A MODERATE DEVIATION PRINCIPLE FOR EMPIRICAL BOOTSTRAP MEASURE

MIKHAIL ERMAKOV

Abstract. We establish two moderate deviation principles (MDP) in the bootstrap setting. We prove MDP for the joint distribution of the empirical measure and the empirical bootstrap measure (empirical measure obtaining by the bootstrap procedure). We derive MDP for the conditional distribution of the empirical bootstrap measure given the empirical probability measure.For most common statistical functionals (in particular differentiable and homogeneous functionals) we show that their asymptotics of moderate deviation probabilities in the cases of empirical measure and bootstrap empirical bootstrap measure coincides. However the moderate deviation zones are different.

1. **Introduction.**

Let
- S be a Hausdorff topological space;
- $\mathscr{F}$ the σ-field of Borel sets in S;
- Λ the space of all probability measures on $(S, \mathscr{F})$.

Let $X_1, \ldots, X_n$ be i.i.d. random variables, taking values in S and $P(\in \Lambda)$ be unknown distribution of X_1.

Denote $\hat{P}_n$ the empirical measure (occupation measure) for $X_1, \ldots, X_n$, that is, for any $\mathscr{F}$-measurable set A,

$$\hat{P}_n(A) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \in A)$$

In 1979, in a landmark paper Efron [12] proposed to analyze the distributions of statistics $V(X_1, \ldots, X_n)$ with the help of the bootstrap procedure. In the bootstrap procedure we consider the empirical measure $\hat{P}_n$ as an estimator of the probability measure (pm) P and simulate the distribution of statistics $V(X_1, \ldots, X_n)$ on the base of pm $\hat{P}_n$. In other words, we simulate independent copies $(X^*_{1i}, \ldots, X^*_{ni})_{i \in [1,k]}$ of i.i.d random variables such that X^*_{11} is distributed according to $\hat{P}_n$. After that the empirical distribution of $(V(X^*_{1i}, \ldots, X^*_{ni}))_{i \in [1,k]}$ is postulated as an estimate of the distribution of $V(X_1, \ldots, X_n)$.

It is of interest to estimate large and moderate deviation probabilities of $V(X_1, \ldots, X_n)$. Such problems emerge constantly in confidence estimation and hypothesis testing. The significant levels in the confidence estimation and the type I error probabilities in hypothesis testing are (usually) of small values and thus are compatible with LDP - MDP analysis. Hence it appears natural to compare $V(X_1, \ldots, X_n)$ and $V(X^*_1, \ldots, X^*_n)$ in terms of LDP - MDP approach.

In this paper we carry out such an MDP based comparison in the following setup.

1991 *Mathematics Subject Classification.* 60F10, 62C12, 62E25.
Key words and phrases. Large deviations, moderate deviations, bootstrap, empirical measure.

We represent $V(X_1,\dots,X_n)$ and $V(X_1^*,\dots,X_n^*)$ as functionals of $\hat P_n$ and $\hat P_n^*$, where P_n^* is the empirical probability measure of $X_1^*,\dots,X_n^*$, i.e.

$$V(X_1,\dots,X_n) = T(\hat P_n),$$
$$V(X_1^*,\dots,X_n^*) = T(P_n^*)$$

Thus we reduce the problem to an MDP study for $T(P_n^*)-T(\hat P_n)$ and $T(\hat P_n)-T(P)$.

The LDP - MDP analysis for empirical measures generated i.i.d. random objects is well known from Sanov [24], Groeneboom, Oosterhoff and Ruymgaart [17], Borovkov and Mogulskii [5], Dembo and Zeitouni [10], Eichelsbacher and Schmock [13], Arcones [2], de Acosta [8], Ermakov [15] (see also references therein). The results there are obtained under rather general assumptions.

Our goal is twofold.

1. We develop MDP technique from the above mentioned papers for

$$(P_n^* - \hat P_n) \times (\hat P_n - P).$$

and implement the above result for the MDP comparison of

$$T(\hat P_n) - T(P) \text{ and } T(P_n^*) - T(\hat P_n).$$

2. We establish the MDP for a conditional distribution of the empirical bootstrap measure P_n^* given empirical probability measure $\hat P_n$.

We notice that the MDP for the joint

"empirical bootstrap + empirical probability"

measures is valid in a "smaller time zone" than the MDP for empirical measure only. On the other hand, the time zone for the above-mentioned conditional MDP is essentially larger with probability close to one. The first statement shows instability of a bootstrap procedure provided that the empirical measure belongs to the MDP zone. The second statement confirms the wellknown fact that the bootstrap statistics have more stable properties (see Hall [18], Wood [28], DasGupta [9]).

The LDP for the empirical bootstrap measure has been studied in Chaganty [7] and Chaganty, Karandikar [6] using weak convergence. In contrast to that, for the MDP analysis we use τ_Φ-topology (see, Arcones [2]) enabling treatment of statistics having unbounded influence functions.

Due to involved structure of the rate function, the LDP result for $P_n^* \times \hat P_n$ is far from being "applicable" even for simple statistical cases (as exceptions, see special cases in Chaganty [7]). In contrast to that, the MDP provides readily derived asymptotics which are compatible with a majority of widespread statistics and thus the MDP effectively serves $T(\hat P_n) - T(P)$ and $T(P_n^*) - T(\hat P_n)$. In particular we show that the asymptotics of moderate deviation probabilities of $T(\hat P_n) - T(P)$ and $T(P_n^*) - T(\hat P_n)$ are calculated easily for the statistical functionals having the Hadamard derivatives.

The assumption of differentiability is the standard tool for the proof of asymptotic normality of statistics $T(\hat P_n)$ (see Serfling [25], van der Vaart and Wellner [27]) and, in implicit form, were also used for the study of moderate deviation probabilities (see Aleskeviciene [1], Jureckova, Kallenberg and Veraverbeke [20]; Inglot, Kallenberg and Ledwina [19], Arcones [3]). The moderate deviations of statistics were studied in Ermakov [15] for the case of Freshet derivative and Gao and Zhao [16] for the case of Hadamard derivatives. In [16] the statistical functionals were considered as functionals of empirical distribution functions and the technique of large deviations of stochastic processes was implemented. We consider the statistical functionals as the functionals of empirical probability measures or empirical bootstrap measure.

The remainder of the paper is organized as follows. In section 2 we present MDP for empirical bootstrap measure, empirical measure and conditional MDP for empirical bootstrap measure given empirical measure. The moderate deviation probabilities of statistical functionals are studied in section 3. The proofs of MDP and conditional MDP are given in sections 4 and 5 respectively.

2. MDP for empirical and empirical bootstrap measures

2.1. **Notations.** Throughout the paper, the following notations are implemented:
- $Q_2 \times Q_1$ the Cartesian product of probability measures $Q_2, Q_1 \in \Lambda$;
- $\Lambda^2 = \Lambda \times \Lambda$ denotes the set of all measures $Q_2 \times Q_1$ with $Q_2, Q_1 \in \Lambda$;
- C, c are generic positive constants;
- $\chi(A)$ is the indicator of event A;
- $[t]$ is the integral part of real number t;
- $\int$ always denotes $\int_S$.

2.2. **τ_Φ-topology.** We begin with the definition.

Let us fix a decreasing sequence of positive numbers $(b_n)_{n\ge1}$ with properties:

$$\left.\begin{array}{l} b_n \to 0 \\ nb_n^2 \to \infty \\ \frac{b_n}{b_{n+1}} \to 1 \end{array}\right\} n \to \infty. \tag{2.1}$$

Denote Φ the set of measurable functions $f : S \to \mathbb{R}$ with the following property:

$$\lim_{n\to\infty} \frac{1}{nb_n^2} \log(nP(|f(X)| > b_n^{-1})) = -\infty. \tag{2.2}$$

Let

$$\Lambda_\Phi = \Big\{ P \in \Lambda : \int |f(X)| dP < \infty, \quad for all f \in \Phi \Big\}.$$

The **coarsest** topology in Λ_Φ providing the continuous mapping

$$Q \Rightarrow \int f\, dQ, \text{ for all } f \in \Phi,\ Q \in \Lambda_\Phi$$

is known as the τ_Φ-topology (henceforth, all topological concepts refer to the τ_Φ-topology). Denote σ_Φ the smallest σ-field that makes all these mapping measurable.

For any set $\Omega \subset \Lambda_\Phi$ the notations: $\mathfrak{clo}(\Omega)$ and $\mathfrak{int}(\Omega)$ are used for the closure and interior of Ω respectively.

The τ_Φ-topology in Λ_Φ^2 is the corresponding product topology. For the set $\Phi_0 (\subset \Phi)$ of all real bounded measurable functions, the τ_Φ-topology coincides with the τ-topology (see GOR [17], Dembo and Zeitouni [10], Eichelsbacher and Schmock [13]). For $P, Q \in \Lambda$ and $P, Q \in \Lambda_\Phi$ we define the sets Λ_0 and $\Lambda_{0\Phi}$ respectively of of all differences $P - Q$. The τ_Φ-topologies in $\Lambda_{0\Phi}$ and $\Lambda_{0\Phi}^2$ are defined in a standard way as well as $\mathfrak{clo}(\bar\Omega_0)$ and $\mathfrak{int}(\bar\Omega_0)$ the closure and interior of $\bar\Omega_0 \subset \Lambda_{0\Phi}^2$.

2.3. **Rate functions.** For $G \in \Lambda_0$, let

$$\rho_0^2(G|P) = \begin{cases} \frac{1}{2}\int \left(\frac{dG}{dP}\right)^2 dP, & G \ll P \\ \infty, & \text{otherwise} \end{cases}$$

be the rate function (in statistical terms, $2\rho_0^2(G|P)$ is the Fisher information) which arises naturally in the MDP analysis of empirical measures $\hat P_n$ (see Borovkov and Mogulskii [5]; Gao and Zhao [16], Arcones [2] and Ermakov [15]).

In the bootstrap setting, a rate function (we shall denote it by ρ_{0b}^2) is constituted from two ones:

$$\rho_{0b}^2(\bar G|P) = \rho_0^2(G_2|P) + \rho_0^2(G_1|P),$$

where $\bar G = G_2 \times G_1 \in \Lambda^2_{0\Phi}$

2.4. **MDP for empirical bootstrap measure.** For any set $A \in \mathscr{F}$ and any signed measure $G \in \Lambda_0$ denote $|G|(A) = \sup\{G(B) - G(D) :\, B \subseteq A, D \subseteq A\}$. The measure $|G|$ is the variation of signed measure G.

Let the signed measures $H, H_n \in \Lambda_{0\Phi}$ satisfy the following assumptions.

A. We have $P_n = P + b_n H_n \in \Lambda_\Phi, P + b_n H \in \Lambda_\Phi$ and $H_n \to H$ as $n \to \infty$ in the τ_Φ-topology.

B1. For any $f \in \Phi$

$$\limsup_{n\to\infty} \sup_m (nb_n^2)^{-1} \log \left(nb_n \int \chi(|f(x)| > b_n^{-1})\, d|H_m| \right) = -\infty.$$

Define the signed measure $O \in \Lambda_{0\Phi}$ such that $O(A) = 0$ for all measurable sets $A \in \Im$. For each $G \in \Lambda_{0\Phi}$ denote $\tilde G = O \times G$.

Theorem 2.1. *Assume A and B1. Let $\bar\Omega_0 \subset \Lambda^2_{0\Phi}$ be σ_Φ measurable set of $\Lambda^2_{0\Phi}$. Then the Moderate Deviation Principle (MDP) holds*

$$\liminf_{n\to\infty} (nb_n^2)^{-1} \log P_n((P_n^* - \hat P_n) \times (\hat P_n - P_0) \in b_n \bar\Omega_0) \ge -\rho^2_{0b}(\mathfrak{int}(\bar\Omega_0 - \tilde H), P) \quad (2.3)$$

and

$$\limsup_{n\to\infty} (nb_n^2)^{-1} \log P_n((P_n^* - \hat P_n) \times (\hat P_n - P) \in b_n \bar\Omega_0) \le -\rho^2_{0b}(\mathfrak{cl}(\bar\Omega_0 - \tilde H), P). \quad (2.4)$$

Remark 2.1. In hypothesis testing, the type II error probabilities are often analyzed for the alternatives P_n converging to the hypothesis P. Theorem 2.1 allows to study moderate deviation probabilities in this setup. The analysis of importance sampling efficiency is also based on MDP with a sequence pm's P_n converging to pm P(see Ermakov [15]). Naturally, if we suppose that H_n, H are absent, we get the usual form of MDP. Bolthausen [4] has proved the Donsker-Varadhan LDP [11] when the laws of random variables converge weakly and a uniform exponential integration condition is satisfied. Theorem 2.1 and further Theorems can be considered as versions of these results.

The modern form of LDP-MDP (see de Acosta [8], Gao and Zhao [16], Leonard and Najim [21]) covers the case of unmeasurable sets $\bar\Omega_0$ and is given in terms of outer and inner probabilities. Let $(\Upsilon, \Im, P)$ be a probability space. The outer probability of an arbitrary subset $B \subset \Upsilon$ is

$$(P)^*(B) = \inf\{P(A); B \subseteq A, A \in \sigma_{\Lambda_{0\Theta_h}}\}$$

and $(P)_*(B) = 1 - (P)^*(\Lambda_{0\Theta_h} \setminus B)$ is the inner probability. All Theorems of the paper hold also for this setup. In Theorem 2.1 it suffices to replace pm's P_n in (2.3) with $(P_n)_*$ and pm's P_n in (2.4) with $(P_n)_*$.

The bootstrap procedures are often implemented with sample size $k \ne n$. In Theorem 2.2 given below the results are extended to this setting. Let $X_1^*, \dots, X_k^*$ be i.i.d.r.v.'s having pm $\hat P_n$. Denote P_k^* the empirical measure of $X_1^*, \dots, X_k^*$. Suppose that $k = k(n)$, $k/n \to \nu > 0$ as $n \to \infty$.

For any $\bar G = G_2 \times G_1 \in \Lambda_0^2$ define the rate function

$$\rho^2_{0\nu}(\bar G :\, P) = \nu\, \rho_0^2(G_2 \,: P) + \rho_0^2(G_1 \,: P).$$

For any $\bar\Omega_0 \subset \Lambda_0^2$ denote $\rho_{0\nu}(\bar\Omega_0 \,: P) = \inf\{\, \rho_{0\nu}(\bar G \,: P) \,: \bar G \in \bar\Omega_0 \,\}$.

Theorem 2.2. *Assume A and B1. Then the Moderate Deviation Principle (MDP) holds*

$$\liminf_{n\to\infty} (nb_n^2)^{-1} \log P_n((P_k^* - \hat P_n) \times (\hat P_n - P_0) \in b_n \bar\Omega_0) \ge -\rho^2_{0\nu}(\mathfrak{int}(\bar\Omega_0 - \tilde H), P) \quad (2.5)$$

and

$$\limsup_{n\to\infty}(nb_n^2)^{-1}\log P_n((P_k^* - \hat{P}_n)\times(\hat{P}_n - P) \in b_n\bar{\Omega}_0) \le -\rho_{0\nu}^2(\mathfrak{cl}(\bar{\Omega}_0 - \tilde{H}), P). \quad (2.6)$$

The proof of Theorem 2.2 is akin to that of Theorem 2.1 and is omitted. From now on, we assume $k = n$.

2.5. **MDP for empirical measure.** In section a version of Arcones [2] MDP for empirical processes is given in the setup for empirical pm. These results were established in Ermakov [15] and are given here for the comparison with the bootstrap setup.

The MDP for the empirical probability measures holds for wider zones of moderate deviations. Define the set Ψ of measurable functions $f : S \to R^1$ such that

$$\lim_{n\to\infty}(nd_n^2)^{-1}\log(nP(|f(X)| > nd_n)) = -\infty \quad (2.7)$$

where $d_n \to 0, nd_n^2 \to \infty, d_{n+1}/d_n \to 1$ as $n \to \infty$.

Assume the following.

B2. For any $f \in \Psi$

$$\lim_{n\to\infty}(nd_n^2)^{-1}\sup_m \log\left(nd_n\int \chi(|f(x)| > nd_n)\, d|H_m|\right) = -\infty. \quad (2.8)$$

Using the reasoning of Lemma 2.5 in Eichelsbacher and Löwe [14], we get that B1 or B2 implies

$$sup_m \int f^2 d|H_m| < \infty \quad (2.9)$$

and (2.2) or (2.7) implies

$$\int f^2 dP < \infty. \quad (2.10)$$

In Lemma 2.5 in [14] (2.10) has been proved, if d_n is decreasing and $n^{1/2}d_n$ is increasing. Since $d_n/d_{n-1} \to 1$ as $n \to \infty$ we can choose a subsequence d_{n_k} such that $n_k^{1/2}d_{n_k}$ is increasing and $d_{n_k}/d_{n_{k-1}} \to 1$ as $k \to \infty$. After that we can choose a subsequence $d_{n_{k_i}}$ such that $d_{n_{k_i}}$ is decreasing and $d_{n_{k_i}}/d_{n_{k_{i-1}}} \to 1$ as $i \to \infty$. Implementing to the subsequence $d_{n_{k_i}}$ the same reasoning as in the proof of Lemma 2.5 in [14] we get (2.10).

Theorem 2.3. *Assume A with $\Phi = \Psi$ and B2. Let Ω_0 be σ_Ψ measurable set of $\Lambda_{0\Psi}$. Then the MDP holds*

$$\liminf_{n\to\infty}(nd_n^2)^{-1}\log P_n(\hat{P}_n \in P + d_n\Omega_0) \ge -\rho_0^2(\mathfrak{int}(\Omega_0 - H), P_0) \quad (2.11)$$

and

$$\limsup_{n\to\infty}(nd_n^2)^{-1}\log P_n(\hat{P}_n \in P + d_n\Omega_0) \le -\rho_0^2(\mathfrak{cl}(\Omega_0 - H), P_0) \quad (2.12)$$

The Theorem 2.4 given below shows that the MDP for the empirical bootstrap measure can not be valid if (2.2) is replaced by (2.7).

Theorem 2.4. *Let the random variable $Y = |f(X)|$ satisfy (2.2). Let the sequences r_n and e_n be such that $b_n^{-1} < r_n, b_n^{-1}e_n \to \infty, ne_n/r_n \to \infty$ as $n \to \infty$ and*

$$\lim_{n\to\infty}(ne_n^2)^{-1}\log(nP(Y > r_n)) = 0, \quad (2.13)$$

$$\lim_{n\to\infty}(r_ne_n)^{-1}\log\frac{ne_n}{r_n} = 0. \quad (2.14)$$

Let $Y_1,\dots,Y_n$ be independent copies of Y and let $Y_1^*,\dots,Y_n^*$ be obtained from $Y_1,\dots,Y_n$ using the bootstrap procedure. Then

$$\lim_{n\to\infty}(ne_n^2)^{-1}\log P\left(\sum_{i=1}^n Y_i^* > ne_n\right) = 0. \tag{2.15}$$

The proof of Theorem 2.5 is given in Appendix.

Example. Let $P(Y > t) = \exp\{-t^\gamma\}, 0 < \gamma < 1$. Then $b_n = o(n^{-\frac{1}{2+\gamma}})$. By straightforward calculations we get that (2.13), (2.14) hold for any sequence $r_n = n^{\frac{1}{2+\gamma}} f_n$, $e_n = n^{-\frac{1}{2+\gamma}} f_n^{\frac{\gamma}{2}-\delta}$ where $(\log n)^{\frac{1}{1+\frac{\gamma}{2}-\delta}} << f_n << n^{\frac{\gamma}{(2+\gamma)(1+\delta)}}$ and $0 < \delta < \frac{\gamma}{2}$. Thus the moderate deviation zone in Theorem 2.1 can not be improved essentially for such an asymptotic of $P(Y > t)$.

2.6. **MDP for conditional bootstrap measure.** Theorem 2.5 given below shows that the MDP holds almost sure (a.s.) for the conditional distribution of the empirical bootstrap measure given the empirical probability measure. We call this version of MDP the conditional MDP. In this model we allow the sample size $k = k_n$ of bootstrap procedures to have values different from n.

Almost sure version of conditional LDP for the bootstraped sample mean has been established Li, Rosalski and Al-Mutairi [22]. Chaganty and Karandikar [6] have proved conditional LDP for empirical bootstrap measure in the case of weak topology.

For a sequence of arbitrary random variables $Z_n : S \to R^1$ (Z_n may not be Borel measurable) we say that $\liminf_{n\to\infty} Z_n \ge c$ inner almost surely ($a.s_*$ if there exist a sequence Δ_n of measurable random variables $\Delta_n \le Z_n$ such that $P(\liminf_{n\to\infty} \Delta_n \ge c) = 1$.

We say that $\limsup_{n\to\infty} Z_n \le C$ outer almost surely ($a.s^*$.) if $\liminf_{n\to\infty} -Z_n \ge -c\ a.s_*$.

We say that $\limsup_{n\to\infty} Z_n = -\infty$ outer almost surely ($a.s^*$.) if $\liminf_{n\to\infty} -Z_n \ge -c\ a.s_*$ for any $C > 0$.

Let $X_1^*,\dots,X_{k_n}^*$ be i.i.d.r.v.'s having pm $\hat P_n$. Denote $P_{k_n}^*$ the empirical probability measure of $X_1^*,\dots,X_{k_n}^*$. Suppose that $\frac{k_n}{n} < c < \infty$ and $k_n \to \infty$ as $n \to \infty$.

For each $t > 2$ define the set $\Theta = \Theta_t$ of real functions $f : S \to R^1$ such that $E[|f(X)|^t] < \infty$.

For decreasing function $h : R_+^1 \to R_+^1$ and $t \ge 2$ define the set $\Theta = \Theta_{2,h}$ of real functions f such that

$$P(|f(X)| > s^{-1}) < h(s), \quad s > 0 \tag{2.16}$$

and

$$E[f^2(X)] < \infty. \tag{2.17}$$

All results are given bellow in terms of τ_Θ-topology.

Theorem 2.5. *Let a sequence $a_n > 0, a_n \to 0, a_{n+1}/a_n \to 1, k_n a_n^2 \to \infty$ as $n \to \infty$ be given. Let*

$$\sum_{n=1}^\infty h(ca_n) < \infty \tag{2.18}$$

for any $c > 0$.

Let $\Omega_0 \subset \Lambda_{0\Theta_{2,h}}$. Then there hold

$$\liminf_{n\to\infty}(k_n a_n^2)^{-1}\ln(\hat P_n)_*(P_{k_n}^* \in \hat P_n + a_n\Omega_0) \ge -\rho_0^2(\mathfrak{int}(\Omega_0), P) \quad a.\,s_*. \tag{2.19}$$

and

$$\limsup_{n\to\infty}(k_n a_n^2)^{-1}\ln(\hat P_n)^*(P_{k_n}^* \in \hat P_n + a_n\Omega_0) \le -\rho_0^2(\mathfrak{cl}(\Omega_0), P) \quad a.\,s^*. \tag{2.20}$$

where the closure and the interior of the set Ω_0 *in (2.19) and (2.20) are considered with respect to* $\tau_{\Theta_{2,h}}$ *-topology. The outer probability measure* $(\hat{P}_n)^*$ *and the inner probability measure* $(\hat{P}_n)_*$ *are considered with respect to* $\sigma_{\Theta_{2,h}}$*-algebra.*

Let $\Omega_0 \subset \Lambda_{0\Theta_t}, t > 2$ *and let* $a_n = o(n^{-1/t})$*. Then (2.19) and (2.20) are valid with the closure and the interior of the set* Ω_0 *with respect to* τ_{Θ_t}*-topology. The outer probability measure* $(\hat{P}_n)^*$ *and the inner probability measure* $(\hat{P}_n)_*$ *are considered with respect to* σ_{Θ_t}*-algebra.*

Example. Let $E[\exp\{c|f(X_1)|^\gamma\}] < \infty$ with $\gamma > 0$ for all $f \in \Theta$. Then we have the following asymptotics

$$b_n = o\left(n^{-\frac{1}{1+\gamma}}\right), \quad d_n = o\left(n^{-\frac{1-\gamma}{2-\gamma}}\right)$$

and

$$a_n = o\left(|\log n|^{-\gamma}\right).$$

Thus the conditional MDP for empirical bootstrap measure holds for the wider zone than the usual MDP for the empirical measure.

Theorem 2.6 given below gives rates of convergence in the conditional MDP.

Theorem 2.6. *Let a sequence* $a_n > 0, a_n \to 0, a_{n+1}/a_n \to 1, k_n a_n^2 \to \infty$ *as* $n \to \infty$ *be given. Let function* $h : R^1_+ \to R^1_+$ *be such that*

$$\lim_{n\to\infty} nh\left(\frac{a_n}{c}\right) = 0 \tag{2.21}$$

for any $c > 0$.

Let $\Omega_0 \subset \Lambda_{0\Theta_{t,h}}, t > 2$*. Then for any* $\epsilon > 0$ *and* $n > n_0(\epsilon, \{k_i\}_{i=1}^\infty, \Omega_0)$ *there hold*

$$(k_n a_n^2)^{-1} \log(\hat{P}_n)_*(P^*_{k_n} \in \hat{P}_n + a_n\Omega_0) \ge -\rho_0^2(\mathfrak{int}_{\Theta_{t,h}}(\Omega_0), P) - \epsilon \tag{2.22}$$

and, if $\rho_0^2(\mathfrak{cl}_{\Theta_{t,h}}(\Omega_0), P) < \infty$ *additionally,*

$$(k_n a_n^2)^{-1} \log(\hat{P}_n)^*(P^*_{k_n} \in \hat{P}_n + a_n\Omega_0) \le -\rho_0^2(\mathfrak{cl}(\Omega_0), P) + \epsilon \tag{2.23}$$

on the sets of events having the inner probabilities more than $\kappa_n = \kappa_n(\epsilon, \Omega_0) = 1 - C(\epsilon, \Omega_0)[\beta_{1n} + \beta_{2n}]$ *where* $\beta_{1n} = nh(\frac{a_n}{\epsilon C_1(\epsilon,\Omega_0)})$ *and* $\beta_{2n} = C_2(\epsilon, \Omega_0)n^{1-t}$.

If $\rho_0^2(\mathfrak{cl}(\Omega_0), P) = \infty$*, for any* $L > 0$

$$(k_n a_n^2)^{-1} \log(\hat{P}_n)^*(P^*_{k_n} \in \hat{P}_n + a_n\Omega_0) \le -L \tag{2.24}$$

on the sets of events having the inner probabilities more than $\kappa_{1n} = \kappa_{1n}(L, \Omega_0) = 1 - C(L, \Omega_0)[\beta_{1n} + \beta_{2n}]$ *where* $\beta_{1n} = nh(\frac{a_n}{C_1(L,\Omega_0)})$ *and* $\beta_{2n} = C_2(L, \Omega_0)n^{1-t}$.

Remark 2.2. The method of the proof of Theorem 2.6 is the following. Let $f \in \Theta_{t,h}$. Denote $Y_i = |f(X_i)|, 1 \le i \le n$ and let $Y^{(1)} \le Y^{(2)} \le \ldots \le Y^{(n)}$ be the order statistics of $Y_1, \ldots, Y_n$. Using $P(\max_{1\le i\le n} f(X_i)| \le C_1(\Omega_0)\epsilon a_n^{-1}) \ge 1 - nh(\frac{a_n}{C_1(\Omega_0)\epsilon})$ we, with probability $1 - \beta_{1n}$, prove conditional if $|f(X_i)| < C_1(\Omega_0)\epsilon a_n^{-1}, 1 \le i \le n$ holds. The rate function in the conditional MDP is defined by $\frac{1}{n}\sum_{s=1}^n f^2(X_s)$. Thus we need to estimate the rate of convergence of $\frac{1}{n}\sum_{s=1}^n f^2(X_s)$ to $Ef^2(X)$. These estimates causes the second term β_{2n} in κ_n.

3. Moderate deviation probabilities of statistical functionals

In section we compare the asymptotics of moderate deviation probabilities of statistics $T(\hat{P}_n) - T(P_n)$ and $T(P_n^*) - T(\hat{P}_n)$. We suppose that the functional $T : \Lambda \to R^1$ has the Hadamard derivative or homogeneous.

3.1. **Functionals having the Hadamard derivatives.** For all $r > 0$ define the set $\Gamma_{0r} = \{G : \rho_0^2(G : P) < r, G \in \Lambda_0\}$.

Let Y be a metric linear topological space with metric ρ. We say that the functional $T : \Lambda_{0\Sigma} \to Y$ has the Hadamard derivative $T' : \Lambda_{0\Sigma} \to Y$ if the following assumption C_Σ holds with $\Sigma = \Psi$, $\Sigma = \Phi$ or $\Sigma = \Theta$.

C_Σ. For any $r > 0$ for each $G \in \Gamma_{0r}$ and any sequence $G_k \in \Gamma_{0r}$ converging to G in τ_Σ-topology there holds

$$\limsup_{n\to\infty} \rho(u_k^{-1}(T(P_0 + u_k G_k) - T(P_0)) - T'(G), 0) = 0 \tag{3.1}$$

for all sequences $u_k \to 0$ as $k \to \infty$ and $u_k \neq 0, 1 \le k < \infty$.

Theorem 3.1. *Assume $A, B2, C_\Psi$. Let the functional $T(P)$ be continuous in τ_Ψ-topology. Then, for any set $\Omega \subset Y$, there holds*

$$\begin{aligned}&\liminf_{n\to\infty}(nb_n^2)^{-1} \ln P_n(T(\hat P_n) - T(P_n) \in b_n\Omega)\\ &\ge -\inf\{\rho_0^2(G : P_0) : T'(G) \in \mathfrak{int}(\Omega), G \in \Lambda_0\}\end{aligned} \tag{3.2}$$

and

$$\begin{aligned}&\limsup_{n\to\infty}(nb_n^2)^{-1} \ln P_n(T(\hat P_n) - T(P_n) \in b_n\Omega)\\ &\le -\inf\{\rho_0^2(G : P_0) : T'(G) \in \mathfrak{cl}(\Omega), G \in \Lambda_0\}.\end{aligned} \tag{3.3}$$

If $T'(G)$ is continuous in τ_Ψ-topology, then, for any $\delta > 0$

$$\limsup_{n\to\infty}(nb_n^2)^{-1} \ln(P_n)^*(\rho(b_n^1(T(\hat P_n) - T(P_n) - T'(\hat P_n - P_0)), 0) \ge \delta) = -\infty \tag{3.4}$$

The Hadamard differentiability of statistical functionals in the Kolmogorov metric (supnorm of difference of distribution functions) are the standard tool for the proofs of asymptotic normality (see van der Vaart and Wellner [27] Ch 3.9 and references therein). Gao and Zhao [16] extended this approach on the moderate deviation zone. The Kolmogorov metrics is continuous w.r.t. τ_Ψ-topology (see Groeneboom, Oosterhoff and Ruymgaart [17] Lemma 2.1 and Ermakov [15] Lemma 4.1). If $S = [a, b] \subseteq R^1$, by contraction principle (see Theorem 4.2.1 in Dembo and Zeitouni [10]) this implies that MDP holds for the set of empirical distribution functions lying in the Banach space of all right continuous with left-hand limits functions $z : [a, b] \to R^1$ equipped with the uniform norm. T Thus Theorem 3.1 in [16] can be replaced with Theorem 3.1 of this paper for the study of moderate deviations of estimators. At the same time Theorems 3.2 and 3.3 allow to get the results for the bootstrap setup as well.

Theorem 3.2. *Assume $A, B1, C_\Phi$. Let the functional $T(P)$ be continuous in τ_Φ-topology. Then, for any set $\bar\Omega \subset Y \times Y$, there holds*

$$\begin{aligned}&\liminf_{n\to\infty}(nd_n^2)^{-1} \ln P_n((T(P_n^*) - T(\hat P_n), T(\hat P_n) - T(P_n) \in b_n\bar\Omega)\\ &\ge -\inf\{\rho_{0b}^2(\bar G : P_0) : (T'(G_2), T'(G_1)) \in \mathfrak{int}(\bar\Omega), G_2 \times G_1 \in \Lambda_0^2\}\end{aligned} \tag{3.5}$$

and

$$\begin{aligned}&\limsup_{n\to\infty}(nd_n^2)^{-1} \ln P_n((T(P_n^*) - T(\hat P_n), T(\hat P_n) - T(P_n) \in b_n\bar\Omega)\\ &\le -\inf\{\rho_{0b}^2(\bar G : P_0) : (T'(G_2), T'(G_1)) \in \mathfrak{cl}(\bar\Omega), G_2 \times G_1 \in \Lambda_0^2\}\end{aligned} \tag{3.6}$$

If $T'(G)$ is continuous in τ_Φ-topology, then, for any $\delta > 0$

$$\begin{aligned}&\limsup_{n\to\infty}(nd_n^2)^{-1} \ln(P_n)^*(\rho(d_n^{-1}(T(P_n^*) - T(\hat P_n) - T'(P_n^* - \hat P_n), 0) > \delta,\\ &\rho(b_n^1(T(\hat P_n) - T(P_n) - T'(\hat P_n - P_0)), 0) \ge \delta) = -\infty\end{aligned} \tag{3.7}$$

Theorem 3.3. *Let a sequence $a_n > 0, a_n \to 0, a_{n+1}/a_n \to 1, k_n a_n^2 \to \infty$ as $n \to \infty$ be given. Let a decreasing function $h : R^1_+ \to R^1_+$ satisfy (2.18). Let the functional $T(P)$ be continuous in $\tau_{\Theta_{2,h}}$-topology and let $C_{\Theta_{2,h}}$ be valid.*

Then, for any set $\Omega \subset Y$, there holds

$$\begin{aligned}&\liminf_{n\to\infty}(k_n a_n^2)^{-1}\ln(\hat P_n)_*(T(P^*_{k_n}) - T(\hat P_n) \in b_n\Omega)\\ &\ge -\inf\{\rho_0^2(G : P_0) : T'(G) \in \mathfrak{int}(\Omega), G \in \Lambda_0\} \quad a.\,s_*.\end{aligned} \tag{3.8}$$

and

$$\begin{aligned}&\limsup_{n\to\infty}(k_n a_n^2)^{-1}\ln(\hat P_n)^*(T(P^*_{k_n}) - T(\hat P_n) \in b_n\Omega)\\ &\le -\inf\{\rho_0^2(G : P_0) : T'(G) \in \mathfrak{cl}(\Omega), G \in \Lambda_0\} \quad a.\,s^*.\end{aligned} \tag{3.9}$$

If $T'(G)$ is continuous in $\tau_{\Theta_{2h}}$-topology, then, for any $\delta > 0$,

$$\limsup_{n\to\infty}(k_n a_n^2)^{-1}\ln(\hat P_n)^*(\rho(b_n^{-1}(T(P^*_n) - T(\hat P_n) - T'(P^*_n - \hat P_n), 0) > \delta) = -\infty \ a.s^*. \tag{3.10}$$

Let the functional $T(P)$ be continuous in τ_{Θ_t}-topology, $t > 2$ and let C_{Θ_t} be valid. Let $\Omega \subset Y$ and let $a_n = o(n^{-1/t})$. Then (3.8), and (3.9) hold. If $T'(G)$ is continuous in τ_{Θ_t}-topology, then (3.10) is valid as well.

Theorem 3.4. *Let a sequence $a_n > 0, a_n \to 0, a_{n+1}/a_n \to 1, k_n a_n^2 \to \infty$ as $n \to \infty$ be given. Let a decreasing function $h : R^1_+ \to R^1_+$ satisfy (2.21). Let the functional $T(P)$ be continuous in $\tau_{\Theta_{t,h}}$-topology and let $C_{\Theta_{t,h}}$ be valid with $t > 2$.*

Let $\Omega \subset Y$. Then, for any $\epsilon > 0$ and $n > n_0(\epsilon, \{k_i\}_{i=1}^\infty, \Omega, T)$, there hold

$$\begin{aligned}&\liminf_{n\to\infty}(k_n a_n^2)^{-1}\ln(\hat P_n)_*(T(P^*_{k_n}) - T(\hat P_n) \in b_n\Omega)\\ &\ge -\inf\{\rho_0^2(G : P_0) : T'(G) \in \mathfrak{int}(\Omega), G \in \Lambda\} - \epsilon\end{aligned} \tag{3.11}$$

and, if $\inf\{\rho_0^2(G : P_0) : T'(G) \in \mathfrak{cl}(\Omega), G \in \Lambda_0\} < \infty$ additionally,

$$\limsup_{n\to\infty}(k_n a_n^2) - \inf\{\rho_0^2(G : P_0) : T'(G) \in \mathfrak{cl}(\Omega), G \in \Lambda_0\} + \epsilon \tag{3.12}$$

on the sets of events having the inner probabilities more than $\kappa_n = \kappa_n(\epsilon, \Omega, T) = 1 - C(\epsilon, \Omega)[\beta_{1n} + \beta_{2n}]$ where $\beta_{1n} = nh(\frac{a_n}{\epsilon C_1(\epsilon,\Omega,T)})$ and $\beta_{2n} = C_2(\epsilon, \Omega, T)n^{1-t}$.

If $T'(G)$ is continuous in τ_Θ-topology, then, for any $\delta > 0$ and any $L > 0$ there exists $n_0 = n_0(L, \delta, \{k_i\}_{i=1}^\infty, T)$ such that for all $n > n_0$

$$\limsup_{n\to\infty}(k_n a_n^2)^{-1}\ln(\hat P_n)^*(\rho(b_n^{-1}(T(P^*_n) - T(\hat P_n) - T'(P^*_n - \hat P_n), 0) > \delta) < -L \tag{3.13}$$

on the sets of events having the inner probabilities more than $\bar\kappa_n = \bar\kappa_n(L, \delta) = 1 - C(L, \delta)[\bar\beta_{1n} + \bar\beta_{2n}]$ where $\bar\beta_{1n} = nh\left(\frac{a_n}{C_1(L,\delta,T)}\right)$ and $\bar\beta_{2n} = C_2(L, \delta, T)n^{1-t}$.

For the proof of Theorem 3.1 it suffices to implement the contraction principle of Theorem 4.2.23 in [10] to the sequence of functions $f_k(G) = b_k^{-1}(T(P_n + b_k G) - T(P_n))$. In Theorem 4.2.23 in [10] it is assumed that

$$\limsup_{k\to\infty}\sup_{G\in\Gamma_{0r}}\rho(f_k(G), T'(G)) = 0. \tag{3.14}$$

Since Γ_r is compact and sequentially compact set in the τ_Ψ-topology (see EIchelsbacher and Schmock [13] Lemma 2.1) then (3.14) follows from (3.1).

The proof of (3.4) is akin to the proof of similar statement (3.4) of Theorem 3.1 in Gao and Zhao [16] and Theorem 3.9.4 in van der Vaart and Wellner [27]. We consider the mapping $\phi_k : \Lambda_{0\Theta} \to Y \times Y$ with $\phi_k(G) = (f_k(G), T'(G))$ for all $G \in \Lambda_{0\Theta}$. By (3.2),(3.3), we get that $\phi_n(\hat P_n - P_0)$ satisfies MDP with the rate function

$$\bar\rho^2(y_1, y_2) = \inf\{\rho_0^2(G : P_0) : T'(G) = y_1 = y_2\} \quad (y_1, y_2) \in Y \times Y.$$

Hence, by the classical contraction principle (see Theorem 4.2.1 in Dembo and Zeitouni [10]), we get (3.4).

The reasoning in the proofs of Theorems 3.2, 3.3 and 3.4 are similar.

3.2. **Examples.** In what follows, $Y = R^1$ and we shall suppose that the assumptions of Theorems 3.1 and 3.2 are satisfied in the case of moderate deviations of empirical measure and bootstrap measure respectively. In this case C_Σ can be rewritten in the following form.

$C1_\Sigma$ There exists $h : S \to R^1, E[h(X_1)] = 0$ such that, for any $r > 0$ for each $G \in \Gamma_{0r}$ and any sequence $G_k \in \Gamma_{0r}$ converging to G in τ_Σ-topology there holds

$$\lim_{k\to\infty} u_k^{-1}(T(P_0 + u_k G_k) - T(P_0)) - \int h dG = 0 \tag{3.15}$$

for all sequences $u_k \to 0$ as $k \to 0$ and $u_k \neq 0, 1 \le k < \infty$.

By Theorem 3.2, we get

$$\begin{aligned}
&\lim_{n\to\infty} (nb_n^2)^{-1} \log P_n(T(P_n^*) - T(\hat{P}_n) > b_n) \\
&= \lim_{n\to\infty} (nb_n^2)^{-1} \log P_n \left(\int h\, d(P_n^* - \hat{P}_n) > b_n \right) \\
&= -\frac{1}{2} \inf \left\{ \int (g_2^2 + g_1^2)\, dP : \int g_2 h\, dP > 1,\ g_1, g_2 \in L_2(P) \right\} \\
&= -\frac{1}{2} \left(\int h^2\, dP \right)^{-1}
\end{aligned} \tag{3.16}$$

and, by Theorem 3.1,

$$\begin{aligned}
&\lim_{n\to\infty} (nd_n^2)^{-1} \log P_n(T(\hat{P}_n) - T(P_n) > d_n) \\
&= \lim_{n\to\infty} (nb_n^2)^{-1} \log P_n \left(\int h\, d(\hat{P}_n - P_n) > d_n \right) \\
&= -\frac{1}{2} \inf \left\{ \int g^2\, dP : \int gh\, dP > 1,\ g \in L_2(P) \right\} = -\frac{1}{2} \left(\int h^2 dP \right)^{-1}.
\end{aligned} \tag{3.17}$$

By Theorem 3.3, we get

$$\begin{aligned}
&\liminf_{n\to\infty} (k_n a_n^2)^{-1} \log (\hat{P}_n)_* (T(P_{k_n}^*) - T(\hat{P}_n) > a_n) \\
&= \limsup_{n\to\infty} (k_n a_n^2)^{-1} \log (\hat{P}_n)^* (T(P_{k_n}^*) - T(\hat{P}_n) > a_n) = -\frac{1}{2} \left(\int h^2 dP \right)^{-1} \quad \text{a.s.}
\end{aligned} \tag{3.18}$$

Thus, the asymptotics of moderate deviations probabilities of $T(P_n^*) - T(\hat{P}_n)$ and $T(\hat{P}_n) - T(P_n)$ coincide. At the same time

$$\begin{aligned}
&\lim_{n\to\infty} (nb_n^2)^{-1} \log P_n(T(P_n^*) - T(P_n) > b_n) \\
&= \lim_{n\to\infty} (nb_n^2)^{-1} \log P_n \left(\int hd(P_n^* - P_n) > b_n \right) \\
&= \frac{1}{2} \inf \left\{ \int (g_2^2 + g_1^2)\, dP : \int (g_2 - g_1) h\, dP > 1,\ g_1, g_2 \in L_2(P) \right\} \\
&= -\frac{1}{4} \left(\int h^2 dP \right)^{-1}.
\end{aligned} \tag{3.19}$$

3.3. **Homogeneous functionals.** Let $N : \Lambda_{0\Phi} \to R^1$ be a seminorm continuous in the τ_Φ-topology. Define the set $\Omega_0 = \{G : N(G) > 1, G \in \Lambda_{0\Phi}\}$ and let the signed measure $H \in \mathfrak{cl}(\Omega_0)$ be such that $\rho_0^2(H : P) = \frac{1}{2}\int h^2\, dP = \rho_0^2(\Omega_0 : P)$ with $h = \frac{dH}{dP}$. Then we have

$$\lim_{n\to\infty} (nd_n^2)^{-1} \log P(N(\hat{P}_n - P) > d_n) = -\rho_0^2(\Omega_0 : P) = -\frac{1}{2}\int h^2 dP, \tag{3.20}$$

$$\begin{aligned}&\lim_{n\to\infty} (nb_n^2)^{-1} \log P(N(P_n^* - \hat{P}_n) > b_n) =\\ &-\frac{1}{2}\inf\left\{\int (g_2^2 + g_1^2)dP :\ N(G_2) \ge 1; g_1 = \frac{dG_1}{dP}, g_2 = \frac{dG_2}{dP}; G_2, G_1 \in \Lambda_{0\Phi}\right\} =\\ &-\rho_0^2(\Omega_0 : P) = -\frac{1}{2}\int h^2 dP.\end{aligned} \tag{3.21}$$

and

$$\lim_{n\to\infty} (nb_n^2)^{-1} \log P(N(P_n^* - P) > b_n) = -\frac{1}{2}\rho_0^2(\Omega_0 : P) = -\frac{1}{4}\int h^2 dP. \tag{3.22}$$

In particular, the statements (3.20) and (3.22) are valid for the functionals

$$N(Q-P) = N_1(Q-P,P) = \max\{|F(x) - F_0(x)|q(F_0(x)) : x \in S\} \tag{3.23}$$

and

$$N(Q-P) = N_2(Q-P,P) = \left(\int_S (F(x) - F_0(x))^2 q(F_0(x)) dF_0(x)\right)^{1/2} \tag{3.24}$$

respectively. Here q is a bounded weight function, $S = R^1$ and F, F_0 are the distribution functions of Q, P respectively. If $q \equiv 1$, $N_1(\hat{P}_n - P, P)$ and $N_2^2(\hat{P}_n - P, P)$ are Kolmogorov and omega-squared test statistics respectively. The functionals N_1, N_2 depend on the probability measure P additionally. Thus (3.21) holds only in the case of $q \equiv 1$. Let us show that, if q is continuous in $[0,1]$ the presence of weight function q does not influence seriously on the asymptotic (3.21), that is,

$$\begin{aligned}&\lim_{n\to\infty} (nb_n^2)^{-1} \log P(N_i(P_n^* - \hat{P}_n, \hat{P}_n) > b_n) =\\ &\lim_{n\to\infty} (nb_n^2)^{-1} \log P(N_i(P_n^* - \hat{P}_n, P) > b_n) =\\ &-\rho_0^2(\Omega_0 : P) = -\frac{1}{2}\int h^2\, dP\end{aligned} \tag{3.25}$$

with $i = 1, 2$. Note that, if q is continuous in [0,1], the following assumption holds.

C1. There exists function $\omega(t), \omega(t)/t \to 0$ as $t \to 0$ such that, for all $P, Q, R \in \Lambda_\Phi$

$$|N(Q-P,P) - N(Q-P,R)| \le \omega(\sup_x |\bar{F}(x) - F_0(x)|)$$

where $\bar{F}$ stands for the distribution function of R.

Let $\hat{F}_n$ be the distribution function of $\hat{P}_n$. Then, by Theorem 2.3,

$$P(\omega(\sup_x |\hat{F}_n(x) - F_0(x)|) > cb_n) \le \exp\{-CnC_n b_n^2\}$$

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

Hence, estimating similarly to the proof of (3.17) in [15] we get (3.25).

Let us find the asymptotic

$$\lim_{n\to\infty} (nb_n^2)^{-1} \log P(N^\gamma(P_n^* - P) - N^\gamma(\hat{P}_n - P) > b_n) \doteq -\frac{1}{2}J$$

with $\gamma > 0$.

By Theorem 2.1, we get

$$J = \inf\left\{ \int (r^2+g^2)\,dP : N^\gamma(G+R) - N^\gamma(G) \ge 1; \right.$$
$$\left. g = \frac{dG}{dP}, r = \frac{dR}{dP}; G, R \in \Lambda_{0\Phi} \right\} \doteq \inf V(G,R). \tag{3.26}$$

Since $N(G+R) \le N(G) + N(R)$, we get

$$J \ge \inf\left\{ \int (r^2+g^2)\,dP : (N(G)+N(R))^\gamma - N^\gamma(G) \ge 1; \right.$$
$$\left. g = \frac{dG}{dP}, r = \frac{dR}{dP}; G, R \in \Lambda_{0\Phi} \right\} \doteq \inf U(G,R). \tag{3.27}$$

It is easy to see that for the fixed G

$$\arg\inf_R U(G,R) = \lambda H, \quad \lambda = \lambda(G) > 0 \tag{3.28}$$

where the signed measure $H \in \mathrm{cl}(\Omega_0)$ is the same as in example 3.3.

Let $r = \lambda h$ be fixed and let us consider the problem of minimization of $U(G, \lambda H)$ with respect to G. We begin with the dual problem. Let $N(R) = d$ =const and one needs to find

$$\sup\left\{ (N(G)+d)^\gamma - N^\gamma(G) : \int g^2\,dP = 1, g = \frac{dG}{dP}, G \in \Lambda_{0\Phi} \right\}$$

Let $\gamma \ge 1$. Since the function $(x+d)^\gamma - x^\gamma$ is increasing the supremum is attained on the charge $G_0 = c\tilde{G}$ where $\tilde{G} = \arg\sup\{N(G) : \int g^2\,dP = 1, g = \frac{dG}{dP}, G \in \Lambda_{0\Phi}\}$ and $\tilde{g} = \frac{d\tilde{G}}{dP} = h/\rho_0$. Therefore $\inf\{U(G,R) : G, R \in \Lambda_{0\Phi}\}$ is attained on the signed measures G, R having the densities $g = ah$, $r = dh$ with $a, d \in R^1$. However $V(aH, dH) = U(aH, dH)$. Hence we get

$$J = \inf\{d^2 + a^2 : (d+a)^\gamma - a^\gamma > 1\} \int h^2(s)\,dP. \tag{3.29}$$

If $\gamma < 1$, then $\arg\sup\{(x+d)^\gamma - x^\gamma : x \ge 0\} = 0$. Therefore $\inf\{U(G,R) : G, R \in \Lambda_{0\Phi}\} = d^\gamma$ and

$$J = \inf\left\{ \int r^2 dP : N^\gamma(R) \ge 1, r = \frac{dR}{dP} \right\} = 2\rho_0^2(\Omega_0, P) = \int h^2\,dP.$$

4. Proof of Theorem 2.1

For each $r > 0$ define the set $\Gamma_r = \{\bar{G} \in \Lambda_0^2 : \rho_{0b}^2(\bar{G} : P) \le r\}$.

Lemma 4.1. *Let (2.7) hold. Then*

i. $\Gamma_r \subset \Lambda_{0\Psi}^2$,

ii. the set Γ_r *is* τ_Ψ*-compact and sequentially* τ_Ψ*-compact set in* $\Lambda_{0\Psi}^2$,

iii. the τ *and* τ_Ψ*- topologies coincide in* Γ_r.

Proof. The reasoning are akin to the proof of Lemma 2.1 in Eichelsbacher and Schmock [13]. For any charge $\bar{G} = G_1 \times G_2 \in \Gamma_r$, any measurable set $A \subseteq S$ and any $\phi_1, \phi_2 \in \Psi$ we have

$$\int_A |\phi_1|\,d|G_1| + \int_A |\phi_2|\,d|G_2| \le$$
$$\alpha\left(\int_A \phi_1^2\,dP + \int_A \phi_2^2\,dP \right) + \alpha^{-1}\left(\int_A \left(\frac{dG_1}{dP}\right)^2 dP + \int_A \left(\frac{dG_2}{dP}\right)^2 dP \right) \tag{4.1}$$

for all $\alpha > 0$. By (2.10), this implies i if $A = S$.

Fix $\epsilon > 0$. Let $\alpha = r/\epsilon$ and let $n = n(\epsilon)$ be such that

$$\frac{r}{\epsilon}\left(\int_{|\phi_1|>n} \phi_1^2\, dP + \int_{|\phi_2|>n} \phi_2^2\, dP\right) < \epsilon$$

Then

$$\alpha^{-1}\left(\int_{|\phi_1|>n} \left(\frac{dG_1}{dP}\right)^2 dP + \int_{|\phi_2|>n} \left(\frac{dG_2}{dP}\right)^2 dP\right) \le \epsilon$$

Hence, by (4.1), we get

$$\int |\phi_1|\, d|G_1| + \int |\phi_2|\, d|G_2| - \int_{|\phi_1|<n} |\phi_1|\, d|G_1| - \int_{|\phi_2|<n} |\phi_2|\, d|G_2| < 2\epsilon$$

Therefore the map $\Gamma_r \ni \bar{G} = G_1 \times G_2 \to \int |\phi_1|\, d|G_1| + \int |\phi_2|\, d|G_2|$ is τ-continuous as the uniform limit of functions

$$\int_{|\phi_1|<n} \phi_1\, dG_1 + \int_{|\phi_2|<n} \phi_2\, dG_2.$$

This implies that the τ and τ_Ψ-topologies coincide in Γ_r. Since the sets Γ_{0r} and $\Gamma_r \subset \Gamma_{0r}^2$ are τ-compact and sequentially τ-compact these sets are τ_Ψ-compact and sequentially τ_Ψ-compact as well. This completes the proof of Lemma 4.1.

The same reasoning of the proof of Lemma 4.1 can be repeated in the case of τ_Φ-topology. Thus the sets Γ_{0r} are τ_Φ- compact as well.

In Lemmas 4.2-4.5 given bellow we suppose that the assumptions of Theorem 2.1 hold.

For any $u, v \in R^k$ denote $u'v$ the inner product of u and v. For any $f \in \Phi$ and any signed measure $G \in \Lambda_{0\Phi}$ denote $< f, G >= \int f dG$.

Let $f_1, \dots, f_{k_1}, g_1, \dots, g_{k_2} \in \Phi$ and $G \in \Lambda_{0\Phi}$. Let $E[f_i(X)] = 0, E[g_j(X)] = 0$ with $1 \le i \le k_1, 1 \le j \le k_2$. Define the covariance matrices $R_f = \{E[f_i(X)f_j(X)]\}_{i,j=1}^{k_1}$ and $R_g = \{E[g_i(X)g_j(X)]\}_{i,j=1}^{k_2}$. Denote $\vec{f} = \{f_i\}_{i=1}^{k_1}$, $\vec{g} = \{g_i\}_{i=1}^{k_2}$ and $\bar{g}_i = \frac{1}{n}\sum_{l=1}^n g_i(X_l), 1 \le i \le k_2$.

By a version of Dawson-Gartner Theorem (see [10] Theorem 4.6.9 and [21]), Theorem 2.1 follows from Lemma 4.2 given below. Note that the de Acosta [2] approach (see section 5) also allows to deduce Theorem 2.1 from Lemma 4.2.

Lemma 4.2. *For the random vectors* $\vec{U}_n(\vec{X}) = \big(\frac{1}{n}\sum_{i=1}^n f_1(X_i), \dots, \frac{1}{n}\sum_{i=1}^n f_{k_1}(X_i)\,,$ $\frac{1}{n}\sum_{i=1}^n g_1(X_i^*) - \bar{g}_1, \dots, \frac{1}{n}\sum_{i=1}^n g_{k_2}(X_i^*) - \bar{g}_{k_2}\big)$ *the MDP holds, that is, for any* $\Omega \subset R^{k_1+k_2}$

$$\liminf_{n\to\infty}(nb_n^2)^{-1}\log P_n(\vec{U}_n(\vec{X}) \in b_n\Omega) \ge -\inf_{x\in int(\Omega)} x'I_{f,g}x \qquad (4.2)$$

and

$$\limsup_{n\to\infty}(nb_n^2)^{-1}\log P_n(\vec{U}_n(\vec{X}) \in b_n\Omega) \le -\inf_{x\in cl(\Omega)} x'I_{f,g}x \qquad (4.3)$$

where for any $x = (y, z) \in R^{k_1+k_2}$, $y \in R^{k_1}$ *and* $z \in R^{k_2}$

$$x'I_{f,g}x = \sup_{t\in R^{k_1}, s\in R^{k_2}} \left(t'y + s'z - < t'f, H > -\frac{1}{2}t'R_f t - \frac{1}{2}s'R_g s\right).$$

Note that, if there exist R_f^{-1} *and* R_g^{-1}, *then*

$$x'I_{fg}x = \frac{1}{2}((y- < f, H >)'R_f^{-1}(y- < f, H >) + \frac{1}{2}z'R_g^{-1}z.$$

Lemma 4.2 follows from Lemmas 4.3 and 4.4 given below.

Lemma 4.3. *We have*

$$\lim_{n\to\infty}(nb_n^2)^{-1}\log P_n(\max_{1\le i\le k_1}\max_{1\le l\le n}|f_i(X_l)|>b_n^{-1})=-\infty \tag{4.4}$$

and

$$\lim_{n\to\infty}(nb_n^2)^{-1}\log P_n(\max_{1\le i\le k_2}\max_{1\le l\le n}|g_i(X_l^*)|>b_n^{-1})=-\infty. \tag{4.5}$$

Proof. We have

$$P_n(\max_{1\le i\le k_1}\max_{1\le l\le n}|f_i(X_l)|>b_n^{-1})\le n\sum_{i=1}^{k_1}P_n(|f_i(X_1)|>b_n^{-1})\le$$

$$n\sum_{i=1}^{k_1}P(|f_i(X_1)|>b_n^{-1})+nb_n\sum_{i=1}^{k_1}\int\chi(|f_i(X_1)|>b_n^{-1})\,d|H_n|$$

By (2.1) and B1, this implies (4.4).

Since $g_1,\dots,g_{k_2}\in\Phi$, the same statement hold for these functions as well and we get

$$P_n(\max_{1\le i\le k_2}\max_{1\le j\le n}|g_i(X_j)|>b_n^{-1})=O(\exp\{-Cnb_n^2\})$$

for each $C>0$. This implies (4.5).

For any $h\in\Phi$ denote $h_n(x)=h(x)\chi(|h(x)|<b_n^{-1})$. Denote $\vec f_n=\{f_{in}\}_{i=1}^{k_1}$ and $\vec g_n=\{g_{in}\}_{i=1}^{k}$. Define the random vector $\tilde U_n(\vec X)=\big(\frac1n\sum_{i=1}^n f_{1n}(X_i),\dots,\frac1n\sum_{i=1}^n f_{k_1n}(X_i),$ $\frac1n\sum_{i=1}^n g_{1n}(X_i^*)-\bar g_{1n},\dots,\frac1n\sum_{i=1}^n g_{k_2n}(X_i^*)-\bar g_{k_2n}\big)$ where $\bar g_{in}=\frac1n\sum_{l=1}^n g_{in}(X_l),1\le i\le k_2$. Define the events $W_n=\{X_1,\dots,X_n:\max_{1\le i\le k_1}\max_{1\le j\le n}|f_i(X_j)|<b_n^{-1},\max_{1\le i\le k_2}\max_{1\le j\le n}|g_i(X_j)|<b_n^{-1}\}$. Denote $\bar W_n$ the complement of the event W_n.

By Lemma 4.2, we get

$$\begin{aligned}P_n(\vec U_n(\vec X)\in b_n\Omega)&\le P_n(\vec U_n(\vec X)\in b_n\Omega|\bar W_n)P(\bar W_n)+P(W_n)\\&<P_n(\vec U_n(\vec X)\in b_n\Omega|\bar W_n)\exp\{o(nb_n^2)\}+\exp\{-Cnb_n^2(1+o(1))\}\end{aligned} \tag{4.6}$$

and

$$P_n(\vec U_n(\vec X)\in b_n\Omega)\ge P_n(\vec U_n(\vec X)\in b_n\Omega|\bar W_n)P(\bar W_n)>P_n(\vec U_n(\vec X)\in b_n\Omega|\bar W_n)\exp\{o(nb_n^2)\}. \tag{4.7}$$

where the constant C in (4.6) can be chosen arbitrary

Therefore Lemma 4.2 follows from Lemma 4.4 given below.

Lemma 4.4. *For the random vectors $\tilde U_n(\vec X)$ the MDP holds, that is, (4.2) and (4.3) are valid with $\vec U_n(\vec X)=\tilde U_n(\vec X)$.*

By Gartner-Ellis Theorem (see Dembo and Zeitouni [10]) Lemma 4.4 follows from Lemma 4.5 given below.

Lemma 4.5. *Let $f_i\in\Phi, g_j\in\Phi$ for all $1\le i\le k_1, 1\le j\le k_2$. Then*

$$\begin{aligned}&\lim_{n\to\infty}(nb_n^2)^{-1}\log E_n\left[\exp\left\{b_n\sum_{l=1}^n t'\vec f_n(X_l)+b_n\sum_{l=1}^n s'(\vec g_n(X_l^*)-\bar g_n)\right\}\right]=\\&<t'\vec f,H>-\frac12 t'R_ft-\frac12 s'R_gs\end{aligned} \tag{4.8}$$

where $\bar g_n=(\bar g_{1n},\dots,\bar g_{k_2n})$.

Proof. We begin with the proof of upper bound in (4.8). We have

$$\begin{aligned}
I_n = E_n\left[\exp\left\{b_n\sum_{l=1}^n t'\vec{f}_n(X_l) + b_n\sum_{l=1}^n s'(\vec{g}_n(X_l^*)-\bar{g}_n)\right\}\right] = \\
E_n\left[\exp\left\{b_n\sum_{l=1}^n t'\vec{f}_n(X_l)\right\}\prod_{l=1}^n E_{\hat{P}_n}[\exp\{s'(\vec{g}_n(X_l^*)-\bar{g}_n)\}]\right] = \\
E_n\left[\exp\left\{b_n\sum_{l=1}^n t'\vec{f}_n(X_l)\right\}\left(\frac{1}{n}\sum_{l=1}^n \exp\{b_n s'(\vec{g}_n(X_l)-\bar{g}_n)\}\right)^n\right] \le \\
E_n\left[\exp\left\{b_n\sum_{l=1}^n t'\vec{f}_n(X_l)\right\}\left(1+\frac{b_n^2}{2n}\sum_{l=1}^n (s'(\vec{g}_n(X_l)-\bar{g}_n))^2 + \right.\right. \qquad (4.9)\\
\left.\left. +C(s,k_2)\frac{b_n^3}{6n}\sum_{l=1}^n |s'(\vec{g}_n(X_l)-\bar{g}_n)|^3\right)^n\right] \le \\
E_n\left[\exp\left\{b_n\sum_{l=1}^n t'\vec{f}_n(X_l) + \frac{b_n^2}{2}\sum_{l=1}^n (s'(\vec{g}_n(X_l)-\bar{g}_n))^2 + \right.\right.\\
\left.\left. C(s,k_2)b_n^3\sum_{l=1}^n |s'(\vec{g}_n(X_l)-\bar{g}_n)|^3\right\}\right] \doteq I_{1n}.
\end{aligned}$$

The first inequality in (4.9) follows from the Taylor formula and

$$|s'(\vec{g_n}(x)-\bar{g}_n)| \le |s|\ |\vec{g_n}(x)-\bar{g}_n| < |s|2k_2^{1/2}b_n^{-1} \qquad (4.10)$$

Denote $\phi_n(X_l) = s'(\vec{g}_n(X_l) - E_n[\vec{g}_n(X_1)])$ with $1 \le l \le n$.

By straightforward calculations, we get

$$\sum_{l=1}^n (s'(\vec{g}_n(X_l)-\bar{g}_n))^2 = \sum_{l=1}^n \phi_n^2(X_l) - n(s'\bar{g}_n - E_n[s'\vec{g}_n(X_1)])^2. \qquad (4.11)$$

We have

$$\begin{aligned}
\sum_{l=1}^n |s'(\vec{g}_n(X_l)-\bar{g}_n)|^3 \le 8\sum_{l=1}^n |\phi_n(X_l)|^3 + \\
8n|s'(\bar{g}_n - E_n[\vec{g}_n(X_1)])|^3 \doteq 8V_1 + 8nV_2.
\end{aligned} \qquad (4.12)$$

Since

$$\begin{aligned}
|s'(\vec{g}_n(X_1) - Eg_n(X_1))|^3 \le |s|^{3/2}|\vec{g}_n(X_1) - E_n g_n(X_1)|^{3/2} \\
= |s|^{3/2}\left(\sum_{j=1}^{k_2}(g_{jn}(X_1) - E_n[g_{jn}(X_1))^2\right)^{3/2} < 8|s|^3 k_2^{3/2} b_n^{-3}
\end{aligned} \qquad (4.13)$$

we get

$$\begin{aligned}
b_n^3|V_1| = b_n^3\sum_{l=1}^n |\phi_n(X_l)|^3\chi(|\phi_n(X_l)| \le \epsilon b_n^{-1}|s|) + \\
b_n^3\sum_{l=1}^n |\phi_n(X_l)|^3\chi(|\phi_n(X_l)| \ge \epsilon b_n^{-1}|s|) \le \\
\epsilon|s|b_n^2\sum_{l=1}^n \phi_n^2(X_l) + 8|s|^3 k_2^{3/2}\sum_{l=1}^n \chi(|\phi_n(X_l)| \ge \epsilon b_n^{-1}|s|).
\end{aligned} \qquad (4.14)$$

By the Jensen's inequality, we get

$$V_2 = n^{-3}\left|\sum_{l=1}^n \phi_n(X_l)\right|^3 \le n^{-1}\sum_{l=1}^n |\phi_n(X_l)|^3 = n^{-1}V_1. \tag{4.15}$$

By (4.11)-(4.15), we get

$$I_{1n} \le E_n\left[\exp\left\{b_n\sum_{l=1}^n t'\vec{f}_n(X_l) + \frac{b_n^2}{2}(1-2C(s,k_2)\epsilon_n)\sum_{l=1}^n(\phi_n^2(X_l)\right.\right.$$
$$\left.\left.-\frac{b_n^2}{2n}\left(\sum_{l=1}^n \phi_n(X_l)\right)^2 + C(s,k_2)|s|^3\sum_{i=1}^n \chi(|\phi_n(X_l)| \ge \epsilon b_n^{-1}|s|)\right)\right\}\right] \doteq E_n[W_n] \tag{4.16}$$

where $\epsilon = \epsilon_n \to 0$ as $n \to \infty$.

For each $r > 0$ define the events $A_n = A_{nr} \doteq \{X_1,\dots,X_n : s'\bar{g}_n - E_n[s'g_n(X_1)] < rb_n\}$. Denote $\bar{A}_n$ the complement of A_n.

We can write

$$\tilde{I}_n = E_n[W_n\chi(A_n)] + E_n[W_n\chi(\bar{A}_n)] \doteq U_{1n} + U_{2n}. \tag{4.17}$$

Let A_n hold. Then we get

$$\frac{r^2b_n^4}{2n}\left(\sum_{l=1}^n \phi_n(X_l)\right)^2 = \frac{nb_n^2}{2}(s'\bar{g}_n - E_n[s'\vec{g}_n(X)])^2 < \frac{nr^2b_n^4}{2}.$$

Therefore

$$\log[U_{1n}] \le \log E_n\left[\exp\left\{b_n\sum_{l=1}^n t'\vec{f}_n(X_l) + \frac{b_n^2}{2}\sum_{l=1}^n \phi_n^2(X_l)(1+2C(s,k_2)\epsilon)+\right.\right.$$
$$\left.\left.C(s,k_2)|s|^3\sum_{l=1}^n \chi(|\phi_n(X_l)| \ge \epsilon b_n^{-1}) + O(nr^2b_n^4)\right\}\right] =$$
$$n\log E_n\left[\exp\left\{b_nt'\vec{f}_n(X_1) + \frac{b_n^2}{2}\phi_n^2(X_1)(1+2C(s,k_2)\epsilon)+\right.\right.$$
$$\left.\left.C(s,k_2)|s|^3\chi(|\phi_n(X_1)| \ge \epsilon b_n^{-1}) + O(r^2b_n^4)\right\}\right]. \tag{4.18}$$

Expanding in the Taylor series, we get

$$\log U_{1n} \le n\log E_n\left[1 + b_nt'\vec{f}_n(X_1) + \frac{b_n^2}{2}(t'\vec{f}_n(X_1))^2 + \right.$$
$$\left.\frac{b_n^2}{2}\phi_n^2(X_1)(1+2C(s,k_2)\epsilon) + C(s,t,k_1,k_2)\omega_n + O(r^2b_n^4)\right] \tag{4.19}$$

where $\omega_n = \omega_{1n} + \omega_{2n} + \omega_{3n} + \omega_{4n} + \omega_{5n}$ with

$$\omega_{1n} = \frac{b_n^3}{6}|t'\vec{f}_n(X_1)|^3, \quad \omega_{2n} = 3\frac{b_n^3}{2}|t'\vec{f}_n(X_1)|\phi_n^2(X_1),$$
$$\omega_{3n} = \frac{b_n^4}{8}\phi_n^4(X_1), \quad \omega_{4n} = \frac{b_n^4}{12}(t'\vec{f}_n(X_1))^2\phi_n^2(X_1),$$
$$\omega_{5n} = \chi(|\phi_n(X_1)| \ge \epsilon b_n^{-1}).$$

We have

$$\omega_{1n} \le b_n^3|t'\vec{f}_n(X_1)|^3\chi(|t'\vec{f}_n(X_1)| < \epsilon b_n^{-1}) + \chi(\epsilon b_n^{-1} < |t'\vec{f}_n(X_1)| < b_n^{-1}) \doteq \omega_{1n1} + \omega_{1n2},$$
$$\omega_{2n} \le b_n^3|t'\vec{f}_n(X_1)|\phi_n^2(X_1)\chi(|t'\vec{f}_n(X_1)| < \epsilon b_n^{-1}) +$$
$$C(s,t,k_1,k_2)\chi(\epsilon b_n^{-1} < |t'\vec{f}_n(X_1)| < b_n^{-1}) \doteq \omega_{2n1} + \omega_{2n2},$$
$$\omega_{3n} \le b_n^4\phi_n^4(X_1)\chi(\phi_n(X_1) < \epsilon b_n^{-1}) + C\chi(\epsilon b_n^{-1} < \phi_n(X_1) < cb_n^{-1}) \doteq \omega_{3n1} + \omega_{3n2},$$

$$\omega_{4n} \le b_n^4 (t'\vec{f}_n(X_1))^2 \phi_n^2(X_1) \chi(|t'\vec{f}_n(X_1)| < \epsilon b_n^{-1}) + \\ c\chi(\epsilon b_n^{-1} < |t'\vec{f}_n(X_1)| < b_n^{-1}) \doteq \omega_{4n1} + \omega_{4n2}.$$

By (2.1), we get

$$E_n[\omega_{1n1}] \le c\epsilon |t| b_n^2 E_n (t'\vec{f}_n(X_1))^2, \quad E_n[\omega_{2n1}] \le c\epsilon |t| b_n^2 E_n \phi_n^2(X_1),$$
$$E_n[\omega_{3n1}] \le c\epsilon^2 |s|^2 b_n^2 E_n \phi_n^2(X_1), \quad E_n[\omega_{4n1}] \le c\epsilon^2 |t|^2 b_n^2 E_n \phi_n^2(X_1)$$

and

$$E_n[\omega_{5n}] \le \epsilon^{-2} b_n^2 E_n[\phi_n^2(X_1)\chi(|\phi_n(X_i)| \ge \epsilon b_n^{-1})] = o(\epsilon^{-2} b_n^2), \tag{4.20}$$

$$E_n[\chi(\epsilon b_n^{-1} < |t'\vec{f}_n(X_1)| < b_n^{-1})] \le \\ \epsilon^{-2} b_n^2 E_n[|t'\vec{f}_n(X_1)|^2 \chi(\epsilon b_n^{-1} < |t'\vec{f}_n(X_1)|)] = o(\epsilon^{-2} b_n^2) \tag{4.21}$$

where the last equalities in (4.20), (4.21) hold by A and (2.9), (2.10).

Hence we get $E_n[\omega_n] = o(b_n^2)$.

Therefore we get

$$\log(U_{1n}) \le -\frac{nb_n^2}{2}\left(2 < t'\vec{f}, H > -t'R_f t - s'R_g s\right)(1 + O(1)) \doteq v_n. \tag{4.22}$$

By the Hoelder's inequality, we get

$$U_{2n} \le (E_n[W_n^{1+\delta}])^{\frac{1}{1+\delta}} (P(\bar{A}_n))^{\frac{\delta}{1+\delta}}. \tag{4.23}$$

By (4.16), we get

$$E_n[W_n^{1+\delta}] \le E_n\left[\exp\left\{(1+\delta)\left(b_n \sum_{i=1}^n t'\vec{f}_n(X_i) + b_n^2 \sum_{i=1}^n \phi_n^2(X_i)(1 + 2C(s,k_2)\epsilon) + \\ 2C(s,k_2)\sum_{i=1}^n \chi(\phi_n(X_i) > \epsilon b_n^{-1})\right)\right\}\right].$$

Hence, repeating the estimates of U_{1n}, we get

$$E_n[W_n^{1+\delta}] \le \exp\left\{-\frac{(1+\delta)nb_n^2}{2}(2 < t'\vec{f}, H > -t'R_f t - s'R_g s)(1 + O(1))\right\} \tag{4.24}$$

Note that (2.1) implies (2.7) and (2.7) implies

$$\lim_{n\to\infty} (nr^2 b_n^2)^{-1} \log(nP(|f(X)| > rnb_n)) = -\infty$$

for each $r > 1$.

Hence, by Theorem 2.4 in Arcones [2], we get

$$\log P_n(\bar{A}_n) \le -cr^2 nb_n^2 \tag{4.25}$$

By (4.23),(4.24),(4.25) we get that

$$U_{2n} = o(U_{1n}) \tag{4.26}$$

if r sufficiently large. This completes the proof of upper bound for I_n.

The proof of lower bound is based on similar estimates. Define the events

$$B_n = \{x_1, \ldots, x_n : |f_{ni}(x_s)| < \epsilon b_n^{-1}, |g_{nj}(x_s)| < \epsilon b_n^{-1}, 1 \le s \le n, 1 \le i \le k_1, 1 \le j \le k_2\}.$$

By (2.1),(4.20), (4.21), we get

$$P_n(|f_{ni}(X_1)| > \epsilon b_n^{-1}) < \epsilon^{-2} b_n^2 E_n[f_{ni}^2(X_1)\chi(|f_{ni}(X_1)| > \epsilon b_n^{-1})] = o(\epsilon^{-2} b_n^2).$$

Estimating $P_n(g_{ni}(X_1)| > \epsilon b_n^{-1})$ similarly, we get

$$P(B_n) = \prod_{i=1}^{k_1}(1 - P(|f_{ni}(X_1)| > \epsilon b_n^{-1}))^n \prod_{i=1}^{k_2}(1 - P(|g_{ni}(X_1)| > \epsilon b_n^{-1}))^n = \exp\{-o(nb_n^2)\}.$$

Hence

$$I_n \ge E_n \left[\exp\left\{b_n\sum_{i=1}^n t'\vec f_n(X_i)\right\}\left(\frac{1}{n}\sum_{i=1}^n \exp\{b_n s'(\vec g_n(X_i)-\bar g_n)\}\right)^n \chi(B_n)\right] \le$$
$$E_n\left(\exp\left\{b_n\sum_{i=1}^n t'\vec f_n(X_i)\right\}\left(\frac{1}{n}\sum_{i=1}^n \exp\{b_n s'(\vec g_n(X_i)-\bar g_n)\}\right)^n \Bigg| B_n\right] P(B_n) =$$
$$E_n\left[\exp\left\{b_n\sum_{i=1}^n t'\vec f_n(X_i)\right\}\left(\frac{1}{n}\sum_{i=1}^n \exp\{b_n s'(\vec g_n(X_i)-\bar g_n)\}\right)^n \Bigg| B_n\right] \exp\{-o(nb_n^2)\}$$
$$\doteq I_{2n}\exp\{-o(nb_n^2)\}. \tag{4.27}$$

Expanding in the Taylor series, we get

$$I_{2n} \ge E_n\left[\exp\left\{b_n\sum_{i=1}^n t'\vec f_n(X_i)\right\}\left(1+\frac{b_n^2}{2n}\sum_{i=1}^n (s'(\vec g_n(X_i)-\bar g_n))^2 -\right.\right.$$
$$\left.\left. C(s,k_2)\frac{b_n^3}{n}\sum_{i=1}^n |s'(\vec g_n(X_i)-\bar g_n)|^3\right)^n \Bigg| B_n\right] \ge$$
$$E_n\left[\exp\left\{b_n\sum_{i=1}^n t'\vec f_n(X_i)\right\}\left(1+\frac{b_n^2}{2n}(1-2\epsilon)\sum_{i=1}^n (s'(\vec g_n(X_i)-\bar g_n))^2\right)^n \Bigg| B_n\right] \doteq I_{3n} \tag{4.28}$$

where the last inequality follows from

$$\sum_{i=1}^n |s'(\vec g_n(X_i)-\bar g_n)|^3 \le 2\epsilon b_n^{-1}\sum_{i=1}^n (s'(\vec g_n(X_i)-\bar g_n))^2$$

Since $\ln(1+x) \ge 1+x-x^2$ with $x>0$ we get

$$I_{3n} = E_n\left[\exp\left\{b_n\sum_{i=1}^n t'\vec f_n(X_i)\right\}\exp\left\{n\ln\left(1+\frac{b_n^2}{2}(1-2\epsilon)\sum_{i=1}^n (s'(\vec g_n(X_i)-\bar g_n))^2\right)\right\}\Bigg| B_n\right] \ge$$
$$E_n\left[\exp\left\{b_n\sum_{i=1}^n t'\vec f_n(X_i) + \frac{b_n^2}{2}(1-2\epsilon)\sum_{i=1}^n (s'(\vec g_n(X_i)-\bar g_n))^2 -\right.\right.$$
$$\left.\left.\frac{b_n^4}{4n}\left(\sum_{i=1}^n (s'(\vec g_n(X_i)-\bar g_n))^2\right)^2\right\}\Bigg| B_n\right] \ge$$
$$E_n\left[\exp\left\{b_n\sum_{i=1}^n t'\vec f_n(X_i) +\right.\right.$$
$$\left.\left.\frac{b_n^2}{2}(1-2\epsilon-4\epsilon^2)\sum_{i=1}^n (s'(\vec g_n(X_i)-\bar g_n))^2\right\}\Bigg| B_n\right] \doteq I_{4n} \tag{4.29}$$

where the last inequality follows from

$$\frac{b_n^4}{4n}\left(\sum_{i=1}^n (s'(\vec g_n(X_i)-\bar g_n))^2\right)^2 \le \epsilon^2 b_n^2\sum_{i=1}^n (s'(\vec g_n(X_i)-\bar g_n))^2.$$

Arguing similarly to the proof of upper bound we get

$$(nb_n^2)^{-1}\ln I_{4n} = -\frac{nb_n^2}{2}\left(-2 < t'\vec{f}, H > -t'R_f t - (1-2\epsilon-2\epsilon^2)s'R_g s\right)(1+O(1)). \tag{4.30}$$

Since the choice of $\epsilon > 0$ is arbitrary, this completes the proof of lower bound and the proof of Lemma 4.5.

5. Proofs of Theorems 2.5 and 2.6

We begin with the proof of Theorem 2.6. The reasoning is akin to the proof of the Sanov Theorem in de Acosta [8].

Lemma 5.1. . *Let (2.17)-(2.23) hold. Then*
i. $\Gamma_{0r} \subset \Lambda_{0\Theta}$,
ii. the set Γ_{0r} *is* τ_Θ*-compact and sequentially* τ_Θ*-compact set in* $\Lambda_{0\Theta}$,
iii. the τ *and* τ_Θ*- topologies coincide in* Γ_{0r}.

The proof of Lemma 5.1 is akin to the proof of Lemma 4.1 and is omitted. It suffices to note only that (2.17) holds.

We begin with the proof of upper bound in (2.23). Denote $\eta = \rho_0^2(\mathfrak{cl}(\Omega_0), P)$ and fix $\delta, 0 < 2\delta < \eta$. IT is clear that $\Gamma_{0,\eta-\delta} \subset \Lambda_{0\Theta} \setminus \Omega_0$

For any $f_1, \dots, f_l \in \Theta$, $G \in \Lambda_{0,\Theta}$ and $\gamma > 0$ denote

$$U(f_1,\dots,f_l,G,\gamma) = \left\{R : \left|\int f_i d(R-G)\right| < \gamma, R \in \Lambda_{0\Theta}, 1 \le i \le l\right\}.$$

Define the linear space

$$\tilde\Lambda_{0\Theta} = \{G : G = \sum_{i=1}^k \lambda_i G_i, G_i \in \Lambda_{0\Theta}, \lambda_i \in R^1, 1 \le i \le k, k = 1,2,\dots\}.$$

Define τ_Θ- topology in $\tilde\Lambda_{0\Theta}$. It is clear that $\Lambda_{0\Theta} \subset \tilde\Lambda_{0\Theta}$.

Since $\tilde\Lambda_{0\Theta}$ is the Hausdorff linear topological space, the space $\tilde\Lambda_{0\Theta}$ is regular space (see Theorem B2 in [10]). Thus for each $G \in \Gamma_{0,\eta-\delta}$ there exists open set $U(f_1,\dots,f_l,G,\gamma) \subset \Lambda_{0\Theta} \setminus \mathrm{cl}(\Omega_0)$. The set $\Gamma_{0,\eta-\delta}$ is compact. Therefore there exists finite covering of $\Gamma_{0,\eta-\delta}$ by the sets $U_1 = U(f_{11},\dots,f_{1l_1},G_1,c_1),\dots,U_m = U(f_{m1},\dots,f_{ml_m},G_m,c_m)$. Denote $U = \cup_{i=1}^m U_i$.

Thus for the proof of (2.23) it suffices to estimate

$$\hat P_n(P_n^* \notin P + a_n U) \ge (\hat P_n)^*(P_n^* \in P + a_n\Omega).$$

This is the finite dimensional problem.

For each $i, j, 1 \le j \le l_i, 1 \le i \le m$ define the signed measure F_{ij} having the density $\frac{dF_{ij}}{dP} = f_{ij} - E[f_{ij}(X)]$. Define the linear spaces

$$L = \{F : F = \sum_{i=1}^k \sum_{j=1}^{l_i} \lambda_{ij}F_{ij}, \lambda_{ij} \in R^1, 1 \le j \le l_i, 1 \le i \le m\},$$

and

$$\tilde l = \left\{f : f = \frac{dF}{dP}, F \in L\right\}.$$

Define the sets $\hat\Gamma_{0c} = \left\{f : f = \frac{dF}{dP}, F \in \Gamma_{0c} \cap L\right\},\ c > 0$.

There exists a finite number of functions $q_1,\dots,q_l \in \hat\Gamma_{\eta-2\delta}$ such that $E[q_i(X)] = 0, E[q_i^2(X)] = 2(\eta-2\delta), 1 \le i \le l$ and

$$\Gamma_{\eta-2\delta} \cap L \subset \cap_{i=1}^l V(q_i) \cap L \subset \Gamma_{\eta-\delta} \cap L \tag{5.1}$$

with

$$V_i = V(q_i) = \left\{ G : \left| \int q_i dG \right| < 2(\eta - 2\delta), G \in \Lambda_{0\Theta} \right\}.$$

Denote

$$V = \cap_{i=1}^k V_i.$$

Since $\Gamma_{\eta-\delta} \cap L \subset U \cap L$ we get $V \subset U$. Hence

$$\Omega_0 \subset W = \Lambda_{0\Theta} \setminus V.$$

Therefore it suffices to estimate the right hand-side of

$$\log(\hat{P}_n)^*(P^*_{k_n} \in \hat{P}_n + a_n\Omega_0) \le \log \hat{P}_n(P^*_{k_n} \in \hat{P}_n + a_n W).$$

We have

$$\begin{aligned} &\hat{P}_n(P^*_{k_n} \in \hat{P}_n + a_n W) \le \sum_{i=1}^k \hat{P}_n(P^*_{k_n} \notin \hat{P}_n + a_n U_i) = \\ &\sum_{i=1}^k \hat{P}_n\left(\int q_i \, d(P^*_{k_n} - \hat{P}_n) - 2a_n(\eta - 2\delta) > 0 \right). \end{aligned} \tag{5.2}$$

Thus it suffices to show that, for each $f \in \Theta, E[f(X)] = 0, E[f^2(X)] = \eta - 2\delta$ and $n > n_0(\epsilon, f)$,

$$(k_n a_n^2)^{-1} \log \hat{P}_n \left(\int f d(P^*_{k_n} - \hat{P}_n) > 2a_n(\eta - 2\delta) \right) \le -2\frac{(\eta - 2\delta)^2}{Var[f(X_1)]}(1-\epsilon) = (\eta-2\delta)(1-\epsilon) \tag{5.3}$$

with probability $\kappa_n(\epsilon, U(f, q))$.

Denote $s^2 \doteq s_f^2 \doteq s_n^2 = \frac{1}{n}\sum_{i=1}^n f^2(X_i) - \bar{f}^2$ with $\bar{f} = \frac{1}{n}\sum_{i=1}^n f(X_s)$. We put $\gamma = \frac{\sqrt{2}s\epsilon}{324\sigma}$ where $\sigma^2 = Var[f(X_1)] = \eta - 2\delta$.

By Theorem 28 Ch.4 in Petrov [23] we get $P(|s_n^2 - \sigma^2| > \epsilon) < \beta_{2n}(f)$ with $\beta_{2n}(f) = C_1(f, \epsilon)n^{1-t}$. Thus, to prove (5.3), we can suppose that

$$|s_n^2 - \sigma^2| < \epsilon. \tag{5.4}$$

Define the sets of events $A_{nf} = \{X_1, \dots, X_n : \max_{1 \le s \le n} |f(X_s)| < \sigma\gamma a_n^{-1}\}$. We have

$$P(A_{nf}) < 1 - nP(|f(X_1)| > \sigma\gamma a_n^{-1}) = 1 - nh\left(\frac{a_n}{\sigma\gamma}\right) \doteq 1 - \beta_{2n}. \tag{5.5}$$

Note that, by (2.21), $nh\left(\frac{a_n}{\sigma\gamma}\right) \to 0$ as $n \to \infty$. Therefore it suffices to prove (5.3), if A_{nf} hold.

The further reasoning is based on slightly simplified version of Theorem 3.2 in [26]. This version of Theorem is given below.

Let $Y_{1n}, \dots, Y_{k_n,n}$ be i.i.d.r.v'.'s having pm P_n, $E[Y_{1n}] = 0, Var[Y_{1n}] = \sigma^2, |Y_{in}| < \sigma\gamma a_n^{-1}$. Denote

$$S_n = \frac{1}{\sqrt{k_n}\sigma} \sum_{i=1}^{k_n} Y_{in}.$$

Suppose that

$$a_n^{-2} z^{-2} \log E[\exp\{z a_n \sigma^{-1} Y_{1n}\}] < C \quad \text{if} \quad |z| < \kappa. \tag{5.6}$$

and

$$\omega = \frac{\sqrt{2}\kappa}{36 \max\{1, C\}} > 1. \tag{5.7}$$

Denote $\Delta = \omega a_n k_n^{1/2}$.

Theorem 5.1. *Assume (5.6,5.7). Then*

$$P(S_n > k_n^{1/2}a_n) = (1-\Phi(k_n^{1/2}a_n))\exp\{L(k_n^{1/2}a_n)\}\left(1+\theta f_1(k_n^{1/2}a_n)\frac{k_n^{1/2}a_n+1}{\Delta}\right) \tag{5.8}$$

with

$$f_1(k_n^{1/2}a_n) = \frac{60(1+10\Delta^2\exp\{-(1-\omega_n^{-1})\sqrt{\Delta}\})}{1-\omega_n^{-1}}.$$

and

$$-\frac{k_na_n^2}{3\omega} < L(k_n^{1/2}a_n) < \frac{k_na_n^2}{2}\frac{1}{1+\omega}. \tag{5.9}$$

Note that, if $\omega > 16$ and $a_nk_n^{1/2} > 100$ then

$$|\theta_1 f_1(k_n^{1/2}a_n)|\frac{k_n^{1/2}a_n+1}{\Delta} < 6. \tag{5.10}$$

If $|z| < \kappa$ and $|f(X_i)| < \sigma\gamma a_n^{-1}, 1 \le i \le n$, we have

$$\begin{aligned}&\log E_{\hat{P}_n}\{\exp\{za_n(f(X_1^*)-\bar{f})/s\}\} = \\ &\log\left[\frac{1}{n}\sum_{l=1}^n \exp\{za_n(f(X_i)-\bar{f})/s\}\right] = \\ &\log\left(1+\frac{z^2a_n^2}{2}+\frac{\theta^3z^3a_n^3s^{-3}}{6n}\sum_{i=1}^n(f(X_i)-\bar{f})^3\exp\{\theta za_n(f(X_i)-\bar{f})/s\}\right) \doteq \tau_n\end{aligned} \tag{5.11}$$

with $0 < \theta < 1$.

Since

$$\exp\{\theta za_n(f(X_1)-\bar{f})/s\} < \exp\{2\gamma\kappa\theta\sigma s^{-1}\} \doteq R,$$

using $\ln(1+x) < x, x > 0$, we get

$$\tau_n < \log\left(1+\frac{z^2a_n^2}{2}(1+\gamma\kappa\sigma Rs^{-1})\right) < \frac{z^2a_n^2}{2}(1+\gamma\kappa\sigma Rs^{-1}) = z^2a_n^2D \tag{5.12}$$

with $D = \frac{1+\gamma\kappa R\sigma s^{-1}}{2}$.

If

$$\kappa = \frac{s}{2\gamma\sigma}, \tag{5.13}$$

then $R < 3$ and $D < 2$. Therefore

$$\omega > \frac{9}{2\epsilon}, \quad L(k_n^{1/2}a_n) \le \frac{k_n^{1/2}a_n^2}{2}\frac{\epsilon}{9/2+\epsilon}.$$

Hence, by (5.8),(5.10), we get

$$\begin{aligned}&(k_na_n^2)^{-1}\log\hat{P}_n\left(\int fd(P_{k_n}^*-\hat{P}_n) > 2a_n(\eta-2\delta)\right) \le \\ &-\frac{1}{2}s^{-2}(\eta-2\delta)^2\left(1-\frac{\epsilon}{9/2+\epsilon}\right)+(\log 7-\frac{1}{2}\log(2\pi s^{-2}(1+\epsilon)))(k_na_n^2)^{-1} \le \\ &-\frac{1}{2}s^{-2}(\eta-2\delta)^2\left(1-\frac{\epsilon}{2}\right)+C(k_na_n^2)^{-1} \\ &= -\frac{1}{2}s^{-2}(\eta-2\delta)^2(1-\frac{\epsilon}{2})+C(k_na_n^2)^{-1} \\ &\le -\frac{1}{2}s^{-2}(\eta-2\delta)^2\left(1-\frac{\epsilon}{2}\right)+C(k_na_n^2)^{-1}.\end{aligned} \tag{5.14}$$

This implies (5.3) if (5.4) and $|f(X_i)| < \sigma\gamma a_n^{-1}, 1 \le i \le n$ hold. This completes the proof of (2.23).

If $\rho_0^2(\mathfrak{cl}(\Omega_0), P) = \infty$, we put $\eta = L$. After that we implement the same reasoning for the proof of (2.24).

The proof of lower bound (2.22) is based on standard reasoning (see Sanov [24], Dembo and Zeitouni [10], de Acosta [8] and references therein) and estimates of Theorem 5.1. For any $\delta > 0$ there exists open set $U = U(f_1, \dots, f_l, G, \gamma_1, \dots, \gamma_l)$ such that $U \subset \mathrm{int}(\Omega_0)$ and $\rho_0^2(U, P) < \eta + \delta$, $\rho_0^2(G, P) < \eta + \delta$. Hence it suffices to find the lower bound for the asymptotic

$$(k_n a_n^2)^{-1} \log \hat{P}_n(P_k^* \in \hat{P}_n + a_n U).$$

Similarly to the proof of upper bound we can suppose that the signed measure G has the density $g = \frac{dG}{dP} = \sum_{i=1}^l \lambda_i f_i, f_i \in \Theta$. Thus the problem became the finite dimensional problem as well.

Let us fix $\lambda, 0 < \lambda < 1$ such that $\lambda G \in U$. Note that value of λ can be chosen arbitrary from some vicinity of 1. Define the set $U_1 = U \cap U(g, G, 2(1-\lambda)^2||g||^2)$. It is clear that we can choose λ so that $\rho_0^2(U_1 : P) \le \lambda^2||g||^2$

Lemma 5.2. *There exist simplex* $\tilde{U} \subset U_1$ *bounded the hyperplane* $\Pi = \{R : \int g\, dR = 2\lambda^2||g||^2, R \in \Lambda_{0\Theta}\}$ *and hyperplanes* $\Pi_i = \{R : \int g_i dR = c_i, R \in \Lambda_{0\Theta}\}$ *with* $g_i \in \Theta, 1 \le i \le l$ *such that* $\rho_0^2(\Pi_i : P) \ge \lambda^2||g||^2 = \rho_0^2(\Pi : P)$.

Let Lemma 5.2 be valid. Suppose A_{bf} holds with $f = g$ and $f = g_i, 1 \le i \le l$. Then, applying Theorem 5.1, we get

$$\begin{aligned}
&\hat{P}_n(P_{k_n}^* \in \hat{P}_n + a_n U_1) \ge \hat{P}_n(P_{k_n}^* \in \hat{P}_n + a_n \tilde{U}) \ge \\
&\hat{P}_n\left(\int g d(P_{k_n}^* - \hat{P}_n) > 2\lambda^2||g||^2 a_n\right) - \\
&\sum_{i=1}^l \hat{P}_n\left(\int g_i (dP_{k_n}^* - \hat{P}_n) > a_n c_i\right) \ge \\
&\hat{P}_n\left(\int g d(P_{k_n}^* - \hat{P}_n) > 2\lambda^2||g||^2 a_n\right) - \sum_{i=1}^l \exp\{-\rho_0^2(\Pi_i : P) a_n^2 k_n (1 + \epsilon_n)\}.
\end{aligned} \tag{5.15}$$

with $\epsilon_n \to 0$ as $n \to \infty$.

Thus it remains to implement Theorem 5.1 to the first addendum in the right-hand side of (5.15).

By (5.8) and (5.9), we get

$$\begin{aligned}
&(a_n^2 k_n)^{-1} \log \hat{P}_n\left(\int g\, dP_{k_n}^* - \hat{P}_n) > a_n||g||\right) \ge \\
&-\frac{1}{2}||g||^2\left(1 + \frac{1}{3\omega}\right) + c(k_n a_n^2)^{-1} = -\frac{1}{2}||g||^2\left(1 + \frac{s}{9\sigma}\epsilon\right) + c(k_n a_n^2)^{-1}.
\end{aligned} \tag{5.16}$$

This completes the proof of lower bound.

Proof of Lemma 5.2. The problem is reduced to the following. There is given a parallelepiped U_1 in R^{l+1} and $0 \notin U_1$, $\rho(0, U_1) = \inf_{x \in U_1} |x|$. The point u lies on the face Π of parallelepiped U_1 and $\rho(0, u) = \rho(0, U_1)$. One needs to point out simplex $V \subset U_1$ such that $\Pi \cap V$ is the face of V, $u \in \Pi \cap V$ and for any hyperplane Π_1 passing through another face of V it holds $\rho(0, \Pi_1) > \rho(0, u)$. Let the distance of u from any face other than Π exceeds r_0. A simple trigonometric reasoning shows that the simplex V can be defined as follows. We take the vertex $v = (1 + \frac{1}{2}r^2)u$ of V where $r << r_0$ and all other vertices $v_i, 1 \le i \le l$ belong Π and $|v_i - u| = r$.

For the proof of this statement it suffices to consider the case $l = 1$. Let us draw through v the line L intersecting the line Π at the point w and such that w

is orthogonal to L. Then $|u-v| = |w-u|^2|u|^{-1}(1+o(1))$. Therefore, if the line $L_1, v \in L_1$ intersect Π at the point $z = c|w-u|^2|u|^{-1}, c < 1$ then $\rho(0, L_1) > |u|$. This completes the proof of Lemma 5.2.

Proof of Theorem 2.5. The reasoning is based on estimates of Theorem 2.6. We begin with the proof of upper bound (2.20) in the case of $\tau_{\Theta_{2h}}$-topology. Suppose that $\rho_0^2(\mathfrak{cl}(\Omega_0), P) < \infty$. If $\rho_0^2(\mathfrak{cl}(\Omega_0), P) = \infty$, the reasoning are similar. It suffices to prove that for any $\epsilon > 0$

$$(k_n a_n^2)^{-1} \log(\hat{P}_n)^*(P^*_{k_n} \in \hat{P}_n + a_n\Omega_0) \le -\rho_0^2(\mathfrak{cl}(\Omega_0), P) + \epsilon \quad a.\,s. \tag{5.17}$$

By the Strong Law of Large Numbers and (2.17), for any $f \in \Theta$, we get

$$s_n^2(f) \to \sigma^2(f) \quad a.\,s. \tag{5.18}$$

with $\sigma^2(f) < \infty$.

By (2.18) and (2.16), for any $\delta > 0$, we get

$$\begin{aligned} P(\max_{i\ge l} a_i|f(X_i)| \le \delta) &= \prod_{i=l}^{\infty}(1 - P(|f(X_i)| > \delta a_s^{-1})) \\ &\ge \prod_{i=l}^{\infty}(1 - h(a_i/\delta)) \ge \exp\{\sum_{i=l}^{\infty} h(a_i/\delta)\} = 1 + o(1). \end{aligned} \tag{5.19}$$

as $l \to \infty$.

For any k,

$$P(\max_{1\le i\le k} a_n|f(X_i)| > \delta) = o(1) \tag{5.20}$$

as $n \to \infty$.

Note that $\max_{i\ge k} a_i|f(X_i)| < \delta$ implies $\max_{k\le i\le n}|f(X_i)| < \delta a_n^{-1}$. Therefore, by (5.19) and (5.20), we get

$$\max_{1\le s\le n} |f(X_s)| < \delta a_n^{-1} \quad a.\,s. \tag{5.21}$$

Using (5.18),(5.21), we can implement the same technique for the proof of (5.3) as in the proof of (2.22) in Theorem 2.6. This completes the proof of (2.20

For the proof of (2.20) in the case of τ_{Θ_t}-topology it suffices to show that, for any $\delta > 0$

$$I_k \doteq P(\max_{i>k} a_i|f(X_i)| > \delta) = o(1) \tag{5.22}$$

as $k \to \infty$.

We have

$$\begin{aligned} I_k &\le \sum_{i=k}^{\infty} P(f(X_i) > \delta a_i^{-1}) \\ &= \sum_{i=k+1}^{\infty} (n-k)P(\delta a_{i-1}^{-1} < |f(X_s)| \le \delta a_i^{-1}) \doteq J_k. \end{aligned} \tag{5.23}$$

Define the function $u(x) = \delta a_{i-1}^{-1} + \delta(a_i^{-1} - a_{i-1}^{-1})(x - i + 1)$ if $x \in [a_{i-1}^{-1}, a_i^{-1})$. Define the inverse function $v(y) = \inf\{t : u(t) = y, t \in R^1\}$. Define the distribution function $F(x) = P(|f(X)| < x), x \in R^1_+$.

Then

$$J_k \le 2\int_{a_k^{-1}}^{\infty} v(x)dF(x) \le 2\int_{a_k^{-1}}^{\infty} x^t dF(x) = o(1) \tag{5.24}$$

as $k \to \infty$. This implies (5.22).

The proof of lower bound (2.19) is based on similar reasoning and is omitted.

6. Appendix

Proof of Theorem 2.4. One needs to show that

$$-\log P\left(\sum_{i=1}^n Y_i^* > ne_n\right) = o(ne_n^2) \tag{6.1}$$

Define the events $A_{ni} = U_{ni} \cup V_{ni}, 1 \le i \le n$ with $U_{ni} = \{Y_i : |Y_i| < b_n^{-1}\}$ and $V_{ni} = \{Y_i : r_n < Y_i\}$. Denote $A_n = \cap_{i=1}^n A_{ni}$.

By (2.2), we get

$$\begin{aligned} &P(A_n) > 1 - P(\max_{1\le i\le n} |Y_i| > b_n^{-1}) > \\ &1 - nP(|Y_1| > b_n^{-1}) = 1 + o(1). \end{aligned} \tag{6.2}$$

Denote P_{cn} the conditional probability measure Y_1 given $Y_1 \in A_{n1}$.

By (6.2), we get

$$\begin{aligned} &P\left(\sum_{i=1}^n Y_i^* > ne_n\right) \ge P\left(\sum_{i=1}^n Y_i^* > ne_n|A_n\right) P(A_n) = \\ &P_{cn}\left(\sum_{i=1}^n Y_i^* > ne_n\right)(1+o(1)). \end{aligned} \tag{6.3}$$

Thus it suffices to prove (6.1) with pm P replaced by pm P_{cn}. Denote $p_n = P_{cn}(Y_1 > r_n)$. By (2.2), we get $np_n \to 0$ as $n \to \infty$. Define the events $W_n(k_n) = \{\, Y_1, \dots, Y_n : n - k_n$ random variables $Y_1, \dots, Y_n$ belong $(0, b_n^{-1})$ and k_n random variables $Y_1, \dots, Y_n$ belong $(r_n, \infty)\,\}$. Suppose that $k = k_n \to \infty$ as $n \to \infty$ and

$$\lim_{n\to\infty} k_n n p_n = 0, \quad \lim_{n\to\infty} (r_n e_n)^{-1} \log \frac{ne_n}{r_n k_n} = 0. \tag{6.4}$$

By the Stirling formula, we get

$$\begin{aligned} &v_n \doteq P_{cn}(W_n(k)) = \frac{n!}{(n-k)!k!} p_n^k (1-p_n)^{n-k} = \\ &(2\pi)^{-1/2} \exp\{(n+1/2)\log n - (n-k+1/2)\log(n-k) - (k+1/2)\log k + \\ &k \log p_n + (n-k)\log(1-p_n)\}(1+o(1)) = \\ &\exp\left\{-(n-k+1/2)\log\frac{n-k}{n(1-p_n)} - k\log\frac{k}{np_n}(1+o(1))\right\} = \\ &\exp\{-n(1-k/n)(-k/n+p_n)(1+o(1)) - k\log[k/(np_n)](1+o(1))\} = \\ &\exp\{(k - np_n - k\log(k/(np_n))(1+o(1))\} = \\ &\exp\left\{-k\log\frac{k}{np_n}(1+o(1))\right\}. \end{aligned} \tag{6.5}$$

It follows from (2.13),(6.5) that we can choose $k = k_n$, such that

$$|\log v_n| = O(k_n|\log(np_n)|) = o(ne_n^2). \tag{6.6}$$

Define the random variable l_n which equals the number of $Y_i^*, 1 \le i \le n$ such that $Y_i^* \in (r_n, \infty)$. Denote $u_n = c\frac{ne_n}{r_n} = c\frac{ne_n^2}{r_n e_n}$ with $c > 1$ and denote $m_n = [u_n]$. Suppose that $\frac{u_n}{k_n} \to \infty$ as $n \to \infty$. Then estimating similarly to (6.5) we get

$$P_c(l_n > u_n|W_n(k_n)) = \exp\left\{-u_n \log\frac{u_n}{k_n}(1+o(1))\right\} \tag{6.7}$$

Denote $c_1 = c - 1$. Denote $Y^{1*} \le \ldots \le Y^{n*}$ the order statistics of $Y_1^*, \ldots, Y_n^*$.

The event $\{Y_1^*,\dots,Y_n^* : \sum_{i=1}^n Y_i^* > ne_n\}$ contains the event

$$U_n = \left\{ Y_1^*,\dots,Y_n^* : \sum_{j=1}^{n-m_n} Y^{j*} > -c_1 ne_n, |Y^{j*}| < b_n^{-1}, \right.$$
$$\left. 1 \le j \le n-m_n, Y^{t*} > r_n, n-m_n < t \le n \right\}$$

since, if U_n holds,

$$\sum_{t=n-m_n-1}^{n} Y^{t*} > r_n m_n = c r_n \frac{ne_n}{r_n} = cne_n.$$

Hence it suffices to show that

$$\log P_c(U_n) = o(ne_n^2). \tag{6.8}$$

We have

$$P_c(U_n) \ge P_c(l_n = m_n) P_c\left(\sum_{i=1}^{n-m_n} Y_i^* > -c_1 ne_n, |Y_i^*| < b_n^{-1}, 1 \le i \le n-m_n \right) \ge$$
$$P_c(l_n = m_n | W_n(k_n)) P_c(W_n(k_n)) P_c\left(\sum_{i=1}^{n-m_n} Y_i^* > -c_1 ne_n, |Y_i^*| < b_n^{-1}, 1 \le i \le n-m_n \right). \tag{6.9}$$

Denote $q_n = P_c(|Y_1| < b_n^{-1})$. Define the conditional probability measure P_{b_n} of the random variable Y_1 given $|Y_1| < b_n^{-1}$.

We have

$$\begin{aligned} P_c(|Y_1^*| < b_n^{-1}) &= \sum_{i=1}^n \frac{n!}{(n-i)!i!} q_n^i (1-q_n)^{n-i} \frac{i}{n} = \\ &= q_n \sum_{i=1}^n \frac{(n-1)!}{(n-i)!(i-1)!} q_n^{i-1}(1-q_n)^{n-i} = q_n \end{aligned} \tag{6.10}$$

We have

$$\begin{aligned} &P_c\left(\sum_{i=1}^{n-m_n} Y_i^* > -c_1 ne_n \,|\, |Y_i^*| < b_n^{-1}, 1 \le i \le n-m_n \right) = \\ &1 - P_c\left(\sum_{i=1}^{n-m_n} Y_i^* < -c_1 ne_n \,|\, |Y_i^*| < b_n^{-1}, 1 \le i \le n-m_n \right). \end{aligned} \tag{6.11}$$

By Chebyshev inequality, using (6.10), we get

$$P_c\left(\sum_{i=1}^{n-m_n} Y_i^* < -c_1 ne_n \,|\, |Y_i^*| < b_n^{-1}, 1 \le i \le n-m_n \right) \le$$
$$\frac{n-m_n}{c_1^2 (n-m_n)^2 e_n^2} E_c[\mathrm{Var}_{\hat P_n}(Y_1^* \,|\, |Y_1^*| < b_n^{-1})] =$$
$$\frac{q_n^2}{c_1^2(n-m_n)e_n^2} \sum_{t=0}^n C_n^t q_n^t (1-q_n)^{n-t} E_{b_n}\left[(n-t)^{-1} \sum_{i=1}^{n-t} \left(Y_i - (n-t)^{-1} \sum_{j=1}^{n-t} Y_j \right)^2 \right] =$$
$$\frac{q_n^2}{c_1^2(n-m_n)e_n^2} \sum_{t=0}^n C_n^t q_n^t (1-q_n)^{n-t} \frac{t-1}{t} \mathrm{Var}_{b_n}[Y] \le \frac{q_n^2}{c_1^2(n-m_n)e_n^2} \mathrm{Var}_{b_n}[Y]. \tag{6.12}$$

and

$$\lim_{n\to\infty} q_n^2 \mathrm{Var}_{b_n}[Y] = \mathrm{Var}[Y]. \tag{6.13}$$

By (6.5,6.7), we get

$$P_c(l_n = m_n|W_n(k_n))P_c(W_n(k_n)) = \exp\left\{-\frac{cne_n^2}{r_ne_n}\log\frac{ne_n}{r_nk_n} - ck_n\log\frac{k_n}{np_n}(1+o(1))\right\} = \exp\{-o(ne_n^2)\} \tag{6.14}$$

where the last equality follows from (6.4,6.6). Now (6.8) follows from (6.11-6.14). This completes the proof of Theorem 2.4.

References

[1] Aleskeviciene, A. K. (1991). Large and moderate deviations for L-statistics. Lithuanian Math. J. 31 145156.

[2] M.A. Arcones. Moderate deviations of empirical processes. In *Stochastic Inequalities and Applications,* 189-212 ed.E.Gine, C.Houdre and D.Nualart. Birkhauser Boston, 2003.

[3] Arcones, M. A. (2006). Large deviations for M-estimators. Ann. Inst. Statist. Math. 58 2152. MR2281205

[4] E.Bolthausen On the probability of large deviations in Banach spaces. Ann.Probab 12 (1984) 427-435

[5] A.A. Borovkov and A.A. Mogulskii. On probabilities of large deviations in topological spaces. II. *Siberian Mathematical Journal,* 21(5):12-26, 1980.

[6] N.R. Chaganty and R.L. Karandikar. Some properties of the Kullback-Leibler number. *Sankhya,A.*, 58:69-80, 1996.

[7] N.R. Chaganty. Large deviations for joint distributions and statistical applications. *Sankhya,A.*, 59:147-166, 1997.

[8] A.de Acosta. On large deviations of empirical measures in the τ-topology. *J. Appl.Probab.* 31A: 41-47, 1994.

[9] R. Dasgupta: Bootstrap of deviation probabilities with applications. J. Multivariate Analysis 101(9): 2137-2148 (2010)

[10] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* Jones and Bartlett, Boston, 1993.

[11] M.D. Donsker and S.R.S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time III, Comm. Pure Appl. Math. 29 (1976), 389-461. MR 55:1492

[12] B. Efron. Bootstrap methods: another look at the jackknife. Ann.Stat. 7 1979 1-26

[13] P.Eichelsbacher and U. Schmock. Large deviations of U- empirical measures in strong topologies and applications. *Annals de L'Institute Henri Poincare. Probability and Statistics,* 38:779-797, 2002.

[14] P. Eichelsbacher and M. Löwe. Moderate deviations for i.i.d. random variables. *ESAIM: Probability and Statistics,* 7:207-216,2003.

[15] M.S. Ermakov.Importance sampling for simulation of moderate deviation probabilities of statistics. *Statistics & Decision,* 25:265-284, 2007.

[16] F.Gao and X.Zhao. Delta method in large deviations and moderate deviations for estimators. The Annals of Statistics 2011, Vol. 39, No. 2, 12111240

[17] P. Groeneboom, J. Oosterhoff and F.H. Ruymgaart. Large deviation theorems for empirical probability measures. *Annals of Probability,* 7: 553-586, 1979.

[18] Hall, P. (1990) On the relative performance of bootstrap and Edgeworth approximations of a distribution function. J. Multivariate Anal., 35, 108-129

[19] T. Inglot, W.C.M. Kallenberg and T. Ledwina. Strong moderate deviation theorems. *Annals of Probability,* 20: 987-1003, 1992.

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

[21] C.Leonard, J.Najim. An extension of Sanov's Theorem. Application to the Gibbs conditioning principle. *Bernoulli* 8:721-743, 2002.

[22] D.Li, A.Rosalski and D.K.Al-Mutairi. A large deviation principle for bootstrapped sample means. Proc.Amer.Math.Soc. (2001) 130 2133-2138

[23] V.V.Petrov Sums of independent random variables. Springer Verlag, New York, 1975.

[24] I. Sanov. On the probability of large deviations of random variables. *Mathematicheskii Sbornik*, 42: 70-95 1957(In Russian) English translation:*Selected Translations in Mathematical Statistics and Probability*, 1: 213-244, 1961.

[25] R.J. Serfling. *Approximation theorems of mathematical statistics.* J.Wiley New York, 1980.

[26] L. Saulis and V.Statulevichius. Limit Theorems for Large Deviations. MOkslas Publishers. Vilnius, 1989.

[27] van der Vaart, A. W. and Wellner, J. A. (1996). Weak Convergence and Empirical Processes with Applications to Statistics. Springer, New York. MR1385671

[28] A.Wood (2000) Bootstrap relative errors and subexponential destributions. Bernoulli 6 809-834

Ermakov Mikhail
Mechanical Engineering Problems Institute
Russian Academy of Sciences
Bolshoy pr.,V.O., 61

199178 St.Petersburg
Russia
e-mail: erm2512@mail.ru