Large Deviations for Past-Dependent Recursions

F. Klebaner *,
Department of Statistic, University of Melbourne,
Parkville, Victoria, Australia 3052
R. Liptser,
Electrical Engineering Systems,
Tel Aviv University, 69978 - Ramat Aviv, Tel Aviv, Israel

May 31, 1999

Abstract

The Large Deviation Principle is established for stochastic models defined by past-dependent non linear recursions with small noise. In the Markov case we use the result to obtain an explicit expression for the asymptotics of exit time.

Key words: Large Deviations, Contraction Principle, Exit Time.

1 Introduction.

The simplest example of a stochastic model defined by past-dependent recursion with small noise is a linear model

$$X_k^\varepsilon = \sum_{i=1}^m a_i X_{k-1}^\varepsilon + \varepsilon\xi_k \tag{1.1}$$

subject to fixed $X_i^\varepsilon = x_i, i = 0, 1, ..., m-1$, where ε is small parameter and $(\xi_k)_{k\ge m}$ is an i.i.d. sequence of random variables. In the present paper we consider a general non linear model:

$$X_k^\varepsilon = f(X_{k-1}^\varepsilon, ..., X_{k-m}^\varepsilon, \varepsilon\xi_k), \tag{1.2}$$

*Research supported by the Australian Research Council

where $f(z_1, ..., z_m, y)$ is continuous function. Note that (ref1.2) includes (ref1.1) as a special case. For $m = 1$ model (1.2) defines a discrete time Markov process. When $\varepsilon \to 0$ random variables X_k^ε converge to deterministic ones, say, X_k and $X_k, k \geq 1$ are determined by the recursion

$$X_k = f(X_{k-1}, ..., X_{k-m}, 0), \tag{1.3}$$

subject to the same initial condition. Furthermore $(X_k^\varepsilon)_{k\geq m}$ converges to $(X_k)_{k\geq m}$ in the metric $\rho(x, y) = \sum_{j\geq m} 2^{-j} \frac{|x_j - y_j|}{1+|x_j - y_j|}$. This fact provides the motivation to consider the large deviation principle (LDP) for family $(X_k^\varepsilon)_{k\geq m}$ in the metric space (R^∞, ρ). For Markov case ($m = 1$) the LDP was considered in [5], [7] and [8]. The choice of the metric space (R^∞, ρ) is a natural one for obtaining the LDP for the family $(X_k^\varepsilon)_{k\geq m}$. Recursion (1.2) defines continuous mapping $(\varepsilon\xi_k)_{k\geq m} \to (X_k^\varepsilon)_{k\geq m}$ in the metric ρ. This implies that the LDP for $(X_k^\varepsilon)_{k\geq m}$ follows from the LDP for $(\varepsilon\xi_k)_{k\geq m}$ by the continuous mapping method of Freidlin [3] and contraction principle of Varadhan [11]. Since $(\xi_k)_{k\geq m}$ is an i.i.d. sequence, the LDP for $(\varepsilon\xi_k)_{k\geq m}$ holds if it holds for the family $\varepsilon\xi$, where ξ is a copy of ξ_m. It must be mentioned that not only the rate function but also a norming factor $q(\varepsilon)$ depend on the distribution of ξ.

In Section 4, sufficient conditions for the LDP for the family $\varepsilon\xi$ are given. Section 3 contains examples for which rate functions can be explicitly calculated, and in the Markov case asymptotics for the probability of $\{\max_{1\leq k\leq M} |X_k^\varepsilon| \geq 1\}$ is found (Theorem 3.1). Main results are formulated in Section 2. One of them gives the asymptotics of the exit time from the interval $[-1, 1]$ for the Markov family $(X_k^\varepsilon)_{k\geq 1}$.

2 Main results.

Following Varadhan [11], family $(X^\varepsilon)_{k\geq m}$ is said to satisfy the LDP in the metric space (R^∞, ρ) with norming factor $q(\varepsilon)$ (a function decreasing to zero as $\varepsilon \downarrow 0$), if

(0) there exists a function $J = J(\overline{u}), \overline{u} = (u_1, u_2, ...) \in R^\infty$ which takes values in $[0, \infty]$ such that for every $\alpha \geq 0$ the set $\Phi(\alpha) = \{\overline{u} \in R^\infty : J(\overline{u}) \leq \alpha\}$ is compact in (R^∞, ρ);

(1) For every closed set $F \in (R^\infty, \rho)$

$$\overline{\lim}_{\varepsilon\to 0} q(\varepsilon) \log P((X_k^\varepsilon)_{k\geq m} \in F) \leq -\inf_{\overline{u}\in F} J(\overline{u});$$

(2) For every open set $G \in (R^\infty, \rho)$

$$\underline{\lim}_{\varepsilon\to 0} q(\varepsilon) \log P((X_k^\varepsilon)_{k\geq m} \in G) \geq -\inf_{\overline{u}\in G} J(\overline{u}).$$

The function $J = J(\overline{u})$, satisfying (0), is called rate function.

As was mentioned in the introduction, the LDP for $(X_k^\varepsilon)_{k\geq m}$ is implied by the LDP for family $\varepsilon\xi$ (ξ is copy of ξ_m). Therefore sufficient conditions for the required LDP are formulated in term of the distribution of ξ. Here and in the sequel we assume the following conditions.

$$E\xi = 0, \tag{2.1}$$

and Cramer's condition is satisfied:

$$Ee^{t\xi} < \infty,\ t \in R, \tag{2.2}$$

and the cumulant function

$$H(t) = \log Ee^{t\xi} \tag{2.3}$$

is twice continuously differentiable with

$$H'(0) = 0 \quad \text{and} \quad H''(t) \geq 0. \tag{2.4}$$

For $H(t)$ define Fenchel-Legendre's transform:

$$L(v) = \sup_{t\in R}[tv - H(t)]. \tag{2.5}$$

Now we define a norming factor $q(\varepsilon)$. Assume there exist a function $q(\varepsilon)$, decreasing to 0 as $\varepsilon \downarrow 0$, and a non negative function $I(v), v \in R$ such that for every $v \in R$

$$\lim_{\varepsilon\to 0} q(\varepsilon)L(\frac{v}{\varepsilon}) = I(v). \tag{2.6}$$

Additional assumptions on the function $I(v)$ are:

$$I(0) = 0 \quad \text{and} \quad \lim_{|v|\to\infty} I(v) = \infty; \tag{2.7}$$

if for some v_0 $I(v_0) = \infty$, then for $v_0 > 0$ ($v_0 < 0$) $I(v)$ is left (right) continuous function); if $I(v) < \infty$, then the value $t_v^\varepsilon = \operatorname{argmax}(t\frac{v}{\varepsilon} - H(t))$ is finite and such that

$$\overline{\lim}_{\varepsilon\to 0}\frac{q(\varepsilon)}{\varepsilon}|t_v^\varepsilon| < \infty \quad \text{and} \quad \lim_{\varepsilon\to 0}\varepsilon^2 H''(t_v^\varepsilon) = 0. \tag{2.8}$$

The main results are given in the theorem below.

Theorem 2.1 *Let (2.1)-(2.3) and (2.5)-(2.8) be satisfied. Then 1) the family* $\varepsilon\xi$

obeys the LDP with rate function $I(v)$ given in (2.6);

2) the family $(\varepsilon\xi_k)_{k\geq 1}$ obeys the LDP in the metric space (R^∞, ρ) with the rate function ($\overline{v} = (v_1, v_2, ...) \in R^\infty$)

$$I_\infty(\overline{v}) = \sum_{k=1}^{\infty} I(v_k); \tag{2.9}$$

3) the family $(X_k^\varepsilon)_{k\geq m}$ obeys the LDP in the metric space (R^∞, ρ) with the rate function ($\overline{u} = (u_m, u_{m+1}, ...) \in R^\infty$)

$$J_\infty(\overline{u}) = \begin{cases} \sum_{k=m}^{\infty} \inf_{v_k : u_k = f(u_{k-1},...,u_{k-m},v_k)} I(v_k), & u_i = x_i, i = 0, ..., m-1 \\ \infty, & \textit{otherwise}, \end{cases} \tag{2.10}$$

where $\inf(\emptyset) = \infty$.

Remark 1. Verification of conditions (2.6) and (2.8) can be simplified if random variable ξ can be represented as a sum of two independent random variables $\xi = \xi^i + \xi^{ii}$ such that for each of them the Cramer's condition (2.2) is satisfied. $H^i(t)$, $H^{ii}(t)$ and $L^i(v)$, $L^{ii}(v)$ are their cumulant functions and Fenchel-Legendre transforms respectively. If for ξ^i all the formulated above assumptions are satisfied, and in particular the rate function $I^i(v)$ and the norming factor $q^i(\varepsilon)$ are defined, then under the condition

$$\lim_{\varepsilon\to 0} q^i(\varepsilon) L^{ii}(\frac{v}{\varepsilon}) = -\infty, \ v \neq 0, \tag{2.11}$$

the statements of the theorem remain true with $I(v) \equiv I^i(v)$. Condition (2.11) always holds for random variables with finite support.

Remark 2. The LDP for the model (1.2) can be verified not only in the case of continuous function $f(z_1, ..., z_m, y)$, but in a more general setting. If for every set $z_1, ..., z_m$ a cumulant function

$$H_\varepsilon(t, z_1, ..., z_m) = \log E \exp\Big(tf(z_1, ..., z_m, \varepsilon\xi_1)\Big),$$

is well defined and is continuous in $z_1, ..., z_m$ for every fixed t and ε, moreover there exists a norming factor $q(\varepsilon)$

$$\lim_{\varepsilon\to 0} q(\varepsilon) \sup_{t\in R} [tu - H_\varepsilon(z_1, ..., z_m)] = I(u, z_1, ..., z_m),$$

then under some technical conditions on $I(u, z_1, ..., z_m)$ the LDP for the family $(X_k^\varepsilon)_{k\geq m}$ holds and rate function is given by formula

$$J_\infty(u_1, u_2, ...) = \sum_{k\geq m} I(u_k, u_{k-1}, ..., u_{k-m}).$$

We do not persue this direction here since in this setting the conditions for the existence of the norming factor and the rate function are not as natural in Theorem 2.1. Thus we limit ourselves by considering model (1.2) with continuous function $f(z_1, ..., z_m, y)$ only.

The asymptotics for exit time is the next result. Let

$$X_k^\varepsilon = af(X_{k-1}^\varepsilon) + \varepsilon\xi_k \tag{2.12}$$

subject to $X_0^\varepsilon = x_0 \in [-1, 1]$, where $f(z)$ is continuous function with $f(z) = z$ for $|z| \leq 1$, and where ξ_1 is Gaussian random variable with parameters $(0, 1)$ (or $\xi_1 = \xi_1^i + \xi_1^{ii}$ with independent random variables ξ_1^i and ξ_1^{ii}, where ξ_1^i is $(0, 1)$-Gaussian random variable, all conditions from Remark 1 are satisfied). Let τ^ε be the exit time from the interval $[-1, 1]$:

$$\tau^\varepsilon = \min\{k \geq 1 : |X_k^\varepsilon| \geq 1\}. \tag{2.13}$$

Theorem 2.2 *1) If $|a| < 1$, then for any $x_0 \in [-1, 1]$ the exit time obeys the following asymptotics:*

$$\lim_{\varepsilon\to 0} \varepsilon^2 \log E\tau^\varepsilon = \frac{1}{2}(1 - a^2).$$

2) If $|a| \geq 1$ and $x_0 = 0$, then

$$\lim_{\varepsilon\to 0} \varepsilon^2 \log P(\tau^\varepsilon \leq M) = -\frac{1}{2\sum_{j=0}^{M-1} a^{2j}}, \quad M = 0, 1, ...$$

A proof of the first statement of the theorem could be obtained by applying a version of Kifer's result in [6], which is a discrete time version of the Freidlin-Wentzell result on the asymptotics of the exit time for diffusion processes [4] with a small diffusion coefficient. However, formally we can not do this since in [6] X_k^ε take values in a compact while in our case $X_k^\varepsilon \in R$. Therefore we give a selfcontained proof, repeating some details from [6].

3 Examples and Applications.

1. We give two examples of random variable ξ for which

$$\begin{aligned} q(\varepsilon) &= \varepsilon^2 \quad \text{and} \quad I(v) = \frac{1}{2}v^2 \\ q(\varepsilon) &= \frac{\varepsilon}{|\log \varepsilon|} \quad \text{and} \quad I(v) = |v|. \end{aligned}$$

In the first case ξ is $(0,1)$-Gaussian random variable with cumulant function $H(t) = \frac{t^2}{2}$. It is easy to verify all the assumptions in this case. In the second case, ξ is equal to difference of two independent copies of Poisson random variable with parameter 1. Then the cumulant function $H(t) = e^t + e^{-t} - 2$. For $v \neq 0$, we have $t_v^\varepsilon = \operatorname{sign}(v) \log \frac{|v|}{\varepsilon} + O(\varepsilon^2)$. Consequently, (2.6) holds with abovementioned norming factor. Furthermore condition (2.8) holds too.

It is interesting to note that for $\xi = \xi^i + \xi^{ii}$, where ξ^i and ξ^{ii} are independent random variables from these examples with $q^i(\varepsilon) = \varepsilon^2$, $q^{ii}(\varepsilon) = \frac{\varepsilon}{|\log \varepsilon|}$ respectively, we get

$$q(\varepsilon) = q^{ii}(\varepsilon) \quad \text{and} \quad I(v) = |v|.$$

2. If function $f(z_1, ..., z_m, y)$, appearing in (ref1.2), is linear in y:

$$f(z_1, ..., z_m, y) = a(z_1, ..., z_m) + b(z_1, ..., z_m)y, \tag{3.1}$$

where $b(z_1, ..., z_m)$ can be zero for some values of the argument $(z_1, ..., z_m)$, then the rate function $J_\infty(\overline{u})$ is given by the simple formula (with the rule $0/0 = 0$)

$$J_\infty(\overline{u}) = \begin{cases} \sum_{k=m}^\infty I\left(\frac{u_k - a(u_{k-1}, ..., u_{k-m})}{b(u_{k-1}, ..., u_{k-m})}\right), & u_0 = x_0, ..., u_{m-1} = x_{m-1} \\ \infty, & \text{otherwise.} \end{cases}$$

In the case of the Markov model $X_k^\varepsilon = a(X_{k-1}^\varepsilon) + b(X_{k-1}^\varepsilon)\varepsilon\xi_k$, with $(0,1)$-Gaussian random variable ξ_1 an analogy of Freidlin-Wentzell's result [4] for the diffusion $dX_t^\varepsilon = a(X_t^\varepsilon)dt + \varepsilon b(X_t^\varepsilon)\varepsilon dW_t$ (W_t is Wiener process) holds. Namely,

$$J_\infty(\overline{u}) = \begin{cases} \frac{1}{2}\sum_{k=1}^\infty \frac{[u_k - a(u_{k-1})]^2}{b^2(u_{k-1})}, & u_0 = x_0 \\ \infty, & u_0 \neq x_0. \end{cases}$$

3. The next result plays an important role in proving Theorem 2.2, it also has an independent interest. Notation P_{x_0} will be used for designating '$X_0^\varepsilon = x_0$'.

Theorem 3.1. *Let the assumptions of Theorem 2.2 be fulfilled and $M \geq 1$.*

1) If $|a| < 1$, *then*

$$\lim_{M\to\infty}\lim_{\varepsilon\to 0}\sup_{|x_0|\le 1}|\varepsilon^2\log P_{x_0}(\max_{1\le k\le M}|X_k^\varepsilon|\ge 1)+\frac{1}{2}(1-a^2)|=0.$$

2) If $|a|\ge 1$ *and* $x_0=0$, *then*

$$\lim_{\varepsilon\to 0}\varepsilon^2\log P(\max_{1\le k\le M}|X_k^\varepsilon|\ge 1)=-\frac{1}{2\sum_{j=0}^{M-1}a^{2j}}.$$

Proof: 1) Evidently

$$\sup_{|x_0|\le 1}P_{x_0}(\max_{1\le k\le M}|X_k^\varepsilon|\ge 1)=P_\alpha(\max_{1\le k\le M}|X_k^\varepsilon|\ge 1),$$

where $\alpha=1$ or -1. By symmetry

$$\inf_{|x_0|\le 1}P_{x_0}(\max_{1\le k\le M}|X_k^\varepsilon|\ge 1)=P_0(\max_{1\le k\le M}|X_k^\varepsilon|\ge 1).$$

Consequently the desired statement holds if for $x_0=0$ and $x_0=\pm 1$

$$\lim_{\varepsilon\to 0}\varepsilon^2\log P_{x_0}(\max_{1\le k\le M}|X_k^\varepsilon|\ge 1)=-\frac{1}{2}(1-a^2). \tag{3.2}$$

The family $(X_k^\varepsilon)_{k\ge 1}$, defined by recursion (2.12) with initial condition $X_0^\varepsilon=x_0$, obeys the LDP with the rate function

$$J_\infty(\overline{u})=\frac{1}{2}\sum_{k=1}^{\infty}[u_k-af(u_{k-1})]^2,\ u_0=x_0$$

Hence by (1) and (2) of Section 2

$$\begin{aligned}\overline{\lim}_{\varepsilon\to 0}\varepsilon^2\log P_{x_0}(\max_{1\le k\le M}|X_k^\varepsilon|\ge 1) &\le -\frac{1}{2}\inf_{\mathcal{M}_1}\sum_{k=1}^{M}(u_k-af(u_{k-1}))^2\\ \underline{\lim}_{\varepsilon\to 0}\varepsilon^2\log P_{x_0}(\max_{1\le k\le M}|X_k^\varepsilon|> 1) &\ge -\frac{1}{2}\inf_{\mathcal{M}_2}\sum_{k=1}^{M}(u_k-af(u_{k-1}))^2,\end{aligned}$$

where $\mathcal{M}_1=(u\in R^M:\max_{1\le k\le M}|u_k|\ge 1)$ and $\mathcal{M}_2=(u\in R^M:\max_{1\le k\le M}|u_k|> 1)$. Obviously $\inf_{\mathcal{M}_1}=\inf_{\mathcal{M}_2}$ and so we obtain

$$\lim_{\varepsilon\to 0}\varepsilon^2\log P_{x_0}(\max_{1\le k\le M}|X_k^\varepsilon|\ge 1)=-\frac{1}{2}\inf_{\mathcal{M}_1}\sum_{k=1}^{M}(u_k-af(u_{k-1}))^2\ (=-\frac{1}{2}W_M(x_0)).$$

Put $\tau = \min\{1 \le k \le M : |u_k| \ge 1\}$. Then

$$W_M(x_0) = \min_{\tau \le M} \min_{|u_k|<1,\ 1\le k\le \tau-1;\ |u|_\tau=1} \sum_{k=1}^{\tau}(u_k - au_{k-1})^2 \tag{3.3}$$

To define the right hand side of (3.3), consider a control problem. Let $u_k, k = 0, 1, ..., \tau$ be controlled sequence, defined by recursion $u_k = au_{k-1} + w_k$ subject to initial $u_0 = x_0, (|x_0| \le 1)$ and final condition $|u_\tau| = 1$ and control $w_k, k = 1, ..., \tau$. We use an obvious relation

$$\min_{|u_k|<1,\ 1\le k\le \tau-1;\ |u|_\tau=1} \sum_{k=1}^{\tau}(u_k - au_{k-1})^2 = \min_{w_k, k=1,...,\tau} \sum_{k=1}^{\tau} w_k^2$$

with constrations: $w_k, k = 1, ..., \tau : |u_k| < 1, k \le \tau - 1, |u_\tau| = 1$. By virtue of representation $u_\tau = a^\tau x_0 + \sum_{k=1}^{\tau} a^{\tau-k} w_k$ and Cauchy-Scwartz's inequality

$$|u_\tau - a^\tau x_0| \le \sqrt{\sum_{k=1}^{\tau} a^{2(\tau-k)} \sum_{k=1}^{\tau} w_k^2}$$

we get a lower bound:

$$\begin{aligned} \sum_{k=1}^{\tau} w_k^2 &\ge \frac{(u_\tau - a^\tau x_0)^2}{\sum_{k=1}^{\tau} a^{2(\tau-k)}} \\ &= \frac{(u_\tau - a^\tau x_0)^2}{\sum_{j=0}^{\tau-1} a^{2j}} = \frac{(u_\tau - a^\tau x_0)^2(1-a^2)}{1-a^{2\tau}}. \end{aligned} \tag{3.4}$$

This lower bound is attainable on the control $w_k^\circ = a^{(\tau-k)} \frac{(u_\tau - a^\tau x_0)(1-a^2)}{1-a^{2\tau}},\ 1 \le k \le \tau$. Take first $a^\tau x_0 \ge 0$ and $u_\tau = 1$. Then $w_k^\circ = a^{(\tau-k)} \frac{(1-a^\tau x_0)(1-a^2)}{1-a^{2\tau}},\ 1 \le k \le \tau$. To the control sequence w_k° corresponds the controlled process

$$\begin{aligned} u_k^\circ &= a^\tau x_0 + \frac{(1-a^\tau x_0)(1-a^2)}{1-a^{2\tau}} \sum_{j=1}^{k} a^{(\tau+k-2j)} \\ &= a^\tau x_0 + (1-a^\tau x_0)\frac{1-a^{2k}}{1-a^{2\tau}},\ 1 \le k \le \tau \end{aligned}$$

which has a property $|u_k^\circ| < 1,\ k \le \tau - 1$ and $u_\tau^\circ = 1$. Therefore the abovementioned control is admissible and optimal. In the same way the case $a^\tau x_0 < 0,\ u_\tau = -1$ is investigated. Since the lower boud in (3.4) is attainable, we obtain

$$\min_{|u_k|<1,\ 1\le k\le \tau-1;\ |u|_\tau=1} \sum_{k=1}^{\tau}(u_k - au_{k-1})^2 = \frac{(u_\tau - a^\tau x_0)^2(1-a^2)}{1-a^{2\tau}} \tag{3.5}$$

hence,

$$\frac{1}{2}W_M(x_0) = \frac{1}{2}(1-a^2)\min_{\tau\le M}\frac{(u_\tau - a^\tau x_0)^2}{1-a^{2\tau}}.$$

This implies the statement of the theorem since

$$\lim_{M\to\infty}\min_{\tau\le M}\frac{(u_\tau - a^\tau x_0)^2}{1-a^{2\tau}} = 1.$$

2) Repeating previous computations we arrive to

$$\lim_{\varepsilon\to 0}\varepsilon^2\log P_0(\max_{1\le k\le M}|X_k^\varepsilon|\ge 1) = -\frac{1}{2}W_M(0)) = -\frac{1}{2\sum_{j=0}^{M-1}a^{2j}}.$$

4 LDP for $\varepsilon\xi$.

In this Section we prove statement 1) of Theorem 2.1. There are different approaches for proving the LDP (see e.g. [1], [2], [9]). In our case considered (cf. Theorem 1.3 in [10]), it is enough to establish the exponential tightness:

$$\lim_{c\to\infty}\overline{\lim}_{\varepsilon\to 0}q(\varepsilon)\log P(|\varepsilon\xi|\ge c) = -\infty, \tag{4.1}$$

and the local LDP: for every $v\in R$

$$\begin{aligned}\lim_{\delta\to 0}\underline{\lim}_{\varepsilon\to 0}q(\varepsilon)\log P(|\varepsilon\xi_1 - v|\le\delta) &= \lim_{\delta\to 0}\overline{\lim}_{\varepsilon\to 0}q(\varepsilon)\log P(|\varepsilon\xi_1 - v|\le\delta).\\ &= -I(v).\end{aligned} \tag{4.2}$$

These two requirements are equivalent to the LDP for the family $(\varepsilon\xi)$ with the rate function $I(v)$ and the same norming factor. Below we verify (4.1) and (4.2). Since (4.1) holds when

$$\begin{aligned}\lim_{c\to\infty}\overline{\lim}_{\varepsilon\to 0}q(\varepsilon)\log P(\varepsilon\xi\ge c) &= -\infty\\ \lim_{c\to\infty}\overline{\lim}_{\varepsilon\to 0}q(\varepsilon)\log P(\varepsilon\xi\le -c) &= -\infty,\end{aligned} \tag{4.3}$$

both relations in (4.3) are verified separately. By Chernoff's inequality

$$P(\varepsilon\xi > c)\le\exp\big(-(tc)/\varepsilon + H(t)\big).$$

On the other hand, taking into account (2.4), one can conclude that for $c>0$:

$$\sup_{t>0}[t(c/\varepsilon) - H(t)] = \sup_{t\in R}[t(c/\varepsilon) - H(t)].$$

Hence by (2.7)

$$\overline{\lim}_{\varepsilon\to 0} q(\varepsilon)\log P(\varepsilon\xi > c) \le -I(c) \to -\infty,\ c\to\infty,$$

that is the first part in (4.3) holds. The second part is proved in the same way.

To check the local LDP put $Z = \exp(t\xi - H(t))$. Due to $EZ = 1$ we get an obvious inequality

$$EI(|\xi - u/\varepsilon| \le \delta/\varepsilon)Z \le 1 \tag{4.4}$$

which remains true when Z is replaced by its lower bound on the set $\{|\xi - u/\varepsilon| \le \delta/\varepsilon\}$: $\underline{Z} = \exp\big(-(\delta/\varepsilon)|t| + t(u/\varepsilon) - H(t)\big)$. Thus,

$$q(\varepsilon)\log P(|\xi - v/\varepsilon| \le \delta/\varepsilon) \le q(\varepsilon)(\delta/\varepsilon)|t| - q(\varepsilon)[t(v/\varepsilon) - H(t)]. \tag{4.5}$$

Using (4.5) for every finite t, and taking $t = t_v^\varepsilon$ when $I(v) < \infty$, we find by virtue of the first part of (2.8) that

$$\overline{\lim}_{\delta\to 0}\overline{\lim}_{\varepsilon\to 0} q(\varepsilon)\log P(|\xi - v/\varepsilon| \le \delta/\varepsilon) \le -I(v). \tag{4.6}$$

Let now for some v_0 $I(v_0) = \infty$. We show that '$-\infty$' is the upper bound for $\overline{\lim}_{\delta\to 0}\overline{\lim}_{\varepsilon\to 0}\log P(|\varepsilon\xi - v| \le \delta)$. Let $v_0 > 0$. Then by the triangular inequality $|\varepsilon\xi - v_0| \ge |\varepsilon\xi - v_0 + \gamma| - \gamma$ with $0 < \gamma < v_0$ we get

$$P(|\varepsilon\xi - v_0| \le \delta) \le P(|\varepsilon\xi - (v_0 - \gamma)| \le \delta + \gamma. \tag{4.7}$$

By the assumption $I(v)$ is left continuous function at point v_0 and so, $I(v_0 - \gamma) < \infty$. Then by (4.7) and (4.5)

$$\begin{aligned} q(\varepsilon)\log P(|\xi - v_0/\varepsilon| &\le \delta/\varepsilon) \\ &\le q(\varepsilon)\log P(|\xi - (v_0-\gamma)/\varepsilon| \le (\delta+\gamma)/\varepsilon) \\ &\le q(\varepsilon)(\delta + \gamma/\varepsilon)|t| - q(\varepsilon)[t((v_0-\gamma)/\varepsilon) - H(t)] \end{aligned} \tag{4.8}$$

and by the first part (2.8) we obtain

$$\overline{\lim}_{\delta\to 0}\overline{\lim}_{\varepsilon\to 0} q(\varepsilon)\log P(|\xi - v_0/\varepsilon| \le \delta/\varepsilon) \le \lim_{\gamma\to 0}\Big[\text{const.}\gamma - I(v_0-\gamma)\Big] = -\infty.$$

For $v_0 < 0$ the upper bound '$-\infty$' is derived in the same way.

It is clear that in proving the lower bound only the case $I(v) < \infty$ has to be considered. Put $\Lambda_t(y) = \exp(ty - H(t))$ and denote by $P(y)$ the distribution function of ξ. Since

$E\Lambda_t(\xi) = 1$, $Q_t(y)$ with $d_y Q_t(y) = \Lambda_t(y) dP(y)$ is a distribution function too. It obeys the following properties:

$$\int_R y dQ_t(y) = H'(t) \quad \text{and} \quad \int_R [y - H'(t)]^2 dQ_t(y) = H''(t). \tag{4.9}$$

Taking $t = t_v^\varepsilon$ we find $dP(y) = \exp\big(- t_v^\varepsilon y + H(t_v^\varepsilon)\big) dQ_{t_v^\varepsilon}(y)$. Hence

$$\begin{aligned} P(|\varepsilon\xi - v| \le \delta) &= \int_{|y-(v/\varepsilon)|\le(\delta/\varepsilon)} \exp\big(- t_v^\varepsilon y + H(t_u^\varepsilon)\big) dQ_{t_v^\varepsilon}(x) \\ &\ge \exp\big(- |t_v^\varepsilon|(\delta/\varepsilon) - t_v^\varepsilon v + H(t_v^\varepsilon)\big) \\ &\quad \times \int_{|x-(v/\varepsilon)|\le(\delta/\varepsilon)} dQ_{t_u^\varepsilon}(x). \end{aligned} \tag{4.10}$$

The desired lower bound

$$\underline{\lim}_{\delta\to 0}\underline{\lim}_{\varepsilon\to 0} \log P(|\xi - v/\varepsilon| \le \delta/\varepsilon) \ge -I(v) \tag{4.11}$$

follows from (4.10) and the lower bound

$$\underline{\lim}_{\varepsilon\to 0} \int_{|x-(v/\varepsilon)|\le(\delta/\varepsilon)} dQ_{t_u^\varepsilon}(x) \ge 1, \;\; \delta > 0.$$

The latter holds by Chebyshev's inequality, (4.9) and the second part of condition (2.8) for every $\delta > 0$ we get

$$\begin{aligned} \int_{|x-(v/\varepsilon)|\le(\delta/\varepsilon)} dQ_{t_v^\varepsilon}(x) &= 1 - \int_{|x-(u/\varepsilon)|>(\delta/\varepsilon)} dQ_{t_v^\varepsilon}(x) \\ &\ge 1 - \frac{\varepsilon^2}{\delta^2}\int_R (x-u/\varepsilon)^2 dQ_{t_v^\varepsilon}(x) \\ &= 1 - \frac{\varepsilon^2}{\delta^2} H''(t_v^\varepsilon) \\ &\to 1, \quad \varepsilon \to 0. \end{aligned}$$

5 LDP's for $(\varepsilon\xi_k)_{k\ge 1}$ and $(X_k^\varepsilon)_{k\ge m}$.

In this Section, the statements 2) and 3) of Theorem 2.1 and the Remark to it are established.

Proof of Statement 2).For $n > 1$, let us check the LDP for the family $(\varepsilon\xi_k)_{1\le k\le n}$ in the metric space (R^n, ρ^n), where for $x, y \in R^n$ $\rho^n(x,y) = \sum_{k=1}^n |x_k - y_k|$. Since $(\varepsilon\xi_k)_{1\le k\le n}$ is a vector of i.i.d. random variables the LDP holds with the rate function (see [12])

$$I_n(v^n) = \sum_{k=1}^n I(v_k) \tag{5.1}$$

Next by Dawson-Gärtner's theorem (see [13] or [1]), the LDP for family $(\varepsilon\xi_k)_{k\geq 1}$ holds too and what is more the rate function is defined as

$$I(\overline{v}) = \sum_{k=1}^{\infty} I(v_k). \tag{5.2}$$

Proof of Statement 3). The continuity of the mapping $(\varepsilon\xi_k)_{k\geq m} \to (X_k^\varepsilon)_{k\geq m}$ in the metric ρ is obvious. Therefore by the contraction principle (continuous mapping method) (see [3] and [11]) the family $(X_k^\varepsilon)_{k\geq m}$ obeys the LDP with the same norming factor and the rate function

$$J_\infty(\overline{u}) = \inf_{(v_k, k\geq m: u_k = f(u_{k-1},\dots,u_{k-m},v_k))} I_\infty(\overline{v}),$$

where $\inf\{\emptyset\} = \infty$ and $I_\infty(\overline{v})$ is defined in (2.9) and it remains to note that

$$\inf_{(v_k, k\geq m: u_k = f(u_{k-1},\dots,u_{k-m},v_k)} I_\infty(\overline{v}) = \sum_{k=m}^{\infty} \inf_{(v_k, k\geq m: u_k = f(u_{k-1},\dots,u_{k-m},v_k))} I(v_k).$$

Remark to Theorem 2.1 holds since by (2.11) the random variable ξ_1^{ii} is exponentially negligible with respect to the norming factor $q^i(\varepsilon)$: for any $\delta > 0$

$$\lim_{\varepsilon\to 0} q^i(\varepsilon) \log P(|\varepsilon\xi_1^{ii}| > \delta) = -\infty.$$

6 Asymptotics of exit time.

In this Section we prove Theorem 2.2.

1) Let M be an integer. It is clear that $\varepsilon^2 \log E\tau^\varepsilon$ and $\varepsilon^2 \log E\frac{\tau^\varepsilon}{M}$ have the same asymptotics as $\varepsilon \to 0$. The following bounds for $E\frac{\tau^\varepsilon}{M}$ hold:

$$\sum_{n\geq 0} P(\tau^\varepsilon > Mn) - 2 \leq E\frac{\tau^\varepsilon}{M} \leq \sum_{n\geq 0} P(\tau^\varepsilon > Mn) + 1. \tag{6.1}$$

In fact, if $[z]$ is integer part of z, that is $[z] = n$ for $n \leq z < n+1$, then, since $[z] \leq z \leq [z]+1$, we find

$$\begin{aligned} E\frac{\tau^\varepsilon}{M} &\leq E\Big[\frac{\tau^\varepsilon}{M}\Big] + 1 \\ &= \sum_{n=1}^{\infty} P\Big(\Big[\frac{\tau^\varepsilon}{M}\Big] \geq n\Big) + 1 \\ &\leq \sum_{n=1}^{\infty} P\Big(\tau^\varepsilon \geq Mn\Big) + 1 \\ &\leq \sum_{n=0}^{\infty} P\Big(\tau^\varepsilon > Mn\Big) + 1 \end{aligned}$$

and similarly,

$$
\begin{aligned}
E\frac{\tau^\varepsilon}{M} &\geq E\Big[\frac{\tau^\varepsilon}{M}\Big] \\
&= \sum_{n=1}^{\infty} P\Big(\Big[\frac{\tau^\varepsilon}{M}\Big] \geq n\Big) \\
&\geq \sum_{n=1}^{\infty} P\Big(\tau^\varepsilon \geq M(n+1)\Big) \\
&\geq \sum_{n=0}^{\infty} P\Big(\tau > Mn\Big) - 2.
\end{aligned}
$$

Estimates in (6.1) allow to conclude that $\lim_{\varepsilon\to 0}\varepsilon^2 \log E\frac{\tau^\varepsilon}{M} = \frac{1}{2}(1-a^2)$, if

$$
\lim_{M\to\infty}\lim_{\varepsilon\to 0}\varepsilon^2 \log \sum_{n=0}^{\infty} P\Big(\tau > Mn\Big) = \frac{1}{2}(1-a^2). \tag{6.2}
$$

To establish (6.2) we use the Markov property of the process $(X^\varepsilon_k)_{k\geq 1}$. By the definition of τ^ε it follows that sets $\{\tau^\varepsilon > Mn\}$ and $\{\max_{1\leq k\leq Mn}|X^\varepsilon_k| < 1\}$ coincide. Hence

$$
\begin{aligned}
P(\tau^\varepsilon > Mn) &= P(\max_{1\leq k\leq Mn}|X^\varepsilon_k| < 1) \\
&= E\Big(I(\max_{1\leq k\leq M(n-1)}|X^\varepsilon_k| < 1)P_{X^\varepsilon_{M(n-1)}}(\max_{M(n-1)<k\leq Mn}|X^\varepsilon_k| < 1)\Big).
\end{aligned}
$$

In turn, due to homogeneity of $(X^\varepsilon_k)_{k\geq 1}$, we get recurrent inequalities:

$$
P(\tau^\varepsilon > Mn)\begin{cases} \leq P(\tau^\varepsilon > M(n-1))\Big(1 - \inf_{|x|\leq 1} P_x(\max_{1\leq k\leq M}|X^\varepsilon_k| \geq 1)\Big) \\ \geq P(\tau^\varepsilon > M(n-1))\Big(1 - \sup_{|x|\leq 1} P_x(\max_{1\leq k\leq M}|X^\varepsilon_k| \geq 1)\Big). \end{cases}
$$

Since $P(\tau^\varepsilon > 0) = 1$,

$$
P(\tau^\varepsilon > Mn)\begin{cases} \leq \Big(1 - \inf_{|x|\leq 1} P_x(\max_{1\leq k\leq M}|X^\varepsilon_k| \geq 1)\Big)^n \\ \geq \Big(1 - \sup_{|x|\leq 1} P_x(\max_{1\leq k\leq M}|X^\varepsilon_k| \geq 1)\Big)^n. \end{cases}
$$

Therefore

$$
\sum_{n=0}^{\infty} P(\tau^\varepsilon > Mn)\begin{cases} \leq \frac{1}{\inf_{|x|\leq 1} P_x\big(\max_{1\leq k\leq M}|X^\varepsilon_k|\geq 1\big)} \\ \geq \frac{1}{\sup_{|x|\leq 1} P_x\big(\max_{1\leq k\leq M}|X^\varepsilon_k|\geq 1\big)} \end{cases}
$$

and so, (6.2) is implied by Theorem 3.1.

2) Since sets $\{\tau^\varepsilon \leq M\}$ and $\{\max_{1\leq k\leq M}|X^\varepsilon_k| \geq 1\}$ coincide, we derive the desired statement from the second statement of Theorem 3.1.

Acknowledgment. Authors would like to thank anonymous reviewer for his comments which lead to correction of mistakes and misprints and improved the paper.

References

[1] A. Dembo and O. Zeitouni *Large Deviations Techniques and Applications.* Jones and Bartlet. 1993.

[2] J.D. Deuschel and D.W. Stroock *Large Deviations.* Academic Press. Boston. 1989.

[3] M.I. Freidlin "Action functional for a class of stochastic processes". *Theory of Probab. Appl.*, **17**, 1972, pp. 511–515.

[4] M.I. Freidlin and A.D. Wentzell A.D. (1984). *Random Perturbation s of Dynamical Systems.* N.Y. Springer. 1984.

[5] Y. Kifer "A discrete time version of the Wentzell-Freidlin theory". *Ann. Probab.*, **18**, 1990, 1676–1692.

[6] Y. Kifer, Y. *Perturbations of Dynamical Systems.* Birkhauser, Boston. 1988.

[7] F.C. Klebaner and O. Zeitouni "The exit problem for a class of density-dependent branching systems". *Ann. Appl. Probab.*, **4**, No 4, 1994, pp. 1188–1205.

[8] G.J. Morrow and S. Sawyer "Large deviation results for a class of Markov chains arising from population genetics". *Ann. Probab.*, **17**, 1989, pp. 1124–1146.

[9] A. Shwarz and A. Weiss *Large Deviations for Performance Analysis, communication and Computing: I, II.* AT&T Bell Laboratories. 1994.

[10] R.Sh. Liptser and A.A. Pukhalskii "Limit theorems on large deviations for semimartingales". *Stochastic and Stochastics Reports*, **38**, 1992, pp. 201–249.

[11] S.R.S Varadhan *Large Deviations and Applications.* SIAM. Philadelphia. 1984.

[12] J. Lynch and J. Sethuraman "Large deviations for processes with independent increments". *Ann. Prob.*, **15**, 1987, pp. 610-627.

[13] D.A. Dawson and J. Gärtner "Large deviations from the McKean-Vlasov limit for weakly interacting diffusions", *Stochastics*, **20**, 1987, pp. 247-308.