LARGE DEVIATIONS OF PRODUCTS OF EMPIRICAL MEASURES AND U-EMPIRICAL MEASURES IN STRONG TOPOLOGIES

PETER EICHELSBACHER AND UWE SCHMOCK*

Abstract. We prove a large deviation principle (LDP) for m-fold products of empirical measures and for U-empirical measures, where the state space $(S, \mathcal{S})$ of the underlying i. i. d. sequence of random variables is an arbitrary measurable space. For this LDP we choose a suitable subset $\mathcal{M}_\Phi(S^m)$ of all probability measures on $(S^m, \mathcal{S}^{\otimes m})$ and endow it with a topology, which is stronger than the τ-topology and makes the map $\nu \mapsto \int_{S^m} f\,d\nu$ continuous even for certain unbounded f. An improved form of a LDP for U- and V-statistics is obtained as a particular application. Furthermore, we improve the Gibbs conditioning principle for interacting ensembles of particles.

1. Introduction, statement of results, and applications

Let $\{X_i\}_{i\in\mathbb{N}}$ be a sequence of independent, identically distributed random variables taking values in a measurable space $(S, \mathcal{S})$. For simplicity we assume that they are defined on the product space $(\Omega, \mathcal{A}, \mathbb{P}) = (S^{\mathbb{N}}, \mathcal{S}^{\otimes\mathbb{N}}, \mu^{\otimes\mathbb{N}})$, where μ is the distribution of X_1. For every $n \in \mathbb{N}$ the empirical measure is defined by $L_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$. The large deviations of $\{L_n\}_{n\in\mathbb{N}}$ have been studied in several papers. Sanov [21] considered the problem when $S = \mathbb{R}$ and the space of probability measures $\mathcal{M}_1(S)$ is endowed with the weak topology. Donsker and Varadhan [11] and Bahadur and Zabell [3] (see also Azencott [2]) proved – with different approaches – a large deviation principle for $\{L_n\}_{n\in\mathbb{N}}$ when S is a Polish space with Borel σ-algebra $\mathcal{S}$, again in the weak topology of $\mathcal{M}_1(S)$. Groeneboom, Oosterhoff and Ruymgaart [17] obtained a stronger result in which S is a Hausdorff topological space with Borel σ-algebra $\mathcal{S}$ and $\mathcal{M}_1(S)$ is endowed with the τ-topology, that is, the coarsest topology which makes the maps $\mathcal{M}_1(S) \ni \nu \mapsto \int_S f\,d\nu$ continuous for all f in the space $B(S)$ of bounded, real-valued, $\mathcal{S}$-measurable functions on S. De Acosta [1] proved the large deviation principle for $\{L_n\}_{n\in\mathbb{N}}$ in the τ-topology setting when $(S, \mathcal{S})$ in an arbitrary measurable space.

According to the strong law of large numbers,

$$\mathbb{P}\biggl(\lim_{n\to\infty}\int_S f\,dL_n = \int_S f\,d\mu\biggr) = 1, \qquad f \in L_1(S, \mathcal{S}, \mu), \tag{1.1}$$

hence it should be possible to extend de Acosta's results to an even stronger topology. For topologies generated by a collection Φ of $\mathcal{S}$-measurable functions $f\colon S \to \mathbb{R}$

1991 *Mathematics Subject Classification.* 60F10 (primary); 62H10 and 28A35 (secondary).

Key words and phrases. Large deviations, empirical measures, Sanov's theorem, strong topology, U-statistics, von Mises statistics, Gibbs conditioning principle.

*Research partially supported by the Swiss National Foundation, Contract No. 21-298333.90.

with all exponential moments, i. e. $\int_S \exp(\alpha|f|)\,d\mu < \infty$ for all $\alpha > 0$, such an extension of the large deviation principle for $\{L_n\}_{n\in\mathbb{N}}$ is included in our results, see Theorem 1.8 with $m = 1$. For a Polish space S with Borel σ-algebra $\mathcal{S}$, a similar extension of such a large deviation principle was obtained by Wu [26, Section 2.4.4 and Section 4.1]. The collection Φ generating his topology consists of all measurable $f\colon S \to \mathbb{R}$ which are dominated in absolute value by some multiple of a *single* measurable reference function $\psi\colon S \to [1,\infty]$ with $\int_S \exp(\alpha\psi)\,d\mu < \infty$ for all $\alpha > 0$.

1.1. **Motivating examples.** Given an integer $m \geq 2$, we can consider the set $\mathcal{M}_1(S^m)$ of probability measures on the product space S^m, equipped with the product σ-algebra $\mathcal{S}^{\otimes m}$. Naturally, the τ-topology on $\mathcal{M}_1(S^m)$ is defined to be the one induced by $B(S^m)$, the set of all bounded $\mathcal{S}^{\otimes m}$-measurable functions on S^m. A special feature of the τ-topologies on $\mathcal{M}_1(S)$ and $\mathcal{M}_1(S^m)$ is the possible discontinuity of the map $\mathcal{M}_1(S) \ni \nu \mapsto \nu^{\otimes m} \in \mathcal{M}_1(S^m)$. An example suggested by Y. Peres (see [7, Exercise 7.3.18]) shows that this is not an artifact of certain exotic measurable spaces. Indeed, his example shows that for the compact unit interval $S = [0,1]$ equipped with the Borel σ-algebra the product-measure operation $\mathcal{M}_1([0,1]) \ni \nu \mapsto \nu \otimes \nu$ is discontinuous at the Lebesgue measure. The following example may serve as a further illustration that products of "empirical measures" can exhibit a strange behaviour with respect to the τ-topology.

Example 1.2. Let the circle $S = \mathbb{R}/\mathbb{Z}$ be equipped with the Borel σ-algebra $\mathcal{S}$ and let μ denote the Lebesgue–Borel measure on $(S,\mathcal{S})$. For every $x \in \mathbb{R}$ define the shift modulo 1 (or rotation) θ_x on S by $\theta_x(y) = x + y \bmod 1$ for all $y \in S$. Using these, define

$$S \ni \omega \mapsto L_n(\omega) = \frac{1}{2^n}\sum_{i=0}^{2^n-1} \delta_{\theta_{i2^{-n}}(\omega)} \in \mathcal{M}_1(S), \quad n \in \mathbb{N}_0.$$

Since S is compact, it is easy to verify that $\{L_n(\omega)\}_{n\in\mathbb{N}_0}$ and $\{L_n(\omega)\otimes L_n(\omega)\}_{n\in\mathbb{N}_0}$ converge weakly to μ and $\mu\otimes\mu$, respectively, for every $\omega \in S$.

Next we want to show (1.1), which is the strong law of large numbers. Define the sub-σ-algebra $\mathcal{S}_n$ of $\mathcal{S}$ by $\mathcal{S}_n = \{\, A \in \mathcal{S} \mid A = \theta_{2^{-n}}(A) \,\}$ for every $n \in \mathbb{N}_0$. Note that $\mathcal{S}_{n+1} \subset \mathcal{S}_n$ and $E_\mu[f\,|\,\mathcal{S}_n] = \int_S f\,dL_n$ for every $f \in L_1(S,\mathcal{S},\mu)$ and $n \in \mathbb{N}_0$. Therefore, $\{\int_S f\,dL_n\}_{n\in\mathbb{N}_0}$ is a reversed martingale relative to $\{\mathcal{S}_n\}_{n\in\mathbb{N}_0}$ and it converges μ-almost surely to $g \equiv E_\mu[f\,|\,\mathcal{S}_\infty]$ with $\mathcal{S}_\infty \equiv \bigcap_{n\in\mathbb{N}_0}\mathcal{S}_n$, see for example [20, Corollary V-3-12]. Consider the Fourier coefficients $\hat{g}(n) \equiv \int_S g e_n\,d\mu$ for $n \in \mathbb{Z}$, where $e_n(t) \equiv \exp(-2\pi i n t)$. Given $n \in \mathbb{Z}\setminus\{0\}$, there exist $k \in \mathbb{N}_0$ and an odd $l \in \mathbb{Z}$ such that $n = 2^k l$. Since $g \circ \theta_{2^{-(k+1)}} = g$ and $e_{2^k l} \circ \theta_{2^{-(k+1)}} = -e_{2^k l}$, all coefficients $\hat{g}(n)$ with $n \in \mathbb{Z}\setminus\{0\}$ vanish and, therefore, $g = \hat{g}(0) = \int_S f\,d\mu$ μ-almost surely [12, Chap. 1.5, Theorem 1].

To show that the product measures $\{L_n \otimes L_n\}_{n\in\mathbb{N}_0}$ can go astray, consider the $\mathcal{S}\otimes\mathcal{S}$-measurable set $A \equiv \{\,(x,y) \in S^2 \mid x - y \in \mathbb{Q}\,\}$. By Fubini's theorem, $(\mu\otimes\mu)(A) = 0$. On the other hand, the support of $L_n(\omega)\otimes L_n(\omega)$, which is $\{\,(\theta_{i2^{-n}}(\omega),\theta_{j2^{-n}}(\omega)) \mid i,j \in \{0,1,\dots,2^n-1\}\,\}$, is contained in A for every $n \in \mathbb{N}_0$ and $\omega \in S$. Therefore, the strong law of large numbers does not hold for the $\mathcal{S}\otimes\mathcal{S}$-measurable indicator function 1_A.

The preceding example makes use of the heavy dependence of $\theta_{i2^{-n}}(\omega)$ for $i \in \{0,\dots,2^n-1\}$. We want to show that in the opposite case, where $L_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$

is given in terms of an i. i. d. sequence $\{X_i\}_{i\in\mathbb{N}}$ taking values in an arbitrary measurable space $(S,\mathcal{S})$ as introduced at the beginning of this section, the products $L_n^{\otimes m}$ are typically close to $\mu^{\otimes m}$. Furthermore – this is our main objective – we want to study the large deviations of $L_n^{\otimes m}$ in topologies which may be stronger than the usual τ-topology on $\mathcal{M}_1(S^m)$. Note that, due to the possible discontinuity of the map $\mathcal{M}_1(S)\ni\nu\mapsto\nu^{\otimes m}$, such results do not follow easily via the contraction principle in our general setting; see Remark 1.23, however.

1.2. **Statement of the large deviation principles.** To formulate our results, we need some additional notations. Fix $m\in\mathbb{N}$. Let Φ be a set of $\mathcal{S}^{\otimes m}$-measurable functions $\varphi\colon S^m\to\mathbb{R}$ containing $B(S^m)$. Define

$$\mathcal{M}_\Phi(S^m)=\left\{\nu\in\mathcal{M}_1(S^m)\;\middle|\;\int_{S^m}|\varphi|\,d\nu<\infty\text{ for every }\varphi\in\Phi\right\}.$$

Let τ_Φ denote the coarsest topology on $\mathcal{M}_\Phi(S^m)$ such that the map $\mathcal{M}_\Phi(S^m)\ni\nu\mapsto\int_{S^m}\varphi\,d\nu$ is continuous for every $\varphi\in\Phi$. If $\Phi=B(S^m)$, then $\mathcal{M}_\Phi(S^m)=\mathcal{M}_1(S^m)$ and τ_Φ coincides with the usual τ-topology introduced above. The σ-algebra on $\mathcal{M}_1(S^m)$ is defined to be the smallest one such that $\mathcal{M}_\Phi(S^m)$ and all the maps $\mathcal{M}_1(S^m)\ni\nu\mapsto\int_{S^m}f\,d\nu$ with $f\in B(S^m)$ are measurable. Note that $\mathcal{M}_\Phi(S^m)$ is convex and contains the Dirac measure δ_x for every $x\in S^m$. We consider the following three exponential moment conditions for Φ:

Condition 1.3. *For every $\varphi\in\Phi$ there exists at least one $\alpha_\varphi>0$ such that*

$$\int_{S^m}\exp(\alpha_\varphi|\varphi|)\,d\mu^{\otimes m}<\infty.$$

Condition 1.4. *For every $\varphi\in\Phi$ and every $\alpha>0$,*

$$\int_{S^m}\exp(\alpha|\varphi|)\,d\mu^{\otimes m}<\infty.$$

Condition 1.5. *For every $\varphi\in\Phi$ there exists at least one $\alpha_\varphi>0$ such that*

$$\int_{S^m}\exp(\alpha_\varphi|\varphi\circ\pi_\tau|)\,d\mu^{\otimes m}<\infty$$

for every map $\tau\colon\{1,\dots,m\}\to\{1,\dots,m\}$, where $\pi_\tau\colon S^m\to S^m$ is defined by $\pi_\tau(s)=(s_{\tau(1)},\dots,s_{\tau(m)})$ for every $s=(s_1,\dots,s_m)\in S^m$.

Note that all three conditions are satisfied in the case $\Phi=B(S^m)$. Conditions 1.4 and 1.5 are needed to prove a full large deviation principle in the τ_Φ-topology for the m-fold products $L_n^{\otimes m}\colon\Omega\to\mathcal{M}_1(S^m)$ of the empirical measures, expressible as

$$L_n^{\otimes m}=\frac{1}{n^m}\sum_{i_1,\dots,i_m=1}^{n}\delta_{(X_{i_1},\dots,X_{i_m})}\tag{1.6}$$

for $n\in\mathbb{N}$. For applications to U-statistics, it is also of interest to prove a large deviation principle for the measures $L_n^m\colon\Omega\to\mathcal{M}_1(S^m)$ with $n\geq m$, which are defined by

$$L_n^m=\frac{1}{n_{(m)}}\sum_{(i_1,\dots,i_m)\in I_{m,n}}\delta_{(X_{i_1},\dots,X_{i_m})},\tag{1.7}$$

where $n_{(m)}\equiv\prod_{k=0}^{m-1}(n-k)$ and where $I_{m,n}\subset\{1,\dots,n\}^m$ contains all m-tuples with pairwise different components. Due to their application, we call the measures

$\{L_n^m\}_{n\geq m}$ the U-empirical measures of order m; see (1.12). Our large deviations result for these measures only requires Condition 1.4. Note that $L_n^{\otimes m}$ and L_n^m take values in $\mathcal{M}_\Phi(S^m)$, hence both mappings turn out to be measurable.

Let us recall the definition of the relative entropy $H(\nu\,|\,\tilde\nu)$ of $\nu\in\mathcal{M}_1(S)$ with respect to $\tilde\nu\in\mathcal{M}_1(S)$:

$$H(\nu\,|\,\tilde\nu)=\begin{cases}\int_S f\log f\,d\tilde\nu, & \text{if } \nu\ll\tilde\nu \text{ and } f=\frac{d\nu}{d\tilde\nu},\\ \infty & \text{otherwise.}\end{cases}$$

The rate function $J_m\colon \mathcal{M}_1(S^m)\to[0,\infty]$ is defined by

$$J_m(\nu)=\begin{cases}\frac{1}{m}H(\nu\,|\,\mu^{\otimes m}), & \text{if } \nu=\nu_1^{\otimes m},\\ \infty & \text{otherwise,}\end{cases}$$

where ν_1 denotes the first marginal of ν. For every $B\subset\mathcal{M}_1(S^m)$ define $J_m(B)=\inf_{\nu\in B}J_m(\nu)$. Note that, if there exists $A\in\mathcal{S}$ satisfying $0<\mu(A)<1$, then the rate function J_m is non-convex for all $m\geq 2$. To see this, define the conditional probability measures $\nu=\mu(\,\cdot\,|A)$ and $\tilde\nu=\mu(\,\cdot\,|A^{\mathrm{c}})$ and show that $\frac12(\nu^{\otimes m}+\tilde\nu^{\otimes m})$ is not a product measure. For every $r\in[0,\infty)$ define the level set

$$K_{m,r}=\{\,\nu\in\mathcal{M}_1(S^m)\mid J_m(\nu)\leq r\,\}.$$

Our main results, which we prove in Section 2, are the following two theorems:

Theorem 1.8 (Large deviations of U-empirical measures of order m)**.**

(a) *For every measurable $B\subset\mathcal{M}_1(S^m)$,*

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(L_n^m\in B)\geq -J_m(\mathrm{int}_{\tau_\Phi}(B)),\tag{1.9}$$

where $\mathrm{int}_{\tau_\Phi}(B)$ denotes the interior of the set $B\cap\mathcal{M}_\Phi(S^m)$ with respect to the τ_Φ-topology.

(b) *If Condition* 1.3 *holds, then $K_{m,r}\subset\mathcal{M}_\Phi(S^m)$ for every $r\in[0,\infty)$.*

(c) *If Condition* 1.4 *holds, then $K_{m,r}$ is τ_Φ-compact and sequentially τ_Φ-compact for every $r\in[0,\infty)$.*

(d) *If Condition* 1.4 *holds, then*

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}(L_n^m\in B)\leq -J_m(\mathrm{cl}_{\tau_\Phi}(B))\tag{1.10}$$

for every measurable $B\subset\mathcal{M}_1(S^m)$, where $\mathrm{cl}_{\tau_\Phi}(B)$ denotes the closure of the set $B\cap\mathcal{M}_\Phi(S^m)$ with respect to the τ_Φ-topology.

Theorem 1.11 (Large deviations of m-fold products of empirical measures)**.**

(a) *If Condition* 1.5 *holds, then* (1.9) *is true with $L_n^{\otimes m}$ in place of L_n^m.*

(b) *If Conditions* 1.4 *and* 1.5 *hold, then* (1.10) *is true with $L_n^{\otimes m}$ in place of L_n^m.*

Theorem 1.8, specialized to $m=1$, yields the aforementioned large deviation principle for the empirical measures $\{L_n\}_{n\in\mathbb{N}}$ in the τ_Φ-topology. Note that the lower bound in Theorem 1.8(a) does not make use of Condition 1.3. But if Condition 1.3 does not hold, then the lower bound might loose its strength because measures $\nu\in B$ with $J_m(\nu)<\infty$ might not be contained in $\mathcal{M}_\Phi(S^m)$.

For the proofs of the lower bounds we adopt the method outlined in [10, Exercise 3.2.23(ii)] and combine it with the law of large numbers for U-statistics and V-statistics, respectively. To prove the upper bounds, we basically use de Acosta's approach contained in [1]; but we have to cope with the unbounded functions in Φ, which cause the τ_Φ-topology to be finer than the usual τ-topology. Furthermore, instead of using Cramér's theorem, we prefer to derive Theorem 1.11(b) from the upper bound of the Gärtner–Ellis theorem in $\mathbb{R}^d$. The reason for this is the presence of the "diagonal" terms in (1.6), which are those with at least two equal components in $(i_1, \dots, i_m)$. The influence of these terms vanishes for n tending to infinity, and the Gärtner–Ellis theorem is well adopted to handle this.

There are two non-trivial problems in transferring de Acosta's approach to the case $m \geq 2$: The dependence of the terms in (1.7) and the non-convexity of the rate function J_m. The first problem is treated by using the existence of almost regular partitions of complete hypergraphs, see Baranyai [4]. This result allows us to conveniently decompose the sum in (1.7) into an n-dependent number of partial sums, each of which consists of independent terms. To circumvent the non-convexity of J_m, we use the convexity of $H(\,\cdot\,|\,\mu^{\otimes m})$ on $\mathcal{M}_1(S^m)$ and the fact that $L_n^{\otimes m}$ is a product measure as well as a good approximation for L_n^m. Finally we have to show in Lemma 2.8 that the infimum of $H(\,\cdot\,|\,\mu^{\otimes m})$ over all "η-approximate product measures" of a τ_Φ-closed set C tends to the infimum over all product measures in C as $\eta \downarrow 0$.

1.3. **Application to U- and V-statistics.** Next we want to derive large deviation results for U-statistics and V-statistics (also called von Mises statistics) from Theorem 1.8 and Theorem 1.11, respectively. Many statistics in common use are members of these two classes and many other statistics may be approximated by a member of one of these classes. Particularly in the field of non-parametric statistics, the significance of the two classes is made plain, see for example [9, Chap. 1], [19], and [23, Chap. 5]. For a measurable map $h\colon S^m \to \mathbb{R}^d$ the U- and V-statistics of degree m with kernel function h are defined by

$$U_n^m(h) = \int_{S^m} h\,dL_n^m \quad \text{and} \quad V_n^m(h) = \int_{S^m} h\,dL_n^{\otimes m} \tag{1.12}$$

for all $n \geq m$ or all $n \in \mathbb{N}$, respectively. Let $h_1, \dots, h_d$ denote the component functions of h and define $\Phi_h = B(S^m) \cup \{h_1, \dots, h_d\}$. Then the statistics defined in (1.12) are compositions of L_n^m and $L_n^{\otimes m}$, respectively, with the τ_{Φ_h}-continuous functional $\mathcal{M}_{\Phi_h}(S^m) \ni \nu \mapsto \int_{S^m} h\,d\nu$, and the contraction principle [7, Theorem 4.2.1] immediately leads to the following theorem which extends results obtained in [14] and [13, Section 5]. Note that the following theorem uses weaker moment conditions.

Theorem 1.13. *Assume that $\int_{S^m} \exp(\alpha|h_i|)\,d\mu^{\otimes m} < \infty$ for all $i \in \{1, \dots, d\}$ and $\alpha > 0$. Then the following holds:*

(a) *The U-statistics $\{U_n^m(h)\}_{n\in\mathbb{N}}$ satisfy a full large deviation principle with the good rate function $I_{h,m}\colon \mathbb{R}^d \to [0,\infty]$ given by*

$$I_{h,m}(x) = \inf\Big\{ J_m(\nu) \;\Big|\; \nu \in \mathcal{M}_{\Phi_h}(S^m), \int_{S^m} h\,d\nu = x \Big\}. \tag{1.14}$$

(b) *If there exists at least one $\alpha_h > 0$ such that*

$$\int_{S^m} \exp(\alpha_h |h_i \circ \pi_\tau|)\,d\mu^{\otimes m} < \infty$$

for every $i \in \{1, \dots, d\}$ and every map $\tau\colon \{1, \dots, m\} \to \{1, \dots, m\}$, then the V-statistics $\{V_n^m(h)\}_{n\in\mathbb{N}}$ also satisfy a full large deviation principle with the good rate function $I_{h,m}$ given by (1.14).

Remark 1.15. If Condition 1.3 holds for Φ_h, then it follows from Theorem 1.8(b) that $K_{m,\infty} \equiv \bigcup_{r>0} K_{m,r}$ is contained in $\mathcal{M}_{\Phi_h}(S^m)$. Since $J_m(\nu) = \infty$ for all $\nu \in \mathcal{M}_{\Phi_h}(S^m) \setminus K_{m,\infty}$, it follows that the rate function $I_{h,m}$ defined by (1.14) coincides with the one in [13, Theorem 5.1].

1.4. **Application to the Gibbs conditioning principle.** As before, let $(S, \mathcal{S})$ be an arbitrary measurable space, and let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of S-valued, independent random variables, each one distributed according to the law μ. In statistical mechanics, $X_1, \dots, X_n$ may denote some attribute of n independent particles. Given $m \in \mathbb{N}$ and a measurable subset A of $\mathcal{M}_1(S^m)$, it is of interest to compute the law of $(X_1, \dots, X_m)$ under the constraint $L_n^m \in A$ and to determine the possible limits as $n \to \infty$, provided that $\mathbb{P}(L_n^m \in A) > 0$ for all sufficiently large $n \geq m$. In statistical mechanics, the distribution under the alluded conditioning is called the micro-canonical distribution of the system. Under suitable assumptions, the limiting conditional distribution is a Gibbs measure, which minimizes the relative entropy under the given constraint. This justifies the name "Gibbs conditioning principle". It was studied by Stroock and Zeitouni [24], for example, who give some remarks on the history of the problem, too. We will mostly refer to the exposition given in [7, Section 7.3]. For refinements see also the recent article by Dembo and Zeitouni [8], where they study the distribution of $(X_1, \dots, X_{k(n)})$ under a condition on L_n for a sequence $\{k(n)\}_{n\in\mathbb{N}}$ increasing with n. When convexity is present in the constraint, then Csiszár [6] obtains convergence in information to the generalized entropy-minimizing projection, which implies convergence in variation distance.

For every $f \in L^1(S^m, \mathcal{S}^{\otimes m}, \mu^{\otimes m})$ and $n \geq m$ satisfying $\mathbb{P}(L_n^m \in A) > 0$, we obtain, due to the exchangeability of $X_1, \dots, X_n$ under the condition $L_n^m \in A$,

$$\mathbb{E}[f(X_1, \dots, X_m) \,|\, L_n^m \in A] = \mathbb{E}[f(X_{i_1}, \dots, X_{i_m}) \,|\, L_n^m \in A]$$

for every $(i_1, \dots, i_m) \in I_{m,n}$ appearing in (1.7). Therefore,

$$\mathbb{E}[f(X_1, \dots, X_m) \,|\, L_n^m \in A] = \mathbb{E}\Big[\int_{S^m} f \, dL_n^m \,\Big|\, L_n^m \in A\Big].$$

Thus, computing the conditional distribution of $(X_1, \dots, X_m)$ under the condition $L_n^m \in A$ is equivalent to the computation of the conditional distribution of L_n^m under the same condition.

In the above-mentioned references (except [6]) Polish state spaces are assumed. The condition always involves events depending on L_n (and not L_n^m), even when the mean of the marginals of an m-tuple empirical measure is considered like in [7, Exercise 7.3.5]. When modelling an m-particle interaction with a potential $U\colon S^m \to \mathbb{R}$, the τ-continuity of $\mathcal{M}_1(S) \ni \nu \mapsto \int_{S^m} U \, d\nu^{\otimes m}$ has to be ensured, see [7, remark on p. 298]. Boundedness and continuity of U are sufficient for this property [7, Lemma 7.3.12]; the continuity of U could be relaxed [5, Example 2.9].

Using Theorem 1.8, we want to show that conditioning on the U-empirical measures $\{L_n^m\}_{n\geq m}$ is possible. This enables us to treat $\mathcal{S}^{\otimes m}$-measurable potentials $U\colon S^m \to \mathbb{R}$, which have exponential moments of all orders. In particular, we do not need any topological structure of the measurable space $(S, \mathcal{S})$ for this part.

As before, let Φ be a set of measurable functions $\varphi\colon S^m \to \mathbb{R}$ containing $B(S^m)$. For the above-mentioned set A one would like to take, for example, the level set $\{\, \nu \in \mathcal{M}_\Phi(S^m) \mid \Psi(\nu) = 0 \,\}$ of an energy functional $\Psi\colon \mathcal{M}_\Phi(S^m) \to \mathbb{R}$. To use large deviation results and to avoid problems with the case $\mathbb{P}(L_n^m \in A) = 0$, we work with the usual approximating set-up. Let $\{A_\delta\}_{\delta>0}$ be a nested collection of measurable subsets of $\mathcal{M}_\Phi(S^m)$, where nested means that $A_\delta \subset A_{\delta'}$ whenever $\delta \le \delta'$. In addition, let $\{F_\delta\}_{\delta>0}$ be nested τ_Φ-closed subsets of $\mathcal{M}_\Phi(S^m)$, not necessarily measurable, such that $A_\delta \subset F_\delta$ for all $\delta > 0$. Define $F_0 = \bigcap_{\delta>0} F_\delta$ and $A_0 = \bigcap_{\delta>0} A_\delta$.

Condition 1.16. *There exists a $\hat{\nu} \in A_0$ (not necessarily unique) such that $J_m(\hat{\nu}) = J_m(F_0) < \infty$ and*

$$\lim_{n\to\infty} \hat{\nu}_1^{\otimes\mathbb{N}}(\{L_n^m \in A_\delta\}) = 1 \tag{1.17}$$

for every $\delta > 0$, where $\hat{\nu}_1$ denotes the first marginal of $\hat{\nu}$.

If Condition 1.3 holds and if there exists a $\hat{\nu} \in A_0$ such that $J_m(\hat{\nu}) = J_m(F_0) < \infty$ and $\hat{\nu} \in \operatorname{int}_{\tau_\Phi}(A_\delta)$ for every $\delta > 0$, then Theorem 1.8(b) and Lemma 2.6(a) imply (1.17). The following theorem, which we prove in Section 3, shows that the U-empirical measures concentrate in every measurable τ_Φ-neighbourhood of the rate-function minimizing measures under the given constraint.

Theorem 1.18. *Assume Conditions* 1.4 *and* 1.16. *Then the following holds:*

(a) *The set $\mathcal{E} \equiv \{\, \nu \in F_0 \mid J_m(\nu) = J_m(F_0) \,\}$ is nonempty, τ_Φ-compact and sequentially τ_Φ-compact.*

(b) *If $\Gamma \subset \mathcal{M}_1(S^m)$ is measurable and $\mathcal{E} \subset \operatorname{int}_{\tau_\Phi}(\Gamma)$, then*

$$\limsup_{\delta\downarrow 0}\ \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(L_n^m \notin \Gamma \,|\, L_n^m \in A_\delta) < 0.$$

Remark 1.19. For $l \in \{1,\dots,m\}$ and a measurable subset Γ' of $\mathcal{M}_1(S^l)$, Theorem 1.18(b) can also apply. If $\Gamma \equiv \pi_{m,l}^{-1}(\Gamma')$, where $\pi_{m,l}\colon \mathcal{M}_1(S^m) \to \mathcal{M}_1(S^l)$ is the projection onto the first l marginals, then $\{L_n^l \notin \Gamma'\} = \{L_n^m \notin \Gamma\}$ for all $n \ge m$.

To apply Theorem 1.18, let $U\colon S^m \to \mathbb{R}$ be an $\mathcal{S}^{\otimes m}$-measurable function such that $\int_{S^m} \exp(\alpha|U|)\, d\mu^{\otimes m} < \infty$ for all $\alpha > 0$. Then $\Phi \equiv B(S^m) \cup \{U\}$ satisfies Condition 1.4. Furthermore, the map $\Psi\colon \mathcal{M}_\Phi(S^m) \to \mathbb{R}$ with $\Psi(\nu) \equiv \int_{S^m} U\, d\nu - 1$ is measurable and τ_Φ-continuous. The sets $A_\delta \equiv F_\delta \equiv \{\, \nu \in \mathcal{M}_\Phi(S^m) \mid |\Psi(\nu)| \le \delta \,\}$ with $\delta > 0$ are, therefore, nested, measurable and τ_Φ-closed. Obviously, $A_0 = F_0 = \{\, \nu \in \mathcal{M}_\Phi(S^m) \mid \Psi(\nu) = 0 \,\}$ and $A_0 \subset \operatorname{int}_{\tau_\Phi}(A_\delta)$ for every $\delta > 0$. If there exists a $\nu \in A_0$ with $r \equiv J_m(\nu) < \infty$, then $A_0 \cap K_{m,r}$ is τ_Φ-compact by Theorem 1.8(c). Hence, the lower τ_Φ-semicontinuous rate function J_m attains the infimum $J_m(A_0)$ at a $\hat{\nu} \in A_0$ and Condition 1.16 holds. Note that $\hat{\nu}$ is necessarily a product measure.

It remains to find further conditions on U such that there exists a $\nu \in A_0$ satisfying $J_m(\nu) < \infty$. Furthermore, it is of interest to know under which assumptions the measure $\hat{\nu}$ from Condition 1.16 is a uniquely determined Gibbs measure. Since our main results do not contribute to this discussion, we do not pursue this topic further. For $m = 2$ it is treated in [7, Subsection 7.3.2].

1.5. **Extension of the large deviation principles and concluding remarks.** For the sake of completeness, we want to state the extensions of Theorems 1.8 and 1.11 to arbitrary subsets of $\mathcal{M}_1(S^m)$. For every subset A of $S^{\mathbb{N}}$ let $\mathbb{P}^*(A)$ and $\mathbb{P}_*(A)$ denote the outer and the inner $\mathbb{P}$-measure of A, respectively.

Theorem 1.20. *Let B be an arbitrary subset of $\mathcal{M}_1(S^m)$.*

(a) *With $\operatorname{int}_{\tau_\Phi}(B)$ denoting the τ_Φ-interior of $B \cap \mathcal{M}_\Phi(S^m)$,*

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_*((L_n^m)^{-1}(B)) \ge -J_m(\operatorname{int}_{\tau_\Phi}(B)). \tag{1.21}$$

(b) *If Condition* 1.4 *holds, then*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}^*((L_n^m)^{-1}(B)) \le -J_m(\mathrm{cl}_{\tau_\Phi}(B)), \tag{1.22}$$

where $\mathrm{cl}_{\tau_\Phi}(B)$ *denotes the* τ_Φ*-closure of* $B \cap \mathcal{M}_\Phi(S^m)$.

(c) *If Condition* 1.5 *holds, then* (1.21) *is true with* $L_n^{\otimes m}$ *in place of* L_n^m.

(d) *If Conditions* 1.4 *and* 1.5 *hold, then* (1.22) *is true with* $L_n^{\otimes m}$ *in place of* L_n^m.

Using Theorem 1.8(c) instead of [1, Lemma 2.1] and the corresponding upper and lower bounds of Theorems 1.8 and 1.11 instead of [1, Theorem 1.1], the proof of [1, Theorem 1.2] can be transferred to the present situation without problems.

Remark 1.23. If S is a Polish space with Borel σ-algebra $\mathcal{S}$, then we have the weak topologies on $\mathcal{M}_1(S)$ and $\mathcal{M}_1(S^m)$ available and we can use the continuity of the map $\nu \mapsto \nu^{\otimes m}$ with respect to the weak topologies [15, Chap. 3, Proposition 4.6]. Therefore, if the empirical measures $\{L_n\}_{n\in\mathbb{N}}$ satisfy a large deviation principle in the weak topology, then the contraction principle implies that the products $\{L_n^{\otimes m}\}_{n\in\mathbb{N}}$ satisfy a large deviation principle in the weak topology on $\mathcal{M}_1(S^m)$. If there exist constants $\beta, M \in [1,\infty)$ and a reference measure $\tilde{\mu} \in \mathcal{M}_1(S^m)$ such that the inequality

$$\sup_{n\in\mathbb{N}} \left(\mathbb{E}\left[\exp\left(n \int_{S^m} V \, dL_n^{\otimes m} \right) \right] \right)^{1/n} \le M \int_{S^m} \exp(\beta V) \, d\tilde{\mu}$$

holds for all bounded measurable functions $V\colon S^m \to [0,\infty)$, then we can use [10, Lemma 3.2.19 and Theorem 3.2.21] to infer that $\{L_n^{\otimes m}\}_{n\in\mathbb{N}}$ actually satisfy a large deviation principle in the τ-topology on $\mathcal{M}_1(S^m)$. We will use this approach in a forthcoming paper for random variables $\{X_i\}_{i\in\mathbb{N}}$ which are dependent or not identically distributed.

Remark 1.24. There does *not* exist a scale $\{\varepsilon_n\}_{n\in\mathbb{N}_0}$ with $\varepsilon_n \downarrow 0$ such that the random measures $\{L_n\}_{n\in\mathbb{N}_0}$ from Example 1.2 satisfy a large deviation upper bound of the form

$$\limsup_{n\to\infty} \varepsilon_n \log \mu(L_n \in C) \le -\inf_{\nu\in C} I(\nu)$$

for all τ-closed measurable $C \subset \mathcal{M}_1(S)$, where $I\colon \mathcal{M}_1(S) \to [0,\infty]$ with $I(\mu) = 0$ and $I(\nu) = \infty$ for $\nu \ne \mu$ is the rate function which governs the large deviations of $\{L_n\}_{n\in\mathbb{N}_0}$ with respect to the weak topology on every scale $\{\varepsilon_n\}_{n\in\mathbb{N}_0}$ with $\varepsilon_n \downarrow 0$. To substantiate this claim, consider the set $C \equiv \{\, \nu \in \mathcal{M}_1(S) \mid \nu(A) \ge \mu(A) + 1/2 \,\}$, where we construct the set $A \in \mathcal{S}$ as follows: Choose a subsequence $\{\varepsilon_{n_k}\}_{k\in\mathbb{N}}$ such that $\sum_{k\in\mathbb{N}} \varepsilon_{n_k} \le 1/2$ and define $A = \bigcup_{k\in\mathbb{N}} A_k$, where

$$A_k = \bigcup_{l=0}^{2^{n_k}-1} [l2^{-n_k}, (l+\varepsilon_{n_k})2^{-n_k}).$$

Then $\mu(A_k) = \varepsilon_{n_k}$ and $\mu(A) \le \sum_{k\in\mathbb{N}} \varepsilon_{n_k} \le 1/2$ as well as $L_{n_k}(A) \ge L_{n_k}(A_k) = 1$ on A_k for every $k \in \mathbb{N}$. Hence, as $k \to \infty$,

$$\varepsilon_{n_k} \log \mu(\{L_{n_k}(A) \ge \mu(A) + 1/2\}) \ge \varepsilon_{n_k} \log \mu(A_k) = \varepsilon_{n_k} \log \varepsilon_{n_k} \to 0.$$

Therefore, we are still looking for an example of a sequence of random (if possible, empirical) measures $\{L_n\}_{n\in\mathbb{N}}$ which satisfy a large deviation principle in the τ-topology, but for which the products $\{L_n^{\otimes m}\}_{n\in\mathbb{N}}$ for some $m \ge 2$ do *not* satisfy the corresponding large deviation principle in the τ-topology on $\mathcal{M}_1(S^m)$.

2. Proofs of the large deviation principles

For every $r \in [0,\infty)$ define the level set of the relative entropy by

$$K'_{m,r} = \{\, \nu \in \mathcal{M}_1(S^m) \mid H(\nu \,|\, \mu^{\otimes m}) \le mr \,\}$$

and note that $K_{m,r} = C_m \cap K'_{m,r}$, where $C_m \equiv \{\, \nu \in \mathcal{M}_1(S^m) \mid \nu = \nu_1^{\otimes m} \,\}$ is the set of product measures. Therefore, part (b) of the following lemma implies part (b) of Theorem 1.8.

Lemma 2.1. *Let $r \in [0,\infty)$.*

(a) *The set $K'_{m,r}$ is τ-compact and sequentially τ-compact.*
(b) *If Condition* 1.3 *holds, then $K'_{m,r} \subset \mathcal{M}_\Phi(S^m)$.*
(c) *If Condition* 1.4 *holds, then the identity on $K'_{m,r}$ is τ-τ_Φ-continuous, hence both topologies coincide on $K'_{m,r}$ and $K'_{m,r}$ is τ_Φ-compact and also sequentially τ_Φ-compact.*

Proof. (a) See [1, Lemma 2.1] for the τ-compactness and [16, Theorem 2.6] for the sequential τ-compactness.

(b) By convexity, $z \le e^{z-1}$ for all $z \in \mathbb{R}$. Substituting $z = x - t$ yields $x \le e^{x-t-1} + t$ for all $t, x \in \mathbb{R}$. Multiplication with $y = e^t$ gives the well-known estimate

$$xy \le e^{x-1} + y \log y \quad \text{for all } x \in \mathbb{R} \text{ and } y \in [0,\infty). \tag{2.2}$$

If $\nu \in \mathcal{M}_1(S^m)$ satisfies $H(\nu \,|\, \mu^{\otimes m}) \le mr$, then there exists a density f of ν with respect to $\mu^{\otimes m}$. Given $\alpha > 0$, $A \in \mathcal{S}^{\otimes m}$ and $\varphi \in \Phi$, the estimate (2.2) leads to

$$\int_A |\varphi| \, d\nu = \frac{1}{\alpha} \int_A (\alpha|\varphi|) f \, d\mu^{\otimes m} \le \frac{1}{\alpha e} \int_A e^{\alpha|\varphi|} \, d\mu^{\otimes m} + \frac{1}{\alpha} \int_A f \log f \, d\mu^{\otimes m}. \tag{2.3}$$

Since $y \log y \ge -1/e$ for all $y \in [0,\infty)$,

$$\int_A |\varphi| \, d\nu \le \frac{1}{\alpha} \biggl(\frac{1}{e} \int_A e^{\alpha|\varphi|} \, d\mu^{\otimes m} + \frac{1}{e} + rm \biggr). \tag{2.4}$$

Choosing $A = S^m$ and $\alpha = \alpha_\varphi$ in (2.4), part (b) follows.

(c) Given $\varphi \in \Phi$ and $\varepsilon > 0$, define $\alpha = (1/e + rm)/\varepsilon$. By Condition 1.4 and the dominated convergence theorem, there exists $n_0 \in \mathbb{N}$ such that

$$\frac{1}{\alpha e} \int_{\{|\varphi|>n\}} e^{\alpha|\varphi|} \, d\mu^{\otimes m} \le \varepsilon$$

for all $n \ge n_0$. Using (2.4), it follows that

$$\biggl| \int_{S^m} \varphi \, d\nu - \int_{\{|\varphi| \le n\}} \varphi \, d\nu \biggr| \le \int_{\{|\varphi|>n\}} |\varphi| \, d\nu \le 2\varepsilon$$

for all $n \ge n_0$ and all $\nu \in K'_{m,r}$, hence $K'_{m,r} \ni \nu \mapsto \int_{S^m} \varphi \, d\nu$ is τ-continuous, because it is the uniform limit of the τ-continuous functions $K'_{m,r} \ni \nu \mapsto \int_{\{|\varphi| \le n\}} \varphi \, d\nu$ as $n \to \infty$. Hence, the identity on $K'_{m,r}$ is τ-τ_Φ-continuous and the τ_Φ-compactness of $K'_{m,r}$ follows from part (a). Since the identity is bijective, it is a τ-τ_Φ-homeomorphism on $K'_{m,r}$ [18, Chap. 5, Theorem 8], hence both topologies coincide on $K'_{m,r}$. Therefore, the sequential τ_Φ-compactness also follows from part (a). □

Proof of Theorem 1.8(c)*:* The set C_m of all product measures is τ-closed because

$$C_m = \bigcap_{A_1,\dots,A_m \in \mathcal{S}} \biggl\{ \nu \in \mathcal{M}_1(S^m) \biggm| \nu(A_1 \times \cdots \times A_m) = \prod_{i=1}^m \nu_1(A_i) \biggr\}. \tag{2.5}$$

Hence, by Lemma 2.1(a), the set $K_{m,r} = C_m \cap K'_{m,r}$ is τ-compact and sequentially τ-compact. The τ_Φ-compactness and the sequential τ_Φ-compactness of $K_{m,r}$ now follow from Lemma 2.1(c). □

Lemma 2.6. *Let $\nu \in \mathcal{M}_1(S)$.*

(a) *If $\varphi \in L_1(S^m, \mathcal{S}^{\otimes m}, \nu^{\otimes m})$, then*

$$\lim_{n\to\infty} \int_{S^m} \varphi \, dL_n^m = \int_{S^m} \varphi \, d\nu^{\otimes m} \quad \nu^{\otimes \mathbb{N}}\text{-almost surely.}$$

(b) *If $\varphi \circ \pi_\tau \in L_1(S^m, \mathcal{S}^{\otimes m}, \nu^{\otimes m})$ for every $\tau \colon \{1, \dots, m\} \to \{1, \dots, m\}$, then*

$$\lim_{n\to\infty} \int_{S^m} \varphi \, dL_n^{\otimes m} = \int_{S^m} \varphi \, d\nu^{\otimes m} \quad \nu^{\otimes \mathbb{N}}\text{-almost surely.}$$

Proof. Part (a) is a reformulation of the strong law of large numbers for U-statistics, see for example [23, Theorem 5.4.A]. Part (b) is the strong law of large numbers for V-statistics, see for example [9, Theorem 1.3.6]. □

Proof of Theorem 1.8(a)*:* Consider a measure $\nu \in \operatorname{int}_{\tau_\Phi}(B)$ with $J_m(\nu) < \infty$. Then $\Phi \subset L_1(S^m, \mathcal{S}^{\otimes m}, \nu)$, $\nu = \nu_1^{\otimes m}$ and $J_m(\nu) = H(\nu_1 \,|\, \mu)$. Define $f = d\nu_1/d\mu$ and $F_n(s) = \prod_{i=1}^n f(s_i)$ for all $s = (s_i)_{i\in\mathbb{N}} \in S^{\mathbb{N}}$. By the definition of the τ_Φ-topology, there exist $\varepsilon > 0$, $k \in \mathbb{N}$ and $\varphi_1, \dots, \varphi_k \in \Phi$ such that the τ_Φ-open set

$$C \equiv \left\{ \tilde\nu \in \mathcal{M}_\Phi(S^m) \,\middle|\, \left| \int_{S^m} \varphi_i \, d\tilde\nu - \int_{S^m} \varphi_i \, d\nu \right| < \varepsilon \text{ for every } i \in \{1, \dots, k\} \right\}$$

is contained in $\operatorname{int}_{\tau_\Phi}(B)$. Note that C is a measurable subset of $\mathcal{M}_1(S^m)$ because $\mathcal{M}_\Phi(S^m)$ is measurable by definition. Define $D_n = \{L_n^m \in C,\, F_n > 0\}$ and note that $a_n \equiv \nu_1^{\otimes\mathbb{N}}(D_n) = \nu_1^{\otimes\mathbb{N}}(\{L_n^m \in C\})$ for every $n \in \mathbb{N}$. It follows from Lemma 2.6(a) that

$$\lim_{n\to\infty} \nu_1^{\otimes\mathbb{N}}(\{L_n^m \in C\}) = 1. \tag{2.7}$$

Choose $n_0 \in \mathbb{N}$ such that $a_n > 0$ for all $n \ge n_0$. Then, for every $n \ge n_0$,

$$\mathbb{P}(L_n^m \in B) \ge \mathbb{P}(L_n^m \in C) \ge \int_{D_n} \frac{1}{F_n} \, d\nu_1^{\otimes\mathbb{N}}.$$

Using Jensen's inequality, we obtain

$$\begin{aligned} \log \int_{D_n} \frac{1}{F_n} \, d\nu_1^{\otimes\mathbb{N}} &\ge \log a_n - \frac{1}{a_n} \int_{D_n} \log F_n \, d\nu_1^{\otimes\mathbb{N}} \\ &= \log a_n - \frac{1}{a_n} \int_{D_n} F_n \log F_n \, d\mu^{\otimes\mathbb{N}}. \end{aligned}$$

Since $x \log x \ge -1/e$ for all $x \in [0, \infty)$, it follows that

$$\begin{aligned} \int_{D_n} F_n \log F_n \, d\mu^{\otimes\mathbb{N}} &\le \frac{1}{e} + \int_{S^{\mathbb{N}}} F_n \log F_n \, d\mu^{\otimes\mathbb{N}} \\ &= \frac{1}{e} + nH(\nu_1 \,|\, \mu) = \frac{1}{e} + nJ_m(\nu). \end{aligned}$$

The last three displays together yield

$$\log \mathbb{P}(L_n^m \in B) \ge \log a_n - \frac{1}{e a_n} - \frac{nJ_m(\nu)}{a_n}.$$

Using (2.7), Theorem 1.8(a) follows. □

With a small modification the same proof applies for the large-deviations lower bound for products of empirical measures.

Proof of Theorem 1.11(a)*:* If $\nu \in \mathcal{M}_1(S^m)$ with $J_m(\nu) < \infty$, then $\nu = \nu_1^{\otimes m}$. Using (2.3) and Condition 1.5, it follows that $\varphi \circ \pi_\tau \in L_1(S^m, \mathcal{S}^{\otimes m}, \nu_1^{\otimes m})$ for all $\varphi \in \Phi$ and $\tau\colon \{1, \dots, m\} \to \{1, \dots, m\}$. Therefore, Lemma 2.6(b) is applicable and replaces Lemma 2.6(a) in the preceding proof. □

Before we can prove the upper bounds, we need two additional lemmas. For every $\eta \geq 0$ define the set of all η-approximate product measures by

$$\mathcal{A}_{m,\eta} = \bigcap_{A_1,\dots,A_m \in \mathcal{S}} \{\, \nu \in \mathcal{M}_1(S^m) \bigm| |\nu(A_1 \times \cdots \times A_m) - \nu_1^{\otimes m}(A_1 \times \cdots \times A_m)| \leq \eta \,\}.$$

Note that every $\mathcal{A}_{m,\eta}$ is τ-closed in $\mathcal{M}_1(S^m)$ and that $\nu = \nu_1^{\otimes m}$ for every $\nu \in \mathcal{A}_{m,0}$.

Lemma 2.8. *If Condition* 1.4 *holds, then*

$$\lim_{\eta \downarrow 0} \inf_{\nu \in C \cap \mathcal{A}_{m,\eta}} \frac{1}{m} H(\nu \,|\, \mu^{\otimes m}) = J_m(C) \tag{2.9}$$

for every τ_Φ-closed subset C of $\mathcal{M}_\Phi(S^m)$.

Proof. For every $\eta \geq 0$ define the greatest lower bound l_η by

$$l_\eta = \inf_{\nu \in C \cap \mathcal{A}_{m,\eta}} \frac{1}{m} H(\nu \,|\, \mu^{\otimes m}).$$

If $\eta \leq \eta'$, then $\mathcal{A}_{m,\eta} \subset \mathcal{A}_{m,\eta'}$ and $l_\eta \geq l_{\eta'}$. Therefore, the limit in (2.9) exists and $l \equiv \lim_{\eta\downarrow 0} l_\eta \leq l_0$. Note that l_0 equals the right-hand side of (2.9).

To prove that $l \geq l_0$, it suffices to consider the case $l < \infty$. Choose $l' \in (l, \infty)$. By Lemma 2.1(c) the set $C_{\eta,l'} \equiv C \cap \mathcal{A}_{m,\eta} \cap K'_{m,l'}$ is τ_Φ-compact for every $\eta \geq 0$. Furthermore, the sets $\{C_{\eta,l'}\}_{\eta \geq 0}$ are decreasing as $\eta \downarrow 0$ and $C_{\eta,l'} \neq \varnothing$ for every $\eta > 0$ because $l' > l$. Hence, there exists $\nu \in \bigcap_{\eta > 0} C_{\eta,l'}$. Obviously, $\nu \in \mathcal{A}_{m,0}$ and, therefore, $\nu \in C_{0,l'}$. This means $l' \geq l_0$, hence $l \geq l_0$. □

Let $\mathcal{F}$ denote the family of all finite, nonempty subsets of Φ. For every $F \in \mathcal{F}$ define

$$\Pi_F\colon \mathcal{M}_\Phi(S^m) \to \mathbb{R}^F \quad \text{by} \quad \Pi_F(\nu) = \biggl(\int_{S^m} \varphi \, d\nu\biggr)_{\varphi \in F}.$$

For $F' \subset F$ with $F' \neq \varnothing$ let $\Pi_{F,F'}\colon \mathbb{R}^F \to \mathbb{R}^{F'}$ denote the canonical projection. For $F \in \mathcal{F}$ and $y = (y_\varphi)_{\varphi \in F} \in \mathbb{R}^F$ define

$$J_F(y) = \sup_{z \in \mathbb{R}^F} \biggl(\langle y, z \rangle - \log \int_{S^m} \exp\Bigl(\sum_{\varphi \in F} z_\varphi \varphi(x)\Bigr)\, \mu^{\otimes m}(dx)\biggr). \tag{2.10}$$

Lemma 2.11. *If Condition* 1.4 *holds, then*

$$J_F(y) = \inf\{\, H(\nu \,|\, \mu^{\otimes m}) \mid \nu \in \mathcal{M}_\Phi(S^m), \Pi_F(\nu) = y \,\}$$

for every $F \in \mathcal{F}$ and $y \in \mathbb{R}^F$.

Proof. If ψ is in the linear span of Φ, then it follows from Hölder's inequality and Condition 1.4 that $\int_{S^m} \exp(|\psi|)\, d\mu^{\otimes m} < \infty$. Hence, $\int_{S^m} |\psi| \, d\nu < \infty$ by (2.3) for every $\nu \in \mathcal{M}_\Phi(S^m)$ with $H(\nu \,|\, \mu^{\otimes m}) < \infty$. Furthermore, note that Π_F is τ_Φ-continuous and that, for every $\varepsilon > 0$ and $y \in \mathbb{R}^F$ with $J_F(y) < \infty$, the set

$$\{\, \nu \in \mathcal{M}_\Phi(S^m) \mid \|\Pi_F(\nu) - y\|_2 \leq \varepsilon,\, H(\nu \,|\, \mu^{\otimes m}) \leq J_F(y) + \varepsilon \,\}$$

is τ_Φ-compact by Lemma 2.1(c). Using these remarks, the proof of [1, Lemma 2.2] can be adapted. □

Proof of Theorem 1.8(d)*:* Define $C = \mathrm{cl}_{\tau_\Phi}(B)$. It suffices to consider only the case $J_m(C) > 0$. Choose $r \in (0, J_m(C))$. According to Lemma 2.8 there exists $n_0 \in \mathbb{N}$ with $n_0 \geq m$ such that, for all $n \geq n_0$,

$$\inf\{\, H(\nu \,|\, \mu^{\otimes m}) \mid \nu \in C \cap \mathcal{A}_{m,m^2/n} \,\} > mr.$$

Define $C_0 = C \cap \mathcal{A}_{m,m^2/n_0}$, which is τ_Φ-closed. Note that $(L_n^m)_1 = L_n$ and

$$\|L_n^m - L_n^{\otimes m}\|_{\mathrm{var}} \leq 1 - \frac{n_{(m)}}{n^m} \leq \frac{n^m - (n-m)^m}{n^m} \leq \frac{m^2}{n},$$

hence $L_n^m \in \mathcal{A}_{m,m^2/n}$ for all $n \geq m$. Since also $L_n^m \in \mathcal{M}_\Phi(S^m)$, it follows that

$$\{L_n^m \in B\} \subset \{L_n^m \in C_0\}$$

for all $n \geq n_0$. By Lemma 2.1(b), the set $K'_{m,r}$ is contained in $\mathcal{M}_\Phi(S^m)$. Since C_0 is τ_Φ-closed and $C_0 \cap K'_{m,r} = \varnothing$, there exist, for every $\nu \in K'_{m,r}$, an $F_\nu \in \mathcal{F}$ and an open neighbourhood $U_\nu \subset \mathbb{R}^{F_\nu}$ of $\Pi_{F_\nu}(\nu)$ such that $C_0 \cap \Pi_{F_\nu}^{-1}(U_\nu) = \varnothing$. Since $K'_{m,r}$ is τ_Φ-compact by Lemma 2.1(c), there exists a finite subset N of $K'_{m,r}$ such that $\bigcup_{\nu\in N} \Pi_{F_\nu}^{-1}(U_\nu)$ covers $K'_{m,r}$. Define $F = \bigcup_{\nu\in N} F_\nu$. Note that $F \in \mathcal{F}$. For every $\nu \in N$ define $U'_\nu = \Pi_{F,F_\nu}^{-1}(U_\nu)$. Note that $U'_\nu \subset \mathbb{R}^F$ is open and $\Pi_F^{-1}(U'_\nu) = \Pi_{F_\nu}^{-1}(U_\nu)$. Define $U = \bigcup_{\nu\in N} U'_\nu$. Then $\Pi_F^{-1}(U) = \bigcup_{\nu\in N} \Pi_{F_\nu}^{-1}(U_\nu)$, hence $\Pi_F^{-1}(U)$ covers $K'_{m,r}$ and is disjoint from C_0. Define $\varepsilon = \mathrm{dist}(\Pi_F(K'_{m,r}), U^{\mathrm{c}})$. Since $\Pi_F(K'_{m,r})$ is a compact subset of the open set U, it follows that $\varepsilon > 0$ and that

$$A_\varepsilon \equiv \{\, x \in \mathbb{R}^F \mid \mathrm{dist}(x, \Pi_F(K'_{m,r})) < \varepsilon \,\}$$

is an open set contained in U. Therefore, for all $n \geq n_0$,

$$\{L_n^m \in B\} \subset \{L_n^m \in C_0\} \subset \{\Pi_F(L_n^m) \in \mathbb{R}^F \setminus A_\varepsilon\,\}. \tag{2.12}$$

For $n \in \mathbb{N}$ let G_n denote the group of all permutations of $\{1,\dots,n\}$. Consider $k, n \in \mathbb{N}$ satisfying $km \leq n$. For $\sigma \in G_n$ define

$$\Omega \ni \omega \mapsto L_{k,n,\sigma}(\omega) = \frac{1}{k}\sum_{j=0}^{k-1} \delta_{(X_{\sigma(jm+1)}(\omega),\dots,X_{\sigma(jm+m)}(\omega))} \in \mathcal{M}_1(S^m).$$

This map is measurable and $\mathbb{P}L_{k,n,\sigma}^{-1}$ does not depend on σ, because $\{X_i\}_{i\in\mathbb{N}}$ are i. i. d. We write $L_{k,n}$ for $L_{k,n,\sigma}$, if σ is the identity on $\{1,\dots,n\}$. Note that

$$S_k \equiv \Pi_F(L_{k,n}) = \frac{1}{k}\sum_{j=0}^{k-1} (\varphi(X_{jm+1},\dots,X_{jm+m}))_{\varphi\in F}$$

is a mean of k independent and identically distributed $\mathbb{R}^F$-valued random variables. By the finite-dimensional version of Cramér's theorem [25, Section 4],

$$\limsup_{k\to\infty} \frac{1}{k} \log \mathbb{P}(\Pi_F(L_{k,km}) \in \mathbb{R}^F \setminus A_\varepsilon) \leq -J_F(\mathbb{R}^F \setminus A_\varepsilon). \tag{2.13}$$

By Lemma 2.11,

$$\begin{aligned} J_F(\mathbb{R}^F \setminus A_\varepsilon) &= \inf\{\, H(\nu \,|\, \mu^{\otimes m}) \mid \nu \in \mathcal{M}_\Phi(S^m), \Pi_F(\nu) \in \mathbb{R}^F \setminus A_\varepsilon \,\} \\ &\geq \inf\{\, H(\nu \,|\, \mu^{\otimes m}) \mid \nu \in \mathcal{M}_\Phi(S^m) \setminus K'_{m,r} \,\} \geq mr. \end{aligned} \tag{2.14}$$

It remains to transfer the upper bound obtainable from (2.13) and (2.14) into a corresponding result for the U-empirical measures $\{L_n^m\}_{n\ge m}$. Consider $n\in\mathbb{N}$ satisfying $n\ge m$. There are $l_n=\binom{n}{m}$ different ordered m-tuples in the index set $I_{m,n}$ appearing in (1.7). If $m=1$, define $k_n=n$, $p_n=1$ and $q_n=0$. If $m\ge 2$, define $k_n=\lfloor n/m\rfloor$, $p_n=l_n-(k_n-1)\lceil l_n/k_n\rceil$ and $q_n=k_n\lceil l_n/k_n\rceil-l_n$. Note that in both cases $l_n=k_np_n+(k_n-1)q_n$ and $q_n\in\{0,1,\dots,k_n-1\}$. If $m\ge 2$ and $n\ge 2m$, then $l_n\ge n(n-1)/2\ge k_n(k_n-1)$, hence p_n is nonnegative because

$$p_n\ge l_n-(k_n-1)\biggl(\frac{l_n}{k_n}+1\biggr)\ge\frac{l_n}{k_n}-(k_n-1)\ge 0.$$

Therefore, we restrict ourselves to $n\ge 2m$ in the following. Let $I'_{m,n}$ denote the set of all ordered m-tuples in $I_{m,n}$. According to a result on complete uniform hypergraphs due to Baranyai [4, Theorem 1], there exists a partition of $I'_{m,n}$ into p_n sets with k_n elements and q_n sets with k_n-1 elements such that each number from $\{1,\dots,n\}$ is a component of at most one m-tuple in each of the p_n+q_n sets. Hence, there exist two (disjoint) subsets G'_n and G''_n of G_n with $|G'_n|=m!\,p_n$ and $|G''_n|=m!\,q_n$ such that for every m-tuple $(i_1,\dots,i_m)\in I_{m,n}$ there exists exactly one pair $(\sigma,j)\in(G'_n\times\{0,1,\dots,k_n-1\})\cup(G''_n\times\{0,1,\dots,k_n-2\})$ such that $(i_1,\dots,i_m)=(\sigma(jm+1),\dots,\sigma(jm+m))$. Thus, we obtain the representation

$$L_n^m=\frac{k_n}{n_{(m)}}\sum_{\sigma\in G'_n}L_{k_n,n,\sigma}+\frac{k_n-1}{n_{(m)}}\sum_{\sigma\in G''_n}L_{k_n-1,n,\sigma}.$$

Since $K'_{m,r}$ is convex, A_ε is convex, too. Hence, for every $n\ge 2m$,

$$\begin{aligned}\{\Pi_F(L_n^m)\in\mathbb{R}^F\setminus A_\varepsilon\}\subset\bigcup_{\sigma\in G'_n}&\{\Pi_F(L_{k_n,n,\sigma})\in\mathbb{R}^F\setminus A_\varepsilon\}\\ &\cup\bigcup_{\sigma\in G''_n}\{\Pi_F(L_{k_n-1,n,\sigma})\in\mathbb{R}^F\setminus A_\varepsilon\}.\end{aligned}$$

Since the distributions of $L_{k_n,n,\sigma}$ and $L_{k_n-1,n,\sigma}$ do not depend on σ,

$$\begin{aligned}\mathbb{P}(\Pi_F(L_n^m)\in\mathbb{R}^F\setminus A_\varepsilon)\le{}& m!\,p_n\,\mathbb{P}(\Pi_F(L_{k_n,k_nm})\in\mathbb{R}^F\setminus A_\varepsilon)\\ &+m!\,q_n\,\mathbb{P}(\Pi_F(L_{k_n-1,(k_n-1)m})\in\mathbb{R}^F\setminus A_\varepsilon).\end{aligned}$$

Note that $p_n\le n^m$ and $q_n\le n$. Hence, using (2.12), (2.13) and (2.14),

$$\begin{aligned}&\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}(L_n^m\in B)\\ &\quad\le\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}(\Pi_F(L_n^m)\in\mathbb{R}^F\setminus A_\varepsilon)\\ &\quad\le\frac{1}{m}\limsup_{k\to\infty}\frac{1}{k}\log\mathbb{P}(\Pi_F(L_{k,km})\in\mathbb{R}^F\setminus A_\varepsilon)\le -r.\end{aligned}\tag{2.15}$$

Since $r\in(0,J_m(C))$ was arbitrary, the upper bound follows. □

The proof of the upper bound for product measures is similar to the previous proof, but we use the Gärtner–Ellis theorem instead of Cramér's theorem to handle the "diagonal terms" in $L_n^{\otimes m}$. Since $L_n^{\otimes m}$ already is a product measure, we don't need Lemma 2.8 here.

Proof of Theorem 1.11(b)*:* Define $C=\mathrm{cl}_{\tau_\Phi}(B)$. It suffices to consider only the case $J_m(C)>0$. Choose $r\in(0,J_m(C))$. With $C_m\equiv\{\,\nu\in\mathcal{M}_1(S^m)\mid\nu=\nu_1^{\otimes m}\,\}$ define

$C_0 = C \cap C_m$. Due to (2.5) the set C_0 is τ_Φ-closed. Furthermore, $L_n^{\otimes m} \in \mathcal{M}_\Phi(S^m)$ and $L_n^{\otimes m} \in C_m$, hence $\{L_n^{\otimes m} \in B\} \subset \{L_n^{\otimes m} \in C_0\}$ for all $n \in \mathbb{N}$. As in the preceding proof we can find $F \in \mathcal{F}$ and an open convex ε-neighbourhood $A_\varepsilon \subset \mathbb{R}^F$ of $\Pi_F(K'_{m,r})$ such that, for all $n \in \mathbb{N}$,

$$\{L_n^{\otimes m} \in B\} \subset \{L_n^{\otimes m} \in C_0\} \subset \{\Pi_F(L_n^{\otimes m}) \in \mathbb{R}^F \setminus A_\varepsilon\}. \tag{2.16}$$

For every $n \geq m$ we want to define a collection $F_n = \{\varphi_n\}_{\varphi \in F}$ of measurable functions such that

$$\Pi_F(L_n^{\otimes m}) = \Pi_{F_n}(L_n^m). \tag{2.17}$$

One possibility is to define, for all $x = (x_1, \dots, x_m) \in S^m$ and $\varphi \in F$,

$$\varphi_n(x) = \sum_{j=1}^m \frac{n_{(j)}}{n^m} \sum_{\tau \in \mathcal{T}_j} \varphi(x_{\tau(1)}, \dots, x_{\tau(m)}), \tag{2.18}$$

where $\mathcal{T}_j$ denotes the set of all surjectiv maps $\tau \colon \{1, \dots, m\} \to \{1, \dots, j\}$ with $\tau(1) = 1$ and $\tau(k) \leq 1 + \max\{\tau(1), \dots, \tau(k-1)\}$ for all $k \in \{2, \dots, m\}$. When checking (2.17), note that, given $j \in \{1, \dots, m\}$ and $(i_1, \dots, i_m) \in \{1, \dots, n\}^m$ consisting of exactly j different components $k_1, \dots, k_j$ which appear in this order, then there exist $(n-j)_{(m-j)} = n_{(m)}/n_{(j)}$ different choices for $(k_{j+1}, \dots, k_m) \in \{1, \dots, n\}^{m-j}$ such that all components of $(k_1, \dots, k_m)$ are different. On the other hand, there exists exactly one $\tau \in \mathcal{T}_j$ such that $(i_1, \dots, i_m) = (k_{\tau(1)}, \dots, k_{\tau(m)})$.

For $k, n \in \mathbb{N}$ satisfying $km \leq n$ and $\sigma \in G_n$, define $L_{k,n,\sigma}$ and $L_{k,n}$ as in the preceding proof. Define the $\mathbb{R}^F$-valued random variable

$$S_{k,n} = \Pi_{F_n}(L_{k,km}) = \frac{1}{k} \sum_{j=0}^{k-1} (\varphi_n(X_{jm+1}, \dots, X_{jm+m}))_{\varphi \in F}. \tag{2.19}$$

We want to show next that the limiting logarithmic moment generating function

$$\Lambda(z) \equiv \lim_{k \to \infty} \frac{1}{k} \log \mathbb{E}[\exp(k\langle z, S_{k,n_k}\rangle)]$$

exists and is finite for every $z = (z_\varphi)_{\varphi \in F} \in \mathbb{R}^F$, does not depend on $\{n_k\}_{k \in \mathbb{N}}$ as long as the sequence satisfies $km \leq n_k$ for every $k \in \mathbb{N}$, and is given by

$$\Lambda(z) = \log \int_{S^m} \exp\biggl(\sum_{\varphi \in F} z_\varphi \varphi(x)\biggr) \mu^{\otimes m}(dx). \tag{2.20}$$

To prove (2.20), note that the k terms in (2.19) are i. i. d. for given n_k, hence

$$\frac{1}{k} \log \mathbb{E}[\exp(k\langle z, S_{k,n_k}\rangle)] = \log \int_{S^m} \exp\biggl(\sum_{\varphi \in F} z_\varphi \varphi_{n_k}(x)\biggr) \mu^{\otimes m}(dx). \tag{2.21}$$

It follows from (2.18) that $\lim_{n \to \infty} \varphi_n(x) = \varphi(x)$ for every $x \in S^m$, because $|\mathcal{T}_m| = 1$ and $|\bigcup_{j=1}^{m-1} \mathcal{T}_j| \leq m^{m-1}$. Furthermore, for every $x \in S^m$,

$$\exp\biggl(\sum_{\varphi \in F} z_\varphi \varphi_n(x)\biggr) \leq \prod_{\varphi \in F} \exp(|z_\varphi|\,|\varphi(x)|) \prod_{j=1}^{m-1} \prod_{\tau \in \mathcal{T}_j} \exp\biggl(\frac{|z_\varphi|}{n} |\varphi \circ \pi_\tau(x)|\biggr).$$

Using Hölder's inequality, Condition 1.4 and Condition 1.5, it follows that the right-hand side is integrable with respect to $\mu^{\otimes m}$ for all sufficiently large n. Therefore, (2.21) and the dominated convergence theorem imply (2.20) for every $z \in \mathbb{R}^F$. Using the upper bound from the Gärtner–Ellis theorem [7, Theorem 2.3.6], it follows that

$$\limsup_{k\to\infty}\frac{1}{k}\log\mathbb{P}(\Pi_{F_{n_k}}(L_{k,km})\in\mathbb{R}^F\setminus A_\varepsilon)\le-\inf_{y\in\mathbb{R}^F\setminus A_\varepsilon}J_F(y) \tag{2.22}$$

for every sequence $\{n_k\}_{k\in\mathbb{N}}$ satisfying $km\le n_k$ for all $k\in\mathbb{N}$. The remaining part of the proof follows along the lines of the preceding proof by using (2.16), (2.17), (2.22) and F_n instead of F. □

3. Proofs concerning the Gibbs conditioning principle

Proof of Theorem 1.18*:* We follow the proof of [7, Theorem 7.3.3].

(a) Note that $A_0\subset F_0$, so $\hat\nu\in\mathcal{E}$ by Condition 1.16. Since F_0 is τ_Φ-closed and $r\equiv J_m(F_0)<\infty$ by Condition 1.16 as well as $\mathcal{E}=F_0\cap K_{m,r}$, part (a) follows from Theorem 1.8(c).

(b) Define $G=(\mathrm{int}_{\tau_\Phi}(\Gamma))^{\mathrm{c}}$. Note that $\Gamma^{\mathrm{c}}\cap A_\delta$ is measurable and contained in the τ_Φ-closed set $G\cap F_\delta$. By the upper bound from Theorem 1.8(d),

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}(L_n^m\in\Gamma^{\mathrm{c}}\cap A_\delta)\le-\lim_{\delta\downarrow 0}J_m(G\cap F_\delta).$$

By Theorem 1.8(c), the rate function J_m is good. Hence, by [7, Lemma 4.1.6(a)], $r\equiv\lim_{\delta\downarrow 0}J_m(G\cap F_\delta)=J_m(G\cap F_0)$. If $r=\infty$, then $r>J_m(F_0)$ by Condition 1.16. If $r<\infty$, then the good rate function J_m attains the infimum r on the τ_Φ-compact set $G\cap F_0\cap K_{m,r+1}$, which is disjoint from $\mathcal{E}$. Hence, $r>J_m(F_0)$, too. To prove that the probability of the condition $\{L_n^m\in A_\delta\}$ decays at a slower rate, i. e.

$$\lim_{\delta\downarrow 0}\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(L_n^m\in A_\delta)\ge-J_m(F_0),$$

just rewrite the proof of Theorem 1.8(a) for $\hat\nu$, using A_δ instead of C and (1.17) instead of (2.7). □

Acknowledgement. We would like to thank Dr. A. Schied for discussions concerning Lemma 2.1(c); see [22, pp. 59–60] and the reference given there.

References

1. A. de Acosta, *On large deviations of empirical measures in the τ-topology*, J. Appl. Probab. **31A** (1994), 41–47, (special volume).
2. R. Azencott, *Grandes Déviations et Applications*, Lectures Notes in Mathematics, vol. 774, Springer-Verlag, Berlin, 1980.
3. R. R. Bahadur and S. Zabell, *Large deviations of the sample mean in general vector spaces*, Ann. Probab. **7** (1979), 587–621.
4. Z. Baranyai, *On the factorisation of the complete uniform hypergraph*, Infinite and Finite Sets (A. Hajnal, T. Rado, and V. T. Sós, eds.), North-Holland, Amsterdam, 1975, pp. 91–108.
5. E. Bolthausen and U. Schmock, *On the maximum entropy principle for uniformly ergodic Markov chains*, Stochastic Process. Appl. **33** (1989), 1–27.
6. I. Csiszár, *Sanov property, generalized I-projection and a conditional limit theorem*, Ann. Probab. **12** (1984), 768–793.
7. A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Jones and Bartlett, Boston, MA, 1993.
8. ______, *Refinements of the Gibbs conditioning principle*, Probab. Theor. Relat. Fields **104** (1996), 1–14.
9. M. Denker, *Asymptotic Distribution Theory in Nonparametric Statistics*, Vieweg, Braunschweig, 1985.
10. J.-D. Deuschel and D. W. Stroock, *Large Deviations*, Academic Press, San Diego, 1989.

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.
12. H. Dym and H. P. McKean, *Fourier Series and Integrals*, Academic Press, New York and London, 1972.
13. P. Eichelsbacher, *Large deviations for products of empirical probability measures in the τ-topology*, preprint, 1995.
14. P. Eichelsbacher and M. Löwe, *Large deviation principle for m-variate von Mises-statistics and U-statistics*, J. Theoret. Probab. **8** (1995), 807–824.
15. S. N. Ethier and T. G. Kurtz, *Markov Processes, Characterization and Convergence*, John Wiley & Sons, New York, 1986.
16. P. Gänssler, *Compactness and sequential compactness in spaces of measures*, Z. Wahrscheinlichkeitstheorie verw. Geb. **17** (1971), 124–146.
17. P. Groeneboom, J. Oosterhoff, and F. H. Ruymgaart, *Large deviation theorems for empirical probability measures*, Ann. Probab. **7** (1979), 553–586.
18. J. L. Kelley, *General Topology*, D. Van Nostrand, New York, 1955.
19. A. J. Lee, *U-Statistics: Theory and Practice*, Marcel Dekker, New York, 1990.
20. J. Neveu, *Discrete-Parameter Martingales*, North-Holland, Amsterdam, 1975.
21. I. N. Sanov, *On the probability of large deviations of random variables*, Selected Transl. Math. Statist. and Prob. **I** (1961), 214–244.
22. A. Schied, *Große Abweichungen für die Pfade der Super-Brownschen Bewegung*, Ph.D. thesis, Universität Bonn, Germany, 1994.
23. R. J. Serfling, *Approximation Theorems of Mathematical Statistics*, John Wiley & Sons, New York, 1980.
24. D. W. Stroock and O. Zeitouni, *Microcanonical distributions, Gibbs states, and the equivalence of ensembles*, Festschrift in Honour of Frank Spitzer (R. Durett and H. Kesten, eds.), Birkhäuser, Basel, 1991, pp. 399–424.
25. S. R. S. Varadhan, *Large Deviations and Applications*, SIAM, Philadelphia, 1984.
26. L. M. Wu, *Some general methods of large deviations and applications*, preprint, 1993.

(Peter Eichelsbacher) Fakultät für Mathematik, Universität Bielefeld, Universitätsstrasse 1, D-33501 Bielefeld, Germany

E-mail address: `peter@mathematik.uni-bielefeld.de`

(Uwe Schmock) Departement Mathematik, ETH-Zentrum, HG G 51.2, CH-8092 Zürich, Switzerland

E-mail address: `schmock@math.ethz.ch`