Statistics and Probability Letters 123 (2017) 45–55

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Large deviations for the Ornstein–Uhlenbeck process without tears

Bernard Bercu *, Adrien Richou

Université de Bordeaux, Institut de Mathématiques de Bordeaux, UMR 5251, 351 Cours de la Libération, 33405 Talence cedex, France

ARTICLE INFO

Article history:
Received 8 September 2016
Received in revised form 30 November 2016
Accepted 30 November 2016
Available online 13 December 2016

Keywords:
Ornstein–Uhlenbeck process
Maximum likelihood estimates
Large deviations

ABSTRACT

Our goal is to establish large deviations for the maximum likelihood estimator of the drift parameter of the Ornstein–Uhlenbeck process without tears. We propose a new strategy to establish large deviation results which allows us, via a suitable transformation, to circumvent the classical difficulty of non-steepness. Our approach holds in the stable case where the process is positive recurrent as well as in the unstable and explosive cases where the process is respectively null recurrent and transient. It can also be successfully implemented for more complex diffusion processes.

1. Introduction

Consider the Ornstein–Uhlenbeck process observed over the time interval $[0, T]$

$$dX_t = \theta X_t dt + dB_t \tag{1.1}$$

where (B_t) is a standard Brownian motion and the drift θ is an unknown real parameter. For the sake of simplicity, we assume that the initial state $X_0 = 0$. The process is said to be stable if $\theta < 0$, unstable if $\theta = 0$, and explosive if $\theta > 0$. The maximum likelihood estimator of θ is given by

$$\widehat{\theta}_T = \frac{\int_0^T X_t dX_t}{\int_0^T X_t^2 dt} = \frac{X_T^2 - T}{2\int_0^T X_t^2 dt}. \tag{1.2}$$

It is well-known that in the stable, unstable, and explosive cases

$$\lim_{T\to\infty} \widehat{\theta}_T = \theta \quad \text{a.s.}$$

The purpose of this paper is to establish large deviation principles (LDP) for $(\widehat{\theta}_T)$ via fairly easy to handle arguments. In the stable case, Florens-Landais and Pham (1999) proved an LDP for the score function defined, for all $c \in \mathbb{R}$, by

$$\int_0^T X_t dX_t - c\int_0^T X_t^2 dt.$$

* Corresponding author.
E-mail addresses: bernard.bercu@u-bordeaux.fr (B. Bercu), adrien.richou@u-bordeaux.fr (A. Richou).

http://dx.doi.org/10.1016/j.spl.2016.11.030
0167-7152/

Then, they were able to deduce, by contraction, the LDP for $(\widehat{\theta}_T)$. However, one can realize in Lemma 4.3 of Florens-Landais and Pham (1999) that the normalized cumulant generating function of the score function is quite complicated to compute. Moreover, its LDP relies on a sophisticated time varying change of probability.

In the unstable and explosive cases (Bercu et al., 2012), the strategy for proving an LDP for $(\widehat{\theta}_T)$ is also far from being obvious. As a matter of fact, on can observe in Lemma 2.1 of Bercu et al. (2012) that the normalized cumulant generating function is also very complicated to evaluate. Moreover, as the limiting cumulant generating function is not steep, it is also necessary to make use of a sophisticated time varying change of probability.

Our approach is totally different. It will allows us, via a suitable transformation, to circumvent the classical difficulty of non-steepness. The starting point is to establish, thanks to Gärtner–Ellis's theorem Dembo and Zeitouni (1998), an LDP for the couple

$$V_T = \Big(\frac{X_T}{\sqrt{T}}, \frac{S_T}{T}\Big) \tag{1.3}$$

where the energy S_T is given by

$$S_T = \int_0^T X_t^2 dt.$$

Then, we will obtain the LDP for $(\widehat{\theta}_T)$ by a direct use of the contraction principle. We refer the reader to Bercu and Richou (2015) where our approach was already implemented for the stable Ornstein–Uhlenbeck process with shift. We also wish to stress that our strategy could be successfully extended to more complex diffusions such as the Pearson diffusion (Forman and Sorensen, 2008)

$$dX_t = (a + bX_t)dt + \sqrt{\alpha X_t^2 + \beta X_t + \gamma}\, dB_t$$

where α, β and γ are chosen such that the square root is well defined for any X_t in the state space. In particular, our approach could be extended to the Jacobi diffusion (Alfonsi, 2015; Demni and Zani, 2009; Zhao and Gao, 2010)

$$dX_t = (a + bX_t)dt + 2\sqrt{1 - X_t^2}\, dB_t$$

where $a \geq 4 + b$ and $a + b \leq -4$, as well as to the Wright–Fisher diffusion (Alfonsi, 2015)

$$dX_t = (a + bX_t)dt + 2\sqrt{X_t(1 - X_t)}\, dB_t$$

where $a \geq 2$ and $a + b \leq -2$. Furthermore, LDP for the estimators of the unknown parameters of the Cox–Ingersoll–Ross diffusion

$$dX_t = (a + bX_t)dt + 2\sqrt{X_t} dB_t$$

where $a > 2$ and $b < 0$ can be found in Du Roy de Chaumaray (2016) and Zani (2002). It still remains to investigate the explosive case $b > 0$.

The paper is organized as follows. In Section 2, we establish an LDP for the couple given by (1.3) and we deduce by contraction the LDP for $(\widehat{\theta}_T)$ in the stable, unstable, and explosive cases. Standard tools for proving LDP such as the Gärtner–Ellis theorem and the contraction principle are recalled in Appendix A, while all technical proofs of Section 2 are postponed to Appendix B.

2. Large deviations

The usual notions of full and weak LDP are as follows.

Definition 2.1. A sequence of random vectors (V_T) of $\mathbb{R}^d$ satisfies an LDP with speed T and rate function I if I is a lower semicontinuous function from $\mathbb{R}^d$ to $[0, +\infty]$ such that,

(i) Upper bound: For any closed set $F \subset \mathbb{R}^d$,

$$\limsup_{T\to\infty} \frac{1}{T} \log \mathbb{P}\big(V_T \in F\big) \leq -\inf_{x\in F} I(x). \tag{2.1}$$

(ii) Lower bound: For any open set $G \subset \mathbb{R}^d$,

$$-\inf_{x\in G} I(x) \leq \liminf_{n\to\infty} \frac{1}{T} \log \mathbb{P}\big(V_T \in G\big). \tag{2.2}$$

Moreover, I is said to be a good rate function if its level sets are compact.

Definition 2.2. A sequence of random vectors (V_T) of $\mathbb{R}^d$ satisfies a weak LDP with speed T and rate function I if I is a lower semicontinuous function from $\mathbb{R}^d$ to $[0, +\infty]$ such that the upper bound (2.1) holds for any compact set, while the lower bound (2.2) is true for any open set.

It is well-known that if (V_T) is exponentially tight and satisfies a weak LDP, then I is a good rate function and the full LDP holds for (V_T), see Lemma 1.2.18 of Dembo and Zeitouni (1998).

2.1. The stable case

First of all, we focus our attention on the easy to handle stable case where the parameter θ is negative in (1.1).

Theorem 2.1. *The couple (V_T), given by (1.3), satisfies an LDP with speed T and good rate function $\mathcal{I}_\theta$ given by*

$$\mathcal{I}_\theta(x,y)=\begin{cases}\dfrac{\theta(1-x^2+\theta y)}{2}+\dfrac{(1+x^2)^2}{8y} & \text{if} \quad y>0,\\ +\infty & \text{if} \quad y\le 0.\end{cases} \tag{2.3}$$

We clearly deduce from (1.2) that

$$\widehat{\theta}_T=f(V_T) \tag{2.4}$$

where f is the continuous function defined, for all $x\in\mathbb{R}$ and for any positive y, by

$$f(x,y)=\frac{x^2-1}{2y}.$$

Hence, an elementary application of the contraction principle given in Appendix A, leads to the following corollary, which was previously established in Florens-Landais and Pham (1999) via a much more complicated strategy, see also Bercu and Rouault (2002).

Corollary 2.1. *The sequence $(\widehat{\theta}_T)$ satisfies an LDP with good rate function*

$$I_\theta(z)=\begin{cases}-\dfrac{(z-\theta)^2}{4z} & \text{if} \quad z\le\dfrac{\theta}{3},\\ 2z-\theta & \text{if} \quad z\ge\dfrac{\theta}{3}.\end{cases} \tag{2.5}$$

Proof. The proofs are given in Appendix B. □

2.2. The unstable case

Hereafter, we carry out our strategy on the unstable case where the parameter $\theta=0$ in (1.1).

Theorem 2.2. *The couple (V_T), given by (1.3), satisfies a weak LDP with speed T and good rate function $\mathcal{I}_0$ given by*

$$\mathcal{I}_0(x,y)=\begin{cases}\dfrac{(1+x^2)^2}{8y} & \text{if} \quad y>0,\\ +\infty & \text{if} \quad y\le 0.\end{cases} \tag{2.6}$$

Despite the lack of exponential tightness, it is possible to establish the following corollary, which was previously proved in Bercu et al. (2012) via a much more complex procedure.

Corollary 2.2. *The sequence $(\widehat{\theta}_T)$ satisfies an LDP with good rate function*

$$I_0(z)=\begin{cases}-\dfrac{z}{4} & \text{if} \quad z\le 0,\\ 2z & \text{if} \quad z\ge 0.\end{cases} \tag{2.7}$$

Proof. The proofs are given in Appendix B. □

2.3. The explosive case

Finally, we deal with the more complicated explosive case where the parameter θ is positive in (1.1).

Theorem 2.3. *The couple (V_T), given by (1.3), satisfies the following bounds.*

(i) *Upper bound: For any compact set $F\subset\mathbb{R}^2$,*

$$\limsup_{T\to\infty}\frac{1}{T}\log\mathbb{P}\big(V_T\in F\big)\le-\inf_{(x,y)\in F}\mathcal{I}_\theta(x,y). \tag{2.8}$$

(ii) *Lower bound: For any open set* $G \subset \mathbb{R}^2$,

$$-\inf_{(x,y)\in G\cap\mathcal{F}} \mathcal{I}_\theta(x,y) \le \liminf_{n\to\infty} \frac{1}{T}\log \mathbb{P}\big(V_T \in G\big), \tag{2.9}$$

where $\mathcal{I}_\theta$ *is the good rate function given by*

$$\mathcal{I}_\theta(x,y) = \begin{cases} \dfrac{\theta(1-x^2+\theta y)}{2} + \dfrac{(1+x^2)^2}{8y} & \text{if} \quad 0 < y < \dfrac{1}{2\theta}(1+x^2), \\ \theta & \text{if} \quad y \ge \dfrac{1}{2\theta}(1+x^2), \\ +\infty & \text{if} \quad y \le 0, \end{cases} \tag{2.10}$$

and $\mathcal{F}$ *is the set of exposed points of* $\mathcal{I}_\theta$ *defined by*

$$\mathcal{F} = \Big\{(x,y) \in \mathbb{R}^2 \text{ such that } 0 < y < \frac{1}{2\theta}(1+x^2)\Big\}. \tag{2.11}$$

Remark 2.1. Let us remark that $\mathcal{I}_\theta$ is a continuous function on $\mathbb{R}\times\mathbb{R}_+^*$ and a constant function on $(\mathbb{R}\times\mathbb{R}_+^*)\setminus\mathcal{F}$. Consequently, we are able to specify (2.9): For any open and connex set $G \subset \mathbb{R}^2$ such that $G\cap\mathcal{F} \neq \emptyset$,

$$\inf_{(x,y)\in G\cap\mathcal{F}} \mathcal{I}_\theta(x,y) = \inf_{(x,y)\in G} \mathcal{I}_\theta(x,y).$$

Despite the weak large deviation result of Theorem 2.3, it is possible to establish the following corollary, which was previously proved in Bercu et al. (2012) via a much more complex procedure.

Corollary 2.3. *The sequence* $(\widehat{\theta}_T)$ *satisfies an LDP with good rate function*

$$I_\theta(z) = \begin{cases} -\dfrac{(z-\theta)^2}{4z} & \text{if} \quad z \le -\theta, \\ \theta & \text{if} \quad |z| < \theta, \\ 0 & \text{if} \quad z = \theta, \\ 2z-\theta & \text{if} \quad z > \theta. \end{cases} \tag{2.12}$$

Proof. The proofs are given in Appendix B. □

Acknowledgments

The authors would like to thank the two anonymous reviewers for their well-pointed comments which helped to fill a gap in the proof of Corollary 2.3.

Appendix A. Gärtner–Ellis theorem and the contraction principle

The most powerful tool for proving LDP is probably the Gärtner–Ellis theorem. Let (V_T) be sequence of random vectors of $\mathbb{R}^d$. Denote by L_T the normalized cumulant generating function of V_T,

$$L_T(a) = \frac{1}{T}\log \mathbb{E}\left[\exp\left(T\langle a, V_T\rangle\right)\right].$$

The existence of the limiting cumulant generating function

$$L(a) = \lim_{T\to\infty} L_T(a)$$

indicates whether or not (V_T) satisfies an LDP. Denote by $\mathcal{D}_L$ the effective domain of L,

$$\mathcal{D}_L = \Big\{a \in \mathbb{R}^d \text{ such that } L(a) < \infty\Big\}.$$

Let I be the Fenchel–Legendre transform of L,

$$I(x) = \sup_{a\in\mathbb{R}^d}\Big\{\langle a, x\rangle - L(a)\Big\}.$$

Theorem A.4 (*Gärtner–Ellis*)**.** *Assume that the function L exists as an extended real number. Then,*

(i) *Upper bound: For any compact set* $F \subset \mathbb{R}^d$,

$$\limsup_{T\to\infty} \frac{1}{T} \log \mathbb{P}\big(V_T \in F\big) \le -\inf_{x\in F} I(x). \tag{A.1}$$

(ii) *Lower bound: For any open set* $G \subset \mathbb{R}^d$,

$$-\inf_{x\in G\cap\mathcal{F}} I(x) \le \liminf_{n\to\infty} \frac{1}{T} \log \mathbb{P}\big(V_T \in G\big), \tag{A.2}$$

where $\mathcal{F}$ *is the set of exposed points of* I *whose exposing hyperplane belongs to the interior of* $\mathcal{D}_L$.

(iii) *If* L *is an essentially smooth, lower semicontinuous function, then the sequence* (V_T) *satisfies a weak LDP with rate function* I. *If, moreover, the origin belongs to the interior of* $\mathcal{D}_L$, (V_T) *satisfies an LDP with good rate function* I.

We refer the reader to the excellent book (Dembo and Zeitouni, 1998) for more insight on the theory of large deviations. In particular, the Gärtner–Ellis is given in Theorem 2.3.6 of Dembo and Zeitouni (1998). Another useful tool is the contraction principle which ensures that an LDP remains valid by continuous mapping, see Theorem 4.2.1 of Dembo and Zeitouni (1998).

Theorem A.5 (*Contraction Principle*)**.** *Assume that a sequence of random vectors* (V_T) *with values in* $E \subset \mathbb{R}^d$ *satisfies an LDP with good rate function* I, *and that* $A_T = f(V_T)$ *where* f *is a continuous function from* E *to* $\mathbb{R}^\delta$. *Then,* (A_T) *also satisfies an LDP with good rate function* J *given, for all* $y \in \mathbb{R}^\delta$, *by*

$$J(y) = \inf\big\{I(x) \text{ with } x \in E \text{ such that } f(x) = y\big\}, \tag{A.3}$$

where the infimum over the empty set is taken to be infinite.

Appendix B. Proofs of LDP results

Let L_T be the normalized cumulant generating function of the couple

$$V_T = \left(\frac{X_T}{\sqrt{T}}, \frac{S_T}{T}\right)$$

defined, for all $(a, b) \in \mathbb{R}^2$, by

$$L_T(a, b) = \frac{1}{T} \log \mathbb{E}\Big[\exp\Big(a\sqrt{T}X_T + bS_T\Big)\Big].$$

The proofs of all the LDP results rely on an accurate evaluation of $L_T(a, b)$ as well as on the existence of the limiting cumulant generating function $L(a, b)$. This is the subject of the following keystone lemma.

Lemma B.1. *In the stable and unstable cases* $\theta \le 0$, *the effective domain of* L *is*

$$\mathcal{D}_L = \left\{(a, b) \in \mathbb{R}^2 \text{ such that } b < \frac{\theta^2}{2}\right\}, \tag{B.1}$$

while, in the explosive case $\theta > 0$, *the effective domain of* L *becomes*

$$\mathcal{D}_L = \Big\{(a, b) \in \mathbb{R}^2 \text{ such that } b < 0\Big\}. \tag{B.2}$$

Moreover, for any $(a, b) \in \mathcal{D}_L$, *we have whatever the value of* θ,

$$L(a, b) = -\frac{1}{2}\left(\theta + \sqrt{\theta^2 - 2b}\right) + \frac{a^2}{2\left(\sqrt{\theta^2 - 2b} - \theta\right)}. \tag{B.3}$$

Remark B.2. One can observe that, as soon as $\theta \ge 0$, the origin does not belong to the interior of $\mathcal{D}_L$.

Proof. We start with the stable and unstable cases. Using the same lines as in Appendix A of Bercu and Richou (2015), we obtain from Girsanov's formula associated with (1.1) that

$$\begin{aligned} L_T(a, b) &= \frac{1}{T} \log \mathbb{E}_\varphi\left[\exp\Big((\theta - \varphi)\int_0^T X_t dX_t - \frac{1}{2}(\theta^2 - \varphi^2)S_T + a\sqrt{T}X_T + bS_T\Big)\right], \\ &= \frac{1}{T} \log \mathbb{E}_\varphi\left[\exp\Big(\frac{(\theta - \varphi)}{2}(X_T^2 - T) + a\sqrt{T}X_T + \frac{1}{2}(2b - \theta^2 + \varphi^2)S_T\Big)\right] \end{aligned}$$

where $\mathbb{E}_\varphi$ stands for the expectation after the usual change of probability,

$$\frac{d\mathbb{P}_\varphi}{d\mathbb{P}_\theta} = \exp\left((\varphi-\theta)\int_0^T X_t dX_t - \frac{1}{2}(\varphi^2-\theta^2)\int_0^T X_t^2\, dt\right).$$

Consequently, if $\theta^2 - 2b > 0$ and $\varphi = \sqrt{\theta^2-2b}$, $L_T(a,b)$ reduces to

$$L_T(a,b) = \frac{\varphi-\theta}{2} + \frac{1}{T}\log \mathbb{E}_\varphi\left[\exp\left(\left(\frac{\theta-\varphi}{2}\right)X_T^2 + a\sqrt{T}X_T\right)\right]. \tag{B.4}$$

Under the new probability $\mathbb{P}_\varphi$, X_T has an $\mathcal{N}(0,\sigma_T^2)$ distribution where

$$\sigma_T^2 = \frac{1}{2\varphi}\left(e^{2\varphi T}-1\right). \tag{B.5}$$

Hence, it follows from straightforward Gaussian calculations that

$$L_T(a,b) = \frac{\varphi-\theta}{2} + \frac{a^2\sigma_T^2}{2\gamma_T} - \frac{1}{2T}\log\gamma_T \tag{B.6}$$

where $\gamma_T = 1 + (\varphi-\theta)\sigma_T^2$. However, we clearly obtain from (B.5) that

$$\lim_{T\to\infty}\frac{1}{T}\log\sigma_T^2 = 2\varphi, \qquad \lim_{T\to\infty}\frac{\gamma_T}{\sigma_T^2} = \varphi-\theta, \qquad \lim_{T\to\infty}\frac{1}{T}\log\gamma_T = 2\varphi.$$

Hence, we deduce from (B.6) that

$$\lim_{T\to\infty} L_T(a,b) = -\frac{1}{2}(\theta+\varphi) + \frac{a^2}{2(\varphi-\theta)}, \tag{B.7}$$

which is exactly the limiting cumulant generating function $L(a,b)$ given by (B.3). In the explosive case $\theta > 0$, calculations are quite the same with the only significant modification that $\varphi = -\sqrt{\theta^2-2b}$ instead of $\sqrt{\theta^2-2b}$. Then, (B.6) holds true with the new parameter φ and

$$\lim_{T\to\infty}\frac{1}{T}\log\gamma_T = 0, \qquad \lim_{T\to\infty}\frac{\gamma_T}{\sigma_T^2} = -(\varphi+\theta).$$

Consequently, (B.3) follows from (B.6), completing the proof of Lemma B.1. □

We shall also make use of normalized cumulant generating function Λ_T of the couple

$$W_T = \left(\frac{X_T^2}{T}, \frac{S_T}{T}\right)$$

defined, for all $(a,b)\in\mathbb{R}^2$, by

$$\Lambda_T(a,b) = \frac{1}{T}\log\mathbb{E}\left[\exp\left(aX_T^2 + bS_T\right)\right].$$

The proofs of LDP results in the unstable and explosive cases require the following lemma on the effective domain of the limiting cumulant generating function $\Lambda(a,b)$ of $\Lambda_T(a,b)$.

Lemma B.2. *If $\theta \geq 0$, the effective domain of Λ is given by*

$$\mathcal{D}_\Lambda = \left\{(a,b)\in\mathbb{R}^2 \text{ such that } \theta^2-2b>0 \text{ and } 2a+\theta < \sqrt{\theta^2-2b}\right\}.$$

Proof. The proof is the same as that of Lemma B.1. □

B.1. The stable case

Proof of Theorem 2.1. The origin belongs to the interior of the domain $\mathcal{D}_L$ given by (B.1). Moreover, the function L, defined in (B.3), is differentiable throughout $\mathcal{D}_L$ and L is steep, which means that L is essentially smooth. Hence, one can immediately deduce from the Gärtner–Ellis theorem that the couple (V_T) satisfies an LDP with speed T and good rate function

$$\mathcal{I}_\theta(x,y) = \sup_{(a,b)\in\mathcal{D}_L}\left\{ax+by-L(a,b)\right\}.$$

It is easy to compute $\mathcal{I}_\theta$. After some straightforward calculations, we obtain the expression given by (2.3), which achieves the proof of Theorem 2.1. □

Proof of Corollary 2.1. Corollary 2.1 follows from Theorem 2.1 together with an elementary application of the contraction principle. We already saw in Section 2 that $\widehat{\theta}_T = f(V_T)$ where f is the continuous function defined, for all $x \in \mathbb{R}$ and $y > 0$, by

$$f(x, y) = \frac{x^2 - 1}{2y}.$$

Consequently, one can immediately deduce from the contraction principle that the sequence $(\widehat{\theta}_T)$ satisfies an LDP with good rate function I_θ given, for all $z \in \mathbb{R}$, by

$$I_\theta(z) = \inf\Big\{\mathcal{I}_\theta(x, y) \text{ with } x \in \mathbb{R}, y > 0 \text{ such that } f(x, y) = z\Big\}. \tag{B.8}$$

Hereafter, it only remains to properly evaluate I_θ. As soon as $1 + 2yz \geq 0$,

$$f(x, y) = z \iff x^2 = 1 + 2yz.$$

Hence, (2.3) together with (B.8) leads to

$$I_\theta(z) = \inf\big\{h(y) \text{ with } 1 + 2yz \geq 0, y > 0\big\} \tag{B.9}$$

where h is the function defined, for any positive y, by

$$h(y) = \frac{\theta y(\theta - 2z)}{2} + \frac{(1 + yz)^2}{2y}. \tag{B.10}$$

We clearly have from (B.10) that h is a convex function as

$$h'(y) = \frac{1}{2}\Big((z - \theta)^2 - \frac{1}{y^2}\Big) \quad \text{and} \quad h''(y) = \frac{1}{y^3}. \tag{B.11}$$

The evaluation of the rate function I_θ depends on the location of its argument. On the one hand, as soon as $z \leq \theta/3$, the border condition $1 + 2yz \geq 0$ plays a prominent role as

$$I_\theta(z) = h\Big(-\frac{1}{2z}\Big) = -\frac{(z - \theta)^2}{4z}.$$

On the other hand, as soon as $z \geq \theta/3$, the border condition $1 + 2yz \geq 0$ does not have to be taken into account as

$$I_\theta(z) = h\Big(\frac{1}{z - \theta}\Big) = 2z - \theta,$$

which completes the proof of Corollary 2.1. □

B.2. The unstable case

Proof of Theorem 2.2. The proof of Theorem 2.2 can be handled exactly as that of Theorem 2.1 by taking the value $\theta = 0$. The function L, given by (B.3), is essentially smooth. However, in contrast to the stable case, the origin does no longer belong to the interior of $\mathcal{D}_L$. It means that the sequence (V_n) is not exponentially tight. This is the reason why we obtain a weak LDP for (V_n) instead of a full LDP, via the weak version of Gärtner–Ellis Theorem A.4. □

Proof of Corollary 2.2. Since Theorem 2.2 provides us a weak LDP for the sequence (V_n), we cannot deduce Corollary 2.2 from a direct application of the contraction principle. Instead of that, we shall prove the LDP for $(\widehat{\theta}_T)$ by considering the rare events $\{\widehat{\theta}_T \leq c\}$ and $\{\widehat{\theta}_T \geq c\}$, for c negative and c positive, respectively. First of all, we have for any negative c,

$$\mathbb{P}(\widehat{\theta}_T \leq c) = \mathbb{P}(f(V_T) \leq c) = \mathbb{P}(V_T \in \Delta_c)$$

where the set Δ_c is given, for $a_c(x) = (x^2 - 1)/2c$, by

$$\Delta_c = \Big\{(x, y) \in \mathbb{R}^2 \text{ such that } |x| \leq 1 \text{ and } y \in [0, a_c(x)]\Big\}.$$

Since Δ_c is a compact set of $\mathbb{R}^2$, Theorem 2.2 implies that

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(\widehat{\theta}_T \leq c) = -\inf_{(x,y)\in\Delta_c} \mathcal{I}_0(x, y).$$

However, the rate function $\mathcal{I}_0$ has no critical points and $\mathcal{I}_0(x, 0) = +\infty$. Hence,

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(\widehat{\theta}_T \leq c) = -\inf_{|x|<1} \mathcal{I}_0(x, a_c(x)) = -\mathcal{I}_0\left(0, -\frac{1}{2c}\right) = \frac{c}{4} = -I_0(c).$$

We now consider the more tedious case where c is positive. We have for any $\alpha > 0$,

$$\mathbb{P}(\widehat{\theta}_T \geq c) = \mathbb{P}\Big(\widehat{\theta}_T \geq c, \frac{|X_T|}{\sqrt{T}} \leq \alpha\Big) + \mathbb{P}\Big(\widehat{\theta}_T \geq c, \frac{|X_T|}{\sqrt{T}} > \alpha\Big). \tag{B.12}$$

One can remark that

$$\mathbb{P}\Big(\widehat{\theta}_T \geq c, \frac{|X_T|}{\sqrt{T}} \leq \alpha\Big) = \mathbb{P}(V_T \in \Delta_{c,\alpha})$$

where $\Delta_{c,\alpha}$ is the compact set of $\mathbb{R}^2$ defined by

$$\Delta_{c,\alpha} = \Big\{(x, y) \in \mathbb{R}^2 \text{ such that } 1 \leq |x| \leq \alpha \text{ and } y \in [0, a_c(x)]\Big\}.$$

Therefore, we deduce from Theorem 2.2 that

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(V_T \in \Delta_{c,\alpha}) = - \inf_{(x,y)\in\Delta_{c,\alpha}} \mathcal{I}_0(x, y).$$

After some straightforward calculations, we obtain that, as soon as $\alpha \geq \sqrt{3}$,

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}\Big(\widehat{\theta}_T \geq c, \frac{|X_T|}{\sqrt{T}} \leq \alpha\Big) = -\mathcal{I}_0\Big(\sqrt{3}, \frac{1}{c}\Big) = -2c = -I_0(c). \tag{B.13}$$

It only remains to prove that the right-hand side of (B.12) is negligible. It follows from Markov's inequality that for any positive λ and μ,

$$\begin{aligned}\mathbb{P}\Big(\widehat{\theta}_T \geq c, \frac{|X_T|}{\sqrt{T}} > \alpha\Big) &= \mathbb{P}\Big(X_T^2 - 2cS_T \geq T, X_T^2 > \alpha^2 T\Big),\\ &\leq \exp\Big(-T(\lambda + \mu\alpha^2)\Big)\mathbb{E}\Big[\exp\Big((\lambda+\mu)X_T^2 - 2\lambda c S_T\Big)\Big],\\ &\leq \exp\Big(-T\Big((\lambda + \mu\alpha^2) - \Lambda_T(\lambda+\mu, -2\lambda c)\Big)\Big).\end{aligned} \tag{B.14}$$

By choosing $\lambda = \mu = c/5$, it is not hard to see that the couple $(2c/5, -2c^2/5)$ belongs to the effective domain $\mathcal{D}_\Lambda$ given in Lemma B.2. Hence, as Λ_T converges simply to Λ on $\mathcal{D}_\Lambda$, we infer from (B.14) that for T large enough,

$$\mathbb{P}\Big(\widehat{\theta}_T \geq c, \frac{|X_T|}{\sqrt{T}} > \alpha\Big) \leq \exp\Big(-T\Big(\frac{c}{5}(1+\alpha^2) - 2\Lambda\Big(\frac{2c}{5}, -\frac{2c^2}{5}\Big)\Big)\Big)$$

which implies that for α and T large enough,

$$\mathbb{P}\Big(\widehat{\theta}_T \geq c, \frac{|X_T|}{\sqrt{T}} > \alpha\Big) \leq \exp(-3cT). \tag{B.15}$$

Therefore, it follows from (B.12), (B.13) and (B.15) that for any positive c,

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(\widehat{\theta}_T \geq c) = -2c = -I_0(c).$$

Finally, in the unstable case, X_T has an $\mathcal{N}(0, T)$ distribution. Hence, the case $c = 0$ is straightforward as

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(\widehat{\theta}_T \geq 0) = \lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(X_T^2 \geq T) = 0 = I_0(0),$$

and by the same token

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(\widehat{\theta}_T \leq 0) = I_0(0).$$

From now on, it remains to deduce the LDP for the sequence $(\widehat{\theta}_T)$ thanks to our tails estimates. First of all, it follows from our tails estimates that $(\widehat{\theta}_T)$ is exponentially tight. Consequently, we only need to establish a weak LDP in order to complete our proof. By applying Theorem 4.1.11 in Dembo and Zeitouni (1998), we just have to show that, for all $c_1, c_2 \in \mathbb{R}$ with $c_1 < c_2$,

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(c_1 < \widehat{\theta}_T < c_2) = - \inf_{c\in]c_1,c_2[} I_0(c). \tag{B.16}$$

Since I_0 is a decreasing function on $]-\infty, 0]$ and an increasing function on $[0, +\infty[$, (B.16) is a direct consequence of tails estimates. As a matter of fact, if $c_1 < c_2 < 0$, we clearly have

$$\mathbb{P}(c_1 < \widehat{\theta}_T < c_2) = \mathbb{P}(\widehat{\theta}_T < c_2) - \mathbb{P}(\widehat{\theta}_T \leq c_1) = \mathbb{P}(\widehat{\theta}_T < c_2)\big(1 + o(1)\big)$$

as T tends to infinity. All other cases can be handled in the same way, which achieves the proof of Corollary 2.2. □

B.3. The explosive case

Proof of Theorem 2.3. The proof of Theorem 2.3 can be handled exactly as that of Theorem 2.1 by taking $\theta > 0$. However, in contrast to the stable case, the origin does no longer belong to the interior of $\mathcal{D}_L$ and the function L, given by (B.3), is not essentially smooth. This is the reason why we are only allowed to apply the weakest version of Gärtner–Ellis Theorem A.4. □

Proof of Corollary 2.3. We shall proceed as in the proof of Corollary 2.2 by considering rare events $\{\widehat{\theta}_T \leq c\}$ and $\{\widehat{\theta}_T \geq c\}$, for $c < \theta$ and $c > \theta$, respectively. First of all, we already saw that for any negative c, $\mathbb{P}(\widehat{\theta}_T \leq c) = \mathbb{P}(V_T \in \Delta_c)$ where Δ_c is the compact set of $\mathbb{R}^2$ given, for $a_c(x) = (x^2-1)/2c$, by

$$\Delta_c = \Big\{(x, y) \in \mathbb{R}^2 \text{ such that } |x| \leq 1 \text{ and } y \in [0, a_c(x)]\Big\}.$$

Since $\Delta_c \cap \mathcal{F} \neq \emptyset$, it follows from Theorem 2.2 together with Remark 2.1 that

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(\widehat{\theta}_T \leq c) = -\inf_{(x,y)\in\Delta_c} \mathcal{I}_\theta(x, y).$$

However, the function $\mathcal{I}_\theta$ has no critical points on $\mathcal{F}$ and $\mathcal{I}_\theta(x, 0) = +\infty$. Consequently,

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(\widehat{\theta}_T \leq c) = -\inf_{|x|<1} \mathcal{I}_\theta(x, a_c(x)) = -\mathcal{I}_\theta\left(0, -\frac{1}{2c}\right).$$

In particular, as soon as $c < -\theta$,

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(\widehat{\theta}_T \leq c) = \frac{(c-\theta)^2}{4c} = -I_\theta(c),$$

while, for $-\theta \leq c < 0$,

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(\widehat{\theta}_T \leq c) = -\theta = -I_\theta(c).$$

From now on, assume that $0 \leq c < \theta$. We have for any $\alpha > 1/2\theta$,

$$\mathbb{P}(\widehat{\theta}_T \leq c) = \mathbb{P}\Big(\widehat{\theta}_T \leq c, \frac{S_T}{T} \leq \alpha\Big) + \mathbb{P}\Big(\widehat{\theta}_T \leq c, \frac{S_T}{T} > \alpha\Big). \tag{B.17}$$

One can remark that

$$\mathbb{P}\Big(\widehat{\theta}_T \leq c, \frac{S_T}{T} \leq \alpha\Big) = \mathbb{P}(V_T \in \Delta_{c,\alpha})$$

where $\Delta_{c,\alpha}$ is the compact set of $\mathbb{R}^2$ defined, for $c > 0$, by

$$\Delta_{c,\alpha} = \Big\{(x, y) \in \mathbb{R}^2 \text{ such that } 0 \leq y \leq \alpha \text{ and } y \geq a_c(x)\Big\},$$

and, for $c = 0$, by

$$\Delta_{c,\alpha} = \Big\{(x, y) \in \mathbb{R}^2 \text{ such that } 0 \leq y \leq \alpha \text{ and } |x| \leq 1\Big\}.$$

Since $\Delta_{c,\alpha} \cap \mathcal{F} \neq \emptyset$, we obtain from Theorem 2.2 together with Remark 2.1 that

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(V_T \in \Delta_{c,\alpha}) = -\inf_{(x,y)\in\Delta_{c,\alpha}} \mathcal{I}_\theta(x, y).$$

After some straightforward calculations, we find that

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}\Big(\widehat{\theta}_T \leq c, \frac{S_T}{T} \leq \alpha\Big) = -\mathcal{I}_\theta\Big(0, \frac{1}{2\theta}\Big) = -\theta = -I_\theta(c). \tag{B.18}$$

It now remains to show that the remainder term of (B.17) is negligible. It follows from Markov's inequality that for any negative λ and for any positive μ,

$$\begin{aligned}\mathbb{P}\Big(\widehat{\theta}_T \leq c, \frac{S_T}{T} > \alpha\Big) &= \mathbb{P}\Big(X_T^2 - 2cS_T \leq T, S_T > \alpha T\Big),\\ &\leq \exp\Big(-T(\lambda+\mu\alpha)\Big)\mathbb{E}\Big[\exp\Big(\lambda X_T^2 + (\mu - 2\lambda c)S_T\Big)\Big],\\ &\leq \exp\Big(-T\Big((\lambda+\mu\alpha) - \Lambda_T(\lambda, \mu - 2\lambda c)\Big)\Big). \end{aligned} \tag{B.19}$$

By setting $\lambda = (c-\theta)/2$ and $\mu = (c-\theta)^2/4$, one can check that the couple $(\lambda, \mu - 2\lambda c)$ belongs to the effective domain $\mathcal{D}_\Lambda$ given in Lemma B.2. Hence, we obtain from (B.19) that for α and T large enough,

$$\mathbb{P}\Big(\widehat{\theta}_T \leq c, \frac{S_T}{T} > \alpha\Big) \leq \exp(-2\theta T). \tag{B.20}$$

As a consequence, we deduce from (B.17), (B.18) and (B.20) that for any $0 \leq c < \theta$,

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(\widehat{\theta}_T \leq c) = -\theta = -I_\theta(c).$$

Finally, we shall investigate the case $c > \theta$. We have for any $\alpha > 0$,

$$\mathbb{P}(\widehat{\theta}_T \geq c) = \mathbb{P}\Big(\widehat{\theta}_T \geq c, \frac{|X_T|}{\sqrt{T}} \leq \alpha\Big) + \mathbb{P}\Big(\widehat{\theta}_T \geq c, \frac{|X_T|}{\sqrt{T}} > \alpha\Big). \tag{B.21}$$

As in the proof of Corollary 2.2,

$$\mathbb{P}\Big(\widehat{\theta}_T \geq c, \frac{|X_T|}{\sqrt{T}} \leq \alpha\Big) = \mathbb{P}(V_T \in \Delta_{c,\alpha})$$

where $\Delta_{c,\alpha}$ is the compact set of $\mathbb{R}^2$ defined by

$$\Delta_{c,\alpha} = \Big\{(x,y) \in \mathbb{R}^2 \text{ such that } 1 \leq |x| \leq \alpha \text{ and } y \in [0, a_c(x)]\Big\}.$$

Since $\Delta_{c,\alpha} \cap \mathcal{F} \neq \emptyset$, it follows from Theorem 2.2 together with Remark 2.1 that

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(V_T \in \Delta_{c,\alpha}) = -\inf_{(x,y)\in\Delta_{c,\alpha}} \mathcal{I}_\theta(x,y).$$

Furthermore, denote $\alpha_c(\theta) = \sqrt{c+\theta}/\sqrt{c-\theta}$. After some straightforward calculations, we obtain that, as soon as $\alpha \geq \alpha_c(\theta)$,

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}\Big(\widehat{\theta}_T \geq c, \frac{|X_T|}{\sqrt{T}} \leq \alpha\Big) = -\mathcal{I}_\theta\Big(\alpha_c(\theta), \frac{1}{c-\theta}\Big) = \theta - 2c = -I_\theta(c). \tag{B.22}$$

Using once again Markov's inequality, we have for any positive λ and μ,

$$\begin{aligned}\mathbb{P}\Big(\widehat{\theta}_T \geq c, \frac{|X_T|}{\sqrt{T}} > \alpha\Big) &= \mathbb{P}\Big(X_T^2 - 2cS_T \geq T, X_T^2 > \alpha^2 T\Big),\\ &\leq \exp\Big(-T(\lambda+\mu\alpha^2)\Big)\mathbb{E}\Big[\exp\Big((\lambda+\mu)X_T^2 - 2\lambda c S_T\Big)\Big],\\ &\leq \exp\Big(-T\Big((\lambda+\mu\alpha^2) - \Lambda_T(\lambda+\mu, -2\lambda c)\Big)\Big). \end{aligned} \tag{B.23}$$

By choosing $\lambda = (c^2-\theta^2)/4c$ and $\mu = (c-\theta)^2/8c$, it is not hard to see that the couple $(\lambda+\mu, -2c\lambda)$ belongs to the effective domain $\mathcal{D}_\Lambda$ given in Lemma B.2. Hence, we obtain from (B.23) that for α and T large enough,

$$\mathbb{P}\Big(\widehat{\theta}_T \geq c, \frac{|X_T|}{\sqrt{T}} > \alpha\Big) \leq \exp(-2(2c-\theta)T). \tag{B.24}$$

Therefore, it follows from (B.21), (B.22) and (B.24) that for any positive $c > \theta$,

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(\widehat{\theta}_T \geq c) = \theta - 2c = -I_\theta(c).$$

Hereafter, it remains to deduce the LDP for the sequence $(\widehat{\theta}_T)$ thanks to our tails estimates. Once again, it follows from our tails estimates that $(\widehat{\theta}_T)$ is exponentially tight. Consequently, we only need to establish a weak LDP in order to complete our proof. By applying Theorem 4.1.11 in Dembo and Zeitouni (1998), we just have to show that, for all $c_1, c_2 \in \mathbb{R}$ with $c_1 < c_2$ different than θ,

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(c_1 < \widehat{\theta}_T < c_2) = -\inf_{c\in]c_1,c_2[} I_\theta(c). \tag{B.25}$$

In contrast to the previous unstable case, the rate function I_θ is constant on $[-\theta, \theta[$. Hence, (B.25) is a direct consequence of our tails estimates only when $]c_1, c_2[\not\subset [-\theta, \theta[$. We now focus our attention on the more tricky case $]c_1, c_2[\in [-\theta, \theta[$ by assuming, for example, that $-\theta \le c_1 < c_2 < 0$. As in the previous calculation, $\mathbb{P}(c_1 < \widehat{\theta}_T < c_2) = \mathbb{P}(V_T \in \Delta_{c_1,c_2})$ where Δ_{c_1,c_2} is the relative compact set of $\mathbb{R}^2$ given by

$$\Delta_{c_1,c_2} = \Big\{(x, y) \in \mathbb{R}^2 \text{ such that } |x| < 1 \text{ and } y \in]a_{c_1}(x), a_{c_2}(x)[\Big\}.$$

Since $\Delta_{c_1,c_2} \cap \mathcal{F} \ne \emptyset$, we deduce from Theorem 2.2 together with Remark 2.1 that

$$\begin{aligned}\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}(c_1 < \widehat{\theta}_T < c_2) &= -\inf_{(x,y)\in\Delta_{c_1,c_2}} \mathcal{I}_\theta(x, y),\\ &= -\mathcal{I}_\theta\Big(0, -\frac{1}{2c_2}\Big) = -\theta,\\ &= -\inf_{c\in]c_1,c_2[} I_\theta(c).\end{aligned}$$

The two other cases $0 \le c_1 < c_2 < \theta$ and $-\theta \le c_1 < 0 \le c_2 < \theta$, can be handled in the same way, which completes the proof of Corollary 2.3. □

References

Alfonsi, A., 2015. Affine Diffusions and Related Processes: Simulation, Theory and Applications. In: Bocconi and Spinger Series, vol. 6. Bocconi University Press, Milan.

Bercu, B., Coutin, L., Savy, N., 2012. Sharp large deviations for the non-stationary Ornstein-Uhlenbeck process. Stochastic Process. Appl. 122 (10), 3393–3424.

Bercu, B., Richou, A., 2015. Large deviations for the Ornstein-Uhlenbeck process with shift. Adv. Appl. Probab. 47, 1–22.

Bercu, B., Rouault, A., 2002. Sharp large deviations for the Ornstein-Uhlenbeck process. SIAM Theory Probab. Appl. 46 (1), 1–19.

Dembo, A., Zeitouni, O., 1998. Large Deviations Techniques and Applications, second ed. In: Applications of Mathematics (New York), vol. 38. Springer-Verlag, New York.

Demni, N., Zani, M., 2009. Large deviations for statistics of the Jacobi process. Stochastic Process. Appl. 119 (2), 518–533.

Du Roy de Chaumaray, M., 2016. Large deviations for the squared radial Ornstein-Uhlenbeck process. Teor. Veroyatn. Primen. 61 (3), 509–546.

Florens-Landais, D., Pham, H., 1999. Large deviations in estimation of an Ornstein-Uhlenbeck model. J. Appl. Probab. 36 (1), 60–77.

Forman, J.L., Sorensen, M., 2008. The Pearson diffusions: a class of statistically tractable diffusion processes. Scand. J. Stat. 35 (3), 438–465.

Zani, M., 2002. Large deviations for the squared radial Ornstein-Uhlenbeck processes. Stochastic Process. Appl. 102 (1), 25–42.

Zhao, S., Gao, F., 2010. Large deviations in testing Jacobi model. Statist. Probab. Lett. 80, 34–41.