Journal of Approximation Theory **118**, 66–80 (2002)
doi:10.1006/jath.2002.3710

Stable Summation of Orthogonal Series with Noisy Coefficients

Peter Mathé[1]

Weierstraß Institute for Applied Analysis and Stochastics, Mohrenstraße 39, D-10117 Berlin, Germany
E-mail: mathe@wias-berlin.de

and

Sergei V. Pereverzev

Ukrainian Academy of Sciences, Institute of Mathematics, Tereshenkivska Str. 3, Kiev 4, Ukraine
E-mail: serg-p@mail.kar.net

Communicated by Rolf J. Nessel

Received March 5, 2001; accepted in revised form June 26, 2002

We study the recovery of continuous functions from Fourier coefficients with respect to certain given orthonormal systems, blurred by noise. For deterministic noise this is a classical ill-posed problem. Emphasis is laid on a priori smoothness assumptions on the solution, which allows to apply regularization to reach the best possible accuracy. Results are obtained for systems obeying norm growth conditions. In the white noise setting mild additional assumptions have to be made to have accurate bounds. We finish our study with the recovery of functions from noisy coefficients with respect to the Haar system. © 2002 Elsevier Science (USA)

Key Words: orthogonal series; stable summation; Gaussian white noise; Haar system.

1. INTRODUCTION

The problem of stable summation of Fourier series with respect to a given orthonormal system of functions $\{\varphi_k,\ k = 1, 2, \dots\}$ under small changes in the coefficients, measured in l_2 is a classical example of an ill-posed problem if we want to measure the error in the sup-norm $||\cdot||_\infty$. To be more specific we study functions belonging to $C[0, 1]$, the space of continuous functions on $[0, 1]$. The problem may now be formulated as follows. We want to

[1] To whom correspondence should be addressed.

0021-9045/02 $35.00

recover a continuous function y from (Fourier) coefficients with respect to a given orthonormal system $\{\varphi_k,\ k=1,2,\ldots\}$, but instead of $y_k := \langle y, \varphi_k \rangle$, $k=1,2,\ldots$ we are given only a noisy sequence of numbers $y_\delta := (y_{\delta,k})_{k=1}^{\infty}$, determined by

$$y_{\delta,k} := y_k + \delta\xi_k, \qquad k = 1,2,\ldots, \tag{1}$$

where $\xi := (\xi_k)_{k=1}^{\infty}$ is the noise. Usually it is assumed that the noise is deterministic and has $\|\xi\|_{l_2} \leqslant 1$. This classical ill-posed problem was studied in [12, Chap. 6], in an appendix to the textbook [4], in papers by Aliev [2] and many others. In all cases the application of Tikhonov regularization was considered. Standard assumptions on the smoothness of the true solution were expressed in terms of spaces W_2^μ, associated with the given system $\{\varphi_k,\ k=1,2,\ldots\}$, i.e.,

$$W_2^\mu := \left\{ y \in L_2[0,1], \quad \|y\|_\mu^2 := \sum_{k=1}^{\infty} k^{2\mu} |\langle y, \varphi_k \rangle|^2 < \infty \right\}.$$

Typically, functions belonging to some W_2^μ, will have certain derivatives.

Example 1.1. For the trigonometric system and integer μ, the space W_2^μ consists of 1-periodic functions having square summable derivatives up to the order μ.

If we let $\{\varphi_k,\ k \in \mathbb{N}\}$ be the (properly scaled) system of Legendre polynomials, then Rafal'son [10] and Tomin [13] have shown, that the respective space W_2^μ consists of all functions $y(t)$, for which the derivatives $y^{(i)},\ i=1,\ldots,\mu-1$ are absolutely continuous on each subinterval $[a,b] \subset (0,1)$ and

$$\int_0^1 |y^{(\mu)}(t)|^2 t^\mu (1-t)^\mu \, dt < \infty.$$

This means, that the highest derivative $y^{(\mu)}$ may have singularities at the end points 0 or 1.

On the other hand, under such assumptions, Il'in and Pozniak [4] for the trigonometric system and Aliev [2] for the more general case of any system with uniformly bounded norm $\|\varphi_k\|_\infty \leqslant C,\ k=1,2,\ldots,$ proved that for $s \in (\frac{1}{4}, \mu - \frac{1}{2})$ Tikhonov regularization $T^{\alpha,s}(y_\delta) := \sum_{k=1}^{\infty} \frac{y_{\delta,k}}{1+\alpha k^{2s}} \varphi_k$ yields $\|y - T^{\alpha,s}(y_\delta)\|_\infty \leqslant C(\sqrt{\alpha} + \delta/\alpha)$. We emphasize that it is one of the major topics within the theory of ill-posed problems to investigate how to choose the regularization parameter α, or the discretization level n, below, as a function of the noise level δ. So, from the previous estimate one can see that the optimal choice for α is $\alpha_0 = \delta^{2/3}$ for which

$$||y - T^{\alpha_0,s}(y_\delta)||_\infty \leqslant C\delta^{1/3}. \tag{2}$$

Here and throughout the paper C denotes generic constants that can vary from appearance to appearance. To the best of our knowledge, we refer also to the survey by Liskovets [8], this is the culmination of all previous work on this particular problem. Still some questions remain open.

First, estimate (2) does not take into account the given smoothness. Moreover, it does not indicate the actual degree of ill-posedness of the problem which is measured by the lack of accuracy due to ill-posedness. It is common belief, that this degree depends on the growth of $||\varphi_k||_\infty$ when $k \to \infty$. We will study this problem for two different classes of orthonormal systems, satisfying norm growth assumptions.

We say, that an orthonormal system $\{\varphi_k,\ k = 1, 2, \ldots\}$ belongs to class (M_ν), if

$$\sum_{k=1}^{\infty} \frac{||\varphi_k||_\infty^2}{k^{2\nu}} < \infty \tag{M_ν}$$

for some $\nu > 0$. On the other hand, if it obeys

$$||\varphi_k||_\infty \asymp k^\beta, \qquad k = 1, 2, \ldots, \tag{K^β}$$

now for some $\beta \geqslant 0$, then we agree to say, that it belongs to class (K^β).

It is immediate, that systems $\{\varphi_1, \varphi_2, \ldots\}$, which obey (K^β), will satisfy (M_ν), with $\nu = \beta + \frac{1}{2} + \varepsilon$ for any small $\varepsilon > 0$. The trigonometric system has property (K^β) for $\beta = 0$, whereas the system of Legendre polynomials requires to take $\beta = 1$. One way to obtain systems from M_ν is to consider lacunary sequences $\{\varphi_k = \psi_{n_k}, n_k = \theta(k)\}$, where $\theta(k)$ increases. If $\{\psi_k,\ k = 1, 2, \ldots\}$ obeys (K^β), then $\{\psi_{n_k},\ k = 1, 2, \ldots\}$ satisfies (M_ν), if $\theta(k)^\beta k^{-\nu}$ is square summable.

We now return to the problem of stable summation. Below it will turn out, that the solution to the problem of stable summation heavily depends as well on the classes of orthonormal systems as on the kind of noise. Within the deterministic noise framework, systems which enjoy property (M_ν) fit best. The point-wise growth as in (K^β) does not exactly fit, but still a slight modification of our arguments allows to provide error estimates.

For white noise, the situation is different, since the behavior of the noise is very much dependent on geometric properties of the underlying space. Here we consider as noise $\xi = (\xi_k)_{k=1}^\infty$ independent standard Gaussian random variables. In the Gaussian white noise setting we have to make an additional assumption on the orthogonal system, expressed in terms of Lipschitz properties, namely, that there are some $\rho < \infty$ and $C < \infty$, for which

$$|\varphi_k(s) - \varphi_k(t)| \leqslant Ck^\rho ||\varphi_k||_\infty |s - t|, \qquad s, t \in [0, 1]. \tag{3}$$

Note that this assumption is fulfilled for the trigonometric system with $\rho = 1$. For algebraic polynomials φ_k of degree k Markov's inequality, see [6, Theorem 3.5.8], asserts $||\varphi_k'||_\infty \leqslant 2k^2 ||\varphi_k||_\infty$, such that (3) is satisfied with $\rho = 2$.

We shall study regularizing properties of summation methods

$$T_n^a(y_\delta) := \sum_{k=1}^n \alpha_k^n y_{\delta,k} \varphi_k,$$

for certain triangular arrays $a := (\alpha_k^n),\ k = 1, \ldots, n,\ n \in \mathbb{N}$. Such summation methods are called λ-methods, see, e.g. [6, Chap. 2.2.4] and play a stabilizing role in the direct (well-posed) problem of recovery.

The quality of the summation methods $T_n^a(y_\delta)$ will depend on the truncation level n and on properties of $a := (\alpha_k^n),\ k = 1, \ldots, n,\ n \in \mathbb{N}$, more precisely we shall assume, that there is C and some λ, such that

$$|1 - \alpha_k^n| \leqslant C \left(\frac{k}{n}\right)^\lambda, \qquad 1 \leqslant k \leqslant n, \quad n \in \mathbb{N}. \tag{4}$$

We agree to call T_n^a of degree λ, if this is the case. We note in passing, that as a consequence, the array a is uniformly bounded. The following examples show that assumption (4) is rather natural.

EXAMPLE 1.2. These methods cover the most prominent methods of regularization, one obtained from *Tikhonov regularization* with parameters α and smoothness s, i.e., $T_n^{\alpha,s}(y_\delta) := \sum_{k=1}^n \frac{y_{\delta,k}}{1+\alpha k^{2s}} \varphi_k$, and the other by *self-regularization* $S_n(y_\delta) := \sum_{k=1}^n y_{\delta,k} \varphi_k$. It is easily seen, that self-regularization S_n has arbitrary degree, while Tikhonov regularization $T_n^{\alpha,s}$ with $\alpha = \alpha(n) \leqslant Cn^{-\rho}$ has degree $2s$, if $\rho \geqslant 2s$ and 0 otherwise.

The *Féjer method* of summation, where $\alpha_k^n := (1 - \frac{k-1}{n}),\ k \leqslant n$, meets (4) with $\lambda = 1$.

Moreover, we also indicate the *Bernstein–Rogosinsky method*, applied to the trigonometric system for which $n = 2m + 1$ usually. In this method we let $\alpha_{2l}^n := \alpha_{2l+1}^n := \cos \frac{\pi l}{2m},\ l = 0, 1, \ldots, m$. Since

$$|1 - \alpha_{2l}^n| = |1 - \alpha_{2l+1}^n| = \left|1 - \cos \frac{\pi l}{2m}\right| = 2 \sin^2 \frac{\pi l}{4m} \leqslant \frac{\pi^2}{8} \left(\frac{l}{m}\right)^2,$$

one can see that the Bernstein–Rogosinsky method has degree 2.

As will be transparent below, given an orthonormal system and a priori smoothness μ, any such summation method will provide the same optimal order of accuracy as $\delta \to 0$, if λ in (4) is large enough. This best possible accuracy will be different, for classes (M_ν) and (K^β) as well as for different assumptions on the noise.

2. SYSTEMS FROM CLASS (M_ν)

We first note the following inequality, which will be useful below. Among others it implies, that for $\mu > \nu$ functions from W_2^μ are continuous, if the system $\{\varphi_k,\ k = 1, 2, \ldots\}$ consists of continuous functions.

PROPOSITION 2.1. *For $\mu > \nu$ we have*

$$\left|\left|\sum_{k=n+1}^{\infty} \langle y, \varphi_k \rangle \varphi_k \right|\right|_\infty \leqslant (n+1)^{-(\mu-\nu)} ||y||_\mu \left(\sum_{k=1}^{\infty} \frac{||\varphi_k||_\infty^2}{k^{2\nu}} \right)^{1/2}.$$

Proof. Let us recall $y_k = \langle y, \varphi_k \rangle$. An application of the Cauchy–Schwarz inequality provides

$$\begin{aligned} \left|\left|\sum_{k=n+1}^{\infty} y_k \varphi_k \right|\right|_\infty &\leqslant \sum_{k=n+1}^{\infty} k^{\nu-\mu} k^{\mu} |y_k| k^{-\nu} ||\varphi_k||_\infty \\ &\leqslant (n+1)^{\nu-\mu} ||y||_\mu \left(\sum_{k=n+1}^{\infty} \frac{||\varphi_k||_\infty^2}{k^{2\nu}} \right)^{1/2}, \end{aligned}$$

from which the assertion easily follows. ∎

The main result in this section is

THEOREM 2.1. *Let $\{\varphi_k,\ k = 1, 2, \ldots\}$ be of class* (M_ν). *Suppose we are given noisy observations* (1) *and have a priori knowledge $\mu > \nu$ of the exact solution.*

Let T_n^a be any summation method of degree $\lambda \geqslant \mu - \nu$.

1. Deterministic noise: *For $n \asymp \delta^{-1/\mu}$ we have*

$$\sup_{||y||_\mu \leqslant 1} \sup_{||\xi||_{l_2} \leqslant 1} ||y - T_n^a(y_\delta)||_\infty \leqslant C \delta^{(\mu-\nu)/\mu}. \tag{5}$$

2. Gaussian white noise: *For systems obeying* (3) *and a choice of $n \asymp \left(\delta\sqrt{\log(1/\delta)}\right)^{-1/\mu}$ we have the following bound*:

$$\sup_{\|y\|_\mu \leqslant 1} \mathbf{E}\|y - T_n^a(y_\delta)\|_\infty \leqslant C(\delta\sqrt{\log(1/\delta)})^{\frac{\mu-\nu}{\mu}}, \tag{6}$$

where $\mathbf{E}$ *denotes the expectation with respect to the Gaussian white noise* ξ.

Proof. For any summation method $T_n^a(y_\delta)$ we decompose the error as

$$\|y - T_n^a(y_\delta)\|_\infty \leqslant \left\|\sum_{k=1}^{n} (1-\alpha_k^n) y_k \varphi_k\right\|_\infty + \left\|\sum_{k=n+1}^{\infty} y_k\varphi_k\right\|_\infty + \delta\left\|\sum_{k=1}^{n} \alpha_k^n \xi_k \varphi_k\right\|_\infty. \tag{7}$$

The first two summands are independent of the noise. For deterministic noise the middle summand on the right-hand side above was estimated in Proposition 2.1 and yields asymptotically $\delta^{(\mu-\nu)/\mu}$ by our choice of n. The first summand can be estimated similarly as in Proposition 2.1 as

$$\begin{aligned}\left\|\sum_{k=1}^{n} (1-\alpha_k^n) y_k \varphi_k\right\|_\infty &\leqslant \sum_{k=1}^{n} |1-\alpha_k^n||y_k|\|\varphi_k\|_\infty \\ &\leqslant \frac{C}{n^{\mu-\nu}} \sum_{k=1}^{n} k^{\mu-\nu}|y_k|\|\varphi_k\|_\infty \\ &\leqslant Cn^{\nu-\mu}\|y\|_\mu \left(\sum_{k=1}^{n} \frac{\|\varphi_k\|_\infty^2}{k^{2\nu}}\right)^{1/2} \asymp \delta^{(\mu-\nu)/\mu}\|y\|_\mu.\end{aligned}$$

For deterministic noise the last summand in (7) can be estimated as

$$\delta\left\|\sum_{k=1}^{n} \alpha_k^n \xi_k \varphi_k\right\|_\infty \leqslant C\delta n^\nu \left(\sum_{k=1}^{n} \frac{\|\varphi_k\|_\infty^2}{k^{2\nu}}\right)^{1/2} \asymp \delta^{(\mu-\nu)/\mu},$$

such that we arrive at (5).

It remains to treat the Gaussian white noise case. This is more elaborate. We have to bound $\mathbf{E}\|\sum_{k=1}^n \alpha_k^n \xi_k \varphi_k\|_\infty$, which will be done using Dudley's Theorem [7, Theorem 11.17]. This asserts, that for Gaussian processes $(X_t)_{t\in T}$ we have

$$\mathbf{E}\sup_{t\in T} |X_t| \leqslant 24\int_0^D \sqrt{\log N(T, d_X, \varepsilon)}\, d\varepsilon, \tag{8}$$

where $N(T, d_X, \varepsilon)$ denotes the minimal number of ε–balls required to cover T in the metric $d_X(s,t) := (\mathbf{E}|X_s - X_t|^2)^{1/2}$. D denotes the diameter of

(T, d_X). Estimate (8) is applied to the process

$$X_t := \sum_{k=1}^{n} \alpha_k^n \xi_k \varphi_k(t), \qquad t \in [0,1].$$

Since $\{\xi_k,\ k = 1,2,\ldots\}$ are i.i.d. standard normal variables, we can explicitly compute the metric

$$d_X(s,t) = \left(\sum_{k=1}^{n} (\alpha_k^n)^2 |\varphi_k(s) - \varphi_k(t)|^2\right)^{1/2}, \qquad s,t \in [0,1],$$

such that it is easy to bound the diameter by

$$D \leqslant 2\left(\sum_{k=1}^{n} (\alpha_k^n)^2 ||\varphi_k||_\infty^2\right)^{1/2}.$$

We can also bound distances in $d_X(s,t)$ by a multiple of $|s-t|$, using assumption (3). Indeed,

$$d_X(s,t) \leqslant \left(\sum_{k=1}^{n} C^2 (\alpha_k^n)^2 k^{2\rho} ||\varphi_k||_\infty^2\right)^{1/2} |s-t|, \qquad s,t \in [0,1].$$

If we denote $B := C(\sum_{k=1}^n (\alpha_k^n)^2 k^{2\rho} ||\varphi_k||_\infty^2)^{1/2}$, then Dudley's estimate (8) yields

$$\mathbf{E}\left|\left|\sum_{k=1}^{n} \alpha_k^n \xi_k \varphi_k\right|\right|_\infty \leqslant C \int_0^D \sqrt{\log(B/\varepsilon)}\, d\varepsilon.$$

We let $\bar{D} := \max\{1, D\}$ and bound the right-hand side, using the Cauchy–Schwarz inequality, by

$$\int_0^D \sqrt{\log(B/\varepsilon)}\, d\varepsilon \leqslant \sqrt{D}\left(\int_0^{\bar{D}} \log(B/\varepsilon)\, d\varepsilon\right)^{1/2} \leqslant \bar{D}\sqrt{1 + \log(B)}.$$

Assumption (M_ν) yields $B \leqslant Cn^{(\rho+\nu)}$ as well as $D \leqslant Cn^\nu$, which finally implies

$$\mathbf{E}\left|\left|\sum_{k=1}^{n} \alpha_k^n \xi_k \varphi_k\right|\right|_\infty \leqslant Cn^\nu \sqrt{\log(n)}. \tag{9}$$

Combining the estimate for the noise free term with (9) we obtain for the proper choice of n

$$\mathbf{E}\|y - T_n^a(y_\delta)\|_\infty \leqslant C\|y\|_\mu (n^{\nu-\mu} + \delta n^\nu \sqrt{\log(n)})$$
$$\leqslant C\|y\|_\mu (\delta\sqrt{\log(1/\delta)})^{\frac{\mu-\nu}{\mu}},$$

which completes the proof of (6) and of the theorem. ∎

3. SYSTEMS FROM CLASS (K^β)

Again we start with a tail estimate, ensuring that W_2^μ consists of continuous functions, if only the orthogonal system was continuous.

PROPOSITION 3.1. *For $\mu > \beta + \frac{1}{2}$ we have*

$$\left\| \sum_{k=n+1}^{\infty} y_k \varphi_k \right\|_\infty \leqslant C n^{-\mu+\beta+1/2} \|y\|_\mu. \tag{10}$$

We omit the proof, because it is straightforward. The main result is

THEOREM 3.1. *Let $\{\varphi_k,\ k = 1, 2, \ldots\}$ obey assumption (K^β). Suppose we are given noisy observations (1) and have a priori knowledge μ of the exact solution for $\mu > \beta + \frac{1}{2}$. Let $T_n^a(y_\delta)$ be any summation method of degree $\lambda \geqslant \mu$.*
1. Deterministic noise: *With $n \asymp \delta^{-1/\mu}$ noisy data we have the following error estimate*:

$$\sup_{\|y\|_\mu \leqslant 1} \ \sup_{\|\xi\|_{l_2} \leqslant 1} \|y - T_n^a(y_\delta)\|_\infty \leqslant C\delta^{(\mu-\beta-1/2)/\mu}. \tag{11}$$

2. Gaussian white noise: *For systems obeying (3) and a choice of $n \asymp (\delta\sqrt{\log(1/\delta)})^{-1/\mu}$ we have*

$$\sup_{\|y\|_\mu \leqslant 1} \mathbf{E}\, \|y - T_n^a(y_\delta)\|_\infty \leqslant C(\delta\sqrt{\log(1/\delta)})^{\frac{\mu-\beta-1/2}{\mu}}. \tag{12}$$

We note that the bounds given above are stronger, than the ones we would obtain, simply using that (K^β) implies (M_ν) for $\nu = \beta + \frac{1}{2} + \varepsilon$ and applying Theorem 2.1.

Proof. Again we start with the error decomposition (7) and use estimate (10) and the assumption on n to bound the middle term.

Moreover, we have the following Nikolski type inequality for any polynomials with respect to the system $\{\varphi_k\}$.

$$\left\|\sum_{k=1}^{n} u_k\varphi_k\right\|_\infty \leqslant C\sum_{k=1}^{n}|u_k|k^\beta$$
$$\leqslant C\left(\sum_{k=1}^{n}|u_k|^2\right)^{1/2}\left(\sum_{k=1}^{n}k^{2\beta}\right)^{1/2} \leqslant Cn^{\beta+1/2}\left(\sum_{k=1}^{n}|u_k|^2\right)^{1/2}.$$

We apply this to $||\sum_{k=1}^{n}(1-\alpha_k^n)y_k\varphi_k||_\infty$ and $||\sum_{k=1}^{n}\alpha_k^n\xi_k\varphi_k||_\infty$, separately and obtain

$$\left\|\sum_{k=1}^{n}(1-\alpha_k^n)y_k\varphi_k\right\|_\infty \leqslant Cn^{\beta+1/2}\left(\sum_{k=1}^{n}|1-\alpha_k^n|^2|y_k|^2\right)^{1/2}$$
$$\leqslant Cn^{\beta+1/2}\left(\max_{1\leqslant k\leqslant n}\{|1-\alpha_k^n|k^{-\mu}\}\right)||y||_\mu$$
$$\leqslant Cn^{-\mu+\beta+1/2}||y||_\mu,$$

and $||\sum_{k=1}^{n}\alpha_k^n\xi_k\varphi_k||_\infty \leqslant Cn^{\beta+1/2}$. Recoursing the above estimates into the error decomposition (7) we obtain for $n \asymp \delta^{-1/\mu}$ and $||y||_\mu \leqslant 1$

$$\left\|y-\sum_{k=1}^{n}\alpha_k^n y_{\delta,k}\varphi_k\right\|_\infty \leqslant Cn^{\beta+1/2}\{n^{-\mu}+\delta\} \leqslant C\delta^{\frac{\mu-\beta-1/2}{\mu}},$$

which completes the proof of estimate (11). The proof for random noise differs from the one for system from class (M_ν) only in the bound for the diameter D, which can be estimated by

$$D \leqslant 2\sqrt{n}\max_{k\leqslant n}|\alpha_k^n|||\varphi_k||_\infty \leqslant Cn^{\beta+1/2}.$$

This allows to prove (12) and to complete the proof of the theorem. ∎

We mention, that the use of the Nikolski type argument evolved during discussions with V. Temlyakov, Univ. of South Carolina.

Remark 3.1. Estimation under the presence of Gaussian white noise has also been studied in [12], and more recently by Tsybakov [14]. For trigonometric systems $\{\varphi_k, = 1, 2, \ldots\}$ and $y \in W_2^\mu$ Tsybakov indicated the best possible order for the expected value of the error, measured in $C[0,1]$. This turns out to be of the order $(\delta\sqrt{\log(1/\delta)})^{\frac{\mu-1/2}{\mu}}$. He proved that this order cannot be improved even if the trigonometric system is replaced by any other orthonormal system, still assuming $y \in W_2^{\mu}$, associated to the trigonometric system. It is worth mentioning, that in this paper, discretized Tikhonov regularization was used and that the regularization parameter α was chosen adapting to the unknown smoothness. Therefore it had to be required, that the number n of observations was at least

$$n \geqslant \delta^{-2/(\min\{1,\mu_0 - 1/2\})}, \tag{13}$$

where μ_0, the minimal smoothness, was supposed to be known.

We note that the number n of observations, which provide the optimal order does not depend on properties of the system $\{\varphi_k, \ k = 1, 2, \ldots\}$. Therefore, since the trigonometric system obeys (K^{β}) for $\beta = 0$, we obtain the best order of accuracy for this system with a number of observations much less than the one in (13).

Further we note that for deterministic noise the order-optimality of the error estimate (11) for the trigonometric system ($\beta = 0$) can be proved within the general scheme for proving error estimates in the worst case, we refer to [9]. We stress that to this end Shadrin's inequality [11] must be used instead of the standard interpolation inequality.

The constants C in the bounds (5), (6), (11) and (12) can be easily estimated in the course of the proof. For example, as C in (5) one can take $C = C_v(2C_4 + 1)$, where C_4 is the constant from (4), and C_v is the sum in $(\mathbf{M}_v)$.

4. THE HAAR SYSTEM

In this section we study the recovery of functions based on noisy coefficients with respect to the Haar system $\{\chi_{m,j}\}$, which is briefly introduced as follows, see [5, Chap. 3]. We let $\chi_{0,0}(t) \equiv 1$ and for naturals $k = 1, 2, \ldots$ and $j = 0, 1, 2, \ldots, 2^k - 1$, we let

$$\chi_{k,j}(t) := \begin{cases} 2^{k/2} & \text{if } j/2^k \leqslant t < (j + 1/2)/2^k, \\ -2^{k/2} & \text{if } (j + 1/2)/2^k \leqslant t < (j + 1)/2^k, \\ 0 & \text{else.} \end{cases}$$

We note that at each level k and for each $t \in (0, 1)$ there is at most one $j = j(k)$ for which $\chi_{k,j} \neq 0$. In this case it takes absolute value $2^{k/2}$. The Haar system is an orthonormal basis in $L_2[0, 1]$. By assigning $(k,j) \to n := 2^k + j$ we obtain a system as in the previous sections. With this identification we see for any function $y = \sum_k \sum_j y_{k,j}\chi_{k,j}$

$$\|y\|_\mu^2 = \sum_{n=1}^{\infty} n^{2\mu}|\langle y, \chi_n\rangle|^2 \asymp \sum_{k=0}^{\infty} 2^{2k\mu} \sum_{j=0}^{2^k-1} |\langle y, \chi_{k,j}\rangle|^2 = \|y\|^2_{B^\mu_{2,2}}. \tag{14}$$

This establishes the usual connection between the Sobolev norm $\|y\|_\mu$ from the previous sections and the Besov norm. Below we shall work with the Besov norm $\|y\|_{B^\mu_{2,2}}$. For brevity we let $y_k := (y_{k,j})_{j=0}^{2^k-1} \in \mathbb{R}^{2^k},\ k = 0, 1, 2, \ldots,$ and $\|y_k\|_p,\ 1 \leqslant p \leqslant \infty$, will mean the standard l_p-norm of the vector y_k in $\mathbb{R}^{2^k}$.

We recall some properties of the Haar system, cf. [5, Chap. 3]. If, for a continuous function y we let $S_m(y) := \sum_{k=0}^{m} \sum_{j=0}^{2^k-1} \langle y, \chi_{k,j}\rangle \chi_{k,j}$, then we have convergence $\|y - S_m(y)\|_\infty \to 0$ as $m \to \infty$. But, since the Haar system consists of discontinuous functions, it can be shown that

$$\limsup_{m\to\infty} \max_{0 \leqslant j < 2^m} |\langle y, \chi_{m,j}\rangle| 2^{3m/2} > 0,$$

unless y was constant. Therefore, we need to study functions from spaces W_2^μ for $\mu < \frac{3}{2}$ only.

Our subsequent arguments will be based on the following important observation. For any vector $u = \{u_{k,j}\}$ and summation region $I \subset \{0, 1, \ldots\}$ we have

$$\begin{aligned} \left|\sum_{k\in I} \sum_{j=0}^{2^k-1} u_{k,j}\chi_{k,j}(t)\right| &= \left|\sum_{k\in I} u_{k,j(k)}\chi_{k,j(k)}(t)\right| \\ &\leqslant \sum_{k\in I} 2^{k/2}|u_{k,j(k)}| \leqslant \sum_{k\in I} 2^{k/2}\|u_k\|_\infty. \end{aligned} \tag{15}$$

For the Haar system, $\mu > \frac{1}{2}$ is sufficient for a tail estimate as in Propositions 2.1 and 3.1; and we will restrict ourselves to $\frac{1}{2} < \mu < \frac{3}{2}$ below. However, since the Haar system does not consist of continuous functions we cannot guarantee, that functions from W_2^μ are continuous.

PROPOSITION 4.1. *For $\mu > \frac{1}{2}$ and any m it holds the estimate*

$$\left\|\sum_{k=m+1}^{\infty} \sum_{j=0}^{2^k-1} y_{k,j}\chi_{k,j}(t)\right\|_\infty \leqslant C(2^{-m})^{(\mu-1/2)}\|y\|_\mu.$$

Proof. Using (15) we infer for any fixed $t \in (0,1)$

$$\left| \sum_{k>m} \sum_{j=0}^{2^k-1} y_{k,j}\chi_{k,j}(t) \right| \leqslant \sum_{k>m} 2^{k/2} \, ||y_k||_\infty \leqslant \sum_{k>m} 2^{k/2} 2^{-k\mu} 2^{k\mu} ||y_k||_2$$
$$\leqslant \left(\sum_{k>m} 2^{(1-2\mu)k} \right)^{1/2} ||y||_\mu \leqslant C(2^m)^{(1/2-\mu)} ||y||_\mu. \qquad \blacksquare$$

Remark 4.1. Since $||\chi_{k,j}||_\infty = 2^{k/2}$, the Haar system obeys (K^β) with $\beta = \frac{1}{2}$. Therefore, Theorem 3.1 does not apply for $\frac{1}{2} < \mu < 1$.

In order to establish the main result of this section we need to rewrite the assumptions on the observations y_δ. We assume that we are given noisy observations

$$y_{\delta,k,j} \coloneqq \langle y, \chi_{k,j} \rangle + \delta \xi_{k,j}, \qquad j = 0, \ldots, 2^k - 1, \; k = 0, 1, 2 \ldots. \tag{16}$$

For deterministic noise $\xi = \{\xi_{k,j}\}$ we assume $||\xi_{k,j}||_{l_2} \leqslant 1$. Again, for random noise we assume, that all $\xi_{k,j}$ are i.i.d. standard normal variables.

Remark 4.2. The recovery of continuous functions from noisy coefficients with respect to the Haar system was studied by Agayan and Bayadyan [1]. The assumption of these authors made on the functions does not directly fit our classes W_2^μ. In terms of the Besov norms these assumptions can be expressed as belonging to $B_{p,1}^{1/2}$, which, for $p \geqslant 2$, is slightly more restrictive than belonging to $W_2^{1/2} = B_{2,2}^{1/2}$; this still ensures, that tails as in Proposition 4.1 tend to 0. In any case, these authors establish convergence of some proper Tikhonov regularization, but no rate of convergence can be deduced. Our assumption is more restrictive, but allows uniform estimates of the accuracy.

It remains to rewrite the properties of stable summation methods for the present setup. For a sequence $a = (\alpha_{k,j}), \; j = 0, 1, \ldots, 2^k - 1, \; k = 0, 1, \ldots$ and truncation level m we recall (with a slight abuse of notation) that T_m^a has degree λ, if $|1 - \alpha_{k,j}| \leqslant C 2^{\lambda(k-m)}$, for $k \leqslant m$ and $j < 2^k$. Now we are ready to state the main result of this section.

THEOREM 4.1. *Let $\frac{1}{2} < \mu < \frac{3}{2}$ and T_m^a be any summation method of degree $\lambda > \mu - \frac{1}{2}$. We have for $m = \frac{1}{\mu}\log_2(1/\delta) + o(1)$ as $\delta \to 0$ and*

1. For deterministic noise

$$\sup_{\|y\|_\mu \leqslant 1} \sup_{\|\xi\|_{l_2} \leqslant 1} \|y - T_m^a(y_\delta)\|_\infty \leqslant C\delta^{(\mu-1/2)/\mu}. \tag{17}$$

2. For Gaussian white noise: *we have*

$$\sup_{\|y\|_\mu \leqslant 1} \mathbf{E}\,\|y - T_m^a(y_\delta)\|_\infty \leqslant C(\delta\sqrt{\log(1/\delta)})^{(\mu-1/2)/\mu}. \tag{18}$$

Proof. Again the proof uses the basic decomposition of the error, corresponding to (7), which here rewrites as

$$\begin{aligned}\|y - T_m^a(y_\delta)\|_\infty \leqslant & \left\| \sum_{k=0}^{m} = 73) \sum_{j=0}^{2^k-1} (1-\alpha_{k,j}) y_{k,j}\chi_{k,j} \right\|_\infty \\ & + \left\| \sum_{k=m+1}^{\infty} \sum_{j=0}^{2^k-1} y_{k,j}\chi_{k,j} \right\|_\infty + \delta \left\| \sum_{k=0}^{m} \sum_{j=0}^{2^k-1} \alpha_{k,j}\xi_{k,j}\chi_{k,j} \right\|_\infty. \end{aligned} \tag{19}$$

The middle summand in this decomposition can be bounded from above using Proposition 4.1 by $(2^{-m})^{(\mu-1/2)} \asymp \delta^{(\mu-1/2)/\mu}$, by the choice of m. The first summand is independent of the noise. It can be estimated using (15) and taking into account the degree of the summation method T_m^a as follows:

$$\begin{aligned}\left\| \sum_{k=0}^{m} \sum_{j=0}^{2^k-1} (1-\alpha_{k,j}) y_{k,j}\chi_{k,j} \right\|_\infty & \leqslant \sum_{k=0}^{m} 2^{k/2} \max_{0 \leqslant j < 2^k} |1-\alpha_{k,j}||y_{k,j}| \\ & \leqslant C \sum_{k=0}^{m} 2^{k/2} 2^{\lambda(k-m)} 2^{-k\mu} 2^{k\mu} \|y_k\|_2 \\ & \leqslant C \left(\sum_{k=0}^{m} 2^{k(1-2\mu+2\lambda)-2\lambda m} \right)^{1/2} \|y\|_\mu \\ & \leqslant C(2^{-m})^{(\mu-1/2)} \|y\|_\mu \leqslant C\delta^{(\mu-1/2)/\mu} \|y\|_\mu. \end{aligned}$$

It remains to estimate the noise term. For deterministic noise we have by (15) the estimate

$$\left\| \sum_{k=0}^{m} \sum_{j=0}^{2^k-1} \alpha_{k,j}\xi_{k,j}\alpha_{k,j} \right\|_\infty \leqslant \sum_{k=0}^{m} 2^{k/2} \|\xi_k\|_2 \leqslant \left(\sum_{k=0}^{m} 2^k \right)^{1/2},$$

which in turn is bounded by $C2^{m/2}$. Since $\delta 2^{m/2} \asymp \delta^{(\mu-1/2)/\mu}$, the proof of estimate (17) can be accomplished.

For Gaussian white noise we cannot use an argument, similar to the one from the previous sections, since the Haar system does not obey any Lipschitz property. Instead, direct calculations, based on (15) yield

$$\mathbf{E}\left\|\sum_{k=0}^{m}\sum_{j=0}^{2^k-1}\alpha_{k,j}\xi_{k,j}\chi_{k,j}\right\|_{\infty} \leqslant C\sum_{k=0}^{m}2^{k/2}\mathbf{E}\max_{0\leqslant j<2^k}|\xi_{k,j}|.$$

It is well known, that for i.i.d. standard normal variables $(\gamma_j)_{j=1}^{N}$ we have $\mathbf{E}\max_{0\leqslant j<N}|\gamma_j|\leqslant C\sqrt{\log(N+1)}$, see, e.g., [7, Chap. 3.3]; thus $\mathbf{E}\max_{0\leqslant j<2^k}|\xi_{k,j}|\leqslant C\sqrt{k+1}$. This implies

$$\mathbf{E}\left\|\sum_{k=0}^{m}\sum_{j=0}^{2^k-1}\alpha_{k,j}\xi_{k,j}\chi_{k,j}\right\|_{\infty} \leqslant C\sum_{k=1}^{m}2^{k/2}\sqrt{k}\leqslant C\sqrt{m}\sum_{k=1}^{m}2^{k/2}\leqslant C\sqrt{2^m m},$$

which by the choice of m allows to accomplish the proof of the theorem. ∎

Remark 4.3. We note that in Theorem 4.1 the number m for deterministic noise and for Gaussian white noise is the same. Actually, one should take m such that $2^m \approx (\delta)^{(-1/\mu)}$ in the first case, while but for Gaussian white noise m should be such that $2^m \approx (\delta\sqrt{\log(1/\delta)})^{(-1/\mu)}$. So, in the both cases m has the order indicated in Theorem 4.1.

Remark 4.4. Function estimation under Gaussian white noise, based on observations (16) was studied by Donoho and Johnstone [3], who proved an L_2-rate $(\delta\sqrt{\log(1/\delta)})^{\mu/(\mu+1/2)}$, which is slightly better than the estimate in the second assertion of Theorem 4.1, due to a relaxed error criterion.

We also note that Theorem 4.1 indicates the same accuracy as for the trigonometric system, although the Haar system is not uniformly bounded.

REFERENCES

1. S. S. Agayan and G. L. Bayadyan, Stable summation of Fourier–Haar series with approximate coefficients, *Mat. Zametki*, **39** (1986), 136–143, 158.
2. B. Aliev, Estimation of the regularization method for the problem of the summability of the Fourier series, *Dokl. Akad. Nauk Tadzhik. SSR*, 21 (1978), 3–6.
3. D. L. Donoho and I. M. Johnstone, Minimax estimation via wavelet shrinkage, *Ann. Statist.*, 26 (1998), 879–921.
4. V. A. Il'in and E. G. Poznjak, Foundations of Mathematical Analysis, Nauka, Moscow, 1973.

5. B. S. Kashin and A. A. Saakyan, "Orthogonal Series," Amer. Math. Soc., Providence, RI, 1989 (*Transl. from the Russian by Ralph P. Boas. Transl. Ben Silver, Ed.*).
6. N. P. Korneichuk, "Exact Constants in Approximation Theory," Encyclopedia of Mathematics & its Applications, Cambridge Univ. Press, Cambridge, UK, 1991.
7. M. Ledoux and M. Talagrand, Isoperimetry and processes, "*Probability in Banach Spaces*," Springer-Verlag, Berlin, 1991.
8. O. A. Liskovets, Theory and methods of solving ill-posed problems, *in* "Mathematical Analysis," Vol. 20, pp. 116–178, 264, Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Informatsii, Moscow, 1982.
9. F. Natterer, Error bounds for Tikhonov regularization in Hilbert Scales, *Appl. Anal.* 18 (1984), 29–37.
10. S. Z. Rafal'son, The approximation of functions by Fourier-Jacobi sums, *Izv. Vysš. Učebn. Zaved. Matematika* **4** (1968), 54–62.
11. A. Yu. Shadrin, Inequalities of Kolmogorov type and estimates of spline-interpolation for periodic functions of the class W_2^m, *Mat. Zametki* 48 (1990), 132–139.
12. A. N. Tikhonov and V. Y. Arsenin, "Solutions of Ill-Posed Problems," V. H. Winston & Sons, Washington, DC, John Wiley & Sons, New York, 1977 (Translated from the Russian, Preface by translation editor Fritz John, Scripta Series in Mathematics).
13. N. G. Tomin, An application of the interpolation of linear operators to questions of the convergence of series of Fourier coefficients with respect to classical orthogonal polynomials, *Dokl. Akad. Nauk SSSR* **212** (1973), 1074–1077.
14. A. B. Tsybakov, Pointwise and sup–norm sharp adaptive estimation of functions on the Sobolev classes, *Ann. Statist.* 26 (1998), 2420–2469.