LARGE DEVIATIONS FOR PARTIAL SUMS U-PROCESSES IN DEPENDENT CASES

PETER EICHELSBACHER

ABSTRACT. The large deviation principle (LDP) is known to hold for partial sums U–processes of real-valued kernel functions of i.i.d. random variables X_i. We prove an LDP when the X_i are independent but not identically distributed or fulfill some Markov dependence or mixing conditions. Moreover we give a general condition which suffices for the LDP to carry over from the partial sums empirical processes LDP to the partial sums U-processes LDP for kernel functions satisfying an appropriate exponential tail condition.

1. INTRODUCTION

Let $\{X_n, n \in \mathbb{N}\}$ be a sequence of random variables on a probability space $(\Omega, \mathcal{A}, P)$ with Polish state space S. Dembo and Zajic [3] proved an LDP for

$$L_n(t) = \frac{1}{n} \sum_{i=1}^{[nt]} \delta_{X_i}, \quad t \in [0,1], \tag{1.1}$$

where the X_i fulfill Markov dependence or mixing conditions. They observed that the special conditions, which guarantee the LDP to hold for the empirical measures of Polish space valued random variables suffice for the LDP to carry over to the corresponding partial sums processes $L_n(\cdot)$. One application of these results considers the calculation of how large delays are built up in a single server queue in the presence of dependent customer interarrival times. For other applications see [3] and [19].

In [13] the LDP is proved for partial sums U-processes, i.e.

$$U_n(t) = \frac{1}{\binom{n}{m}} \sum_{C_m^{[nt]}} h(X_{i_1}, \dots, X_{i_m}), \quad t \in [0,1], \tag{1.2}$$

where the X_i are i. i. d. random variables and h is a measurable, $\mathbb{R}^d$-valued function, called kernel function, which is symmetric, i.e. invariant under permutation of its arguments. For $k, m \in \mathbb{N}$ C_m^k denotes the set $\{(i_1, \dots, i_m) : 1 \leq i_1 < \cdots < i_m \leq k\}$. The LDP holds in $L_\infty([0,1], (\mathbb{R}^d, \|\cdot\|))$, the space of almost surely bounded, measurable $\mathbb{R}^d$-valued functions on $[0,1]$, where $[0,1]$ is equipped with

1991 *Mathematics Subject Classification.* 60F10 (primary).

Key words and phrases. Large deviations, partial sums, U-processes, Markov chains, hypermixing, strong mixing.

Lebesgue measure and the underlying topology is the supremum norm topology. As a payoff, applications in the linear case were extended to higher order statistics. Thus LDP results for multi-server queues or cost functions with several variables were obtainable.
The LDP for U-statistics of degree $m \in \mathbb{N}$, i.e.

$$U_n := U_n(1) = \frac{1}{\binom{n}{m}} \sum_{C_m^n} h(X_{i_1}, \dots, X_{i_m}),$$

with h and C_m^n as above, is proved in the i. i. d. case for $\mathbb{R}^d$-valued h and Polish space valued X_i in [12] and under weaker moment conditions in [10]. In [13] the LDP for $\{U_n(\cdot)\}_{n\in\mathbb{N}}$ is proved under the same exponential tail conditions for the kernel function as for U-statistics.

In this paper we discuss the LDP for the partial sums U-processes for dependent X_i as well as the LDP for another class of processes, the so called V-processes or von-Mises processes, given by

$$V_n(t) = \frac{1}{n^m} \sum_{D_m^{[n\,t]}} h(X_{i_1}, \dots, X_{i_m}), \quad t \in [0,1], \tag{1.3}$$

where $D_m^n := \{(i_1, \dots, i_m) : 1 \le i_1, \dots, i_m \le n\}$ and h is as above. The corresponding statistics are given by the V-statistics or von Mises-statistics given by

$$V_n = \frac{1}{n^m} \sum_{D_m^n} h(X_{i_1}, \dots, X_{i_m}).$$

Due to the symmetry of h, U_n and V_n mainly differ by the diagonal terms, where at least two of the indices $i_1, \dots, i_m$ are equal.
For weakly dependent sequences $\{X_n, n \in \mathbb{N}\}$ invariance principles and the weak convergence of appropriately normalized sample path U-processes and von Mises-processes to Brownian motion or limit processes expressible as multiple Wiener integrals were studied in [6]. Notice that U-processes indexed by a class of measurable functions are studied for example in [1].

We will proceed as follows: as in the i.i.d. situation first we will prove that the m-fold product $L_n^{\otimes m}(\cdot)$ of $L_n(\cdot)$ satisfies a large deviation principle in a suitable topology, if $L_n(\cdot)$ satisfies an LDP for the underlying sequence X_i. Next we will apply a generalized version of the well-known contraction principle using an approximation of the kernel function worked out in [13]. On the other hand a special moment condition (2.4, stated in Section 2) has to be verified to guarantee that the partial sums U-process with the approximating kernel has the same large deviation behaviour as the partial sums U-process we are interested in.

In our examples we will check Condition 2.4 for a suitable modification of $U_n(\cdot)$. The remaining part of the proof is then to verify that these modifications have a

large deviation behaviour like $U_n(\cdot)$. In the i.i.d. case this is done via the well-known Hoeffding decomposition of a U-statistic into a dependent mean of i.i.d. means. For dependent sequences this is not as easy.
We will prove a general Theorem which allows us to transfer the LDP for $\{L_n(\cdot), n \in \mathbb{N}\}$ with respect to a suitable topology to the LDP for modifications of $\{U_n(\cdot), n \in \mathbb{N}\}$ (on $L_\infty([0,1], (\mathbb{R}^d, \|\cdot\|)))$, whenever Condition 2.4 can be verified and the kernel function satisfies appropriate exponential tail conditions. After that we have to assume some more tail conditions to get the LDP for the partial sums U-processes as well as for the partial sums V-processes.

This paper is divided into four sections. In Section 2 we present the main results. In Section 3 the main Theorem will be proved. In Section 4 we will prove a LDP for some independent but not identically distributed sequences or Markov dependence or mixing conditions. To this end we will apply some techniques of [16] which are already developed to get LDP results for U-statistics for weakly dependent sequences.

2. Statement of the results

Let $\{X_n, n \in \mathbb{N}\}$ be a sequence of random variables on a probability space $(\Omega, \mathcal{A}, P)$ with Polish state space S. Denote by $\mathcal{S}$ the Borel σ-Algebra on S.
We denote by $\mathcal{M}(S)$, $\mathcal{M}^+(S)$ and $\mathcal{M}^m(S)$, respectively, the sets of Borel measures on S which are signed, positive and positive having total mass m, respectively. These spaces are equipped with the topology of weak convergence.
The topological dual of $\mathcal{M}(S)$ is identified with the set $C_b(S)$ of continuous and bounded mappings from S to $\mathbb{R}$ via $\langle f, \mu\rangle = \int_S f d\mu$. $\mathcal{M}^+(S)$ becomes a Polish spaces when equipped with the metric topology induced by

$$\beta(\mu, \nu) := \sup\{|\int_S f d\mu - \int_S f d\nu| : f \in C_b(S), ||f||_\infty + ||f||_L \leq 1\},$$

where $||\cdot||_\infty$ denotes the supremum norm,

$$||f||_L := \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x,y)},$$

and d denotes the metric on S (see [3], Lemma A.1.).
Let us denote by $D([0,1], (\mathcal{M}^+(S), \beta))$ the space of all maps on $[0,1]$ with values in $(\mathcal{M}^+(S), \beta)$ continuous from the right and having left limits equipped with the uniform metric topology induced by

$$d_\infty(y(\cdot), z(\cdot)) := \sup_{t \in [0,1]} \beta(y(t), z(t)).$$

We will assume that the underlying sequence $\{X_n, n \in \mathbb{N}\}$ has the property that the partial sums processes $\{L_n(\cdot), n \in \mathbb{N}\}$ satisfy the LDP in $D([0,1], (\mathcal{M}^+(S), \beta))$ with a convex rate function $I(\nu(\cdot))$. Notice that in all cases of interest a nice representation of $I(\nu(\cdot))$ is given on its support $\mathcal{AC}_0$, where $\mathcal{AC}_0$ is defined as the set of all maps $\nu : [0,1] \to \mathcal{M}^+(S)$ that are absolutely continuous with respect to the variation norm $||\cdot||_{\text{var}}$, satisfy $\nu(t) - \nu(s) \in \mathcal{M}^{t-s}(S)$ for all $t \geq s$, while $\nu(0) = 0$, and possess a weak derivative for almost all t (the latter means that for almost every t the expression $\langle f, \nu(t+h) - \nu(t)\rangle / h$ converges to a limit $\langle f, \dot\nu(t)\rangle$ for every $f \in C_b(S)$). Following [3], the sequence $\{L_n(\cdot), n \in \mathbb{N}\}$ satisfies the LDP in $D([0,1], (\mathcal{M}^+(S), \beta))$ with rate function $I(\nu(\cdot))$, if the following assumption is fulfilled.

Condition 2.1. *Fix $m \in \mathbb{N}$ and $0 = t_0 < t_1 < \cdots < t_m \leq 1$. Denote by $L_n = (L_n(t_1), L_n(t_2) - L_n(t_1), \ldots, L_n(t_m) - L_n(t_{m-1}))$. Then $\{L_n, n \in \mathbb{N}\}$ satisfies the LDP in $D([0,1], (\mathcal{M}^+(S), \beta))^m$ with the good rate function*

$$J_m(z) = \sum_{i=1}^m (t_i - t_{i-1}) J\big(\frac{z_i}{t_i - t_{i-1}}\big),$$

where $z = (z_1, \ldots, z_m)$ and $J(\cdot)$ is the convex good rate function associated with the LDP of $\{L_n(1), n \in \mathbb{N}\}$.

If Condition 2.1 is fulfilled, the rate function $I(\nu(\cdot))$ has the form

$$I(\nu(\cdot)) = \int_0^1 J(\dot\nu) dt \quad \text{if } \nu \in \mathcal{A}C_0, \tag{2.2}$$

where

$$J(\nu) = \sup_{f \in C_b(S)} \{ \int_S f \, d\nu - \Lambda(f) \}, \tag{2.3}$$

and

$$\Lambda(f) := \lim_{n\to\infty} \frac{1}{n} \log E[\exp(\sum_{i=1}^n \int_S f \, dX_i)].$$

Condition 2.1 especially implies that for any fixed $t \in [0,1]$ the sequence $\{L_n(t), n \in \mathbb{N}\}$ satisfies the LDP with rate function $t\, J(\cdot/t)$.

For $m \geq 2$ we denote by $D([0,1], (\mathcal{M}^+(S^m), \beta^m))$ the space of all $(\mathcal{M}^+(S^m), \beta^m)$-valued cadlag-functions. S^m as well as $\mathcal{M}^+(S^m)$ are Polish spaces, since S is Polish. We take β^m to be the metric on $\mathcal{M}^+(S^m)$ defined analogously to β. Again we equip the space $D([0,1], (\mathcal{M}^+(S^m), \beta^m))$ with the uniform metric topology induced by $d_\infty^m(y(\cdot), z(\cdot)) := \sup_{t\in[0,1]} \beta^m(y(t), z(t))$.
Denote by $||\cdot||$ the Euclidean norm on $\mathbb{R}^d$ and by $\mathcal{A}C_0([0,1], (\mathbb{R}^d, ||\cdot||))$ the space of all absolutely continuous $\mathbb{R}^d$-valued functions ϕ on $[0,1]$ with $\phi(0) = 0$.

We consider the following conditions for a sequence $\{R_n\}_{n\in\mathbb{N}} \subseteq \mathcal{M}^t(\mathcal{M}^t(S^m))$, $t \in [0,1]$. Let $B(S^m)$ denote the space of all bounded, real-valued, $\mathcal{S}^{\otimes m}$-measurable functions on S^m.

Condition 2.4. *There exist constants $\beta, M \in [1,\infty)$ and a reference measure $\mu \in \mathcal{M}^1(S)$ such that the inequality*

$$\sup_{n\in\mathbb{N}}\left(\int_{\mathcal{M}^1(S^m)} \exp\left(n\int_{S^m}\varphi\, d\nu\right) R_n(d\nu)\right)^{1/n} \le M\int_{S^m}\exp(\beta\varphi)\, d\mu^{\otimes m}$$

holds for all $\varphi:\ S^m \to [0,\infty)$ with $\varphi \in B(S^m)$.

Remark 2.5. Condition 2.4 describes the amount of dependence of the underlying process under which a LDP for the laws of the partial sums empirical measure is preserved under the transformation leading to the partial sums U-process. The condition is also involved in transferring a LDP for empirical measures to a stronger topology on the set of probability measures (c.f. [7], Lemma 3.2.19 and Theorem 3.2.21) as well as transferring a LDP for empirical measures to products of empirical measures and U-empirical measures, respectively (c.f. [16]).

Let $h:\ S^m \to \mathbb{R}^d$ be a measurable and symmetric function and define the corresponding partial sums U-process $U_n(\cdot)$ and the partial sums V-process $V_n(\cdot)$ as in (1.2) and (1.3). We will prove the LDP for $\{U_n(\cdot), n\in\mathbb{N}\}$ assuming

Condition 2.6. *For every $\alpha > 0$,*

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

where μ is the reference measure of Condition 2.4.

Proving the LDP for $\{V_n(\cdot), n\in\mathbb{N}\}$ we will assume

Condition 2.7. *For every $\alpha > 0$,*

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

for every map $\sigma:\ \{1,\ldots,m\}\to\{1,\ldots,m\}$, where $\pi_\sigma:\ S^m\to S^m$ is defined by $\pi_\sigma(s) = (s_{\sigma(1)},\ldots,s_{\sigma(m)})$ for every $s=(s_1,\ldots,s_m)\in S^m$, and μ is the reference measure of Condition 2.4.

In the applications we will check Condition 2.4 for a suitable modification $M_n(t)$ of $L_n^{\otimes m}(t)$ for any fixed $t\in[0,1]$. The remaining part is to verify that these modifications have a large deviation behaviour as $L_n^{\otimes m}(t)$. Recall that for fixed $t\in[0,1]$ the sequences $\{L_n^{\otimes m}(t), n\in\mathbb{N}\}$ and $\{M_n(t), n\in\mathbb{N}\}$ are called exponentially equivalent with respect to the metric β^m if for each $\delta>0$

$$\limsup_{n\to\infty}\frac{1}{n}\log P(\beta^m(L_n^{\otimes m}(t), M_n(t)) > \delta) = -\infty.$$

In all dependent situations we have in mind, that the modifications $M_n(\cdot)$ are "partial" sums of $L_n^{\otimes m}$. To be more precise define

$$M_n(t) := \frac{1}{|A_{m,[nt]}|}\sum_{(i_1,\ldots,i_m)\in A_{m,[nt]}}\delta_{(X_{i_1}(\omega),\ldots,X_{i_m}(\omega))}, \tag{2.8}$$

where $A_{m,n} \subset D_m^n$ has the property

$$\lim_{n\to\infty} \frac{n^m - |A_{n,m}|}{n^m} = 0, \tag{2.9}$$

if $n \to \infty$. Finally let $\widetilde{M}_n(\cdot)$ be the polygonal approximation of $M_n(\cdot)$:

$$\widetilde{M}_n(t) := M_n(t) + (nt - [nt]) \frac{1}{|A_{n,m}|} \sum_{A_{m-1,[nt]}} \delta_{(X_{i_1},\dots,X_{i_{m-1}},X_{[nt]+1})}.$$

By definition we get $\widetilde{M}_n(\cdot) \in \mathcal{C}_0([0,1],(\mathcal{M}^+(S^m),\beta^m))$, the space of all continuous maps equipped with the uniform metric topology induced by d_∞^m. Now our main result can be stated as follows:

Theorem 2.10. *Assume that the sequence $\{X_n, n \in \mathbb{N}\}$ satisfies Condition 2.1. Let $\{M_n(t), n \in \mathbb{N}\}$ be exponentially equivalent (with respect to β^m) to $L_n^{\otimes m}(t)$ for every $t \in [0,1]$. Assume that Condition 2.4 holds for a reference measure $\mu \in \mathcal{M}^1(S)$ and $M_n(t)$ for every $t \in [0,1]$.*

Assuming 2.6 holds for h with respect to the reference measure μ then the sequences $\{\int_{S^m} h\, dM_n(\cdot), n \in \mathbb{N}\}$ and $\{\int_{S^m} h\, d\widetilde{M}_n(\cdot), n \in \mathbb{N}\}$ satisfy the LDP in $L_\infty([0,1],(\mathbb{R}^d, \|\cdot\|))$ with the good rate function

$$I_\infty(\phi) = \inf\{\int_0^1 I(\dot\nu)dt, \nu \in \mathcal{AC}_0 \cap K_\infty \quad \textit{and} \quad \int_{S^m} h d\nu^m(\cdot) = \phi(\cdot)\} \tag{2.11}$$

and speed n, if $\phi \in \mathcal{AC}_0([0,1],(\mathbb{R}^d, ||\cdot||))$ and $I_\infty = \infty$ otherwise. Here

$$K_\infty := \bigcup_{L \geq 0} \{\nu(\cdot) : I(\nu(\cdot)) \leq L\}.$$

Remark 2.12. In Section 4 we will formulate the moment conditions under which the LDP for the partial sums U-processes and V-processes can be deduced from Theorem 2.10.

Remark 2.13. Under the conditions mentioned in Remark 2.12 we can extend the results to the laws of

$$V_\varepsilon(t) = \varepsilon^m \sum_{D_m^{[t/\varepsilon]}} h(X_{i_1},\dots,X_{i_m}), \quad t \in [0,1],$$

and

$$U_\varepsilon(t) = \frac{1}{\binom{1/\varepsilon}{m}} \sum_{C_m^{[t/\varepsilon]}} h(X_{i_1},\dots,X_{i_m}), \quad t \in [0,1],$$

for $\varepsilon > 0$. More precisely we get the LDP for the family $\{U_\varepsilon(\cdot), \varepsilon > 0\}$ in $L_\infty([0,1],(\mathbb{R}^d, \|\cdot\|))$ with the good rate function $I_\infty(\cdot)$ given by (2.11) and speed ε^{-1} whenever it can be established for $U_n(\cdot)$. The same is true for $\{V_\varepsilon(\cdot), \varepsilon > 0\}$.

Finally we want to show that for $m = 1$ and $h(x) = x$ we get the same rate function as given in [4]. Let us assume that the assumptions of Theorem 2.10 are fulfilled, thus we get the LDP for

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

with rate function

$$\inf\{\int_0^1 I(\dot{\nu})dt, \nu \in \mathcal{AC}_0 \cap K_\infty \quad \text{and} \quad \int_S x d\nu(\cdot) = \phi(\cdot)\}. \tag{2.14}$$

We will show that this is the usual rate function for $\{S_n(\cdot), n \in \mathbb{N}\}$. Let $\Lambda^*(x) := \sup_{\lambda \in \mathbb{R}^d}\{\langle \lambda, x\rangle - \Lambda(\lambda)\}$, $x \in \mathbb{R}^d$, where

$$\Lambda(\lambda) := \lim_{n\to\infty} \frac{1}{n} \log E[\exp(\sum_{i=1}^{n} \langle \lambda, X_i \rangle)].$$

Corollary 2.15. *Whenever*

$$\inf\{J(\nu(1)),\, \nu(1) \in \cap K_{1,\infty} \quad \textit{and} \quad \int_S x d\nu(1) = \phi(1)\} = \Lambda^*(\phi(1)),$$

where $K_{1,\infty} = \bigcup_{L\geq 0}\{\varrho \in \mathcal{M}^1(S) :\ J(\varrho) \leq L\}$, *we get*

$$\int_0^1 \Lambda^*(\dot{\phi}(t))dt = \inf\{\int_0^1 I(\dot{\nu})dt, \nu \in \mathcal{AC}_0 \cap K_\infty \quad \textit{and} \quad \int_S x d\nu(\cdot) = \phi(\cdot)\}.$$

3. Proof of the Theorem

First we will collect some preliminaries. The following Lemma is an easy observation proved in [14], Example 2.1.

Lemma 3.1. *Let S be Polish and d_v the metric induced by the total variation norm $\|\cdot\|_v$ on $\mathcal{M}^1(S)$. Then a LDP for a sequence of probability measures on $\mathcal{M}^1(S)$ with respect to the weak topology is transfered to every d_v-exponentially equivalent sequence of probability measures on $\mathcal{M}^1(S)$.*

Lemma 3.2. *Suppose the sequence $\{X_n, n \in \mathbb{N}\}$ satisfies Condition 2.1. Let us define*

$$L_n^{\otimes m}(\cdot) := \frac{1}{n^m} \sum_{D_m^{[n\cdot]}} \delta_{(X_{i_1},\dots,X_{i_m})}$$

and

$$L_n^m(\cdot) := \frac{1}{n_{(m)}} \sum_{(i_1,\dots,i_m)\in I_{m,[n\cdot]}} \delta_{(X_{i_1},\dots,X_{i_m})},$$

where $n_{(m)} := \prod_{k=0}^{m-1}(n-k)$ and where $I_{m,n} \subset \{1,\dots,n\}^m$ contains all m-tuples with pairwise different components. Then the sequences $\{L_n^{\otimes m}(\cdot), n \in \mathbb{N}\}$ and $\{L_n^m(\cdot), n \in \mathbb{N}\}$ satisfy the LDP in $D([0,1], (\mathcal{M}^+(S^m), \beta^m))$ with the good rate function $I^m(\nu(\cdot)) = I(\nu_1(\cdot))$ if $\nu_1^m(\cdot) = \nu(\cdot)$ and $\nu_1(\cdot) \in \mathcal{AC}_0$ and $I^m(\nu(\cdot)) = \infty$ otherwise.

Proof of Lemma 3.2. In view of the contraction principle [5], Theorem 4.2.23, we have to check that $\nu(\cdot) \mapsto \nu^m(\cdot)$ maps $D([0,1], (\mathcal{M}^+(S), \beta))$ into $D([0,1], (\mathcal{M}^+(S^m), \beta^m))$ and is continuous with respect to the topologies of uniform convergence. This is already proven in [13], Lemma 3.6. Since moreover for every fixed $t \in [0,1]$

$$||L_n^{\otimes m}(t) - L_n^m(t)||_{\text{var}} \leq \frac{C}{n^m}$$

with a constant C, which does only depend on m, we get the LDP for $\{L_n^m(\cdot), n \in \mathbb{N}\}$ applying Lemma 3.1 $\square$

Remark 3.3. Since for every fixed $t \in [0,1]$ $\{M_n(t), n \in \mathbb{N}\}$ is exponentially equivalent to $\{L_n^{\otimes m}(t), n \in \mathbb{N}\}$ by the assumption of Theorem 2.10, we get

$$\limsup_{n \to \infty} \frac{1}{n} \log P(d_\infty^m(L_n^{\otimes m}(\cdot), M_n(\cdot)) > \delta) = -\infty.$$

Therefore the sequence $\{M_n(\cdot), n \in \mathbb{N}\}$ satisfies the LDP in the same space with the same rate $I^m(\nu(\cdot))$. Since

$$||M_n(t) - \widetilde{M}_n(t)||_{\text{var}} \leq \frac{C}{n^m}$$

with a constant C, which only depends on m, we get the LDP for $\{\widetilde{M}_n(\cdot), n \in \mathbb{N}\}$ applying Lemma 3.1. Note that $\{\nu(\cdot) : I^m(\nu(\cdot)) < \infty\} \subset \mathcal{C}_0([0,1], (\mathcal{M}^+(S^m), \beta^m))$ and $P(\widetilde{M}_n(\cdot) \in \mathcal{C}_0([0,1], (\mathcal{M}^+(S^m), \beta^m))) = 1$. Thus the LDP for $\{\widetilde{M}_n(\cdot), n \in \mathbb{N}\}$ holds in the space $\mathcal{C}_0([0,1], (\mathcal{M}^+(S^m), \beta^m))$ by [5], Lemma 4.1.5.

Lemma 3.4. *With the notations of Lemma 3.2 let* $K_\infty^m := \bigcup_{L \geq 0} \{\nu(\cdot) : I^m(\nu(\cdot)) \leq L\}$. *If* $\nu(\cdot) \in K_\infty^m \cap \mathcal{AC}_0$ *then* $\int h d\nu(\cdot) \in \mathcal{AC}_0([0,1], (\mathbb{R}^d, ||\cdot||))$ *for all* $h : S^m \to \mathbb{R}^d$, *satisfying Condition 2.6. Thus* $I_\infty(\phi) < \infty$ *implies* $\phi \in \mathcal{AC}_0([0,1], (\mathbb{R}^d, ||\cdot||))$. *If* $\nu(\cdot) \in K_\infty^m \cap \mathcal{C}_0([0,1], (\mathcal{M}^+(S^m), \beta^m))$, *then* $\int h d\nu(\cdot) \in \mathcal{C}_0([0,1], (\mathbb{R}^d, ||\cdot||))$ *for all* $h : S^m \to \mathbb{R}^d$, *satisfying Condition 2.6.*

The proof of Lemma 3.4 is given after the proof of Theorem 2.10.

Proof of Theorem 2.10. 1.Step:
As in the proof of Theorem 2.4 in [13], we define a function F by $F(\varrho(\cdot)) = \int h \, d\varrho(\cdot)$. We want to prove a LDP for $F(M_n(\cdot))$. By Remark 3.3 the sequence $\{M_n(\cdot), n \in \mathbb{N}\}$ satisfies a LDP. Therefore we would like to apply the contraction principle (cf. [5], Theorem 4.2.23, and [18], Theorem 2.1). Since for arbitrary kernel functions h the integral $\int h d\varrho(t)$ is generally not defined for all $\varrho(t) \in \mathcal{M}^+(S^m)$, first we have to describe the images under F carefully. If $\varrho(\cdot) \in K_\infty^m$ we get $I^m(\varrho(\cdot)) = I(\nu(\cdot)) \leq K$ for some $K \geq 0$ and $\nu^m(\cdot) = \varrho(\cdot)$. For any fixed $t \in [0,1]$ the sequence $\{L_n(t), n \in \mathbb{N}\}$ satisfies the LDP in $\mathcal{M}^+(S)$, endowed with the weak topology, with the good rate function $t\, J(\nu/t)$, $\nu \in \mathcal{M}^t(S)$. Thus by assumption $\{M_n(t), n \in \mathbb{N}\}$ satisfies the LDP in $\mathcal{M}^+(S^m)$, endowed with the weak topology, too. The rate function has

the form $t\, J^m(\varrho) := t\, J(\nu)$, if $(\frac{\nu}{t})^m = \varrho$. Since $M_n(t)$ satisfies Condition 2.4, we get by [7], Lemma 3.2.7 and Lemma 3.2.19, that

$$t\, H(\varrho|\mu^m) \leq \beta(t\, J^m(\varrho) + \log(2M)), \quad \varrho \in \mathcal{M}_1(S^m), \tag{3.5}$$

and $\beta,\ M$ and μ as in Condition 2.4. Since by Lemma 3 in [3] we get

$$I(\nu(\cdot)) = \sup_{0=t_0<t_1<\ldots<t_k\leq 1, k\in\mathbb{N}} \sum_{i=1}^{k} (t_i - t_{i-1}) J(\frac{\nu(t_i)-\nu(t_{i-1})}{t_i - t_{i-1}}), \tag{3.6}$$

the convexity of $J(\cdot)$ yields

$$\begin{aligned} t_i\, H(\frac{\nu^m(t_i)}{t_i^m}|\mu^m) &\leq \beta(t_i\, J^m(\frac{\nu^m(t_i)}{t_i^m}) + \log(2M)) = \beta(t_i\, J(\frac{\nu(t_i)}{t_i}) + \log(2M)) \\ &\leq \beta(\sum_{j=1}^{i}(t_j - t_{j-1})\, J(\frac{\nu(t_j)-\nu(t_{j-1})}{t_j - t_{j-1}}) + \log(2M)) \\ &\leq \beta(K + \log(2M)) \end{aligned}$$

for each $t_i \in (0,1]$, $i \in \mathbb{N}$. Using the well-known inequality $xy \leq e^x + y\log y$ for $x, y \geq 0$ we get

$$t \int ||h|| \frac{d\varrho(t)}{t} \leq t \int \exp(||h||/t) d\mu^m + t\, H(\frac{\nu^m(t)}{t}|\mu^m)$$

and hence $\int ||h|| d\rho(t) < \infty$ for all $t \in [0,1]$ by Condition 2.6, whenever $I^m(\varrho(\cdot)) \leq K$.

Therefore the contraction map

$$F : \mathcal{C}_0([0,1], (\mathcal{M}^+(S^m), \beta^m)) \cap K_\infty^m \to \mathcal{C}_0([0,1], (\mathbb{R}^d, ||\cdot||))$$

with $F(\varrho(\cdot)) = \int h d\varrho(\cdot)$ is well defined on $K_\infty^m \cap \mathcal{C}_0([0,1], (\mathcal{M}^+(S^m), \beta^m))$ by an application of Lemma 3.4. By Remark 3.3, we have the LDP for $\{M_n(\cdot), n \in \mathbb{N}\}$ in $\mathcal{C}_0([0,1], (\mathcal{M}^+(S^m), \beta^m)) \cap K_\infty^m$.

2.Step:
If h is a continuous and bounded kernel function with $||h||_{BL} := ||h||_\infty + ||h||_L \leq C$ for any finite constant C we get the LDP for $\{\int hd\, M_n(\cdot), n \in \mathbb{N}\}$ (as well as for $\{F(\widetilde{M}_n(\cdot)), n \in \mathbb{N}\}$) as a consequence of Lemma 3.2 using the contraction principle (cf. Theorem 4.2.1 in [5] and Lemma 4.1.5 in [5]).

3.Step:
The main step in the proof will be an application of an extended version of the contraction principle (cf. [5], Theorem 4.2.23 and [18], Theorem 2.1). So let h be an arbitrary, measurable, and symmetric function. For $0 < L < \infty$ let

$$h_L(x_1, \ldots, x_m) := 1_{\{||h|| \leq L\}} h(x_1, \ldots, x_m)$$

be h truncated at height L. According to Lusin's theorem there exists a compact set $C \subseteq S^m$ with measure $\mu^m(C) \geq 1 - e^{-L^2}$ such that $h_L|_C$ is continuous. Due to the Stone-Weierstrass theorem we then can find a function $g_L \in BL(C) = \{f : C \subset S^m \to \mathbb{R}^d : ||f||_\infty + ||f||_L \leq 1\}$ such that

$$||g_L - h_L|_C|| \leq \frac{1}{L}$$

and $\sup_{x \in C} g_L(x) \leq \sup_{x \in C} h_L|_C(x) \leq L$. By the extension theorem of Kirszbraun and McShane (cf. [9], Theorem 7.3) g_L admits an extension $f_L \in BL(S^m)$ with $||f_L||_{BL} = ||g_L||_{BL}$. Now we define
$F_L : \mathcal{C}_0([0,1], (\mathcal{M}^+(S^m), \beta^m)) \cap K_\infty^m \to \mathcal{C}_0([0,1], (\mathbb{R}^d, ||\cdot||))$ by

$$F_L(\varrho(\cdot)) = \int f_L d\varrho(\cdot)$$

for $L > 0$. Note that F_L is well defined by the same arguments as above. Moreover F_L is continuous with respect to the topologies of uniform convergence by definition of β^m.

4.Step:
We have to check that

$$\limsup_{L \to \infty} \sup\{\sup_{t \in [0,1]} ||F(\varrho(t)) - F_L(\varrho(t))|| : I^m(\varrho(\cdot)) \leq K\} = 0 \tag{3.7}$$

for all $K > 0$. Here again (3.5) comes into play: For each $t \in (0,1]$ we get $t\,H(\frac{\varrho(t)}{t}|\mu^m) \leq K$ whenever $I^m(\varrho(\cdot)) \leq K$. Therefore with $\varrho(0) = 0$ we obtain $\{\varrho(\cdot) : I^m(\varrho(\cdot)) \leq K\} \subseteq \{\varrho(\cdot) : tH(\frac{\varrho(t)}{t}|\mu^m) \leq K, \forall t \in [0,1]\} =: N_K$. Now we can adopt the proof of Theorem 2.4 in [13] to get (3.7). For the sake of completeness we give the proof again: The triangle inequality yields $||h - f_L|| \leq ||h - h_L|| + ||h_L - f_L||$. A simple adaption of the proof of Lemma 5.1 in [8] shows that

$$\limsup_{L \to \infty} \sup_{\varrho(\cdot) \in N_K} \int ||h - h_L||\, d\varrho(\cdot) = 0$$

uniformly in t and for all $K > 0$. So (3.7) follows if we can prove

$$\limsup_{L \to \infty} \sup_{\varrho(\cdot) \in N_K} \int ||f_L - h_L||\, d\varrho(\cdot) = 0 \tag{3.8}$$

uniformly in t for all $K > 0$. To this end put $||h_L - f_L|| =: f_L^{(1)} + f_L^{(2)}$ with $f_L^{(1)} \leq 1/L$ and $\mu^m(f_L^{(2)} \neq 0) \leq e^{-L^2}$ and $f_L^{(1)}, f_L^{(2)} \geq 0$.
Clearly

$$\limsup_{L \to \infty} \sup_{\varrho(\cdot) \in N_K} \int f_L^{(1)} d\varrho(\cdot) = 0$$

uniformly in t and on each level set of $H(\cdot\,|\,\mu)$. With $x\,y \leq e^{\sigma\,x} + \frac{y}{\sigma} \log \frac{y}{\sigma}$ for all $x, y \geq 0$ and $\sigma > 0$ we obtain for $\sigma \geq 1$ as in the proof of Lemma 5.1 in [8]

$$
\begin{aligned}
t\int f_L^{(2)} d\frac{\varrho(t)}{t} &= t\int_{\{f_L^{(2)}\neq 0\}} f_L^{(2)} \frac{d\big(\frac{\varrho(t)}{t}\big)}{d\mu^m} d\mu^m \\
&\le \int_{\{f_L^{(2)}\neq 0\}} \exp\big(\sigma f_L^{(2)}\big) d\mu^m + \frac{1}{\sigma}\int_{S^m} t\left|\log \frac{d\big(\frac{\varrho(t)}{t}\big)}{d\mu^m}\right| d\big(\frac{\varrho(t)}{t}\big) \qquad (3.9)\\
&\le \exp(L(\sigma - L)) + \frac{1}{\sigma}\int_{S^m} t\left|\log \frac{d\big(\frac{\varrho(t)}{t}\big)}{d\mu^m}\right| d\big(\frac{\varrho(t)}{t}\big).
\end{aligned}
$$

Now the first term in (3.9) converges to zero for every fixed σ by letting $L\to\infty$ (independent of t and N_K). Since the second term on the right in (3.9) converges uniformly in t on N_K for each K we have proved (3.8) and therefore (3.7) by first letting $L\to\infty$ and then $\sigma\to\infty$.

5.Step:
The last step in applying the contraction principle is to verify the exponential equivalence between $F(\widetilde{M}_n(\cdot))$ and $F_L(\widetilde{M}_n(\cdot))$. We have to check

$$
\lim_{L\to\infty}\limsup_{n\to\infty}\frac{1}{n}\log P\big(\sup_{t\in[0,1]}||F(\widetilde{M}_n(t)) - F_L(\widetilde{M}_n)|| \ge \varepsilon\big) = -\infty \qquad (3.10)
$$

for every $\varepsilon > 0$. By definition the values of $W_n(\cdot) := \int_{S^m} h dM_n(\cdot)$ and $W_n(f_L)(\cdot) := \int_{S^m} f_L dM_n(\cdot)$ are constant on each interval $(i/n, (i+1)/n]$ for $i\in\{0,\dots,n-1\}$ and for all $t\in(i/n,(i+1)/n]$ the value at t is the value at i/n. In each intervall the oscillation $||F(\widetilde{M}_n(t) - F_L(\widetilde{M}_n)||$ is dominated by $\sup_{j\in\{i,i+1\}}||W_n(t_j) - W_n(f_L)(t_j)||$, where we have set $t_j := j/n$. Thus

$$
P\big(\sup_{t\in[0,1]}||F(\widetilde{M}_n(t)) - F_L(\widetilde{M}_n(t))|| \ge \varepsilon\big) \le \sum_{i=0}^{n-1} P\big(||W_n(t_i) - W_n(f_L)(t_i)|| \ge \varepsilon\big).
$$

By the Chebyshev-Markov inequality we obtain for any positive Θ

$$
\begin{aligned}
&P\big(\|W_n(t_i) - W_n(h_L)(t_i)\| > \frac{\varepsilon}{2}\big) \le \\
&\qquad\qquad \exp\left(-\frac{\Theta\varepsilon}{2}\right)\mathbb{E}_P\left(\exp(\Theta\|W_n(t_i) - W_n(h_L)(t_i)\|)\right).
\end{aligned}
$$

Applying Condition 2.4 to the right hand side we get, substituting Θ by $\Theta\, n$,

$$
\limsup_{n\to\infty}\frac{1}{n}\log P\big(\|W_n(h - h_L)(t_i)\| > \frac{\varepsilon}{2}\big) \le -\frac{\Theta\,\varepsilon}{2} + \log M\int_{S^m}\exp(\Theta\beta\|h - h_L\|)\, d\mu^m.
$$

By the arbitrariness of Θ we conclude

$$
\limsup_{L\to\infty}\limsup_{n\to\infty}\frac{1}{n}\log P\big(\|W_n(h - h_L)(t_i)\| > \frac{\varepsilon}{2}\big) = -\infty.
$$

Recall $\|h_L - f_L\| = f_L^{(1)} + f_L^{(2)}$. If we take L large such that $\frac{1}{L} < \frac{\varepsilon}{4}$ we get $P\big(\|W_n(f_L^{(1)})(t_i)\| > \frac{\varepsilon}{4}\big) = 0$ and thus it remains to show that

$$\limsup_{L\to\infty} \limsup_{n\to\infty} \frac{1}{n} \log P\big(\|W_n(f_L^{(2)})(t_i)\| > \frac{\varepsilon}{4}\big) = -\infty$$

for every positive ε. Again applying the Chebyshev-Markov inequality and Condition 2.4 we get

$$\limsup_{n\to\infty} \frac{1}{n} \log P\big(\|W_n(f_L^{(2)})(t_i)\| > \frac{\varepsilon}{4}\big) \le \log M - \frac{\Theta\varepsilon}{4} + \log\big(1 + \exp(L(\Theta - L))\big)$$

for every $\Theta, \varepsilon > 0$. So letting $L \to \infty$ and then $\Theta \to \infty$ proves (3.10). Thus we get the LDP for $\{F(\widetilde{M}_n(\cdot)), n \in \mathbb{N}\}$ in $L_\infty([0,1], (\mathbb{R}^d, \|\cdot\|))$ by Lemma 4.1.5 in [5]. $\square$

Proof of Lemma 3.4. The proof follows the ideas of the proof of Lemma 3.4 in [13]. Since we are only interested in the case $\nu(\cdot) \in K_\infty^m$, by the same arguments as in the proof of Theorem 2.10 the integrals $\int h d\nu(t)$ are defined for all $t \in [0,1]$. Now fix $\nu(\cdot) \in \mathcal{AC}_0$: By definition this means that for each $\varepsilon > 0$ there is a $\delta > 0$ such that for any finite sequences $0 \le s_1 < t_1 \le s_2 < t_2 \le \cdots \le s_k < t_k \le 1$

$$\sum_{l=1}^k (t_l - s_l) \le \delta \quad \text{implies} \quad \sum_{l=1}^k ||\nu(t_l) - \nu(s_l)||_{\text{var}} \le \varepsilon.$$

For $h_L := 1_{\{||h|| \le L\}} h$ and for a given $\varepsilon > 0$ we can choose $L > 0$ such that $|| \int h_L d\nu(t) - \int h d\nu(t)|| \le \varepsilon$ uniformly in $t \in [0,1]$, see (3.7). If we choose the sequence $0 \le s_1 < t_1 \le s_2 < t_2 \le \cdots \le s_k < t_k \le 1$ such that $\sum_{l=1}^k ||\nu(t_l) - \nu(s_l)||_{\text{var}} \le \frac{\varepsilon}{L}$, we get with $||\varrho_1 - \varrho_2||_{\text{var}} = \sup\{|\int f d\varrho_1 - \int f d\varrho_2| : ||f|| \le 1\}$:

$$\begin{aligned}\sum_{l=1}^k |\int h d\nu(t_l) - \int h d\nu(s_l)| &\le \sum_{l=1}^k \{|\int h d\nu(t_l) - \int h_L d\nu(t_l)| \\ &+ |\int h_L d\nu(t_l) - \int h_L d\nu(s_l)| + |\int h_L d\nu(s_l) - \int h d\nu(s_l)|\} < 3\varepsilon,\end{aligned} \tag{3.11}$$

and therefore $\int h d\nu(\cdot) \in \mathcal{AC}_0([0,1], (\mathbb{R}^d, ||\cdot||))$. The second statement of the lemma follows by the same steps with just replacing h_L by f_L as defined in the proof of Theorem 2.10. $\square$

Proof of Corollary 2.15. Assume that the $\int_0^1 \Lambda^*(\dot\phi(t))dt$ is finite. Using the identity (3.6) we get that (2.14) is greater than or equal to

$$\inf\{\sum_{i=1}^k (t_i - t_{i-1}) J(\frac{\nu(t_i) - \nu(t_{i-1})}{t_i - t_{i-1}}); \int x d\nu(t_i) = \phi(t_i), 1 \le i \le k\}$$

for arbitrary $0 = t_0 < t_1 < \ldots < t_k \le 1$. Let

$$J_k(\nu(t_1, \ldots, t_k)) := \sum_{i=1}^k (t_i - t_{i-1}) J(\frac{\nu(t_i) - \nu(t_{i-1})}{t_i - t_{i-1}}).$$

In the case $k = 1, t_1 = 1$ we have by assumption that

$$\inf\{J(\nu(1)); \int x d\nu(1) = \phi(1), \nu(1) \in K_{1,\infty}\} = \Lambda^*(\phi(1)).$$

Since $\mathcal{M}^+(S) \ni \varrho = (\varrho_1, \dots, \varrho_k) \mapsto (\int x d\varrho_1, \dots, \int x d\varrho_k)$ is linear and $\Lambda^*(\cdot)$ is convex we get by a simple calculation that

$$\begin{aligned} &\inf\{J_k(\nu(t_1, \dots, t_k)); \int x d\nu(t_i) = \phi(t_i), 1 \le i \le k\} \\ &= \sum_{i=1}^{k} (t_i - t_{i-1}) \Lambda^*(\frac{\phi(t_i) - \phi(t_{i-1})}{t_i - t_{i-1}}). \end{aligned} \tag{3.12}$$

Since this is true for every $0 = t_0 < t_1 < \dots < t_k \le 1$ we get by Lemma 5.1.6 in [5] that $\int_0^1 \Lambda^*(\dot\phi(t))dt$ is smaller than (2.14). In the case that $\int_0^1 \Lambda^*(\dot\phi(t))dt = \infty$ the same arguments as in the proof of Lemma 3.4 show that $\{\nu(\cdot) : \nu(\cdot) \in K_\infty$ and $\int x d\nu(\cdot) = \phi(\cdot)\} = \emptyset$. For the proof of the other inequality we can be adopt the proof of Corollary 2.7 in [13]. □

4. EXAMPLES

4.1. **Markov Chains.** Let $\pi : \mathcal{S} \times \mathcal{S} \to [0, 1]$ be a probability transition kernel and let $\{P_\sigma\}_{\sigma \in S}$ be the family of Markovian measures on the sequence space $(\Omega, \mathcal{F}) = (S^{\mathbb{N}_0}, \mathcal{S}^{\mathbb{N}_0})$ such that for every $\sigma \in S$, the projections $\{X_i, i \in \mathbb{N}\}$ from Ω to S form a Markov chain with transition kernel π and $P_\sigma(X_0 = \sigma) = 1$. We assume that there exist $N \in \mathbb{N}$, $l \in \{1, 2, \dots, N\}$ and $M \in [1, \infty)$ satisfying

$$\pi^l(\sigma, \cdot) \le \frac{M}{N} \sum_{m=1}^{N} \pi^m(\tau, \cdot) \tag{4.1}$$

for all $\sigma, \tau \in S$. As in [7, (4.1.39)] define the rate function $J_1 : \mathcal{M}^1(S) \to [0, \infty]$ by

$$J_1(\mu) = - \inf_{u \in B(S,[1,\infty))} \int_S \log \frac{\pi u}{u} \, d\mu.$$

Let $\mu \in \mathcal{M}^1(S)$ be given by

$$\mu = \frac{1}{N} \sum_{k=1}^{N} \pi^k(\sigma_0, \cdot) \tag{4.2}$$

with an arbitrary chosen $\sigma_0 \in S$. Let $m \ge 2$. For $n \ge ml$ define

$$\begin{aligned} A_{m,n} := \{\, &(i_1, \dots, i_m) \in \{l, l+1, \dots, n\}^m \mid \\ &|i_j - i_k| \ge l \text{ for all } j, k \in \{1, \dots, d\} \text{ with } j \ne k \,\}. \end{aligned} \tag{4.3}$$

Denote by $M_n(t)$ the corresponding partial sums process with respect to $A_{m,[nt]}$. We will prove:

Theorem 4.4. *Assume that Condition 2.6 holds for h and μ (defined by 4.2) uniformly in $\sigma_0 \in S$, then for every $m \in \mathbb{N}$ and every $\sigma \in S$ the measures*

$\{P_\sigma \circ (\int h dM_n(\cdot))^{-1}, n \in \mathbb{N}, \}$ *satisfy a LDP on* $L_\infty([0,1], (\mathbb{R}^d, \|\cdot\|))$ *with a good rate function* $I_\infty(\phi)$*, defined as in (2.11), where*

$$I(\nu(\cdot)) = \int_0^1 J_1(\dot{\nu}) dt \quad \textit{if } \nu \in \mathcal{AC}_0. \tag{4.5}$$

Remark 4.6. As mentioned in [7], Exercise 4.1.48, (4.1) is more than enough to guarantee that there exists precisely one $\nu \in \mathcal{M}_1(S)$ which is π-invariant: $\nu = \nu\pi$. Futhermore, analogously to Exercise 4.1.51(ii) in [7] we can assume that Condition 2.4 holds for h and ν.

Proof. With (4.1) the Markov chain satisfies Condition 2.1, thus the partial sums process $\{L_n(\cdot), n \in \mathbb{N}\}$ satisfies the LDP in $(D([0,1], (\mathcal{M}^+(S), \beta)), d_\infty)$ with rate function (4.5) (Theorem 3(a) in [3]). Since $|A_{m,[nt]}| \geq (\max\{1, [nt] - (2m-1)l\})^m$ for $[nt] \geq ml$ it follows that

$$\|L_n^{\otimes m}(t) - M_n(t)\|_{\text{var}} \leq \frac{[nt]^m - |A_{m,[nt]}|}{[nt]^m} \to 0$$

as $n \to \infty$ and $M_n(t)$ has the property (2.9). The measures $\{P_\sigma \circ (L_n^{\otimes m}(t))^{-1}\}_{\sigma \in S, [nt] \geq ml}$ and $\{P_\sigma \circ (M_n(t))^{-1}\}_{\sigma \in S, [nt] \geq ml}$ are uniformly exponentially equivalent in the sense that

$$\limsup_{n\to\infty} \frac{1}{n} \log \sup_{\sigma \in S} P_\sigma(\beta^m(L_n^{\otimes m}(t), M_n(t)) > \varepsilon) = -\infty$$

for every $\varepsilon > 0$ as a consequence of Lemma 3.1. Thus the remaining part of the proof is to check that $M_n(t)$ satisfies Condition 2.4 with μ defined in (4.2). Theorem 3.6 in [16] gives us the proof of this condition. We get the following estimate:

$$\begin{aligned} &\sup_{n \geq 4ml} \left(\sup_{\sigma \in S} \mathbb{E}_\sigma \left(\exp \left(n \int_{\mathcal{M}_1(S^m)} \varphi \, dM_n(t) \right) \right) \right)^{1/n} \\ &\qquad \leq M^m \int_{S^m} \exp(4^m m! \, l\varphi) \, d\mu^{\otimes m} \end{aligned} \tag{4.7}$$

for every $\varphi: S^m \to [0, \infty]$, bounded and measurable. For the sake of completeness we will give a sketch of the proof. With $n \geq ml$ and $A_{m,n}$ given by (4.3), let

$$A'_{m,n} = \{ (i_1, \ldots, i_m) \in A_{m,n} \mid i_j \geq i_{j-1} + l \text{ for all } j \in \{2, \ldots, d\} \}$$

denote the subset of all ordered m-tuples of $A_{m,n}$. We can find sets $C_n(r) \subset A'_{m,n}$ with the two properties:

(a) Every $i \in \{l, l+1, \ldots, n\}$ occurs at most once in at most one m-tuple contained in $C_n(r)$.
(b) If $i, i' \in \{l, l+1, \ldots, n\}$ are components of a m-tuple contained in $C_n(r)$ and $i \neq i'$ then $|i - i'| \geq l$,

such that

$$A'_{m,n} \subset \bigcup_{r \in B_{m,n}} C_n(r) \tag{4.8}$$

with $B_{m,n} = \{l, l+1, \dots, 3l-1\} \times \{2l, 2l+1, \dots, 2n\}^{m-1}$. Define $B'_{m,n} = \{ r \in B_{m,n} \mid C_n(r) \neq \emptyset \}$ and let $S(m)$ denote the set of all permutations of $\{1, \dots, m\}$. For every $\pi \in S(m)$ and $r \in B'_{m,n}$ define $C_n(\pi, r) = \{ (i_{\pi(1)}, \dots, i_{\pi(m)}) \mid (i_1, \dots, i_m) \in C_n(r) \}$. Every $C_n(\pi, r)$ has the corresponding properties (a) and (b) and it follows from (4.8) that

$$A_{m,n} \subset \bigcup_{\pi \in S(m)} \bigcup_{r \in B'_{m,n}} C_n(\pi, r). \tag{4.9}$$

Starting with the expectation on the left-hand side of (4.7), Hölder's inequality yields, for every $n \geq ml$ and $\sigma \in S$,

$$\begin{aligned}
&\mathbb{E}_\sigma \left(\exp\left(n \int_{\mathcal{M}^1(S^m)} \varphi \, dL_{n,m} \right) \right) \\
&\quad = \mathbb{E}_\sigma \left(\exp\left(\frac{n}{|A_{m,n}|} \sum_{(i_1,\dots,i_m) \in A_{m,n}} \varphi(X_{i_1}, \dots, X_{i_m}) \right) \right) \\
&\quad \leq \prod_{\pi \in S(m)} \prod_{r \in B'_{m,n}} \left(\mathbb{E}_\sigma \left(\exp\left(\frac{nm!\,|B'_{m,n}|}{|A_{m,n}|} \right. \right. \right. \\
&\qquad\qquad \left. \left. \left. \times \sum_{(i_1,\dots,i_m) \in C_n(\pi,r)} \varphi(X_{i_1}, \dots, X_{i_m}) \right) \right) \right)^{1/(m!|B'_{m,n}|)}
\end{aligned} \tag{4.10}$$

where we have used (4.9) for the last step. Since $|B'_{m,n}| \leq |B_{m,n}| \leq 2l(2n)^{m-1}$ and $|A_{m,n}| \geq (n-(2m-1)l)^m \geq (n/2)^m$ for $n \geq 4ml$, it follows that

$$\frac{nm!\,|B'_{m,n}|}{|A_{m,n}|} \leq 4^m m!\, l \quad \text{for } n \geq 4ml. \tag{4.11}$$

Given $\pi \in S(m)$ and $r \in B'_{m,n}$, it follows from (a) and (b) for $C_n(\pi, r)$, that the m-tuples in $C_n(\pi, r)$ consist of $p := m|C_n(\pi, r)|$ different $q_1, \dots, q_p \in \{l, l+1, \dots, n\}$, which we may label such that $q_j \geq q_{j-1} + l$ for all $j \in \{2, 3, \dots, p\}$. It follows from (4.1) and (4.2) that $\pi^{q_j - q_{j-1}}(\sigma, \cdot) \leq M\mu$ for every $j \in \{1, \dots, p\}$, where $q_0 := 0$. Hence for every $\sigma \in S$, $P_\sigma(X_{q_1}, \dots, X_{q_p})^{-1} \leq M^p \mu^{\otimes p}$ on $\mathcal{S}^{\otimes p}$. Using (4.11) it follows that

$$\begin{aligned}
&\mathbb{E}_\sigma \left(\exp\left(\frac{nm!\,|B'_{m,n}|}{|A_{m,n}|} \sum_{(i_1,\dots,i_m) \in C_n(\pi,r)} \varphi(X_{i_1}, \dots, X_{i_m}) \right) \right) \\
&\qquad \leq \left(M^d \int_{S^m} \exp(4^m m!\, l\varphi) \, d\mu^{\otimes m} \right)^{|C_n(\pi,r)|}.
\end{aligned} \tag{4.12}$$

Since $|C_n(\pi, r)| \leq \lfloor n/ml \rfloor \leq n$, the estimates (4.10) and (4.12) imply (4.7). $\square$

The LDP of $\int h dM_n(\cdot)$ can be transferred to the LDP of the partial sums U-processes under the same tail conditions, i.e.

$$\sup_{\sigma \in S} E_\sigma(\exp(\alpha ||h||)) < \infty$$

for all $\alpha > 0$. Actually, we need this condition in the proof of Theorem 4.4 only for the elements $(i_1, \dots, i_m)$ in the sets $C_n(\pi, r)$. In comparision to condition $(\hat{U})$ in [7], the stronger assumption for the U-processes LDP seem to be a natural generalization. Nevertheless, in the multivariate case the assumptions get a little more involved. The proof of the LDP for the partial sums U-processes now is a simple exercise. We only have to prove that the additional assumptions suffice to get the exponential equivalence of $\{U_n(\cdot), n \in \mathbb{N}\}$ and $\{\int h dM_n(\cdot), n \in \mathbb{N}\}$ in $L_\infty([0,1], (\mathbb{R}^d, \|\cdot\|))$, i.e.

$$\limsup_{n\to\infty} \frac{1}{n} \log P\big(\sup_{t\in[0,1]} \|U_n(t) - \int h dM_n(\cdot)\| > \delta\big) = -\infty.$$

Applying the Chebyshev-Markov inequality, the proof of Lemma 3.2 in [13] gives us the result.

Remark 4.13. By the same arguments we can transfer the LDP to the partial sums V-processes, assuming that Condition 2.7 holds uniformly in $\sigma \in S$ for μ defined in (4.2). We get the LDP for the laws of $V_\varepsilon(\cdot)$ and $U_\varepsilon(\cdot)$ analogously (see Theorem 2.6 in [13] and its proof).

As a Corollary we obtain the result of [3], Theorem 3(b), the $\mathbb{R}^d$-valued case for Markov chains which fulfill (4.1):

Corollary 4.14. *If the Markov-chain fulfills Condition 4.1 and*

$$\sup_{\sigma\in S} \int_S \exp(\alpha\|x\|)\pi(\sigma, dx) < \infty$$

for all $\alpha > 0$ (condition $\hat{U}$ in [7]), then the sequence $\{\mu_{\sigma,n}, n \in \mathbb{N}\}$, $\sigma \in S$, of laws of $Z_n(t) = 1/n \sum_{i=1}^{[nt]} X_i$, $t \in [0,1]$, satisfies for each $\sigma \in S$ the LDP in $L_\infty([0,1], (\mathbb{R}^d, \|\cdot\|))$ with the good rate function

$$I(\phi) = \int_0^1 \Lambda^*(\dot{\phi}(t))\, dt \quad \textit{if} \quad \phi \in \mathcal{AC}([0,1], (\mathbb{R}^d, \|\cdot\|)),\ \phi(0) = 0,$$

$I(\phi) = \infty$ otherwise. Here, for $x \in \mathbb{R}^d$, $\Lambda^*(x) := \sup_{\lambda\in\mathbb{R}^d}\{\langle\lambda, x\rangle - \Lambda(\lambda)\}$, where

$$\Lambda(\lambda) := \lim_{n\to\infty} \frac{1}{n} \log E[\exp(\sum_{i=1}^n \langle\lambda, X_i\rangle)].$$

Proof. With Exercise 2.1.20 (i) and Exercise 4.1.56 in [7], we get

$$\inf\{J_1(\nu(1)), \nu(1) \in \cap K_{1,\infty} \quad \text{and} \quad \int_S x d\nu(1) = \phi(1)\} = \Lambda^*(\phi(1)),$$

where $K_{1,\infty} = \bigcup_{L\geq 0}\{\varrho \in \mathcal{M}^1(S) :\ J_1(\varrho) \leq L\}$. Thus we are done by Corollary 2.15. $\square$

4.2. **Stationary, hypermixing sequences.** Let $\{X_i, i \in \mathbb{N}\}$ be a stationary sequence of random variables which take values in S. We assume that the hypermixing conditions (H1) and (H2) of [5] are fulfilled. We will treat only the case $m = 2$ and will suppose that an additional assumption is fulfilled:

Assumption 4.15. There exist $l_2 \in \mathbb{N}$ and $\gamma, \beta \in [1, \infty)$ such that for all $i \in \mathbb{N}$ and $f \in B(S^2, [0, \infty))$,

$$E_P\Bigg[\prod_{j=1}^{\lfloor i/2-l_2\rfloor} f(X_{i-j}, X_j)\Bigg] \leq \gamma E_P\Bigg[\prod_{j=1}^{\lfloor i/2-l_2\rfloor} f(X_{i-j}, \widetilde{X}_j)^{\beta}\Bigg]^{1/\beta},$$

where the process $\{\widetilde{X}_j\}_{j\in\mathbb{N}}$ is an independent copy of $\{X_j\}_{j\in\mathbb{N}}$.

Remark 4.16. There are some nice examples (see Exercise 6.4.19 in [5] and the continuous time Ornstein-Uhlenbeck process discussed in [7]), where (H1), (H2) and (4.15) are fulfilled, but (4.1) not. Notice moreover, that there are only a few results concerning U-Processes with mixing increments, see for example [2], where a central limit theorem for β-mixing sequences is studied.

For a fixed $l \in \mathbb{N}$ define

$$A_{2,n} := \{\, (i,j) \in \{1, \dots, n\}^2 \mid |i-j| \geq l \text{ for all } i \neq j \,\}. \tag{4.17}$$

Denote by $M_n(\cdot)$ the corresponding partial sums process with respect to $A_{2,[nt]}$. We will prove:

Theorem 4.18. *Assume that Condition 2.6 holds for h and $\mathcal{L}(X_1)$, then the measures $\{P \circ (\int h dM_n(\cdot))^{-1}, n \in \mathbb{N}\}$ and $\{P \circ (U_n(\cdot))^{-1}, n \in \mathbb{N}\}$ satisfy a LDP on $L_\infty([0,1], (\mathbb{R}^d, \|\cdot\|))$ with the good rate function $I_\infty(\phi)$, defined as in (2.11), where*

$$I(\nu(\cdot)) = \int_0^1 J(\dot{\nu})dt \quad \text{if } \nu \in \mathcal{AC}_0, \tag{4.19}$$

and J is defined by (2.3).

Proof. By Theorem 4 in [3] the partial sums process $\{L_n(\cdot), n \in \mathbb{N}\}$ satisfies the LDP in $(D([0,1], (\mathcal{M}^+(S), \beta)), d_\infty)$ with rate function (4.19), when condition (H-2) is fulfilled. We obtain

$$\|L_n^{\otimes 2}(t) - M_n(t)\|_{\text{var}} \to 0$$

as $n \to \infty$, therefore $M_n(t)$ has the property (2.9) and the measures $\{P \circ (L_n^{\otimes 2}(t))^{-1}\}_{[nt]\geq 2l}$ are exponentially equivalent to $\{P \circ (M_n(t))^{-1}\}_{[nt]\geq 2l}$. We have to check Condition 2.4. Using (H1) and Assumption 4.15, by the proof of Theorem 3.24 in [16] we get the following estimate:

$$\frac{1}{n}\log E_P\Bigg(\exp\Bigg(\frac{n}{|A_{[nt],l}|}\sum_{(i,j)\in A_{[nt],l}} \varphi(X_i, X_j)\Bigg)\Bigg)$$
$$\leq \log\gamma + \log E_P\big(\exp(16\beta k\alpha\varphi(X_1, \widetilde{X}_1))\big)$$

for every $\varphi : S^2 \to [0,\infty]$, bounded and measurable and the theorem is proved. For the sake of completeness again we will give a sketch of the calculations. Without loss of generality $\varphi : S^2 \mapsto [0,\infty)$ is a bounded and symmetric function. Remark that $|A_{2,n}| \geq n^2/4$ for $n \geq 8l$ and thus $\frac{n(n-l)}{|A_{2,n}|} \leq 4$ for $n \geq 8l$. Now, for $n \geq 8l$ we get

$$
\begin{aligned}
&E_P\biggl(\exp\biggl(\frac{n}{|A_{n,l}|}\sum_{(i,j)\in A_{n,l}}\varphi(X_i,X_j)\biggr)\biggr)\\
&\quad \leq E_P\biggl(\exp\biggl(\frac{n(n-l)}{|A_{n,l}|}\sum_{j=l}^{n-l}\varphi(X_{n+1-j},X_j)\biggr)\biggr)^{1/n-l}\\
&\qquad \times \biggl(\prod_{\substack{i=l+2\\ i\neq n+1}}^{2n-l} E_P\biggl(\exp\biggl(\frac{n2(n-l)}{|A_{n,l}|}\sum_{j=\max\{1,i-n\}}^{\min\{i-1,n\}}\varphi(X_{i-j},X_j)\\
&\qquad\qquad \times 1_{\{(i,j)\,:\,|i-2j|\geq l\}}\biggr)\biggr)\biggr)^{1/2(n-l)}\\
&\quad \leq \biggl(\prod_{i=l+2}^{n+1} E_P\biggl(\exp\biggl(8\sum_{j=1}^{i-1}\varphi(X_{i-j},X_j)1_{\{(i,j)\,:\,|i-2j|\geq l\}}\biggr)\biggr)\biggr)^{1/n-l},
\end{aligned}
$$

where we have applied Hölder's inequality and used the stationarity of the sequence $\{X_i\}_{i\in\mathbb{N}}$. We obtain by Assumption 4.15 for $l_2 - l$

$$
\begin{aligned}
&E_P\biggl(\exp\biggl(8\sum_{j=1}^{i-1}\varphi(X_{i-j},X_j)1_{\{(i,j)\,:\,|i-2j|>l\}}\biggr)\biggr)\\
&\qquad \leq \gamma E_P\biggl(\exp\biggl(16\beta\sum_{j=1}^{\lfloor\frac{i-l}{2}\rfloor}\varphi(X_{i-j},\widetilde{X}_j)\biggr)\biggr)^{1/\beta}.
\end{aligned}
$$

For each k with $k-1 \geq l$ we obtain

$$
\begin{aligned}
&E_P\bigl(\exp(\varphi(X_1,\widetilde{X}_k)+\varphi(X_k,\widetilde{X}_1))\bigr)\\
&\quad \leq \int_{S^2}\biggl(\int_S \exp(\alpha\varphi(x_1,\widetilde{x}_k)\mu_1(dx_1)\biggr)^{1/\alpha}\\
&\qquad \times \biggl(\int_S \exp(\alpha\varphi(x_k,\widetilde{x}_1)\mu_k(dx_k)\biggr)^{1/\alpha}\quad \mu(d\widetilde{x}_1,d\widetilde{x}_k)
\end{aligned}
$$

by Assumption (H1), where μ_1 and μ_k, respectively, denote the marginal distribution on the first and k-th component, respectively, and $\mu(d\widetilde{x}_1,d\widetilde{x}_k)$ is the two dimensional marginal distribution on the first and k-th component. Applying Assumption (H1) again, we get

$$
\begin{aligned}
&E_P\bigl(\exp(\varphi(X_1,\widetilde{X}_k)+\varphi(X_k,\widetilde{X}_1))\bigr)\\
&\quad \leq E_P\bigl(\exp(\alpha\varphi(X_1,\widetilde{X}_k))\bigr)^{1/\alpha}E_P\bigl(\exp(\alpha\varphi(X_k,\widetilde{X}_1))\bigr)^{1/\alpha}.
\end{aligned}
$$

Thus with $k \in \mathbb{N}$ and $m > 0$ such that $km = \lfloor \frac{i-l}{2} \rfloor$ we get

$$E_P\Big(\exp\Big(16\beta \sum_{j=1}^{\lfloor \frac{i-l}{2} \rfloor} \varphi(X_{i-j}, \widetilde{X}_j)\Big)\Big)^{1/\beta} \leq E_P\big(\exp(16\beta k\alpha\varphi(X_1, \widetilde{X}_1))\big)^{m/\alpha\beta}.$$

We arrive at

$$\begin{aligned}
&\frac{1}{n} \log E_P\Big(\exp\Big(\frac{n}{|A_{n,l}|} \sum_{(i,j)\in A_{n,l}} \varphi(X_i, X_j)\Big)\Big) \\
&\leq \frac{1}{n} \log \gamma^{n-l} + \frac{1}{n} \log\big(E_P(\exp(16\beta k\alpha\varphi(X_1, \widetilde{X}_1)))\big)^{\frac{n-l-1}{2k\alpha\beta}} \\
&\leq \log\gamma + \log E_P\big(\exp(16\beta k\alpha\varphi(X_1, \widetilde{X}_1))\big).
\end{aligned}$$

□

As a corollary we obtain Theorem 5 in [3], the $\mathbb{R}^d$-valued case for (H1)- and (H2)-mixing sequences.

Corollary 4.20. *Suppose that the stationary sequence $\{X_i, i \in \mathbb{N}\}$, taking values in $\mathbb{R}^d$, satisfies the mixing conditions (H1) and (H2) and $E(\exp(\alpha\|X\|)) < \infty$ for all $\alpha > 0$. Then the sequence $\{\mu_n, n \in \mathbb{N}\}$ of laws of $Z_n(t) = 1/n \sum_{i=1}^{[nt]} X_i$, $t \in [0,1]$, satisfies the LDP in $L_\infty([0,1], (\mathbb{R}^d, \|\cdot\|))$ with the good rate function*

$$I(\phi) = \int_0^1 \Lambda^*(\dot{\phi}(t))\, dt \quad \text{if} \quad \phi \in \mathcal{AC}([0,1], (\mathbb{R}^d, \|\cdot\|)),\ \phi(0) = 0,$$

$I(\phi) = \infty$ otherwise, where Λ^* is defined as in Corollary (4.14).

Proof. We have to check

$$\inf\{J(\nu(1)),\ \nu(1) \in \cap K_{1,\infty} \quad \text{and} \quad \int_S x d\nu(1) = \phi(1)\} = \Lambda^*(\phi(1)),$$

where $K_{1,\infty} = \bigcup_{L\geq 0}\{\varrho \in \mathcal{M}^1(S) : J(\varrho) \leq L\}$. To this end we can actually go along the proof of Theorem 5 in [4]. There the identification is checked. Here we only have to apply the first part of Proposition 1 in [4]. The proof is completed by applying Corollary 2.15. □

4.3. **Independent but not identically distributed sequences.** Let $\{X_i, i \in \mathbb{N}\}$ be a sequence of independent random variables with values in $\mathbb{R}^d$ and laws $\mathcal{L}(X_i) = \nu_i$. Let $\Lambda_k(\lambda) = \log E[\exp\langle\lambda, X_k\rangle]$ are such that for all $\lambda \in \mathbb{R}^d$ the limit

$$\Lambda(\lambda) = \lim_{n\to\infty} \frac{1}{m} \sum_{i=1}^{m} \Lambda_i(\lambda) \tag{4.21}$$

exists and $\Lambda(\cdot)$ is finite and differentiable. Then $\{L_n(1), n \in \mathbb{N}\}$ satisfies the LDP in $\mathcal{M}^1(\mathbb{R}^d)$ equipped with the weak topology with the convex rate function $J(\nu)$ defined in (2.3) (see [11], where the LDP is proved in an even stronger topology). Moreover $\{L_n(\cdot), n \in \mathbb{N}\}$ satisfies the LDP in $D([0,1], (\mathcal{M}^+(\mathbb{R}^d), \beta))$ equipped with the sup norm topology with the convex good rate function defined in (2.2) (see [11]).

Consider the following additional assumption: assume that $\nu_i \ll \mu$ for all $i \in \mathbb{N}$ and $\mu \in \mathcal{M}^1(S)$. Moreover we assume for $f_i := \frac{d\nu_i}{d\mu}$ that there exists a $q > 1$ such that

$$\sup_{i \in \mathbb{N}} \|f_i\|_q < \infty, \tag{4.22}$$

where $\|\cdot\|_q$ denotes the q-norm in $L_1(S, \mathcal{S}, \mu)$.

Theorem 4.23. *Consider a sequence $\{X_i, i \in \mathbb{N}\}$ as above. Assume that Condition 2.6 holds for h and μ. Then the measures $\{P \circ (U_n(\cdot))^{-1}, n \in \mathbb{N}\}$ satisfy a LDP on $L_\infty([0,1], (\mathbb{R}^d, \|\cdot\|))$ with a good rate function $I_\infty(\phi)$, defined as in (2.11).*

Proof. The partial sums process $\{L_n(\cdot), n \in \mathbb{N}\}$ satisfies the LDP in $(D([0,1], (\mathcal{M}^+(\mathbb{R}^d))), d_\infty)$ with rate function (2.2). We will check Condition 2.4 for $M_n(t) = L_n^m(t)$. For every $\varphi \in B(\mathbb{R}^{md})$ and the laws of $L_n^m(t)$ we get

$$\mathbb{E}_P \left(\exp \left(n \int_{S^m} \varphi \, dL_n^m(t) \right) \right) \leq \mathbb{E}_P \left(\exp(nB(X_{i_1}, \ldots, X_{i_n})) \right)$$

via Hoeffding's result in [17, Section 5], where

$$\begin{aligned} &B(X_1, \ldots, X_n) \\ &\quad := \frac{1}{k}[\varphi(X_1, \ldots, X_m) + \varphi(X_{m+1}, \ldots, X_{2m}) + \cdots + \varphi(X_{km-m+1}, \ldots, X_{km})] \end{aligned}$$

and $k := [\frac{n}{m}]$. Since $B(\cdot)$ is an average of independent random variables and

$$\begin{aligned} &\mathbb{E}_P(\exp(m\varphi(X_1, \ldots, X_m))) \\ &\quad \leq \left(\int_{S^m} \exp(pm\varphi(X_1, \ldots, X_m)) \, d\mu^{\otimes m} \right)^{1/p} \left(\int_{S^m} \left(\prod_{i=1}^m f_i(x_i) \right)^q d\mu^{\otimes m} \right)^{1/q} \end{aligned}$$

by Hölder's inequality, Condition 2.4 is fulfilled by our assumption $\sup_{i \in \mathbb{N}} \|f_i\|_q < \infty$ for a $q > 1$ and Condition 2.6. □

As a Corollary we finally obtain Corollary 1 in [3], the $\mathbb{R}^d$-valued case for independent but not identically distributed sequences.

Corollary 4.24. *Suppose that the X_i are independent $\mathbb{R}^d$-valued and $\Lambda_k(\lambda)$ are such that for all $\lambda \in \mathbb{R}^d$ $\Lambda(\lambda)$ (see (4.21)) exists and is finite and differentiable. Then the sequence $\{\mu_n, n \in \mathbb{N}\}$ of laws of $Z_n(t) = 1/n \sum_{i=1}^{[nt]} X_i$, $t \in [0,1]$, satisfies the LDP in $L_\infty([0,1], (\mathbb{R}^d, \|\cdot\|))$ with the good rate function*

$$I(\phi) = \int_0^1 \Lambda^*(\dot{\phi}(t)) \, dt \quad \text{if} \quad \phi \in \mathcal{AC}([0,1], (\mathbb{R}^d, ||\cdot||)), \, \phi(0) = 0,$$

$I(\phi) = \infty$ otherwise, where Λ^* is defined as in Corollary (4.14).

Proof. We have to check

$$\inf\{J(\nu(1)),\, \nu(1) \in \cap K_{1,\infty} \quad \text{and} \quad \int_S x d\nu(1) = \phi(1)\} = \Lambda^*(\phi(1)),$$

where $K_{1,\infty} = \bigcup_{L\geq 0}\{\varrho \in \mathcal{M}^1(S) : J(\varrho) \leq L\}$. Going along the line of Exercise 2.1.20 and 3.3.12 in [7], we only have to prove that there exists a lower semicontinuous function $f : [0,\infty) \to [0,\infty]$ such that $\lim_{x\to\infty} f(x)/x = \infty$ and

$$\sup_{n\in\mathbb{N}} \frac{1}{n} \log E\Big[\exp\big(\sum_{i=1}^n f(||X_i||)\big)\Big] < \infty$$

(see Proposition 1 in [4]). We choose $f(x) = 1/2 \sup_{\alpha\in[0,\infty)}\{\alpha x - \log E[\exp(\alpha x)]\}$. This function has the properties as desired using Lemma 2.2.20 in [5] and the inequality

$$E[\exp(f(||X_i||))] \leq 4 < \infty$$

(see inequality (5.1.6) in [5]): by the independence of the X_i we get

$$E\Big[\exp\big(\sum_{i=1}^n f(||X_i||)\big)\Big] \leq \prod_{i=1}^n E[\exp(f(||X_i||))] \leq 4^n.$$

$\square$

Finally we want to transfer the LDP to the partial sums V-processes. Here we can get a result under weaker conditions:

Assume that there exists at least one $\alpha > 0$ such that

$$\int_{S^m} \exp(\alpha\|h \circ \pi_\sigma\|)\, d\mu < \infty \tag{4.25}$$

for every map $\sigma : \{1,\dots,m\} \to \{1,\dots,m\}$, where $\pi_\sigma : S^m \to S^m$ is defined by $\pi_\sigma(s) = (s_{\sigma(1)},\dots,s_{\sigma(m)})$ for every $s = (s_1,\dots,s_m) \in S^m$, and μ as above.

Consider the case $t = 1$. For every $n \geq m$ we want to define a measurable function h_n such that in the notation of the proof of Theorem 2.10 we get

$$F(L_n^{\otimes m}) = F_n(L_n^m), \tag{4.26}$$

where $F_n(L_n^m) := \int h_n dL_n^m$. One possible choice is given in [15], i.e. for all $x = (x_1,\dots,x_m) \in S^m$ we set

$$h_n(x) = \sum_{j=1}^m \frac{n_{(j)}}{n^m} \sum_{\tau\in\mathcal{T}_j} h(x_{\tau(1)},\dots,x_{\tau(m)}), \tag{4.27}$$

where $\mathcal{T}_j$ denotes the set of all surjective maps $\tau : \{1,\dots,m\} \to \{1,\dots,j\}$ with $\tau(1) = 1$ and $\tau(k) \leq 1+\max\{\tau(1),\dots,\tau(k-1)\}$ for all $k \in \{2,\dots,m\}$ (see proof of Theorem 1.11(b) in [15]). Define $\widetilde{\alpha} = (m^{m-1}+m^2)^{-1}\alpha$ with α as in Condition 4.25.

By using the exponential Chebychev-Markov inequality, independence, Hoeffding's formula [17, Section 5] and Hölder's inequality we get

$$\begin{aligned}
&\mathbb{P}(\|U_n(1) - V_n(1)\| \geq \varepsilon) \\
&\quad = \mathbb{P}(\|F(L_n^m(1)) - F(L_n^{\otimes m}(1))\| \geq \varepsilon) \\
&\quad \leq e^{-\widetilde{\alpha}\varepsilon n^2}\, \mathbb{E}[\exp(\widetilde{\alpha} n^2 \|F(L_n^m(1)) - F_n(L_n^m(1))\|)] \\
&\quad \leq e^{-\widetilde{\alpha}\varepsilon n^2}\, \mathbb{E}[\exp(p\widetilde{\alpha} mn \|h(x) - h_n(x)\|)]^{[n/m]1/p} \\
&\qquad \left(\mathbb{E}\left(\prod_{i=1}^m f_i(x_i) \right)^q d\mu^{\otimes m} \right)^{[n/m]1/q}.
\end{aligned} \tag{4.28}$$

Note that $1 - n_{(m)}/n^m \leq m^2/n$. Using (4.27), it follows that

$$n\|h(x) - h_n(x)\| \leq m^2 \|h(x)\| + \sum_{j=1}^{m-1} \sum_{\tau \in \mathcal{T}_j} \|h \circ \pi_\tau(x)\|$$

for all $x \in S^m$. Using this estimate, $|\bigcup_{j=1}^{m-1} \mathcal{T}_j| \leq m^{m-1}$, Hölder's inequality and Condition 4.25 as well as the assumption $\sup_{i \in \mathbb{N}} \|f_i\|_q < \infty$ for a $q > 1$, it follows that the product of the integrals in (4.28) is bounded by a constant which does not depend on n. Hence, for any sequence $\{n_k\}_{k \in \mathbb{N}}$ satisfying $km \leq n_k$ for all $k \in \mathbb{N}$,

$$\limsup_{k \to \infty} \frac{1}{n} \log \mathbb{P}(\|F(L_n^m(1)) - F_n(L_n^m(1))\| \geq \varepsilon) = -\infty.$$

For $t \in [0,1)$ the calculations are identical. Thus we have proved that $\{U_n(\cdot), n \in \mathbb{N}\}$ and $\{V_n(\cdot), n \in \mathbb{N}\}$ are exponentially equivalent under the weaker moment Condition 4.25.

References

1. M. A. Arcones and E. Giné, *Limit theorems for U-processes*, Ann. Probab. **21** (1991), 1494–1542.
2. M. A. Arcones and B. Yu, *Central limit theorems for empirical and U-processes of stationary mixing sequences*, J. Theoret. Probab. **7** (1994), no. 1, 47–71.
3. A. Dembo and T. Zajic, *Large deviations: From emprical mean and measure to partial sums process*, Stochastic Process. Appl. **57** (1995), 191–224.
4. ———, *Uniform large and moderate deviations for functional empirical processes*, preprint, 1995.
5. A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Jones and Bartlett, Boston, MA, 1993.
6. M. Denker and G. Keller, *On U-statistics and v. Mises' statistics for weakly dependent processes*, Z. Wahrsch. Verw. Gebiete **64** (1983), 505–522.
7. J.-D. Deuschel and D. W. Stroock, *Large Deviations*, Academic Press, San Diego, 1989.
8. 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.
9. R. M. Dudley, *Probabilities and metrics*, Aarhus Universitet, Lecture notes series No. 45, Aarhus Universitet, Mathematisk Institut, 1976.
10. P. Eichelsbacher, *Large deviations for products of empirical probability measures in the τ-topology*, J. Theoret. Probab., (to appear).

11. ______, *Partial sums large deviations for independent random variables*, in preparation, 1997.
12. P. Eichelsbacher and M. Löwe, *Large deviation principle for m-variate von Mises-statistics and U-statistics*, J. Theoret. Probab. **8** (1995), 807–824.
13. ______, *Large deviations for partial sums U-processes*, SFB-preprint 95-115, 1995.
14. P. Eichelsbacher and U. Schmock, *Exponential approximations in completely regular topological spaces and extensions of Sanov's theorem*, SFB-preprint 96-109, 1996.
15. ______, *Large deviations of products of empirical measures and U-empirical measures in strong topologies*, SFB-preprint 96-027, 1996.
16. ______, *Large deviations for products of empirical measure of dependent random sequences in strong topologies*, preprint, 1997.
17. W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc. **58** (1963), 13–30.
18. V. Pérez-Abreu and C. Tudor, *Large deviations for a class of chaos expansions*, J. Theoret. Probab. **7** (1994), no. 4, 757–765.
19. T. Zajic, *Large exceedances for uniformly recurrent Markov-additive processes and strong mixing processes*, J. Appl. Probab. **32** (1995), no. 3, 679–691.

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

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