PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 19, Fasc. 1 (1999), pp. 1–13

LARGE DEVIATIONS FOR EXTREMES OF U-PROCESSES

BY

PETER EICHELSBACHER (BIELEFELD)

Abstract. We prove a large deviations principle (LDP) for partial sums U-processes indexed by the half line. The LDP can be formulated on a suitable subset of the set of all absolutely continuous paths. We endow the space with a topology, which is stronger than the usual topology of uniform convergence on compact intervals. An LDP for the maximum of the sample path of the U-processes is obtained as a particular application.

1. Introduction, statement of the result and applications. A sequence of probability measures $\{\mu_n, n \in N\}$ on a topological space $\mathscr{X}$ equipped with σ-field $\mathscr{B}$ is said to *satisfy the large deviations principle* (LDP) with speed $1/n$ and good rate function $I(\cdot)$ if the level sets $\{x: I(x) \leqslant \alpha\}$ are compact for all $\alpha < \infty$, and for all $\Gamma \in \mathscr{B}$ the lower bound

$$\liminf_{n \to \infty} \frac{1}{n} \log \mu_n(\Gamma) \geqslant - \inf_{x \in \mathrm{int}(\Gamma)} I(x),$$

and the upper bound

$$\limsup_{n \to \infty} \frac{1}{n} \log \mu_n(\Gamma) \leqslant - \inf_{x \in \mathrm{cl}(\Gamma)} I(x)$$

hold. Here $\mathrm{int}(\Gamma)$ and $\mathrm{cl}(\Gamma)$ denote the interior and closure of Γ, respectively. We say that a sequence of random variables *satisfies the LDP* when the sequence of measures induced by these variables satisfies the LDP.

For a sequence of R^d-valued i.i.d. random variables X_i with a finite moment generating function Borovkov [2], Mogulskii [13] and Varadhan [15] investigated the large deviation behaviour of the partial sums process

$$S_n(t) = \frac{1}{n} \sum_{i=1}^{[nt]} X_i, \quad t \in [0, 1],$$

for different scalings. Denote by $L_\infty([0, 1], R^d)$ the space of (equivalence classes modulo equality a.e. of) bounded measurable functions on $[0, 1]$, equipped with the uniform topology. The large deviations principle for $\{S_n(\cdot), n \in N\}$

was established in $L_\infty([0,1], \boldsymbol{R}^d)$ with good rate function $I(\cdot)$ defined by

$$I(\phi) = \int_0^1 \Lambda^*(\dot{\phi})\, dt \tag{1.1}$$

if $\phi \in \mathscr{A}C_0([0,1], \boldsymbol{R}^d)$ and $I(\phi) = \infty$ otherwise (cf. [4], Theorem 5.1). Here $\mathscr{A}C_0([0,1], \boldsymbol{R}^d)$ denotes the subspace of absolutely continuous functions ϕ on $[0,1]$ with $\phi(0) = 0$ and Λ^* denotes the convex dual (Fenchel–Legendre transform) of

$$\Lambda(\theta) = \log \boldsymbol{E} \exp(\langle \theta, X_1 \rangle),$$

that is

$$\Lambda^*(x) := \sup_{\theta \in \boldsymbol{R}^d} \{\langle \theta, x \rangle - \Lambda(\theta)\}.$$

The result can easily be adapted to a time interval $[0, T]$. Applying Theorem 4.6.1 of [4], called the *projective limit approach*, for $T \in \boldsymbol{N}$ yields an LDP for $\{S_n(\cdot), n \in \boldsymbol{N}\}$ in $\mathscr{A}C_0(\boldsymbol{R}_+, \boldsymbol{R}^d)$, the subspace of absolutely continuous functions ϕ on $\boldsymbol{R}_+$ with $\phi(0) = 0$, equipped with the topology of uniform convergence on compact subsets of $\boldsymbol{R}_+$.

Once the LDP with a good rate function is established for a sequence $\{\mu_n, n \in \boldsymbol{N}\}$, the basic contraction principle (cf. [4], Theorem 4.2.1) yields the LDP for $\{\mu_n \circ f^{-1}, n \in \boldsymbol{N}\}$, where f is any continuous function. Applying this principle to the function f: $\mathscr{A}C_0([0,1], \boldsymbol{R}^d) \to \boldsymbol{R}^d$, defined by

$$f(\phi(\cdot)) := \sup_{t \in [0,1]} \phi(t),$$

we obtain the LDP for the sequence $\{\sup_{t \in [0,1]} S_n(t), n \in \boldsymbol{N}\}$ with a good rate function given by

$$J(y) = \inf\{I(\phi): f(\phi(\cdot)) = y\}, \qquad y \in \boldsymbol{R}^d$$

([4], Theorem 4.2.1), sometimes referred to as the *drawback* rate function. However, the function f_+: $\mathscr{A}C_0(\boldsymbol{R}_+, \boldsymbol{R}^d) \to \boldsymbol{R}^d$, defined by

$$f_+(\phi(\cdot)) := \sup_{t \in \boldsymbol{R}_+} \phi(t), \tag{1.2}$$

is not continuous in the topology of uniform convergence on compact subsets of $\boldsymbol{R}_+$ on any supporting subspace. Dobrushin and Pechersky introduced in [5] and [6] a finer topology (a gauge topology) which allows one to prove an LDP via the contraction principle for typical quantities of interest in queueing theory (the steady-state queue-length at a deterministic buffer with inputs given by a real-valued stationary sequence). In their topology the restriction of the mapping $\phi(\cdot) \mapsto \sup_{t \in \boldsymbol{R}_+} (\phi(t) - t)$ to a subspace of non-decreasing paths $\phi(\cdot)$ with $\lim_{t \to \infty} \phi(t)/t = \mu < 1$ is continuous. In [14] O'Connell strengthened the LDP for the polygonal approximation of $\{S_n(\cdot), n \in \boldsymbol{N}\}$ to the following topology: Consider the set of paths

$$\mathscr{Y} = \bigcap_j \left\{ \phi \in \mathscr{C}(\boldsymbol{R}_+, \boldsymbol{R}^d) \colon \lim_{t \to \infty} \frac{\phi^j(t)}{1+t} \text{ exists} \right\},$$

where $\mathscr{C}(\boldsymbol{R}_+, \boldsymbol{R}^d)$ is the space of continuous functions ϕ on $\boldsymbol{R}_+$ and $\phi = (\phi^1, \ldots, \phi^d)$. Let us equip $\mathscr{Y}$ with the norm

$$\|\phi\|_u = \sup_j \sup_t \left| \frac{\phi^j(t)}{1+t} \right|.$$

Note that $\mathscr{Y}$ can be identified with the Polish space $\mathscr{C}(\boldsymbol{R}_+^*, \boldsymbol{R}^d)$ of continuous functions on the extended (and compactified) real line, equipped with the supremum norm, via the bijective mapping $\phi(t) \mapsto \phi(t)/(1+t)$. In particular, $\mathscr{Y}$ is a Polish space. O'Connell shows that the polygonal approximation of $\{S_n(\cdot), n \in \boldsymbol{N}\}$ satisfies the LDP in $\mathscr{Y}$ with good rate function $I(\cdot)$. This result provides a new tool for looking at large deviations for queueing systems in equilibrium (see [14] and references therein). One of the main advantages of O'Connell's approach is that f_+ is a continuous function with respect to the new topology.

The aim of this paper is to extend the LDP when going from linear statistics to higher order statistics, namely the partial sums U-processes, that is

$$U_n(t) = \frac{1}{\binom{n}{m}} \sum_{C_m^{[nt]}} h(X_{i_1}, \ldots, X_{i_m}), \qquad t \in [0, 1].$$

Here the X_i are i.i.d. random variables and h is a measurable, symmetric $\boldsymbol{R}^d$-valued function, called a *kernel function*, where 'symmetric' means that h is invariant under all permutations of its arguments. C_m^k with $k, m \in \boldsymbol{N}$ denotes the set $\{(i_1, \ldots, i_m) \colon 1 \leqslant i_1 < \ldots < i_m \leqslant k\}$. In Eichelsbacher and Löwe [11] the LDP was proved for $\{U_n(\cdot), n \in \boldsymbol{N}\}$, when $L_\infty([0, 1], \boldsymbol{R}^d)$ is equipped with the uniform topology. By applying the projective limit approach ([4], Theorem 4.6.1) the LDP holds for $\{U_n(\cdot), n \in \boldsymbol{N}\}$ in $\mathscr{A}C_0(\boldsymbol{R}_+, \boldsymbol{R}^d)$, equipped with the topology of uniform convergence on compact subsets of $\boldsymbol{R}_+$. In [10] the LDP for $\{U_n(\cdot), n \in \boldsymbol{N}\}$ was proved in a topology, which is an extension of the uniform topology by Orlicz functionals. The aim of the paper is to prove the LDP for $\{U_n(\cdot), n \in \boldsymbol{N}\}$ in $\mathscr{Y}$.

To this end we give a suitable representation of the rate function $I_\infty(\phi)$ for the LDP of $\{U_n(\cdot), n \in \boldsymbol{N}\}$ for any $\phi \in \mathscr{A}C_0(\boldsymbol{R}_+, \boldsymbol{R}^d)$. Moreover, we prove that the convex dual of the so-called *free energy* of a U-statistics yields a nice lower bound for $I_\infty(\cdot)$. In the main technique used to strengthen the LDP to the topology induced by the norm $\|\cdot\|_u$ the *inverse contraction principle* ([4], Theorem 4.2.4 and Corollary 4.2.6) is applied, by which it suffices to prove exponential tightness in the space $(\mathscr{Y}, \|\cdot\|_u)$. That is to check that for every $\alpha < \infty$ there exists a compact set $K_\alpha \subset \mathscr{Y}$ such that

(1.3) $$\limsup_{n\to\infty} \frac{1}{n} \log P(\tilde{U}_n(\cdot) \in K_\alpha^c) < -\alpha,$$

where $\tilde{U}_n(\cdot)$ is the polygonal approximation of $U_n(\cdot)$, defined in (1.5). To formulate our main result we need some more definitions. Let $\{X_n, n \in \boldsymbol{N}\}$ be a sequence of i.i.d. random variables on a probability space $(\Omega, \mathscr{A}, P)$ with Polish state space S and common law μ. Denote by $\mathscr{S}$ the Borel σ-algebra in S. We denote by $\mathscr{M}^+(S)$ and $\mathscr{M}_k(S)$ the sets of Borel measures on S which are positive and positive having total mass k, respectively. These spaces are equipped with the topology of weak convergence. Recall the definition of the relative entropy $H(\nu \mid \mu)$ of $\nu \in \mathscr{M}_1(S)$ with respect to $\mu \in \mathscr{M}_1(S)$:

$$H(\nu \mid \mu) := \begin{cases} \int_S f \log f \, d\mu & \text{if } \nu \ll \mu \text{ and } f = d\nu/d\mu, \\ +\infty & \text{otherwise.} \end{cases}$$

Denote by $\|\cdot\|$ a norm in $\boldsymbol{R}^d$. Consider the following condition:

Condition 1.1 (the strong Cramér condition):

$$\int_{S^m} \exp(\theta \|h(x)\|)\, \mu^m(dx) < \infty \quad \text{for all } \theta > 0,$$

where μ^m denotes the m-fold product measure. We define

(1.4) $$I_\infty(\phi) = \inf \{ \int_0^\infty H(\dot{\nu} \mid \mu)\, dt,\ \nu \in \mathscr{A}C_0 \cap K_\infty \text{ and } \int_{S^m} h\, d\nu^m(\cdot) = \phi(\cdot) \}$$

if $\phi \in \mathscr{A}C_0(\boldsymbol{R}_+, \boldsymbol{R}^d)$, and $I_\infty(\phi) = \infty$ otherwise. Here

$$K_\infty := \bigcup_{L \geqslant 0} \{\nu(\cdot): \int_0^\infty H(\dot{\nu} \mid \mu)\, dt \leqslant L\},$$

and $\mathscr{A}C_0$ is defined to be the set of all maps ν: $\boldsymbol{R}_+ \to \mathscr{M}^+(S)$ which are absolutely continuous with respect to the variation norm $\|\cdot\|_{\text{var}}$, satisfy $\nu(t) - \nu(s) \in \mathscr{M}_{t-s}(S)$ for all $t \geqslant s$, while $\nu(0) = 0$, and the maps ν have a weak derivative for almost all t. The latter means that for almost every $t \in \boldsymbol{R}_+$ the expression $\langle f, \nu(t+h) - \nu(t) \rangle / h$ converges to a limit $\langle f, \dot{\nu}(t) \rangle$ for every $f \in C_b(S)$, where $\langle \cdot, \cdot \rangle$ denotes the Euclidean scalar product and $C_b(S)$ denotes the set of bounded and continuous $\boldsymbol{R}^d$-valued functions on S (cf. [3]). Using the main result of [11] we obtain

Theorem 1.2. *If Condition* 1.1 *is satisfied, the LDP holds for* $\{U_n(\cdot), n \in \boldsymbol{N}\}$ *in* $L_\infty(\boldsymbol{R}_+, \boldsymbol{R}^d)$, *equipped with the topology of uniform convergence on compact subsets of* $\boldsymbol{R}_+$, *with good rate function* $I_\infty(\cdot)$ *defined in* (1.4).

We consider the polygonal approximation of $U_n(\cdot)$, that is

(1.5) $$\tilde{U}_n(t) := U_n(t) + (nt - [nt]) \frac{1}{\binom{n}{m}} \sum_{C_{m-1}^{[nt]}} h(X_{i_1}, \ldots, X_{i_{m-1}}, X_{[nt]+1}).$$

Note that $\tilde{U}_n(\cdot)$ is continuous. It carries the same information as $U_n(\cdot)$. Moreover, by [11], Lemma 3.2, $\{U_n(\cdot), n \in N\}$ and $\{\tilde{U}_n(\cdot), n \in N\}$ are exponentially equivalent in $L_\infty([0, 1], \boldsymbol{R}^d)$, equipped with the uniform topology. Hence the LDP behaviour is the same with respect to this topology. Analogously to Theorem 1.2 we get:

COROLLARY 1.3. *If Condition* 1.1 *is satisfied, the LDP holds for* $\{\tilde{U}_n(\cdot), n \in N\}$ *in* $L_\infty(\boldsymbol{R}_+, \boldsymbol{R}^d)$, *equipped with the topology of uniform convergence on compact subsets of* $\boldsymbol{R}_+$, *with good rate function* $I_\infty(\cdot)$ *defined in* (1.4).

The main result in this paper is the following:

THEOREM 1.4. *If Condition* 1.1 *is satisfied, then* $\{\tilde{U}_n(\cdot), n \in N\}$ *satisfies the LDP in* $(\mathscr{Y}, \|\cdot\|_u)$, *with good rate function* $I_\infty(\cdot)$ *given by* (1.4).

Remark 1.5 (*V-processes*). Notice that by similar arguments we get the same result for a partial sums V-processes

$$V_n(t) = \frac{1}{n^m} \sum_{1 \leqslant i_1, \ldots, i_m \leqslant [nt]} h(X_{i_1}, \ldots, X_{i_m}), \qquad t \in \boldsymbol{R}_+.$$

We only have to assume in addition a weak Cramér condition for the diagonals: there exists at least one $\theta > 0$ such that

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

for every map τ: $\{1, \ldots, m\} \to \{1, \ldots, m\}$, where π_τ: $S^m \to S^m$ is defined by

$$\pi_\tau(s) = (s_{\tau(1)}, \ldots, s_{\tau(m)}) \quad \text{for every } s = (s_1, \ldots, s_m) \in S^m.$$

We omit the details.

To illustrate how Theorem 1.4 can be applied, we will continue to go through the extremes of the sample path process: consider the function f_+ defined by (1.2). If $E(h) = \mu$, then the LDP holds in the subspace

$$\mathscr{Y}(\mu) := \left\{ \phi \in \mathscr{C}(\boldsymbol{R}_+, \boldsymbol{R}^d) \colon \lim_{t \to \infty} \frac{\phi(t)}{1+t} = \mu \right\}.$$

Now, if $\mu < 1$, the restriction of f_+ to $\mathscr{Y}(\mu)$ is finite and continuous. Hence $\{\sup_{t \in \boldsymbol{R}_+} \tilde{U}_n(t), n \in N\}$ satisfies an LDP with good rate function given by

$$\inf\{I_\infty(\phi) \colon f_+(\phi(\cdot)) = y\}, \qquad y \in \boldsymbol{R}^d.$$

Let the kernel function h be given by $h(x, y) = xy$ (the *sample second moment*) and assume that μ is an arbitrary probability distribution such that Condition 1.1 is fulfilled. Since for $m = 1$, by Corollary 2.7 in [11],

$$\inf\left\{\int_0^\infty H(\dot{\nu} \mid \mu)\, dt, \ \nu \in \mathscr{A}C_0 \cap K_\infty \text{ and } \int_S x \, d\nu(\cdot) = \phi(\cdot)\right\} = \int_0^\infty \Lambda^*(\dot{\phi})\, dt,$$

the partial sums U-process with kernel h fulfills the LDP with rate function

$$I_\infty(\phi) = \int_0^\infty \Lambda^*\big(\sqrt{\dot\phi(t)}\big)\,dt \quad \text{if } \phi \geqslant 0 \text{ and } \sqrt{\phi(t)} \in \mathscr{A}C_0(\boldsymbol{R}_+, \boldsymbol{R}^d),$$

and $I_\infty(\phi) = \infty$ otherwise (it follows from the contraction principle). Hence

$$(1.6) \qquad \inf\{I_\infty(\phi)\colon f_+(\phi(\cdot)) = y\} = \inf\Big\{\int_0^\infty \Lambda^*\big(\sqrt{\dot\phi(t)}\big)\colon f_+(\phi(\cdot)) = y\Big\}$$

$$= \inf_{\tau>0}\inf\Big\{\int_0^\tau \Lambda^*\big(\sqrt{\dot\phi(t)}\big)\colon \phi(\tau) = y\Big\} = \inf_\tau \tau\Lambda^*\Big(\frac{\sqrt{y}}{\tau}\Big).$$

Similar calculations have previously been demonstrated by several authors (cf. [8], [14] and references therein). The calculations heavily depend on the special kernel function. For arbitrary kernel functions h such a nice representation of the rate function is, of course, out of reach.

2. Preliminaries. First we state a representation of the rate function $I_\infty(\cdot)$ defined in (1.4) and deduce a lower bound for $I_\infty(\cdot)$. For $k \in N$ and $0 = t_0 < t_1 < \ldots < t_k < \infty$, set $\varrho_0 = 0$ and define

$$(2.1) \qquad H_k(\varrho \mid \mu) := \sum_{i=1}^k (t_i - t_{i-1})\, H\left(\frac{\varrho_{t_i} - \varrho_{t_{i-1}}}{t_i - t_{i-1}} \,\middle|\, \mu\right)$$

for $\varrho := (\varrho_{t_1}, \ldots, \varrho_{t_k}) \in (\mathscr{M}^+(S))^k$. Since $H(\nu \mid \mu) = +\infty$ for $\nu \notin \mathscr{M}_1(S)$, we may concentrate on those sequences of measures ϱ with $\varrho_{t_j}(S) = t_j$, $1 \leqslant j \leqslant k$. The next result follows from Lemma 2.2 in [10]:

LEMMA 2.1. *Assuming that Condition* 1.1 *holds, we get the representation*

$$(2.2) \qquad I_\infty(\phi) = \sup_{0 \leqslant t_1 < \ldots < t_k < \infty} \inf\Big\{H_k(\varrho \mid \mu)\colon \int h\, d\varrho_{t_j}^m = \phi(t_j),$$
$$1 \leqslant j \leqslant k,\ \varrho \in K_\infty(t)\Big\},$$

where

$$K_\infty(t) := \bigcup_{L \geqslant 0} \{\varrho \in (\mathscr{M}^+(S))^k\colon H(\varrho_{t_j}/t_j \mid \mu) \leqslant L,\ 1 \leqslant j \leqslant k\}$$

and $t = (t_1, \ldots, t_k)$.

Proof. For the sake of completeness we sketch the proof. It is an easy adaptation of the proof of Lemma 2.2 in [10]. Denote the right-hand side of (2.2) by $J(\phi)$. Assume that $J(\phi) < \infty$. We can find a sequence $\varrho = (\varrho_{t_1}, \ldots, \varrho_{t_k})$ such that

$$\varrho \in K_\infty(t) \cap \Big\{\int h\, d\varrho_{t_i}^m = \phi(t_i),\ 1 \leqslant i \leqslant k\Big\} \quad \text{and} \quad H_k(\varrho \mid \mu) \nearrow J(\phi)$$

(to see why, note that $H_k(\varrho \mid \mu)$ is increasing for a nested sequence $\{(t_1, \ldots, t_k)\}_k$). We may assume that the sequence $\{(t_1, \ldots, t_k)\}_k$ is increasing (nested) and the sets $\{(t_1, \ldots, t_k)\}_k$ become dense in $\boldsymbol{R}_+$. Note that

$$H_k(\varrho \mid \mu) = \int_0^\infty H\big(\dot{\tilde\varrho}(s) \mid \mu\big)\,ds,$$

where $\tilde{\varrho}(\cdot)$ denotes the polygonalization of ϱ. There exists a subsequence v_n of these polygonalized measure-valued functions that converges in the topology generated by the uniform (on compact sets) weak convergence on $\boldsymbol{R}_+$ to some measure-valued function v and

$$J(\phi) = \liminf_{n\to\infty} \int_0^\infty H(\dot{v}_n \mid \mu)\, ds \geqslant \int_0^\infty H(\dot{v} \mid \mu)\, ds.$$

Therefore we get $I_\infty(\phi) \leqslant J(\phi)$ if we can show that v satisfies

$$\int h\, dv^m(t) = \phi(t), \qquad t \in \boldsymbol{R}_+.$$

Clearly, $v(\cdot) \in K_\infty$. For each $t \in \bigcup_k \{(t_1, \ldots, t_k)\}$, by truncating h at height L and an application of Lusin's theorem and Tietze's extension theorem (and denoting this map by h_L) we get

$$\left\|\int h\, dv^m(t) - \phi(t)\right\| \leqslant \left\|\int h\, dv^m(t) - \int h_L\, dv^m(t)\right\| + \left\|\int h_L\, dv^m(t) - \int h_L\, dv_n^m(t)\right\| \leqslant 2\varepsilon$$

by the weak convergence of the $v_n(t)$ to $v(t)$. For any other $s \in \boldsymbol{R}_+$ we obtain for $t \in \boldsymbol{R}_+$

$$\begin{aligned}\left\|\int h\, dv^m(s) - \phi(s)\right\| \leqslant{}& \left\|\int h\, dv^m(s) - \int h_L\, dv^m(s)\right\| + \left\|\int h_L\, dv^m(s) - \int h_L\, dv^m(t)\right\| \\ &+ \left\|\int h_L\, dv^m(t) - \int h_L\, dv_n^m(t)\right\| + \left\|\int h_L\, dv_n^m(t) - \int h\, dv_n^m(t)\right\| \\ &+ \left\|\int h\, dv_n^m(t) - \phi(t)\right\| + \|\phi(t) - \phi(s)\|.\end{aligned}$$

Therefore, first choosing t in some set $\{(t_1, \ldots, t_k)\}$ close enough to s such that the second term on the right-hand side is bounded by a given $\varepsilon > 0$ and $\|\phi(t) - \phi(s)\| < \varepsilon$ (since $v(t)$ weakly converges to $v(s)$ on an appropriate compact interval and ϕ is continuous) and then choosing L large enough bounds the first and fourth terms on the right $\leqslant \varepsilon$. Finally, choose n large enough so that the third term is $\leqslant \varepsilon$ and so that $\int h\, dv_n^m(t)$ is equal to $\phi(t)$. Putting this together allows us to bound $\left\|\int h\, dv^m(s) - \phi(s)\right\|$ by 6ε. Note that, for any $(t_1, \ldots, t_k)$, with $\varrho = (\varrho_{t_1}, \ldots, \varrho_{t_k})$,

$$\inf\left\{H_k(\varrho \mid \mu)\colon \int h\varrho_{t_j}^m = \phi(t_j),\ 1 \leqslant j \leqslant k,\ \varrho \in K_\infty(t)\right\} \leqslant H_k(\varrho \mid \mu) \leqslant \int_0^\infty H(\dot{\varrho} \mid \mu)\, ds.$$

Therefore, $I_\infty(\phi) = J(\phi)$ holds in the case $J(\phi) < \infty$. In the case $J(\phi) = \infty$, the set on the right-hand side in (2.2) over which the infimum is taken must be empty, for otherwise $J(\phi) < \infty$. Therefore, (2.2) holds also in this case. ∎

A conclusion of the LDP result for U-statistics is the existence of the free energy function of $U_n(1)$, defined by $C_U(x) := \lim_{n\to\infty} C_{U,n}(x)$, where

$$C_{U,n}(x) := \frac{1}{n} \log \boldsymbol{E}_{\mu^n}\big(\exp(n \langle x,\, U_n(1)\rangle)\big)$$

(cf. Corollary 5.8 in [9]). Moreover, using Theorem 4.5.10 in [4], we obtain

$$C_U(x) = \sup_{y \in \mathbf{R}^d} \{\langle x, y\rangle - I_{U_n(1)}(y)\}$$

for $x \in \mathbf{R}^d$, where $I_{U_n(1)}(\cdot)$ denotes the LDP rate function of the U-statistic $U_n(1)$. Set

$$C^*(x) := \sup_{y \in \mathbf{R}^d} \{\langle y, x\rangle - C_U(y)\}, \qquad x \in \mathbf{R}^d.$$

Then

$$C^*(\cdot) \leqslant I_{U_n(1)}(\cdot).$$

C^* is called the *affine regularization* (cf. [9], Corollary 5.8). The following observation follows easily from Lemma 2.4 in [10] and our Lemma 2.1. We omit the proof.

LEMMA 2.2. *Assuming that Condition* 1.1 *holds, we get the following lower bound for* $I_\infty(\cdot)$:

$$I_\infty(\phi) \geqslant \int_0^\infty C^*(\dot{\phi})\,dt.$$

The drawback rate function in the contraction principle has the following property:

LEMMA 2.3. *Let* $\{\mu_n, n \in \mathbf{N}\}$ *be a sequence of probability measures on a metric space* $\mathscr{X}$ *satisfying the LDP with good rate function* $I(\cdot)$. *Let* $f\colon \mathscr{X} \to \mathscr{Y}$ *be a continuous map into another metric space* $\mathscr{Y}$. *Then the good rate function of the sequence* $\{\mu_n \circ f^{-1}, n \in \mathbf{N}\}$ *defined by*

$$J(y) := \inf\{I(x)\colon f(x) = y\}, \qquad y \in \mathscr{Y},$$

has the following property: *If* $I(x) = 0$ *iff* $x = \bar{x}$, *then*

$$J(y) = 0 \Leftrightarrow y = \bar{y} := f(\bar{x}).$$

Proof. The proof is a simple adaptation of the proof of Theorem 5.2 (ii) in [7]. If $I(\bar{x}) = 0$, it follows that $J(\bar{y}) = 0$. Conversely, if $J(y) = 0$, there exists a sequence $\{x_n, n \in \mathbf{N}\}$ in $\mathscr{X}$ such that $f(x_n) = y$ for each $n \in \mathbf{N}$ and $I(x_n) \to 0$ for $n \to \infty$. Hence the sequence $\{x_n, n \in \mathbf{N}\}$ is in one of the level sets of $I(\cdot)$. Thus there exists a subsequence $\{x_{n_l}, l \in \mathbf{N}\}$ such that $\lim_{l\to\infty} x_{n_l} = x$ and, by the lower semicontinuity of $I(\cdot)$, we have $I(x) = 0$. Using the continuity of f, we obtain $f(x) = y$. But, by assumption, $I(x) = 0$ iff $x = \bar{x}$, and thus $y = f(\bar{x}) = \bar{y}$. ∎

Moreover, main ingredients in our proofs are the following results established in [12], Section 5, and in [4], Corollary 4.2.6:

LEMMA 2.4 (Hoeffding). *For* $s > 0$

$$\mathbf{E}_P\big(\exp(s\,\|U_n(1)\|)\big) \leqslant \mathbf{E}_P\Big(\exp\Big(\frac{s}{k}\,\|h\|\Big)\Big)^k, \quad \textit{where } k := \Big\lfloor \frac{n}{m} \Big\rfloor.$$

LEMMA 2.5 (the inverse contraction principle). *Let $\{\mu_n,\ n\in N\}$ be an exponentially tight sequence of probability measures on a topological space $\mathscr{X}$ equipped with the topology τ_1, that is, the sequence satisfies* (1.3) *for every $\alpha < \infty$. If $\{\mu_n,\ n\in N\}$ satisfies an LDP with respect to a Hausdorff topology τ_2 on $\mathscr{X}$ that is coarser than τ_1, then the same LDP holds with respect to the topology τ_1.*

3. Proof of the results.

Proof of Theorem 1.2. By Theorem 4.6.1 in [4], the LDP follows immediately. We only have to check the representation (1.4) of the good rate function. Lemma 2.2 in [10] yields that the rate function for the LDP of $\{U_n(\cdot),\ n\in N\}$ in $L_\infty([0,\ T],\ \boldsymbol{R}^d)$ for $T>0$ fixed has the representation

$$\sup_{0\leqslant t_1<\ldots<t_k\leqslant T} \inf\{H_k(\varrho\mid\mu)\colon \int h\,d\varrho_{t_j}^m = \phi(t_j),\ 1\leqslant j\leqslant k,\ \varrho\in K_\infty(t)\}.$$

Hence applying (4.6.2) in [4] yields that the right-hand side of (2.2) is the rate function for the LDP of $\{U_n(\cdot),\ n\in N\}$ in $L_\infty(\boldsymbol{R}_+,\ \boldsymbol{R}^d)$, endowed with the projective limit topology. Now we apply Lemma 2.1 and the proof is given. ∎

Proof of Corollary 1.3. By Lemma 3.2 in [11], $\{U_n(\cdot),\ n\in N\}$ and $\{\tilde{U}_n(\cdot),\ n\in N\}$ are exponentially equivalent in $L_\infty([0,\ 1],\ \boldsymbol{R}^d)$, equipped with the uniform topology. The same proof works on $L_\infty([0,\ T],\ \boldsymbol{R}^d)$ for a fixed $T>0$. Therefore applying Theorem 4.6.1 in [4] yields the LDP for $\{\tilde{U}_n(\cdot),\ n\in N\}$ in $L_\infty(\boldsymbol{R}_+,\ \boldsymbol{R}^d)$, equipped with the topology of uniform convergence on compact subsets $\boldsymbol{R}_+$. The rate function is the right-hand side of (2.2). Hence applying Lemma 2.1 gives the result. ∎

Proof of Theorem 1.4. Denote by $\mathscr{D}(I_\infty) := \{\phi\colon I_\infty(\phi)<\infty\}$ the effective domain of $I_\infty(\cdot)$. First we prove that $\mathscr{D}(I_\infty)\subset\mathscr{Y}$. By the definition of $\mathscr{Y}$ we can consider the $\boldsymbol{R}$-valued case. Moreover, by considering $\tilde{U}_n(t)-\boldsymbol{E}(\tilde{U}_n(t)) \approx \tilde{U}_n(t)-t^m\boldsymbol{E}(h)$ we can, without loss of generality, assume that $\boldsymbol{E}(h)=0$ (using the contraction principle). Let $\phi(\cdot)$ be chosen such that $I_\infty(\phi)\leqslant\alpha$. By convexity of C^* we infer from the proof of Lemma 5.1.6, (5.1.11), in [4] that

$$\int_0^\infty C^*(\dot{\phi})\,dt = \sup_{0=t_0<t_1<t_2<\ldots<t_k<\infty} \sum_{i=1}^{k} (t_i-t_{i-1})\,C^*\left(\frac{\phi(t_i)-\phi(t_{i-1})}{t_i-t_{i-1}}\right) \tag{3.1}$$

holds. Therefore, applying Jensen's inequality and Lemma 2.2 yields

$$tC^*(\phi(t)/t)\leqslant\int_0^\infty C^*(\dot{\phi})\,dt\leqslant I_\infty(\phi)\leqslant\alpha \quad \text{for each } t>0.$$

If Condition 1.1 is fulfilled, $C_U(x)$ is finite for every $x\in\boldsymbol{R}$, and therefore we observe that, for all $x\in\boldsymbol{R}$ and all $\varrho>0$,

$$C^*(x)\geqslant\varrho|x|-\sup_{|\lambda|=\varrho}\{C_U(\lambda)\}.$$

Consequently,

$$\lim_{t\to\infty}\left|\frac{\phi(t)}{1+t}\right| \leqslant \lim_{t\to\infty}\frac{\alpha}{\varrho t}+\frac{1}{\varrho}\sup_{|\lambda|=\varrho}\{C_U(\lambda)\}.$$

Applying Lemma 2.3 to the rate function $I_{U_n(1)}(\cdot)$ for the sequence $\{U_n(1), n\in N\}$ (see [9], Theorem 2 and its proof), we obtain $I_{U_n(1)}(E(h))=0$. Thus $C_U(x)\leqslant xE(h)$, and by our assumption we obtain $\mathscr{D}(I_\infty)\subset\mathscr{Y}$.

Next we want to check that $P(\tilde{U}_n\in\mathscr{Y})=1$: this follows by using (1.5) and the LIL for U-statistics (see, for example, [1], Corollary 3.5).

Now, by the Dawson–Gärtner theorem for projective limits and by Lemma 4.1.5 in [4] it follows that $\{\tilde{U}_n(\cdot), n\in N\}$ satisfies the LDP in $\mathscr{Y}$ when equipped with the topology of uniform convergence on compact intervals. To strengthen this to the topology induced by the norm $\|\cdot\|_u$, we use the inverse contraction principle, by which it suffices to prove exponential tightness in the space $(\mathscr{Y}, \|\cdot\|_u)$. For each t, denote by $\mathscr{C}([0, t], \boldsymbol{R}^d)$ the projection of $\mathscr{C}(\boldsymbol{R}_+, \boldsymbol{R}^d)$ onto the interval $[0, t]$, equipped with the uniform topology, and by $\phi[0, t](\cdot)$ for $\phi\in\mathscr{C}(\boldsymbol{R}_+, \boldsymbol{R}^d)$ denote the restriction of ϕ to the interval $[0, t]$. Goodness of the rate function $I_\infty(\cdot)$ implies that the sequence $\{\tilde{U}_n[0, 1](\cdot), n\in N\}$ is exponentially tight in the uniform topology on $\mathscr{A}C([0, 1], \boldsymbol{R}^d)$. In other words, for each $\alpha>0$ there exists a compact set K_α in $\mathscr{A}C([0, 1], \boldsymbol{R}^d)$ such that

$$\limsup_{n\to\infty}\frac{1}{n}\log P(\tilde{U}_n[0, 1](\cdot)\notin K_\alpha)\leqslant -\alpha.$$

It follows that for each $t>0$

$$K_\alpha(t):=\{\phi\in\mathscr{C}([0, t], \boldsymbol{R}^d): \{s\mapsto\phi(st)\}\in K_\alpha\}$$

is compact in $\mathscr{C}([0, t], \boldsymbol{R}^d)$, and for each $0<\varepsilon<\alpha$

$$\begin{aligned}(3.2)\qquad \limsup_{n\to\infty}\frac{1}{n}\log P\Big(\bigcup_{t>1}\{\tilde{U}_n[0, t](\cdot)\notin K_\alpha(t)\}\Big) &\leqslant \limsup_{n\to\infty}\frac{1}{n}\log P\Big(\bigcup_{t>1}\{\tilde{U}_{nt}[0, 1](\cdot)\notin K_\alpha(1)\}\Big)\\ &\leqslant \limsup_{n\to\infty}\frac{1}{n}\log\sum_{k>n}e^{-(\alpha-\varepsilon)k}\leqslant -\alpha+\varepsilon.\end{aligned}$$

Since ε is arbitrary, for each $\alpha>0$ we have

$$(3.3)\qquad \limsup_{n\to\infty}\frac{1}{n}\log P\Big(\bigcup_{t>1}\{\tilde{U}_n[0, t](\cdot)\notin K_\alpha(t)\}\Big)\leqslant -\alpha.$$

For $\alpha, t>0$, define

$$d_\alpha(t)=\begin{cases}\alpha^2 & \text{for } t\leqslant\alpha^2,\\ t^{-1/2} & \text{for } t>\alpha^2,\end{cases}$$

and consider the sets

$$D_\alpha := \bigcap_{j=1}^{d} \left\{ \phi \in \mathscr{Y} : \left| \frac{\phi^j(t)}{\beta(t)} \right| \leqslant d_\alpha(t) \text{ for all } t,\ \phi[0,t](\cdot) \in K_\alpha(t) \text{ for all } t > 1 \right\}.$$

The exponential tightness of $\{\tilde{U}_n(\cdot),\ n \in N\}$ in $(\mathscr{Y}, \|\cdot\|_u)$ will be established by the following two lemmas. The first one is exactly Lemma 1 in [14]. For the sake of completeness we give the proof.

LEMMA 3.1. *For each $\alpha > 0$, D_α is compact in $(\mathscr{Y}, \|\cdot\|_u)$.*

Proof. Let ϕ_n be a sequence in D_α. By Tychonoff's theorem, the set $\bigcap_{t>1} K_\alpha(t)$ is compact in $\mathscr{Y}$ when equipped with the topology of uniform convergence on compact intervals, so there exists a subsequence $n(k)$ such that $\phi_{n(k)}$ converges to some $\phi \in \bigcap_{t>1} K_\alpha(t)$ in this topology. It follows that, for each $T > 0$ and for each j,

$$\lim_{k\to\infty} \sup_{t \leqslant T} \left| \frac{\phi^j_{n(k)}(t)}{1+t} - \frac{\phi^j(t)}{1+t} \right| = 0.$$

Note that this implies, for each t and j,

$$\left| \frac{\phi^j(t)}{1+t} \right| \leqslant d_\alpha(t),$$

and so $\phi \in D_\alpha$. Now for each $\varepsilon > 0$ (sufficiently small), we have, for k sufficiently large,

$$(3.4) \qquad \|\phi_{n(k)} - \phi\|_u \leqslant \sup_{t \leqslant 1/\varepsilon^2} \left| \frac{\phi^j_{n(k)}(t)}{1+t} - \frac{\phi^j(t)}{1+t} \right| + \sup_{t > 1/\varepsilon^2} \left| \frac{\phi^j_{n(k)}(t)}{1+t} - \frac{\phi^j(t)}{1+t} \right|$$
$$\leqslant \varepsilon + 2d_\alpha(1/\varepsilon^2) = 3\varepsilon.$$

The set D_α is therefore sequentially compact, and hence compact, in $(\mathscr{Y}, \|\cdot\|_u)$. ∎

LEMMA 3.2. *If the assumption of Theorem 1.4 is satisfied, then*

$$\lim_{\alpha\to\infty} \limsup_{n\to\infty} \frac{1}{n} \log P\big(\tilde{U}_n(\cdot) \notin D_\alpha\big) = -\infty.$$

Proof. The proof is an adaptation of the proof of Lemma 2 in [14]. Denote by $U^j_n(\cdot)$ the j-th coordinate function, $j \in \{1, \ldots, d\}$. For some $\theta > 0$ and for each j we have

$$(3.5) \quad P\Big(\bigcup_{t \leqslant \alpha^2} \{|\tilde{U}^j_n(t)| > \alpha^2(1+t)\}\Big) \leqslant P\Big(\bigcup_{i=0}^{n-1} \bigcup_{k=0}^{[\alpha^2]} \{|\tilde{U}^j_n(k+i/n)| > (1+k)\alpha^2\}\Big)$$
$$\leqslant n \sum_{k=0}^{[\alpha^2]} C(\theta) \exp\big(-\theta n(k+1)\alpha^2\big) \leqslant nC(\theta)(\alpha^2+1)\exp(-\theta n\alpha^2).$$

Here we have used Chebyshev's inequality, Condition 1.1 as well as Lemma 2.4:

$$(3.6)\quad \begin{aligned} &\boldsymbol{P}\big(|U^j_{nk+i}(1)| > (1+k)\,\alpha^2\big) \\ &\leqslant \exp\big(-\theta n(k+1)\,\alpha^2\big)\,\boldsymbol{E}\big(\exp\big(\theta n(k+1)\,|U^j_{nk+i}(1)|\big)\big) \\ &\leqslant \exp\big(-\theta n(k+1)\,\alpha^2\big)\,\boldsymbol{E}\big(\exp(\theta m\,\|h\|)\big)^{n/m} \leqslant C(\theta)^{n/m}\exp\big(-\theta n(k+1)\,\alpha^2\big). \end{aligned}$$

It follows that

$$(3.7)\qquad \limsup_{n\to\infty}\frac{1}{n}\log\boldsymbol{P}\big(\bigcup_{t\leqslant\alpha^2}\{|\tilde{U}^j_n(t)| > (1+t)\,\alpha^2\}\big) \leqslant -\theta\alpha^2.$$

We also have, for each j and some $\theta > 0$,

$$(3.8)\quad \begin{aligned} &\boldsymbol{P}\big(\bigcup_{t>\alpha^2}\{|\tilde{U}^j_n(t)| > (1+t)\,d_\alpha(t)\}\big) \\ &\leqslant \boldsymbol{P}\big(\bigcup_{i=0}^{n-1}\bigcup_{k=[\alpha^2]}^{\infty}\{|\tilde{U}^j_n(k+i/n)| > (1+k)\,d_\alpha(k)\}\big) \\ &\leqslant n\sum_{k=[\alpha^2]}^{\infty} C(\theta)^{n/m}\exp\big(-\theta n(1+k)\,d_\alpha(k)\big) \leqslant nC(\theta)^{n/m}D\exp\big(-\theta n\sqrt{\alpha^2-1}/2\big). \end{aligned}$$

Again we have used Chebyshev's inequality, Condition 1.1 and Lemma 2.4. Moreover, we have used the inequality

$$\sum_{k\geqslant k_0}\exp\big(-\varrho\sqrt{k}\big) \leqslant D\exp\big(-\varrho\sqrt{k_0-1}/2\big).$$

It follows that

$$(3.9)\qquad \limsup_{n\to\infty}\frac{1}{n}\log\boldsymbol{P}\big(\bigcup_{t>\alpha^2}\{|\tilde{U}^j_n(t)| > (1+t)\,d_\alpha(t)\}\big) \leqslant -\theta\sqrt{\alpha^2-1}/2.$$

The statement can now be obtained from (3.3), (3.7) and (3.9), via the principle of the largest term. ▪

This concludes the proof of Theorem 1.4. ▪

REFERENCES

[1] M. A. Arcones and E. Giné, *U-processes indexed by Vapnik–Červonenkis classes of functions with applications to asymptotics and bootstrap of U-statistics with estimated parameters*, Stochastic Process. Appl. 52 (1994), pp. 17–38.

[2] A. A. Borovkov, *Boundary-value problems for random walks and large deviations in function spaces*, Theory Probab. Appl. 12 (1967), pp. 575–595.

[3] A. Dembo and T. Zajic, *Large deviations: From empirical mean and measure to partial sums process*, Stochastic Process. Appl. 57 (1995), pp. 191–224.

[4] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Springer, New York 1998.

[5] R. L. Dobrushin and E. A. Pechersky, *Large deviations for tandem queueing systems*, J. Appl. Math. Stochastic Anal. 7 (1994), pp. 301–330.

[6] — *Large deviations for random processes with independent increments on infinite intervals*, preprint, 1995.

[7] 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), pp. 389–461.

[8] N. G. Duffield and N. O'Connell, *Large deviations and overflow probabilities for the general single-server queue, with applications*, Math. Proc. Cambridge Philos. Soc. 118 (1995), pp. 363–374.

[9] P. Eichelsbacher, *Large deviations for products of empirical probability measures in the τ-topology*, J. Theoret. Probab. 10, No. 4 (1997), pp. 903–920.

[10] — *Partial sums U-processes large deviations in stronger topologies*, Statist. Probab. Lett. 38 (1998), pp. 207–214.

[11] — and M. Löwe, *Large deviations for partial sums U-processes*, Theory Probab. Appl. 43 (1998), pp. 97–115.

[12] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc. 58 (1963), pp. 13–30.

[13] A. A. Mogulskii, *Large deviations for trajectories of multidimensional random walks*, Theory Probab. Appl. 21 (1976), pp. 300–315.

[14] N. O'Connell, *A large deviations principle with queueing applications*, preprint, 1997.

[15] S. R. S. Varadhan, *Asymptotic probabilities and differential equations*, Comm. Pure Appl. Math. 19 (1966), pp. 261–286.

Fakultät für Mathematik, Universität Bielefeld
Postfach 10-0131, D-33501 Bielefeld, Germany
E-mail address, Peter Eichelsbacher: peter@mathematik.uni-bielefeld.de

Received on 16.7.1998