Small Deviations in L_2-norm for Gaussian Dependent Sequences

Seok Young Hong * Mikhail Lifshits † Alexander Nazarov ‡

May 6, 2016

Abstract

Let $U = (U_k)_{k\in\mathbb{Z}}$ be a centered Gaussian stationary sequence satisfying some minor regularity condition. We study the asymptotic behavior of its weighted ℓ_2-norm small deviation probabilities. It is shown that

$$\ln \mathbb{P}\left(\sum_{k\in\mathbb{Z}} d_k^2 U_k^2 \le \varepsilon^2\right) \sim -M\varepsilon^{-\frac{2}{2p-1}}, \qquad \text{as } \varepsilon \to 0,$$

whenever

$$d_k \sim d_\pm |k|^{-p} \quad \text{for some } p > \frac{1}{2}, \quad k \to \pm\infty,$$

using the arguments based on the spectral theory of pseudo-differential operators by M. Birman and M. Solomyak. The constant M reflects the dependence structure of U in a non-trivial way, and marks the difference with the well-studied case of the i.i.d. sequences.

1 Introduction

Let $(Y(t))_{t\in T}$ be a centered Gaussian process defined on some parametric measure space (T, μ). Many studies have been devoted to the asymptotic behavior of its small deviation probabilities

$$\mathbb{P}\left(||Y||_2^2 = \int_T |Y(t)|^2 \mu(dt) \le \varepsilon^2\right), \qquad \text{as } \varepsilon \to 0,$$

see e.g. [9, 11, 12, 16, 22, 23, 24], to mention just a small sample. Since by the Karhunen–Loève expansion (see for instance [1, Section 1.4])

$$||Y||_2^2 = \sum_{k=1}^{\infty} d_k^2 X_k^2$$

where $(X_k)_{k\ge 0}$ is a standard Gaussian i.i.d. sequence and d_k^2 are the eigenvalues of the covariance operator of Y, the small deviation probability may be written as

$$\mathbb{P}\left(\sum_{k=1}^{\infty} d_k^2 X_k^2 \le \varepsilon^2\right), \qquad \text{as } \varepsilon \to 0.$$

*Statistical Laboratory, Faculty of Mathematics, University of Cambridge, UK. email syh30@cam.ac.uk.
†St.Petersburg State University, Russia, and Linköping University, Sweden. email mikhail@lifshits.org.
‡St.Petersburg Department of Steklov Institute of Mathematics and St.Petersburg State University, Russia. email al.il.nazarov@gmail.com.

Sharp evaluation of this asymptotics is available when the limiting behavior of the eigenvalues d_k^2 is understood well enough. Moreover, a considerable amount of results is known also for the case where (X_k) is an i.i.d. non-Gaussian sequence, see e.g. [9, 26, 27]. The importance of small deviation probabilities in a broader context and the wide spectrum of their applications are described in the surveys [18, 19]; for an extensive up-to-date bibliography see [20].

In this paper, we move towards a different direction and examine the asymptotic behavior of the small deviation probabilities of *dependent sequences*. That is,

$$\mathbb{P}\left(\sum_{k=1}^{\infty} d_k^2 U_k^2 \leq \varepsilon^2\right), \qquad \text{as } \varepsilon \to 0, \tag{1.1}$$

for some stationary centered Gaussian random sequence $U = (U_k)_{k\in\mathbb{Z}}$ that is dependent and only satisfies some mild regularity condition.

The motivation for looking at this small deviation problem under dependence (1.1) is twofold. First, it is an interesting mathematical question in its own right. The existing literature on small deviation probability for sums of random variables has been strictly confined to the i.i.d. framework, so the dependent case is still an open field of research. Second, there are several potential statistical applications where this extension could be found useful. In functional statistics literature, it is well-known that the convergence rates of nonparametric estimators depend upon the asymptotics of the associated small deviation probabilities, see e.g. [10], [21] and references therein. Yet in many practical situations where the functional variable of interest is discrete-valued, strict independence assumption between the coordinate variables is too restrictive, so the extent to which the existing small deviation results can be feasible is limited and the asymptotics of (1.1) should be understood. We refer the reader to [14] for more details.

Consider a random vector $Z \in \ell_2(\mathbb{Z})$ defined by its coordinates $Z_k = d_k U_k$, $k \in \mathbb{Z}$, where the positive coefficients d_k satisfy the assumption

$$d_k \sim d_{\pm}|k|^{-p}, \quad \text{for some } p > \frac{1}{2}, \quad k \to \pm\infty, \tag{1.2}$$

where at least one of the numbers $d_{\pm}$ is strictly positive. This assumption is typical of the literature on small deviations of Gaussian processes and related matters; see for example [16, 17, 25].

We are interested in the asymptotics of the small deviation probabilities

$$\mathbb{P}\left(||Z||_2 \leq \varepsilon\right) = \mathbb{P}\left(\sum_{k\in\mathbb{Z}} d_k^2 U_k^2 \leq \varepsilon^2\right), \qquad \text{as } \varepsilon \to 0. \tag{1.3}$$

In particular, one wonders to what extent this asymptotics is the same as that for the i.i.d. Gaussian sequence having the same variance with U_k.

One example of mild dependence structure one can think of would be *linear regularity* (in the sense of [6, Chapter VII, p.248] and [15, Chapter 17, p.303]). We say that a stationary sequence $U = (U_k)_{k\in\mathbb{Z}}$ is linearly regular if

$$H_{-\infty} := \bigcap_{m\in\mathbb{Z}} H_m = \{0\},$$

where H_m denotes the closed linear span of $\{U_k\}_{k\leq m}$. It is a type of asymptotic independence condition that roughly means the process has no significant influence from the distant past.

When the process is Gaussian, linear regularity is implied by the class of mixing-type conditions, a popular notion of dependence under which probability theories have been extensively studied in the literature; see e.g. [5] and [8] for the precise definition and a comprehensive review.

Since a consequence of the Wold decomposition theorem suggests that any stationary linearly regular Gaussian sequence admits a causal moving average representation (cf. [6, Chapter VII, Theorem 13]):

$$U_k = \sum_{m=0}^{\infty} a_m X_{k-m} = \sum_{j=-\infty}^{k} a_{k-j} X_j,$$

where $\sum_{m=0}^{\infty} a_m^2 < \infty$ and $(X_j)_{j\in\mathbb{Z}}$ is an i.i.d. standard Gaussian sequence, it follows that many popular dependent processes such as strongly mixing sequences do have such representations.

In the sequel we shall consider a more general assumption than causality, and postulate that

$$U_k = \sum_{m=-\infty}^{\infty} a_m X_{k-m} = \sum_{j=-\infty}^{\infty} a_{k-j} X_j, \tag{1.4}$$

where $(a_m) \in \ell_2(\mathbb{Z})$, and (X_j) is i.i.d. standard Gaussian as above. In fact, this representation exists iff the stationary sequence (U_k) has a spectral density (cf. Remark 2.1 below) but we will not develop this point of view any further.

Our main result is as follows:

Theorem 1.1 *Let a stationary centered Gaussian sequence $(U_k)_{k\in\mathbb{Z}}$ admit a representation* (1.4) *and let the coefficients $(d_k)_{k\in\mathbb{Z}}$ have the asymptotics* (1.2). *For $p < 1$ suppose in addition that $(a_m) \in \ell_r(\mathbb{Z})$ with some $r < 2$. Then*

$$\ln \mathbb{P}\left(\sum_{k\in\mathbb{Z}} d_k^2 U_k^2 \le \varepsilon^2\right) \sim -B_p \left(\frac{C}{\varepsilon^2}\right)^{\frac{1}{2p-1}}, \qquad \text{as } \varepsilon \to 0, \tag{1.5}$$

with the constants

$$B_p = \frac{2p-1}{2}\left(\frac{\pi}{2p\sin\left(\frac{\pi}{2p}\right)}\right)^{\frac{2p}{2p-1}},$$

$$C = \left(\frac{1}{2\pi}\int_0^{2\pi}\Big|\sum_{m=-\infty}^{\infty} a_m\, e^{\mathbf{i}mx}\Big|^{1/p} dx\right)^{2p} \left(d_-^{1/p} + d_+^{1/p}\right)^{2p}. \tag{1.6}$$

Remark 1.2 The power term in the logarithmic small deviation asymptotics is the same as that in the i.i.d. case (characterized by $a_m = a_0 \mathbf{1}_{\{m=0\}}$), but the constant C in front of it depends on the sequence (a_m) in a nontrivial way, no matter how weak the linear dependence in (U_k) is (in other words, how fast a_m decays).

Remark 1.3 We do not know whether the extra assumption on (a_m) for $p < 1$ is essential or purely technical.

Remark 1.4 For sharper results on small deviations, one would need to know a sharper spectral asymptotics (the so-called two-term asymptotics). This seems to be a much harder problem in general.

Remark 1.5 Similar technique can be applied in the study of the weighted L_2-norm small deviations for *continuous time* stationary processes. This will be done elsewhere.

2 Proof of Theorem 1.1

Recall that we have a random vector $Z = (d_k U_k) \in \ell_2(\mathbb{Z})$ and a random vector with independent coordinates $X = (X_j)$, $j \in \mathbb{Z}$. It follows from the definitions that

$$Z = \mathbf{D}U = \mathbf{D}\mathbf{A}X,$$

where $\mathbf{D}$ is the diagonal matrix with elements $d_{kj} = d_k \mathbf{1}_{\{k=j\}}$ and $\mathbf{A}$ is the Toeplitz matrix with elements $a_{kj} = a_{k-j}$. Therefore, the covariance operator of Z that maps $\ell_2(\mathbb{Z})$ into $\ell_2(\mathbb{Z})$ can be expressed as

$$K_Z = cov(Z) = (\mathbf{D}\mathbf{A})(\mathbf{A}^*\mathbf{D}),$$

and by the Karhunen–Loève expansion (see [1, Section 1.4]),

$$||Z||^2 = \sum_{n=1}^{\infty} \lambda_n \, \xi_n^2 \ ,$$

where $(\xi_n)_{n\in\mathbb{N}}$ is an i.i.d. standard Gaussian sequence and $(\lambda_n)_{n\in\mathbb{N}}$ are the eigenvalues of K_Z.

We remark that the small deviations (1.3) depend heavily on the asymptotic behavior of λ_n. In particular, if we can show that

$$\lambda_n \sim C\, n^{-2p}, \qquad \text{as } n \to \infty, \tag{2.7}$$

then (1.5) will follow from [9, p.67] or [28], and [23]. The decay rate for λ_n would then be the same as that of d_n^2, and the constant C in front of the power rate would depend on the sequence (a_m) in a non-trivial way, cf. (1.6).

Therefore it now remains to prove the eigenvalue asymptotics (2.7), and to specify the constant C.

Since all separable Hilbert spaces are isomorphic, we may replace $\ell_2(\mathbb{Z})$ with the more appropriate space $L_2([0, 2\pi], \nu)$ with $\nu(dy) = \frac{dy}{2\pi}$, equipped with the standard exponential basis $e_m(x) = \exp(\mathbf{i}\, mx)$, $m \in \mathbb{Z}$.

Notice that in this space $\mathbf{A}$ becomes the multiplication operator $\mathbf{A}f = \mathfrak{a}f$ related to the function

$$\mathfrak{a}(x) = \sum_{m=-\infty}^{\infty} a_m \, e_m(x),$$

while $\mathbf{D}$ becomes the convolution operator

$$(\mathbf{D}f)(x) = \int_0^{2\pi} \mathfrak{D}(x-y) f(y) \, \nu(dy)$$

with the kernel

$$\mathfrak{D}(x) = \sum_k d_k \, e_k(x). \tag{2.8}$$

Indeed, if $f = \sum_j f_j \, e_j$, then

$$\mathfrak{a}f = \sum_{m,j} a_m f_j \, e_{m+j} = \sum_k \left(\sum_j a_{k-j} f_j \right) e_k$$

and

$$\int_0^{2\pi} \mathfrak{D}(x-y)f(y)\,\nu(dy) \;=\; \sum_{j,k} d_k f_j \int_0^{2\pi} e_k(x-y)e_j(y)\,\nu(dy)$$

$$= \sum_{j,k} d_k f_j e_k(x) \int_0^{2\pi} e_{j-k}(y)\,\nu(dy) = \sum_k d_k f_k e_k(x).$$

Remark 2.1 *Interestingly, $|\mathfrak{a}(\cdot)|^2$ is the spectral density of the stationary sequence (U_k).*

In our spectral analysis, we will first slightly reinforce condition (1.2) by assuming that (d_k) is *exactly* equal to the non-isotropic power function

$$d_k = d(\mathrm{sgn}(k))\ |k|^{-p}, \tag{2.9}$$

where $d(\pm 1) = d_\pm$ are two constants and $d_0 = 0$.

In the sequel, our main argument will be a reduction of the operator $\mathbf{A}^*\mathbf{D}$ to a special case of the pseudo-differential operators (ΨDO) studied by M. Birman and M. Solomyak (hereafter BS) in [2, 3][1], see also [7].

The following exposition provides an interpretation of [2] and [3] adapted to our case. The aim of the papers BS is the spectral analysis of the following operator (in their notation)

$$(\mathbf{F}u)(x) = b(x) \int_{\mathbb{R}^m} \mathfrak{F}(x, x-y)c(y)u(y)dy.$$

Here and elsewhere by spectral analysis of an operator, we understand the study of the asymptotic behavior of its singular values.

In our case the space dimension $m = 1$, and we can assume that the function $\mathfrak{F}$ depends only on the second argument, i.e.

$$(\mathbf{F}u)(x) = b(x) \int_{\mathbb{R}} \mathfrak{F}(x-y)c(y)u(y)dy. \tag{2.10}$$

The kernel $\mathfrak{F}(\cdot)$ in [2] is of specific Fourier transform form, namely,

$$\mathfrak{F} = (\zeta \cdot \mathrm{d})\check{}. \tag{2.11}$$

Here $\zeta(\cdot)$ is any smooth function that vanishes on a neighborhood of zero and equals to one on a neighborhood of infinity, while $\mathrm{d}(\cdot)$ in the one-dimensional case is a homogeneous function as in (2.9) but considered in continuous time, i.e., in the notation of BS

$$\mathrm{d}(\xi) = \mathrm{d}(\mathrm{sgn}(\xi))\ |\xi|^{-\alpha}, \qquad \xi \in \mathbb{R}\backslash\{0\}, \tag{2.12}$$

where $\mathrm{d}(\pm 1) = \mathrm{d}_\pm$ are two constants. For us, the homogeneity index α in (2.12) is p. Notice immediately that the "mysterious" formula (2.11) is, apart from the inessential smoothing term ζ, a version of our former kernel definition (2.8) for continuous time.

[1]The referee mentioned [4] which also provides estimates relevant to small deviations. However, these estimates are not sharp enough to establish the asymptotic behavior of singular values up to equivalence that we need here.

BS consider the operator $\mathbf{F}$ either on $\mathbb{R}^m$ or on a cube. The latter means that the weights b and c in (2.10) are supported by a cube. In our case the weight function $b(\cdot)$ from (2.10) is $\overline{\mathfrak{a}(\cdot)}$, and the function $c(\cdot)$ is the indicator on the interval $[0, 2\pi]$ that plays the role of a cube. Moreover, the index $\mu = \frac{m}{\alpha}$ used by BS for the description of singular values behavior is $\frac{1}{p}$ in our notation. Notice that [2] distinguishes three cases $\mu > 1$, $\mu = 1$ and $\mu < 1$, which in our notation are $p \in (\frac{1}{2}, 1)$, $p = 1$ and $p > 1$, respectively.

The weight size restrictions in [2] are $b \in L_{q_1}$, $c \in L_{q_2}$. Our assumptions give $q_1 = 2$ for $p \geq 1$ and $q_1 = \frac{r}{r-1} > 2$ for $p < 1$ (the latter fact is due to the Hausdorff–Young inequality, see, e.g. [13, § 8.5]). Without loss of generality we can suppose $\frac{r}{r-1} < \frac{1}{p}$. Further, $q_2 \geq 1$ may be taken arbitrarily.

The main results of BS are stated in Theorems 1 and 2 of [2]. Let us first check the weight assumptions of Theorem 1 in [2].

If $p > 1$, then $\mu < 1$ and Theorem 1(b) applies with $q_1 = q_2 = 2$.

If $p = 1$, then $\mu = 1$ and Theorem 1(c) applies with $q_1 = 2$ and any $q_2 > 2$. This case is relevant to Wiener process and its relatives such as Brownian bridge, OU-process etc.

If $p \in (\frac{1}{2}, 1)$, then $2 > \mu > 1$, and Theorem 1(a) applies with $q_1 > 2$ and $q_2 > 2$ chosen from the relation $\frac{1}{q_2} = p - \frac{r-1}{r}$, as required in Theorem 1(a).

Theorem 2 in [2] is disregarded because it requires some extra assumptions and only applies to the case of infinite q_1 or q_2.

Now let us proceed to follow the BS result. They denote the singular values of $\mathbf{F}$ by $s_n(\mathbf{F})$ and study the corresponding *distribution function*

$$N_{\mathbf{F}}(s) := \#\{n : s_n(\mathbf{F}) \geq s\}$$

and its asymptotics at zero. This is indeed an equivalent setting because

$$N_{\mathbf{F}}(s) \sim \Delta \cdot s^{-1/p}, \quad \text{as } s \to 0 \quad \Longleftrightarrow \quad s_n(\mathbf{F}) \sim \Delta^p \cdot n^{-p}, \quad \text{as } n \to \infty. \tag{2.13}$$

Next, BS introduce the following notations

$$\Delta_\mu := \limsup_{s \to 0_+} s^\mu N_{\mathbf{F}}(s), \qquad \delta_\mu := \liminf_{s \to 0_+} s^\mu N_{\mathbf{F}}(s). \tag{2.14}$$

In their Theorem 2 of [2] BS prove that $\Delta_\mu = \delta_\mu$ and find the common value for the upper and the lower limit

$$\lim_{s \to 0_+} s^\mu N_{\mathbf{F}}(s) = \Delta_\mu = \delta_\mu.$$

Namely, they introduce the "operator symbol" $G(s, \xi)$, see formula (14) of [2]. In the one-dimensional case the symbol is a scalar defined by

$$G(x, \xi) = \overline{\mathfrak{a}}(x) \mathrm{d}(\xi) = \overline{\mathfrak{a}}(x) \cdot \mathbf{1}_{[0,2\pi]}(x) \cdot \mathrm{d}(\mathrm{sgn}(\xi))\, |\xi|^{-p}.$$

Further, formula (18) of [2] suggests that in our case (recall that $\mu = \frac{1}{p}$)

$$\begin{aligned}
\Delta_\mu &= (2\pi)^{-1}\int\limits_0^{2\pi}\int\limits_{\mathbb{R}\setminus\{0\}} \mathbf{1}_{\{|G(x,\xi)|\geq 1\}} d\xi dx \\
&= (2\pi)^{-1}\int\limits_0^{2\pi}\int\limits_{\mathbb{R}\setminus\{0\}} \mathbf{1}_{\{|\mathfrak{a}(x)|\,|\mathrm{d}(\mathrm{sgn}(\xi))|\,|\xi|^{-p}\geq 1\}} d\xi dx \\
&= (2\pi)^{-1}\int\limits_0^{2\pi}\int\limits_{\mathbb{R}\setminus\{0\}} \mathbf{1}_{\{|\mathfrak{a}(x)|^{1/p}\,|\mathrm{d}(\mathrm{sgn}(\xi))|^{1/p}\geq |\xi|\}} d\xi dx \\
&= (2\pi)^{-1}\int\limits_0^{2\pi} |\mathfrak{a}(x)|^{1/p} dx \left(|\mathrm{d}(-1)|^{1/p} + |\mathrm{d}(1)|^{1/p}\right).
\end{aligned}$$

Now we compare the spectral behavior of the operator of our interest $\mathbf{A}^*\mathbf{D}$ with that of the operator $\mathbf{F}$ in (2.10), assuming that the parameters $d_\pm$ in (2.9) coincide with their counterparts $\mathrm{d}_\pm$ in (2.12), and substituting $b = \overline{\mathfrak{a}} \cdot \mathbf{1}_{[0,2\pi]}$ and $c = \mathbf{1}_{[0,2\pi]}$ in (2.10).

Let us prove that

$$N_{\mathbf{A}^*\mathbf{D}}(s) \sim N_{\mathbf{F}}(s), \quad \text{as } s \to 0. \tag{2.15}$$

Notice that since we are working on the interval of length 2π, it is sufficient to consider only the restriction of our periodical function $\mathfrak{D}$ to $[-2\pi, 2\pi]$.

Let h be the cut-off function equal to one on $[\frac{3\pi}{2}, 2\pi]$ and zero on $[-2\pi, \pi]$. Then it follows that the function $h_0(x) := 1 - h(x) - h(-x)$ equals to one on $[-\pi, \pi]$ and vanishes outside of the interval $[-\frac{3\pi}{2}, \frac{3\pi}{2}]$.

Comparing the kernels of two operators, we have the following decomposition

$$\mathfrak{D}(x) - \mathfrak{F}(x) = \mathfrak{D}(x)\ \big(h(x) + h(-x)\big) + \mathfrak{D}_1(x), \qquad x \in [-2\pi, 2\pi]. \tag{2.16}$$

We claim that the function $\mathfrak{D}_1 := \mathfrak{D} \cdot h_0(x) - \mathfrak{F}$ satisfies

$$\widehat{\mathfrak{D}_1}(\xi) = o(|\xi|^{-p}) \quad \text{as } |\xi| \to \infty, \tag{2.17}$$

where $\widehat{\mathfrak{D}_1}$ denotes the Fourier transform of $\mathfrak{D}_1$. Indeed, we have

$$\widehat{\mathfrak{D} \cdot h_0}(\xi) = \sum_{k\neq 0} d(\mathrm{sgn}(k))\ |k|^{-p}\widehat{h}_0(\xi - k),$$

and then by spliting the series into two sums,

$$\widehat{\mathfrak{D} \cdot h_0}(\xi) = \Sigma_1 + \Sigma_2 := \left(\sum_{|k-\xi|\leq\sqrt{\xi}} + \sum_{|k-\xi|>\sqrt{\xi}}\right) d(\mathrm{sgn}(k))\ |k|^{-p}\widehat{h}_0(\xi - k).$$

Since $\widehat{h}_0$ rapidly decays at infinity, we have $\Sigma_2 = o(|\xi|^{-p})$ as $|\xi| \to \infty$. Further,

$$\begin{aligned}
\Sigma_1 &= d(\mathrm{sgn}(\xi))\ |\xi|^{-p} \sum_{|k-\xi|\leq\sqrt{\xi}} \widehat{h}_0(\xi - k) + o(|\xi|^{-p}) \\
&= d(\mathrm{sgn}(\xi))\ |\xi|^{-p} \sum_k \widehat{h}_0(\xi - k) + o(|\xi|^{-p}) = d(\mathrm{sgn}(\xi))\ |\xi|^{-p} + o(|\xi|^{-p})
\end{aligned}$$

by the Poisson summation formula (see, e.g., [29, Ch. II, Sect. 13]), so that (2.17) follows.

Decomposition (2.16) generates the corresponding operator representation

$$\mathbf{A}^*\mathbf{D} - \mathbf{F} = (\mathbf{D}_+ + \mathbf{D}_-) + \mathbf{D}_1.$$

By corollary 4) in [3], relation (2.17) gives $\lim_{s\to 0_+} s^{1/p} N_{\mathbf{D}_1}(s) = 0$. Further, since $\mathfrak{D}$ is 2π-periodic, the singular values of $\mathbf{D}_+$ coincide with the singular values of the operator

$$\overline{\mathfrak{a}}(x+\pi)\mathbf{1}_{[0,\pi]}(x) \int\limits_{\mathbb{R}} \mathfrak{D}(x-y)h(x+2\pi-y)\mathbf{1}_{[\pi,2\pi]}(y)u(y)dy.$$

For this operator, we have supp(b) $= [0,\pi]$ and supp(c) $= [\pi, 2\pi]$ in terms of (2.10), and Lemma 3 in [3] gives $\lim_{s\to 0_+} s^{1/p} N_{\mathbf{D}_+}(s) = 0$. By the same reason, $\lim_{s\to 0_+} s^{1/p} N_{\mathbf{D}_-}(s) = 0$, yielding (2.15).

Using the equivalence in (2.13), we obtain

$$\begin{aligned} s_n(\mathbf{A}^*\mathbf{D}) &\sim \Delta_\mu^p n^{-p} \\ &= \left(\frac{1}{2\pi}\int\limits_0^{2\pi} |\mathfrak{a}(x)|^{1/p} dx\right)^p \left(|d(-1)|^{1/p} + |d(1)|^{1/p}\right)^p n^{-p}. \end{aligned}$$

Since $\lambda_n = s_n^2(\mathbf{A}^*\mathbf{D})$ by the definition of singular values, it follows that

$$\lambda_n \sim \left(\frac{1}{2\pi}\int\limits_0^{2\pi} |\mathfrak{a}(x)|^{1/p} dx\right)^{2p} \left(|d(-1)|^{1/p} + |d(1)|^{1/p}\right)^{2p} n^{-2p}, \qquad n \to \infty,$$

as required in (2.7), and the conclusion for small deviations follows.

So far, the result of the theorem is obtained only for the *homogeneous* coefficients (2.9). However, since any finite number of terms in the sequence (d_k) is irrelevant for small deviation probability asymptotics, by monotonicity of the quadratic form $\sum_{k\in\mathbb{Z}} d_k^2 U_k^2$ in (d_k), it follows that (1.5) also holds for any (d_k) satisfying (1.2). $\square$

Acknowledgment. We are grateful to Professor G. Rozenblum for useful advice and for a number of provided references and to the anonymous referee for careful reading, constructive critics and for a number of suggested improvements.

M.A. Lifshits and A.I. Nazarov are supported by RFBR grant 16-01-00258. S.Y. Hong acknowledges financial support from the ERC grant 2008-AdG-NAMSEF.

References

[1] R.B. Ash, M.F. Gardner. *Topics in Stochastic Processes.* Academic Press, New York, 1975.

[2] M.Š. Birman, M.Z. Solomjak. *Asymptotics of the spectrum of pseudodifferential operators with anisotropic-homogeneous symbols*, Vestnik LGU (1977), no 13, 13–21 (Russian); English transl.: Vestnik Leningrad Univ. Math. **10** (1982), 237–247.

[3] M.Š. Birman, M.Z. Solomjak. *Asymptotics of the spectrum of pseudodifferential operators with anisotropic-homogeneous symbols. II*, Vestnik LGU (1979), no 13, 5–10 (Russian); English transl.: Vestnik Leningrad Univ. Math. **12** (1980), 155–161.

[4] M.Š. Birman, M.Z. Solomjak. *Estimates of singular numbers of integral operators*, Uspekhi Mat. Nauk **32** (1977), no 1, 17–84 (Russian); English transl.: Russian Math. Surveys **32** (1977), 15–89.

[5] R.C. Bradley. *Introduction to Strong Mixing Conditions*, Vol. 1–3. Kendrick Press, Heber City, 2007.

[6] A.V. Bulinskii, A.N. Shiryaev. *Theory of Random Processes*, Fizmatlit, 2003 (in Russian).

[7] M. Dauge, D. Robert. *Weyl's formula for a class of pseudodifferential operators with negative order on $L_2(\mathbb{R}^n)$*. In: Proc. Conf. "Pseudo-differential operators", Oberwolfach, 1986; Lecture Notes in Math., v. 1256, Springer-Verlag, 1987, 90–122.

[8] P. Doukhan. *Mixing: Properties and Examples.* Lecture Notes in Statistics 85, Springer-Verlag, 1994.

[9] T. Dunker, M.A. Lifshits, W. Linde. *Small deviations of sums of independent variables.* In: Proc. Conf. High Dimensional Probability; Ser. Progress in Probability, v. 43, Birkhäuser, 1998, 59–74.

[10] F. Ferraty, P. Vieu. *Nonparametric Functional Data Analysis: Theory and Practice*, Springer, New York, 2006.

[11] F. Gao, J. Hannig, T.-Y. Lee, and F. Torcaso. *Laplace transforms via Hadamard factorization with applications to small ball probabilities*, Electronic J. Probab. **8** (2003), paper 13.

[12] F. Gao, J. Hannig, T.-Y. Lee, and F. Torcaso. *Exact L^2-small balls of Gaussian processes*, J. Theoret. Probab. **17** (2004), no. 2, 503–520.

[13] G.H. Hardy, J.E. Littlewood, G. Pólya. *Inequalities*, Cambridge University Press, Cambridge, 1934.

[14] S.Y. Hong, O. Linton. *Asymptotic properties of a Nadaraya–Watson type estimator for regression functions of infinite order.* Preprint `https://arxiv.org/abs/1604.06380`.

[15] I.A. Ibragimov, Yu.V. Linnik. *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff, Groningen, 1971.

[16] A.I. Karol', A.I. Nazarov, Ya.Yu. Nikitin. *Small ball probabilities for Gaussian random fields and tensor products of compact operators*, Trans. Amer. Math. Soc. **360** (2008), no. 3, 1443–1474.

[17] W.V. Li. *Comparison results for the lower tail of Gaussian seminorms*, J. Theor. Probab. **5** (1992), 1–31.

[18] W.V. Li, Q.-M. Shao. *Gaussian processes: inequalities, small ball probabilities and applications*, In: Stochastic Processes: Theory and Methods, Handbook of Statistics (C.R. Rao and D. Shanbhag, eds.), vol. 19, North-Holland/Elsevier, Amsterdam, 2001, pp. 533–597.

[19] M.A. Lifshits. *Asymptotic behavior of small ball probabilities*, In: Probab. Theory and Math. Statist. Proc. VII International Vilnius Conference (1998) (B. Grigelionis, ed.), VSP/TEV. Vilnius, 1999, pp. 453–468.

[20] M.A. Lifshits. *Bibliography of small deviation probabilities*, On the small deviation website `http://www.proba.jussieu.fr/pageperso/smalldev/biblio.pdf`

[21] A. Mas. *Lower bound in regression for functional data by small ball probability representation in Hilbert space*, Electronic J. Statist. **6** (2012), 1745–1778.

[22] A.I. Nazarov. *Exact L_2-small ball asymptotics of Gaussian processes and the spectrum of boundary-value problems*, J. Theor. Probab. **22** (2009), no. 3, 640–665.

[23] A.I. Nazarov. *Log-level comparison principle for small ball probabilities.* Statist. & Probab. Letters **79** (2009), no. 4, 481–486.

[24] A.I. Nazarov, Ya.Yu. Nikitin. *Exact L_2-small ball behavior of integrated Gaussian processes and spectral asymptotics of boundary value problems.* Probab. Theor. Rel. Fields **129** (2004), no. 4, 469–494.

[25] A. Papageorgiou, G.W. Wasilkowski. *On the average complexity of multivariate problems*, J. Complexity **6** (1990), 1–23.

[26] L.V. Rozovsky. *Small deviation probabilities for sums of independent positive random variables*, J. Math. Sci. **147** (2007), no. 4, 6935–6945.

[27] L.V. Rozovsky. *On the behavior of the log Laplace transform of series of weighted non-negative random variables at infinity*, Statist. & Probab. Letters **80** (2010), 764–770.

[28] V.M. Zolotarev. *Asymptotic behavior of Gaussian measure in ℓ_2*, J. Math. Sci. **35** (1986), 2330–2334.

[29] A. Zygmund. *Trigonometrical Series*, Vol.1, Cambridge University Press, Cambridge, 1959.