Moderate deviations for mean-field Gibbs measures

PETER EICHELSBACHER[1] and TIM ZAJIC[2]

[1]*Fakultät für Mathematik, Ruhr-Universität Bochum, NA 3/68, D-44780 Bochum, Germany.*
E-mail: peter.eichelsbacher@ruhr-uni-bochum.de
[2]*School of Mathematics, University of Minnesota, Minneapolis MN 55455, USA.*
E-mail: zajic@ima.umn.edu

We present a moderate-deviations principle around non-degenerate attractors of the empirical measure of random variables distributed according to a mean-field Gibbs measure. We state a result for a large class of densities of the Gibbs measure. This result is an application of a rank-dependent moderate-deviations principle for a collection of U-empirical measures. The results are applied to diffusion processes with mean-field interaction leading to a McKean–Vlasov limit, and to the Curie–Weiss model.

Keywords: Curie–Weiss model; decoupling; Gibbs measures; Langevin dynamics; mean field; moderate deviations; U-statistics

1. Introduction

Let S be a Polish space furnished with the Borel σ-algebra $\mathcal{S}$ and $\mathcal{M}_1(S)$ be the set of probability measures on $(S, \mathcal{S})$. For a function $\Gamma : \mathcal{M}_1(S) \to \mathbb{R}$ and for $\mu \in \mathcal{M}_1(S)$ we consider the *Gibbs measure* P_Γ^n defined by

$$P_\Gamma^n(\mathrm{d}X) := \frac{1}{Z_\Gamma^n} \exp\left\{ n\Gamma\left(\frac{1}{n} \sum_{i=1}^n \delta_{x_i} \right) \right\} \mathrm{d}\mu^{\otimes n}(X),$$

where $X = (x_1, \ldots, x_n) \in S^n$ and

$$Z_\Gamma^n := \int \exp\left\{ n\Gamma\left(\frac{1}{n} \sum_{i=1}^n \delta_{x_i} \right) \right\} \mathrm{d}\mu^{\otimes n}(X).$$

We are interested in the asymptotic behaviour of P_Γ^n when n goes to infinity. In this paper we are concerned with *moderate fluctuations* (moderate-deviations principle) of the empirical measure

$$\mu_n := \frac{1}{n} \sum_{i=1}^n \delta_{x_i}$$

under P_Γ^n. Fluctuations for mean-field interacting particles have been widely studied; see, for

example, Ben Arous and Brunaud (1989), Bolthausen (1986), Kusuoka and Tamura (1984), McKean (1975), Sznitman (1984) and Tanaka (1984).

If Γ is bounded and continuous on $\mathcal{M}_1(S)$ furnished with the strong topology (which is defined as the coarsest topology on $\mathcal{M}_1(S)$ such that the map $\mathcal{M}_1(S) \ni \nu \mapsto \int_S \varphi \, d\nu$ is continuous for every bounded measurable function $\varphi : S \to \mathbb{R}$) it is known (see Ben Arous and Brunaud 1989) that μ_n under P_Γ^n satisfies a *large-deviations principle* (LDP) – for a definition see Dembo and Zeitouni (1998) and below – with a good rate function

$$H(\nu) := H(\nu|\mu) - \Gamma(\nu) - \inf_{\tilde{\nu}\in\mathcal{M}_1(S)}\big\{H(\tilde{\nu}|\mu) - \Gamma(\tilde{\nu})\big\}, \tag{1.1}$$

where $H(\nu|\mu)$ is the *relative entropy* of ν with respect to μ:

$$H(\nu|\mu) = \begin{cases} \displaystyle\int \log\frac{d\nu}{d\mu} \, d\nu & \nu \ll \mu, \\ \infty & \text{otherwise.} \end{cases} \tag{1.2}$$

In the so-called *non-degenerate* case (for a precise definition see (1.5) and thereafter) μ_n converges to a convex combination of the minimizers of H; see Ben Arous and Brunaud (1989) and Kusuoka and Tamura (1984). If H achieves its minimal value at a unique probability measure μ^* which is a non-degenerate minimum (this is, roughly speaking, the property that H is strictly convex in a neighbourhood of μ^*), then Gaussian fluctuations around μ^* are expected. The most general assumptions on Γ to obtain Gaussian behaviour can be found in Ben Arous and Brunaud (1989), Guionnet (1999) and Kusuoka and Tamura (1984). The authors assume that Γ is of the form

$$\Gamma(\mu) = \sum_{k=2}^{r} \int V_k(x_1, \ldots, x_k) \, d\mu^{\otimes k}(x_1, \ldots, x_k), \tag{1.3}$$

for a finite integer r and symmetric functions V_k on S^k, so that there exist a compact metric space C_k, a signed measure ν_k on C_k with bounded total variation and a real-valued bounded continuous function $g_k : C_k \times S \to \mathbb{R}$ such that, for any $k \in \{2, \ldots, r\}$ and any $(x_1, \ldots, x_k) \in S^k$,

$$V_k(x_1, \ldots, x_k) = \int g_k(\tau, x_1) \, g_k(\tau, x_2) \cdots g_k(\tau, x_k) \, d\nu_k(\tau). \tag{1.4}$$

Guionnet (1999) assumed that V_2 is symmetric and bounded and satisfies a special *regularity* condition, and hence considered a larger class of functions Γ.

The *Hessian* Ξ of Γ is the symmetric operator in the subspace $L_0^2(\mu^*) = \{\varphi \in L^2(\mu^*) : \int \varphi \, d\mu^* = 0\}$ of $L^2(\mu^*)$ so that, for any $\varphi \in L_0^2(\mu^*)$ such that, for $\varepsilon \in \mathbb{R}$ small enough, $(1 + \varepsilon\,\varphi).\mu^* \in \mathcal{M}_1(S)$,

$$\langle\varphi, \Xi\varphi\rangle_{L_0^2(\mu^*)} := \lim_{\varepsilon\searrow 0}\frac{1}{2\varepsilon^2}\big\{\Gamma((1+\varepsilon\,\varphi).\mu^*) + \Gamma((1-\varepsilon\,\varphi).\mu^*) - 2\Gamma(\mu^*)\big\}. \tag{1.5}$$

We remark that, for Γ given in (1.3) with symmetric V_k,

$$\langle \varphi, \Xi\varphi \rangle_{L_0^2(\mu^*)} = \sum_{k=2}^{r} k(k-1) \int_{S^k} \varphi(x_1)\varphi(x_2) V_k(x_1, x_2, y_1, \ldots, y_{k-2}) \mathrm{d}(\mu^*)^{\otimes k}.$$

If μ^* is a minimizer of H, it is called *non-degenerate* if $\mathrm{Id} - \Xi$ is positive definite on $L_0^2(\mu^*)$.

We will assume the following:

Condition 1.1. *H in* (1.1) *achieves its minimum value at a unique probability measure μ^*, and μ^* is a non-degenerate minimum.*

Throughout this paper we denote the space of bounded measurable functions on S by $B(S)$ and the space of bounded continuous functions by $C_b(S)$.

Let us recall the definition of the moderate-deviations principle (MDP). A sequence of probability measures $\{\mu_n\}_{n\in\mathbb{N}}$ on a topological space $\mathcal{X}$ equipped with σ-field $\mathcal{B}$ is said to satisfy the LDP with speed $a_n \downarrow 0$ and good rate function $I(\cdot)$ if the level sets $\{x : I(x) \leq \alpha\}$ are compact for all $\alpha < \infty$, and for all $B \in \mathcal{B}$ the lower bound

$$\liminf_{n\to\infty} a_n \log \mu_n(B) \geq - \inf_{x\in\mathrm{int}(\Gamma)} I(x),$$

and the upper bound

$$\limsup_{n\to\infty} a_n \log \mu_n(B) \leq - \inf_{x\in\mathrm{cl}(\Gamma)} I(x)$$

hold, where $\mathrm{int}(B)$ and $\mathrm{cl}(B)$ denote the interior and closure of B, respectively. We say that a sequence of random variables satisfies the LDP provided the sequence of measures induced by these variables satisfies the LDP. Let $\{b_n\}_{n\in\mathbb{N}} \subset (0, \infty)$ be a sequence satisfying

$$\lim_{n\to\infty} \frac{b_n}{n} = 0 \quad \text{and} \quad \lim_{n\to\infty} \frac{n}{b_n^2} = 0. \tag{1.6}$$

If $\mathcal{X}$ is a topological vector space, then a sequence of random variables $\{Z_n\}_{n\in\mathbb{N}}$ satisfies the MDP with speed n/b_n^2 and with good rate function $I(\cdot)$ if the sequence $\{(n/b_n)Z_n\}_{n\in\mathbb{N}}$ satisfies the LDP in $\mathcal{X}$ with the good rate function $I(\cdot)$ and with speed n/b_n^2.

We will prove an MDP for the laws of $(1/n)\sum_{i=1}^{n}(\delta_{X_i} - \mu^*)$ under P_Γ^n if Condition 1.1 is fulfilled with V_k satisfying (1.4) for $k \geq 3$ and a symmetric, bounded V_2 – that is, an LDP of

$$M^n := M^n(\mu^*) := \frac{1}{b_n} \sum_{i=1}^{n} (\delta_{x_i} - \mu^*) \tag{1.7}$$

under P_Γ^n. The assumption of regularity of V_2 in Guionnet (1999) is relaxed here. Note that a bounded and continuous V_2 is regular but a bounded measurable V_2 can fail to be regular by Dembo and Zeitouni (1998, Exercise 7.3.18).

Consider the set $\mathcal{M}(S)$ of signed measures on $(S, \mathcal{S})$ with finite total variation endowed with the strong topology. Let $\mu \in \mathcal{M}_1(S)$ be the law of an independent and identically distributed (i.i.d.) sequence $\{X_i\}_{i\in\mathbb{N}}$. In de Acosta (1994) it is proved that the sequence

$$\hat{\mu}_n := \frac{1}{b_n} \sum_{i=1}^{n} (\delta_{X_i} - \mu)$$

satisfies an LDP with speed n/b_n^2 in $\mathcal{M}(S)$ with respect to the strong topology for every (b_n) satisfying (1.6) with good rate function

$$I(\nu) = \frac{1}{2} \int_S \left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu} \right)^2 \mathrm{d}\mu,$$

if $\nu(S) = 0$ and $\nu \ll \mu$ and $+\infty$ otherwise.

Let us consider a Hamiltonian $\Gamma : \mathcal{M}(S) \to \mathbb{R}$ given by

$$\Gamma(\nu) = \int W(x, y) \, \mathrm{d}\nu(x) \, \mathrm{d}\nu(y).$$

If Γ is assumed to be continuous with respect to the strong topology on $\mathcal{M}(S)$ we may hope to obtain an LDP for the empirical measure $\hat{\mu}_n$ under P_Γ^n by applying the results of de Acosta (1994) in the i.i.d. case and Dembo and Zeitouni (1998, Theorem 4.3.1 and Exercise 4.3.11). There are at least two reasons why such an approach is not very useful. First, the map $\mathcal{M}(S) \ni \nu \mapsto \nu \otimes \nu$ is *not continuous* even in the case where $\mathcal{M}(S)$ and $\mathcal{M}(S^2)$ are endowed with the weak topology (see Fuglede 1960, Lemma 1.2.4, p. 148). Thus, in general, even the class of bounded and continuous functions W is not included. Moreover, applying the results of de Acosta (1994) and Dembo and Zeitouni (1998 Theorem 4.3.1 and Exercise 4.3.11) is only of use for the bivariate case because otherwise the corresponding Gibbs measures are not of the mean-field type we consider in this paper.

We therefore propose another method of proving the MDP for mean-field interacting particle systems with bounded interaction. Besides applying standard techniques from large-deviations theory (contraction principle, Varadhan's lemma, Dawson–Gärtner projective limit approach) the main tool in our proofs is to establish the MDP for a collection of so-called rank-dependent U-statistics and U-empirical measures. For example, in the case $\Gamma(\mu) = \int W(x, y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y)$, $\mu \in \mathcal{M}_1(S)$, with bounded measurable W we will need the LDP for

$$\left(\frac{1}{b_n} \sum_{i=1}^{n} (f(X_i) - \mu^*(f)), \frac{1}{b_n^2} \sum_{i \neq j} W(X_i, X_j) \right), \tag{1.8}$$

where $\{X_i\}_{i \in \mathbb{N}}$ is an i.i.d. sequence and $f : S \to \mathbb{R}$ is a bounded measurable function. In Eichelsbacher and Schmock (2001) an LDP is proved for the second term in (1.8). The proof of an LDP for the sequence (1.8) is far from obvious because of the dependency of the entries in (1.8).

The paper is organized as follows. In Section 2 we state our main results. We also state some exponential estimates deduced from results in Ben Arous and Brunaud (1989), Guionnet (1999) and Kusuoka and Tamura (1984) which are proven by the control on U-statistics developed in Arcones and Giné (1993) and de la Peña (1992) (decoupling and randomization techniques; see de la Peña and Giné 1999). In Section 3 we state the rank-dependent MDP for vectors of U-empirical measures in a general setting. In Section 4 we

prove the MDP stated in Section 3, and in Section 5 we prove our main results. Finally, in Section 6, we consider applications to the Curie–Weiss model and to diffusion processes with mean-field interaction, considered in Ben Arous and Brunaud (1989), Ben Arous and Zeitouni (1999), Kusuoka and Tamura (1984) and McKean (1975).

2. Moderate deviations for Gibbs measures

In this section we state the LDP for $M^n(\mu^*)$ under P_Γ^n in $\mathcal{M}(S)$ endowed with the strong topology, where μ^* is defined in Condition 1.1. First, we consider the case in which Γ is given by

$$\Gamma(\mu) = \int W(x, y)\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y), \qquad \mu \in \mathcal{M}_1(S). \tag{2.1}$$

(Later, we allow for more general types of polynomial interaction.) We denote the Hessian Ξ in this case by $2W$. Our first result is the following theorem:

Theorem 2.1. (Moderate deviations for bivariate potentials). *If Condition* 1.1 *holds and W in* (2.1) *is symmetric, bounded and measurable, then the sequence $\{M^n(\mu^*)\}_{n=1}^\infty$ satisfies the LDP under P_Γ^n in $\mathcal{M}(S)$, equipped with the strong topology, with speed n/b_n^2 and with good convex rate function*

$$I(\nu) = \begin{cases} \dfrac{1}{2}\displaystyle\int_S \left(\dfrac{\mathrm{d}\nu}{\mathrm{d}\mu^*}\right)^2 \mathrm{d}\mu^* - \int_{S^2} W\,\mathrm{d}\nu\,\mathrm{d}\nu & \nu(S) = 0, \\ \infty & otherwise. \end{cases} \tag{2.2}$$

In the course of proving this theorem we will see that we obtain another form of the rate function (2.2) using the projective limit approach (see (5.8)).

In the next theorem we will consider polynomial interactions. We will assume the following:

Condition 2.1. *Let Γ have the form*

$$\Gamma(\mu) = \sum_{k=2}^m \int_{S^k} W_k(x_1, \ldots, x_k)\mathrm{d}\mu^{\otimes k}(x_1, \ldots, x_k). \tag{2.3}$$

Let W_2 be a bounded measurable and symmetric function. For every $k \geqslant 3$, let W_k be in $\widetilde{\otimes}_\pi^k C_b(S)$, the k-fold projective tensor product of $C_b(S)$, and symmetric.

Remark 2.1. If Γ satisfies Condition 2.1, by Corollary 1.6 in Ben Arous and Brunaud (1989), there exists a compact metric space C and a signed measure ν with finite total variation and a bounded continuous function $h : C \times S \to \mathbb{R}$ such that

$$\sum_{k=3}^m \int_{S^k} W_k(x_1, \ldots, x_k)\mathrm{d}\mu^{\otimes k}(x_1, \ldots, x_k) = \int_{S^m} W\mathrm{d}\mu^{\otimes m}(x_1, \ldots, x_m),$$

with

$$W = \int_C h^{\otimes m}(\tau)\, \nu(\mathrm{d}\tau). \qquad (2.4)$$

Theorem 2.2 (Moderate deviations for polynomial interaction). *Assume that Γ satisfies Conditions* 1.1 *and* 2.1. *We then have the LDP holding for $\{M^n(\mu^*)\}_{n=1}^{\infty}$ in $\mathcal{M}(S)$ equipped with the strong topology with speed n/b_n^2 and with good convex rate function given by*

$$I(\nu) = \begin{cases} \dfrac{1}{2}\displaystyle\int_S \left(\dfrac{\mathrm{d}\nu}{\mathrm{d}\mu^*}\right)^2 \mathrm{d}\mu^* - \displaystyle\sum_{k=2}^{m} \dfrac{k(k-1)}{2} \int_{S^k} W_k \mathrm{d}(\mu^*)^{\otimes k-2} \mathrm{d}\nu\, \mathrm{d}\nu & \nu(S) = 0, \\ \infty & \text{otherwise}. \end{cases} \qquad (2.5)$$

The following corollary presents a special case of Theorem 2.2 which we will apply in Section 6.2.

Corollary 2.3. *Suppose Condition* 1.1 *holds and $W_k = 0$, $k = 2, \ldots, m-1$. Denoting $V = W_m$, the sequence $\{M^n(\mu^*)\}_{n \geqslant 1}$ satisfies the LDP on $\mathcal{M}(S)$ equipped with the strong topology with speed n/b_n^2 and good convex rate function*

$$I(\nu) = \begin{cases} \dfrac{1}{2}\displaystyle\int_S \left(\dfrac{\mathrm{d}\nu}{\mathrm{d}\mu^*}\right)^2 \mathrm{d}\mu^* - \dfrac{m(m-1)}{2} \int_{S^m} V\, \mathrm{d}(\mu^*)^{\otimes m-2} \mathrm{d}\nu\, \mathrm{d}\nu & \nu(S) = 0, \\ \infty & \text{otherwise}. \end{cases}$$

To be able to apply Varadhan's lemma we have to obtain strong upper bounds on exponential moments of U-statistics. The following is based on results of de la Peña (1992) and Arcones and Giné (1993) as well as the considerations in Eichelsbacher and Zajic (2002) and Guionnet (1999). We need to prove the following lemma.

Lemma 2.4. *Assume that Γ satisfies Conditions* 1.1 *and* 2.1. *Then, for some $\gamma > 1$, one obtains, in the notation of* (3.9) *below,*

$$\sup_{n \geqslant 1} \mathrm{E}_{(\mu^*)^{\otimes n}}\left[\exp\left(\gamma \frac{b_n^2}{n} \sum_{k=2}^{r} M_n^{k,2}(W_k) \right) \right] < \infty. \qquad (2.6)$$

Before we give the proof of the lemma let us remark that under the conditions for the potential functions W, $(W_k)_{k \geqslant 2}$ and V in Theorems 2.1 and 2.2 and Corollary 2.3 respectively, the LDP of μ_n under P_Γ^n holds. This follows basically from Theorems 1.7 and 1.19 and Lemma 1.21 in Eichelsbacher and Schmock (2002). The result is the following:

Theorem 2.5. *Assume that the interaction Γ is given by the right-hand side of* (2.3) *with measurable and bounded functions W_k. If $\mathcal{M}_1(S)$ is furnished with the strong topology then μ_n under P_Γ^n satisfies an LDP with good rate function defined in* (1.1).

Proof. The interaction can be written as a function $\tilde{\Gamma}$ on $\mathcal{M}_1(S^m)$:

$$\tilde{\Gamma}(\mu) = \sum_{k=2}^{m} \int_{S^k} W_k(x_1, \ldots, x_k) \mathrm{d}\mu_k(x_1, \ldots, x_k),$$

where μ_k denotes the marginal distribution of μ on the first k components. If the spaces $\mathcal{M}_1(S^k)$, $k = 2, \ldots, m$, are furnished with the strong topology as well, by our assumptions $\tilde{\Gamma}$ is a continuous function. Hence, applying Theorem 2.9 in Eichelsbacher and Schmock (2002), the sequence $\mu_n^{\otimes m}$ under P_Γ^n satisfies and LDP. Since the projection map $\mathcal{M}_1(S^m) \ni \mu \mapsto \mu_1 \in \mathcal{M}_1(S)$ is continuous, the contraction principle (see Dembo and Zeitouni 1998, Theorem 4.2.1) implies the result. $\qquad\square$

Remark 2.2. Actually the results in Eichelsbacher and Schmock (2002) give an LDP for μ_n under P_Γ^n even for unbounded functions W_k which satisfy some exponential moment conditions; see Lemma 1.21 in Eichelsbacher and Schmock (2002).

Proof of Lemma 2.4. We would like to be able to replace W_2 by a function $\tilde{W}_2$ of the form (1.4) and maintain the positive definiteness of $\mathrm{Id} - \Xi$. If W_2 were regular we could apply Guionnet (1999, Lemma 3.1). However, the proof of the latter lemma continues to hold for bounded measurable W_2 if, in the proof of the latter lemma, we apply Sanov's theorem for products of empirical measures with a sufficiently strong topology (see Eichelsbacher and Schmock 2002, Theorems 1.7 and 1.10, for such a topology) and apply the contraction principle followed by Varadhan's lemma to Theorems 1.7 and 1.10 of Eichelsbacher and Schmock (2002) in order to obtain inequality (43) of Guionnet (1999). For details, see Eichelsbacher and Zajic (2002). Let $\tilde{W}_k = W_k$, for $3 \leqslant k \leqslant r$, and $\tilde{\Gamma} = \Gamma$ with the W_i replaced by $\tilde{W}_i$, $2 \leqslant i \leqslant r$.

By Remark 2.1, $\tilde{\Gamma}$ can be represented as $\tilde{\Gamma}(\mu) = \langle W, \mu^{\otimes m} \rangle$ with W given in (2.4). Consider $\mathcal{B}(\tilde{\Gamma}) := L^m(C, \nu)$ and $T : \mathcal{M}_1(S) \to \mathcal{B}(\tilde{\Gamma})$ defined by $T(\mu) = \int_S h(x, \cdot) \mathrm{d}\mu(x)$; then we define

$$B_\delta(\mu^*) := \{\mu \in \mathcal{M}_1(S) : \|T(\mu - \mu^*)\| < \delta\}$$

with $\|\cdot\| = \|\cdot\|_{L^m(\nu)}$. Then $B_\delta(\mu^*)$ is an open ball around μ^* in the strong topology. Proceeding as in Kusuoka and Tamura (1984, Section 4), we have that

$$\mathrm{E}_{(\mu^*)^{\otimes n}} \left[1_{(B_\delta(\mu^*))^c} \exp \left(\frac{b_n^2}{n} \sum_{k=2}^{m} M_n^{k,2}(W_k) \right) \right]$$

$$= \mathrm{E}_{\mu^{\otimes n}} \left[1_{(B_\delta(\mu^*))^c} \exp \left(\frac{b_n^2}{n} \sum_{k=2}^{m} L_n^k(W_k) \right) \right] \exp\left(n(H(\mu^*|\mu) - \Gamma(\mu^*)) \right)$$

so that the LDP implies that, for any $\delta > 0$,

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathrm{E}_{(\mu^*)^{\otimes n}} \left[1_{(B_\delta(\mu^*))^c} \exp \left(\frac{b_n^2}{n} \sum_{k=2}^{m} M_n^{k,2}(W_k) \right) \right] \leqslant - \inf_{\nu \in (B_\delta(\mu^*))^c} H(\nu).$$

Using the fact that H is a good rate function and that μ^* is a unique minimizer, we obtain that there exists a positive constant c such that, for n large enough,

$$\left| \mathrm{E}_{(\mu^*)^{\otimes n}} \left[\exp \left(\frac{b_n^2}{n} \sum_{k=2}^m M_n^{k,2}(W_k) \right) \right] - \mathrm{E}_{(\mu^*)^{\otimes n}} \left[1_{(B_\delta(\mu^*))} \exp \left(\frac{b_n^2}{n} \sum_{k=2}^m M_n^{k,2}(W_k) \right) \right] \right| \leqslant \exp(-cn).$$

Hence it remains to prove that

$$\limsup_{n \to \infty} \mathrm{E}_{(\mu^*)^{\otimes n}} \left[1_{(B_\delta(\mu^*))} \exp \left(\gamma \frac{b_n^2}{n} \sum_{k=2}^m M_n^{k,2}(\tilde{W}_k) \right) \right] < \infty.$$

But this follows from Kusuoka and Tamura (1984, Lemma 3.2, (3.28)) and Ben Arous and Brunaud (1989, Theorem (B)iv, (3.4)).   □

3. Rank-dependent moderate deviations

Let us state the moderate-deviations result for vectors of U-empirical measures of different ranks. Let $(S, \mathcal{S}, \mu)$ be a probability space, let $(\Omega, \mathcal{A}, \mathbb{P}) \equiv (S^{\mathbb{N}}, \mathcal{S}^{\otimes \mathbb{N}}, \mu^{\otimes \mathbb{N}})$ be the product space, and let $\{X_i\}_{i \in \mathbb{N}}$ be the coordinate projections from Ω to S forming an i.i.d. sequence with $\mathcal{L}(X_i) = \mu$. Let $(E, \|\cdot\|_E)$ be a separable real Banach space with Borel σ-algebra $\mathcal{E}$. Let $m \in \mathbb{N}$ and $\mathcal{M}(S^m)$ denote the set of signed measures with finite total variation on the product space $(S^m, \mathcal{S}^{\otimes m})$. Let Φ be a collection of $\mathcal{S}^{\otimes m}$–$\mathcal{E}$-measurable and Bochner $\mu^{\otimes m}$-integrable functions $\varphi : S^m \to E$ containing the set $B(S^m, E)$ of all bounded measurable ones. We define the Φ-restricted set of signed measures by

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

where $|\nu|$ denotes the total variation measure corresponding to ν. Let $\tau^{\Phi}(E)$ be the coarsest topology on $\mathcal{M}^{\Phi}(S^m)$ such that $\mathcal{M}^{\Phi}(S^m) \ni \nu \mapsto \int_{S^m} \varphi \, d\nu$ is continuous for every $\varphi \in \Phi$. If $\Phi = B(S^m, E)$, then we write $\tau(E)$ instead of $\tau^{\Phi}(E)$. The σ-algebra on $\mathcal{M}(S^m)$ is defined to be the smallest one containing $\mathcal{M}^{\Phi}(S^m)$ such that all the maps $\mathcal{M}(S^m) \ni \nu \mapsto \int_{S^m} \varphi \, d\nu$ with $\varphi \in B(S^m, E)$ are measurable.

The *U-empirical measure of order m* is defined by

$$L_n^m = \frac{1}{n_{(m)}} \sum_{(i_1, \ldots, i_m) \in I(m,n)} \delta_{(X_{i_1}, \ldots, X_{i_m})} \tag{3.1}$$

for all integers $n \geqslant m$, where $n_{(m)} \equiv \prod_{k=0}^{m-1}(n-k)$ and the set $I(m, n) \subset \{1, \ldots, n\}^m$ consists of all m-tuples with pairwise different components. Given a measurable function φ on S^m, the *U-statistic* of order m with kernel function φ is defined, for every $n \geqslant m$, as

$$U_n^m(\varphi) \equiv \int_{S^m} \varphi \, dL_n^m = \frac{1}{n_{(m)}} \sum_{(i_1, \ldots, i_m) \in I(m,n)} \varphi(X_{i_1}, \ldots, X_{i_m}). \tag{3.2}$$

For the moderate deviations upper bound in Theorem 3.1 below, we have to formulate moment conditions for the unbounded $\varphi \in \Phi$. This in turn requires a decomposition, which

for symmetric φ is closely related to the Hoeffding decomposition for the corresponding U-statistic.

Given $\varphi \in L_1(\mu^{\otimes m}, E)$ and a non-empty subset A of $\{1, \ldots, m\}$, define $\varphi_A \in L_1(\mu^{\otimes |A|}, E)$ by μ-integrating $\varphi(s_1, \ldots, s_m)$ with respect to every s_i with $i \in \{1, \ldots, m\} \setminus A$. By convention, $\varphi_\varnothing \equiv \int_{S^m} \varphi \, d\mu^{\otimes m} \in E$. Furthermore, define $\tilde{\varphi}_A \in L_1(\mu^{\otimes |A|}, E)$ by

$$\tilde{\varphi}_A((s_i)_{i \in A}) = \sum_{B \subset A} (-1)^{|A \setminus B|} \varphi_B((s_i)_{i \in B}), \tag{3.3}$$

for every non-empty $A \subset \{1, \ldots, m\}$, and let $\tilde{\varphi}_\varnothing \equiv \varphi_\varnothing$. Note that

$$\tilde{\varphi}_A((s_i)_{i \in A}) = \int_{S^m} \varphi \, d\left(\bigotimes_{i=1}^m (1_A(i)(\delta_{s_i} - \mu) + 1_{A^c}(i)\mu) \right). \tag{3.4}$$

According to the inclusion–exclusion principle or the Möbius inversion formula,

$$\varphi(s_1, \ldots, s_m) = \sum_{A \subset \{1, \ldots, m\}} \tilde{\varphi}_A((s_i)_{i \in A})$$

for $\mu^{\otimes m}$-almost all $(s_1, \ldots, s_m) \in S^m$. Hence, for every $n \geqslant m$,

$$\int_{S^m} \varphi \, dL_n^m = \tilde{\varphi}_0 + \sum_{a=1}^m \int_{S^a} \tilde{\varphi}_a \, dL_n^a \tag{3.5}$$

$\mathbb{P}$-almost surely, where, for every $a \in \{0, 1, \ldots, m\}$,

$$\tilde{\varphi}_a \equiv \sum_{\substack{A \subset \{1, \ldots, m\} \\ |A| = a}} \tilde{\varphi}_A. \tag{3.6}$$

Due to (3.4), every $\tilde{\varphi}_A$ with non-empty $A \subset \{1, \ldots, m\}$ is μ-degenerate.

We can now state the moment conditions on $\varphi \in \Phi$ in terms of the corresponding μ-degenerate functions $\tilde{\varphi}_1, \ldots, \tilde{\varphi}_m$ given via (3.3) and (3.6). Assume that $\{b_n\}_{n \in \mathbb{N}} \subset (0, \infty)$ satisfies (1.6). We make an effort to balance the moment conditions with the type of Banach space E and the growth rate of $\{b_n\}_{n \in \mathbb{N}}$.

Condition 3.1. *For some rank $r \in \{1, \ldots, m\}$, Φ satisfies the following:*

  (i) *The separable real Banach space E is of type $p \in [1, 2]$.*
  (ii) *$\Phi \subset L_2(\mu^{\otimes m}, E)$.*
  (iii) *If $r = 1$ then, for every $\varphi \in \Phi$,*
     (a) *$\int_S \|\tilde{\varphi}_1\|_E^2 d\mu < \infty$, and*
     (b) *$\lim_{n \to \infty} (n/b_n^2) \log (n \mathbb{P}(\|\tilde{\varphi}_1(X_i)\|_E \geqslant b_n)) = -\infty$.*
  (iv) *For every $a \in \{r, \ldots, m\}$ with $a \geqslant 2$ there exists $p_a \in (1, p]$ such that*
     (a) *$p_r = 2$ and $\lim_{n \to \infty} n^{a-r}/b_n^{2a/p_a - r} = \infty$ if $a \geqslant r + 1$, and*
     (b) *for every $\varphi \in \Phi$ there exists at least one $\alpha_{a,\varphi} > 0$ such that*

$$\int_{S^a} \exp(\alpha_{a,\varphi} \|\tilde{\varphi}_a\|_E^{p_a}) \mathrm{d}\mu^{\otimes a} < \infty. \tag{3.7}$$

If $\Phi = B(S^m, E)$, then all corresponding μ-degenerate functions are bounded and Conditions 3.1(iii) and (iv)(b) certainly hold. Unless $r = 1$, the Banach space E has to be of type 2 due to Condition 3.1 (iv)(a). In the case $p_a = 2$, the limit in Condition 3.1(iv)(a) reduces to the first condition in (1.6). For $a \geqslant r + 1$, a choice $p_a < 2$ allows us to balance a weaker moment condition (3.7) with a slower growth of the sequence $\{b_n\}_{n \in \mathbb{N}}$. Note that for the case of rank $r = 1$ we do *not* need E to be of type 2.

Let us define the moderate U-empirical measures of general rank $r \in \{1, \dots, m\}$. Given $a \in \{1, \dots, m\}$ and $A = \{i_1, \dots, i_a\}$ with $1 \leqslant i_1 < i_2 < \dots < i_a \leqslant m$, let $1 \leqslant i_{a+1} < \dots < i_m \leqslant m$ denote the indices in $\{1, \dots, m\} \backslash A$. Define the permutation τ_A of $\{1, \dots, m\}$ such that $\tau_A(i_j) = j$ for every $j \in \{1, \dots, m\}$. For every map $\tau : \{1, \dots, m\} \to \{1, \dots, m\}$ we define $\pi_\tau : S^m \to S^m$ by $\pi_\tau(s) = (s_{\tau(1)}, \dots, s_{\tau(m)})$ for every $s = (s_1, \dots, s_m) \in S^m$. Using τ_A, define the mapping $\gamma_{A,m} : \mathcal{M}(S^a) \to \mathcal{M}(S^m)$ by $\gamma_{A,m}(\nu) = (\nu \otimes \mu^{\otimes m-a}) \pi_{\tau_A}^{-1}$ for all $\nu \in \mathcal{M}(S^a)$.

The marginal measure of $\gamma_{A,m}(\nu)$, corresponding to the ordered indices in A, is then given by ν; all other one-component marginals equal μ. Related to (3.3), define $\tilde{\gamma}_{A,m} : \mathcal{M}(S^a) \to \mathcal{M}(S^m)$ by

$$\tilde{\gamma}_{A,m}(\nu) = (-1)^{|A|} \mu^{\otimes m} + \sum_{B \subset A, B \neq \varnothing} (-1)^{|A \backslash B|} \gamma_{B,m}(\nu_{A,B}), \tag{3.8}$$

where $\nu_{A,B}$ denotes the marginal $\nu_{j_1, \dots, j_b}$ of $\nu \in \mathcal{M}_1(S^a)$, when $B = \{i_{j_1}, \dots, i_{j_b}\}$ with $1 \leqslant j_1 < \dots < j_b \leqslant a$. For $n \geqslant m$ define the moderate U-empirical measure $M_n^{m,r}$ of rank $r \in \{1, \dots, m\}$ by

$$M_n^{m,r} = \left(\frac{n}{b_n}\right)^r \left(L_n^m - \mu^{\otimes m} - \sum_{\substack{A \subset \{1, \dots, m\} \\ 1 \leqslant |A| \leqslant r-1}} \tilde{\gamma}_{A,m}(L_n^{|A|}) \right). \tag{3.9}$$

Using (3.3)–(3.6), it follows from these definitions that, for every $\varphi \in L_1(\mu^{\otimes m}, E)$,

$$\int_{S^m} \varphi \, \mathrm{d}M_n^{m,r} = \left(\frac{n}{b_n}\right)^r \sum_{a=r}^m \int_{S^a} \tilde{\varphi}_a \, \mathrm{d}L_n^a \qquad \mathbb{P}\text{-a.s.}, \tag{3.10}$$

which means that $M_n^{m,r}$ extracts from φ the components of higher rank.

Let us assume that for every unbounded $\varphi \in \Phi$, every non-void $A \subsetneq \{1, \dots, m\}$ and every $(x_i)_{i \in A} \in S^A$, the function $S^{\{1, \dots, m\} \backslash A} \ni (x_i)_{i \in \{1, \dots, m\} \backslash A} \mapsto \varphi(x_1, \dots, x_m)$ is Bochner $\mu^{\otimes m-|A|}$-integrable. This avoids measurability problems.

For $m = (m_1, m_2) \in \mathbb{N}^2$ and $r = (r_1, r_2) \in \mathbb{N}^2$ we define the rate function $I_{m,r} : \mathcal{M}(S^{m_1}) \times \mathcal{M}(S^{m_2}) \to [0, \infty]$ for the moderate deviations of rank $r \in \{1, \dots, m_1\} \times \{1, \dots, m_2\}$ by

$$I_{m,r}(\nu_1, \nu_2) = \frac{1}{2} \int_S \left(\frac{\mathrm{d}\tilde{\nu}}{\mathrm{d}\mu}\right)^2 \mathrm{d}\mu \tag{3.11}$$

if there exists a $\tilde{\nu} \in \mathcal{M}(S)$ satisfying $\tilde{\nu}(S) = 0$ and $\tilde{\nu} \ll \mu$ such that, for $i = 1, 2$,

$$\nu_i = \sum_{\substack{A \subset \{1,\ldots,m_i\} \\ |A|=r_i}} \gamma_{A,m_i}(\tilde{\boldsymbol{\nu}}^{\otimes r_i}) = \sum_{\substack{A \subset \{1,\ldots,m_i\} \\ |A|=r_i}} \bigotimes_{j \in \{1,\ldots,m_i\}} (1_A(j)\tilde{\nu} + 1_{A^c}(j)\mu); \qquad (3.12)$$

and we define $I_{m,r}(\nu_1, \nu_2) = \infty$ otherwise. In the case $r_i = 1$, equation (3.12) reduces to $\nu_i = \sum_{j=1}^{m} \mu^{\otimes j-1} \otimes \tilde{\nu} \otimes \mu^{\otimes m-j}$ and every one-component marginal of ν_i is equal to $\tilde{\nu}$. For every $l \in [0, \infty)$, let $K(I_{m,r}, l) = \{(\nu_1, \nu_2) \in \mathcal{M}(S^{m_1}) \times \mathcal{M}(S^{m_2})|I_{m,r}(\nu_1, \nu_2) \leq l\}$ denote the corresponding level set. For $B \subset \mathcal{M}(S^{m_1}) \times \mathcal{M}(S^{m_2})$, define $I_{m,r}(B) = \inf_{\nu \in B} I_{m,r}(\nu)$. The space $\mathcal{M}(S^{m_1}) \times \mathcal{M}(S^{m_2})$ is equipped with the product topology and the appropriate product σ-field. For $\Phi = (\Phi_1, \Phi_2) \subset L_2(\mu^{\otimes m_1}, E) \times L_2(\mu^{\otimes m_2}, E)$, in the following the $\tau^{\Phi}(E)$-topology denotes the corresponding product topology.

Theorem 3.1 Rank-dependent moderate deviations. *The following assertions hold for every* $r = (r_1, r_2) \in \{1, \ldots, m_1\} \times \{1, \ldots, m_2\}$:

(a) *If* $\Phi = (\Phi_1, \Phi_2) \subset L_2(\mu^{\otimes m_1}, E) \times L_2(\mu^{\otimes m_2}, E)$, *then the level set* $K(I_{m,r}, l) \subset \mathcal{M}^{\Phi_1}(S^{m_1}) \times \mathcal{M}^{\Phi_2}(S^{m_2})$ *for every* $l \in [0, \infty)$.

(b) *If* $\Phi = (\Phi_1, \Phi_2) \subset L_2(\mu^{\otimes m_1}, E) \times L_2(\mu^{\otimes m_2}, E)$, *then* $K(I_{m,r}, l)$ *is* $\tau^{\Phi}(E)$-compact for every $l \in [0, \infty)$.

(c) *Assume that there exists a* $p \in (1, 2]$ *such that the Banach space* $(E, \|\cdot\|_E)$ *is of type* p *and* $\lim_{n \to \infty} n/b_n^p = 0$. *If* $\Phi_i \subset L_p(\mu^{\otimes m_i}, E)$ *in the case* $m_i = r_i$, $i = 1, 2$, *or if* $\Phi_i \subset L_2(\mu^{\otimes m_i}, E)$ *in the case* $r_i < m_i$, $i = 1, 2$, *then*

$$\liminf_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{P}((M_n^{m_1,r_1}, M_n^{m_2,r_2}) \in B) \geq -I_{m,r}(int_{\tau^{\Phi}(E)}(B)) \qquad (3.13)$$

for every measurable $B \subset \mathcal{M}(S^{m_1}) \times \mathcal{M}(S^{m_2})$, *where* $int_{\tau^{\Phi}(E)}(B)$ *denotes the interior of the set* $B \cap (\mathcal{M}^{\Phi_1}5(S^{m_1}) \times \mathcal{M}^{\Phi_2}(S^{m_2}))$ *with respect to the* $\tau^{\Phi}(E)$-*topology.*

(d) *If* Φ_i *satisfies Condition 3.1 for* r_i, $i = 1, 2$, *then*

$$\limsup_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{P}((M_n^{m_1,r_1}, M_n^{m_2,r_2}) \in B) \leq -I_{m,r}(cl_{\tau^{\Phi}(E)}(B)) \qquad (3.14)$$

for every measurable $B \subset \mathcal{M}(S^{m_1}) \times \mathcal{M}(S^{m_2})$, *where* $cl_{\tau^{\Phi}(E)}(B)$ *denotes the closure of the set* $B \cap (\mathcal{M}^{\Phi_1}(S^{m_1}) \times \mathcal{M}^{\Phi_2}(S^{m_2}))$ *with respect to the* $\tau^{\Phi}(E)$-*topology.*

Remark 3.1. It is obvious from the proof of Theorem 3.1 that we have an MDP for a vector of a finite number of U-empirical measures.

The following corollary is an immediate consequence:

Corollary 3.2. *For* $r = (1, 2)$ *and* $m = (1, k)$, *the rate function is*

$$I(\nu, \nu_2) = \frac{1}{2} \int_S \left(\frac{d\nu}{d\mu}\right)^2 d\mu$$

if $\nu(S) = 0$, $\nu \ll \mu$ *and* ν_2 *is given by (3.12) with* $r_2 = 2$ *and* $m_2 = k$, *and* $+\infty$ *otherwise. In the case* $k = 2$ *we obtain* $\nu_2 = \nu \otimes \nu$.

4. Proof of the rank-dependent moderate deviations

In the proof of Theorem 3.1 we will several times apply the following Bernstein-type inequality proved in Eichelsbacher and Schmock (2001):

Theorem 4.1. (Bernstein-type inequality). *Let $(E, \|\cdot\|_E)$ be a separable Banach space of type $p \in (1, 2]$ and let $r \in \mathbb{N}$. Then there exist constants $c_3, c_4 \in [1, \infty)$ such that for every symmetric, Bochner $\mu^{\otimes r}$-integrable, completely μ-degenerate function $\varphi : S^r \to E$ with an $\alpha > 0$ satisfying*

$$\tilde{a} \equiv \int_{S^r} \exp(\alpha \|\varphi\|_E^p) \, \mathrm{d}\mu^{\otimes r} < \infty, \tag{4.1}$$

there exists a constant $c_\varphi \in (0, \infty)$ such that, for all $x > 0$ and all integers $n \geq r$,

$$\mathbb{P}(\|n^\varrho U_n^r(\varphi)\|_E \geq x) \leq c_3 \exp\left(-\frac{x^{p/r}}{c_4 \sigma^{p/r} + c_\varphi (x^{p/r}/n)^{1/(r+1)}}\right), \tag{4.2}$$

where $\sigma \equiv \|\varphi\|_{L_p(\mu^{\otimes r}, E)}$ and $\varrho \equiv r(1 - 1/p)$.

Furthermore, we will use the following approximation result (see, for example, Eichelsbacher and Schmock 2001).

Lemma 4.2. *Let $\sigma > 0$, $p \geq 1$, $r \in \mathbb{N}$, and let $\varphi \in L_p(\mu^{\otimes r}, E)$ be a symmetric and completely μ-degenerate function. Then there exist $k \in \mathbb{N}$, vectors $\beta_1, \ldots, \beta_k \in E$ and bounded measurable functions $f_1, \ldots, f_k : S \to \mathbb{R}$ with $\int_S f_i \, \mathrm{d}\mu = 0$ for all $i \in \{1, \ldots, k\}$ such that $\varphi_\sigma \equiv \sum_{i=1}^k \beta_i f_i^{\otimes r}$ satisfies $\|\varphi - \varphi_\sigma\|_{L_p(\mu^{\otimes r}, E)} \leq \sigma$, where $f_i^{\otimes r}(s_1, \ldots, s_r) \equiv \prod_{j=1}^r f_i(s_j)$ for all $s_1, \ldots, s_r \in S$.*

Let $I_{m_i, r_i} : \mathcal{M}(S^{m_i}) \to [0, \infty]$ be defined by

$$I_{m_i, r_i}(\nu_i) = \frac{1}{2} \int_S \left(\frac{\mathrm{d}\tilde{\nu}}{\mathrm{d}\mu}\right)^2 \mathrm{d}\mu \tag{4.3}$$

if there exists a $\tilde{\nu} \in \mathcal{M}(S)$ satisfying $\tilde{\nu}(S) = 0$ and $\tilde{\nu} \ll \mu$ such that ν_i satisfies (3.12), and $I_{m_i, r_i}(\nu_i) = \infty$ otherwise. Then it is easy to see that

$$K(I_{m,r}, l) \subset K(I_{m_1, r_1}, l) \times K(I_{m_2, r_2}, l) \tag{4.4}$$

for every $l \in [0, \infty)$.

Proof of Theorem 3.1. (a) and (b). Using (4.4), (a) follows from Theorem 2.14(a) in Eichelsbacher and Schmock (2001).

By Theorem 2.14(b) in Eichelsbacher and Schmock (2001) we know that $K(I_{m_i, r_i}, l)$ is $\tau^\Phi(E)$-compact for $i = 1, 2$. Using (4.4), to prove (b) it therefore suffices to show that $I_{m,r}$ is lower semicontinuous, which is left to the reader. $\qquad \square$

Proof of Theorem 3.1. (c). Consider a measure $\nu = (\nu_1, \nu_2) \in \text{int}_{\tau^\Phi(E)}(B)$ with $I_{m,r}(\nu) < \infty$. Then $\nu \in \mathcal{M}^{\Phi_1}(S_1^m) \times \mathcal{M}^{\Phi_2}(S_2^m)$ by the definition of the $\tau^\Phi(E)$-interior of the set B, hence $\int_{S_i^m} \|\varphi_i\|_E \, d|\nu_i| < \infty$ for every $\varphi_i \in \Phi_i$. By the definition of the $\tau^\Phi(E)$-product topology, there exist $\eta > 0$, $k_1, k_2 \in \mathbb{N}$ and $\varphi_{1,1}, \ldots, \varphi_{1,k_1} \in \Phi_1$, $\varphi_{2,1}, \ldots, \varphi_{2,k_2} \in \Phi_2$ such that the $\tau^\Phi(E)$-open set

$$C((\nu_1, \nu_2), 2\eta) := C(\nu_1, 2\eta) \times C(\nu_2, 2\eta)$$

with

$$C(\nu_i, 2\eta) = \left\{ \nu_i' \in \mathcal{M}^{\Phi_i}(S^{m_i}) \Big| \left\| \int_{S_i^m} \varphi_{i,j} \, d(\nu_i - \nu_i') \right\|_E < 2\eta \text{ for every } j \in \{1, \ldots, k_i\} \right\}$$

is contained in $\text{int}_{\tau^\Phi(E)}(B)$. Since $I_{m,r}(\nu) < \infty$, a measure $\tilde{\nu} \in \mathcal{M}(S)$ for the representation (3.12) and a density $\tilde{g} \equiv d\tilde{\nu}/d\mu$ exist, but both might not be unique, as the discussion after (3.12) showed. The density $\tilde{g}$ satisfies $\int_S \tilde{g} \, d\mu = 0$ and $\int_S \tilde{g}^2 \, d\mu = 2 I_{m,r}(\nu) < \infty$. Furthermore, a density of ν_i with respect to $\mu^{\otimes m_i}$ is given by

$$\frac{d\nu_i}{d\mu^{\otimes m_i}}(s_1, \ldots, s_{m_i}) = \sum_{\substack{A \subset \{1, \ldots, m_i\} \\ |A| = r_i}} \prod_{i \in A} \frac{d\tilde{\nu}}{d\mu}(s_i), \qquad -\mu^{\otimes m_i}\text{-a.s.,} \tag{4.5}$$

and, for every $\varphi_i \in \Phi_i$,

$$\int_{S^{m_i}} \|\varphi(s_1, \ldots, s_{m_i})\|_E \prod_{j=1}^{m_i} |\tilde{g}(s_j)| \mu^{\otimes m_i}(ds_1, \ldots, ds_{m_i}) = \int_{S^{m_i}} \|\varphi\|_E \, d|\nu_i| < \infty$$

in the case $m_i = r_i$; otherwise, when $\Phi_i \subset L_2(\mu^{\otimes m_i}, E)$,

$$\int_{S^{m_i}} \|\varphi(s_1, \ldots, s_{m_i})\|_E \sum_{\substack{A \subset \{1, \ldots, m_i\} \\ |A| = r_i}} \prod_{j \in A} |\tilde{g}(s_j)| \mu^{\otimes m_i}(ds_1, \ldots, ds_{m_i})$$

$$\leq \binom{m_i}{r_i} \|\varphi_i\|_{L_2(\mu^{\otimes m_i}, E)} (2 I_{m,r}(\nu))_i^r < \infty$$

by the Cauchy–Schwarz inequality. Define $\tilde{g}^\pm(s) = \max\{\pm\tilde{g}(s), 0\}$ for $s \in S$. It follows from the dominated convergence theorem that there exist constants $c^\pm \in [0, \|\tilde{g}^\pm\|_{L_\infty(\mu)}) \cup \{0\}$ such that the truncated function $g \equiv \min\{c^+, \max\{-c^-, \tilde{g}\}\}$ satisfies $\int_S g \, d\mu = 0$ and

$$\int_{S^{m_i}} \|\varphi_{j,i}\|_E \, d|\nu_i - \bar{\nu}_i| \leq \eta, \qquad j \in \{1, \ldots, k_i\}, i \in \{1, 2\}, \tag{4.6}$$

where $\bar{\nu}_i \in \mathcal{M}(S^{m_i})$ is determined by

$$\frac{\mathrm{d}\bar{\nu}_i}{\mathrm{d}\mu^{\otimes m_i}}(s_1, \ldots, s_{m_i}) = \sum_{\substack{A \subset \{1, \ldots, m_i\} \\ |A| = r_i}} \prod_{j \in A} g(s_j), \qquad (s_1, \ldots, s_{m_i}) \in S^{m_i}.$$

Since g is bounded and $\Phi_i \subset L_1(\mu^{\otimes m_i}, E)$, the measure $\bar{\nu}_i$ is actually in $\mathcal{M}^{\Phi_i}(S^{m_i})$.

Note that $C((\bar{\nu}_1, \bar{\nu}_2), \eta) \subset C((\nu_1, \nu_2), 2\eta)$ by (4.6). Furthermore, $C((\bar{\nu}_1, \bar{\nu}_2), \eta)$ is a measurable subset of $\mathcal{M}(S^{m_1}) \times \mathcal{M}(S^{m_2})$ because $\mathcal{M}^{\Phi_i}(S^{m_i})$ is measurable for every i by definition. Define the function $F_n(s) = \prod_{j=1}^{n}(1 + (b_n/n)g(s_j))$ for all $s = (s_j)_{j \in \mathbb{N}} \in S^{\mathbb{N}}$. Due to (1.6) there exists, for every $\varepsilon \in (0, 1)$, an $n_\varepsilon \geq m$ such that $b_n c^-/n \leq \varepsilon$ for all $n \geq n_\varepsilon$. Define $\mathbb{P}_n \in \mathcal{M}_1(\Omega)$ by $\mathrm{d}\mathbb{P}_n/\mathrm{d}\mathbb{P} = F_n$ for all these n. Then

$$\mathbb{P}((M_n^{m_1, r_1}, M_n^{m_2, r_2}) \in B) \geq \mathbb{P}((M_n^{m_1, r_1}, M_n^{m_2, r_2}) \in C((\bar{\nu}_1, \bar{\nu}_2), \eta)) = \int_{D_n} \frac{1}{F_n} \, \mathrm{d}\mathbb{P}_n, \qquad (4.7)$$

where $D_n \equiv \{(M_n^{m_1, r_1}, M_n^{m_2, r_2}) \in C((\bar{\nu}_1, \bar{\nu}_2), \eta)\}$. Defining $a_n = \mathbb{P}_n(D_n)$ and using Jensen's inequality, we obtain, for every $n \geq n_\varepsilon$,

$$\log \int_{D_n} \frac{1}{F_n} \, \mathrm{d}\mathbb{P}_n = \log a_n - \frac{1}{a_n} \int_{D_n} F_n \log F_n \, \mathrm{d}\mathbb{P}. \qquad (4.8)$$

Since $x \log x \geq -1/\mathrm{e}$ for $x \geq 0$ and $\int_S g \, \mathrm{d}\mu = 0$, it follows that, for all $n \geq n_\varepsilon$,

$$\int_{D_n} F_n \log F_n \, \mathrm{d}\mathbb{P} \leq \frac{1}{\mathrm{e}} + \frac{b_n^2}{(1-\varepsilon)n} I_{m,r}(\nu), \qquad (4.9)$$

where we have used the estimate $(1 + x)\log(1 + x) \leq x + x^2/(2(1 - \varepsilon))$ for all $x \geq -\varepsilon$ for the second inequality. If we can show that $\lim_{n \to \infty} a_n = 1$, then (1.6), (4.7), (4.8) and (4.9) together imply

$$\liminf_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{P}((M_n^{m_1, r_1}, M_n^{m_2, r_2}) \in B) \geq -\frac{I_{m,r}(\nu)}{1 - \varepsilon},$$

which in turn implies (3.13). Since, for every $i \in \{1, 2\}$,

$$\lim_{n \to \infty} \mathbb{P}_n(M_n^{m_i, r_i} \in C(\bar{\nu}_i, \eta)) = 1$$

(see the proof of Theorem 2.14(c) in Eichelsbacher and Schmock 2001, pp. 13 and 14) we obtain $\lim_{n \to \infty} a_n = 1$. The proof in Eichelsbacher and Schmock (2001) depends strongly on quite sharp moment inequalities for Banach-space-valued U-statistics using decoupling inequalities and the hypercontractive method in moment inequalities. $\square$

Proof of Theorem 3.1 (d). Let C denote the $\tau^\Phi(E)$-closure of $B \cap (\mathcal{M}^{\Phi_1}(S^{m_1}) \times \mathcal{M}^{\Phi_2}(S^{m_2}))$. The pair of moderate U-empirical measures $\{(M_n^{m_1, r_1}, M_n^{m_2, r_2})\}_{n \geq (m_1 \wedge m_2)}$ take values in $\mathcal{M}^{\Phi_1}(S^{m_1}) \times \mathcal{M}^{\Phi_2}(S^{m_2})$. It suffices to consider only the case $I_{m,r}(C) > 0$. Choose $l \in (0, I_{m,r}(C))$. Let $\mathcal{F}_i$ denote the family of all finite, non-empty subsets of Φ_i, $i = 1, 2$. For every $(F_1, F_2) \in \mathcal{F}_1 \times \mathcal{F}_2$ define

$$\Pi_{(F_1,F_2)} : \mathcal{M}^{\Phi_1}(S^{m_1}) \times \mathcal{M}^{\Phi_2}(S^{m_2}) \to E^{F_1} \times E^{F_2}$$

by

$$\Pi_{(F_1,F_2)}(\nu_1, \nu_2) = \left(\int_{S^{m_1}} \varphi \, d\nu_1\right)_{\varphi \in F_1} \times \left(\int_{S^{m_2}} \tilde{\varphi} \, d\nu_2\right)_{\tilde{\varphi} \in F_2}.$$

Π^{F_1} (Π^{F_2}) denotes the restriction of $\Pi_{(F_1,F_2)}$ to the first m_1 (last m_2) coordinates. Note that E^{F_i} with $\|y\|_{E^{F_i}} := \sum_{\varphi_i \in F_i} \|y_{\varphi_i}\|_E$ for $y = (y_{\varphi_i})_{\varphi_i \in F_i} \in E^{F_i}$, $i = 1, 2$, is a Banach space, as is the product space $E^{F_1} \times E^{F_2}$. Following the proof of Theorem 2.14(d) in Eichelsbacher and Schmock (2001), there exist an $F = (F_1, F_2) \in \mathcal{F}_1 \times \mathcal{F}_2$ and an open set U in $E^{F_1} \times E^{F_2}$ such that $\Pi_{(F_1,F_2)}^{-1}(U)$ covers $K(I_{m,r}, l)$ and is disjoint from C. The construction of U uses the closeness of C as well as the compactness of $K(I_{m,r}, l)$. Define $\varepsilon := \mathrm{dist}(\Pi_{(F_1,F_2)}(K(I_{m,r}, l)), U^c)$, where dist is a distance in $E^{F_1} \times E^{F_2}$. We choose the supremum of distances dist_i, $i = 1, 2$, in E^{F_i}, which metrizes the product topology in $E^{F_1} \times E^{F_2}$. Since $\Pi_{(F_1,F_2)}(K(I_{m,r}, l))$ is a compact subset of the open set U, it follows that $\varepsilon > 0$ and that

$$A_\varepsilon := \{(x, y) \in E^{F_1} \times E^{F_2} : \mathrm{dist}((x, y), \Pi_{(F_1,F_2)}(K(I_{m,r}, l))) < \varepsilon\}$$

is an open set contained in U. We rewrite A_ε as

$$A_\varepsilon = \{(x, y) \in E^{F_1} \times E^{F_2} : \mathrm{dist}_1(x, \Pi^{F_1}(\nu_1)) < \varepsilon, \, \mathrm{dist}_2(y, \Pi^{F_2}(\nu_2)) < \varepsilon,$$
$$\text{with } (\nu_1, \nu_2) \in K(I_{m,r}, l)\}.$$

Thus we can find $(F_1, F_2) \in \mathcal{F}_1 \times \mathcal{F}_2$ and an open (possibly non-convex) ε-neighbourhood $A_\varepsilon \subset E^{F_1} \times E^{F_2}$ of $\Pi_{(F_1,F_2)}(K(I_{m,r}, l))$ such that

$$\{(M_n^{m_1,r_1}, M_n^{m_2,r_2}) \in B\} \subset \{(M_n^{m_1,r_1}, M_n^{m_2,r_2}) \in C\}$$
$$\subset \{\Pi_{(F_1,F_2)}(M_n^{m_1,r_1}, M_n^{m_2,r_2}) \in E^{F_1} \times E^{F_2} \backslash A_\varepsilon\}, \qquad (4.10)$$

for all $n \geqslant m_1 \wedge m_2$. Using (3.10), it follows that

$$\mathbb{P}\big(\Pi_{(F_1,F_2)}(M_n^{m_1,r_1}, M_n^{m_2,r_2}) \in E^{F_1} \times E^{F_2} \backslash A_\varepsilon\big)$$

$$\leqslant \sum_{a=r_1+1}^{m_1} \mathbb{P}\left(\left(\frac{n}{b_n}\right)^{r_1} \|\Pi_{F_{1,a}}(L_n^a)\|_{E^{F_1}} \geqslant \frac{\varepsilon}{4m}\right)$$

$$+ \sum_{a=r_2+1}^{m_2} \mathbb{P}\left(\left(\frac{n}{b_n}\right)^{r_2} \|\Pi_{F_{2,a}}(L_n^a)\|_{E^{F_2}} \geqslant \frac{\varepsilon}{4m}\right) \qquad (4.11)$$

$$+ \mathbb{P}\left(\left(\left(\frac{n}{b_n}\right)^{r_1} \Pi_{F_{1,r_1}}(L_n^{r_1}), \left(\frac{n}{b_n}\right)^{r_2} \Pi_{F_{2,r_2}}(L_n^{r_2})\right) \in E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/2}\right),$$

where $F_{i,a} \equiv \{\tilde{\varphi}_{i,a}\}_{\varphi_i \in F_i}$ for $a \in \{r_i, r_i + 1, \ldots, m_i\}$, $i = 1, 2$, with the notation from (3.6). We will assume in the following (without loss of generality) that F_i, and therefore F_{i,r_i} up to F_{i,m_i}, consist of symmetric functions. In order to see that the terms with $a \geqslant r_i + 1$ in (4.11) do not contribute to the moderate-deviations upper bound (3.14), it suffices to prove that

$$\limsup_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\left(\frac{n}{b_n}\right)^{r_i} \|U_n^a(\tilde{\varphi}_{i,a})\|_E \geq \eta\right) = -\infty \tag{4.12}$$

for every $a \in \{r_i + 1, \ldots, m_i\}$, $\eta > 0$ and $\tilde{\varphi}_{i,a} \in F_{i,a}$. This follows from an application of the Bernstein-type inequality (4.2) (for details, see Eichelsbacher and Schmock 2001). Choose $\sigma_i > 0$ so that $\varepsilon/(4|F_i|\sigma_i) > (\max\{c_4, 2\}l)^{r_i/2}$, for $i = 1, 2$, with c_4 as in (4.2). For every $\tilde{\varphi}_{i,r_i}$ corresponding to a $\varphi_i \in F_i$, we can find a function $\tilde{\varphi}_{i,r_i,\sigma_i}$ as described in Lemma 4.2 such that $\|\tilde{\varphi}_{i,r_i} - \tilde{\varphi}_{i,r_i,\sigma_i}\|_{L_2(\mu^{\otimes r_i}, E)} \leq \sigma_i$. Define $F_{i,r_i,\sigma_i} = \{\tilde{\varphi}_{i,r_i,\sigma_i}\}_{\varphi_i \in F_i}$. Then

$$\mathbb{P}\left(\left(\left(\frac{n}{b_n}\right)^{r_1} \Pi_{F_{1,r_1}}(L_n^{r_1}), \left(\frac{n}{b_n}\right)^{r_2} \Pi_{F_{2,r_2}}(L_n^{r_2})\right) \in E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/2}\right)$$

$$\leq \sum_{i=1}^2 \sum_{\varphi_i \in F_i} \mathbb{P}\left(\left(\frac{n}{b_n}\right)^{r_i} \|U_n^{r_i}(\tilde{\varphi}_{i,r} - \tilde{\varphi}_{i,r_i,\sigma_i})\|_E \geq \frac{\varepsilon}{8|F_i|}\right)$$

$$+ \mathbb{P}\left(\left(\left(\frac{n}{b_n}\right)^{r_1} \Pi_{F_{1,r_1,\sigma_1}}(L_n^{r_1}), \left(\frac{n}{b_n}\right)^{r_2} \Pi_{F_{2,r_2,\sigma_2}}(L_n^{r_2})\right) \in E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4}\right). \tag{4.13}$$

In order to show that, for every $\varphi_i \in F_i$,

$$\limsup_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\left(\frac{n}{b_n}\right)^{r_i} \|U_n^{r_i}(\tilde{\varphi}_{i,r} - \tilde{\varphi}_{i,r_i,\sigma_i})\|_E \geq \frac{\varepsilon}{8|F_i|}\right) \leq -l, \tag{4.14}$$

apply the Bernstein-type inequality (4.2) (for details, see Eichelsbacker and Schmock 2001); (4.14) is determined by the choice of the σ_i, which measures the quality of the approximation.

To treat the last term in (4.13), we need to prove that

$$\limsup_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\left(\frac{n}{b_n}\right)^{r_i} \|\Pi_{F_{i,r_i,\sigma_i}}(L_n^{r_i} - L_n^{\otimes r_i})\|_{E^{F_i}} \geq \frac{\varepsilon}{8}\right) = -\infty \tag{4.15}$$

for $i = 1, 2$. Due to the form of the functions in F_{i,r_i,σ_i}, as described in Lemma 4.2, it suffices to show that

$$\limsup_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\left(\frac{n}{b_n}\right)^{r_i} |U_n^{r_i}(f^{\otimes r_i}) - (U_n^1(f))^{r_i}| \geq \eta\right) = -\infty \tag{4.16}$$

for every $\eta > 0$ and every bounded measurable function $f : S \to \mathbb{R}$ satisfying $\int_S f \, d\mu = 0$. When proving (4.16) we assume that $r_i \geq 2$, $i = 1, 2$, and we use the arguments in the proof of Theorem 2.14 in Eichelsbacher and Schmock (2001). We skip the details.

To treat the last term in (4.11) in the case $r_i \in \{2, 3, \ldots, m_i\}$, combining (4.11)–(4.15), it remains to show that

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\left(\left(\frac{n}{b_n}\right)^{r_1} \Pi_{F_{1,r_1,\sigma_1}}(L_n^{\otimes r_1}), \left(\frac{n}{b_n}\right)^{r_2} \Pi_{F_{2,r_2,\sigma_2}}(L_n^{\otimes r_2})\right) \in E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4}\right) \leqslant -l.$$

(4.17)

For this purpose, we assume without loss of generality that there exist k_1, $k_2 \in \mathbb{N}$ such that every $\tilde{\varphi}_{i,r_i,\sigma_i}$ with $\varphi_i \in F_i$ used in (4.13) and described in Lemma 4.2 is of the form $\tilde{\varphi}_{i,r_i,\sigma_i} = \sum_{j=1}^{k_i} \beta_{\varphi_i,j} f_{\varphi_i,j}^{\otimes r_i}$ where $\beta_{\varphi_i,j} \in E$ and $f_{\varphi_i,j} : S \to \mathbb{R}$ is bounded and μ-degenerate for all $j \in \{1, \ldots, k_i\}$ and $\varphi_i \in F_i$. Define the map $\psi_i : S \to (\mathbb{R}^{k_i})^{F_i}$ by $\psi_i(s) = (f_{\varphi_i,1}(s), \ldots, f_{\varphi_i,k_i}(s))_{\varphi_i \in F_i}$ and the map $T : (\mathbb{R}^{k_1})^{F_1} \times (\mathbb{R}^{k_2})^{F_2} \to E^{F_1} \times E^{F_2}$ by

$$T((x_{\varphi_1,1}, \ldots, x_{\varphi_1,k_1})_{\varphi_1 \in F_1}, (x_{\varphi_2,1}, \ldots, x_{\varphi_2,k_2})_{\varphi_2 \in F_2})$$

$$= \left(\left(\sum_{j=1}^{k_1} \beta_{\varphi_1,j} x_{\varphi_1,j}^{r_1}\right)_{\varphi_1 \in F_1}, \left(\sum_{j=1}^{k_1} \beta_{\varphi_2,j} x_{\varphi_2,j}^{r_2}\right)_{\varphi_2 \in F_2}\right).$$

Then, for every $n \in \mathbb{N}$,

$$\left\{\left(\left(\frac{n}{b_n}\right)^{r_1} \Pi_{F_{1,r_1,\sigma_1}}(L_n^{\otimes r_1}), \left(\frac{n}{b_n}\right)^{r_2} \Pi_{F_{2,r_2,\sigma_2}}(L_n^{\otimes r_2})\right) \in E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4}\right\}$$

$$= \left\{\left(\int_S \psi_1 \, \mathrm{d}M_n^{1,1}, \int_S \psi_2 \, \mathrm{d}M_n^{1,1}\right) \in T^{-1}\left(E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4}\right)\right\}. \quad (4.18)$$

Since T is continuous, $T^{-1}(E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4})$ is closed in $(\mathbb{R}^{k_1})^{F_1} \times (\mathbb{R}^{k_2})^{F_2}$ and the moderate-deviations upper bound in the τ-topology on $\mathcal{M}(S)$ from de Acosta (1994, Theorem 3.1) implies that

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\left(\int_S \psi_1 \, \mathrm{d}M_n^{1,1}, \int_S \psi_2 \, \mathrm{d}M_n^{1,1}\right) \in T^{-1}(E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4})\right)$$

$$\leqslant -\inf\left\{I_{1,1}(\tilde{\nu})|\tilde{\nu} \in \mathcal{M}\right\}(S), \left(\int_S \psi_1 \, \mathrm{d}\tilde{\nu}, \int_S \psi_2 \, \mathrm{d}\tilde{\nu}\right) \in T^{-1}(E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4})\right\} \quad (4.19)$$

with $I_{1,1}$ as in (4.3). Note that

$$T\left(\int_S \psi_1 \, \mathrm{d}\tilde{\nu}, \int_S \psi_2 \, \mathrm{d}\tilde{\nu}\right) = (\Pi_{F_{1,r_1,\sigma_1}}(\tilde{\nu}^{\otimes r_1}), \Pi_{F_{2,r_2,\sigma_2}}(\tilde{\nu}^{\otimes r_2}))$$

for all $\tilde{\nu} \in \mathcal{M}(S)$. Hence, to derive (4.17) from (4.18) and (4.19), it suffices to show that $(\Pi_{F_{1,r_1,\sigma_1}}(\tilde{\nu}^{\otimes r_1}), \Pi_{F_{2,r_2,\sigma_2}}(\tilde{\nu}^{\otimes r_2})) \in A_{\varepsilon/4}$ for all $\tilde{\nu} \in K(I_{1,1}, l)$.

Consider any $\tilde{\nu} \in K(I_{1,1}, l)$. Define $\nu_i \in \mathcal{M}(S^{m_i})$ by (3.12). According to (3.11) and (4.3), $I_{m_i,r_i}(\nu_i) = I_{1,1}(\tilde{\nu})$, hence $\nu_i \in K(I_{m_i,r_i}, l) \subset \mathcal{M}^{\Phi_i}(S^{m_i})$. With the arguments given in the proof of Theorem 2.14(d) in Eichelsbacher and Schmock (2001), it follows that $\|\Pi^{F_i}(\nu_i) - \Pi_{F_{i,r_i,\sigma_i}}(\tilde{\nu}^{\otimes r_i})\|_{E^{F_i}} \leqslant \sigma_i |F_i| (2l)^{r_i/2} < \varepsilon/4$. This shows that

$$(\Pi_{F_{1,r_1,\sigma_1}}(\tilde{\nu}^{\otimes r_1}), \Pi_{F_{2,r_2,\sigma_2}}(\tilde{\nu}^{\otimes r_2})) \in A_{\varepsilon/4},$$

which finishes the proof of the moderate-deviations upper bound (3.14) for $r_i \in \{2, 3, \ldots, m_i\}$.

Let us now consider the case $r_1 = r_2 = 1$. Note that

$$\left(\frac{n}{b_n} \Pi_{F_{1,1}}(L_n^1), \frac{n}{b_n} \Pi_{F_{2,1}}(L_n^1) \right) = \left(\Pi_{F_{1,1}}\big(M_n^{1,1}\big), \Pi_{F_{2,1}}\big(M_n^{1,1}\big) \right)$$

is the sum of n independent and identically distributed $E^{F_1} \times E^{F_2}$-valued random vectors of mean zero. Condition 3.1(iii)(a) and Hölder's inequality imply that, for every $i = 1, 2$, $\int_S \|(\tilde{\varphi}_i)_{\varphi_i \in F_i}\|_{E^{F_i}}^2 \, \mathrm{d}\mu < \infty$ and therefore

$$\int_S \|(\tilde{\varphi}_1)_{\varphi_1 \in F_1}, (\tilde{\varphi}_2)_{\varphi_2 \in F_2}\|_E^{2 \, F_1} \times E^{F_2} \, \mathrm{d}\mu < \infty.$$

Condition 3.1(iii)(b) for Φ_i, $i = 1, 2$, implies that

$$\lim_{n \to \infty} \frac{n}{b_n^2} \log \left(n\mathbb{P}(\|(\tilde{\varphi}_1(X_i))_{\varphi_1 \in F_1}, (\tilde{\varphi}_2(X_i))_{\varphi_2 \in F_2}\|_{E^{F_1} \times E^{F_2}} \geqslant b_n) \right) = -\infty.$$

Therefore, it follows from Eichelsbacher and Schmock (2001, Theorem 2.17) and from Arcones (1999, Theorem 2.4) that

$$\limsup_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{P}((\Pi_{F_{1,1}}(M_n^{1,1}), \Pi_{F_{2,1}}(M_n^{1,1})) \in E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4})$$

$$\leqslant -\inf\{I_{1,1}(\tilde{\nu}) | \tilde{\nu} \in \mathcal{M}(S), (\Pi_{F_{1,1}}(\tilde{\nu}), \Pi_{F_{2,1}}(\tilde{\nu})) \in E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4}\}.$$

It suffices to show that $(\Pi_{F_{1,1}}(\tilde{\nu}), \Pi_{F_{2,1}}(\tilde{\nu})) \in A_{\varepsilon/4}$ for all $\tilde{\nu} \in K(I_{1,1}, l)$. Consider any $\tilde{\nu} \in K(I_{1,1}, l)$. Define $\nu_i \in \mathcal{M}(S^{m_i})$ by (3.12). We know that $I_{m_i, r_i}(\nu_i) = I_{1,1}(\tilde{\nu})$ and therefore $(\nu_1, \nu_2) \in K(I_{m,(1,1)}, l)$. Thus we obtain

$$\limsup_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{P}\Big((\Pi_{F_{1,1}}(M_n^{1,1}), \Pi_{F_{2,1}}(M_n^{1,1})) \in E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4} \Big) \leqslant -l$$

and, together with (4.10) and (4.11), this estimate implies (3.14) for $r_1 = r_2 = 1$.

Finally, let us consider the case $r_1 = 1$ and $r_2 \geqslant 2$. Here it remains to show that

$$\limsup_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\left(\left(\frac{n}{b_n}\right) \Pi_{F_{1,1}}(L_n), \left(\frac{n}{b_n}\right)^{r_2} \Pi_{F_{2,r_2,\sigma_2}}(L_n^{\otimes r_2}) \right) \in E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4} \right) \leqslant -l.$$

$$(4.20)$$

For this purpose define the map $\psi_2 : S \to (\mathbb{R}^{k_2})^{F_2}$ as above and the map $\hat{T} : E^{F_1} \times (\mathbb{R}^{k_2})^{F_2} \to E^{F_1} \times E^{F_2}$ by

$$\hat{T}(y, (x_{\varphi_2,1}, \ldots, x_{\varphi_2,k_2})_{\varphi_2 \in F_2}) = \left(y, \left(\sum_{j=1}^{k_1} \beta_{\varphi_2,j} x_{\varphi_2,j}^{r_2} \right) \varphi_2 \in F_2 \right).$$

Then, for every $n \in \mathbb{N}$, by

$$\left\{ \left(\left(\frac{n}{b_n}\right) \Pi_{F_{1,1}}(L_n), \left(\frac{n}{b_n}\right)^{r_2} \Pi_{F_{2,r_2,\sigma_2}}(L_n^{\otimes r_2}) \right) \in E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4} \right\}$$

$$= \left\{ \left(\left(\frac{n}{b_n}\right) \Pi_{F_{1,1}}(L_n), \int_S \psi_2 \, \mathrm{d}M_n^{1,1} \right) \in \hat{T}^{-1}(E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4}) \right\}. \tag{4.21}$$

Since $\hat{T}$ is continuous, $\hat{T}^{-1}(E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4})$ is closed in $E^{F_1} \times (\mathbb{R}^{k_2})^{F_2}$ and the moderate-deviations upper bound in the τ^Φ-topology on $\mathcal{M}(S)$ from Eichelsbacher and Schmock (2001, Theorem 2.17) and Arcones (1999, Theorem 2.4) implies that

$$\limsup_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\left(\Pi_{F_{1,1}}(M_n^{1,1}), \int_S \psi_2 \, \mathrm{d}M_n^{1,1} \right) \in \hat{T}^{-1}(E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4}) \right)$$

$$\leq -\inf\left\{ I_{1,1}(\tilde{\nu}) | \tilde{\nu} \in \mathcal{M}(S), \left(\Pi F_{1,1}(\tilde{\nu}), \int_S \psi_2 \, \mathrm{d}\tilde{\nu} \right) \in \hat{T}^{-1}(E^{F_1} \times E^{F_2} \backslash A_{\varepsilon/4}) \right\}. \tag{4.22}$$

Note that

$$\hat{T}\left(\Pi_{F_{1,1}}(\tilde{\nu}), \int_S \psi_2 \, \mathrm{d}\tilde{\nu} \right) = \left(\Pi_{F_{1,1}}(\tilde{\nu}), \Pi_{F_{2,r_2,\sigma_2}}(\tilde{\nu}^{\otimes r_2}) \right)$$

for all $\tilde{\nu} \in \mathcal{M}(S)$. We are done if we can show that $\left(\Pi_{F_{1,1}}(\tilde{\nu}), \Pi_{F_{2,r_2,\sigma_2}}(\tilde{\nu}^{\otimes r_2}) \right) \in A_{\varepsilon/4}$ for all $\tilde{\nu} \in K(I_{1,1}, l)$. But this follows from our previous results: consider any $\tilde{\nu} \in K(I_{1,1}, l)$ and define $\nu_i \in \mathcal{M}(S^{m_i})$ by (3.12). We know already that $\|\Pi^{F_2}(\nu_2) - \Pi_{F_{2,r_2,\sigma_2}}(\tilde{\nu}^{\otimes r_2})\|_{E^{F_2}} < \varepsilon/4$ for a $\nu_2 \in K(I_{m_2,r_2}, l)$ and that $\Pi_{F_{1,1}}(\tilde{\nu}) \in \Pi^{F_1}(K(I_{m,r}, l))$. This finishes the proof of the moderate-deviations upper bound. $\qquad \square$

5. Proofs of the mean-field moderate deviations

We begin our examination of the LDP of $\{M^n\}_{n \geq 1}$ for Γ given by (2.1) by considering the LDP of $\{M^n(\phi)\}_{n \geq 1}$ where $\phi \in B(S)$ is fixed. The following proposition follows from Theorem 3.1, Corollary 3.2, the contraction principle and Varadhan's lemma.

Proposition 5.1. *If Condition* 1.1 *holds and* $W : S^2 \to \mathbb{R}$ *is symmetric, for every fixed* $\phi \in B(S)$ *the sequence* $\{M^n(\phi)\}_{n=1}^\infty$ *satisfies the LDP in* $\mathbb{R}$ *with speed* n/b_n^2 *and with good rate function*

$$I_\phi(y) = \inf\left\{ \frac{1}{2} \int_S \left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*} \right)^2 \mathrm{d}\mu^* - \int_{S^2} W \, \mathrm{d}\nu \, \mathrm{d}\nu : \int_S \phi \, \mathrm{d}\nu = y, \nu(S) = 0 \right\}.$$

Proof. To obtain the MDP we first demonstrate the upper bound. We apply Theorem 3.1 for $r = (1, 2)$, $m = (1, 2)$, $\Phi = (B(S), B(S^2))$, and $\mathcal{M}(S) \times \mathcal{M}(S^2)$ endowed with the product topology of the strong topologies of the components. By Corollary 3.2 and the contraction principle (see Dembo and Zeitouni 1998, Theorem 4.2.1), the LDP holds for every $\phi \in B(S)$

for $\{(M_n^{1,1}(\phi),\, M_n^{2,2}(W))\}_{n=1}^{\infty} = \{(M^n(\phi),\, M_n^{2,2}(W))\}_{n=1}^{\infty}$ under the distribution $\mu^{*\otimes\infty}$ in $\mathbb{R}^2$ with good rate function

$$I_{W,\phi}(y,\, x) = \inf\left\{\frac{1}{2}\int_S \left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*}\right)^2 \mathrm{d}\mu^* \mid \nu(S) = 0,\, \int_{S^2} W\mathrm{d}\nu\,\mathrm{d}\nu = x,\, \int_S \phi\mathrm{d}\nu = y\right\}.$$

For $C \subseteq \mathbb{R}$ a fixed closed set, define the upper semicontinuous function $\Phi : \mathbb{R}^2 \to \overline{\mathbb{R}}$ by

$$\Phi(y,\, x) = \begin{cases} x & y \in C, \\ -\infty & \text{otherwise.} \end{cases} \tag{5.1}$$

We obtain

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log P_\Gamma^n(M^n(\phi) \in C)$$

$$= \limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathrm{E}_{P_\Gamma^n}\left[\exp\left(\frac{b_n^2}{n}\Phi(M^n(\phi),\, 0)\right)\right]$$

$$= \limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathrm{E}_{\mu^{\otimes n}}\left[\exp\left(\frac{b_n^2}{n}\Phi\left(M^n(\phi),\, \left(\frac{n}{b_n}\right)^2 L_n^2(W)\right)\right)\right] \Big/ Z_\Gamma^n$$

$$= \limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathrm{E}_{\mu^{*\otimes n}}\left[\exp\left(\frac{b_n^2}{n}\Phi(M^n(\phi),\, M_n^{2,2}(W))\right)\right] \Big/ (Z_\Gamma^n)^*,$$

where $(Z_\Gamma^n)^* = \int \exp\{n\Gamma((1/n)\sum_i \delta_{x_i})\}\, \mathrm{d}(\mu^*)^{\otimes n}(X)$. By Varadhan's lemma (see, in particular, Dembo and Zeitouni 1998, Lemmas 4.3.4, 4.3.6 and 4.3.8) we obtain the upper bound

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log P_\Gamma^n(M^n(\phi) \in C)$$

$$\leqslant \sup_{(y,x)\in\mathbb{R}^2} \{\Phi(y,\, x) - I_{W,\phi}(y,\, x)\} - \sup_{x\in\mathbb{R}}\{x - I_W(x)\}$$

$$= -\inf_{(y,x)\in C\times R}\{I_{W,\phi}(y,\, x) - x\} + \inf_{x\in R}\{I_W(x) - x\}. \tag{5.2}$$

Here

$$I_W(x) = \inf\left\{\frac{1}{2}\int_S \left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*}\right)^2 \mathrm{d}\mu^* \mid \nu(S) = 0,\, \int_{S^2} W\mathrm{d}\nu\,\mathrm{d}\nu = x\right\},$$

so that the second term in (5.2) becomes

$$\inf\left\{\frac{1}{2}\int_S \left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*}\right)^2 \mathrm{d}\mu^* - \int_{S^2} W\mathrm{d}\nu\,\mathrm{d}\nu \mid \nu(S) = 0\right\}. \tag{5.3}$$

To apply Varadhan's lemma with the function $\Phi(\cdot,\, \cdot)$ we have to check the following moment condition: for some $\gamma > 1$,

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathrm{E}_{(\mu^*)^{\otimes n}}\left[\exp\left(\gamma\frac{b_n^2}{n}\,\Phi(M^n(\phi),\,M_n^{2,2}(W))\right)\right] < \infty. \tag{5.4}$$

That (5.4) holds follows from Hölder's inequality and Lemma 2.12.

Since $\mathrm{Id} - 2\,W$ is positive definite on $L_0^2(\mu^*)$, we have that (5.3) is zero. More explicitly, the bounded operator defined by

$$(\mathrm{Id} - 2\,W)(g) = \int_S g^2(x)\mathrm{d}\mu^*(x) - \int_{S^2} 2\,W(x,\,y)g(x)g(y)\mathrm{d}\mu^*(x)\mathrm{d}\mu^*(y),$$

for $g \in L_0^2(\mu^*)$, is positive definite (this operator appears, for example, in (1.4) of Kusuoka and Tamura 1984). Hence, we have, for some $c > 0$, that

$$(\mathrm{Id} - 2\,W)(g) \geqslant c \qquad \forall g \in L_0^2(\mu^*)\colon \|g\|_{L^2(\mu^*)} = 1,$$

which implies

$$(\mathrm{Id} - 2\,W)(g) \geqslant c\int_S g^2(x)\mathrm{d}\mu^*(x) \qquad \forall\, g \in L_0^2(\mu^*).$$

Choose, in particular, $\nu \in \mathcal{M}(S)$ such that $\nu(S) = 0$ and $\mathrm{d}\nu/\mathrm{d}\mu^* \in L_0^2(\mu^*)$, we have that

$$(\mathrm{Id} - 2\,W)\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*}$$

$$= \int_S \left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*}\right)^2(x)\mathrm{d}\mu^*(x) - \int_{S^2} 2\,W(x,\,y)\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*}(x)\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*}(y)\,\mathrm{d}\mu^*(x)\mathrm{d}\mu^*(y)$$

$$= \int_S \left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*}\right)^2(x)\mathrm{d}\mu^*(x) - \int_{S^2} 2\,W(x,\,y)\,\mathrm{d}\nu(x)\mathrm{d}\nu(y)$$

$$\geqslant c\int_S \left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*}\right)^2(x)\,\mathrm{d}\mu^*(x).$$

From this we see that the quantity (5.3) must be zero. That the first term in the right-hand side of (5.2) equals $-\inf_{y\in C} I_\phi(y)$ is clear. Hence we have the desired upper bound. The lower bound is obtained in an analogous manner (using the same exponential moment estimations as in the proof of the upper bound) and we omit the proof.

To see that the proposed rate function is lower semicontinuous, note that $I_{W,\phi}(y,x)$, being a rate function, is lower semicontinuous, while the function x is continuous. Hence subtracting the latter from the former, lower semicontinuity is maintained.

To see that the proposed rate function is good, consider a fixed level set

$$\left\{y : I_\phi(y) = \inf_{x\in\mathbb{R}}\{I_{W,\phi}(y,\,x) - x\} \leqslant \alpha\right\},$$

where $\alpha \geqslant 0$ is fixed. Suppose this set is not compact. There then exists a sequence y^n, with $y^n \to \infty$, such that $I_\phi(y^n) \leqslant \alpha$ (note that we may pick the y^n positive as $I_\phi(y) = I_\phi(-y)$). Hence, for each n there exists ν such that

$$\frac{1}{2}\int_S \left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*}\right)^2 \mathrm{d}\mu^* - \int_{S^2} W \mathrm{d}\nu \, \mathrm{d}\nu \leqslant 2\alpha \tag{5.5}$$

and

$$\int_S \phi \mathrm{d}\nu = y^n.$$

However, as $\mathrm{Id} - 2\,W$ is positive definite on $L_0^2(\mu^*)$ we have, for some $c > 0$,

$$\frac{1}{2}\int_S \left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*}\right)^2 \mathrm{d}_t\mu^* - \int_{S^2} W \mathrm{d}\nu \, \mathrm{d}\nu \geqslant c\int_S \left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*}\right)^2 \mathrm{d}_t\mu^* \tag{5.6}$$

and, by Hölder's inequality, we have

$$\infty \leftarrow y^n = \int_S \phi \mathrm{d}\nu^* \leqslant \sqrt{\int_S \phi^2 \mathrm{d}\mu^*}\sqrt{\int_S \frac{\mathrm{d}\nu}{\mathrm{d}\mu^*}\mathrm{d}\mu^*}. \tag{5.7}$$

As $y^n > 0$ we cannot have $\sqrt{\int_S \phi^2 \mathrm{d}\mu^*} = 0$ and it suffices to note that (5.5), (5.6) and (5.7) lead to an absurdity. ☐

The following extension of Proposition 5.1 is also true. We omit the proof as it is quite similar to that of the previous proposition. Fix $\phi_i \in B(S)$, $i = 1, \ldots, d$, where d is a positive integer.

Proposition 5.2. *If Condition* 1.1 *holds and* W *is symmetric, then the sequence* $\{(M^n(\phi_1), \ldots, M^n(\phi_d))\}_{n=1}^\infty$ *satisfies the MDP in* $\mathbb{R}^d$ *with speed* n/b_n^2 *and good rate function*

$$I_{\phi_1,\ldots,\phi_d}(y) = \inf_{x \in \mathbb{R}}\{I_{W,\phi_1,\ldots,\phi_d}(y, x) - x\},$$

where

$$I_{W,\phi_1,\ldots,\phi_d}(y, x) = \inf\left\{\frac{1}{2}\int_S \left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*}\right)^2 \mathrm{d}\mu^* : \nu(S) = 0, \int_{S^2} W \mathrm{d}\nu \, \mathrm{d}\nu = x, \int_S \phi_i \mathrm{d}\nu = y_i, i = 1, \ldots, d\right\}.$$

The following theorem now follows from results regarding large deviations for projective limits due to Dawson and Gärtner: we have in mind here Dembo and Zeitouni (1998, Theorem 4.6.9).

Theorem 5.3. *If Condition* 1.1 *holds and* W *is symmetric, then the sequence* $\{M^n\}_{n=1}^\infty$ *satisfies the LDP in* $\mathcal{M}(S)$ *equipped with the strong topology with speed* n/b_n^2 *and with good rate function*

$$I(\nu) = \sup_{d \in \mathbb{Z}_+} \sup_{\{\phi_1,\ldots,\phi_d \in B(S)\}} I_{\phi_1,\ldots,\phi_d}(\langle \phi_1, \nu \rangle, \ldots, \langle \phi_d, \nu \rangle), \tag{5.8}$$

where $\langle \phi_i, \nu \rangle$ *denotes* $\int_S \phi_i \, \mathrm{d}\nu$.

Remark 5.1. Note that Theorem 5.3 provides a proof of Theorem 2.1 modulo the form of the rate function.

Proof of Theorem 2.1. Given uniqueness of rate functions and Proposition 5.2, it suffices to show that the relevant upper and lower bounds hold and that $I(\nu)$ is lower semicontinuous. By the contraction principle and Corollary 3.2, the MDP holds for the sequence $\{(M^n, M_n^{2,2}(W))\}_{n=1}^{\infty}$ under the distribution $\mu^{*\otimes\infty}$ in $\mathcal{M}(S) \times \mathbb{R}$, endowed with the topology given by the product of the strong topology on $\mathcal{M}(S)$ and the usual topology on $\mathbb{R}$. The good rate function is

$$I_W(\nu, x) = \frac{1}{2}\int_S \left(\frac{d\nu}{d\mu^*}\right)^2 d\mu^*, \quad \nu(S) = 0, \quad \int_{S^2} W \, d\nu \, d\nu = x.$$

For $C \subseteq \mathcal{M}(S)$ a fixed closed set, define the upper semicontinuous function Φ by

$$\Phi(\nu, x) = \begin{cases} x & \nu \in C, \\ -\infty & \text{otherwise.} \end{cases}$$

By Varadhan's lemma (in particular, Dembo and Zeitouni 1998, Lemma 4.3.4, Lemma 4.3.6) and (5.4) we have

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log P_\Gamma^n(M^n \in C) = \limsup_{n\to\infty} \frac{n}{b_n^2} \log E_{P_\Gamma^n}\left[\exp\left(\frac{b_n^2}{n}\Phi(M^n, 0)\right)\right]$$

$$= \limsup_{n\to\infty} \frac{n}{b_n^2} \log E_{\mu^{\otimes n}}\left[\exp\left(\frac{b_n^2}{n}\Phi\left(M^n, \left(\frac{n}{b_n}\right)^2 L_n^2(W)\right)\right)\right] / Z_\Gamma^n$$

$$= \limsup_{n\to\infty} \frac{n}{b_n^2} \log E_{\mu^{*\otimes n}}\left[\exp\left(\frac{b_n^2}{n}\Phi\left(M^n, M_n^{2,2}(W)\right)\right)\right] / (Z_\Gamma^n)^*$$

$$\leq \sup_{(\nu,x)\in(S)\times\mathbb{R}}\{\Phi(\nu, x) - I_W(\nu, x)\} - \sup_{x\in\mathbb{R}}\{x - I_W(x)\}$$

$$= -\inf_{(\nu,x)\in C\times\mathbb{R}}\{I_W(\nu, x) - x\} + \inf_{x\in\mathbb{R}}\{I_W(x) - x\}.$$

Here $I_W(x)$ is as in the proof of Proposition 5.1 and $\inf_{x\in\mathbb{R}}\{I_W(x) - x\} = 0$. The lower bound is obtained in an analogous manner.

To see that the proposed rate function is lower semicontinuous, note that $I_W(\cdot, x)$, a rate function, is lower semicontinuous and hence, as $-x$ is clearly lower semicontinuous, the sum of these two functions is also lower semicontinuous. $\qquad\square$

Proof of Theorem 2.2. We proceed as in the proof of Theorem 2.1. By Corollary 3.2 and the contraction principle (see Dembo and Zeitouni 1998, Theorem 4.2.1), the LDP holds for every $\phi \in B(S)$ for

$$\{(M^n(\phi), M_n^{2,2}(W_2), M_n^{3,2}(W_3), \ldots, M_n^{m,2}(W_m))\}_{n=1}^{\infty}$$

under the distribution $\mu^{*\otimes\infty}$ in $\mathbb{R}^m$. Again by the contraction principle the LDP holds for

$$\left\{ \left(M^n(\phi), \sum_{k=2}^m M_n^{k,2}(W_k) \right) \right\}_{n=1}^\infty$$

with good rate function

$I_{W,\phi}(y, x)$

$$= \inf \left\{ \frac{1}{2} \int_S \left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*} \right)^2 \mathrm{d}\mu^* \,\middle|\, \nu(S) = 0, \sum_{k=2}^m \frac{k(k-1)}{2} \int_{S^k} W_k \,\mathrm{d}\nu \,\mathrm{d}\nu \,\mathrm{d}(\mu^*)^{\otimes k-2} = x, \int_S \phi \,\mathrm{d}\nu = y \right\}.$$

For $C \subseteq \mathbb{R}$ a fixed closed set and Φ given by (5.1), the upper bound is given by

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log P_\Gamma^n(M^n(\phi) \in C) \le \sup_{(y,x)\in\mathbb{R}^2} \{ \Phi(y, x) - I_{W,\phi}(y, x) \} - \sup_{x\in\mathbb{R}} \{ x - I_W(x) \}$$

$$= - \inf_{(y,x)\in C\times R} \{ I_{W,\phi}(y, x) - x \} + \inf_{x\in R} \{ I_W(x) - x \}. \tag{5.9}$$

Here

$$I_W(x) = \inf \left\{ \frac{1}{2} \int_S \left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu^*} \right)^2 \mathrm{d}\mu^* \,\middle|\, \nu(S) = 0, \sum_{k=2}^m \frac{k(k-1)}{2} \int_{S^k} W_k \,\mathrm{d}\nu \,\mathrm{d}\nu = x \right\}.$$

Analogously to the proof of Proposition 5.1, the second term in (5.9) becomes 0, using the positive definiteness of $\mathrm{Id} - \Xi$. To be in the situation of Lemma 4.3.6 in Dembo and Zeitouni (1998) for the function $\Phi(\cdot, \cdot)$ we have to check the following moment condition (see Dembo and Zeitouni 1998 Lemma 4.3.8): for some $\gamma > 1$

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathrm{E}_{(\mu^*)^{\otimes n}} \left[\exp \left(\gamma \frac{b_n^2}{n} \Phi \left(M^n(\phi), \sum_{k=2}^m M_n^{k,2}(W_k) \right) \right) \right] < \infty.$$

But this is true by Lemma 2.4. Now we proceed as in the proof of Proposition 5 to obtain the lower bound and the properties of the rate function. $\square$

6. Applications

6.1. Curie–Weiss model

In the Curie–Weiss model, $S = \{-1, +1\}$, $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$ and

$$\Gamma(\nu) = \Gamma_{\beta,h}(\nu) := \beta \langle \nu, x \rangle^2 + h \langle \nu, x \rangle,$$

hence

$$P_\Gamma^n(\mathrm{d}X) = \frac{1}{Z_\Gamma^n} \exp \left\{ \beta \frac{1}{n} \sum_{i,j=1}^n x_i x_j + h \sum_{j=1}^n x_j \right\} \mathrm{d}\mu^{\otimes n}(X);$$

$\beta > 0$ is called inverse temperature and $h > 0$ the strength of an external magnetic field. Define

$$H(x) = -\beta x^2 - hx + \frac{1 - x}{2} \log(1 - x) + \frac{1 + x}{2} \log(1 + x),$$

if $|x| \leq 1$, and $+\infty$ otherwise; the rate function of the LDP is given by $H(x) - \inf_x H(x)$. It is well known (see, for example, Ellis 1985, Section IV.4) that in the case $h = 0$ and $\beta < \frac{1}{2}$ or in the case $h \neq 0$ there is a unique minimizer μ^* of $H(\cdot)$ which is non-degenerate. In the case $h = 0$ and $\beta > \frac{1}{2}$ there are two non-degenerate minima μ_1^*, μ_2^*. The case $h = 0$ and $\beta = \frac{1}{2}$ is degenerate. In all cases and for any minimum, $\mu^*(1)$ satisfies the relation

$$\log\left(\frac{\mu^*(1)}{1 - \mu^*(1)}\right) = 4\beta(2\mu^*(1) - 1) + 2h\mu^*(1),$$

with $\mu^*(1) = \frac{1}{2}$ for $h = 0$ and $\beta \leq \frac{1}{2}$. For $0 < \beta < \frac{1}{2}$ and $h = 0$ in Ellis and Newman (1978) the central limit theorem is proved:

$$\frac{\sum_{i=1}^n X_i}{\sqrt{n}} \to N(0, \sigma^2(\beta))$$

in distribution with respect to the Curie–Weiss finite-volume Gibbs states with $\sigma^2(\beta) = (4(1 - 2\beta))^{-1}$.

Theorem 2.1 presents a MDP for the Curie–Weiss model. The rate function is

$$I(x) = 2x^2 - 4\beta x^2.$$

6.2. Diffusion processes with mean-field interaction

6.2.1. McKean type

We now consider an application of Corollary 2.3 to the empirical distribution of the stochastic differential equation

$$dX_n^i(t, w) = dB^i(t, w) + \frac{1}{n} \sum_{j=1}^n \operatorname{grad} \Phi(X_n^i(t, w) - X_n^j(t, w))dt$$

$$X_n^i(0, w) = Y^i(w), \qquad i = 1, \ldots, n,$$

where $B^i(t, w)$, $i = 1, 2, \ldots$, are independent $\mathbb{R}^d$-valued Brownian motions, Y^i are i.i.d. $\mathbb{R}^d$-valued random variables with $E[\|Y_1\|^2] < \infty$ and Φ is a rapidly decreasing smooth function on $\mathbb{R}^d$ satisfying $\Phi(z) = \Phi(-z)$, $z \in \mathbb{R}^d$. With S the Polish space of $\mathbb{R}^d$-valued continuous functions on $[0, T]$, $T \in \mathbb{R}$, denoted by $C([0, T], \mathbb{R}^d)$, it is known (cf. Kusuoka and Tamura 1984; McKean 1967) that, for any bounded continuous function f on S^k, $k \geq 1$,

$$\lim_{n \to \infty} \mathrm{E}[f(X_n^1(\cdot, w), \ldots, X_n^k(\cdot, w))] = \int_{S^k} f(w^1, \ldots, w^k) \prod_{i=1}^k R_0(\mathrm{d}w^i).$$

Here R_0 denotes the distribution on S induced by the solution of the following stochastic differential equation of *McKean type*:

$$\mathrm{d}X(t, w) = \mathrm{d}B^1(t, w) + \left(\int_{\mathbb{R}^d} \mathrm{grad}\, \Phi(X(t, w) - z) u_t(\mathrm{d}z) \right) \mathrm{d}t$$

$$X(0, w) = Y^1(w),$$

where $u_t(\mathrm{d}z)$ is the probability distribution of $X(t, w)$.

Suppose that μ is the probability on S induced by $Y^1(w) + B^1(t, w)$, $0 \leqslant t \leqslant T$. Let ν_n, $n \in \mathbb{N}$, be the probability distribution on S^n induced by $(X_n^1(t, w), \ldots, X_n^n(t, w))$, $0 \leqslant t \leqslant T$. Then by the Cameron–Martin formula one obtains the mean-field character of the model:

$$\frac{\mathrm{d}\nu_n}{\mathrm{d}\mu^{\otimes n}}(w_n) = \exp \left(\frac{1}{n} \sum_{i,j=1}^n \int_0^T \mathrm{grad}\, \Phi(w^i(t) - w^j(t))\, \mathrm{d}w^i(t) \right.$$

$$\left. - \frac{1}{2n^2} \sum_{i=1}^n \int_0^T \left\| \sum_{j=1}^n \mathrm{grad}(w^i(t) - w^j(t)) \right\|^2 \mathrm{d}t \right),$$

where $w^n = (w^1, \ldots, w^n) \in S^n$ and the first integral is an Itô integral.

Theorem 6.1. *The sequence of empirical measures* $\{M^n(R_0)\}_{n \geqslant 1}$ *satisfies the LDP with speed* n/b_n^2 *and with good convex rate function*

$$I(\nu) = \frac{1}{2} \int_S \left(\frac{\mathrm{d}\nu}{\mathrm{d}R_0} \right)^2 \mathrm{d}R_0 - 3 \int V \,\mathrm{d}\nu \,\mathrm{d}\nu \,\mathrm{d}R_0,$$

where

$$V(x, y, z) := \frac{1}{3} \{ V_2(x, y) + V_2(y, z) + V_2(z, x) \} + \frac{1}{3} \{ V_3(x, y, z) + V_3(y, z, x) + V_3(z, x, y) \},$$

$$V_2(x, y) = \frac{1}{2} \left\{ \Phi(x(T) - y(T)) - \Phi(x(0) - y(0)) - \int_0^T \Delta\Phi(x(t) - y(t))\mathrm{d}t \right\},$$

$$V_3(x, y, z) = -\frac{1}{2} \left\{ \int_0^T \mathrm{grad}\, \Phi(x(t) - y(t))\, \mathrm{grad}\, \Phi(x(t) - z(t))\, \mathrm{d}t.$$

Proof. We apply Corollary 2.3 with $m = 3$. Condition 1.1 follows from Theorems 3 and 4 and the proof of Theorem 5 in Kusuoka and Tamura (1984). The boundedness is easily

verified. V even satisfies the conditions assumed in Ben Arous and Brunaud (1989), which are like (1.4). □

6.2.2. *Langevin dynamics with McKean–Vlasov limit*

Next we apply the results to the Langevin dynamics of a system of interacting particles leading to a McKean–Vlasov limit. Let $X_n^i(t, \cdot)$ satisfy the system of stochastic differential equations as in Ben Arous and Brunaud (1989):

$$\mathrm{d}X_n^i(t, \omega) = \mathrm{d}B_n^i(t, \omega) - \nabla U(X_n^i(t, \omega))\, \mathrm{d}t + \frac{1}{n} \sum_{i \leqslant j \leqslant n} \nabla_1 V(X_n^i(t, \omega), X_n^j(t, \omega))\, \mathrm{d}t.$$

Again the $(B_n^i)_{1 \leqslant i \leqslant n}$ are n independent $\mathbb{R}^d$-valued Brownian motions, $U : \mathbb{R}^d \to \mathbb{R}^d$ is a C_b^2 function and $V \in \tilde{\otimes}_\pi^2 C_b^2(\mathbb{R}^d, \mathbb{R})$, where $\tilde{\otimes}_\pi^2$ denotes the two-fold projective tensor product of $C_b^2(\mathbb{R}^d, \mathbb{R})$. V is assumed to admit the representation

$$V = \int_C g(\cdot, \tau)^{\otimes 2}\, v(\mathrm{d}\tau)$$

with (C, v) a compact space and $g : \mathbb{R}^d \times C \to \mathbb{R}$ differentiable in its first coordinate for all $\tau \in C$. It is known that V has such a representation with continuous g; see Ben Arous and Brunaud (1998, Corollary 1.6). Finally, we assume for the initial law of $(X_n^i(0, \cdot))_{1 \leqslant i \leqslant n}$ that

$$\mu_0^n = \mathrm{law}(\{X_n^i(0, \cdot)\}_{1 \leqslant i \leqslant n})$$

$$= Z_n^{-1} \exp\left(n^{1-r} \sum_{1 \leqslant i_1 \leqslant \ldots \leqslant i_r} S(x_{i_1}, \ldots, x_{i_r}) \right) \mathrm{d}\mu_0^{\otimes n}(\mathrm{d}x_1, \ldots, \mathrm{d}x_n)$$

with $S \in \tilde{\otimes}_\pi^r C_b(\mathbb{R}^d, \mathbb{R})$, $r \geqslant 2$. Denote by $v_{k,n} \in \mathcal{M}_1(C([0, T], (\mathbb{R}^d)^k))$ the law of $(X_n^1, \ldots, X_n^k)$. The empirical measure of the X_n^i converges to Dirac measure concentrated in the solution of the following SDE. Let B_t denote an $\mathbb{R}^d$-valued Brownian motion, independent of X_0, and for $\mu \in \mathcal{M}_1(\mathbb{R}^d)$ let

$$\mathrm{d}X_t = \mathrm{d}B_t - \nabla U(X_t)\, \mathrm{d}t + \int \nabla_1 V(X_t, y)\, u_t(\mathrm{d}y)\, \mathrm{d}t$$

with u_t the law of X_t, $u_0 = \mu$. Then it is known that under our assumptions the process X_t exists and is well defined. It is the *McKean–Vlasov* nonlinear diffusion; see Funaki (1984). Let $P_T(\mu) \in \mathcal{M}_1(C([0, T], \mathbb{R}^d)) =: S$ denote its law. We are now ready to state the result for this model:

Theorem 6.2. *Let $s = r \vee 3$. The sequence of empirical measures $\{M^n(P_T(\mu))\}_{n \geqslant 1}$ satisfies the LDP with speed n/b_n^2 and with good convex rate function*

$$I(v) = \frac{1}{2} \int_S \left(\frac{\mathrm{d}v}{\mathrm{d}P_T(\mu)} \right)^2 \mathrm{d}P_T(\mu) - \frac{s(s-1)}{2} \int_{S^s} \gamma_{0,T} \mathrm{d}v\, \mathrm{d}v\, \mathrm{d}(P_T(\mu))^{\otimes s-2},$$

where $\gamma_{0,T} : S^s \to \mathbb{R}$ *is a symmetric function defined by* (2.5) *in Ben Arous and Brunaud* (1989) *(depending on V and S).*

Proof. By calculations in Ben Arous and Brunaud (1989) we see that the law $\nu_{n,n}$ is of mean-field Gibbs type with respect to $P_T(\mu)^{\otimes n}$. The potential function $\gamma_{0,T}$ is in the space $\tilde{\otimes}^s_\pi C_b(S)$ (see Ben Arous and Brunaud 1989, (3.6) and Corollary 3.21). That $\gamma_{0,T}$ possesses a unique non-degenerate minimizer of the LDP rate function which is $P_T(\mu)$ is proved in Ben Arous and Brunaud (1989, Corollary 3.10 and Theorem 2.7). $\square$

Acknowledgement

References

Arcones, M. (1999) The large deviation principle for empirical processes. Preprint.

Arcones, M.A. and Giné, E. (1993) Limit theorems for *U*-processes. *Ann. Probab.*, **21**, 1494–1542.

Ben Arous, G. and Brunaud, M. (1989) Méthode de Laplace: étude variationnelle des fluctuations de diffusions de type 'champ moyen'. *Stochastics*, **31**, 79–144.

Ben Arous, G. and Zeitouni, O. (1999) Increasing propagation of chaos for mean field models. *Ann. Inst. H. Poincaré Probab. Statist.*, **35**(1), 85–102.

Bolthausen, E. (1986) Laplace approximations for sums of independent random vectors. *Probab. Theory Related fields*, **72**, 305–318.

de Acosta, A. (1994) Projective systems in large deviation theory II: some applications. In J. Hoffman-Jørgensen, J. Kuelbs and M. Marcus (eds), *Probability in Banach Spaces*, 9, Progr. Probab. 35, pp. 241–250. Boston: Birkhäuser.

de la Peña, V. (1992) Decoupling and Khintchine's inequalities for *U*-statistics. *Ann. Probab.*, **20**, 1877–1892.

del la Peña, V. and Giné, E. (1999) *Decoupling: from Dependence to independence*. New York: Springer-Verlag.

Dembo, A. and Zeitouni, O. (1998) *Large Deviations Techniques and Aplications*. New York: Springer-Verlag.

Eichelsbacher, P. and Schmock, U. (2001) Rank-dependent moderate deviations of *U*-empirical measures in strong topologies. *Probab. Theory Related Fields*. To appear.

Eichelsbacher, P. and Schmock, U. (2002) Large deviations for products of empirical measures in strong topologies and applications. *Ann. Inst. H. Poincaré Probab. Statist.*, **38**, 779–797.

Eichelsbacher, P. and Zajic, T. (2002) Refinements of the CLT for mean field Gibbs measures. Preprint.

Ellis, R.S. (1985) *Entropy, Large Deviations, and Statistical Mechanics*. New York: Springer-Verlag.

Ellis, R.S. and Newman, C.M. (1978) Limit theorems for sums of dependent random variables occurring in statistical mechanics. *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, **44**, 117–139.

Fuglede, B. (1960) On the theory of potentials in locally compact spaces. *Acta Math.*, **103**, 139–215.

Funaki, T. (1984) A certain class of diffusion processes associated with nonlinear parabolic equations. *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, **67**, 331–348.

Guionnet, A. (1999) Stability of precise Laplace's method under approximations: applications. *ESAIM Probab. Statist.*, **3**, 67–88.

Kusuoka, S. and Tamura, Y. (1984) Gibbs measures for mean field potentials. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, **31**(1), 223–245.

McKean, Jr., H.P. (1967) Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pp. 41–57. Air Force Office Sci. Res., Arlington, Va.

McKean, H.P. (1975) Fluctuations in the kinetic theory of gases. *Comm. Pure Appl. Math.*, **28**, 435–455.

Sznitman, A.-S. (1984) Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated. *J. Funct. Anal.*, **56**, 311–336.

Tanaka, H. (1984) Limit theorems for certain diffusion processes with interaction. In K. Itô (ed.), *Stochastic Analysis*, North-Holland Math. Library 32, pp. 469–488. Amsterdam: North-Holland.