Electronic Journal of Statistics
Vol. 6 (2012) 2150–2184
ISSN: 1935-7524
DOI: 10.1214/12-EJS742

The sharp lower bound of asymptotic efficiency of estimators in the zone of moderate deviation probabilities

Mikhail Ermakov

Institute of Problems of Mechanical Engineering, RAS and
St. Petersburg State University, St. Petersburg, RUSSIA
e-mail: erm2512@gmail.com

Abstract: For the zone of moderate deviation probabilities the local asymptotic minimax lower bound of asymptotic efficiency of estimators is established. The estimation parameter is multidimensional. The lower bound admits the interpretation as the lower bound of asymptotic efficiency in confidence estimation.

AMS 2000 subject classifications: Primary 62F10, 62F25.
Keywords and phrases: Asymptotic efficiency, confidence estimation, large deviations.

Received February 2012.

1. Introduction

The local asymptotic minimax Theorem [16, 18, 22, 30, 31] allows to study the asymptotic efficiency of estimators in the zone of Central Limit Theorem (CLT) approximation. We do not have information that the values of estimators lie in this zone. Therefore the investigation of asymptotic efficiency of estimators in the zones of large and moderate deviation probabilities is interesting as well.

In the zone of large deviation probabilities the analysis of estimator quality is based on the Bahadur asymptotic efficiency (see [3, 18, 31, 25] and references therein). The moderate deviation probabilities of statistics is also the subject of numerous publications (see [5, 1, 8, 15, 27, 19, 20, 11, 14] and references therein). However their asymptotic efficiency was studied only in [10, 26].

The study of Bahadur asymptotic efficiency of estimators is a rather difficult. This problem is often replaced with the study of local Bahadur asymptotic efficiency. The local Bahadur asymptotic efficiency is a particular case of asymptotic efficiency in the moderate deviation zone.

Let $X_1, \dots, X_n$ be independent sample of random variable X having the probability measure $P_\theta, \theta \in R^1$. Let $b_n > 0, b_n \to 0, nb_n^2 \to \infty$ as $n \to \infty$. Let $\theta_0 \in R^1$. Then (see [10]) for any estimator $\hat\theta_n$

$$\liminf_{n\to\infty} \inf_{\theta=\theta_0,\theta_0+2b_n} (nb_n^2/2)^{-1} \ln P_\theta(|\hat\theta_n - \theta| > b_n) \ge -I(\theta_0).$$

---

*Research partially supported by RFFI Grant 11-10-00577.

Here we suppose that there exists the finite Fisher information $I(\theta_0)$. We note that the calculation of moderate deviation probabilities is the simpler problem [11, 14] than the calculation of large deviation probabilities.

For one-dimensional parameter the sharp lower bound for the asymptotic of moderate deviation probabilities of estimators has been established in [10]. This lower bound represents a version of the local asymptotic minimax Theorem [16, 18, 22, 30, 31] in the moderate deviation zone. The goal of this paper is to obtain similar results for the multidimensional parameter. Thus one can say that the local asymptotic minimax Theorem works in a wider zone than the zone of CLT approximation.

The study of large and moderate deviation probabilities of estimators is closely related to the problem of confidence estimation. For the large samples the asymptotic normality of estimators is the key property allowing to construct the confidence sets. The inequalities of the Berry-Esseen type and the Edgeworth expansions (see [13, 5, 27, 15] and references therein) show that the convergence rate to the normal distribution has the order $n^{-1/2}$ (here n is a sample size). The coverage errors α of confidence sets have usually small values ($\alpha = 0.1; 0.05; 0.01$ are the standard values in practice). For such a slow rate of convergence and for such small values of α the implementation of normal approximation requires additional arguments if the sample sizes is several hundreds observations or smaller.

Thus the problem of asymptotic efficiency of estimators in large and moderate deviation zones can be considered as the problem of asymptotic efficiency of confidence estimation.

The variances of estimators are usually unknown. Therefore it is natural to determine the lower bounds of asymptotic efficiency for the pivotal statistics [21, 17]. This is also the goal of the paper.

We make use of the letters C and c as generic notation for positive constants. Denote $\chi(A)$ the indicator of set A, $[a]$ - the integral part of a. For any $u, v \in R^d$ denote $u'v$ the inner product of u, v and u' the transposed vector of u. For positive sequences a_n denote $a_n \asymp b_n$, if $c < a_n/b_n < C$, and denote $a_n >>> b_n$ if $a_n/b_n \to \infty$ as $n \to \infty$. For any set of events $B_{\dots}$ denote $A_{\dots}$ the complementary event to $B_{\dots}$. For any set $D \subset R^d$ denote ∂D the boundary of D.

2. Main Results

Let $X_1, \dots, X_n$ be i.i.d.r.v.'s having a probability measure (p.m.) $P_\theta, \theta \in \Theta \subseteq R^d$, defined on a probability space (S, Υ). Assume that p.m.'s $P_\theta, \theta \in \Theta$, are absolutely continuous w.r.t. p.m. ν defined on the same probability space (S, Υ). Denote $f(x,\theta) = \frac{dP_\theta}{d\nu}(x), x \in S$. For any $\theta_1, \theta_2 \in \Theta$ denote $P^a_{\theta_1,\theta_2}$ and $P^s_{\theta_1,\theta_2}$ respectively absolutely continuous and singular components of p.m. P_{θ_1} w.r.t. P_{θ_2}. For all $x \in S$ such that $f(x,\theta_1) \neq 0$ denote

$$g(x,\theta_1,\theta_2) = (f(x,\theta_2)/f(x,\theta_1))^{1/2} - 1.$$

The statistical experiment $\Psi = \{(S,\Upsilon), P_\theta, \theta \in \Theta\}$ has the finite Fisher information at the point $\theta \in \Theta$ if there exists the vector function $\phi_\theta(x) = (\phi_{\theta,1}(x),\dots,\phi_{\theta,d}(x))', x \in S, \phi_{\theta,i} \in L_2(P_\theta), 1 \le i \le d$ such that

$$\int_S \left(g(x,\theta,\theta+u) - \frac{1}{2}u'\phi_\theta(x) \right)^2 dP_\theta = o(|u|^2), \quad P^s_{\theta+u,\theta}(S) = o(|u|^2) \qquad (2.1)$$

as $u \to 0$.

The Fisher information matrix at the point θ equals

$$I(\theta) = \int_S \phi_\theta \phi_\theta' \, dP_\theta.$$

For any p.m.'s $P_{\theta_1}, P_{\theta_2}, \theta_1, \theta_2 \in R^d$ the Hellinger distance equals

$$\rho(P_{\theta_1}, P_{\theta_2}) = \rho(\theta_1,\theta_2) = \left(\int_S (f^{1/2}(x,\theta_1) - f^{1/2}(x,\theta_2))^2 \, d\nu \right)^{1/2}.$$

Let Θ be an open set and let $0 < \lambda \le 1$.

We make the following assumptions.

Assumption 2.1. *For all $\theta \in \Theta$ there exists the positively definite Fisher information matrix $I(\theta)$.*

Assumption 2.2. *For all $\theta, \theta+u \in \Theta$ the following inequalities hold*

$$\int_S (g(x,\theta,\theta+u) - \frac{1}{2}u'\phi_\theta(x))^2 \, dP_{\theta_0} < C|u|^{2+\lambda}, \quad P^s_{\theta+u,\theta}(S) < C|u|^{2+\lambda}, \quad (2.2)$$

$$|4\rho^2(\theta,\theta+u) - u'I(\theta)u| < C|u|^{2+\lambda}, \qquad (2.3)$$

$$\int_S |\phi_\theta(x)|^{2+\lambda} \, dP_\theta < C < \infty, \qquad (2.4)$$

$$h'I(\theta)h - h'I(\theta+u)h < C|h|^2|u|^\lambda. \qquad (2.5)$$

The constants C in (2.2-2.5) do not depend on $\theta, \theta+u \in \Theta$.

We say that a set $\Omega \subset R^d$ is central-symmetric if $x \in \Omega$ implies $-x \in \Omega$.

We make also the following assumptions

Assumption 2.3. *The set Ω is convex and central-symmetric.*

The risk asymptotic is depend on the geometry of the set

$$M = \{x : |x| = \inf_{y \in \partial\Omega} |y|, \quad x \in \partial\Omega\}.$$

Assumption 2.4. *There exists a neighborhood V of the set M such that $\partial\Omega \cap V$ is C^2-manifold.*

Assumption 2.5. *The principal curvatures of $\partial\Omega$ at each point of M are negative.*

Denote ζ-a Gaussian random vector in R^d such that $E\zeta = 0, E[\zeta\zeta'] = I$. Here I is the unit matrix.

Theorem 2.1. *Let Assumptions 2.1- 2.5 be valid. Let* Θ_0 *be bounded open subset of* Θ *and let* $\partial\Theta_0 \subset \Theta$. *Let* $nb_n^2 \to \infty, nb_n^{2+\lambda} \to 0$, $b_n - b_{n-1} = o(n^{-1}b_n^{-1})$ *as* $n \to \infty$. *Then for any estimator* $\hat{\theta}_n = \hat{\theta}_n(X_1, \dots, X_n)$ *we have*

$$\liminf_{n\to\infty} \inf_{\theta_0\in\Theta_0} \sup_{|\theta-\theta_0|<C_nb_n} \frac{P_\theta(I^{1/2}(\theta_0)(\hat{\theta}_n - \theta) \notin b_n\Omega)}{P(\zeta \notin n^{1/2}b_n\Omega)} \geq 1 \tag{2.6}$$

with $C_n \to \infty$ *as* $n \to \infty$.

If $b_n = n^{-1/2}$, Theorem 2.1 is a particular case of the Local Asymptotic Minimax Theorem [16, 18, 22, 30, 31]. Wolfowitz [32] was the first who pointed out the relationship between the lower bounds of (2.6)-type and the problem of asymptotic efficiency in the confidence estimation.

In [10] the statement (2.6) has been established for $\theta \in \Theta \subseteq R^1$ if (2.2)-(2.4) are valid. If $d = 1$, the inequality (2.5) follows from (2.3). Note that (2.5) is fulfilled evidently in the case of location parameter. If (2.5) is not valid, we could not take $I^{1/2}(\theta_0)$ as the constant normalized matrix in (2.6).

The statement (2.6) of Theorem 2.1 contains the infimum over $\theta_0 \in \Theta_0$. In the Local Asymptotic Minimax Theorem [16, 18, 22, 30, 31] the value of θ_0 is fixed. This Theorem is valid if the finite Fisher information $I(\theta_0)$ exists at the fixed point θ_0. The one-dimensional version of Theorem 2.1 was proved also for the fixed point θ_0 (see [10]). The assumptions of one-dimensional version of Theorem 2.1 suppose that the finite Fisher information $I(\theta_0)$ exists at the fixed point θ_0 and (2.2)-(2.4) hold at the point θ_0 as well. We can prove multidimensional version of Theorem 2.1 for the fixed point θ_0 only if the finite Fisher information $I(\theta_0)$ exists in some vicinity of the point θ_0 and (2.2)-(2.5) hold uniformly in some vicinity of the point θ_0.

It suffices to suppose in Theorem 2.1 that assumptions 2.4 and 2.5 hold in some vicinity of the set M.

In confidence estimation the set Ω is usually a ball Ω_r having the center zero and the radius $r > 0$. In this case Theorem 2.1 can be rewritten in a more evident form.

Corollary 2.1. *Let Assumptions of Theorem 2.1 be valid. Let* $\Omega = \Omega_r$.*Then for any estimator* $\hat{\theta}_n = \hat{\theta}_n(X_1, \dots, X_n)$ *we have*

$$\liminf_{n\to\infty} \inf_{\theta_0\in\Theta_0} \sup_{|\theta-\theta_0|<C_nb_n} \frac{P_\theta(I^{1/2}(\theta_0)(\hat{\theta}_n - \theta) \notin b_n\Omega_r)}{2^{d/2-1}\Gamma(d/2)(n^{1/2}b_nr)^{d-2}\exp\{-nb_n^2r^2/2\}} \geq 1$$

with $C_n \to \infty$ *as* $n \to \infty$. *Here* $\Gamma(\cdot)$ *is Euler's gamma function.*

If Ω is the ellipsoid $\Omega_{\sigma,r} = \{\theta : \sum_{i=1}^d \sigma_i^2\theta_i^2 > r^2, \theta = \{\theta_i\}_{i=1}^d, \theta_i \in R^1\}, \sigma = \{\sigma_i\}_{i=1}^d, \sigma_1 = \sigma_2 = \dots = \sigma_k > \sigma_{k+1} > \dots > \sigma_d > 0$, we get the following asymptotic (see [23]) in the denominator of (2.6)

$$P(\zeta \notin n^{1/2}b_n\Omega_{\sigma,r}) = C_k(n^{1/2}b_nr)^{k-2}\exp\{-nb_n^2r^2/(2\sigma_1^2)\}(1 + o(1)).$$

Here $C_k = 2^{1-k/2}\sigma_1^{1-k}(\Gamma(k/2))^{-1}\prod_{i=k+1}^{d}(1-\sigma_r^2/\sigma_1^2)^{-1/2}$.

The assumptions of Theorem 2.1 are rather weak. The sharp asymptotics of moderate deviation probabilities of likelihood ratio were established under the more restrictive assumptions (see [5, 7, 8, 29] and references therein). The proofs of the lower bounds for the moderate deviation probabilities do not require such strong assumptions (see [2, 10]) and they are usually proved more easily than the upper bounds.

The assumptions of Theorem 2.1 are different from the traditional assumption of local asymptotic normality [16, 18, 22, 30, 31]. Thus Theorem 2.1 could not be straightforwardly extended on the models having this property. At the same time the assumptions 2.1, 2.2 represent a slightly more stable form of usual assumptions arising in the proof of local asymptotic normality. This allows us to make use of the technique arising in the proofs of local asymptotic normality and to get the results similar to (2.6) for other models of estimation. This problem will be considered in the sequel.

For the semiparametric estimation the local asymptotic minimax lower bounds in the zone of moderate deviation probabilities have been established in [12]. In [12] the statistical functionals take the values in R^1. The results were based on the assumptions that (2.2)-(2.4) hold uniformly for the families of "least-favourable" distributions. In the case of multidimensional parameter there arises only one additional assumption (2.5). Thus the difference is not very significant.

In confidence estimation of parameter θ the density $f(x,\theta,\psi)$ may depend on additional nuisance parameter $\psi \in \Psi \supset R^{d_1}$. The covariance matrix $H(\theta,\psi)$ of the limit distribution of $n^{1/2}(\hat\theta_n - \theta)$ may also depend on unknown values of parameters θ,ψ. In this case the construction of confidence sets is based on the pivotal statistics $\sqrt{n}H^{-1/2}(\hat\theta,\hat\psi)(\hat\theta_n-\theta)$ or $\sqrt{n}H^{-1/2}(X_1,\dots,X_n)(\hat\theta_n-\theta)$, where $\hat H_n \doteq H(\hat\theta,\hat\psi)$ and $\hat H_n \doteq H(X_1,\dots,X_n)$ are the estimators of $H(\theta,\psi)$. Here $\hat\psi_n$ is an estimator of the nuisance parameter ψ.

The lower bound for asymptotic efficiency of the pivotal statistics is given below in Theorem 2.2.

For all $x \in S$ and all $\theta, \theta+u \in \Theta, \psi \in \Psi$ such that $f(x,\theta,\psi) \neq 0$ denote

$$g(x,\theta,\theta+u) = g(x,\theta,\theta+u,\psi) = (f(x,\theta+u,\psi)/f(x,\theta,\psi))^{1/2} - 1.$$

Make the following assumptions.

Assumption 2.6. *For all $\theta \in \Theta$ and all $\psi \in \Psi$ there exists the positively definite Fisher information matrix*

$$I_\psi(\theta) = \int_S \phi_{\theta,\psi}\phi'_{\theta'\psi}\; dP_{\theta,\psi}$$

where $\phi_\theta = \phi_{\theta,\psi}$ satisfies

$$\int_S \left(g(x,\theta,\theta+u,\psi) - \frac{1}{2}u'\phi_{\theta,\psi}(x)\right)^2 dP_\theta = o(|u|^2), \qquad P^s_{\theta+u,\theta,\psi}(S) = o(|u|^2)$$

as $u \to 0$. Here $P^s_{\theta+u,\theta,\psi}$ is the singular component of p.m. $P_{\theta+u,\psi}$ w.r.t. $P_{\theta,\psi}$.

Assumption 2.7. *For any fixed $\psi \in \Psi$ the statements (2.2)-(2.5) hold with $P_\theta = P_{\theta,\psi}$. The constants C in (2.2)-(2.5) do not depend on $\theta \in \Theta$ and $\psi \in \Psi$.*

Assumption 2.8. *For all $\theta \in \Theta$ and $\psi \in \Psi$ the matrix $H(\theta,\psi)$ is positively definite.*

Assumption 2.9. *For all $\theta, \theta+u \in \Theta$ and $\psi, \psi+v \in \Psi$ the following inequality holds*

$$|h'H(\theta,\psi)h - h'H(\theta+u,\psi+v)h| \le C|h|^2(|u|^\gamma + |v|^\gamma), \quad h \in R^d,$$

with $\gamma \ge \lambda$.

Assumption 2.10. *The boundary $\partial\Omega$ is C^2-manifold.*

Assumption 2.11. *The principal curvatures at each point of $\partial\Omega$ are negative.*

Assumption 2.12. *For all $\theta, \theta+u \in \Theta$ and $\psi, \psi+v \in \Psi$ the following inequality holds*

$$|h'H(\theta,\psi)h - h'H(\theta+u,\psi+v)h| \le C|h|^2(|u|^\gamma + |v|^\gamma), \quad h \in R^d,$$

with $\gamma \ge \lambda$.

Assumption 2.13. *For any $C > 0$ there exists $n_0(C)$ such that, for all $n > n_0(C)$, there holds*

$$\sup_{\theta\in\Theta,\,\psi\in\Psi} (nb_n^2)^{-1} \log P_{\theta,\psi}(|\hat\psi_n - \psi| > a_n) < -C.$$

Here the sequence $a_n > 0$ is such that $a_n^\gamma b_n^{-\lambda} \to \infty$ and $nb_n^2 a_n^\gamma \to 0$ as $n \to \infty$.

In these assumptions we do not suppose that $H(\theta,\psi)$ is covariance matrix of limit distribution of $n^{1/2}(\hat\theta_n - \theta)$.

For any matrix D denote $||D|| = \sup\{|\eta' D\eta| : |\eta| = 1, \eta \in R^d\}$.

If $\hat H = H(X_1,\dots,X_n)$, the assumptions 2.12, 2.13 are replaced with the assumption 2.14.

Assumption 2.14. *There exists a sequence $a_n > 0$ such that $nb_n^2 a_n \to 0$ as $n \to \infty$ and*

$$\limsup_{n\to\infty} \sup_{\theta\in\Theta,\,\psi\in\Psi} (nb_n^2)^{-1} \log P_{\theta,\psi}(||H(X_1,\dots,X_n) - H(\theta,\psi_n)|| > a_n) = -\infty.$$

Theorem 2.2. *Let Assumptions 2.3, 2.6-2.13 be valid. Let Θ and Ψ be bounded open sets. Let the set $\Theta_0 \subset \Theta$ be open and let $\partial\Theta_0 \subset \Theta$. Let $nb_n^2 \to \infty, nb_n^{2+\lambda} \to 0$, $b_n - b_{n-1} = o(n^{-1}b_n^{-1})$ as $n \to \infty$. Then for any estimator $\hat\theta_n$ there holds*

$$\liminf_{n\to\infty} \inf_{\theta_0\in\Theta_0,\,\psi\in\Psi} \sup_{|\theta-\theta_0|<C_n b_n} \frac{P_{\theta,\psi}(\hat H^{-1/2}(\hat\theta_n - \theta) \notin b_n\Omega)}{P(H^{-1/2}(\theta_0,\psi)I^{-1/2}(\theta_0,\psi)\zeta \notin n^{1/2}b_n\Omega)} \ge 1 \tag{2.7}$$

with $C_n \to \infty$ as $n \to \infty$.

If $\hat H = H(X_1,\dots,X_n)$ and Assumptions 2.3, 2.6-2.11, 2.14 are valid, the statement (2.7) holds as well.

Theorem 2.2 is deduced easily from Theorem 2.1 in section 4.

The plan of the proof of Theorem 2.1 is the following. In section 3 we outline the basic steps of the proof of Theorem 2.1. After that the proof is given for the set Ω with the most simple geometry. For the set Ω with arbitrary geometry we point out the differences in the proof at the end of section 3. The key Lemmas 3.1, 3.2 are proved in section 5. The proof of Lemma 3.2 is based on new Theorems 5.1 and 5.2 on large deviation probabilities of sums of independent random vectors. The proofs of Theorems 5.1 and 5.2 are given in section 6. Section 7 contains the proofs of technical Lemmas of sections 3 and 5.

3. Proof of Theorem 2.1

3.1. Notation

To simplify the notation we suppose that θ_0 equals zero. The estimates of all reminder terms are uniform with respect to $\theta_0 \in \Theta_0$. Assume that the matrix $I(\theta_0)$ is the unit.

For any $\theta_1, \theta_2 \in \Theta$ denote

$$\xi_s(\theta_1,\theta_2) = \ln \frac{f(X_s,\theta_2)}{f(X_s,\theta_1)}, \quad \tau_s(\theta_1) = \{\tau_{ks}(\theta_1)\}_1^d = \phi_{\theta_1}(X_s)$$

with $1 \le s \le n$.

We will often omit $\theta = \theta_0$ in notation. For example, we shall write $\xi_s(\theta) = \xi_s(\theta_0,\theta), \tau_s = \tau_s(\theta_0)$. The index s will be omitted for $s = 1$. For example, $\tau = \tau_1(\theta_0)$.

Denote

$$\psi_n = n^{-1/2} I^{-1/2}(\theta_0) \sum_{s=1}^n \tau_s.$$

Note, that $(\theta-\theta_0)' \sum_{s=1}^n \tau_s$ is the stochastic part of the linear approximation of logarithm of likelihood ratio.

3.2. Plan of the proof

The reasoning is based on the standard proof of local asymptotic minimax lower bound [16, 18, 22, 30, 31]. In particular we make use of the fact that the minimax risk exceeds the Bayes one and study the asymptotic of Bayes risks. However, in this setup, the estimates of residual terms of asymptotics of posterior Bayes risks should have the order $o(\exp\{-cnb_n^2\})$. This does not allow to implement the technique of local asymptotic normality

$$\sum_{s=1}^n \xi_s(u_n) - n^{1/2} u_n' I^{1/2} \psi_n + \frac{1}{2} n u_n' I u_n = o_P(1) \tag{3.1}$$

in the zone $|u_n| \le Cb_n$ of moderate deviation probabilities. This is the basic reason of differences in the proof.

Instead of (3.1) we are compelled to prove that, for any $\epsilon > 0$, there holds

$$P\left(\sup_{u\in U_n}\left\{\sum_{s=1}^n \xi_s(u) - n^{1/2}u'I^{1/2}\psi_n + \frac{1}{2}nu'Iu\right\} > \epsilon\right) = o(\exp\{-cnb_n^2\}) \quad (3.2)$$

where U_n is a fairly broad set of parameters. Therefore, the main problem is how to narrow down the set U_n.

The following two facts allowed us to solve this problem.

- The normalized values of posterior Bayes risks tend to a constant in probability.
- In the zone of moderate deviation probabilities the normal approximation [4, 24] holds for the sets of events $\psi_n \in n^{1/2}\Gamma_{ni}$ where the domain Γ_{ni} has a diameter $o(n^{-1}b_n^{-1})$.

Thus we can find the asymptotic of posterior Bayes risks independently for each an event $\psi_n \in n^{1/2}\Gamma_{ni}$, summarize them over i and then to get the lower bound. Fixing the set Γ_{ni} allows us to replace the proof of (3.2) with the statement

$$\begin{aligned}&P\left(\sup_{u\in U_n}\left\{\sum_{s=1}^n \xi_s(u) - n^{1/2}u'I^{1/2}\psi_n + \frac{1}{2}nu'Iu,\right\} > \epsilon, \psi_n \in n^{1/2}\Gamma_{ni}, A_{1ni}\right)\\ &= o\left(\int_{n^{1/2}\Gamma_{ni}} \exp\{-x^2/2\}dx\right)\end{aligned} \quad (3.3)$$

where $P(A_{1ni}) = 1 + o(1)$.

To narrow down the sets U_n we define the lattice Λ_n in the cube $K_{v_n}, v_n = Cb_n$ and split Λ_n into subsets Λ_{nile}. The set Λ_{nile} is the lattice in the union of a finite number of very narrow parallelepipeds K_{nij} whose orientation is given by the position of the set Γ_{ni} relative to θ_0. The problem of Bayes risk minimization is solved independently for each set Λ_{nile} and the results are added.

Note that the proof of (3.3) with $U_n = \Lambda_{nile}$ is based on the "chaining method" together with the inequality

$$\begin{aligned}&P\left(\sum_{s=1}^n \xi_s(\theta_1,\theta_2) - (\theta_2-\theta_1)'\sum_{s=1}^n \tau_{s\theta_1} + \frac{1}{2}n(\theta_2-\theta_1)'I(\theta_2-\theta_1) > \epsilon,\right.\\ &\left.\psi_n \in n^{1/2}\Gamma_{ni}, A_{1n}\right) \le C|\theta_2-\theta_1|^2 b_n^{\lambda}\int_{n^{1/2}\Gamma_{ni}} \exp\{-x^2/2\}dx.\end{aligned} \quad (3.4)$$

To prove (3.4) we implement simultaneously Chebyshev inequality to the first sum in the left-hand side of (3.4) and Theorem on large deviation probabilities for ψ_n. Thus we prove some anisotropic version of Theorem on large deviation probabilities (see Theorem 5.2).

3.3. Notation

Denote $v_n = Cb_n$. Define a sequence $\delta_{1n} = c_{1n}(nb_n)^{-1}$, with $c_{1n} \to 0, c_{1n}^{-3}nb_n^{2+\lambda} \to 0$ as $n \to \infty$. In the cube $K_{v_n} = [-v_n, v_n]^d$ we define a lattice $\Lambda_n = \{h : h = (j_1\delta_{1n}, \dots, j_d\delta_{1n}), -l_n \le j_k \le l_n = [v_n/\delta_{1n}], 1 \le k \le d\}$. Thus $l_n \asymp c_{1n}^{-1}nb_n^2$.

We split the cube $K_{\kappa v_n}, 0 < \kappa < 1$ into the small cubes $\Gamma_{ni} = x_{ni} + (-c_{2n}\delta_{1n}, c_{2n}\delta_{1n}]^d$, where $c_{2n} \to \infty, c_{2n}\delta_{1n} = o(n^{-1}b_n^{-1}), c_{2n}^3 c_{1n}^{-3} n b_n^{2+\lambda} \to 0$ as $n \to \infty, 1 \le i \le m_n = [(\kappa c_{2n}^{-1} C c_{1n}^{-1})^d n^d b_n^{2d}], x_{ni} \in K_{v_n}$.

Suppose C is chosen such that $b_n\Omega \subset K_{(1-\kappa)v_n}$.

For each $x_{ni}, 1 \le i \le m_n$ we define the partition of the cube K_{v_n} on the subsets

$$\begin{aligned} &K_{nij} = K(\theta_{nij}) = \{x : x = \lambda x_{ni} + u + \theta_{nij}, u = \{u_k\}_{k=1}^d, \\ &u \perp x_{ni}, |u_k| \le c_{3n}\delta_{1n}, \lambda \in R^1, u \in R^d\} \cap K_{v_n}, 1 \le j \le m_{1ni}, \end{aligned}$$

where $c_{3n}/c_{2n} \to \infty, c_{3n}\delta_{1n} = o(n^{-1}b_n^{-1}), c_{3n}^3 c_{1n}^{-3} n b_n^{2+\lambda} \to 0$ as $n \to \infty$.

Let us fix i. Suppose x_{ni} is parallel to $e_1 = (1, 0, \dots, 0)'$. This does not cause serious differences in the reasoning. Denote Π_1 the subspace orthogonal to e_1. Suppose the points $\theta_{nij}, 1 \le j \le m_{1ni}$ are chosen so that they form a lattice in $\Pi_1 \cap K_{v_n}$. Define the sets

$$\Lambda_n(\theta_{nij}) = K(\theta_{nij}) \cap \Lambda_n, 1 \le j \le m_{1ni}, \quad \Theta_{ni} = \{\theta : \theta = \theta_{nij}, 1 \le j \le m_{1ni}\}.$$

The risk asymptotic is depend on the set

$$M = \{x : |x| = \inf_{y \in \partial\Omega} |y|, \quad x \in \partial\Omega\}.$$

We begin with the proof of Theorem 2.1 for the two-point case $M = \{-y, y\}, y \in \partial\Omega$. For arbitrary geometry of the set M we are compelled to make use of a rather cumbersome constructions. At the same time the basic part of the proof is the same.

Let θ_{nij_0} be such that $b_n y \in K(\theta_{nij_0})$ Then $-b_n y \in K(-\theta_{nij_0})$. Let us split Θ_{ni} into the subsets

$$\begin{aligned} \Theta_i(k_1, \dots, k_{d-d_1}) = \{\theta : \theta = \theta_{nij_0} + (-1)^{t_2} 2k_2 c_{3n}\delta_{1n} e_2 \\ + \dots + (-1)^{t_d} 2k_d c_{3n}\delta_{1n} e_d;\ t_2, \dots t_d = \pm 1\} \end{aligned}$$

where $0 \le k_2, \dots, k_d < C_{1n}$ with $C_{1n}c_{3n}c_{1n} \to \infty, nC_{1n}^3 c_{3n}^3 c_{1n}^3 b_n^{2+\lambda} \to 0$ as $n \to \infty$.

Denote

$$\tilde{K}_{ni}(k_1, \dots, k_{d-1}) = \cup_{\theta \in \Theta_i(k_1, \dots, k_{d-1})} K(\theta).$$

It will be convenient to number the sets $\tilde{K}_{ni}(k_1, \dots, k_{d-d_1})$ denoting their $\tilde{K}_{ni1}, \dots, \tilde{K}_{nim_{2ni}}$. Denote

$$\Theta_{nie} = \Theta_{ni} \cap \tilde{K}_{nie}, \quad \Lambda_{nie} = \tilde{K}_{nie} \cap \Lambda_n, \quad 1 \le e \le m_{2ni}.$$

Thus Θ_{nie} contains $k = 2^{d-1}$ points, that is, $\Theta_{nie} = \{\theta_j\}_{j=1}^k$.

3.4. Proof for the simple geometry of the set Ω

In this case the problem of risk minimization on Λ_n is reduced to the same problems on the subsets Λ_{nie}. Thus we have

$$\begin{aligned}
&\inf_{\hat{\theta}_n} \sup_{\theta \in K_{v_n}} P_\theta(\hat{\theta}_n - \theta \notin b_n\Omega) \\
&\ge \inf_{\hat{\theta}_n} (2l_n)^{-d} \sum_{i=1}^{m_n} \sum_{\theta \in \Lambda_n} P_\theta(\hat{\theta}_n - \theta \notin b_n\Omega, \psi_n \in n^{1/2}\Gamma_{ni}) \\
&\ge (2l_n)^{-d} \sum_{i=1}^{m_n} \sum_{e=1}^{m_{2ni}} \inf_{\hat{\theta}_n} \sum_{\theta \in \Lambda_{nie}} P_\theta(\hat{\theta}_n - \theta \notin b_n\Omega, \psi_n \in n^{1/2}\Gamma_{ni}).
\end{aligned} \tag{3.5}$$

Therefore we can minimize the Bayes risk on each subset Λ_{nie} independently and make use of the own linear approximation (3.1) of logarithms of likelihood ratio on each set $U_n = \Lambda_{nie}$.

For the arbitrary geometry of the set M the additional summation over index $l, 1 \le l \le m_{3ni}$ caused the different points of M arises in (3.5). Thus the right-hand side of (3.5) is the following

$$(2l_n)^{-d} \sum_{i=1}^{m_n} \sum_{l=1}^{m_{3ni}} \sum_{e=1}^{m_{2nil}} \inf_{\hat{\theta}_n} \sum_{\theta \in \Lambda_{nile}} P_\theta(\hat{\theta}_n - \theta \notin b_n\Omega, \psi_n \in n^{1/2}\Gamma_{ni}). \tag{3.6}$$

The definition of the sets Λ_{nile} is akin to Λ_{nie}. The statement (3.5) with the right-hand side (3.6) is the basic difference of the proof for the arbitrary geometry of M. For the completeness of the proof we shall write the index l in the further reasoning. This index should be omitted for the two-point case.

The plan of the further proof is the following. First the basic reasoning will be given. After that we define the partitions of Λ_n into the sets Λ_{nile} for the arbitrary geometry of M. The basic reasoning is given on the set of events $A_{1n} \doteq A_{1nile}$ such that

$$P(A_{1nile}) = 1 + O(nb_n^{2+\lambda}). \tag{3.7}$$

The definition of the set A_{1n} is rather cumbersome. To simplify the understanding of the proof we have postponed the definition of the set A_{1n} to the end of section.

For each $\theta \in \Lambda_{nile}$ denote

$$S_{n\theta} = \sum_{s=1}^{n} \xi_s(\theta) - \theta' \sum_{s=1}^{n} \tau_s + 2n\rho^2(0,\theta)$$

and define the events

$$B_{n\theta} = \{X_1, \dots, X_n : S_{n\theta} > \epsilon_{1n}\}$$

where $\epsilon_{1n} \to 0, \epsilon_{1n}^{-2} c_{1n}^{-3} n b_n^{2+\lambda} \to 0$ as $n \to \infty$.

Denote $B_{nile} = \cup_{\theta\in\Lambda_{nile}} B_{n\theta}$. For any $\theta_{nij} \in \Theta_{nile}$ denote $B_{ni}(\theta_{nij}) = \cup_{\theta\in\Lambda(\theta_{nij})} B_{n\theta}$.

We have

$$
\begin{aligned}
&\inf_{\hat\theta_n} \sum_{\theta\in\Lambda_{nile}} P_\theta(\hat\theta_n - \theta \notin b_n\Omega, \psi_n \in n^{1/2}\Gamma_{ni})\\
&\ge \inf_{\hat\theta_n} \sum_{\theta\in\Lambda_{nile}} E\left[\chi(\hat\theta_n - \theta \notin b_n\Omega)\exp\left\{\sum_{s=1}^n \xi_s(\theta)\right\}, \psi_n \in n^{1/2}\Gamma_{ni}, A_{1n}\right]\\
&\ge E\left[\inf_t \sum_{\theta\in\Lambda_{nile}} \chi(t-\theta\notin b_n\Omega)\exp\left\{\theta\sum_{s=1}^n \tau_s - \frac{1}{2}n\theta' I\theta + o(1)\right\},\right.\\
&\left.\psi_n \in n^{1/2}\Gamma_{ni}, A_{nile}|A_{1n}\right] P(A_{1n}) = R_n.
\end{aligned}
\tag{3.8}
$$

Denote $\Delta_n = \exp\{\psi_n'\psi_n/2\}, y = y_\theta = n^{1/2}\theta - \psi_n$. Then, using $nb_n\delta_n \to 0, nb_n^{2+\lambda} \to 0$ as $n\to\infty$, we get

$$
\begin{aligned}
(2l_n)^{-d}R_n &\ge (2l_n)^{-d}E\Bigg[\Delta_n \inf_t \sum_{\theta\in\Lambda_{nile}} \chi(t-y_\theta-\psi_n\notin n^{1/2}b_n\Omega)\exp\left\{-\frac{1}{2}y_\theta' I y_\theta\right\},\\
&\qquad \psi_n \in n^{1/2}\Gamma_{ni}, A_{nile}|A_{1n}\Bigg](1+o(1))\\
&\qquad = (2v_n)^{-d}E\Bigg[\Delta_n \inf_t \int_{n^{1/2}K_{nile}-\psi_n} \chi(t-y\notin n^{1/2}b_n\Omega)\exp\left\{-\frac{1}{2}y' I y\right\}dy,\\
&\qquad \psi_n \in n^{1/2}\Gamma_{ni}, A_{nile}|A_{1n}\}\Bigg](1+o(1)) \doteq (2v_n)^{-d} I_{nile}(1+o(1)).
\end{aligned}
\tag{3.9}
$$

For each $\kappa\in(0,1)$ denote

$$
\begin{aligned}
K_{ni\kappa}(\theta_{nij}) = \{&x : x = \lambda x_{ni} + u + \theta_{nij}, u = \{u_k\}_1^d, |u_k| \le (c_{3n} - Cc_{2n})\delta_{1n},\\
&u \perp x_{ni}, \lambda\in R^1\} \cap K_{(1-\kappa)v_n},
\end{aligned}
$$

$$
K_{nile\kappa} = \cup_{\theta\in\Theta_{nile}} K_{ni\kappa}(\theta).
$$

If $\psi_n \in n^{1/2}\Gamma_{ni} \subset K_{\kappa v_n}$, then $n^{1/2}K_{nile\kappa} \subset n^{1/2}K_{nile} - \psi_n$ and therefore

$$
I_{nile} \ge U_{nile}\bar J_{nile}(1+o(1)) \tag{3.10}
$$

with

$$
U_{nile} = E\left[\Delta_n, \psi_n \in \Gamma_{ni}, A_{nile}|A_{1n}\right],
$$

$$
\bar J_{nile} \doteq \inf_t J_{nile}(t) \doteq \inf_t \int_{n^{1/2}K_{nile\kappa}} \chi(t-y\notin n^{1/2}b_n\Omega)\exp\left\{-\frac{1}{2}y' I y\right\}dy.
$$

Lemma 3.1. *We have*

$$
\bar J_{nile} = J_{nile}(0). \tag{3.11}
$$

Summing over l and e, by (3.11), we get

$$\sum_{l=1}^{m_{3ni}} \sum_{e=1}^{m_{2nil}} \bar{J}_{nile\kappa} \ge P(I^{1/2}(\theta_0)\zeta \notin n^{1/2} b_n \Omega)(1 + o(1)). \tag{3.12}$$

We have

$$\begin{aligned} U_{nile} &= E\left[\Delta_n, \psi_n \in n^{1/2}\Gamma_{ni} | A_{1n}\right] \\ &\quad - E\left[\Delta_n, \psi_n \in n^{1/2}\Gamma_{ni}, B_{nile} | A_{1n}\right] \doteq U_{1ni} - U_{2nile}. \end{aligned} \tag{3.13}$$

Lemma 3.2. *For all $i, 1 \le i \le m_n$, we have*

$$U_{1ni} = mes(\Gamma_{ni})(1 + o(1)), \tag{3.14}$$

$$U_{2nile} = o(mes(\Gamma_{ni})) \tag{3.15}$$

as $n \to \infty$.

Summing over i, by Lemma 3.2, we get

$$\sum_{i=1}^{m_n} U_{nile} \ge \mathrm{mes}(K_{\kappa v_n})(1 + o(1)) = (2\kappa v_n)^d (1 + o(1)). \tag{3.16}$$

By (3.12, 3.16), we get

$$\sum_{i=1}^{m_n} \sum_{l=1}^{m_{3ni}} \sum_{e=1}^{m_{4ni}} \bar{J}_{nile\kappa} U_{nile} \ge (2\kappa v_n)^d P(I^{1/2}(\theta_0)\zeta \notin n^{1/2} b_n \Omega)(1 + o(1)). \tag{3.17}$$

Since $\kappa, 0 < \kappa < 1$, is arbitrary, (3.5), (3.8)-(3.10), (3.17) together imply Theorem 2.1.

3.5. Constructions for the arbitrary geometry of the set Ω

Let us allocate in M connectivity components $M_1, \dots, M_{s_1}$ having the greatest dimension. These components define the asymptotic of lower bound of risks. Denote $\tilde{M} = \cup_{i=1}^{s_1} M_i$. Define the linear manifold N having the smallest dimension d_1 such that $\tilde{M} \subset N$. Define in R^d the coordinate system, such that N is induced the first d_1 coordinates. Denote $e_1, \dots, e_d$ the vectors of the coordinate system.

Denote $y_{nij} \doteq y(\theta_{nij}) \doteq \{x : x = \lambda x_{ni} + \theta_{nij}, \lambda > 0\} \cap b_n \partial\Omega, 1 \le j \le m_{ni}$. Define the sets $Y_{ni} = \{y : y = y_{nij}, 1 \le j \le m_{1ni}\}$. We allocate in Y_{ni} the subset $\tilde{Y}_{ni}$ of all points y_{nij} such that $K(\theta_{nij}) \cap b_n \tilde{M}$ is not empty.

For each $y_{nij} \in \tilde{Y}_{ni}$ we set $z_{nij} \in b_n \tilde{M}$ such that

$$|y_{nij} - z_{nij}| = \inf_{z \in b_n \tilde{M}} |y_{nij} - z|.$$

Define the set $\tilde{Z}_{ni} = \{z : z = z_{nij}, y_{nij} \in \tilde{Y}_{ni}\}$. Denote m_{4ni} the number of points of $\tilde{Z}_{ni}$.

We split $\tilde{Z}_{ni}$ into subsets of points $\tilde{Z}_{nil} = \{z_{nil1}, \dots, z_{nild_1}\}, 1 \le l \le m_{3ni}$ such that the vectors $z_{nil1}, \dots, z_{nild_1}$ induce N. Note that $t < d_1$ points could not enter in these partitions since m_{4ni} may not be a multiple of d_1. However their exception is not essential for the further reasoning. Moreover, for the existence of such a partition we may have to define different constants c_{3n} in the definition of different sets K_{nij}. However, this does not affect significantly on the subsequent proof and we omit the reasoning.

For each z_{nile} define the point $y_{nile}, y_{nile} \in \tilde{Y}_{ni}$ such that $|y_{nile} - z_{nile}| \le c_{3n}\delta_{1n}$.

For each set $\tilde{Z}_{nil} \doteq \{z_{ni_1j_1}, \dots, z_{ni_{d_1}j_{d_1}}\} = \{z_{nil1}, \dots, z_{nild_1}\}$ we make the following. For each point $\theta_{ni_sj_s}, 1 \le s \le d_1$ we define the linear manifold $L_{i_sj_s} = \{z : z = \theta_{ni_sj_s} + \lambda_1 e_{d_1+1} + \dots + \lambda_{d-d_1} e_d,\ \lambda_1, \dots, \lambda_{d-d_1} \in R^1\}$. We split $\Theta_{ni} \cap L_{i_sj_s}$ into the subsets

$$\begin{aligned}\Theta_{i_sj_s}(k_1, \dots, k_{d-d_1}) = \{\theta : \theta = \theta_{ni_sj_s} &+ (-1)^{t_1} 2k_1 c_{3n}\delta_{1n} e_{d_1+1} \\ &+ \dots + (-1)^{t_{d-d_1}} 2k_{d-d_1} c_{3n}\delta_{1n} e_d;\ t_1, \dots t_{d-d_1} = \pm 1\}\end{aligned}$$

where $0 \le k_1, \dots, k_{d-d_1} < C_{1n}$ with $C_{1n}c_{3n}c_{1n} \to \infty, nb_n^{2+\lambda} C_{1n}^3 c_{3n}^3 c_{1n}^3 \to 0$ as $n \to \infty$. Denote

$$\tilde{K}_{i_sj_s}(k_1, \dots, k_{d-d_1}) = \cup_{\theta \in \Theta_{i_sj_s}(k_1, \dots, k_{d-d_1})} K(\theta).$$

Denote $m_{2nil}(i_s, j_s)$ the number of sets $\tilde{K}_{i_sj_s}(k_1, \dots, k_{d-d_1})$.

Without loss of generality we can assume that $m_{2nil}(i_1, j_1) = m_{2nil}(i_2, j_2) = \dots = m_{2nil}(i_d, j_d) \doteq m_{2nil}, 1 \le l \le m_{3ni}$. This can always be achieved by choosing different constants c_{3n} defining the sets K_{nij}. Denote

$$\bar{K}_{nil}(k_1, \dots, k_{d-d_1}) = \cup_{s=1}^{d_1} \tilde{K}_{i_sj_s}(k_1, \dots, k_{d-d_1}).$$

It will be convenient to number the sets $\bar{K}_{nil}(k_1, \dots, k_{d-d_1})$ denoting them $\bar{K}_{nil1}, \dots, \bar{K}_{nilm_{2nil}}$. Denote

$$\Theta_{nile} = \Theta_{ni} \cap \bar{K}_{nile}, \quad \Lambda_{nile} = \bar{K}_{nile} \cap \Lambda_n, \quad 1 \le e \le m_{2nil}.$$

Thus Θ_{nile} contains $d_1 2^{d-d_1}$ points, that is, $\Theta_{nile} = \{\theta_{sj}\}_{s=1,j=1}^{d-d_1,k}, k = 2^{d-d_1}$.

The further proof of Theorem 2.1 follows to the reasoning for the two-point $\{y, -y\}$ geometry of set M given above.

3.6. Definition of the set A_{1n} and Estimate of $P(A_{1n})$

Now the definition of the set $A_{1n} = A_{1nile}$ and the complementary set $B_{1n} = B_{1nile} = D_{nile} \cup B_{4nile} \cup B_{3nile}$ will be given. The definitions of the sets $D_{nile}, B_{4nile}, B_{3nile}$ are given bellow.

For all $s, 1 \le s \le n$, denote $D_{ns}(\theta_{nij}) = \{X_s : f(X_s, 0) \ne 0, f(X_s, \theta) = 0, \theta \ne 0, \theta \in \Lambda_n(\theta_{nij}\}, D_n(\theta_{nij}) = \cup_{s=1}^n D_{ns}(\theta_{nij}), D_{nile} = \cup_{\theta \in \Theta_{nile}} D_n(\theta)$.

Now we define the set $B_{2nile} \subset B_{4nile}$. For any $\theta_1, \theta_2 \in \Theta$ denote $\eta_s(\theta_1, \theta_2) = g(X_s, \theta_1, \theta_2)$ with $1 \le s \le n$. Define the sets of events $B_{2s}(\theta_1, \theta_2) = \{X_s : |\eta_s(\theta_1, \theta_2)| \ge \epsilon\}, B_{2s}(\theta_2) = B_{2s}(0, \theta_2)$ with $0 < \epsilon < \frac{1}{3}$.

For any $\theta \in \Theta_{nile}$ denote $B_{2nis}(\theta) = \cup_{\theta' \in \Lambda_n(\theta)} B_{2s}(\theta')$, $B_{2ni}(\theta) = \cup_{s=1}^n B_{2nis}(\theta)$. Denote $B_{2niles} = \cup_{\theta \in \Theta_{nile}} B_{2nis}(\theta)$, $B_{2nile} = \cup_{s=1}^n B_{2niles}$.

The estimates of $P(B_{2nile})$ are based on the "chaining method". For simplicity we suppose that $l_n = 2^m$. This does not cause serious differences in the reasoning. For each $\theta \in \Theta_{nile}$ we define the sets $\Psi_j = \Psi_j(\theta), 1 \le j \le m$ of points $h_k = \theta + k\delta_{1n}e_1, h_k \in \Lambda_{nile}$, such that $|k|$ is divisible by 2^{m-j} and is not divisible by $2^{m-j+1}, -l_{1n} \le k \le l_{1n}$. Denote $\Psi_{m+1} = \Psi_{m+1}(\theta) = \Lambda_n(\theta) \setminus \cup_{k=1}^m \Psi_k(\theta)$. Denote $\Psi_0(\theta) = \{\theta_0\}$.

We say that the points $h \in \Psi_j$ and $h_1 \in \Psi_{j-1}$ are neighbors if h_1 is the nearest point of Ψ_{j-1} for h. For any $h \in \Psi_j$ we denote $\Pi(h) = \{h_1 : h_1 \in \Psi_{j-1}$ and h, h_1- are neighbors $\}$.

For any $\theta \in \Theta_{nile}$ for each $h \in \Psi_j(\theta), 2 \le j \le m+1$, and all $s, 1 \le s \le n$ define the events

$$V_{hs}(\theta) = \{X_1 : |\eta_s(h_1, h)| > \epsilon j^{-2}, \eta_s(0, h_1) + 1 > \frac{1}{3} - \epsilon \sum_{k=0}^{j} k^{-2}, h_1 \in \Pi(h)\}.$$

Denote

$$B_{4nis}(\theta) = B_{2s}(\theta) \cup \cup_{2 \le j \le m+1} \cup_{h \in \Psi_j(\theta)} V_{hs}(\theta), \quad B_{4niles} = \cup_{\theta \in \Theta_{nile}} B_{4nis}(\theta)$$

and $B_{4nile} = \cup_{s=1}^n B_{4niles}(\theta)$. It is clear that $B_{2nis}(\theta) \subset B_{4nis}(\theta)$.

Lemma 3.3. *We have*

$$P(B_{2nile} \cup D_{nile}) \le P(B_{4nile} \cup D_{nile}) = o(1). \tag{3.18}$$

Define the event $B_{3ns} = \{X_s : |\tau_s| > \epsilon v_n^{-1}\}$. For any $\theta \in \Theta_{nile}$ for each $h \in \Psi_j(\theta), 1 \le j \le m+1$, and all $s, 1 \le s \le n$ define the events

$$B_{3nhs} = \{X_s : |\tau_{sh} - \tau_s| > \epsilon b_n^{-1} 2^{j/2}\}.$$

Denote

$$B_{3nis}(\theta) = B_{3ns} \cup \cup_{2 \le j \le m+1} \cup_{h \in \Psi_j(\theta)} B_{3nhs}(\theta), \quad B_{3niles} = \cup_{\theta \in \Theta_{nile}} B_{3nis}(\theta).$$

and $B_{3nile}(\theta) = \cup_{s=1}^n B_{3niles}$

Lemma 3.4. *We have*

$$P(B_{3nile} \cap A_{4nile}) = o(1).$$

For any $\theta \in \Theta_{nile}$ denote $B_{1ns}(\theta) = B_{4ns}(\theta) \cup B_{3ns}(\theta) \cup D_{ns}(\theta)$. Denote $B_{1n}(\theta) = \cup_{s=1}^n B_{1ns}(\theta), B_{1n} \doteq B_{1nile} = \cup_{\theta \in \Theta_{nile}} B_{1n}(\theta)$.

By Lemmas 3.3 and 3.4, we get (3.7).

4. Proof of Theorem 2.2

Denote $\epsilon_n = \hat{H}^{-1/2} - H^{-1/2}(\theta,\psi)$.

Suppose $\hat{H} = H(\hat{\theta}_n, \hat{\psi}_n)$ and Assumption 2.13 holds.

Choose a sequence δ_n such that $\delta_n = o(n^{-1}b_n^{-1})$ and $\delta_n a_n^{-\gamma} b_n^{-1} \to \infty$ as $n \to \infty$.

By Assumption 2.12,

$$H^{-1/2}(\theta,\psi)(\hat{\theta}_n - \theta) \in (b_n + \delta_n)\Omega, \quad |\hat{\psi}_n - \psi| < ca_n,$$

imply

$$|\epsilon_n| \le C|\hat{\theta}_n - \theta|^\gamma + C|\hat{\psi}_n - \psi|^\gamma < Ca_n^\gamma.$$

Hence

$$H^{-1/2}(\theta,\psi)(\hat{\theta}_n - \theta) \in (b_n + \delta_n)\Omega \ \text{ and } \ |\hat{\psi}_n - \psi| < ca_n$$

imply

$$|\epsilon_n(\hat{\theta}_n - \theta)| \le C\delta_n.$$

Therefore

$$H^{-1/2}(\theta,\psi)(\hat{\theta}_n - \theta) \notin (b_n + \delta_n)\Omega, \quad |\hat{\psi}_n - \psi| < ca_n$$

imply

$$\hat{H}^{-1/2}(\hat{\theta}_n - \theta) \notin b_n\Omega, \quad |\hat{\psi}_n - \psi| < ca_n.$$

Hence, for any $C > 0$ and all $n > n_0(C)$, we have

$$\begin{aligned}
&P_{\theta,\psi}(\hat{H}^{-1/2}(\hat{\theta}_n - \theta) \notin b_n\Omega)\\
&\ge P_{\theta,\psi}(\hat{H}^{-1/2}(\hat{\theta}_n - \theta) \notin b_n\Omega,\ |\hat{\psi}_n - \psi| < ca_n)\\
&\ge P_{\theta,\psi}(H^{-1/2}(\theta,\psi)(\hat{\theta}_n - \theta) \notin (b_n + \delta_n)\Omega,\ |\hat{\psi}_n - \psi| < ca_n)\\
&\ge P_{\theta,\psi}(H^{-1/2}(\theta,\psi)(\hat{\theta}_n - \theta) \notin (b_n + \delta_n)\Omega) - \exp\{-Cnb_n^2\}
\end{aligned} \tag{4.1}$$

where the last inequality follows from Assumption 2.13.

It remains only to implement Theorem 2.1 to the right-hand side of (4.1) to get (2.7).

Suppose that $\hat{H}_n = H(X_1,\dots,X_n)$ and Assumption 2.14 holds. Choose a sequence δ_n such that $\delta_n = o(n^{-1}b_n^{-1})$ and $\delta_n a_n^{-1} b_n^{-1} \to \infty$ as $n \to \infty$.

Note that

$$\hat{H}_n^{-1/2}(\hat{\theta}_n - \theta) \in b_n\Omega \ \text{ and } \ ||\epsilon_n|| < a_n$$

implies

$$H^{-1/2}(\theta,\psi)(\hat{\theta}_n - \theta) \in (b_n + \delta_n)\Omega \ \text{ and } \ ||\epsilon_n|| < a_n.$$

Therefore we can implement similar reasoning and obtain Theorem 2.2 in this case.

5. Proofs of Lemmas 3.1 and 3.2

We begin with the proof of Lemma 3.2.

5.1. Proof of Lemma 3.2

Proof. The proof of (3.14) is based on some version of Osypov-van Bahr Theorems [4, 24] on large deviation probabilities.

Let Z be random vector in R^d such that $E[Z] = 0, \mathrm{Var}(Z) = I$, where I is unit matrix. Let $P(|Z| < \epsilon b_n^{-1}) = 1$, where $\epsilon > 0$. Suppose $E|Z|^{2+\lambda} < C < \infty$. Let $Z_1, \ldots, Z_n$ be independent copies of Z. Denote $S_n = n^{-1/2}(Z_1 + \cdots + Z_n)$.

Denote μ_n the probability measure of Gaussian random vector ζ with $E[\zeta] = 0$ and covariance matrix nI. For any Borel set W denote W_δ the δ- vicinity of $W, \delta > 0$.

Theorem 5.1. *Let the set W belong to a ball in R^d having the radius $r = o(\epsilon_n n^{1/2} b_n)$ where $\epsilon_n \to 0$ as $n \to 0$. Let $nb_n^2 \to \infty, nb_n^{2+\lambda} \to 0$ as $n \to \infty$. Let $W = W_1 \setminus W_2$ where W_1, W_2 are the convex sets. Then*

$$P(S_n \in W) = \mu_n(W)(1 + O(b_n^\lambda)) + O(b_n^\lambda)\mu_n(W_{c_n})$$

where $c_n = o(n^{-1/2} b_n^{\lambda-1})$.

The differences in the statements of Theorem 5.1 and Osypov - van Bahr Theorem [4, 24] are caused the differences in the assumptions. In [4, 24] the results have been proved if $E[\exp\{c|Z|\}] < \infty$.

Let us check up that the assumptions of Theorem 5.1 are fulfilled for the random vector $Z = I^{-1/2}(\theta_0)\tau\chi(A_{1n1})$.

Lemma 5.1. *We have*

$$E[\tau, A_{1n1}] = O(b_n^{1+\lambda}), \tag{5.1}$$

$$E[\tau\tau', A_{1n1}] = I(\theta_0) + O(b_n^\lambda). \tag{5.2}$$

Lemma 5.1 and Theorem 5.1 imply (3.14).

Let us prove (3.15).

Lemma 5.2. *Uniformly in $\theta \in \Lambda_{nile}$ we have*

$$E_\theta[S_{n\theta}|A_{1n}] = o(1). \tag{5.3}$$

Let ϵ_{1n} be such that

$$\sup_{\theta \in \Lambda_{nile}} |E[S_{n\theta}|A_{1n}] \le \frac{\epsilon_{1n}}{4}.$$

Let $h \in \Psi_j, h_1 \in \Pi(h), 2 \le j \le m+1$. We have

$$S_{nh} - E[S_{nh}|A_{1n}] = S_{nh_1} + S_{1nh} + S_{2nh} - E[S_{nh_1} + S_{1nh} + S_{2nh}|A_{1n}]$$

where

$$S_{1nh} = \sum_{s=1}^n \xi_s(h_1, h) - \bar{h}' \sum_{s=1}^n \tau_{sh_1},$$

$$S_{2nh} = \bar{h}' \sum_{s=1}^{n} (\tau_{sh_1} - \tau_s)$$

with $\bar{h} = h - h_1$.

Denote

$$B_{0n} = \{X_1, \ldots, X_n : \sup_{h \in \Psi_1} S_{nh} > \epsilon_{1n}/4\}.$$

For any $h \in \Psi_j, 2 \le j \le m+1$ denote

$$B_{5nh} = \{X_1, \ldots, X_n : j^2(S_{1nh} - E[S_{1nh}|A_{1n}]) > \epsilon_{1n}/4\},$$

$$B_{6nh} = \{X_1, \ldots, X_n : j^2(S_{2nh} - E[S_{2nh}|A_{1n}]) > \epsilon_{1n}/4\}.$$

Denote $B_n = B_{0n} \cup (\cup_{\theta \in \Lambda_{nile} \setminus \Psi_1} (B_{5n\theta} \cup B_{6n\theta}))$. Note that $B_n \supseteq B_{nile}$. Hence

$$U_{2nile} \le U_{3nile} \doteq E\left[\Delta_n, \psi_n \in n^{1/2}\Gamma_{ni}, B_n | A_{1n}\right]. \tag{5.4}$$

Denote $r_{ni} = \inf_{x \in \Gamma_{ni}} |x|$.

We have

$$U_{3nile} \le C \exp\{n r_{ni}^2/2\} \left(V_{0n} + \sum_{\theta \in \Lambda_{1nile}} (V_{5n\theta} + V_{6n\theta}) \right) \tag{5.5}$$

where $\Lambda_{1nile} = \Lambda_{nile} \setminus \Theta_{nile}$,

$$V_{en\theta} = P\left(\psi_n \in n^{1/2}\Gamma_{ni}, B_{en\theta} \,|\, A_{1n}\right), \quad e = 5, 6,$$

$$V_{0n} = P\left(\psi_n \in n^{1/2}\Gamma_{ni}, B_{0n} \,|\, A_{1n}\right).$$

Lemma 5.3. *Let ζ Gaussian random vector having the covariance matrix $I(\theta_0)$ and let $E[\zeta] = 0$. Then for any $h \in \Psi_j, h_1 \in \Pi(h)$ we have*

$$V_{0n} \le C n b_n^{2+\lambda} \epsilon_{1n}^{-2} P(\zeta \in n^{1/2}\Gamma_{ni}), \tag{5.6}$$

$$V_{5nh} \le C n |\bar{h}|^2 b_n^{\lambda} \epsilon_{1n}^{-2} j^4 P(\zeta \in n^{1/2}\Gamma_{ni}), \tag{5.7}$$

$$V_{6nh} \le C n |\bar{h}|^2 b_n^{\lambda} \epsilon_{1n}^{-2} j^4 P(\zeta \in n^{1/2}\Gamma_{ni}). \tag{5.8}$$

The number of points of $\Psi_j, 1 \le j \le m$, equals 2^j and, if $h \in \Psi_j$, then $\bar{h} = b_n 2^{-j}$. The number of points of Ψ_{m+1} equals $C c_{3n}^{d-1} 2^m$ and, if $h \in \Psi_{m+1}$, then $|\bar{h}| \le C c_{3n} \delta_{1n}$. Hence, by (5.5) and Lemma 5.3, we get

$$\begin{aligned} U_{3nile} &\le C n \epsilon_{1n}^{-2} \exp\{n r_{ni}^2/2\} P(\zeta \in n^{1/2}\Gamma_{ni}) \\ &\quad \times \left(b_n^{2+\lambda} + b_n^{\lambda} \left(\sum_{j=1}^{m} 2^j (b_n 2^{-j})^2 j^4 + c_{3n}^{d+1} m^4 2^m \delta_{1n}^2 \right) \right). \end{aligned} \tag{5.9}$$

Note that m satisfies $\delta_{1n} = v_n 2^{-m}$ or $2^m = Cc_{1n}^{-1} n b_n^2 (1 + o(1))$. Hence

$$\begin{aligned} n\epsilon_{1n}^{-2} b_n^{\lambda} c_{3n}^{d+1} m^4 2^m \delta_{1n}^2 &= Cn\epsilon_{1n}^{-2} b_n^{\lambda} c_{3n}^{d+1} c_{1n}^{-1} n b_n^2 m^4 c_{1n}^{-2} n^{-2} b_n^{-2} \\ &= C\epsilon_{1n}^{-2} b_n^{\lambda} c_{3n}^{d+1} c_{1n}^{-3} m^4 = o(1). \end{aligned} \tag{5.10}$$

By (5.9, 5.10), we get

$$U_{3nile} = o(mes(\Gamma_{ni})). \tag{5.11}$$

By (5.4) and (5.11), we get (3.15). This completes the proof of (3.15) □

Proof. Proof of Lemma 5.3 is based on Theorem 5.2.

Theorem 5.2. *Let $V = (X, Z)$ be a random vector $V = (X, Z)$ where random variable X and random vector $Z = (Z_1, \dots, Z_d)$ are such that $E[V] = 0$. Let*

$$P(|X| < \epsilon) = 1, \quad E[|X|^2] < Cb_n^{2+\lambda}, \tag{5.12}$$

$$P(|Z| < \epsilon b_n^{-1}) = 1, \quad E[|Z|^{2+\lambda}] < C < \infty, \tag{5.13}$$

$$E[XZ_k] = O(b_n^{1+\lambda}), \quad 1 \le k \le d \tag{5.14}$$

with $0 < \epsilon < 1$. Suppose the covariance matrix of random vector Z is positively definite.

Let $V_1 = (X_1, Z_1), \dots, V_n = (X_n, Z_n)$ be independent copies of the random vector V. Let U be a bounded set in R^d being a difference of two convex sets.

Denote $S_{nX} = n^{-1/2}(X_1 + \cdots + X_n)$ and $S_n = n^{-1/2}(Z_1 + \cdots + Z_n)$. Denote Y the Gaussian random vector having the same covariance matrix as the random vector Z.

Then, for all sufficiently large n, we have

$$I \doteq P(S_{nX} > \epsilon_{1n}, S_n \in nb_n v + r_n U) \le CP(S_{nX} > \epsilon_{1n}) P(Y \in nb_n v + r_n U)$$

where ϵ_{1n}, r_n are chosen such that $nb_n^{2+\lambda} c_{n1}^{-3} \epsilon_{1n}^{-2} \to 0$ as $n \to \infty$ and $r_n > c_{n1} n^{-1/2} b_n^{-1}$.

It is clear that ϵ_{1n}, r_n can be chosen such that $\epsilon_{1n} \to 0, r_n n^{1/2} b_n \to 0$ as $n \to \infty$. In the proof of (5.7, 5.8) we suppose that ϵ_{1n} and r_n satisfy these assumptions.

For the estimates of V_{5nh} in (5.7) we implement Theorem 5.2 with $Z = \tau$ and

$$X = \varphi(h_1, h) = \xi(h_1, h) - \bar{h}'\tau_{h_1} - \sum_{k=1}^{d} \rho_{kh_1h}\tau_k.$$

Here $\tau = \{\tau_k\}_{k=1}^d$ and $\rho_{h_1h} = \{\rho_{kh_1h}\}_{k=1}^d = r_{h_1h}(E[\tau\tau'|A_{1n1}])^{-1}$ with $r_{h_1h} = \{r_{kh_1h}\}_{k=1}^d, r_{kh_1h} = E[(\xi(h_1, h) - \bar{h}'\tau_{h_1})\tau_k|A_{1n1}]$.

Thus S_{1nh} is replaced by

$$S_{nx} = S_{1nh} - \sum_{s=1}^{n}\sum_{k=1}^{d} \rho_{kh_1h}\tau_{ks} = \sum_{s=1}^{n} \varphi_s(h_1, h).$$

It is easy to see that $E[\varphi(h_1, h)\tau_k | A_{1n1}] = 0, 1 \le k \le d$. This implies (5.14).
Now we show that

$$\sum_{s=1}^{n} \sum_{k=1}^{d} \rho_{kh_1h} \tau_{ks} = o(1) \tag{5.15}$$

if $\psi_n \in n^{1/2}\Gamma_{ni}$ This justifies such a replacement.

By Lemma 5.4 given below, we get $|r_{kh_1h}| \le C|\bar{h}|^{1+\lambda/2}$, if $2 \le k \le d$. Hence, since $\psi_n \in n^{1/2}\Gamma_{ni}$, we get

$$r_{kh_1h} \sum_{s=1}^{n} \tau_{ks} = O(|\bar{h}|^{1+\lambda/2} b_n^{-1}) = o(1) \tag{5.16}$$

with $2 \le k \le d$.

Lemma 5.4. *Let $h \in \Psi_j(\theta), 1 \le j \le m+1,\ h_1 \in \Pi(h)$ and let $v \perp \bar{h}, u \in R^d$. Then*

$$E[(\xi(h_1, h) - \bar{h}'\tau_{h_1})(v'\tau), A_{1n1}] = O(|v||\bar{h}|^{1+\lambda/2}).$$

By Lemma 5.5 given below $|r_{1h_1h}| \le C|\bar{h}|b_n^{\lambda}$. Hence, since $\psi_n \in n^{1/2}\Gamma_{ni}$, we have

$$r_{1h_1h} \sum_{s=1}^{n} \tau_{1s} = O(n|\bar{h}|b_n^{1+\lambda}) = o(1). \tag{5.17}$$

By (2.5), (5.16), (5.17), we get (5.15).

Lemma 5.5. *Let $h \in \Psi_j(\theta), 1 \le j \le m+1,\ h_1 \in \Pi(h)$ and let $v \parallel \bar{h}$. Then*

$$E[(\xi(h_1, h) - \bar{h}'\tau_{h_1})(v'\tau), A_{1n1}] = O(|v||\bar{h}|b_n^{\lambda}). \tag{5.18}$$

Note that

$$2\eta(h_1, h) - 2\eta^2(h_1, h) \le \xi(h_1, h) \le 2\eta(h_1, h) < 2\epsilon \tag{5.19}$$

if A_{1n1} holds.

By (5.19) and Lemma 5.6 given below, we get (5.12).

Lemma 5.6. *For all $\theta \in \Lambda_{nile}$ we have*

$$E[(\xi(\theta) - \theta'\tau)^2, A_{1n1}] = O(|\theta|^{2+\lambda}). \tag{5.20}$$

Let $h \in \Psi_j(\theta), 1 \le j \le m+1,\ h_1 \in \Pi(h)$. Then

$$E[(\xi(h_1, h) - \bar{h}'\tau_{h_1})^2, A_{1n1}] = O(|\bar{h}|^{2+\lambda}). \tag{5.21}$$

This completes the proof of (5.7).

The proof of (5.6) is akin to the proof of (5.7) and is omitted.

For the estimates of V_{6nh} in (5.8) we choose $Z = \tau$ and

$$X \doteq \bar{h}'(\tau_{h_1} - \tau) - \sum_{k=1}^{d} \bar{\rho}_{kh_1h} \tau_k.$$

Here $\tau = \{\tau_k\}_{k=1}^d$ and $\bar\rho_{kh_1h} = \{\bar\rho_{kh_1h}\}_{k=1}^d = \bar r_{h_1h}(E[\tau\tau'|A_{1n1}])^{-1}$ with $\bar r_{h_1h} = \{\bar r_{kh_1h}\}_{k=1}^d, \bar r_{kh_1h} = E[\bar h'(\tau_{h_1} - \tau)\tau_k|A_{1n1}], 1 \le k \le d$.

Using the same reasoning as in the proof of (5.7) and Lemmas 5.7, 5.8 given below we get (5.8).

Lemma 5.7. *Let $u, h \in R^d$. Then*

$$E[(u'(\tau - \tau_h))^2, A_{1n1}] = O(|u|^2|h|^\lambda). \tag{5.22}$$

Lemma 5.8. *Let $h \in \Psi_j(\theta), 1 \le j \le m+1, h_1 \in \Pi(h)$. Let $v \perp \bar h, v \in R^d$. Then*

$$E[\bar h'(\tau_{h_1} - \tau)(v'\tau), A_{1n1}] = O(|v||\bar h||h_1|^{\lambda/2}). \tag{5.23}$$

If $v \parallel \bar h$, we have

$$E[\bar h'(\tau_{h_1} - \tau)(v'\tau), A_{1n1}] = O(|v||\bar h||h_1|^{\lambda}). \tag{5.24}$$

□

5.2. Proof of Lemma 3.1

Proof. The set Λ_{nile} is defined by the set of the points $\Theta_{nile} = \{\theta_{sj}\}_{s,j=1}^{d_1,k}, k = 2^{d-d_1}$. The reasoning first will be given for $|t| < c < \infty$. Denote $n^{1/2}y_{sj}(t) \in (n^{1/2}b_n\partial\Omega - t) \cap (n^{1/2}K(\theta_{sj}))$ the point to which $n^{1/2}y_{sj} = n^{1/2}y(\theta_{sj})$ will pass at the shift t. Denote $n^{1/2}y_{s+d_1,j}(t) \in (n^{1/2}b_n\partial\Omega - t) \cap (n^{1/2}K(\theta_{sj}))$ the point to which $n^{1/2}y_{d_1+s} = -n^{1/2}y_{sj}$ will pass at the shift t.

Lemma 5.9. *The following inequality holds*

$$\sum_{s=1}^{2d_1}\sum_{j=1}^{k} \exp\left\{-\frac{1}{2}n|y_{sj}(t)|^2\right\} \ge 2\sum_{s=1}^{d_1}\sum_{j=1}^{k} \exp\left\{-\frac{1}{2}n|y_{sj}|^2\right\}. \tag{5.25}$$

Proof. For a while we fix $s \le d_1$ and j. We slightly modify the coordinate system for the further reasoning. Suppose $x_{ni} = (1, \beta_2, \dots, \beta_d)$ and $y_{sj} = (b_n, 0, \dots, 0, \delta_{d_1+1,n}n^{-1/2}, \dots, \delta_{dn}n^{-1/2})(1 + o(n^{-1/2}b_n^{-1}))$ with $\delta_{kn} \in R^1, d_1 + 1 \le k \le d$.

Define the line $y = n^{1/2}(y_{sj} + ux_{ni}), \quad u \in R^1$, that is,

$$y_1 = n^{1/2}b_n + u, y_2 = \beta_2 u, \dots, x_{d_1} = \beta_{d_1}u,$$

$$y_{d_1+1} = \delta_{d_1+1,n} + \beta_{d_1+1}u, \dots, y_d = \delta_{d,n} + \beta_d u, \quad |\delta_{kn}| < C, d_1+1 \le k \le d, u \in R^1.$$

Denote $\delta_{kn} = 0$ for $1 < k \le d_1$.

Since the reasoning is given in a sufficiently small vicinity of point $n^{1/2}y_{sj}$ the surface $n^{1/2}b_n\partial\Omega$ admits the approximation in this vicinity by an ellipsoid

$$(x_1 - n^{1/2}b_n)^2 + \alpha_2 x_2^2 + \dots + \alpha_d x_d^2 = nb_n^2$$

where $-\alpha_2, \dots, -\alpha_d$ are the principal curvatures of the surface $\partial\Omega$ at the point $(1, 0, \dots, 0)$. Thus, in the further reasoning, we can replace the set $n^{1/2} b_n \partial\Omega$ with the ellipsoid. After the shift $t = (t_1, \dots, t_d)$ the ellipsoid is defined by the equation

$$(x_1 - n^{1/2} b_n + t_1)^2 + \alpha_2 (x_2 + t_2)^2 + \dots + \alpha_d (x_d + t_d)^2 = n b_n^2.$$

It intersects the line $y = n^{1/2}(\theta_{sj} + u x_{ni}), u \in R^1$ at the point $n^{1/2} y_{sj}(t)$ having the coordinates

$$n^{1/2} y_1(t) = n^{1/2} b_n - t_1 + \omega_{1n}, n^{1/2} y_k(t) = \delta_{kn} - \beta_2 t_1 + \beta_2 \omega_{1n}, \quad 1 < k \le d. \tag{5.26}$$

with

$$\omega_{1n} = -(2n^{1/2} b_n)^{-1} (\alpha_2 (\delta_{2n} + t_2 - \beta_2 t_1)^2 + \dots + \alpha_d (\delta_{dn} + t_d - \beta_d t_1)^2)(1 + o(1)). \tag{5.27}$$

Arguing similarly we get that the ellipsoid intersects the line $y = n^{1/2}(-y_{sj} + u x_{ni}), u \in R^1$ at the point $n^{1/2} y_{s+d_1,j}(t)$ having the coordinates

$$n^{1/2} y_1'(t) = -n^{1/2} b_n - t_1 + \omega_{2n}, \quad n^{1/2} y_s'(t) = -\delta_{kn} - \beta_k t_1 + \beta_k \omega_{2n} \quad 1 < k \le d_1 \tag{5.28}$$

with

$$\omega_{2n} = (2n^{1/2} b_n)^{-1} (\alpha_2 (-\delta_{2n} + t_2 - \beta_2 t_1)^2 + \dots + \alpha_d (-\delta_{dn} + t_d - \beta_d t_1)^2)(1 + o(1)). \tag{5.29}$$

Substituting (5.26, 5.28) in (5.25) we find that, if $t_1 >>> n^{-1/2} b_n^{-1}$, then

$$\begin{aligned} &\max\{\exp\{-n(y_1(t)^2)/2\}, \exp\{-n(y_1'(t)^2)/2\}\} \\ &>>> \exp\{-(n b_n^2 + \delta_{d_1+1}^2 \dots + \delta_d^2)/2\}. \end{aligned}$$

Thus we can suppose that $t_1 < c n^{-1/2} b_n^{-1}$ and neglect the terms $\beta_i t_1, 2 \le i \le d$ in (5.27, 5.29).

Using (5.26, 5.28), we get

$$\begin{aligned} &\exp\left\{-\frac{1}{2} n |y_{sj}(t)|^2\right\} + \exp\left\{-\frac{1}{2} n |y_{s+d_1,j}(t)|^2\right\} \\ &= \exp\{-n|y_{sj}|^2/2\} \left(\exp\left\{ n^{1/2} b_n t_1 + \sum_{k=d_1+1}^{d} \alpha_k t_k \delta_{kn} \right\} \right. \\ &\left. + \exp\left\{ -n^{1/2} b_n t_1 - \sum_{k=d_1+1}^{d} \alpha_k t_k \delta_{kn} \right\} \right) \exp\left\{ \frac{1}{2} \sum_{k=d_1+1}^{d} \alpha_k t_k^2 \right\} (1 + o(1)). \end{aligned}$$

Taking the points $y_{sj}, 1 \le j \le 2^{d-d_1}$, with all possible values $\pm\delta_{kn}, d_1 < k \le d$

and summing up over them $\exp\{-\frac{|y^2_{sj}(t)|^2}{2}\}$, we get

$$\begin{aligned}&\exp\left\{-\frac{nb_n^2+\delta^2_{d_1+1,n}+\cdots\delta^2_{dn}}{2}\right\}\\&\times(\exp\{n^{1/2}b_nt_1\}+\exp\{-n^{1/2}b_nt_1\})\\&\times\prod_{k=d_1+1}^{d}(\exp\{\alpha_kt_k\delta_{kn}\}+\exp\{-\alpha_kt_k\delta_{kn}\})(1+o(1)).\end{aligned}\tag{5.30}$$

Since $\exp\{v\}+\exp\{-v\}-2\ge 0$ for $v\in R^1$, then (5.30) implies (5.25) for $|t|<C$.

In essence, we have considered only the case $u=0$. Any point $y_u=n^{1/2}(y_{sj}+ux_{ni}), 0<u<<<1$, passes to the point $n^{1/2}(y_{sj}(t)+ux_{ni})\in(R^d\setminus(n^{1/2}b_n\Omega-t))\cap(n^{1/2}K(\theta_{sj}))$ at the shift t. Thus for any point $y_u, 0<u<<<1$ we can write a similar inequality (5.25). Since the shift t is negligible, we get

$$\text{mes}((n^{1/2}b_n\partial\Omega)\cap K(\theta_{sj}))=\text{mes}((n^{1/2}b_n\partial\Omega-t)\cap K(\theta_{sj}))(1+o(1)).\tag{5.31}$$

This implies $\bar J_{nile}(t)\ge J_{nile}(0)$. □

Let us consider the case $c<<|t|<<Cn^{1/2}b_n$. Note that, since all the principal curvatures in all points of $\partial\Omega$ are negative, we can conclude $n^{1/2}b_n\Omega$ into an ellipsoid

$$\Xi=\{x=\{x_i\}_{i=1}^d: x_1^2+\cdots+x_{d_1}^2+\bar\alpha_{d_1+1}x^2_{d_1+1}+\cdots+\bar\alpha_dx_d^2=nb_n^2\}$$

passing through the points y_{nile} and $-y_{nile}, 1\le e\le d_1$ and such that $\bar\alpha_k<1, d_1+1\le k\le d$. Denote by $y_{sj}(t)\in(n^{1/2}b_n\partial\Omega-t)\cap\{y:y=\theta_{sj}+x_{ni}u, u\in R^1\}$ and denote by $\bar y_{sj}(t)\in(\Xi-t)\cap\{y:y=\theta_{sj}+x_{ni}u,u\in R^1\}$ the point to which the y_{sj} passes at the shift t.

It is easy to see that

$$\sum_{s=1}^{2d_1}\sum_{j=1}^{k}\exp\left\{-\frac{|y_{sj}(t)|^2}{2}\right\}\ge\sum_{s=1}^{2d_1}\sum_{j=1}^{k}\exp\left\{-\frac{|\bar y_{sj}(t)|^2}{2}\right\}.\tag{5.32}$$

For the points $\bar y_{sj}(t)$ we can derive estimates similar to the case $|t|<C<\infty$ and can get

$$\sum_{s=1}^{2d_1}\sum_{j=1}^{k}\exp\left\{-\frac{|\bar y_{nils}(t)|^2}{2}\right\}\ge\sum_{s=1}^{2d_1}\sum_{j=1}^{k}\exp\left\{-\frac{|y_{nils}|^2}{2}\right\}.\tag{5.33}$$

The statement (5.33) implies $J(t)>J(0)$ for $c<<|t|<<Cn^{1/2}b_n$.

Finally, after the shift $t, |t|\asymp n^{1/2}b_n$ one of the points y_{nile} or $-y_{nile}, 1\le e\le d_1$ will be located at a distance of the order $n^{1/2}b_n$ outside the ellipsoid Ξ and hence outside $n^{1/2}b_n\Omega$. This implies $J(t)>J(0)$. □

6. Proofs of Theorems 5.1 and 5.2

The proof of Theorem 5.1 contains only some new technical details in comparison with the proof of similar Theorem in [24]. The proof of Theorem 5.2 is based on a fairly new analytical technique (see [6, 9]) and is more interesting. Thus we begin with the proof of Theorem 5.2.

6.1. Proof of Theorem 5.2

Proof. We begin with auxillary estimates of moments of random variable X and random vector Z. We have

$$E[|X||Z|^2] \le (E|X|^{\frac{2+\lambda}{\lambda}})^{\frac{\lambda}{2+\lambda}}(E|Z|^{2+\lambda})^{\frac{2}{2+\lambda}} \le C(E[X^2])^{\frac{\lambda}{2+\lambda}} \le Cb_n^{\lambda}, \tag{6.1}$$

$$E[X^2|Z|] \le Cb_n^{-1}E[X^2] \le Cb_n^{1+\lambda}, \tag{6.2}$$

$$E[X^2|Z|^2] \le Cb_n^{-2}E[X^2] \le Cb_n^{\lambda}, \tag{6.3}$$

$$E[X^2|Z|^3] \le Cb_n^{-3}E[X^2] \le Cb_n^{\lambda-1}, \tag{6.4}$$

$$E[X^2|Z|^3] \le CE[|Z|^3] \le Cb_n^{\lambda-1}E[|Z|^{2+\lambda}] \le Cb_n^{\lambda-1}. \tag{6.5}$$

For each $x = \{x_1, \dots, x_d\} \in R^d$ denote $||x|| = \max_{1\le i\le d}|x_i|$. For any $z \in R^d$ and any $A \subset R^d$ denote $||A - z|| = \inf_{x\in A}||x - z||$. For any $\epsilon > 0$ denote $A_\epsilon = \{x : ||A - x|| \le \epsilon, x \in R^d\}$.

Define twice continuously differential functions $f_{1n} : R^1 \to R^1$ such that

$$f_{1n}(x) = \begin{cases} 1 & \text{if} \quad |x| > \epsilon_{1n} \\ 0 & \text{if} \quad |x| < \epsilon_{1n}/2 \end{cases}$$

and $0 \le f_{1n}(x) \le 1, |\frac{\partial f_{1n}(x)}{\partial x_{i_1}\partial x_{i_2}}| \le C\epsilon_{1n}^{-2}, 1 \le i_1, i_2 \le d, x \in R^d$.

Denote $c_n = c_{n1}n^{-1/2}b_n^{-1}$. We slightly modify the setup of Theorem 5.2 in the proof. The reasoning will be given for $r_n = 1$. Theorem 5.2 follows from the reasoning if we put $r_n = c_n$.

Define three- times continuously differentiable functions $f_{2n} : R^d \to R^1$ such that

$$f_{2n}(x) = \begin{cases} 1 & \text{if} \quad x \in n^{1/2}b_nv + U \\ 0 & \text{if} \quad x \notin n^{1/2}b_nv + U_{c_n} \end{cases}$$

and $0 \le f_{2n}(x) \le 1, |\frac{\partial^3 f_{2n}(x)}{\partial x_{i_1}\partial x_{i_2}\partial x_{i_3}}| \le Cc_n^{-3}, 1 \le i_1, i_2, i_3 \le d$ if $x \in R^d$.

Denote

$$S_{knX} = X_1 + \dots + X_{k-1} + X_{k+1} + \dots + X_n$$

and

$$W_{kn} = n^{-1/2}(Z_1 + \dots + Z_{k-1} + Y_{k+1} + \dots + Y_n).$$

Hereafter $Y_1, \dots, Y_n$ are independent copies of random vector Y. Random variables $Y, Y_1, \dots, Y_n$ are independent of $X_1, \dots, X_n, Z_1, \dots, Z_n$.

For any $\gamma > 0$ denote

$$G_n(\gamma) = sup\ E[f_{1n}(S_{nX}), S_{nZ} \in n^{1/2} b_n v + U_\gamma]$$

where the supremum is taken over all distributions of (X, Z) satisfying the assumptions of Theorem 5.2.

Lemma 6.1. *Let assumptions of Theorem 5.2 be satisfied. Then*

$$\begin{aligned} &E[f_{1n}(S_{nX}), S_{nZ} \in n^{1/2} b_n v + U] \\ &\le E[f_{1n}(S_{nX})] P(Y \in n^{1/2} b_n v + U_{c_n}) + C n b_n^{2+\lambda} c_{n1}^{-3} \epsilon_{1n}^{-2} G_{n-1}(\gamma_n) \end{aligned} \tag{6.6}$$

for $n > n_0$. Here $\gamma_n = \epsilon b_n^{-1}(n-1)^{-1/2} + (n(n-1)^{-1/2} b_n - (n-1)^{1/2} b_{n-1}) + C/n + c_n$ where C depends on U.

Proof. We have

$$E[f_{1n}(S_{nX}) f_{2n}(S_{nZ})] \le E[f_{1n}(S_{nX}) f_{2n}(Y)] + \Delta$$

where

$$\Delta = |E[f_{1n}(S_{nX}) f_{2n}(S_{nZ})] - E[f_{1n}(S_{nX}) f_{2n}(Y)]|.$$

It is clear that $\Delta \le \Delta_1 + \dots + \Delta_n$ where

$$\begin{aligned} \Delta_k = |&E[f_{1n}(S_{knX} + X_k) f_{2n}(W_{kn} + n^{-1/2} Z_k)] \\ &- E[f_{1n}(S_{knX} + X_k) f_{2n}(W_{kn} + n^{-1/2} Y)]| \end{aligned}$$

for $1 \le k \le n$.

Using the Taylor expansion of f_{1n} and f_{2n}, we get

$$\begin{aligned} \Delta_k &= |E[f_{1n}(S_{knX} + X_k)(f_{2n}(W_{kn} + n^{-1/2} Z)] - f_{2n}(W_{kn} + n^{-1/2} Y))]| \\ &\le \left| E\left[\left(f_{1n}(S_{knX}) + f'_{1n}(S_{knX}) X_k + \frac{1}{2} \int_0^1 f''_{1n}(S_{knX} + \omega X_k)(1-\omega)\, d\omega X_k^2 \right) \right.\right. \\ &\times \left(n^{-1/2}(Z_k - Y)' f'_{2n}(W_{kn}) + \frac{1}{2} n^{-1} (Z'_k f''_{2n}(W_{kn}) Z_k - Y' f''_{2n}(W_{kn}) Y) \right. \\ &\left.\left.\left. + \frac{1}{6} n^{-3/2} \int_0^1 (1-\omega)^2 (f'''_{2n}(W_{kn} + \omega Z_k) Z_k^3 - f'''_{2n}(W_{kn} + \omega Y) Y^3)\, d\omega \right) \right] \right| . \end{aligned} \tag{6.7}$$

After opening the brackets in the right-hand side of (6.7) it remains to estimate each term independently. The estimates are performed in the same way, using (5.12, 5.13, 5.14, 6.1 - 6.5). Therefore, we estimate only three of them.

Using (6.4), we get

$$\begin{aligned} &\left| n^{-3/2} E\left[\int_0^1 f''_{1n}(S_{knX} + \omega X_k)(1-\omega_1)\, d\omega_1 X_k^2 \right.\right. \\ &\left.\left. \times \int_0^1 (1-\omega)^2 (f'''_{2n}(W_{kn} + \omega_2 Z_k) Z_k^3 - f'''_{2n}(W_{kn} + \omega_2 Y) Y^3)\, d\omega_2 \right] \right| \\ &\le C n^{-3/2} c_n^{-3} \epsilon_{1n}^{-2} b_n^{\lambda-1} G_{kn}(\gamma_n) \le C \epsilon_{1n}^{-2} c_{n1}^{-3} b_n^{2+\lambda} G_{kn}(\gamma_n). \end{aligned} \tag{6.8}$$

The first inequality in (6.8) is due to the following reasoning

$$
\begin{aligned}
& W_{kn} + n^{-1/2} Z \in n^{1/2} b_n v + U_{c_n} \Rightarrow W_{kn} \in n^{1/2} b_n v + U_{\epsilon n^{-1/2} b_n^{-1} + c_n} \\
& \Rightarrow n^{1/2}(n-1)^{-1/2} W_{kn} \in (n-1)^{1/2} b_{n-1} v + (n(n-1)^{-1/2} b_n - (n-1)^{1/2} b_{n-1}) v \\
& + n^{1/2}(n-1)^{-1/2} U_{\epsilon n^{-1/2} b_n^{-1} + c_n} \\
& \Rightarrow n^{1/2}(n-1)^{-1/2} W_{kn} \in (n-1)^{1/2} b_{n-1} v + U_{\gamma_n}.
\end{aligned}
$$

Using (6.1), we get

$$
\begin{aligned}
& E[|f'_{1n}(S_{k,n-1,X}) X_k n^{-1} f''_{2n}(W_{kn}) Z_k^2|] \\
& \le C n^{-1} b_n^{\lambda} c_n^{-2} \epsilon_{1n}^{-1} G_{kn}(\gamma_n) \le C b_n^{2+\lambda} \epsilon_{1n}^{-1} c_{n1}^{-2} G_{kn}(\gamma_n).
\end{aligned}
$$

Using (5.14), we get

$$
\begin{aligned}
& n^{-1/2} E[f'_{1n}(S_{knX}) X_k (Z_k - Y) f'_{2n}(W_{kn})] \\
& = n^{-1/2} E[X_k Z_k] E[f'_{1n}(S_{knX}) f'_{2n}(W_{kn})] \le C n^{-1/2} b_n^{1+\lambda} \epsilon_{1n}^{-1} c_{n1}^{-1} G_{kn}(\gamma_n).
\end{aligned}
$$

$\square$

We begin the proof of Theorem 5.2 with auxilliary estimates. The first one is

$$
\begin{aligned}
P(Y \in n^{1/2} b_n + U_{c_n}) & \le \exp\{C c_n n^{1/2} b_n\} P(Y \in n^{1/2} b_n + U) \\
& \le a_0 P(Y \in n^{1/2} b_n + U).
\end{aligned}
$$

Note that

$$
Y \in (n-1)^{1/2} b_{n-1} v + U_{\gamma_n} \Rightarrow Y \in n^{1/2} b_n v + U_{\omega_n}
$$

with $\omega_n = \gamma_n + n^{1/2} b_n - (n-1)^{1/2} b_{n-1}$.

Therefore

$$
\begin{aligned}
& P(Y \in (n-1)^{1/2} b_{n-1} v + U_{\gamma_n}) \le P(Y \in n^{1/2} b_n v + U_{\omega_n}) \\
& \le C \exp\{n^{1/2} b_n \omega_n\} P(Y \in n^{1/2} b_n v + U) \le a_1 P(Y \in n^{1/2} b_n v + U).
\end{aligned}
$$

The further reasoning is based on an induction on n. We take a sufficiently large $n = n_0$ such that $C n_0 \epsilon_{1n_0}^{-2} c_{n_0,1}^{-3} b_{n_0}^{2+\lambda} < a$ with $a a_0 a_1 < 1$. We take C_{n_0} such that

$$
C_{n_0} P(Y \in n_0^{1/2} b_{n_0} + U) E[f_{1n}(S_{n_0 X})] \ge 1.
$$

Then

$$
E[f_{1n}(S_{n_0 X}), S_{n_0 Z} \in n_0^{1/2} b_{n_0} v + U] \le C_{n_0} P(Y \in n_0^{1/2} b_{n_0} + U) E[f_{1n}(S_{n_0 X})].
$$

Suppose Theorem 5.2 was proved for $n - 1 \ge n_0$. Let us prove it for n. We show that

$$
E[f_{1n}(S_{nX}), S_{nZ} \in n^{1/2} b_n v + U] \le C_n P(Y \in n^{1/2} b_n + U) E[f_1(S_{nX})] \quad (6.9)
$$

where $C_n = a_0 + C_{n-1}aa_1$. Then, since C_n form geometric progression with exponent $aa_0a_1 < 1$, Theorem 5.2 follows from (6.9).

Applying (6.6) and the inductive assumption, we get

$$\begin{aligned}
&E[f_{1n}(S_{nX}), S_{nZ} \in n^{1/2}b_n v + U] \le P(Y \in n^{1/2}b_n + U_{c_{1n}})E[f_{1n}(S_{nX})]\\
&+ Cnb_n^{2+\lambda}c_{n1}^{-3}\epsilon_{1n}^{-2}C_{n-1}E[f_{1n}(S_{nX})]P(Y \in (n-1)^{1/2}b_{n-1} + U_{\gamma_n})\\
&\le (a_0 + C_{n-1}aa_1)E[f_{1n}(S_{nX})]P(Y \in n^{1/2}b_n + U).
\end{aligned}$$

□

6.2. Proof of Theorem 5.1

Proof. In the proofs of Theorem 5.1 and Osypov Theorem [24] the basic reasoning coincide. The difference is only in the preliminary estimates. On these estimates the basic reasoning are based on.

Denote $\phi(h) = E[\exp\{h'X\}]$. Define random vector X_h having the conjugate distribution

$$F_h(dx) = F(dx, h) = \phi^{-1}(h)\exp\{h'x\}F(dx).$$

Denote

$$m(h) = E_h[X_h], \quad \sigma(h) = \mathrm{Var}[X_h].$$

For any $v \in R^d$ denote $h(v)$ the solution of the equation

$$m(h) = v. \tag{6.10}$$

Lemma 6.2. *For all $v, |v| < \epsilon b_n, \epsilon > 0$ there exists the solution $h(v)$ of equation (6.10) and the following relations hold*

$$\phi(h) = 1 + |h|^2/2 + O(|h|^3 b_n^{\lambda-1}), \tag{6.11}$$

$$m(h) = h + O(|h|^2 b_n^{\lambda-1}), \tag{6.12}$$

$$h(v) = v + O(|v|^2 b_n^{\lambda-1}), \tag{6.13}$$

$$\sigma(h) = I(1 + O(|h|^2 b_n^{\lambda-1})). \tag{6.14}$$

Proof of Lemma 6.2. Using the Taylor expansion, we get

$$\phi(h) = 1 + \frac{1}{2}\int (h'x)^2\,dF(x) + O\left(|h|^3 \int |x|^3\,dF(x)\right) = 1 + \frac{1}{2}|h|^2 + O(|h|^3 b_n^{\lambda-1}), \tag{6.15}$$

$$\begin{aligned}
m(h) &= \phi^{-1}(h)\int x\exp\{h'x\}\,dF(x)\\
&= \int x(h'x)dF(x)(1 - |h|^2/2 + O(|h|^3 b_n^{\lambda-1}) + O\left(\int x(h'x)^2 dF(x)\right)\\
&= h + O(|h|^2 + |h|^2 b_n^{\lambda-1}).
\end{aligned} \tag{6.16}$$

Substituting (6.16) in (6.10), we get (6.13). Estimating similarly to (6.16), we get (6.14).

Denote

$$\Lambda(h,v) = -(h,v) + \ln \phi(h).$$

By (6.11, 6.13), we get

$$\ln \phi(h(v)) = \frac{1}{2} h^2(v)(1 + O(b_n^\lambda)). \tag{6.17}$$

By (6.14), we get

$$\det{}^{-1/2} \sigma(h(v)) = 1 + O(b_n^\lambda). \tag{6.18}$$

By (6.13) and (6.17) we get

$$\Lambda(h(v),v) = |v|^2(1 + O(|v| b_n^{\lambda-1})) - \frac{1}{2}|v|^2(1 + O(b_n^\lambda)) = \frac{1}{2}|v|^2 + O(|v|^2 b_n^\lambda). \tag{6.19}$$

The estimates (6.11-6.14) and (6.17-6.19) are the versions of similar estimates in [24]. Using these estimates we get Theorem 5.1 on the base of the same reasoning as in [24]. This reasoning is omitted □

7. Proofs of Lemmas 3.3, 3.4, 5.1, 5.2 and 5.4-5.8

The Lemmas will be proved in the following order: 3.3, 3.4, 5.1, 5.2, 5.6, 5.4, 5.7, 5.5, 5.8.

Proof. The proof of Lemma 3.3 is based on the following reasoning. Let $h \in \Psi_j(\theta)$ and $h_1 \in \Pi(h)$. By (2.2) and (2.4), we get

$$\begin{aligned} P_{h_1}(|\eta(h_1,h)| > \epsilon) &\le P_{h_1}(|\eta(h_1,h) - \frac{1}{2}\bar{h}'\tau_{h_1}| > \epsilon/2) + P_{h_1}(|\bar{h}'\tau_{h_1}| > \epsilon/2) \\ &< 4\epsilon^{-2} E_{h_1}[(\eta(h_1,h) - \frac{1}{2}\bar{h}'\tau_{h_1})^2] + 2^{2+\lambda}\epsilon^{-2-\lambda}|\bar{h}|^{2+\lambda} E_{h_1}|\tau_{h_1}|^{2+\lambda} \le C|\bar{h}|^{2+\lambda}. \end{aligned} \tag{7.1}$$

By straightforward computations, using (7.1), for $1 \le j \le m$, we get

$$P(V_h(\theta)) \le C P_{h_1}(|\eta(h_1,h)| > \epsilon j^{-2}) \le C\epsilon^{-2} j^4 |\bar{h}|^{2+\lambda} \le C j^4 \left(\frac{b_n}{2^j}\right)^{2+\lambda}. \tag{7.2}$$

In the case $j = m+1$ the constant C in (7.2) is replaced with $C c_{3n}^{d-1}$.

By (7.2), we get

$$P(B_{4n}(\theta)) < Cn \sum_{j=1}^{m} 2^j \left(\frac{b_n}{2^j}\right)^{2+\lambda} j^4 + Cn c_{3n}^{d-1} 2^m c_{3n}^{2+\lambda} \delta_{1n}^{2+\lambda} m^4. \tag{7.3}$$

Note that $2^m = Cc_{1n}^{-1} nb_n^2(1+o(1))$. Therefore, using inequality $n^{-\lambda}b_n^{-\lambda} < nb_n^{2+\lambda}$, we get

$$\begin{aligned} P(B_{4n}(\theta)) &< Cnb_n^{2+\lambda} + Cnc_{3n}^{d+1+\lambda}2^{-m(1+\lambda)}m^4 b_n^{2+\lambda} \\ &\le Cnb_n^{2+\lambda}\epsilon^{-2-\lambda} + CC_n c_{3n}^{d+2+\lambda} n^{-\lambda} b_n^{-\lambda} m^4 = O(nb_n^{2+\lambda}) = o(1) \end{aligned} \tag{7.4}$$

if c_{3n} tends to infinity sufficiently slowly.

Since $P^{(s)}_{h,h_1}(S) < C|\bar{h}|^{2+\lambda}$, then, arguing similarly to (7.2)-(7.4), we get

$$\begin{aligned} P(D_{nile}) &\le Cn \sum_{j=1}^{m+1} \sum_{h \in \Psi_j(\theta)} P^{(s)}_{h,h_1}(S) \\ &\le Cn \sum_{j=1}^{m} 2^j (b_n 2^{-j})^{2+\lambda} + Cnc_{3n}^{d+1+\lambda} 2^m \delta_{1n}^{2+\lambda} = o(1). \end{aligned} \tag{7.5}$$

Now (7.4, 7.5) implies (3.18). □

Proof. The proof of Lemma 3.4 is based on the following reasoning. Applying the Chebyshev inequality and using (2.4), we get

$$P(B_{3n1}) \le \epsilon^{-2-\lambda} b_n^{2+\lambda} E[|\tau|^{2+\lambda}] < Cb_n^{2+\lambda}. \tag{7.6}$$

Let $h \in \Psi_j(\theta), 1 \le j \le m+1$. By the Chebyshev inequality, we get

$$\begin{aligned} &P(|\tau_{sh} - \tau_s| > \epsilon b_n^{-1} 2^{j/2} | A_{4n1}) \\ &< C2^{-j(2+\lambda)/2} b_n^{2+\lambda} \epsilon^{-2-\lambda} (E[|\tau_h|^{2+\lambda} | A_{4n1}] + E[|\tau|^{2+\lambda}]) \\ &< C2^{-j(2+\lambda)/2} b_n^{2+\lambda} \epsilon^{-2-\lambda} (E_h[|\tau_h|^{2+\lambda}] + E[|\tau|^{2+\lambda}]) \\ &\le C2^{-j(2+\lambda)/2} b_n^{2+\lambda}. \end{aligned} \tag{7.7}$$

By (7.6), (7.7), we get

$$\begin{aligned} P(B_{3nile}) &< Cn \sum_{j=1}^{m} 2^j b_n^{2+\lambda} 2^{-j(2+\lambda)/2} + Cnc_{3n}^{d-1} 2^m 2^{-m(2+\lambda)/2} b_n^{2+\lambda} \\ &< Cnb_n^{2+\lambda} = o(1). \end{aligned} \tag{7.8}$$

By (7.4), (7.5) and (7.8), we get

$$P(B_{1nile}) < Cnb_n^{2+\lambda}. \tag{7.9}$$

□

Proof. The proof of Lemma 5.1 is based on the following reasoning. Since $E[\tau] = 0$, we have

$$\begin{aligned} |E[\tau, A_{1n1}]| &= |E[\tau, B_{1n1}]| \\ &\le E[|\tau|, |\tau| > b_n^{-1}] + E[|\tau|, B_{1n1} \cap \{|\tau| \le b_n^{-1}\})] \\ &\le b_n^{1+\lambda} E|\tau|^{2+\lambda} + b_n^{-1} P(B_{1n1}) = O(b_n^{1+\lambda}) \end{aligned} \tag{7.10}$$

where the last equality follows from (2.4), (7.4), (7.6). This implies (5.1).

The proof of (5.2) is similar and is omitted. □

The considerable part of the subsequent estimates is based on the following Lemma.

Lemma 7.1. *Let* $h \in \Psi_j(\theta), h_1 \in \Pi(h), 1 \le j \le m+1, \theta \in \Theta_{nile}$. *Then, for any* $a \ge 0, b \ge 0, a+b \ge 2+\lambda$, *there holds*

$$E_{h_1}[|\bar{h}\tau_{h_1}|^a |\eta(h_1,h)|^b, A_{1n1}] \le C|\bar{h}|^{2+\lambda}.$$

Proof. By (2.2) and (2.4), we get

$$\begin{aligned}
&E_{h_1}[|\bar{h}\tau_{h_1}|^a |\eta(h_1,h)|^b, A_{1n1}] \le CE_{h_1}[|\bar{h}\tau_{h_1}|^{a+b}, A_{1n1}] + CE_{h_1}[|\eta(h_1,h)|^{a+b}, A_{1n1}]\\
&\le CE_{h_1}[|\bar{h}\tau_{h_1}|^{a+b}, A_{1n1}] + CE_{h_1}[|\eta(h_1,h) - \bar{h}\tau_{h_1}|^{a+b}, A_{1n1}]\\
&\le CE_{h_1}[|\bar{h}\tau_{h_1}|^{2+\lambda}, A_{1n1}] + CE_{h_1}[|\eta(h_1,h) - \bar{h}\tau_{h_1}|^{2}, A_{1n1}] \le C|\bar{h}|^{2+\lambda}.
\end{aligned}$$

□

Proof. The proof of Lemma 5.2 is based on the following reasoning. Using the Taylor expansion of ξ_n, we get

$$S_{n\theta} = \sum_{s=1}^n (2\eta_{ns}(\theta) - \theta'\tau_s) - \sum_{s=1}^n \eta_{ns}^2(\theta) + \frac{2}{3}\sum_{s=1}^n \frac{\eta_{ns}^3(\theta)}{(1+\kappa\eta_{ns}(\theta))^3} + 2n\rho^2(0,\theta) \quad (7.11)$$

where $0 \le \kappa \le 1$.

Since $E[\eta_n^2(\theta)] = \rho^2(0,\theta)$ and $2E[\eta_n(\theta)] = -E[\eta_n^2(\theta)] = -\rho^2(0,\theta)$, then, by virtue of (2.3), we get

$$E[(2\eta_n(\theta) - \theta'\tau) - \eta_{ns}^2(\theta) + \frac{1}{2}\theta' I\theta] = O(|\theta|^{2+\lambda}). \quad (7.12)$$

By (7.4), (7.9), we get

$$\begin{aligned}
E[|\eta_n(\theta)|, B_{1n1}) &\le E[|\eta_n(\theta)|, |\eta_n(\theta)| > \epsilon] + E[|\eta_n(\theta)|, B_{1n1} \setminus \{|\eta_n(\theta)| < \epsilon\}]\\
&\le E[|\eta_n(\theta)|, |\eta_n(\theta)| > \epsilon] + \epsilon P(B_{1n1}) \qquad (7.13)\\
&\le E[|\eta_n(\theta)|, |\eta_n(\theta)| > \epsilon] + Cb_n^{2+\lambda}.
\end{aligned}$$

By (2.2, 2.4), we get

$$\begin{aligned}
&E[|\eta_n(\theta)|, |\eta_n(\theta)| > \epsilon]\\
&\le E[|\eta_n(\theta)|, |\eta_n(\theta)| > \epsilon, |\eta_n(\theta) - \frac{1}{2}\theta'\tau| < \epsilon/2] + E[|\eta_n(\theta)|, |\eta_n(\theta)| > \epsilon, |\theta\tau| < \epsilon/2]\\
&\le CE[|\theta'\tau|, |\eta_n(\theta)| > \epsilon, |\eta_n(\theta) - \frac{1}{2}\theta'\tau| < \epsilon/2] + 4\epsilon^{-1}E[(\eta_n(\theta) - \frac{1}{2}\theta'\tau)^2]\\
&\le C\epsilon^{-1-\lambda}E[|\theta'\tau|^{2+\lambda}] + Cb_n^{2+\lambda} \le Cb_n^{2+\lambda}.
\end{aligned}$$

(7.14)

By (7.13) and (7.14), we get

$$E[\eta_n(\theta)|B_{1n1}] \le C b_n^{2+\lambda}. \tag{7.15}$$

Arguing similarly to (7.13, 7.14), we get

$$E[\eta_n^2(\theta), B_{1n1}] = O(b_n^{2+\lambda}). \tag{7.16}$$

By (7.12), (7.9), (7.10), (7.15), (7.16), we get

$$E[(2\eta_n(\theta) - \frac{1}{2}\theta'\tau) - \eta_{ns}^2(\theta) + \frac{1}{2}\theta' I\theta, B_{1n1}] = O(|b_n|^{2+\lambda}). \tag{7.17}$$

By Lemma 7.1, we get

$$E\left[\left|\frac{\eta_n^3(\theta)}{(1+\kappa\eta_n(\theta))^3}\right|, A_{1n1}\right] \le CE[|\eta_n^3(\theta)|, A_{1n1}] \le C|\theta|^{2+\lambda}. \tag{7.18}$$

By (7.11), (7.12), (7.17), (7.18) we get (5.3). ☐

Proof. The proof of Lemma 5.6 is based on the following reasoning. We have

$$\begin{aligned} &E[(\xi(\theta) - \theta'\tau)^2, A_{1n1}] \le CE[(\eta_n(\theta) - \frac{1}{2}\theta'\tau)^2] \\ &+ CE[\eta_n^4(\theta), A_{1n1}] + CE[\eta_n^6(\theta), A_{1n1}]. \end{aligned} \tag{7.19}$$

By Lemma 7.1, we get

$$E[\eta_n^4(\theta), A_{1n1}] = O(|\theta|^{2+\lambda}), \quad E[\eta_n^6(\theta), A_{1n1}] = O(|\theta|^{2+\lambda}). \tag{7.20}$$

By (2.2), (7.19), (7.20) we get (5.20).

Estimating similarly to (7.19), (7.20), we get

$$E[(\xi(h_1,h) - \frac{1}{2}\bar h'\tau_{h_1})^2, A_{1n1}] \le CE_{h_1}[(\xi(h_1,h) - \frac{1}{2}\bar h'\tau_{h_1})^2, A_{1n1}] \le C|\bar h|^{2+\lambda}.$$

This implies (5.21). ☐

Proof. The proof of Lemma 5.4 is based on the following reasoning. Applying the Cauchy inequality, by (5.22), we get

$$\begin{aligned} &E[(\xi(h_1,h) - \bar h'\tau_{h_1})(v'\tau), A_{1n1}] \\ &\le (E[(\xi(h_1,h) - \bar h'\tau_{h_1})^2, A_{1n1}])^{1/2}(E[(v'\tau)^2, A_{1n1}])^{1/2} \le C|v||\bar h|^{1+\lambda/2}. \end{aligned} \tag{7.21}$$

☐

Proof. The proof of Lemma 5.7 is based on the following reasoning. Using the obvious inequality $(a+b)^2 - 2b^2 \le 2a^2$, putting $a = \eta(0,u) + \frac{1}{2}u'\tau - \eta(h,h+u) + \frac{1}{2}u'\tau_h$ and $b = \eta(h,h+u) - \eta(0,u)$, we get

$$\begin{aligned} &E[(u'(\tau - \tau_h))^2, A_{1n1}] - 2E[(\eta(h,h+u) - \eta(0,u))^2, A_{1n1}] \\ &\le 2E[(\eta(h,h+u) - \frac{1}{2}u'\tau_h - \eta(0,u) + \frac{1}{2}u'\tau)^2, A_{1n1}] \doteq J. \end{aligned} \tag{7.22}$$

Using the inequality $2a^2 \le 4(a+b)^2 + 4b^2$, putting $a = \eta(h,h+u) - \frac{1}{2}u'\tau_h - \eta(0,u) + \frac{1}{2}u'\tau$ and $b = \eta(0,u) - \frac{1}{2}u'\tau$, by (2.2), we get

$$\begin{aligned} J &\le 4E[(\eta(h,h+u) - \frac{1}{2}u'\tau_h)^2, A_{1n1}] + 4E[(\eta(0,u) - \frac{1}{2}u'\tau)^2, A_{1n1}] \\ &\le CE_h[(\eta(h,h+u) - \frac{1}{2}u'\tau_h)^2] + C|u|^{2+\lambda} \le C|u|^{2+\lambda}. \end{aligned} \tag{7.23}$$

Thus, for the proof of (5.22), it suffices to show that

$$J_1 \doteq E[(\eta(h,h+u) - \eta(0,u))^2, A_{1n1}] = O(|u|^2|h|^\lambda).$$

By straightforward calculations, we get

$$\begin{aligned} &(\eta(h,h+u) - \eta(0,u))^2 \\ &= (\eta(0,h+u) - \eta(0,h) - \eta(0,u) - \eta(0,h)\eta(0,u))^2(\eta(0,h)+1)^{-2}. \end{aligned}$$

Therefore we have

$$\begin{aligned} J_1 &= E[(\eta(0,h+u) - \eta(0,h) - \eta(0,u) - \eta(0,h)\eta(0,u))^2(\eta(0,h)+1)^{-2}, A_{1n1}] \\ &\le CE[(\eta(0,h+u) - \eta(0,h) - \eta(0,u) - \eta(0,h)\eta(0,u))^2, A_{1n1}] \\ &\le CE[(\eta(0,h+u) - \frac{1}{2}(h+u)'\tau - (\eta(0,h) - \frac{1}{2}h'\tau) - (\eta(0,u) - \frac{1}{2}u'\tau))^2, A_{1n1}] \\ &+ CE[\eta^2(0,h)\eta^2(0,u)), A_{1n1}] \doteq J_{11} + J_{12}. \end{aligned} \tag{7.24}$$

Applying (2.2), we derive

$$\begin{aligned} J_{11} &\le CE[(\eta(0,h+u) - \frac{1}{2}(h+u)'\tau)^2] + CE[(\eta(0,h) - \frac{1}{2}h'\tau)^2] \\ &+ CE[(\eta(0,u) - \frac{1}{2}u'\tau)^2] \le C|h+u|^{2+\lambda} + C|h|^{2+\lambda}. \end{aligned} \tag{7.25}$$

By Lemma 7.1, we get

$$J_{12} \le CE[\eta^4(0,h), A_{1n1}] + CE[\eta^4(0,u), A_{1n1}] \le C(|u|^{2+\lambda} + |h|^{2+\lambda}). \tag{7.26}$$

By (7.24-7.26, 7.23, 7.22), we get

$$E[(u'(\tau - \frac{1}{2}\tau_h))^2, A_{1n1}] \le C(|h+u|^{2+\lambda} + |u|^{2+\lambda} + |h|^{2+\lambda}).$$

Putting $|u| = c_0|h|$ and $C_1 = C((1+c_0)^{2+\lambda} + c_0^{2+\lambda} + c_0^2)c_0^{-2}$, we get

$$E[(u'(\tau - \tau_h))^2, A_{1n1}] \le C_1|u|^2|h|^\lambda.$$

□

Proof. The proof of Lemma 5.5 is based on the following reasoning. Denote

$$\begin{aligned} W &\doteq E[(h_1'\tau)(\xi(h_1,h) - \bar{h}'\tau_{h_1})|A_{1n1}] = E[(h_1'(\tau - \tau_{h_1}))(\xi(h_1,h) - \bar{h}'\tau_{h_1})|A_{1n1}] \\ &\quad + E[(h_1'\tau_{h_1})(\xi(h_1,h) - \bar{h}'\tau_{h_1})|A_{1n1}] \doteq W_{11} + W_{12}. \end{aligned} \tag{7.27}$$

By (5.22), (5.21), we get

$$\begin{aligned} W_{11} &\le (E[(h_1'(\tau - \tau_{h_1}))^2|A_{1n1}])^{1/2}(E[(\xi(h_1,h) - \bar{h}'\tau_{h_1})^2|A_{1n1}])^{1/2} \\ &\le C|h_1|^{1+\lambda/2}|\bar{h}|^{1+\lambda/2}. \end{aligned} \tag{7.28}$$

We have

$$\begin{aligned} W_{12} &= E_{h_1}[(1+\eta(h_1,0))^2(h_1'\tau_{h_1})(\xi(h_1,h) - \bar{h}'\tau_{h_1})|A_{1n1}] \\ &= E_{h_1}[(h_1'\tau_{h_1})(\xi(h_1,h) - \bar{h}'\tau_{h_1})|A_{1n1}] \\ &\quad + 2E_{h_1}[\eta(h_1,0)(h_1'\tau_{h_1})(\xi(h_1,h) - \bar{h}'\tau_{h_1})|A_{1n1}] \\ &\quad + E_{h_1}[\eta^2(h_1,0)(h_1'\tau_{h_1})(\xi(h_1,h) - \bar{h}'\tau_{h_1})|A_{1n1}] \doteq W_{121} + W_{122} + W_{123}. \end{aligned} \tag{7.29}$$

By (7.11), we get

$$\begin{aligned} W_{121} &= E_{h_1}[h_1'\tau_{h_1}(2\eta(h_1,h) - \bar{h}\tau_{h_1}), A_{1n1}] - E_{h_1}[h_1'\tau_{h_1}\eta^2(h_1,h), A_{1n1}] \\ &\quad + \frac{2}{3}E_{h_1}\left[h_1'\tau_{h_1}\frac{\eta^3(h_1,h)}{(1+\kappa\eta(h_1,h))^3}, A_{1n1}\right] \doteq W_{1211} + W_{1212} + W_{1213}. \end{aligned} \tag{7.30}$$

By (2.2), (2.3), we get

$$\begin{aligned} O(|\bar{h}|^{2+\lambda}) &= E_{h_1}[(\eta(h_1,h) - \frac{1}{2}\bar{h}'\tau_{h_1})^2] \\ &= \rho^2(h_1,h) - E_{h_1}[\eta(h_1,h)\bar{h}'\tau_{h_1}] + \frac{1}{4}\bar{h}I(h_1)\bar{h} \\ &= \frac{1}{2}\bar{h}'I(h_1)\bar{h}(1+|\bar{h}|^{\lambda}) - E_{h_1}[\eta(h_1,h)\bar{h}\tau_{h_1}]. \end{aligned} \tag{7.31}$$

Since $h_1 \parallel \bar{h}$, by (7.31), we get

$$E_{h_1}[h_1'\tau_{h_1}\eta(h_1,h)] = \frac{1}{2}h_1'I(h_1)\bar{h}(1+O(|\bar{h}|^{\lambda})). \tag{7.32}$$

Applying the Holder's inequality, we get

$$\begin{aligned} &E_{h_1}[h_1'\tau_{h_1}(\eta(h_1,h) - \frac{1}{2}\bar{h}'\tau_{h_1}), B_{1n1}] \\ &\le (E_{h_1}[(h_1'\tau_{h_1})^{2+\lambda}])^{\frac{1}{2+\lambda}}(E_{h_1}[(\eta(h_1,h) - \frac{1}{2}\bar{h}\tau_{h_1})^2])^{1/2}(P_{h_1}(B_{1n1}))^{\frac{\lambda}{2(2+\lambda)}} \\ &= O(|h_1||\bar{h}|^{1+\lambda/2}b_n^{\lambda/2}). \end{aligned} \tag{7.33}$$

By (7.32), (7.33), (5.2), we get

$$W_{1211} = O(|h_1'||\bar{h}|b_n^{\lambda}). \tag{7.34}$$

By Lemma 7.1, we get

$$W_{1212} + W_{1213} = O(|h_1||\bar{h}|^{1+\lambda}). \tag{7.35}$$

By (7.30), (7.34), (7.35), we get

$$W_{121} = O(|h_1'||\bar{h}|b_n^{\lambda}). \tag{7.36}$$

Using Lemma 7.1 and (7.11), we get

$$W_{122} + W_{123} = O(|\bar{h}|^{1+\lambda}|h_1|). \tag{7.37}$$

By (7.29), (7.36), (7.37), we get

$$W_{12} = O(|h_1'||\bar{h}|b_n^{\lambda}). \tag{7.38}$$

By (7.27), (7.28), (7.38), we get (5.18). □

Proof. The proof of Lemma 5.8 is based on the following reasoning. We begin with the proof of (5.23). Using (5.22), we get

$$E[\bar{h}'(\tau - \tau_{h_1})\tau_k, A_{1n1}] \le (E[\bar{h}'(\tau - \tau_{h_1})^2, A_{1n1}])^{1/2}(E[\tau_k^2])^{1/2} < C|\bar{h}||h_1|^{\lambda/2}.$$

The proof of (5.24) is based on the following reasoning. By (5.22), we get

$$\begin{aligned} O(|\bar{h}|^2 b_n^{\lambda}) &= E[(\bar{h}(\tau - \tau_{h_1}))^2, A_{1n1}] = E[(\bar{h}\tau)^2, A_{1n1}] - \\ &- 2E[(\bar{h}\tau)(\bar{h}\tau_h), A_{1n1}] + E[(\bar{h}\tau_{h_1})^2, A_{1n1}] \doteq J_1 - 2J_2 + J_3. \end{aligned} \tag{7.39}$$

We have

$$\begin{aligned} J_3 &= E_{h_1}[(\eta(h_1,0)+1)^2(\bar{h}\tau_{h_1})^2, A_{1n1}] \\ &= E_{h_1}[\eta^2(h_1,0)(\bar{h}\tau_{h_1})^2, A_{1n1}] + 2E_{h_1}[\eta(h_1,0)(\bar{h}\tau_{h_1})^2, A_{1n1}] \\ &+ E_{h_1}[(\bar{h}\tau_{h_1})^2, A_{1n1}] = J_{31} + 2J_{32} + J_{33}. \end{aligned} \tag{7.40}$$

By Lemma 7.1, we get

$$J_{31} + 2J_{32} \le C|\bar{h}|^2|h|^{\lambda}. \tag{7.41}$$

Estimating similarly to the proof of (5.1), (5.2), we get

$$J_{33} = \bar{h}' I(h)\bar{h} + O(|\bar{h}|^2 b_n^{\lambda}). \tag{7.42}$$

By (7.40)-(7.42), we get

$$J_3 = \bar{h}_1' I(h_1)\bar{h}_1 + O(|\bar{h}|^2 b_n^{\lambda}). \tag{7.43}$$

By (7.39), (5.2), (7.43), we get

$$J_2 = \bar{h}_1' I\bar{h}_1 + O(|\bar{h}|^2 b_n^{\lambda}). \tag{7.44}$$

By (7.44), (5.2), we get

$$J_1 - J_2 = O(|\bar{h}|^2 b_n^{\lambda}).$$

This implies (5.24). □

Acknowledgments

I thank the referees and the Associate Editor for their careful reading of the paper and valuable comments.

References

[1] Aleshkevichiene, A.K. (1990). Probabilities of large deviations for U-statistics and Von Mises functionals. *Theor. Probab. Appl.* **35** 1–13. MR1050049

[2] Archangelskii, A.N. (1989). Lower bounds for probabilities of large deviations for sums of independent random variables. *Theory Probab. Appl.* **34** 625–635. MR1036703

[3] Bahadur, R.R. (1960). Asymptotic efficiency of tests and estimates. *Sankhya.* **22** 229–252. MR0293767

[4] von Bahr, B. (1967). Multidimensional integral limit theorems for large deviations, *Arc. Math.* **7** 89–99. MR0224129

[5] Barndorff-Nielsen, O.E. and Cox, D.R. (1989). *Asymptotic Techniques for Use in Statistics.* London: Chapman and Hall. MR1010226

[6] Bentkus, V. and Rachkauskas, A. (1990). On probabilities of large deviations in Banach spaces. *Probab. Th. Rel. Fields.* **86** 131–154. MR1065276

[7] Borovkov, A.A. and Mogulskii, A.A. (1992). Large deviations and testing of statistical hypothesis. *Proc. Inst. Math. Russian Acad. Sci., Siberian Division.* English transl: *Siber. Adv. Math* **2** 1992 No 3,4 **3**, 1993 No 1,2.

[8] Butler, R.W. (2007). *Saddlepoint Approximations with Applications.* Cambridge, UK: Cambridge Univ. Press. MR2357347

[9] Ermakov, M.S. (2001). Large deviation probabilities in Banach spaces. *Zapiski Nauchn Seminarov POMI RAN.* **270** 42–64. (English Translation: *Journal of Mathematical Sciences* (2003) **118** 5555–5569.) MR1874389

[10] Ermakov, M.S. (2003). Asymptotically efficient statistical inference for moderate deviation probabilities. *Theory Probab. Appl.* **48**(4) 676–700. MR2143210

[11] Ermakov, M. (2007). Importance sampling for simulation of moderate deviation probabilities of statistics. *Statistics & Decision* **25** 265–284. MR2441748

[12] Ermakov, M.S. (2008). On semiparametric statistical inferences in the moderate deviation zone. *J. Math. Sci.* **152** 869–874. MR2742905

[13] Feller, W. (1966). *An Introduction to Probability Theory and its Applications.* New York: Wiley MR0210154

[14] Gao, F. and Zhao, X. (2011). Delta method in large deviations and moderate deviations for estimators. *Ann. Statist.* **39** 1211–1240. MR2816352

[15] Gupta, A.D. (2008). *Asymptotic Theory of Statistics and Probability.* Berlin: Springer. MR2664452

[16] Hajek, J. (1972). Local asymptotic minimax and admisibility in estimation. *Proc.Sixth Berkeley Symp. on Math. Statist. and Probab.* **1** 175–194, Berkeley, California Univ. Press. MR0400513

[17] van den Heuvel, E.R. and Klaassen, C.A.J. (1997). Spread, estimators and nuisance parameters. *Bernoulli* **3** 323–328. MR1468308

[18] Ibragimov, I.A. and Hasminskii, R.Z. (1981). *Statistical Estimation: Asymptotic Theory.* Berlin: Springer. MR0620321

[19] Inglot, T., Kallenberg, W.C.M. and Ledwina, T. (1992). Strong moderate deviation theorems. *Ann.Probab.* **20** 987–1003. MR1159582

[20] Jureckova, J., Kallenberg, W.C.M. and Veraverbeke, N. (1988). Moderate and Cramer-type large deviation theorems for M-estimators. *Statistics & Probability Letters* **6** 191–199. MR0919374

[21] Klaassen, C.A.J. (1989). Estimators and spread. *The Annals of Statistics* **17**, 859–867. MR0994272

[22] Le Cam, L. (1972). Limits of Experiments Berlin: Springer. *Proc.Sixth Berkeley Symp. on Math. Statist. and Probab.* **1** 245–261, Berkeley, California Univ. Press. MR0415819

[23] Lifshits, M.A. (1995). *Gaussian Random Functions.* Dordrecht: Kluwer Academic Publishers. MR1472736

[24] Osypov, L.V. (1981). On large deviations for sums of random vectors in R^k. *J. Mult. Anal.* **11** 115–126. MR0618780

[25] Puhalskii, A. and Spokoiny, V. (1998). On large-deviation efficiency in statistical inference. *Bernoulli* **4** 203–272. MR1632979

[26] Radavichius, M. (1991). From asymptotic efficiency in minimax sense to Bahadur efficiency. *New Trends in Probab. and Statist.* V. Sazonov and T. Shervashidze (Eds) Vilnius, VSP/Mokslas **1** 629–635. MR1200944

[27] Reid, N. (2003). Asymptotics and the theory of inference. *Ann. Statist* **31** 1695–1731. MR2036388

[28] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics.* New York: Wiley. MR0595165

[29] Skovgaard, I. M. (2001). Likelihood asymptotics. *Scandinavian Journal of Statistics* **28** 3–32.

[30] Strasser, H. (1985). *Mathematical Theory of Statistics.* Berlin: W. de Gruyter. MR0812467

[31] Van der Vaart, A.W. (1998). *Asymptotic Statistics.* Cambridge, UK: Cambridge Univ. Press. MR1652247

[32] Wolfowitz, J. (1965). Asymptotic efficiency of the maximum likelihood estimator. *Theor. Probab.Appl.* **10** 267–281. MR0184338