, k_0]^d \cap \mathbb{Z}^d} c_k \exp\left(\frac{2\pi i}{t} x \cdot k\right), (c_k) \in \ell^2, x \in \mathbb{R}^d \right\}.$$

We can take Λ arbitrarily and give the correspondence between Λ and the basis arbitrarily, and denote the basis still by Φ . Note that Λ is finite, hence is relatively-separated.

If we take $R = mt$ where $m \in \mathbb{Z}^+$, it is a Riesz basis obviously since the basis is orthogonal. Moreover, for any weight function $u(x)$ that monotonically increases with $|x|$, we can easily get that $\|\phi_\lambda(\cdot)u(\cdot - \lambda)\|_{L^\infty(k+[0,1]^d)} \leq u(R/2 + 1 + \max_{\lambda \in \Lambda} \{\|\lambda\|_\infty\})$ is finite for $|k| \leq \lceil R/2 \rceil$ while $\|\phi_\lambda(\cdot)u(\cdot - \lambda)\|_{L^\infty(k+[0,1]^d)} = 0$ for $|k| > \lceil R/2 \rceil$. Then $\|\Phi\|_{\infty,1,u} \leq \max\{(2k_0 + 1)^d, (R + 2)^d\} \cdot u(R/2 + 1 + \max_{\lambda \in \Lambda} \{\|\lambda\|_\infty\})$ holds.

Two examples of probability density function fulfilling the assumptions (B.0)–(B.1).

Example 10 Beta distribution of one dimension

$$w_1(x) = B_{a,b,\alpha,\beta}(x) = \begin{cases} \frac{1}{B(\alpha,\beta)(b-a)^{\alpha+\beta-1}} (x-a)^\alpha (b-x)^\beta, & x \in [a, b] \\ 0, & \text{otherwise} \end{cases}$$

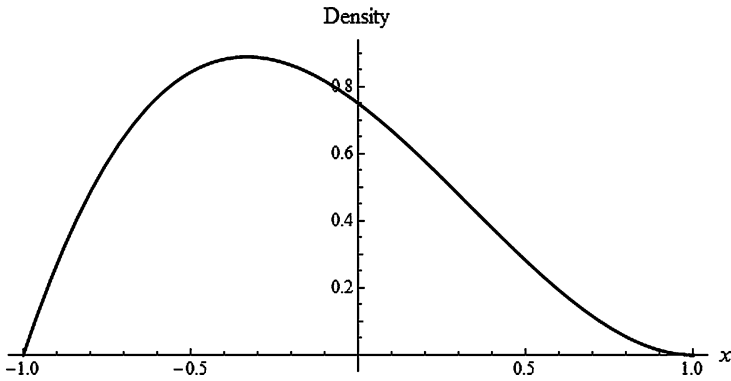


Fig. 2 Beta distribution

where $\alpha, \beta \geq 1, a < -\frac{1}{2}, b > \frac{1}{2}$ and $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ is the Beta function. We can check that $K_w = B_{a,b,\alpha,\beta}(\frac{\alpha-1}{\alpha+\beta-2}) < \infty$ and $\rho = \frac{\min\{B_{a,b,\alpha,\beta}(-\frac{1}{2}), B_{a,b,\alpha,\beta}(\frac{1}{2})\}}{K_w} > 0$. Specially, we take $\alpha = \beta = 1, a \rightarrow -\frac{1}{2}-$ and $b \rightarrow \frac{1}{2}+$, then w_R degenerates into the uniform distribution on C_R , just like the distribution used in [3] and [8]. Figure 2 is the plot of the P.D.F. of a typical Beta distribution with $a = -1, b = 1, \alpha = 1, \beta = 2$.

Example 11 Normal distribution centered at 0:

$$w_1(x) = N_\sigma(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

where $\sigma > 0$. We have that $K_w = N_\sigma(0) = \frac{1}{\sqrt{2\pi\sigma^2}} < \infty$ and $\rho = \frac{N_\sigma(\frac{1}{2})}{N_\sigma(0)} = \exp(-\frac{1}{8\sigma^2}) > 0$. For fixed $\delta \in (0, \frac{1}{2+24D_3})$, in order to optimize the lower bound A of the sampling inequality, the optimal choice σ satisfies

$$\left(\frac{1}{2} - \delta\right)(1 - 4\sigma^2) + 48D_3\delta \exp\left(\frac{1}{8\sigma^2}\right)\sigma^2 = 0 \tag{12}$$

It is easy to check that there exists a unique σ satisfying Eq. (12). We can guarantee that A is positive when σ take the optimal choice and v is small enough by the assumption $0 < \delta < \frac{1}{2+24D_3}$. Figure 3 is the plot of the P.D.F. of standard normal distribution.

4 Spectrum Estimation Based on Hilbert–Schmidt Operator

In the following, we will mostly work with the notation

$$P_\Phi = \sum_{\lambda \in \Lambda} \phi_\lambda \otimes \tilde{\phi}_\lambda, \tag{13}$$

where unconditional convergence in the strong operator topology is provided by the frame properties for all $f \in L^2(\mathbb{R}^n)$. The tensor product notation $v \otimes w$ refers to the rank-one operator $(v \otimes w)f \mapsto \langle f, w \rangle v$. By our assumptions (A.0)–(A.2), it is obvious that P_Φ is the orthogonal projection onto $V^2(\Lambda, \Phi)$.

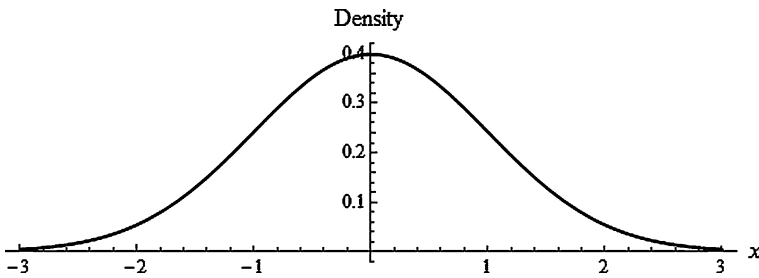


Fig. 3 Standard normal distribution

Given $R > 0$, we write $Q_R : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ for the orthogonal projection operator $f \mapsto f \cdot \chi_{C_R}$. We introduce the localization operator

$$A_R = P_\Phi \circ Q_R \circ P_\Phi.$$

We will show that A_R is self-adjoint and Hilbert–Schmidt, and thus has a basis of eigenvectors with associated ℓ^2 spectrum. We denote the span of the eigenvectors associated to the largest N eigenvalues by \mathcal{P}_N , and then establish a random sampling theorem for that space. Next we will built the connection between the sampling of the elements in $V_{R,\delta}^2(\Lambda, \Phi)$ and that in \mathcal{P}_N . How to choose the proper value of N depends on the decay of the spectrum of A_R . The localization of the above operator is the closest connection with the energy condition of $V_{R,\delta}^2(\Lambda, \Phi)$.

Lemma 12 A_R is a positive-semidefinite Hilbert–Schmidt operator.

Proof First of all, A_R is positive-semidefinite since

$$\langle A_R f, f \rangle = \langle Q_R P_\Phi f, P_\Phi f \rangle = \|Q_R P_\Phi f\|_2^2,$$

by using the self-adjointness of Q_R .

We next prove that $Q_R P_\Phi$ is Hilbert–Schmidt operator, which will imply that A_R is Hilbert–Schmidt operator.

Let $\forall \lambda \in \Lambda$, we have

$$Q_R \circ (\phi_\lambda \otimes \tilde{\phi}_\lambda) = (Q_R \phi_\lambda) \otimes \tilde{\phi}_\lambda.$$

By (13) and continuity of Q_R , we have

$$Q_R P_\Phi = \sum_{\lambda \in \Lambda} (Q_R \phi_\lambda) \otimes \tilde{\phi}_\lambda$$

with unconditional convergence in the strong operator topology.

Next we will prove that

$$\sum_{\lambda \in \Lambda} \|Q_R \phi_\lambda\|_2^2 < \infty. \tag{14}$$

In fact, we have

$$\begin{aligned} \sum_{\lambda \in \Lambda} \|Q_R \phi_\lambda\|_2^2 &= \sum_{\lambda \in \Lambda} \|\phi_\lambda\|_{L^2(C_R)}^2 \\ &\leq \sum_{k \in (\mathbb{Z} \cap [-\frac{R}{2}, \frac{R}{2}])^d} \sum_{\lambda \in \Lambda} \|\phi_\lambda\|_{L^2(k+[0,1]^d)}^2 \\ &\leq \sum_{k \in (\mathbb{Z} \cap [-\frac{R}{2}, \frac{R}{2}])^d} \|\Phi\|_{2,2,u_0}^2 \leq (R+2)^d \|\Phi\|_{2,2,u_0}^2 < \infty. \end{aligned}$$

Thus, we have

$$\begin{aligned} (Q_R P_\Phi)(Q_R P_\Phi)^* &= (Q_R P_\Phi) \sum_{\lambda \in \Lambda} \tilde{\phi}_\lambda \otimes (Q_R \phi_\lambda) \\ &= \sum_{\lambda \in \Lambda} (Q_R P_\Phi \tilde{\phi}_\lambda) \otimes (Q_R \phi_\lambda) \\ &= \sum_{\lambda \in \Lambda} (Q_R \tilde{\phi}_\lambda) \otimes (Q_R \phi_\lambda) \end{aligned}$$

with convergence in the strong operator topology. Similarly to the proof of Eq. (14), we see that

$$\sum_{\lambda \in \Lambda} \|Q_R \tilde{\phi}_\lambda\|_2^2 < \infty.$$

Combing with (14) and the Cauchy–Schwarz inequality, we can get that

$$\begin{aligned} \text{trace}(Q_R P_\Phi)(Q_R P_\Phi)^* &\leq \sum_{\lambda \in \Lambda} \text{trace}(|(Q_R \tilde{\phi}_\lambda) \otimes (Q_R \phi_\lambda)|) \\ &= \sum_{\lambda \in \Lambda} \|Q_R \tilde{\phi}_\lambda\| \|Q_R \phi_\lambda\| < \infty. \end{aligned}$$

Hence $Q_R P_\Phi$ is Hilbert–Schmidt operator. □

It follows that A_R has an orthogonal normalized basis of eigenvectors, whose associated eigenvalues are non-negative and square-summable. So the cluster points of the spectrum of A_R are either $\{0\}$ or \emptyset , and thus they can be sorted in decreasing order. Denoting the non-zero eigenvalues by $(\mu_n)_{n \in I}$ sorted in decreasing order and the associated eigenfunctions by $(\psi_n)_{n \in I}$, with I being either \mathbb{N}_+ or $\{1, \dots, M\}$ for some integer M . Then A_R is given as the sum

$$\sum_{n \in I} \mu_n \psi_n \otimes \psi_n.$$

We can easily check that $\psi_n \in V^2(\Phi, \Lambda)$, as well as $P_\Phi(\psi_n) = \psi_n$. Since A_R is a composition of orthogonal projections, we have $\mu_1 \leq \|A_R\|_{op} \leq 1$. Let

$$\mathcal{P}_N = \text{span}_{n \in \mathbb{N}_+, n \leq N} \{\psi_n\}.$$

For the sampling problem we will need some information on the spectrum of A_R . Let

$$N(R) = \max\{n \in \mathbb{N} : \mu_n \geq 1/2\},$$

and $N(R) = 0$ when $\mu_1 < 1/2$. Then whenever $N(R) > 0$, we have $\mu_{N(R)} \geq 1/2 > \mu_{N(R)+1}$.

An upper bound for $N(R)$ is provided in the following lemma.

Lemma 13 *Let β and R_0 be that defined in Table 1. Then for all $R > R_0$, the inequalities $0 < N(R) \leq D(\Lambda)\beta^d R^{d^2/a'}$ hold, where $D(\Lambda)$ stands for the relative covering index of Λ , defined in (2).*

Proof We use the minimax formula for a given order of eigenvalue

$$\mu_m = \inf\{\sup\{\langle A_R f, f \rangle : f \perp \mathcal{H}, \|f\|_2 = 1\} : \mathcal{H} \subset L^2(\mathbb{R}^n), \dim(\mathcal{H}) < m\}.$$

Now fix $S \geq R + 2$, and denote $\mathcal{H}_S := \text{span}\{\tilde{\phi}_\lambda : \|\lambda\|_\infty \leq S/2\}$. It follows that

$$\dim(\mathcal{H}_S) \leq (\lfloor S \rfloor + 1)^d D(\Lambda).$$

For any unit vector f in \mathcal{H}_S^\perp , we obtain

$$\langle A_R f, f \rangle = \sum_{\lambda, \lambda' \in \Lambda} \langle f, \tilde{\phi}_\lambda \rangle \langle Q_R \phi_\lambda, \phi_{\lambda'} \rangle \overline{\langle f, \tilde{\phi}_{\lambda'} \rangle}.$$

We denote $c_\lambda = \langle f, \tilde{\phi}_\lambda \rangle$, $c = (c_\lambda)_{\lambda \in \Lambda}$,

$$a_{\lambda, \lambda'} = \begin{cases} \langle Q_R \phi_\lambda, \phi_{\lambda'} \rangle, & \min\{\|\lambda\|_\infty, \|\lambda'\|_\infty\} > \frac{S}{2} \\ 0, & \text{otherwise} \end{cases}$$

and an infinite, positive semidefinite matrix

$$\mathcal{A} = (a_{\lambda, \lambda'})_{\lambda, \lambda' \in \Lambda},$$

by recalling the assumption that $f \perp \tilde{\phi}_\lambda$ for $\|\lambda\|_\infty \leq S/2$. In particular, we get

$$\langle A_R f, f \rangle = c \mathcal{A} c^* \leq \|c\|_2^2 \|\mathcal{A}\|_{\text{op}} \leq \|\tilde{\Phi}\|_{2,1,u_0}^2 \|\mathcal{A}\|_{HS}$$

by using Proposition 22 in Appendix. To estimate the Hilbert–Schmidt norm of the matrix, we have

$$\begin{aligned} \|\mathcal{A}\|_{HS}^2 &= \sum_{\substack{\lambda, \lambda' \in \Lambda \\ \|\lambda\|_\infty > S/2 \\ \|\lambda'\|_\infty \geq S/2}} |a_{\lambda, \lambda'}|^2 \\ &\leq \sum_{\substack{\lambda \in \Lambda \\ \|\lambda\|_\infty > S/2}} \sum_{\lambda' \in \Lambda} |\langle Q_R \phi_\lambda, \phi_{\lambda'} \rangle|^2 \\ &\leq \sum_{\substack{\lambda \in \Lambda \\ \|\lambda\|_\infty > S/2}} \sum_{\lambda' \in \Lambda} \|Q_R \phi_\lambda\|_2^2 \|\phi_{\lambda'}\|_2^2 \\ &= \sum_{\substack{\lambda \in \Lambda \\ \|\lambda\|_\infty > S/2}} \|Q_R \phi_\lambda\|_2^2 \sum_{\lambda' \in \Lambda} \|\phi_{\lambda'}\|_2^2 \\ &\leq \sum_{\substack{\lambda \in \Lambda \\ \|\lambda\|_\infty > S/2}} \|\Phi\|_{2,1,u_0}^2 \|Q_R \phi_\lambda\|_2^2 \end{aligned}$$

where Proposition 22 in Appendix is used in the above argument again.

Now we will estimate $\sum_{\lambda, \|\lambda\|_\infty > S/2} \|\mathcal{Q}_R \phi_\lambda\|_2^2$ by the following:

$$\begin{aligned} & \sum_{\substack{\lambda \in A \\ \|\lambda\|_\infty > \frac{S}{2}}} \|\mathcal{Q}_R \phi_\lambda\|_2^2 \\ &= \sum_{\substack{\lambda \in A \\ \|\lambda\|_\infty > \frac{S}{2}}} \|\phi_\lambda\|_{L^2(C_R)} \\ &\leq \sum_{k \in (\mathbb{Z} \cap [-\frac{R}{2}, \frac{R}{2}])^d} \sum_{\substack{\lambda \in A \\ \|\lambda\|_\infty > \frac{S}{2}}} \|\phi_\lambda\|_{L^2(k+[0,1]^d)} \\ &\leq \sum_{k \in (\mathbb{Z} \cap [-\frac{R}{2}, \frac{R}{2}])^d} \sum_{\substack{\lambda \in A \\ \|\lambda\|_\infty > \frac{S}{2}}} \|\phi_\lambda u_\alpha(\cdot - \lambda)\|_{L^2(k+[0,1]^d)} \sup_{x \in k+[0,1]^d} (u_\alpha(x - \lambda))^{-2} \\ &\leq \left(\inf_{\|x\|_\infty \geq \frac{S}{2} - \frac{R}{2} - 1} u_\alpha(x) \right)^{-2} (R + 2)^d \|\Phi\|_{2,2,u_\alpha}^2. \end{aligned}$$

Therefore we have

$$\begin{aligned} \langle A_R f, f \rangle^2 &\leq \|\tilde{\Phi}\|_{2,1,u_0}^4 \|A\|_{HS}^2 \\ &\leq \|\tilde{\Phi}\|_{2,1,u_0}^4 \|\Phi\|_{2,1,u_0}^2 \|\Phi\|_{2,2,u_\alpha}^2 (R + 2)^d \left(u_\alpha \left(\frac{S}{2} - \frac{R}{2} - 1 \right) \right)^{-2} \\ &= D_1 (R + 2)^d \left(\frac{S}{2} - \frac{R}{2} \right)^{-2\alpha}. \end{aligned}$$

For $S = (\beta - \frac{3}{2})R^{d/\alpha'} + 2$ with β in Table 1, one finds that $R > R_0 \geq 2$ and the definition of α' yield that $R \leq R^{d/\alpha'}$, and hence

$$\begin{aligned} D_1 (R + 2)^d \left(\frac{S}{2} - \frac{R}{2} \right)^{-2\alpha} &\leq D_1 (R + 2)^d \left(\frac{\beta - \frac{1}{2}}{2} R^{d/\alpha'} + 1 - \frac{R}{2} \right)^{-2\alpha} \\ &< D_1 (R + 2)^d \left(\frac{\beta - \frac{3}{2}}{2} R^{d/\alpha'} - \frac{1}{2} R^{d/\alpha'} \right)^{-2\alpha} \\ &\leq D_1 (R + 2)^d \left(\frac{\beta - \frac{5}{2}}{2} \right)^{-2\alpha} R^{-\frac{2\alpha d}{\alpha'}} \\ &\leq \frac{1}{4}. \end{aligned}$$

Hence we have shown

$$\langle A_R f, f \rangle < 1/2$$

for all unit vectors $f \in \mathcal{H}_S^\perp$.

If $R > 2$, then

$$\begin{aligned} N &= \left(\left\lfloor \left(\beta - \frac{3}{2} \right) R^{d/\alpha'} + 2 \right\rfloor + 1 \right)^d D(\Lambda) \\ &\leq \left(\left(\beta - \frac{3}{2} \right) R^{d/\alpha'} + 3 \right)^d D(\Lambda) \\ &\leq \beta^d R^{d^2/\alpha'} D(\Lambda), \end{aligned}$$

and the minimax estimate yields $\mu_N < 1/2$.

On the other hand, if $R > R_0 \geq 2\|\lambda^*\|_\infty + 2$, then

$$\begin{aligned} \mu_1 &\geq \frac{\langle A_R \phi_{\lambda^*}, \phi_\lambda \rangle}{\|\phi_{\lambda^*}\|_{L^2(\mathbb{R}^d)}^2} \\ &= \frac{\langle Q_R \phi_{\lambda^*}, \phi_\lambda \rangle}{\|\phi_{\lambda^*}\|_{L^2(\mathbb{R}^d)}^2} \\ &= 1 - \frac{\|\phi_{\lambda^*}\|_{L^2(\mathbb{R}^d \setminus C_R)}^2}{\|\phi_{\lambda^*}\|_{L^2(\mathbb{R}^d)}^2} \\ &\geq 1 - \frac{\sum_{k \in (\mathbb{Z} \cap [-\frac{R}{2}, \frac{R}{2}])^d} \|\phi_{\lambda^*}\|_{L^2(k+[0,1]^d)}^2}{\|\phi_{\lambda^*}\|_{L^2(\mathbb{R}^d)}^2} \\ &\geq 1 - \frac{\sum_{k \in (\mathbb{Z} \cap [-\frac{R}{2}, \frac{R}{2}])^d} \|\phi_{\lambda^*} u_\alpha(\cdot - \lambda^*)\|_{L^2(k+[0,1]^d)} (\inf_{x \in k+[0,1]^d} u_\alpha(x - \lambda^*))^{-2}}{\|\phi_{\lambda^*}\|_{L^2(\mathbb{R}^d)}^2} \\ &\geq 1 - \frac{\|\Phi\|_{2,2,u_\alpha}^2 (u_\alpha(\lfloor \frac{R}{2} \rfloor - \|\lambda^*\|_\infty))^{-2}}{\|\phi_{\lambda^*}\|_{L^2(\mathbb{R}^d)}^2} > \frac{1}{2}. \end{aligned}$$

So for all $R > R_0$ the following inequality holds

$$0 < N(R) \leq \beta^d R^{d^2/\alpha'} D(\Lambda). \quad \square$$

5 Non-uniform Random Sampling in Finite Sums of Eigenspaces \mathcal{P}_N

Recall from the previous section that $\lambda_1 \geq \lambda_2 \geq \dots$ are the eigenvalues of A_R , with corresponding eigenfunctions ψ_1, ψ_2, \dots . The span of the first N eigenfunctions is denoted by \mathcal{P}_N . We let $\Delta_N = \text{diag}(\lambda_1, \dots, \lambda_N)$.

The goal of this section is to show a non-uniform random sampling statement for \mathcal{P}_N . We will show it by applying a matrix Bernstein inequality which will be stated in Theorem 14, using the following notation: For $A \in \mathbb{C}^{N \times N}$, we let $\|A\|$ denote the operator norm with respect to the Euclidean norm. Further, the inequality $A \leq B$ for two matrices A, B of equal size means that $B - A$ is positive semidefinite.

Theorem 14 (Matrix Bernstein inequality [22, Theorem 1.4]) *Let X_j be a sequence of independent, random self-adjoint $N \times N$ -matrices. Suppose that $\mathbb{E}X_j = 0$, and $\|X_j\| \leq B$ a.s.*

Then for all $t > 0$,

$$\mathbb{P}\left(\lambda_{\max}\left(\sum_{i=1}^s X_j\right) \geq t\right) \leq N \exp\left(-\frac{t^2/2}{\sigma^2 + Bt/3}\right)$$

holds where $\lambda_{\max}(U)$ is the largest singular value of a matrix U so that $\|U\| = \lambda_{\max}(U^*U)^{1/2}$ is the operator norm, and $\sigma^2 = \|\sum_{j=1}^s \mathbb{E}(X_j^2)\|$.

The random matrices under consideration are constructed as follows: For each $j \in \mathbb{N}$ and $k, l \in \{1, \dots, N\}$, we introduce the $N \times N$ rank-one random matrix M_j defined by

$$(M_j)_{k,l} = \psi_k(x_j)\overline{\psi_l(x_j)}. \tag{15}$$

Here the x_j 's denote i.i.d. random variables, and are distributed in \mathbb{R}^d with probability density function w_R . Let

$$X_j = M_j - \mathbb{E}(M_j). \tag{16}$$

We can now formulate and prove the non-uniform random sampling statement for \mathcal{P}_N by using Theorem 14.

Theorem 15 *Let $(x_j)_{j \in \mathbb{N}} \subset \mathbb{R}^d$ be a sequence of independent and identically distributed random variables with probability density function w_R . Then, for all $\nu \geq 0$ and $s \in \mathbb{N}$:*

$$\mathbb{P}\left(\inf_{f \in \mathcal{P}_N, \|f\|_2=1} \frac{1}{s} \sum_{j=1}^s (|f(x_j)|^2 - \|v_R f\|_2^2) \leq -R^{-d}\nu\right) \leq N \exp\left(-\frac{\nu^2 s}{2D_2 R^d (K_w + \frac{\nu}{3})}\right) \tag{17}$$

holds where v_R denotes the square root of w_R and $\|v_R f\|_2^2$ is the expectation of $|f(x_j)|^2$.

Proof By the definition of M_j defined by (15), we have

$$\begin{aligned} (\mathbb{E}(M_j))_{k,\ell} &= \int_{\mathbb{R}^d} w_R(x)\psi_\ell(x)\overline{\psi_k(x)}dx \\ &= \int_{\mathbb{R}^d} v_R(x)\psi_\ell(x)\overline{v_R(x)}\overline{\psi_k(x)}dx \\ &= \langle v_R\psi_\ell, v_R\psi_k \rangle. \end{aligned}$$

Furthermore, for any unit vector $f \in \mathcal{P}_N$ we can rewrite it by $f = \sum_{n=1}^N c_n \psi_n$ with a unit vector $(c_n)_n$ of coefficients, and

$$\langle c, M_j c \rangle = |f(x_j)|^2,$$

and similarly

$$\begin{aligned} \langle c, \mathbb{E}M_j c \rangle &= \sum_k \sum_\ell c_k \overline{c_\ell} (\mathbb{E}M_j)_{k,\ell} \\ &= \sum_k \sum_\ell c_k \overline{c_\ell} \langle v_R\psi_k, v_R\psi_\ell \rangle \end{aligned}$$

$$\begin{aligned}
 &= \left\langle v_R \sum_k c_k \psi_k, v_R \sum_\ell c_\ell \psi_\ell \right\rangle \\
 &= \|v_R f\|_2^2,
 \end{aligned}$$

which implies that both M_j and $\mathbb{E}M_j$ are positive semi-definite.

Thus we have

$$\begin{aligned}
 &\inf_{f \in \mathcal{P}_N, \|f\|_2=1} \frac{1}{s} \sum_{j=1}^s (|f(x_j)|^2 - \|v_R f\|_2^2) \\
 &= \inf_{\|c\|_2=1} \frac{1}{s} \sum_{j=1}^s (\langle c, M_j c \rangle - \langle c, \mathbb{E}M_j c \rangle) \\
 &= \lambda_{\min} \left(\frac{1}{s} \sum_{j=1}^s X_j \right),
 \end{aligned}$$

where λ_{\min} represents the smallest eigenvalue of a self-adjoint matrix.

Next we estimate $\|X_j\|$. For all $\|c\|_2 = 1$ and $f = \sum_{k=1}^s c_k \psi_k$, we have

$$\langle c, M_j c \rangle = |f(x_j)|^2 \leq \|f\|_\infty^2.$$

Since we can ignore a zero-measure subset from the existence of probability density function of x_j , as well as

$$\begin{aligned}
 \langle c, \mathbb{E}M_j c \rangle &= \|v_R f\|_2^2 \\
 &\leq \|v_R\|_2^2 \|f\|_\infty^2 = \|f\|_\infty^2.
 \end{aligned}$$

Thus we have

$$\begin{aligned}
 |\langle c, X_j c \rangle| &= |\langle c, M_j c \rangle - \langle c, \mathbb{E}M_j c \rangle| \\
 &\leq \|f\|_\infty^2 \leq D_2
 \end{aligned}$$

by Proposition 24 in Appendix. Therefore $\|X_j\| \leq D_2$ holds.

Finally we will estimate σ^2 . In fact, we have that

$$\mathbb{E}(X_j^2) = \mathbb{E}(M_j^2) - [\mathbb{E}(M_j)]^2 \leq \mathbb{E}(M_j^2),$$

and

$$\begin{aligned}
 (M_j^2)_{km} &= \sum_{\ell=1}^N (M_j)_{k\ell} (M_j)_{\ell m} \\
 &= \sum_{\ell=1}^N \psi_k(x_j) \overline{\psi_\ell(x_j)} \psi_\ell(x_j) \overline{\psi_m(x_j)} \\
 &= \sum_{\ell=1}^N |\psi_\ell(x_j)|^2 (M_j)_{km}.
 \end{aligned}$$

Thus $M_j^2 = \sum_{\ell=1}^N |\psi_\ell(x_j)|^2 (M_j)_{km}$. Moreover, we have the following estimation that

$$\begin{aligned} \sum_{\ell=1}^N |\psi_\ell(x_j)|^2 &= \sum_{\ell=1}^N |\langle \psi_\ell, v_{x_j} \rangle|^2 \\ &\leq \sum_{\ell=1}^{\infty} |\langle \psi_\ell, v_{x_j} \rangle|^2 \\ &= \|v_{x_j}\|_2^2 \leq D_2 \end{aligned}$$

almost everywhere by Proposition 24 in Appendix.

Therefore

$$M_j^2 \leq D_2 M_j,$$

as well as

$$\mathbb{E}(M_j^2) \leq D_2 \mathbb{E}(M_j).$$

By the arguments above, we can get that

$$\begin{aligned} \sigma^2 &= \left\| \sum_{j=1}^s \mathbb{E}(X_j^2) \right\| \\ &\leq \left\| \sum_{j=1}^s \mathbb{E}(M_j^2) \right\| \\ &\leq D_2 \left\| \sum_{j=1}^s \mathbb{E}(M_j) \right\| \\ &\leq s D_2 \sup_{\|f\|_2=1} \|v_R f\|_2^2 \\ &\leq s D_2 \|v_R\|_\infty^2 = s D_2 R^{-d} K_w. \end{aligned}$$

Now the statement of the theorem follows from the matrix Bernstein inequality formulated in Theorem 14, with proper estimation constants provided above. □

6 Proofs of the Main Results

It remains to transfer the random sampling statements from the spaces \mathcal{P}_N to the set $V_{R,\delta}(\Phi)$. The following lemma is a first step in this direction, by providing a norm estimate for the projection onto \mathcal{P}_N .

Lemma 16 *Let $N \in \mathbb{N}$ and $\gamma \in \mathbb{R}$ with $\lambda_N \geq \gamma \geq \lambda_{N+1}$. Denote the orthogonal projection onto \mathcal{P}_N by E_N , and the orthogonal projection onto the orthogonal complement by F_N . Then for all $f \in V_{R,\delta}(\Phi)$, we have*

$$\|E_N f\|_2^2 \geq \left(1 - \frac{\delta}{1 - \gamma}\right) \|f\|_2^2,$$

$$\begin{aligned} \|Q_R E_N f\|_2^2 &\geq \gamma \left(1 - \frac{\delta}{1-\gamma}\right) \|f\|_2^2, \\ \|F_N f\|_2^2 &\leq \frac{\delta}{1-\gamma} \|f\|_2^2. \end{aligned}$$

If $N = N(R) \neq 0$, then these estimation can be simplified into

$$\begin{aligned} \|E_N f\|_2^2 &\geq (1 - 2\delta) \|f\|_2^2, \\ \|Q_R E_N f\|_2^2 &\geq \left(\frac{1}{2} - \delta\right) \|f\|_2^2, \\ \|F_N f\|_2^2 &\leq 2\delta \|f\|_2^2. \end{aligned}$$

Proof Let $f \in V_{R,\delta}(\Phi)$, w.l.o.g. $\|f\|_2 = 1$. Since $f = P_\Phi f$, we obtain

$$1 - \delta \leq \|Q_R f\|_2^2 = \|Q_R P_\Phi f\|_2^2 = \langle Q_R P_\Phi f, Q_R P_\Phi f \rangle = \langle A_R f, f \rangle = \sum_j |\langle f, \psi_j \rangle|^2 \lambda_j.$$

Let $c_j = \langle f, \psi_j \rangle$, and define

$$A = \|E_N f\|_2^2 = \sum_{j=1}^N |c_j|^2,$$

and $B = 1 - A = \|F_N f\|_2^2$. Then $\sum_{j=N+1}^\infty |c_j|^2 \leq \|F_N f\|_2^2 = 1 - A$. Using $\gamma \geq \lambda_{N+1} \geq \lambda_{N+2} > \dots$ and $\lambda_j \leq 1$, we find

$$\begin{aligned} A &= \sum_{j=1}^N |\langle f, \psi_j \rangle|^2 \geq \sum_{j=1}^N |\langle f, \psi_j \rangle|^2 \lambda_j \\ &= \sum_{j=1}^\infty |c_j|^2 \lambda_j - \sum_{j=N+1}^\infty |c_j|^2 \lambda_j \\ &\geq 1 - \delta - \gamma \left(\sum_{j=N+1}^\infty |c_j|^2 \right) \\ &\geq 1 - \delta - \gamma(1 - A). \end{aligned}$$

Solving this inequality for A yields $A \geq 1 - \frac{\delta}{1-\gamma}$, which implies $B \leq \frac{\delta}{1-\gamma}$. Finally, $\gamma \leq \lambda_N$ holds, which implies

$$\|Q_R E_N f\|_2^2 = \sum_{j=1}^N \lambda_j |c_j|^2 \geq \gamma A \geq \gamma \left(1 - \frac{\delta}{1-\gamma}\right).$$

For $N = N(R) \neq 0$, we may pick $\gamma = 1/2$, which results in the estimates given for this case. □

To connect the sampling theorem between $V_{R,\delta}^2(\Lambda, \Phi)$ and the eigenspace \mathcal{P}_N , we need a statement below.

Lemma 17 Let $X = \{x_j\}_{j=1}^s$ be a finite subset of \mathbb{R}^d and $\alpha \in [\mu_{N+1}, \mu_N]$. Assume that the inequality

$$\frac{1}{s} \sum_{j=1}^s (|g(x_j)|^2 - \|v_R g\|_2^2) \geq -R^{-d} v \|g\|_2^2$$

holds for all $g \in \mathcal{P}_N$. Then the inequality

$$\sum_{j=1}^s |f(x_j)|^2 \geq \hat{A} \|f\|_2^2$$

holds for all $f \in V_{R,\delta}^2(\Lambda, \Phi)$, where

$$\hat{A} = sR^{-d} \left(K_w^2 \rho \alpha \left(1 - \frac{\delta}{1-\alpha} \right) - v \right) - 2D_3 N_0(X) \frac{\delta}{1-\alpha}.$$

Proof Denoting the orthogonal projection of a square-integrable function f to \mathcal{P}_N by Ef . By Lemma 16 above and Proposition 23 in Appendix, we have

$$\begin{aligned} \sum_{j=1}^s |f(x_j)|^2 &\geq \sum_{j=1}^s |Ef(x_j)|^2 - D_3 N_0(X) \|f - Ef\|_2^2 \\ &\geq \sum_{j=1}^s |Ef(x_j)|^2 - 2D_3 N_0(X) \frac{\delta}{1-\alpha} \|f\|_2^2 \\ &\geq s \|v_R Ef\|_{L^2(C_R)}^2 - \frac{vs}{R^d} \|Ef\|_2^2 - 2D_3 N_0(X) \frac{\delta}{1-\alpha} \|f\|_2^2 \\ &\geq s K_w R^{-d} \rho \|Ef\|_{L^2(C_R)}^2 - \frac{vs}{R^d} \|Ef\|_2^2 - 2D_3 N_0(X) \frac{\delta}{1-\alpha} \|f\|_2^2 \\ &\geq K_w R^{-d} \rho s \alpha \left(1 - \frac{\delta}{1-\alpha} \right) \|f\|_2^2 - \frac{vs}{R^d} \|f\|_2^2 - 2D_3 N_0(X) \frac{\delta}{1-\alpha} \|f\|_2^2. \end{aligned}$$

Then we get the conclusion. □

Before proving the main result, we have a following statement about absolute covering index.

Lemma 18 Suppose $X = \{x_j\}_{j=1}^s$ are independent and identically distributed random variables in \mathbb{R}^d . Let $P \geq \sup_{k \in \mathbb{Z}^d} \mathbb{P}(x_j \in k + [0, 1]^d)$ and $a > eP$ with $as \geq 1$. Then we have

$$\mathbb{P}(N_0(X) > as) \leq \frac{1}{P} \left(\frac{eP}{a} \right)^{as}$$

where $N_0(X)$ represents the absolute covering index of X stated in (3).

Proof Let $P_k = \mathbb{P}(x_j \in k + [0, 1]^d)$ for $k \in \mathbb{Z}^d$. Combing with Chebyshev's inequality, then we have that

$$\begin{aligned} \mathbb{P}(N_0(X) > as) &\leq \sum_{k \in \mathbb{Z}^d} \mathbb{P}(\#\{x_j\} \cap k + [0, 1]^d > as) \\ &= \sum_{k \in \mathbb{Z}^d} \mathbb{P}\left(\sum_{j=1}^s \chi_{k+[0,1]^d}(x_j) > as\right) \\ &= \sum_{k \in \mathbb{Z}^d} \mathbb{P}\left(\exp\left(b_k \sum_{j=1}^s \chi_{k+[0,1]^d}(x_j)\right) > e^{b_k as}\right) \\ &\leq \sum_{k \in \mathbb{Z}^d} e^{b_k as} \mathbb{E} \exp\left(b_k \sum_{j=1}^s \chi_{k+[0,1]^d}(x_j)\right) \\ &= \sum_{k \in \mathbb{Z}^d} e^{b_k as} \prod_{j=1}^s \mathbb{E} \exp(b_k \chi_{k+[0,1]^d}(x_j)) \\ &= \sum_{k \in \mathbb{Z}^d} e^{b_k as} (1 + (e^{b_k} - 1)P_k)^s \\ &\leq \sum_{k \in \mathbb{Z}^d} e^{b_k as} (\exp((e^{b_k} - 1)P_k))^s. \end{aligned}$$

Taking the optimal choice $b_k = \log\left(\frac{a}{P_k}\right)$, we obtain that

$$\begin{aligned} \mathbb{P}(N_0(X) > as) &\leq \sum_{k \in \mathbb{Z}^d} \exp\left(-s\left(a \log\left(\frac{a}{P_k}\right) - (a - P_k)\right)\right) \\ &= \sum_{k \in \mathbb{Z}^d} \exp\left(-sa \log\left(\frac{a}{P_k}\right) + sa - sP_k\right) \\ &\leq \sum_{k \in \mathbb{Z}^d} \exp\left(as\left(1 - \log\left(\frac{a}{P_k}\right)\right)\right) \\ &= \sum_{k \in \mathbb{Z}^d} \exp\left(as\left(\log\left(\frac{eP_k}{a}\right)\right)\right) \\ &= \sum_{k \in \mathbb{Z}^d} \left(\frac{eP_k}{a}\right)^{as} = \sum_{k \in \mathbb{Z}^d} \left(\frac{eP_k}{a}\right) \left(\frac{eP_k}{a}\right)^{as-1} \\ &\leq \sum_{k \in \mathbb{Z}^d} \left(\frac{eP_k}{a}\right) \left(\frac{eP}{a}\right)^{as-1} = \frac{1}{P} \left(\frac{eP}{a}\right)^{as}. \quad \square \end{aligned}$$

Remark 19 For a distribution with probability density function $w(x)$, we can take $P = \|w\|_\infty$ so that for a and s satisfying the conditions of Lemma 18. So we have

$$\mathbb{P}(N_0(X) > as) \leq \frac{1}{\|w\|_\infty} \left(\frac{e\|w\|_\infty}{a}\right)^{as}.$$

Remark 20 If we take a special example that the distribution is uniform distribution on C_R , then P can be taken as $\lfloor R \rfloor^{-d}$. So we have

$$\mathbb{P}(N_0(X) > as) \leq \lfloor R \rfloor^d \left(\frac{e \lfloor R \rfloor^{-d}}{a} \right)^{as} \leq R^d \exp(-s(a \log(a \lfloor R \rfloor^d) - a)).$$

It provide a different estimate from [3, Lemma 8]. Note that we demand a more strictly conditions than that in [3], but our estimate might be better than that in [3]. And our result is suitable for a large kinds of non-uniform probability distribution.

Then we can see the following statement.

Theorem 21 *Let $(x_j)_{j \in \mathbb{N}} \subset \mathbb{R}^d$ be a sequence of independent random variables with probability density function w_R . Assume that $R > R_0$, and furthermore*

$$\delta < \frac{\rho}{2(\rho + 12D_3)}, \quad v < K_w \left(\rho \left(\frac{1}{2} - \delta \right) - 12D_3\delta \right).$$

Then, for any $s \in \mathbb{N}$ and $s \geq \frac{R^d}{3K_w}$,

$$A = K_w \rho \left(\frac{1}{2} - \delta \right) - v - 12K_w D_3 \delta$$

is strictly positive, and the sampling estimate

$$sAR^{-d} \|f\|_2^2 \leq \sum_{j=1}^s |f(x_j)|^2 \leq sD_2 \|f\|_2^2, \quad \forall f \in V_{R,\delta}^2(\Lambda, \Phi) \tag{18}$$

holds with probability at least

$$1 - D(\Lambda)R^{d2/\alpha'} \beta^d \exp\left(-\frac{v^2s}{2D_2R^d(K_w + \frac{v}{3})}\right) - \frac{R^d}{K_w} \left(\frac{e}{3}\right)^{3sR^{-d}K_w}. \tag{19}$$

Proof Define the random variable N_0 as the absolute covering index of x_1, \dots, x_s . Fix $N = N(R)$, and consider the events

$$V_1 = \left\{ \inf_{f \in \mathcal{P}_N, \|f\|_2=1} \frac{1}{s} \sum_{j=1}^s (|f(x_j)|^2 - \|v_R f\|_2^2) \leq -vR^{-d} \right\}$$

and

$$V_2 = \{N_0 \geq 3R^{-d}K_w\}.$$

Following Lemma 17 and taking $\alpha = \frac{1}{2}$, we see that for all (x_1, \dots, x_s) in the complement of $V_1 \cup V_2$ and all $f \in V_{R,\delta}(\Phi)$,

$$\frac{1}{s} \sum_{j=1}^s |f(x_j)|^2 \geq AR^{-d} \|f\|_2^2.$$

Theorem 15 combined with Lemma 13 imply that V_1 occurs with probability at most

$$D(\Lambda)R^{d^2/\alpha'}\beta^d \exp\left(-\frac{\nu^2s}{2D_2R^d(K_w + \frac{\nu}{3})}\right).$$

Furthermore, Lemma 18 yields that V_2 occurs with probability at most

$$\frac{R^d}{K_w} \left(\frac{e}{3}\right)^{3sR^{-d}K_w}.$$

Thus the lower estimate in (18) occurs at least with the probability given in (19), whereas the upper estimate follows Proposition 24 in Appendix. \square

Proof of Theorem 1 By the definitions of M and β in Table 1, we have $D(\Lambda)R^{d^2/\alpha'}\beta^d \leq D(\Lambda)R^{d^2/\alpha'}M$ as well as $\frac{R^d}{K_w} \leq D(\Lambda)R^{d^2/\alpha'}M$. And by the assumptions in (7), we have $(\frac{\epsilon}{3})^{3sR^{-d}K_w} \leq \exp(-\frac{\nu^2s}{2D_2R^d(K_w + \frac{\nu}{3})})$. Hence, as soon as

$$s \geq R^d \frac{2D_2(K_w + \frac{\nu}{3})}{\nu^2} \log\left(\frac{2D(\Lambda)MR^{d^2/\alpha'}}{\epsilon}\right)$$

holds we have that $D(\Lambda)R^{d^2/\alpha'}M \exp(-\frac{\nu^2s}{2D_2R^d(K_w + \frac{\nu}{3})}) \leq \epsilon/2$. The theorem is proved. \square

Proof of Corollary 4 and 5 Obviously $\ell = Sf$, therefore $S^*\ell = S^*Sf$. Now we estimate the spectrum of S^*S . For all $f \in V^2(\Lambda, \Phi)$, we can see that

$$\langle f, S^*Sf \rangle = \langle Sf, Sf \rangle = \|\ell\|_2^2 = \sum_{j=1}^s |f(x_j)|^2.$$

By Theorem 1, we have that

$$sAR^{-d}\langle f, f \rangle \leq \langle f, S^*Sf \rangle = \sum_{j=1}^s |f(x_j)|^2 \leq sD_2\langle f, f \rangle$$

hold under the conditions in Theorem 1 with probability at least $1 - \epsilon$. This implies that $\|S^*S\| \leq sD_2$, as well as S^*S is invertible and $\|(S^*S)^{-1}\| \leq (sAR^{-d})^{-1}$ in this case. Then Corollary 4 is proved.

In this case, the conditional number

$$\begin{aligned} \kappa(S^*S) &= \|S^*S\| \|(S^*S)^{-1}\| \\ &\leq sD_2(AsR^{-d})^{-1} \\ &= \frac{D_2R^d}{A} \\ &= \frac{D_2R^d}{K_w\rho(\frac{1}{2} - \delta) - \nu - 12K_wD_3\delta}. \end{aligned}$$

Let $\delta \rightarrow 0$, then Corollary 5 is proved. \square

Proof of Corollary 6 By the conclusions of Theorem 1, we have that $AR^{-d}\|f\|_2^2 \leq \frac{1}{s} \sum_{i=1}^s |f(x_i)|^2 \leq D_2\|f\|_2^2$ for all $f \in V_{R,\delta}^2(\Lambda, \Phi)$ hold with probability $1 - \epsilon$ under the conditions in Theorem 1. Therefore we have that

$$(1 - D_2)\|f\|_2^2 \leq \|f\|_2^2 - \frac{1}{s} \sum_{i=1}^s |f(x_i)|^2 \leq (1 - AR^{-d})\|f\|_2^2$$

in this case. The corollary is proved. □

7 Reconstruction Algorithm from Non-uniformly Random Sampling

Realistically one can sample f only on a bounded set; furthermore, every function vanishes at infinity, thus samples far out do not contribute anything significant to sampling and reconstruction. We can learn about f only if the samples are taken in the “essential support” of f . Let Λ is a finite set so we can write Φ as $\{\phi_1, \phi_2, \dots, \phi_i, \dots, \phi_n\}$ where $n = |\Lambda|$. Suppose the generators are bounded and the **support** of every ϕ_i is in C_R . So that for any small $\delta > 0$, we have $V_{R,\delta}^2(\Lambda, \Phi) = V^2(\Lambda, \Phi)$ since all the energy of these functions is all in C_R . In addition, we assume that Φ is a Riesz basis of $V^2(\Lambda, \Phi)$.

For a signal $f = \sum_{i=1}^n c_i \phi_i \in V^2(\Lambda, \Phi)$, denote

$$F = (f(x_1) \cdots f(x_j) \cdots f(x_s))$$

to be our sample value,

$$C = (c_1 \cdots c_i \cdots c_n)$$

to be the coefficients, and the matrix

$$U = (\phi_i(x_j))_{\substack{i=1,\dots,n \\ j=1,\dots,s}} \tag{20}$$

Then we can easily see that $F = CU$.

Note that Φ is a Riesz basis, we have that $c_i = \langle f, \tilde{\phi}_i \rangle$. So that by Proposition 22 in Appendix, we have

$$\|C\|_{\ell^2(\Lambda)}^2 \leq \|\tilde{\Phi}\|_{2,1,u_0}^2 \|f\|_2^2.$$

On the other hand, use the same method as Proposition 22 in Appendix, we have

$$\begin{aligned} \|f\|_2^2 &= \left\| \sum_{i=1}^n c_i \phi_i \right\|_2^2 \\ &= \sum_{k \in \mathbb{Z}^d} \left\| \sum_{i=1}^n c_i \phi_i \right\|_{L^2(k+[0,1]^d)}^2 \\ &\leq \sum_{k \in \mathbb{Z}^d} \left(\sum_{i=1}^n |c_i| \|\phi_i\|_{L^2(k+[0,1]^d)} \right)^2 \\ &\leq \|\Phi\|_{2,1,u_0}^2 \|C\|_{\ell^2(\Lambda)}^2. \end{aligned}$$

Table 2 Reconstruction algorithm

Algorithm:	Reconstruction from non-uniform random sampling
Given	$R, \Phi = (\phi_i)_{i=1}^n, w_1$
Input	1. ϵ, ν
Init	2. Choose s by (8)
	3. Generate random sampling point $\{x_1, \dots, x_s\}$ by w_R
	4. Calculate U by (20)
	5. Calculate $(UU^*)^{-1}$
Input	6. $F = (f(x_1), \dots, f(x_n))$
	7. Calculate $C = (FU^*)(UU^*)^{-1}$
Output	8. $f = \sum_{i=1}^n c_i \phi_i$

By Corollary 4, for ν small enough and s large enough, we have that

$$A_1 \|f\|_2^2 \leq \sum_{j=1}^n |f(x_j)|^2 \leq B_1 \|f\|_2^2$$

hold with probability at least $1 - \epsilon$ where $A_1 = \frac{s}{R^d} (\frac{K_w \rho}{2} - \nu)$ and $B_1 = s D_2$. Combining with the arguments above, we have that

$$A_2 \|C\|_{\ell^2(\Lambda)} \leq \|f\|_2^2 \leq B_2 \|C\|_{\ell^2(\Lambda)},$$

where $A_2 = \|\tilde{\Phi}\|_{2,1,u_0}^{-2}$ and $B_2 = \|\Phi\|_{2,1,u_0}^2$. Therefore

$$A_1 A_2 \|C\|_{\ell^2(\Lambda)} \leq \sum_{j=1}^n |f(x_j)|^2 \leq B_1 B_2 \|C\|_{\ell^2(\Lambda)}$$

hold with probability at least $1 - \epsilon$. Similar to the arguments which prove Corollary 4 and Corollary 5, we get that UU^* is invertible and the conditional number of UU^* which is denoted by $\kappa(UU^*)$ would not larger than $\frac{B_1 B_2}{A_1 A_2} = \frac{2R^d D_2 \|\Phi\|_{2,1,u_0}^2 \|\tilde{\Phi}\|_{2,1,u_0}^2}{K_w \rho - 2\nu}$ with probability at least $1 - \epsilon$. That is,

$$\mathbb{P}\left(\kappa(UU^*) > \frac{2R^d D_2 \|\Phi\|_{2,1,u_0}^2 \|\tilde{\Phi}\|_{2,1,u_0}^2}{K_w \rho - 2\nu}\right) \leq \epsilon.$$

Then one can reconstruct the original function from the sample by

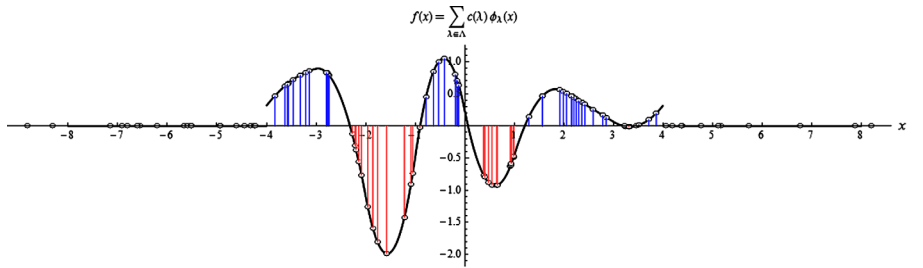
$$C = (FU^*)(UU^*)^{-1}$$

with probability with at least $1 - \epsilon$.

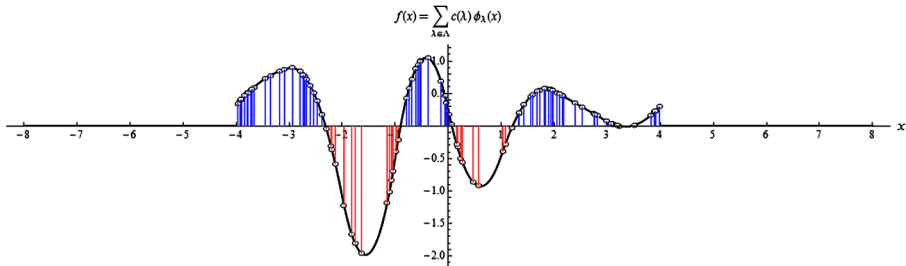
The reconstruction algorithm from non-uniform random sampling is shown in Table 2.

7.1 Numerical Test

Recall the spline functions mentioned in Example 8. Let us consider it as a numerical model for the generators in $V_{R,\delta}^2(\Lambda, \Phi)$. Let the spline knots be integer points and $m = 3$. We restrict the splines in C_R , where R is an even positive integer. We will review that space



(a) Random sampling with normal distribution in \mathbb{R}



(b) Random sampling with uniform distribution in C_R

Fig. 4 Different distributions of sampling points

by the following definition. Let $d = 1$ and R be an even positive integer. We use the spline function

$$\phi_0 = \chi_{[-\frac{1}{2}, \frac{1}{2}]} * \chi_{[-\frac{1}{2}, \frac{1}{2}]} * \chi_{[-\frac{1}{2}, \frac{1}{2}]},$$

where “*” represents the convolution product. The generators are presented by

$$\phi_\lambda(x) = \phi_0(x - \lambda)\chi_{C_R}(x), \quad \lambda \in \frac{1}{2} + \mathbb{Z}.$$

Let $\Lambda = \{\lambda : \phi_\lambda \neq 0\}$. We can easily check that $\Lambda = (\frac{1}{2} + \mathbb{Z}) \cap [-\frac{R}{2} - 1, \frac{R}{2} + 1]$ and $\lambda^* = 1/2$.

We will test two kinds of probability distributions in sampling. One is the normal distribution with mean 0 and variance σ like Example 10. In this case, we can get that the optimal choice of σ to optimize the lower bound of sampling inequality is $1/2$ taking $\delta \rightarrow 0$, so that $K_w = \sqrt{\frac{2}{\pi}}$ and $\rho = \exp(-\frac{1}{2})$. Note that we actually abandon those sampling points out of C_R in practice. Another distribution is the uniform distribution in C_R , just like that used in [3] and [8]. In this case, we have $K_w = 1$ and $\rho = 1$. Figure 4 plots the examples of different distributions of sampling points with $R = 8$.

Then we can give a totally practical trial to get the actual success rate. We use the algorithm in Table 2, using the numerical pseudo-inverse in Step 5, just like the “PseudoInverse” function in Wolfram Mathematica or “pinv” function in MATLAB. We take the space Φ as above while the distribution of sampling is normal distribution with a specific σ defined in Example 10. More specific, we take $c(\lambda)$ periodically of period 8. The signal is shown in Fig. 4 too.

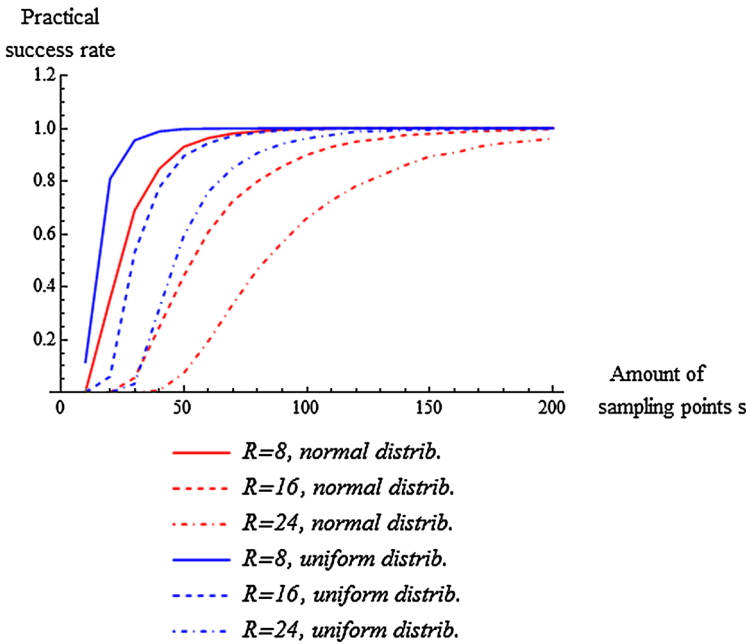


Fig. 5 The relation between success rate and sampling rate

For a test, we will run the algorithm and calculate the Euclidean distance between the C we calculated numerically in Step 7 and the real C of the original function f which is regarded as the reconstruction error. One test is regard as successful test when the reconstruction error is less than 10^{-9} .

Firstly, we take into two parts: one is using the normal distribution mentioned above with $\sigma = 1/2$, taking $R = 8, 16, 24$ and $s = 20, 30, \dots, 100$ respectively; another is using the uniform distribution in C_R taking the same parameters. We take 10000 trials for each case, then test the actual success rate in those cases. Figure 5 plots the success rate respectively to every value of s in each case.

From Fig. 5, we can see that: (i) the success rate increases with s increases; (ii) the success rate decreases with R increases and other conditions unchanged; (iii) due to the nonuniformness of the distribution, we need more sampling points or less R to reach the same success rate comparing with the uniform distribution. The results of the numerical tests match with Theorem 1 and Corollary 5.

Secondly, we fix $R = 8$, take $s = 40, 60$ as well as $\sigma = 1/80, 2/80, \dots, 120/80$ respectively and take 10000 trials for each case with the same normal distribution, then test the actual success rate in those cases. Figure 6 shows the success rate with the σ 's.

From Fig. 6, we can see that for a certain s , the success rate first increases then decreases with σ increasing. We get the best choice that optimize the actual success rate is that $\sigma = 40/80 = 1/2$ by the data of $s = 60$. At the same time, we can improve the success rate by increasing sampling points under the same σ . The results of the numerical tests match with Theorem 1 and Corollary 5.

Next, we explore the “threshold” of s , or the smallest s to have the actual sampling success rate reached 95% for different R 's. We fix $\sigma = 1/2$ and $R = 8, 16, 32, 64$ respectively, take different values of sampling rates around the threshold we pre-estimated while taking

Fig. 6 The relation between the practical success rate and σ

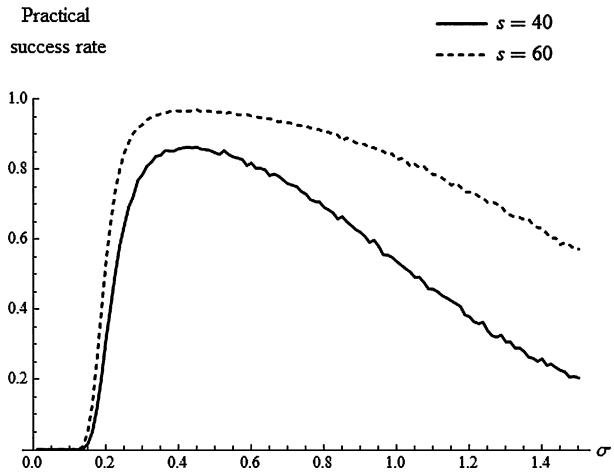


Table 3 Thresholds of s with different R 's

R	8	16	32	64
Thresholds of s	55	121	258	527

10000 trials for each case of sampling rate, and look for the “threshold” of s . Table 3 is the thresholds with those R 's.

We can see that every value of the threshold is a large twice times than the front value from Table 3. So the $O(R^d \log R^d)$ sampling rate in Theorem 1 is match with numerical test.

8 Conclusions

In this paper, we considered the non-uniform random sampling and reconstruction problem on signal spaces with finite rate of innovation. Under very mild assumptions on the generators and the probability distribution function, we show that for R sufficiently large, taking $O(R^d \log(R^d))$ many random samples yields a sampling set for $V_{R,\delta}^2(\Phi)$ with high probability. As one of applications of our main results, we provided a reconstruction algorithm from non-uniform random sampling with high probability and implemented it when the generators are compact-supported.

Acknowledgements The authors would like to thank the reviewers for their valuable comments and suggestions that led to the improvement of this paper. The corresponding author Jun Xian is partially supported by the National Natural Science Foundation of China (11422102, 11631015, 11871481); the Guangdong Provincial Government of China through the Computational Science Innovative Research Team program, China; and the Guangdong Province Key Laboratory of Computational Science, China.

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

Appendix. Some Properties of the Signal Spaces $V^2(\Lambda, \Phi)$

We assume that $\Phi := \{\phi_\lambda\}_{\lambda \in \Lambda}$ is a Riesz basis of $V^2(\Lambda, \Phi)$, therefore there exists a dual Riesz basis $\tilde{\Phi} = \{\tilde{\phi}_\lambda\}_{\lambda \in \Lambda}$. Thus we have that for all $f \in V^2(\Lambda, \Phi)$, such that

$$f = \sum_{\lambda \in \Lambda} \langle f, \tilde{\phi}_\lambda \rangle \phi_\lambda = \sum_{\lambda \in \Lambda} \langle f, \phi_\lambda \rangle \tilde{\phi}_\lambda. \tag{21}$$

We can easily see that for all $1 \leq q' \leq q \leq \infty, \infty \geq p' \geq p \geq 1$ and $u'(x) \leq u(x), \forall x \in \mathbb{R}^d$, we have $\|\Phi\|_{q', p', u'} \leq \|\Phi\|_{q, p, u}$. Moreover, by [19, Theorem 4.1], we have that if $\|\Phi\|_{q, p, u} < \infty$, then so is the same norm of the dual Riesz basis, or to say $\|\tilde{\Phi}\|_{q, p, u} < \infty$.

For convenience of the main proof, we will show some properties of the space $V^2(\Lambda, \Phi)$ under the assumptions of the generators in Sect. 2.2.

Proposition 22 *If $f \in L^2(\mathbb{R}^d)$, $\Phi = \{\phi_\lambda\}$ satisfying (A.0) and $\|\Phi\|_{2,1,u_0} < \infty$, then we have:*

$$\sum_{\lambda \in \Lambda} |\langle f, \phi_\lambda \rangle|^2 \leq \|\Phi\|_{2,1,u_0}^2 \|f\|_{L^2(\mathbb{R}^d)}^2 \tag{22}$$

and

$$\sum_{\lambda \in \Lambda} |\langle f, \tilde{\phi}_\lambda \rangle|^2 \leq \|\tilde{\Phi}\|_{2,1,u_0}^2 \|f\|_{L^2(\mathbb{R}^d)}^2. \tag{23}$$

Proof We firstly prove (22). In fact, we have

$$\begin{aligned} \sum_{\lambda \in \Lambda} |\langle f, \phi_\lambda \rangle|^2 &= \sum_{\lambda \in \Lambda} \left| \sum_{k \in \mathbb{Z}^d} \int_{k+[0,1]^d} f(x) \phi_\lambda(x) dx \right|^2 \\ &\leq \sum_{\lambda \in \Lambda} \left(\sum_{k \in \mathbb{Z}^d} \int_{k+[0,1]^d} |f(x)| |\phi_\lambda(x)| dx \right)^2 \\ &\leq \sum_{\lambda \in \Lambda} \left(\sum_{k \in \mathbb{Z}^d} \|f(x)\|_{L^2(k+[0,1]^d)} \|\phi_\lambda(x)\|_{L^2(k+[0,1]^d)} dx \right)^2 \\ &= \sum_{\lambda \in \Lambda} \left(\sum_{k \in \mathbb{Z}^d} \|\phi_\lambda(x)\|_{L^2(k+[0,1]^d)}^{1/2} \times \|\phi_\lambda(x)\|_{L^2(k+[0,1]^d)}^{1/2} \|f(x)\|_{L^2(k+[0,1]^d)} \right)^2 \\ &\leq \sum_{\lambda \in \Lambda} \left(\sum_{k \in \mathbb{Z}^d} \|\phi_\lambda(x)\|_{L^2(k+[0,1]^d)} \right) \left(\sum_{k \in \mathbb{Z}^d} \|\phi_\lambda(x)\|_{L^2(k+[0,1]^d)} \|f(x)\|_{L^2(k+[0,1]^d)}^2 \right) \\ &\leq \|\Phi\|_{2,1,u_0} \sum_{\lambda \in \Lambda} \sum_{k \in \mathbb{Z}^d} \|\phi_\lambda(x)\|_{L^2(k+[0,1]^d)} \|f(x)\|_{L^2(k+[0,1]^d)}^2 \\ &= \|\Phi\|_{2,1,u_0} \sum_{k \in \mathbb{Z}^d} \|f(x)\|_{L^2(k+[0,1]^d)}^2 \sum_{\lambda \in \Lambda} \|\phi_\lambda(x)\|_{L^2(k+[0,1]^d)} \\ &\leq \|\Phi\|_{2,1,u_0}^2 \sum_{k \in \mathbb{Z}^d} \|f(x)\|_{L^2(k+[0,1]^d)}^2 = \|\Phi\|_{2,1,u_0}^2 \|f\|_{L^2(\mathbb{R}^d)}^2. \end{aligned}$$

A similar method is used in the proof of [18, Theorem 2.4(iii)]. Similar to the proof above, we can prove (23) as well. □

Proposition 23 For every $f \in V^2(\Lambda, \Phi)$ and every subset $\Gamma \subset \mathbb{R}^d$ where the absolute covering index $N_0(\Gamma)$ defined in (3) is finite, we have

$$\sum_{\gamma \in \Gamma} |f(\gamma)|^2 \leq D_3 N_0(\Gamma) \|f\|_2^2. \tag{24}$$

Proof First we consider Γ with $N_0(\Gamma) = 1$, that is, for any $k \in \mathbb{Z}^d$, there is at most one point in $\Gamma \cap (k + [0, 1]^d)$. Let $c(\lambda) = \langle f, \tilde{\phi}_\lambda \rangle$, then $f = \sum_{\lambda \in \Lambda} c(\lambda) \phi_\lambda$. Similar to the proof of Proposition 22, we have

$$\begin{aligned} \sum_{\gamma \in \Gamma} |f(\gamma)|^2 &= \sum_{\gamma \in \Gamma} \left| \sum_{\lambda \in \Lambda} c(\lambda) \phi_\lambda(\gamma) \right|^2 \\ &\leq \sum_{\gamma \in \Gamma} \left(\sum_{\lambda \in \Lambda} |c(\lambda)| |\phi_\lambda(\gamma)| \right)^2 \\ &\leq \sum_{k \in \mathbb{Z}^d} \left(\sum_{\lambda \in \Lambda} |c(\lambda)| \|\phi_\lambda\|_{L^\infty(k+[0,1]^d)} \right)^2 \\ &= \sum_{k \in \mathbb{Z}^d} \left(\sum_{\lambda \in \Lambda} \|\phi_\lambda\|_{L^\infty(k+[0,1]^d)}^{1/2} \times \|\phi_\lambda\|_{L^\infty(k+[0,1]^d)}^{1/2} |c(\lambda)| \right)^2 \\ &\leq \sum_{k \in \mathbb{Z}^d} \left(\sum_{\lambda \in \Lambda} \|\phi_\lambda\|_{L^\infty(k+[0,1]^d)} \right) \left(\sum_{\lambda \in \Lambda} \|\phi_\lambda\|_{L^\infty(k+[0,1]^d)} |c(\lambda)|^2 \right) \\ &\leq \|\Phi\|_{\infty,1,u_0} \sum_{k \in \mathbb{Z}^d} \sum_{\lambda \in \Lambda} \|\phi_\lambda\|_{L^\infty(k+[0,1]^d)} |c(\lambda)|^2 \\ &\leq \|\Phi\|_{\infty,1,u_0}^2 \sum_{\lambda \in \Lambda} |c(\lambda)|^2 \\ &\leq \|\Phi\|_{\infty,1,u_0}^2 \|\tilde{\Phi}\|_{2,1,u_0}^2 \|f\|_2^2 = D_3 \|f\|_2^2, \end{aligned}$$

where the last inequality is derived from the conclusion of Proposition 22.

For a general subset Γ , we can split it into at most $N_0(\Gamma)$ non-intersect parts each of which has absolute covering index 1. Then sum the function value of each part respectively, sum all of them all together and we get the final conclusion. \square

Proposition 24 $V^2(\Lambda, \Phi)$ is a reproducing kernel space, or to say there exists a family $(v_x)_{x \in \mathbb{R}^d} \subset V^2(\Lambda, \Phi)$ such that $f(x) = \langle f, v_x \rangle, \forall f \in V^2(\Lambda, \Phi)$ and $\|v_x\|_2^2 \leq D_2$ almost everywhere. Moreover, we have $\|f\|_\infty^2 \leq D_2 \|f\|_2^2$ for all $f \in V^2(\Lambda, \Phi)$.

Proof We define that $v_x(y) = \sum_{\lambda \in \Lambda} \tilde{\phi}_\lambda(y) \overline{\phi_\lambda(x)}$, and $w_x(y) = \sum_{\lambda \in \Lambda} \phi_\lambda(y) \overline{\tilde{\phi}_\lambda(x)}$. For all $f \in V^2(\Lambda, \Phi)$, we have that

$$\begin{aligned} f(x) &= \sum_{\lambda \in \Lambda} \langle f, \tilde{\phi}_\lambda \rangle \phi_\lambda(x) \\ &= \sum_{\lambda \in \Lambda} \int_{\mathbb{R}^d} f(y) \overline{\tilde{\phi}_\lambda(y)} dx \phi_\lambda(x) \\ &= \int_{\mathbb{R}^d} f(y) \overline{v_x(y)} dy = \langle f, v_x \rangle. \end{aligned}$$

Similarly, we have

$$\begin{aligned} f(x) &= \sum_{\lambda \in \Lambda} \langle f, \phi_\lambda \rangle \tilde{\phi}_\lambda(x) \\ &= \int_{\mathbb{R}^d} f(y) \overline{w_x(y)} dy = \langle f, w_x \rangle. \end{aligned}$$

Therefore, $\langle f, w_x - v_x \rangle = 0$ holds for all $f \in V^2(\Lambda, \Phi)$. However, both v_x and w_x are in $V^2(\Lambda, \Phi)$ by definition, then $w_x - v_x \in V^2(\Lambda, \Phi)$. So $\langle w_x - v_x, w_x - v_x \rangle = 0$, or $w_x - v_x = 0$. Therefore $w_x = v_x$, which implies $v_x(y) = \sum_{\lambda \in \Lambda} \phi_\lambda(y) \overline{\phi_\lambda(x)}$ as well.

We now show that $v(x)_{x \in \mathbb{R}^d}$ is a reproducing kernel. For fixed $x \in \mathbb{R}^d$, we have

$$\begin{aligned} \|v_x\|_2^2 &= \int_{\mathbb{R}^d} v_x(y) \overline{v_x(y)} dy \\ &= \int_{\mathbb{R}^d} \sum_{\lambda \in \Lambda} \sum_{\lambda' \in \Lambda} \tilde{\phi}_\lambda(y) \overline{\phi_\lambda(x)} \overline{\phi_{\lambda'}(y)} \tilde{\phi}_{\lambda'}(x) dy. \end{aligned}$$

By $\|\Phi\|_{\infty,1,u_0} < \infty$, we can change the order of the sum and the integration. So we can obtain that

$$\begin{aligned} \|v_x\|_2^2 &= \int_{\mathbb{R}^d} \sum_{\lambda \in \Lambda} \sum_{\lambda' \in \Lambda} \tilde{\phi}_\lambda(y) \overline{\phi_\lambda(x)} \overline{\phi_{\lambda'}(y)} \tilde{\phi}_{\lambda'}(x) dy \\ &= \sum_{\lambda \in \Lambda} \sum_{\lambda' \in \Lambda} \tilde{\phi}_{\lambda'}(x) \overline{\phi_\lambda(x)} \int_{\mathbb{R}^d} \tilde{\phi}_\lambda(y) \overline{\phi_{\lambda'}(y)} dy \\ &= \sum_{\lambda \in \Lambda} \sum_{\lambda' \in \Lambda} \tilde{\phi}_{\lambda'}(x) \overline{\phi_\lambda(x)} \langle \tilde{\phi}_\lambda, \phi_{\lambda'} \rangle \\ &= \sum_{\lambda \in \Lambda} \tilde{\phi}_\lambda(x) \overline{\phi_\lambda(x)} \\ &\leq \left(\sum_{\lambda \in \Lambda} |\tilde{\phi}_\lambda(x)|^2 \right)^{\frac{1}{2}} \left(\sum_{\lambda \in \Lambda} |\phi_\lambda(x)|^2 \right)^{\frac{1}{2}} \\ &\leq \|\Phi\|_{\infty,2,u_0} \|\tilde{\Phi}\|_{\infty,2,u_0} = D_2. \end{aligned}$$

Thus $\text{ess sup}_{x \in \mathbb{R}^d} \|v_x\|_2^2 \leq D_2 < \infty$. Therefore $(v_x)_{x \in \mathbb{R}^d}$ is a reproducing kernel of $V^2(\Phi, \Lambda)$. Moreover, we have

$$\begin{aligned} \|f\|_\infty^2 &= \text{ess sup}_{x \in \mathbb{R}^d} |f(x)|^2 \\ &= \text{ess sup}_{x \in \mathbb{R}^d} |\langle f, v_x \rangle|^2 \\ &\leq \text{ess sup}_{x \in \mathbb{R}^d} \|f\|_2^2 \|v_x\|_2^2 \\ &\leq D_2 \|f\|_2^2. \quad \square \end{aligned}$$

References

1. Aldroubi, A., Gröchenig, K.: Nonuniform sampling and reconstruction in shift-invariant spaces. *SIAM Rev.* **43**(4), 585–620 (2001)

2. Baechler, G., Scholefield, A., Baboulaz, L., Vetterli, M.: Sampling and exact reconstruction of pulses with variable width. *IEEE Trans. Signal Process.* **65**(10), 2629–2644 (2017)
3. Bass, R.F., Gröchenig, K.: Relevant sampling of band-limited functions. III. *J. Math.* **57**(1), 43–58 (2013)
4. Bernardi, O., Giménez, O.: A linear algorithm for the random sampling from regular languages. *Algoritmica* **62**(1–2), 130–145 (2012)
5. Bi, N., Nashed, M.Z., Sun, Q.Y.: Reconstructing signals with finite rate of innovation from noisy samples. *Acta Appl. Math.* **107**(1–3), 339–372 (2009)
6. Fang, K.T., Ma, C.X., Winker, P.: Centered L_2 -discrepancy of random sampling and Latin hypercube design, and construction of uniform designs. *Math. Comput.* **71**(237), 275–296 (2002)
7. Führ, H., Xian, J.: Quantifying invariance properties of shift-invariant spaces. *Appl. Comput. Harmon. Anal.* **36**, 514–521 (2014)
8. Führ, H., Xian, J.: Relevant sampling in finitely generated shift-invariant spaces. *J. Approx. Theory* **240**, 1–15 (2019)
9. Gröchenig, K., Schwab, H.: Fast local reconstruction methods for nonuniform sampling in shift-invariant spaces. *SIAM J. Matrix Anal. Appl.* **24**(4), 899–913 (2003)
10. Li, Y.X., Wen, J.M., Xian, J.: Reconstruction from convolution random sampling in local shift invariant spaces. *Inverse Problems* (2019). <https://iopscience.iop.org/article/10.1088/1361-6420/ab40f7/meta>
11. Moskowitz, B., Caffisch, R.E.: Smoothness and dimension reduction in quasi-Monte Carlo methods. *Math. Comput. Model.* **23**(8), 37–54 (1996)
12. Mulletti, S., Seelamantula, C.S.: Ellipse fitting using the finite rate of innovation sampling principle. *IEEE Trans. Image Process.* **25**(3), 1451–1464 (2016)
13. Pederghana, M., García, S.G.: Smart sampling and incremental function learning for very large high dimensional data. *Neural Netw.* **78**, 75–87 (2016)
14. Romero, V.J., Burkardt, J.V., Gunzburger, M.D., Peterson, J.S.: Comparison of pure and “Latinized” centroidal Voronoi tessellation against various other statistical sampling methods. *Reliab. Eng. Syst. Saf.* **91**(10–11), 1266–1280 (2006)
15. Rudresh, S., Seelamantula, C.S.: Finite-rate-of-innovation-based super-resolution radar imaging. *IEEE Trans. Signal Process.* **65**(19), 5021–5033 (2017)
16. Schumaker, L.L.: *Spline Functions: Basic Theory*. John Wiley & Sons, New York (1981)
17. Sun, Q.Y.: Nonuniform average sampling and reconstruction of signals with finite rate of innovation. *SIAM J. Math. Anal.* **38**(5), 1389–1422 (2006)
18. Sun, Q.Y.: Wiener’s lemma for infinite matrices. *Trans. Am. Math. Soc.* **359**(7), 3099–3123 (2007)
19. Sun, Q.Y.: Frames in spaces with finite rate of innovation. *Adv. Comput. Math.* **28**(4), 301–329 (2008)
20. Sun, W.C.: Local sampling theorems for spaces generated by splines with arbitrary knots. *Math. Comput.* **78**(265), 225–239 (2009)
21. Sun, W.C., Zhou, X.W.: Characterization of local sampling sequences for spline subspaces. *Adv. Comput. Math.* **30**(2), 153–175 (2009)
22. Tropp, J.A.: User-friendly tail bounds for sums of random matrices. *Found. Comput. Math.* **12**(4), 389–434 (2012)
23. Vetterli, M., Marziliano, P., Blu, T.: Sampling signals with finite rate of innovation. *IEEE Trans. Signal Process.* **50**(6), 1417–1428 (2002)
24. 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)
25. Yang, J.B.: Random sampling and reconstruction in multiply generated shift-invariant spaces. *Anal. Appl.* **17**(2) 323–347 (2019)