Large and moderate deviations for functionals depending of infinite variables of an i.i.d. sequence.

Hacène DJELLOUT, Arnaud GUILLIN and Liming WU

{djellout,guillin,wuliming}@ucfma.univ-bpclermont.fr
Laboratoire de Mathématiques Appliquées, CNRS-UMR 6620
Université Blaise Pascal, 63177 AUBIERE, France

Abstract : Let $\{\xi_n,\ n \in \mathbb{Z}\}$ be a sequence of E-valued independent and identically distributed random variable, where E is a polish space, and let $X = (X_n = \Phi(\xi_{n+\cdot}))_{n\in\mathbb{Z}}$ for some measurable mapping Φ from $E^{\mathbb{Z}}$ to F, another polish space. Under some assumptions on the law of ξ_0 and on the mapping Φ, we establish the process-level large deviation principle and the moderate deviations of X, and we identify the rate functions. For this we use an improved version of the generalized contraction principle and an (perhaps new) exponential inequality. The application to the linear case of moving average processes are considered.

1991 AMS Subject Classification : 60F10.

Key Words : large deviations; moderate deviations; moving average processes; generalized contraction principle.

1 Introduction.

Let $\{\xi_n,\ n \in \mathbb{Z}\}$ be a doubly infinite sequence of E-valued independent and identically distributed random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where E is a polish space with metric d_E . This sequence can be realized as the coordinates $\xi_n(\omega) = \omega(n)$ of the product topological space $\Omega = E^{\mathbb{Z}}$ equipped with the product probability measure $\mathbb{P} = \alpha^{\mathbb{Z}}$, where α is the common law of ξ_n.

Let $\Phi : E^{\mathbb{Z}} \to F$ be a measurable mapping where F is a polish space with metric d_F, and consider the process

$$X_n = \Phi(\xi_{n+\cdot}) = \Phi(..., \xi_{i+n}, ...) \quad, \forall n \in \mathbb{Z} \quad \mathbb{P} - a.s. \tag{1.1}$$

It appears often in filtrage, time series analysis, statistics and dynamical systems in the following fashions:

1. $\xi = (\xi_n)$ is the noise, and $X = (X_n)$ is the (non-linear) filtrage.
2. The stationary process (X_n) represents the received message at instant n, which is assumed to be of form (1.1).
3. A wide class of dynamical systems can be described as (1.1).

We deal in this paper with the process level Large Deviation Principle and the moderate deviations of (X_n). The special case of the moving average, when Φ is linear (more exactly when $X_n := \sum_{j=-\infty}^{+\infty} a_j \xi_{n+j}$), has attracted much attention and many works. Several works have been realized by Burton and Dehling [BD90], Jiang, Rao, and Wang [JRW92], [JRW95], Djellout and Guillin [DG98] and recently by Wu [Wu99] on the linear case, under different assumptions on the law ξ_0, and the spectral density function of X, see Wu [Wu99] and section 4. for relevant reference and more details.

Introduce now some notations. Let $\tau^k\omega = \omega(k+\cdot)$, $k \in \mathbb{Z}$, be the shifts on Ω. Consider the process level occupation measures of the i.i.d. sequence

$$O_n(\omega) := \frac{1}{n}\sum_{k=1}^{n} \delta_{\tau^k\omega}, \forall n \geq 1,$$

which are random elements in $M_1(\Omega)$, the space of all probability measures on Ω. The well known result due to Donsker-Varadhan [DV85] says that $\mathbb{P}(O_n \in \cdot)$ satisfies the LDP on $M_1(\Omega)$ equipped with the usual weak convergence topology, with speed n and with the rate function given by the entropy functional below

$$H(Q) = \begin{cases} \mathbb{E}^Q \log \dfrac{Q_{\omega(-\infty,0]}(dx)}{\alpha(dx)}, & if\ Q \in M_1^s(\Omega); Q_{\omega(-\infty,0]}(dx) \ll \alpha(dx),\ Q-a.s. \\ +\infty & otherwise \end{cases}$$

where $Q_{\omega(-\infty,0]}(dx) := Q(\omega(1) \in dx | \omega(k), k \leq 0)$ is the regular conditional distribution, $M_1^s(\Omega)$ is the space of those elements $Q \in M_1(\Omega)$ such that $Q \circ \tau^{-1} = Q$ (or Q is stationary).

Let

$$\bar{\Phi} = (\Phi_n)\ ,\ \Phi_n(\omega) := \Phi(\omega(n+\cdot)),\ (\Phi \ \text{ is identified as } \ \Phi_0).$$

Define the process level occupation measure R_n by

$$R_n = \frac{1}{n}\sum_{k=1}^{n} \delta_{X_{k+\cdot}} = \frac{1}{n}\sum_{k=1}^{n} \delta_{\bar{\Phi}(\tau^k\omega)} = O_n \circ \bar{\Phi}^{-1} \quad \mathbb{P}-a.s.$$

which are random elements in $M_1(F^{\mathbb{Z}})$.

The first purpose of this paper is to establish the process-level large deviation principle of X, i.e, the LDP of (R_n).

Let us regard roughly the feature of this question. At first when $\Phi : E^{\mathbb{Z}} \to F$ is continuous w.r.t the product topology of $E^{\mathbb{Z}}$, then by the contraction principle, the LDP of (R_n) follows from that of (O_n) with speed n and with the rate function given by

$$J(\hat{Q}) = \inf\left\{H(Q) | Q \in M_1^s(\Omega),\ H(Q) < \infty,\ \hat{Q} = Q \circ \bar{\Phi}^{-1}\right\}\ .$$

We will consider here the case that Φ is not necessarily continuous. All of our results rely on the assumption of the existence of a sequence of continuous mappings Φ^N such that $\|\nabla(\Phi - \Phi^N)\|_{\infty,1} \to 0$ (see section 3 for details on notation, and it is equivalent to $\sum_{j\in\mathbb{Z}} |a_j| < +\infty$ in the linear case). Under the hypothesis $\mathbb{E}^\alpha(e^{\delta d_E(\xi_0,o)}) < +\infty$ for some $\delta > 0$, ($o \in E$ some fixed point), we establish the large deviation principle for $\mathbb{P}(R_n \in \cdot)$ in $M_1(F^{\mathbb{Z}})$ equipped with the weak convergence topology, using a generalized contraction principle, proved in [Wu99]. The second purpose of this paper is to establish the moderate deviations of (X_n).

Let us present now the structure of this paper. In section 2, we set some preliminary results (as the generalized contraction principle), especially a key exponential inequality will be proved here. Then in section 3, we give the main results of this paper, the large deviation principle of (R_n), and the moderate deviation principle of (X_n), and we prove these results. Section 4 is devoted to the moving average case.

2 Preliminary results.

First, recall the following [DS89] , [DZ93]:

Definition : *Let (μ_n) be a family of probability measures on a polish space E. We say that (μ_n) satisfies the large deviation principle (in short : LDP) on the topological space E with speed $\lambda(n)$ (a sequence tending to infinity) and with the rate function I if*

- **(i)** *$\forall L > 0$, the level set $[I \leq L]$ is compact in E, (or I is a good rate function).*
- **(ii)** *For each open subset $O \subset E$,*

$$\liminf_{n\to\infty} \frac{1}{\lambda(n)} \log \mu_n(O) \geq -\inf_{x\in O} I(x) \ ;$$

- **(iii)** *For each closed subset $C \subset E$,*

$$\limsup_{n\to\infty} \frac{1}{\lambda(n)} \log \mu_n(C) \leq -\inf_{x\in C} I(x) \ .$$

We give now two lemmas taken from [Wu99, Th. 1.1 and Th. 1.2], which are very useful for our purpose :

Lemma 2.1 (Wu) *If $(Z_n^{(N)}), (Z_n)$ are E-valued r.v defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that*

- **(i)** *For each N, $\mathbb{P}(Z_n^{(N)} \in \cdot)$ satisfies LDP on E with speed $\lambda(n)$ and the rate function $I^{(N)}$;*
- **(ii)** *$(Z_n^{(N)})$ is an exponential good approximation of (Z_n),i.e.,*

$$\lim_{N\to\infty} \limsup_{n\to\infty} \frac{1}{\lambda(n)} \log \mathbb{P}\Big(d_E(Z_n^{(N)}, Z_n) > \delta\Big) = -\infty, \quad \forall \delta > 0 \ .$$

Then $\mathbb{P}(Z_n \in \cdot)$ satisfies LDP on E with speed $\lambda(n)$ and rate function I given by

$$I(x) := \sup_{\delta>0} \liminf_{N\to\infty} \inf_{B(x,\delta)} I^{(N)} = \sup_{\delta>0} \limsup_{N\to\infty} \inf_{B(x,\delta)} I^{(N)} \ .$$

where $B(x,\delta) = \{y \in E; d_E(x,y) < \delta\}$ is the ball centered at x with radius δ.

Remark :
The main difference from [DZ93, Th.4.2.16] is : their technical assumption of (ii) and of the inf compactness of I is dropped, and the inf-compactness of I becomes now a consequence.

Lemma 2.2 (Wu) *Let E, F be two Polish spaces and d_F a metric compatible with the topology of F. Given a family of probability measures (μ_n) on E such that (μ_n) satisfies LDP on E with speed $\lambda(n)$ and good rate function I, and a measurable application $f : E \to F$.*
If there is a sequence of continuous mappings $f^N : E \to F$ such that

$$\lim_{N\to\infty} \limsup_{n\to\infty} \frac{1}{\lambda(n)} \log \mu_n(d_F(f^N, f) > \delta) = -\infty, \quad \forall \delta > 0,$$

then there is a continuous function $\widetilde{f} : [I < \infty] \subset F$ *such that*

$$\sup_{x \in [I \le L]} d_F(f^N(x), \widetilde{f}) \to 0, \quad as\ N \to \infty,\ \forall L > 0, \tag{2.1}$$

and $\mu_n o f^{-1}$ *satisfies LDP on* F *with speed* $\lambda(n)$ *and good rate function* $I_{\widetilde{f}}$ *given by*

$$I_{\widetilde{f}}(y) := \inf\Big\{ I(x) | I(x) < +\infty; \widetilde{f}(x) = y \Big\};\ \forall y \in F.$$

Moreover $[I_{\widetilde{f}} \le L] = \widetilde{f}([I \le L]),\ \forall L \in \mathbb{R}^+$.

Remark :
This result extends Theorem 4.2.23 of Dembo-Zeitouni [DZ93]. In their Theorem, (2.1) is an assumption with $\widetilde{f} = f$, while it becomes now a consequence of the large deviation negligeability in this new version. Note that in practice f is often only well defined μ_n *a.s.* and the conclusion means that f admits a continuous version $\widetilde{f}$ on $[I < \infty]$ (which is often of zero μ_n measure however), in some sense.

We need also the following exponential inequality, which will be crucial for the proof of our main results:

Lemma 2.3 *Let* $f : (\Omega = E^{\mathbb{Z}}, \mathbb{P} = \alpha^{\mathbb{Z}}) \to \mathbb{R}$ *be a measurable function. We suppose that there is* $\delta > 0$ *such that* $\mathbb{E}^\alpha(e^{\delta d(\xi_0, o)}) < +\infty$ *(where* $o \in E$ *is some fixed point); then*

$$\mathbb{P}\Big(|f - \mathbb{E}(f)| > r\Big) \le 2 \exp\left(\frac{-r^2}{2C(\delta)\sum_{m \in \mathbb{Z}} \|\partial_m f\|_\infty^2} \vee \frac{-\delta r}{4 \sup_{m \in \mathbb{Z}} \|\partial_m f\|_\infty} \right) \tag{2.2}$$

where

$$\|\partial_m f\|_\infty := \left\| \frac{f(\xi_{]-\infty, m-1]}, \omega_m, \omega_{[m+1, +\infty[}) - f(\xi_{]-\infty, m-1]}, \xi_m, \omega_{[m+1, +\infty[})}{d_E(\omega_m, \xi_m)} \right\|_{L^\infty(\mathbb{P}(d\xi) \otimes \mathbb{P}(d\omega))}$$

and

$$C(\delta) = \mathbb{E}\left(\Big(d(\xi_0, o) + \mathbb{E}d(\xi_0, o) \Big)^2 e^{\frac{\delta}{2}(d(\xi_0, o) + \mathbb{E}d(\xi_0, o))} \right).$$

Proof : We can assume that $\sum_{m \in \mathbb{Z}} \|\partial_m f\|_\infty^2 < +\infty$ (trivial otherwise).
On the extended filtered probability space

$$\begin{cases} \bar{\Omega} := E^{\mathbb{Z}} \times E^{\mathbb{Z}} = \{(\omega_k, \xi_k)_{k \in \mathbb{Z}} | \omega_k \in E; \xi_k \in E\}\ , \\ \bar{\mathbb{P}} := \alpha^{\mathbb{Z}} \times \alpha^{\mathbb{Z}} = \mathbb{P} \times \mathbb{P}\ ; \\ \bar{\mathcal{F}}_n := \sigma((\omega_k, \xi_k); k \le n)\ ; \end{cases}$$

we consider $f(\xi, \omega) = f(\xi)$ as a function on $\bar{\Omega}$. Assume at first that f is bounded.
So, on $(\bar{\Omega}, \bar{\mathcal{F}}, (\bar{\mathcal{F}}_n), \bar{\mathbb{P}})$, we can define the martingale $M_n = \mathbb{E}^{\bar{\mathbb{P}}}(f(\xi)|\bar{\mathcal{F}}_n)$ and $M_\infty = f$. Consider the martingale difference : $m_n = M_n - M_{n-1}$. By Kolmogorov's 0-1 law $\mathbb{E}^{\mathbb{P}}(f) = \mathbb{E}^{\bar{\mathbb{P}}}(f|\bar{\mathcal{F}}_{-\infty})$ and by the martingale convergence we have

$$f - \mathbb{E}^{\mathbb{P}}(f) = M_\infty - \mathbb{E}^{\bar{\mathbb{P}}}(M_\infty) = \sum_{k \in \mathbb{Z}} m_k, \quad (L^2(\bar{\mathbb{P}})\text{−convergence})\ .$$

We have now to estimate

$$\left|\mathbb{E}^{\bar{\mathbb{P}}}(e^{\sum_{k\in\mathbf{Z}} m_k})\right| \le \prod_{k\in\mathbb{Z}} \|\mathbb{E}^{k-1}(e^{\lambda m_k})\|_\infty$$

where $\mathbb{E}^{k-1}(\cdot) = \mathbb{E}^{\bar{\mathbb{P}}}(\cdot|\bar{\mathcal{F}}_{k-1})$. By Taylor formula, we have

$$\mathbb{E}^{k-1}(e^{\lambda m_k}) \le 1 + \frac{\lambda^2}{2}\mathbb{E}^{k-1}(m_k^2 e^{|\lambda m_k|}).$$

Our key observation is: m_k depends only on ξ and

$$\begin{aligned}
|m_k(\xi)| &= \left|\int f(\xi_{(-\infty,k-1]},\xi_k,\omega_{k+1},...)\mathbb{P}(d\omega) - \int f(\xi_{(-\infty,k-1]},\omega_k,\omega_{k+1},...)\mathbb{P}(d\omega)\right| \\
&\le \int \mathbb{P}(d\omega)\|\partial_k f\|_\infty d(\xi_k,\omega_k) \\
&\le \|\partial_k f\|_\infty\Big(d(\xi_k,o) + \mathbb{E}^{\mathbb{P}} d(\omega_k,o)\Big) \\
&= \|\partial_k f\|_\infty (d(\xi_k,o)+C),
\end{aligned}$$

where $C = \mathbb{E}^{\mathbb{P}}(d_E(\xi_0,o)) < \infty$ and $o \in E$ is some fixed point. Using this inequality and the fact that $1+x \le e^x$ for all $x \in \mathbb{R}$, we get for all $|\lambda| \le \frac{\delta}{2\sup_k \|\partial_k f\|_\infty} := \delta'$

$$\begin{aligned}
\left|\mathbb{E}^{k-1}\Big(e^{\lambda m_k}\Big)\right| &\le 1 + \frac{\lambda^2}{2}\mathbb{E}\Big(\|\partial_k f\|_\infty^2 (d(\xi_k,o)+C)^2 e^{|\lambda|(d(\xi_k,o)+C)\ \|\partial_k f\|_\infty}\Big) \\
&\le \exp\left(\frac{\lambda^2}{2}C(\delta)\|\partial_k f\|_\infty^2\right),
\end{aligned}$$

where

$$C(\delta) = \mathbb{E}\left((d(\xi_0,o)+C)^2 e^{\frac{\delta}{2}(d(\xi_0,o)+C)}\right).$$

Thus

$$\mathbb{E}^{\mathbb{P}}\left(e^{\lambda(f-\mathbb{E}(f))}\right) \le \exp\left(\frac{\lambda^2}{2}C(\delta)\sum_{k\in\mathbb{Z}}\|\partial_k f\|_\infty^2\right), \quad \forall|\lambda| \le \delta' . \tag{2.3}$$

By Chebychev's inequality, we have for all $r \ge 0,\ 0 < \lambda \le \delta'$

$$\begin{aligned}
\mathbb{P}(f - \mathbb{E}(f) > r) &\le e^{-\lambda r}\mathbb{E}^{\mathbb{P}}\left(e^{\lambda(f-\mathbb{E}(f))}\right) \\
&\le \exp\left(-\lambda r + \frac{\lambda^2}{2}C(\delta)\sum_k \|\partial_k f\|_\infty^2\right).
\end{aligned}$$

We now have to optimize this inequality regarding only on λ. We consider the following two cases.

1. When $r \le \frac{\delta C(\delta)}{2\sup_k\|\partial_k f\|_\infty}\sum_k \|\partial_k f\|_\infty^2$, choose $\lambda = \frac{r}{C(\delta)\sum_k \|\partial_k f\|_\infty^2} \le \delta'$, for which

$$\exp\left(-\lambda r + \frac{\lambda^2}{2}C(\delta)\sum_k \|\partial_k f\|_\infty^2\right) = \exp\left(-\frac{r^2}{2C(\delta)\sum_{m\in\mathbb{Z}}\|\partial_m f\|_\infty^2}\right).$$

2. When $r > \frac{\delta C(\delta)}{2\sup_k \|\partial_k f\|_\infty} \sum_k \|\partial_k f\|_\infty^2$, we choose $\lambda = \frac{\delta}{2\sup_k \|\partial_k f\|_\infty}$, for which

$$\exp\left(-\lambda r + \frac{\lambda^2}{2} C(\delta) \sum_k \|\partial_k f\|_\infty^2\right) \leq \exp\left(-\frac{\delta r}{4\sup_{m\in\mathbb{Z}} \|\partial_m f\|_\infty}\right).$$

These two estimates together yield

$$\mathbb{P}\Big(f - \mathbb{E}(f) > r\Big) \leq \exp\left(-\min\left(\frac{r^2}{2C(\delta)\sum_{m\in\mathbb{Z}} \|\partial_m f\|_\infty^2}; \frac{\delta r}{4\sup_{m\in\mathbb{Z}} \|\partial_m f\|_\infty}\right)\right) .$$

Since the same inequalities holds for $-f$, we get

$$\mathbb{P}\Big(|f - \mathbb{E}(f)| > r\Big) \leq 2\exp\left(\frac{-r^2}{2C(\delta)\sum_{m\in\mathbb{Z}} \|\partial_m f\|_\infty^2} \vee \frac{-\delta r}{4\sup_{m\in\mathbb{Z}} \|\partial_m f\|_\infty}\right) \tag{2.4}$$

In the general case where f is unbounded, we take $f_n = (f \wedge n) \vee (-n)$. Notice that $\|\partial_m f_n\|_\infty \leq \|\partial_m f\|_\infty$. At first, we show that

$$\{\mathbb{E}f_n,\ n \geq 0\} \quad \text{is} \quad \text{bounded.}$$

To this purpose, suppose that $\{\mathbb{E}f_n,\ n \geq 0\}$ is not bounded, then there is a subsequence f_{n_k} such that

$$\mathbb{E}f_{n_k} \to \pm\infty \ .$$

The inequality (2.4) applied to f_{n_k} implies that

$$f_{n_k} \to \pm\infty \ \text{ in probability.}$$

This is in contradiction with the fact that $f_{n_k} \to f \ \mathbb{P} - a.s.$, proving the boundedness of $\{\mathbb{E}f_n,\ n \geq 0\}$.
Therefore $\{f_n,\ n \geq 0\}$ is uniformly integrable by (2.3). Consequently

$$\mathbb{E}f_n \to \mathbb{E}f \ .$$

Now by (2.3) and Fatou's Lemma,

$$\begin{aligned}
\mathbb{E}e^{\lambda(f-\mathbb{E}f)} &\leq \liminf_{n\to\infty} \mathbb{E}\exp\Big(\lambda(f_n - \mathbb{E}f_n)\Big) \\
&\leq \liminf_{n\to\infty} \exp\left(\frac{\lambda^2}{2} C(\delta) \sum_{k\in\mathbb{Z}} \|\partial_k f_n\|_\infty^2\right) \\
&\leq \exp\left(\frac{\lambda^2}{2} C(\delta) \sum_{k\in\mathbb{Z}} \|\partial_k f\|_\infty^2\right) .
\end{aligned}$$

We get so (2.2) like precedently.

3 Main Results.

3.1 Large deviations

We have the following:

Theorem 3.1 *Assume that there exists $\delta > 0$ such that $\mathbb{E}^\alpha(e^{\delta d(\xi_0,o)}) < \infty$. We suppose also that there exists a sequence of continuous mappings $\Phi^N : E^{[-N,N]} \to F$, such that*

$$\|\nabla(\Phi - \Phi^N)\|_{\infty,1} = \sum_{m\in\mathbb{Z}} \|\partial_m d_F(\Phi, \Phi^N)\|_\infty \longrightarrow 0, \tag{3.1}$$

then there is an application $\widetilde{\Phi} : \Omega \to F^{\mathbb{Z}}$ ($\tilde{\Phi} = (\tilde{\Phi}_n)$ with the same notations as those introduced for Φ) such that

$$\begin{cases} \widetilde{\Phi}_0(\tau^l\omega) = \widetilde{\Phi}_l(\omega),\ \forall l \in \mathbb{Z},\ \omega \in \Omega\ ; \\ \sup_{\{Q;H(Q)\leq L\}} \int d_F(\Phi_l^{(N)}, \widetilde{\Phi}_l) \wedge 1\ dQ \to 0 \text{ as } N \to \infty\ \forall L > 0,\ l \in \mathbb{Z}; \\ Q \to Q \circ \widetilde{\Phi}^{-1} \text{ is continuous from } [H < +\infty] \text{ to } M_1(F^{\mathbb{Z}}); \end{cases} \tag{3.2}$$

and $\mathbb{P}(R_n \in \cdot)$ satisfies LDP on $M_1(F^{\mathbb{Z}})$ equipped with the weak convergence topology with speed n and with the good rate function given by

$$J(\widehat{Q}) = \inf\left\{H(Q)|Q \in M_1^s(\Omega); H(Q) < +\infty; \widehat{Q} = Q \circ \widetilde{\Phi}^{-1}\right\}\ . \tag{3.3}$$

Remark :
Perhaps condition (3.1) seems to imply the continuity of Φ, but in fact it is far from to be the case. As an example, see the moving average process in section 4. And in section 4, we will see that condition (3.1) is quite sharp for the LDP.

Proof :
We define $R_n^N = \frac{1}{n}\sum_{k=1}^n \delta_{\bar{\Phi}^N(\tau^k\omega)} = O_n \circ (\bar{\Phi}^N)^{-1}$, with $\bar{\Phi}^N = (\Phi_n^N)$ as previous notation. Since Φ^N is continuous from $E^{[-N,N]}$ to F, the application $Q \to Q \circ (\bar{\Phi}^N)^{-1}$ is continuous from $M_1(\Omega)$ to $M_1(F^{\mathbb{Z}})$.

It is well known that the following metric on $M_1(F^{[-l,l]})$

$$d_{[-l,l]}\left(Q_{[-l,l]}, Q'_{[-l,l]}\right) := \sup\left\{\left|\int_{F^{[-l,l]}} f dQ_{[-l,l]} - \int_{F^{[-l,l]}} f dQ'_{[-l,l]}\right|;\ |f(x)| \leq 1,\right.$$
$$\left.|f(x) - f(y)| \leq d_{[-l,l]}(x, y) = \sum_{j=-l}^{l} d_F(x_j, y_j), \forall x, y \in F^{[-l,l]}\right\}$$

where $Q_{[-l,l]} = Q((x_{-l}, .., x_0, .., x_l) \in \cdot)$, is compatible with the weak convergence topology of $M_1(F^{[-l,l]})$. Now we define a metric on $M_1(F^{\mathbb{Z}})$ by

$$d^W(Q, Q') = \sum_{l=0}^{+\infty} \frac{1}{2^{l+1}} \frac{d_{[-l,l]}(Q_{[-l,l]}, Q'_{[-l,l]})}{1 + d_{[-l,l]}(Q_{[-l,l]}, Q'_{[-l,l]})}. \tag{3.4}$$

Note that this distance is compatible with the weak convergence topology on $M_1(F^{\mathbb{Z}})$.

The $d_{[-l,l]}$-distance between the marginal laws of R_n and R_n^N restricted to $M_1(F^{[-l,l]})$ is less than (see [Wu99, Lemma 4.1]),

$$\frac{1}{n}\sum_{k=1}^n d_{[-l,l]}\Big((\Phi(\omega(k + j + \cdot)))_{-l\leq j\leq l}, (\Phi^N(\omega(k + j + \cdot)))_{-l\leq j\leq l}\Big)$$

$$= \sum_{j=-l}^{l} \frac{1}{n} \sum_{k=1}^{n} d_F \left(\Phi(\omega(k+j+\cdot)), \Phi^N(\omega(k+j+\cdot)) \right) .$$

Substituting it into (3.4), we get that for each $l \geq 0$

$$d^W(R_n, R_n^N) \leq \frac{1}{2^l} + \sum_{j=-l}^{l} \frac{1}{n} \sum_{k=1}^{n} d_F \left(\Phi(\omega(k+j+\cdot)), \Phi^N(\omega(k+j+\cdot)) \right).$$

For the LDP of $\mathbb{P}(R_n \in \cdot)$, by lemma 2.2, it is enough to establish the negligeability between R_n and R_n^N, i.e.

$$\text{(3.5)} \qquad \lim_{N\to\infty} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P} \left(d^W \left(\frac{1}{n} \sum_{k=1}^{n} \delta_{\bar{\Phi}^{(N)}(\tau^k\omega)}, \frac{1}{n} \sum_{k=1}^{n} \delta_{\bar{\Phi}(\tau^k\omega)} \right) > \delta \right) = -\infty \ , \ \forall \delta > 0.$$

To this end, for each δ fixed, choose $l \geq 1$ so that $\frac{1}{2^l} < \delta$. Hence by the last inequality and the shift invariance of $\mathbb{P}$,

$$\begin{aligned} \mathbb{P}(d^W(R_n, R_n^N) > \delta) \quad &\leq \quad \sum_{j=-l}^{l} \mathbb{P} \left(\frac{1}{n} \sum_{k=1}^{n} d_F(\Phi(\omega(k+j+\cdot)), \Phi^N(\omega(k+j+\cdot))) > \frac{\delta - \frac{1}{2^l}}{2l+1} \right) \\ &= \quad (2l+1)\mathbb{P} \left(\frac{1}{n} \sum_{k=1}^{n} d_F \left(\Phi(\omega(k+\cdot)), \Phi^N(\omega(k+\cdot)) \right) > \frac{\delta - \frac{1}{2^l}}{2l+1} \right). \end{aligned}$$

Define

$$f = \sum_{k=1}^{n} d_F \left(\Phi(\omega(k+\cdot)), \Phi^N(\omega(k+\cdot)) \right).$$

We have

$$\text{(3.6)} \qquad \begin{aligned} \sup_m \|\partial_m f\|_\infty \quad &= \sup_m \sum_{k=1}^{n} \|\partial_{m-k} d_F(\Phi, \Phi^N)\|_\infty \\ &\leq \|\nabla(\Phi - \Phi^N)\|_{\infty,1} \end{aligned}$$

and

$$\text{(3.7)} \qquad \begin{aligned} \sum_{k\in\mathbb{Z}} \|\partial_m f\|_\infty^2 \quad &\leq \sum_{m\in\mathbb{Z}} \left(\sum_{k=1}^{n} \|\partial_{m-k} d_F(\Phi, \Phi^N)\|_\infty \right)^2 \\ &\leq \|\nabla(\Phi - \Phi^N)\|_{\infty,1} \sum_{m\in\mathbb{Z}} \sum_{k=1}^{n} \|\partial_{m-k} d_F(\Phi, \Phi^N)\|_\infty \\ &= n \ \|\nabla(\Phi - \Phi^N)\|_{\infty,1}^2. \end{aligned}$$

Hence by lemma 2.3, we have for any $r > 0$,

$$\begin{aligned} &\mathbb{P} \left(\sum_{k=1}^{n} \left[d_F \Big(\Phi(\xi_{k+\cdot}), \Phi^N(\xi_{k+\cdot}) \Big) - \mathbb{E} d_F \Big(\Phi(\xi_{k+\cdot}), \Phi^N(\xi_{k+\cdot}) \Big) \right] > rn \right) \\ &\qquad\qquad \leq \exp \left(- \left(\frac{nr^2}{2C(\delta)\|\nabla(\Phi - \Phi^N)\|_{\infty,1}^2} \wedge \frac{\delta nr}{4\|\nabla(\Phi - \Phi^N)\|_{\infty,1}} \right) \right). \end{aligned}$$

Since for sufficiently large N,

$$\begin{aligned} \mathbb{E} \left(d_F(\Phi(\xi_{k+\cdot}), \Phi^N(\xi_{k+\cdot})) \right) \quad &\leq \quad 2\mathbb{E} \left(\sum_{m\in\mathbb{Z}} \|\partial_m d_F(\Phi, \Phi^N)\|_\infty \cdot d(\xi_m, o) \right) \\ &\leq \quad 2\|\nabla(\Phi - \Phi^N)\|_{\infty,1} \mathbb{E} \left(d(\xi_0, o) \right) < r \ ; \end{aligned}$$

we obtain

$$\lim_{n\to\infty}\frac{1}{n}\log \mathbb{P}\left(\sum_{k=1}^{n} d_F\left(\Phi(\xi_{k+\cdot}),\Phi^N(\xi_{k+\cdot})\right) > 2rn\right) \le \frac{-r^2}{2C(\delta)\|\nabla(\Phi-\Phi^N)\|^2_{\infty,1}} \vee \frac{-\delta r}{4\|\nabla(\Phi-\Phi^N)\|_{\infty,1}} \ ;$$

where the right hand side, depending only on N, tends to $-\infty$ as N goes to infinity, because $\|\nabla(\Phi-\Phi^N)\|_{\infty,1}\to 0$. It ends the proof of (3.5).

The identification of the rate function needs more effort. By Lemma 2.2, the application $\psi^N(Q) = Q\circ(\bar{\Phi}^N)^{-1}$ will converge to some continuous application $\psi(Q)$ from $\{Q\in M_1^s(\Omega); H(Q)<\infty\}$ to $M_1(F^{\mathbb{Z}})$, but the limit application $\psi(Q)$ is not always of the form $Q\circ G^{-1}$ for some mapping $G:\Omega\to F^{\mathbb{Z}}$. To overcome this technical difficulty, we do appeal to the inequality (2.3) .

In fact, for all $N',N''\ge N$ and for

$$\lambda(N) = \left(\frac{1}{C(\delta)\sup\limits_{N',N''\ge 1}\|\nabla(\Phi^{N'}-\Phi^{N''})\|_{\infty,1}}\right)^{1/2}$$

by Varadhan's Laplace principle,

$$\sup\left\{\lambda(N)\int d_F(\Phi^{N'},\Phi^{N''})\wedge 1 \ dQ - H(Q) \mid Q\in M_1(\Omega), H(Q)<+\infty\right\}$$

$$= \lim_{n\to\infty}\frac{1}{n}\log \mathbb{E}\exp\left(\lambda(N)\sum_{k=1}^{n} d_F\left(\Phi^{N'}(\xi_{k+\cdot}),\Phi^{N''}(\xi_{k+\cdot})\right)\wedge 1\right)$$

which is less than $\frac{1}{2}$ by (2.3) and a similar calculation as in (3.6) and (3.7).
It follows that for any $L\in\mathbb{N}$

$$\sup_{N',N''\ge N}\ \sup_{\{Q;H(Q)\le L\}}\int d_F(\Phi^{N'},\Phi^{N''})\wedge 1 \ dQ \to 0, \quad (as\ N\to\infty).$$

Put

$$q(N',N'') := \sum_{L=0}^{\infty}\frac{1}{2^L}\frac{\sup_{\{Q;H(Q)\le L\}}\int d_F(\Phi^{N'},\Phi^{N''})\wedge 1 \ dQ}{1+\sup_{\{Q;H(Q)\le L\}}\int d_F(\Phi^{N'},\Phi^{N''})\wedge 1 \ dQ} \ .$$

We get by the dominated convergence $\lim_{N',N''\to\infty} q(N',N'')=0$.
Therefore, we can find a subsequence (N_k) such that $q(N_k,N_{k+1})<\frac{1}{2^k}$.

Now consider $A=\left\{\omega\in\Omega |\ \lim_{k\to\infty}\Phi^{(N_k)}(\omega)=\widetilde{\Phi}_0(\omega) \ \text{ exists}\right\}$,

$$\Omega^0=\bigcap_{l\in\mathbb{Z}}\tau^{-l}A,\ \widetilde{\Phi}_l(\omega)=\widetilde{\Phi}_0(\tau^l\omega),$$

$$\widetilde{\Phi}(\omega)=(\widetilde{\Phi}_l(\omega))_{l\in\mathbb{Z}},\forall\omega\in\Omega^0,\ \text{ and } \widetilde{\Phi}(\omega)=(..,0,..,0,..),\ if\ \omega\notin\Omega^0.$$

By Borel-Cantelli and our choice of (N_k), for any Q with $H(Q)<\infty$, $Q(A)=1$. Then $Q(\tau^{-l}A)=1$, $\forall\ l\in\mathbb{Z}$, by the shift invariance of Q. Consequently $Q(\Omega^o)=1$.

Letting $N'=N,\ N''=N_k$ and $k\to\infty$, we get for any $l\in\mathbb{Z}$, as $Q\in M_1^s(\Omega)$,

$$\sup_{Q\in[H\le L]}\int d_F(\Phi_l^{(N)},\widetilde{\Phi}_l)\wedge 1\ dQ = \sup_{Q\in[H\le L]}\int d_F(\Phi^{(N)},\widetilde{\Phi})\wedge 1\ dQ \to 0 \quad as \quad N\to\infty\ .$$

Consequently for any Q with $H(Q) < +\infty$, $Q\circ(\Phi^N)^{-1}$ converge to $Q\circ\widetilde{\Phi}^{-1}$. By lemma 2.1, $Q\to Q\circ\widetilde{\Phi}^{-1}$ must be continuous from $[H<+\infty]$ to $M_1(F^{\mathbb{Z}})$, and the rate function governing the LDP of R_n must be given by (3.3). So we have finished the proof of the theorem.

By Theorem 3.1 and the contraction principle, for any bounded continuous function G on $F^{\mathbb{Z}}$ with values in a separable Banach space, $\mathbb{P}(\int GdR_n\in\cdot)$ satisfies the Large Deviation Principle on B with speed n and with rate function

$$J_G(z)=\inf\left\{H(Q)|Q\in M_1^s(\Omega),\ H(Q)<\infty,\ \int G(\tilde{\Phi})dQ=z\right\}.$$

With the same proof as that of Theorem 3.1, we have the following

Corollary 3.2 *Let $F=(B,\|\cdot\|)$ a separable Banach space. Assume that for all $\lambda>0$ we have $\mathbb{E}^\alpha(e^{\lambda d(\xi_0,o)})<\infty$. We suppose also that there exists a sequence of continuous mappings $\Phi^N: E^{[-N,N]}\to B$, such that $\sup_N \mathbb{E}^{\mathbb{P}}(e^{\lambda\|\Phi^N\|})<\infty$ for all $\lambda>0$, and*

$$\|\nabla(\Phi-\Phi^N)\|_{\infty,1}=\sum_{m\in\mathbb{Z}}\|\partial_m(\|\Phi-\Phi^N\|)\|_\infty\longrightarrow 0, \tag{3.8}$$

then $\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^n\Phi(\xi_{k+\cdot})\in\cdot\right)$ satisfies a large deviation principle on B with the rate function given by

$$J_\Phi(z)=\inf\left\{H(Q)|\tilde{\Phi}_0\in L^1(dQ)\ \text{and}\ \int\tilde{\Phi}_0\ dQ=z\right\}.$$

3.2 Moderate deviations

Let $\Phi: E^{\mathbb{Z}}\longrightarrow(B,\|\cdot\|)$ be a measurable function, where $(B,\|\cdot\|)$ is a separable Banach space. Let $\{b_n\}_{n\ge0}$ be a sequence of positive numbers such that

$$\frac{b_n}{\sqrt{n}}\to+\infty,\qquad \frac{b_n}{n}\to 0.$$

In the following theorem, we are looking for the large deviation principle on $(B,\|\cdot\|)$ of

$$M_n(\Phi):=\frac{1}{b_n}\sum_{k=1}^n\Big(\Phi(\xi_{k+\cdot})-\mathbb{E}\Phi(\xi_{k+\cdot})\Big)\ .$$

Theorem 3.3 *Assume that for some $\delta>0$ we have $\mathbb{E}^\alpha(e^{\delta d(\xi_0,o)})<\infty$. We suppose also that there exists a sequence of mappings $\Phi^N: E^{[-N,N]}\to B$, such that $\mathbb{E}^{\mathbb{P}}(e^{\beta\|\Phi^N\|})<\infty$ for some $\beta>0$ (depending on N), and*

$$\|\nabla(\Phi-\Phi^N)\|_{\infty,1}=\sum_{m\in\mathbb{Z}}\|\partial_m(\|\Phi-\Phi^N\|)\|_\infty\longrightarrow 0, \tag{3.9}$$

$$M_n(\Phi^N)\to 0 \textit{ in probability} \text{ and } \lim_{N\to\infty}\limsup_{n\to\infty}\mathbb{E}\|M_n(\Phi-\Phi^N)\|=0. \tag{3.10}$$

Then

$$\sigma^2(y) := \lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left(\sum_{k=1}^{n}\langle y, \Phi(\xi_{k+\cdot}) - \mathbb{E}\Phi(\xi_{k+\cdot})\rangle\right)^2 \in \mathbb{R}, \tag{3.11}$$

exists for every $y \in B'$, and $\mathbb{P}\left(\frac{1}{b_n}\sum_{k=1}^n (\Phi(\xi_{k+\cdot}) - \mathbb{E}\Phi(\xi_{k+\cdot})) \in \cdot\right)$ satisfies the large deviation principle on B, with speed $\frac{b_n^2}{n}$ and with the good convex rate function $I(x)$, given by

$$I(x) = \sup_{y\in B'}\left\{\langle x, y\rangle - \frac{1}{2}\sigma^2(y)\right\} \tag{3.12}$$

where B' is the dual Banach space of B.

Remarks :
(i) In the literature, the large deviation principles in Theorem 3.3 are often called Moderate Deviation Principles (in abrigde MDP, see e.g [DZ]).
(ii) If $B = \mathbb{R}^d$, $M_n(\Phi^N) \to 0$ in $L^2(\mathbb{P})$ (a obvious fact) and by (3.17) in the proof below, $\lim_{N\to\infty}\sup_n \mathbb{E}|M_n(\Phi - \Phi^N)|^2 = 0$. Then the second condition of (3.10) is automatically satisfied, too. In other words, (3.10) is superfluous in the finite dimensional setting.

Proof :
We separate its proof into four steps.

1) It is known (see [Che97]) that for each N fixed,

$$\begin{aligned}\lim_{n\to\infty}\frac{n}{b_n^2}\log\mathbb{E}\left(e^{\frac{b_n^2}{n}\langle y, M_n(\Phi^N)\rangle}\right) &= \frac{1}{2}\lim_{n\to\infty}\frac{1}{n}\mathbb{E}\left(\langle y, b_n M_n(\Phi^N)\rangle\right)^2 \\ &:= \frac{1}{2}\sigma_N^2(y) \in \mathbb{R},\ \forall y \in B',\end{aligned} \tag{3.13}$$

and that $M_n(\Phi^N)$ satisfies the MDP with the good rate function

$$I^N(x) = \sup_{y\in B'}\left\{\langle x, y\rangle - \frac{1}{2}\sigma_N^2(y)\right\}\ .$$

2) We denote

$$f_n^N := b_n M_n(\Phi - \Phi^N).$$

By Lemma 2.3, and by some similar calculation as in (3.6) and (3.7), we have for $rb_n > \mathbb{E}\|f_n^N\|$,

$$\mathbb{P}\left(\|f_n^N\| > rb_n\right) \le \exp\left(\frac{-\left(rb_n - \mathbb{E}\|f_n^N\|\right)^2}{2C(\delta)n\|\nabla(\Phi-\Phi^N)\|_{\infty,1}} \vee \frac{-\delta\left(rb_n - \mathbb{E}\|f_n^N\|\right)}{4\|\nabla(\Phi-\Phi^N)\|_{\infty,1}}\right).$$

By (3.9) and (3.10), we deduce that

$$\lim_{N\to\infty}\limsup_{n\to\infty}\frac{n}{b_n^2}\log\mathbb{P}\left(\|f_n^N\| > rb_n\right) = -\infty\ ,\ \forall r > 0\ .$$

The MDP is then established by lemma 2.1, with the good rate function given by

$$\begin{aligned}\widetilde{I}(x) &= \sup_{\delta>0}\liminf_{N\to\infty}\inf_{B(x,\delta)} I^N \\ &= \sup_{\delta>0}\limsup_{N\to\infty}\inf_{B(x,\delta)} I^N\ .\end{aligned} \tag{3.14}$$

3) We have now to prove the identification of the rate function. Firstly, we show that $\forall y \in B'$

$$\sigma^2(y) := \lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left(\sum_{k=1}^n \langle y, \Phi(\xi_{k+\cdot}) - \mathbb{E}\Phi(\xi_{k+\cdot})\rangle\right)^2 \text{ exists and}$$

$$\sigma^2(y) = \lim_{N\to+\infty} \sigma_N^2(y) \in \mathbb{R}. \tag{3.15}$$

By (2.3) again, we have as in (3.6) and (3.7) that for all $|\lambda|$ small enough,

$$\begin{aligned}\mathbb{E}\left(e^{\lambda\langle y, f_n^N\rangle}\right) &\le \exp\left(\frac{\lambda^2}{2}\|y\|_{B'}^2 C(\delta) n \|\nabla(\Phi - \Phi^N)\|_{\infty,1}^2\right) \\ &= 1 + \frac{\lambda^2}{2}\|y\|_{B'}^2 C(\delta)\|\nabla(\Phi - \Phi^N)\|_{\infty,1}^2 + o\left(\frac{\lambda^2}{2}\right) .\end{aligned} \tag{3.16}$$

But

$$\mathbb{E}\left(e^{\lambda\langle y, f_n^N\rangle}\right) = 1 + \frac{\lambda^2}{2}\mathbb{E}\left(\langle y, f_n^N\rangle\right)^2 + o\left(\frac{\lambda^2}{2}\right),$$

we deduce that

$$\mathbb{E}\left(\langle y, f_n^N\rangle\right)^2 \le nC(\delta)\|y\|_{B'}^2\|\nabla(\Phi - \Phi^N)\|_{\infty,1}^2 .$$

So we have

$$\sup_n \frac{1}{n}\mathbb{E}\left(\langle y, f_n^N\rangle\right)^2 = \sup_n \frac{b_n^2}{n}\mathbb{E}\left(\langle y, M_n(\Phi - \Phi^N)\rangle\right)^2 \to 0 \quad as\ N \to +\infty . \tag{3.17}$$

Whence the limit $\sigma^2(y)$ in (3.15) exists and $\sigma_N^2(y) \longrightarrow \sigma^2(y),\ \forall y \in B'$.

Now we claim that

$$\lim_{n\to\infty} \frac{n}{b_n^2}\log \mathbb{E}\exp\left(\frac{b_n^2}{n}\langle y, M_n(\Phi)\rangle\right) = \frac{1}{2}\sigma^2(y), \quad \forall y \in B' . \tag{3.18}$$

For fixed $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, by the Hölder inequality we have that

$$\begin{aligned}\log \mathbb{E}\exp\left(\frac{b_n^2}{n}\langle y, M_n(\Phi)\rangle\right) &\le \frac{1}{q}\log \mathbb{E}\exp\left(q\frac{b_n^2}{n}\langle y, M_n(\Phi - \Phi^N)\rangle\right) \\ &+ \frac{1}{p}\log \mathbb{E}\exp\left(p\frac{b_n^2}{n}\langle y, M_n(\Phi^N)\rangle\right)\end{aligned}$$

for every $y \in B'$. From (3.13) and (3.16) it follows that

$$\limsup_{n\to\infty} \frac{n}{b_n^2}\log \mathbb{E}\left(e^{\frac{b_n^2}{n}\langle y, M_n(\Phi)\rangle}\right) \le \frac{1}{2p}\sigma_N^2(py) + \frac{q}{2}\|\nabla(\Phi - \Phi^N)\|_{\infty,1}^2 C(\delta)\|y\|_{B'}^2.$$

Letting $N \to \infty$ and using (3.15), we get

$$\limsup_{n\to\infty} \frac{n}{b_n^2}\log \mathbb{E}\left(e^{\frac{b_n^2}{n}\langle y, M_n(\Phi)\rangle}\right) \le \frac{1}{2p}\sigma^2(py) = \frac{p}{2}\sigma^2(y). \tag{3.19}$$

Similarly, by the Hölder inequality, we have

$$\begin{aligned}\log \mathbb{E}\exp\left(\frac{b_n^2}{pn}\langle y, M_n(\Phi^N)\rangle\right) &\le \frac{1}{q}\log \mathbb{E}\exp\left(\frac{qb_n^2}{pn}\langle y, M_n(\Phi^N-\Phi)\rangle\right)\\ &+ \frac{1}{p}\log \mathbb{E}\exp\left(\frac{b_n^2}{n}\langle y, M_n(\Phi)\rangle\right)\end{aligned}$$

for every $y \in B'$. From (3.13) and (3.16) it follows that

$$\frac{1}{2}\sigma_N^2\left(\frac{y}{p}\right) \le \liminf_{n\to\infty} \frac{n}{pb_n^2}\log \mathbb{E}\left(e^{\frac{b_n^2}{n}\langle y, M_n(\Phi)\rangle}\right) + \frac{q}{2p^2}\|\nabla(\Phi-\Phi^N)\|_{\infty,1}^2 C(\delta)\|y\|_{B'}^2.$$

Letting $N \to \infty$ and using (3.15), we obtain

$$\frac{1}{2p}\sigma^2(y) \le \liminf_{n\to\infty} \frac{n}{b_n^2}\log \mathbb{E}\left(e^{\frac{b_n^2}{n}\langle y, M_n(\Phi)\rangle}\right). \tag{3.20}$$

Letting $p \to 1$ in (3.19) and (3.20) yields (3.18).

So by (3.18) and the Laplace principle [DS89, Th. 2.1.10, p.43], we have

$$\frac{1}{2}\sigma^2(y) = \lim_{n\to\infty} \frac{n}{b_n^2}\log \mathbb{E}\left(e^{\frac{b_n^2}{n}\langle y, M_n(\Phi)\rangle}\right) = \sup_{x\in B}\left\{\langle x, y\rangle - \widetilde{I}(x)\right\}. \tag{3.21}$$

4) We have now to show that $\widetilde{I}(x)$ defined in (3.14) is convex.

$$\widetilde{I}\left(\frac{1}{2}(x_1+x_2)\right) = \sup_{\delta>0}\limsup_{N\to\infty}\inf_{B(\frac{1}{2}(x_1+x_2),\delta)} I^N$$

$$\begin{aligned}\inf_{B(\frac{1}{2}(x_1+x_2),\delta)} I^N &\le \inf_{y_1\in B(x_1,\delta), y_2\in B(x_2,\delta)} I^N\left(\frac{1}{2}(y_1+y_2)\right)\\ &\le \frac{1}{2}\inf_{y_1\in B(x_1,\delta), y_2\in B(x_2,\delta)}\left(I^N(y_1)+I^N(y_2)\right)\\ &= \frac{1}{2}\left(\inf_{B(x_1,\delta)} I^N + \inf_{B(x_2,\delta)} I^N\right)\end{aligned}$$

So

$$\limsup_{N\to\infty}\inf_{B(\frac{1}{2}(x_1+x_2),\delta)} I^N \le \frac{1}{2}\left(\limsup_{N\to\infty}\inf_{B(x_1,\delta)} I^N + \limsup_{N\to\infty}\inf_{B(x_2,\delta)} I^N\right)$$

Letting $\delta \downarrow 0$, we get

$$\widetilde{I}\left(\frac{1}{2}(x_1+x_2)\right) \le \frac{1}{2}\left(\widetilde{I}(x_1)+\widetilde{I}(x_2)\right).$$

Since $\widetilde{I}$ is inf-compact and convex, by Fenchel's theorem and (3.21), we get for all $x \in \mathbb{R}$

$$\widetilde{I}(x) = \sup_{y\in B'}\left\{\langle x, y\rangle - \frac{1}{2}\sigma^2(y)\right\},$$

which is exactly the rate function given by (3.12).

4 The moving average case

We now consider an example where Φ is linear, i.e. the moving average process. Consider

$$X_n = \Phi(\xi_{n+\cdot}) = \sum_{j\in\mathbb{Z}} a_j \xi_{n+j}$$

where $(a_n)_{n\in\mathbb{Z}}$ is some doubly infinite real sequence, and (ξ_k) are i.i.d. centered random variables with values in a separable Banach space $(B, \|\cdot\|)$. $X = (X_n)$ is called the moving average process.

When $B = \mathbb{R}^d$, Burton and Dehling [BD90] obtained the large deviation of the empirical means

$$\frac{S_n}{n} = \frac{1}{n}\sum_{k=1}^{n} X_k$$

under the conditions that $\sum_{n\in\mathbb{Z}} |a_n| < \infty$ and $\mathbb{E}\exp(\delta|\xi_0|) < +\infty$ for all positive δ. Jiang, Rao and Wang [JRW95] showed that the lower bound of S_n is valid without the condition $\mathbb{E}\exp(\delta|\xi_0|) < +\infty$ and that the upper bound (with perhaps a different rate function) is true if $\mathbb{E}\exp(\delta|\xi_0|) < +\infty$ for some positive δ. Moreover their result still holds in a separable Banach space B, under the assumption $\mathbb{E}\exp(q_K(B)) < \infty$ for some balanced convex compact subset K of B (q_K denotes the Minkowski functional).

We will show the corresponding level 3 LDP for X_n which seems to be unknown. Denote $\bar{\Phi}(\omega) = (\Phi_n(\omega))$, where $\Phi_n(\omega) = \sum_{j\in\mathbb{Z}} a_j\omega(n+j)$, and (using notations of previous sections)

$$R_n^{\bar{\Phi}} := \frac{1}{n}\sum_{k=1}^{n} \delta_{X_{k+\cdot}} = O_n \circ \bar{\Phi}^{-1}.$$

We have the following

Theorem 4.1 *Assume* $\sum_i |a_i| < \infty$ *and* $\mathbb{E}\exp(\delta\|\xi_0\|) < \infty$ *for some positive* δ*, then* $\mathbb{P}\left(R_n^{\bar{\Phi}} \in \cdot\right)$ *satisfies a LDP on* $M_1(B^{\mathbb{Z}})$ *equipped with the weak convergence topology, with speed* n *and with the rate function given by*

$$J(\hat{Q}) = \inf\left\{H(Q) | Q \in M_1^s(\Omega),\ H(Q) < \infty,\ \hat{Q} = Q\circ\bar{\Phi}^{-1}\right\}. \tag{4.1}$$

Proof :
We have to find some approximation functions which satisfy conditions of Theorem 3.1. So define

$$\Phi_n^N(\omega) := \sum_{j=-N}^{N} a_j\omega(n+j) := \sum_{j\in\mathbb{Z}} a_j^N \omega(n+j)$$

where

$$a_j^N = \begin{cases} a_j & if\ |j| \le N \\ 0 & if\ |j| > N \end{cases}$$

By simple calculations,

$$\begin{aligned}\|\nabla(\Phi - \Phi^N)\|_{\infty,1} &= \sum_{k\in\mathbb{Z}} |a_{m-k} - a_{m-k}^N| \\ &= \sum_{|j|\ge N} |a_j| \\ &\longrightarrow 0\ as\ N \to +\infty\ .\end{aligned}$$

Thus the LDP of $\mathbb{P}(R_n^{\bar{\Phi}} \in \cdot)$ follows from Theorem 3.1.
We still have to establish the identification of the rate function (4.1). To this end, we follow Remark 1.5 of Wu [Wu99]. If $H(Q) < \infty$, the integrability condition $\mathbb{E}\exp(\delta|\xi_0|) < \infty$ implies that $\mathbb{E}^Q|\xi_0| < \infty$, and then

$$\mathbb{E}^Q \sum_{j\in\mathbb{Z}} |a_j\xi_{n+j}| \leq \sum_{j\in\mathbb{Z}} |a_j|\mathbb{E}^Q|\xi_0| < \infty,$$

and we consequently have (using Theorem 3.1 and its notations) that $\tilde{\Phi} = \Phi$, $Q - a.s.$ and (4.1) follows from the expression (3.3).

Remarks :
(i) Note that corollary 3.2 applied to the linear case gives exactly the results of corollary 3 of [JRW95].
(ii) Condition (3.1) in the nonlinear case is quite sharp. It is equivalent to the absolute summability $\sum_{j\in\mathbb{Z}} |a_j| < +\infty$ in the linear case. Notice that last condition is rather sharp even for the central limit theorem (CLT). In fact, when $a_n \geq 0$, the best condition for the CLT of $\sum_{k=1}^n X_k$ in [HH80], becomes exactly $\sum_{n\in\mathbb{Z}} a_n < +\infty$. In this non-negative case, the summability $\sum_{n\in\mathbb{Z}} a_n < +\infty$ is also sharp for the level 1 and level 3 LDP of (X_n), see Bryc and Dembo [BD95].
(iii) The results of section 3.2 on Moderate deviations Th. 3.3, applied to the moving average case, is Th. 2.1 of [JRW92] in Banach space.

References

[BD90] R.M. Burton and H. Dehling. Large deviations for some weakly dependent random processes. *Statistics and Probability Letters*, 9:397–401, 1990.

[BD95] W. Bryc and A. Dembo. On large deviations of empirical measures for stationnary gaussian processes. *Stochastic Processes and Their Applications*, 58:23–34, 1995.

[Che97] X. Chen. Moderate deviations for m-dependent random variables with Banach space values. *Statistics and Probability Letters*, 35:123–134, 1997.

[DG98] H. Djellout and A. Guillin. Large deviations and moderate deviations for moving average processes. *Preprint*, 1998.

[DS89] J.D. Deuschel and D.W. Stroock. *Large deviations.* Academic Press, Boston, 1989.

[DV85] M.D. Donsker and S.R.S. Varadhan. Large deviations for stationary gaussian processes. *Communications in Mathematical Physics*, 97:187–210, 1985.

[DZ93] A. Dembo and O. Zeitouni. *Large deviations techniques and their applications.* Jones and Bartlett, Boston, MA, 1993.

[HH80] P. Hall and C.C. Heyde. *Martingale limit theory and its application.* Academic Press, New York, 1980.

[JRW92] T. Jiang, M.B. Rao, and X. Wang. Moderate deviations for some weakly dependent random processes. *Statistics and Probability Letters*, 15:71–76, 1992.

[JRW95] T. Jiang, M.B. Rao, and X. Wang. Large deviations for moving average processes. *Stochastic Processes and Their Applications*, 59:309–320, 1995.

[Wu99] L. Wu. On large deviations for moving average processe. *Preprint*, 1999.