Bernoulli **21**(4), 2015, 2430–2456
DOI: 10.3150/14-BEJ650

On ADF goodness-of-fit tests for perturbed dynamical systems

YURY A. KUTOYANTS

Laboratoire de Statistique et Processus, Université du Maine, 72085 Le Mans, France and Laboratory of Quantitative Finance, Higher School of Economics, Moscow, Russia. E-mail: kutoyants@univ-lemans.fr

We consider the problem of construction of goodness-of-fit tests for diffusion processes with a *small noise*. The basic hypothesis is composite parametric and our goal is to obtain asymptotically distribution-free tests. We propose two solutions. The first one is based on a change of time, and the second test is obtained using a linear transformation of the "natural" statistics.

Keywords: asymptotically distribution free test; Cramér–von Mises tests; diffusion processes; goodness of fit test; perturbed dynamical systems

1. Introduction

We consider the following problem. Suppose that we observe a trajectory $X^\varepsilon = \{X_t, 0 \le t \le T\}$ of the following diffusion process:

$$\mathrm{d}X_t = S(t, X_t)\,\mathrm{d}t + \varepsilon\sigma(t, X_t)\,\mathrm{d}W_t, \qquad X_0 = x_0, 0 \le t \le T, \tag{1.1}$$

where $W_t, 0 \le t \le T$ is a Wiener process, $\sigma(t, x)$ is known smooth function, the initial value x_0 is deterministic and the trend coefficient $S(t, x)$ is a unknown function. Here $\varepsilon \in (0, 1)$ is a given parameter. We have to test the composite (parametric) hypothesis

$$\mathscr{H}_0 : \mathrm{d}X_t = S(\vartheta, t, X_t)\,\mathrm{d}t + \varepsilon\sigma(t, X_t)\,\mathrm{d}W_t, \qquad X_0 = x_0, 0 \le t \le T \tag{1.2}$$

against alternative $\mathscr{H}_1$: not $\mathscr{H}_0$. Here $S(\vartheta, t, x)$ is a known smooth function of ϑ and x. The parameter $\vartheta \in \Theta$ is unknown and the set $\Theta \subset \mathbb{R}^d$ is open and bounded. Let us fix some value $\alpha \in (0, 1)$ and consider the class of tests of asymptotic ($\varepsilon \to 0$) size α:

$$\mathcal{K}_\alpha = \bigl\{\bar\psi_\varepsilon : \mathbf{E}_\vartheta \bar\psi_\varepsilon = \alpha + \mathrm{o}(1)\bigr\} \qquad \text{for all } \vartheta \in \Theta.$$

The test $\bar\psi_\varepsilon = \bar\psi_\varepsilon(X^\varepsilon)$ is the probability to reject the hypothesis $\mathscr{H}_0$ and $\mathbf{E}_\vartheta$ stands for the mathematical expectation under hypothesis $\mathscr{H}_0$.

Our goal is to find goodness-of-fit (GoF) tests which are *asymptotically distribution free* (ADF), that is, we look for a test statistics whose limit distributions under null hypothesis do not depend on the underlying model given by the functions $S(\vartheta, t, x)$, $\sigma(t, x)$ and the parameter ϑ. This work is a continuation of the study Kutoyants [9], where an ADF test was proposed in the case of simple basic hypothesis.

1350-7265

The behaviour of stochastic systems governed by such equations (called *perturbed dynamical systems*) is well studied, see, for example, Freidlin and Wentzell [3] and the references therein. Estimation theory (parametric and non-parametric) for such models of observations is also well developped, see, for example, Kutoyants [8] and Yoshida [17,18].

Let us remind the well-known basic results in this problem for the i.i.d. model. We start with the simple hypothesis. Suppose that we observe n i.i.d. r.v.'s $(X_1, \dots, X_n) = X^n$ with a continuous distribution function $F(x)$, and the basic hypothesis is

$$\mathscr{H}_0 : F(x) \equiv F_0(x), \qquad x \in \mathbb{R}.$$

Then the Cramér–von Mises statistic is

$$D_n = n \int \bigl[\hat{F}_n(x) - F_0(x)\bigr]^2 \,\mathrm{d}F_0(x), \qquad \hat{F}_n(x) = \frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_{\{X_j < x\}},$$

where $\hat{F}_n(x)$ is the empirical distribution function. Denote by $\mathcal{K}_\alpha$ the class of tests of asymptotic $(n \to \infty)$ size $\alpha \in (0, 1)$, that is,

$$\mathcal{K}_\alpha = \bigl\{\bar{\psi} : \mathbf{E}_0 \bar{\psi} = \alpha + \mathrm{o}(1)\bigr\}.$$

We have the convergence (under hypothesis $\mathscr{H}_0$)

$$B_n(x) = \sqrt{n}\bigl(\hat{F}_n(x) - F_0(x)\bigr) \Longrightarrow B\bigl(F_0(x)\bigr),$$

where $B(\cdot)$ is a Brownian bridge process. Hence, it can be shown that

$$D_n \Longrightarrow \delta \equiv \int_0^1 B(s)^2 \,\mathrm{d}s,$$

and the *Cramér–von Mises* test

$$\psi_n\bigl(X^n\bigr) = \mathbb{1}_{\{D_n > c_\alpha\}} \in \mathcal{K}_\alpha, \qquad \mathbf{P}\{\delta > c_\alpha\} = \alpha$$

is *asymptotically distribution-free* (ADF).

The situation changes in the case of parametric basic hypothesis:

$$\mathscr{H}_0 : F(x) = F(\vartheta, x), \qquad \vartheta \in \Theta,$$

where $\Theta = (\alpha, \beta)$. If we introduce the similar statistic

$$\hat{D}_n = n \int_{-\infty}^{\infty} \bigl[\hat{F}_n(x) - F(\hat{\vartheta}_n, x)\bigr]^2 \,\mathrm{d}F(\hat{\vartheta}_n, x),$$

where $\hat{\vartheta}_n$ is the maximum likelihood estimator (MLE), then (under regularity conditions) we have

$$U_n(x) = \sqrt{n}\bigl(\hat{F}_n(x) - F(\hat{\vartheta}_n, x)\bigr) = B_n(x) - \sqrt{n}(\hat{\vartheta}_n - \vartheta)\dot{F}(\vartheta, x) + \mathrm{o}(1).$$

For the MLE, we can use its representation

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \frac{\dot{\boldsymbol{\ell}}(\vartheta, X_j)}{\mathrm{I}(\vartheta)} + \mathrm{o}(1), \qquad \ell(\vartheta, x) = \ln f(\vartheta, x).$$

All this allows us to write the limit $U(\cdot)$ of the statistic $U_n(\cdot)$ as follows:

$$\begin{aligned} U_n(x) &\Longrightarrow B\bigl(F(\vartheta, x)\bigr) - \int \frac{\dot{\ell}(\vartheta, y)}{\sqrt{\mathrm{I}(\vartheta)}} \,\mathrm{d}B\bigl(F(\vartheta, y)\bigr) \int_{-\infty}^{x} \frac{\dot{\ell}(\vartheta, y)}{\sqrt{\mathrm{I}(\vartheta)}} \,\mathrm{d}F(\vartheta, y) \\ &= B(t) - \int_0^1 h(\vartheta, v) \,\mathrm{d}B(v) \int_0^t h(\vartheta, v) \,\mathrm{d}v \equiv U(\vartheta, t), \end{aligned}$$

where $t = F(\vartheta, x)$ and we put $h(\vartheta, t) = \mathrm{I}(\vartheta)^{-1/2} \dot{\ell}(\vartheta, F_\vartheta^{-1}(t))$.

If $\vartheta \in \Theta \subset \mathbb{R}^d$, then we obtain a similar equation

$$U(\vartheta, t) = B(t) - \left\langle \int_0^1 \mathbf{h}(\vartheta, v) \,\mathrm{d}B(v), \int_0^t \mathbf{h}(\vartheta, v) \,\mathrm{d}v \right\rangle, \tag{1.3}$$

where $\langle \cdot, \cdot \rangle$ is the scalar product in $\mathbb{R}^d$.

This presentation of the limit process $U(\vartheta, t)$ can be found in Darling [2]. Of course, the test $\hat{\psi}_n = \mathbb{1}_{\{\hat{D}_n > c_\alpha\}}$ is not ADF and the choice of the threshold c_α can be a difficult problem. One way to avoid this problem is, for example, to find a transformation $L_W[U](t) = w(t)$, where $w(\cdot)$ is the Wiener process. This transformation allows to write the equality

$$\Delta = \int_{-\infty}^{\infty} L_W[U]\bigl(F(\vartheta, x)\bigr)^2 \,\mathrm{d}F(\vartheta, x) = \int_0^1 w(t)^2 \,\mathrm{d}t.$$

Hence, if we prove the convergence

$$\tilde{D}_n = \int_{-\infty}^{\infty} L_W[U_n](x)^2 \,\mathrm{d}F(\hat{\vartheta}_n, x) \Longrightarrow \Delta,$$

then the test $\tilde{\psi}_n = \mathbb{1}_{\{\tilde{D}_n > c_\alpha\}}$, with $\mathbf{P}(\Delta > c_\alpha) = \alpha$ is ADF. Such transformation was proposed in Khmaladze [6].

In the present work, we consider a similar problem for the model of observations (1.1) with parametric basic hypothesis (1.2). Note that several problems of GoF testing for the model of observations (1.1) with simple basic hypothesis $\Theta = \{\vartheta_0\}$ were studied in Dachian and Kutoyants [1], Iacus and Kutoyants [5], Kutoyants [9]. The tests considered there are mainly based on the normalized difference $\varepsilon^{-1}(X_t - x_t)$, where $x_t = x_t(\vartheta_0)$ is a solution of equation (1.2) for $\varepsilon = 0$. This statistic is in some sense similar to the normalized difference $\sqrt{n}(\hat{F}_n(x) - F_0(x))$ used in the GoF problems for i.i.d. models. We propose two GoF ADF tests. Note that the construction of the first test is in some sense close to the one considered in Kutoyants [11] and based on the score function process. These tests are originated by the different processes but after our first transformation of the normalized difference $\varepsilon^{-1}(X_t - x_t(\hat{\vartheta}_\varepsilon))$ we obtain the same integrals to calculate as those in Kutoyants [11].

Let us remind the related results in the case of simple hypothesis (see Kutoyants [9]). Suppose that the observed homogeneous diffusion process under null hypothesis is

$$\mathrm{d}X_t = S_0(X_t)\,\mathrm{d}t + \varepsilon\sigma(X_t)\,\mathrm{d}W_t, \qquad X_0 = x_0, 0 \le t \le T,$$

where $S_0(x)$ is a known smooth function. Denote $x_t = X_t|_{\varepsilon=0}$. We have $X_t \to x_t$ as $\varepsilon \to 0$ and we construct a GoF test based on statistic $v_\varepsilon(t) = \varepsilon^{-1}(X_t - x_t)$. The limit of this statistic is a Gaussian process. This process can be transformed into the Wiener process as follows: introduce the statistic

$$\delta_\varepsilon = \left[\int_0^T \left(\frac{\sigma(x_t)}{S_0(x_t)}\right)^2 \mathrm{d}t\right]^{-2} \int_0^T \left(\frac{X_t - x_t}{\varepsilon S_0(x_t)^2}\right)^2 \sigma(x_t)^2\,\mathrm{d}t.$$

The following convergence:

$$\delta_\varepsilon \Longrightarrow \Delta = \int_0^1 w(s)^2\,\mathrm{d}s$$

was proved and therefore the test $\hat{\psi}_\varepsilon = \mathbb{1}_{\{\delta_\varepsilon > c_\alpha\}}$ with $\mathbf{P}(\Delta > c_\alpha) = \alpha$ is ADF.

Consider now the hypotheses testing problem (1.1) and (1.2). The solution x_t of equation (1.2) for $\varepsilon = 0$ depends on $\vartheta \in \Theta \subset \mathbb{R}^d$, that is, $x_t = x_t(\vartheta)$. The statistic $\hat{v}_\varepsilon(t) = \varepsilon^{-1}(X_t - x_t(\hat{\vartheta}_\varepsilon))$ (here $\hat{\vartheta}_\varepsilon$ is the MLE) is in some sense similar to $U_n(\cdot)$. Denote by $\hat{v}(t)$ the limit of $\hat{v}_\varepsilon(t)$ as $\varepsilon \to 0$ and suppose that we know the transformation $L_U[\hat{v}](\cdot)$ of $\hat{v}(\cdot)$ into the Gaussian process

$$U(\vartheta, t) = W(t) - \left\langle \int_0^1 \mathbf{h}(\vartheta, s)\,\mathrm{d}W(s), \int_0^t \mathbf{h}(\vartheta, s)\,\mathrm{d}s \right\rangle, \qquad 0 \le t \le 1$$

with a vector-function $h(\vartheta, s)$ satisfying

$$\int_0^1 \mathbf{h}(\vartheta, s)\mathbf{h}(\vartheta, s)^*\,\mathrm{d}s = \mathbb{J}.$$

Here $\mathbb{J}$ is the $d \times d$ unit matrix.

The next steps are two transformations of $U(\cdot)$: one transformation into the Brownian bridge $L_B[U](s) = B(s)$ and another one into the Wiener process $L_W[U](s) = w(s)$, respectively. This allows us to construct the ADF GoF tets as follows: let us introduce (formally) the statistics

$$\delta_\varepsilon = \int_0^T \big(L_B\big[L_U[\hat{v}_\varepsilon]\big](t)\big)^2\,\mathrm{d}t, \qquad \Delta_\varepsilon = \int_0^T \big(L_W\big[L_U[\hat{v}_\varepsilon](t)\big]\big)^2\,\mathrm{d}t,$$

and suppose that we have proved the convergences

$$\delta_\varepsilon \Longrightarrow \delta = \int_0^1 B(s)^2\,\mathrm{d}s, \qquad \Delta_\varepsilon \Longrightarrow \Delta = \int_0^1 w(s)^2\,\mathrm{d}s.$$

Then the tests

$$\hat{\psi}_\varepsilon = \mathbb{1}_{\{\delta_\varepsilon > d_\alpha\}}, \qquad \mathbf{P}(\delta > d_\alpha) = \alpha, \qquad \hat{\Psi}_\varepsilon = \mathbb{1}_{\{\Delta_\varepsilon > c_\alpha\}}, \qquad \mathbf{P}(\Delta > c_\alpha) = \alpha,$$

belong to the class $\mathcal{K}_\alpha$ and are ADF. Our objective is to realize this program.

A similar result for ergodic diffusion processes is contained in Kutoyants [12] (simple basic hypothesis) and Kleptsyna and Kutoyants [7] (parametric basic hypothesis).

2. Auxiliary results

We have the following stochastic differential equation:

$$\mathrm{d}X_t = S(\vartheta, t, X_t)\,\mathrm{d}t + \varepsilon\sigma(t, X_t)\,\mathrm{d}W_t, \qquad X_0 = x_0, 0 \le t \le T, \tag{2.1}$$

where $\vartheta \in \Theta$, Θ is an open bounded subset of $\mathbb{R}^d$ and ε is a *small parameter*, that is, we study this equation in the asymptotics of *small noise* $\varepsilon \to 0$.

Introduce the Lipschitz condition and that of linear growth:

$\mathcal{C}1$. *The functions* $S(\vartheta, t, x)$ *and* $\sigma(t, x)$ *satisfy the relations*

$$\begin{aligned} \bigl|S(\vartheta, t, x) - S(\vartheta, t, y)\bigr| + \bigl|\sigma(t, x) - \sigma(t, y)\bigr| &\le L|x - y|, \\ \bigl|S(\vartheta, t, x)\bigr| + \bigl|\sigma(t, x)\bigr| &\le L\bigl(1 + |x|\bigr). \end{aligned}$$

Recall that by these conditions the stochastic differential equation (2.1) has a unique strong solution (Liptser and Shiryaev [14]), and moreover this solution $X^\varepsilon = \{X_t, 0 \le t \le T\}$ converges uniformly, with respect to t, to the solution $x^T = \{x_t, 0 \le t \le T\}$ of the ordinary differential equation

$$\frac{\mathrm{d}x_t}{\mathrm{d}t} = S(\vartheta, t, x_t), \qquad x_0, 0 \le t \le T. \tag{2.2}$$

Observe that $x_t = x_t(\vartheta)$ (for the proof see Freidlin and Wentzell [3], Kutoyants [8]).

$\mathcal{C}2$. *The diffusion coefficient* $\sigma(t, x)^2$ *is bounded away from zero*

$$\inf_{0 \le t \le T, x} \sigma(t, x)^2 > 0.$$

Conditions $\mathcal{C}1$ and $\mathcal{C}2$ provide the equivalence of the measures $\{\mathbf{P}_\vartheta^{(\varepsilon)}, \vartheta \in \Theta\}$ induced on the measurable space $(\mathscr{C}_T, \mathfrak{B}_T)$ by the solutions of equation (2.1) (Liptser and Shiryaev [14]). Here $\mathscr{C}_T$ is the space of continuous functions on $[0, T]$ with uniform metrics and $\mathfrak{B}_T$ is the Borelian σ-algebra of its subsets. The likelihood ratio function is

$$L\bigl(\vartheta, X^\varepsilon\bigr) = \exp\left\{\int_0^T \frac{S(\vartheta, t, X_t)}{\varepsilon^2\sigma(t, X_t)^2}\,\mathrm{d}X_t - \int_0^T \frac{S(\vartheta, t, X_t)^2}{2\varepsilon^2\sigma(t, X_t)^2}\,\mathrm{d}t\right\}, \qquad \vartheta \in \Theta,$$

and the maximum likelihood estimator (MLE) $\hat{\vartheta}_\varepsilon$ is defined by the equation

$$L\bigl(\hat{\vartheta}_\varepsilon, X^\varepsilon\bigr) = \sup_{\vartheta \in \Theta} L\bigl(\vartheta, X^\varepsilon\bigr).$$

The following regularity conditions (smoothness and identifiability) provides us necessary properties of the MLE. Below $x_t = x_t(\vartheta_0)$.

$\mathcal{C}3$. *The functions $S(\vartheta, t, x)$ and $\sigma(t, x)$ have two continuous bounded derivatives w.r.t. x and the function $S(\vartheta, t, x)$ has two continuous bounded derivatives w.r.t. ϑ.*

For any $\nu > 0$

$$\inf_{\vartheta_0\in\Theta}\inf_{|\vartheta-\vartheta_0|>\nu}\int_0^T\left(\frac{S(\vartheta,t,x_t)-S(\vartheta_0,t,x_t)}{\sigma(t,x_t)}\right)^2\mathrm{d}t>0$$

and the information matrix $(d \times d)$

$$\mathbb{I}(\vartheta_0)=\int_0^T\frac{\dot{\mathbf{S}}(\vartheta_0,t,x_t)\dot{\mathbf{S}}(\vartheta_0,t,x_t)^*}{\sigma(t,x_t)^2}\,\mathrm{d}t$$

is uniformly non-degenerate:

$$\inf_{\vartheta_0\in\Theta}\inf_{|\lambda|=1}\lambda^*\mathbb{I}(\vartheta_0)\lambda>0.$$

We denote by a prime the derivatives w.r.t. x and t, and by a dot those w.r.t. ϑ, that is, for a function $f = f(\vartheta, t, x)$ we write

$$f'(\vartheta,t,x)=\frac{\partial f(\vartheta,t,x)}{\partial x},$$
$$f_t'(\vartheta,t,x)=\frac{\partial f(\vartheta,t,x)}{\partial t},$$
$$\dot{f}(\vartheta,t,x)=\frac{\partial f(\vartheta,t,x)}{\partial\vartheta}.$$

Of course, in the case of $d > 1$ the derivative $\dot{\mathbf{f}}(\vartheta, t, x)$ is a column vector.

If the conditions $\mathcal{C}2$ and $\mathcal{C}3$ hold, then the MLE admits the representation

$$\varepsilon^{-1}(\hat{\vartheta}_\varepsilon-\vartheta)=\mathbb{I}(\vartheta)^{-1}\int_0^T\frac{\dot{\mathbf{S}}(\vartheta,t,x_t)}{\sigma(t,x_t)}\,\mathrm{d}W_t+\mathrm{o}(1). \tag{2.3}$$

Here, $x_t = x_t(\vartheta)$. For the proof see, Kutoyants [9].

Note that $X_t = X_t(\varepsilon)$ (solution of equation (2.1)) under condition $\mathcal{C}3$ is continuously differentiable w.r.t. ε. Denote the derivatives

$$X_t^{(1)}=\frac{\partial X_t}{\partial\varepsilon},\qquad x_t^{(1)}=\frac{\partial X_t}{\partial\varepsilon}\bigg|_{\varepsilon=0},\qquad 0\le t\le T.$$

The equations for $X_t^{(1)}$ and $x_t^{(1)}$ are

$$\mathrm{d}X_t^{(1)}=S'(\vartheta,t,X_t)X_t^{(1)}\,\mathrm{d}t+\left[\varepsilon\sigma'(t,X_t)X_t^{(1)}+\sigma(t,X_t)\right]\mathrm{d}W_t,\qquad X_0^{(1)}=0$$

and

$$\mathrm{d}x_t^{(1)}=S'(\vartheta,t,x_t)x_t^{(1)}\,\mathrm{d}t+\sigma(t,x_t)\,\mathrm{d}W_t,\qquad x_0^{(1)}=0, \tag{2.4}$$

respectively. Hence $x_t^{(1)}$, $0 \le t \le T$ is a Gaussian process and it can be written as

$$x_t^{(1)} = \int_0^t \exp\left\{\int_s^t S'(\vartheta, v, x_v)\,\mathrm{d}v\right\}\sigma(s, x_s)\,\mathrm{d}W_s. \tag{2.5}$$

Denote

$$\psi(t) = \exp\left\{\int_0^t S'(\vartheta, v, x_v)\,\mathrm{d}v\right\}, \qquad \psi_\varepsilon(t) = \exp\left\{\int_0^t S'(\hat{\vartheta}_\varepsilon, v, X_v)\,\mathrm{d}v\right\}.$$

We can write

$$\begin{aligned}
\frac{X_t - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} &= \frac{X_t - x_t(\vartheta)}{\varepsilon} + \frac{x_t(\vartheta) - x_t(\hat{\vartheta}_\varepsilon)}{\varepsilon} \\
&= X_t^{(1)} - \left\langle \frac{(\hat{\vartheta}_\varepsilon - \vartheta)}{\varepsilon}, \dot{\mathbf{x}}_t(\vartheta)\right\rangle + \mathrm{o}(1) \\
&= x_t^{(1)} - \left\langle \mathbb{I}(\vartheta)^{-1}\int_0^T \frac{\dot{\mathbf{S}}(\vartheta, s, x_s)}{\sigma(s, x_s)}\,\mathrm{d}W_s, \dot{\mathbf{x}}_t(\vartheta)\right\rangle + \mathrm{o}(1) \\
&= \psi(t)V(t) + \mathrm{o}(1),
\end{aligned}$$

where

$$V(t) = \psi(t)^{-1}x_t^{(1)} - \psi(t)^{-1}\left\langle \mathbb{I}(\vartheta)^{-1}\int_0^T \frac{\dot{\mathbf{S}}(\vartheta, s, x_s)}{\sigma(s, x_s)}\,\mathrm{d}W_s, \dot{\mathbf{x}}_t(\vartheta)\right\rangle.$$

Introduce the random process

$$U(\vartheta, t) = \int_0^t \frac{\psi(s)}{\sigma(s, x_s)}\,\mathrm{d}V(s).$$

Lemma 1. *We have the equality*

$$U(\vartheta, t) = W_t - \left\langle \int_0^T \mathbf{h}(\vartheta, s)\,\mathrm{d}W_s, \int_0^t \mathbf{h}(\vartheta, s)\,\mathrm{d}s\right\rangle, \qquad 0 \le t \le T, \tag{2.6}$$

where

$$\mathbf{h}(\vartheta, t) = \mathbb{I}(\vartheta)^{-1/2}\frac{\dot{\mathbf{S}}(\vartheta, t, x_t)}{\sigma(t, x_t)} \tag{2.7}$$

is a vector-valued function.

Proof. The solution of equation (2.4) can be written (see (2.5)) as

$$x_t^{(1)} = \int_0^t \frac{\psi(t)\sigma(s, x_s)}{\psi(s)}\,\mathrm{d}W_s.$$

For the vector $\dot{\mathbf{x}}_t(\vartheta)$, we can write

$$\dot{\mathbf{x}}_t(\vartheta) = \int_0^t S'(\vartheta, s, x_s)\dot{\mathbf{x}}_s(\vartheta)\,\mathrm{d}s + \int_0^t \dot{\mathbf{S}}(\vartheta, s, x_s)\,\mathrm{d}s.$$

The solution of this equation is

$$\dot{\mathbf{x}}_t(\vartheta) = \psi(t)\int_0^t \frac{\dot{\mathbf{S}}(\vartheta, s, x_s)}{\psi(s)}\,\mathrm{d}s.$$

Introduce two stochastic processes

$$v_1(t) = \psi(t)^{-1}x_t^{(1)} = \int_0^t \psi(s)^{-1}\sigma(s, x_s)\,\mathrm{d}W_s$$

and

$$\mathbf{v}_2(t) = \psi(t)^{-1}\dot{\mathbf{x}}_t(\vartheta) = \int_0^t \psi(s)^{-1}\dot{\mathbf{S}}(\vartheta, s, x_s)\,\mathrm{d}s.$$

Then we can write

$$\begin{aligned}
U(\vartheta, t) &= \int_0^t \frac{\psi(s)}{\sigma(s, x_s)}\,\mathrm{d}V(s) \\
&= \int_0^t \frac{\psi(s)}{\sigma(s, x_s)}\,\mathrm{d}v_1(s) \\
&\quad - \left\langle \mathbb{I}(\vartheta)^{-1}\int_0^T \frac{\dot{\mathbf{S}}(\vartheta, s, x_s)}{\sigma(s, x_s)}\,\mathrm{d}W_s, \int_0^t \frac{\psi(s)}{\sigma(s, x_s)}\,\mathrm{d}\mathbf{v}_2(s)\right\rangle \\
&= W(t) - \left\langle \mathbb{I}(\vartheta)^{-1/2}\int_0^T \frac{\dot{\mathbf{S}}(\vartheta, s, x_s)}{\sigma(s, x_s)}\,\mathrm{d}W_s, \mathbb{I}(\vartheta)^{-1/2}\int_0^t \frac{\dot{\mathbf{S}}(\vartheta, s, x_s)}{\sigma(s, x_s)}\,\mathrm{d}s\right\rangle \\
&= W_t - \left\langle \int_0^T \mathbf{h}(\vartheta, s)\,\mathrm{d}W_s, \int_0^t \mathbf{h}(\vartheta, s)\,\mathrm{d}s\right\rangle.
\end{aligned}$$

Introduce the random process

$$u(\vartheta, r) = T^{-1/2}U(\vartheta, rT), \qquad 0 \le r \le 1$$

and denote

$$\begin{aligned}
\mathbb{I}_1(\vartheta) &= \int_0^1 \frac{\dot{\mathbf{S}}(\vartheta, rT, x_{rT})\dot{\mathbf{S}}(\vartheta, rT, x_{rT})^*}{\sigma(rT, x_{rT})^2}\,\mathrm{d}r, \\
\tilde{\mathbf{h}}(\vartheta, r) &= \mathbb{I}_1(\vartheta)^{-1/2}\frac{\dot{\mathbf{S}}(\vartheta, rT, x_{rT})}{\sigma(rT, x_{rT})}, \qquad w_r = T^{-1/2}W_{rT}.
\end{aligned}$$

Then we can write

$$u(\vartheta, r) = w_r - \left\langle \int_0^1 \tilde{\mathbf{h}}(\vartheta, q)\, \mathrm{d}w_q, \int_0^r \tilde{\mathbf{h}}(\vartheta, q)\, \mathrm{d}q \right\rangle, \qquad 0 \le r \le 1, \tag{2.8}$$

and therefore

$$\int_0^1 \tilde{\mathbf{h}}(\vartheta, q)\tilde{\mathbf{h}}(\vartheta, q)^*\, \mathrm{d}q = \mathbb{J}. \qquad \square$$

Note that $u(\cdot)$ is in some sense a *universal limit* which appears in the problems of goodness of fit testing for stochastic processes. For example, the same limit is obtained in the case of ergodic diffusion process and in the case of inhomogeneous Poisson process (Kutoyants [10]). The main difference with the i.i.d. case is due to the Wiener process here, while in the i.i.d. case the Brownian bridge $B(t)$, $0 \le t \le 1$ appears (see (1.3)). Of course, we can immediately replace $B(t)$ by a Wiener process $B(t) = W_t - W_1 t$ and this will increase the dimension of the vector $h(\vartheta, \cdot)$. In the case of vector-valued parameter ϑ, this change is not essential and will slightly modify calculations of the test statistics for the first type test. At the same time if the parameter ϑ is one-dimensional, then we can easily construct the second-type goodness-of-fit test for stochastic processes and it remains unclear how to construct such tests in the i.i.d. case. The difference will be explained in Section 3.2.

In the construction of a GoF test, we will use another condition.

$\mathcal{C}4$. *The functions* $S(\vartheta, t, x)$, $\dot{\mathbf{S}}(\vartheta, t, x)$ *and* $\sigma(t, x)$ *have continuous bounded derivatives w.r.t.* $t \in [0, T]$.

3. Main results

Suppose that we observe a trajectory $X^\varepsilon = (X_t, 0 \le t \le T)$ of the following diffusion process:

$$\mathrm{d}X_t = S(t, X_t)\, \mathrm{d}t + \varepsilon\sigma(t, X_t)\, \mathrm{d}W_t, \qquad X_0 = x_0, 0 \le t \le T. \tag{3.1}$$

We have to test the basic parametric hypothesis

$$\mathcal{H}_0 : S(t, x) = S(\vartheta, t, x), \qquad 0 \le t \le T, \vartheta \in \Theta,$$

that is, that the observed process (3.1) has the stochastic differential

$$\mathrm{d}X_t = S(\vartheta, t, X_t)\, \mathrm{d}t + \varepsilon\sigma(t, X_t)\, \mathrm{d}W_t, \qquad X_0 = x_0, 0 \le t \le T \tag{3.2}$$

with some $\vartheta \in \Theta$. Here $S(\vartheta, t, x)$ and $\sigma(t, x)$ are known strictly positive smooth functions and $\Theta \subset R^d$ is an open convex set. We have to test this hypothesis in the asymptotics of a *small noise* (as $\varepsilon \to 0$).

Our goal is to construct such statistics $v_\varepsilon[X^\varepsilon](\cdot)$, $V_\varepsilon[X^\varepsilon](\cdot)$ that (under hypothesis $\mathcal{H}_0$)

$$\delta_\varepsilon=\int_0^T v_\varepsilon\big[X^\varepsilon\big](t)^2\,\mathrm{d}t \quad\Longrightarrow\quad \delta=\int_0^1 B(s)^2\,\mathrm{d}s,$$
$$\Delta_\varepsilon=\int_0^T V_\varepsilon\big[X^\varepsilon\big](t)^2\,\mathrm{d}t \quad\Longrightarrow\quad \Delta=\int_0^1 w(s)^2\,\mathrm{d}s,$$

where $B(\cdot)$ and $w(\cdot)$ are the Brownian bridge and the Wiener process, respectively. Then we introduce the tests

$$\hat\psi_\varepsilon=\mathbb{1}_{\{\delta_\varepsilon>d_\alpha\}},\qquad \hat\Psi_\varepsilon=\mathbb{1}_{\{\Delta_\varepsilon>c_\alpha\}}$$

with the thresholds c_α and d_α satisfying the equations

$$\mathbf{P}(\delta>d_\alpha)=\alpha,\qquad \mathbf{P}(\Delta>c_\alpha)=\alpha. \tag{3.3}$$

These tests will belong to the class

$$\mathcal{K}_\alpha=\Big\{\bar\psi_\varepsilon:\lim_{\varepsilon\to 0}\mathbf{E}_\vartheta\bar\psi_\varepsilon=\alpha,\forall\vartheta\in\Theta\Big\}$$

and will be ADF.

We propose these tests in the Sections 3.1 and 3.2 below. We call $\hat\psi_\varepsilon$ *the first* test and $\hat\Psi_\varepsilon$ *the second* test.

3.1. First test

The construction of the first ADF GoF test is based on the following well known property. Suppose that we have a Gaussian process $U(t), 0\le t\le T$ satisfying the equation

$$U(t)=w(t)-\int_0^T h(s)\,\mathrm{d}w(s)\int_0^t h(s)\,\mathrm{d}s,\qquad \int_0^T h(s)^2\,\mathrm{d}s=1.$$

Introduce the process

$$\begin{aligned}b(t)&=\int_0^t h(s)\,\mathrm{d}U(s)\\&=\int_0^t h(s)\,\mathrm{d}w(s)-\int_0^T h(s)\,\mathrm{d}w(s)\int_0^t h(s)^2\,\mathrm{d}s.\end{aligned}$$

It is easy to see that $b(0)=b(T)=0$ and

$$\mathbf{E}\big[b(t)b(s)\big]=\int_0^{t\wedge s} h(v)^2\,\mathrm{d}v-\int_0^t h(v)^2\,\mathrm{d}v\int_0^s h(v)^2\,\mathrm{d}v.$$

Let us put

$$\tau=\int_0^s h(v)^2\,\mathrm{d}v,\qquad b(t)=B(\tau),\qquad 0\le\tau\le 1.$$

Then

$$\begin{aligned}\delta&=\int_0^T\left(\int_0^t h(s)\,\mathrm{d}U(s)\right)^2 h(t)^2\,\mathrm{d}t\\&=\int_0^T b(t)^2h(t)^2\,\mathrm{d}t=\int_0^1 B(\tau)^2\,\mathrm{d}\tau.\end{aligned}$$

Suppose that the parameter ϑ is one-dimensional, $\vartheta\in\Theta=(a,b)$ and that we already proved the convergence (see Lemma 1)

$$U_\varepsilon(t)=\int_0^t\frac{\psi_\varepsilon(s)}{\sigma(s,X_s)}\,\mathrm{d}\left(\frac{X_s-x_s(\hat\vartheta_\varepsilon)}{\varepsilon\psi_\varepsilon(s)}\right)\longrightarrow U(\vartheta,t),\qquad 0\le t\le T,$$

where

$$U(\vartheta,t)=w(t)-\int_0^T h(\vartheta,s)\,\mathrm{d}w(s)\int_0^t h(\vartheta,s)\,\mathrm{d}s,\qquad \int_0^T h(\vartheta,s)^2\,\mathrm{d}s=1.$$

Recall that

$$h(\vartheta,s)=\mathrm{I}(\vartheta)^{-1/2}\frac{\dot S(\vartheta,s,x_s)}{\sigma(s,x_s)},\qquad \mathrm{I}(\vartheta)=\int_0^T\frac{\dot S(\vartheta,s,x_s)^2}{\sigma(s,x_s)^2}\,\mathrm{d}s.$$

Introduce (formally) the statistic

$$\hat\delta_\varepsilon=\int_0^T\left(\int_0^t h(\hat\vartheta_\varepsilon,s)\,\mathrm{d}U_\varepsilon(s)\right)^2 h(\hat\vartheta_\varepsilon,t)^2\,\mathrm{d}t.$$

If we prove that

$$\begin{aligned}&\int_0^T\left(\int_0^t h(\hat\vartheta_\varepsilon,s)\,\mathrm{d}U_\varepsilon(s)\right)^2 h(\hat\vartheta_\varepsilon,t)^2\,\mathrm{d}t\\&\quad\Longrightarrow\int_0^T\left(\int_0^t h(\vartheta,s)\,\mathrm{d}U(\vartheta,s)\right)^2 h(\vartheta,t)^2\,\mathrm{d}t\end{aligned}$$

then the test $\hat\psi_\varepsilon=\mathbb{1}_{\{\delta_\varepsilon>c_\alpha\}}$ will be ADF.

The main technical problem in carrying out this program is to define the stochastic integral

$$\int_0^t h(\hat\vartheta_\varepsilon,s)\,\mathrm{d}U_\varepsilon(s)$$

containing the MLE $\hat\vartheta_\varepsilon = \hat\vartheta_\varepsilon(X_t, 0 \le t \le T)$. We will proceed as follows: First, we formally differentiate and integrate and then we take the final expressions, which do not contain stochastic integrals, as starting statistics.

Introduce the statistics

$$
\begin{aligned}
D(\vartheta, s, X_s) &= S\bigl(\vartheta, s, x_s(\vartheta)\bigr) + S'(\vartheta, s, X_s)\bigl(X_s - x_s(\vartheta)\bigr),\\
R\bigl(\vartheta, t, X^t\bigr) &= \int_{x_0}^{X_t} \frac{\dot S(\vartheta, t, y)}{\sqrt{\mathrm{I}(\vartheta)}\sigma(t, y)^2}\,\mathrm{d}y\\
&\quad - \int_0^t \int_{x_0}^{X_s} \frac{\dot S_s'(\vartheta, s, y)\sigma(s, y) - 2\dot S(\vartheta, s, y)\sigma_s'(s, y)}{\sqrt{\mathrm{I}(\vartheta)}\sigma(s, y)^3}\,\mathrm{d}y\,\mathrm{d}s,\\
Q\bigl(\vartheta, t, X^t\bigr) &= \int_0^t \frac{\dot S(\vartheta, s, X_s) D(\vartheta, s, X_s)}{\sqrt{\mathrm{I}(\vartheta)}\sigma(s, X_s)^2}\,\mathrm{d}s,\\
K_\varepsilon(\vartheta, t) &= \varepsilon^{-1}\bigl[R\bigl(\vartheta, t, X^t\bigr) - Q\bigl(\vartheta, t, X^t\bigr)\bigr],\\
\delta_\varepsilon &= \int_0^T K_\varepsilon(\hat\vartheta_\varepsilon, t)^2 h_\varepsilon(\hat\vartheta_\varepsilon, t)^2\,\mathrm{d}t.
\end{aligned}
$$

The first test is given in the following theorem.

Theorem 1. *Suppose that the conditions $\mathcal{C}1$–$\mathcal{C}4$ hold. Then the test*

$$
\hat\psi_\varepsilon = \mathbb{1}_{\{\delta_\varepsilon > c_\alpha\}}, \qquad \mathbf{P}\{\delta > c_\alpha\} = \alpha
$$

is ADF and belongs to $\mathcal{K}_\varepsilon$.

Proof. We can write (formally)

$$
\begin{aligned}
U_\varepsilon(t) &= \int_0^t \frac{\psi_\varepsilon(s)}{\sigma(s, X_s)}\,\mathrm{d}V_\varepsilon(s)\\
&= \int_0^t \frac{\psi_\varepsilon(s)}{\sigma(s, X_s)}\,\mathrm{d}\left(\frac{X_s - x_s(\hat\vartheta_\varepsilon)}{\psi_\varepsilon(s)\varepsilon}\right)\\
&= \int_0^t \frac{\mathrm{d}X_s}{\varepsilon\sigma(s, X_s)} - \int_0^t \left[\frac{S(\hat\vartheta_\varepsilon, s, x_s(\hat\vartheta_\varepsilon))}{\varepsilon\sigma(s, X_s)} + \frac{S'(\hat\vartheta_\varepsilon, s, X_s)(X_s - x_s(\hat\vartheta_\varepsilon))}{\varepsilon\sigma(s, X_s)}\right]\mathrm{d}s\\
&= \int_0^t \frac{\mathrm{d}X_s}{\varepsilon\sigma(s, X_s)} - \int_0^t \frac{D(\hat\vartheta_\varepsilon, s, X_s)}{\varepsilon\sigma(s, X_s)}\,\mathrm{d}s,
\end{aligned}
\tag{3.4}
$$

where we have used the equality

$$
\mathrm{d}x_s(\hat\vartheta_\varepsilon) = S\bigl(\hat\vartheta_\varepsilon, s, x_s(\hat\vartheta_\varepsilon)\bigr)\,\mathrm{d}s.
$$

Hence (formally), we obtain the following expression.

$$\int_0^t h_\varepsilon(\hat{\vartheta}_\varepsilon, s)\,\mathrm{d}U_\varepsilon(s) = \int_0^t \frac{\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s)}{\sqrt{\mathrm{I}(\hat{\vartheta}_\varepsilon)}\varepsilon\sigma(s, X_s)^2}\,\mathrm{d}X_s - \int_0^t \frac{\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s)D(\hat{\vartheta}_\varepsilon, s, X_s)}{\sqrt{\mathrm{I}(\hat{\vartheta}_\varepsilon)}\varepsilon\sigma(s, X_s)^2}\,\mathrm{d}s.$$

The estimator $\hat{\vartheta}_\varepsilon = \hat{\vartheta}_\varepsilon(X_t, 0 \le t \le T)$ and therefore the stochastic integral is not well defined because the integrand $\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s)$ is not a non-anticipative random function. Note that in the linear case $S(\vartheta, t, x) = \vartheta Q(s, x)$ we have no such problem (see example below). This difficulty can be avoided in general case by at least two ways: The first one is to replace the stochastic integral by it's *robust version* as we show below. The second possibility is to use a consistent estimator $\bar{\vartheta}_{\nu_\varepsilon}$ of the parameter ϑ constructed after the observations $X^{\nu_\varepsilon} = (X_t, 0 \le t \le \nu_\varepsilon)$, where $\nu_\varepsilon \to 0$ but sufficiently slowly. With this estimator, we can calculate the integral

$$\int_{\nu_\varepsilon}^t \frac{\dot{S}(\bar{\vartheta}_{\nu_\varepsilon}, s, X_s)}{\sigma(s, X_s)^2}\,\mathrm{d}X_s$$

without any problem, and all limits will be the same. Such construction is discussed for a different problem in Kutoyants and Zhou [13].

Introduce the function

$$M(\vartheta, t, x) = \int_{x_0}^x \frac{\dot{S}(\vartheta, t, y)}{\sigma(t, y)^2}\,\mathrm{d}y.$$

Then by the Itô formula

$$\mathrm{d}M(\vartheta, t, X_t) = M_t'(\vartheta, t, X_t)\,\mathrm{d}t + \frac{\varepsilon^2\sigma(t, X_t)^2}{2}M_{xx}''(\vartheta, t, X_t)\,\mathrm{d}t + M_x'(\vartheta, t, X_t)\,\mathrm{d}X_t$$

and therefore

$$\begin{aligned}
&\int_0^t \frac{\dot{S}(\vartheta, s, X_s)}{\sigma(s, X_s)^2}\,\mathrm{d}X_s \\
&\quad = M(\vartheta, t, X_t) - \int_0^t \left[M_s'(\vartheta, s, X_s) + \frac{\varepsilon^2\sigma(s, X_s)^2}{2}M_{xx}''(\vartheta, s, X_s)\right]\mathrm{d}s \\
&\quad = \int_{x_0}^{X_t} \frac{\dot{S}(\vartheta, t, y)}{\sigma(t, y)^2}\,\mathrm{d}y - \int_0^t \int_{x_0}^{X_s} \frac{\dot{S}_s'(\vartheta, s, y)}{\sigma(s, y)^2}\,\mathrm{d}y \\
&\qquad + \int_0^t \int_{x_0}^{X_s} \frac{2\dot{S}(\vartheta, s, y)\sigma_s'(s, y)}{\sigma(s, y)^3}\,\mathrm{d}s - \frac{\varepsilon^2}{2}\int_0^t \sigma(s, X_s)^2 M_{xx}''(\vartheta, s, X_s)\,\mathrm{d}s.
\end{aligned}$$

Note that the contribution of the term

$$\varepsilon^2 \int_0^t \sigma(s, X_s)^2 M''_{xx}(\hat{\vartheta}_\varepsilon, s, X_s)\, \mathrm{d}s$$

is asymptotically ($\varepsilon \to 0$) negligible. Therefore,

$$K_\varepsilon(\vartheta, t) = \varepsilon^{-1}\big[R\big(\vartheta, t, X^t\big) - Q\big(\vartheta, t, X^t\big)\big]$$

is asymptotically equivalent to

$$\tilde{K}_\varepsilon(\vartheta, t) = \int_0^t h_\varepsilon(\vartheta, s)\, \mathrm{d}U_\varepsilon(s).$$

The difference is in the dropped term of order $\mathrm{O}(\varepsilon)$.

We have to verify the convergence of the integrals

$$\delta_\varepsilon = \int_0^T \frac{K_\varepsilon(\hat{\vartheta}_\varepsilon, t)^2 \dot{S}(\hat{\vartheta}_\varepsilon, t, X_t)^2}{\mathrm{I}(\hat{\vartheta}_\varepsilon)\sigma(t, X_t)^2}\, \mathrm{d}t \longrightarrow \int_0^T \frac{K(\vartheta, t)^2 \dot{S}(\vartheta, t, x_t)^2}{\mathrm{I}(\vartheta)\sigma(t, x_t)^2}\, \mathrm{d}t.$$

Regularity conditions $\mathcal{C}1$–$\mathcal{C}3$ give the uniform convergences

$$\sup_{0\le t\le T} \big|X_t - x_t(\vartheta)\big| \longrightarrow 0, \qquad \mathrm{I}(\hat{\vartheta}_\varepsilon) \longrightarrow \mathrm{I}(\vartheta),$$

$$\sup_{0\le t\le T} \big|h_\varepsilon(\hat{\vartheta}_\varepsilon, t) - h(\hat{\vartheta}_\varepsilon, t)\big| = \sup_{0\le t\le T} \left| \frac{\dot{S}(\hat{\vartheta}_\varepsilon, t, X_t)}{\sqrt{\mathrm{I}(\hat{\vartheta}_\varepsilon)}\sigma(t, X_t)} - \frac{\dot{S}(\vartheta, t, x_t)}{\sqrt{\mathrm{I}(\vartheta)}\sigma(t, x_t)} \right| \longrightarrow 0.$$

Introduce two processes

$$Y_\varepsilon\big(\hat{\vartheta}_\varepsilon, t, X^t\big) = \int_0^t \frac{\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s)[S(\vartheta, s, X_s) - D(\hat{\vartheta}_\varepsilon, s, X_s)]}{\sigma(s, X_s)^2}\, \mathrm{d}s,$$

$$Z\big(\hat{\vartheta}_\varepsilon, t, X^t\big) = R\big(\hat{\vartheta}_\varepsilon, t, X^t\big) - \int_0^t \frac{\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s) S(\vartheta, s, X_s)}{\sigma(s, X_s)^2}\, \mathrm{d}s.$$

Then

$$K_\varepsilon(t) = \varepsilon^{-1}\big[Y_\varepsilon\big(\hat{\vartheta}_\varepsilon, t, X^t\big) + Z\big(\hat{\vartheta}_\varepsilon, t, X^t\big)\big].$$

We have

$$\begin{aligned} &S(\vartheta, s, X_s) - D(\hat{\vartheta}_\varepsilon, s, X_s) \\ &\quad = S(\vartheta, s, X_s) - S(\hat{\vartheta}_\varepsilon, s, X_s) + S(\hat{\vartheta}_\varepsilon, s, X_s) \\ &\qquad - S\big(\hat{\vartheta}_\varepsilon, s, x_s(\hat{\vartheta}_\varepsilon)\big) - S'(\hat{\vartheta}_\varepsilon, s, X_s)\big[X_s - x_s(\hat{\vartheta}_\varepsilon)\big] \\ &\quad = -(\hat{\vartheta}_\varepsilon - \vartheta)\dot{S}(\tilde{\vartheta}, s, X_s) \end{aligned}$$

$$
\begin{aligned}
&+\big[S'(\hat\vartheta_\varepsilon,s,\tilde X_s)-S'(\hat\vartheta_\varepsilon,s,X_s)\big]\big[X_s-x_s(\hat\vartheta_\varepsilon)\big]\\
&=-(\hat\vartheta_\varepsilon-\vartheta)\dot S(\tilde\vartheta,s,X_s)+\mathrm{O}\big(\varepsilon^2\big).
\end{aligned}
$$

Therefore

$$
\varepsilon^{-1}Y_\varepsilon\big(\hat\vartheta_\varepsilon,t,X^t\big)=-\frac{(\hat\vartheta_\varepsilon-\vartheta)}{\varepsilon}\int_0^t\frac{\dot S(\hat\vartheta_\varepsilon,s,X_s)^2}{\sigma(s,X_s)^2}\,\mathrm{d}s+\mathrm{o}(1).
$$

Further,

$$
\varepsilon^{-1}\big(Z\big(\hat\vartheta_\varepsilon,t,X^t\big)-Z\big(\vartheta,t,X^t\big)\big)=\frac{(\hat\vartheta_\varepsilon-\vartheta)}{\varepsilon}\dot Z\big(\vartheta,t,X^t\big)+\mathrm{o}(1),
$$

where

$$
\begin{aligned}
\dot Z\big(\vartheta,t,X^t\big)=&\int_{x_0}^{X_t}\frac{\ddot S(\vartheta,t,y)}{\sigma(t,y)^2}\,\mathrm{d}y-\int_0^t\frac{\ddot S(\vartheta,s,X_s)S(\vartheta,s,X_s)}{\sigma(s,X_s)^2}\,\mathrm{d}s\\
&-\int_0^t\int_{x_0}^{X_s}\frac{\ddot S_s'(\vartheta,s,y)\sigma(s,y)-2\ddot S(\vartheta,s,y)\sigma_s'(s,y)}{\sigma(s,y)^2}\,\mathrm{d}y\,\mathrm{d}s.
\end{aligned}
$$

We have uniform convergence of X_t to x_t w.r.t. t. Hence,

$$
\sup_{0\le t\le T}\big|\dot Z\big(\vartheta,t,X^t\big)-\dot Z\big(\vartheta,t,x^t\big)\big|\to 0.
$$

Note that for any continuously differentiable function $g(s,x)$ w.r.t. s we have the relation

$$
\int_{x_0}^{x_t}g(t,y)\,\mathrm{d}y-\int_0^t g(s,x_s)S(\vartheta,s,x_s)\,\mathrm{d}s-\int_0^t\int_{x_0}^{x_s}g_s'(s,y)\,\mathrm{d}y\,\mathrm{d}s=0
$$

since

$$
\int_0^t g(s,x_s)S(\vartheta,s,x_s)\,\mathrm{d}s=\int_0^t g(s,x_s)\,\mathrm{d}x_s
$$

and

$$
\begin{aligned}
\int_0^t g(t,x_s)\,\mathrm{d}x_s-\int_0^t g(s,x_s)\,\mathrm{d}x_s&=\int_0^t\int_s^t\frac{\partial g(v,x_s)}{\partial v}\,\mathrm{d}v\,\mathrm{d}x_s\\
&=\int_0^t\int_0^t\mathbb{1}_{\{v:x_v>x_s\}}\frac{\partial g(v,x_s)}{\partial v}\,\mathrm{d}v\,\mathrm{d}x_s\\
&=\int_0^t\int_{x_0}^{x_v}g_v'(v,y)\,\mathrm{d}y\,\mathrm{d}v.
\end{aligned}
$$

Hence, $\dot Z(\vartheta,t,x^t)\equiv 0$ for all $t\in[0,T]$.

By the Itô formula,

$$\begin{aligned}\frac{Z(\vartheta,t,X^t)}{\varepsilon}&=\frac{R(\vartheta,t,X^t)}{\varepsilon}-\int_0^t\frac{\dot S(\vartheta,s,X_s)S(\vartheta,s,X_s)}{\varepsilon\sigma(s,X_s)^2}\,\mathrm{d}s\\&=\int_0^t\frac{\dot S(\vartheta,s,X_s)}{\varepsilon\sigma(s,X_s)^2}\,\mathrm{d}X_s-\int_0^t\frac{\dot S(\vartheta,s,X_s)S(\vartheta,s,X_s)}{\varepsilon\sigma(s,X_s)^2}\,\mathrm{d}s\\&\quad+\frac{\varepsilon}{2}\int_0^t\sigma(s,X_s)^2M''_{xx}(\vartheta,s,X_s)\,\mathrm{d}s\\&=\int_0^t\frac{\dot S(\vartheta,s,X_s)}{\sigma(s,X_s)}\,\mathrm{d}W_s+\mathrm{O}(\varepsilon).\end{aligned}$$

Therefore, we obtain the convergence

$$K_\varepsilon(t)\longrightarrow K(\vartheta,t).$$

This convergence can be shown to be uniform w.r.t. t. This proves the convergence $\delta_\varepsilon\to\delta$. Therefore the Theorem 1 is proved. □

Let us study the behaviour of the power function under the alternative. Suppose that the observed diffusion process (1.1) has the trend coefficient $S(t,x)$ which does not belong to the parametric family. This family we described as follows:

$$\mathcal{F}=\{S(\cdot):S(\vartheta,t,x_t(\vartheta)),0\le t\le T,\vartheta\in\Theta\}.$$

Here $x_t(\vartheta)$, $0\le t\le T$ is the solution of equation (2.2).

We introduce a slightly more strong condition of separability of the basic hypothesis and the alternative. Suppose that the function $S(t,x)$ satisfies conditions $\mathcal{C}1$, $\mathcal{C}2$ and denote by y_t, $0\le t\le T$ the solution of the ordinary differential equation obtained for $(\varepsilon=0)$

$$\frac{\mathrm{d}y_t}{\mathrm{d}t}=S(t,y_t),\qquad y_0=x_0.$$

Then

$$\begin{aligned}\varepsilon^{-1}\big(X_t-x_t(\hat\vartheta_\varepsilon)\big)&=\varepsilon^{-1}(X_t-y_t)+\varepsilon^{-1}\big(y_t-x_t(\hat\vartheta_\varepsilon)\big)\\&=y_t^{(1)}+\varepsilon^{-1}\big(y_t-x_t(\vartheta_*)\big)-\varepsilon^{-1}(\hat\vartheta_\varepsilon-\vartheta_*)\dot x_t(\vartheta_*)+\mathrm{o}(1),\end{aligned}$$

where $y_t^{(1)}$ is a solution of the equation

$$\mathrm{d}y_t^{(1)}=S'(t,y_t)y_t^{(1)}\,\mathrm{d}t+\sigma(t,y_t)\,\mathrm{d}W_t,\qquad y_0^{(1)}=0$$

and ϑ_* is defined by the relation

$$\inf_{\vartheta\in\Theta}\int_0^T\left(\frac{S(\vartheta,t,y_t)-S(t,y_t)}{\sigma(t,y_t)}\right)^2\mathrm{d}t=\int_0^T\left(\frac{S(\vartheta_*,t,y_t)-S(t,y_t)}{\sigma(t,y_t)}\right)^2\mathrm{d}t.\tag{3.5}$$

Suppose that this equation has a unique solution ϑ_*. Note that $\varepsilon^{-1}(\hat{\vartheta}_\varepsilon - \vartheta_*)$ is tight (see Kutoyants [8] for details). Moreover, we also suppose that the basic hypothesis and the alternative are separated in the following sense:

$$\inf_{\vartheta\in\Theta}\int_0^T\left(\frac{S(\vartheta,t,y_t)-S(t,y_t)}{\sigma(t,y_t)}\right)^2\,\mathrm{d}t>0.$$

First, formally, we write

$$\begin{aligned}
&\int_0^t h_\varepsilon(\hat{\vartheta}_\varepsilon,s)\,\mathrm{d}U_\varepsilon(s)\\
&\quad=\int_0^t\frac{\dot{S}(\hat{\vartheta}_\varepsilon,s,X_s)}{\sqrt{\mathrm{I}(\hat{\vartheta}_\varepsilon)}\sigma(s,X_s)}\,\mathrm{d}W_s-\int_0^t\frac{\dot{S}(\hat{\vartheta}_\varepsilon,s,X_s)[S(s,X_s)-D(\hat{\vartheta}_\varepsilon,s,X_s)]}{\sqrt{\mathrm{I}(\hat{\vartheta}_\varepsilon)}\varepsilon\sigma(s,X_s)^2}\,\mathrm{d}s\\
&\quad=\int_0^t\frac{\dot{S}(\vartheta_*,s,y_s)}{\sqrt{\mathrm{I}(\vartheta_*)}\sigma(s,y_s)}\,\mathrm{d}W_s-\int_0^t\frac{\dot{S}(\vartheta_*,s,X_s)[S(s,X_s)-S(\vartheta_*,s,X_s)]}{\sqrt{\mathrm{I}(\vartheta_*)}\varepsilon\sigma(s,X_s)^2}\,\mathrm{d}s.
\end{aligned}$$

Further

$$\begin{aligned}
&S(s,X_s)-D(\hat{\vartheta}_\varepsilon,s,X_s)\\
&\quad=S(s,X_s)-S\bigl(\hat{\vartheta}_\varepsilon,s,x_s(\vartheta)\bigr)-S'(\hat{\vartheta}_\varepsilon,s,X_s)\bigl(X_s-x_s(\hat{\vartheta}_\varepsilon)\bigr)\\
&\quad=S(s,X_s)-S(\hat{\vartheta}_\varepsilon,s,X_s)+\mathrm{O}\bigl(\varepsilon^2\bigr)\\
&\quad=S(s,X_s)-S(\vartheta_*,s,X_s)+S(\vartheta_*,s,X_s)-S(\hat{\vartheta}_\varepsilon,s,X_s)+\mathrm{O}\bigl(\varepsilon^2\bigr)\\
&\quad=S(s,X_s)-S(\vartheta_*,s,X_s)+(\hat{\vartheta}_\varepsilon-\vartheta_*)\dot{S}(\vartheta_*,s,X_s)+\mathrm{O}\bigl(\varepsilon^2\bigr).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\int_0^t h_\varepsilon(\hat{\vartheta}_\varepsilon,s)\,\mathrm{d}U_\varepsilon(s)&=\int_0^t\frac{\dot{S}(\vartheta_*,s,y_s)}{\sqrt{\mathrm{I}(\vartheta_*)}\sigma(s,y_s)}\,\mathrm{d}W_s-\int_0^t\frac{(\hat{\vartheta}_\varepsilon-\vartheta_*)\dot{S}(\vartheta_*,s,y_s)^2}{\varepsilon\sqrt{\mathrm{I}(\vartheta_*)}\sigma(s,y_s)^2}\,\mathrm{d}s\\
&\quad-\varepsilon^{-1}\int_0^t\frac{\dot{S}(\vartheta_*,s,y_s)[S(s,y_s)-S(\vartheta_*,s,y_s)]}{\sqrt{\mathrm{I}(\vartheta_*)}\sigma(s,y_s)^2}\,\mathrm{d}s+\mathrm{O}\bigl(\varepsilon^2\bigr)\\
&=I_1(t)-I_2(t)-\varepsilon^{-1}I_3(t)+\mathrm{O}\bigl(\varepsilon^2\bigr)
\end{aligned}$$

with an obvious notation. For the statistic δ_ε we have the relations

$$\sqrt{\delta_\varepsilon}\ge\varepsilon^{-1}\bigl\|I_3(\cdot)h(\cdot)\bigr\|-\bigl\|I_1(\cdot)h(\cdot)\bigr\|-\bigl\|I_2(\cdot)h(\cdot)\bigr\|+\mathrm{O}(\varepsilon),\tag{3.6}$$

where $h(\cdot)=h(\vartheta_*,s)$ and $\|\cdot\|$ is the $L_2(0,T)$ norm. Recall that the quantities $\|I_1(\cdot)h(\cdot)\|$ and $\|I_2(\cdot)h(\cdot)\|$ are bounded in probability.

Introduce the condition

$\mathcal{C}5$. *The functions* $S(\vartheta, t, x)$, $S(t, x)$ *and* $\sigma(t, x)$ *are such that*

$$\begin{aligned}&\|I_3(\cdot)h(\cdot)\|^2\\&\quad=\int_0^T\left(\int_0^t\frac{\dot S(\vartheta_*, s, y_s)[S(s, y_s)-S(\vartheta_*, s, y_s)]}{\mathrm{I}(\vartheta_*)\sigma(s, y_s)^2}\,\mathrm{d}s\right)^2\frac{\dot S(\vartheta_*, t)^2}{\sigma(t, y_t)^2}\,\mathrm{d}t>0.\end{aligned}$$

This condition provides consistency of the test.

Theorem 2. *Let conditions* $\mathcal{C}1$–$\mathcal{C}5$ *hold. Then the test* $\hat\psi_\varepsilon$ *is consistent.*

Proof. The proof follows from the convergence $\delta_\varepsilon \to \infty$ under alternative (see (3.6)). □

Note that if ϑ_* is an interior point of Θ, then

$$\int_0^T\frac{\dot S(\vartheta_*, s, y_s)[S(s, y_s)-S(\vartheta_*, s, y_s)]}{\sigma(s, y_s)^2}\,\mathrm{d}s=0.$$

If condition $\mathcal{C}5$ does not hold, then

$$\int_0^t\frac{\dot S(\vartheta_*, s, y_s)[S(s, y_s)-S(\vartheta_*, s, y_s)]}{\sigma(s, y_s)^2}\,\mathrm{d}s\equiv 0,\qquad\text{for all } t\in[0, T].$$

This equality is possible if

$$\dot S(\vartheta_*, s, y_s)\big[S(s, y_s)-S(\vartheta_*, s, y_s)\big]\equiv 0,\qquad\text{for all } s\in[0, T].$$

An example of such *invisible* alternative can be constructed as follows: Suppose that the function $S(\vartheta, s, x)$ does not depend on ϑ for $s \in [0, T/2]$, that is, $S(\vartheta, s, x) = S_*(s, x)$ for all $\vartheta \in \Theta$. Then $\dot S(\vartheta_*, s, y_s) \equiv 0$ for $s \in [0, T/2]$. Therefore if $S(s, y_s) = S(\vartheta_*, s, y_s)$ for $s \in [T/2, T]$ and a corresponding ϑ_* then condition $\mathcal{C}5$ does not hold, but we can have $S(s, y_s) \neq S_*(s, y_s)$ for $s \in [0, T/2]$. This implies that the test $\hat\psi_\varepsilon$ is not consistent for this alternative.

3.2. Second test

The second test is based on the following well-known transformation. Suppose that we have a Gaussian process $U(t)$, $0 \le t \le 1$ and $d \times d$ matrix $\mathbb{N}(t)$ defined by the relations

$$U(t)=W_t-\left\langle\int_0^1\mathbf{h}(s)\,\mathrm{d}W_s, \int_0^t\mathbf{h}(s)\,\mathrm{d}s\right\rangle,\tag{3.7}$$

$$\mathbb{N}(t)=\int_t^1\mathbf{h}(s)\mathbf{h}(s)^*\,\mathrm{d}s,\qquad\mathbb{N}(0)=\mathbb{J},\tag{3.8}$$

where $\mathbb{J}$ is the $d \times d$ unit matrix and $\mathbf{h}(t)$ is a continuous vector-valued function.

Lemma 2. *Suppose that the matrix $\mathbb{N}(t)$ is non-degenerate for all $t \in [0,1)$. Then*

$$U(t) + \int_0^t \mathbf{h}(s)^* \mathbb{N}(s)^{-1} \int_0^s \mathbf{h}(v)\, \mathrm{d}U(v)\, \mathrm{d}s = w(t), \qquad 0 \le t \le 1, \tag{3.9}$$

where $w(\cdot)$ is a Wiener process.

Proof. This formula was obtained by Khmaladze [6]. The proof there is based on two results: a result of Hitsuda [4] and another one of Shepp [16]. Observe that there are many publications dealing with this transformation (see, e.g., the paper Maglaperidze *et al.* [15] and the references therein). Another direct proof is given in Kleptsyna and Kutoyants [7]. □

Note that representation (3.7) and (3.8) implies that

$$\int_0^1 \mathbf{h}(s)\, \mathrm{d}U(s) = 0. \tag{3.10}$$

Suppose that $\vartheta \in \Theta$. Here Θ is an open bounded subset of $\mathscr{R}^d$. Now $\mathbf{h}(\vartheta, s)$, $\mathbf{R}(\vartheta, t, X^t)$ and $\mathbf{Q}(\vartheta, t, X^t)$ are d-vectors and the Fisher information $\mathbb{I}(\vartheta)$ is a $d \times d$ matrix.

Introduce the following stochastic processes:

$$\bar{\mathbf{h}}_\varepsilon(\vartheta, t) = \frac{\dot{\mathbf{S}}(\vartheta, t, X_t)}{\sigma(t, X_t)},$$

$$\bar{\mathbb{N}}(\vartheta, t) = \int_t^T \frac{\dot{\mathbf{S}}(\vartheta, s, x_s)\dot{\mathbf{S}}(\vartheta, s, x_s)^*}{\sigma(s, x_s)^2}\, \mathrm{d}s,$$

$$\bar{\mathbb{N}}_\varepsilon(\vartheta, t) = \int_t^T \frac{\dot{\mathbf{S}}(\vartheta, s, X_s)\dot{\mathbf{S}}(\vartheta, s, X_s)^*}{\sigma(s, X_s)^2}\, \mathrm{d}s,$$

and put

$$\Delta_\varepsilon = \frac{1}{T^2} \int_0^T W_\varepsilon(t)^2\, \mathrm{d}t.$$

Here

$$\begin{aligned} W_\varepsilon(t) = {} & \int_0^t \frac{\mathrm{d}X_s}{\varepsilon\sigma(s, X_s)} - \int_0^t \frac{D(\hat{\vartheta}_\varepsilon, s, X_s)}{\varepsilon\sigma(s, X_s)}\, \mathrm{d}s \\ & + \varepsilon^{-1} \int_0^t \bar{\mathbf{h}}_\varepsilon(\hat{\vartheta}_\varepsilon, s)^* \bar{\mathbb{N}}_\varepsilon(\hat{\vartheta}_\varepsilon, s)_+^{-1} \bigl[\mathbf{R}\bigl(\hat{\vartheta}_\varepsilon, s, X^s\bigr) - \mathbf{Q}\bigl(\hat{\vartheta}_\varepsilon, s, X^s\bigr)\bigr]\, \mathrm{d}s. \end{aligned} \tag{3.11}$$

We use the following convention for the matrix $\bar{\mathbb{N}}$:

$$\bar{\mathbb{N}}_+^{-1} = \begin{cases} \bar{\mathbb{N}}^{-1}, & \text{if } \bar{\mathbb{N}} \text{ is non-degenerate,} \\ 0, & \text{if } \bar{\mathbb{N}} \text{ is degenerate.} \end{cases}$$

We have the following result.

Theorem 3. *Suppose that conditions $\mathcal{C}2$–$\mathcal{C}4$ hold and the matrix $\bar{\mathbb{N}}(\vartheta, t)$ is uniformly in $\vartheta \in \Theta$ non-degenerate for all $t \in [0, 1)$. Then the test*

$$\hat{\Psi}_\varepsilon = \mathbb{1}_{\{\Delta_\varepsilon > c_\alpha\}}, \qquad \mathbf{P}\left(\int_0^1 w(s)^2 \,\mathrm{d}s > c_\alpha\right) = \alpha$$

is ADF and belongs to $\mathcal{K}_\alpha$.

Proof. We have to show that under hypothesis $\mathcal{H}_0$ the convergence

$$\Delta_\varepsilon \Longrightarrow \Delta = \int_0^1 w(s)^2 \,\mathrm{d}s \tag{3.12}$$

holds.

The construction of the ADF GoF test is based on Lemmas 1 and 2. We have the similar to (2.6) presentation (3.7) with $\mathbf{h}(\vartheta, t)$ defined in (2.7). Let us denote $U_\varepsilon(\cdot)$, $\mathbf{h}_\varepsilon(\hat{\vartheta}_\varepsilon, \cdot)$, and $\mathbb{N}_\varepsilon(\cdot)$ the *empirical versions* of $U(\cdot)$, $\mathbf{h}(\vartheta, \cdot)$ and

$$\mathbb{N}(\vartheta, t) = \mathbb{I}(\vartheta)^{-1} \int_t^T \frac{\dot{\mathbf{S}}(\vartheta, s, x_s)\dot{\mathbf{S}}(\vartheta, s, x_s)^*}{\sigma(s, x_s)^2} \,\mathrm{d}s, \qquad \mathbb{N}(\vartheta, 0) = \mathbb{J},$$

respectively:

$$\begin{aligned}
U_\varepsilon(t) &= \int_0^t \frac{\psi_\varepsilon(s)}{\sigma(s, X_s)} \,\mathrm{d}V_\varepsilon(s),\\
V_\varepsilon(t) &= \frac{X_t - x_t(\hat{\vartheta}_\varepsilon)}{\psi_\varepsilon(t)\varepsilon},\\
\mathbf{h}_\varepsilon(\hat{\vartheta}_\varepsilon, t) &= \mathbb{I}_\varepsilon(\hat{\vartheta}_\varepsilon)^{-1/2} \frac{\dot{\mathbf{S}}(\hat{\vartheta}_\varepsilon, t, X_t)}{\sigma(t, X_t)},\\
\mathbb{I}_\varepsilon(\hat{\vartheta}_\varepsilon) &= \int_0^T \frac{\dot{\mathbf{S}}(\hat{\vartheta}_\varepsilon, s, X_s)\dot{\mathbf{S}}(\hat{\vartheta}_\varepsilon, s, X_s)^*}{\sigma(s, X_s)^2} \,\mathrm{d}s,\\
\mathbb{N}_\varepsilon(\hat{\vartheta}_\varepsilon, t) &= \mathbb{I}_\varepsilon(\hat{\vartheta}_\varepsilon)^{-1} \int_t^T \frac{\dot{\mathbf{S}}(\hat{\vartheta}_\varepsilon, s, X_s)\dot{\mathbf{S}}(\hat{\vartheta}_\varepsilon, s, X_s)^*}{\sigma(s, X_s)^2} \,\mathrm{d}s.
\end{aligned}$$

Recall that there is a problem of definition of the integral for $U_\varepsilon(\cdot)$ because the integrand depends on the future. As convergence is uniform w.r.t. $t \in [0, T - \nu]$:

$$\mathbf{h}_\varepsilon(\hat{\vartheta}_\varepsilon, t) \longrightarrow \mathbf{h}(\vartheta, t), \qquad \mathbb{I}_\varepsilon(\hat{\vartheta}_\varepsilon) \longrightarrow \mathbb{I}(\vartheta), \qquad \mathbb{N}_\varepsilon(\hat{\vartheta}_\varepsilon, t) \longrightarrow \mathbb{N}(\vartheta, t).$$

The required limits can be obtained.

Introduce (formally) the statistic

$$W_\varepsilon^\star(t) = U_\varepsilon(t) + \int_0^t \mathbf{h}_\varepsilon(\hat{\vartheta}_\varepsilon, s)^* \mathbb{N}_\varepsilon(\hat{\vartheta}_\varepsilon, s)_+^{-1} \int_0^s \mathbf{h}_\varepsilon(\hat{\vartheta}_\varepsilon, v) \,\mathrm{d}U_\varepsilon(v) \,\mathrm{d}s. \tag{3.13}$$

Observe that

$$\begin{aligned}&\mathbf{h}(\vartheta,s)^*\mathbb{N}(\vartheta,s)^{-1}\mathbf{h}(\vartheta,v)\\&\quad=\frac{\dot{\mathbf{S}}(\vartheta,s,x_s)^*}{\sigma(s,x_s)}\left(\int_s^T\frac{\dot{\mathbf{S}}(\vartheta,r,x_r)\dot{\mathbf{S}}(\vartheta,r,x_r)^*}{\sigma(r,x_r)^2}\,\mathrm{d}r\right)^{-1}\frac{\dot{\mathbf{S}}(\vartheta,v,x_v)}{\sigma(v,x_v)}.\end{aligned}$$

Therefore this term does not depend on the information matrix $\mathbb{I}(\vartheta)$ and we can replace the statistics $\mathbf{h}_\varepsilon(\hat{\vartheta}_\varepsilon,s)$ and $\mathbb{N}_\varepsilon(\hat{\vartheta}_\varepsilon,s)$ in (3.13) by $\bar{\mathbf{h}}_\varepsilon(\hat{\vartheta}_\varepsilon,s)$ and $\bar{\mathbb{N}}_\varepsilon(\hat{\vartheta}_\varepsilon,s)$.

For the process $U_\varepsilon(\cdot)$, we have equality (3.4) (formally)

$$U_\varepsilon(t)=\int_0^t\frac{\mathrm{d}X_s}{\varepsilon\sigma(s,X_s)}-\int_0^t\frac{D(\hat{\vartheta}_\varepsilon,s,X_s)}{\varepsilon\sigma(s,X_s)}\,\mathrm{d}s.$$

Hence, we obtain the vector-valued integral

$$\int_0^t\bar{\mathbf{h}}_\varepsilon(\hat{\vartheta}_\varepsilon,s)\,\mathrm{d}U_\varepsilon(s)=\int_0^t\frac{\dot{\mathbf{S}}(\hat{\vartheta}_\varepsilon,s,X_s)}{\varepsilon\sigma(s,X_s)^2}\,\mathrm{d}X_s-\int_0^t\frac{\dot{\mathbf{S}}(\hat{\vartheta}_\varepsilon,s,X_s)D(\hat{\vartheta}_\varepsilon,s,X_s)}{\varepsilon\sigma(s,X_s)^2}\,\mathrm{d}s.$$

Introduce the vector-function

$$\mathbf{M}(\vartheta,t,x)=\int_{x_0}^x\frac{\dot{\mathbf{S}}(\vartheta,t,y)}{\sigma(t,y)^2}\,\mathrm{d}y.$$

Then by the Itô formula

$$\begin{aligned}\int_0^t\frac{\dot{\mathbf{S}}(\vartheta,s,X_s)}{\sigma(s,X_s)^2}\,\mathrm{d}X_s&=\int_{x_0}^{X_t}\frac{\dot{\mathbf{S}}(\vartheta,t,y)}{\sigma(t,y)^2}\,\mathrm{d}y-\int_0^t\int_{x_0}^{X_s}\frac{\dot{\mathbf{S}}'_s(\vartheta,s,y)}{\sigma(s,y)^2}\,\mathrm{d}s\\&\quad+\int_0^t\int_{x_0}^{X_s}\frac{2\dot{\mathbf{S}}(\vartheta,s,y)\sigma'_s(s,y)}{\sigma(s,y)^3}\,\mathrm{d}s+\mathrm{O}(\varepsilon^2).\end{aligned}$$

Put

$$\mathbf{K}_\varepsilon(t)=\int_0^t\bar{\mathbf{h}}_\varepsilon(\hat{\vartheta}_\varepsilon,s)\,\mathrm{d}U_\varepsilon(s)=\varepsilon^{-1}\big[\mathbf{R}\big(\hat{\vartheta}_\varepsilon,t,X^t\big)-\mathbf{Q}\big(\hat{\vartheta}_\varepsilon,t,X^t\big)\big].$$

Note that we have dropped the term of order $\mathrm{O}(\varepsilon^2)$.

Then formal expression (3.13) for $W_\varepsilon^\star(t)$ can be replaced by (3.11)

$$\begin{aligned}W_\varepsilon(t)&=\int_0^t\frac{\mathrm{d}X_s}{\varepsilon\sigma(s,X_s)}-\int_0^t\frac{D(\hat{\vartheta}_\varepsilon,s,X_s)}{\varepsilon\sigma(s,X_s)}\,\mathrm{d}s\\&\quad+\varepsilon^{-1}\int_0^t\bar{\mathbf{h}}_\varepsilon(\hat{\vartheta}_\varepsilon,s)^*\bar{\mathbb{N}}_\varepsilon(\hat{\vartheta}_\varepsilon,s)_+^{-1}\big[R\big(\hat{\vartheta}_\varepsilon,s,X^s\big)-Q\big(\hat{\vartheta}_\varepsilon,s,X^s\big)\big]\,\mathrm{d}s.\end{aligned}$$

For the first two terms of $W_\varepsilon(t)$ we have

$$\begin{aligned}
U_\varepsilon(t) &= \int_0^t \frac{\mathrm{d}X_s}{\varepsilon\sigma(s,X_s)} - \int_0^t \frac{D(\hat\vartheta_\varepsilon,s,X_s)}{\varepsilon\sigma(s,X_s)}\,\mathrm{d}s \\
&= W_t + \int_0^t \frac{S(\vartheta,s,X_s) - S(\hat\vartheta_\varepsilon,s,x_s(\hat\vartheta_\varepsilon)) - S'(\hat\vartheta_\varepsilon,s,X_s)(X_s - x_s(\hat\vartheta_\varepsilon))}{\varepsilon\sigma(s,X_s)}\,\mathrm{d}s \\
&= W_t - \left\langle \frac{\hat\vartheta_\varepsilon - \vartheta}{\varepsilon}, \int_0^t \frac{\dot S(\tilde\vartheta,s,X_s)}{\sigma(s,X_s)}\,\mathrm{d}s \right\rangle \\
&\quad + \int_0^t \frac{[S'(\hat\vartheta_\varepsilon,s,\tilde X_s) - S'(\hat\vartheta_\varepsilon,s,X_s)](X_s - x_s(\hat\vartheta_\varepsilon))}{\varepsilon\sigma(s,X_s)}\,\mathrm{d}s \\
&= W_t - \left\langle \mathrm{I}(\vartheta)^{-1} \int_0^T \frac{\dot{\mathbf{S}}(\vartheta,s,x_s)}{\sigma(s,x_s)}\,\mathrm{d}W_s, \int_0^t \frac{\dot{\mathbf{S}}(\vartheta,s,x_s)}{\sigma(s,x_s)}\,\mathrm{d}s \right\rangle + \mathrm{o}(1) \\
&= U(\vartheta,t) + \mathrm{o}(1).
\end{aligned}$$

Here $|\tilde\vartheta - \vartheta| \le |\hat\vartheta_\varepsilon|$ and

$$\begin{aligned}
|\tilde X_s - X_s| &\le \big|x_s(\hat\vartheta_\varepsilon) - X_s\big| \\
&\le \big|x_s(\hat\vartheta_\varepsilon) - x_s(\vartheta)\big| + \big|x_s(\vartheta) - X_s\big| \to 0.
\end{aligned}$$

This convergence is uniform w.r.t. $s \in [0,T]$. Hence,

$$\sup_{0\le t\le T} \big|U_\varepsilon(t) - U(\vartheta,t)\big| \longrightarrow 0.$$

Further, similar arguments give the uniform convergence w.r.t. $t \in [0,T]$

$$\bar{\mathbf{h}}_\varepsilon(\hat\vartheta_\varepsilon,t) = \frac{\dot{\mathbf{S}}(\hat\vartheta_\varepsilon,t,X_t)}{\sigma(t,X_t)} \to \bar{\mathbf{h}}(\vartheta,t), \qquad \bar{\mathbb{N}}_\varepsilon(\hat\vartheta_\varepsilon,t) \to \bar{\mathbb{N}}(\vartheta,t).$$

We have to show that $\mathbf{K}_\varepsilon(t) \longrightarrow \mathbf{K}(\vartheta,t)$, where

$$\mathbf{K}(\vartheta,t) = \int_0^t \bar{\mathbf{h}}(\vartheta,s)\,\mathrm{d}W_s - \int_0^T \bar{\mathbf{h}}(\vartheta,s)\,\mathrm{d}W_s \int_0^t \bar{\mathbf{h}}(\vartheta,s)\bar{\mathbf{h}}(\vartheta,s)^*\,\mathrm{d}s.$$

Denote

$$\begin{aligned}
\mathbf{Y}_\varepsilon\big(\hat\vartheta_\varepsilon,t,X^t\big) &= \int_0^t \frac{\dot{\mathbf{S}}(\hat\vartheta_\varepsilon,s,X_s)[S(\vartheta,s,X_s) - D(\hat\vartheta_\varepsilon,s,X_s)]}{\sigma(s,X_s)^2}\,\mathrm{d}s, \\
\mathbf{Z}\big(\hat\vartheta_\varepsilon,t,X^t\big) &= \mathbf{R}\big(\hat\vartheta_\varepsilon,t,X^t\big) - \int_0^t \frac{\dot{\mathbf{S}}(\hat\vartheta_\varepsilon,s,X_s)S(\vartheta,s,X_s)}{\sigma(s,X_s)^2}\,\mathrm{d}s.
\end{aligned}$$

Then

$$\mathbf{K}_\varepsilon(t)=\varepsilon^{-1}\big[\mathbf{Y}_\varepsilon\big(\hat\vartheta_\varepsilon,t,X^t\big)+\mathbf{Z}\big(\hat\vartheta_\varepsilon,t,X^t\big)\big].$$

We have

$$\begin{aligned}
&S(\vartheta,s,X_s)-D(\hat\vartheta_\varepsilon,s,X_s)\\
&\quad=S(\vartheta,s,X_s)-S(\hat\vartheta_\varepsilon,s,X_s)+S(\hat\vartheta_\varepsilon,s,X_s)\\
&\qquad-S\big(\hat\vartheta_\varepsilon,s,x_s(\hat\vartheta_\varepsilon)\big)-S'(\hat\vartheta_\varepsilon,s,X_s)\big[X_s-x_s(\hat\vartheta_\varepsilon)\big]\\
&\quad=-\big\langle(\hat\vartheta_\varepsilon-\vartheta),\dot{\mathbf{S}}(\tilde\vartheta,s,X_s)\big\rangle\\
&\qquad+\big[S'(\hat\vartheta_\varepsilon,s,\tilde X_s)-S'(\hat\vartheta_\varepsilon,s,X_s)\big]\big[X_s-x_s(\hat\vartheta_\varepsilon)\big]\\
&\quad=-\big\langle(\hat\vartheta_\varepsilon-\vartheta),\dot{\mathbf{S}}(\tilde\vartheta,s,X_s)\big\rangle+\mathrm{O}\big(\varepsilon^2\big).
\end{aligned}$$

Therefore

$$\varepsilon^{-1}\mathbf{Y}_\varepsilon\big(\hat\vartheta_\varepsilon,t,X^t\big)=-\frac{(\hat\vartheta_\varepsilon-\vartheta)}{\varepsilon}\int_0^t\frac{\dot{\mathbf{S}}(\hat\vartheta_\varepsilon,s,X_s)\dot{\mathbf{S}}(\hat\vartheta_\varepsilon,s,X_s)^*}{\sigma(s,X_s)^2}\,\mathrm{d}s.$$

Further,

$$\varepsilon^{-1}\big(\mathbf{Z}\big(\hat\vartheta_\varepsilon,t,X^t\big)-\mathbf{Z}\big(\vartheta,t,X^t\big)\big)=\frac{\hat\vartheta_\varepsilon-\vartheta}{\varepsilon}\dot{\mathbb{Z}}\big(\vartheta,t,X^t\big)+\mathrm{o}(1),$$

where

$$\begin{aligned}
\dot{\mathbb{Z}}\big(\vartheta,t,X^t\big)&=\int_{x_0}^{X_t}\frac{\ddot{\mathbb{S}}(\vartheta,t,y)}{\sigma(t,y)^2}\,\mathrm{d}y-\int_0^t\frac{\ddot{\mathbb{S}}(\vartheta,s,X_s)S(\vartheta,s,X_s)}{\sigma(s,X_s)^2}\,\mathrm{d}s\\
&\quad-\int_0^t\int_{x_0}^{X_s}\frac{\ddot{\mathbb{S}}_s'(\vartheta,s,y)\sigma(s,y)-2\ddot{\mathbb{S}}(\vartheta,s,y)\sigma_s'(s,y)}{\sigma(s,y)^2}\,\mathrm{d}y\,\mathrm{d}s.
\end{aligned}$$

Here $\ddot{\mathbb{S}}(\cdot)$ is the matrix of second derivatives w.r.t. ϑ. We have uniform convergence of X_t to x_t w.r.t. t, hence

$$\sup_{0\le t\le T}\big|\dot{\mathbb{Z}}\big(\vartheta,t,X^t\big)-\dot{\mathbb{Z}}\big(\vartheta,t,x^t\big)\big|\to 0.$$

Observe that for any continuously differentiable function $g(s,x)$ w.r.t. s we have

$$\int_{x_0}^{x_t}g(t,y)\,\mathrm{d}y-\int_0^t g(s,x_s)S(\vartheta,s,x_s)\,\mathrm{d}s-\int_0^t\int_{x_0}^{x_s}g_s'(s,y)\,\mathrm{d}y\,\mathrm{d}s=0$$

since

$$\int_0^t g(s,x_s)S(\vartheta,s,x_s)\,\mathrm{d}s=\int_0^t g(s,x_s)\,\mathrm{d}x_s$$

and

$$\begin{aligned}
&\int_0^t g(t,x_s)\,\mathrm{d}x_s - \int_0^t g(s,x_s)\,\mathrm{d}x_s \\
&\quad = \int_0^t \int_s^t \frac{\partial g(v,x_s)}{\partial v}\,\mathrm{d}v\,\mathrm{d}x_s \\
&\quad = \int_0^t \int_0^t \mathbb{1}_{\{v:x_v>x_s\}} \frac{\partial g(v,x_s)}{\partial v}\,\mathrm{d}v\,\mathrm{d}x_s = \int_0^t \int_{x_0}^{x_v} g_v'(v,y)\,\mathrm{d}y\,\mathrm{d}v.
\end{aligned}$$

Hence, $\dot{\mathbb{Z}}(\vartheta,t,x^t) \equiv 0$ for all $t \in [0,T]$.

By the Itô formula

$$\begin{aligned}
\frac{\mathbf{Z}(\vartheta,t,X^t)}{\varepsilon} &= \frac{\mathbf{R}(\vartheta,t,X^t)}{\varepsilon} - \int_0^t \frac{\dot{\mathbf{S}}(\vartheta,s,X_s)S(\vartheta,s,X_s)}{\varepsilon\sigma(s,X_s)^2}\,\mathrm{d}s \\
&= \int_0^t \frac{\dot{\mathbf{S}}(\vartheta,s,X_s)}{\varepsilon\sigma(s,X_s)^2}\,\mathrm{d}X_s - \int_0^t \frac{\dot{\mathbf{S}}(\vartheta,s,X_s)S(\vartheta,s,X_s)}{\varepsilon\sigma(s,X_s)^2}\,\mathrm{d}s \\
&\quad + \frac{\varepsilon}{2}\int_0^t \sigma(s,X_s)^2 \mathbf{M}_{xx}''(\vartheta,s,X_s)\,\mathrm{d}s \\
&= \int_0^t \frac{\dot{\mathbf{S}}(\vartheta,s,X_s)}{\sigma(s,X_s)}\,\mathrm{d}W_s + \mathrm{O}(\varepsilon).
\end{aligned}$$

Therefore, we obtain the convergence

$$\begin{aligned}
\mathbf{K}_\varepsilon(t) &= \varepsilon^{-1}\bigl(\mathbf{R}\bigl(\hat{\vartheta}_\varepsilon,t,X^t\bigr) - \mathbf{Q}\bigl(\hat{\vartheta}_\varepsilon,t,X^t\bigr)\bigr) \\
&= \varepsilon^{-1}\bigl(\mathbf{Y}\bigl(\hat{\vartheta}_\varepsilon,t,X^t\bigr) + \mathbf{Z}\bigl(\hat{\vartheta}_\varepsilon,t,X^t\bigr)\bigr) \longrightarrow \mathbf{K}(\vartheta,t).
\end{aligned}$$

Further, the matrix $\bar{\mathbb{N}}_\varepsilon(\hat{\vartheta}_\varepsilon,s)$ converges uniformly in $s \in [0,T]$ to the matrix $\bar{N}(\vartheta,s)$. Therefore, for $\nu > 0$ we have uniform on $s \in [0,T-\nu]$ convergence of $\bar{\mathbb{N}}_\varepsilon(\hat{\vartheta}_\varepsilon,s)_+^{-1}$ to $N(\vartheta,s)^{-1}$. Introduce the random function

$$y_\varepsilon(s) = \varepsilon^{-1}\bar{\mathbf{h}}_\varepsilon(\hat{\vartheta}_\varepsilon,s)^*\bar{\mathbb{N}}_\varepsilon(\hat{\vartheta}_\varepsilon,s)_+^{-1}\bigl[\mathbf{R}\bigl(\hat{\vartheta}_\varepsilon,s,X^s\bigr) - \mathbf{Q}\bigl(\hat{\vartheta}_\varepsilon,s,X^s\bigr)\bigr].$$

It is shown that we have convergence

$$\sup_{0\le s\le T-\nu}\bigl|y_\varepsilon(s) - y(\vartheta,s)\bigr| \longrightarrow 0,$$

where

$$y(\vartheta,s) = \bar{\mathbf{h}}(\vartheta,s)^*\bar{\mathbb{N}}(\vartheta,s)^{-1}\mathbf{K}(\vartheta,s).$$

Hence we also have convergence for all $t \in [0,1)$

$$W_\varepsilon(t) \longrightarrow U(\vartheta,t) + \int_0^t \bar{\mathbf{h}}(\vartheta,s)^*\bar{N}(\vartheta,s)^{-1}\mathbf{K}(\vartheta,s)\,\mathrm{d}s = w(t).$$

A similar argument can show that for any $0 \le t_1 < \cdots < t_k \le T$ we have convergence of the vectors

$$\big(W_\varepsilon(t_1), \ldots, W_\varepsilon(t_k)\big) \Longrightarrow \big(w(t_1), \ldots, w(t_k)\big).$$

Further, a direct but cumbersome calculation allows us to write the estimate

$$\mathbf{E}_\vartheta \big|W_\varepsilon(t_1) - W_\varepsilon(t_2)\big|^2 \le C|t_2 - t_1|, \qquad t_1, t_2 \in [0, T - \nu].$$

These two conditions provide weak convergence of the integrals

$$\int_0^{T-\nu} W_\varepsilon(t)^2 \, \mathrm{d}t \Longrightarrow \int_0^{T-\nu} w(t)^2 \, \mathrm{d}t$$

for any $\nu > 0$. It can be shown that for any $\eta > 0$ there exist $\nu > 0$ such that

$$\int_{T-\nu}^{T} \mathbf{E}_\vartheta \, W_\varepsilon(t)^2 \, \mathrm{d}t \le \eta.$$

The proof is close to that given in Maglaperidze *et al.* [15] for a similar integral. □

4. Examples

Example 1. We consider the simplest case which allows us to have an ADF GoF test for each ε, that is, no need to study statistics as $\varepsilon \to 0$. Observe that a similar situation is discussed in Khmaladze [6] but for a different problem.

Suppose that the observed diffusion process (under hypothesis) is

$$\mathrm{d}X_t = \vartheta \, \mathrm{d}t + \varepsilon \, \mathrm{d}W_t, \qquad X_0 = 0, 0 \le t \le 1. \tag{4.1}$$

Then

$$h(\vartheta, t) = 1, \qquad \mathrm{I}(\vartheta) = 1, \qquad N(\vartheta, t) = 1 - t,$$
$$\hat\vartheta_\varepsilon = X_1, \qquad \varepsilon^{-1}(\hat\vartheta_\varepsilon - \vartheta) = W_1 \sim \mathcal{N}(0, 1).$$

Further

$$x_t(\vartheta) = \vartheta t, \qquad x_t^{(1)}(\vartheta) = W_t, \qquad U(\vartheta, t) = W_t - W_1 t,$$
$$V_\varepsilon(t) = U_\varepsilon(t) = \varepsilon^{-1}(X_t - X_1 t) = W_t - W_1 t = B(t).$$

Therefore,

$$W_\varepsilon(t) = \varepsilon^{-1}(X_t - X_1 t) + \varepsilon^{-1} \int_0^t (1 - s)^{-1}[X_s - X_1 s] \, \mathrm{d}s$$

and under the basic hypothesis we have

$$W_\varepsilon(t) = B(t) + \int_0^t \frac{B(s)}{1-s}\, \mathrm{d}s = w(t).$$

Therefore,

$$\Delta_\varepsilon = \int_0^1 W_\varepsilon(t)^2\, \mathrm{d}t = \int_0^1 w(t)^2\, \mathrm{d}t$$

and the test $\hat{\Psi}_\varepsilon = \mathbb{1}_{\{\Delta_\varepsilon > c_\alpha\}} \in \mathcal{K}_\alpha$ satisfies the equality

$$\mathbf{E}_\vartheta \hat{\Psi}_\varepsilon = \mathbf{P}\left\{\int_0^1 w(t)^2\, \mathrm{d}t > c_\alpha\right\} = \alpha.$$

Example 2. Consider the linear case

$$\mathrm{d}X_t = \big\langle\vartheta, \mathbf{H}(t, X_t)\big\rangle\, \mathrm{d}t + \varepsilon\sigma(t, X_t)\, \mathrm{d}W_t, \qquad X_0 = x_0, 0 \le t \le T,$$

where $\vartheta \in \Theta \subset \mathcal{R}^d$ and assume that the functions $\mathbf{H}(t, x)$ and $\sigma(t, x)$ satisfy regularity conditions. In this case, we can take $\bar{\mathbf{h}}_\varepsilon(\vartheta, t) = \bar{\mathbf{h}}_\varepsilon(t)$, that is, this vector-valued function does not depend on ϑ. Hence, the stochastic integral is well defined and the test has a simplified form. We have

$$\bar{\mathbf{h}}_\varepsilon(t) = \frac{\mathbf{H}(t, X_t)}{\sigma(t, X_t)}, \qquad \bar{\mathbb{N}}_\varepsilon(\vartheta, s) = \int_s^T \frac{\mathbf{H}(t, x_t(\vartheta))\mathbf{H}(t, x_t(\vartheta))^*}{\sigma(t, x_t(\vartheta))^2}\, \mathrm{d}s,$$

$$\mathrm{d}U_\varepsilon(t) = \frac{\mathrm{d}X_t}{\varepsilon\sigma\varphi(t, X_t)} - \frac{[\langle\hat{\vartheta}_\varepsilon, \mathbf{H}(t, x_t(\hat{\vartheta}_\varepsilon))\rangle + \langle\hat{\vartheta}_\varepsilon, \mathbf{H}'_x(t, X_t)\rangle(X_t - x_t(\hat{\vartheta}_\varepsilon))]\, \mathrm{d}t}{\varepsilon\sigma(t, X_t)},$$

$$W_\varepsilon(t) = U_\varepsilon(t) + \int_0^t \mathbf{H}(s, X_s)^* \bar{\mathbb{N}}_\varepsilon(\hat{\vartheta}_\varepsilon, s)^{-1} \int_0^s \mathbf{H}(v, X_v)\, \mathrm{d}U_\varepsilon(v)\, \mathrm{d}s$$

and so on.

Acknowledgements

This study was partially supported by Russian Science Foundation (research project No. 14-49-00079). The author thanks the referee for helpful comments.

References

[1] Dachian, S. and Kutoyants, Y.A. (2008). On the goodness-of-fit tests for some continuous time processes. In *Statistical Models and Methods for Biomedical and Technical Systems* (F. Vonta, M. Nikulin, N. Limnios and C. Huber-Carol, eds.). *Stat. Ind. Technol.* 385–403. Boston, MA: Birkhäuser. MR2462767

[2] Darling, D.A. (1955). The Cramér–Smirnov test in the parametric case. *Ann. Math. Statist.* **26** 1–20. MR0067439

[3] Freidlin, M.I. and Wentzell, A.D. (1998). *Random Perturbations of Dynamical Systems*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **260**. New York: Springer. Translated from the 1979 Russian original by Joseph Szücs. MR1652127

[4] Hitsuda, M. (1968). Representation of Gaussian processes equivalent to Wiener process. *Osaka J. Math.* **5** 299–312. MR0243614

[5] Iacus, S.M. and Kutoyants, Yu.A. (2001). Semiparametric hypotheses testing for dynamical systems with small noise. *Math. Methods Statist.* **10** 105–120. MR1841810

[6] Khmaladze, È.V. (1981). A martingale approach in the theory of goodness-of-fit tests. *Theory Probab. Appl.* **26** 240–257.

[7] Kleptsyna, M. and Kutoyants, Y.A. (2014). On asymptotically distribution free tests with parametric hypothesis for ergodic diffusion processes. *Stat. Inference Stoch. Process.* To appear. Available at arXiv:1305.3382.

[8] Kutoyants, Yu. (1994). *Identification of Dynamical Systems with Small Noise*. *Mathematics and Its Applications* **300**. Dordrecht: Kluwer Academic. MR1332492

[9] Kutoyants, Y.A. (2011). Goodness-of-fit tests for perturbed dynamical systems. *J. Statist. Plann. Inference* **141** 1655–1666. MR2763197

[10] Kutoyants, Yu.A. (2014). On ADF goodness-of-fit tests for stochastic processes. In *New Perspectives on Stochastic Modeling and Data Analysis* (J. Bozeman, V. Girardin and C. Skiadas, eds.). To appear.

[11] Kutoyants, Yu.A. (2014). On score-function processes and goodness of fit tests for stochastic processes. Available at arXiv:1403.7715.

[12] Kutoyants, Y.A. (2014). On asymptotic distribution of parameter free tests for ergodic diffusion processes. *Stat. Inference Stoch. Process.* **17** 139–161. MR3219526

[13] Kutoyants, Y.A. and Zhou, L. (2014). On approximation of the backward stochastic differential equation. *J. Statist. Plann. Inference* **150** 111–123. MR3206723

[14] Liptser, R. and Shiryaev, A. (2001). *Statistics of Random Processes*. *Vols. I, II*, 2nd ed. Berlin: Springer.

[15] Maglaperidze, N.O., Tsigroshvili, Z.P. and van Pul, M. (1998). Goodness-of-fit tests for parametric hypotheses on the distribution of point processes. *Math. Methods Statist.* **7** 60–77. MR1626572

[16] Shepp, L.A. (1966). Radon–Nikodým derivatives of Gaussian measures. *Ann. Math. Statist.* **37** 321–354. MR0190999

[17] Yoshida, N. (1993). Asymptotic expansion of Bayes estimators for small diffusions. *Probab. Theory Related Fields* **95** 429–450. MR1217445

[18] Yoshida, N. (1996). Asymptotic expansions for perturbed systems on Wiener space: Maximum likelihood estimators. *J. Multivariate Anal.* **57** 1–36. MR1392575

Received January 2014 and revised June 2014