Bernoulli **5**(6), 1999, 1059–1098

Asymptotic inference for a linear stochastic differential equation with time delay

ALEXANDER A. GUSHCHIN[1] and UWE KÜCHLER[2]

[1]*Steklov Mathematical Institute, Gubkina 8, 117966 Moscow GSP-1, Russia. E-mail: gushchin@mi.ras.ru*
[2]*Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany. E-mail: kuechler@mathematik.hu-berlin.de*

For the stochastic differential equation

$$\mathrm{d}X(t) = \{aX(t) + bX(t-1)\}\,\mathrm{d}t + \mathrm{d}W(t), \qquad t \geqslant 0,$$

the local asymptotic properties of the likelihood function are studied. They depend strongly on the true value of the parameter $\vartheta = (a, b)^*$. Eleven different cases are possible if ϑ runs through $\mathbb{R}^2$. Let $\hat{\vartheta}_T$ be the maximum likelihood estimator of ϑ based on $(X(t),\ t \leqslant T)$. Applications to the asymptotic behaviour of $\hat{\vartheta}_T$ as $T \to \infty$ are given.

Keywords: likelihood function; limit theorems for martingales; local asymptotic mixed normality; local asymptotic normality; local asymptotic properties; local asymptotic quadraticity; maximum likelihood estimator; stochastic differential equations; time delay

1. Introduction

Assume $(W(t),\ t \geqslant 0)$ is a real-valued standard Wiener process, a and b are real numbers and $(X(t),\ t \geqslant -1)$ is a solution of

$$\mathrm{d}X(t) = aX(t)\,\mathrm{d}t + bX(t-1)\,\mathrm{d}t + \mathrm{d}W(t), \qquad t \geqslant 0, \tag{1.1}$$

with some fixed initial condition $X(t) = X_0(t),\ t \in [-1, 0]$, where $X_0(\cdot)$ is a continuous stochastic process independent of $W(\cdot)$. The solution $(X(t),\ t \geqslant -1)$ of (1.1) exists, is pathwise uniquely determined and can be represented as

$$X(t) = x_0(t)X_0(0) + b\int_{-1}^{0} x_0(t-s-1)X_0(s)\,\mathrm{d}s + \int_0^t x_0(t-s)\,\mathrm{d}W(s), \qquad t \geqslant 0. \tag{1.2}$$

Obviously, it has continuous paths for $t \geqslant 0$ with probability one and, conditionally on X_0, X is a Gaussian process. Here $(x_0(t),\ t \geqslant -1)$ denotes the so-called fundamental solution of the deterministic equation

1350–7265

$$\begin{aligned} \dot{x}(t) &= ax(t) + bx(t-1), && t > 0, \\ x(t) &= 1, && t = 0, \\ x(t) &= 0, && t \in [-1, 0). \end{aligned} \tag{1.3}$$

Equation (1.1) is a very special case of linear stochastic differential equations of the type

$$\mathrm{d}X(t) = \int_{-1}^{0} X(t+s)a(\mathrm{d}s)\,\mathrm{d}t + \mathrm{d}M(t), \qquad t \geqslant 0, \tag{1.4}$$

where $a(\cdot)$ is an arbitrary function of finite variation on $[-1, 0]$ and $(M(t),\ t \geqslant 0)$ is, for example, a semimartingale; see Mohammed and Scheutzow (1990).

Assume that the solution $(X(t),\ t \in [-1, T])$ of (1.1), for some finite $T > 0$, has been observed, and that the parameters (a, b) are unknown and have to be estimated. Then we have a parametric problem, which generalizes the statistical problem of estimating the parameter in Langevin's equation

$$\mathrm{d}X(t) = aX(t)\,\mathrm{d}t + \mathrm{d}W(t), \qquad t \geqslant 0 \tag{1.5}$$

(see, for example, Basawa and Prakasa Rao 1980). Estimation problems for stochastic differential equations with time delay have been considered in few papers up to now; see Dietz (1992), Küchler and Kutoyants (1996) and the references therein. The model we consider seems to be of interest for the following reasons. First, it is a relatively simple example exhibiting a variety of qualitatively different local asymptotic properties for different values of the parameter. Second, the model already shows some typical effects appearing in estimation problems for equations with time-delayed terms, for example of type (1.4): a wide range of local asymptotic structures for the likelihood; a close connection between asymptotic properties of the likelihoods and the set of solutions to the corresponding characteristic equation (see (1.9) below); and periodic behaviour of the likelihoods and estimators for some values of the parameter. Third, in contrast to more general delay models, we are able to compute explicitly the rates of convergence and the limit distributions of estimators for every value of the parameter.

The solutions of (1.1) form an exponential family of continuous stochastic processes in the sense of Küchler and Sørensen (1989). Thus the maximum likelihood estimator $\hat{\vartheta}_T$ of $\vartheta = (a, b)^*$ (where * denotes matrix or vector transposition) can be expressed explicitly by

$$\hat{\vartheta}_T = (I_T^0)^{-1} V_T^0,$$

where V_T^0 denotes the vector

$$V_T^0 = \left(\int_0^T X(t)\,\mathrm{d}X(t),\ \int_0^T X(t-1)\,\mathrm{d}X(t) \right)^*$$

and I_T^0 is the observed Fisher information matrix given by

$$I_T^0 = \begin{pmatrix} \int_0^T X^2(t)\,\mathrm{d}t & \int_0^T X(t)X(t-1)\,\mathrm{d}t \\ \int_0^T X(t)X(t-1)\,\mathrm{d}t & \int_0^T X^2(t-1)\,\mathrm{d}t \end{pmatrix}.$$

The estimator $\hat{\vartheta}_T$ is calculated from the log-likelihood function

$$\log \frac{\mathrm{d}P_T^{\vartheta}}{\mathrm{d}P_T^{(0,0)}}(X) = \vartheta^* V_T^0 - \frac{1}{2}\vartheta^* I_T^0 \vartheta, \qquad \vartheta \in \mathbb{R}^2$$

(see, for example, Liptser and Shiryaev 1977). Here $P_T^{(a,b)}$ is the measure on $\mathbb{C}([-1, T])$ generated by the solution $(X(t), t \in [-1, T])$ of (1.1).

The main purpose of this paper is to study local asymptotic properties of the family $(P_T^{\vartheta}, \vartheta \in \mathbb{R}^2)$ and then to draw conclusions for properties of the estimator $\hat{\vartheta}_T$ when $T \to \infty$.

Since the log-likelihoods are quadratic in ϑ for each $T > 0$, it is not surprising that the family (P_T^{ϑ}), $T > 0$, is locally asymptotically quadratic (LAQ) at every $\vartheta_0 \in \mathbb{R}^2$; see Section 2 (for the notion of local asymptotic quadraticity, see Le Cam and Yang 1990; or Jeganathan 1995). Namely, choose $\vartheta_0 = (a, b)^* \in \mathbb{R}^2$ arbitrary but fixed and introduce $\vartheta = \vartheta_0 + \varphi_T \mu$, where $\mu = (\alpha, \beta)^* \in \mathbb{R}^2$ and $\varphi_T = \varphi_T(\vartheta_0)$ is a normalizing regular 2×2 matrix with $\varphi_T \to 0$ as $T \to \infty$. Then we obtain

$$\log \frac{\mathrm{d}P_T^{\vartheta}}{\mathrm{d}P_T^{\vartheta_0}}(X) = \mu^* V_T - \frac{1}{2}\mu^* I_T \mu, \tag{1.6}$$

where

$$V_T^* = \left(\int_0^T X(t)\,\mathrm{d}W(t), \int_0^T X(t-1)\,\mathrm{d}W(t)\right)\varphi_T \tag{1.7}$$

and

$$I_T = \varphi_T^* I_T^0 \varphi_T. \tag{1.8}$$

In view of (1.6), to prove local asymptotic quadraticity at ϑ_0 one has to choose the matrices $\varphi_T(\vartheta_0)$ in such a way that (a) the vectors (V_T, I_T) are bounded in probability as $T \to \infty$; (b) if (V_{T_n}, I_{T_n}) converges in distribution to a limit (V_∞, I_∞) for a subsequence $\{T_n\} \to \infty$, then

$$\mathrm{E}\exp(\mu^* V_\infty - \tfrac{1}{2}\mu^* I_\infty \mu) = 1$$

for every $\mu \in \mathbb{R}^2$; (c) if I_{T_n} converges in distribution to a limit I_∞ for a subsequence $\{T_n\} \to \infty$, then I_∞ is almost surely positive definite. Recall also that the important special cases of local asymptotic quadraticity are local asymptotic mixed normality and local asymptotic normality. Local asymptotic mixed normality at ϑ_0 means that (V_T, I_T) converges in distribution to $(I_\infty^{1/2} Z, I_\infty)$ as $T \to \infty$, where the matrix I_∞ is almost surely positive definite and Z is a standard Gaussian vector independent of I_∞. If, moreover, I_∞ is non-random, then we have local asymptotic normality at ϑ_0.

Note that condition (c) is important since otherwise we are not in a position even to establish asymptotic properties of $\hat{\vartheta}_T$ (cf. Dietz 1992). Actually, this condition plays a decisive role in determining φ_T. In general, (c) cannot be reached with matrices φ_T being diagonal. We construct φ_T as the product of two quadratic matrices $\varphi_T^{(1)}$ and $\varphi_T^{(2)}$, $\varphi_T = \varphi_T^{(1)}\varphi_T^{(2)}$, where $\varphi_T^{(1)}$ converges to a non-singular limit as $T \to \infty$ (the dependence on T cannot be avoided in general) and $\varphi_T^{(2)}$ is diagonal with elements tending to zero, in most cases at different rates.

It is obvious from (1.7), (1.8) and (1.2) that the properties of the fundamental solution $x_0(t)$ for $t \to \infty$ very much influence the limit properties of (V_T, I_T). Recall that for Langevin's equation ($b = 0$) we have $x_0(t) = \mathrm{e}^{at}$, the solution $(X(t),\ t \geqslant 0)$ is the Ornstein–Uhlenbeck process and there are exactly three relevant cases in considering local asymptotic properties ($a < 0$, $a = 0$, $a > 0$). In our case the picture turns out to be much richer. To specify φ_T and to study the limit behaviour of (V_T, I_T) we have to distinguish 11 different cases for ϑ_0. These cases will be introduced as follows.

The behaviour of $x_0(\cdot)$ is connected with the set Λ of (complex) solutions of the so-called characteristic equation for (1.3),

$$\lambda - a - b\,\mathrm{e}^{-\lambda} = 0. \tag{1.9}$$

Note that a complex number λ solves (1.9) if and only if $(\mathrm{e}^{\lambda t},\ t \geqslant -1)$ solves $\dot{x}(t) = ax(t) + bx(t-1),\ t \geqslant 0$.

It is easy to see that the set Λ of solutions of (1.9) is countably infinite (if $b \neq 0$) and that for every $c \in \mathbb{R}$ the set $\Lambda_c := \{\lambda \in \Lambda | \operatorname{Re}\lambda \geqslant c\}$ is finite. In particular, $v_0 := \max\{\operatorname{Re}\lambda | \lambda \in \Lambda\} < \infty$. Define $v_1 := \max\{\operatorname{Re}\lambda | \lambda \in \Lambda,\ \operatorname{Re}\lambda < v_0\}$ $(\max\varnothing = -\infty)$. One verifies easily that if $\lambda \in \Lambda$ then $\overline{\lambda} \in \Lambda$ and no other $\mu \in \Lambda$ with $\operatorname{Re}\mu = \operatorname{Re}\lambda$ exists. The equation (1.9) has at most two real solutions. If there exists a real solution v then the real part of every non-real solution is strictly less than v. Consequently, the only possible real solutions are v_0 (if there is exactly one) or v_0 and v_1 (if there are two).

We have $v_0 \in \Lambda$ if and only if

$$b \geqslant v(a) := -\mathrm{e}^{a-1}, \tag{1.10}$$

otherwise there exists a unique λ_0 in Λ with $\operatorname{Re}\lambda_0 = v_0$ and $\xi_0 := \operatorname{Im}\lambda_0 > 0$. Furthermore, in this case we have $\xi_0 < \pi$. Moreover, a second real solution exists, $v_1 \in \Lambda$, if and only if $v(a) < b < 0$.

For every λ in Λ, denote by $m(\lambda)$ the multiplicity of λ as a solution of (1.9). We have $m(\lambda) = 1$ for all $\lambda \in \Lambda$ with the only exception that if $b = v(a)$, then $\lambda = a - 1$ is the unique real solution of (1.9) and $m(\lambda) = 2$.

Additional information about the solutions of the equation (1.9) can be found in Hayes (1950).

The following lemma is crucial for this paper. It is based on the inverse Laplace transform and Cauchy's residue theorem and it can be found in a slightly different form in Myschkis (1972); see also Hale and Verduyn Lunel (1993). The proof will be sketched in Section 5.

Lemma 1.1. *For all $c < v_0$ the fundamental solution $x_0(\cdot)$ of* (1.3) *can be represented in the form*

$$x_0(t) = \psi_0(t)\mathrm{e}^{v_0 t} + \sum_{\substack{\lambda_k \in \Lambda_c \\ \operatorname{Re}\lambda_k < v_0}} c_k\,\mathrm{e}^{\lambda_k t} + o(\mathrm{e}^{\gamma t}), \qquad \textit{for } t \to \infty, \tag{1.11}$$

where $\gamma < c$ *and* c_k *are some constants. Here* $\psi_0(t)$ *equals*

$$\psi_0(t) = \begin{cases} \dfrac{1}{v_0 - a + 1} & \textit{if } v_0 \in \Lambda,\ m(v_0) = 1, \\ 2t + \dfrac{2}{3} & \textit{if } v_0 \in \Lambda,\ m(v_0) = 2, \\ A_0 \cos(\xi_0 t) + B_0 \sin(\xi_0 t) & \textit{if } v_0 \notin \Lambda, \end{cases}$$

with

$$A_0 = \frac{2(v_0 - a + 1)}{(v_0 - a + 1)^2 + \xi_0^2}, \qquad B_0 = \frac{2\xi_0}{(v_0 - a + 1)^2 + \xi_0^2}.$$

Remarks.

(1) Note that the three cases for ψ_0 correspond to $b > v(a)$, $b = v(a)$ and $b < v(a)$, respectively.

(2) Recall that for Langevin's equation ($b = 0$) we have $b > v(a)$ for every $a \in \mathbb{R}$. In this case $\Lambda = \{a\}$ and therefore $v_0 = a$ and $x_0(t) = \mathrm{e}^{at}$.

(3) If $v_0 \in \Lambda$, $m(v_0) = 1$ (and $b \neq 0$ to avoid the case from the previous remark), then for our purposes it is necessary to separate a further term from the sum in (1.11). We obtain

$$x_0(t) = \frac{1}{v_0 - a + 1}\mathrm{e}^{v_0 t} + \psi_1(t)\mathrm{e}^{v_1 t} + o(\mathrm{e}^{\gamma t}), \qquad \text{for } t \to \infty, \tag{1.12}$$

where $\gamma < v_1$,

$$\psi_1(t) = \begin{cases} \dfrac{1}{v_1 - a + 1} & \text{if } v_1 \in \Lambda, \\ A_1 \cos(\xi_1 t) + B_1 \sin(\xi_1 t) & \text{if } v_1 \notin \Lambda, \end{cases}$$

and

$$A_1 = \frac{2(v_1 - a + 1)}{(v_1 - a + 1)^2 + \xi_1^2}, \qquad B_1 = \frac{2\xi_1}{(v_1 - a + 1)^2 + \xi_1^2}.$$

Here ξ_1 denotes the uniquely determined positive number such that $\lambda_1 = v_1 + \mathrm{i}\xi_1 \in \Lambda$. (We note that $\xi_1 \in (\pi, 2\pi)$ in this case.)

The proof follows the line of the proof of Lemma 1.1 (see Section 5) in an obvious way.

As was mentioned above, 11 cases can be distinguished for the limit properties of (V_T, I_T); Table 1 represents these cases. The first column describes these cases in terms of v_0 and v_1, and the relations (1.11) and (1.12) make clear a connection between our classification and asymptotic properties of $x_0(\cdot)$. The second column characterizes the cases

Table 1. The 11 cases in terms of $(v_0,\ v_1)$ and $(a,\ b)$

<table>
<tr><td colspan="5">$v_0 < 0$</td><td>$a < 1,\ u(a) < b < -a$</td><td>N</td></tr>
<tr><td rowspan="3">$v_0 = 0$</td><td rowspan="2">$v_0 \in \Lambda$</td><td colspan="3">$m(v_0) = 1$</td><td>$a < 1,\ b = -a$</td><td>Q1</td></tr>
<tr><td colspan="3">$m(v_0) = 2$</td><td>$a = 1,\ b = -a$</td><td>Q2</td></tr>
<tr><td colspan="4">$v_0 \notin \Lambda$</td><td>$a < 1,\ b = u(a)$</td><td>Q3</td></tr>
<tr><td rowspan="7">$v_0 > 0$</td><td rowspan="6">$v_0 \in \Lambda$</td><td rowspan="5">$m(v_0) = 1$</td><td colspan="2">$v_1 < 0$</td><td>$-a < b < w(a)$</td><td>M1</td></tr>
<tr><td rowspan="2">$v_1 = 0$</td><td>$v_1 \in \Lambda$</td><td>$a > 1,\ b = -a$</td><td>Q4</td></tr>
<tr><td>$v_1 \notin \Lambda$</td><td>$b = w(a)$</td><td>Q5</td></tr>
<tr><td rowspan="2">$v_1 > 0$</td><td>$v_1 \in \Lambda$</td><td>$a > 1,\ v(a) < b < -a$</td><td>M2</td></tr>
<tr><td>$v_1 \notin \Lambda$</td><td>$b > w(a)$</td><td>P1</td></tr>
<tr><td>$m(v_0) = 2$</td><td colspan="2"></td><td>$a > 1,\ b = v(a)$</td><td>M3</td></tr>
<tr><td colspan="4">$v_0 \notin \Lambda$</td><td>$a < 1,\ b < u(a)$ or $a \geqslant 1,\ b < v(a)$</td><td>P2</td></tr>
</table>

in terms of a and b. The last column gives a designation for each case which will be used in the rest of this paper. The functions $u(a)$, $a < 1$, and $w(a)$, $a \in \mathbb{R}$, are defined as follows: introduce a parametric curve $(a(\xi), b(\xi))$, $\xi > 0$, $\xi \neq \pi, 2\pi, \ldots,$ in $\mathbb{R}^2$ by

$$a(\xi) = \xi \cot \xi, \qquad b(\xi) = -\xi / \sin \xi;$$

then $b = u(a)$ and $b = w(a)$ are the branches of this curve corresponding to $\xi \in (0, \pi)$ and $\xi \in (\pi, 2\pi)$ respectively; see also Figure 1.

In the following we wish to give a first impression of what happens in the 11 cases. The first subdivision on the left-hand side of Table 1 follows the Ornstein–Uhlenbeck case (where $v_0 = a$): $v_0 < 0$, $v_0 = 0$ and $v_0 > 0$.

The first case, $v_0 < 0$ holds if and only if there exists a stationary solution of (1.1). This solution is Gaussian and uniquely determined (see Küchler and Mensch 1992). In this case the statistical properties of our model are classical: the local asymptotic normality property holds. The form of $\psi_0(\cdot)$ does not influence the asymptotic properties of V_T and I_T if $v_0 < 0$. But it does if $v_0 = 0$ or $v_0 > 0$.

If $v_0 = 0$, the underlying experiment is only LAQ (as if $a = 0$ in the Ornstein–Uhlenbeck case) in all three cases Q1–Q3. But the normalizing matrix φ_T and the corresponding limit experiment now principally depend on the form of $\psi_0(\cdot)$, represented by three different expressions in Lemma 1.1.

Now let us consider the case $v_0 > 0$. The form of $\psi_0(\cdot)$ is essential again for our purposes. If $v_0 \notin \Lambda$, then we obtain a periodic behaviour of (V_T, I_T) in a certain sense. We call this the periodically locally asymptotically mixed normal (PLAMN) property, to emphasize the fact that the cluster points of (V_T, I_T) have the same structure as in the locally asymptotically mixed normal (LAMN) case but (V_T, I_T) converges in distribution only if T runs to infinity through a sequence in such a way that, for a certain $\Delta > 0$, the fractional part of T/Δ tends to a limit. If $v_0 \in \Lambda$ and $m(v_0) = 2$, then the model is LAMN. This is the only case where the matrix $\varphi_T^{(1)}$ has to be chosen dependent on T.

If $v_0 \in \Lambda$ and $m(v_0) = 1$, we also have to take into consideration the second term on the

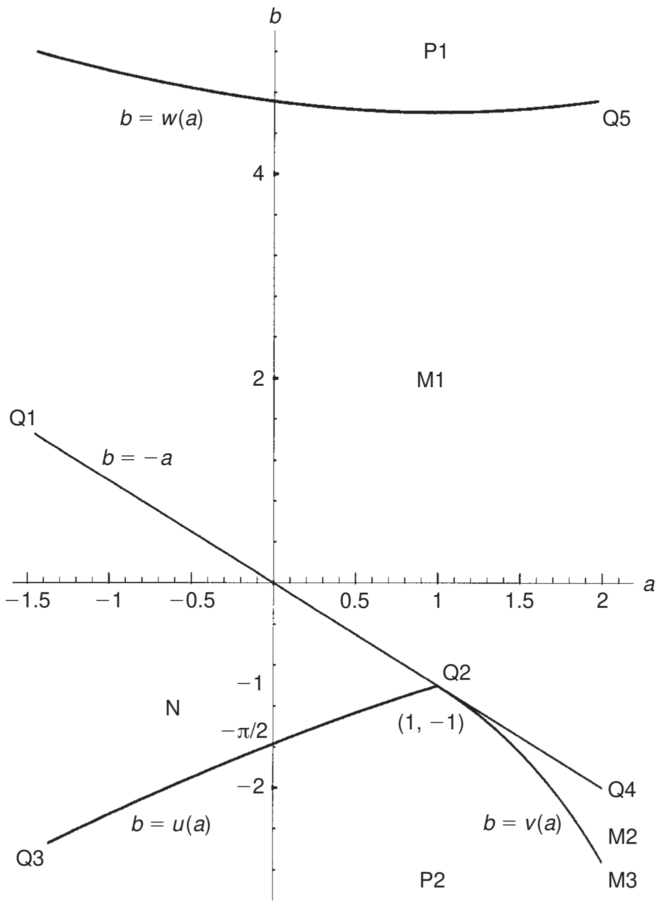

Figure 1. The different cases for $\vartheta_0 = (a, b)^*$

right-hand side of (1.12) and obtain five difference cases. Indeed, the limit behaviour depends on the sign of v_1 and, if $v_1 \geqslant 0$, on whether $v_1 \in \Lambda$ or not.

If $v_1 < 0$, or $v_1 > 0$ and $v_1 \in \Lambda$, the underlying experiment is LAMN. But if $v_1 > 0$ and $v_1 \notin \Lambda$, a periodic behaviour of (V_T, I_T) occurs again and the experiment is PLAMN.

If $v_1 = 0$, the model is LAQ, whether $v_1 \in \Lambda$ or not. Both cases have similar limit experiments.

Finally, we note that the LAMN (or PLAMN) property fails only if ϑ_0 belongs to the lines $b = -a$, $b = u(a)$ or $b = w(a)$.

2. Local asymptotic properties

In the preceding section we have introduced a series of cases for which the fundamental solution $x_0(\cdot)$ and (V_T, I_T) have different asymptotic properties. Here we shall study the asymptotic properties of (V_T, I_T) as $T \to \infty$ in more detail. The proofs are given in Section 4.

The symbols $\xrightarrow{P}$ and $\xrightarrow{d}$ henceforth denote convergence in probability and in distribution, respectively. We shall use the symbol $\xrightarrow{\mathcal{L}}$ to denote the convergence in distribution in the space $\mathbb{C}^d([0, 1])$ of continuous functions on [0, 1] with values in $\mathbb{R}^d$. Sometimes we shall use the abbreviated notation $\int_0^1 \tilde{W}_1 \,\mathrm{d}\tilde{W}_1$ instead of $\int_0^1 \tilde{W}_1(t)\,\mathrm{d}\tilde{W}_1(t)$ or $\int_0^1 XY\,\mathrm{d}t$ instead of $\int_0^1 X(t)Y(t)\,\mathrm{d}t$ etc. The concrete meaning will be clear from the context.

Some details are presented in Table 2. In the first column all 11 cases are listed in the order in which they will be considered. The next three columns describe the choice of ψ_T. Recall that

Table 2. The choice of φ_T and a description of convergence to I_∞

Case	$\varphi_T^{(1)}$	$\varphi_{11}(T)$	$\varphi_{22}(T)$	I_∞	Prop.
N	$\mathcal{I}_2$	$T^{-1/2}$	$T^{-1/2}$	$\begin{pmatrix} c & c \\ c & c \end{pmatrix}$	2.1
M1	$\mathcal{J}_2$	e^{-v_0T}	$T^{-1/2}$	$\begin{pmatrix} p & 0 \\ 0 & c \end{pmatrix}$	2.2
M2	$\mathcal{J}_2$	e^{-v_0T}	e^{-v_1T}	$\begin{pmatrix} p & p \\ p & p \end{pmatrix}$	2.3
M3	$\begin{pmatrix} 1 & 1 \\ 0 & -(1+T^{-1})\mathrm{e}^{v_0} \end{pmatrix}$	$T^{-1}\,\mathrm{e}^{-v_0T}$	$T\,\mathrm{e}^{-_0T}$	$\begin{pmatrix} p & p \\ p & p \end{pmatrix}$	2.4
P1	$\mathcal{J}_2$	e^{-v_0T}	e^{-v_1T}	$\begin{pmatrix} p & p^* \\ p^* & p^* \end{pmatrix}$	2.5
P2	$\mathcal{I}_2$	e^{-v_0T}	e^{-v_0T}	$\begin{pmatrix} p^* & p^* \\ p^* & p^* \end{pmatrix}$	2.6
Q1	$\mathcal{J}_2$	T^{-1}	$T^{-1/2}$	$\begin{pmatrix} d & 0 \\ 0 & c \end{pmatrix}$	2.7
Q2	$\mathcal{J}_2$	T^{-2}	T^{-1}	$\begin{pmatrix} d & d \\ d & d \end{pmatrix}$	2.8
Q3	$\mathcal{I}_2$	T^{-1}	T^{-1}	$\begin{pmatrix} d & d \\ d & d \end{pmatrix}$	2.9
Q4	$\mathcal{J}_2$	e^{-v_0T}	T^{-1}	$\begin{pmatrix} p & 0 \\ 0 & d \end{pmatrix}$	2.10
Q5	$\mathcal{J}_2$	e^{-v_0T}	T^{-1}	$\begin{pmatrix} p & 0 \\ 0 & d \end{pmatrix}$	2.11

$$\varphi_T = \varphi_T^{(1)}\varphi_T^{(2)} \qquad \text{and} \qquad \varphi_T^{(2)} = \begin{pmatrix} \varphi_{11}(T) & 0 \\ 0 & \varphi_{22}(T) \end{pmatrix}. \tag{2.1}$$

It turns out that we can choose $\varphi_T^{(1)}$ so that it does not depend on T in every case except M3; moreover, $\varphi_T^{(1)}$ equals

$$\mathcal{I}_2 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \mathcal{J}_2 := \begin{pmatrix} 1 & 1 \\ 0 & -\mathrm{e}^{v_0} \end{pmatrix}$$

(of course, the value of v_0 depends on ϑ_0).

It will be proved in Propositions 2.1, 2.2–2.4, 2.7–2.11 that in cases N, M1–M3 and Q1–Q5, under this choice of φ_T, we have $(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty)$. It is implicitly assumed that

$$\mathrm{E}\exp(\mu^* V_\infty - \tfrac{1}{2}\mu^* I_\infty \mu) = 1$$

for every $\mu \in \mathbb{R}^2$ and I_∞ is a non-singular matrix; the proof of this is either trivial or routine. In the fifth column of Table 2 we describe the matrix I_∞ and the type of convergence of I_T to I_∞ in a symbolic manner. The elements of the matrix in this column have the following meaning. The symbol '0' means that the corresponding element of I_∞ is 0. The symbol 'c' means that the corresponding element of I_∞ is a (non-zero) constant. In both cases, the corresponding element of I_T converges to this constant in probability. The symbol 'p' means that the corresponding element of I_∞ is random but there is still the convergence in probability of the corresponding element of I_T to that of I_∞. Finally, the symbol 'd' means that we have only the convergence in distribution of the corresponding element of I_T to that of I_∞ but not the convergence in probability. In cases P1 and P2 studied in Propositions 2.5 and 2.6, we have a periodic behaviour of (V_T, I_T) in a certain sense. There we use the symbol 'p^*' to indicate that we have the convergence in probability of the corresponding elements of I_T to a random limit but only when T runs to infinity through certain grids.

The last column of the table indicates the number of the proposition in which the corresponding case is considered.

In the following we shall treat every case mentioned above in a separate proposition. Recall that V_T and I_T are given by (1.7) and (1.8). The process $X(\cdot)$ is defined by (1.2) for some fixed a and b and the matrices φ_T are constructed in (2.1). For every proposition below, the parameter $\vartheta_0 = (a, b)^*$ is assumed to belong to the set described by Table 1 in accordance with the case under consideration. The definitions of $\varphi_T^{(1)}$, $\varphi_{11}(T)$ and $\varphi_{22}(T)$ are taken from Table 2. Unless otherwise specified, all limits are taken as $T \to \infty$.

Let us start with the simplest case, $v_0 < 0$. This is the only case where $\int_0^\infty x_0^2(t)\,\mathrm{d}t < \infty$ and a stationary solution of (1.1) exists.

Proposition 2.1. *In case* N *the family* $(P^\vartheta, \vartheta \in \mathbb{R}^2)$ *is locally asymptotically normal at every* ϑ_0:

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty),$$

where

$$I_\infty = \begin{pmatrix} \int_0^\infty x_0^2(t)\,\mathrm{d}t & \int_0^\infty x_0(t)x_0(t+1)\,\mathrm{d}t \\ \int_0^\infty x_0(t)x_0(t+1)\,\mathrm{d}t & \int_0^\infty x_0^2(t)\,\mathrm{d}t \end{pmatrix}$$

and $V_\infty \sim \mathcal{N}(0, I_\infty)$.

Now let us treat cases M1–M3.

Proposition 2.2. *In case* M1 *the family* $(P^\vartheta, \vartheta \in \mathbb{R}^2)$ *is LAMN at every* ϑ_0:

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty),$$

where $(V_\infty, I_\infty) \stackrel{d}{=} (I_\infty^{1/2} Z, I_\infty)$ *and the vector* Z *is independent of* I_∞ *and distributed as* $\mathcal{N}(0, \mathcal{I}_2)$. *The matrix* I_∞ *is given by*

$$I_\infty = \begin{pmatrix} \dfrac{U_0^2}{2v_0(v_0 - a + 1)^2} & 0 \\ 0 & \int_0^\infty (x_0(t) - \mathrm{e}^{v_0} x_0(t-1))^2\,\mathrm{d}t \end{pmatrix},$$

where

$$U_0 = X_0(0) + b\int_{-1}^0 \mathrm{e}^{-v_0(s+1)} X_0(s)\,\mathrm{d}s + \int_0^\infty \mathrm{e}^{-v_0 s}\,\mathrm{d}W(s).$$

Proposition 2.3. *In case* M2 *the family* $(P^\vartheta, \vartheta \in \mathbb{R}^2)$ *is LAMN at every* ϑ_0:

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty),$$

where $(V_\infty, I_\infty) \stackrel{d}{=} (I_\infty^{1/2} Z, I_\infty)$ *and the vector* Z *is independent of* I_∞ *and distributed as* $\mathcal{N}(0, \mathcal{I}_2)$. *The matrix* I_∞ *is given by*

$$I_\infty = \begin{pmatrix} \dfrac{U_0^2}{2v_0(v_0 - a + 1)^2} & \dfrac{U_0 U_1(\mathrm{e}^{v_0 - v_1} - 1)}{(v_0 + v_1)(v_0 - a + 1)(a - v_1 - 1)} \\ \dfrac{U_0 U_1(\mathrm{e}^{v_0 - v_1} - 1)}{(v_0 + v_1)(v_0 - a + 1)(a - v_1 - 1)} & \dfrac{U_1^2(\mathrm{e}^{v_0 - v_1} - 1)^2}{2v_1(a - v_1 - 1)^2} \end{pmatrix},$$

where U_0 *is defined in Proposition* 2.2 *and*

$$U_1 = X_0(0) + b\int_{-1}^0 \mathrm{e}^{-v_1(s+1)} X_0(s)\,\mathrm{d}s + \int_0^\infty \mathrm{e}^{-v_1 s}\,\mathrm{d}W(s).$$

Proposition 2.4. *In case* M3 *the family* $(P^\vartheta, \vartheta \in \mathbb{R}^2)$ *is LAMN at* ϑ_0:

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty),$$

where $(V_\infty, I_\infty) \stackrel{d}{=} (I_\infty^{1/2} Z, I_\infty)$ *and the vector* Z *is independent of* I_∞ *and distributed as* $\mathcal{N}(0, \mathcal{I}_2)$*. The matrix* I_∞ *is given by*

$$I_\infty = \begin{pmatrix} \dfrac{2U_0^2}{v_0} & \dfrac{U_0(\frac{4}{3}U_0 + 2U_2)}{v_0} + \dfrac{U_0^2}{v_0^2} \\ \dfrac{U_0(\frac{4}{3}U_0 + 2U_2)}{v_0} + \dfrac{U_0^2}{v_0^2} & \dfrac{(\frac{4}{3}U_0 + 2U_2)^2}{2v_0} + \dfrac{U_0(\frac{4}{3}U_0 + 2U_2)}{v_0^2} + \dfrac{U_0^2}{v_0^3} \end{pmatrix},$$

where U_0 *is defined in Proposition* 2.2 *and*

$$U_2 = b\int_{-1}^0 (s+1)\mathrm{e}^{-v_0(s+1)} X_0(s)\,\mathrm{d}s + \int_0^\infty s\,\mathrm{e}^{-v_0 s}\,\mathrm{d}W(s).$$

The next two propositions treat cases P1 and P2. Recall that if $v_i \notin \Lambda$, then ξ_i denotes the positive imaginary part of $\lambda \in \Lambda$ with $\operatorname{Re}\lambda = v_i$, $i = 0, 1$.

Proposition 2.5. *In case* P1 *the family* $(P^\vartheta, \vartheta \in \mathbb{R}^2)$ *is 'PLAMN' at every* ϑ_0 *in the following sense: for* $T_n = u + n\Delta$, *where* $u \in [0, \Delta)$ *is fixed,* $\Delta = 2\pi/\xi_1$, $n \geqslant 0$,

$$(V_{T_n}, I_{T_n}) \xrightarrow{d} (V_\infty(u), I_\infty(u)), \qquad n \to \infty,$$

where $(V_\infty(u), I_\infty(u)) \stackrel{d}{=} (I_\infty^{1/2}(u)Z, I_\infty(u))$ *and the random vector* Z *is independent of* $I_\infty(u)$ *and distributed as* $\mathcal{N}(0, \mathcal{I}_2)$*. The matrix* $I_\infty(u)$ *is given by*

$$I_\infty(u) = \begin{pmatrix} \dfrac{U_0^2}{2v_0(v_0 - a + 1)^2} & \dfrac{U_0}{v_0 - a + 1}\displaystyle\int_0^\infty \mathrm{e}^{-(v_0+v_1)t} U(u-t)\,\mathrm{d}t \\ \dfrac{U_0}{v_0 - a + 1}\displaystyle\int_0^\infty \mathrm{e}^{-(v_0+v_1)t} U(u-t)\,\mathrm{d}t & \displaystyle\int_0^\infty \mathrm{e}^{-2v_1 t} U^2(u-t)\,\mathrm{d}t \end{pmatrix}.$$

Here U_0 *is defined as in Proposition* 2.2,

$$U(t) = X_0(0)\phi(t) + b\int_{-1}^0 \phi(t-s-1)\mathrm{e}^{-v_1(s+1)} X_0(s)\,\mathrm{d}s + \int_0^\infty \phi(t-s)\mathrm{e}^{-v_1 s}\,\mathrm{d}W(s),$$
$$\phi(t) = A\cos(\xi_1 t) + B\sin(\xi_1 t)$$

and

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} - \mathrm{e}^{v_0 - v_1}\begin{pmatrix} \cos\xi_1 & -\sin\xi_1 \\ \sin\xi_1 & \cos\xi_1 \end{pmatrix}\begin{pmatrix} A_1 \\ B_1 \end{pmatrix},$$

where

$$A_1 = \frac{2(v_1 - a + 1)}{(v_1 - a + 1)^2 + \xi_1^2}, \qquad B_1 = \frac{2\xi_1}{(v_1 - a + 1)^2 + \xi_1^2}.$$

Proposition 2.6. *In case* P2 *the family* $(P^\vartheta, \vartheta \in \mathbb{R}^2)$ *is 'PLAMN' at every* ϑ_0 *in the following sense: for* $T_n = u + n\Delta$, *where* $u \in [0, \Delta)$ *is fixed,* $\Delta = \pi/\xi_0$, $n \geqslant 0$,

$$(V_{T_n}, I_{T_n}) \xrightarrow{d} (V_\infty(u), I_\infty(u)), \qquad n \to \infty,$$

where $(V_\infty(u), I_\infty(u)) \stackrel{d}{=} (I_\infty^{1/2}(u)Z, I_\infty(u))$ *and the random vector* Z *is independent of* $I_\infty(u)$ *and distributed as* $\mathcal{N}(0, \mathcal{I}_2)$. *The matrix* $I_\infty(u)$ *is given by*

$$I_\infty(u) = \begin{pmatrix} \int_0^\infty e^{-2v_0 t} U_0^2(u-t)\,dt & \int_0^\infty e^{-2v_0 t} U_0(u-t)U_2(u-t)\,dt \\ \int_0^\infty e^{-2v_0 t} U_0(u-t)U_2(u-t)\,dt & \int_0^\infty e^{-2v_0 t} U_2^2(u-t)\,dt \end{pmatrix},$$

where

$$U_i(t) = X_0(0)\phi_i(t) + b\int_{-1}^0 \phi_i(t-s-1)e^{-v_0(s+1)}X_0(s)\,ds + \int_0^\infty \phi_i(t-s)e^{-v_0 s}\,dW(s),$$

$$\phi_i(t) = A_i\cos(\xi_0 t) + B_i\sin(\xi_0 t), \qquad i = 0, 2,$$

$$A_0 = \frac{2(v_0 - a + 1)}{(v_0 - a + 1)^2 + \xi_0^2}, \qquad B_0 = \frac{2\xi_0}{(v_0 - a + 1)^2 + \xi_0^2},$$

$$\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} = e^{-v_0}\begin{pmatrix} \cos\xi_0 & -\sin\xi_0 \\ \sin\xi_0 & \cos\xi_0 \end{pmatrix}\begin{pmatrix} A_0 \\ B_0 \end{pmatrix}.$$

So far we have treated all the cases for which local asymptotic normality, local asymptotic mixed normality or periodic local asymptotic mixed normality holds. There remain five cases, where local asymptotic quadraticity is valid.

Proposition 2.7. *In case* Q1 *the family* $(P^\vartheta, \vartheta \in \mathbb{R}^2)$ *is LAQ at every* ϑ_0:

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty),$$

where (V_∞, I_∞) *is given by*

$$V_\infty = \left(\frac{1}{1-a}\int_0^1 \tilde{W}(t)\,d\tilde{W}(t),\ Z\right)^*$$

and

$$I_\infty = \begin{pmatrix} \frac{1}{(1-a)^2}\int_0^1 \tilde{W}^2(t)\,dt & 0 \\ 0 & \sigma^2 \end{pmatrix}.$$

Here $\sigma^2 = \int_0^\infty (x_0(t) - x_0(t-1))^2\,dt$, $(\tilde{W}(t),\ t \in [0, 1])$ *denotes a standard Wiener process and the random variable* Z *is independent of* $\tilde{W}(\cdot)$ *and distributed as* $\mathcal{N}(0, \sigma^2)$.

Proposition 2.8. *In case* Q2 *the family* $(P^\vartheta,\ \vartheta \in \mathbb{R}^2)$ *is LAQ at* $\vartheta_0 = (1,\ -1)^*$:

$$(V_T,\ I_T) \xrightarrow{d} (V_\infty,\ I_\infty),$$

where $(V_\infty,\ I_\infty)$ *is given by*

$$V_\infty = 2\left(\int_0^1 \tilde{X}(t)\,\mathrm{d}\tilde{W}(t),\ \int_0^1 \tilde{W}(t)\,\mathrm{d}\tilde{W}(t)\right)^*,$$

$$I_\infty = 4\begin{pmatrix} \int_0^1 \tilde{X}^2(t)\,\mathrm{d}t & \int_0^1 \tilde{X}(t)\tilde{W}(t)\,\mathrm{d}t \\ \int_0^1 \tilde{X}(t)\tilde{W}(t)\,\mathrm{d}t & \int_0^1 \tilde{W}^2(t)\,\mathrm{d}t \end{pmatrix}.$$

Here $(\tilde{W}(t),\ t \in [0,\ 1])$ *is a standard Wiener process and* $\tilde{X}(t) = \int_0^t \tilde{W}(s)\,\mathrm{d}s$.

Proposition 2.9. *In case* Q3 *the family* $(P^\vartheta,\ \vartheta \in \mathbb{R}^2)$ *is LAQ at every* ϑ_0:

$$(V_T,\ I_T) \xrightarrow{d} (V_\infty,\ I_\infty),$$

where

$$V_\infty = \frac{1}{2}\begin{pmatrix} A_0\int_0^1 \tilde{W}_1\,\mathrm{d}\tilde{W}_1 + A_0\int_0^1 \tilde{W}_2\,\mathrm{d}\tilde{W}_2 + B_0\int_0^1 \tilde{W}_1\,\mathrm{d}\tilde{W}_2 - B_0\int_0^1 \tilde{W}_2\,\mathrm{d}\tilde{W}_1 \\ A_2\int_0^1 \tilde{W}_1\,\mathrm{d}\tilde{W}_1 + A_2\int_0^1 \tilde{W}_2\,\mathrm{d}\tilde{W}_2 + B_2\int_0^1 \tilde{W}_1\,\mathrm{d}\tilde{W}_2 - B_2\int_0^1 \tilde{W}_2\,\mathrm{d}\tilde{W}_1. \end{pmatrix}$$

and

$$I_\infty = \frac{A_0^2 + B_0^2}{4}\begin{pmatrix} 1 & \cos\xi_0 \\ \cos\xi_0 & 1 \end{pmatrix}\int_0^1 (\tilde{W}_1^2 + \tilde{W}_2^2)\,\mathrm{d}t.$$

Here $(\tilde{W}_i(t),\ t \in [0,\ 1])$, $i = 1,\ 2$, *are two independent standard Wiener processes and*

$$A_0 = \frac{2(1-a)}{(1-a)^2 + \xi_0^2}, \qquad B_0 = \frac{2\xi_0}{(1-a)^2 + \xi_0^2},$$

$$\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} = \begin{pmatrix} \cos\xi_0 & -\sin\xi_0 \\ \sin\xi_0 & \cos\xi_0 \end{pmatrix}\begin{pmatrix} A_0 \\ B_0 \end{pmatrix}.$$

Proposition 2.10. *In case* Q4 *the family* $(P^\vartheta,\ \vartheta \in \mathbb{R}^2)$ *is LAQ at every* ϑ_0:

$$(V_T,\ I_T) \xrightarrow{d} (V_\infty,\ I_\infty),$$

where $(V_\infty,\ I_\infty)$ *is given by*

$$V_\infty = \left(\frac{U_0 Z}{\sqrt{2v_0}(v_0 - a + 1)}, \frac{e^{v_0} - 1}{a - 1} \int_0^1 \tilde{W}(t)\, d\tilde{W}(t) \right)^*$$

and

$$I_\infty = \begin{pmatrix} \dfrac{U_0^2}{2v_0(v_0 - a + 1)^2} & 0 \\ 0 & \dfrac{(e^{v_0} - 1)^2}{(a - 1)^2} \displaystyle\int_0^1 \tilde{W}^2(t)\, dt \end{pmatrix}.$$

Here U_0 is the same as in Proposition 2.2, Z and $(\tilde{W}(t),\ t \in [0, 1])$ are a standard normal distributed random variable and a standard Wiener process respectively, and U_0, Z and $\tilde{W}(\cdot)$ are independent.

Proposition 2.11. *In case Q5 the family $(P^\vartheta,\ \vartheta \in \mathbb{R}^2)$ is LAQ at ϑ_0:*

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty),$$

where (V_∞, I_∞) is given by

$$V_\infty = \begin{pmatrix} \dfrac{U_0 Z}{\sqrt{2v_0}(v_0 - a + 1)} \\ \dfrac{1}{2}\left(A\displaystyle\int_0^1 \tilde{W}_1\, d\tilde{W}_1 + A\int_0^1 \tilde{W}_2\, d\tilde{W}_2 + B\int_0^1 \tilde{W}_1\, d\tilde{W}_2 - B\int_0^1 \tilde{W}_2\, d\tilde{W}_1 \right) \end{pmatrix}$$

and

$$I_\infty = \begin{pmatrix} \dfrac{U_0^2}{2v_0(v_0 - a + 1)^2} & 0 \\ 0 & \dfrac{1}{4}(A^2 + B^2)\displaystyle\int_0^1 (\tilde{W}_1^2 + \tilde{W}_2^2)\, dt \end{pmatrix}.$$

Here U_0 is the same as in Proposition 2.2, Z and $(\tilde{W}_i(t),\ t \in [0, 1])$, $i = 1, 2$, are a standard normal distributed random variable and standard Wiener processes respectively, U_0, Z, $\tilde{W}_1(\cdot)$, $\tilde{W}_2(\cdot)$ are independent, and

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} - e^{v_0} \begin{pmatrix} \cos\xi_1 & -\sin\xi_1 \\ \sin\xi_1 & \cos\xi_1 \end{pmatrix} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix},$$

where

$$A_1 = \frac{2(1 - a)}{(1 - a)^2 + \xi_1^2}, \qquad B_1 = \frac{2\xi_1}{(1 - a)^2 + \xi_1^2}.$$

3. Asymptotic properties of the maximum likelihood estimator

Assume that we observe $(X(t),\ t \leqslant T)$ continuously, where $X(t)$ is a solution of (1.1) and the parameters a and b are unknown. The maximum likelihood estimator $\hat{\vartheta}_T$ of the true parameter $\vartheta_0 = (a_0, b_0)^*$ is given by

$$\hat{\vartheta}_T = \arg\max_{\vartheta \in \mathbb{R}^2} \ell^0_T(\vartheta) = (I^0_T)^{-1} V^0_T,$$

where

$$\ell^0_T(\vartheta) = \vartheta^* V^0_T - \tfrac{1}{2}\vartheta^* I^0_T \vartheta, \qquad \vartheta \in \mathbb{R}^2,$$

$$V^0_T = \left(\int_0^T X(t)\,\mathrm{d}X(t),\ \int_0^T X(t-1)\,\mathrm{d}X(T) \right)^*$$

and

$$I^0_T = \begin{pmatrix} \displaystyle\int_0^T X^2(t)\,\mathrm{d}t & \displaystyle\int_0^T X(t)X(t-1)\,\mathrm{d}t \\ \displaystyle\int_0^T X(t)X(t-1)\,\mathrm{d}t & \displaystyle\int_0^T X^2(t-1)\,\mathrm{d}t \end{pmatrix}.$$

Choose an arbitrary non-singular 2×2-matrix φ_T and introduce a new parameter $\mu = (\alpha, \beta)^* \in \mathbb{R}^2$ given by

$$\vartheta = \vartheta_0 + \varphi_T \mu.$$

Then

$$\hat{\vartheta}_T = \vartheta_0 + \varphi_T \hat{\mu}_T,$$

where $\hat{\mu}_T$ is defined by

$$\hat{\mu}_T = \arg\max_{\mu \in \mathbb{R}^2} \ell_T(\mu) = I_T^{-1} V_T$$

with

$$\ell_T(\mu) = \mu^* V_T - \frac{1}{2}\mu^* I_T \mu,$$

$$V_T = \varphi_T^* \left(\int_0^T X(t)\,\mathrm{d}W(t),\ \int_0^T X(t-1)\,\mathrm{d}W(t) \right)^*$$

and

$$I_T = \varphi_T^* I^0_T \varphi_T.$$

From Section 2 we know that under appropriate choice of φ_T we have (in the notation of Section 2)

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty) \tag{3.1}$$

or

$$(V_{u+n\Delta}, I_{u+n\Delta}) \xrightarrow{d} (V_\infty(u), I_\infty(u)) \tag{3.2}$$

with $\det I_\infty \neq 0$ and $\det I_\infty(u) \neq 0$ for every $u \in [0, \Delta)$, respectively. Consequently we obtain

$$\varphi_T^{-1}(\hat{\vartheta}_T - \vartheta_0) = \hat{\mu}_T \xrightarrow{d} I_\infty^{-1} V_\infty$$

or

$$\varphi_{u+n\Delta}^{-1}(\hat{\vartheta}_{u+n\Delta} - \vartheta_0) = \hat{\mu}_{u+n\Delta} \xrightarrow{d} I_\infty^{-1}(u) V_\infty(u)$$

for every $u \in [0, \Delta)$, respectively.

Thus we can draw conclusions concerning the asymptotic behaviour of $\hat{\vartheta}_T$ for $T \to \infty$. But some more properties follow from (3.1) and (3.2) by standard arguments. Indeed, if the LAMN property holds (for example, in cases N, M1–M3) then we have the local asymptotic minimax bound for an arbitrary estimator $\tilde{\vartheta}_T$,

$$\lim_{r\to\infty} \liminf_{T\to\infty} \sup_{\|\varphi_T^{-1}(\vartheta-\vartheta_0)\|\leqslant r} \mathrm{E}_\vartheta w\{\varphi_T^{-1}(\tilde{\vartheta}_T - \vartheta)\} \geqslant \mathrm{E}w\{(I_\infty)^{-1} V_\infty\}$$

$$= \mathrm{E}w\{(I_\infty)^{-1/2} Z\}, \tag{3.3}$$

where Z is an $\mathcal{N}(0, \mathcal{I}_2)$-vector independent of I_∞ and $w : \mathbb{R}^2 \to [0, \infty)$ is a bowl-shaped loss function. The maximum likelihood estimator $\hat{\vartheta}_T$ attains this bound, at least for bounded w. Moreover, the estimator $\hat{\vartheta}_T$ is asymptotically efficient in the convolution theorem sense (for example, see, Le Cam and Yang 1990; Jeganathan 1995).

In other cases, for example, if only the LAQ property holds, it follows that there exists a lower asymptotic minimax bound but possibly of different form (see Shiryaev and Spokoiny 1999; Greenwood and Wefelmeyer 1993). This bound need not be attainable. It is known that the maximum likelihood estimator is asymptotically generalized Bayesian with respect to the uniform distribution on $\mathbb{R}^2$ (Shiryaev and Spokoiny 1999).

For some class of estimators $\tilde{\vartheta}_T$ satisfying certain conditions of regularity, for example that the limit distribution of the randomly normed deviation

$$I_T \varphi_T^{-1}(\tilde{\vartheta}_T - \vartheta_0)$$

exists and is unbiased, the covariance matrix of this limit distribution is bounded from below by the corresponding covariance matrix for the maximum likelihood estimator $\hat{\vartheta}_T$ which is equal to $\mathrm{E}I_\infty$; see Gushchin (1995).

We have seen that the maximum likelihood estimator, after a certain matrix normalization, converges in distribution to some limit. In cases N, P2 and Q3 we have that φ_T is equal to $\varphi_{22}(T)\mathcal{I}_2$, and thus the normalization by the number $\varphi_{22}^{-1}(T)$ yields the same limit distribution.

In all other cases $\varphi_T^{(1)}$ is an upper triangular matrix and $\varphi_{11}(T) = o(\varphi_{22}(T))$. This reflects some singularities in the local structure of our model, which have not been mentioned so far. Before studying these singularities, we introduce the following notation. Let

$$\Theta := \{(a, b) \in \mathbb{R}^2 : a \leqslant 1,\ b \geqslant -a \text{ or } a \geqslant 1,\ b \geqslant v(a)\}$$

$$= \{(a, b) \in \mathbb{R}^2 : v_0 = v_0(a, b) \geqslant 0 \text{ and } v_0 \in \Lambda\}.$$

For any $\gamma \geqslant 0$ put $\Theta_\gamma := \{(a, b) \in \Theta : v_0(a, b) = \gamma\}$. It is easy to see that Θ_γ is a straight half-line: $\Theta_\gamma = \{(a, b) \in \mathbb{R}^2 : b = \mathrm{e}^\gamma(\gamma - a),\ a \leqslant 1 + \gamma\}$. The sets Θ_γ, $\gamma \geqslant 0$, form a family of disjoint rays covering Θ and having the curve $b = v(a)$, $a \geqslant 1$, as envelope. Note also that $\tilde{\Theta}_\gamma := \{(a, b) \in \mathbb{R}^2 : a + b\,\mathrm{e}^{-\gamma} = \gamma\} \subseteq \Theta$ is the tangent line to $v(\cdot)$ at the point $(1 + \gamma, \mathrm{e}^{-\gamma})$, that the points in $\tilde{\Theta}_\gamma \backslash \Theta_\gamma$ correspond to case Q4 if $\gamma = 0$ and to case M2 if $\gamma > 0$, and that $v_1 = v_1(a, b) = \gamma$ if $(a, b) \in \tilde{\Theta}_\gamma \backslash \Theta_\gamma$.

In the rest of this section we shall assume that the true value $\vartheta_0 = (a, b)^*$ of the parameter corresponds to one of cases M1–M3, P1, Q1, Q2, Q4, or Q5, that is, $\vartheta_0 \in \Theta$. Real solutions of the characteristic equation (1.9) exist and, as before, v_0 denotes the maximal one. By construction, $\vartheta_0 \in \Theta_{v_0}$.

First, we note that the normalization of $\hat{\vartheta}_T - \vartheta_0$ by the scalar $\varphi_{22}^{-1}(T)$ leads to a non-trivial limit distribution which is concentrated on the straight line passing through the origin and parallel to $\tilde{\Theta}_{v_0}$. (This and subsequent remarks are modified in an obvious way in the periodic case P1.) Indeed, we obtain

$$\begin{aligned}
\varphi_{22}^{-1}(T)(\hat{\vartheta}_T - \vartheta_0) &= \varphi_{22}^{-1}(T)\varphi_T\hat{\mu}_T = \varphi_{22}^{-1}(T)\varphi_T^{(1)}\varphi_T^{(2)}\hat{\mu}_T \\
&= \varphi_T^{(1)}\begin{pmatrix} \varphi_{22}^{-1}(T)\varphi_{11}(T) & 0 \\ 0 & 1 \end{pmatrix}\hat{\mu}_T \\
&\xrightarrow{d} \begin{pmatrix} 1 & 1 \\ 0 & -\mathrm{e}^{v_0} \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} I_\infty^{-1} V_\infty \\
&= \begin{pmatrix} 0 & 1 \\ 0 & -\mathrm{e}^{v_0} \end{pmatrix} I_\infty^{-1} V_\infty. \qquad (3.4)
\end{aligned}$$

In particular, the rate of convergence of $\hat{\vartheta}_T$ to ϑ_0 is $\varphi_{22}^{-1}(T)$ and $\hat{\vartheta}_T$ lies near the straight line $\tilde{\Theta}_{v_0}$ in the sense that the distance between $\hat{\vartheta}_T$ and $\tilde{\Theta}_{v_0}$ is of smaller order than $\varphi_{22}(T)$.

In this connection it is of some interest to see what happens if one of the parameters or a linear combination of them is known. Here we shall concentrate on the maximum likelihood estimators. The corresponding arguments concerning local asymptotic properties are similar and omitted.

Assume that the parameter ϑ belongs to a straight line Θ' which meets ϑ_0. The limit behaviour of the maximum likelihood estimator

$$\hat{\vartheta}'_T = \arg\max_{\vartheta \in \Theta'} \ell_T^0(\vartheta)$$

is essentially different in the following two cases: (1) $\Theta' \neq \tilde{\Theta}_{v_0}$; (2) $\Theta' = \tilde{\Theta}_{v_0}$. Denote by $\mathfrak{M}_T$ the image of Θ' by the map

$$\vartheta \to \varphi_T^{-1}(\vartheta - \vartheta_0)$$

and by $\hat{\mu}'_T$ the maximum likelihood estimator

$$\hat{\mu}'_T = \arg\max_{\mu \in \mathfrak{M}_T} \ell_T(\mu).$$

Then

$$\hat{\vartheta}'_T = \vartheta_0 + \varphi_T \hat{\mu}'_T,$$

Case (1). It is easy to see that $\mathfrak{M}_T$ is a straight line passing through (0, 0) with slope tending to zero if $T \to \infty$. So we obtain, with the notation $\mu = (\alpha, \beta)^*$,

$$\hat{\mu}'_T \xrightarrow{d} \arg\max_{\beta=0, \alpha \in \mathbb{R}} \left(\mu^* V_\infty - \frac{1}{2} \mu^* I_\infty \mu \right). \tag{3.5}$$

This means that the limit distribution of $\hat{\mu}'_T$ is the distribution of the vector $\Phi(1, 0)^*$, where $\Phi = V_{\infty,1}/I_{\infty,11}$ ($V_{\infty,i}$ are the elements of V_∞ and $I_{\infty,ii}$ are the diagonal elements of I_∞, $i = 1, 2$). Since $\varphi_{11}^{-1}(T)(1, \mathrm{e}^{-v_0})\varphi_T \to (1, 0)$, we obtain

$$\varphi_{11}^{-1}(T)(1, \mathrm{e}^{-v_0})(\hat{\vartheta}'_T - \vartheta_0) \xrightarrow{d} \Phi,$$

hence

$$\varphi_{11}^{-1}(T)(\hat{\vartheta}'_T - \vartheta_0) \xrightarrow{d} \Phi(1 + c, -\mathrm{e}^{v_0} c)^*,$$

for some real c. The rate of convergence of $\hat{\vartheta}'_T$ to ϑ_0 equals $\varphi_{11}^{-1}(T)$.

Case (2). Let us assume additionally that we are not in case M3. Then $\mathfrak{M}_T = \{(\alpha, \beta) : \alpha = 0\}$ and we obtain

$$\hat{\mu}'_T \xrightarrow{d} \arg\max_{\alpha=0, \beta \in \mathbb{R}} \left(\mu^* V_\infty - \frac{1}{2} \mu^* I_\infty \mu \right).$$

Thus the limit distribution of $\hat{\mu}'_T$ is the distribution of the vector $\Psi(0, 1)^*$, where $\Psi = V_{\infty,2}/I_{\infty,22}$. Now it is easy to see that

$$\varphi_{22}^{-1}(T)(\hat{\vartheta}'_T - \vartheta_0) \xrightarrow{d} \Psi(1, -\mathrm{e}^{v_0})^*. \tag{3.6}$$

Therefore, the rate of convergence of $\hat{\vartheta}'_T$ to ϑ_0 is $\varphi_{22}^{-1}(T)$. Moreover, if I_∞ is diagonal (this happens in cases M1, Q1, Q4, and Q5) then

$$\varphi_{22}^{-1}(T)(\hat{\vartheta}_T - \hat{\vartheta}'_T) \xrightarrow{P} 0 \tag{3.7}$$

(compare (3.4) and (3.6)). Furthermore, $I_{\infty,22}$ is non-random in cases M1 and Q1. Hence the submodel $(P_\vartheta,\ \vartheta \in \tilde{\Theta}_{v_0})$ is locally asymptotically normal in these two cases. Applying the asymptotic minimax theorem to this submodel, we obtain the following local asymptotic minimax bound for an arbitrary estimator $\tilde{\vartheta}_T$ in the global model:

$$\lim_{r\to\infty} \liminf_{T\to\infty} \sup_{T^{1/2}\|\vartheta-\vartheta_0\|\leqslant r} \mathrm{E}_\vartheta w\{T^{1/2}\|\tilde{\vartheta}_T - \vartheta\|\} \geqslant \mathrm{E}w\{|\sigma|^{-1}(1+\mathrm{e}^{2v_0})^{1/2} Z\},$$

where Z is a standard normal variable, $\sigma^2 = \int_0^\infty (x_0(t) - \mathrm{e}^{v_0} x_0(t-1))^2\,\mathrm{d}t$ and $w : \mathbb{R} \to [0, \infty)$ is a bowl-shaped loss function; here ϑ_0 satisfies $a < 1$, $-a \leqslant b < w(a)$ or $a \geqslant 1$, $-a < b < w(a)$ (corresponding to cases M1 and Q1, respectively). Note that a similar estimate can be obtained from (3.3) also in case N. Due to (3.7), the maximum likelihood estimator $\hat{\vartheta}_T$ attains this bound, at least for bounded w.

Finally, let us consider case M3. Here we have $\varphi_{11}(T) = T^{-1}\,\mathrm{e}^{-v_0 T}$ and $\varphi_{22}(T) = T\,\mathrm{e}^{-v_0 T}$. Thus $\mathfrak{M}_T$ is the straight line passing through (0, 0) with slope $1/T$. The estimator $\hat{\mu}'_T$ has the limit distribution as in (3.5) above. This implies

$$\mathrm{e}^{v_0 T}(\hat{\vartheta}'_T - \vartheta_0) \xrightarrow{d} \Phi(1, -\mathrm{e}^{v_0})^*.$$

So here the rate of convergence of $\hat{\vartheta}'_T$ to ϑ_0 is intermediate between $\varphi_{11}^{-1}(T)$ and $\varphi_{22}^{-1}(T)$.

4. Proofs

The main goal of this section is to prove Propositions 2.1–2.11, that is, to prove the weak convergence of (V_T, I_T) to the corresponding limit. Unless otherwise specified, all limits are taken as $T \to \infty$.

Let us start with some general remarks. With the exception of case M3, we have

$$V_T = \left(\varphi_{11}(T)\int_0^T X(t)\,\mathrm{d}W(t),\ \varphi_{22}(T)\int_0^T Y(t)\,\mathrm{d}W(t)\right)^* \tag{4.1}$$

and

$$I_T = \begin{pmatrix} \varphi_{11}^2(T)\int_0^T X^2(t)\,\mathrm{d}t & \varphi_{11}(T)\varphi_{22}(T)\int_0^T X(t)Y(t)\,\mathrm{d}t \\ \varphi_{11}(T)\varphi_{22}(T)\int_0^T X(t)Y(t)\,\mathrm{d}t & \varphi_{22}^2(T)\int_0^T Y^2(t)\,\mathrm{d}t \end{pmatrix}, \tag{4.2}$$

where

$$Y(t) = \begin{cases} X(t-1) & \text{in cases N, P2, Q3,} \\ X(t) - \mathrm{e}^{v_0} X(t-1) & \text{otherwise.} \end{cases} \tag{4.3}$$

Note that the process $Y(t)$ defined in this way has a representation similar to (1.2), where the function $x_0(t)$ is replaced by the linear combination of $x_0(t)$ and $x_0(t-1)$ corresponding to (4.3); this representation holds for $t \geqslant 1$.

More generally, we will consider the following representation of a continuous process $(Y(t),\ t \geqslant 0)$ based on a function $y(\cdot)$:

$$Y(t) = y(t)X_0(0) + b\int_{-1}^{0} y(t-s-1)X_0(s)\,\mathrm{d}s + \int_0^t y(t-s)\,\mathrm{d}W(s), \qquad t \geqslant t_0, \tag{4.4}$$

where $t_0 \geqslant 1$ and $y = (y(t),\ t \geqslant 0)$ is a deterministic continuous function. Before proving Propositions 2.1–2.11, we shall study some properties of processes with representation (4.4).

Our first lemma summarizes in an appropriate form some simple facts used over and over throughout this section. The proof is trivial and therefore omitted.

Lemma 4.1. *Assume* $Y_1(t)$, $Y_2(t)$ *and* $Z(t)$, $t \geqslant 0$, *are adapted continuous processes,* $Y(t) = Y_1(t) + Y_2(t)$, $t \geqslant 0$, *and* $(W(t),\ t \geqslant 0)$ *is a standard Wiener process. Moreover, let* $\delta(T)$ *and* $\varepsilon(T)$ *be normalizing functions such that*

$$\delta^2(T)\int_0^T Y_1^2(t)\,\mathrm{d}t, \qquad T \geqslant 0, \qquad \text{and} \qquad \varepsilon^2(T)\int_0^T Z^2(t)\,\mathrm{d}t, \qquad T \geqslant 0,$$

are bounded in probability and

$$\delta^2(T)\int_0^T Y_2^2(t)\,\mathrm{d}t \xrightarrow{P} 0.$$

Then

$$\delta(T)\left\{\int_0^T Y(t)\,\mathrm{d}W(t) - \int_0^T Y_1(t)\,\mathrm{d}W(t)\right\} \xrightarrow{P} 0,$$

$$\delta^2(T)\left\{\int_0^T Y^2(t)\,\mathrm{d}t - \int_0^T Y_1^2(t)\,\mathrm{d}t\right\} \xrightarrow{P} 0.$$

$$\delta(T)\varepsilon(T)\left\{\int_0^T Y(t)Z(t)\,\mathrm{d}t - \int_0^T Y_1(t)Z(t)\,\mathrm{d}t\right\} \xrightarrow{P} 0.$$

Let $(Y(t),\ t \geqslant 0)$ be a process with representation (4.4). Sometimes the first term on the right-hand side of (4.4) is small in the sense of Lemma 4.1, that is, it can be chosen as $Y_2(t)$. The next lemma shows that then the second term on the right-hand side of (4.4) is also small in the same sense.

Lemma 4.2. *Put*

$$z(t) = \int_{-1}^{0} y(t-s-1)X_0(s)\,\mathrm{d}s, \qquad t \geqslant t_0.$$

Then

$$\int_{t_0}^{T} z^2(t)\,\mathrm{d}t \leqslant \int_{-1}^{0} X_0^2(s)\,\mathrm{d}s \int_{0}^{T} y^2(t)\,\mathrm{d}t.$$

To prove this, use Fubini's theorem and the Cauchy–Schwarz inequality.

In Lemmas 4.3, 4.5, 4.7, 4.8 and Corollary 4.4 we assume that $Y(\cdot)$, $Y_1(\cdot)$, $Y_2(\cdot)$ are continuous processes having representation (4.4) with functions $y(\cdot)$, $y_1(\cdot)$, $y_2(\cdot)$, respectively.

Lemma 4.3. *Assume that* $y = (y(t),\, t \geqslant 0)$ *is a square-integrable function. Then*

$$T^{-1}\int_0^T Y(t)\,\mathrm{d}t \xrightarrow{P} 0,$$

$$T^{-1}\int_0^T Y^2(t)\,\mathrm{d}t \xrightarrow{P} \sigma^2 := \int_0^\infty y^2(t)\,\mathrm{d}t.$$

Proof. According to Lemmas 4.1 and 4.2, it is sufficient to prove the assertion for $X_0(s) \equiv 0$. We introduce the stationary process $Z(t) = \int_{-\infty}^{t} y(t-s)\,\mathrm{d}W(s)$, $t \geqslant 0$, where $W(\cdot)$ is extended to $(-\infty, 0)$ as a Wiener process independently of $(W(s),\, s \geqslant 0)$.

Obviously, we have

$$T^{-1}\mathrm{E}\int_{t_0}^{T} (Z(t) - Y(t))^2\,\mathrm{d}t = T^{-1}\int_{t_0}^{T}\int_{t}^{\infty} y^2(s)\,\mathrm{d}s\,\mathrm{d}t \to 0. \tag{4.5}$$

Applying the law of large numbers to the Gaussian stationary process $Z(\cdot)$, which is ergodic, we obtain

$$T^{-1}\int_0^T Z(t)\,\mathrm{d}t \xrightarrow{P} \mathrm{E}Z(0) = 0, \qquad T^{-1}\int_0^T Z^2(t)\,\mathrm{d}t \xrightarrow{P} \mathrm{E}Z^2(0) = \sigma^2.$$

Now the claim follows from (4.5). □

Corollary 4.4. *If* $\int_0^\infty y_i^2(t)\,\mathrm{d}t < \infty$, $i = 1, 2$, *then*

$$T^{-1}\int_0^T Y_1(t)Y_2(t)\,\mathrm{d}t \xrightarrow{P} \int_0^\infty y_1(t)y_2(t)\,\mathrm{d}t.$$

Lemma 4.5. *Suppose that* $y(t) = t^\alpha \mathrm{e}^{wt}$ *for some* $\alpha = 0, 1, 2, \ldots$ *and* $w > 0$. *Then with probability one*

$$\lim_{t\to\infty} t^{-\alpha}\,\mathrm{e}^{-wt}\,Y(t) = U,$$

where

$$U = X_0(0) + b\int_{-1}^{0} \mathrm{e}^{-w(s+1)}X_0(s)\,\mathrm{d}s + \int_0^\infty \mathrm{e}^{-ws}\,\mathrm{d}W(s).$$

Proof. Using the representation (4.4) of $Y(\cdot)$, the appearance of the first two terms of U is quite obvious. Furthermore,

$$t^{-\alpha}\,\mathrm{e}^{-wt}\int_0^t y(t-s)\,\mathrm{d}W(s) = \int_0^t \left(1-\frac{s}{t}\right)^{\alpha}\mathrm{e}^{-ws}\,\mathrm{d}W(s)$$

$$= \int_0^t \mathrm{e}^{-ws}\,\mathrm{d}W(s) + \sum_{k=1}^{\alpha}(-1)^k\binom{\alpha}{k}t^{-k}\int_0^t s^k\,\mathrm{e}^{-ws}\,\mathrm{d}W(s).$$

It remains to note that, with probability one,

$$\lim_{t\to\infty}\int_0^t \mathrm{e}^{-ws}\,\mathrm{d}W(s) = \int_0^\infty \mathrm{e}^{-ws}\,\mathrm{d}W(s)$$

by Lévy's theorem and

$$\lim_{t\to\infty} t^{-k}\int_0^t s^k\,\mathrm{e}^{-ws}\,\mathrm{d}W(s) = 0$$

by the strong law of large numbers for martingales; see, for example, Liptser and Shiryaev (1989, Chapter 2, §6, Theorem 10). □

Lemma 4.6. *Let $Z_1(\cdot)$ and $Z_2(\cdot)$ be two continuous processes such that with probability one*

$$\lim_{t\to\infty} t^{-\alpha_i}\,\mathrm{e}^{-w_i t}Z_i(t) = U_i,$$

for some $\alpha_i \in \mathbb{R}$, $w_i > 0$, and some random variables U_i almost surely finite, $i = 1, 2$. Then

$$T^{-\alpha_1-\alpha_2}\,\mathrm{e}^{-(w_1+w_2)T}\int_0^T Z_1(t)Z_2(t)\,\mathrm{d}t \xrightarrow{P} \frac{U_1U_2}{w_1+w_2}$$

and

$$T^{-\alpha_1-1/2}\,\mathrm{e}^{-w_1T}\int_0^T |Z_1(t)|\,\mathrm{d}t \xrightarrow{P} 0.$$

Remark. In fact, we have the almost sure convergence in the assertions of Lemma 4.6. For the proof, apply L'Hôpital's rule.

Lemma 4.7. *Suppose that a continuous process $Y(\cdot)$ has representation* (4.4) *with a bounded $y(\cdot)$. If $Z(\cdot)$ is a continuous process such that with probability one*

$$\lim_{t\to\infty}\mathrm{e}^{-wt}Z(t) = U$$

for some $w > 0$ and some random variable U almost surely finite, then

$$T^{-1}\,\mathrm{e}^{-wT}\int_0^T Y(t)Z(t)\,\mathrm{d}t \xrightarrow{P} 0.$$

If, moreover, $y(\cdot)$ *is square-integrable on* $[0, \infty)$*, then this convergence holds for* $T^{-1/2}$ *instead of* T^{-1}.

Proof. We can use Lemmas 4.2 and 4.6 to apply Lemma 4.1 and therefore we can assume that $X_0(0) \equiv 0$. Applying these lemmas again, we can substitute $Z(t)$ by $\mathrm{e}^{wt}U$. Thus it remains to prove that

$$T^{-1}\,\mathrm{e}^{-wT}\int_0^T \mathrm{e}^{wt}\,Y(t)\,\mathrm{d}t \xrightarrow{P} 0$$

(or $T^{-1/2}$ instead of T^{-1} if $\int_0^\infty y^2(t)\,\mathrm{d}t < \infty$). Now observe that

$$\begin{aligned}\mathrm{E}\left|\int_0^T \mathrm{e}^{wt}\,Y(t)\,\mathrm{d}t\right| &\leqslant \int_0^T \mathrm{e}^{wt}\mathrm{E}|Y(t)|\mathrm{d}t \\ &\leqslant \int_0^T \mathrm{e}^{wt}(\mathrm{E}|Y(t)|^2)^{1/2}\,\mathrm{d}t\end{aligned}$$

and that

$$\mathrm{E}|Y(t)|^2 = \int_0^t y^2(s)\,\mathrm{d}s, \qquad t \geqslant t_0,$$

which implies the assertion. □

Lemma 4.8. *Assume that* $y(t) = \phi(t)\mathrm{e}^{wt}$*, where* $\phi(t) = \cos(\xi t)$ *or* $\phi(t) = \sin(\xi t)$ *and* $w > 0$*. Then with probability one*

$$\lim_{t\to\infty}\{\mathrm{e}^{-wt}\,Y(t) - U(t)\} = 0,$$

where

$$U(t) = X_0(0)\phi(t) + b\int_{-1}^0 \phi(t-s-1)\mathrm{e}^{-w(s+1)}X_0(s)\,\mathrm{d}s + \int_0^\infty \phi(t-s)\mathrm{e}^{-ws}\,\mathrm{d}W(s)$$

is a continuous periodic process.

Proof. Note that

$$\mathrm{e}^{-wt}\,Y(t) - U(t) = -\int_t^\infty \phi(t-s)\mathrm{e}^{-ws}\,\mathrm{d}W(s), \qquad t \geqslant t_0.$$

If $\phi(t) = \cos(\xi t)$ then

$$\int_t^\infty \phi(t-s)\mathrm{e}^{-ws}\,\mathrm{d}W(s) = \cos(\xi t)\int_t^\infty \cos(\xi s)\mathrm{e}^{-ws}\,\mathrm{d}W(s) + \sin(\xi t)\int_t^\infty \sin(\xi s)\mathrm{e}^{-ws}\,\mathrm{d}W(s),$$

which obviously tends almost surely to zero. The case $\phi(t) = \sin(\xi t)$ can be treated similarly. □

Lemma 4.9. *Let $Z(\ldots,)$ be a continuous process such that with probability one*

$$\lim_{t\to\infty}\{\mathrm{e}^{-wt}Z(t)-U(t)\}=0,$$

where $U(\cdot)$ is a continuous periodic process on $\mathbb{R}$ and $w>0$. Then

$$\mathrm{e}^{-2wT}\int_0^T Z^2(t)\,\mathrm{d}t-\int_0^\infty \mathrm{e}^{-2wt}U^2(T-t)\,\mathrm{d}t\xrightarrow{P}0.$$

Proof. Applying Lemma 4.1, we can replace $Z(t)$ by $\mathrm{e}^{wt}U(t)$ and observe that

$$\mathrm{e}^{-2wT}\int_0^T \mathrm{e}^{2wt}U^2(t)\,\mathrm{d}t=\int_0^T \mathrm{e}^{-2wt}U^2(T-t)\,\mathrm{d}t$$

and

$$\int_T^\infty \mathrm{e}^{-2wt}U^2(T-t)\,\mathrm{d}t\xrightarrow{P}0.$$ □

Lemma 4.10. *Let $Z_1(\cdot)$ and $Z_2(\cdot)$ be two continuous processes such that with probability one*

$$\lim_{t\to\infty}\mathrm{e}^{-w_1t}Z_1(t)=U_1$$

and

$$\lim_{t\to\infty}\{\mathrm{e}^{-w_2t}Z_2(t)-U_2(t)\}=0,$$

where w_1, $w_2>0$, U_1 is a finite random variable and $U_2(\cdot)$ is a continuous periodic process on $\mathbb{R}$. Then

$$\mathrm{e}^{-(w_1+w_2)T}\int_0^T Z_1(t)Z_2(t)\,\mathrm{d}t-U_1\int_0^\infty \mathrm{e}^{-(w_1+w_2)t}U_2(T-t)\,\mathrm{d}t\xrightarrow{P}0.$$

Proof. The proof is analogous to that of Lemma 4.9. □

Now we are in a position to prove Proposition 2.1–2.8.

Proof of Proposition 2.1. According to (4.1)–(4.3),

$$V_T=\left(T^{-1/2}\int_0^T X(t)\,\mathrm{d}W(t),\ T^{-1/2}\int_0^T X(t-1)\,\mathrm{d}W(t)\right)^*$$

and

$$I_T=\begin{pmatrix} T^{-1}\int_0^T X^2(t)\,\mathrm{d}t & T^{-1}\int_0^T X(t)X(t-1)\,\mathrm{d}t \\ T^{-1}\int_0^T X(t)X(t-1)\,\mathrm{d}t & T^{-1}\int_0^T X^2(t-1)\,\mathrm{d}t \end{pmatrix}.$$

The process $(X(t),\ t \geqslant 0)$ has representation (4.4) with $y(t) = x_0(t)$, $t \geqslant 0$, and $(X(t-1),\ t \geqslant 0)$ has this representation with $y(t) = x_0(t-1)$, $t \geqslant 0$.

By assumption, $v_0 < 0$, that is, $\int_0^\infty x_0^2(t)\,\mathrm{d}t < \infty$ holds. Thus we can apply Lemma 4.3 and Corollary 4.4 to obtain $I_T \xrightarrow{P} I_\infty$. Now the claim follows from the central limit theorem; see, for example, Basawa and Prakasa Rao (1980, Theorem 2.1, Appendix 2, p. 405). □

Proof of Proposition 2.2. According to (4.1)–(4.3),

$$V_T = \left(\mathrm{e}^{-v_0 T} \int_0^T X(t)\,\mathrm{d}W(t),\ T^{-1/2} \int_0^T Y(t)\,\mathrm{d}W(t) \right)^*$$

and

$$I_T = \begin{pmatrix} \mathrm{e}^{-2v_0 T} \int_0^T X^2(t)\,\mathrm{d}t & T^{-1/2}\,\mathrm{e}^{-v_0 T} \int_0^T X(t)Y(t)\,\mathrm{d}t \\ T^{-1/2}\,\mathrm{e}^{-v_0 T} \int_0^T X(t)Y(t)\,\mathrm{d}t & T^{-1} \int_0^T Y^2(t)\,\mathrm{d}t \end{pmatrix},$$

where

$$Y(t) = X(t) - \mathrm{e}^{v_0} X(t-1), \qquad t \geqslant 0. \tag{4.6}$$

Note that $Y(\cdot)$ has representation (4.4) with $y(t) = x_0(t) - \mathrm{e}^{v_0} x_0(t-1)$. In the case considered $v_0 > 0$ and $v_1 < 0$. It follows from Lemma 1.1 that

$$x_0(t) = \frac{1}{v_0 - a + 1}\,\mathrm{e}^{v_0 t} + o(\mathrm{e}^{\gamma t}) \tag{4.7}$$

for some $\gamma < 0$, and this implies $y(t) = o(\mathrm{e}^{\gamma t})$ and therefore $\int_0^\infty y^2(t)\,\mathrm{d}t < \infty$. Now Lemmas 4.5, 4.6 and 4.1 imply that

$$\mathrm{e}^{-2v_0 T} \int_0^T X^2(t)\,\mathrm{d}t \xrightarrow{P} \frac{U_0^2}{2v_0(v_0 - a + 1)^2},$$

it follows from Lemmas 4.5, 4.7 and 4.1 that

$$T^{-1/2}\,\mathrm{e}^{-v_0 T} \int_0^T X(t)Y(t)\,\mathrm{d}t \xrightarrow{P} 0,$$

and Lemma 4.3 implies that

$$T^{-1} \int_0^T Y^2(t)\,\mathrm{d}t \xrightarrow{P} \int_0^\infty y^2(t)\,\mathrm{d}t.$$

Summarizing these results, we obtain the convergence in probability of I_T to I_∞. The joint convergence of (V_T, I_T) to (V_∞, I_∞) follows from the stable limit theorem for martingales; see Jacod and Shiryaev (1987, Theorem VIII.5.42 and Example VIII.5.38) or Touati (1991, Theorem 1). □

Proof of Proposition 2.3. According to (4.1)–(4.3),

$$V_T = \left(e^{-v_0 T} \int_0^T X(t) \, dW(t), \, e^{-v_1 T} \int_0^T Y(t) \, dW(t) \right)^*$$

and

$$I_T = \begin{pmatrix} e^{-2v_0 T} \int_0^T X^2(t) \, dt & e^{-(v_0+v_1)T} \int_0^T X(t)Y(t) \, dt \\ e^{-(v_0+v_1)T} \int_0^T X(t)Y(t) \, dt & e^{-2v_1 T} \int_0^T Y^2(t) \, dt \end{pmatrix},$$

where $Y(t)$ is defined as in (4.6) above and has representation (4.4) with $y(t) = x_0(t) - e^{v_0} x_0(t-1)$. As in the previous proposition, it is sufficient to check that $I_T \xrightarrow{P} I_\infty$.

Since $v_0 > v_1 > 0$, $v_0 \in \Lambda$ and $v_1 \in \Lambda$ in the case considered, it follows from Lemma 1.1 and (1.12) that (4.7) holds for some $\gamma < v_0$ and

$$y(t) = \frac{e^{v_0 - v_1} - 1}{a - v_1 - 1} e^{v_1 t} + o(e^{\gamma_1 t})$$

for some $\gamma_1 < v_1$.

Using Lemmas 4.5, 4.6 and 4.1, we obtain

$$e^{-2v_0 T} \int_0^T X^2(t) \, dt \xrightarrow{P} \frac{U_0^2}{2v_0(v_0 - a + 1)^2},$$

$$e^{-(v_0+v_1)T} \int_0^T X(t)Y(t) \, dt \xrightarrow{P} \frac{U_0 U_1 (e^{v_0 - v_1} - 1)}{(v_0 + v_1)(v_0 - a + 1)(a - v_1 - 1)}$$

and

$$e^{-2v_1 T} \int_0^T Y^2(t) \, dt \xrightarrow{P} \frac{U_1^2 (e^{v_0 - v_1} - 1)^2}{2v_1(a - v_1 - 1)^2},$$

which yields the desired convergence. □

Proof of Proposition 2.4. By the choice of φ_T, we have

$$V_T = \left(T^{-1} e^{-v_0 T} \int_0^T X(t) \, dW(t), \, T e^{-v_0 T} \int_0^T (Y(t) - T^{-1} Z(t)) \, dW(t) \right)^*$$

and

$$I_T = \begin{pmatrix} T^{-2} e^{-2v_0 T} \int_0^T X^2(t) \, dt & e^{-2v_0 T} \int_0^T X(t)(Y(t) - T^{-1} Z(t)) \, dt \\ e^{-2v_0 T} \int_0^T X(t)(Y(t) - T^{-1} Z(t)) \, dt & T^2 e^{-2v_0 T} \int_0^T (Y(t) - T^{-1} Z(t))^2 \, dt \end{pmatrix},$$

where $Z(t) = \mathrm{e}^{v_0} X(t-1)$ and $Y(t)$ is defined as in (4.6). Obviously $Z(\cdot)$ has representation (4.4) with $z(t) = \mathrm{e}^{v_0} x(t-1)$. It follows from (1.11) that

$$x_0(t) = (2t + \tfrac{2}{3})\mathrm{e}^{v_0 t} + o(\mathrm{e}^{\gamma t}),$$
$$y(t) = 2\,\mathrm{e}^{v_0 t} + o(\mathrm{e}^{\gamma t}),$$
$$z(t) = (2t - \tfrac{4}{3})\mathrm{e}^{v_0 t} + o(\mathrm{e}^{\gamma t})$$

for some $0 < \gamma < v_0$.

Put

$$\hat{y}(t) = 2\,\mathrm{e}^{v_0 t}, \qquad \hat{z}(t) = (2t - \tfrac{4}{3})\mathrm{e}^{v_0 t},$$

and let $\hat{Y}(t)$ and $\hat{Z}(t)$ be continuous processes having representation (4.4) with the functions $\hat{y}(t)$ and $\hat{z}(t)$ respectively, $\hat{X}(t) = \hat{Y}(t) + \hat{Z}(t)$. It can be easily checked that

$$\mathrm{e}^{-2\gamma T} \int_0^T (X(t) - \hat{X}(t))^2 \,\mathrm{d}t \xrightarrow{P} 0 \tag{4.8}$$

$$\mathrm{e}^{-2\gamma T} \int_0^T (Y(t) - \hat{Y}(t))^2 \,\mathrm{d}t \xrightarrow{P} 0 \tag{4.9}$$

and

$$\mathrm{e}^{-2\gamma T} \int_1^T t^{-2} (Z(t) - \hat{Z}(t))^2 \,\mathrm{d}t \xrightarrow{P} 0. \tag{4.10}$$

Lemma 4.5 implies that with probability one

$$\lim_{t \to \infty} t^{-1}\,\mathrm{e}^{-v_0 t} \hat{X}(t) = 2U_0 \tag{4.11}$$

and

$$\lim_{t \to \infty} t^{-1}\,\mathrm{e}^{-v_0 t} \hat{Z}(t) = 2U_0; \tag{4.12}$$

the same proof as in Lemma 4.5 shows that

$$\lim_{t \to \infty} t\,\mathrm{e}^{-v_0 t} (\hat{Y}(t) - t^{-1} \hat{Z}(t)) = \tfrac{4}{3} U_0 + 2U_2. \tag{4.13}$$

By Lemma 4.6 we obtain

$$T^{-2}\,\mathrm{e}^{-2v_0 T} \int_0^T \hat{X}^2(t) \,\mathrm{d}t \xrightarrow{P} \frac{2U_0^2}{v_0}, \tag{4.14}$$

$$\mathrm{e}^{-2v_0 T} \int_1^T \hat{X}(t)(\hat{Y}(t) - t^{-1}\hat{Z}(t)) \,\mathrm{d}t \xrightarrow{P} \frac{U_0(\tfrac{4}{3}U_0 + 2U_2)}{v_0}, \tag{4.15}$$

and

$$T^2\,\mathrm{e}^{-2v_0 T} \int_1^T (\hat{Y}(t) - t^{-1}\hat{Z}(t))^2 \,\mathrm{d}t \xrightarrow{P} \frac{(\tfrac{4}{3}U_0 + 2U_2)^2}{2v_0}. \tag{4.16}$$

It follows from (4.11)–(4.13) by L'Hôpital's rule that

$$\mathrm{e}^{-2v_0 T}\int_1^T \hat{X}(t)\hat{Z}(t)(t^{-1} - T^{-1})\,\mathrm{d}t \xrightarrow{P} \frac{U_0^2}{v_0^2}, \tag{4.17}$$

$$T^2\,\mathrm{e}^{-2v_0 T}\int_1^T \hat{Z}^2(t)(t^{-1} - T^{-1})^2\,\mathrm{d}t \xrightarrow{P} \frac{U_0^2}{v_0^3}, \tag{4.18}$$

$$T^2\,\mathrm{e}^{-2v_0 T}\int_1^T (\hat{Y}(t) - t^{-1}\hat{Z}(t))\hat{Z}(t)(t^{-1} - T^{-1})\,\mathrm{d}t \xrightarrow{P} \frac{U_0(\frac{4}{3}U_0 + 2U_2)}{2v_0^2}. \tag{4.19}$$

It follows from (4.8)–(4.10) that we can replace $\hat{X}(t)$, $\hat{Y}(t)$ and $\hat{Z}(t)$ by $X(t)$, $Y(t)$ and $Z(t)$ respectively in relations (4.14)–(4.19). This implies the convergence $I_T \xrightarrow{P} I_\infty$. Now the claim follows from the stable limit theorem for martingales as in Propositions 2.2 and 2.3. □

Proof of Proposition 2.5. According to (4.1)–(4.3),

$$V_T = \left(\mathrm{e}^{-v_0 T}\int_0^T X(t)\,\mathrm{d}W(t),\ \mathrm{e}^{-v_1 T}\int_0^T Y(t)\,\mathrm{d}W(t)\right)^*$$

and

$$I_T = \begin{pmatrix} \mathrm{e}^{-2v_0 T}\int_0^T X^2(t)\,\mathrm{d}t & \mathrm{e}^{-(v_0+v_1)T}\int_0^T X(t)Y(t)\,\mathrm{d}t \\ \mathrm{e}^{-(v_0+v_1)T}\int_0^T X(t)Y(t)\,\mathrm{d}t & \mathrm{e}^{-2v_1 T}\int_0^T Y^2(t)\,\mathrm{d}t \end{pmatrix},$$

where

$$Y(t) = X(t) - \mathrm{e}^{v_0} X(t-1), \qquad t \geqslant 0.$$

Note that $Y(\cdot)$ has representation (4.4) with $y(t) = x_0(t) - \mathrm{e}^{v_0} x_0(t-1)$. It follows from (1.12) that

$$x_0(t) = \frac{1}{v_0 + 1 - a}\,\mathrm{e}^{v_0 t} + \{A_1\cos(\xi_1 t) + B_1\sin(\xi_1 t)\}\mathrm{e}^{v_1 t} + o(\mathrm{e}^{\gamma t})$$

for some $\gamma < v_1$ and, hence,

$$y(t) = \phi(t)\mathrm{e}^{v_1 t} + o(\mathrm{e}^{\gamma t}).$$

Applying Lemmas 4.5, 4.6, 4.8, 4.9, 4.10 and 4.1, we obtain

$$I_T - I_\infty(T) \xrightarrow{P} 0.$$

Now we complete the proof similarly to the previous case. The matrix-valued process $I_\infty(T)$ is periodic with period $\Delta = 2\pi/\xi_1$, and the claim follows from the stable limit theorem for martingales as in previous propositions. □

Proof of Proposition 2.6. According to (4.1)–(4.3),

$$V_T = \left(\mathrm{e}^{-v_0 T} \int_0^T X(t)\,\mathrm{d}W(t),\ \mathrm{e}^{-v_0 T} \int_0^T X(t-1)\,\mathrm{d}W(t) \right)^*$$

and

$$I_T = \begin{pmatrix} \mathrm{e}^{-2v_0 T} \int_0^T X^2(t)\,\mathrm{d}t & \mathrm{e}^{-2v_0 T} \int_0^T X(t)X(t-1)\,\mathrm{d}t \\ \mathrm{e}^{-2v_0 T} \int_0^T X(t)X(t-1)\,\mathrm{d}t & \mathrm{e}^{-2v_0 T} \int_0^T X^2(t-1)\,\mathrm{d}t \end{pmatrix}.$$

The process $(X(t),\ t \geqslant 0)$ has representation (4.4) with $y(t) = x_0(t),\ t \geqslant 0$, and $(X(t-1),\ t \geqslant 0)$ has this representation with $y(t) = x_0(t-1),\ t \geqslant 0$. It follows from (1.11) that

$$x_0(t) = \phi_0(t)\mathrm{e}^{v_0 t} + o(\mathrm{e}^{\gamma t})$$

for some $\gamma < v_0$ because of $\phi_0 = \psi_0$ by definition. Hence,

$$x_0(t-1) = \phi_2(t)\mathrm{e}^{v_0 t} + o(\mathrm{e}^{\gamma t}).$$

From Lemmas 4.8, 4.9 and 4.1 we now have

$$I_T - I_\infty(T) \xrightarrow{P} 0.$$

Obviously, the matrix-valued process $I_\infty(T)$ is periodic with period $\Delta = \pi/\xi_0$, and we complete the proof similarly to the previous proposition. □

Proof of Proposition 2.7. According to (4.1)–(4.3),

$$V_T = \left(T^{-1} \int_0^T X(t)\,\mathrm{d}W(t),\ T^{-1/2} \int_0^T Y(t)\,\mathrm{d}W(t) \right)^*$$

and

$$I_T = \begin{pmatrix} T^{-2} \int_0^T X^2(t)\,\mathrm{d}t & T^{-3/2} \int_0^T X(t)Y(t)\,\mathrm{d}t \\ T^{-3/2} \int_0^T X(t)Y(t)\,\mathrm{d}t & T^{-1} \int_0^T Y^2(t)\,\mathrm{d}t \end{pmatrix},$$

where

$$Y(t) = X(t) - X(t-1), \qquad t \geqslant 0.$$

Note that here we have

$$x_0(t) = \frac{1}{1-a} + o(\mathrm{e}^{\gamma t})$$

for some $\gamma < 0$, and $X(\cdot)$ and $Y(\cdot)$ have representation (4.4) with the functions $x_0(t)$ and $y(t) = x_0(t) - x_0(t-1)$, respectively. Obviously $y(t) = o(\mathrm{e}^{\gamma t})$.

Consider the processes

$$W^T(s) = T^{-1/2} W(Ts),$$

$$X^T(s) = (1-a)^{-1} T^{-1} \int_0^{Ts} W(t)\,\mathrm{d}W(t) = (1-a)^{-1} \int_0^s W^T(t)\,\mathrm{d}W^T(t),$$

$$Y^T(s) = T^{-1/2} \int_0^{Ts} Y(t)\,\mathrm{d}W(t) = \int_0^s Y(Tt)\,\mathrm{d}W^T(t), \qquad s \in [0, 1].$$

These processes are continuous local martingales. Since

$$\int_0^s Y^2(Tt)\,\mathrm{d}t = T^{-1} \int_0^{Ts} Y^2(t)\,\mathrm{d}t \xrightarrow{P} \sigma^2 s,$$

$$\int_0^s Y(Tt)\,\mathrm{d}t = T^{-1} \int_0^{Ts} Y(t)\,\mathrm{d}t \xrightarrow{P} 0$$

by Lemma 4.3, the functional central limit theorem for martingales (Jacod and Shiryaev 1987, Theorem VIII.3.11) implies that

$$(W^T, Y^T) \xrightarrow{\mathcal{L}} (\tilde{W}, \sigma \tilde{W}_1),$$

where $(\tilde{W}_1(t),\, t \in [0, 1])$ is a standard Wiener process independent of $\tilde{W}(\cdot)$. Since $X^T(s) = ((W^T(s))^2 - s)/2(1-a)$ by Itô's formula, we also have

$$(X^T, Y^T) \xrightarrow{\mathcal{L}} (\tilde{X}, \sigma \tilde{W}_1), \tag{4.20}$$

where

$$\tilde{X}(s) = (1-a)^{-1} \int_0^s \tilde{W}(t)\,\mathrm{d}\tilde{W}(t).$$

Moreover, the convergence (4.20) implies the joint functional convergence of (X^T, Y^T) together with their quadratic (co)variations; see Jacod and Shiryaev (1987, Theorem VI.6.1). In particular, $(\hat{V}_T, \hat{I}_T) \xrightarrow{d} (V_\infty, I_\infty)$, where

$$\tilde{V}_T = (X^T(1), Y^T(1))^* = \left((1-a)^{-1} T^{-1} \int_0^T W(t)\,\mathrm{d}W(t),\ T^{-1/2} \int_0^T Y(t)\,\mathrm{d}W(t) \right)^*$$

and

$$\hat{I}_T = \begin{pmatrix} (1-a)^{-2} T^{-2} \int_0^T W^2(t)\,\mathrm{d}t & (1-a)^{-2} T^{-3/2} \int_0^T W(t)Y(t)\,\mathrm{d}t \\ (1-a)^{-2} T^{-3/2} \int_0^T W(t)Y(t)\,\mathrm{d}t & T^{-1} \int_0^T Y^2(t)\,\mathrm{d}t \end{pmatrix}.$$

But, evidently, $V_T - \hat{V}_T \xrightarrow{P} 0$ and $I_T - \hat{I}_T \xrightarrow{P} 0$ by Lemmas 4.1 and 4.2. □

Proof of Proposition 2.8. According to (4.1)–(4.3),

$$V_T^* = \left(T^{-2}\int_0^T X(t)\,\mathrm{d}W(t),\ T^{-1}\int_0^T Y(t)\,\mathrm{d}W(t) \right) \tag{4.21}$$

and

$$I_T = \begin{pmatrix} T^{-4}\int_0^T X^2(t)\,\mathrm{d}t & T^{-3}\int_0^T X(t)Y(t)\,\mathrm{d}t \\ T^{-3}\int_0^T X(t)Y(t)\,\mathrm{d}t & T^{-2}\int_0^T Y^2(t)\,\mathrm{d}t \end{pmatrix}, \tag{4.22}$$

where

$$Y(t) = X(t) - X(t-1), \qquad t \geqslant 0.$$

Here we have

$$x_0(t) = (2t + \tfrac{2}{3}) + o(\mathrm{e}^{\gamma t})$$

for some $\gamma < 0$ and $X(\cdot)$ and $Y(\cdot)$ have representation (4.4) with the functions $x_0(t)$ and $y(t) = x_0(t) - x_0(t-1)$. Obviously $y(t) = 2 + o(\mathrm{e}^{\gamma t})$.

Let $\hat{V}_T$ and $\hat{I}_T$ be defined by (4.21) and (4.22) respectively after replacing $X(t)$ by $\hat{X}(t) = 2\int_0^t (t-s)\,\mathrm{d}W(s)$ and $Y(t)$ by $2W(t)$. We have $V_T - \hat{V}_T \xrightarrow{P} 0$ and $I_T - \hat{I}_T \xrightarrow{P} 0$ by Lemmas 4.1 and 4.2. Now it remains to note that $\hat{X}(t) = 2\int_0^t W(s)\,\mathrm{d}s$ by Itô's formula and

$$(\hat{V}_T, \hat{I}_T) \stackrel{d}{=} (V_\infty, I_\infty), \qquad \text{for all } T > 0,$$

in view of the self-similarity of the Wiener process. □

To prove the remaining propositions we need an additional result. In the next lemma and corollary, for each integer n we consider a d-dimensional process $M^n = (M_t^n)_{t\in[0,1]}$ on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t^n)_{t\in[0,1]}, P)$, whose components $M^{n,i}$ are continuous local martingales; $M_0^n = 0$. We also consider a d-dimensional process $M = (M_t)_{t\in[0,1]}$ with the same properties, on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, P)$. We denote by N^n the $\mathbb{R}^d \otimes \mathbb{R}^d$-valued process whose components $N^{n,ij}$ are defined as stochastic integrals $N_t^{n,ij} = \int_0^t M_s^{n,i}\,\mathrm{d}M_s^{n,j}$, $t \in [0, 1]$, and we associate the process N with M similarly. For the notion of stable convergence, we refer to Jacod and Shiryaev (1987, Chapter VIII, §5c).

Lemma 4.11. *Assume that*:

(i) $M^n \xrightarrow{\mathcal{L}} M$;

(ii) *for every finite subdivision* $\tau = \{0 = t_0 < t_1 < \cdots < t_m = 1\}$ *of* $[0, 1]$, *the vectors* $(\xi^n, M_{t_1}^n, \ldots, M_{t_m}^n)$ *converge* $\mathcal{G}$*-stably to the vector* $(\xi, M_{t_1}, \ldots, M_{t_m})$, *where* $\mathcal{G}$ *is a sub-*

σ-algebra of $\mathscr{F}$, and ξ^n and ξ are random variables.

Then, for every finite subdivision τ of [0, 1]*, the vectors* $(\xi^n, M^n_{t_1}, \ldots, M^n_{t_m}, N^n_{t_1}, \ldots, N^n_{t_m})$ *converge $\mathscr{G}$-stably to the vector* $(\xi, M_{t_1}, \ldots, M_{t_m}, N_{t_1}, \ldots, N_{t_m})$.

Proof. The proof is an easy consequence of the following fact. Let $\nu = \{0 = u_0 < u_1 < \cdots < u_k = 1\}$ be a subdivision of [0, 1]. Put

$$S^{n,ij}_t(\nu) = \sum_{p=0}^{k-1} M^{n,i}_{u_p}(M^{n,j}_{t\wedge u_{p+1}} - M^{n,j}_{t\wedge u_p}).$$

Then, for any $\varepsilon > 0$, there exists a $\delta > 0$ such that for all divisions ν of [0, 1] satisfying $|\nu| = \sup_{1\leqslant p\leqslant k}(u_p - u_{p-1}) \leqslant \delta$, we have

$$\sup_n P\left(\sup_{t\in[0,1]} |S^n_t(\nu) - N^n_t| \geqslant \varepsilon\right) \leqslant \varepsilon.$$

This can be shown from (i) following the lines of the proof of Lemma VI.6.13 in Jacod and Shiryaev (1987); moreover, the proof is much simpler in our case since $M^{n,j}$ are assumed to be continuous local martingales. □

Corollary 4.12. *Let the assumptions of Lemma 4.11 hold. Denote by* $[N^n, N^n]$ *the* $(\mathbb{R}^d \otimes \mathbb{R}^d) \otimes (\mathbb{R}^d \otimes \mathbb{R}^d)$*-valued process whose components are the quadratic covariations* $[N^{n,ij}, N^{n,kl}]$; $[N, N]$ *is defined similarly. Then the vectors* $(\xi^n, N^n_1, [N^n, N^n]_1)$ *converge $\mathscr{G}$-stably to the vector* $(\xi, N_1, [N, N]_1)$.

Proof. Note that $[N^{n,ij}, N^{n,kl}]_t = N^{n,ij}_t N^{n,kl}_t - \int_0^t N^{n,ij}_s \, dN^{n,kl}_s - \int_0^t N^{n,kl}_s \, dN^{n,ij}_s$ by Itô's formula, so the claim follows from Lemma 4.11 applied to the processes N^n. □

Remark. If $\xi^n \equiv \xi$ and $\mathcal{G} = \{\varnothing, \Omega\}$, the assertions of Lemma 4.11 and Corollary 4.12 are very special cases of theorems on convergence of stochastic integrals; see Jakubowski *et al.* (1989) and Kurtz and Protter (1991), cf. also Jacod and Shiryaev (1987, Theorem VI.6.1).

Proof of Proposition 2.9. According to (4.1)–(4.3),

$$V_T = \left(T^{-1}\int_0^T X(t)\, dW(t),\ T^{-1}\int_0^T X(t-1)\, dW(t)\right)^*$$

and

$$I_T = \begin{pmatrix} T^{-2}\int_0^T X^2(t)\, dt & T^{-2}\int_0^T X(t)X(t-1)\, dt \\ T^{-2}\int_0^T X(t)X(t-1)\, dt & T^{-2}\int_0^T X^2(t-1)\, dt \end{pmatrix}.$$

Because of (1.2) the process $(X(t),\ t \geqslant 0)$ has representation (4.4) with the function $x_0(t)$, $t \geqslant 0$, and $(X(t-1),\ t \geqslant 0)$ has this representation with the function $y(t) = x_0(t-1)$, $t \geqslant 0$.

By Lemma 1.1 we have

$$x_0(t) = A_0 \cos(\xi_0 t) + B_0 \sin(\xi_0 t) + o(\mathrm{e}^{\gamma t}) \tag{4.23}$$

and

$$x_0(t-1) = A_2 \cos(\xi_0 t) + B_2 \sin(\xi_0 t) + o(\mathrm{e}^{\gamma t})$$

for some $\gamma < 0$. We introduce $(X_1(t),\ t \geqslant 0)$ and $(X_2(t),\ t \geqslant 0)$ by

$$X_1(t) = \int_0^t \cos(\xi_0 s)\,\mathrm{d}W(s) \quad \text{and} \quad X_2(t) = \int_0^t \sin(\xi_0 s)\,\mathrm{d}W(s).$$

The solution $(X(t),\ t \geqslant 0)$ has representation

$$X(t) = \int_0^t x_0(t-s)\,\mathrm{d}W(s) + x_0(t)X_0(0) + b\int_{-1}^0 x_0(t-s-1)X_0(s)\,\mathrm{d}s, \qquad t \geqslant 0. \tag{4.24}$$

Inserting (4.23) into the first term, we obtain

$$X(t) = A_0 \cos(\theta_0 t)X_1(t) + A_0 \sin(\xi_0 t)X_2(t) + B_0 \sin(\xi_0 t)X_1(t) - B_0 \cos(\xi_0 t)X_2(t) + \overline{X}(t),$$

where $\overline{X}(t)$ is the sum of the last two terms in (4.24) and the contribution arising from the remainder term in (4.23); $T^{-2}\int_0^T \overline{X}^2(t)\,\mathrm{d}t \xrightarrow{P} 0$ by Lemmas 4.2 and 4.3.

Similarly, we obtain

$$X(t-1) = A_2 \cos(\xi_0 t)X_1(t) + A_2 \sin(\xi_0 t)X_2(t) + B_2 \sin(\xi_0 t)X_1(t) - B_2 \cos(\xi_0 t)X_2(t) + \overline{Y}(t),$$

where $T^{-2}\int_0^T \overline{Y}^2(t)\,\mathrm{d}t \xrightarrow{P} 0$.

Consider the following processes on the interval [0, 1]:

$$W^T(s) = T^{-1/2}W(Ts),$$

$$X_1^T(s) = T^{-1/2}X_1(Ts) = \int_0^s \cos(\xi_0 Tt)\,\mathrm{d}W^T(t),$$

$$X_2^T(s) = T^{-1/2}X_2(Ts) = \int_0^s \sin(\xi_0 Tt)\,\mathrm{d}W^T(t),$$

$$X^T(s) = A_0 \cos(\xi_0 Ts)X_1^T(s) + A_0 \sin(\xi_0 Ts)X_2^T(s) + B_0 \sin(\xi_0 Ts)X_1^T(s) - B_0 \cos(\xi_0 Ts)X_2^T(s),$$

$$Y^T(s) = A_2 \cos(\xi_0 Ts)X_1^T(s) + A_2 \sin(\xi_0 Ts)X_2^T(s) + B_2 \sin(\xi_0 Ts)X_1^T(s) - B_2 \cos(\xi_0 Ts)X_2^T(s).$$

Then

$$X(t) = T^{1/2}X^T(t/T) + \overline{X}(t), \qquad X(t-1) = T^{1/2}Y^T(t/T) + \overline{Y}(t),$$

and by Lemma 4.1 it is enough to check that

$$(\hat{V}_T,\ \hat{I}_T) \xrightarrow{d} (V_\infty,\ I_\infty), \tag{4.25}$$

where

$$\hat{V}_T = \left(\int_0^1 X^T(t)\,\mathrm{d}W^T(t),\ \int_0^1 Y^T(t)\,\mathrm{d}W^T(t) \right)^*$$

and

$$\hat{I}_T = \begin{pmatrix} \displaystyle\int_0^1 \{X^T(t)\}^2\,\mathrm{d}t & \displaystyle\int_0^1 X^T(t)Y^T(t)\,\mathrm{d}t \\ \displaystyle\int_0^1 X^T(t)Y^T(t)\,\mathrm{d}t & \displaystyle\int_0^T \{Y^T(t)\}^2\,\mathrm{d}t \end{pmatrix}.$$

But

$$\int_0^1 X^T\,\mathrm{d}W^T = A_0\int_0^1 X_1^T\,\mathrm{d}X_1^T + A_0\int_0^1 X_2^T\,\mathrm{d}X_2^T + B_0\int_0^1 X_1^T\,\mathrm{d}X_2^T - B_0\int_0^1 X_2^T\,\mathrm{d}X_1^T$$

and

$$\int_0^1 Y^T\,\mathrm{d}W^T = A_2\int_0^1 X_1^T\,\mathrm{d}X_1^T + A_2\int_0^1 X_2^T\,\mathrm{d}X_2^T + B_2\int_0^1 X_1^T\,\mathrm{d}X_2^T - B_2\int_0^1 X_2^T\,\mathrm{d}X_1^T$$

are represented as linear combinations of the stochastic integrals $\int X_i^T\,\mathrm{d}X_j^T$, $i, j = 1, 2$. Since

$$(X_1^T, X_2^T) \xrightarrow{\mathcal{L}} \frac{1}{\sqrt{2}}(\tilde{W}_1, \tilde{W}_2)$$

by the functional central limit theorem, the claim follows from Corollary 4.12. □

Proof of Proposition 2.10. According to (4.1)–(4.3),

$$V_T = \left(\mathrm{e}^{-v_0 T}\int_0^T X(t)\,\mathrm{d}W(t),\ T^{-1}\int_0^T Y(t)\,\mathrm{d}W(t) \right)^*$$

and

$$I_T = \begin{pmatrix} \mathrm{e}^{-2v_0 T}\displaystyle\int_0^T X^2(t)\,\mathrm{d}t & T^{-1}\,\mathrm{e}^{-v_0 T}\displaystyle\int_0^T X(t)Y(t)\,\mathrm{d}t \\ T^{-1}\,\mathrm{e}^{-v_0 T}\displaystyle\int_0^T X(t)Y(t)\,\mathrm{d}t & T^{-2}\displaystyle\int_0^T Y^2(t)\,\mathrm{d}t \end{pmatrix},$$

where

$$Y(t) = X(t) - \mathrm{e}^{v_0}X(t-1), \qquad t \geqslant 0,$$

and $Y(\cdot)$ has representation (4.4) with $y(t) = x_0(t) - \mathrm{e}^{v_0}x_0(t-1)$. It follows from (1.12) that

$$x_0(t) = \frac{1}{v_0 - a + 1}\,\mathrm{e}^{v_0 t} - \frac{1}{a-1} + o(\mathrm{e}^{\gamma t})$$

and

$$y(t) = \frac{e^{v_0} - 1}{a - 1} + o(e^{\gamma t})$$

for some $\gamma < 0$. Due to Lemmas 4.5, 4.6, 4.7 and 4.1, this implies

$$e^{-2v_0 T} \int_0^T X^2(t)\,dt \xrightarrow{P} \frac{U_0^2}{2v_0(v_0 - a + 1)^2}, \tag{4.26}$$

$$T^{-1/2} e^{-v_0 T} \int_0^T |X(t)|\,dt \xrightarrow{P} 0, \tag{4.27}$$

and

$$T^{-1} e^{-v_0 T} \int_0^T X(t)Y(t)\,dt \xrightarrow{P} 0. \tag{4.28}$$

Introduce the following processes on the interval [0, 1]:

$$W^T(s) = T^{-1/2} W(Ts),$$

$$X^T(s) = e^{-v_0 T} \int_0^{Ts} X(t)\,dW(t) = T^{1/2} e^{-v_0 T} \int_0^s X(Tt)\,dW^T(t),$$

$$Y^T(s) = T^{-1} \int_0^{Ts} W(t)\,dW(t) = \int_0^s W^T(t)\,dW^T(t),$$

which are continuous local martingales with respect to the filtration $\mathcal{F}_s^T = \sigma\{X_0(t),\ t \in [-1, 0];\ W(t),\ t \in [0, Ts]\}$. Let $\tau = \{0 = t_0 < t_1 < \cdots < t_m = 1\}$ be a subdivision of [0, 1]. It follows from (4.26) and (4.27) that

$$T e^{-2v_0 T} \int_0^1 X^2(Tt)\,dt \xrightarrow{P} \frac{U_0^2}{2v_0(v_0 - a + 1)^2}$$

and

$$T^{1/2} e^{-v_0 T} \int_0^s X(Tt)\,dt \xrightarrow{P} 0, \qquad s \in [0, 1];$$

therefore, we can apply the stable limit theorem for martingales (Jacod and Shiryaev 1987, Theorem VIII.5.42; or Touati 1991, Theorem 1) to the process $X^T(\cdot)$ and to the stopped processes $W^T(t_1 \wedge \cdot), \ldots, W^T(t_m \wedge \cdot)$, which yields that the vectors $(X^T(1), W^T(t_1), \ldots, W^T(t_m))$ converge $\mathscr{F}$-stably (where $\mathcal{F} = \sigma\{X_0(t),\ t \in [-1, 0];\ W(t),\ t \geqslant 0\}$) to the vector

$$\left(\frac{U_0 Z}{\sqrt{2v_0}(v_0 - a + 1)}, \tilde{W}(t_1), \ldots, \tilde{W}(t_m) \right)$$

as $T \to \infty$. Clearly, $W^T \xrightarrow{\mathcal{L}} \tilde{W}$. Applying Corollary 4.12, we obtain the $\mathscr{F}$-stable convergence of the vector

$$\left(e^{-v_0 T} \int_0^T X(t)\,dW(t),\ T^{-1} \int_0^T W(t)\,dW(t),\ T^{-2} \int_0^T W^2(t)\,dt \right)$$

to the vector

$$\left(\frac{U_0 Z}{\sqrt{2v_0}(v_0 - a + 1)}, \int_0^1 \tilde{W}(t)\,\mathrm{d}\tilde{W}(t), \int_0^1 \tilde{W}^2(t)\,\mathrm{d}t\right).$$

By Lemmas 4.1 and 4.2, we have the $\mathscr{F}$-stable convergence of the vector

$$\left(\mathrm{e}^{-v_0 T}\int_0^T X(t)\,\mathrm{d}W(t),\ T^{-1}\int_0^T Y(t)\,\mathrm{d}W(t),\ T^{-2}\int_0^T Y^2(t)\,\mathrm{d}t\right)$$

to the vector

$$\left(\frac{U_0 Z}{\sqrt{2v_0}(v_0 - a + 1)}, \frac{\mathrm{e}^{v_0} - 1}{a-1}\int_0^1 \tilde{W}(t)\,\mathrm{d}\tilde{W}(t), \frac{(\mathrm{e}^{v_0} - 1)^2}{(a-1)^2}\int_0^1 \tilde{W}^2(t)\,\mathrm{d}t\right).$$

Now the convergence $(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty)$ follows from the properties of the stable convergence and relations (4.26) and (4.28). □

Proof of Proposition 2.11. The proof follows the same lines as that of Proposition 2.10. Here we have

$$x_0(t) = \frac{1}{v_0 - a + 1}\,\mathrm{e}^{v_0 t} + A_1\cos(\xi_1 t) + B_1 \sin(\xi_1 t) + o(\mathrm{e}^{\gamma t})$$

and

$$y(t) = A\cos(\xi_1 t) + B\sin(\xi_1 t) + o(\mathrm{e}^{\gamma t})$$

for some $\gamma < 0$; in particuar, (4.26), (4.27) and (4.28) are still true.

Introduce the processes $W^T(s)$ and $X^T(s)$, $s \in [0, 1]$, as in the proof of Proposition 2.10 and the processes

$$X_1^T(s) = \int_0^s \cos(\xi_1 Tt)\,\mathrm{d}W^T(t), \qquad X_2^T(t) = \int_0^s \sin(\xi_1 Tt)\,\mathrm{d}W^T(t),$$

$$Y^T(s) = A\cos(\xi_1 Ts)X_1^T(s) + A\sin(\xi_1 Ts)X_2^T(s) + B\sin(\xi_1 Ts)X_1^T(s) - B\cos(\xi_1 Ts)X_2^T(s).$$

Note that

$$\int_0^1 Y^T\,\mathrm{d}W^T = A\int_0^1 X_1^T\,\mathrm{d}X_1^T + A\int_0^1 X_2^T\,\mathrm{d}X_2^T + B\int_0^1 X_1^T\,\mathrm{d}X_2^T - B\int_0^1 X_2^T\,\mathrm{d}X_1^T \tag{4.29}$$

and

$$(X_1^T, X_2^T) \xrightarrow{\mathcal{L}} \frac{1}{\sqrt{2}}(\tilde{W}_1, \tilde{W}_2). \tag{4.30}$$

In view of (4.26) and (4.27), we have

$$T\,\mathrm{e}^{-2v_0 T}\int_0^1 X^2(Tt)\,\mathrm{d}t \xrightarrow{P} \frac{U_0^2}{2v_0(v_0 - a + 1)^2},$$

$$T^{1/2}\,\mathrm{e}^{-v_0 T}\int_0^s X(Tt)\cos(\xi_1 Tt)\,\mathrm{d}t \xrightarrow{P} 0,$$

$$T^{1/2}\,\mathrm{e}^{-v_0 T}\int_0^s X(Tt)\sin(\xi_1 Tt)\,\mathrm{d}t \xrightarrow{P} 0,$$

where $s \in [0, 1]$. Let us again apply the stable limit theorem for martingales but now to the process $X^T(\cdot)$ and to the stopped processes $X_1^T(t_1 \wedge \cdot), \ldots, X_1^T(t_m \wedge \cdot), X_2^T(t_1 \wedge \cdot), \ldots, X_2^T(t_m \wedge \cdot)$, where $\tau = \{0 = t_0 < t_1 < \cdots < t_m = 1\}$ is a subdivision of [0, 1], which yields the $\mathscr{F}$-stable convergence of the vectors

$$(X^T(1), X_1^T(t_1), \ldots, X_1^T(t_m), X_2^T(t_1), \ldots, X_2^T(t_m))$$

to the vector

$$\left(\frac{U_0 Z}{\sqrt{2v_0}(v_0 - a + 1)}, \frac{1}{\sqrt{2}}\tilde{W}_1(t_1), \ldots, \frac{1}{\sqrt{2}}\tilde{W}_1(t_m), \frac{1}{\sqrt{2}}\tilde{W}_2(t_1), \ldots, \frac{1}{\sqrt{2}}\tilde{W}_2(t_m)\right).$$

In view of (4.29) and (4.30), applying Corollary 4.12, we obtain the $\mathscr{F}$-stable convergence of the vector

$$\left(X^T(1), \int_0^1 Y^T(t)\,\mathrm{d}W^T(t), \int_0^1 \{Y^T(t)\}^2\,\mathrm{d}t\right)$$

to the vector

$$\left(\frac{U_0 Z}{\sqrt{2v_0}(v_0 - a + 1)}, \frac{1}{2}\left(A\int_0^1 \tilde{W}_1\,\mathrm{d}\tilde{W}_1 + A\int_0^1 \tilde{W}_2\,\mathrm{d}\tilde{W}_2 + B\int_0^1 \tilde{W}_1\,\mathrm{d}\tilde{W}_2 - B\int_0^1 \tilde{W}_2\,\mathrm{d}\tilde{W}_1\right),\right.$$
$$\left.\frac{1}{4}(A^2 + B^2)\int_0^1 (\tilde{W}_1^2 + \tilde{W}_2^2)\,\mathrm{d}t\right).$$

But $X^T(1) = \mathrm{e}^{-v_0 T}\int_0^T X(t)\,\mathrm{d}W(t)$ by the definition of X^T, $Y(t) = T^{1/2}Y^T(t/T) + \overline{Y}(t)$, where $T^{-2}\int_0^T \overline{Y}^2(t)\,\mathrm{d}t \xrightarrow{P} 0$ as in the proof of Proposition 2.9, hence

$$T^{-1}\int_0^T Y(t)\,\mathrm{d}W(t) - \int_0^1 Y^T(t)\,\mathrm{d}W^T(t) = T^{-1}\int_0^T \overline{Y}(t)\,\mathrm{d}W(t) \xrightarrow{P} 0$$

and, similarly,

$$T^{-2}\int_0^T Y^2(t)\,\mathrm{d}t - \int_0^1 \{Y^T(t)\}^2\,\mathrm{d}t \xrightarrow{P} 0.$$

So we have the $\mathscr{F}$-stable convergence of the vector

$$\left(\mathrm{e}^{-v_0 T}\int_0^T X(t)\,\mathrm{d}W(t), T^{-1}\int_0^T Y(t)\,\mathrm{d}W(t), T^{-2}\int_0^T Y^2(t)\,\mathrm{d}t\right)$$

to the vector

$$\left(\frac{U_0 Z}{\sqrt{2v_0}(v_0 - a + 1)}, \frac{1}{2}\left(A\int_0^1 \tilde{W}_1\,\mathrm{d}\tilde{W}_1 + A\int_0^1 \tilde{W}_2\,\mathrm{d}\tilde{W}_2 + B\int_0^1 \tilde{W}_1\,\mathrm{d}\tilde{W}_2 - B\int_0^1 \tilde{W}_2\,\mathrm{d}\tilde{W}_1\right),\right.$$
$$\left.\frac{1}{4}(A^2 + B^2)\int_0^1 (\tilde{W}_1^2 + \tilde{W}_2^2)\,\mathrm{d}t\right),$$

and we finish the proof as in the previous proposition. □

5. Appendix

In this section we present the proof of Lemma 1.1. We took the idea from Myschkis (1972); see also Hale and Verduyn Lunel (1993).

Proof of Lemma 1.1. Equation (1.3) is equivalent to

$$x_0(t) = 1 + a\int_0^t x_0(s)\,\mathrm{d}s + b\int_0^t x_0(s-1)\,\mathrm{d}s, \qquad t \geqslant 0.$$

Thus we have the inequality

$$|x_0(t)| \leqslant 1 + (|a| + |b|)\int_0^t |x_0(s)|\,\mathrm{d}s, \qquad t \geqslant 0.$$

From a Gronwall-type lemma (Liptser and Shiryaev 1977, Lemma 4.13) it follows that

$$|x_0(t)| \leqslant \mathrm{e}^{ct}, \qquad t \geqslant 0,$$

with $c = |a| + |b|$. Thus the Laplace transform

$$\hat{x}_0(\lambda) = \int_0^\infty \mathrm{e}^{-\lambda t} x_0(t)\,\mathrm{d}t$$

exists at least for all λ with $\mathrm{Re}\,\lambda > c$ and can be calculated from (1.3) as

$$\hat{x}_0(\lambda) = h^{-1}(\lambda), \quad \mathrm{Re}\,\lambda > c, \qquad \text{where } h(\lambda) = \lambda - a - b\,\mathrm{e}^{-\lambda}. \tag{5.1}$$

The inversion formula yields, for every $v > c$,

$$x_0(t) = \lim_{w\to\infty} \frac{1}{2\pi\mathrm{i}}\int_{v-\mathrm{i}w}^{v+\mathrm{i}w} \mathrm{e}^{\mu t}\hat{x}_0(\mu)\,\mathrm{d}\mu, \qquad t \geqslant 0.$$

If $\lambda \in \Lambda$, then $|\lambda| \leqslant |a| + |b|\mathrm{e}^{-\mathrm{Re}\,\lambda}$. This implies $|v_0| \leqslant |a| + |b|\mathrm{e}^{-v_0}$ and, consequently, $v_0 \leqslant |a| + |b| = c$. Now choose a real $u < v_0$ and fix a $u_0 < u$ such that $\mathrm{Re}\,\lambda \notin [u_0, u)$ for every $\lambda \in \Lambda$. Then, by using Cauchy's residue theorem, we obtain

$$x_0(t) = \sum_{\lambda\in\Lambda:\mathrm{Re}\,\lambda \geqslant u} \mathrm{Res}\,\chi_t(\lambda) + \lim_{w\to\infty} \frac{1}{2\pi\mathrm{i}}\int_{u_0-\mathrm{i}w}^{u_0+\mathrm{i}w} \chi_t(\mu)\,\mathrm{d}\mu, \qquad t \geqslant 0, \tag{5.2}$$

where $\chi_t(\lambda) = \mathrm{e}^{\lambda t}h^{-1}(\lambda)$, $t \geqslant 0$, $\lambda \notin \Lambda$. Here we have used the fact that $|\chi_t(\lambda)|$ tends to zero uniformly on $\lambda \in [u_0 + \mathrm{i}w, v + \mathrm{i}w]$ and on $\lambda \in [u_0 - \mathrm{i}w, v - \mathrm{i}w]$ if $|w| \to \infty$.

Now observe that either $v_0 \in \Lambda$ (if $b \geqslant v(a)$) or $\lambda_0 = v_0 + i\xi_0 \in \Lambda$ for some $\xi_0 > 0$ (if $b < v(a)$). The explicit calculation of the residues in v_0 in the first case and in $\overline{\lambda_0}$ and λ_0 in the second case yields the form $\psi_0(\cdot)$ given in Lemma 1.1. The limit in (5.2) can be estimated by $K\,e^{u_0 t}$ for some $K > 0$, thus it is $o(e^{\gamma t})$ for some $\gamma < u < v_0$.

Acknowledgements

The authors acknowledge their gratitude to the Deutsche Forschungsgemeinschaft, to the Sonderforschungsbereich 373 at the Humboldt University in Berlin, and to the EU Human Capital and Mobility Programme for supporting our collaboration on the topic of this paper. They also express their thanks to D. Kramkov and V. Spokoiny for helpful discussions and to anonymous referees for interesting comments.

References

Basawa, I.V. and Prakasa Rao, B.L.S. (1980) *Statistical Inference for Stochastic Processes*. London: Academic Press.

Dietz, H.M. (1992) A non-Markovian relative to the Ornstein–Uhlenbeck process and some of its local statistical properties. *Scand. J. Statist.*, **19**, 363–379.

Greenwood, P.E. and Wefelmeyer, W. (1993) Asymptotic minimax results for stochastic process families with critical points. *Stochastic Process. Appl.*, **44**, 107–116.

Gushchin, A.A. (1995) On asymptotic optimality of estimators of parameters under the LAQ condition. *Theory Probab. Appl.*, **40**, 261–272.

Hale, J.K. and Verduyn Lunel, S.M. (1993) *Introduction to Functional Differential Equations*. New York: Springer-Verlag.

Hayes, N.D. (1950) Roots of the transcendental equation associated with a certain differential-difference equation. *J. London Math. Soc.*, **25**, 226–232.

Jacod, J. and Shiryaev, A.N. (1987) *Limit Theorems for Stochastic Processes*. Berlin: Springer-Verlag.

Jakubowski, A., Mémin, J. and Pages G. (1989) Convergence en loi des suites d'intégrales stochastiques sur l'espace $\mathbb{D}^1$ de Skorokhod. *Probab. Theory Related Fields*, **81**, 111–137.

Jeganathan, P. (1995) Some aspects of asymptotic theory with applications to time series models. *Econometric Theory*, **11**, 818–887.

Küchler, U. and Kutoyants, Yu.A. (1996) Delay estimation for stationary diffusion-type process. Preprint.

Küchler, U. and Mensch, B. (1992). Langevin's stochastic differential equations extended by a time-delayed term. *Stochastics Stochastics Rep.*, **40**, 23–42.

Küchler, U. and Sørensen, M. (1989) Exponential families of stochastic processes: A unifying semimartingale approach. *Internat. Statist. Rev.*, **57**, 123–144.

Kurtz, T.G. and Protter, Ph. (1991) Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.*, **19**, 1035–1070.

Le Cam, L. and Lo Yang, G. (1990) *Asymptotics in Statistics: Some Basic Concepts*. New York: Springer-Verlag.

Liptser, R.S. and Shiryaev, A.N. (1977) *Statistics of Random Processes*, Vol. 1. New York: Springer-Verlag.

Liptser, R.S. and Shiryaev, A.N. (1989) *Theory of Martingales*. Dordrecht: Kluwer.

Mohammed, S.-E.A. and Scheutzow, M.K.R. (1990) Lyapunov exponents and stationary solutions for affine stochastic delay equations. *Stochastics Stochastics Rep.*, **29**, 259–283.

Myschkis, A.D. (1972) *Linear Differential Equations with Delayed Argument*. Moscow: Nauka (in Russian).

Shiryaev, A.N. and Spokoiny, V.G. (1999) *Statistical Experiments and Decisions: Asymptotic Theory.* Singapore: World Scientific.

Touati, A. (1991) On the functional convergence in distribution of sequences of semimartingales to a mixture of Brownian motions. *Theory Probab. Appl.*, **36**, 752–771.

Received June 1997 and revised March 1998