October 1, 2013

SECOND-ORDER CONTINUOUS-TIME NON-STATIONARY GAUSSIAN AUTOREGRESSION

N. LIN AND S. V. LOTOTSKY

ABSTRACT. The objective of the paper is to identify and investigate all possible types of asymptotic behavior for the maximum likelihood estimators of the unknown parameters in the second-order linear stochastic ordinary differential equation driven by Gaussian white noise. The emphasis is on the non-ergodic case, when the roots of the corresponding characteristic equation are not both in the left half-plane.

1. INTRODUCTION

Many problems, in particular, in mechanics and electric circuit theory, lead to a second-order ordinary differential equation with constant coefficients

$$(1.1) \qquad \ddot{X}(t) = \theta_1 \dot{X}(t) + \theta_2 X(t) + f(t),$$

with $X(t)$ representing, for example, the position of the moving object or the value of the electric current and $f(t)$ representing the external force. Here and below, $\dot{X}(t) = dX(t)/dt$ and $\ddot{X}(t) = d^2 X(t)/dt^2$. When the external force is Gaussian white noise, (1.1) becomes the stochastic ordinary differential equation

$$(1.2) \qquad \ddot{X}(t) = \theta_1 \dot{X}(t) + \theta_2 X(t) + \sigma \dot{W}(t), \ t > 0,$$

in which $W = W(t)$ is standard Brownian motion; we assume that the initial conditions $X(0)$ and $\dot{X}(0)$ are non-random. A natural statistical question is estimation of the real-valued parameters θ_1, θ_2, and σ from the continuous time observations $(X(t), \dot{X}(t), 0 < t < T)$ of the solution of (1.2). With mechanical or electrical applications in mind, observation of $(X(t), \dot{X}(t))$ corresponds to measuring the position and velocity of the object or measuring the current and voltage in the electric circuit. Since $\sigma^2 t$ is the quadratic variation of $\dot{X}(t)$ and therefore can be, in principle, determined exactly from the continuous time observations, we will concentrate on estimating θ_1 and θ_2.

Equation (1.2) is often referred to as a continuous time auto-regression of second order, or CAR(2), being a particular case of CAR(N)

$$(1.3) \qquad X^{(N)} = \sum_{k=0}^{N-1} \theta_{N-k} X^{(k)} + \sigma \dot{W};$$

see [4]. A rigorous interpretation of equation (1.2) is the system

$$(1.4) \qquad dX(t) = \dot{X}(t)dt, \ d\dot{X}(t) = \big(\theta_2 X(t) + \theta_1 \dot{X}(t)\big)dt + \sigma dW(t).$$

In the matrix-vector form, system (1.4) becomes a particular case of the multi-dimensional Ornstein-Uhlenbeck process

$$(1.5) \qquad d\boldsymbol{X}(t) = \Theta \boldsymbol{X}(t)dt + \boldsymbol{\sigma} dW(t),$$

with

$$\boldsymbol{X}(t) = \left(\begin{array}{c} X(t) \\ \dot{X}(t) \end{array} \right), \ \Theta = \left(\begin{array}{cc} 0 & 1 \\ \theta_2 & \theta_1 \end{array} \right), \ \boldsymbol{\sigma} = \left(\begin{array}{c} 0 \\ \sigma \end{array} \right).$$

The following estimator of Θ is studied in [2] using the continuous time observations of the vector process $\boldsymbol{X}$:

$$(1.6) \qquad \widehat{\Theta}_T = \left(\int_0^T \left(d\boldsymbol{X}(t)\boldsymbol{X}^\top(t) \right) \right) \left(\int_0^T \boldsymbol{X}(t)\boldsymbol{X}^\top(t)dt \right)^{-1}.$$

The estimator is strongly consistent as $T \to \infty$, and this implies strong consistency of the maximum likelihood estimators for θ_1 and θ_2 in (1.2) for all $(\theta_1, \theta_2) \in \mathbb{R}^2$; see [2] or Theorem 3.1 below.

When the process $\boldsymbol{X}$ defined by (1.5) is ergodic (equivalently, when all eigenvalues of the matrix Θ are in the left half-plane), it is known [1, Theorem 4.6.2] that the estimator $\widehat{\Theta}_T$ is asymptotically normal with rate $T^{1/2}$. Some non-ergodic models of the type (1.5) have also been studied (see, for example, [17]). The key object in the analysis of a maximum likelihood estimator (MLE) is the normalized log-likelihood ratio. A desirable property of this normalized log-likelihood ratio is local asymptotic normality (LAN), as it implies certain efficiency of the MLE; see [8, Chapter II]. While the LAN property is typically associated with ergodic models, it can also hold in some non-ergodic models [10]. Still, in many non-ergodic models, LAN is replaced with LAMN (local asymptotic mixed normality), leading to a different kind of efficiency of the MLE; see [13, Chapter 5].

So far, analysis of (1.5) in general and (1.2) in particular was aimed at identifying the particular cases that fit in either LAN or LAMN framework. It turns out that many of the non-ergodic regimes of (1.2) lead to new asymptotic forms of the normalized log-likelihood ratio and to new types of asymptotic behavior of the MLE. In the current paper, we present a complete asymptotic analysis of the MLE and the likelihood ratio for all possible values of the unknown parameters $(\theta_1, \theta_2) \in \mathbb{R}^2$.

The paper is organized as follows. Section 3 states the main results and discusses the results in the broader context of statistical estimation. After some preliminary work in Section 4, the proofs are in Section 5, followed by a brief summary in Section 6. Section 6 also discusses some questions related to the discrete-time analogues of (1.3).

Throughout the paper, we fix a stochastic basis $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ with a standard Brownian motion $W = W(t)$. Other common notations are $\mathbb{E}$ for the expectation with respect to $\mathbb{P}$, $\sqrt{-1}$ for the imaginary unit, $^\top$ for the transpose of a vector or a matrix, $\dot{g}(t)$, $\ddot{g}(t)$ for the first and second time derivatives of the function g, $\stackrel{d}{=}$ for equality in distribution of random variables, $\stackrel{\mathcal{L}}{=}$ for equality in law of random processes, and $0_{\text{a.s.}}(T)$ to denote a process converging to zero with probability one as $T \to \infty$.

2. Summary of the main results

To write the likelihood ratio for equation (1.2), define the vectors

$$
(2.1) \qquad \boldsymbol{\theta} = \begin{pmatrix} \theta_2 \\ \theta_1 \end{pmatrix}, \quad \boldsymbol{X}(t) = \begin{pmatrix} X(t) \\ \dot{X}(t) \end{pmatrix}.
$$

Then equation (1.2) becomes

$$
(2.2) \qquad d\dot{X} = \boldsymbol{\theta}^\top \boldsymbol{X} \, dt + \sigma dW(t).
$$

Since $X(t) = X(0) + \int_0^t \dot{X}(s) ds$, it follows from (2.2) that

(1) $\dot{X} = \dot{X}(t)$ is a diffusion-type process in the sense of Liptser and Shiryaev ([16, Definition 4.2.7]) and generates the measure $P_T^{\boldsymbol{\theta}}$ on the space of continuous functions on $[0, T]$;

(2) For every $\boldsymbol{\theta} \in \mathbb{R}^2$, the measure $P_T^{\boldsymbol{\theta}}$ is absolutely continuous with respect to the measure $P_T^{\boldsymbol{0}}$ generated by the process $\dot{X}(0) + \sigma W(t)$, $0 \le t \le T$, and

$$
(2.3) \qquad \frac{dP_T^{\boldsymbol{\theta}}}{dP_T^{\boldsymbol{0}}}(\dot{X})
$$
$$
= \exp \left(\frac{1}{\sigma^2} \int_0^T (\theta_2 X(t) + \theta_1 \dot{X}(t)) d\dot{X}(t) - \frac{1}{2\sigma^2} \int_0^T (\theta_2 X(t) + \theta_1 \dot{X}(t))^2 dt \right);
$$

see [16, Theorem 7.6].

The expressions for the maximum likelihood estimators $\widehat{\theta}_{1,T}$ of θ_1 and $\widehat{\theta}_{2,T}$ of θ_2, using continuous-time observations of $\boldsymbol{X}(t) = \big(X(t), \dot{X}(t)\big)^\top$, $0 \le t \le T$, now follow from (2.3):

$$
(2.4) \qquad
\begin{aligned}
\widehat{\theta}_{1,T} &= \frac{\int_0^T X^2(t) dt \int_0^T \dot{X}(t) d\dot{X}(t) - \int_0^T X(t) \dot{X}(t) dt \int_0^T X(t) d\dot{X}(t)}{\int_0^T X^2(t) dt \int_0^T \dot{X}^2(t) dt - \left(\int_0^T X(t) \dot{X}(t) dt \right)^2}, \\
\widehat{\theta}_{2,T} &= \frac{\int_0^T \dot{X}^2(t) dt \int_0^T X(t) d\dot{X}(t) - \int_0^T X(t) \dot{X}(t) dt \int_0^T \dot{X}(t) d\dot{X}(t)}{\int_0^T X^2(t) dt \int_0^T \dot{X}^2(t) dt - \left(\int_0^T X(t) \dot{X}(t) dt \right)^2}.
\end{aligned}
$$

Similar to (2.1) we write

$$
\widehat{\boldsymbol{\theta}}_T = \begin{pmatrix} \widehat{\theta}_{2,T} \\ \widehat{\theta}_{1,T} \end{pmatrix}.
$$

The amount of integration in (2.4) can be reduced using rules of the usual and stochastic calculus and keeping in mind that the processes X is continuously differentiable and the process $\dot{X}$ is a continuous semi-martingale with quadratic variation equal to

$\sigma^2 t$:

$$\int_0^T X(t)\dot{X}(t)dt = \int_0^T X(t)dX(t) = \frac{X^2(T) - X^2(0)}{2},$$

$$\int_0^T \dot{X}(t)d\dot{X}(t) = \frac{\dot{X}^2(T) - \sigma^2 T - \dot{X}^2(0)}{2},$$

$$\int_0^T X(t)d\dot{X}(t) = X(T)\dot{X}(T) - X(0)\dot{X}(0) - \int_0^T \dot{X}^2(t)dt.$$

With the continuous time observations, the value of σ can be assumed known because the quadratic variation process of $\dot{X}$ at time t, an observable quantity, is $\sigma^2 t$. Note also that σ does not appear in the formulas (2.4).

The measure $P_T^{\mathbf{0}}$ in (2.3) corresponds to the solution of (2.2) with $\theta_1 = \theta_2 = 0$. Sometimes it is more convenient to work with the reference measure coming from some other solution of (1.2), for example, the actual observations. Accordingly, let us fix $\boldsymbol{\theta} \in \mathbb{R}^2$ and the corresponding observation process $\boldsymbol{X} = \boldsymbol{X}(t)$, $0 \le t \le T$, satisfying (1.4). Then, for every $\boldsymbol{\vartheta} \in \mathbb{R}^2$,

$$(2.5) \qquad \begin{aligned} \ln \frac{dP_T^{\boldsymbol{\vartheta}}}{dP_T^{\boldsymbol{\theta}}}(\dot{X}) &= \frac{1}{\sigma^2} \int_0^T \left(\boldsymbol{\vartheta}^\top \boldsymbol{X}(t) - \boldsymbol{\theta}^\top \boldsymbol{X}(t) \right) d\dot{X}(t) \\ &\quad - \frac{1}{2\sigma^2} \int_0^T \left(\left(\boldsymbol{\vartheta}^\top \boldsymbol{X}(t) \right)^2 - \left(\boldsymbol{\theta}^\top \boldsymbol{X}(t) \right)^2 \right) dt; \end{aligned}$$

see [16, Theorem 7.19]. With $\boldsymbol{\theta}$ and $\boldsymbol{X}$ fixed, we write

$$L_T(\boldsymbol{\vartheta}) = \ln \frac{dP_T^{\boldsymbol{\vartheta}}}{dP_T^{\boldsymbol{\theta}}}(\dot{X}).$$

Define the matrix

$$\Psi_T = \int_0^T \boldsymbol{X}(t)\boldsymbol{X}^\top(t)dt = \begin{pmatrix} \int_0^T X^2(t)dt & \int_0^T X(t)\dot{X}(t)dt \\ \int_0^T X(t)\dot{X}(t)dt & \int_0^T \dot{X}^2(t)dt \end{pmatrix}.$$

Then (2.2) and (2.5) imply

$$L_T(\boldsymbol{\vartheta}) = \frac{1}{\sigma} \int_0^T (\boldsymbol{\vartheta} - \boldsymbol{\theta})^\top \boldsymbol{X}(t)dW(t) - \frac{1}{2\sigma^2} \int_0^T (\boldsymbol{\vartheta} - \boldsymbol{\theta})^\top \boldsymbol{X}(t)\boldsymbol{X}^\top(t)(\boldsymbol{\vartheta} - \boldsymbol{\theta})dt$$

$$(2.6) \qquad = \frac{1}{\sigma} \int_0^T (\boldsymbol{\vartheta} - \boldsymbol{\theta})^\top \boldsymbol{X}(t)dW(t) - \frac{1}{2\sigma^2}(\boldsymbol{\vartheta} - \boldsymbol{\theta})^\top \Psi(T)(\boldsymbol{\vartheta} - \boldsymbol{\theta}).$$

We address the following questions about the estimators $\widehat{\theta}_{1,T}, \widehat{\theta}_{2,T}$:

(1) **rate of convergence**, that is, finding positive deterministic functions $v_i(T)$, $i = 1, 2$, such that, as $T \to \infty$, $v_i(T) \nearrow +\infty$ and $v_i(T)(\widehat{\theta}_{i,T} - \theta_i)$ converge in distribution to non-degenerate random variables, and identifying the corresponding limit distributions;

(2) existence of a **normal limit with a random rate** (NLRR), that is, finding a random matrix $R = R_T$ such that, as $T \to \infty$, $R_T(\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})$ converges in distribution to a Gaussian random vector. The hope is that at least one of the two things happens: (a) the random rate leads to a normal limit when a

deterministic rate does not; (b) neither the matrix R_T nor the parameters of the limit distribution depend explicitly on θ_1, θ_2.

(3) **local asymptotic structure** of the normalized log-likelihood ratio, that is, analyzing

$$(2.7) \qquad \ell_T(\boldsymbol{u}) = L_T(\boldsymbol{\theta} + A_T \boldsymbol{u}),$$

as $T \to \infty$, where $\boldsymbol{\theta}$ is fixed, $A_T \in \mathbb{R}^{2 \times 2}$ is a suitably chosen deterministic matrix with $\lim_{T \to \infty} |A_T| = 0$ (any matrix norm will work), and $\boldsymbol{u} \in \mathbb{R}^2$.

Let us recall some definitions related to the normalized log-likelihood ratio (2.7). It follows from (2.6) that

$$(2.8) \qquad \ell_T(\boldsymbol{u}) = \frac{1}{\sigma} \int_0^T \left(\boldsymbol{u}^\top A_T^\top \boldsymbol{X}(t) \right) dW(t) - \frac{1}{2\sigma^2} \boldsymbol{u}^\top A_T^\top \Psi_T A_T \boldsymbol{u}.$$

With a suitable choice of the matrix A_T, there exists a non-trivial limit in distribution

$$(2.9) \qquad \ell_\infty(\boldsymbol{u}) = \lim_{T \to \infty} \ell_T(\boldsymbol{u}).$$

Three particular cases of ℓ_T satisfying (2.9) are of special interest:

- `Local Asymptotic Normality` (LAN), if there exists a bivariate normal vector $\boldsymbol{\xi}$ with mean zero and non-degenerate covariance matrix $\Sigma_{\boldsymbol{\xi}}$ such that, for every $\boldsymbol{u} \in \mathbb{R}^2$,

$$\ell_\infty(\boldsymbol{u}) = \frac{1}{\sigma} \boldsymbol{u}^\top \boldsymbol{\xi} - \frac{1}{2\sigma^2} \boldsymbol{u}^\top \Sigma_{\boldsymbol{\xi}} \boldsymbol{u}.$$

- `Local Asymptotic Mixed Normality` (LAMN) if there exist a bivariate normal vector $\boldsymbol{\eta}$ with zero mean and unit covariance matrix, and a random symmetric positive definite matrix $B \in \mathbb{R}^{2 \times 2}$ such that B and $\boldsymbol{\eta}$ are independent and, for every $\boldsymbol{u} \in \mathbb{R}^2$,

$$(2.10) \qquad \ell_\infty(\boldsymbol{u}) = \frac{1}{\sigma} \boldsymbol{u}^\top B^{1/2} \boldsymbol{\eta} - \frac{1}{2\sigma^2} \boldsymbol{u}^\top B \boldsymbol{u}.$$

If (2.10) holds with a degenerate matrix B, we refer to ℓ_T as DLAMN (degenerate locally asymptotically mixed normal).

- `Local Asymptotic Brownian Functional` structure (LABF) if

$$(2.11) \qquad \ell_\infty(\boldsymbol{u}) = \frac{1}{\sigma} \int_0^1 \boldsymbol{u}^\top G(t) d\boldsymbol{w}(t) - \frac{1}{2\sigma^2} \int_0^1 \boldsymbol{u}^\top G(t) G^\top(t) \boldsymbol{u} \, dt,$$

where $G \in \mathbb{R}^{2 \times 2}$ is an adapted process, $\boldsymbol{w} \in \mathbb{R}^2$ is a standard Brownian motion, and the pair $(G, \boldsymbol{w})$ is a Gaussian process.

The definitions of LAN, LAMN, and LABF extend to more general likelihood ratios and to any finite number of unknown parameters. For details see [8, Chapter II] (LAN), [3, Chapter 1] (LAMN), and [11, Section 2] (unified approach to LAN, LAMN, and LABF).

The LAN and LAMN properties of ℓ_T imply certain asymptotic efficiency of the corresponding maximum likelihood estimator (MLE):

(1) If ℓ_T is LAN, then the corresponding MLE is asymptotically efficient in the sense of achieving the lower bound in the Cramer-Rao inequality; for details see [8, Theorem II.12.1].

(2) If ℓ_T is LAMN, then the corresponding MLE has the maximal concentration property; for details, see [3, Theorem 2.2.1]. Note that the result requires non-degeneracy of the matrix B and therefore does not immediately extend to DLAMN.

In the LABF case, there are results about asymptotic efficiency of Bayessian estimators [11, Section 3, Proposition 10] and sequential estimators [7, Theorem 2].

To put our results in perspective, let us recall the estimation problem of the drift θ in the CAR(1) model, which is the one-dimensional OU process $Y = Y(t)$ defined by

$$(2.12) \qquad dY(t) = \theta Y(t)dt + \sigma dW(t).$$

Here is a summary of the results. For details, see [6, 12].

- The maximum likelihood estimator $\widehat{\theta}_T$ of θ using the observations of $Y(t)$, $0 \le t \le T$, is

$$(2.13) \qquad \widehat{\theta}_T = \frac{\int_0^T Y(t)dY(t)}{\int_0^T Y^2(t)dt};$$

  the estimator is strongly consistent as $T \to \infty$: $\lim_{T\to\infty} \widehat{\theta}_T = \theta$ with probability one for all $\theta \in \mathbb{R}$.
- If $\theta < 0$ (asymptotically stable or ergodic case), then

$$(2.14) \qquad \lim_{T\to\infty} \sqrt{|\theta|T}(\widehat{\theta}_T - \theta) \overset{d}{=} \sqrt{2}\,|\theta|\,\xi,$$

  where ξ is a standard normal random variable.
- If $\theta = 0$ (neutrally stable case), then

$$(2.15) \qquad \lim_{T\to\infty} T(\widehat{\theta}_T - \theta) \overset{d}{=} \frac{w^2(1) - 1}{2\int_0^1 w^2(s)ds},$$

  where $w = w(s)$, $0 \le s \le 1$, is a standard Brownian motion.
- If $\theta > 0$ (unstable or explosive case), then

$$\lim_{T\to\infty} e^{\theta T}(\widehat{\theta}_T - \theta) \overset{d}{=} 2\theta\,\frac{\eta}{\xi + c},$$

  where $\xi = \sqrt{2\theta}\int_0^\infty e^{-\theta t}dW(t)$ is a standard normal random variable, η is a standard normal random variable independent of ξ, and $c = \sqrt{2\theta}\,Y(0)/\sigma$. In particular, if $Y(0) = 0$, then the limit has the Cauchy distribution and does not depend on σ.
- If $\theta \ne 0$, then NLRR holds:

$$(2.16) \qquad \lim_{T\to\infty} \left(\int_0^T Y^2(t)dt\right)^{1/2} (\widehat{\theta}_T - \theta) \overset{d}{=} \sigma\,\eta,$$

  where η is a standard normal random variable.
- The normalized log-likelihood ratio

$$\ell_T(u) = \frac{u}{\sigma}\,\frac{\int_0^T Y(t)dW(t)}{\left(\mathbb{E}\int_0^T Y^2(t)dt\right)^{1/2}} - \frac{u^2}{2\sigma^2}\,\frac{\int_0^T Y^2(t)dt}{\mathbb{E}\int_0^T Y^2(t)dt}, \quad u \in \mathbb{R},$$

is LAN if $\theta < 0$, LABF if $\theta = 0$, and LAMN if $\theta > 0$.

Note that (a) the limit distributions in (2.14), (2.15), and (2.16) do not depend on the initial condition; (b) equality (2.16) illustrates the attractive features of NLRR: the rate

$$R(T) = \left(\int_0^T Y^2(t) dt \right)^{1/2}$$

does not explicitly depend on θ and the limit distribution does not depend on θ or the initial conditions.

Table 1 summarizes the results, where $F_d(w)$ denotes a generic functional of the standard d-dimensional Brownian motion and Ch denotes a Cauchy-type distribution (ratio of two independent normal random variables).

TABLE 1. Estimation in CAR(1)

Parameter	Rate	LD	ℓ_T	NLRR
$\theta < 0$	$\sqrt{T}$	$\mathcal{N}$	LAN	Yes
$\theta = 0$	T	$F_1(w)$	LABF	No
$\theta > 0$	$e^{\theta T}$	Ch	LAMN	Yes

Let us now turn to equation (1.2). Asymptotic behavior of estimators (2.4) depends on the roots p, q of the characteristic equation

(2.17) $$r^2 - \theta_1 r - \theta_2 = 0.$$

There are nine cases to consider:

(1) The asymptotically stable (ergodic) case, when $\theta_1 < 0$ and $\theta_2 < 0$. All in all, there are three possibilities for the roots: $q < p < 0$, $q = p < 0$, and $p = \lambda + \sqrt{-1}\,\nu$, $q = \lambda - \sqrt{-1}\,\nu$, with $\lambda < 0$, but the asymptotic behavior of the estimators is the same in all three cases.

(2) Six non-ergodic cases with real p, q: $q < p = 0$, $q < 0 < p$, $p > q = 0$, $p > q > 0$, $p = q > 0$, and $p = q = 0$.

(3) The harmonic oscillator, when $p = \sqrt{-1}\,\nu$, $q = -\sqrt{-1}\,\nu$, $\nu > 0$;

(4) Unstable oscillations, when $p = \lambda + \sqrt{-1}\,\nu$, $q = \lambda - \sqrt{-1}\,\nu$, and $\lambda > 0, \nu > 0$.

Table 2 summarizes the results for CAR(2). The detailed statements are in Section 3. In Table 2, we use the same notations as in Table 1. In particular, $F_d(w)$ denotes a functional of the standard d-dimensional Brownian motion and Ch is a Cauchy-type distribution.

There are obvious similarities between CAR(1) and CAR(2) in the ergodic case and, perhaps less obvious, similarities in the neutrally stable case ($\theta = 0$ in CAR(1) compared to $p = \sqrt{-1}\nu$ in CAR(2)) and in the exponentially unstable cases ($\theta > 0$ in CAR(1) compared to real $p, q > 0$ in CAR(2)). In both CAR(1) and CAR(2), the rate $\sqrt{T}$ corresponds to normal distribution in the limit, exponential rate corresponds to a Cauchy-type distribution, and any rate polynomial in T leads to some functional of

TABLE 2. Estimation in CAR(2)

Case	v_1	v_2	LD$_1$	LD$_2$	NLRR	ℓ_T
$\theta_1 < 0,\ \theta_2 < 0$	$\sqrt{T}$	$\sqrt{T}$	$\mathcal{N}$	$\mathcal{N}$	Yes	LAN
$q < 0 < p$	$\sqrt{T}$	$\sqrt{T}$	$\mathcal{N}$	$\mathcal{N}$	Yes	DLAMN
$0 < q < p$	e^{qT}	e^{qT}	Ch	Ch	Yes	DLAMN
$0 < q = p$	$T^{-1}e^{qT}$	$T^{-1}e^{qT}$	Ch	Ch	Yes	DLAMN
$q < p = 0$	$\sqrt{T}$	T	$\mathcal{N}$	$F_1(w)$	Yes $(\widehat{\theta}_{1,T})$	LABF/LAN
$q = 0 < p$	T	T	$F_1(w)$	$F_1(w)$	No	DLAMN
$q = p = 0$	T	T^2	$F_1(w)$	$F_1(w)$	No	LABF
$\Re(p) = 0$	T	T	$F_2(w)$	$F_2(w)$	No	LABF
$\Re(p) = \lambda > 0$	$e^{\lambda T}$	$e^{\lambda T}$	Many	Many	Yes	LAMN family

the standard Brownian motion. In the case of the positive double root ($p = q > 0$) the rate is slightly slower than exponential, but the limit distribution is still of the Cauchy type. Also of interest are (a) several appearances of DLAMN instead of LAMN, (b) an unusual combination of asymptotic normality of the estimators and DLAMN (rather than LAN) of ℓ_T when $q < 0 < p$, (c) relative compactness rather than convergence in distribution for both the estimators and the normalized log-likelihood ratio when $p = \lambda + \sqrt{-1}\nu,\ \lambda > 0$.

It is instructive to compare CAR(2) with $p = \lambda + \sqrt{-1}\nu,\ \lambda > 0, \nu > 0$, and the example considered in [17, Section 4.1]:

$$(2.18) \qquad \begin{pmatrix} dX_1(t) \\ dX_2(t) \end{pmatrix} = \begin{pmatrix} \lambda & -\nu \\ \nu & \lambda \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} dt + \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix}.$$

While the eigenvalues of the matrix in (2.18) are also $\lambda \pm \sqrt{-1}\nu$, the special structure of the model ensures that the normalized local log-likelihood ratio is LAMN and the MLEs of λ and ν, when normalized by $\sqrt{2}\,\lambda e^{\lambda T}$, converge to a joint limit (which, for zero initial conditions, is the bivariate t_2-distribution).

3. ASYMPTOTIC PROPERTIES OF THE MLE AND THE NORMALIZED LOG-LIKELIHOOD RATIO

Strong consistency of (2.4) is a consequence of a more general result by Basak and Lee [2]. For the sake of completeness, here are the statement and the proof.

Theorem 3.1. *With probability one,* $\lim_{T\to\infty}\widehat{\theta}_{1,T} = \theta_1$ *and* $\lim_{T\to\infty}\widehat{\theta}_{2,T} = \theta_2$ *for every* $\boldsymbol{\theta} \in \mathbb{R}^2$.

Proof. With $\widehat{\Theta}_T$ defined in (1.6), we find using (1.2) that

$$\widehat{\Theta}_T = \begin{pmatrix} 0 & 1 \\ \widehat{\theta}_{2,T} & \widehat{\theta}_{1,T} \end{pmatrix}.$$

The statement of the theorem now follows from [2, Theorem 2.1 and Remark 3.1]. $\square$

Next, we present the theorem describing the limit distributions of $\widehat{\theta}_{1,T}$ and $\widehat{\theta}_{2,T}$. The proofs are in Section 5. To keep visual track of the formulas, it is convenient to think

of the process X in (1.2) as a dimensionless quantity and to measure t in the units $[t]$ of time. Table 3 summarizes the resulting dimensions of all the variables and parameters in the problem.

TABLE 3. Dimensions in CAR(2)

Quantity	t	$X(t)$	$W(t)$	$\dot{X}(t)$, θ_1, $\widehat{\theta}_{1,T}$, p, q	σ	$\ddot{X}(t)$, θ_2, $\widehat{\theta}_{2,T}$
Units	$[t]$	None	$[t]^{1/2}$	$[t]^{-1}$	$[t]^{-3/2}$	$[t]^{-2}$

One caveat: the auxiliary Brownian motion $w(s)$, $0 \le s \le 1$, and its parameter s, as well as all random variables appearing in the limit distributions, are dimensionless.

Theorem 3.2 (Rate of convergence and limit distributions).
I. ERGODIC CASE. *Assume that $\theta_1 < 0$ and $\theta_2 < 0$, and let η_1, η_2 be iid standard normal random variables. Then*

$$\lim_{T \to \infty} \sqrt{T|\theta_1|}(\widehat{\theta}_{1,T} - \theta_1) \stackrel{d}{=} \sqrt{2}|\theta_1|\,\eta_1, \quad \lim_{T \to \infty} \sqrt{T|\theta_1|}(\widehat{\theta}_{2,T} - \theta_2) \stackrel{d}{=} \sqrt{2|\theta_2|}\,|\theta_1|\,\eta_2.$$

For the rest of the theorem, denote by p and q the roots of equation (2.17).

II. NON-ERGODIC CASE: DISTINCT REAL ROOTS. *Let ξ, η be iid standard normal random variables, and let $w = w(s)$, $0 \le s \le 1$, be a standard Brownian motion independent of η.*
  (A) *If $p > 0$ and $q < 0$, then*

$$(3.1) \qquad \lim_{T \to \infty} \sqrt{|q|T}\,(\widehat{\theta}_{1,T} - \theta_1) \stackrel{d}{=} -\frac{1}{p} \lim_{T \to \infty} \sqrt{|q|T}\,(\widehat{\theta}_{2,T} - \theta_2) \stackrel{d}{=} \sqrt{2}\,|q|\,\eta.$$

  (B) *If $p > q > 0$, then*

$$(3.2) \qquad \lim_{T \to \infty} e^{qT}(\widehat{\theta}_{1,T} - \theta_1) \stackrel{d}{=} -\frac{1}{p} \lim_{T \to \infty} e^{qT}\,(\widehat{\theta}_{2,T} - \theta_2) \stackrel{d}{=} \frac{2(p+q)q}{p-q}\,\frac{\eta}{\xi + c},$$

where $c = \sqrt{2q}\,(\dot{X}(0) - pX(0))/\sigma$.
  (C) *If $p = 0$ and $q < 0$, then $\theta_1 = q$, $\theta_2 = 0$, and*

$$(3.3) \qquad \lim_{T \to \infty} \sqrt{|\theta_1|T}(\widehat{\theta}_{1,T} - \theta_1) \stackrel{d}{=} \sqrt{2}\,|\theta_1|\,\eta, \quad \lim_{T \to \infty} T\widehat{\theta}_{2,T} \stackrel{d}{=} |\theta_1|\,\frac{w^2(1) - 1}{2\int_0^1 w^2(s)ds}.$$

  (D) *If $p > 0$ and $q = 0$. Then $\theta_1 = p$, $\theta_2 = 0$, and*

$$(3.4) \qquad \lim_{T \to \infty} \theta_1 T\,(\widehat{\theta}_{1,T} - \theta_1) \stackrel{d}{=} -\lim_{T \to \infty} T\widehat{\theta}_{2,T} \stackrel{d}{=} \theta_1 \frac{w^2(1) - 1}{2\int_0^1 w^2(s)ds}.$$

III. NON-ERGODIC CASE: A DOUBLE ROOT.
  (A) *If $p = q > 0$, then $\theta_1 = 2q$, $\theta_2 = -q^2$, and*

$$(3.5) \qquad \lim_{T \to \infty} \frac{e^{qT}}{qT}(\widehat{\theta}_{1,T} - \theta_1) \stackrel{d}{=} -\frac{1}{q} \lim_{T \to \infty} \frac{e^{qT}}{qT}\,(\widehat{\theta}_{2,T} - \theta_2) \stackrel{d}{=} 4\sqrt{2}\,q\frac{\eta}{\xi + c},$$

where $c = \sqrt{2q}\,(\dot{X}(0) - pX(0))/\sigma$ and ξ, η are iid standard normal random variables.

(B) *If $p = q = 0$, then $\theta_1 = 0$, $\theta_2 = 0$, and*

(3.6)
$$\lim_{T\to\infty} T\widehat{\theta}_{1,T} \overset{d}{=} \frac{2\mathfrak{z}_3\big(w^2(1) - 1\big) - 2\mathfrak{z}_1^2\big(w(1)\mathfrak{z}_1 - \mathfrak{z}_2\big)}{4\mathfrak{z}_2\mathfrak{z}_3 - \mathfrak{z}_1^4},$$

$$\lim_{T\to\infty} T^2\widehat{\theta}_{2,T} \overset{d}{=} \frac{4\mathfrak{z}_2\big(w(1)\mathfrak{z}_1 - \mathfrak{z}_2\big) - \mathfrak{z}_1^2\big(w^2(1) - 1\big)}{4\mathfrak{z}_2\mathfrak{z}_3 - \mathfrak{z}_1^4},$$

where $w = w(s)$, $0 \le s \le 1$, is a standard Brownian motion, and

$$\mathfrak{z}_1 = \int_0^1 w(s)ds, \quad \mathfrak{z}_2 = \int_0^1 w^2(s)ds, \quad \mathfrak{z}_3 = \int_0^1 \left(\int_0^t w(s)ds\right)^2 dt.$$

IV. NON-ERGODIC CASE: COMPLEX ROOTS.
(A) *If $p = \sqrt{-1}\nu$, $\nu > 0$, then $\theta_1 = 0$, $\theta_2 = -\nu^2$, and*

(3.7)
$$\lim_{T\to\infty} T\widehat{\theta}_{1,T} \overset{d}{=} \frac{2 - w_1^2(1) - w_2^2(1)}{\int_0^1 w_1^2(t)dt + \int_0^1 w_2^2(t)dt},$$

$$\lim_{T\to\infty} T(\widehat{\theta}_{2,T} - \theta_2) \overset{d}{=} 2\nu\frac{\int_0^1 w_1(t)dw_2(t) - \int_0^1 w_2(t)dw_1(t)}{\int_0^1 w_1^2(t)dt + \int_0^1 w_2^2(t)dt},$$

where w_1, w_2 are independent standard Brownian motions.
(B) *If $p = \lambda + \sqrt{-1}\nu$, $\lambda > 0$, $\nu > 0$, then, for $i = 1, 2$, the families $\{e^{\lambda T}(\widehat{\theta}_{i,T} - \theta_i), T > 0\}$ are relatively compact, and all the limit distributions are the form*

$$\frac{\bar{\xi}_c\bar{\eta}_c + \bar{\xi}_s\bar{\eta}_s}{\bar{\xi}_c^2 + \bar{\xi}_s^2},$$

where $(\bar{\xi}_c, \bar{\xi}_s)$ and $(\bar{\eta}_c, \bar{\eta}_s)$ are independent bivariate normal vectors, $\mathbb{E}\bar{\eta}_c = \mathbb{E}\bar{\eta}_s = 0$, and the mean values of $\bar{\xi}_c$ and $\bar{\xi}_s$ depend on the initial conditions $X(0), \dot{X}(0)$.

One general conclusion of Theorem 3.2 is that, if $p > 0$ and $q < p$, then it is the value of the smaller root q that determines asymptotic behavior of the estimators. This result comes as a surprise: the asymptotic behavior of both estimators is dictated by the non-dominant mode, even though this mode is "invisible" with probability one. Indeed, the solution of (1.2) is a Gaussian process

(3.8)
$$X(t) = X(0)x_1(t) + \dot{X}(0)x_2(t) + \sigma\int_0^t x_2(t - s)dW(s),$$

where the functions $x_1(t), x_2(t)$ form the fundamental system of solutions for the equation

(3.9)
$$\ddot{x}(t) - \theta_1\dot{x}(t) - \theta_2 x(t) = 0.$$

In other words, $x_1(0) = 1, \dot{x}_1(0) = 0, x_2(0) = 0, \dot{x}_2(0) = 1$, and both $x_1 = x_1(t)$ and $x_2 = x_2(t)$ satisfy (3.9). The roots of the characteristic equation (2.17) are

(3.10)
$$p = \frac{\theta_1 + \sqrt{\theta_1^2 + 4\theta_2}}{2}, \quad q = \frac{\theta_1 - \sqrt{\theta_1^2 + 4\theta_2}}{2}.$$

Then, with the usual modifications for complex, p, q,

$$(3.11) \qquad x_1(t) = \begin{cases} \dfrac{qe^{pt} - pe^{qt}}{q - p}, & \text{if } p > q, \\ (1 - qt)e^{qt}, & \text{if } p = q; \end{cases} \qquad x_2(t) = \begin{cases} \dfrac{e^{pt} - e^{qt}}{p - q}, & \text{if } p > q, \\ te^{qt}, & \text{if } p = q. \end{cases}$$

When the roots (3.10) are real and distinct, equation (3.9) has two Lyapunov exponents, p and q, and it follows from (3.8) that if $q < p$, then

$$\mathbb{P}\left(\lim_{t \to \infty} \frac{1}{t} \ln |X(t)| = q\right) = \mathbb{P}\left(\lim_{t \to \infty} \frac{1}{t} \ln |\dot{X}(t)| = q\right) = 0$$

for all initial conditions $X(0)$, $\dot{X}(0)$. Thus, Theorem 3.2 shows that if the larger Lyapunov exponent of (3.9) is positive, then the asymptotic behavior of the estimators is determined by the smaller Lyapunov exponent.

As another illustration of the effects of the exponentially unstable mode, note that if $\theta_2 = 0$, then (1.2) becomes

$$d\dot{X} = \theta_1 \dot{X} + \sigma \dot{W}.$$

In other words, $\dot{X}$ is a CAR(1) process. If $\theta_1 < 0$, then asymptotics of $\widehat{\theta}_{1,T}$ and $\widehat{\theta}_{2,T}$ are similar to the CAR(1) situation in ergodic and neutrally stable cases, respectively. If $\theta_1 \geq 0$, then there is no clear similarity with CAR(1).

When $p > 0$, the normalized limits of $\widehat{\theta}_{1,T} - \theta_1$ and $\widehat{\theta}_{2,T} - \theta_2$ are negative multiples of each other. Some correlation is also present when $p = \lambda + \sqrt{-1}\nu$, $\lambda \geq 0, \nu > 0$. This type of correlation in non-ergodic multi-parameter models has been observed before; see, for example, [17, Section 4.1]. Still, as the case $q < 0 = p$ shows, lack of ergodicity does not necessarily imply correlation of the limits.

One can verify qualitative consistency of the results of Theorem 3.2 by considering various limiting regimes for p and q. For example, passing to limit $p \searrow q$ in (3.2) suggests that the rate in the case of the positive double root should be slower than exponential.

Next, we study the possibility of NLRR, that is, existence of a random matrix $R = R_T$ such that $R_T(\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})$ converges in distribution to a bivariate normal vector.

Theorem 3.3 (Normal Limit with a Random Rate (NLRR)). *Denote by p and q the roots of equation (2.17). Normal limit with a random rate is possible in the following six cases: ergodic; $p > q > 0$; $p = q > 0$; $q < 0 < p$; $q < 0 = p$ (for $\widehat{\theta}_{1,T}$ only); $p = \lambda + \sqrt{-1}\nu$, $\lambda > 0$, $\nu > 0$.*

ERGODIC CASE. Assume that $\theta_1 < 0$, $\theta_2 < 0$, and let η_1, η_2 be iid standard normal random variables. Then

$$\lim_{T \to \infty} \left(\int_0^T \dot{X}^2(t)dt\right)^{1/2} (\widehat{\theta}_{1,T} - \theta_1) \overset{d}{=} \sigma\, \eta_1,$$

$$\lim_{T \to \infty} \left(\int_0^T X^2(t)dt\right)^{1/2} (\widehat{\theta}_{2,T} - \theta_2) \overset{d}{=} \sigma\, \eta_2.$$

For the rest of the theorem, denote by η a standard normal random variable.

DISTINCT POSITIVE ROOTS. *Assume that $p > q > 0$ and define*

$$(3.12) \qquad r(T) = \left(\int_0^T \left(\dot{X}(t) - pX(t) \right)^2 dt \right)^{1/2}.$$

Then

$$\lim_{T \to \infty} r(T)(\widehat{\theta}_{1,T} - \theta_1) \overset{d}{=} -\frac{1}{p} \lim_{T \to \infty} r(T)(\widehat{\theta}_{2,T} - \theta_2) \overset{d}{=} \frac{p+q}{p-q} \sigma \eta.$$

POSITIVE DOUBLE ROOT. *Assume that $p = q > 0$ and define*

$$r(T) = \frac{1}{T^2} \left(\int_0^T X^2(t) dt \right)^{1/2}.$$

Then

$$\lim_{T \to \infty} r(T)(\widehat{\theta}_{1,T} - \theta_1) \overset{d}{=} -\frac{1}{p} \lim_{T \to \infty} r(T)(\widehat{\theta}_{2,T} - \theta_2) \overset{d}{=} 2\sqrt{2} p \sigma \eta.$$

ROOTS OF OPPOSITE SIGN. *Assume that $q < 0 < p$ and let $r(T)$ be as in (3.12).*
Then

$$\lim_{T \to \infty} r(T)(\widehat{\theta}_{1,T} - \theta_1) \overset{d}{=} -\frac{1}{p} \lim_{T \to \infty} r(T)(\widehat{\theta}_{2,T} - \theta_2) \overset{d}{=} \sigma \eta.$$

ZERO ROOT. *Assume that $q < 0$, $p = 0$, and define*

$$r(T) = T^{-3/2} \int_0^T X^2(t) dt.$$

Then

$$\lim_{T \to \infty} r(T)(\widehat{\theta}_{1,T} - \theta_1) \overset{d}{=} \frac{\sigma^2}{\sqrt{2}|q|^{3/2}} \eta.$$

COMPLEX ROOTS. *Assume that $p = \lambda + \sqrt{-1}\nu$, $\lambda > 0$, $\nu > 0$ and define matrices*

$$A_T = \begin{pmatrix} \nu & \lambda \\ 0 & -1 \end{pmatrix} e^{-\lambda T}, \ \ B(x,y) = \frac{1}{x^2 + y^2} \begin{pmatrix} x & y \\ -y & x \end{pmatrix}.$$

Then there exists a bivariate normal vector $(\mathfrak{u}_s, \mathfrak{u}_c)$ such that the family

$$B(\mathfrak{u}_s, \mathfrak{u}_c) A_T^\top \Psi_T \left(\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta} \right), \ T > 0,$$

is relatively compact, and all partial limits are bivariate normal random vectors inde-
pendent of $(\mathfrak{u}_s, \mathfrak{u}_c)$.

Theorem 3.3 suggests two negative conclusions: (a) $\Psi_T^{1/2}$ is usually not the correct random normalization (which is especially striking when $p = \lambda + \sqrt{-1}\nu$); (b) existence of NLRR in non-ergodic CAR(2) is not as helpful as in CAR(1), because the rates and/or the limit distributions depend on the unknown parameters.

Finally, we describe the asymptotic structure of the normalized log-likelihood ratio ℓ_T, defined in (2.8).

Theorem 3.4 (The structure of ℓ_T)**.** *Denote by p and q the roots of equation (2.17), by $\mathrm{diag}(x,y)$ the diagonal matrix with x and y on the main diagonal, and by $\boldsymbol{b}_p$ the column vector*

$$\boldsymbol{b}_p = \begin{pmatrix} 1 \\ p \end{pmatrix}.$$

If $\theta_1 < 0$, $\theta_2 < 0$ (ergodic case), and $A_T = \mathrm{diag}(T^{-1/2}, T^{-1/2})$, then ℓ_T is LAN.

If $p > 0$, $q < p$, and $A_T = e^{-pT} \boldsymbol{b}_p \boldsymbol{b}_p^\top$, then ℓ_T is degenerate LAMN and the matrix B in (2.10) is

$$(3.13) \qquad B = \frac{(1+p^2)^2}{2p} \sigma^2 \zeta^2 \, \boldsymbol{b}_p \boldsymbol{b}_p^\top,$$

where ζ is a standard normal random variable.

If $q < 0 = p$ and $A_T = \mathrm{diag}(T^{-1}, T^{-1/2})$ then ℓ_T is mixed LABF/LAN:

$$\ell_\infty(\boldsymbol{u}) = \boldsymbol{u}^\top \boldsymbol{\xi} - \frac{1}{2} \boldsymbol{u}^\top B \boldsymbol{u},$$

where

$$\boldsymbol{\xi} = \begin{pmatrix} \sigma |q|^{-1} \int_0^1 w(s) dw(s) \\ (2|q|)^{-1/2} \sigma \eta \end{pmatrix}, \quad B = \begin{pmatrix} \sigma^2 |q|^{-2} \int_0^1 w^2(s) ds & 0 \\ 0 & \sigma^2 (2|q|)^{-1} \end{pmatrix},$$

η is a standard normal random variable, w is a standard Brownian motion, and η and w are independent.

If $p = q > 0$ and $A_T = T^{-1} e^{-pT} \boldsymbol{b}_p \boldsymbol{b}_p^\top$, then ℓ_T is degenerate LAMN and the matrix B in (2.10) is given by (3.13).

If $p = q = 0$ and $A_T = \mathrm{diag}(T^{-2}, T^{-1})$, then ℓ_T is LABF and the matrix $G(t)$ in (2.11) is

$$G(t) = \begin{pmatrix} \sigma \int_0^t w(s) ds & 0 \\ \sigma w(t) & 0 \end{pmatrix},$$

where w is a standard Brownian motions.

If $p = \sqrt{-1}\nu$, $\nu > 0$, and $A_T = \mathrm{diag}(T^{-1}, T^{-1})$, then ℓ_T is LABF and the matrix $G(t)$ in (2.11) is

$$G(t) = \begin{pmatrix} \sigma w_1(t) & \sigma w_2(t) \\ -\sigma w_2(t) & \sigma w_1(t) \end{pmatrix},$$

where w_1 and w_2 are independent standard Brownian motions.

If $p = \lambda + \sqrt{-1}\nu$, $\lambda, \nu > 0$, and

$$A_T = \begin{pmatrix} \nu & \lambda \\ 0 & -1 \end{pmatrix} e^{-\lambda T},$$

then $\{\ell_T, T > 0\}$ is a relatively compact LAMN family having all partial limits of the form (2.10).

One general conclusion of Theorem 3.4 is that, if $p > 0$ and $q < p$, then it is the value of the larger root p that determines asymptotic behavior of the normalized log-likelihood ratio, which is in sharp contrast with Theorem 3.2.

4. Preparation for the proofs

To study asymptotic behavior of $\widehat{\theta}_{1,T}$ and $\widehat{\theta}_{2,T}$, we need the expressions for the residuals $\widehat{\theta}_{i,T} - \theta_i$, $i = 1, 2$:

(4.1)

$$\widehat{\theta}_{1,T} - \theta_1 = \frac{\left(\int_0^T X^2(t)dt\right)\left(\int_0^T \dot{X}(t)\sigma dW(t)\right) - \left(\int_0^T X(t)\dot{X}(t)dt\right)\left(\int_0^T X(t)\sigma dW(t)\right)}{\left(\int_0^T \dot{X}^2(t)dt\right)\left(\int_0^T X^2(t)dt\right) - \left(\int_0^T X(t)\dot{X}(t)dt\right)^2},$$

$$\widehat{\theta}_{2,T} - \theta_2 = \frac{\left(\int_0^T \dot{X}^2(t)dt\right)\left(\int_0^T X(t)\sigma dW(t)\right) - \left(\int_0^T X(t)\dot{X}(t)dt\right)\left(\int_0^T \dot{X}(t)\sigma dW(t)\right)}{\left(\int_0^T \dot{X}^2(t)dt\right)\left(\int_0^T X^2(t)dt\right) - \left(\int_0^T X(t)\dot{X}(t)dt\right)^2}.$$

These equalities follow directly from (2.4) and (1.2).

Equation (1.2) has a closed-form solution (3.8), meaning that (4.1) can be written in terms of integrals of the type $\int_0^t f(s)dW(s)$. Unfortunately, this direct approach quickly leads to intractable expressions. Therefore, despite availability of explicit formulas, a more sophisticated approach to the analysis of (4.1) is necessary. In the ergodic case, the ergodic theorem provides all the necessary tools, and when $\theta_1 = \theta_2 = 0$, the expressions are simplified using self-similarity of the standard Brownian motion. Other cases benefit from the following construction.

Given two square-integrable on $[0, T]$ functions f, g, define

(4.2) $$D(T; f, g) = \left(\int_0^T f^2(t)dt\right)\left(\int_0^T g^2(t)dt\right) - \left(\int_0^T f(t)g(t)dt\right)^2,$$

(4.3) $$N(T; f, g) = \left(\int_0^T f^2(t)dt\right)\left(\int_0^T g(t)\sigma dW(t)\right)$$
$$- \left(\int_0^T f(t)g(t)dt\right)\left(\int_0^T f(t)\sigma dW(t)\right).$$

Clearly, $D(T; f, g) = D(T; g, f)$, but in general $N(T; f, g) \neq N(T; g, f)$. Then formulas (4.1) become

(4.4) $$\widehat{\theta}_{1,T} - \theta_1 = \frac{N(T; X, \dot{X})}{D(T; X, \dot{X})}, \quad \widehat{\theta}_{2,T} - \theta_2 = \frac{N(T; \dot{X}, X)}{D(T; X, \dot{X})}.$$

The general idea of the proof of Theorem 3.2 is to find the asymptotic behavior of $D(T; X, \dot{X})$, $N(T; X, \dot{X})$, and $N(T; \dot{X}, X)$, as $T \to \infty$. To keep track of the results, note that, according to Table 3, $D(T; X, \dot{X})$ is dimensionless, $N(T; X, \dot{X})$ is measured in $[t]^{-1}$, and $N(T; \dot{X}, X)$ is measured in $[t]^{-2}$.

For every real numbers $\alpha, \beta, \gamma, \kappa$ and every square-integrable functions f, g, we have the following identities:

(4.5) $$D(T; \alpha f + \beta g, \gamma f + \kappa g) = (\alpha\kappa - \beta\gamma)^2 D(T; f, g),$$

(4.6) $$N(T; \alpha f + \beta g, \gamma f + \kappa g) = (\alpha^2\kappa - \alpha\beta\gamma)N(T; f, g) + (\beta^2\gamma - \alpha\beta\kappa)N(T; g, f).$$

Recall that $0_{\text{a.s.}}(t)$ denotes a continuous random process converging to zero with probability one as $t \to \infty$. We will often use the following result: if $f(t) > 0$ is a

continuous process, $F(T) = \int_0^T f(s)ds$, and $\lim_{T\to\infty} F(T) = \infty$ with probability one, then

$$(4.7) \qquad \frac{1}{F(T)} \int_0^T f(t)\, 0_{\text{a.s.}}(t)dt = 0_{\text{a.s.}}(T).$$

Indeed, (4.7) is immediate if the integral $\int_0^\infty f(t)\, 0_{\text{a.s.}}(t)dt$ converges; otherwise, (4.7) follows after one application of L'Hospital's rule.

Here are some other technical results to be used later. For $r > 0$, define Gaussian random variables

$$(4.8) \qquad \xi_r = \int_0^\infty e^{-rs}dW(s), \quad \eta_r(T) = \int_0^T e^{-r(T-s)}dW(s).$$

We have $\mathbb{E}\eta_r(T) = 0$, $\mathbb{E}\eta_r^2(T) = (1 - e^{-2rT})/(2r)$, and therefore

$$(4.9) \qquad \lim_{T\to\infty} \eta_r(T) \overset{d}{=} \eta_r,$$

where the random variable η_r is normal with mean zero and variance $1/(2r)$. Since

$$(4.10) \qquad \mathbb{E}\xi_q\eta_r(T) = e^{-rT} \int_0^T e^{(r-q)s}ds \to 0, \ T \to \infty,$$

it follows that ξ_q and η_r are independent for every $q, r > 0$.

A more sophisticated version of the above observations is the following result.

Proposition 4.1. *Let $M = (M_1(t), \ldots, M_d(t))$, $0 \le t \le 1$, be a d-dimensional continuous Gaussian martingale with $M(0) = 0$, and let $M_T = (M_{T,1}(t), \ldots, M_{T,d}(t))$, $T \ge 0$, $0 \le t \le 1$, be a family of continuous square-integrable d-dimensional martingales such that $M_T(0) = 0$ for all T.*
(a) If

$$\lim_{T\to\infty} \langle M_{T,i}, M_{T,j} \rangle(t_0) = \langle M_i, M_j \rangle(t_0)$$

in probability for some $t_0 \in [0,1]$ and every $i, j = 1, \ldots, d$, then $\lim_{T\to\infty} M_T(t_0) \overset{d}{=} M(t_0)$.
(b) If

$$\lim_{T\to\infty} \langle M_{T,i}, M_{T,j} \rangle(t) = \langle M_i, M_j \rangle(t)$$

in probability for every $t \in [0,1]$ and $i, j = 1, \ldots, d$, then $\lim_{T\to\infty} M_T \overset{\mathcal{L}}{=} M$ in the topology of continuous functions on $[0,1]$.

Proof. Modulo a non-essential (in this case) difference between a sequence and a family indexed by the positive reals, part (a) is a particular case of Theorem VIII.3.6 in [9], and part (b) is a particular case of Theorem VIII.3.11 in [9]. $\square$

Corollary 4.2. *Let $\eta_r(T)$ and η_r be from (4.8) and (4.9), respectively, and let $\eta_r^\perp$ be an independent copy of η_r. Then*

$$(4.11) \qquad \lim_{T\to\infty} \left(\eta_r(T), \frac{1}{\sqrt{T}} \int_0^T \eta_r(t)dW(t) \right) \overset{d}{=} (\eta_r, \eta_r^\perp).$$

Proof. This follows from Proposition 4.1, part (a), with

$$M_{T,1}(t) = \int_0^{Tt} e^{-r(T-s)} dW(s), \quad M_{T,2} = \frac{1}{\sqrt{T}} \int_0^{Tt} \eta_r(s) dW(s),$$

$M = (w_1(t), w_2(t))/\sqrt{2r}$, and $t_0 = 1$, where w_1 and w_2 are independent standard Brownian motions. First, note that, as $T \to \infty$,

$$\langle M_{T,1}\rangle(1) = \int_0^T e^{-2r(T-s)} ds \to \frac{1}{2r} = \langle M_1\rangle(1).$$

Next, observe that, for every $r > 0$, the process $\eta_r(t)$, $t \geq 0$, is ergodic and the invariant measure of the process is Gaussian with mean zero and variance $1/(2r)$. The ergodic theorem implies

$$(4.12) \qquad \frac{1}{T} \int_0^T |\eta_r|^n(t) dt = \mathbb{E}|\eta_r|^n + 0_{\text{a.s.}}(T), \quad n > 0.$$

In particular, as $T \to \infty$,

$$\langle M_{2,T}\rangle(1) = \frac{1}{T} \int_0^T \eta_r^2(t) dt \to \frac{1}{2r} = \langle M_2\rangle(1).$$

Finally,

$$\langle M_{1,T}, M_{2,T}\rangle(1) = \frac{1}{\sqrt{T}} \int_0^T e^{-r(T-s)} \eta_r(s) ds,$$

and, by the Hölder inequality,

$$|\langle M_{1,T}, M_{2,T}\rangle(t)| \leq \frac{1}{T^{1/6}} \left(\int_0^T e^{3r(T-s)/2} ds \right)^{2/3} \left(\frac{1}{T} \int_0^T |\eta_r(s)|^3 ds \right)^{1/3}.$$

Using (4.12) with $n = 3$,

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T |\eta_r(s)|^3 ds = \mathbb{E}|\eta_r|^3 \quad \text{with probability one,}$$

and therefore

$$\lim_{T \to \infty} \langle M_{1,T}, M_{2,T}\rangle(1) = 0 = \langle M_1, M_2\rangle(1) \quad \text{with probability one,}$$

which completes the proof of the corollary. $\square$

Asymptotic analysis of certain stochastic integrals can benefit from the law of iterated logarithm. Recall that if f is locally square-integrable adapted process, $M(t) = \int_0^t f(s) dW(s)$ and $\langle M\rangle(t) = \int_0^t f^2(s) ds \nearrow +\infty$, $t \to \infty$, with probability one, then there exists a standard Brownian motion $\bar{W}$ such that

$$M(t) = \bar{W}(\langle M\rangle(t)).$$

The law of iterated logarithm for $\bar{W}$ implies

$$(4.13) \qquad \frac{M(T)}{\sqrt{\langle M\rangle(T)} \, (\ln\langle M\rangle(T))^\varepsilon} = 0_{\text{a.s.}}(T)$$

for every $\varepsilon > 0$. For example,

$$(4.14) \qquad T^{-\varepsilon} \eta_r(T) = 0_{\text{a.s.}}(T), \quad \varepsilon > 0, \quad r > 0.$$

To conclude the general discussion we establish an integration by parts formula. As a motivation, recall that analysis of CAR(1) in the exponentially unstable case leads to the function $\tilde{V}_p(t) = \int_0^t e^{p(t-s)} dW(s)$, $p > 0$, for which integration by parts shows that

$$(4.15) \qquad e^{-pT}\left(\int_0^T \tilde{V}_p(t) dW(t) - e^{pT}\xi_p \eta_p(T)\right) = 0_{\text{a.s.}}(T).$$

Since exponentially unstable solutions of equation (3.9) are of the form $e^{pt}f(t)$, where the function f grows at most polynomially, we generalize (4.15) as follows.

Proposition 4.3. *Given a deterministic (for simplicity) and locally square-integrable function f, define $S_f(t) = \int_0^t f(s) dW(s)$.*
Let functions φ and ψ be such that

$$(4.16) \qquad \int_0^{+\infty} \psi^2(t) dt = +\infty, \quad e^{pT}|\varphi(T)| + e^{-pT}|\psi(T)| \le C(1 + T^r), \; T \ge 0,$$

for some $p, C, r > 0$. Then, for all $q > 0$,

$$(4.17) \qquad e^{-qT}\left(\int_0^T S_\varphi(t) dS_\psi(t) - S_\varphi(T)S_\psi(T)\right) = 0_{\text{a.s.}}(T).$$

Proof. By the Itô formula,

$$(4.18) \qquad S_\varphi(T)S_\psi(T) - \int_0^T S_\varphi(t) dS_\psi(t) = \int_0^T \varphi(t)\psi(t) dt + \int_0^T S_\psi(t) dS_\varphi(t).$$

It follows from (4.16) that

$$|\varphi(t)\psi(t)| \le C^2(1 + t^r)^2 \le 2C^2(1 + t^{2r}),$$

and then

$$e^{-qT}\int_0^T |\varphi(t)\psi(t)| dt \le 2C^2 e^{-qT}(T + T^{2r+1}) \to 0, \; T \to \infty.$$

To estimate the second term on the right-hand side of (4.18), recall that, if $M = M(t)$ is a continuous square-integrable martingale, then, by the strong law of large numbers, a finite limit

$$\lim_{T\to\infty} \frac{M(T)}{1 + \langle M\rangle(T)}$$

exists with probability one ([15, Corollary 2 to Theorem 2.6.10]). As a result, if $F = F(t)$ is a function such that $F(T)\langle M\rangle(T) = 0_{\text{a.s.}}(T)$, then $F(T)M(T) = 0_{\text{a.s.}}(T)$.
Next, consider

$$N(t) = \int_0^t S_\psi(s) dS_\varphi(s) = \int_0^t S_\psi(s)\varphi(s) dW(s) \quad \text{with} \quad \langle N\rangle(t) = \int_0^t S_\psi^2(s)\varphi^2(s) ds.$$

We need to show that $e^{-qT}\langle N\rangle(T) = 0_{\text{a.s.}}(T)$. By (4.16),

$$\varphi^2(t) \le 2C^2 e^{-2pt}(1 + t^{2r}), \; \psi^2(t) \le 2C^2 e^{2pt}(1 + t^{2r}),$$

so that

$$(4.19) \qquad \langle S_\psi\rangle(t) = \int_0^t \psi^2(s) ds \le 2C^2 t e^{2pt}(1 + t^{2r}),$$

and, with $q > 0$,

$$\lim_{t\to\infty} e^{-qt}\langle S_\psi\rangle(t)\varphi^2(t) \leq \lim_{t\to\infty} 4C^4 e^{-qt}t(1+t^{2r})^2 = 0.$$

By assumption, $\sup_t\langle S_\psi\rangle(t) = +\infty$. We then use (4.13) and (4.19) to conclude that

$$\frac{S_\psi^2(T)}{\langle S_\psi\rangle(T)\,T^\varepsilon} = 0_{\text{a.s.}}(T),\ \ \varepsilon > 0.$$

Therefore,

$$e^{-qT}S_\psi^2(T)\varphi^2(T) = T^\varepsilon e^{-qT}\langle S_\psi\rangle(T)\varphi^2(T)\frac{S_\psi^2(T)}{\langle S_\psi\rangle(T)\,T^\varepsilon} = 0_{\text{a.s.}}(T),$$

which, by L'Hospital's rule, implies $e^{-qT}\langle N\rangle(T) = 0_{\text{a.s.}}(T)$.
   This completes the proof of Proposition 4.3. $\qquad\square$

We will use Proposition 4.3 with $q = p$ to simplify various stochastic integrals. As a quick illustration, let us verify (4.15). Take $\varphi(t) = e^{-pt}$, $\psi(t) = e^{pt}$. Then, together with (4.8), equality (4.17) implies

$$e^{-pT}\int_0^T\left(\int_0^t e^{p(t-s)}dW(s)\right)dW(t) = \left(\int_0^T e^{-p(T-t)}dW(t)\right)\left(\int_0^T e^{-pt}dW(t)\right)$$
$$+0_{\text{a.s.}}(T) = \eta_p(T)\,\xi_p + 0_{\text{a.s.}}(T).$$

5. Proofs of Theorems 3.2–3.4

Proof of Theorem 3.2.
I. The ergodic case. See [1, Remark 2 after Theorem 4.6.2].

II. Non-ergodic case: Distinct real roots. If $p \neq q$, then (3.8) and (3.11) imply

(5.1) $$X(t) = V_p(t) - V_q(t),\ \ \dot{X}(t) = pV_p(t) - qV_q(t),$$

where

(5.2)
$$V_p(t) = e^{pt}U_p(t),\ U_p = \frac{\dot{X}(0) - X(0)q}{p-q} + \frac{\sigma}{p-q}\int_0^t e^{-ps}dW(s),$$
$$V_q(t) = e^{qt}U_q(t),\ U_q = \frac{\dot{X}(0) - X(0)p}{p-q} + \frac{\sigma}{p-q}\int_0^t e^{-qs}dW(s).$$

By (4.5) and (4.6)

(5.3)
$$D(T;X,\dot{X}) = (p-q)^2 D(T;V_p,V_q),$$
$$N(T;X,\dot{X}) = (p-q)\Big(N(T;V_p,V_q) + N(T;V_q,V_p)\Big),$$
$$N(T;\dot{X},X) = -(p-q)\Big(pN(T;V_p,V_q) + qN(T;V_q,V_p)\Big).$$

To streamline computations even further, introduce the notations

$$\mathbb{V}_{r,T} = \int_0^T V_r^2(t)dt, \ r = p, q,$$

$$\mathbb{V}_{pq,T} = \int_0^T V_p(t)V_q(t)dt, \ q < p,$$

$$\mathbb{S}_{r,T} = \int_0^T V_r(t)dW(t), \ r = p, q.$$

Then, by (4.2) and (4.3),

$$(5.4) \qquad\qquad D(T; V_p, V_q) = \mathbb{V}_{p,T}\mathbb{V}_{q,T} - \mathbb{V}_{pq,T}^2,$$

$$(5.5) \qquad\qquad N(T; V_p, V_q) = \sigma\big(\mathbb{V}_{p,T}\mathbb{S}_{q,T} - \mathbb{V}_{pq,T}\mathbb{S}_{p,T}\big),$$

$$(5.6) \qquad\qquad N(T; V_q, V_p) = \sigma\big(\mathbb{V}_{q,T}\mathbb{S}_{p,T} - \mathbb{V}_{pq,T}\mathbb{S}_{q,T}\big),$$

and it remains to investigate asymptotic behavior of the processes $\mathbb{V}$ and $\mathbb{S}$ as $T \to \infty$.

II(A). ROOTS OF OPPOSITE SIGN: $q < 0 < p$. Using notations (4.8), define Gaussian random variable ζ_p by

$$(5.7) \qquad \zeta_p = \frac{\dot{X}(0) - X(0)q}{p - q} + \frac{\sigma}{p - q}\int_0^{+\infty} e^{-ps}dW(s) = \frac{\dot{X}(0) - X(0)q}{p - q} + \frac{\sigma}{p - q}\xi_p$$

and note that

$$(5.8) \qquad V_p(t) = e^{pt}\big(\zeta_p + 0_{\text{a.s.}}(t)\big), \ V_q(t) = \frac{\sigma}{p - q}\eta_{|q|}(t) + \frac{\dot{X}(0) - X(0)p}{p - q}e^{qt}.$$

Then

$$\mathbb{V}_{p,T} = \int_0^T V_p^2(t)dt = \left(\frac{\zeta_p^2}{2p} + 0_{\text{a.s.}}(T)\right)e^{2pT} \text{ (by (5.8))},$$

$$\mathbb{V}_{q,T} = \int_0^T V_q^2(t)dt = \frac{\sigma^2 T}{2|q|(p - q)^2}\big(1 + 0_{\text{a.s.}}(T)\big) \text{ (by (5.8) and (4.12) with } n = 2),$$

$$(5.9) \qquad \mathbb{V}_{pq,T} = \int_0^T V_p(t)V_q(t)dt = T^\varepsilon e^{pT}0_{\text{a.s.}}(T), \ \varepsilon > 0 \text{ (see below)},$$

$$\mathbb{S}_{p,T} = \int_0^T V_p(t)dW(t) = T^\varepsilon e^{pT}0_{\text{a.s.}}(T), \ \varepsilon > 0 \text{ (by (5.8) and (4.13))},$$

$$\lim_{T\to\infty}\frac{\mathbb{S}_{q,T}}{\sqrt{T}} = \lim_{T\to\infty}\frac{1}{\sqrt{T}}\int_0^T V_q(t)dW(t) \overset{d}{=} \frac{\sigma}{p - q}\eta_{|q|}^\perp \text{ (by (5.8) and (4.11))}.$$

To establish (5.9), use (5.8), followed by L'Hospital's rule and (4.14):

$$\lim_{T\to\infty}T^{-\varepsilon}e^{-pT}\int_0^T \big|V_p(t)V_q(t)\big|dt = \lim_{T\to\infty}T^{-\varepsilon}e^{-pT}\int_0^T \big|e^{pt}\big(\zeta_p + 0_{\text{a.s.}}(t)\big)V_q(t)\big|dt$$

$$= \lim_{T\to\infty}\frac{\big|\sigma\zeta_p\eta_{|q|}(T)\big|}{p(p - q)T^\varepsilon} = 0.$$

Taking $\varepsilon \le 1/4$,

$$D_T(T; V_{p,T}, V_{q,T}) = \mathbb{V}_{p,T}\mathbb{V}_{q,T} - \mathbb{V}_{pq,T}^2 = \frac{Te^{2pT}}{(p-q)^2}\left(\frac{\zeta_p^2}{2p}\frac{\sigma^2}{2|q|} + 0_{\text{a.s.}}(T)\right),$$

$$\lim_{T\to\infty}\frac{e^{-2pT}}{\sqrt{T}}N(T; V_p, V_q) = \lim_{T\to\infty}\frac{\sigma e^{-2pT}}{\sqrt{T}}\left(\mathbb{V}_{p,T}\mathbb{S}_{q,T} - \mathbb{V}_{pq,T}\mathbb{S}_{p,T}\right) \overset{d}{=} \frac{\sigma^2}{p-q}\frac{\zeta_p^2}{2p}\eta_{|q|}^{\perp},$$

$$\lim_{T\to\infty}\frac{e^{-2pT}}{\sqrt{T}}N(T; V_q, V_p) = \lim_{T\to\infty}\frac{\sigma e^{-2pT}}{\sqrt{T}}\left(\mathbb{V}_{q,T}\mathbb{S}_{p,T} - \mathbb{V}_{pq,T}\mathbb{S}_{q,T}\right) \overset{d}{=} 0.$$

Then both equalities in (3.1) follow from (5.3) and (4.4).

II(B). DISTINCT POSITIVE ROOTS: $0 < q < p$. This time, we have

(5.10) $V_r(t) = e^{rt}\left(\zeta_r + 0_{\text{a.s.}}(T)\right), \ r = p, q,$

where ζ_p is given by (5.7) and

$$\zeta_q = \frac{\dot{X}(0) - X(0)p}{p-q} + \frac{\sigma}{p-q}\int_0^{+\infty}e^{-qs}dW(s) = \frac{\dot{X}(0) - X(0)p}{p-q} + \frac{\sigma}{p-q}\xi_q.$$

Then

$$\mathbb{V}_{r,T} = \int_0^T V_r^2(t)dt = \left(\frac{\zeta_r^2}{2r} + 0_{\text{a.s.}}(T)\right)e^{2pT}, \ r = p, q \ \text{(by (5.10))},$$

$$\mathbb{V}_{pq,T} = \int_0^T V_p(t)V_q(t)dt = e^{(p+q)T}\frac{\zeta_p\zeta_q}{p+q}\left(1 + 0_{\text{a.s.}}(T)\right) \ \text{(by (5.10))},$$

$$\lim_{T\to\infty}e^{-rT}\mathbb{S}_{r,T} = \lim_{T\to\infty}e^{-rT}\int_0^T V_r(t)dW(t) \overset{d}{=} \zeta_r\eta_r, \ r = p, q \ \text{(by (5.2) and Prop. 4.3).}$$

Now combine and (5.4)–(5.6) with the above asymptotics of $\mathbb{V}$ and $\mathbb{S}$ to get

$$D(T; V_p, V_q) = \mathbb{V}_{p,T}\mathbb{V}_{q,T} - \mathbb{V}_{pq,T}^2 = \frac{\zeta_p^2\zeta_q^2(p-q)^2}{4pq(p+q)^2}e^{2(p+q)T}\left(1 + 0_{\text{a.s.}}(T)\right),$$

$$\lim_{T\to\infty}e^{-(q+2p)T}N(T; V_p, V_q) = \lim_{T\to\infty}\sigma e^{-(q+2p)T}\left(\mathbb{V}_{p,T}\mathbb{S}_{q,T} - \mathbb{V}_{pq,T}\mathbb{S}_{p,T}\right)$$

$$\overset{d}{=} \sigma\zeta_p^2\zeta_q\left(\frac{\eta_q}{2p} - \frac{\eta_p}{p+q}\right),$$

$$\lim_{T\to\infty}e^{-(q+2p)T}N(T; V_q, V_p) = \lim_{T\to\infty}\sigma e^{-(q+2p)T}\left(\mathbb{V}_{q,T}\mathbb{S}_{p,T} - \mathbb{V}_{pq,T}\mathbb{S}_{q,T}\right) \overset{d}{=} 0.$$

It remains to observe that, similar to (4.9),

$$\lim_{T\to\infty}(\eta_p(T), \eta_q(T)) \overset{d}{=} (\eta_p, \eta_q), \ p, q > 0,$$

and (η_p, η_q) is a Gaussian vector with

$$\mathbb{E}\eta_p\eta_q = \lim_{T\to\infty}\mathbb{E}\eta_p(T)\eta_q(T) = \lim_{T\to\infty}\int_0^T e^{-(p+q)(T-s)}ds = \frac{1}{p+q}.$$

Then

$$\frac{\eta_q}{2p} - \frac{\eta_p}{p+q}$$

is a Gaussian random variable with mean zero and variance

$$\frac{1}{2q \cdot 4p^2} + \frac{1}{2p(p+q)^2} - \frac{2}{2p \cdot (p+q)^2} = \frac{(p+q)^2 - 4pq}{8p^2q(p+q)^2} = \frac{(p-q)^2}{8p^2q(p+q)^2};$$

similar to (4.10), the pair (η_p, η_q) is independent of the pair (ζ_p, ζ_q). Then both equalities in (3.2) follow from (5.3) and (4.4).

II(c). LARGER ROOT IS ZERO: $q < p = 0$. Formulas (5.1) and (5.2) with $p = 0$ lead to

(5.11) $$X(t) = V_0(t) - V_q(t), \quad \dot{X}(t) = -qV_q(t),$$

and

(5.12) $$V_0(t) = U_0(t) = \frac{\dot{X}(0) - X(0)q}{|q|} + \frac{\sigma}{|q|}W(t), \quad V_q = \frac{\dot{X}(0)}{|q|}e^{qt} + \frac{\sigma}{|q|}\eta_{|q|}(t).$$

Then

$$\mathbb{V}_{0,T} = \int_0^T V_0^2(t)dt = \frac{\sigma^2 T^2}{q^2}\left(\frac{1}{T^2}\int_0^T W^2(t)dt + 0_{\text{a.s.}}(T)\right) \text{ (by (5.12))},$$

$$\mathbb{V}_{q,T} = \int_0^T V_q^2(t)dt = \frac{\sigma^2 T}{2|q|^3}(1 + 0_{\text{a.s.}}(T)) \text{ (by (4.12))},$$

(5.13) $$\mathbb{V}_{0q,T} = \int_0^T V_0(t)V_q(t)dt = T^{1+\varepsilon}0_{\text{a.s.}}(T), \; \varepsilon > 0 \text{ (see below)},$$

$$\mathbb{S}_{0,T} = \int_0^T U_0(t)dW(t) = \frac{\sigma T}{|q|}\left(\frac{W^2(T) - T}{2T} + 0_{\text{a.s.}}(T)\right) \text{ (by (5.12))},$$

$$\lim_{T \to \infty} \frac{\mathbb{S}_{q,T}}{\sqrt{T}} = \lim_{T \to \infty} \frac{1}{\sqrt{T}}\int_0^T V_q(t)dW(t) \stackrel{d}{=} \frac{\sigma}{|q|}\eta_{|q|}^{\perp} \text{ (by (4.11))}.$$

To establish (5.13), set $\bar{V}_q(t) = \int_0^t V_q(s)ds$ and integrate by parts

$$\int_0^T V_0(t)V_q(t)dt = V_0(T)\bar{V}_q(T) - \int_0^T \bar{V}_q(t)dV_0(t).$$

By (5.11),

$$\bar{V}_q(t) = \frac{X(t) - X(0)}{|q|} = \frac{V_0(t) - V_q(t) - X(0)}{|q|}.$$

Using (5.12), (4.13), and (4.14),

$$V_0(T)\bar{V}_q(T) = T^{(1/2)+\varepsilon}0_{\text{a.s.}}(T), \; \varepsilon > 0.$$

Similarly,

$$\int_0^T \left((V_0(t) - V_q(t) - X(0)\right)dW(t) = T^{1+\varepsilon}0_{\text{a.s.}}(T), \; \varepsilon > 0,$$

because

$$2\int_0^T W(t)dW(t) = W^2(T) - T = T^{1+\varepsilon}0_{\text{a.s.}}(T), \; \varepsilon > 0 \text{ (by (4.13))}$$

and

$$\int_0^T \eta_r(t)dW(t) = T^{(1/2)+\varepsilon}0_{\text{a.s.}}(T), \ \varepsilon > 0, r > 0 \ (\text{by (4.12) and (4.13)}).$$

Using self-similarity of the standard Brownian motion, namely, $W(T\cdot) \stackrel{\mathcal{L}}{=} \sqrt{T}w(\cdot)$, we conclude that

(5.14) $$\frac{W^2(T) - T}{2T} \stackrel{d}{=} \frac{w^2(1) - 1}{2}, \quad \frac{1}{T^2}\int_0^T W^2(t)dt \stackrel{d}{=} \int_0^1 w^2(s)ds,$$

Taking $\varepsilon \leq 1/2$ in (5.13),

$$D(T; V_0, V_q) = \mathbb{V}_{0,T}\mathbb{V}_{q,T} - \mathbb{V}_{0q,T}^2 = \frac{\sigma^4 T^3}{2|q|^5}\left(\frac{1}{T^2}\int_0^T W^2(t)dt + 0_{\text{a.s.}}(T)\right),$$

$$\lim_{T\to\infty} T^{-5/2}N(T; V_0, V_q) = \lim_{T\to\infty} \frac{\sigma}{T^{5/2}}\left(\mathbb{V}_{0,T}\mathbb{S}_{q,T} - \mathbb{V}_{0q,T}\mathbb{S}_{0,T}\right) \stackrel{d}{=} \frac{\sigma^4}{q^2}\eta_{|q|}^\perp \int_0^1 w^2(s)ds,$$

$$\lim_{T\to\infty} T^{-2}N(T; V_q, V_0) = \lim_{T\to\infty} \frac{\sigma}{T^2}\left(\mathbb{V}_{q,T}\mathbb{S}_{0,T} - \mathbb{V}_{0q,T}\mathbb{S}_{q,T}\right) \stackrel{d}{=} \frac{\sigma^4}{4q^2}\left(w^2(1) - 1\right).$$

Then both equalities in (3.3) follow from (5.3) and (4.4).

To show independence of η and w, take a standard Brownian motion $\tilde{w} = \tilde{w}(t)$, $t \in [0,1]$, that is independent of w and apply Proposition 4.1, part (b), with

$$M(t) = (w(t), \ \tilde{w}(t)), \ M_T(t) = \left(\frac{1}{\sqrt{T}}\int_0^{Tt} dW(s), \ \frac{\sqrt{2|q|}}{\sqrt{T}}\int_0^{Tt} \eta_{|q|}(s)dW(s)\right).$$

II(D). SMALLER ROOT IS ZERO: $q = 0 < p$. Formulas (5.1) and (5.2) with $q = 0$ lead to

$$X(t) = V_p(t) - V_0(t), \ \dot{X}(t) = pV_p(t),$$

(5.15) $$V_0(t) = U_0(t) = \frac{\dot{X}(0) - X(0)p}{p} + \frac{\sigma}{p}W(t), \ V_p(t) = e^{pt}\left(\zeta_p + 0_{\text{a.s.}}(t)\right),$$

$$\zeta_p = \frac{\dot{X}(0)}{p} + \frac{\sigma}{p}\int_0^\infty e^{-pt}dW(t).$$

Then

$$\mathbb{V}_{p,T} = \int_0^T V_p^2(t)dt = \frac{e^{2pT}}{2p}\left(\zeta_p^2 + 0_{\text{a.s.}}(T)\right) \ (\text{by (5.15)}),$$

$$\mathbb{V}_{0,T} = \int_0^T V_0^2(t)dt = \frac{\sigma^2 T^2}{p^2}\left(\frac{1}{T^2}\int_0^T W^2(t)dt + 0_{\text{a.s.}}(T)\right) \ (\text{by (5.15)}),$$

$$\mathbb{V}_{p0,T} = \int_0^T V_p(t)V_0(t)dt = Te^{pT}0_{\text{a.s.}}(T) \ (\text{by L'Hospital's rule}),$$

$$\mathbb{S}_{p,T} = \int_0^T V_p(t)dW(t) = e^{pT}\left(\zeta_p \eta_p(T) + 0_{\text{a.s.}}(T)\right) \ (\text{by Proposition 4.3}),$$

$$\mathbb{S}_{0,T} = \int_0^T U_0(t)dW(t) = \frac{\sigma T}{p}\left(\frac{W^2(T) - T}{2T} + 0_{\text{a.s.}}(T)\right) \ (\text{by (5.15)}).$$

As a result,

$$D(T; V_p, V_0) = \mathbb{V}_{p,T}\mathbb{V}_{0,T} - \mathbb{V}_{p0,T}^2 = \frac{\sigma^2 T^2 e^{2pT}}{2p^3}\left(\zeta_p^2 \frac{1}{T^2}\int_0^T W^2(t)dt + 0_{\text{a.s.}}(T)\right),$$

$$N(T; V_p, V_0) = \sigma\left(\mathbb{V}_{p,T}\mathbb{S}_{0,T} - \mathbb{V}_{p0,T}\mathbb{S}_{p,T}\right) = \frac{\sigma^2}{2p^2}Te^{2pT}\left(\zeta_p^2\frac{W^2(T)-T}{2T} + 0_{\text{a.s.}}(T)\right),$$

$$N(T; V_0, V_p) = \sigma\left(\mathbb{V}_{0,T}\mathbb{S}_{p,T} - \mathbb{V}_{p0,T}\mathbb{S}_{0,T}\right)$$

$$= \frac{\sigma^2 T^2}{p^2}e^{pT}\left(\zeta_p\eta_p(T)\frac{1}{T^2}\int_0^T W^2(t)dt + 0_{\text{a.s.}}(T)\right) = Te^{2pT}0_{\text{a.s.}}(T),$$

and both equalities in (3.4) follow from (5.14), (5.3), and (4.4).

III(A). POSITIVE DOUBLE ROOT: $p = q > 0$. With $x_1(t) = (1 - qt)e^{qt}$, $x_2(t) = te^{qt}$, (3.8) becomes

$$X(t) = X(0)(1 - qt)e^{qt} + \dot{X}(0)te^{qt} + \sigma\int_0^t (t-s)e^{q(t-s)}dW(s),$$

(5.16) $\qquad \dot{X}(t) = qX(t) + Q(t),$ where

$$Q(t) = \left(\dot{X}(0) - X(0)q + \sigma\int_0^t e^{-qs}dW(s)\right)e^{qt}.$$

If we define

$$\zeta = \dot{X}(0) - X(0)q + \sigma\int_0^{+\infty} e^{-qs}dW(s),$$

then

(5.17) $\qquad X(t) = te^{qt}\left(\zeta + 0_{\text{a.s.}}(t)\right),\ \ Q(t) = e^{qt}\left(\zeta + 0_{\text{a.s.}}(t)\right).$

By (4.5) and (5.16),

$$D(T; X, \dot{X}) = D(T; X, qX + Q) = D(T; X, Q).$$

For $f(t) = te^{qt}$, $g(t) = e^{qt}$, direct computations show that

$$D(T; f, g) = \frac{e^{4qT}}{16q^4}(1 + \epsilon(T)),\ \ \lim_{T\to\infty}\epsilon(T) = 0.$$

As a result, (5.17) implies

(5.18) $\qquad D(T; X, \dot{X}) = e^{4qT}\left(\frac{\zeta^4}{16q^4} + 0_{\text{a.s.}}(T)\right).$

Next, define

$$\eta_{q,1}(t) = \int_0^t e^{-q(t-s)}dW(s),\ \ \eta_{q,2}(t) = \frac{1}{t}\int_0^t s\,e^{-q(t-s)}dW(s),$$

and observe that

(5.19) $\qquad \lim_{T\to\infty} T\left(\eta_{q,1}(T) - \eta_{q,2}(T)\right) \stackrel{d}{=} \frac{\eta}{2q^{3/2}},$

where η is a standard Gaussian random variable, independent of ζ. Indeed, for every $T > 0$, $\big(\eta_{q,1}(T), \eta_{q,2}(T)\big)$ is a two-dimensional Gaussian vector with zero mean and

$$\lim_{T\to\infty} T^2 \mathbb{E}\big(\eta_{q,1}(T) - \eta_{q,2}(T)\big)^2 = \lim_{T\to\infty} \int_0^T (T-s)^2 e^{-2q(T-s)} ds = \int_0^{+\infty} t^2 e^{-2qt} dt = \frac{1}{4q^3};$$

independence between η and ζ follows from the computation similar to (4.10).

It follows from (5.17) that

$$\int_0^T X^2(t)dt = T^2 e^{2qT}\left(\frac{\zeta^2}{2q} + 0_{\mathrm{a.s.}}(T)\right), \quad \int_0^T Q^2(t)dt = e^{2qT}\left(\frac{\zeta^2}{2q} + 0_{\mathrm{a.s.}}(T)\right),$$

$$\int_0^T X(t)Q(t)dt = T e^{2qt}\left(\frac{\zeta^2}{2q} + 0_{\mathrm{a.s.}}(T)\right), \quad \int_0^T Q(t)dW(t) = e^{qT}\big(\zeta\,\eta_{q,1}(T) + 0_{\mathrm{a.s.}}(T)\big),$$

$$\int_0^T X(t)dW(t) = T e^{qT}\big(\zeta\,\eta_{q,2}(T) + 0_{\mathrm{a.s.}}(T)\big);$$

the last two equalities also rely on (5.16) and Proposition 4.3. Then

$$N(T; X, Q) = \sigma\left(\int_0^T X^2(t)dt \int_0^T Q(t)dW(t) - \int_0^T X(t)Q(t)dt \int_0^T X(t)dW(t)\right)$$

$$= \frac{\sigma}{2q} T^2 e^{3qT}\big(\eta_{q,1}(T) - \eta_{q,2}(T)\big)\big(\zeta^3 + 0_{\mathrm{a.s.}}(T)\big),$$

$$N(T; Q, X) = \sigma\left(\int_0^T Q^2(t)dt \int_0^T X(t)dW(t) - \int_0^T X(t)Q(t)dt \int_0^T Q(t)dW(t)\right)$$

$$= \frac{\sigma}{2q} T e^{3qT}\big(\eta_{q,2}(T) - \eta_{q,1}(T)\big)\big(\zeta^3 + 0_{\mathrm{a.s.}}(T)\big).$$

It remains to observe that, by (5.16) and (4.6),

$$N(T; X, \dot X) = N(T; X, Q), \ \ N(T; \dot X, X) = -qN(T; X, Q) + N(T; Q, X).$$

Then both equalities in (3.5) follow from (4.4) and (5.19).

III(B). ZERO DOUBLE ROOT: $p = q = 0$. In this case the result follows directly from (4.1) using self-similarity of the standard Brownian motion: $W(T\cdot) \overset{\mathcal{L}}{=} \sqrt{T} w(\cdot)$. Recall the notations

$$\mathfrak{z}_1 = \int_0^1 w(s)ds, \quad \mathfrak{z}_2 = \int_0^1 w^2(s)ds, \quad \mathfrak{z}_3 = \int_0^1 \left(\int_0^t w(s)ds\right)^2 dt.$$

With $\theta_1 = \theta_2 = 0$, equation (1.2) becomes $d\dot X = \sigma\, dW$, and therefore

$$\dot X(t) = \dot X(0) + \sigma\, W(t), \ \ X(t) = X(0) + \dot X(0)t + \sigma \int_0^t W(s)ds.$$

Then

$$D(T; X, \dot{X}) \stackrel{d}{=} \frac{\sigma^4 T^6}{4} \Big(4\mathfrak{z}_2\mathfrak{z}_3 - \mathfrak{z}_1^4 + 0_{\text{a.s.}}(T) \Big),$$

$$N(T; X, \dot{X}) \stackrel{d}{=} \frac{\sigma^4 T^5}{2} \Big(\mathfrak{z}_3 \big(w^2(1) - 1 \big) - \mathfrak{z}_1^2 \big(w(1)\mathfrak{z}_1 - \mathfrak{z}_2 \big) + 0_{\text{a.s.}}(T) \Big),$$

$$N(T; \dot{X}, X) \stackrel{d}{=} \frac{\sigma^4 T^4}{4} \Big(4\mathfrak{z}_2 \big(w(1)\mathfrak{z}_1 - \mathfrak{z}_2 \big) - \mathfrak{z}_1^2 \big(w^2(1) - 1 \big) + 0_{\text{a.s.}}(T) \Big),$$

leading to (3.6). If $X(0) = \dot{X}(0) = 0$, then equalities in (3.6) hold for every $T > 0$.

IV(A). COMPLEX ROOTS: $p = \sqrt{-1}\nu$, $\nu > 0$. This case is the subject of [14]. Below we outline the main steps.

By (3.8),

$$X(t) = X(0)\cos(\nu t) + \frac{\dot{X}(0)}{\nu}\sin(\nu t) + \frac{\sigma}{\nu}\int_0^t \sin(\nu(t-s))dW(s),$$

$$\dot{X}(t) = -X(0)\nu\sin(\nu t) + \dot{X}(0)\cos(\nu t) + \sigma\int_0^t \cos(\nu(t-s))dW(s).$$

By Proposition 4.1, as $T \to \infty$, the pair

$$\left(\frac{\sqrt{2}}{\sqrt{T}}\int_0^{Tt} \sin(\nu s)\,dW(s), \ \frac{\sqrt{2}}{\sqrt{T}}\int_0^{Tt} \cos(\nu s)\,dW(s), \ t \in [0, 1] \right)$$

converges in distribution to a two-dimensional standard Brownian motion (w_1, w_2). Then

$$\lim_{T\to\infty} \frac{1}{T^2}\int_0^T X^2(t)dt \stackrel{d}{=} \frac{\sigma^2}{2\nu^2}\int_0^1 (w_1^2(s) + w_2^2(s))ds,$$

$$\lim_{T\to\infty} \frac{1}{T^2}\int_0^T \dot{X}^2(t)dt \stackrel{d}{=} \frac{\sigma^2}{2}\int_0^1 (w_1^2(s) + w_2^2(s))ds,$$

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T X(t)dW(t) \stackrel{d}{=} \frac{\sigma}{2\nu}\int_0^1 \big(w_1(s)dw_2(s) - w_2(s)dw_1(s) \big),$$

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \dot{X}(t)dW(t) \stackrel{d}{=} \frac{\sigma}{2}\int_0^1 \big(w_1(s)dw_1(s) + w_2(s)dw_2(s) \big).$$

Finally, (4.13) implies $X^2(T) = T^2 0_{\text{a.s.}}(T)$. Then both equalities in (3.7) follow from (4.1). For details, see [14, Sections 2–4].

IV(B). COMPLEX CONJUGATE ROOTS: $p = \lambda + \sqrt{-1}\nu$, $\lambda > 0$, $\nu > 0$. The fundamental system of solutions of (3.9) in this case is

$$x_1(t) = \frac{e^{\lambda t}}{\nu}\left(\nu\cos\nu t - \lambda\sin\nu t \right), \ x_2(t) = \frac{1}{\nu}e^{\lambda t}\sin\nu t.$$

Therefore,

$$(5.20) \qquad X(t) = \frac{e^{\lambda t}}{\nu} \left((\dot{X}(0) - X(0)\lambda) \sin \nu t + X(0)\nu \cos \nu t \right)$$

$$+ \frac{\sigma}{\nu} \int_0^t e^{\lambda(t-s)} \sin \nu(t-s) \, dW(s),$$

$$(5.21) \qquad \dot{X}(t) = \lambda X(t) + \nu Y(t), \text{ where}$$

$$(5.22) \qquad Y(t) = \frac{e^{\lambda t}}{\nu} \left((\dot{X}(0) - X(0)\lambda) \cos \nu t - X(0)\nu \sin \nu t \right)$$

$$+ \frac{\sigma}{\nu} \int_0^t e^{\lambda(t-s)} \cos \nu(t-s) dW(s).$$

Define Gaussian random variables

$$(5.23) \qquad \mathfrak{u}_c = \frac{\dot{X}(0) - X(0)\lambda}{\nu} + \frac{\sigma}{\nu} \int_0^{+\infty} e^{-\lambda t} \cos \nu t \, dW(t),$$

$$\mathfrak{u}_s = -X(0) + \frac{\sigma}{\nu} \int_0^{+\infty} e^{-\lambda t} \sin \nu t \, dW(t),$$

and the functions

$$V_{c,d}(t) = e^{\lambda t} \cos \nu t, \ V_{s,d}(t) = e^{\lambda t} \sin \nu t;$$

subscript d stands for *deterministic*. Then

$$(5.24) \ \ X(t) = \mathfrak{u}_c V_{s,d}(t) - \mathfrak{u}_s V_{c,d}(t) + e^{\lambda t} 0_{\text{a.s.}}(t), \ Y(t) = \mathfrak{u}_s V_{s,d}(t) + \mathfrak{u}_c V_{c,d}(t) + e^{\lambda t} 0_{\text{a.s.}}(t).$$

Direct computations show that

$$D(T; V_{c,d}, V_{s,d}) = \frac{\nu^2 e^{4\lambda T}}{16\lambda^2(\lambda^2 + \nu^2)} (1 + \epsilon(T)), \ \lim_{T \to \infty} \epsilon(T) = 0.$$

By (4.5) and (5.21),

$$D(T; X, \dot{X}) = D(T; X, \lambda X + \nu Y) = \nu^2 D(T; X, Y).$$

Combining this with (5.24),

$$(5.25) \qquad \begin{aligned} D(T; X, \dot{X}) &= \nu^2 (\mathfrak{u}_c^2 + \mathfrak{u}_s^2)^2 D(T; V_{c,d}, V_{s,d}) + e^{4\lambda T} 0_{\text{a.s.}}(T) \\ &= \frac{\nu^4 \, e^{4\lambda T}}{16\lambda^2(\lambda^2 + \nu^2)} \left((\mathfrak{u}_c^2 + \mathfrak{u}_s^2)^2 + 0_{\text{a.s.}}(T) \right). \end{aligned}$$

As a side remark, note that if, in (5.25), we pass to the limit $\nu \to 0$ and set $\lambda = q$, we get (5.18).

Next, define Gaussian random variables

$$(5.26) \qquad \mathfrak{h}_c(T) = \int_0^T e^{\lambda(t-T)} \cos \nu t \, dW(t), \ \mathfrak{h}_s(T) = \int_0^T e^{\lambda(t-T)} \sin \nu t \, dW(t).$$

For each $T > 0$, the vector $(\mathfrak{h}_c(T), \mathfrak{h}_s(T))$ is bivariate normal with zero mean,

$$\sigma_c^2(T) = \mathbb{E}\mathfrak{h}_c^2(T) = \frac{1}{4\lambda}\left(1 + \frac{\lambda}{\sqrt{\lambda^2 + \nu^2}}\cos(2\nu T - \phi)\right) + o(1),$$

$$\sigma_s^2(T) = \mathbb{E}\mathfrak{h}_s^2(T) = \frac{1}{4\lambda}\left(1 - \frac{\lambda}{\sqrt{\lambda^2 + \nu^2}}\cos(2\nu T - \phi)\right) + o(1),$$

$$\sigma_{cs}(T) = \mathbb{E}\mathfrak{h}_c(T)\mathfrak{h}_s(T) = \frac{1}{4\sqrt{\lambda^2 + \nu^2}}\sin(2\nu T - \phi) + o(1),$$

where $\cos\phi = \lambda/\sqrt{\lambda^2 + \nu^2}$, $\sin\phi = \nu/\sqrt{\lambda^2 + \nu^2}$, and $o(1)$ denotes a non-random function $\epsilon = \epsilon(T)$ such that $\lim_{T\to\infty}\epsilon(T) = 0$. Thus, the family $\{(\mathfrak{h}_c(T), \mathfrak{h}_s(T)), \ T > 0\}$ is relatively compact, with limit points being bivariate Gaussian vectors $(\mathfrak{h}_c, \mathfrak{h}_s)$. Furthermore, straightforward computations show that (a) similar to (4.10), each vector $(\mathfrak{h}_c, \mathfrak{h}_s)$ is independent of $(\mathfrak{u}_c, \mathfrak{u}_s)$; (b) by definition of $N(T; f, g)$,

$$N(T; V_{c,d}, V_{s,d}) = \sigma e^{3\lambda T}\big(\sigma_c^2(T)\mathfrak{h}_s(T) - \sigma_{cs}\mathfrak{h}_c(T)\big),$$

$$N(T; V_{s,d}, V_{c,d}) = \sigma e^{3\lambda T}\big(\sigma_s^2(T)\mathfrak{h}_c(T) - \sigma_{cs}\mathfrak{h}_s(T)\big).$$

To continue, we use (4.6) and (5.21) to write

(5.27) $N(T; X, \dot{X}) = \nu N(T; X, Y), \quad N(T; \dot{X}, X) = -\nu\lambda N(T; X, Y) + \nu^2 N(T; Y, X).$

Now combine (5.20) and (5.22)–(5.24) with (4.6) and Proposition 4.3 to find $N(T; X, Y)$ and $N(T; Y, X)$ in terms of $N(T; V_{c,d}, V_{s,d})$ and $N(T; V_{s,d}, V_{c,d})$:

(5.28) $N(T; X, Y) = (\mathfrak{u}_c^2 + \mathfrak{u}_s^2)\big(\mathfrak{u}_c N(T; V_{s,d}, V_{c,d}) + \mathfrak{u}_s N(T; V_{c,d}, V_{s,d})\big) + e^{3\lambda T}0_{\text{a.s.}}(T),$

(5.29)
$$N(T; Y, X) = (\mathfrak{u}_c^2 + \mathfrak{u}_s^2)\big(-\mathfrak{u}_s N(T; V_{s,d}, V_{c,d}) + \mathfrak{u}_c N(T; V_{c,d}, V_{s,d})\big) + e^{3\lambda T}0_{\text{a.s.}}(T).$$

It remains to substitute (5.28) and (5.29) into (5.27) and combine the result with (5.25) and (4.4).

This completes the proof of Theorem 3.2.

Proof of Theorem 3.3. The proof is straightforward analysis of the computations in the proof of Theorem 3.2, with the goal to find suitable normalization leading to the Gaussian limit. Let us illustrate this in the most interesting case, when $p = \lambda + \sqrt{-1}\nu$ and the corresponding rate matrix R_T is not diagonal. In this case, (5.20)–(5.24) and (5.26), together with Proposition 4.3, imply

(5.30)
$$\begin{pmatrix} \int_0^T X(s)dW(s) \\ \int_0^T \dot{X}(s)dW(s) \end{pmatrix} = -\frac{e^{\lambda T}}{\nu}\begin{pmatrix} -1 & 0 \\ -\lambda & \nu \end{pmatrix}\begin{pmatrix} \mathfrak{u}_s & -\mathfrak{u}_c \\ \mathfrak{u}_c & \mathfrak{u}_s \end{pmatrix}\begin{pmatrix} \mathfrak{h}_c(T) + 0_{\text{a.s.}}(T) \\ \mathfrak{h}_s(T) + 0_{\text{a.s.}}(T) \end{pmatrix}$$

$$= \big(A_T^\top\big)^{-1}\big(B(\mathfrak{u}_s, \mathfrak{u}_c)\big)^{-1}\begin{pmatrix} \mathfrak{h}_c(T) + 0_{\text{a.s.}}(T) \\ \mathfrak{h}_s(T) + 0_{\text{a.s.}}(T) \end{pmatrix}.$$

On the other hand, according to (4.1),

$$\begin{pmatrix} \widehat{\theta}_{2,T} - \theta_2 \\ \widehat{\theta}_{1,T} - \theta_1 \end{pmatrix} = \Psi_T^{-1}\begin{pmatrix} \int_0^T X(s)dW(s) \\ \int_0^T \dot{X}(s)dW(s) \end{pmatrix}.$$

That is,

$$B(\mathfrak{u}_s, \mathfrak{u}_c)A_T^\top \Psi_T \left(\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}\right) = \left(\begin{array}{c} \mathfrak{h}_c(T) + 0_{\mathrm{a.s.}}(T) \\ \mathfrak{h}_s(T) + 0_{\mathrm{a.s.}}(T) \end{array} \right).$$

This completes the proof of Theorem 3.3.

Proof of Theorem 3.4. The proof is straightforward analysis of the computations in the proof of Theorem 3.2, this time with an emphasis on the asymptotic behavior of the matrix Ψ_T and the vector $\int_0^T \boldsymbol{X}(t)dW(t)$. The analysis is easy in the ergodic case and also when $p = 0$ or $p = \sqrt{-1}\nu$. When $p = \lambda + \sqrt{-1}\nu$, the proof is essentially complete after (5.30). It remain to consider $p > 0$.

If $p > 0$ and $q < p$, then

$$\int_0^T \boldsymbol{X}(t)dW(t) = e^{pT}\zeta_p\big(\eta_p(T) + 0_{\mathrm{a.s.}}(T)\big)\boldsymbol{b}_p, \quad \Psi_T = \frac{\zeta_p^2}{2p} e^{2pT}\Big(\boldsymbol{b}_p\,\boldsymbol{b}_p^\top + 0_{\mathrm{a.s.}}(T)\Big).$$

If $p > 0$ and $q = p$, then

$$\int_0^T \boldsymbol{X}(t)dW(t) = Te^{pT}\zeta\big(\eta_{p,2}(T) + 0_{\mathrm{a.s.}}(T)\big)\boldsymbol{b}_p, \quad \Psi_T = \frac{\zeta^2}{2p} T^2 e^{2pT}\Big(\boldsymbol{b}_p\,\boldsymbol{b}_p^\top + 0_{\mathrm{a.s.}}(T)\Big).$$

Note also that

$$\big(\boldsymbol{b}_p\,\boldsymbol{b}_p^\top\big)^2 = (1 + p^2)\boldsymbol{b}_p\,\boldsymbol{b}_p^\top.$$

This completes the proof of Theorem 3.4.

6. SUMMARY AND DISCUSSION

The maximum likelihood estimator (2.13) in CAR(1), that is, the one-dimensional OU process (2.12), has three types of asymptotic regimes, depending on the sign of the parameter θ. For CAR(2), that is, the second-order equation (1.2) with two unknown parameters, while the MLE still has the same form and is strongly consistent for all values of the parameters, the number of different asymptotic regimes is nine. This jump in complexity underlines the challenges related to the analysis of the general estimation problem, either for the N-th order linear equation (CAR(N)) or an N-by-N system. In fact, our results suggest that the "obvious" normalization of the estimator $\widehat{\Theta}_T$ by either $\Psi_T^{1/2}$ or $\big(\mathbb{E}\Psi_T\big)^{1/2}$ does not always work, and, for $N > 2$, a complete classification similar to Table 2 might be not worth the effort.

If equation (2.17) has real roots and one of them is positive, then the following two features seem to be common:

(1) The rate of convergence of the estimators is determined by the smaller root of equation (2.17);

(2) The asymptotic behavior of the normalized log-likelihood ratio is determined by the larger root of equation (2.17).

We expect similar features in the case $N > 2$ as well.

The problem is much easier for homogeneous equations, that is, multi-dimensional analogues of the geometric Brownian motion. In those models, estimation of the drift matrix in any number of dimensions leads to the LAN situation as long as the diffusion matrix is non-degenerate; see [10] for details.

Equation (1.2) can be discretized exactly, resulting in the finite-difference equation

$$(6.1) \qquad\qquad X_{n+2} + AX_{n+1} + BX_n = \xi_n$$

for the sequence $X_n = X(n\Delta t)$, $n \geq 0$, where the numbers A and B depend on θ_1, θ_2, and Δt; the random variables $\xi_n, n \geq 0$, are not independent: see [4] or [14, Section 5] for details. A similar discretization is possible for the general CAR(N) (1.3). See [4] for analysis of (6.1) in the ergodic case.

A DAR(N), that is, N-th order discrete time auto-regression driven by a white noise sequence, can only be an approximation of the continuous-time model (1.3). See [5] for analysis of DAR(N) when the underlying deterministic dynamic is not exponentially unstable.

7. Acknowledgement

References

[1] M. Arató, *Linear stochastic systems with constant coefficients: a statistical approach*, Lecture Notes in Control and Information Sciences, vol. 45, Springer-Verlag, Berlin, 1982.

[2] G. K. Basak and P. Lee, *Asymptotic properties of an estimator of the drift coefficients of multidimensional Ornstein-Uhlenbeck processes that are not necessarily stable*, Electron. J. Stat. **2** (2008), 1309–1344.

[3] I. V. Basawa and D. J. Scott, *Asymptotic optimal inference for nonergodic models*, Lecture Notes in Statistics, vol. 17, Springer-Verlag, New York, 1983.

[4] P. J. Brockwell, R. A. Davis, and V. Yang, *Continuous-time Gaussian autoregression*, Statistica Sinica **17** (2007), no. 1, 63–80.

[5] N. H. Chan and C. Z. Wei, *Limiting distributions of least squares estimates of unstable autoregressive processes*, Ann. Statist. **16** (1988), no. 1, 367–401.

[6] P. D. Feigin, *Maximum likelihood estimation for continuous-time stochastic processes*, Advances in Appl. Probability **8** (1976), no. 4, 712–736.

[7] P. E. Greenwood and W. Wefelmeyer, *Asymptotic minimax results for stochastic process families with critical points*, Stochastic Process. Appl. **44** (1993), no. 1, 107–116.

[8] I. A. Ibragimov and R. Z. Khas'minskiĭ, *Statistical estimation: Asymptotic theory*, Applications of Mathematics, vol. 16, Springer, 1981.

[9] J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes, 2nd ed.*, Grundlehren der Mathematischen Wissenschaften, vol. 288, Springer, 2003.

[10] A. Jankunas and R. Z. Khasminskii, *Estimation of parameters of linear homogeneous stochastic differential equations*, Stochastic Process. Appl. **72** (1997), no. 2, 205–219.

[11] P. Jeganathan, *Some aspects of asymptotic theory with applications to time series models*, Econometric Theory **11** (1995), no. 5, 818–887.

[12] Yu. A. Kutoyants, *Statistical inference for ergodic diffusion processes*, Springer, 2004.

[13] L. Le Cam and G. L. Yang, *Asymptotics in statistics: Some basic concepts*, second ed., Springer Series in Statistics, Springer-Verlag, New York, 2000.

[14] N. Lin and S. V. Lototsky, *Undamped harmonic oscillator driven by additive Gaussian white noise: A statistical analysis*, Commun. Stoch. Anal. **5** (2011), no. 1, 233–250.

[15] R. Sh. Liptser and A. N. Shiryaev, *Theory of martingales*, Mathematics and its Applications (Soviet Series), vol. 49, Kluwer Academic Publishers, Dordrecht, 1989.

[16] _____, *Statistics of random processes, I: General theory, 2nd ed.*, Applications of Mathematics, vol. 5, Springer, 2001.

[17] H. Luschgy, *Local asymptotic mixed normality for semimartingale experiments*, Probab. Theory Related Fields **92** (1992), no. 2, 151–176.

Current address, N. Lin: Department of Mathematics, USC, Los Angeles, CA 90089 USA, tel. (+1) 213 821 1480; fax: (+1) 213 740 2424
    E-mail address, N. Lin: `nlin@usc.edu`

Current address, S. V. Lototsky (corresponding author): Department of Mathematics, USC, Los Angeles, CA 90089 USA, tel. (+1) 213 740 2389; fax: (+1) 213 740 2424
    E-mail address, S. V. Lototsky: `lototsky@usc.edu`
    URL: `http://www-rcf.usc.edu/~lototsky`