PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 20, Fasc. 2 (2000), pp. 273–285

LARGE DEVIATION PRINCIPLE FOR SET-VALUED UNION PROCESSES

BY

TOMASZ SCHREIBER* (TORUŃ)

Abstract. The purpose of the paper is to establish a large deviation principle for a certain class of increasing set-valued processes obeying Markovian dynamics. The obtained result is then applied to investigate the asymptotics of the sequence of successive convex hulls generated by uniform samples on a d-dimensional ball.

Key words: large deviations, random closed sets, Markov process, convex hull.

1. INTRODUCTION AND MAIN RESULTS

Consider a compact metric space E and denote by $\mathscr{F}(E)$ the family of all its closed (and hence compact) subsets. We endow $\mathscr{F}(E)$ with the topology induced by the Hausdorff distance ρ_E. It is known that the resulting space $(\mathscr{F}(E), \rho_E)$ is compact (see Chapter 1 in [11]). It is convenient to assume that $\emptyset \in \mathscr{F}(E)$ and to set $\rho_E(\emptyset, A) = 1$ for all nonempty $A \in \mathscr{F}(E)$.

The random elements taking values in $(\mathscr{F}(E), \mathscr{B}_E)$, where $\mathscr{B}_E$ is the Borel σ-field corresponding to ρ_E, will be referred to as random closed sets (for extensive reference see [8], [11] or [12]).

Let $\mathscr{D}$ be a certain subclass of $\mathscr{F}(E)$ closed with respect to finite unions and limits in ρ_E, i.e.

(K1) if $A, B \in \mathscr{D}$, then also $A \cup B \in \mathscr{D}$;

(K2) if $A_1, A_2, \ldots \in \mathscr{D}$ and $\lim_{n\to\infty} \rho_E(A_n, A) = 0$, then $A \in \mathscr{D}$.

In particular, we conclude from (K2) that $(\mathscr{D}, \rho_E)$ is compact.

In this paper we investigate a general class of growing $\mathscr{D}$-valued processes which can be represented as successive unions of random closed sets obeying Markovian dynamics in the following sense. Let $\pi(\cdot|\cdot)$ be a certain stochastic kernel on $\mathscr{D}$ given $\mathscr{D}$, i.e. π is required to be a measurable mapping from $\mathscr{D}$ to the space $\mathscr{P}(\mathscr{D})$ of all the Borel probability measures on $\mathscr{D}$ endowed with the

* Nicolas Copernicus University, Toruń.

usual weak topology. For each $H_0 \in \mathscr{D}$, on a probability space $(\Omega, \mathfrak{F}, \boldsymbol{P}_{H_0})$, construct recursively the sequence of random closed sets $(Z_i)_{i=0}^{\infty}$ taking

$$\boldsymbol{P}_{H_0}(Z_0 = H_0) = 1$$

and

$$\boldsymbol{P}_{H_0}(Z_{j+1} \in \mathscr{E} \mid Z_0, \ldots, Z_j) = \pi\big(\mathscr{E} \mid \bigcup_{i=0}^{j} Z_i\big) \tag{1}$$

for $j \geqslant 0$ and $\mathscr{E} \subseteq \mathscr{D}$. Note that the sequence $\big(\bigcup_{i=0}^{n} Z_i\big)_{n=0}^{\infty}$ is a Markov chain. For each $n \in N$ define the piecewise constant $\mathscr{D}$-valued random process $(X_t^n)_{0 \leqslant t \leqslant 1}$ as follows:

$$X_t^n := \bigcup_{j \leqslant nt} Z_j. \tag{2}$$

We shall call this process the *union process* associated with $Z_0, \ldots, Z_n$. The purpose of the paper is to establish and prove the large deviation principle for the sequence X^n.

It is convenient to consider the union processes as random elements taking values in the space $\mathscr{U} = \mathscr{U}(\mathscr{D})$ of all nondecreasing (with respect to inclusion) right continuous $\mathscr{D}$-valued functions defined on $[0, 1]$. Identifying each function $U \in \mathscr{U}$ with the closed set

$$\Gamma(U) = \{(x, t) \in E \times [0, 1] \mid x \in U(t)\}$$

we construct an embedding Γ of $\mathscr{U}$ onto a closed (and hence compact) subset of the compact space $(\mathscr{F}(E \times [0, 1]), \rho_{(E \times [0,1])})$. Let us endow $\mathscr{U}$ with the following metric ϱ induced by this embedding:

$$\varrho(U, V) = \rho_{(E \times [0,1])}\big(\Gamma(U), \Gamma(V)\big). \tag{3}$$

Clearly, the resulting space $(\mathscr{U}, \varrho)$ is compact.

We impose on the transition kernel $\pi(\cdot \mid \cdot)$ the regularity conditions given in the sequel. For the notational convenience let us agree to write '$\pi(Z$ satisfies $\mathscr{R} \mid A)$' instead of '$\pi(\{C \in \mathscr{D} \mid C$ satisfies $\mathscr{R}\} \mid A)$', where $A \in \mathscr{D}$ and $\mathscr{R}$ is a certain property.

(C1) If $\lim_{n \to \infty} \rho_E(A_n, A) = 0$ for $A, A_1, A_2, \ldots \in \mathscr{D}$, then for each closed family $\mathscr{E} \subseteq \mathscr{D}$ we have $\lim_{n \to \infty} \pi(\mathscr{E} \mid A_n) = \pi(\mathscr{E} \mid A)$.

(C2) Let $A, B \in \mathscr{D}$, $A \subseteq B$ and suppose that $\pi(Z \subseteq B \mid B) > 0$. Then for each $\varepsilon > 0$ there exists $m \geqslant 0$ such that

$$\inf\Big\{\boldsymbol{P}_C\Big(\rho_E\Big(\bigcup_{i=0}^{m} Z_i, B\Big) \leqslant \varepsilon\Big) \mid C \in \mathscr{D}, \rho_E(C, A) \leqslant \varepsilon\Big\} > 0.$$

Roughly speaking, condition (C2) requires that the process of successive unions $\bigcup_{i=0}^{n} Z_i$ starting from the neighbourhood of some $A \in \mathscr{D}$ reaches with positive probability the appropriate neighbourhood of B for any $B \in \mathscr{D}$ containing

A and stable in the sense that $\pi(Z \subseteq B \mid B) > 0$. Note the unidirectional character of (C2) (transitions are allowed from subsets to supersets only) due to the monotonicity of the successive unions process.

For each $H_0 \in \mathscr{D}$ we define the *rate function* $I_{H_0}: \mathscr{U} \to \boldsymbol{R}_+ \cup \{\infty\}$ by

$$I_{H_0}(U) := -\int_{[0,1]} \log \pi\big((Z \cup H_0) \subseteq U(t) \mid U(t)\big)\, dt. \tag{4}$$

In Lemma 1 we show that the name *rate function* used for I is justified in the sense of the following definition (see Section 1.1 in [4] or the definition of a good rate function in Chapter 2 of [3]).

DEFINITION 1. A nonnegative function J defined on a Polish metric space $\mathscr{X}$ is called a *rate function* if its level sets $\{x \in \mathscr{X} \mid J(x) \leqslant M\}$ are compact for $0 \leqslant M < \infty$.

Note that in particular each rate function is lower semicontinuous. Further, in view of the compactness of $\mathscr{U}$, each lower semicontinuous function on $\mathscr{U}$ is a rate function.

The main result of the paper is

THEOREM 1. *Under conditions* (C1) *and* (C2) *the sequence X^n with the initial condition $Z_0 = H_0$ satisfies the large deviation principle on $\mathscr{U}$ with the rate function I_{H_0}, i.e. for each open set $\mathscr{G} \subseteq \mathscr{F}$*

$$\liminf_{n\to\infty} \frac{1}{n} \log \boldsymbol{P}_{H_0}(X^n \in \mathscr{G}) \geqslant -\inf_{U \in \mathscr{G}} I_{H_0}(U)$$

and for each closed set $\mathscr{H} \subseteq \mathscr{F}$

$$\limsup_{n\to\infty} \frac{1}{n} \log \boldsymbol{P}_{H_0}(X^n \in \mathscr{H}) \leqslant -\inf_{U \in \mathscr{H}} I_{H_0}(U).$$

The next section contains the proof of the above statements. In the final section, as an example of application, we use Theorem 1 to prove the large deviation principle for sequences of successive convex hulls of uniform samples on a d-dimensional ball (see Theorem 2).

It is to be emphasised that all the definitions and results presented above can be easily extended to a more general case with the setting space E which is not necessarily compact, but is metrisable, separable and locally compact. The topology induced by the Hausdorff metric is then to be replaced by the vague (Fell) topology on $\mathscr{F}(E)$ (for the definition see Chapter 1 in [11] or Section 1.1 in [12]). However, since the resulting topological space can be embedded in a natural way into the space $\mathscr{F}(\bar{E})$ of closed subsets of the one-point compactification $\bar{E}$ of E (see the remark on Theorem 1.4.1 in [11]), we have decided to confine ourselves to the apparently less general case of compact E, which allows us to simplify the presentation of certain details of the proofs due to the convenient form of the Hausdorff metric.

2. PROOFS

2.1. Lower semicontinuity of I_{H_0}. To justify the name *rate function* used for I_{H_0} in (4) we prove the following lemma:

LEMMA 1. *For $H_0 \in \mathscr{D}$ the function I_{H_0} given by* (4) *is lower semicontinuous.*

Proof. We shall show that the mapping $\mathscr{F} \ni A \mapsto \pi\big((Z \cup H_0) \subseteq A \mid A\big)$ is upper semicontinuous, i.e. if $\lim_{n\to\infty} \rho_E(A_n, A) = 0$, then

$$\limsup_{n\to\infty} \pi\big((Z \cup H_0) \subseteq A_n \mid A_n\big) \leqslant \pi\big((Z \cup H_0) \subseteq A \mid A\big). \tag{5}$$

From condition (C1) we conclude in particular that the sequence of probability measures $\pi(\cdot \mid A_n)$ converges weakly to $\pi(\cdot \mid A)$. Applying Skorokhod's representation theorem we construct random sets $\Xi_1, \Xi_2, \ldots$ and Ξ distributed according to $\pi(\cdot \mid A_1), \pi(\cdot \mid A_2), \ldots$ and $\pi(\cdot \mid A)$, respectively, and such that almost surely $\lim_{n\to\infty} \rho_E(\Xi_n, \Xi) = 0$, and hence $\lim_{n\to\infty} \rho_E(\Xi_n \cup H_0, \Xi \cup H_0) = 0$. Further, note that we have $(\Xi_n \cup H_0) \nsubseteq A_n$ for n large enough whenever $(\Xi \cup H_0) \nsubseteq A$, so that

$$\limsup_{n\to\infty} \mathbf{1}_{((\Xi_n \cup H_0) \subseteq A_n)} \leqslant \mathbf{1}_{((\Xi \cup H_0) \subseteq A)}.$$

Taking the expectations of both sides and applying Fatou's lemma we obtain (5).

To proceed take an arbitrary sequence $(U_n)_{n=1}^{\infty} \subset \mathscr{U}$ convergent to a certain $U \in \mathscr{U}$ and observe that for each $t \in (0, 1]$ at which U is continuous with respect to ρ_E we have $\lim_{n\to\infty} \rho_E\big(U_n(t), U(t)\big) = 0$, so by (5) we have

$$\limsup_{n\to\infty} \pi\big((Z \cup H_0) \subseteq U_n(t) \mid U_n(t)\big) \leqslant \pi\big((Z \cup H_0) \subseteq U(t) \mid U(t)\big). \tag{6}$$

However, since U is nondecreasing, the number of its discontinuity points is at most countable. Indeed, let $\mathcal{O}$ be a countable open set basis of E. Then each closed set $A \subseteq E$ is uniquely determined by the subfamily $\mathcal{O}_A = \{G \in \mathcal{O} \mid G \cap A \neq \emptyset\}$. Now, let $t_1 \in (0, 1)$ be a discontinuity point of U. Then, obviously, there exists $G_{t_1} \in \mathcal{O}$ such that $G_{t_1} \notin \mathcal{O}_{U(t)}$ for $t < t_1$ and $G_{t_1} \in \mathcal{O}_{U(t)}$ for $t \geqslant t_1$. Further, if t_1 and t_2 are two different discontinuity points, then $G_{t_1} \neq G_{t_2}$. This proves that U can be discontinuous at a countable number of points only.

Thus, we can integrate both sides of (6) over $[0, 1]$ neglecting the discontinuity points and apply Fatou's lemma to get

$$I_{H_0}(U) \leqslant \liminf_{n\to\infty} I_{H_0}(U_n),$$

as required. ∎

2.2. Proof of Theorem 1. The scheme of the proof is the following. First we apply Theorem 1.3.7 of [4], originally due to O'Brien and Verwaat [14] and Pukhalskii [15], formulated below.

PROPOSITION 1. *Let $(n') \subseteq N$ be a certain sequence of natural numbers. If a sequence of random elements $Y^{n'}$ taking values in a Polish metric space $\mathscr{X}$ is exponentially tight, i.e. for each $R > 0$ there exists a compact set $K \subset \mathscr{X}$ such that*

$$\limsup_{n' \to \infty} \frac{1}{n'} \log P(Y^{n'} \notin K) \leqslant -R,$$

then there exists a further subsequence $(n'') \subseteq (n')$ such that $Y^{n''}$ satisfies on $\mathscr{X}$ the large deviation principle with a certain rate function J, i.e. for every open set $\mathscr{G} \subset \mathscr{X}$

$$\liminf_{n'' \to \infty} \frac{1}{n''} \log P(Y^{n''} \in \mathscr{G}) \geqslant -\inf_{x \in \mathscr{X}} J(x)$$

and for each closed set $\mathscr{H} \subset \mathscr{X}$

$$\limsup_{n'' \to \infty} \frac{1}{n''} \log P(Y^{n''} \in \mathscr{H}) \leqslant -\inf_{x \in \mathscr{X}} J(x).$$

To proceed we fix some $H_0 \in \mathscr{D}$ and a subsequence (n'). Note that since $\mathscr{U}$ is compact, it is immediately obvious that $X^{n'}$ is exponentially tight. Therefore there exists a further subsequence $(n'') \subseteq (n')$ and a rate function J_{H_0} on $\mathscr{U}$ such that the sequence $X^{n''}$ under P_{H_0} satisfies the large deviation principle on $\mathscr{U}$ with the rate function J_{H_0}. Since the subsequence (n') was chosen arbitrary, the proof of Theorem 1 will be complete if we succeed to show that $J_{H_0} = I_{H_0}$, where I_{H_0} is defined by (4). We do this in three steps. First, in Lemma 2, we prove that $J_{H_0}(U) \geqslant I_{H_0}(U)$ for all $U \in \mathscr{U}$. Then, in Lemma 3, we show that the converse inequality (and hence equality) holds for all piecewise constant $U \in \mathscr{U}$. Finally, in Lemma 4 we extend the latter result onto the whole $\mathscr{U}$, thus completing the proof.

LEMMA 2. *For each $U \in \mathscr{U}$ we have $J_{H_0}(U) \geqslant I_{H_0}(U)$.*

Proof. Fix some $U \in \mathscr{U}$ and choose an arbitrary $\varepsilon > 0$. We claim that there exists an increasing sequence $0 = t_0 < t_1 < \ldots < t_k = 1$ such that

$$\max_{i=0}^{k-1} \sup_{t,s \in [t_i, t_{i+1})} \rho_E(U(t), U(s)) \leqslant \varepsilon/2. \tag{7}$$

To see this define for each $t \in [0, 1]$

$$\psi(t) := \min\{s > t \mid s = 1 \text{ or } \rho_E(U(s), U(t)) \geqslant \varepsilon/2\}.$$

Note that the correctness of this definition follows from the right continuity of U. Now set

$$t_0 := 0 \quad \text{and} \quad t_{i+1} := \psi(t_i) \text{ for } i \geqslant 0.$$

It remains to show that there exists some k for which $t_k = 1$. If it were not the case, we would have an infinite sequence $U(t_0) \subset U(t_1) \subset U(t_2) \subset \ldots$ with the property that $\rho_E(U(t_i), U(t_j)) \geqslant \varepsilon/2$ for $i \neq j$ (because for $i<j$ we have $U(t_{i+1}) \subseteq U(t_j)$ so that $\rho_E(U(t_i), U(t_j)) \geqslant \rho_E(U(t_i), U(t_{i+1})) \geqslant \varepsilon/2$). Since $\mathscr{D}$ is compact, it cannot happen, so (7) holds.

To proceed take some $\delta > 0$ such that $\delta < \min_{i=0}^{k-1}(t_{i+1}-t_i)/2$ and $\delta < \varepsilon/2$ and define $\Delta(\delta)$ to be the open 'δ-sausage' around U, i.e.

$$\Delta(\delta) := \{V \in \mathscr{U} \mid \varrho(V, U) < \delta\},$$

where ϱ is the metric on $\mathscr{U}$ given by (3). We will investigate the asymptotic behaviour of the quantity

$$L_{n''} := \frac{1}{n''} \log P_{H_0}(X^{n''} \in \Delta(\delta)).$$

Define for $A \in \mathscr{D}$

$$P_A(\varepsilon) := \sup\{\pi(Z \subseteq A^{(\varepsilon)} \mid C) \mid H_0 \subseteq C \in \mathscr{D},\ \rho_E(A, C) < \varepsilon\}$$

with $\sup \emptyset = 0$ and $A^{(\varepsilon)} = \{x \in E \mid \operatorname{dist}(x, A) < \varepsilon\}$. Further, for all $0 \leqslant i < k$ choose arbitrary $\tau_i \in [t_i, t_{i+1})$. Then for sufficiently large n'' we have

$$(8) \quad P_{H_0}(X^{n''} \in \Delta(\delta)) \leqslant (P_{U(\tau_0)}(\varepsilon))^{[(t_1-t_0-2\delta)n'']-1} \ldots (P_{U(\tau_{k-1})}(\varepsilon))^{[(t_k-t_{k-1}-2\delta)n'']-1}$$

with $[\alpha]$ denoting the greatest integer not exceeding α. Indeed, if $X^{n''} \in \Delta(\delta)$, then in particular

$$\rho_E(U(\tau_i), X_t^{n''}) \leqslant \sup_{s \in [t_i, t_{i+1})} \rho_E(U(\tau_i), U(s)) + \inf_{s \in [t_i, t_{i+1})} \rho_E(U(s), X_t^{n''}) \leqslant \varepsilon/2 + \delta < \varepsilon$$

for $t \in [t_i+\delta, t_{i+1}-\delta]$. Thus, during the whole period $[t_i+\delta, t_{i+1}-\delta]$ the process $X^{n''}$, performing at least $[(t_{i+1}-t_i-2\delta)n'']-1$ transitions, each with probability at most $P_{U(\tau_i)}(\varepsilon)$, remains in the ε-neighbourhood of $U(\tau_i)$, which yields (8). Letting $n'' \to \infty$ we conclude from (8)

$$\liminf_{n'' \to \infty} L_{n''} \leqslant \sum_{i=1}^{k-1} (t_{i+1}-t_i-2\delta) \log P_{U(\tau_i)}(\varepsilon).$$

Thus, since $X^{n''}$ satisfies the large deviation principle with the rate function J_{H_0}, taking into account that $\Delta(\delta)$ is open, we get

$$J_{H_0}(U) \geqslant \inf_{V \in \Delta(\delta)} J_{H_0}(V) \geqslant -\liminf_{n'' \to \infty} L_{n''} \geqslant -\sum_{i=0}^{k-1} (t_{i+1}-t_1-2\delta) \log P_{U(\tau_i)}(\varepsilon).$$

Since $\tau_i \in [t_i, t_{i+1})$ were arbitrary, taking $\delta \to 0$ we conclude that

$$(9) \qquad J_{H_0}(U) \geqslant - \int_{[0,1]} \log P_{U(t)}(\varepsilon)\, dt.$$

We will show that for each $A \in \mathscr{D}$

$$(10) \qquad \limsup_{\varepsilon \to 0} P_A(\varepsilon) \leqslant \pi\big((Z \cup H_0) \subseteq A \mid A\big).$$

Clearly, we can confine ourselves to the case $H_0 \subseteq A$, for otherwise both sides equal 0. Choose some $\eta > 0$. Let $\varepsilon_k \to 0$ for $k \to \infty$ and let $H_0 \subseteq A_k \in \mathscr{D}$ be such that $\rho_E(A_k, A) < \varepsilon_k$ and

$$\pi(Z \subseteq A^{(\varepsilon_k)} \mid A_k) > P_A(\varepsilon_k) - \eta.$$

Further, take an arbitrary $\theta > 0$. Then, for k such that $\varepsilon_k < \theta$ we have

$$P_A(\varepsilon_k) < \pi(Z \subseteq A^{[\theta]} \mid A_k) + \eta$$

with $A^{[\theta]} = \{x \in E \mid \operatorname{dist}(x, A) \leqslant \theta\}$. Letting $k \to \infty$ we get from (C1)

$$\limsup_{\varepsilon \to 0} P_A(\varepsilon) \leqslant \pi(Z \subseteq A^{[\theta]} \mid A) + \eta.$$

Taking in turn $\eta, \theta \to 0$ we obtain (10).

Finally, combining (10) with (9) and applying Fatou's lemma we conclude that

$$J_{H_0}(U) \geqslant - \int_{[0,1]} \log \pi\big((Z \cup H_0) \subseteq U(t) \mid U(t)\big) = I_{H_0}(U),$$

as required. ■

We pass now to the second step of the proof showing that the inequality converse to that established in the previous lemma is satisfied for all piecewise constant functions $U \in \mathscr{U}$.

LEMMA 3. *Let $U \in \mathscr{U}$ be piecewise constant. Then $J_{H_0}(U) \leqslant I_{H_0}(U)$.*

Proof. To simplify the notation we assume without loss of generality that U is of the form

$$U(t) = \begin{cases} A & \text{if } 0 \leqslant t < \alpha, \\ B & \text{otherwise,} \end{cases}$$

where $A, B \in \mathscr{D}$, $A \subseteq B$ and $0 < \alpha < 1$. Also, we can require that $H_0 \subseteq A$, for otherwise $I_{H_0}(U) = +\infty$ and the assertion of the lemma becomes obvious. For the same reasons we assume that $\pi(Z \subseteq A \mid A) > 0$ and $\pi(Z \subseteq B \mid B) > 0$. Take some ε such that

$$0 < \varepsilon < \frac{\min(\alpha, 1-\alpha)}{2},$$

and define

$$\bar{\Delta}(\varepsilon) := \{V \in \mathscr{U} \mid \varrho(V, U) \leqslant \varepsilon\}$$

to be the closed 'ε-sausage' around U. Let

$$\Lambda_{n''}(\varepsilon) := \frac{1}{n''} \log \boldsymbol{P}_{H_0}\big(X^{n''} \in \bar{\Delta}(\varepsilon)\big).$$

Choose $m \in \boldsymbol{N}$ such that

$$p_{H_0 \to A}(\varepsilon) := \boldsymbol{P}_{H_0}\Big(\rho_E\big(\bigcup_{j=0}^{m} Z_j, A\big) \leqslant \varepsilon\Big) > 0$$

and

$$p_{A \to B}(\varepsilon) := \inf\Big\{\boldsymbol{P}_C\Big(\rho_E\big(\bigcup_{j=0}^{m} Z_j, B\big) \leqslant \varepsilon\Big) \mid C \in \mathscr{D}, \rho_E(C, A) \leqslant \varepsilon\Big\} > 0.$$

The existence of such m follows from condition (C2) (recall that $H_0 \subseteq A \subseteq B$ and both $\pi(Z \subseteq A \mid A)$ and $\pi(Z \subseteq B \mid B)$ are positive). Further, let

$$p_A(\varepsilon) := \inf\{\pi(Z \subseteq A \mid C) \mid C \in \mathscr{D}, \rho_E(A, C) \leqslant \varepsilon\}$$

and

$$p_B(\varepsilon) := \inf\{\pi(Z \subseteq B \mid C) \mid C \in \mathscr{D}, \rho_E(B, C) \leqslant \varepsilon\}.$$

Take $k \in \boldsymbol{N}$ such that $m/k \leqslant \varepsilon$. Then, for $n'' > k$ we have

$$\text{(11)} \qquad \boldsymbol{P}_{H_0}\big(X^{n''} \in \bar{\Delta}(2\varepsilon)\big) \geqslant p_{H_0 \to A}(\varepsilon)\, \big(p_A(\varepsilon)\big)^{\alpha n''}\, p_{A \to B}(\varepsilon)\, \big(p_B(\varepsilon)\big)^{(1-\alpha)n''+1}.$$

To see this observe that the right-hand side of the above inequality does not exceed the probability of the following event $\mho(\varepsilon) \in \mathfrak{F}$:

$\mho(\varepsilon)$: *During the first m steps the process $X^{n''}$ passes from $X_0^{n''} = H_0$ to some $X_{(m/n'')}^{n''}$ such that $\rho_E(X_m^{n''}, A) \leqslant \varepsilon$. Then, for $m \leqslant j \leqslant \alpha n''$, $X_{(j/n'')}^{n''}$ remains in the ε-neighbourhood of A. During the further m steps it performs a transition to a state which lies in the ε-neighbourhood of B. Finally, $\rho_E(X_{(j/n'')}^{n''}, B) \leqslant \varepsilon$ for $[\alpha n''] + m \leqslant j \leqslant n''$.*

Since $X^{n''}$ is nondecreasing, in view of the definition of $\bar{\Delta}(2\varepsilon)$ and because of (3) the event $\mho(\varepsilon)$ entails $X^{n''} \in \bar{\Delta}(2\varepsilon)$. Hence (11) is established. Passing with n'' to infinity we easily obtain

$$\limsup_{n'' \to \infty} \Lambda_{n''}(2\varepsilon) \geqslant \alpha \log p_A(\varepsilon) + (1-\alpha) \log p_B(\varepsilon).$$

Using the fact that $X^{n''}$ satisfies the large deviation principle with the rate function J_{H_0} and taking into account that $\bar{\Delta}(2\varepsilon)$ is closed we conclude that

$$\text{(12)} \qquad \inf_{V \in \bar{\Delta}(2\varepsilon)} J_{H_0}(V) \leqslant -\limsup_{n'' \to \infty} \Lambda_{n''}(2\varepsilon) \leqslant -\big(\alpha \log p_A(\varepsilon) + (1-\alpha) \log p_B(\varepsilon)\big).$$

Condition (C1) together with the definitions of $p_A(\varepsilon)$ and $p_B(\varepsilon)$ yields by standard arguments

$$\lim_{\varepsilon\to 0} p_A(\varepsilon) = \pi(Z \subseteq A \mid A) = \pi\big((Z \cup H_0) \subseteq A \mid A\big)$$

and

$$\lim_{\varepsilon\to 0} p_B(\varepsilon) = \pi(Z \subseteq B \mid B) = \pi\big((Z \cup H_0) \subseteq B \mid B\big).$$

Thus, letting $\varepsilon \to 0$ we conclude from (12) and the lower semicontinuity of J_{H_0} that

$$J_{H_0}(U) \leqslant -\alpha \log \pi\big((Z\cup H_0) \subseteq A \mid A\big) - (1-\alpha)\log \pi\big((Z \cup H_0) \subseteq B \mid B\big) = I_{H_0}(U),$$

as required. ∎

The last step of the proof of Theorem 1 is to extend the above result for all $U \in \mathscr{U}$.

LEMMA 4. *For each* $U \in \mathscr{U}$ *we have* $J_{H_0}(U) \leqslant I_{H_0}(U)$.

Proof. Take an arbitrary $\varepsilon > 0$ and let $0 = t_0 < t_1 < \ldots < t_k = 1$ be as in (7). For each $0 \leqslant i < k$ choose $\sigma_i \in [t_i, t_{i+1})$ so that

$$(13)\qquad -(t_{i+1}-t_i)\log \pi\big((Z\cup H_0) \subseteq U(\sigma_i) \mid U(\sigma_i)\big) \leqslant -\int_{[t_i,t_{i+1})} \log \pi\big((Z \cup H_0) \subseteq U(t) \mid U(t)\big)\,dt.$$

Define the piecewise constant function $U^\varepsilon \in \mathscr{M}$ by

$$U^\varepsilon(t) := U(\sigma_i) \quad \text{for } t \in [t_i, t_{i+1}),\ 0 \leqslant i < k.$$

Then, by (7), $\varrho(U, U^\varepsilon) \leqslant \varepsilon/2$ and, by (13), $I_{H_0}(U^\varepsilon) \leqslant I_{H_0}(U)$. Thus, letting $\varepsilon \to 0$ we obtain, by the lower semicontinuity of J_{H_0} and by Lemma 3,

$$J_{H_0}(U) \leqslant \liminf_{n''\to\infty} J_{H_0}(U^\varepsilon) \leqslant \liminf_{n''\to\infty} I_{H_0}(U^\varepsilon) \leqslant I_{H_0}(U).$$

This yields the assertion of the lemma. ∎

Combining Lemmas 2, 3 and 4 completes the proof of Theorem 1. ∎

3. CONVEX HULLS OF UNIFORM SAMPLES

Let $\zeta, \zeta_1, \zeta_2, \ldots$ be a sequence of i.i.d. random vectors uniformly distributed on the d-dimensional unit ball $B_d \subset \boldsymbol{R}^d$ and define C_n to be the convex hull of $\{\zeta_1, \ldots, \zeta_n\}$:

$$C_n := \operatorname{conv}(\{\zeta_1, \zeta_2, \ldots, \zeta_n\}).$$

It can be easily verified that C_n converges almost surely to B_d (for instance, in the Hausdorff metric). Since the pioneering papers of Rényi and Sulanke [16] and Efron [5] the speed of this convergence and related questions have been thoroughly investigated in the literature by various authors. For an extensive reference of these results see e.g. [17] and the references therein. Over the last decades important progress has been made in the particular two-dimensional case where very strong limit theorems for some functionals of C_n (such as volume and perimeter) have been proven since the paper of Groeneboom [6] followed by other authors (e.g. [1], [2], [7], [9]). Several results, though weaker, have also been obtained in higher dimensions (see [10] and the references therein). However, there remains a number of open questions.

The asymptotic behaviour of the sequence of successive convex hulls can be studied by means of the growing random processes $(\Theta^n)_{t\in[0,1]}$ given for $n\in N$ by

$$\Theta_t^n := C_{[nt]}.$$

In this example we aim at applying the general results presented in the previous section to establish the large deviation principle for Θ^n.

Let us define

$$C(B_d) := \{C \subset B_d \mid C \text{ is closed and convex}\}$$

and endow this space with the usual Hausdorff metric denoted by ρ^C. For the formal presentation of our theorem we need to impose a topological structure on the family $\mathscr{C}$ of all nondecreasing right continuous $C(B_d)$-valued functions defined on $[0, 1]$. We do it exactly in the same way as we did in the case of $\mathscr{U}$ (see (3)), thus making $\mathscr{C}$ a compact metric space.

As an application of Theorem 1 we prove

THEOREM 2. *For $\Phi\in\mathscr{C}$ define*

$$\hat{I}(\Phi) := \log\lambda(B_d) - \int_{[0,1]} \log\lambda(\Phi(t))\,dt, \tag{14}$$

where λ is the d-dimensional Lebesgue measure. Then $\hat{I}$ is a rate function and the sequence Θ^n satisfies the large deviation principle on $\mathscr{C}$ with the rate function $\hat{I}$.

Proof. Since it is not convenient for us to deal directly with the convex sets, because this class is not closed with respect to set-theoretic unions, we choose to represent each convex set C by the subgraph of the restriction of its support function (see e.g. Section 1.7 in [18]) to the unit sphere $S_{d-1} = \partial B_d$. Namely, we associate with each $C\in C(B_d)$ the closed set $\boldsymbol{h}(C)$ given by

$$S_{d-1}\times[-1, 1] \supset \boldsymbol{h}(C) := \{(u, y)\in S_{d-1}\times[-1, 1] \mid y \leqslant \max_{x\in C}\langle x, u\rangle\},$$

where $\langle\cdot,\cdot\rangle$ denotes the scalar product. By convention, we set $\boldsymbol{h}(\emptyset) := \emptyset$.

Since $\boldsymbol{h}(C)$ determines C uniquely, we can identify C with $\boldsymbol{h}(C)$, thus obtaining an embedding of $\boldsymbol{C}(B_d)$ into the space $\mathscr{F}(S_{d-1}\times[-1,1])$ of all the closed subsets of $S_{d-1}\times[-1,1]$. It is not difficult to verify that this embedding is in fact a homeomorphism of compact spaces $(\boldsymbol{C}(B_d), \rho^C)$ and $\boldsymbol{h}(\boldsymbol{C}(B_d)) \subseteq (\mathscr{F}(S_{d-1}\times[-1,1]), \rho_{(S_{d-1}\times[-1,1])})$. Hence, defining $\mathscr{D}$ as the image of $\boldsymbol{h}$, i.e. $\mathscr{D} := \{\boldsymbol{h}(C) \mid C\in \boldsymbol{C}(B_d)\}$, we note that $\mathscr{D}$ satisfies condition (K2). To see that also (K1) is fulfilled check that

$$\boldsymbol{h}(\operatorname{conv}(A\cup B)) = \boldsymbol{h}(A)\cup \boldsymbol{h}(B) \tag{15}$$

for $A, B\in \boldsymbol{C}(B_d)$. Formula (15) allows us to replace, using $\boldsymbol{h}$, the operation of taking the convex hull of union of two sets by the operation of set-theoretic union. In particular, we obtain

$$\boldsymbol{h}(C_n) = \bigcup_{i=1}^{n} \boldsymbol{h}(\{\zeta_i\}). \tag{16}$$

This identity allows us to establish a handy representation for C_n and Θ^n, which fits well into the general setting of Theorem 1. Namely, let almost surely $Z_0 := \emptyset$ and $Z_i := \boldsymbol{h}(\{\zeta_i\})$. In view of (1) this corresponds to choosing the stochastic kernel π independent of the second variable and given by

$$\pi(\mathscr{E} \mid B) = \boldsymbol{P}(Z_i\in \mathscr{E}) \tag{17}$$

for $\mathscr{E} \subset \mathscr{D}$ and $B\in\mathscr{D}$. Further, (16) translates into $\boldsymbol{h}(C_n) = \bigcup_{i=0}^{n} Z_i$, and therefore

$$\boldsymbol{h}(\Theta_t^n) = X_t^n \tag{18}$$

for all $t\in[0, 1]$, where $(X^n)_{t\in[0,1]}$ is the union process defined in (2).

Consider $\mathscr{U} = \mathscr{U}(\mathscr{D})$ defined as usually (see the discussion following the definition of the union process (2)). Recall the definition of $\mathscr{C}$ given before the formulation of Theorem 2 and observe that the mapping $\hat{\boldsymbol{h}}\colon \mathscr{C} \to \mathscr{U}(\mathscr{D})$ given by $[\hat{\boldsymbol{h}}(F)](t) := \boldsymbol{h}(F(t))$ for $F\in\mathscr{C}$ establishes a homeomorphism of the compact spaces $\mathscr{U}$ and $\mathscr{C}$ endowed with respective topologies. Note also that (18) translates into $\hat{\boldsymbol{h}}(\Theta^n) = X^n$. Therefore, to prove Theorem 2 it suffices to show that the sequence X^n satisfies on $\mathscr{U}$ the large deviation principle with a certain rate function J such that

$$J(\hat{\boldsymbol{h}}(\Phi)) = \hat{I}(\Phi), \qquad \Phi\in\mathscr{C}, \tag{19}$$

with $\hat{I}$ defined as in (14).

We will proceed as follows. First we shall argue that the regularity conditions (C1) and (C2) are satisfied. Then we will apply Theorem 1 to establish the large deviation principle for X^n. Finally, we will show that (19) holds, thus completing the proof.

Condition (C1) is obvious because π does not depend on its second variable (see (17)). To establish (C2) fix $\varepsilon > 0$, take some $A, B\in\mathscr{D}$, $A \subseteq B$, and let

$C_A, C_B \in \boldsymbol{C}(B_d)$, $C_A \subseteq C_B$, be the corresponding convex sets, i.e. $A = \boldsymbol{h}(C_A)$ and $B = \boldsymbol{h}(C_B)$. Choose $\eta > 0$ such that $\rho_{(S_{d-1}\times[-1,1])}(\boldsymbol{h}(C), B) \leqslant \varepsilon$ for each $C \in \boldsymbol{C}(B_d)$ such that $\rho^{\boldsymbol{C}}(C, C_B) \leqslant \eta$ (such a choice is possible because $\boldsymbol{h}$ is continuous). Since $\pi(Z \subseteq \emptyset \mid \emptyset) = 0$, we can assume without loss of generality that C_B is nonempty. The boundary ∂C_B is compact, so we can cover it with a finite number $m(\delta)$ of open balls $K_1, \ldots, K_{m(\delta)}$ with common radius $\delta > 0$ such that $\rho^{\boldsymbol{C}}(C, C_B) \leqslant \eta$ for each $C \in \boldsymbol{C}(B_d)$, $C \subseteq \operatorname{conv}(K_1 \cup \ldots \cup K_{m(\delta)})$, such that $C \cap K_i \neq \emptyset$ for all $i = 1, \ldots, m(\delta)$. Then, for each $D \in \mathscr{D}$ such that $\rho_{(S_{d-1}\times[-1,1])}(D, A) \leqslant \varepsilon$ we have

$$\boldsymbol{P}_D\Big(\rho_{(S_{d-1}\times[-1,1])}\Big(\bigcup_{i=0}^{m(\delta)} Z_i, B\Big) \leqslant \varepsilon\Big) > \boldsymbol{P}_D\big(\zeta_i \in K_i \text{ for all } i = 1, \ldots, m(\delta)\big) > 0.$$

Since this bound does not depend on A, condition (C2) holds true.

Thus, applying Theorem 1 we conclude that the sequence X^n satisfies on $\mathscr{U}$ the large deviation principle with the rate function

$$I_{\emptyset}(U) = -\int_{[0,1]} \log \pi\big(Z \subseteq U(t) \mid U(t)\big)\,dt = -\int_{[0,1]} \log \boldsymbol{P}\big(Z \subseteq U(t)\big)\,dt.$$

Taking into account the identifications made in the course of the proof we see that this identity translates into

$$I_{\emptyset}(U) = -\int_{[0,1]} \log \boldsymbol{P}\big(\zeta \in \boldsymbol{h}^{-1}(U(t))\big)\,dt. \tag{20}$$

Clearly, for each $C \in \boldsymbol{C}(B_d)$ we have $\boldsymbol{P}(\zeta \in C) = \lambda(C)/\lambda(B_d)$. This proves (19) with $J = I_{\emptyset}$. The proof of Theorem 2 is thus complete. ∎

Acknowledgments. The author would like to express his gratitude to Professor A.V. Nagaev for his valuable suggestions and comments.

REFERENCES

[1] H. Bräker and T. Hsing, *On the area and perimeter of a random convex hull in a bounded convex set*, Probab. Theory Related Fields 111 (1998), pp. 517–550.

[2] A. J. Cabo and P. Groeneboom, *Limit theorems for functionals of convex hulls*, Probab. Theory Related Fields 100 (1994), pp. 31–55.

[3] J.-D. Deuschel and D. W. Stroock, *Large Deviations*, Academic Press, Inc., Boston 1989.

[4] P. Dupuis and R. S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations*, Wiley, Chichester 1997.

[5] B. Efron, *The convex hull of a random set of points*, Biometrika 52 (1965), pp. 331–343.

[6] P. Groeneboom, *Limit theorems for convex hulls*, Probab. Theory Related Fields 79 (1988), pp. 327–368.

[7] T. Hsing, *On the asymptotic distribution of the area outside a random convex hull in a disk*, Ann. Appl. Probab. 4 (2) (1994), pp. 478–493.

[8] W. S. Kendall, J. Mecke and D. Stoyan, *Stochastic Geometry and Its Applications*, Akademie-Verlag, Berlin 1987.

[9] I. M. Khamdamov and A. V. Nagaev, *Limiting distributions for functionals of the convex hull generated by uniformly distributed variables* (in Russian), Dokl. Akad. Nauk UzSSR, Tashkent, 7 (1991), pp. 8–9.

[10] K. H. Küfer, *On the approximation of a ball by random polytopes*, Adv. Appl. Probab. 26 (1994), pp. 876–892.

[11] G. Matheron, *Random Sets and Integral Geometry*, Wiley, New York 1975.

[12] I. S. Molchanov, *Limit Theorems for Unions of Random Closed Sets*, Lecture Notes in Math. 1561, Springer, Berlin 1993.

[13] I. S. Molchanov, *On the convergence of random processes generated by polyhedral approximation of convex compacts*, Theory Probab. Appl. 40 (1995), pp. 383–390.

[14] G. L. O'Brien and W. Verwaat, *Capacities, large deviations and* loglog *laws*, in: *Stable Processes and Related Topics*, Progress in Probability, Vol. 25, Birkhäuser, Boston 1991, pp. 43–83.

[15] A. A. Pukhalskii, *On functional principle of large deviations*, in: *New Trends in Probability and Statistics*, V. Sazonov and T. Shervashidze (Eds.), VSP-Mokslal, Zeist, The Netherlands, 1991, pp. 198–218.

[16] A. Rényi and R. Sulanke, *Über die konvexe Hülle von n zufällig gewählten Punkten* (I and II), Z. Wahrscheinlichkeitstheorie verw. Gebiete 2 (1963), pp. 75–84; ibidem 3 (1964), pp. 138–147.

[17] R. Schneider, *Random approximation of convex sets*, Journal of Microscopy 151 (1988), pp. 211–227.

[18] R. Schneider, *Convex bodies: the Brunn–Minkowski theory*, Encyclopaedia of Mathematics and Its Applications 44, Cambridge University Press, 1993.

Nicholas Copernicus University
Toruń, Poland
E-mail: tomeks@mat.uni.torun.pl

Received on 23.2.2000