Canad. Math. Bull. Vol. **54** (3), 2011 pp. 566–576
doi:10.4153/CMB-2011-029-8

Non-uniform Randomized Sampling for Multivariate Approximation by High Order Parzen Windows

Xiang-Jun Zhou, Lei Shi, and Ding-Xuan Zhou

Abstract. We consider approximation of multivariate functions in Sobolev spaces by high order Parzen windows in a non-uniform sampling setting. Sampling points are neither i.i.d. nor regular, but are noised from regular grids by non-uniform shifts of a probability density function. Sample function values at sampling points are drawn according to probability measures with expected values being values of the approximated function. The approximation orders are estimated by means of regularity of the approximated function, the density function, and the order of the Parzen windows, under suitable choices of the scaling parameter.

1 Introduction and Formal Setting

We will consider approximation of functions on $\mathbb{R}^n$ from samples of type $\mathbf{z} = \{(x_i, y_i)\}_{i\in\mathbf{Z}^n}$. If f^* is a function to be approximated, then in our setting described below, the sample function value y_i at the sampling point x_i satisfies $y_i \approx f^*(x_i)$. In randomized sampling, the sampling points $\{x_i\}$ are governed by some probability distributions.

In this paper we continue our study [11] on randomized sampling of functions on $\mathbb{R}^n$ by means of high order Parzen windows

$$f_{\mathbf{z},\sigma}(x) = \sum_{i\in\mathbb{Z}^n} y_i \Phi\left(\frac{x}{\sigma}, \frac{x_i}{\sigma}\right), \qquad x \in \mathbb{R}^n.$$

Here $\Phi\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a window function and $\sigma > 0$ is a window width. Under some conditions on the approximated function f^*, window function Φ, and noise controlling $y_i - f^*(x_i)$, the error in $L^2(\mathbb{R}^n)$ between f^* and $f_{\mathbf{z},\sigma}$ (with normalization) is analyzed in [11] when $\{x_i\}_{i\in\mathbb{Z}^n}$ is drawn randomly from a sequence of probability densities $\{p(\cdot - hi)\}_{i\in\mathbb{Z}^n}$ for some fixed density function p on $\mathbb{R}^n$ and some grid size $h > 0$. In that setting, $\{x_i\}$ is uniform in expectation: $\mathbf{E}(x_i) = hi + \mathbf{E}(p)$ for each $i \in \mathbb{Z}^n$.

The purpose of this paper is to establish improvements of the above result in two directions. First, we consider a non-uniform setting in the sense that the sampling

Received by the editors September 11, 2008; revised February 12, 2009.
Published electronically March 5, 2011.
The work described in this paper is partially supported by the City University of Hong Kong [Project No. 7002126], the National Science Fund for Distinguished Young Scholars of China [Project No. 10529101], and the National Basic Research Program of China [Project No. 973-2006CB303102].
AMS subject classification: **68T05**, 62J02.
Keywords: multivariate approximation, Sobolev spaces, non-uniform randomized sampling, high order Parzen windows, convergence rates.

points $\{x_i\}_{i\in\mathbb{Z}^n}$ are drawn according to the probability density functions $\{p(\cdot - t_i)\}$ with non-uniform nodes $\{t_i\}$ satisfying $\sup_{i\in\mathbb{Z}^n}|t_i - hi| \le \Delta$. The quantity $\Delta > 0$ measures the degree of non-uniformality as in the literature of non-uniform sampling [1]. Secondly, we estimate bounds for the error in Sobolev spaces $H^s(\mathbb{R}^n)$ by allowing $s \ge 0$, not only in the space $L^2(\mathbb{R}^n)$ with $s = 0$. To this end, the basic window function Φ is defined in terms of the index s. Throughout this paper, $0 \le s < J$ are two integers. For $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}_+^n$ and $x = (x^1, \dots, x^n) \in \mathbb{R}^n$, we denote $x^\alpha = \Pi_{j=1}^n (x^j)^{\alpha_j}$, $|\alpha| = \alpha_1 + \cdots + \alpha_n$, and $D^\alpha\Phi(x,u) = \frac{\partial^\alpha}{\partial x^\alpha}(\Phi(\cdot, u))(x)$.

Definition 1.1 A function $\Phi\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is called a *basic window function* of type (J, s) if it satisfies

(i) $\int_{\mathbb{R}^n} \Phi(x,u)\,du \equiv 1$ and $\int_{\mathbb{R}^n} \Phi(x,u)(u-x)^\alpha\,du \equiv 0$ for $0 < |\alpha| < J$;
(ii) for some $q > n + J$ and $c_q > 0$,

$$|D^\alpha\Phi(x,u)| \le \frac{c_q}{(1+|x-u|)^q} \qquad \forall x, u \in \mathbb{R}^n, 0 \le |\alpha| \le s. \tag{1.1}$$

Condition (i) above is called the vanishing moment condition in the literature of multivariate approximation [3] or wavelets [4].

As in the literature of Shannon sampling [8] or online learning [9], we shall assume throughout this paper that the sample $\mathbf{z} = \{(x_i, y_i)\}_{i\in\mathbb{Z}^n}$ is drawn independently from a sampling sequence $\{\rho^{(i)}\}_{i\in\mathbb{Z}^n}$ of probability measures on $Z := \mathbb{R}^n \times \mathbb{R}$ associated with (M, h, Δ, p).

Definition 1.2 Let $M, \Delta > 0$, $0 < h \le 1$ and p be a probability density function on $\mathbb{R}^n$. Each $x \in \mathbb{R}^n$ is assigned a Borel probability measure ρ_x supported on $[-M, M]$, and $\{t_i\}_{i\in\mathbb{Z}^n} \subset \mathbb{R}^n$ is a sequence satisfying $|t_i - hi| \le \Delta$ for each $i \in \mathbb{Z}^n$. We call $\{\rho^{(i)}\}_{i\in\mathbb{Z}^n}$ a *sampling sequence* associated with (M, h, Δ, p) if for each i, the marginal distribution $\rho_X^{(i)}$ of $\rho^{(i)}$ on $\mathbb{R}^n$ has density $p(\cdot - t_i)$ and the conditional distribution of $\rho^{(i)}$ at each $x \in \mathbb{R}^n$ equals ρ_x.

Define a function f^* to be approximated by

$$f^*(x) = \int_{\mathbb{R}} y\,d\rho_x, \qquad x \in \mathbb{R}^n.$$

2 Main Result

Our main result provides bounds for the approximation of the function f^* on $\mathbb{R}^n$ by (normalized) $f_{\mathbf{z},\sigma}$ in the Sobolev space $H^s(\mathbb{R}^n)$ with norm

$$\|f\|_{H^s(\mathbb{R}^n)} = \sum_{|\alpha|\le s} \|D^\alpha f\|_{L^2(\mathbb{R}^n)}.$$

Denote the variance of ρ_x by σ_x^2.

Theorem 2.1 *Let Φ be a basic window function of type (J, s). Assume $f^* \in C^{J+s}(\mathbb{R}^n)$, $p \in C^{J+s}(\mathbb{R}^n)$, and that for some $\eta > 2n$ and $c_\eta > 0$, each of the functions $D^\alpha f^*, D^\alpha p$ with $|\alpha| \le J + s$ and σ_x^2 satisfies the decay condition*

$$|f(x)| \le \frac{c_\eta}{(1+|x|)^\eta} \qquad \forall\, x \in \mathbb{R}^n. \tag{2.1}$$

If $\Delta \le h^{\frac{n(J-s)}{n+2J}}$, then by taking $\sigma = h^{\frac{n}{n+2J}}$, for any $0 < \delta < 1$, with confidence $1-\delta$, we have

$$\left\| h^{\frac{2nJ}{n+2J}} f_{\mathbf{z},\sigma} - f^* \right\|_{H^s(\mathbb{R}^n)} \le \widetilde{C}_{q,n,\eta,J,s} \left\{ (1+M)c_q + c_q c_\eta^2 + c_\eta \|f^*\|_{H^s(\mathbb{R}^n)} \right\} h^{\frac{n(J-s)}{n+2J}} \log \frac{2}{\delta},$$

where $\widetilde{C}_{q,n,\eta,J,s}$ is the constant depending only on q, n, η, J, s.

Theorem 2.1 will be proved in Section 4 with the constant $\widetilde{C}_{q,n,\eta,J,s}$ given explicitly.

The idea of approximation in Sobolev spaces can be used to learn norms of gradients $\frac{\partial f^*}{\partial x^i}$ for variable selection and inner products $\langle \frac{\partial f^*}{\partial x^i}, \frac{\partial f^*}{\partial x^j} \rangle$ for studying covariances [2, 5]. We shall discuss details somewhere else.

It would be interesting to study the randomized sampling for dependent samples [10].

3 Sample Error and Approximation Error

The sample error refers to the difference between $f_{\mathbf{z},\sigma}$ and its data-free limit defined by

$$f_\sigma(x) = \int_{\mathbb{R}^n} \Phi\left(\frac{x}{\sigma}, \frac{u}{\sigma}\right) f^*(u) \sum_{i\in\mathbb{Z}^n} p(u - t_i)\,du, \qquad x \in \mathbb{R}^n. \tag{3.1}$$

It can be bounded by the following probability inequality for random variables with values in a Hilbert space (see [7] or [6]).

Lemma 3.1 *Let $\mathcal{H}$ be a Hilbert space, and let $\{\xi_i\}_{i\in\mathbb{Z}^n}$ be independent random variables on Z with values in $\mathcal{H}$. Assume that for each i, $\|\xi_i\| \le \widetilde{M} < \infty$ almost surely. Denote $\widetilde{\sigma}^2 = \sum_{i\in\mathbb{Z}^n} \mathbf{E}(\|\xi_i\|^2)$. Then for any $0 < \delta < 1$, with confidence $1-\delta$,*

$$\Big\| \sum_{i\in\mathbb{Z}^n} [\xi_i - \mathbf{E}(\xi_i)] \Big\| \le 2\widetilde{M}\log(2/\delta) + \sqrt{2\widetilde{\sigma}^2\log(2/\delta)}.$$

In our setting, we take $\{\xi_i\}_{i\in\mathbb{Z}^n}$ to be the random variable on Z with values in the Hilbert space $H^s(\mathbb{R}^n)$ given by

$$\begin{aligned}\xi_i &= y_i\Phi\left(\frac{\cdot}{\sigma}, \frac{x_i}{\sigma}\right) - \int_Z y\Phi\left(\frac{\cdot}{\sigma}, \frac{x}{\sigma}\right) d\rho^{(i)} \\ &= y_i\Phi\left(\frac{\cdot}{\sigma}, \frac{x_i}{\sigma}\right) - \int_{\mathbb{R}^n} \Phi\left(\frac{\cdot}{\sigma}, \frac{x}{\sigma}\right) p(x - t_i) f^*(x)\,dx.\end{aligned}$$

Then we can apply Lemma 3.1 to obtain bounds for sample error. To this end, we need the Sobolev space norm of the function $\Phi\left(\frac{\cdot}{\sigma}, \frac{x}{\sigma}\right)$.

Lemma 3.2 *Let $0 < \sigma \le 1$. If (1.1) holds, then for any $x \in \mathbb{R}^n$, we have*

$$\left\|\Phi\left(\frac{\cdot}{\sigma}, \frac{x}{\sigma}\right)\right\|_{H^s(\mathbb{R}^n)} \le \frac{2c_q(s+1)^n \pi^{n/4}}{\sqrt{(2q-n)\Gamma(n/2)}}\sigma^{\frac{n}{2}-s},$$

where $\Gamma(t)$ is the Gamma function defined by $\Gamma(t) = \int_0^\infty r^{t-1}e^{-r}dr$ for $t > 0$.

Proof For any $\alpha \in \mathbb{Z}_+^n$ with $0 \le |\alpha| \le s$, we have

$$\left\|D^\alpha\left(\Phi\left(\frac{\cdot}{\sigma}, \frac{x}{\sigma}\right)\right)\right\|^2_{L^2(\mathbb{R}^n)} = \left\|\sigma^{-|\alpha|}D^\alpha\Phi\left(\frac{\cdot}{\sigma}, \frac{x}{\sigma}\right)\right\|^2_{L^2(\mathbb{R}^n)}$$
$$\le \sigma^{-2|\alpha|}\int_{\mathbb{R}^n}\frac{c_q^2}{\left(1+\frac{|u-x|}{\sigma}\right)^{2q}}du \le \sigma^{n-2|\alpha|}\frac{2\pi^{n/2}c_q^2}{(2q-n)\Gamma(n/2)}.$$

Since $0 < \sigma \le 1$, we have $\sigma^{-2|\alpha|} \le \sigma^{-2s}$, and the desired bound follows. ■

Proposition 3.3 *Let Φ be a basic window function of type (J, s), and let $\mathbf{z}$ be a sample. For any $0 < \delta < 1$, with confidence $1 - \delta$, we have*

$$\|f_{\mathbf{z},\sigma} - f_\sigma\|_{H^s(\mathbb{R}^n)} \le$$
$$\frac{8c_q(s+1)^n\pi^{n/4}\sigma^{\frac{n}{2}-s}}{\sqrt{(2q-n)\Gamma(n/2)}}\log\frac{2}{\delta}\left\{M + \left(\int_{\mathbb{R}^n}\left[\sigma_x^2 + 2|f^*(x)|^2\right]\sum_{i\in\mathbb{Z}^n}p(x-t_i)dx\right)^{\frac{1}{2}}\right\}.$$

Proof Since ρ_x is supported on $[-M, M]$, $|f^*(x)| \le M$ for each x. By Lemma 3.2, the random variable $\xi_i = y_i\Phi\left(\frac{\cdot}{\sigma}, \frac{x_i}{\sigma}\right) - \int_{\mathbb{R}^n}\Phi\left(\frac{\cdot}{\sigma}, \frac{x}{\sigma}\right)p(x-t_i)f^*(x)dx$ with values in $H^s(\mathbb{R}^n)$ satisfies

$$\|\xi_i\|_{H^s(\mathbb{R}^n)} \le M\left\|\Phi\left(\frac{\cdot}{\sigma}, \frac{x_i}{\sigma}\right)\right\|_{H^s(\mathbb{R}^n)} + \int_{\mathbb{R}^n}\left\|\Phi\left(\frac{\cdot}{\sigma}, \frac{x}{\sigma}\right)\right\|_{H^s(\mathbb{R}^n)}p(x-t_i)|f^*(x)|dx$$
$$\le \frac{4Mc_q(s+1)^n\pi^{n/4}}{\sqrt{(2q-n)\Gamma(n/2)}}\sigma^{\frac{n}{2}-s}.$$

Write ξ_i as

$$\xi_i = \big(y_i - f^*(x_i)\big)\Phi\left(\frac{\cdot}{\sigma}, \frac{x_i}{\sigma}\right) + f^*(x_i)\Phi\left(\frac{\cdot}{\sigma}, \frac{x_i}{\sigma}\right) - \int_{\mathbb{R}^n}\Phi\left(\frac{\cdot}{\sigma}, \frac{x}{\sigma}\right)p(x-t_i)f^*(x)dx.$$

Then we see from Lemma 3.2 that for each $i \in \mathbb{Z}^n$,

$$\|\xi_i\|_{H^s(\mathbb{R}^n)} \le$$
$$\left\{|y_i - f^*(x_i)| + |f^*(x_i)| + \int_{\mathbb{R}^n}|f^*(x)|p(x-t_i)dx\right\}\frac{2c_q(s+1)^n\pi^{n/4}}{\sqrt{(2q-n)\Gamma(n/2)}}\sigma^{\frac{n}{2}-s}.$$

https://doi.org/10.4153/CMB-2011-029-8 Published online by Cambridge University Press

It follows from $\left(\int_{\mathbb{R}^n} |f^*(x)| p(x-t_i)dx\right)^2 \leq \int_{\mathbb{R}^n} |f^*(x)|^2 p(x-t_i)dx$ that

$$\sum_{i\in\mathbb{Z}^n} \mathbf{E}\left(\|\xi_i\|^2_{H^s(\mathbb{R}^n)}\right) \leq \frac{12c_q^2(s+1)^{2n}\pi^{n/2}}{(2q-n)\Gamma(n/2)} \left\{ \int_{\mathbb{R}^n} \left[\sigma_x^2 + 2|f^*(x)|^2\right] \sum_{i\in\mathbb{Z}^n} p(x-t_i)dx \right\} \sigma^{n-2s}.$$

Then our conclusion follows from Lemma 3.1. ■

Now we turn to the approximation error. We show that in the space $H^s(\mathbb{R}^n)$, f_σ tends to $f^*\sum_{i\in\mathbb{Z}^n} p(\cdot - t_i)$ as σ becomes small.

Proposition 3.4 *Let Φ be a basic window function of type (J,s) defined as Definition 1.1. Define f_σ by (3.1). If the function $f^*\sum_{i\in\mathbb{Z}^n} p(\cdot - t_i)$ lies in $H^{J+s}(\mathbb{R}^n)$, then for any $\sigma \leq 1$, we have*

$$\left\| f_\sigma - \sigma^n f^* \sum_{i\in\mathbb{Z}^n} p(\cdot - t_i)\right\|_{H^s(\mathbb{R}^n)} \leq \frac{2\cdot 3^{J+s} c_q \pi^{n/2} s!}{(q-J-n)\Gamma(n/2)J!} \left\| f^* \sum_{i\in\mathbb{Z}^n} p(\cdot - t_i)\right\|_{H^{J+s}(\mathbb{R}^n)} \sigma^{n+J-s}. \tag{3.2}$$

Proof Denote $g(x) = f^*(x)\sum_{i\in\mathbb{Z}^n} p(x-t_i)$. We apply a Taylor expansion to the function

$$f_\sigma(x) = \int_{\mathbb{R}^n} \Phi\left(\frac{x}{\sigma}, \frac{u}{\sigma}\right) g(u)du.$$

Let $u, x \in \mathbb{R}^n$. Define a univariate function $h : [0,1] \to \mathbb{R}$ as

$$h(t) = g(x + t(u-x)), \qquad t \in [0,1].$$

Then we have

$$g(u) = h(1) = \sum_{\ell=0}^{J-1} \frac{h^{(\ell)}(0)}{\ell!} + \frac{1}{(J-1)!}\int_0^1 (1-v)^{J-1} h^{(J)}(v)dv.$$

Since $h^{(\ell)}(t) = \sum_{|\alpha|=\ell} (u-x)^\alpha D^\alpha g(x+t(u-x))$, we see

$$f_\sigma(x) = \int_{\mathbb{R}^n} \Phi(\frac{x}{\sigma}, \frac{u}{\sigma}) \sum_{\ell=0}^{J-1} \frac{1}{\ell!} \sum_{|\alpha|=\ell} (u-x)^\alpha D^\alpha g(x)du$$
$$+ \int_{\mathbb{R}^n} \Phi(\frac{x}{\sigma}, \frac{u}{\sigma}) \frac{1}{(J-1)!} \int_0^1 (1-v)^{J-1} \sum_{|\alpha|=J} (u-x)^\alpha D^\alpha g(x+v(u-x))dvdu.$$

Since $\int_{\mathbb{R}^n} \Phi(\frac{x}{\sigma}, \frac{u}{\sigma})du = \sigma^n$, we know from property (i) of the basic window function Φ that

$$f_\sigma(x) = \sigma^n g(x) + \frac{1}{(J-1)!} \int_0^1 (1-\nu)^{J-1} \sum_{|\alpha|=J} \left\{ \int_{\mathbb{R}^n} \Phi(\frac{x}{\sigma}, \frac{u}{\sigma})(u-x)^\alpha D^\alpha g(x+\nu(u-x))du \right\} d\nu.$$

To compute the $H^s(\mathbb{R}^n)$-norm of $f_\sigma - \sigma^n g$, we take $\beta \in \mathbb{Z}_+^n$ satisfying $|\beta| \le s$. We have

$$D^\beta \left\{ \int_{\mathbb{R}^n} \Phi(\frac{x}{\sigma}, \frac{u}{\sigma})(u-x)^\alpha D^\alpha g(x+\nu(u-x))du \right\} = \sum_{j,k,l\in\mathbb{Z}_+^n,\, j+k+l=\beta} \frac{\beta!}{j!k!l!} \int_{\mathbb{R}^n} \sigma^{-|j|} D^j \Phi(\frac{x}{\sigma}, \frac{u}{\sigma}) \frac{\alpha!}{(\alpha-k)!} (-1)^{|k|} (u-x)^{\alpha-k} (1-\nu)^{|l|} D^{\alpha+l} g(x+\nu(u-x))du.$$

Thus $\left\| D^\beta \left(f_\sigma - \sigma^n g \right) \right\|_{L^2(\mathbb{R}^n)}$ is bounded by

$$\sum_{|\alpha|=J} \sum_{j+k+l=\beta} \frac{\alpha!}{(\alpha-k)!} \frac{\beta!}{j!k!l!} \sigma^{-|j|} \frac{1}{(J-1)!} \int_0^1 (1-\nu)^{J+|l|-1} J_{\alpha,j,k,l}(\nu) d\nu,$$

where for $\nu \in (0,1)$,

$$J_{\alpha,j,k,l}(\nu) := \left\| \int_{\mathbb{R}^n} D^j \Phi\Big(\frac{x}{\sigma}, \frac{u}{\sigma}\Big) (u-x)^{\alpha-k} D^{\alpha+l} g(x+\nu(u-x))du \right\|_{L^2(\mathbb{R}^n)}.$$

We need to estimate $J_{\alpha,j,k,l}(\nu)$. By the Schwarz inequality

$$\big(J_{\alpha,j,k,l}(\nu)\big)^2 \le \int_{\mathbb{R}^n} \left\{ \int_{\mathbb{R}^n} \left| D^j \Phi\Big(\frac{x}{\sigma}, \frac{u}{\sigma}\Big) \right| \left| (u-x)^{\alpha-k} \right| du \int_{\mathbb{R}^n} \left| D^j \Phi\Big(\frac{x}{\sigma}, \frac{u}{\sigma}\Big) \right| \left| (u-x)^{\alpha-k} \right| \left| D^{\alpha+l} g(x+\nu(u-x)) \right|^2 du \right\} dx.$$

Decay condition (1.1) of Φ tells us that

$$\int_{\mathbb{R}^n} \left| D^j \Phi\Big(\frac{x}{\sigma}, \frac{u}{\sigma}\Big) \right| \left| (u-x)^{\alpha-k} \right| du \le \sigma^{n+|\alpha-k|} \int_{\mathbb{R}^n} \frac{c_q |u|^{|\alpha-k|}}{(1+|u|)^q} du \le \frac{2c_q \pi^{n/2} \sigma^{n+J-|k|}}{(q+|k|-J-n)\Gamma(n/2)}.$$

https://doi.org/10.4153/CMB-2011-029-8 Published online by Cambridge University Press

Make a variable change $w = x + v(u - x)$. We see that

$$\int_{\mathbb{R}^n} \left|D^j \Phi\Big(\frac{x}{\sigma}, \frac{u}{\sigma}\Big)\right| \left|(u-x)^{\alpha-k}\right| \left|D^{\alpha+l} g(x + v(u-x))\right|^2 du \leq$$
$$\int_{\mathbb{R}^n} \frac{c_q}{(1 + |\frac{w-x}{\sigma v}|)^q} \left|(\frac{w-x}{v})^{\alpha-k}\right| \left|D^{\alpha+l} g(w)\right|^2 v^{-n} dw.$$

It follows from the Schwarz inequality that

$$\big(J_{\alpha,j,k,l}(v)\big)^2 \leq$$
$$\frac{2c_q \pi^{n/2} \sigma^{n+J-|k|}}{(q + |k| - J - n)\Gamma(n/2)} \int_{\mathbb{R}^n} \left|D^{\alpha+l} g(w)\right|^2 \sigma^{n+|\alpha-k|} \int_{\mathbb{R}^n} \frac{c_q |u|^{|\alpha-k|}}{(1+|u|)^q} du dw,$$

which is bounded by $\left(\frac{2c_q \pi^{n/2} \sigma^{n+J-|k|}}{(q+|k|-J-n)\Gamma(n/2)}\right)^2 \|D^{\alpha+l} g\|^2_{L^2(\mathbb{R}^n)}$. Therefore,

$$\left\|D^\beta\left(f_\sigma - \sigma^n g\right)\right\|_{L^2(\mathbb{R}^n)}$$
$$\leq \sum_{|\alpha|=J} \sum_{j+k+l=\beta} \frac{\alpha!}{(\alpha-k)!} \frac{\beta!}{j!k!l!} \frac{1}{(J-1)!} \frac{1}{J+|l|} \frac{2c_q \pi^{n/2} \sigma^{n+J-|j|-|k|}}{(q+|k|-J-n)\Gamma(n/2)} \left\|D^{\alpha+l} g\right\|_{L^2(\mathbb{R}^n)}$$
$$\leq \frac{2 \cdot 3^{J+s} c_q \pi^{n/2} s! \sigma^{n+J-s}}{(q-J-n)\Gamma(n/2) J!} \sum_{|\alpha|=J} \sum_{l \leq \beta} \left\|D^{\alpha+l} g\right\|_{L^2(\mathbb{R}^n)}.$$

Hence, bound (3.2) for $\|f_\sigma - \sigma^n g\|_{H^s(\mathbb{R}^n)} = \sum_{|\beta| \leq s} \left\|D^\beta\left(f_\sigma - \sigma^n g\right)\right\|_{L^2(\mathbb{R}^n)}$ follows. ■

4 Deriving Convergence Rates

Combining Propositions 3.3 and 3.4, we obtain bounds for the total error.

Proposition 4.1 *Under the assumption of Theorem* 2.1*, for* $0 < \sigma, h \leq 1$ *we have with confidence* $1 - \delta$*,*

$$\left\|\frac{h^n f_{\mathbf{z},\sigma}}{\sigma^n} - f^*\right\|_{H^s(\mathbb{R}^n)} \leq C(1+\Delta)^\eta \log\frac{2}{\delta} h^{\frac{n}{2}} \sigma^{-\frac{n}{2}-s} \tag{4.1}$$
$$\max\left\{1, \sigma^{J+\frac{n}{2}} h^{-\frac{n}{2}}, \Delta\sigma^{\frac{n}{2}+s} h^{-\frac{n}{2}}, \Delta^{n+1}\sigma^{\frac{n}{2}+s} h^{-\frac{n}{2}}, \sigma^{\frac{n}{2}+s} h^{\frac{n}{2}+1}\right\}$$

where C *is a constant independent of* δ*,* σ*,* h*, or* Δ*.*

Proof Let us first refine the bound in Proposition 3.3. Since σ_x^2 and p satisfy decay condition (2.1), we have for $i \in \mathbb{Z}^n$,

$$\int_{\mathbb{R}^n} \sigma_x^2 p(x - t_i) dx \leq \int_{\mathbb{R}^n} \frac{c_\eta}{(1+|x|)^\eta} \frac{c_\eta}{(1+|x-t_i|)^\eta} dx.$$

https://doi.org/10.4153/CMB-2011-029-8 Published online by Cambridge University Press

We divide $\mathbb{R}^n$ into two domains, one with $|x-hi| \geq 2\Delta$ and the other with $|x-hi| < 2\Delta$.

When $|x-hi| \geq 2\Delta$, we have $|x-t_i| \geq |x-hi|-|t_i-hi| \geq |x-hi|-\Delta \geq \frac{1}{2}|x-hi|$. It follows that

$$\begin{aligned}\int_{|x-hi|\geq 2\Delta} \frac{c_\eta}{(1+|x|)^\eta} \frac{c_\eta}{(1+|x-t_i|)^\eta} dx &\leq \int_{\mathbb{R}^n} \frac{c_\eta}{(1+|x|)^\eta} \frac{c_\eta}{(1+\frac{1}{2}|x-hi|)^\eta} dx \\ &= \int_{|x|\geq \frac{1}{2}h|i|} \frac{c_\eta}{(1+|x|)^\eta} \frac{c_\eta}{(1+\frac{1}{2}|x-hi|)^\eta} dx + \int_{|x|<\frac{1}{2}h|i|} \frac{c_\eta}{(1+|x|)^\eta} \frac{c_\eta}{(1+\frac{1}{2}|x-hi|)^\eta} dx \\ &\leq \frac{c_\eta}{(1+\frac{1}{2}h|i|)^\eta} \int_{\mathbb{R}^n} \frac{c_\eta}{(1+\frac{1}{2}|x-hi|)^\eta} dx + \frac{c_\eta}{(1+\frac{1}{4}h|i|)^\eta} \int_{\mathbb{R}^n} \frac{c_\eta}{(1+|x|)^\eta} dx \\ &\leq \frac{2^{n+1}c_\eta^2 \pi^{n/2}}{(\eta-n)\Gamma(n/2)} \Big(1+\frac{1}{2}h|i|\Big)^{-\eta} + \frac{2c_\eta^2\pi^{n/2}}{(\eta-n)\Gamma(n/2)}(1+\frac{1}{4}h|i|)^{-\eta}.\end{aligned}$$

When $|x-hi| < 2\Delta$, we have $|x| > h|i| - 2\Delta$. Thus if $|i| > \frac{4\Delta}{h}$, we have $|x| > \frac{1}{2}h|i|$ and

$$\int_{|x-hi|<2\Delta} \frac{c_\eta}{(1+|x|)^\eta} \frac{c_\eta}{(1+|x-t_i|)^\eta} dx \leq \frac{c_\eta}{(1+\frac{1}{2}h|i|)^\eta} \int_{\mathbb{R}^n} \frac{c_\eta}{(1+|x-t_i|)^\eta} dx$$

is bounded by $\frac{2c_\eta^2\pi^{n/2}}{(\eta-n)\Gamma(n/2)}(1+\frac{1}{2}h|i|)^{-\eta}$.

If $|i| \leq \frac{4\Delta}{h}$, we see that

$$\int_{|x-hi|<2\Delta} \frac{c_\eta}{(1+|x|)^\eta} \frac{c_\eta}{(1+|x-t_i|)^\eta} dx \leq \int_{\mathbb{R}^n} \frac{c_\eta^2}{(1+|x|)^\eta} dx \leq \frac{2c_\eta^2\pi^{n/2}}{(\eta-n)\Gamma(n/2)}$$

is bounded by $\frac{2c_\eta^2\pi^{n/2}}{(\eta-n)\Gamma(n/2)}(1+\frac{1}{2}h|i|)^{-\eta}(1+2\Delta)^\eta$.

Applying the above estimates together with the following bound from [11]

$$\sum_{i\in\mathbb{Z}^n} (1+\frac{1}{2}h|i|)^{-\eta} \leq (\sqrt{n}+1)^n + (4/h)^n \frac{2\pi^{n/2}}{(\eta-n)\Gamma(n/2)},$$

we see that $\sum_{i\in\mathbb{Z}^n} \int_{\mathbb{R}^n} \sigma_x^2 p(x-t_i)dx$ is bounded by

$$\frac{2c_\eta^2\pi^{n/2}}{(\eta-n)\Gamma(n/2)} \left\{2^n + 2^\eta + 1 + (1+2\Delta)^\eta\right\} \left\{(\sqrt{n}+1)^n + (8/h)^n \frac{2\pi^{n/2}}{(\eta-n)\Gamma(n/2)}\right\}.$$

In the same way, $\sum_{i\in\mathbb{Z}^n} \int_{\mathbb{R}^n} |f^*(x)|^2 p(x-t_i)dx$ is bounded by the above same expression with c_η^2 replaced by c_η^3.

Therefore from Proposition 3.3 we see that with confidence $1-\delta$,

$$\|f_{\mathbf{z},\sigma} - f_\sigma\|_{H^s(\mathbb{R}^n)} \leq \frac{8c_q(s+1)^n\pi^{n/4}}{\sqrt{(2q-n)\Gamma(n/2)}} \left\{M + c_{\eta,n}(1+\Delta)^{\eta/2}\right\} h^{-\frac{n}{2}} \sigma^{\frac{n}{2}-s} \log\frac{2}{\delta}, \tag{4.2}$$

where $c_{\eta,n}$ is a constant given by

$$c_{\eta,n} = \frac{(c_\eta^{3/2} + c_\eta)\pi^{n/4}2^{2+\eta/2}}{\sqrt{(\eta-n)\Gamma(n/2)}} \left\{ (\sqrt{n}+1)^{n/2} + \frac{8^{n/2}\pi^{n/4}}{\sqrt{(\eta-n)\Gamma(n/2)}} \right\}.$$

In order to use Proposition 3.4, we need to bound the norm

$$\Bigl\| f^* \sum_{i\in\mathbb{Z}^n} p(\,\cdot\, - t_i) \Bigr\|_{H^{J+s}(\mathbb{R}^n)}.$$

Applying decay condition (2.1) again, we see that $\| f^* \sum_{i\in\mathbb{Z}^n} p(\,\cdot\, - t_i)\|_{H^{J+s}(\mathbb{R}^n)}$ is bounded by

$$c_\eta^2 \frac{\{2(J+s)+1\}^n \pi^{n/4}}{\sqrt{(2\eta-n)\Gamma(n/2)}} (1+\Delta)^\eta 2^{4+\eta} \left\{ (\sqrt{n}+1)^n + \frac{8^n \pi^{n/2}}{(\eta-n)\Gamma(n/2)} \right\} h^{-n}.$$

It follows from Proposition 3.4 that

$$\Bigl\| f_\sigma - \sigma^n f^* \sum_{i\in\mathbb{Z}^n} p(\,\cdot\, - t_i) \Bigr\|_{H^s(\mathbb{R}^n)} \le \frac{\{2(J+s)+1\}^n 3^{J+s+1} s! c_q c_\eta^2 \pi^{3n/4}}{(q-J-n)J!\sqrt{(2\eta-n)\Gamma(n/2)}\Gamma(n/2)} (1+\Delta)^\eta 2^{4+\eta} \left\{ (\sqrt{n}+1)^n + \frac{8^n\pi^{n/2}}{(\eta-n)\Gamma(n/2)} \right\} \sigma^{n+J-s} h^{-n}. \tag{4.3}$$

Finally, we must study the difference between $\sum_{i\in\mathbb{Z}^n} p(\,\cdot\, - t_i)$ and $\sum_{i\in\mathbb{Z}^n} p(\,\cdot\, - hi)$ by the restriction $\sup_i |t_i - hi| \le \Delta$. For $x \in \mathbb{R}^n$,

$$\begin{aligned}
\Bigl| \sum_{i\in\mathbb{Z}^n} p(x-hi) - \sum_{i\in\mathbb{Z}^n} p(x-t_i) \Bigr|
&\le \sum_{i\in\mathbb{Z}^n} |p(x-hi) - p(x-hi-(t_i-hi))| \\
&= \sum_{i\in\mathbb{Z}^n} \Bigl| \int_0^1 \sum_{|\alpha|=1} (t_i-hi)^\alpha D^\alpha p(x-hi-u(t_i-hi)) du \Bigr| \\
&\le \sum_{i\in\mathbb{Z}^n} \int_0^1 \frac{n\Delta c_\eta}{(1+|x-hi-u(t_i-hi)|)^\eta} du.
\end{aligned}$$

For every $x \in \mathbb{R}^n$, we can find some $k \in \mathbb{Z}^n$ such that $x - hk \in [-\frac{h}{2}, \frac{h}{2})^n$. Then $|x-hi| = |x-hk+h(k-i)| \ge \frac{1}{2}|k-i|h$. Separate $\sum_{i\in\mathbb{Z}^n}$ into two parts $\sum_{|x-hi|\ge 2\Delta}$ (where $|x-hi-u(t_i-hi)| \ge |x-hi| - \Delta \ge \frac{1}{2}|x-hi|$ for $u\in[0,1]$) and $\sum_{|x-hi|<2\Delta}$.

We have

$$\Big|\sum_{i\in\mathbb{Z}^n} p(x-hi) - \sum_{i\in\mathbb{Z}^n} p(x-t_i)\Big|$$
$$\le \sum_{i\in\mathbb{Z}^n} \frac{n\Delta c_\eta}{(1+\frac{1}{2}|x-hi|)^\eta} + (\frac{4\Delta}{h})^n n\Delta c_\eta$$
$$\le \sum_{i\in\mathbb{Z}^n} \frac{n\Delta c_\eta}{(1+\frac{1}{4}|k-i|h)^\eta} + (\frac{4\Delta}{h})^n n\Delta c_\eta$$
$$\le \Big\{(\sqrt{n}+1)^n + \frac{2\pi^{n/2}8^n}{(\eta-n)\Gamma(n/2)}\Big\} nc_\eta h^{-n}\Delta + 4^n nc_\eta h^{-n}\Delta^{n+1}.$$

Combining an estimation in [11] for $\big|h^n\sum_{i\in\mathbb{Z}^n} h^n p(x-hi)-1\big|$, we have

$$\Big|h^n\sum_{i\in\mathbb{Z}^n} p(x-t_i)-1\Big| \le c_\eta\left\{c'_{\eta,n}\Delta + 4^n n\Delta^{n+1} + (c'_{\eta,n}+n2^n)h^{n+1}\right\},$$

where $c'_{\eta,n}$ is the constant given by

$$c'_{\eta,n} = \Big\{(\sqrt{n}+1)^n + \frac{2\pi^{n/2}8^n}{(\eta-n)\Gamma(n/2)}\Big\} n.$$

This, in connection with (4.2) and (4.3), tells us that with confidence $1-\delta$, the total error $\|\frac{h^n f_{\mathbf{z},\sigma}}{\sigma^n} - f^*\|_{H^s(\mathbb{R}^n)}$ is bounded by

$$\frac{8c_q(s+1)^n\pi^{n/4}}{\sqrt{(2q-n)\Gamma(n/2)}}\left\{M + c_{\eta,n}(1+\Delta)^{\eta/2}\right\} h^{\frac{n}{2}}\sigma^{-\frac{n}{2}-s}\log\frac{2}{\delta}$$
$$+ c_\eta\|f^*\|_{H^s(\mathbb{R}^n)}\{c'_{\eta,n}\Delta + 4^n n\Delta^{n+1} + (c'_{\eta,n}+n2^n)h^{n+1}\}$$
$$+ \frac{\{2(J+s)+1\}^n 3^{J+s+1} s! c_q c_\eta^2 \pi^{3n/4}(1+\Delta)^\eta 2^{4+\eta}}{(q-J-n)J!\sqrt{(2\eta-n)\Gamma(n/2)}\Gamma(n/2)}$$
$$\Big\{(\sqrt{n}+1)^n + \frac{8^n\pi^{n/2}}{(\eta-n)\Gamma(n/2)}\Big\}\sigma^{J-s}.$$

So desired bound (4.1) holds true with the constant C taken to be

$$C = 40c_q(s+1)^n\{M+c_{\eta,n}\} + c_\eta\|f^*\|_{H^s(\mathbb{R}^n)}\{2c'_{\eta,n}+n4^n+n2^n\}$$
$$+ \{2(J+s)+1\}^n 3^{J+s+5} s! c_q c_\eta^2 2^{6+\eta}\left\{(\sqrt{n}+1)^n + 8^n\right\}.$$

This proves Proposition 4.1. ∎

Now we can prove Theorem 2.1 easily.

Proof of Theorem 2.1 Since $\Delta \leq h^{\frac{n(J-s)}{n+2J}} \leq 1$ and $\sigma = h^{\frac{n}{n+2J}}$, we have $\sigma^{J+\frac{n}{2}} h^{-\frac{n}{2}} = 1$ and $\Delta \sigma^{\frac{n}{2}+s} h^{-\frac{n}{2}} = 1$. So we get

$$\max\left\{1, \sigma^{J+\frac{n}{2}} h^{-\frac{n}{2}}, \Delta\sigma^{\frac{n}{2}+s} h^{-\frac{n}{2}}, \Delta^{n+1}\sigma^{\frac{n}{2}+s} h^{-\frac{n}{2}}, \sigma^{\frac{n}{2}+s} h^{\frac{n}{2}+1}\right\} = 1.$$

Then the conclusion of Theorem 2.1 follows from Proposition 4.1 with the constant

$$\widetilde{C}_{q,n,\eta,J,s} = 2^{\eta}\Big\{ 40(s+1)^n\{1+c_{\eta,n}\} + 2c'_{\eta,n} + n4^n + n2^n + \{2(J+s)+1\}^n 3^{J+s+5} s! 2^{6+\eta}\big[(\sqrt{n}+1)^n + 8^n\big]\Big\}.$$

The proof of Theorem 2.1 is complete. ■

References

[1] A. Aldroubi and K. Gröchenig, *Nonuniform sampling and reconstruction in shift-invariant spaces.* SIAM Rev. **43**(2001), no 4, 585–620. doi:10.1137/S0036144501386986

[2] C. Campbell and Y. Ying, *Learning coordinate gradients with multi-task kernels.* In: Proceedings of the 21st Annual Conference on Learning Theory (COLT), 2008.

[3] C. de Boor and R.-Q. Jia, *Controlled approximation and a characterization of the local approximation order.* Proc. Amer. Math. Soc. **95**(1985), no. 4, 547–553.

[4] B. Han and R.-Q. Jia, *Multivariate refinement equations and convergence of subdivision schemes.* SIAM J. Math. Anal. **29**(1998), no. 5, 1177–1999. doi:10.1137/S0036141097294032

[5] S. Mukherjee and D.-X. Zhou, *Learning coordinate covariances via gradients.* J. Mach. Learn. Res. **7**(2006), 519–549.

[6] I. Pinelis, *Optimum bounds for the distributions of martingales in Banach spaces.* Ann. Probab. **22**(1994), no. 4, 1679–1706. doi:10.1214/aop/1176988477

[7] S. Smale and D.-X. Zhou, *Learning theory estimates via integral operators and their approximations.* Constr. Approx. **26**(2007), no. 2, 153–172. doi:10.1007/s00365-006-0659-y

[8] ———, *Shannon sampling and function reconstruction from point values.* Bull. Amer. Math. Soc. **41**(2004), no. 3, 279–305. doi:10.1090/S0273-0979-04-01025-0

[9] ———, *Online learning with Markov sampling.* Anal. Appl. **7**(2009), no. 1, 87–113. doi:10.1142/S0219530509001293

[10] H. Sun and Q. Wu, *Regularized least square regression with dependent samples.* Adv. Comput. Math. **32**(2010), no. 2, 175–189. doi:10.1007/s10444-008-9099-y

[11] X.-J. Zhou and D.-X. Zhou, *High order Parzen windows and randomized sampling.* Adv. Comput. Math. **31**(2009), no. 4, 349–368. doi:10.1007/s10444-008-9073-8

Joint Advanced Research Center in Suzhou, University of Science and Technology of China and City University of Hong Kong, Suzhou, Jiangshu, 215123, China
e-mail: (Xiang-Jun Zhou) 50009229@student.cityu.edu.hk (Lei Shi) sl1983@mail.ustc.edu.cn

Department of Mathematics, City University of Hong Kong, Kowloon, Hong Kong, China
e-mail: (Ding-Xuan Zhou) mazhou@cityu.edu.hk