Mathematical Problems in Engineering
Volume 5, pp. 33–54
Reprints available directly from the publisher
Photocopying permitted by license only

A Model Discrimination Method for Processes with Different Memory Structure

MARK FREIDLIN[†] and RUTH PFEIFFER[*,‡]

Department of Mathematics, University of Maryland, College Park, MD 20742, USA

(Received 14 July 1998; In final form 5 November 1998)

In this paper we develop a test for determining whether the observed sample path comes from a system with hysteresis perturbed by noise, or if it arises from a system governed by an ordinary differential equation with the same noise. A large sample size test is constructed, which is appropriate in many practical situations. Two models are considered as alternatives to the hysteresis model. An asymptotic expression for the cutoff point of the test is found using a version of the central limit theorem.

Keywords: Hysteresis, Stochastic processes, Likelihood ratio test

AMS Subject Classification: 60J70, 60K40, 62F05

1 INTRODUCTION

Hysteresis can be a by-product of fundamental physical mechanisms such as phase transitions, or it can be built into a system deliberately in order to monitor its behavior, as in the case of temperature control via a thermostat. Hysteresis is a genuinely nonlinear phenomenon, which is usually not smooth and therefore not easy to treat in a mathematical way. The simplest example of a hysteresis nonlinearity is given by a relay (or loop) operator that is characterized by two

* Corresponding author. Tel.: 301-405-5069. E-mail: pfru@math.umd.edu.
† The work was partly supported by NSF Grant DMS-9504177.
‡ Supported by NASA Grant NGT-30332.

threshold values x_0, x_1, with $x_0 < x_1$, and two output values (which we take to be 0 and 1). In our setting the system with hysteresis that incorporates feedback is governed by one of two stochastic differential equations, depending on which of the two thresholds x_0 and x_1, $x_1 > x_0$ was crossed last. Two models are considered as alternatives to the hysteresis model. The first model is described by a diffusion process $\mathrm{d}X(t) = -a(X(t))\,\mathrm{d}t + \sigma\,\mathrm{d}W(t)$ with the drift $a(x)$ obtaining just two values, that is $a(x) = a$ for $x < x^*$ and $a(x) = -a$ for $x > x^*$, where x^* is the arithmetic mean of the thresholds in the model with hysteresis. The second alternative model is described by the Ornstein–Uhlenbeck process $\mathrm{d}X(t) = -a \cdot (X(t) - x^*)\,\mathrm{d}t + \sigma\,\mathrm{d}W(t)$. We construct the likelihood ratio test for these problems. An asymptotic expression for the cutoff point of the test is found using a version of the central limit theorem based on a renewal argument. The hysteresis model is presented in detail in Section 2. Section 2 also contains the results for the alternative hypothesis of the first type (the drift with a jump). Simulation results that allow to assess the power of the test are presented in Section 3. The proofs are given Section 4. The case of the Ornstein–Uhlenbeck alternative is considered in Section 5.

2 HYSTERESIS VERSUS DIFFUSION PROCESS WITH DISCONTINUOUS DRIFT

In this section we introduce a simple stochastic model with hysteresis that incorporates feedback. The system is governed by one of two stochastic differential equations depending on which of two thresholds x_0 and x_1 was crossed last. The process $X(t)$ is defined as the continuous Markovian solution of the equations

$$\begin{aligned} \mathrm{d}X(t) &= b_{i(t)}\,\mathrm{d}t + \sigma\,\mathrm{d}W(t), \quad X(0) = x < x_0, \\ i(t) &= \begin{cases} 0, & \tau_{2k} \le t < \tau_{2k+1}, \\ 1, & \tau_{2k+1} \le t < \tau_{2k+2}. \end{cases} \end{aligned} \tag{2.1}$$

where $X(0) = x \le x_0$, $\tau_0 = 0$ and

$$\begin{aligned} \tau_{2k+1} &= \min\{t > \tau_{2k}\colon X(t) = x_1\}, \\ \tau_{2k+2} &= \min\{t > \tau_{2k+1}\colon X(t) = x_0\}. \end{aligned}$$

We assume that $b_0 < 0$ and $b_1 > 0$, therefore, as is easy to check, all the τ_n are finite with probability one and also have finite expectation. Note that the process $X(t)$ itself is not Markovian, while the couple $(X(t), i(t))$ is a Markov process. Figure 1 shows the state space of the stochastic process (X, i), where the lower line corresponds to $i(t) = 0$ and the upper line to $i(t) = 1$. The transition from $i = 0$ to $i = 1$ occurs at $x = x_1$ and the transition from $i = 1$ to $i = 0$ at $x = x_0$.

Assume that the drift coefficients of $X(t)$ differ only by the sign, $b_0 = -b_1 = b$. It can be shown that the case of general drift coefficients can be reduced to this special case by properly shifting the process [8].

Let $(X(0), i(0)) = (x_0, 0)$. Denote by $\tilde{X}(t)$ the process defined by the equation $\mathrm{d}\tilde{X}(t) = \sigma \mathrm{d}W(t)$, $\tilde{X}(0) = x_0$. Applying Girsanov's formula (see, for example, [9, Chapter 3]), the Radon–Nikodym derivative of the measure corresponding to the process $X(t)$ on $[0, T]$ with respect to the measure of the process $\tilde{X}(t)$ is given by

$$
\begin{aligned}
\frac{\mathrm{d}\mu_X}{\mathrm{d}\mu_{\tilde{X}}}(X) &= \exp\left\{ \int_0^T \frac{b_{i(t)}}{\sigma}\,\mathrm{d}W(t) - \int_0^T \frac{b_{i(t)}^2}{2\sigma^2}\,\mathrm{d}t \right\} \\
&= \exp\left\{ \sum_i \left(\int_{\tau_{2i}}^{(\tau_{2i+1}\wedge T)} \frac{b}{\sigma}\,\mathrm{d}W(t) - \int_{\tau_{2i}}^{(\tau_{2i+1}\wedge T)} \frac{b^2}{2\sigma^2}\,\mathrm{d}t \right.\right. \\
&\qquad \left.\left. + \int_{\tau_{2i+1}}^{(\tau_{2i+2}\wedge T)} \frac{-b}{\sigma}\,\mathrm{d}W(t) - \int_{\tau_{2i+1}}^{(\tau_{2i+2}\wedge T)} \frac{b^2}{2\sigma^2}\,\mathrm{d}t \right) \right\}.
\end{aligned}
$$

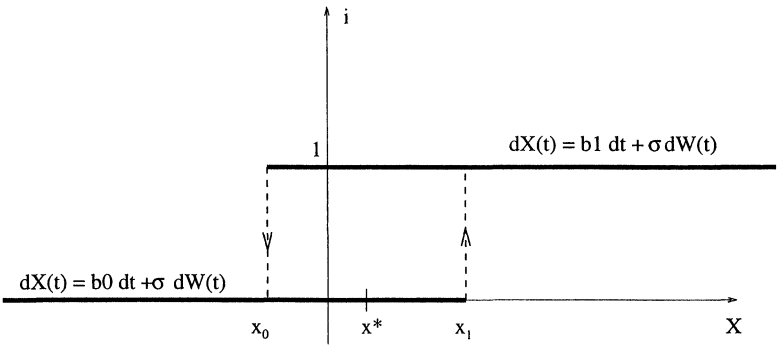

FIGURE 1 State space of the stochastic process (X, i).

We call a cycle for the thresholds x_0, x_1 the time interval $[\tau_{2k}, \tau_{2k+2})$ between two successive intersections of the trajectory $X(t)$ with the level x_0 separated by crossings of the level x_1. The duration of such a cycle is $\tau_{2k+2} - \tau_{2k}$ and it is finite with probability 1. Let $h = x_1 - x_0$ and $p_1(X, T) = \mathrm{d}\mu_X / \mathrm{d}\mu_{\tilde{X}}$.

Then the density can be written in the form

$$p_1(X,T) = \begin{cases} \exp\{2N(T,h)hb/\sigma + (b/\sigma)X(T) - (b^2/2\sigma^2)T\}, \\ \quad \text{if } (\tau_n) = x_0, \\ \exp\{2N(T,h)hb/\sigma + (bh/2\sigma) - (b/\sigma)X(T) - (b^2/2\sigma^2)T\}, \\ \quad \text{if } X(\tau_n) = x_1, \end{cases} \tag{2.2}$$

where $\tau_n = \max\{\tau_k\colon \tau_k \leq T\}$ and $N(T, h)$ is the number of cycles of $X(t)$ in $[0, T]$ for the thresholds x_0, x_1.

We now slightly modify the notation for the model with hysteresis. We assume that the arithmetic mean of the two thresholds, $x^* = (x_0 + x_1)/2$, is known, and $h = x_1 - x_0$ is unknown. The model, also called 'model 1', is now parameterized by x^*, h and b.

First we discuss how to find a test for the case of a known $h = h_0$, and then we show that the result extends to the composite hypothesis when the observations come from a process with hysteresis with some $h \geq h_0 > 0$.

The process with the drift with a jump (referred to as 'model 0', as it is the model of the null hypothesis) satisfies

$$\begin{aligned} \mathrm{d}X(t) &= -a\,\mathrm{d}t + \sigma\,\mathrm{d}W(t), \quad \text{for } X(t) > x^*, \\ \mathrm{d}X(t) &= a\,\mathrm{d}t + \sigma\,\mathrm{d}W(t), \quad\ \ \text{for } X(t) \leq x^*, \end{aligned} \tag{2.3}$$

where $a > 0$, $X(0) = x_0$ and x^* is the midpoint of the interval $[x_0, x_1]$. The equalities (2.3) can be written as

$$\mathrm{d}X(t) = a \cdot (1 - 2I(X(t) > x^*))\,\mathrm{d}t + \sigma\,\mathrm{d}W(t), \quad X(0) = x_0,$$

where $I(A)$ denotes the indicator function of the set A.

Without loss of generality let $x^* = 0$. The Radon–Nikodym derivative of the measure corresponding to this process on $[0, T]$ with respect to the measure for the process $\tilde{X}(t) = \sigma W(t), \tilde{X}(0) = x_0$ has the form

$$\frac{\mathrm{d}\mu_X}{\mathrm{d}\mu_{\tilde{X}}}(X) = \exp\left\{ \int_0^T (a/\sigma)(1 - 2I(X(t) > 0))\, \mathrm{d}X(t) \right. \\ \left. - \int_0^T (a^2/2\sigma^2)(1 - 2I(X(t) > 0))^2\, \mathrm{d}t \right\}.$$

Using that $(1 - 2I(X(t) > 0))^2 = 1$ and applying Tanaka's formula (see, e.g. [9]), one can reduce the density to

$$p_0(X, T) = \frac{\mathrm{d}\mu_X}{\mathrm{d}\mu_{\tilde{X}}}(X) = \exp\left\{ a(L(T, 0) - |X(T)|) - \frac{a^2}{2\sigma^2} T \right\}. \quad (2.4)$$

Here $L(T, 0)$ denotes the local time for the process X at 0 and is defined as

$$L(T, 0) = \lim_{\epsilon \downarrow 0} \frac{1}{2\epsilon} \int_0^T I(|X(s)| < \epsilon)\, \mathrm{d}s$$

(see, for example, [3, p. 52]).

The likelihood ratio for the two models is therefore

$$LR(X, T) = \log \frac{p_0(X, T)}{p_1(X, T)} = aL(T, 0) - \frac{b}{\sigma} 2N(T, h)h \\ - \frac{T}{2\sigma^2}(a^2 - b^2) + R(X, T),$$

where $R(X, T)$ denotes the remainder containing the terms $X(T)$ and $|X(T)|$. We reject the null hypothesis that the observed trajectory comes from model 0, if $LR(X, T) < k(\alpha, T)$, where $k(\alpha, T)$ is chosen such that $\alpha = P_0(LR(X, T) < k(\alpha, T))$. As we test a simple hypothesis versus a simple alternative, the corresponding test is optimal according to Neyman–Pearson Lemma.

Next we show how to determine $k(\alpha, T)$ for large values of T. Let $l_i(0) := L(\tau_{2i}, 0) - L(\tau_{2i-2}, 0), i = 1, \ldots, N(T, h)$, be local time at zero in the ith cycle, and let $d_i = \tau_{2i} - \tau_{2i-2}$, for $i \geq 1$, denote the duration of the ith cycle. By strong Markov property, the vectors $(d_i, l_i(0))$, $i \geq 1$, form an iid sequence under both models. As $L(T, 0)$ is an additive functional, it has the representation $L(T, 0) = \sum_{i=1}^{N(T,h)} l_i(0) + R(0)$, where $R(0) = L(T, 0) - L(\tau_{N(T, h)}, 0)$ denotes the remainder term. The likelihood ratio statistic can now be written in the form

$$LR(X,T) = a\sum_{i=1}^{N(T,h)} l_i(0) - \frac{b}{\sigma}2N(T,h)h - \frac{T}{2\sigma^2}(a^2-b^2) - R(X,T).$$

We show that a central-limit-theorem (CLT) type argument can be applied to get the limiting distribution for the properly standardized statistic $LR(X,T)$. The following theorem will be used to find the cutoff point of the likelihood ratio test statistic.

THEOREM 2.1 *Assume that* $\mathbb{E}\,d_1^2 < \infty$ *and* $\mathbb{E}\,l_1^2 < \infty$. *Let* $\lambda = (\mathbb{E}d_1)^{-1}$. *Then*

$$T^{-1/2}\left(N(T,h) - \lambda T,\ \sum_{i=1}^{N(T,h)} (l_i(0) - \mathbb{E}\,l_i(0))\right) \xrightarrow{\mathcal{D}} N(0,\Sigma^*), \tag{2.5}$$

where $\Sigma^*_{11} = \lambda^3 \operatorname{Var} d_1$, $\Sigma^*_{12} = -\lambda^2 \operatorname{Cov}(d_1, l_i(0))$, $\Sigma^*_{22} = \lambda \operatorname{Var} l_1(0)$, *and*

$$T^{-1/2}\left(N(T,h) - \lambda T,\ \sum_{i=1}^{N(T,h)} l_i(0) - \lambda T\,\mathbb{E}\,l_i(0)\right) \xrightarrow{\mathcal{D}} N\big(0,\tilde{\Sigma}\big), \tag{2.6}$$

where $\tilde{\Sigma}_{11} = \Sigma^*_{11}$, $\tilde{\Sigma}_{12} = \Sigma^*_{12} + \mathbb{E}l_1(0)\Sigma^*_{11}$, $\tilde{\Sigma}_{22} = \Sigma^*_{22} + (\mathbb{E}l_1(0))^2\Sigma^*_{11} + 2\mathbb{E}l_1(0)\Sigma^*_{12}$.

Using Lemmas 1 and 2 from the next section we can show that the remainder term $R(X,T)$ goes to zero under the proper normalization.

Applying this CLT type result, we find the cutoff point for the test statistic through the relation

$$\begin{aligned}\alpha &= \lim_{T\to\infty} P_0\left(\frac{1}{\sigma_{LR}T^{1/2}}(LR(X,T) - \mathbb{E}_0 LR(X,T)) < z_\alpha\right)\\ &= \lim_{T\to\infty} P_0(LR(X,T) < z_\alpha\sigma_{LR}T^{1/2} + \mathbb{E}_0 LR(X,T)),\end{aligned}$$

where z_α denotes the α quantile of the standard normal distribution.

Thus $\lim_{T\to\infty} P_0(LR(X,T) < k(\alpha,T)) = \alpha$ for

$$k(\alpha,T) = \sigma_{LR} z_\alpha T^{1/2} + \mathbb{E}_0 LR(X,T).$$

The expectation is taken to be

$$\mathbb{E}LR(X,T) = \lambda T(a\mathbb{E}_0 l_1(0) - 2bh/\sigma) + \frac{T}{2\sigma^2}(a^2 - b^2),$$

as the remainder term is of negligible order by Lemmas 1 and 2. The theorem also provides us with the variance of $LR(X,T)$:

$$\begin{aligned}\sigma^2_{LR} &= 2ab\lambda^2/\sigma \operatorname{Cov}(d_1, l_1(0))(2abh/\sigma - a^2\mathbb{E}l_1(0)) \\ &\quad + \lambda^3 b^2/\sigma^2 \operatorname{Var} d_1(4b^2h^2/\sigma^2 - 4abh/\sigma\mathbb{E}l_1(0) \\ &\quad + (a\mathbb{E}l_1(0))^2) + \lambda a^2 \operatorname{Var} l_1(0),\end{aligned}$$

where the expectations and λ are calculated under the model of the null hypothesis.

We notice that for large values of T, $k(\alpha, T, h)$ is monotone increasing in h. Therefore if $LR(X,T) < k(\alpha,T,h_0)$ is satisfied, then $LR(X,T) < k(\alpha,T,h)$ holds for any $h > h_0$. We can now extend the test to a composite alternative. To test if the observed sample path comes from the diffusion model, H_0: $p = p_0$, or results from a process with hysteresis, H_1: $p = p_1$ with $h \geq h_0$, we thus take $k(\alpha, T, h_0)$ as the cutoff point for the level α test.

Generally though, the drift coefficients a and b cannot be assumed to be known and have to be replaced by a properly chosen estimates. Note that p_0 and p_1, the densities in the hypotheses, are not in the same parametric family, that is, the density $p_1(X,T,b)$ cannot be approximated arbitrarily close by $p_0(X,T,b)$. Cox [4, 5] showed how to treat this situation of non-tested hypotheses in the case of a classical iid sample. He proposes to replace the parameter in each density by its maximum likelihood estimate (MLE) under the specific model and calculate the expectation of the likelihood statistic using the estimate under the null hypothesis and the value to which the estimate under the alternative converges. We adapt this procedure for our situation. Assume for simplicity that $\sigma = 1$. Let $\hat{a}$ denote the MLE under the model of the null hypothesis:

$$\hat{a} = 1/T \sum_{i=1}^{N(T,h)} l_i(0) + o(T^{1/2}),$$

and $\hat{b}$ the MLE under the model of the alternative:

$$\hat{b} = 2hN(T,h)/T + o(T^{1/2}).$$

Asymptotic normality of the estimates under both models follows immediately from the theorem. Then

$$LR(X,T,\hat{a},\hat{b}) = \frac{1}{2}\left[\frac{\left(\sum_{i=1}^{N(T,h)} l_i(0)\right)^2}{T} - \frac{(2hN(T,h))^2}{T}\right] + o(T^{1/2}).$$

The test statistic that Cox proposes is

$$LR(X,T,\hat{a},\hat{b}) - \mathbb{E}_{0,\hat{a}}LR(X,T,\hat{a},\hat{b}).$$

We first find $\mathbb{E}_{0,a}LR(X,T,\hat{a},\hat{b})$ and then replace a by its estimator.

In the following we write $\lambda(a)$ instead of λ wherever it seems necessary to stress the dependence of λ on the drift parameter of the null hypothesis. Note that by application of the continuous mapping theorem (see, for example, [2, p. 31, Corollary 1]), with $f(x,y) = -4h^2x^2 + y^2$ we get that

$$1/T^{1/2}[LR(X,T,\hat{a},\hat{b}) - \mathbb{E}_{0,a}LR(X,T,\hat{a},\hat{b})] = T^{1/2}\left(\left(\sum_{i=1}^{N(T,h)} l_i(0)/T\right)^2 - (2hN(T,h)/T)^2 - (\lambda\mathbb{E}l_1(0))^2 + (2h\lambda)^2\right) \xrightarrow{\mathcal{D}} N(0,\sigma^2),$$

where σ^2 can be calculated from the theorem. We now replace the parameter in the expectation by its estimator under the null hypothesis. This results in

$$\begin{aligned} &T^{-1/2}[LR(X,T,\hat{a},\hat{b}) - \mathbb{E}_{0,\hat{a}}LR(X,T,\hat{a},\hat{b})] \\ &\quad = 4h^2T^{1/2}(\lambda(\hat{a})^2 - N(T,h)^2/T^2). \end{aligned}$$

To find the variance of the above expression note that

$$\begin{aligned} &T^{1/2}(\lambda(\hat{a})^2 - N(T,h)^2/T^2) \\ &\quad = T^{1/2}(\lambda(\hat{a})^2 - \lambda(a)^2 + \lambda(a)^2 - N(T,h)^2/T^2). \end{aligned}$$

Thus by applying the continuous mapping theorem again with $f(x,y) = -x^2 + \lambda^2(y)$ and noting that $\mathbb{E}\hat{a} = \lambda(a)\mathbb{E}l_1 = a$ under the null hypothesis, we get that

$$T^{1/2}(\lambda^2(\hat{a}) - N(T,h)^2/T^2) \xrightarrow{\mathcal{D}} N(0,\sigma^2(a)),$$

where

$$\sigma^2(a) = 4\lambda^2(a)(\lambda'(a))^2\tilde{\Sigma}_{22}(a) + 4\lambda^2(a)\tilde{\Sigma}_{11}(a) - 4\lambda^2(a)\lambda'(a)\tilde{\Sigma}_{12}(a).$$

The expressions for the $\tilde{\Sigma}_{ij}$ are stated in the theorem.

For practical purposes $\sigma^2(a)$ has to be replaced by its consistent estimate $\sigma^2(\hat{a})$. Now we find the cutoff point for the test statistic and reject the null hypothesis, if $\lambda^2(\hat{a})T - N(T,h)^2 < z_\alpha T^{1/2}\sigma(\hat{a})$.

Atkinson [1] suggests a modification of the method by Cox for iid data: he proposes to replace $\hat{b}$ by its expected value under the null hypothesis. As $\mathbb{E}_0\hat{b} = 2h\lambda(a)$ the estimator for b in the likelihood ratio statistic is $2h\lambda(\hat{a})$. Atkinson shows that the bias for $LR(X,T,\hat{a},b(\hat{a})) - \mathbb{E}_{0,a}LR(X,T,\hat{a},b(\hat{a}))$ is less than the bias for Cox's test statistic $LR(X,T,\hat{a},\hat{b}) - \mathbb{E}_{0,a}LR(X,T,\hat{a},\hat{b})$. Asymptotically both test statistics are equivalent. We did not investigate this approach here, though a comparison of the performance of both methods is desirable and should be included in future work.

3 SIMULATION RESULTS

In this section we consider applications of the likelihood ratio test to simulated data. We begin by discussing an algorithm for producing realizations of the stochastic process with a discontinuous drift component. After describing the algorithm we assess the performance of our test for different combinations of the parameters.

In order to simulate realizations of the process with the jump, we have to find its transition density. Following Karatzas and Shreve [9, Chapter 6], the transition density is found to be

$$p(x,z,t) = \begin{cases} (2\pi t)^{-1/2}\Big[\exp(-(x-z-at)^2/2t) + a\exp(-2az) \\ \qquad \times \int_{x+z}^{\infty}\exp-(v-at)^2/2t)\,\mathrm{d}v\Big], \quad x \geq 0,\ z \geq 0; \\ (2\pi t)^{-1/2}\Big[\exp(2ax - (x-z-at)^2/2t) + a\exp(2az) \\ \qquad \times \int_{x-z}^{\infty}\exp-(v-at)^2/2t)\,\mathrm{d}v\Big], \quad x \geq 0,\ z \leq 0. \end{cases} \tag{3.1}$$

Also notice that the symmetry of the process yields

$$p(x, z, t) = p(-x, -z, t).$$

Thus for $x \leq 0$ the transition density is easily obtained from (3.1). The density of the stationary distribution can also be derived from (3.1) by letting $t \to \infty$ and equals

$$\pi(z) = \begin{cases} a\exp(-2az), & z \geq 0, \\ a\exp(2az), & z \leq 0. \end{cases} \tag{3.2}$$

Let x_i denote the discrete sample points of $X(t)$ for $t = t_0, \ldots, t_n$, where $t_i = Ti/n$, $i = 0, \ldots, n$. After the initial value x_0 is found using the stationary distribution (3.2), the simulations are performed in three steps: first we calculate $p_i = \mathbb{P}(z > 0 \,|\, x_i)$. Then we simulate a uniform $[0, 1]$ random variable and depending if it is smaller or larger than p_i we choose the negative or the positive branch of the transition density for the calculations. We then simulate another random variable uniform on either $(0, p_i)$ or $(0, 1 - p_i)$ and calculate the corresponding quantile from the transition density by Newton–Raphson method. The quantile gives us x_{i+1}.

A question of importance is the following: can values of local time for Brownian motion (and therefore for other one-dimensional diffusion processes) be estimated from observations in discrete time? A positive answer to this question was given by Csörgö and Revesz [6]. They discuss several estimators for Brownian local time $L(x, t)$ for $t \in [0, 1]$ and $x \in R$, but we only present the estimator we actually used in the simulations. Let

$$\hat{L}_n(x, t) := \frac{1}{2n^{1/2}} \sum_{k=1}^{[nt]} I\{W(k/n) \in [x - n^{-1/2}, x + n^{-1/2})\}.$$

Then the following result holds:

THEOREM 3.1 (Csörgö, Revesz) *For any $\epsilon > 0$ with probability one we have*

$$\sup_{(t,x) \in [0,1] \times R} |\hat{L}_n(x, t) - L(t, x)| = o(n^{-1/4+\epsilon}).$$

TABLE I Number of correct decisions (rejections), true model: hysteresis, $h = 0.2$

b	$T = 40, \alpha = 0.05$	$T = 40, \alpha = 0.1$	b	$T = 60, \alpha = 0.05$	$T = 60, \alpha = 0.1$
1.6	31/100	48/100	1.5	12/100	22/100
1.7	72/100	88/100	1.6	61/100	81/100
1.8	97/100	100/100	1.7	94/100	99/100
2.0	100/100	100/100	1.8	100/100	100/100

The extension to $t \in [0, T]$ and $X(t)$ follows immediately.

Simulation results showed that the prescribed significance level was reached for $T \geq 40$ and for $b = 1$ and $\sigma^2 = 1$.

Table I gives the number of correct decisions, that is correct rejections of the null hypothesis, out of 100 repetitions fo the test when the data are generated by a model with hysteresis. We see that the results are sensitive to the size of the drift. This is not surprising because the larger the drift, the more the cycles for the model with hystersis we observe for a fixed time interval. The parameters in the models were replaced by their maximum likelihood estimates. For the simulation results presented in the table σ was chosen to be one.

4 DETAILS AND PROOFS

Before we prove the theorem we calculate the moments of $(d_1, l_1(0))$ under each model using Laplace transforms to show that the assumption in the theorem is satisfied and to find the moments of $LR(X, T)$.

Notice that the absolute value of the process with a two-valued a or $-a$, is equivalent to a diffusion with drift $-a$ on $[0, \infty)$ and reflection at zero. A cycle for the process with reflection is defined as the time interval between two successive intersections of the trajectory with the level $h/2$ separated by touching zero.

We start by investigating the behavior of the process on a half cycle, i.e. the time it takes for a trajectory that starts in $h/2$ to touch 0 and go back to $h/2$. To get the duration of the full cycle, because of the symmetry properties of the process we just double this amount. For the first half of the half-cycle (the time interval between intersection $h/2$ and the first time τ_1 the trajectory reaches zero), the local time process equals zero. The Laplace transform for τ_1 is

$$\mathbb{E}_2 \exp(-\alpha\tau_1) = \exp(ah/2\sigma^2 - h/2(2\alpha/\sigma^2 + a^2/\sigma^4)^{1/2}).$$

The joint Laplace transform for local time and $\tau_2 = \inf\{t \geq 0\colon X(t) = h/2, X(0) = 0\}$ is given by

$$\mathbb{E}_2 \exp(-\sigma\tau_2 - \gamma l_1(0)) = N/D, \tag{4.1a}$$

where

$$\begin{aligned} N &= 2(2\alpha + a^2/\sigma^2)^{1/2}, \\ D &= \exp(ha/2\sigma^2 - h/2(2\alpha + a^2/\sigma^2)^{1/2})(a/\sigma^2 + (2\alpha + a^2/\sigma^2)^{1/2} - \gamma) \\ &\quad + \exp(ha/2\sigma^2 + h/2(2\alpha + a^2/\sigma^2)^{1/2})(\gamma - a/\sigma^2 + (2\alpha + a^2/\sigma^2)^{1/2}). \end{aligned} \tag{4.1b}$$

Therefore the joint Laplace transform for $d_1 = 2(\tau_1 + \tau_2)$ and $l_1(0)$ is given by

$$\mathbb{E}_2 \exp(-\alpha d_1 - \gamma l_1(0)) = \mathbb{E}_2 \exp(-\alpha 2\tau_1)\mathbb{E}_2 \exp(-\alpha 2\tau_2 - \gamma l_1(0)).$$

This provides us with the moments for d_1 and $l_1(0)$:

$$\begin{aligned} \mathbb{E} l_1(0) &= \sigma^2(\exp(ha/\sigma^2) - 1)/2a, \\ \operatorname{Var} l_1(0) &= \sigma^4(\exp(ha/\sigma^2) - 1)^2/4a^2, \\ \mathbb{E}\, d_1(0) &= \sigma^2(\exp(ha/\sigma^2) - 1)/a^2, \\ \operatorname{Var} d_1 &= \frac{\sigma^4}{a^4}[\exp(2ha/\sigma^2) - 5 + 4\exp(ha/\sigma^2)] - \frac{4h\sigma^2}{a^3}\exp(ha/\sigma^2), \\ \operatorname{Cov}(d_1, l_1(0)) &= \frac{\sigma^4}{2a^3}[\exp(2ha/\sigma^2) - 1 - 3hb\exp(ha/\sigma^2)/\sigma^2 + ah/\sigma^2]. \end{aligned} \tag{4.2}$$

The Laplace transform (4.1) is obtained as follows. Let $u(x)$ be the solution of

$$\begin{aligned} &Lu(x) - \alpha u(x) = 0 \quad \text{for } x \in (0, h/2) \\ &u'(0) = \gamma u(0), \qquad u(h/2) = 1. \end{aligned} \tag{4.3}$$

Then $u(x)$ satisfies $u(x)=\mathbb{E}_x u(X(\tau))\exp(Y(\tau)+Z(\tau))=\mathbb{E}_x \exp(Y(\tau)+Z(\tau))$, where $\tau=\inf\{t\geq 0: X(t)=h/2,\ X(0)=0\}$ and X is a diffusion process with operator $L=\sigma^2/2(\mathrm{d}^2/\mathrm{d}x^2)-a(\mathrm{d}/\mathrm{d}x)$ on $[0,\infty)$ and reflection at zero. This process can also be written in the form

$$\mathrm{d}X(t)=\sigma\,\mathrm{d}W(t)-a\,\mathrm{d}t+I(X(t)=0)\,\mathrm{d}L(t,0),\quad X(0)=x,\ L(0,0)=0,$$

where $L(t,0)$ is the local time of the process at zero, the boundary of the domain (see [7, p. 87]). Let $Y(t)=-\alpha t$ and $Z(t)=-\gamma L(t,0)$ with $\alpha,\gamma>0$. Applying the generalized Ito formula (see e.g. [7, p. 96]) to the function $f(x,y,z)=u(x)\exp(y+z)$, we obtain

$$\begin{aligned} &u(X(t))\exp(Y(t)+Z(t))-u(x)\\ &=\int_0^t \sigma\exp(Y(s)+Z(s))u'(X(s))\,\mathrm{d}W(s)\\ &+\int_0^t \exp(Y(s)+Z(s))[u'(X(s))-\gamma u(X(s))]\,\mathrm{d}L(s,0)\\ &+\int_0^t \exp(Y(s)+Z(s))[Lu(X(s)-\alpha u(X(s))]\,\mathrm{d}s. \end{aligned}$$

As u satisfies (4.3) the above expression reduces to

$$u(X(t))\exp(Y(t)+Z(t))-u(x)=\int_0^t \sigma\exp(Y(s)+Z(s))u'(X(s))\,\mathrm{d}W(s).$$

After replacing t by τ and taking expectation, we obtain that

$$u(x)=\mathbb{E}_x\exp(Y(\tau)+Z(\tau))=\mathbb{E}_x\exp(\sigma\tau-\gamma L(\tau,0)).$$

Thus solving (4.3) and taking $x=0$ results in expression (4.1) for the Laplace transform. Under the hysteresis model the vector $(d_i, l_i(0))$ has the joint Laplace transform

$$\begin{aligned} &\mathbb{E}_1\exp(-\alpha d_1-\gamma l_1(0))\\ &=\frac{\exp(2hb/\sigma^2-2h(2\alpha/\sigma^2+b^2/\sigma^2)^{1/2})(2\alpha/\sigma^2+b^2/\sigma^2)}{((2\alpha/\sigma^2+b^2/\sigma^2)^{1/2}+\gamma(1-\exp(-h(2\alpha/\sigma^2+b^2/\sigma^2)^{1/2})))^2}. \end{aligned} \tag{4.4}$$

Thus the moments for d_1 and $l_1(0)$ are

$$\begin{aligned}\mathbb{E}l_1(0) &= 2(1-\exp(-hb/\sigma^2))/b,\\ \operatorname{Var} l_1(0) &= 2\sigma^4(1-\exp(-hb/\sigma^2))^2/b^2,\\ \mathbb{E}\, d_1 &= 2h/b, \qquad \operatorname{Var} d_1 = 2\sigma^2 h/b^3,\\ \operatorname{Cov}(d_1, l_1(0)) &= 2\sigma^2(1-\exp(-hb/\sigma^2) - 2hb\exp(-hb/\sigma^2))/b^3.\end{aligned} \tag{4.5}$$

We are now prepared to prove the joint CLT for the number of cycles and local time at zero in $[0,T]$.

Proof of Theorem 2.1 As the $(d_i, l_i(0))$ form an iid sequence with finite second moments, we have the classical CLT result,

$$T^{-1/2}\left(\sum_{i=1}^{[\lambda T]} d_i - T, \sum_{i=1}^{[\lambda T]}(l_i - \mathbb{E}l_i)\right) \xrightarrow{\mathcal{D}} N(0, \lambda\Sigma), \tag{4.6}$$

where $\Sigma_{11} = \operatorname{Var} d_1$, $\Sigma_{12} = \operatorname{Cov}(d_1, l_1(0))$, $\Sigma_{22} = \operatorname{Var} l_1(0)$.

By [10, Theorem 2],

$$T^{-1/2}\left(\left(\sum_{i=1}^{[\lambda T]} d_i - T\right) - \left(T - \frac{1}{\lambda}N(T,h)\right)\right) \xrightarrow{\mathcal{D}} 0. \tag{4.7}$$

As convergence in distribution to a constant implies convergence in probability, we can apply [2, Theorem 4.1, p. 25] and obtain

$$T^{-1/2}\left(T - \frac{1}{\lambda}N(T,h), \sum_{i=1}^{[\lambda T]}(l_i - \mathbb{E}l_i)\right) \xrightarrow{\mathcal{D}} N(0, \lambda\Sigma). \tag{4.8}$$

By combining (4.6) and (4.7) and applying [10, Theorem 7], which states that

$$T^{-1/2}\left(\left(\sum_{i=1}^{[\lambda T]}(l_i - \mathbb{E}l_i)\right) - \sum_{i=1}^{N(T,h)}(l_i - \mathbb{E}l_i)\right) \xrightarrow{\mathcal{D}} 0$$

and [2, Theorem 4.1] once more, we obtain that

$$T^{-1/2}\left(T - \frac{1}{\lambda}N(T,h),\ \sum_{i=1}^{N(T,h)} (l_i - \mathbb{E}l_i)\right) \xrightarrow{\mathcal{D}} N(0, \lambda\Sigma).$$

The final result for the first part of the theorem is given by

$$T^{-1/2}\left(N(T,h) - \lambda T,\ \sum_{i=1}^{N(T,h)} (l_i - \mathbb{E}l_i)\right) \xrightarrow{\mathcal{D}} N(0, \Sigma^*), \tag{4.9}$$

where the entries Σ^* are given in the statement of the theorem.

As

$$\begin{aligned} T^{-1/2}&\left(N(T,h) - \lambda T,\ \sum_{i=1}^{N(T,h)} l_i(0) - \lambda T\mathbb{E}l_i(0)\right) \\ &= T^{-1/2}\left(N(T,h) - \lambda T,\ \sum_{i=1}^{N(T,h)} (l_i(0) - \mathbb{E}l_1(0))\right. \\ &\qquad \left. + \mathbb{E}l_1(0)(N(T,h) - \lambda T)\right), \end{aligned}$$

we apply the continuous mapping theorem (see e.g. [2, p. 31, Corollary 1]), with $h(x,y) = (x, y + x)$, and (2.5) to get the second statement of the theorem.

To show that the remainder term $R(X,T)$ in the likelihood ratio statistic goes to zero in probability, we use the following two lemmas.

Lemma 1 *Under both models* $|X(T)|/T^{1/2} \to 0$ *in probability for* $T \to \infty$.

Proof Recall that for the process with hysteresis $\mathrm{d}X(t) = b_{i(t)}\,\mathrm{d}t + \sigma\,\mathrm{d}W(t)$, $X(0) = x_0 < 0$, where $i(t) = 0$ or 1. Therefore

$$\begin{aligned} |X(T)| &= \left|\int_0^T b_{i(t)}\,\mathrm{d}t + \sigma W(T) + x\right| = \left|\sum_{i=0}^{N(T,h)} (X(\tau_{2i+2}) - X(\tau_{2i}))\right. \\ &\quad \left. + \int_{\tau_{N(T,h)}}^T b_{i(t)}\,\mathrm{d}t + \sigma W(T) - \sigma W(\tau_{N(T,h)})\right| \\ &\le b(T - \tau_{N(T,h)}) + \sigma|W(T) - W(\tau_{N(T,h)})| \end{aligned}$$

as $X(\tau_{2i+2}) = X(\tau_{2i}) = x_0$.

Analogously we have for the process with a drift change in zero that

$$
\begin{aligned}
|X(T)| &= \Bigg| \sum_{i=0}^{N(T,h)} (X(\tau_{2i+2}) - X(\tau_{2i})) \\
&\quad + \int_{\tau_{N(T,h)}}^{T} a(1 - 2I(X(t) > 0))\,\mathrm{d}t + \sigma W(T) - \sigma W(\tau_{N(T,h)}) \Bigg| \\
&\le a(T - \tau_{N(T,h)}) + \sigma |W(T) - W(\tau_{N(T,h)})|
\end{aligned}
$$

as here also $X(\tau_{2i+2}) = X(\tau_{2i}) = x_0$.

Note that $0 < T - \tau_{2N(T,h)} \le d_{N(T,h)+1}$, where d denotes the duration of the $(N(T,h)+1)$th cycle, and, as we showed before, $\mathbb{E}_i d < \infty$ under the model with hysteresis ($i = 1$) and under the model with a drift change in zero ($i = 0$). This leads to the inequality

$$
P_i((T - \tau_{2N(T,h)})/T^{1/2} \ge \epsilon) \le P_i(d_{N(T,h)+1}/T^{1/2} \ge \epsilon) \le \frac{\mathbb{E}_i d}{\epsilon T^{1/2}} \to 0
$$

for $T \to \infty$ for $i = 0, 1$; $W(T) - W(\tau_{N(T,h)}) \stackrel{\mathcal{D}}{=} W(T - \tau_{N(T,h)})$ and

$$
\begin{aligned}
&P_i(|W(T - \tau_{2N(T,h)})|/T^{1/2} \ge \epsilon) \\
&\quad = \int_0^\infty P_i(|W(s)|/T^{1/2} \ge \epsilon \mid T - \tau_{2N(T,h)})\,\mathrm{d}P_{i,T-\tau_{2N(T,h)}}(s) \\
&\quad \le \frac{1}{T\epsilon} \int_0^\infty s\,\mathrm{d}P_{i,T-\tau_{2N(T,h)}}(s) \le \frac{1}{T\epsilon} \mathbb{E}_i(T - \tau_{2N(T,h)}) \\
&\quad \le \frac{1}{T\epsilon} \mathbb{E}_i\, d \to 0, \quad i = 0, 1.
\end{aligned}
$$

Combining these two results we get that $|X(T)|/T^{1/2} \to 0$ in probability under both models.

LEMMA 2 $(L(T,0) - L(\tau_{N(T,h)}, 0))/T^{1/2} \to 0$ *in probability for* $T \to \infty$ *under both models.*

Proof As local time is nondecreasing in the time argument, $0 \le L(T,0) - L(\tau_{N(T,h)}, 0) \le L(\tau_{N(T,h)+1}, 0) - L(\tau_{N(T,h)}, 0) = l_{N(T,h)+1}(0)$. Thus the expectation of the remainder term is bounded by the expectation of $l_{N(T,h)+1}$, which is finite under both models. Therefore

$$P_i((L(T,0) - L(\tau_{N(T,h)}))/T^{1/2} \geq \epsilon) \\ \leq \frac{\mathbb{E}_i l_{N(T,h)+1}(0)}{\epsilon T^{1/2}} \to 0 \quad \text{for } T \to \infty, \quad i = 0, 1.$$

5 HYSTERESIS VERSUS ORNSTEIN – UHLENBECK PROCESS

In this section we present a criterion to determine if the observed trajectory comes from a process with hysteresis with the thresholds symmetric with respect to $x^* = 0$, or an Ornstein–Uhlenbeck process X, referred to as 'model 3', that satisfies the equation

$$\mathrm{d}X(t) = -aX(t)\,\mathrm{d}t + \sigma\,\mathrm{d}W(t), \quad X(0) = x < x_0. \tag{5.1}$$

To derive the criterion we follow the steps for the case of testing hysteresis versus a Markov diffusion process with a drift with a discontinuity. The Radon–Nikodym derivative of the measure corresponding to the Ornstein–Uhlenbeck process X on $[0, T]$ with respect to the measure for the process $\tilde{X}(t) = \sigma W(t), \tilde{X}(0) = x_0$, is obtained by application of Girsanov's theorem and is given by

$$p_3(X,T) = \frac{\mathrm{d}\mu_X}{\mathrm{d}\mu_{\tilde{X}}}(X) = \exp\left\{\int_0^T -\frac{a}{\sigma^2}X(t)\,\mathrm{d}X(t) - \frac{a^2}{2\sigma^2}\int_0^T X^2(t)\,\mathrm{d}t\right\}$$
$$= \exp\left\{\int_0^T -\frac{a}{\sigma}X(t)\,\mathrm{d}W(t) + \frac{a^2}{2\sigma^2}\int_0^T X^2(t)\,\mathrm{d}t\right\}.$$

Note that application of Ito's formula to the function $f(x) = x^2$ and Ornstein–Uhlenbeck process yields

$$X^2(T) - x^2 = -2a\int_0^T X^2(s)\,\mathrm{d}s + T\sigma^2 + 2\sigma\int_0^T X(s)\,\mathrm{d}W(s). \tag{5.2}$$

Using the above relation, we can find an expression for the stochastic integral, substitute it in the Radon–Nikodym derivative and get

$$p_3(X,T) = \exp\left\{-\frac{a^2}{2\sigma^2}\int_0^T X^2(t)\,\mathrm{d}t + Ta/2 - a(X^2(T) - x^2)/2\sigma\right\}.$$

The likelihood ratio statistic $LR(X,T) = \log(p_3(X,T)/p_1(X,T))$ for the Ornstein–Uhlenbeck process and the process with hysteresis is therefore

$$LR(X,T) = -\frac{a^2}{2\sigma^2}\int_0^T X^2(t)\,\mathrm{d}t - 2\frac{b}{\sigma}N(T,h)h + \frac{a}{2}T + \frac{b^2T}{2\sigma^2} + R(X,T), \tag{5.3}$$

where $R(X,T)$ denotes the remainder term that contains the constants and the terms $X(T)$ and $X^2(T)$.

Let $Y_i = \int_{\tau_{2i-2}}^{\tau_{2i}} X^2(t)\,\mathrm{d}t$ for $i = 1,\ldots,N(T,h)$, where the τ_i are defined as in (2.1). Then

$$LR(X,T) = -\frac{a^2}{2\sigma^2}\sum_{i=1}^{N(T,h)} Y_i - 2\frac{b}{\sigma}N(T,h)h + \frac{a}{2}T + \frac{b^2T}{2\sigma^2} + R(X,T).$$

The following CLT type result holds, that can be proven in the exactly same manner as in Section 2.

PROPOSITION *Let* $\lambda = (\mathbb{E}d_1)^{-1}$. *If* $\mathbb{E}Y_1 < \infty$ *and* $\mathbb{E}\tau_1 < \infty$ *then*

$$T^{-1/2}\left(N(T,h) - \lambda T,\ \sum_{i=1}^{N(T,h)} Y_i(0) - \lambda T\mathbb{E}Y_i(0)\right) \xrightarrow{\mathcal{D}} N(0,\tilde{\Sigma}).$$

The entries of the variance covariance matrix are $\tilde{\Sigma}_{11} = \lambda^3 \mathrm{Var}\, d_1$, $\tilde{\Sigma}_{12} = \lambda^3\,\mathrm{Var}\, d_1 \mathbb{E}Y_1 - \lambda^2\,\mathrm{Cov}(d_1, Y_1)$ *and* $\tilde{\Sigma}_{22} = \lambda\,\mathrm{Var}\, Y_1 + \lambda^3 \times \mathrm{Var}\, d_1(\mathbb{E}Y_1)^2 - 2\lambda^2\,\mathbb{E}Y_1\,\mathrm{Cov}(d_1, Y_1)$.

Remark An expression for the expectation of Y_1 in terms of the first moment of d_1 can be derived directly from (5.2), as for $\tau_1 = \inf\{t \geq 0: X(t) = h/2,\ X(0) = -h/2\}$ we have $\mathbb{E}_3\int_0^{\tau_1} X^2(s)\,\mathrm{d}s = (\sigma^2/2a)\mathbb{E}_3\tau_1$. Because of symmetry of the Ornstein–Uhlenbeck process around zero it follows that τ_1 is equal in distribution to $\tau_2 - \tau_1$ and thus $\mathbb{E}_3 d_1 = \mathbb{E}_3(\tau_2 - \tau_1) + \mathbb{E}_3\tau_1 = 2\mathbb{E}_3\tau_1$. Therefore the following relation holds:

$$\mathbb{E}_3 Y_1 = \mathbb{E}_3\int_0^{d_1} X^2(s)\,\mathrm{d}s = \frac{\sigma^2}{a}\mathbb{E}_3\tau_1 = \frac{\sigma^2}{2a}\mathbb{E}_3 d_1.$$

Squaring (5.2) and taking expectations also yields a bound for the second moment of Y_1 in terms of the second moment of τ_1. Thus the variance of Y_1 and the covariance of Y_1 and τ_1 are well defined.

We reject the hypothesis that the observed trajectory comes from an Ornstein–Uhlenbeck process if $LR(X,T) < k(\alpha,T)$. The cutoff point can now be found through application of the proposition and equals

$$k(\alpha,T) = \sigma_{LR} z_\alpha T^{1/2} + \mathbb{E}_3 LR(X,T),$$

where z_α denotes the α-quantile of the standard normal distribution.

As in the case with the discontinuous drift, we now replace the parameter in each model with its maximum likelihood estimate under the specific model. Recall that under the hysteresis model

$$\hat{b} = 2hN(T,h)/T + o(T^{1/2}).$$

Let $\hat{a}$ denote the maximum likelihood estimate under the Ornstein–Uhlenbeck process. Straightforward calculation yields

$$\hat{a} = \frac{\sigma^2 T}{2\sum_{i=1}^{N(T,h)} Y_i} + o(T^{1/2}).$$

The likelihood ratio statistic with the parameter estimates in place of the parameters equals

$$LR(X,T,\hat{a},\hat{b}) = \frac{1}{2}\left[\frac{\sigma^2 T^2}{4\sum_{i=1}^{N(T,h)} Y_i} - \frac{(2hN(T,h))^2}{T}\right] + o(T).$$

Application of the continuous mapping theorem with $f(x,y) = -4h^2x^2 + \sigma^2/4y$ yields

$$\begin{aligned} &T^{-1/2}[LR(X,T,\hat{a},\hat{b}) - \mathbb{E}_{3,a} LR(X,T,\hat{a},\hat{b})] \\ &= T^{1/2}\left(\frac{T\sigma^2}{4\sum_{i=1}^{N(T,h)} Y_i} - (2hN(T,h)/T)^2 - \frac{\sigma^2}{4\lambda \mathbb{E} Y_1} + (2h\lambda)^2\right) \\ &\xrightarrow{\mathcal{D}} N(0,\sigma^2(a)). \end{aligned}$$

Replacing the parameter a in the expectation by its estimator under the null hypothesis and taking into account that $\mathbb{E}Y_1 = \sigma^2/(2a\lambda(a))$, we have

$$T^{-1/2}[LR(X,T,\hat{a},\hat{b}) - \mathbb{E}_{3,\hat{a}}LR(X,T,\hat{a},\hat{b})] = T^{1/2}[\lambda^2(\hat{a}) - N^2(T,h)/T^2.$$

An expression for the variance can be obtained by application of the continuous mapping theorem with $f(x,y) = -x^2 + \lambda^2(y)$.

In order to determine the cutoff point explicitly, we have to find σ_{LR} and $\mathbb{E}_3 LR(X,T)$, which means that we have to find the moments of (d_1, Y_1) under the measure of the Ornstein–Uhlenbeck process. This is done by calculating the Laplace transform of $(\tau_1, \int_0^{\tau_1} X^2(s)\,\mathrm{d}s)$ as follows.

Let $u(x)$ be the solution of

$$\begin{aligned} &Lu(x) - (\alpha x^2 + \beta)u(x) = 0 \quad \text{for } x \in (-h/2, h/2),\\ &u(h/2) = 1, \end{aligned} \tag{5.4}$$

where $L = \sigma^2/2(\mathrm{d}^2/\mathrm{d}x^2) - ax(\mathrm{d}/\mathrm{d}x)$ is the generator of the Ornstein–Uhlenbeck process. Then $u(x)$ satisfies $u(x) = \mathbb{E}_x \exp(-\alpha\int_0^{\tau_1} X^2(s)\,\mathrm{d}s - \beta\tau_1)$, and for $x = -h/2$ we get the required result. One of the complications in this case is that the solution of the system (5.4) cannot be given in a closed form.

With the transformation $u(x) = y(x)\exp((b/2\sigma^2)x^2)$, (5.4) becomes

$$\begin{aligned} &y''(x) - y(x)[a/\sigma^2 - 2\beta/\sigma^2 - x^2(a^2/\sigma^4 - 2\alpha/\sigma^2)] = 0,\\ &y(h/2) = \exp\left(-\frac{a}{2\sigma^2}(h/2)^2\right). \end{aligned} \tag{5.5}$$

The solution of this equation can be given in terms of a series expansion.

An analytic expression for $\mathbb{E}\tau_1$ and thus for λ can be found by noting that $u(x) = \mathbb{E}_x\tau_1$ is the solution of

$$\begin{aligned} &Lu(x) = -1 \quad \text{for } x < h/2,\\ &u(h/2) = 0. \end{aligned}$$

Then

$$u(x) = 2/\sigma^2 \left(\int_x^{h/2} \exp\left(\frac{ay^2}{\sigma^2}\right) \mathrm{d}y \right) \int_{-\infty}^{x} \exp\left(-\frac{ay^2}{\sigma^2}\right) \mathrm{d}y$$
$$+ 2/\sigma^2 \int_x^{h/2} \int_z^{h/2} \exp\left(\frac{ay^2}{\sigma^2}\right) \exp\left(-\frac{az^2}{\sigma^2}\right) \mathrm{d}z$$

and $\lambda = (2u(-h/2))^{-1}$.

Recall that $\lambda = (\mathbb{E}_3 d_1)^{-1}$ where the expectation is taken under the model of the null hypothesis, respectively. The expectation of the likelihood ratio statistic can now be calculated and equals

$$\begin{aligned} \mathbb{E}_3 LR(X,T) &= -\frac{a^2}{2\sigma^2} T\lambda \mathbb{E}_3 Y_1 - 2\frac{b}{\sigma}\lambda Th + \frac{a}{2}T + \frac{bT}{2\sigma^2} \\ &= T(-a^2\lambda/2\sigma^2 - 2bh\lambda/\sigma + a/2 + b^2/2\sigma^2). \end{aligned}$$

As the second moment of d_1 is finite, the same argument as in Lemma 1 can be used to show that the remainder term $R(X,T)$ converges to zero in probability as T goes to infinity.

We now have $k(\alpha, T) = \sigma_{LR} z_\alpha T^{1/2}\sigma/b + \mathbb{E}_3 LR(X,T)$ where, for the case that the parameters are known,

$$\sigma_{LR}^2 = 2\lambda^2 \operatorname{Cov}(d_1, Y_1)\left(2\frac{b}{\sigma}h - \frac{\sigma^2}{2b}\lambda^{-1}\right)$$
$$+ \lambda^3 \operatorname{Var} d_1 \left(4\frac{b^2}{\sigma^2}h^2 - 2h\frac{\sigma^2}{2b}\lambda^{-1} + \left(\frac{\sigma^2}{2b}\lambda^{-1}\right)^2\right) + \lambda \operatorname{Var} Y_1,$$

and $\mathbb{E}_3 LR(X,T) = T(-a^2\lambda/2\sigma^2 - 2bh\lambda/\sigma + a/2 + b^2/2\sigma^2)$.

References

[1] A.C. Atkinson, A method for discriminating between models (with discussion). *J.R. Statist. Soc. B.* **32** (1970), 323–353.
[2] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.
[3] A.N. Borodin and P. Salminen. *Handbook of Brownian Motion – Facts and Formulae*, Birkhaeuser Verlag, Basel, 1996.
[4] D.R. Cox, Tests of separate families of hypotheses. *Proc. 4th Berkeley Symp. Mathematical Statistics and Probability*, Vol. 1, 1961, pp. 105–123.
[5] D.R. Cox. Further results on tests of separate families of hypotheses. *J.R. Statist. Soc. B.* **24** (1962), 405–424.

[6] M. Csörgö and P. Revesz, Three strong approximations of the local time of a Wiener process and their applications to invariance. Colloq. Math. Soc. Janos Bolyai 36, North-Holland, Amsterdam-New York, 1984, pp. 223–254.

[7] M.I. Freidlin, *Functional Integration and Partial Differential Equations*, Princeton University Press, 1985.

[8] M. Freidlin and R. Pfeiffer, A threshold estimation problem for processes with hysteresis. *Statistics and Probability Letters* **36** (1998), 337–348.

[9] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 1988.

[10] P.W. Glynn and W. Whitt, Ordinary CLT and WLLN Version of $L = \lambda W$. *Math. Oper. Res.* **13** (1988), 674–692.