Uniform concentration inequality for ergodic diffusion processes observed at discrete times. *

L. Galtchouk † S. Pergamenshchikov ‡

September 15, 2011

Abstract

In this paper a concentration inequality is proved for the deviation in the ergodic theorem in the case of discrete time observations of diffusion processes. The proof is based on the geometric ergodicity property for diffusion processes. As an application we consider the nonparametric pointwise estimation problem for the drift coefficient under discrete time observations.

Keywords: Ergodic diffusion processes; Markov chains; Tail distribution; Upper exponential bound; Concentration inequality.
AMS 2000 Subject Classifications: 60F10

*The second author is partially supported by the RFFI-Grant 09-01-00172-a.

†Department of Mathematics, Strasbourg University 7, rue Rene Descartes, 67084, Strasbourg, France, e-mail: leonid.galtchouk@math.unistra.fr

‡Laboratoire de Mathématiques Raphael Salem, Avenue de l'Université, BP. 12, Université de Rouen, F76801, Saint Etienne du Rouvray, Cedex France, e-mail: Serge.Pergamenchtchikov@univ-rouen.fr

1 Introduction

We consider the process $(y_t)_{t\geq 0}$ governed by the stochastic differential equation

$$\mathrm{d}y_t = S(y_t)\,\mathrm{d}t + \sigma(y_t)\mathrm{d}W_t\,, \quad 0 \leq t \leq T\,, \tag{1.1}$$

where $(W_t, \mathcal{F}_t)_{t\geq 0}$ is a standard Wiener process, y_0 is a initial condition and $\vartheta = (S, \sigma)$ are unknown functions. For this model we consider the pointwise estimation problem for the function S at a fixed point $x_0 \in \mathbb{R}$ (i.e. $S(x_0)$), on the basis of the discrete time observations of the process (1.1), i.e.

$$(y_{t_j})_{1\leq j\leq N}\,, \tag{1.2}$$

where $t_j = j\delta$, $N = [T/\delta]$ and δ is some positive fixed observation frequency which will be specified later. Usually, for this problem one uses kernel estimators $\widehat{S}_N(x_0)$ defined as

$$\widehat{S}_N(x_0) = \frac{\sum_{k=1}^N \psi_{h,x_0}(y_{t_k})\,\Delta y_{t_k}}{\sum_{k=1}^N \psi_{h,x_0}(y_{t_k})\,\Delta t_k}\,, \quad \psi_{h,x_0}(y) = \frac{1}{h}\,\Psi\left(\frac{y-x_0}{h}\right)\,, \tag{1.3}$$

where $\Psi(y)$ is a kernel function which equals to zero for $|y| \geq 2$ and will be specified later, $0 < h < 1$ is a bandwidth, $\Delta y_{t_k} = y_{t_k} - y_{t_{k-1}}$ and $\Delta t_k = \delta$.

Main difficulty in this estimator is that the denominator is random. Therefore, to obtain the convergence rate for this estimator we have to study the behavior of the denominator, more precisely, one needs to show that

$$\sum_{k=1}^N \psi_{h,x_0}(y_{t_k})\Delta t_k \approx \pi_\vartheta(\psi_{h,x_0})hT \quad \text{as} \quad T \to \infty\,,$$

where

$$\pi_\vartheta(\psi_{h,x_0}) = \int_{\mathbb{R}} \psi_{h,x_0}(y)\,q_\vartheta(y)\,\mathrm{d}y \tag{1.4}$$

and q_ϑ is the ergodic density defined in (2.2).

Unfortunately, the ergodic theorem does not permit to obtain this kind of result because the times t_k and the bandwidth h depend on T. Usually one obtains such properties through concentration inequalities for the deviation in the ergodic theorem, i.e. one needs to study the limit behavior of the deviation

$$D_T(\phi) = \sum_{k=1}^N \Big(\phi(y_{t_k}) - \pi_\vartheta(\phi)\Big)\,\Delta t_k \tag{1.5}$$

for some functions ϕ which can be dependent on T, for example, $\phi(\cdot) = \psi_{h,x_0}(\cdot)$. More precisely, we need to show, that for any $\varepsilon > 0$ and for any $m > 0$, uniformly over ϑ,

$$\lim_{T\to\infty} T^m\,\mathbf{P}_\vartheta\Big(|D_T(\psi_{h,x_0})| > \varepsilon T\Big) = 0\,, \tag{1.6}$$

where $\mathbf{P}_\vartheta$ is the law of the process $(y_t)_{t\geq 0}$ under the coefficients $\vartheta = (S, \sigma)$. Usually, to get properties of type (1.6) one needs to establish an exponential inequality for the deviations (1.5).

There are a number of papers devoted to concentration inequalities for functions of independent random variables (we refer the reader to [2] and references therein), for functions of dependent random variables (see [4], [5], [14]). For Markov chains such inequalities were obtained in [1]. For continuous time Markov processes an exponential concentration inequality was obtained in [3] (see also references therein). Some applications of concentration inequalities to statistics are presented in [13]. Concentration inequalities for diffusion processes are given in [8], [16], [18].

For statistical applications, we need uniform upper bounds for the tail distribution over functions ϕ like to the exponential bounds in [8]. We can not apply directly the method from [8], since there it is based on the continuous times version of the Ito formula. In this paper we apply this approach through uniform (over the functions S) geometric ergodicity. We recall (see [15]), that the geometric ergodicity yields a geometric rate in the convergence

$$\lim_{t\to\infty} \mathbf{E}_\vartheta(g(y_t)|y_0 = x) = \pi_\vartheta(g)$$

for any integrable functions g and any initial value $x \in \mathbb{R}$. Here $\mathbf{E}_\vartheta$ denotes the expectation with respect to the distribution $\mathbf{P}_\vartheta$. In [10] through the Lyapunov functions method it is shown that the process (1.1) is geometrically ergodic uniformly over functions $\vartheta = (S, \sigma)$ from the functional class Θ defined in (2.1).

The paper is organized as follows. In the next section we formulate the main results. In Section 3 we introduce all the necessary parameters. In Section 4 we show a concentration inequality in ergodic theorem for the continuous observations of the process (1.1). In Section 5 we announce the uniform geometric ergodic property for the process (1.1). In Section 6 we give the Burkhölder inequality for dependent random variables. In Section 7 we prove all main results. The Appendix contains the proofs of some auxiliary results.

2 Main results

First we describe the functional class Θ for functions $\vartheta = (S, \sigma)$ defined in [10]. We start with some real numbers $\mathbf{x}_* \geq 1$, $M > 0$ and $L > 1$ for which we denote by $\Sigma_{L,M}$ the class of functions S from $\mathbf{C}^1(\mathbb{R})$ such that

$$\sup_{|x|\leq \mathbf{x}_*} \left(|S(x)| + |\dot{S}(x)| \right) \leq M$$

and

$$-L \le \inf_{|x|\ge \mathbf{x}_*} \dot{S}(x) \le \sup_{|x|\ge \mathbf{x}_*} \dot{S}(x) \le -L^{-1}\,.$$

Furthermore, for some fixed numbers $0 < \sigma_{\rm min} \le \sigma_{\rm max} < \infty$, we denote by $\mathcal{V}$ the class of the functions σ from $\mathbf{C}^2(\mathbb{R})$ such that

$$\sigma_{min} \le \inf_{x\in\mathbb{R}} \min\left(|\sigma(x)|\,,\,|\dot{\sigma}(x)|\,,\,|\ddot{\sigma}(x)|\right)$$
$$\le \sup_{x\in\mathbb{R}} \max\left(|\sigma(x)|\,,\,|\dot{\sigma}(x)|\,,\,|\ddot{\sigma}(x)|\right) \le \sigma_{max}\,.$$

Finally, we set

$$\Theta = \Sigma_{L,M} \times \mathcal{V}\,. \tag{2.1}$$

It should be noted (see, for example, [11]), that for any $\vartheta = (S,\sigma) \in \Theta$, the equation (1.1) has a unique strong solution which is a ergodic process with the invariant density q_ϑ defined as

$$q_\vartheta(x) = \left(\int_{\mathbb{R}} \sigma^{-2}(z)\, e^{\widetilde{S}(z)} \mathrm{d}z\right)^{-1} \sigma^{-2}(x)\, e^{\widetilde{S}(x)}\,, \tag{2.2}$$

where $\widetilde{S}(x) = 2\int_0^x S_1(v)\mathrm{d}v$ and $S_1(x) = S(x)/\sigma^2(x)$.

Now we describe the functional classes for the functions ϕ. First, for any parameters $\nu_0 > 0$ and $\nu_1 > 0$ we set

$$\mathcal{V}_{\nu_0,\nu_1} = \left\{\phi \in \mathbf{C}(\mathbb{R})\,:\, |\phi|_1 \le \nu_0\,,\, |\phi|_* \le \nu_1\right\}\,, \tag{2.3}$$

where $|\phi|_1 = \int_{\mathbb{R}} |\phi(y)|\,\mathrm{d}y$ and $|\phi|_* = \sup_{y\in\mathbb{R}} |\phi(y)|$.

For any function ϕ from $\mathbf{C}^2(\mathbb{R})$ we denote by $\mathcal{L}_\vartheta(\phi)$ the generator operator for the process (1.1), i.e.

$$\mathcal{L}_\vartheta(\phi)(y) = S(y)\dot{\phi}(y) + \frac{\sigma^2(y)}{2}\ddot{\phi}(y)\,.$$

Using this notation, we set

$$\mu(\phi) = \sup_{\vartheta\in\Theta} \|\mathcal{L}_\vartheta(\phi)\|_* \quad \text{and} \quad \widetilde{\mu}(\phi) = \sup_{\vartheta\in\Theta} |\widetilde{\pi}_\vartheta(\phi)|\,, \tag{2.4}$$

where $\widetilde{\pi}_\vartheta(\phi) = \pi_\vartheta(\mathcal{L}_\vartheta(\phi))$. Now for any vector $\nu = (\nu_0,\nu_1,\nu_2,\nu_3,\nu_4)$ from $\mathbb{R}^5_+$ we set

$$\mathcal{K}_\nu = \left\{\phi \in \mathcal{V}_{\nu_0,\nu_1}\,:\, \|\dot{\phi}\|_* \le \nu_2\,,\, \mu(\phi) \le \nu_3\,,\, \widetilde{\mu}(\phi) \le \nu_4\right\}\,. \tag{2.5}$$

Theorem 2.1. *For any vector* $\nu=(\nu_0,\nu_1,\nu_2,\nu_3,\nu_4)$ *from* $\mathbb{R}^5_+$ *and any* $0<\delta\le 1$ *there exist positive parameters* $z_0=z_0(\delta,\nu)$, $\gamma=\gamma(\delta,\nu)$ *and* $\varkappa=\varkappa(\delta,\nu)$ *such that*

$$\sup_{T\ge 1}\,\sup_{z\ge z_0}\,\sup_{\phi\in\mathcal{K}_\nu}\,\sup_{\vartheta\in\Theta} e^{z\min(\varkappa z\,,\gamma)}\,\mathbf{P}_\vartheta\left(|D_T(\phi)|\,\ge z\sqrt{N}\right)\le 4\,, \tag{2.6}$$

where the parameters z_0, γ *and* $\varkappa$ *are defined in* (3.5)–(3.6).

Now we apply this theorem to the pointwise estimation problem, i.e. for the functions ψ_{h,x_0} defined in (1.3). To this end we assume that the frequency δ in the observations (1.2) is of the following form

$$\delta=\delta_T=\frac{1}{T\,l_T}\,, \tag{2.7}$$

where the function l_T is such that for any $m>0$

$$\lim_{T\to\infty}\frac{l_T}{T^m}=0\quad\text{and}\quad\lim_{T\to\infty}\frac{l_T}{\ln T}=+\infty\,. \tag{2.8}$$

Further, let $\epsilon=\epsilon_T$ be a positive function satisfying the following properties

$$\lim_{T\to\infty}\epsilon_T=0,\quad\lim_{T\to\infty}\frac{l_T}{T\epsilon_T}=0\quad\text{and}\quad\lim_{T\to\infty}\frac{\epsilon^5_T l_T}{\ln T}=+\infty\,. \tag{2.9}$$

We can take, for example, for some $\iota>0$

$$l_T=\ln^{1+6\iota}(T+1)\quad\text{and}\quad\epsilon_T=\frac{1}{\ln^\iota(T+1)}\,.$$

Theorem 2.2. *Assume that the kernel function* Ψ *in* (1.3) *is two continuously differentiable. Moreover, assume that the functions* δ_T *and* l_T *satisfy the properties* (2.7) *and* (2.9). *Then there exist coefficients* $z^*_0=z^*_0(\Psi)>0$ *and* $\gamma^*=\gamma^*(\Psi)>0$ *such that*

$$\limsup_{T\to\infty} e^{a\gamma^*\,l_T}\,\sup_{a\ge a_*}\,\sup_{h\ge T^{-1/2}}\,\sup_{\vartheta\in\Theta}\mathbf{P}_\vartheta\left(|D_T(\psi_{h,x_0})|\,\ge\,a\,T\right)\le 4\,, \tag{2.10}$$

where $a_*=z^*_0/l_T$, *the parameters* z^*_0 *and* γ^* *are given in Section 3.*

This theorem implies immediately the following

Corollary 2.1. *Assume, that all conditions of Theorem 2.2 hold. Then, for any* $m>0$,

$$\limsup_{T\to\infty} T^m\,\sup_{a\ge a_*}\,\sup_{h\ge T^{-1/2}}\,\sup_{\vartheta\in\Theta}\mathbf{P}_\vartheta\left(|D_T(\psi_{h,x_0})|\,\ge\,a\,T\right)=0\,.$$

Now we study the deviation (1.5) for the function

$$\chi_{h,x_0}(y) = \frac{1}{h}\,\chi\left(\frac{y-x_0}{h}\right)\,, \qquad (2.11)$$

where $\chi(y) = \mathbf{1}_{\{|y|\le 1\}}$.

Theorem 2.3. *Assume that the parameter δ has the form* (2.7)*. Then, for any $m>0$, and for any function ϵ_T, satisfying the condtions* (2.8) *and* (2.9)

$$\lim_{T\to\infty} T^m \sup_{h\ge T^{-1/2}} \sup_{\vartheta\in\Theta} \mathbf{P}_\vartheta\left(|D_T(\chi_{h,x_0})| \ge \epsilon_T\, T\right) = 0\,. \qquad (2.12)$$

Remark 2.1. *It is well known that to obtain the optimal rate in the estimation problem for a differentiable function S in the process* (1.1) *one needs to choose the bandwidth h as*

$$h = T^{-1/(2\alpha+1)}$$

with the regularity parameter $\alpha\ge 1$. This means that, really for the pointwise estimation problem, $h\ge T^{-1/3}$. But in the quadratic risk one needs to choose the parameter h as $h=T^{-1/2}$ (see [6]-[7],[9]).

3 Parameters

In this section we introduce all necessary constants and parameters. First, we set

$$\upsilon_1 = e^{\beta_1^2/(4\beta_2)} \quad \text{and} \quad \upsilon_2 = \sqrt{\pi/\beta_2}\, e^{\beta_1^2/(4\beta_2)}\,, \qquad (3.1)$$

where $\beta_1 = 2M/\sigma^2_{\min}$ and $\beta_2 = 1/L\sigma^2_{\max}$. Moreover, as we will see in Appendix, the ergodic density (2.2) is uniformly bounded by q^*, where

$$q^* = \frac{\sigma^2_{max}}{\sigma^2_{min}}\, e^{\beta_1 \mathbf{x}_* + \beta_1^2/(4\beta_2)}\,. \qquad (3.2)$$

Now we set

$$r = r(\nu_0) = \frac{2\nu_0}{\sigma^2_{\min}}\,\left(1+\upsilon_1 + q^*\left(\mathbf{x}_* + \upsilon_2\right)\right)\, e^{\mathbf{x}_*\beta_1}\,, \qquad (3.3)$$

where the parameter ν_0 is defined in (2.3). Now using this function we set

$$\kappa_0 = \kappa_0(\nu_0) = \frac{1}{108\, r^2 (3\rho^2 + y_0^2 + 2\sigma^2_{\max})} \qquad (3.4)$$

where $\rho = \max\left(|y_0|\,,\, \sigma_{\max}\sqrt{L}\,,\, 2(\mathbf{x}_* + ML)\right)$.

Now for any $\delta > 0$ and any parameter vector $\nu = (\nu_0, \nu_1, \nu_2, \nu_3, \nu_4)$ from $\mathbb{R}^5_+$ we set

$$z_0 = z_0(\delta, \nu) = \delta^{3/2} \max \left(2c^*_1\nu_3 \,,\, 2c^*_2\nu_2 \,,\, \nu_4 T^{1/2} \,,\, \nu_1 T^{-1/2}\right) \,,$$

$$\tau = \tau(\delta, \nu) = \delta^{3/2} \max \left(c^*_1\nu_3 \,,\, c^*_2\nu_2\right) \,, \tag{3.5}$$

where

$$c^*_1 = 2e^{\kappa+1}\sqrt{\frac{R(1+\rho)}{\kappa}} \quad \text{and} \quad c^*_2 = \sqrt{2}e\sigma_{max} \,.$$

The parameters R and κ are defined in Theorem 5.1. Finally we set

$$\gamma = \frac{1}{4\tau} \quad \text{and} \quad \varkappa = \varkappa(\delta, \nu) = \frac{9\kappa_0(1-\delta)}{64\delta} \,. \tag{3.6}$$

Now we set

$$M_1 = M + L\left(\mathbf{x}_* + |x_0| + 2\right) \,. \tag{3.7}$$

Now for any inegrated two times continuously differentiable $\mathbb{R} \to \mathbb{R}$ function Ψ we define

$$\mathbf{k}_*(\Psi) = \max\left(|\dot{\Psi}|_1 \,,\, |\ddot{\Psi}|_1 \,,\, \|\Psi\|_* \,,\, \|\dot{\Psi}\|_* \,,\, \|\ddot{\Psi}\|_*\right) \,. \tag{3.8}$$

Using this operator we define the parameters

$$z^*_0 = \lambda_1 \mathbf{k}_*(\Psi) \quad \text{and} \quad \tau^* = \lambda_2 \mathbf{k}_*(\Psi) \,, \tag{3.9}$$

where

$$\lambda_1 = \max\left(2c^*_1 M_1 \,,\, 2c^*_2 \,,\, M_1 q^* \,,\, 1\right) \quad \text{and} \quad \lambda_2 = \max\left(c^*_1 M_1 \,,\, c^*_2\right) \,.$$

Finally, we set

$$\gamma^* = \frac{1}{4\tau^*} \,. \tag{3.10}$$

4 Continuous observations

In this section we study the deviation in the ergodic theorem for the continuous observation case, which in this case is defined as

$$\Delta_T(\phi) = \frac{1}{\sqrt{T}} \int_0^T \left(\phi(y_t) \;-\; \pi_\vartheta(\phi)\right) \mathrm{d}t \,, \tag{4.1}$$

where ϕ is any integrated function, i.e. $|\phi|_1 < \infty$.

Proposition 4.1. *For any* $\nu_0 > 0$ *and* $\nu_1 > 0$

$$\sup_{z\ge 0} \quad e^{\kappa_0 z^2} \quad \sup_{T\ge 1} \quad \sup_{\phi\in\mathcal{V}_{\nu_0,\nu_1}} \quad \sup_{\vartheta\in\Theta} \quad \mathbf{P}_\vartheta\left(|\Delta_T(\phi)| \ge z\right) \le 2\,, \tag{4.2}$$

where the parameter κ_0 *is given in* (3.4).

Proof. Similarly to [8] firstly we show that the deviation (4.1) has an exponential moment, i.e. we show that for the parameter κ_0

$$\sup_{T\ge 1}\sup_{\vartheta\in\Theta} \mathbf{E}_\vartheta e^{\kappa_0\Delta^2_T(\phi)} \le 2\,. \tag{4.3}$$

Indeed, to show this inequality we need to estimate the expectation of any even power for the deviation $\Delta_T(\phi)$. To this end we have to represent this deviation as the sum of a continuous martingale and a negligible term. For this one needs to find a bounded solution for the following differential equation

$$\dot{v}_\vartheta(u) \,+\, 2\frac{S(u)}{\sigma^2(u)}v_\vartheta(u) \,=\, 2\frac{\widetilde{\phi}(u)}{\sigma^2(u)}, \quad \widetilde{\phi}(u) = \phi(u) - \pi_\vartheta(\phi)\,. \tag{4.4}$$

One can check directly that the function

$$v_\vartheta(u) = -2\int_u^\infty \frac{\widetilde{\phi}(y)}{\sigma^2(y)}\,\exp\{2\int_u^y S_1(z)\mathrm{d}z\}\,\mathrm{d}y \tag{4.5}$$

yields such a solution. We recall that the function S_1 is defined in (2.2). Moreover, due to Lemma A.2 from Appendix implies this function is uniform bounded. By applying the Ito formula to the function $V(y) = \int_0^y v_\vartheta(u)\mathrm{d}u$ we following representation

$$\int_0^T \widetilde{\phi}(y_s)\mathrm{d}s = V(y_T) - V(y_0) - \zeta_T\,, \tag{4.6}$$

where $\zeta_T = \int_0^T v_\vartheta(y_s)\sigma(y_s)\mathrm{d}w_s$. Therefore, for any $T \ge 1$ through Lemma A.2 we can estimate $\Delta_T(\phi)$ from above as

$$|\Delta_T(\phi)| \,\le\, r|y_T| \,+\, r|y_0| \,+\, \frac{1}{\sqrt{T}}\,|\zeta_T|\,.$$

Moreover, taking into account (see [12], Lemma 4.11), that for any $m \ge 1$,

$$\mathbf{E}_\vartheta\,\left(\zeta_T\right)^{2m} \le (2m-1)!!\, r^{2m}\sigma^{2m}_{\max}\,T^m\,,$$

we obtain by Proposition A.1 , that for any $m \ge 1$

$$\mathbf{E}_\vartheta|\Delta_T(\phi)|^{2m} \le 3^{2m-1}\left(r^{2m}(\mathbf{E}_\vartheta|y_T|^{2m} + |y_0|^{2m}) + \frac{\mathbf{E}_\vartheta\,\left(\zeta_T\right)^{2m}}{T^m}\right)$$

$$\le (3\,r)^{2m}\left(4(m+1)(2m-1)!!\,\rho^{2m} + y_0^{2m} + (2m-1)!!\,\sigma^{2m}_{\max}\right)\,.$$

Therefore, taking into account the definition of κ_0, we obtain

$$\mathbf{E}_\vartheta e^{\kappa_0 \Delta_T^2(\phi)} = 1 + \sum_{m=1}^{\infty} \frac{\kappa_0^m}{m!} (3\,r)^{2m} \left(4(2m+1)!!\rho^{2m} + y_0^{2m} + (2m-1)!!\sigma_{\max}^{2m}\right)$$

$$\leq 1 + \sum_{m=1}^{\infty} \kappa_0^m (3\,r)^{2m} \left(4(3\rho^2)^m + y_0^{2m} + 2^m \sigma_{\max}^{2m}\right)$$

$$\leq 1 + \sum_{m=1}^{\infty} (1/2)^m = 2\,.$$

From here we obtain the inequality (4.3) and by the Chebychev inequality we come to the upper bound (4.2). Hence Proposition 4.1. $\square$

Remark 4.1. *It should be noted that the inequality* (4.2) *is shown in [8] for the process* (1.1) *with* $\sigma = 1$. *Thus Proposition 4.1 extends teh result from [8] for any diffusion function* σ.

5 Uniform geometric ergodicity

Here we announce a result on geometric ergodicity obtained in [10].

Theorem 5.1. *There exist some constants* $R \geq 1$ *and* $\kappa > 0$ *such that*

$$\sup_{t\geq 0} e^{\kappa t} \sup_{\|g\|_* \leq 1} \sup_{x\in\mathbb{R}} \sup_{\vartheta\in\Theta} \frac{|\mathbf{E}_\vartheta\,(g(y_t)|y_0 = x) - \pi_\vartheta(g)|}{1+|x|} \leq R\,, \tag{5.1}$$

where the parameters R *and* κ *are given in [10].*

6 Burkhölder's inequality

In this section we give the following inequality from [4],[17].

Proposition 6.1. *Let* $(\Omega, \mathcal{F}, (\mathcal{F}_j)_{1\leq j\leq n}, \mathbf{P})$ *be a filtered probability space and* $(X_j, \mathcal{F}_j)_{1\leq j\leq n}$ *be sequence of random variables such that for some* $p \geq 2$

$$\max_{1\leq j\leq n} \mathbf{E}\,|X_j|^p < \infty\,.$$

Define

$$b_{j,n}(p) = \left(\mathbf{E}\,(|X_j| \sum_{k=j}^{n} |\mathbf{E}\,(X_k|\mathcal{F}_j)|)^{p/2}\right)^{2/p}\,.$$

Then

$$\mathbf{E}\,|\sum_{j=1}^{n} X_j\,|^p \le\, (2p)^{p/2}\left(\sum_{j=1}^{n} b_{j,n}(p)\right)^{p/2}\,. \tag{6.1}$$

Proof of this Proposition is given in Appendix.

7 Proofs

7.1 Proof of Theorem 2.1

First note, that by Proposition A.1 and the Hölder inequality we obtain for any $\alpha \ge 1$

$$\sup_{t\ge 0}\sup_{\vartheta\in\Theta} \mathbf{E}_\vartheta\left(|y_t|^\alpha | y_0 = x\right) \,\le\, 4\,(\alpha+1)^{\alpha/2}\,\rho^\alpha\,. \tag{7.1}$$

Now we represent the deviation $D_T(\phi)$ as

$$\begin{aligned} D_T(\phi) &= \int_0^T (\phi(y_t) - \pi_\vartheta(\phi))\mathrm{d}t + \mathbf{A}_{1,T} \,-\, \mathbf{A}_{2,T} \\ &= \sqrt{T}\,\Delta_T(\phi) + \mathbf{A}_{1,T} \,-\, \mathbf{A}_{2,T}\,, \end{aligned} \tag{7.2}$$

where

$$\mathbf{A}_{1,T} = \sum_{j=1}^{N}\int_{t_{j-1}}^{t_j}\left(\phi(y_{t_j}) - \phi(y_t)\right)\,\mathrm{d}t \quad \text{and} \quad \mathbf{A}_{2,T} = \int_{\delta N}^{T} (\phi(y_t) - \pi_\vartheta(\phi))\mathrm{d}t\,.$$

To estimate the term $\mathbf{A}_{1,T}$ we represent through the Ito formula the difference $\phi(y_{t_j}) - \phi(y_t)$ as

$$\begin{aligned} \phi(y_{t_j}) - \phi(y_t) &= \int_t^{t_j} \mathcal{L}_\vartheta(\phi)(y_s)\,\mathrm{d}s + \int_t^{t_j} \dot\phi(y_s)\sigma(y_s)\mathrm{d}W_s \\ &= \widetilde{\pi}_\vartheta(\phi)(t_j - t) + \Psi_j(t) + \int_t^{t_j} \dot\phi(y_s)\sigma(y_s)\mathrm{d}W_s\,, \end{aligned}$$

where

$$\Psi_j(t) = \int_t^{t_j} \psi(y_s)\,\mathrm{d}s\,, \quad \omega_j(t) = \int_t^{t_j} \dot\phi(y_s)\sigma(y_s)\,\mathrm{d}W_s$$

and $\psi(y) = \mathcal{L}_\vartheta(\phi)(y) - \widetilde{\pi}_\vartheta(\phi)$. Now setting

$$X_j = \int_{t_{j-1}}^{t_j} \Psi_j(t)\,\mathrm{d}t \quad \text{and} \quad \eta_j = \int_{t_{j-1}}^{t_j} \omega_j(t)\mathrm{d}t\,,$$

we obtain

$$\mathbf{A}_{1,T} = \widetilde{\pi}_{\vartheta}(\phi)\frac{N\delta^2}{2} + \sum_{j=1}^{N} X_j + \sum_{j=1}^{N} \eta_j \,. \tag{7.3}$$

To estimate the second term in the right-hand part of (7.3), we make use of the Proposition 6.1.We start with verifying its conditions. Putting $\mathcal{F}_s = \sigma\{y_u\,, 0 \le u \le s\}$, we obtain by Theorem 5.1, that for any $t \ge s$ and for any ϕ from the functional class (2.5)

$$|\mathbf{E}_{\vartheta}\left(\psi(y_t)|\mathcal{F}_s\right)| \le\ \mu(\phi)\, R\, \left(1+|y_s|\right)\, e^{-\kappa(t-s)} \le\ \nu_3\, R\, \left(1+|y_s|\right)\, e^{-\kappa(t-s)}\,.$$

Therefore, for any $k > j$,

$$|\mathbf{E}_{\vartheta}(X_k|\mathcal{F}_{t_j})| \le Re^{\kappa}(1+|y_{t_j}|)\,\nu_3\,\delta^2\, e^{-\kappa\delta(k-j)}\,. \tag{7.4}$$

It should be noted also, that the random variables X_j are bounded, i.e.

$$|X_j| \le \nu_3\delta^2\,.$$

To estimatie the probability tail for the sum $\sum_{j=1}^{n} X_j$ we will use the inequality (6.1). For this we need to estmate the coefficients $b_{j,N}(p)$ for any $p \ge 1$. From here, taking into account that $1 - e^{-\kappa\delta} \ge \kappa\delta e^{-\kappa}$ and that for $p \ge 2$

$$\left(\mathbf{E}_{\vartheta}(1+|y_{t_j}|)^{p/2}\right)^{2/p} \le 1 + \left(\mathbf{E}_{\vartheta}|y_{t_j}|^{p/2}\right)^{2/p}\,,$$

we can estimate the coefficient $b_{j,N}(p)$ as

$$b_{j,N}(p)\ \le\ \frac{1}{\kappa}\, R\, e^{2\kappa}\, \varsigma^2\, \left(1 + (\mathbf{E}|y_{t_j}|^{p/2})^{2/p}\right)\,,$$

where $\varsigma^2 = \nu_3^2\delta^3$. Now the inequality (7.1) yields

$$b_{j,N}(p)\ \le R_1\varsigma^2\sqrt{2+p} \le R_1\varsigma^2\sqrt{2p}\,,$$

where

$$R_1 = \frac{1}{\kappa}\, R\, e^{2\kappa}(1+\rho)\,.$$

Using this in (6.1) we obtain, that for any $p > 2$,

$$\mathbf{E}_{\vartheta}\,|\sum_{k=1}^{N} X_k|^p \le (2p)^{p/2}\, N^{p/2}\, R_1^{p/2}\, \varsigma^p\, (2p)^{p/4}$$

$$\le (2\sqrt{R_1}\,\varsigma)^p\, N^{p/2}\, p^p\,.$$

Therefore, by Chebyshev's inequality

$$\mathbf{P}_\vartheta\left(|\sum_{k=1}^N X_k| \ge z\sqrt{N}\right) \le e^{p\ln(a)+p\ln p}$$

with $a = 2\sqrt{R_1}\varsigma/z$. Minimizing now the right-hand part over $p \ge 2$, we obtain for $z \ge 4e\sqrt{R_1}\nu_3\delta^{3/2}$

$$\mathbf{P}_\vartheta\left(|\sum_{k=1}^N X_k| \ge z\sqrt{N}\right) \le e^{-z/\varsigma_1}\,, \tag{7.5}$$

where $\varsigma_1 = 2e\sqrt{R_1}\varsigma$.

Moreover, note that by the Burkholder-Davis-Gundy inequality, for any $\alpha \ge 1$,

$$\mathbf{E}_\vartheta\,|\omega_j(t)|^\alpha \le (\alpha)^{\alpha/2}\,\nu_2^\alpha\sigma_{\max}^\alpha\,(t_j - t)^{\alpha/2}\,.$$

Using this and the the Hölder inequality, we get

$$\mathbf{E}_\vartheta\,|\eta_j|^\alpha \le \delta^{\alpha-1}\int_{t_{j-1}}^{t_j}\mathbf{E}_\vartheta\,|\omega_j(t)|^\alpha\,\mathrm{d}t \le \delta^{3\alpha/2}\,\alpha^{\alpha/2}\,\nu_2^\alpha\sigma_{\max}^\alpha\,.$$

Note, that in this case in the right hand of the inequality (6.1)

$$b_{j,N} = \left(\mathbf{E}_\vartheta\,|\eta_j|^p\right)^{2/p}\,.$$

Therefore, similarly to the inequality (4.5) we find, that for all $z \ge 2\varsigma_2$,

$$\mathbf{P}_\vartheta\left(|\sum_{k=1}^N \eta_k| \ge z\sqrt{N}\right) \le e^{-z/\varsigma_2}\,, \tag{7.6}$$

where $\varsigma_2 = \sqrt{2}e\delta^{3/2}\nu_2\sigma_{\max}$. Now from (7.3), (4.5)–(4.6) it follows that for $z \ge z_0$

$$\mathbf{P}_\vartheta\left(|\mathbf{A}_{1,T}| \ge z\sqrt{N}\right) \le \mathbf{P}_\vartheta\left(|\sum_{k=1}^N X_k| \ge z\sqrt{N}/4\right) + \mathbf{P}_\vartheta\left(|\sum_{k=1}^N \eta_k| \ge z\sqrt{N}/4\right) \le 2\,e^{-z/4\tau}\,, \tag{7.7}$$

when the parameters z_0 and τ are given in (3.5). Moreover, note that due to (2.5) the last term in (7.2) is bounded, i.e.

$$|\mathbf{A}_{2,T}| \le 2\delta\|\phi\|_* \le 2\delta\nu_2 \le z_0\sqrt{N}/4\,.$$

Finally, from (7.2) for $z \ge z_0$ one has

$$\begin{aligned}\mathbf{P}_\vartheta(|D_T(\phi)| \ge z\sqrt{N}) &\le \mathbf{P}_\vartheta\left(\sqrt{T}|\Delta_T(\phi)| + |\mathbf{A}_{1,T}| \ge 3z\sqrt{N}/4\right)\\ &\le \mathbf{P}_\vartheta\left(\sqrt{T}|\Delta_T(\phi)| \ge 3z\sqrt{N}/8\right)\\ &+ \mathbf{P}_\vartheta\left(|\mathbf{A}_{1,T}| \ge 3z\sqrt{N}/8\right)\,.\end{aligned}$$

Taking into account here, that $N/T \ge (1-\delta)/\delta$ for any $0 < \delta < 1$ and $T \ge 1$, we obtain, that

$$\begin{aligned}\mathbf{P}_\vartheta(|D_T(\phi)| \ge z\sqrt{N}) \le \mathbf{P}_\vartheta&\left(|\Delta_T(\phi)| \ge \frac{3z\sqrt{(1-\delta)}}{8\sqrt{\delta}}\right)\\ &+ \mathbf{P}_\vartheta\left(|\mathbf{A}_{1,T}| \ge \frac{3}{8}z\sqrt{N}\right)\,.\end{aligned}$$

Therefore, applying here the inequalities (4.2) and (7.7) we come to the upper bound (2.6) with the parameter $\varkappa$ given in (3.6). Hence Theorem 2.1. $\square$

7.2 Proof of Theorem 2.2

Firstly, note that in this case

$$|\psi_{h,x_0}|_1 = |\Psi|_1\,, \quad \|\psi_{h,x_0}\|_* = \frac{1}{h}\|\Psi\|_* \quad \text{and} \quad \|\dot{\psi}_{h,x_0}\|_* = \frac{1}{h^2}\|\dot{\Psi}\|_*\,.$$

Moreover, taking into account that $|S(y)| \le M + L\mathbf{x}_* + L|y|$, we find that

$$\sup_{|y|\le|x_0|+2} |S(y)| \le M_1\,, \tag{7.8}$$

where M_1 is given in (3.7).
Therefore, in view of the fact that $0 < h < 1$, we can estimate from above the parametrs (2.4) as

$$\mu(\psi_{h,x_0}) \le \mu_* h^{-3} \quad \text{and} \quad \widetilde{\mu}(\psi_{h,x_0}) \le \widetilde{\mu}_* h^{-2}\,, \tag{7.9}$$

where

$$\mu_* = \max\left(\|\dot{\Psi}\|_*\,,\, \|\ddot{\Psi}\|_*\right) M_1 \quad \text{and} \quad \widetilde{\mu}_* = \max\left(|\dot{\Psi}|_1\,,\, |\ddot{\Psi}|_1\right) M_1 q^*\,.$$

Therefore, the function ψ_{h,x_0} belongs to the class (2.5) with the following parameters

$$\nu_0 = |\Psi|_1\,, \quad \nu_1 = \frac{\|\Psi\|_*}{h}\,, \quad \nu_2 = \frac{\|\dot{\Psi}\|_*}{h^2}\,, \quad \nu_3 = \frac{\mu_*}{h^3}\,, \quad \nu_4 = \frac{\widetilde{\mu}_*}{h^2}\,.$$

Therefore, in this case the coefficient (3.4) equals to $\kappa_0(|\Psi|_1)$ and the parameters (3.5) can be represented as

$$z_0 = \frac{\delta^{3/2}}{h^3}\,\max\left(2c_1^*\mu_*\,,\,2c_2^*\|\dot{\Psi}\|_* h\,,\,\widetilde{\mu}_* h T^{1/2}\,,\,\|\Psi\|_* h^2 T^{-1/2}\right)$$

$$\tau = \frac{\delta^{3/2}}{h^3}\,\max\left(c_1^*\mu_*\,,\,c_2^*\|\dot{\Psi}\|_* h\right)\,. \tag{7.10}$$

Therefore, thanks to the condition (2.8) for any $T^{-1/2} \le h \le 1$

$$z_0 \le l_T^{-3/2}\,z_0^* \quad \text{and} \quad \tau \le l_T^{-3/2}\,\tau^*\,, \tag{7.11}$$

where the parameters z_0^* and τ^* are given in (3.9). Note now that, by the condition (2.7)

$$\mathbf{P}_\vartheta\left(|D_T(\psi_{h,x_0})|\,\ge\,a\,T\right) \le \mathbf{P}_\vartheta\left(|D_T(\psi_{h,x_0})|\,\ge\,z_1\sqrt{N}\right)$$

where $z_1 = a/\sqrt{l_T}$. The first inequality in (7.11) implies that $z_1 \ge z_0$ for all $a \ge a_* = z_0^*/l_T$. Moreover, from the last inequality in (7.11) it follows, that for $a \ge a_*$

$$\min\left(\varkappa z_1\,,\,\gamma\right) = \min\left(\varkappa z_1\,,\,\frac{1}{4\tau}\right) \ge \min\left(\varkappa\frac{z_0^*}{l_T\sqrt{l_T}}\,,\,\frac{l_T\sqrt{l_T}}{4\tau^*}\right)\,.$$

Taking into account here the definition of $\varkappa$ in (3.6) and the form for δ given by (2.7) we obtain that for sufficiently large T

$$\min\left(\varkappa\frac{z_0^*}{l_T\sqrt{l_T}}\,,\,\frac{l_T\sqrt{l_T}}{4\tau^*}\right) = \frac{l_T\sqrt{l_T}}{4\tau^*}\,.$$

Thus, through Theorem 2.1 we come to the inequality (2.10). Hence Theorem 2.2 $\square$

7.3 Proof of Theorem 2.3

First we represent the tail probability as

$$\mathbf{P}_\vartheta\left(|D_T(\chi_{h,x_0})|\,\ge\,\epsilon_T\,T\right) = \mathbf{I}_1 + \mathbf{I}_2\,,$$

where

$$\mathbf{I}_1 = \mathbf{P}_\vartheta\left(\sum_{j=1}^N \chi_{h,x_0}(y_{t_j})\Delta t_j \le \left(\pi_\vartheta(\chi_{h,x_0}) - \epsilon_T\right)T\right)$$

and

$$\mathbf{I}_2 = \mathbf{P}_\vartheta \left(\sum_{j=1}^{N} \chi_{h,x_0}(y_{t_j}) \Delta t_j \geq (\pi_\vartheta(\chi_{h,x_0}) + \epsilon_T) \, T \right) .$$

Let us define now the following smoothing indicator functions

$$\Psi_{1,\eta}(u) = \frac{1}{\eta} \int_{-\infty}^{+\infty} \mathbf{1}_{\{|z| \leq 1-\eta\}} \, V\left(\frac{z-u}{\eta}\right) \mathrm{d}z$$

and

$$\Psi_{2,\eta}(u) = \frac{1}{\eta} \int_{-\infty}^{+\infty} \mathbf{1}_{\{|z| \leq 1+\eta\}} \, V\left(\frac{z-u}{\eta}\right) \mathrm{d}z \, ,$$

where η is a smoothing positive parameter which will be specified later, V is a two times continuously differentiable even $\mathbb{R} \to \mathbb{R}$ function such that $V(z) = 0$ for $|z| \geq 1$ and

$$\int_{-1}^{1} V(z) \mathrm{d}z = 1 \, .$$

It is easy to see that, for any $y \in \mathbb{R}$ and $0 < \eta \leq 1/2$,

$$\Psi_{1,\eta}(u)(y) \leq \chi(y) \leq \Psi_{2,\eta}(y)$$

and $\Psi_{2,\eta}(y) = 0$ for $|y| \geq 2$. Moreover, for the functions

$$\psi_{i,h}(y) = \frac{1}{h} \, \Psi_{i,\eta} \left(\frac{y - z_0}{h} \right)$$

using the inequality (A.4), we can estimate the difference between the coore-sponding ergodic intergals (1.4) as

$$|\pi_\vartheta(\chi_{h,x_0}) - \pi_\vartheta(\psi_{i,h})| \leq 4\eta q^* \, .$$

Therefore, choosing here $\eta = \epsilon_T^2$ we obtain, for sufficiently large T,

$$\mathbf{I}_i \leq \mathbf{P}_\vartheta \left(|D_T(\phi_{i,h})| \geq \, \epsilon_T \, T/2 \right) .$$

One can check directly that in this case the operator (3.8) has the following asymptotic ($T \to \infty$) form

$$\mathbf{k}_*(\Psi_{i,\eta}) = \mathrm{O}\left(\eta^{-2}\right) .$$

Therefore, from (3.9) and (7.11) it follows that for $T \to \infty$ and $h \geq T^{-1/2}$

$$z_0(\phi_{i,h}) = \mathrm{O}\left(\eta^{-2} l_T^{-3/2}\right) \quad \text{and} \quad \tau(\phi_{i,h}) = \mathrm{O}\left(\eta^{-2} l_T^{-3/2}\right) ,$$

i.e.

$$z_0(\phi_{i,h}) = \mathrm{O}\left(\frac{1}{\epsilon_T^4 l_T^{3/2}}\right) \quad \text{and} \quad \tau(\phi_{i,h}) = \mathrm{O}\left(\frac{1}{\epsilon_T^4 l_T^{3/2}}\right).$$

Now we have

$$\mathbf{P}_\vartheta\left(|D_T(\psi_{h,x_0})| \geq \epsilon_T T\right) \leq \mathbf{P}_\vartheta\left(|D_T(\psi_{h,x_0})| \geq z_1\sqrt{N}\right),$$

where $z_1 = \epsilon_T/\sqrt{l_T}$. The last equality in (2.9) implies $z_1 \geq z_0$ for sufficiently large T. Moreover, taking into account, that there exists a constant $\mathbf{c}_* > 0$ such that for sufficiently large T

$$\varkappa z_1 \geq \mathbf{c}_* T\sqrt{l_T}\epsilon_T \quad \text{and} \quad \gamma \geq \mathbf{c}_* l_T\sqrt{l_T}\epsilon_T^4,$$

i.e. for sufficiently large T

$$\min\left(\varkappa z_1\,,\,\gamma\right) \geq c_* l_T\sqrt{l_T}\epsilon_T^4.$$

Therefore, by Theorem 2.1 for sufficiently large T

$$\mathbf{P}_\vartheta\left(|D_T(\psi_{h,x_0})| \geq \epsilon_T T\right) \leq 4e^{-\mathbf{c}_* l_T\epsilon_T^5}.$$

Now the last condition in (2.9) yields the equality (2.12). Hence Theorem 2.3.
□

A Appendix

A.1 Proof of Proposition 6.1

We set

$$h_n(t) = \mathbf{E}|S_{n-1} + tX_n|^p \quad \text{with} \quad S_n = \sum_{j=1}^n X_j.$$

By the induction method we assume that for any $1 \leq k \leq n-1$ and $0 \leq t \leq 1$

$$h_k(t) \leq (2p)^{p/2}\, B_k^{p/2}(t)\,, \tag{A.1}$$

where

$$B_k(t) = \sum_{j=1}^{k-1} b_{j,k}(p) + t b_{k,k}(p).$$

Note now that as is shown in [17] (Theorem 2.3)

$$\mathbf{E}|S_n|^p = p(p-1)\sum_{j=1}^n \int_0^1 \mathbf{E}|S_{j-1} + vX_j|^{p-2}(-vX2_j + \Upsilon(j,n))\mathrm{d}v. \tag{A.2}$$

with

$$\Upsilon(j,n) = X_j \sum_{k=j}^{n} \mathbf{E}(X_k|\mathcal{F}_j)\,.$$

Therefore,

$$h_n(t) = p(p-1)\sum_{j=1}^{n-1}\int_0^1 \mathbf{E}|S_{j-1}+vX_j|^{p-2}(-vX2_j + G(i,n,t))\mathrm{d}v$$

$$+\,p(p-1)\int_0^1 \mathbf{E}|S_{n-1}+vtX_n|^{p-2}t^2(1-v)X_n^2\mathrm{d}v\,,$$

where

$$G(j,n,t) = \Upsilon(j,n-1) + tX_j\,\mathbf{E}(X_n|\mathcal{F}_j)\,.$$

Moreover, we can estimate $h_n(t)$ as

$$\frac{h_n(t)}{p2} \le \sum_{j=1}^{n-1}\int_0^1 \mathbf{E}|S_{j-1}+vX_j|^{p-2}\,|G(i,n,t)|\mathrm{d}v$$

$$+\int_0^t \mathbf{E}|S_{n-1}+sX_n|^{p-2}\,X_n^2\mathrm{d}s$$

Now taking into account that for $0\le t\le 1$

$$\left(\mathbf{E}|G(j,n,t)|^{p/2}\right)^{2/p} \;\le\; b_{j,n}(p)\,,$$

we obtain by the Hölder inequality

$$\int_0^1 \mathbf{E}|S_{j-1}+vX_j|^{p-2}\,|G(i,n,t)|\,\mathrm{d}v \;\le\; \int_0^1 h_j^{\alpha}(v)\,b_{j,n}(p)\mathrm{d}v\,,$$

where $\alpha = 1-2/p$. Therefore,

$$\frac{h_n(t)}{p2} \le \sum_{j=1}^{n-1} b_{j,n}(p)\int_0^1 h_j^{\alpha}(v)\mathrm{d}v + b_{n,n}(p)\int_0^t h_n^{\alpha}(s)\mathrm{d}s$$

Now by the induction assumption for any $1\le j\le n-1$

$$b_{j,n}(p)\int_0^1 h_j^{\alpha}(v)\mathrm{d}v \;\le\; (2p)^{(p-2)/2}\int_0^1 B_j^{(p-2)/2}(v)\,\mathrm{d}v\,b_{j,n}(p)\,.$$

Moreover, taking into account that

$$B_j(v) \le \sum_{i=1}^{j-1} b_{i,n} + vb_{j,n}(p)\,,$$

we obtain that

$$\int_0^1 B_j^{(p-2)/2}(v)\,\mathrm{d}v\, b_{j,n}(p) \;\leq\; \frac{2}{p}\left((\sum_{i=1}^{j} b_{i,n})^{p/2} - (\sum_{i=1}^{j-1} b_{i,n})^{p/2} \right).$$

This implies for any $0 \leq t \leq 1$

$$h_n(t) \leq k_n \int_0^t h_n^\alpha(v)\,\mathrm{d}v + f_n \tag{A.3}$$

with

$$k_n = p^2 b_{n,n}(p) \quad \text{and} \quad f_n = \left(2p \sum_{j=1}^{n-1} b_{j,n}(p) \right)^{p/2}.$$

Now by setting

$$Z(t) = \int_0^t h_n^\alpha(s)\mathrm{d}s + \frac{f_n}{k_n}\,,$$

we obtain from (A.3) that

$$\dot{Z}(t) \leq k_n^\alpha Z^\alpha(t)\,.$$

Now introducing

$$g(t) = \dot{Z}(t) - k_n^\alpha Z^\alpha(t)\,,$$

we obtain the differential equation

$$\dot{Z}(t) = k_n^\alpha Z^\alpha(t) + g(t)$$

with $g(t) \leq 0$. From here we obtain

$$Z^{2/p}(t) = Z^{2/p}(0) + \frac{2}{p} k_n^\alpha t + \int_O^t \frac{g(u)}{Z^\alpha(u)}\,\mathrm{d}u \leq\; Z^{2/p}(0) + \frac{2}{p} k_n^\alpha t\,,$$

i.e.

$$Z(t) \leq \left(Z^{2/p}(0) + \frac{2}{p} k_n^\alpha t \right)^{p/2}.$$

Substituting this bound in (A.3) we obtain

$$h_n(t) \leq k_n Z(t) \leq k_n \left(Z^{2/p}(0) + \frac{2}{p} k_n^\alpha t \right)^{p/2}$$

$$= \left(2p \sum_{j=1}^{n-1} b_{j,n}(p) + 2pt b_{n,n}(p) \right)^{p/2}.$$

Hence Proposition 6.1. $\square$

A.2 Uniform bound for the invariant density

Lemma A.1. *The invariant density* (2.2) *is uniformly bounded:*

$$\sup_{x\in\mathbb{R}}\sup_{\vartheta\in\Theta} q_\vartheta(x) \,\le\, q^* < \infty\,, \tag{A.4}$$

where the upper bound q^* is given in (3.2)*.*

Proof. First, note that through the definition of Θ we can check directly that for any $|x|\ge x_*$

$$2\int_0^x S_1(v)\,d\,v \le \beta_1|x| - \beta_2(|x|-\mathbf{x}_*)^2\,, \tag{A.5}$$

where the coefficients β_1 and β_2 are given in (3.1). Therefore, taking into account, that for $|x|\ge x_*$

$$2\int_0^x S_1(v)\,d\,v \le \beta_1\,x_*\,,$$

we obtain that

$$2\sup_{x\in\mathbb{R}}\int_0^x S_1(v)\,d\,v \le \beta_1 x_* + \frac{\beta_1}{4\beta_2}\,.$$

Estimating now the denominator in (2.2) from below as

$$\int_{\mathbb{R}}\sigma^{-2}(z)\,e^{\widetilde{S}(z)}\mathrm{d}z \,\ge\, \int_0^1 \sigma^{-2}(z)\,\mathrm{d}z \ge \frac{1}{\sigma^2_{max}}\,,$$

and taking into account the definition of q^* we come to the upper uniform bound (A.4). Hence Proposition A.1. $\square$

A.3 Moment bound for the process y_t.

Proposition A.1. *For any $m\ge 1$*

$$\sup_{t\ge 0}\sup_{\vartheta\in\Theta}\mathbf{E}_\vartheta|y_t|^{2m} \,\le\, 4(m+1)(2m-1)!!\,\rho^{2m} \le 4(2m)^m\,\rho^{2m}\,,$$

where ρ is given in (3.4)*.*

Proof. First note, that through the Ito formula we can write for the function $z_t(m)=\mathbf{E}_\vartheta y_t^{2m}$ the following intergal equality

$$z_t(m) = z_0(m) + 2m\int_0^t \mathbf{E}_\vartheta y_s^{2m-1}\,S(y_s)\mathrm{d}s$$

$$+\, m(2m-1)\int_0^t \mathbf{E}_\vartheta y_s^{2m-2}\sigma^2(y_s)ds\,,$$

which can be rewritten as the differential equality

$$\dot{z}_t(m) = 2m\mathbf{E}_\vartheta y_s^{2m-1}\, S(y_t) + m(2m-1)\mathbf{E}_\vartheta y_t^{2m-2}\sigma^2(y_t)\,.$$

Taking into account here that $\sup_{x\in\mathbb{R}}\sigma^2(x) \le \sigma^2_{\max}$ we obtain, that for any $m \ge 1$ and $t \ge 0$

$$\dot{z}_t(m) \le 2m\mathbf{E}_\vartheta y_t^{2m-1} S(y_t) + m(2m-1)\sigma^2_{\max} z_t(m-1)\,.$$

Now we need to estimate from above the function $x^{2m-1}S(x)$. Obviously, that for any $K > \mathbf{x}_*$

$$x^{2m-1}S(x) \le K^{2m-1} \sup_{|x|\le K} |S(x)| \mathbf{1}_{\{|x|\le K\}} \,+\, x^{2m}\frac{S(x)}{x}\mathbf{1}_{\{|x|>K\}}\,.$$

Taking into account that $\sup_{|x|>\mathbf{x}_*} |\dot{S}(x)| \le L$, we obtain, for any $x \in [\mathbf{x}_*, K]$,

$$|S(x)| \le |S(\mathbf{x}_*)| + L|x - \mathbf{x}_*| \le M + L(K - \mathbf{x}_*)\,.$$

Similarly, we obtain the same upper bound for $x \in [-K, -\mathbf{x}_*]$. Therefore,

$$\sup_{|x|\le K} |S(x)| \le M \,+\, L\,(K - \mathbf{x}_*).$$

Consider now the case $|x| > K$. We recall, that $\sup_{|x|\ge \mathbf{x}_*} \dot{S}(x) \le -L^{-1}$. Therefore,

$$\frac{S(x)}{x} \le \frac{M}{K} - \frac{K - \mathbf{x}_*}{LK}\,.$$

Choosing $K = 2(\mathbf{x}_* + ML)$ yields

$$\frac{S(x)}{x} \le -\frac{1}{2L}\,.$$

Therefore,

$$\begin{aligned} x^{2m-1}S(x) &\le K^{2m-1}\left(M + L(K - \mathbf{x}_*)\right) - \frac{1}{2L}x^{2m}\mathbf{1}_{\{|x|>K\}} \\ &= K^{2m-1}\left(M + L(K - \mathbf{x}_*)\right) + \frac{\beta}{2}x^{2m}\mathbf{1}_{\{|x|\le K\}} \,-\, \frac{1}{2L}x^{2m} \\ &\le \mathbf{A}_m - \frac{\beta}{2}x^{2m}\,, \end{aligned}$$

where

$$\mathbf{A}_m = \left(2(\mathbf{x}_* + ML)\right)^{2m-1}\left(2M + \mathbf{x}_*\left(L + L^{-1}\right) + 2L^2 M\right)$$

From here it follows, that

$$\dot{z}_t(m) \le 2m\, \mathbf{A}_m - L^{-1}\, m\, z_t(m) + m(2m-1)\sigma^2_{\max}\, z_t(m-1)\,.$$

We can rewrite this inequality as follows

$$\dot{z}_t(m) = -L^{-1} m z_t(m) \,+\, m(2m-1)\sigma_1^2 z_t(m-1) \,+\, \psi_t\,,$$

where $\sup_{t\ge 0}\psi_t \le 2m\,\mathbf{A}_m$. This equality provides

$$\begin{aligned} z_t(m) &= z_0(m)e^{-mL^{-1}\,t} \,+\, m(2m-1)\sigma^2_{\max}\int_0^t e^{-mL^{-1}(t-s)} z_s(m-1)\mathrm{d}s \\ &\quad + \int_0^t e^{-mL^{-1}(t-s)}\psi_s \mathrm{d}s \\ &\le m(2m-1)\sigma^2_{\max}\int_0^t e^{-mL^{-1}(t-s)} z_s(m-1)\mathrm{d}s \,+ \mathbf{B}_m\,, \end{aligned}$$

where $\mathbf{B}_m = y_0^{2m} + 2\mathbf{A}_m L$. Setting $\mathbf{B}_0 = 1$ and resolving this inequality by recurrence yields

$$z_t(m) \le 4\,(2m-1)!!\sum_{j=0}^m \left(\sigma^2_{\max}L\right)^{m-j}\mathbf{B}_j\,.$$

It is easy to see, that

$$\mathbf{B}_m \le 4\left(\max\left(|y_0|^2, 4(\mathbf{x}_* + ML)^2\right)\right)^m\,.$$

Therefore

$$\sup_{t\ge 0} z_t(m) \le 4(m+1)(2m-1)!!\,\rho^{2m} \le 4(2m)^m\,\rho^{2m}\,,$$

where ρ is defined in (3.4). Hence Proposition A.1. □

A.4 Properties of the function (4.5)

Lemma A.2. *For any integrated function ϕ the solution* (4.5) *is uniform bounded, i.e.*

$$\sup_{\vartheta\in\Theta}\sup_{y\in\mathbb{R}} |v_\vartheta(y)| \le r\,,$$

where the upper bound r is introduced in (3.3).

Proof. Firstly we note, that for any ϑ from Θ and any intergated $\mathbb{R} \to \mathbb{R}$ function ϕ

$$|\pi_\vartheta(\phi)| \le q^* \, |\phi|_1 \,.$$

Moreover, by the definition of the parameter β_1 we get

$$2 \sup_{|u|\le \mathbf{x}_*} |S_1(u)| \le \beta_1 \,.$$

Therefore, for $0 \le u \le \mathbf{x}_*$ we can estimate the function υ_ϑ as

$$|\upsilon_\vartheta(u)| \le \frac{2e^{\mathbf{x}_*\beta_1}}{\sigma^2_{\min}} \left((1+q^*\mathbf{x}_*)\,|\phi|_1 + \mathbf{I}(\phi)\right) \,,$$

where β_1 is given in (3.1) and

$$\mathbf{I}(\phi) = \int_{\mathbf{x}_*}^{\infty} \left(|\phi(y)| + q^* \, |\phi|_1\right) \, e^{2\int_{\mathbf{x}_*}^{y} S_1(z)\mathrm{d}z} \, \mathrm{d}y \,.$$

To estimate this term note that similarly to (A.5) we can obtain that for any $y \ge a \ge \mathbf{x}_*$

$$2 \int_a^y S_1(z)\mathrm{d}z \le \beta_1(y-a) - \beta_2(y-a)^2 \,. \tag{A.6}$$

Using this inequlity for $a = \mathbf{x}_*$, we get

$$\begin{aligned}\mathbf{I}(\phi) &\le \int_{\mathbf{x}_*}^{\infty} |\phi(y)| e^{\beta_1(y-\mathbf{x}_*)-\beta_2(y-\mathbf{x}_*)^2} \, \mathrm{d}y + q^* \, |\phi|_1 \int_0^{\infty} e^{\beta_1 z - \beta_2 z^2} \, \mathrm{d}z \\ &\le |\phi|_1 \sup_{z\ge 0} e^{\beta_1 z - \beta_2 z^2} + q^* \, |\phi|_1 \int_0^{\infty} e^{\beta_1 z - \beta_2 z^2} \, \mathrm{d}z \\ &\le |\phi|_1 \left(\upsilon_1 + q^* \upsilon_2\right) \,,\end{aligned}$$

where the parameters υ_1 and υ_2 are introduced in (3.1). Therefore, taking into account the definition (3.3), the last inequality implies

$$\sup_{\vartheta\in\Theta} \sup_{0\le u\le \mathbf{x}_*} |\upsilon_\vartheta(u)| \le r \,. \tag{A.7}$$

If $u \ge \mathbf{x}_*$, then through the inequality (A.6) we estimate the function $\upsilon_\vartheta(u)$ from above as

$$\sup_{\vartheta\in\Theta} \sup_{u\ge \mathbf{x}_*} |\upsilon_\vartheta(u)| \le \frac{2|\phi|_1}{\sigma^2_{\min}} \left(\upsilon_1 + q^* \upsilon_2\right) \le r \,.$$

Let now $u \le 0$. Taking into account that

$$\int_R \frac{\widetilde{\phi}(y)}{\sigma^2(y)} \, \exp\{2 \int_0^y S_1(z)\,\mathrm{d}z\} \, \mathrm{d}y = 0 \,,$$

we can represent the function v_ϑ as

$$v_\vartheta(u) \,=\, 2\int_{|u|}^{\infty} \frac{\widetilde{\phi}(-y)}{\sigma^2(-y)}\, e^{-2\int_{|u|}^{y} S_1(-z)\,\mathrm{d}z}\,\mathrm{d}y\,.$$

Similarly to (A.6), one can check directly, that for any $y \ge a \ge \mathbf{x}_*$

$$-2\int_a^y S_1(-z)\,\mathrm{d}z \le \beta_1(y-a) - \beta_2(y-a)^2\,.$$

Therefore, by the same way as in the proof of (A.7) we can estimate the function $v_\vartheta(u)$ as

$$\sup_{\vartheta\in\Theta}\sup_{u\le 0}|v_\vartheta(u)| \le\, r\,.$$

Hence Lemma A.2. $\square$

References

[1] P. Bertail, S. Clémencon, Sharp bounds for the tails of functionals of Markov chains. *Theor. Veroyatnost i Primenen* **54** N3,(2009), 609-619.

[2] S. Boucheron, G. Lugosi, P. Massart, Concentration inequalities using the entropy method. *The Ann. Probab.* **31**(2003) 1583-1614.

[3] D.A. Cattiaux, A. Guillin, Deviation bounds for additive functionals of Markov process. *ESAIM PS*, Vol 12,(2008) p. 12-29.

[4] J. Dedecker, P. Doukhan, A new covariance inequality and applications. *Stochastic Proc. Appl.* **106**(2003) 63-80.

[5] J. Dedecker, C. Prieur, New dependence coefficients. Examples and applications to statistics. *Probab. Theory and Related Fields* **132**(2005) 203-236.

[6] L. Galtchouk, S. Pergamenshchikov, Sequential nonparametric adaptive estimation of the drift coefficient in diffusion processes. *Mathematical Methods of Statistics* **10**(2001) 316-330.

[7] L. Galtchouk, S. Pergamenshchikov, Asymptotically efficient sequential kernel estimates of the drift coefficient in ergodic diffusion processes. *Statistic Inferences for Stochastic Processes* **9**(2006) 1-16.

[8] L. Galtchouk, S. Pergamenshchikov, Uniform concentration inequality for ergodic diffusion processes. *Stochastic Processes and their applications* **117**(2007) 830-839.

[9] L. Galtchouk, S. Pergamenshchikov, Adaptive sequential estimation for ergodic diffusion processes in quadratic metric. *Journal of Nonparametric Statistics* **23**(2) (2011) 255-285.

[10] L. Galtchouk, S. Pergamenshchikov, Geometric ergodicity for families of homogeneous Markov chains. *http://hal.archives-ouvertes.fr/hal-00455976/fr* (2011)

[11] I.I. Gihman, A.V. Skorohod, *Stochastic differential equations.* Springer, New York, 1972.

[12] R.Sh. Liptser, A.N. Shiryaev, *Statistics of random processes, I,* Springer, New York, 1977.

[13] P. Massart, Some applications of concentration inequalities to statistics. *Ann. Fac. Sci. Toulouse Math.* **9**(2000) 245-303.

[14] V. Maume-Deschamps, Concentration inequalities and estimation of conditional probabilities. *Université de Bourgogne*, Décembre 2004, Prépublication n. 396.

[15] S. Meyn, R. Tweedie, Markov Chains and Stochastic Stability. Springer Verlag, 1993.

[16] S.M. Pergamenshchikov, On large deviation probabilities in ergodic theorem for singularly perturbed stochastic systems. *Weierstrass Institut für Angewandte Analysis und Stochastik, Berlin*, 1998, Preprint n. 414.

[17] Rio, E. (2000) *Théorie asymptotique des processus faiblement dépendants*. In Collection : Mathématiques & Applications, **31**, Springer, Berlin.

[18] A.Yu. Veretennikov, On large deviations for diffusion processes with measurable coefficients. *Uspekhi Mat. Nauk*, **50**(1995) 135-146.