THEORY PROBAB. APPL.
Vol. 46, No. 1

SHARP LARGE DEVIATIONS FOR THE ORNSTEIN–UHLENBECK PROCESS*

B. BERCU† AND A. ROUAULT‡

Abstract. We establish sharp large deviation principles for well-known random variables associated with the Ornstein–Uhlenbeck process, such as the energy, the maximum likelihood estimator of the drift parameter, and the log-likelihood ratio.

Key words. large deviations, Ornstein–Uhlenbeck process, likelihood estimation

PII. S0040585X97978737

1. Introduction.

Consider the Ornstein–Uhlenbeck process

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

where W is a standard Brownian motion and the parameter θ is strictly negative. For the sake of simplicity, we choose the initial state $X_0 = 0$. In this paper, we investigate the sharp large deviation properties for well-known random variables associated with (1.1), such as the energy

$$S_T = \int_0^T X_t^2 \, dt, \tag{1.2}$$

the maximum likelihood estimator of θ

$$\hat{\theta}_T = \frac{\int_0^T X_t \, dX_t}{\int_0^T X_t^2 \, dt} = \frac{X_T^2 - T}{2 \int_0^T X_t^2 \, dt}, \tag{1.3}$$

and the log-likelihood ratio

$$V_T = (\theta_0 - \theta_1) \int_0^T X_t \, dX_t - \frac{1}{2} (\theta_0^2 - \theta_1^2) \int_0^T X_t^2 \, dt \tag{1.4}$$

with θ_0 and θ_1 strictly negative. It was already proven that a.s., as T goes to infinity,

$$\frac{S_T}{T} \to -\frac{1}{2\theta}, \quad \hat{\theta}_T \to \theta, \quad \frac{V_T}{T} \to -\frac{(\theta_0 - \theta_1)^2}{4\theta_0}. \tag{1.5}$$

Fluctuations are also known [1]. More recently, Bryc and Dembo [8] and Florens-Landais and Pham [14] have established large deviation principles for S_T and $\hat{\theta}_T$. Strictly speaking, $\hat{\theta}_T$ is not the maximum likelihood estimator of θ since $\hat{\theta}_T$ may take nonnegative values, whereas the parameter θ is assumed to be strictly negative.

*Received by the editors October 27, 1998.
http://www.siam.org/journals/tvp/46-1/97873.html

†Laboratoire de Statistiques, Bâtiment 425 Mathématiques, Université Paris Sud, 91405 Orsay, France (bernard.bercu@math.u-psud.fr).

‡Département de Mathématiques, Bâtiment Fermat, Université de Versailles, 78035 Versailles, France (rouault@math.uvsq.fr).

Nevertheless, we use this terminology throughout the paper. We present here sharp large deviation principles for the random variables S_T, $\hat{\theta}_T$, and V_T. As usual, we shall say that a family of real random variables (Z_T) satisfies a large deviation principle (LDP), with rate function I, if I is a lower semicontinuous function from $\mathbf{R}$ to $[0, +\infty]$ such that, for any closed set $F \subset \mathbf{R}$, $\limsup_{T\to\infty} T^{-1} \log \mathbf{P}\{Z_T \in F\} \leqq -\inf_{x\in F} I(x)$, while for any open set $G \subset \mathbf{R}$,

$$-\inf_{x\in G} I(x) \leqq \liminf_{T\to\infty} \frac{1}{T} \log \mathbf{P}\{Z_T \in G\}.$$

Moreover, I is a good rate function if its level sets are compact subsets of $\mathbf{R}$. When I has a unique minimum m, which will always be the case here, an LDP for (Z_T) gives the asymptotic behavior of $\mathbf{P}\{Z_T \geqq c\}$ or $\mathbf{P}\{Z_T \leqq c\}$ in a logarithmic scale whenever $c > m$ or $c < m$, respectively. We shall say that a sequence of real random variables (Z_T) satisfies a sharp large deviation principle (SLDP) if, for any real number c, it is possible to give asymptotic expansions of $e^{TI(c)}\mathbf{P}\{Z_T \geqq c\}$ or $e^{TI(c)}\mathbf{P}\{Z_T \leqq c\}$ in powers of T^{-1}. For the sake of simplicity, we present only the first one.

In order to prove an LDP for (1.2), (1.3), or (1.4), the main tool is the normalized cumulant generating function (c.g.f.) of the pair $(\int_0^T X_t\, dX_t, \int_0^T X_t^2\, dt)$

$$\mathcal{L}_T(a,b) = \frac{1}{T} \log \mathbf{E}\Big[\exp\big(\mathcal{Z}_T(a,b)\big)\Big], \tag{1.6}$$

where, for any $(a,b) \in \mathbf{R}^2$,

$$\mathcal{Z}_T(a,b) = a\int_0^T X_t\, dX_t + b\int_0^T X_t^2\, dt. \tag{1.7}$$

It is possible to establish SLDP for all linear combinations of $\int_0^T X_t\, dX_t$ and $\int_0^T X_t^2\, dt$. However, in order to improve the presentation of the paper, we prefer to focus our attention on the random variables (1.2), (1.3), and (1.4).

In discrete time, the analogue of (1.1) is the first order autoregressive process. It was studied, as a particular case, in [4] for LDP and in [5] for SLDP. The proofs used Toeplitz matrices and their asymptotic spectral properties given by the first Szegö theorem for LDP and the strong Szegö theorem for SLDP.

The covariance of the stationary Ornstein–Uhlenbeck process is a Wiener–Hopf operator and, for some "quadratic forms," we can use a similar scheme [23]. Here, we take advantage of the explicit expression of the c.g.f.

An important point for proving an LDP is the determination of the domain Δ of the limit $\mathcal{L}$ of the c.g.f. We shall say that $\mathcal{L}$ is steep if its derivative has an infinite limit at the boundary of Δ (see, e.g., [9, p. 44]). It is a sufficient condition to apply the Gärtner–Ellis theorem. The main difficulty arises when $\mathcal{L}$ is not steep.

Before entering into details of the different cases, let us summarize the general scheme. We have to study probabilities of the form $\mathbf{P}\{\mathcal{Z}_T(a,b) > zT\}$. We follow a similar approach to the one recently given in [5] for discrete time quadratic forms of Gaussian stationary processes. It was inspired by the original work of Bahadur and Rao [2] for the sample mean of a sequence of independent and identically distributed (i.i.d.) random variables. We perform an exponential change of probability of parameter φ to be chosen to "track" z,

$$\frac{d\mathbf{Q}_T}{d\mathbf{P}} = \exp\big(\varphi\mathcal{Z}_T(a,b) - T\mathcal{L}_T(a\varphi, b\varphi)\big). \tag{1.8}$$

Then, we have the decomposition

$$\mathbf{P}\big\{\mathcal{Z}_T(a,b) > zT\big\} = A_T B_T$$

with

$$A_T = \exp\Big[T\big(\mathcal{L}_T(a\varphi, b\varphi) - z\varphi\big)\Big], \tag{1.9}$$

$$B_T = \mathbf{E}_T\left(\exp\big[-\varphi(\mathcal{Z}_T(a,b) - zT)\big]\, \mathbb{1}_{\mathcal{Z}_T(a,b)\geqq zT}\right), \tag{1.10}$$

where $\mathbf{E}_T$ is the expectation under $\mathbf{Q}_T$. On the one hand, an asymptotic expansion for A_T is given in section 2. On the other hand, we expand B_T via an evaluation of the characteristic function Φ_T of $(\mathcal{Z}_T(a,b)-zT)$ (properly normalized) under the new probability $\mathbf{Q}_T$. In general, Φ_T has a pointwise limit. This is, however, not enough to study B_T. Actually, we obtain a complete expansion of Φ_T similar to the one established in the i.i.d. case by Cramer [7] and Esseen [12] and we integrate it using Parseval's theorem. This approach was developed in other contexts with different asymptotics by Ben Arous [3], Bolthausen [6], Dembo, Mayer-Wolf, and Zeitouni [11], Ibragimov [16], Li [19], and Zolotarev [24] (see also the references therein).

In the steep case, the limit distribution is $N(0,1)$ and the convergence of Φ_T is dominated. Thus, we get an SLDP similar to the one obtained in discrete time [5]. In particular, for $c > -1/(2\theta)$, we prove that $\sqrt{T}\, e^{TI(c)}\, \mathbf{P}\{S_T \geqq cT\}$ has a limit given explicitly.

In the nonsteep case, occurring only for (1.3), we find new regimes. For $c \geqq \theta/3$, the rate function I is linear. In the change of probability (1.8), we use a time varying parameter φ_T. This strategy was developed in [10] and [8]. Under this new probability, for $c > \theta/3$, the limit distribution is a centered χ^2 (see also [14]). To establish an SLDP, this gives rise to a new problem since uniform integrability is not guaranteed. Eventually, the SLDP has the same form as in the steep case but with different constants. For $c = \theta/3$, φ_T tends to the boundary of Δ at a different rate and the limit distribution is the convolution of an $N(0,1)$ and a centered χ^2. This yields an SLDP with rate $\sqrt{T}$ instead of T. In particular, $T^{1/4}e^{TI(c)}\mathbf{P}\{\hat{\theta}_T \geqq c\}$ has a limit given explicitly.

The paper is organized as follows. In section 2, we give a sharp description of $\mathcal{L}_T(a,b)$ and $\mathcal{Z}_T(a,b)$. Then, we present the SLDP results: section 3 is devoted to the energy, section 4 to the maximum likelihood estimator, and section 5 to the log-likelihood ratio. Proofs are collected in sections 6 and 7.

2. Main tools. Before presenting specific results for each functional, we give two lemmas which are the core of all our proofs and which allow unified notation.

The convergence of (1.6) was studied in [14]. Lemma 2.1 below gives a new presentation enlightening the role of the limit $\mathcal{L}$ and of the first order term $\mathcal{H}$ for both LDP and SLDP. This decomposition is completely analogous to the one given in discrete time [5].

It is well known that c.g.f.'s of quadratic functionals of Gaussian processes are connected to empirical distributions of eigenvalues of some operator (see, e.g., [15, Chap. 11]).

In our case, Lemma 2.2 gives a decomposition of $\mathcal{Z}_T$ using eigenvalues (see section 7.1 for the operator) and a convergence of their empirical distribution. This last part is similar to the first Szegö theorem and uses the convergence of $\mathcal{L}_T$ to $\mathcal{L}$ for its proof.

The limit is proportional to the image of the Lebesgue measure by the spectral density of the stationary Ornstein–Uhlenbeck process of parameter $\theta < 0$

$$g(x) = \frac{1}{\theta^2 + x^2}$$

associated with the covariance $r(t) = -\exp(\theta|t|)/(2\theta)$.

LEMMA 2.1. *Set* $\Delta = \{(a,b) \in \mathbf{R}^2 \mid \theta^2 - 2b > 0$ *and* $\theta + a < \sqrt{\theta^2 - 2b}\}$ *and let* $\rho(b) = \sqrt{\theta^2 - 2b}$.

(i) *For all* $(a,b) \in \Delta$, *we have*

$$\mathcal{L}_T(a,b) = \mathcal{L}(a,b) + \frac{\mathcal{H}(a,b)}{T} + \frac{1}{T}\mathcal{R}_T(a,b) \tag{2.1}$$

with

$$\begin{aligned} \mathcal{L}(a,b) &= -\frac{1}{2}\big(a + \theta + \rho(b)\big), \\ \mathcal{H}(a,b) &= -\frac{1}{2}\log\left(\frac{1}{2}\big(1 - (a+\theta)\,\rho^{-1}(b)\big)\right), \end{aligned} \tag{2.2}$$

$$\mathcal{R}_T(a,b) = -\frac{1}{2}\log\left(1 + \frac{1 + (a+\theta)\,\rho^{-1}(b)}{1 - (a+\theta)\rho^{-1}(b)}\,e^{-2T\rho(b)}\right). \tag{2.3}$$

(ii) *Moreover, the remainder* $\mathcal{R}_T(a,b)$ *goes exponentially fast to zero*:

$$\mathcal{R}_T(a,b) = \mathcal{O}\big(e^{-2T\rho(b)}\big). \tag{2.4}$$

Denote by $\mathcal{F}$ the class of all continuous functions f on $\mathbf{R}$ such that $f(x) = xh(x)$ with h continuous.

LEMMA 2.2. (i) *There exists a sequence of real numbers* (λ_j^T) *such that* $(\lambda_j^T) \in l^1(\mathbf{N})$,

$$\mathcal{Z}_T(a,b) = -\frac{aT}{2} + \sum_{j=1}^{\infty} \lambda_j^T \varepsilon_j^2, \tag{2.5}$$

where (ε_j) *are independent standard* $N(0,1)$ *random variables.*

(ii) *There exists a fixed compact* $[A,B]$ *such that* $\lambda_j^T \in [A,B]$ *for every* j *and* T.

(iii) *If* $b \neq 0$, *then the empirical measure*

$$\nu_T = \frac{1}{T}\sum_{j=1}^{\infty} \delta_{\lambda_j^T} \tag{2.6}$$

converges for the duality with functions in $\mathcal{F}$ *to* ν *defined, for any continuous function* h *with compact support, by*

$$\langle \nu, h\rangle = \frac{1}{2\pi}\int_{\mathbf{R}} h(bg(x))\,dx. \tag{2.7}$$

Remarks. (1) The sequence (λ_j^T) is finite if $b = 0$ and infinite otherwise.

(2) The case $a = 0$ is already known. Fix $b = 1$ to simplify. The Karhunen–Loève expansion of the process (X_t) directly gives (i). Bryc and Dembo [8] proved (ii) and (iii) (see also [15, Chaps. 8 and 11]). In this case, $A = 0$ and $B = \theta^{-2}$.

(3) If a and b are both nonnegative (respectively, nonpositive), all the (λ_j^T) are nonnegative and we may take $A = 0$ (respectively, $B = 0$).

(4) In the particular case of the likelihood ratio (1.4), decomposition (2.5) is well known (see, e.g., [18] and [22]).

Warning. The functions I, L, L_T given in the following sections are different. They are defined at the beginning of each section. In order to avoid heaviness in the notation, we choose to keep the same symbol for quantities of the same nature.

3. Energy. The LDP for (S_T) was proved by Bryc and Dembo [8] for general centered stationary Gaussian processes. In the particular case of the Ornstein–Uhlenbeck process (1.1), they have the following result.

LEMMA 3.1. *$(T^{-1} S_T)$ satisfies an LDP with good rate function*

$$I(c) = \begin{cases} \dfrac{(2\theta c + 1)^2}{8c} & \text{if } c > 0, \\ +\infty & \text{otherwise.} \end{cases} \tag{3.1}$$

The rate function I has a unique minimum in $c = -1/(2\theta)$ which is the a.s. limit given in (1.5). We are going to improve this result by an SLDP for (S_T) similar to the well-known Bahadur–Rao theorem [2]. Set, for all $a < \theta^2/2$,

$$L(a) = \mathcal{L}(0, a) \quad \text{and} \quad H(a) = \mathcal{H}(0, a). \tag{3.2}$$

One can remark that the rate function I given by (3.1) is the Fenchel–Legendre dual of the function L.

THEOREM 3.1. *$(T^{-1}S_T)$ satisfies an SLDP associated with L and H. More precisely, for all $c > -1/(2\theta)$, there exists a sequence $(b_{c,k})$ such that, for any $p > 0$ and T large enough,*

$$\mathbf{P}\{S_T \geqq cT\} = \frac{e^{-TI(c)+H(a_c)}}{\sigma_c a_c \sqrt{2\pi T}} \left[1 + \sum_{k=1}^{p} \frac{b_{c,k}}{T^k} + \mathcal{O}\left(\frac{1}{T^{p+1}}\right)\right] \tag{3.3}$$

with

$$a_c = \frac{4\theta^2 c^2 - 1}{8c^2}, \qquad H(a_c) = -\frac{1}{2} \log\left(\frac{1}{2}(1 - 2\theta c)\right), \tag{3.4}$$

and $\sigma_c^2 = L''(a_c) = 4c^3$. The coefficients $b_{c,1}, b_{c,2}, \dots, b_{c,p}$ may be explicitly given as functions of the derivatives of L and H at point a_c. For example, the first coefficient $b_{c,1}$ is given by

$$b_{c,1} = \frac{1}{\sigma_c^2}\left(-\frac{h_2}{2} - \frac{h_1^2}{2} + \frac{l_4}{8\sigma_c^2} + \frac{l_3 h_1}{2\sigma_c^2} - \frac{5 l_3^2}{24\sigma_c^4} + \frac{h_1}{a_c} - \frac{l_3}{2a_c\sigma_c^2} - \frac{1}{a_c^2}\right) \tag{3.5}$$

with $l_k = L^{(k)}(a_c)$ and $h_k = H^{(k)}(a_c)$. More precisely,

$$b_{c,1} = \frac{-2c(12\theta^4c^4 + 12\theta^3c^3 + 35\theta^2c^2 + 4\theta c + 2)}{(4\theta^2c^2 - 1)^2}. \tag{3.6}$$

4. Maximum likelihood estimator. Florens-Landais and Pham [14] proved the following LDP for $(\hat{\theta}_T)$.

LEMMA 4.1. $(\hat{\theta}_T)$ *satisfies an LDP with good rate function*

$$I(c) = \begin{cases} -\dfrac{(c-\theta)^2}{4c} & \text{if } c < \dfrac{\theta}{3}, \\ 2c-\theta & \text{otherwise.} \end{cases} \tag{4.1}$$

The rate function I has a unique minimum in $c=\theta$ which is the a.s. limit given in (1.5). The large deviation properties of $(\hat{\theta}_T)$ are related to the ones of

$$Z_T(c) = \int_0^T X_t\, dX_t - c\int_0^T X_t^2\, dt \tag{4.2}$$

with $c \in \mathbf{R}$ since $\mathbf{P}\{\hat{\theta}_T \geqq c\} = \mathbf{P}\{Z_T(c) \geqq 0\}$. One has to keep in mind that the threshold c for $\hat{\theta}_T$ appears as a parameter for $Z_T(c)$. As before, we also improve Lemma 4.1 by an SLDP for $(\hat{\theta}_T)$. Define $\Gamma = \{a \in \mathbf{R} \mid \theta^2 + 2ac > 0 \text{ and } \theta + a < \sqrt{\theta^2+2ac}\}$ and set, for all $a \in \Gamma$,

$$L(a) = \mathcal{L}(a, -ca) \quad \text{and} \quad H(a) = \mathcal{H}(a, -ca). \tag{4.3}$$

The rate function I given by (4.1) is $I(c) = -\inf_a L(a)$. It is easy to check that

$$\Gamma = \begin{cases} \left]-\infty, -\dfrac{\theta^2}{2c}\right[& \text{if } c \leqq \dfrac{\theta}{2}, \\ ]-\infty, a^c[& \text{if } \dfrac{\theta}{2} < c \leqq 0, \\ \left]-\dfrac{\theta^2}{2c}, a^c\right[& \text{if } c > 0 \end{cases} \tag{4.4}$$

with $a^c = 2(c-\theta)$. The main difficulty in comparing this with the previous section is that L is not always steep. Actually, for all $a \in \Gamma$

$$L(a) = -\frac{1}{2}\left(a + \theta + \sqrt{\theta^2+2ac}\right), \quad L'(a) = -\frac{1}{2}\left(1 + \frac{c}{\sqrt{\theta^2+2ac}}\right). \tag{4.5}$$

Consequently, for $c \leqq \theta/2$, L is steep while this is no longer true for $c > \theta/2$ since $L'(a^c) = -\frac{1}{2}((3c-\theta)/(2c-\theta))$. Moreover, $L'(a) = 0$ if and only if $a = a_c$ with $a_c = (c^2-\theta^2)/(2c)$ and $a_c \in \Gamma$ only when $c < \theta/3$. Therefore, $I(c) = -L(a_c)$ if $c < \theta/3$ and $I(c) = -L(a^c)$ otherwise.

THEOREM 4.1. $(\hat{\theta}_T)$ *satisfies an SLDP associated with L and H. More precisely, for all $\theta < c < \theta/3$, there exists a sequence $(b_{c,k})$ such that, for any $p > 0$ and T large enough,*

$$\mathbf{P}\{\hat{\theta}_T \geqq c\} = \frac{e^{-TI(c)+H(a_c)}}{\sigma_c a_c \sqrt{2\pi T}}\left[1 + \sum_{k=1}^{p} \frac{b_{c,k}}{T^k} + \mathcal{O}\left(\frac{1}{T^{p+1}}\right)\right] \tag{4.6}$$

with

$$a_c = \frac{c^2-\theta^2}{2c}, \quad H(a_c) = -\frac{1}{2}\log\frac{(c+\theta)(3c-\theta)}{4c^2}, \tag{4.7}$$

and $\sigma_c^2 = L''(a_c) = -(2c)^{-1}$. *The coefficients* $b_{c,1}, b_{c,2}, \dots, b_{c,p}$ *may be explicitly given as in Theorem* 3.1. *Furthermore, for* $c > \theta/3$ *with* $c \neq 0$, *there exists a sequence* $(d_{c,k})$ *such that, for any* $p > 0$ *and* T *large enough,*

$$\mathbf{P}\{\hat{\theta}_T \geqq c\} = \frac{e^{-TI(c)+K(c)}}{\sigma^c a^c \sqrt{2\pi T}} \left[1 + \sum_{k=1}^{p} \frac{d_{c,k}}{T^k} + \mathcal{O}\left(\frac{1}{T^{p+1}}\right)\right] \tag{4.8}$$

with $a^c = 2(c-\theta)$,

$$(\sigma^c)^2 = L''(a^c) = \frac{c^2}{2(2c-\theta)^3}, \qquad K(c) = -\frac{1}{2}\log\frac{(c-\theta)(3c-\theta)}{4c^2}. \tag{4.9}$$

The coefficients $d_{c,1}, d_{c,2}, \dots, d_{c,p}$ *may be explicitly calculated. For example,*

$$d_{c,1} = \frac{2c^4 - c^3(19+3\theta) + c^2\theta(23+\theta) - 12c\theta^2 + 2\theta^3}{4(c-\theta)(2c-\theta)(3c-\theta)^2}. \tag{4.10}$$

Finally, for $c = 0$, $p > 0$, *and for* T *large enough,*

$$\mathbf{P}\{\hat{\theta}_T \geqq 0\} = 2\frac{e^{-TI(c)}}{\sqrt{2\pi T}\sqrt{-2\theta}} \left[1 + \sum_{k=1}^{p} \frac{(2k)!}{2^{2k}\theta^k T^k k!} + \mathcal{O}\left(\frac{1}{T^{p+1}}\right)\right]. \tag{4.11}$$

Remark. As $\mathbf{P}\{\hat{\theta}_T \geqq 0\} = \mathbf{P}\{X_T^2 \geqq T\}$ and X_T is Gaussian, (4.11) immediately follows from [13] (see relation (7.1), p. 193). It is easy to check that, for $c \to 0$, the main part of (4.8) and the first coefficient $d_{c,1}$ given by (4.10) coincide with the corresponding terms in (4.11).

THEOREM 4.2. *For* $c = \theta/3$, *there exists a sequence* (d_k) *such that, for any* $p > 0$ *and* T *large enough,*

$$\mathbf{P}\{\hat{\theta}_T \geqq c\} = \frac{e^{-TI(c)}}{2\pi T^{1/4}} \frac{\Gamma(1/4)}{a_\theta^{3/4}\sigma_\theta} \left[1 + \sum_{k=1}^{2p} \frac{d_k}{(\sqrt{T})^k} + \mathcal{O}\left(\frac{1}{T^p\sqrt{T}}\right)\right] \tag{4.12}$$

with $a_\theta = -4\theta/3$ *and* $\sigma_\theta^2 = L''(a_\theta) = -3/(2\theta)$. *As before, the coefficients* $d_1, d_2, \dots, d_{2p}$ *may be explicitly calculated.*

5. Log-likelihood ratio. If we wish to test H_0: $\theta = \theta_0$ versus H_1: $\theta = \theta_1$ for the Ornstein–Uhlenbeck process (1.1), then the most powerful test is based on the log-likelihood ratio (V_T).

LEMMA 5.1. *Under the hypothesis* H_0, $(T^{-1}V_T)$ *satisfies an LDP with good rate function*

$$I(c) = \begin{cases} -\dfrac{1}{8}\dfrac{((\theta_0-\theta_1)^2 + 4\theta_0 c)^2}{(\theta_0-\theta_1+2c)(\theta_0^2-\theta_1^2)} & \text{if } \dfrac{2c}{\theta_0-\theta_1} > -1, \\ +\infty & \text{otherwise.} \end{cases} \tag{5.1}$$

The rate function I has a unique minimum in $c = -(\theta_0-\theta_1)^2/(4\theta_0)$ which is the a.s. limit given in (1.5). For all $a \in \mathbf{R}$ such that $\theta_0^2 + (\theta_0^2 - \theta_1^2) \times a > 0$, set

$$\begin{aligned} L(a) &= \mathcal{L}\left(a(\theta_0-\theta_1),\ -\frac{1}{2}a(\theta_0^2-\theta_1^2)\right), \\ H(a) &= \mathcal{L}\left(a(\theta_0-\theta_1),\ -\frac{1}{2}a(\theta_0^2-\theta_1^2)\right). \end{aligned} \tag{5.2}$$

As in section 2, the function L is steep. Therefore, we obtain an SLDP for (V_T) similar to Theorem 3.1.

THEOREM 5.1. *Under the hypothesis H_0, $(T^{-1}V_T)$ satisfies an SLDP associated with L and H. More precisely, for all $c > -(\theta_0 - \theta_1)^2/(4\theta_0)$, there exists a sequence $(b_{c,k})$ such that, for any $p > 0$ and T large enough,*

$$\mathbf{P}\{V_T \geqq cT\} = \frac{e^{-TI(c)+H(a_c)}}{\sigma_c a_c \sqrt{2\pi T}} \left[1 + \sum_{k=1}^{p} \frac{b_{c,k}}{T^k} + \mathcal{O}\Big(\frac{1}{T^{p+1}}\Big)\right] \tag{5.3}$$

with a_c given by $L'(a_c) = c$ and

$$\sigma_c^2 = L''(a_c) = -\frac{(\theta_0 - \theta_1 + 2c)^3}{\theta_0^2 - \theta_1^2}. \tag{5.4}$$

The coefficients $b_{c,1}, b_{c,2}, \ldots, b_{c,p}$ may be explicitly given as in Theorem 3.1.

6. Proofs of the main results.

6.1. Proof of Theorem 3.1. Let L_T be the normalized c.g.f. of S_T.
For all $c > -1/(2\theta)$, set $\mathbf{P}\{S_T \geqq cT\} = A_T B_T$ with

$$A_T = \exp\big[T(L_T(a_c) - ca_c)\big], \tag{6.1}$$

$$B_T = \mathbf{E}_T\Big(\exp\big[-a_c(S_T - cT)\big]\, \mathbb{1}_{\{S_T \geqq cT\}}\Big), \tag{6.2}$$

where $\mathbf{E}_T$ is the expectation after the usual change of probability

$$\frac{d\mathbf{Q}_T}{d\mathbf{P}} = \exp\big(a_c S_T - T L_T(a_c)\big). \tag{6.3}$$

For all $a < \theta^2/2$, set $R_T(a) = \mathcal{R}_T(0, a)$. It follows from part (ii) of Lemma 2.1 that $R_T(a_c) = \mathcal{O}(e^{-T/|c|})$. In addition, we also have from (2.1) together with (3.2) that

$$A_T = \exp\big[-TI(c) + H(a_c)\big] \Big(1 + \mathcal{O}(e^{-T/|c|})\Big). \tag{6.4}$$

It now remains to give an expansion of B_T which can be rewritten as

$$B_T = \mathbf{E}_T\Big(\exp\Big[-a_c\sigma_c\sqrt{T}U_T\Big]\, \mathbb{1}_{\{U_T \geqq 0\}}\Big), \quad \text{where } U_T = \frac{S_T - cT}{\sigma_c\sqrt{T}}. \tag{6.5}$$

LEMMA 6.1. *For $c > -1/(2\theta)$, the distribution of U_T under $\mathbf{Q}_T$ converges, as T goes to infinity, to the $N(0,1)$ distribution. Moreover, there exists a sequence (δ_k) such that, for any $p > 0$ and T large enough,*

$$B_T = \frac{1}{\sqrt{T}} \left(\sum_{k=0}^{p} \frac{\delta_k}{T^k} + \mathcal{O}\Big(\frac{1}{T^{p+1}}\Big)\right). \tag{6.6}$$

The sequence (δ_k) depends only on the derivatives of L and H at point a_c. For example,

$$\delta_0 = \frac{1}{\sigma_c a_c \sqrt{2\pi}} \quad \text{and} \quad \delta_1 = \frac{-2c\delta_1(12\theta^4c^4 + 12\theta^3c^3 + 35\theta^2c^2 + 4\theta c + 2)}{(4\theta^2c^2 - 1)^2}.$$

Proof of Theorem 3.1. Relation (3.3) follows from the conjunction of (6.4) and (6.6), which completes the proof of Theorem 3.1.

6.2. Proof of Theorem 4.1. This proof is completely different from the previous one since the function L is steep for $\theta < c \leqq \theta/2$, while this is no longer true when $c > \theta/2$. The point a_c, given by (4.7), belongs to Γ whenever $c < \theta/3$. First, if $\theta < c \leqq \theta/3$, we prove (4.6) as (3.3) via the usual change of probability.

Next, if $c > \theta/3$, we use a slight modification of the strategy of time varying change of probability proposed in [10] and [8]. Let L_T be the normalized c.g.f. of $Z_T(c)$, where the parameter c is omitted in order to simplify the notation. There is a unique a_T, which belongs to Γ and converges to a^c as $T \to \infty$, solution of

$$L'(a_T) + \frac{H'(a_T)}{T} = 0. \tag{6.7}$$

Set

$$\frac{d\mathbf{Q}_T}{d\mathbf{P}} = \exp\big(a_T Z_T(c) - T L_T(a_T)\big) \tag{6.8}$$

and denote by $\mathbf{E}_T$ the expectation under this new probability. We have the decomposition $\mathbf{P}\{\hat{\theta}_T \geqq c\} = A_T B_T$ with

$$A_T = \exp\big[T L_T(a_T)\big], \tag{6.9}$$

$$B_T = \mathbf{E}_T\Big(\exp\big[-a_T Z_T(c)\big]\,\mathbb{1}_{\{Z_T(c) \geqq 0\}}\Big). \tag{6.10}$$

It follows from the identity (6.7) together with (4.3) that

$$T\big(\rho(a_T) - (a_T + \theta)\big) = \frac{\theta^2 - c\theta + a_T c}{\rho(a_T)(\rho(a_T) + c)}. \tag{6.11}$$

Since $\rho(a_T) = \sqrt{\theta^2 + 2a_T c}$, we have $\lim_{T\to\infty} \rho(a_T) = 2c - \theta$, and (6.11) immediately implies

$$\lim_{T\to\infty} T\big(\rho(a_T) - (a_T + \theta)\big) = \frac{c - \theta}{3c - \theta}. \tag{6.12}$$

From (6.7), there exists a sequence (a_k) such that, for any $p > 0$ and T large enough,

$$a_T = a^c + \sum_{k=1}^{p} \frac{a_k}{T^k} + \mathcal{O}\Big(\frac{1}{T^{p+1}}\Big). \tag{6.13}$$

For example, by (6.12),

$$a_1 = -\frac{2c - \theta}{3c - \theta} \quad \text{and} \quad a_2 = -\frac{c(c^2 - 5\theta c + 2\theta^2)}{2(c - \theta)(3c - \theta)^3}.$$

Moreover, for all $a \in \Gamma$, set $R_T(a) = \mathcal{R}_T(a, -ca)$. Using (4.3), relation (2.1) can be rewritten as

$$L_T(a) = L(a) + \frac{H(a)}{T} + \frac{1}{T} R_T(a) \tag{6.14}$$

with $a \in \Gamma$. Next, by part (ii) of Lemma 2.1, the remainder $R_T(a_T) = \mathcal{O}(Te^{-2(2c-\theta)T})$. Thus, we obtain from (6.9) and (6.14) that

$$A_T = \exp\big[T L(a_T) + H(a_T)\big]\,\big(1 + \mathcal{O}\big(Te^{-2(2c-\theta)T}\big)\big). \tag{6.15}$$

On the one hand, it follows from (6.13) that there exists a sequence (α_k) such that, for any $p > 0$ and T large enough,

$$TL(a_T) = TL(a^c) + \frac{1}{2} + \sum_{k=1}^{p} \frac{\alpha_k}{T^k} + \mathcal{O}\left(\frac{1}{T^{p+1}}\right). \tag{6.16}$$

The sequence (α_k) depends only on (a_k) and on the derivatives of L at point a^c. For example, $\alpha_1 = a_2 L'(a^c) + \frac{1}{2} a_1^2 L''(a^c)$ so that

$$\alpha_1 = \frac{c(2c^3 + c^2(1-3\theta) + \theta c(\theta - 5) + 2\theta^2)}{4(c-\theta)(2c-\theta)(3c-\theta)^2}.$$

On the other hand, we also have from (2.2) and (4.3) that

$$\exp\big[H(a_T)\big] = \sqrt{\frac{2T\rho(a_T)}{T(\rho(a_T) - (a_T + \theta))}}. \tag{6.17}$$

Hence, from (6.12) together with (6.13), there exists a sequence (β_k) such that, for any $p > 0$ and T large enough,

$$\exp\big[H(a_T)\big] = \sqrt{T}\sqrt{\frac{2(2c-\theta)(3c-\theta)}{c-\theta}}\left(1 + \sum_{k=1}^{p} \frac{\beta_k}{T^k} + \mathcal{O}\left(\frac{1}{T^{p+1}}\right)\right). \tag{6.18}$$

The sequence (β_k) can be explicitly given as (α_k). For example,

$$\beta_1 = -\frac{c(c^2 - 3\theta c + \theta^2)}{(c-\theta)(2c-\theta)(3c-\theta)^2}.$$

Finally, from (6.15) together with (6.16) and (6.18) it follows that there exists a sequence (γ_k) such that, for any $p > 0$ and T large enough,

$$A_T = \exp\big[-TI(c)\big]\sqrt{eT}\sqrt{\frac{2(2c-\theta)(3c-\theta)}{c-\theta}}\left(1 + \sum_{k=1}^{p} \frac{\gamma_k}{T^k} + \mathcal{O}\left(\frac{1}{T^{p+1}}\right)\right). \tag{6.19}$$

The sequence (γ_k) depends only on (a_k) and on the derivatives of L at point a^c. For example, $\gamma_1 = \alpha_1 + \beta_1$ so that

$$\gamma_1 = \frac{c(2c^3 - 3c^2(1+\theta) + \theta c(\theta + 7) - 2\theta^2)}{4(c-\theta)(2c-\theta)(3c-\theta)^2}.$$

For $c > \theta/3$, the following lemma gives the asymptotical behavior of the distribution of U_T and an expansion of B_T which can be rewritten as

$$B_T = \mathbf{E}_T\big(\exp[-a_T T U_T]\, \mathbb{1}_{U_T \geq 0}\big), \quad \text{where } U_T = \frac{Z_T(c)}{T}. \tag{6.20}$$

Lemma 6.2. (i) *The distribution of* U_T *under* $\mathbf{Q}_T$ *converges, as* T *goes to infinity, to the distribution of* $\gamma(N^2 - 1)$, *where* N *is an* $N(0,1)$ *random variable and* $\gamma = -L'(a^c) = (3c-\theta)/(2(2c-\theta))$; *i.e., the limit of the characteristic function of* U_T *under* $\mathbf{Q}_T$ *is*

$$\Phi(u) = \frac{\exp(-i\gamma u)}{\sqrt{1 - 2i\gamma u}}. \tag{6.21}$$

(ii) *There exists a sequence* (δ_k) *such that, for any* $p > 0$ *and* T *large enough,*

$$B_T = \sum_{k=1}^{p} \frac{\delta_k}{T^k} + \mathcal{O}\left(\frac{1}{T^{p+1}}\right). \tag{6.22}$$

The sequence (δ_k) *depends only on the Taylor expansion of* a_T *at the neighborhood of* a^c *together with the derivatives of* L *at point* a^c. *For example,*

$$\delta_1 = \frac{1}{a^c \gamma \sqrt{2\pi e}} \quad \textit{and} \quad \delta_2 = \frac{\delta_1}{\gamma} \frac{-8c^3 + 8c^2\theta - 5c\theta^2 + \theta^3}{2(c-\theta)(2c-\theta)(3c-\theta)^2}.$$

Remark. Part (i) was previously proven in [14] (see Lemma 4.6, p. 17). Part (ii) is not a consequence of (i). It requires a more precise study of the convergence of Φ_T towards Φ given in Lemma 7.2 below.

Proof of Theorem 4.1. Relation (4.8) follows from the conjunction of (6.19) and (6.22), which completes the proof of Theorem 4.1.

6.3. Proof of Theorem 4.2. We follow the same approach as that of Theorem 4.1. Assume that $c = \theta/3$ so that $a_c = a^c = a_\theta$ with $a_\theta = -4\theta/3$. There is a unique a_T solution, which belongs to Γ and converges to a_θ as $T \to \infty$, of the equation

$$L'(a_T) + \frac{H'(a_T)}{T} = 0. \tag{6.23}$$

From (2.2) together with (4.3), it is easy to see that

$$\lim_{T\to\infty} \big(\rho(a_T) - (a_T + \theta)\big)\, H'(a_T) = 1. \tag{6.24}$$

Moreover, a Taylor expansion around a_θ gives $L'(a_T) = (a_T - a_\theta)\sigma_\theta^2 + o(a_T - a_\theta)$ with $\sigma_\theta^2 = L''(a_\theta) = -3/(2\theta)$. Thus, we find from (6.23) together with (6.24)

$$\lim_{T\to\infty} T(a_T - a_\theta)^2 = \frac{1}{2\sigma_\theta^2}. \tag{6.25}$$

Next, as in section 6.2, there exists a sequence (a_k) such that, for any $p > 0$ and T large enough,

$$a_T = a_\theta + \sum_{k=1}^{2p} \frac{a_k}{(\sqrt{T})^k} + \mathcal{O}\left(\frac{1}{T^p\sqrt{T}}\right). \tag{6.26}$$

For example, $a_1 = -1/(\sigma_\theta\sqrt{2})$ and $a_2 = -\frac{1}{8}$. In addition, if we use the decomposition $\mathbf{P}\{\hat{\theta}_T \geqq c\} = A_T B_T$ given by (6.9) and (6.10), we obtain, as in (6.19), that for any $p > 0$ and T large enough

$$A_T = \exp\big[-TI(c)\big]\,(eT)^{1/4}\left(-\frac{\theta}{3}\right)^{1/4}\left(1 + \sum_{k=1}^{2p} \frac{\gamma_k}{(\sqrt{T})^k} + \mathcal{O}\left(\frac{1}{T^p\sqrt{T}}\right)\right), \tag{6.27}$$

where the sequence (γ_k) may be explicitly calculated. For $c = \theta/3$, it now remains to give an expansion of B_T which can be rewritten as

$$B_T = \mathbf{E}_T\Big(\exp[-a_T\sqrt{T}U_T]\,\mathbb{1}_{\{U_T \geqq 0\}}\Big), \quad \text{where } U_T = \frac{Z_T(c)}{\sqrt{T}}. \tag{6.28}$$

LEMMA 6.3. (i) *The distribution of* U_T *under* $\mathbf{Q}_T$ *converges, as* T *goes to infinity, to the distribution of* $\sigma_\theta N_1 + \eta_\theta(N_2^2 - 1)$, *where* N_1 *and* N_2 *are independent* $N(0,1)$ *random variables and* $\sigma_\theta^2 = L''(a_\theta) = -3/(2\theta)$ *and* $\eta_\theta = \sigma_\theta/\sqrt{2}$; *i.e., the limit of the characteristic function of* U_T *under* $\mathbf{Q}_T$ *is*

$$\Phi(u) = \frac{\exp(-i\eta_\theta u - \sigma_\theta^2 u^2/2)}{\sqrt{1 - 2i\eta_\theta u}}. \tag{6.29}$$

(ii) *For any* $p > 0$ *and* T *large enough*

$$B_T = \sum_{k=1}^{2p} \frac{\delta_k}{(\sqrt{T})^k} + \mathcal{O}\left(\frac{1}{T^p\sqrt{T}}\right), \tag{6.30}$$

where the sequence (δ_k) *may be explicitly calculated. For example,*

$$\delta_1 = \frac{1}{4\pi a_\theta \eta_\theta}\, e^{-1/4}\Gamma\left(\frac{1}{4}\right).$$

Proof of Theorem 4.2. Relation (4.12) follows from the conjunction of (6.27) and (6.30), which completes the proof of Theorem 4.2.

6.4. Proof of Theorem 5.1. The proof follows exactly along the same lines as that of Theorem 3.1 since the function L associated with the log-likelihood ratio (V_T) is steep.

7. Proofs of lemmas.

7.1. Proof of Lemma 2.2. (i) We refer to Janson [17, Chaps. 2 and 6]. The two random variables $Y_1 = X_T^2/2$ and $Y_2 = \int_0^T X_t^2\,dt$ are in the chaos $H^{:2:} \oplus H^{:0:}$. The operators B_1 and B_2 associated with their bilinear forms are positive, trace class, and their traces are expectations $\mathrm{tr}B_1 = \mathbf{E}Y_1$ and $\mathrm{tr}B_2 = \mathbf{E}Y_2$. By linearity, for any $(a, b) \in \mathbf{R}^2$, $a(Y_1 - \mathbf{E}Y_1) + b(Y_2 - \mathbf{E}Y_2)$ belongs to $H^{:2:}$. Consequently, from Theorem 6.1 of [17, p. 78], we have the decomposition

$$a(Y_1 - \mathbf{E}Y_1) + b(Y_2 - \mathbf{E}Y_2) = \sum_{j=1}^{\infty} \lambda_j^T(\varepsilon_j^2 - 1), \tag{7.1}$$

where (ε_j) are independent $N(0,1)$ random variables and (λ_j^T) are the eigenvalues of $aB_1 + bB_2$. Since this operator is trace class, part (i) of Lemma 2.2 follows from (7.1).

(ii) The spectrum $\sigma(B_2) \subset [0, \theta^{-2}]$ (see [8]) and $\mathbf{E}X_T^2/2$ is uniformly bounded in T.

(iii) From (1.6), (1.7), and (2.5), the normalized c.g.f. of $\mathcal{Z}_T(a, b)$ in α is $\mathcal{L}_T(\alpha a, \alpha b)$. From part (i), it can be rewritten as

$$\mathcal{L}_T(\alpha a, \alpha b) = -\frac{\alpha a}{2} - \frac{1}{2}\,\langle \nu_T, \Psi_\alpha \rangle, \tag{7.2}$$

where ν_T is given by (2.6) and $\Psi_\alpha(x) = \log(1 - 2\alpha x)$ (notice that $\Psi_\alpha \in \mathcal{F}$). From Lemma 2.1, we already know that

$$\lim_{T\to\infty} \mathcal{L}_T(\alpha a, \alpha b) = \mathcal{L}(\alpha a, \alpha b) = -\frac{1}{2}\big(\alpha a + \theta + \rho(\alpha b)\big) \tag{7.3}$$

for α in some interval $[-\alpha_0, \alpha_0]$ depending on a and b, with $\alpha_0 > 0$. Thus, it immediately follows from (7.2) and (7.3) that $\lim_{T\to\infty}\langle \nu_T, \Psi_\alpha\rangle = \theta + \rho(\alpha b)$. Since ν is defined by $\langle \nu, h\rangle = 1/(2\pi)\int_{\mathbf{R}} h(bg(x))\,dx$, where h is any continuous function with compact support, an easy integration by parts gives $\langle \nu, \Psi_\alpha\rangle = \theta + \rho(\alpha b)$ which proves that

$$\lim_{T\to\infty}\langle \nu_T, \Psi_\alpha\rangle = \langle \nu, \Psi_\alpha\rangle \tag{7.4}$$

for every $\alpha \in [-\alpha_0, \alpha_0]$. By use of the above part (ii) together with a slight modification of Lemma 9 of [4], we arrive from (7.4) at the convergence of ν_T towards ν, which completes the proof of Lemma 2.2.

7.2. Proof of Lemma 6.1. If Φ_T is the characteristic function of U_T under $\mathbf{Q}_T$, (6.3) immediately implies

$$\Phi_T(u) = \exp\left[-\frac{iu\sqrt{T}c}{\sigma_c} + T\left(L_T\left(a_c + \frac{iu}{\sigma_c\sqrt{T}}\right) - L_T(a_c)\right)\right] \tag{7.5}$$

so that

$$\left|\Phi_T(u)\right|^2 = \prod_{j=1}^{\infty}\left(1 + \frac{4u^2(\lambda_j^T)^2}{\sigma_c^2 T(1-2a_c\lambda_j^T)^2}\right)^{-1/2}. \tag{7.6}$$

Choose $\varepsilon > 0$ such that $1 - 2a_c\varepsilon > 0$ and let $q_T = \operatorname{card}\{\lambda_j^T \mid \lambda_j^T > \varepsilon\}$. Then, by Lemma 2.2, part (iii), there exists some positive constant η, depending only on ε, such that

$$\liminf_{T\to\infty}\frac{q_T}{T} \geqq \eta. \tag{7.7}$$

Thus, it follows from (7.6) together with (7.7) that, for T large enough,

$$\left|\Phi_T(u)\right|^2 \leqq \left(1 + \frac{\xi u^2}{T}\right)^{-\eta T/2} \tag{7.8}$$

with $\xi > 0$. Thereby, for T large enough, $\Phi_T \in L^2(\mathbf{R})$ and we can use the Parseval formula in (6.5) to get

$$B_T = \frac{1}{2\pi a_c\sigma_c\sqrt{T}}\int_{\mathbf{R}}\left(1 + \frac{iu}{a_c\sigma_c\sqrt{T}}\right)^{-1}\Phi_T(u)\,du. \tag{7.9}$$

For some positive constant s, set $s_T = sT^{1/6}$. We separate $B_T = C_T + D_T$, where

$$C_T = \frac{1}{2\pi a_c\sigma_c\sqrt{T}}\int_{|u|\leqq s_T}\left(1 + \frac{iu}{a_c\sigma_c\sqrt{T}}\right)^{-1}\Phi_T(u)\,du, \tag{7.10}$$

$$D_T = \frac{1}{2\pi a_c\sigma_c\sqrt{T}}\int_{|u|> s_T}\left(1 + \frac{iu}{a_c\sigma_c\sqrt{T}}\right)^{-1}\Phi_T(u)\,du. \tag{7.11}$$

On the one hand, we find from (7.8) that

$$|D_T| = O\left(\left(1 + \frac{\xi s_T^2}{T}\right)^{-\eta T/4}\right) \tag{7.12}$$

so that $|D_T| = \mathcal{O}(e^{-\mu T^{1/3}})$ with $\mu > 0$. On the other hand, it follows from (2.3) that for any $k \in \mathbf{N}$, $R_T^{(k)}(a_c) = \mathcal{O}(T^k e^{-T/|c|})$. Then, using (3.2) and (2.1) we get

$$L_T^{(k)}(a_c) = L^{(k)}(a_c) + \frac{H^{(k)}(a_c)}{T} + \mathcal{O}(T^k e^{-T/|c|}). \tag{7.13}$$

Consequently, via (7.5) together with (7.13) we can prove the following Taylor expansion for Φ_T. Similar expansion was established in the i.i.d. case by Cramer [7] and Esseen [12].

LEMMA 7.1. *For any $p > 0$, there exist integers $q(p)$, $r(p)$, and a sequence $(\varphi_{k,l})$ independent of p such that, for T large enough,*

$$\Phi_T(u) = \exp\left(-\frac{u^2}{2}\right)\left[1 + \frac{1}{\sqrt{T}}\sum_{k=0}^{2p}\sum_{l=k+1}^{q(p)} \frac{\varphi_{k,l}u^l}{(\sqrt{T})^k} + \mathcal{O}\left(\frac{\max(1,|u|^{r(p)})}{T^{p+1}}\right)\right] \tag{7.14}$$

and the remainder $\mathcal{O}$ is uniform as soon as $|u| = \mathcal{O}(T^{1/6})$.

Proof. From (7.5) and (7.13), we have

$$\begin{aligned}\log \Phi_T(u) = &-\frac{u^2}{2} + T\sum_{k=3}^{2p+3}\left(\frac{iu}{\sigma_c\sqrt{T}}\right)^k \frac{L^{(k)}(a_c)}{k!} + \sum_{k=1}^{2p+1}\left(\frac{iu}{\sigma_c\sqrt{T}}\right)^k \frac{H^{(k)}(a_c)}{k!}\\ &+ \mathcal{O}\left(\frac{\max(1,u^{2p+4})}{T^{p+1}}\right).\end{aligned} \tag{7.15}$$

Thus, we follow the same approach as Cramer (see [7, Lemma 2, p. 72]), remarking that in the range $|u| = \mathcal{O}(T^{1/6})$ the quantity $u^l/(\sqrt{T})^k$ remains bounded in (7.14). Lemma 7.1 is proved.

Proof of Lemma 6.1. Relation (6.6) follows from (7.10) and (7.14) together with standard calculus on the $N(0,1)$ distribution.

7.3. Proof of Lemma 6.2. Let Φ_T be the characteristic function of U_T under $\mathbf{Q}_T$. We have from (6.8)

$$\Phi_T(u) = \exp\left[T\left(L_T\left(a_T + \frac{iu}{T}\right) - L_T(a_T)\right)\right] \tag{7.16}$$

so that

$$\left|\Phi_T(u)\right|^2 = \prod_{j=1}^{\infty}\left(1 + \frac{4u^2(\lambda_j^T)^2}{T^2(1-2a_T\lambda_j^T)^2}\right)^{-1/2}. \tag{7.17}$$

As in (7.8), it follows from Lemma 2.2, part (iii) that, for some positive constants η, ξ, and for T large enough,

$$\left|\Phi_T(u)\right|^2 \leqq \left(1 + \frac{\xi u^2}{T^2}\right)^{-\eta T}. \tag{7.18}$$

Thereby, for T large enough, $\Phi_T \in L^2(\mathbf{R})$ and we can use the Parseval formula in (6.20) to obtain that

$$B_T = \frac{1}{2\pi T a_T}\int_{\mathbf{R}}\left(1 + \frac{iu}{Ta_T}\right)^{-1}\Phi_T(u)\,du. \tag{7.19}$$

Let $s_T > 0$ be such that $\sqrt{T} = o(s_T)$ as $T \to \infty$. We split B_T into two terms, $B_T = C_T + D_T$, where

$$C_T = \frac{1}{2\pi T a_T} \int_{|u| \leqq s_T} \left(1 + \frac{iu}{Ta_T}\right)^{-1} \Phi_T(u)\, du, \tag{7.20}$$

$$D_T = \frac{1}{2\pi T a_T} \int_{|u| > s_T} \left(1 + \frac{iu}{Ta_T}\right)^{-1} \Phi_T(u)\, du. \tag{7.21}$$

First, we find as in (7.12) that, if $s_T = o(T)$, $|D_T| = \mathcal{O}(\exp(-\mu s_T^2/T))$ with $\mu > 0$. It now remains to precisely evaluate C_T via the following Taylor expansion for Φ_T.

LEMMA 7.2. *For any $p > 0$, there exist integers $q(p)$, $r(p)$, $s(p)$, and a sequence $(\varphi_{k,l,m})$ independent of p such that, for T large enough,*

$$\Phi_T(u) = \Phi(u) \exp\left(-\frac{\sigma^2 u^2}{2T}\right) \times \left[1 + \sum_{k=1}^{p} \sum_{l=k+1}^{q(p)} \sum_{m=0}^{r(p)} \frac{\varphi_{k,l,m} u^l}{T^k (1 - 2i\gamma u)^m} + \mathcal{O}\left(\frac{\max(1, |u|^{s(p)})}{T^{p+1}}\right)\right],$$

where Φ is given by (6.21),

$$\gamma = -L'(a^c) = \frac{3c - \theta}{2(2c - \theta)} \quad \text{and} \quad \sigma^2 = L''(a^c) = \frac{c^2}{2(2c - \theta)^3}.$$

Moreover, the remainder $\mathcal{O}$ is uniform as soon as $|u| = o(T^{2/3})$.

Remark. One can see in this asymptotic expansion the limit χ^2 distribution Φ together with an independent centered Gaussian distribution with small variance σ^2/T.

Proof. From (6.14) together with (7.16)

$$\Phi_T(u) = \exp\left[T\left(L\left(a_T + \frac{iu}{T}\right) - L(a_T)\right)\right] \times \exp\left[H\left(a_T + \frac{iu}{T}\right) - H(a_T)\right] (1 + \mathcal{O}(e^{-\xi T}))$$

with $\xi > 0$. On the one hand, (4.5) immediately implies

$$T\left(L\left(a_T + \frac{iu}{T}\right) - L(a_T)\right) = -\frac{iu}{2} - \frac{T}{2}\rho(a_T)\left(\left(1 + \frac{iub_T}{T}\right)^{1/2} - 1\right), \tag{7.22}$$

where $b_T = 2c(\rho(a_T))^{-2}$. Consequently, for any $p \geqq 2$

$$\exp\left[T\left(L\left(a_T + \frac{iu}{T}\right) - L(a_T)\right)\right] = \exp\left(-\frac{iu}{4}(2 + \rho(a_T)\, b_T)\right) \times \exp\left(-\frac{T}{2}\rho(a_T) \sum_{k=2}^{p} l_k \left(\frac{iub_T}{T}\right)^k\right)\left(1 + \mathcal{O}\left(\frac{|u|^{p+1}}{T^{p+1}}\right)\right), \tag{7.23}$$

where $l_k = (-1)^{k-1}(2k)!/((2k-1)(2^k k!)^2)$. On the other hand, we obtain from (2.2)

$$\exp\left[H\left(a_T + \frac{iu}{T}\right) - H(a_T)\right] = \sqrt{\frac{\rho(a_T) - (a_T + \theta)}{\rho(a_T) - (a_T + iu/T + \theta)(1 + iub_T/T)^{-1/2}}}. \tag{7.24}$$

If $c_T = T(\rho(a_T) - (a_T + \theta))$ and $d_T(u) = 1 - iu/c_T + (a_T + \theta)iub_T/(2c_T)$, we have for any $p \geqq 2$

$$\exp\left[H\left(a_T + \frac{iu}{T}\right) - H(a_T)\right]$$
$$= \frac{1}{\sqrt{d_T(u)}}\left(1 - \frac{u^2 b_T}{2Tc_T d_T(u)} - \frac{T(a_T + \theta + iuT^{-1})}{c_T d_T(u)}\right.$$
$$\left.\times\left[\sum_{k=2}^{p} h_k\left(\frac{iub_T}{T}\right)^k + \mathcal{O}\left(\frac{|u|^{p+1}}{T^{p+1}}\right)\right]\right)^{-1/2} \tag{7.25}$$

with $h_k = (-1)^k(2k)!/(2^k k!)^2$. As T goes to infinity, the limits of a_T, b_T, c_T, and $d_T(u)$ are a^c, $2c/(2c-\theta)^2$, $(c-\theta)/(3c-\theta)$, and $1 - 2i\gamma u$, respectively. Therefore, we find by (7.23), together with (7.25), the pointwise convergence

$$\lim_{T\to\infty} \Phi_T(u) = \Phi(u) = \frac{\exp(-i\gamma u)}{\sqrt{1 - 2i\gamma u}} \tag{7.26}$$

which achieves the proof of the first part of Lemma 6.2. Finally, after tedious calculus, we prove Lemma 7.2 via a Taylor expansion of the exponential in (7.23) together with a Taylor expansion of the square root in (7.25).

End of the proof of Lemma 6.2. From (7.20) and the above expansion of Φ_T, we have for T large enough

$$2\pi T a_T C_T = \int_{|u|\leqq s_T} \Phi(u)\exp\left(-\frac{\sigma^2 u^2}{2T}\right)\left[1 + \mathcal{O}\left(\frac{\max(1, |u|^{s(p)})}{T^{p+1}}\right)\right] du$$
$$+ \sum_{k=1}^{p}\sum_{l=k+1}^{q(p)}\sum_{m=0}^{r(p)} \psi_{k,l,m}\int_{|u|\leqq s_T} \Phi(u)\exp\left(-\frac{\sigma^2 u^2}{2T}\right)\frac{u^l}{T^k(1-2i\gamma u)^m}\, du, \tag{7.27}$$

where the sequence $(\varphi_{k,l,m})$ is replaced by $(\psi_{k,l,m})$ due to the factor preceding Φ_T in (7.20). Furthermore, for all $u \in \mathbf{R}$, $|\Phi(u)| \leqq 1$ and

$$\int_{|u|>s_T} \exp\left(-\frac{\sigma^2 u^2}{2T}\right) du = \mathcal{O}\left(\exp\left(-\frac{\sigma^2 s_T^2}{2T}\right)\right).$$

Thus, we may change in (7.27) all the integrals by integrals over $\mathbf{R}$ at the cost of $\mathcal{O}(\exp(-\mu s_T^2/T))$ with $\mu > 0$. We complete the proof of (6.22) via the following lemma which gives the values of integrals of the type

$$\int_{\mathbf{R}} \exp\left(-iu\gamma - \frac{\sigma^2 u^2}{2T}\right)\frac{u^\beta}{(1-2i\gamma u)^\alpha}\, du.$$

Let f_α be the density of the gamma distribution with parameters α and $\frac{1}{2}$,

$$f_\alpha(x) = \begin{cases} \exp\left(-\dfrac{x}{2}\right)\dfrac{x^{\alpha-1}}{2^\alpha\Gamma(\alpha)} & \text{if } x > 0, \\ 0 & \text{otherwise.}\end{cases} \tag{7.28}$$

For any $p > 0$ and $0 \leqq k \leqq p$, set

$$v_k(\alpha, \beta) = \frac{2\pi\sigma^{2k} i^\beta}{\gamma^{2k+\beta+1} 2^k k!} f_\alpha^{(2k+\beta)}(1). \tag{7.29}$$

LEMMA 7.3. *For any $p>0$, we have*

$$\int_{\mathbf{R}} \exp\left(-iu\gamma - \frac{\sigma^2 u^2}{2T}\right) \frac{u^\beta}{(1-2i\gamma u)^\alpha}\,du = \sum_{k=0}^{p} \frac{v_k(\alpha,\beta)}{T^k} + \mathcal{O}\left(\frac{1}{T^{p+1}}\right). \tag{7.30}$$

Proof. Let $\widehat{f}_\alpha$ be the characteristic function of f_α, $\widehat{f}_\alpha(u) = 1/(1-2iu)^\alpha$. Denote by N_τ the Gaussian kernel of variance τ. First, we notice that

$$\int_{\mathbf{R}} \exp\left(-iv\xi - \frac{\tau v^2}{2}\right) v^\beta \widehat{f}_\alpha(v)\,dv = 2\pi i^\beta f_\alpha * N_\tau^{(\beta)}(\xi). \tag{7.31}$$

It is rather easy to see that, for every $p>0$,

$$f_\alpha * N_\tau^{(\beta)}(\xi) = \sum_{k=0}^{p} \frac{\tau^k}{2^k k!} f_\alpha^{(2k+\beta)}(\xi) + \mathcal{O}(\tau^{(p+1)}). \tag{7.32}$$

For the sake of completeness, let us prove (7.32). We start with a localization

$$\left| f_\alpha * N_\tau^{(\beta)}(\xi) - \int_{|x|\leqq\delta} f_\alpha(\xi-x)\,N_\tau^{(\beta)}(x)\,dx \right| \leqq \sup_{|x|>\delta} \left|N_\tau^{(\beta)}(x)\right|$$

with $0<\delta<\xi$. First, this supremum is $\mathcal{O}(\exp(-\mu_1\tau^{-1}))$ with $\mu_1>0$. Next, the first term is easily integrated by parts since

$$\int_{|x|\leqq\delta} f_\alpha(\xi-x)\,N_\tau^{(\beta)}(x)\,dx = \int_{|x|\leqq\delta} f_\alpha^{(\beta)}(\xi-x)\,N_\tau(x)\,dx + \mathcal{O}\big(\exp(-\mu_2\tau^{-1})\big)$$

with $\mu_2>0$. Furthermore, as the function $f_\alpha \in \mathcal{C}^\infty$ in $[\xi-\delta, \xi+\delta]$, we have the Taylor expansion

$$f_\alpha^{(\beta)}(\xi-x) = \sum_{k=0}^{2p+1} \frac{(-x)^k}{k!} f_\alpha^{(\beta+k)}(\xi) + \mathcal{O}(x^{2p+2})$$

for every $x\in[-\delta,+\delta]$. This yields

$$\begin{aligned}\int_{|x|\leqq\delta} f_\alpha(\xi-x)N_\tau^{(\beta)}(x)\,dx = \sum_{k=0}^{p} f_\alpha^{(2k+\beta)}(\xi)\,\frac{1}{(2k)!}\int_{|x|\leqq\delta} x^{2k}N_\tau(x)\,dx \\ +\mathcal{O}\left(\int_{|x|\leqq\delta} x^{2p+2}N_\tau(x)\,dx\right).\end{aligned}$$

Therefore, if we remove the localization whose cost is $\mathcal{O}(\exp(-\mu_3\tau^{-1}))$, we find that

$$f_\alpha * N_\tau^{(\beta)}(\xi) = \sum_{k=0}^{p} f_\alpha^{(2k+\beta)}(\xi)\,\frac{1}{(2k)!}\int_{\mathbf{R}} x^{2k}N_\tau(x)\,dx + \mathcal{O}\left(\int_{\mathbf{R}} x^{2p+2}N_\tau(x)\,dx\right)$$

which immediately leads to (7.32). Finally, by a change of variables in (7.31), we obtain

$$\int_{\mathbf{R}} \exp\left(-iu\gamma - \frac{\sigma^2u^2}{2T}\right)\frac{u^\beta}{(1-2i\gamma u)^\alpha}\,du = \frac{2\pi(-i)^\beta}{\gamma^{\beta+1}} f_\alpha * N_\tau^{(\beta)}(1) \tag{7.33}$$

with $\tau = \sigma^2/(T\gamma^2)$, which completes the proof of Lemma 7.3.

7.4. Proof of Lemma 6.3. Let Φ_T be the characteristic function of U_T under $\mathbf{Q}_T$ given by (6.8),

$$\Phi_T(u) = \exp\left[T\left(L_T\left(a_T + \frac{iu}{\sqrt{T}}\right) - L_T(a_T)\right)\right]. \tag{7.34}$$

We prove only the pointwise convergence of Φ_T given in (6.29), as the proof of the Taylor expansion (6.30) follows essentially the same arguments as those of section 7.3. From the definition (4.5) of L, we have

$$\lim_{T\to\infty} T\left(L\left(a_T + \frac{iu}{\sqrt{T}}\right) - L(a_T)\right) = -i\eta_\theta u - \frac{\sigma_\theta^2 u^2}{2}. \tag{7.35}$$

Moreover, we also have from the definition (2.2) of H

$$\lim_{T\to\infty} \exp\left[H\left(a_T + \frac{iu}{\sqrt{T}}\right) - H(a_T)\right] = \frac{1}{\sqrt{1-2i\eta_\theta u}}. \tag{7.36}$$

Therefore, the pointwise convergence (6.29) immediately follows from (2.1), (7.35), and (7.36). The first term in (6.30) is $\delta_1 = (2\pi a_\theta)^{-1}\int_{\mathbf{R}} \Phi(u)\,du$. By a change of variables together with the fact that $2\eta_\theta^2 = \sigma_\theta^2$, we find that

$$\delta_1 = \frac{1}{2\pi a_\theta \eta_\theta}\int_{\mathbf{R}} \frac{1}{\sqrt{1-2iu}} \exp(-(u^2+iu))\,du.$$

Finally, via a contour integral for the Gamma function, we obtain

$$\delta_1 = \frac{1}{4\pi a_\theta \eta_\theta} \exp\left(-\frac{1}{4}\right)\Gamma\left(\frac{1}{4}\right)$$

which completes the proof of Lemma 6.3.

REFERENCES

[1] I. Basawa and D. J. Scott, *Asymptotic Optimal Inference for Non-Ergodic Models*, Springer-Verlag, Berlin, Heidelberg, 1983.

[2] R. Bahadur and R. Ranga Rao, *On deviations of the sample mean*, Ann. Math. Statist., 31 (1960), pp. 1015–1027.

[3] G. Ben Arous, *Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus*, Ann. Sci. École Norm. Sup., 21 (1988), pp. 307–331.

[4] B. Bercu, F. Gamboa, and A. Rouault, *Large deviations for quadratic forms of stationary Gaussian processes*, Stochastic Process. Appl., 71 (1997), pp. 75–90.

[5] B. Bercu, F. Gamboa, and M. Lavielle, *Sharp large deviations for Gaussian quadratic forms with applications*, ESAIM Probab. Statist., 4 (2000), pp. 1–24.

[6] E. Bolthausen, *Laplace approximations for sums of independent random vectors*, Probab. Theory Related Fields, 76 (1987), pp. 167–206.

[7] H. Cramer, *Random Variables and Probability Distributions*, Cambridge University Press, Cambridge, UK, 1970.

[8] W. Bryc and A. Dembo, *Large deviations for quadratic functionals of Gaussian processes*, J. Theoret. Probab., 10 (1997), pp. 307–332.

[9] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Jones and Bartlett, Boston, MA, 1993.

[10] A. Dembo and O. Zeitouni, *Large deviations via parameter dependent change of measure and an application to the lower tail of Gaussian processes*, Progr. Probab., 36 (1995), pp. 111–121.

[11] A. Dembo, E. Mayer-Wolf, and O. Zeitouni, *Exact behavior of Gaussian seminorms*, Statist. Probab. Lett., 23 (1995), pp. 275–280.

[12] C. Esseen, *Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law*, Acta Math., 77 (1945), pp. 1–125.
[13] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, Wiley, New York, 1968.
[14] D. Florens-Landais and H. Pham, *Large deviations in estimation of an Ornstein–Uhlenbeck model*, J. Appl. Probab., 36 (1999), pp. 60–77.
[15] V. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*, University of California Press, Berkeley, Los Angeles, CA, 1958.
[16] I. A. Ibragimov, *Hitting probability of a Gaussian vector with values in a Hilbert space in a sphere of small radius*, J. Soviet Math., 20 (1982), pp. 2164–2174.
[17] S. Janson, *Gaussian Hilbert Spaces*, Cambridge University Press, Cambridge, UK, 1997.
[18] H.-H. Kuo, *Gaussian Measures in Banach Spaces*, Springer-Verlag, Berlin, Heidelberg, 1975.
[19] W. V. Li, *Comparison results for the lower tail of Gaussian seminorms*, J. Theoret. Probab., 5 (1992), pp. 1–31.
[20] R. Liptser and A. Shiryayev, *Statistics of Random Processes*, Vol. 2, Springer-Verlag, Berlin, New York, 1978.
[21] H. Luschgy, *Asymptotic behavior of Neyman–Pearson tests for autoregressive processes*, Scand. J. Statist., 21 (1994), pp. 461–473.
[22] J. Neveu, *Processus aléatoires gaussiens*, Presses de l'Université de Montréal, Montréal, Quebec, Canada, 1968.
[23] M. Zani, *Grandes déviations pour des fonctionnelles issues de la statistique des processus*, Thése, Université Paris-Sud, Paris, France, 2000.
[24] V. M. Zolotarev, *Asymptotic behavior of the Gaussian measure in l_2*, J. Soviet Math., 35 (1986), pp. 2330–2334.