Ann Inst Stat Math (2017) 69:833–864
DOI 10.1007/s10463-016-0564-y

A change detection procedure for an ergodic diffusion process

Koji Tsukuda[1,2]

Received: 21 April 2015 / Revised: 16 December 2015 / Published online: 18 May 2016

Abstract A test procedure based on continuous observation to detect a change in drift parameters of an ergodic diffusion process is proposed. The asymptotic behavior of a random field relating to an estimating equation under the null hypothesis is established using weak convergence theory in separable Hilbert spaces. This result is applied to a change point detection test.

Keywords Change point problems · Diffusion processes · Weak convergences in $L^2(0, 1)$

1 Introduction and notation

1.1 Introduction

Diffusion processes play important roles in several fields, including economics, financial mathematics and population genetics. Many statistical problems corresponding to classical i.i.d. settings can also be considered for diffusion processes (see, e.g.,

The large part of this paper is based on the thesis of the author at SOKENDAI (The Graduate University for Advanced Studies). The revision was done when the author was a member of Kurume University, Fukuoka. The author was a Research Fellow of Japan Society for the Promotion of Science and this work was partly supported by JSPS KAKENHI Grant Number 26-1487 (Grant-in-Aid for JSPS Fellows).

✉ Koji Tsukuda
k.tsukuda@fuji.waseda.jp

1 Faculty of International Research and Education, Waseda University, 1-6-1 Nishi-waseda, Shinjuku-ku, Tokyo 169-8050, Japan

2 Present Address: Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan

the book Kutoyants (2004)). In particular, tests to detect changes in drift parameters of diffusion processes are considered in the studies by Lee et al. (2006), Negri and Nishiyama (2012), and Dehling et al. (2014). De Gregorio and Iacus (2008) and Song and Lee (2009) consider testing for changes in diffusion coefficients using discrete observations and Negri and Nishiyama (2014) considers a way to detect changes in drift and diffusion coefficients at the same time. See also Mihalache (2012) for a sequential change detection for ergodic diffusion processes. These are all change point problems, a topic on which there has been much research: see studies by Csörgő and Horváth (1997), Brodsky and Darkhovsky (2000), and Chen and Gupta (2012) for general surveys.

Let us now roughly explain the problem setting and our approach, leaving the precise description to Sect. 4. Consider an ergodic diffusion process

$$X_t = X_0 + \int_0^t S(X_s, \theta)\mathrm{d}s + \int_0^t \sigma(X_s)\mathrm{d}W_s \tag{1}$$

for $t \in [0, \infty)$ with the state space $I = (l, r)$ for $-\infty \le l < r \le \infty$, where $W.$ is a standard Brownian motion and X_0 is a random variable that is independent of $W.$ and satisfies $\mathbb{E}[(X_0)^2] < \infty$. The problem is to test the following pair of hypotheses:

$$\begin{aligned}
&\mathcal{H}_0 : \exists\theta_0 \in \Theta \; such \; that \; \theta_{(t)} = \theta_0 \; \forall t \in [0, T] \\
&\mathcal{H}_1 : \exists\theta_0, \theta_1 \in \Theta \; and \; \exists u_* \in (0, 1) \; such \; that \theta_{(t)} = \theta_0 \; \forall t \in [0, Tu_*) \; and \\
&\quad \theta_{(t)} = \theta_1 \neq \theta_0 \; \forall t \in [Tu_*, T]
\end{aligned}$$

Based on the continuous time observations $\{X_t; t \in [0, T]\}$ with the asymptotic setting $T \to \infty$, we propose a consistent procedure to test these hypotheses. Similar problem settings have been considered by some previous works, such as Lee et al. (2006) and Negri and Nishiyama (2012). For estimating the drift parameter in (1), the likelihood equation

$$\frac{1}{T}\int_0^T \frac{\dot{S}(X_s, \theta)}{\sigma(X_s)^2}(\mathrm{d}X_s - S(X_s, \theta)\mathrm{d}s) = 0$$

is considered. Define the random field

$$(u, \theta) \rightsquigarrow \mathbb{Z}_T(u, \theta) = \frac{1}{\sqrt{T}}\int_0^T w_s^T(u)\frac{\dot{S}(X_s, \theta)}{\sigma(X_s)^2}(\mathrm{d}X_s - S(X_s, \theta)\mathrm{d}s), \tag{2}$$

where

$$w_s^T : (0, 1) \ni u \mapsto w_s^T(u) = \frac{1\{s \le Tu\} - u}{\sqrt{u(1-u)}}$$

and $s \in [0, T]$. We shall see that, under $\mathcal{H}_0$, the random field $u \rightsquigarrow \mathbb{Z}_T(u, \theta_0)$ converges weakly to a Gaussian field in $L^2(0, 1)$ as T tends to infinity. The denominator of $w_s^T(\cdot)$ converges to 0 as $u \to 0$ or $u \to 1$; so, under $\mathcal{H}_1$, $\mathbb{Z}_T$ becomes large when u_* is close to 0 or 1 and then the test is expected to have a high power if we use the function of $\mathbb{Z}_T$ as a test statistic. This is the main motivation of the work.

The idea of using the partial sum of the estimating equation basically comes from the study by Horváth and Parzen (1994). This work examined the asymptotic behavior of a Fisher score change process, which is a stochastic process relating to the likelihood equation, for general independent observations under the null hypothesis. Negri and Nishiyama (2012) refines the idea and applies it to the detection of changes of drift parameters in an ergodic diffusion process. The proof of the limit theorem of Negri and Nishiyama (2012), especially the proof of asymptotic tightness, is based on the tightness criterion for martingales taking values in ℓ^∞ spaces, which is the set of all bounded real functions endowed with the uniform metric. However, we cannot apply this kind of weak convergence theorem to the current problem because the random field $\mathbb{Z}_T(\cdot, \theta)$ is not bounded owing to the denominator of $w_s^T(\cdot)$. Hence, we regard the random field (2) as an element of $L^2(0, 1)$ and prove the limit theorems in $L^2(0, 1)$. Generally speaking, weak convergences in L^2 are weaker than other often-used results in the Skorokhod topology or the uniform topology. But, for some tests, weak convergences in L^2 are enough: for goodness-of-fit tests see studies by Khmaladze (1979), Mason (1984), and LaRiccia and Mason (1986) and for change point detection tests see studies by Suquet and Viano (1998) and Tsukuda and Nishiyama (2014).

Note that Mihalache (2012) and Dehling et al. (2014) also consider weighted test statistics for the detection of changes of a drift parameter in a diffusion processes. In particular, Mihalache (2012) considers a sequential change detection problem and proposes a weighted CUSUM test statistic. Their result is strong convergence and the limit is $t \rightsquigarrow B(t)/t^\gamma$, $\gamma \in [0, 1/4)$ in the sense of the supremum metric, whereas ours is $t \rightsquigarrow B^\circ(t)/\{t(1-t)\}^\gamma$, $\gamma = 1/2$ in the sense of the $L^2(0, 1)$ metric, where B is a standard Brownian motion and B° is a standard Brownian bridge with dimensions depending on the dimension of the parameter of interest. We believe that using the weight function corresponding to $\gamma = 1/2$ is important even though our result is one of the weak convergences in $L^2(0, 1)$. Dehling et al. (2014) consider another model and propose test statistics using the log likelihood ratio with two results: one is weak convergence with a fixed interval that does not contain 0 and 1, and the other is a Darling-Erdős type result which has the same limit as a result in the study by Horváth (1993). In contrast, our result is convergence in $L^2(0, 1)$ and the interval contains 0 and 1.

To close this subsection, let us describe the organization of this paper. Section 2 introduces the preliminary results that will be used in the following sections. Section 3 includes the limit theorem of a stochastic integral taking values in $L^2(0, 1)$. This result is applied to a change point detection test in Sect. 4. The proofs of the results are in Sect. 5.

1.2 Notation

Let us explain some notations. We shall consider asymptotic behaviors as T tends to infinity and the notations $\to^p$ and $\to^d$ denote convergence in probability and convergence in distribution, respectively. The notation l.i.m. means the limit in mean square, where "mean" indicates the expectation. The notation $1\{\cdot\}$ denotes the indicator function. The binary relation $a \wedge b$ for $a, b \in \mathbb{R}$ means $\min(a, b)$. Let us denote the transpose of a vector or matrix by the superscript $\top$. The finite-dimensional Euclidean norm of a vector x is denoted by $\|x\| = (x^\top x)^{1/2}$. The (i, j) element of matrix A is denoted by $(A)_{(i,j)}$ and the operator norm of matrix A is denoted by $\|A\|_{OP}$, that is,

$$\|A\|_{OP} = \sup_{x \in \mathbb{R}^d, \|x\|=1} \|Ax\| = \sup_{x \in \mathbb{R}^d, \|x\|>0} \frac{\|Ax\|}{\|x\|}.$$

Moreover, the Frobenius norm of matrix A is denoted by $\|A\|$, that is,

$$\|A\| = (\mathrm{tr}(A^\top A))^{1/2} = \left(\sum_i \sum_j \left| (A)_{(i,j)} \right|^2 \right)^{1/2}.$$

Note that

$$\|A\|_{OP} = \max_\sigma \sigma(A) \leq \left(\sum_\sigma (\sigma(A))^2 \right)^{1/2} = \|A\|,$$

where $\sigma(A)$ denotes the singular value of the matrix A. The expectation of a random variable X is denoted by $\mathbb{E}[X]$. In particular, for a random vector or a random matrix X, $\mathbb{E}[X]$ denotes the vector or the matrix in which each element is the expectation of the corresponding element of X.

We introduce a functional space $L^2(S, \mathbb{R}^d, \mathrm{d}s)$, or in abbreviated form $L^2(S)$, where S is a bounded subset of the Euclidean space. Consider the inner product

$$\langle z_1, z_2 \rangle_{L^2(S)} = \int_S z_1(s)^\top z_2(s) \mathrm{d}s,$$

where z_1 and z_2 are d-dimensional vector-valued functions on S and $\mathrm{d}s$ is the Lebesgue measure. The functional space $L^2(S)$ is equivalence classes of square integrable real vector functions on a bounded set S, that is, the set of all measurable functions $z : S \to \mathbb{R}^d$ that satisfy $\|z\|^2_{L^2(S)} = \langle z, z \rangle_{L^2(S)} < \infty$. This space is a separable Hilbert space with respect to L^2 distance $\|z_1 - z_2\|_{L^2(S)}$.

The predictable quadratic variation process of a martingale $t \rightsquigarrow M_t$ is denoted by $t \rightsquigarrow \langle M \rangle_t$.

The derivatives of f with respect to θ_i and x, which will appear in Sects. 4 and 5, are denoted by $\partial_i f$ and f', respectively. Moreover, the gradient vector with respect to θ is denoted by $\dot{f}$.

2 Preliminary results

2.1 On tightness criteria in $L^2(0, 1)$

Let $\mathbb{H}$ be a real separable Hilbert space with inner product $\langle\cdot,\cdot\rangle_{\mathbb{H}}$ and a complete orthonormal system $\{e_i\}_{i=1}^{\infty}$. An $\mathbb{H}$-valued random sequence $\{X_n\}_{n=1}^{\infty}$ is said to be *asymptotically finite dimensional* if for any $\delta, \varepsilon > 0$, there exists a finite subset $\{e_i\}_{i\in I}$ of the complete orthonormal system such that

$$\limsup_{n\to\infty} \mathbb{P}\left(\sum_{j\notin I}\langle X_n, e_j\rangle_{\mathbb{H}}^2 > \delta\right) < \varepsilon.$$

This tightness criterion was established by Prokhorov (1956). The phrase "asymptotically finite dimensional" seems to have been first used by van der Vaart and Wellner (1996) and the following theorem is contained in Section 1.8 of this book.

Theorem 1 (van der Vaart and Wellner (1996), Theorem 1.8.4) *A sequence of random variables $X_n : \Omega_n \to \mathbb{H}$ converges in distribution to a tight random variable X if and only if it is asymptotically finite dimensional and the sequence $\langle X_n, h\rangle_{\mathbb{H}}$ converges in distribution to $\langle X, h\rangle_{\mathbb{H}}$ for every $h \in \mathbb{H}$.*

It should be noted that the measurability of $\{X_\cdot\}$ is not assumed in van der Vaart and Wellner (1996), whereas it is assumed in this paper. A sufficient condition to verify that a given sequence of random elements taking values in $\mathbb{H}$ is asymptotically finite dimensional is given in the following proposition which is due to Prof. Nishiyama.

Proposition 1 *A sequence of random variables $X_n : \Omega \to \mathbb{H}$ is asymptotically finite dimensional if there exists the random variable X such that*

$$\mathbb{E}\big[\|X_n\|_{\mathbb{H}}^2\big] \to \mathbb{E}\big[\|X\|_{\mathbb{H}}^2\big] < \infty \tag{3}$$

and

$$\mathbb{E}\big[\langle X_n, e_j\rangle_{\mathbb{H}}^2\big] \to \mathbb{E}\big[\langle X, e_j\rangle_{\mathbb{H}}^2\big], \quad \forall j \in J, \tag{4}$$

as $n \to \infty$, where $\{e_j : j \in J\}$ is a complete orthonormal system of $\mathbb{H}$.

2.2 On limit theorems for stochastic processes

In this subsection, we introduce two theorems that can be used to prove the consistency and asymptotic normality of Z-estimators, including the maximum likelihood estimator, together with the general theory of Z-estimation. See Remark 2 in Sect. 4 for general results on the Z-estimator (see, for example, van der Vaart 1998).

The following theorem is a uniform law of large numbers for ergodic stochastic processes. For one-dimensional ergodic diffusion processes, a corresponding result with a more general envelope condition for a set of functions instead of (5) can be found in van Zanten (2003).

Theorem 2 (Nishiyama (2011), Theorem 8.4.1(i)) *Let $(\mathcal{X}, \mathcal{A})$ be a measurable space, Θ be a bounded subset of $\mathbb{R}^p$. Consider a set of measurable function $\{f(\cdot, \theta); \theta \in \Theta\}$ on $\mathcal{X}$. Suppose that*

$$|f(x, \theta_1) - f(x, \theta_2)| \leq K(x)\|\theta_1 - \theta_2\|^\gamma \tag{5}$$

for $\forall\theta_1, \theta_2 \in \Theta$, a measurable function K and positive constant γ. Consider an ergodic stochastic process $\{X_t\}_{t\in[0,\infty)}$ which takes its value in $\mathcal{X}$ and let μ be the invariant measure. If all $f(\cdot;\theta)$ and K are integrable with respect to μ, then it holds that

$$\sup_{\theta\in\Theta}\left|\frac{1}{T}\int_0^T f(X_t, \theta)\mathrm{d}t - \int_{\mathcal{X}} \mathrm{f}(\mathrm{x}; \theta)\mu(\mathrm{dx})\right| \to^p 0$$

as $T \to \infty$.

The next theorem is a central limit theorem in ℓ^∞ for martingales, where $\ell^\infty(\Theta)$ is the set of all bounded real-valued functions on Θ. This result is based on Theorems 3.1.1 and 3.4.2 of Nishiyama (2000).

Theorem 3 (Nishiyama (2011), Theorem 8.6.4(i)) *Let (Θ, ρ) be a metric space satisfying the metric entropy condition, $X^{n,\theta}_{\cdot}$ be a continuous time martingale, and T_n be a finite stopping time. Suppose that there exists a sequence of positive random variables $\{K_n\}_{n=1}^\infty$ satisfying*

$$\sqrt{\langle X^{n,\theta_1} - X^{n,\theta_2}\rangle_{T_n}} \leq K_n\rho(\theta_1, \theta_2)$$

for $\forall\theta_1, \theta_2 \in \Theta$ and $K_n = O_p(1)$. If for every $\theta_1, \theta_2 \in \Theta$, $\langle X^{n,\theta_1}, X^{n,\theta_2}\rangle_{T_n}$ converges to a constant $C(\theta_1, \theta_2)$ in probability, then the random field $\theta \rightsquigarrow X^{n,\theta}_{T_n}$ converges weakly in $\ell^\infty(\Theta)$ to a Gaussian field $\theta \rightsquigarrow G(\theta)$ such that $\mathbb{E}[G(\theta)] = 0$ and the covariance is $\mathbb{E}[G(\theta_1, \theta_2)] = C(\theta_1, \theta_2)$. The limit $\theta \rightsquigarrow G(\theta)$ is almost surely continuous with respect to ρ and the semimetric ρ_G defined by $\rho_G(\theta_1, \theta_2) = (\mathbb{E}[|G(\theta_1) - G(\theta_2)|^2])^{1/2}$.

For other approaches to deriving asymptotic properties of the maximum likelihood estimators of drift parameters of diffusion processes, see studies by Lánska (1979) and by Kutoyants (2004).

3 A weak convergence theorem in $L^2(0, 1)$ for a stochastic integral

Choose a measurable space and introduce a filtration. Let us consider a locally square integrable martingale $M_{\cdot}$ whose predictable quadratic variation process is

$$\langle M\rangle_{\cdot} = \int_0^{\cdot} \lambda_s \mathrm{d}s,$$

where $\lambda_\cdot$ is a non-negative adapted process which satisfies

$$\sup_{s\in[0,\infty)} \mathbb{E}[\lambda_s] < \infty.$$

It follows that $M_\cdot$ is a martingale. Define the random field

$$(u,\theta) \rightsquigarrow \mathcal{M}_T(u,\theta) = \frac{1}{\sqrt{T}} \int_0^T w_s^T(u) H_s(\theta) \mathrm{d}M_s,$$

where

$$w_s^T(u) = \frac{1\{s \le Tu\} - u}{\sqrt{u(1-u)}}, \quad \forall u \in (0,1),$$

θ is an element of an open bounded subset Θ of $\mathbb{R}^d$ and $H_\cdot(\theta)$ is a d-dimensional predictable process such that

$$\int_0^T \|H_s(\theta)\|^2 \lambda_s \mathrm{d}s < \infty, \quad a.s. \quad \forall\theta \in \Theta.$$

Note that $\mathcal{M}_T(u,\theta)$ is the terminal value of the martingale

$$\frac{1}{\sqrt{T}} \int_0^\cdot w_s^T(u) H_s(\theta) \mathrm{d}M_s.$$

The following proposition gives a relation between moments.

Proposition 2 *Fix a* $\theta \in \Theta$. (i) *If*

$$\sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^4 \lambda_s^2\right] < \infty \tag{6}$$

holds, then

$$\sup_{s\in[0,\infty)} \left\|\mathbb{E}\left[H_s(\theta) H_s(\theta)^\top \lambda_s\right]\right\|_{OP} < \infty. \tag{7}$$

(ii) *If* (6) *and*

$$\sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^2\right] < \infty \tag{8}$$

hold, then

$$\sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^3 \lambda_s\right] < \infty. \tag{9}$$

(iii) If (6) *holds, then*

$$\sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^2 \lambda_s\right] < \infty. \tag{10}$$

Moreover, (10) *implies that*

$$\mathbb{E}\left[\|\mathcal{M}_T(\cdot,\theta)\|^2_{L^2(0,1)}\right] < \infty, \tag{11}$$

and in particular, $\mathcal{M}_T(\cdot,\theta)$ *almost surely takes its values in* $L^2(0,1)$*.*

The following theorem describes the asymptotic behavior of $u \rightsquigarrow \mathcal{M}_T(u,\theta)$ in $L^2(0,1)$.

Theorem 4 *Fix a* $\theta \in \Theta$*. Suppose that there exists the limit*

$$C(\theta,\eta) = \underset{T\to\infty}{\text{l.i.m.}} \left(\frac{1}{T}\int_0^T H_s(\theta)H_s(\eta)^\top \lambda_s \mathrm{d}s\right) \tag{12}$$

for $\theta,\eta \in \Theta$*. If* (6) *and* (8) *hold, then the random field* $\mathcal{M}_T(\cdot,\theta)$ *converges weakly to*

$$\Gamma(\cdot,\theta) = \frac{C(\theta,\theta)^{1/2} B_d^\circ(\cdot)}{w(\cdot)}$$

in $L^2(0,1)$ *as* $T \to \infty$*, where* $B_d^\circ(\cdot)$ *denotes a* d*-dimensional standard Brownian bridge and* $w(u) = (u(1-u))^{1/2}$ *for* $u \in (0,1)$*.*

The following proposition will be used in the proof of Theorem 4.

Proposition 3 *For any* $u, v \in (0,1)$*,* $\theta \in \Theta$ *and* $h \in L^2(0,1)$*,* (12) *implies*

$$\frac{1}{T}\int_0^T w_s^T(u) w_s^T(v) J_s \mathrm{d}s \to \frac{\mathrm{u}\wedge\mathrm{v}-\mathrm{uv}}{\sqrt{\mathrm{uv}(1-\mathrm{u})(1-\mathrm{v})}}\mathrm{h}(\mathrm{u})^\top \mathrm{C}(\theta,\theta)\mathrm{h}(\mathrm{v}),$$

where $J_s = h(u)^\top \mathbb{E}[H_s(\theta)H_s(\theta)^\top \lambda_s] h(v)$*.*

4 A change detection procedure for an ergodic diffusion process

Let us consider the stochastic differential equation given by (1). The parameter θ is an element of Θ, an open bounded subset of $\mathbb{R}^d$. Suppose that there exists a strong solution to this SDE and that

$$\sup_{s\in[0,\infty)} \mathbb{E}[\sigma(X_s)^2] < \infty.$$

Further, suppose that $X_\cdot$ is ergodic in mean square with respect to an invariant measure μ_θ for some θ, that is, for any μ_θ-integrable function f, it holds that

$$\lim_{T\to\infty} \mathbb{E}\left[\left\|\frac{1}{T}\int_0^T f(X_s)\mathrm{d}s - \int_I f(x)\mu_\theta(\mathrm{d}x)\right\|^2\right] = 0.$$

Remark 1 The previous work, Negri and Nishiyama (2012) assumes ergodicity which guarantees the convergence in probability, so this assumption is stronger than theirs.

Let us denote the true value of θ for X_t by $\theta_{(t)}$. For the model above, we wish to test the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ in Sect. 1.

To estimate the parameter θ, let us consider the estimating equation under $\mathcal{H}_0$

$$\Psi_T(\theta) = \frac{1}{T}\int_0^T \frac{\dot{S}(X_s,\theta)}{\sigma(X_s)^2}(\mathrm{d}X_s - S(X_s,\theta)\mathrm{d}s) = 0. \tag{13}$$

Suppose that there exists a unique solution $\hat{\theta}_T$ of this estimating equation. Let us introduce the following conditions.

(I) The function $(x,\theta) \mapsto S(x,\theta)$ is continuously differentiable with respect to x and third-order continuously differentiable with respect to θ and the order of the derivatives is exchangeable. The function $x \mapsto \sigma(x)$ is continuously differentiable with respect to x. The functions $\sup_{\theta\in\Theta}|S(x,\theta)|$, $\sup_{\theta\in\Theta}|\partial_i S(x,\theta)|$, $\sup_{\theta\in\Theta}|\partial_{ij} S(x,\theta)|$, $\sup_{\theta\in\Theta}|\partial_{ijk} S(x;\theta)|$, $\sigma(x)$ and $\sigma'(x)$ are bounded above by polynomial growth functions of x: that is, for example, it holds that

$$\sup_{\theta\in\Theta}|S(x,\theta)| \le C(1+|x|)^p, \quad \forall x \in \mathbb{R}$$

for some constants C, $p \ge 1$.

(II) $\inf_{x\in\mathbb{R}} \sigma(x) > 0$.

(III) For arbitrary $q \ge 1$, $\sup_{s\in[0,\infty)} \mathbb{E}\left[|X_s|^q\right] < \infty$.

(IV) For all $\theta, \kappa \in \Theta$,

$$\Psi(\theta,\kappa) = \int_I \frac{(S(x,\kappa) - S(x,\theta))\dot{S}(x,\theta)}{\sigma(x)^2}\mu_\kappa(\mathrm{d}x) < \infty. \tag{14}$$

For all $\kappa \in \Theta$ and any $\varepsilon > 0$, $\inf_{\theta:\|\theta-\kappa\|>\varepsilon} \|\Psi(\theta,\kappa)\| > 0$ holds.

(V) For all $\theta, \eta, \kappa \in \Theta$

$$C_\kappa(\theta,\eta) = \int_I \frac{\dot{S}(x,\theta)\dot{S}(x,\eta)^\top}{\sigma(x)^2}\mu_\kappa(\mathrm{d}x) < \infty.$$

The matrix $C_\kappa(\theta,\theta)$ is positive definite for all $\theta, \kappa \in \Theta$.

(VI) There exist positive functions $x \mapsto K(x)$, $K_d(x)$ such as

$$\max_{i,j}|\partial_i\partial_j S(\cdot,\theta_1) - \partial_i\partial_j S(\cdot,\theta_2)| \le K(\cdot)\|\theta_1-\theta_2\|,$$
$$\max_{i,j}|\partial_i\partial_j S'(\cdot,\theta_1) - \partial_i\partial_j S'(\cdot,\theta_2)| \le K_d(\cdot)\|\theta_1-\theta_2\|$$

for $\forall\theta_1, \theta_2 \in N$, where N is a neighborhood of any θ_0. The function $K(x)$ is continuously differentiable with respect to x. The functions $K(x)$ and $K_d(x)$ are bounded above by polynomial growth functions of x.

Remark 2 From (13) and (14), under $\mathcal{H}_0$, it holds that

$$\begin{aligned}
&\|\Psi_T(\theta) - \Psi(\theta, \theta_0)\| \\
&\le \left\| \frac{1}{T} \int_0^T \frac{S(X_s, \theta_0)\dot{S}(X_s, \theta)}{\sigma(X_s)^2} \mathrm{d}s - \int_I \frac{S(x, \theta_0)\dot{S}(x, \theta)}{\sigma(x)^2} \mu_{\theta_0}(\mathrm{d}x) \right\| \\
&\quad + \left\| \frac{1}{T} \int_0^T \frac{\dot{S}(X_s, \theta)}{\sigma(X_s)} \mathrm{d}W_s \right\| \\
&\quad + \left\| \frac{1}{T} \int_0^T \frac{S(X_s, \theta)\dot{S}(X_s, \theta)}{\sigma(X_s)^2} \mathrm{d}s - \int_I \frac{S(x, \theta)\dot{S}(x, \theta)}{\sigma(x)^2} \mu_{\theta_0}(\mathrm{d}x) \right\|.
\end{aligned}$$

Under $\mathcal{H}_0$, the supremum of $\|\Psi_T(\theta) - \Psi(\theta, \theta_0)\|$ with respect to θ converges to 0 in probability by Theorems 2 and 3. This leads to the consistency of $\hat{\theta}_T$: see Theorem 5.9 of van der Vaart (1998). Asymptotic normality, which is Lemma 1 (i) of Negri and Nishiyama (2012) under stronger conditions, follows from the consistency, the Taylor expansion and Theorems 2 and 3: see Theorem 5.21 of van der Vaart (1998) (although Theorems 5.9 and 5.21 of van der Vaart (1998) deal with discrete observations, corresponding results are valid for continuous observations). Moreover, part (ii) of the following lemma, which is Lemma 1 (ii) of Negri and Nishiyama (2012), also holds from a similar argument involving only consistency.

Lemma 1 *Assume conditions* (I–V). (i) *Under* $\mathcal{H}_0$, *it holds that* $\sqrt{T}(\hat{\theta}_T - \theta_0) \to^d N(0, C_{\theta_0}(\theta_0, \theta_0)^{-1})$.
(ii) *Under* $\mathcal{H}_1$, *it holds that* $\hat{\theta}_T \to^p \theta_*$, *where* θ_* *is a value that satisfies* $u_*\Psi(\theta_*, \theta_0) + (1 - u_*)\Psi(\theta_*, \theta_1) = 0$.

Proposition 4 *Assume conditions* (I–III). (i) *Under* $\mathcal{H}_0$, *it holds that*

$$\sup_{s \in [0,\infty)} \mathbb{E}\left[\frac{\|\dot{S}(X_s, \theta_0)\|^4}{\sigma(X_s)^4}\right] < \infty \tag{15}$$

and

$$\sup_{s \in [0,\infty)} \mathbb{E}\left[\frac{\|\dot{S}(X_s, \theta_0)\|^2}{\sigma(X_s)^4}\right] < \infty. \tag{16}$$

(ii) *Under* $\mathcal{H}_1$, (15) *and* (16) *hold if we replace* θ_0 *with* θ_*. *Moreover, it holds that*

$$\sup_{s \in [0,\infty)} \mathbb{E}\left[\frac{\|\dot{S}(X_s, \theta_*)\|^2}{\sigma(X_s)^4}(S(X_s, \theta))^2\right] < \infty$$

for $\theta \in \{\theta_0, \theta_1, \theta_*\}$.

Introduce the random field $\{\mathbb{Z}_T(u, \theta); (u, \theta) \in (0, 1) \times \Theta\}$ given by

$$\mathbb{Z}_T(u, \theta) = \frac{1}{\sqrt{T}} \int_0^T w_s^T(u) \frac{\dot{S}(X_s, \theta)}{\sigma(X_s)^2}(\mathrm{d}X_s - S(X_s, \theta)\mathrm{d}s),$$

where

$$w_s^T(u) = \frac{1\{s \le Tu\} - u}{\sqrt{u(1-u)}}, \quad u \in (0,1).$$

Its "predictable projection" to the true model is

$$\mathbb{Z}_T^p(u,\theta) = \frac{1}{\sqrt{T}} \int_0^T w_s^T(u) \frac{\dot{S}(X_s,\theta)}{\sigma(X_s)^2} (S(X_s,\theta_{(s)}) - S(X_s,\theta))\mathrm{d}s.$$

The difference between $\mathbb{Z}$ and $\mathbb{Z}^p$, which is a martingale random field, is denoted by $\{\mathbb{M}_T(u,\theta); (u,\theta) \in (0,1) \times \Theta\}$:

$$\mathbb{M}_T(u,\theta) = \mathbb{Z}_T(u,\theta) - \mathbb{Z}_T^p(u,\theta) = \frac{1}{\sqrt{T}} \int_0^T w_s^T(u) \frac{\dot{S}(X_s,\theta)}{\sigma(X_s)} \mathrm{d}W_s$$

for $u \in (0,1)$ and $\theta \in \Theta$. Its weak convergence follows from the limit theorem in the preceding section.

Under $\mathcal{H}_0$, it holds that

$$\mathbb{Z}_T^p(\cdot,\theta_0) = 0,$$

so

$$\mathbb{M}_T(\cdot,\theta_0) = \mathbb{Z}_T(\cdot,\theta_0).$$

This relationship motivates the use of functions of $\mathbb{Z}_T$ as test statistics. Since we cannot know θ_0, it is crucial that, under $\mathcal{H}_0$,

$$\left\| \mathbb{Z}_T(\cdot,\hat{\theta}_T) - \mathbb{Z}_T(\cdot,\theta_0) \right\|_{L^2(0,1)} \to^p 0.$$

This will be established by the following two lemmas.

Lemma 2 *Assume conditions* (I–VI). *Under* $\mathcal{H}_0$*,*

$$\left\| \frac{1}{\sqrt{T}} \int_0^T w_s^T(\cdot) \frac{\dot{S}(X_s,\theta)}{\sigma(X_s)} \mathrm{d}W_s|_{\theta=\hat{\theta}_T} - \frac{1}{\sqrt{T}} \int_0^T w_s^T(\cdot) \frac{\dot{S}(X_s,\theta_0)}{\sigma(X_s)} \mathrm{d}W_s \right\|_{L^2(0,1)}$$

converges to 0 in probability as $T \to \infty$*.*

Remark 3 Let us confirm the Itô formula, which will be frequently used in the proof. Let $s \rightsquigarrow X_s$ be a one-dimensional continuous semimartingale whose predictable quadratic variation process is denoted by $s \rightsquigarrow \langle X \rangle_s$. Let the map $x \mapsto f(x)$ be second-order continuously differentiable. Its first and second derivatives are denoted by f' and f'', respectively. It holds that

$$\int_{X_0}^{X_T} f'(x)\mathrm{d}x = f(X_T) - f(X_0) = \int_0^T f'(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^T f''(X_s)\mathrm{d}\langle X \rangle_s.$$

In particular, when we consider the stochastic differential equation

$$X_t = X_0 + \int_0^t S(X_s, \theta_0)\mathrm{d}s + \int_0^t \sigma(X_s)\mathrm{d}W_s,$$

by putting $f' = g(\cdot)/\sigma^2(\cdot)$, it holds that

$$\begin{aligned}\int_{X_0}^{X_T} &\frac{g(x)}{\sigma(x)^2}\mathrm{d}x \\ &= \int_0^T \frac{g(X_s)}{\sigma(X_s)}\mathrm{d}W_s + \int_0^T \left(\frac{g(X_s)S(X_s,\theta_0)}{\sigma(X_s)^2} + \frac{g'(X_s)}{2} - \frac{\sigma'(X_s)g(X_s)}{\sigma(X_s)}\right)\mathrm{d}s.\end{aligned}$$

We shall use $\dot{S}$, $\ddot{S}$, K as g in the proof.

Lemma 3 *Assume conditions* (I–VI). *Under $\mathcal{H}_0$,*

$$\left\|\mathbb{Z}_T^p(\cdot, \hat{\theta}_T)\right\|_{L^2(0,1)}^2$$

converges to 0 in probability as $T \to \infty$.

Next, we discuss the limit behavior of $\mathbb{M}_T(\cdot, \theta_0)$, which almost surely takes values in $L^2(0, 1)$ by Propositions 2 and 4. The following lemma follows from Theorem 4.

Lemma 4 *Assume conditions* (I–VI). *Under $\mathcal{H}_0$, the random field $u \rightsquigarrow \mathbb{M}_T(u, \theta_0)$ converges weakly to $u \rightsquigarrow C_{\theta_0}(\theta_0, \theta_0)^{1/2} B_d^\circ(u)/(u(1-u))^{1/2}$ in $L^2(0, 1)$ as $T \to \infty$, where B_d° is a d-dimensional standard Brownian bridge.*

Remark 4 Lemmas 2 and 4 above yield the weak convergence of $u \rightsquigarrow \mathbb{M}_T(u, \hat{\theta}_T)$ to $u \rightsquigarrow C_{\theta_0}(\theta_0, \theta_0)^{1/2} B_d^\circ(u)/(u(1-u))^{1/2}$ in $L^2(0, 1)$. This corresponds to Lemma 3 of Negri and Nishiyama (2012), which states the weak convergence of $u \rightsquigarrow (u(1-u))^{1/2}\mathbb{M}_T(u, \hat{\theta}_T)$ to $u \rightsquigarrow C_{\theta_0}(\theta_0, \theta_0)^{1/2} B_d^\circ(u)$ in $\ell^\infty([0, 1])$ under $\mathcal{H}_0$. Their Lemma 3 follows from their Lemmas 2 and 5. It seems that weak convergences in L^2 is too weak to show the result corresponding to their Lemma 5, so we take a different approach.

The following proposition will be used in the proofs of Lemmas 2 and 4.

Proposition 5 *Let $x \mapsto f(x)$ be a function satisfying*

$$\sup_{s\in[0,\infty)} \mathbb{E}\left[f(X_s)^2\right] < \infty$$

and

$$\int_I f(x)\mu_{\theta_0}(dx) < \infty.$$

Under $\mathcal{H}_0$, it holds that

$$\mathbb{E}\left[\int_0^1 \left(\frac{1}{T}\int_0^T w_s^T(u) f(X_s) ds\right)^2 du\right] \to 0.$$

We now make some assertions that guarantee the consistency of the test.

Lemma 5 *Assume conditions* (I–VI). (i) *Under $\mathcal{H}_1$,*

$$\left\| \frac{1}{T}\int_0^T w_s^T(\cdot)\frac{\dot{S}(X_s,\theta)}{\sigma(X_s)} dW_s|_{\theta=\hat{\theta}_T} - \frac{1}{T}\int_0^T w_s^T(\cdot)\frac{\dot{S}(X_s,\theta_*)}{\sigma(X_s)} dW_s \right\|_{L^2(0,1)}^2$$

converges to 0 in probability as $T \to \infty$.
(ii) *Under $\mathcal{H}_1$,*

$$\frac{1}{T}\|\mathbb{Z}_T(\cdot,\hat{\theta}_T) - \mathbb{Z}_T(\cdot,\theta_*)\|_{L^2(0,1)}^2$$

converges to 0 in probability as $T \to \infty$.
(iii) Under $\mathcal{H}_1$, it holds that $\|\mathbb{M}_T(\cdot,\theta_*)\|_{L^2(0,1)} = O_p(1)$.

Introduce the test statistic

$$AD_T = \int_0^1 \mathbb{Z}_T(u,\hat{\theta}_T)^\top \hat{C}_T^{-1}\mathbb{Z}_T(u,\hat{\theta}_T)du,$$

where

$$\hat{C}_T = \frac{1}{T}\int_0^T \frac{\dot{S}(X_s,\hat{\theta}_T)\dot{S}(X_s,\hat{\theta}_T)^\top}{\sigma(X_s)^2}\mathrm{d}s.$$

It follows from Theorem 2 that $\hat{C}_T$ converges in probability to $C_{\theta_0}(\theta_0,\theta_0)$ under $\mathcal{H}_0$ and to $u_* C_{\theta_0}(\theta_*,\theta_*) + (1-u_*)C_{\theta_1}(\theta_*,\theta_*)$ under $\mathcal{H}_1$ (see page 915 in the study by Negri and Nishiyama 2012). The continuous mapping theorem and the Slutsky theorem yield part (i) of the following theorem.

Theorem 5 *Assume conditions* (I–VI). (i) *Under $\mathcal{H}_0$, it holds that*

$$AD_T \to^d \int_0^1 \frac{\|B_d^\circ(u)\|^2}{u(1-u)}\mathrm{du}$$

as $T \to \infty$.
(ii) *Under $\mathcal{H}_1$, the test is consistent.*

Remark 5 Theorem 1 (i) by Negri and Nishiyama (2012) shows the convergence in distribution

$$\sup_{u\in[0,1]} (u(1-u)\mathbb{Z}_T(u,\hat{\theta}_T)^\top \hat{C}_T^{-1}\mathbb{Z}_T(u,\hat{\theta}_T)) \to^d \sup_{u\in[0,1]} \|B_d^\circ(u)\|^2$$

as $T \to \infty$ under $\mathcal{H}_0$. This result corresponds to a Kolmogorov–Smirnov type test in goodness-of-fit testing in terms of its limit distribution. On the other hand, the result in Theorem 5 (i) corresponds to an Anderson–Darling type test, which often has better power than a Kolmogorov–Smirnov type test.

5 Proofs

Proof of Proposition 1. It is enough to show that $\forall \epsilon > 0$, there exists a finite subset $\{e_i : i \in I\}$ of the complete orthonormal system such that

$$\limsup_{n\to\infty} \mathbb{E}\left[\sum_{j\notin I} \langle X_n, e_j\rangle_{\mathbb{H}}^2\right] < \epsilon$$

by the Markov inequality. The Parseval identity yields

$$\|X\|_{\mathbb{H}}^2 = \sum_{j\in I} \langle X, e_j\rangle_{\mathbb{H}}^2 + \sum_{j\notin I} \langle X, e_j\rangle_{\mathbb{H}}^2,$$

so it holds that, for any $\epsilon > 0$, there exists a finite subset $I \subset J$ such that

$$\sum_{j\in I} \mathbb{E}\left[\langle X, e_j\rangle_{\mathbb{H}}^2\right] > \mathbb{E}\left[\|X\|_{\mathbb{H}}^2\right] - \epsilon.$$

Hence, it follows from the assumptions that

$$\begin{aligned}\mathbb{E}\left[\sum_{j\notin I} \langle X_n, e_j\rangle_{\mathbb{H}}^2\right] &= \mathbb{E}\left[\|X_n\|_{\mathbb{H}}^2\right] - \mathbb{E}\left[\sum_{j\in I} \langle X_n, e_j\rangle_{\mathbb{H}}^2\right] \\ &\to \mathbb{E}\left[\|X\|_{\mathbb{H}}^2\right] - \mathbb{E}\left[\sum_{j\in I} \langle X, e_j\rangle_{\mathbb{H}}^2\right] < \epsilon\end{aligned}$$

for a large enough finite set I. This completes the proof. $\square$

Proof of Proposition 2. (i) It follows from the property of the operator norm and the Jensen inequality that

$$\begin{aligned}&\sup_{s\in[0,\infty)} \left\|\mathbb{E}\left[H_s(\theta) H_s(\theta)^\top \lambda_s\right]\right\|_{OP}^2 \\ &\leq \sup_{s\in[0,\infty)} \sum_{i=1}^d \sum_{j=1}^d \left|\mathbb{E}\left[(H_s(\theta))_{(i)} (H_s(\theta))_{(j)} \lambda_s\right]\right|^2\end{aligned}$$

$$\leq \sup_{s\in[0,\infty)} \sum_{i=1}^{d}\sum_{j=1}^{d} \mathbb{E}\left[(H_s(\theta))_{(i)}^2\,(H_s(\theta))_{(j)}^2\,\lambda_s^2\right]$$
$$= \sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^4\,\lambda_s^2\right] < \infty.$$

(ii) It follows from the Schwartz inequality that

$$\sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s\|^3\lambda_s\right] \leq \sup_{s\in[0,\infty)} \left(\mathbb{E}\left[\|H_s\|^4\lambda_s^2\right]\mathbb{E}\left[\|H_s\|^2\right]\right)^{1/2} < \infty.$$

(iii) As for the former assertion, (6) implies (10) because of the Schwartz inequality. As for the latter assertion, the left-hand side of (11) is equal to

$$\mathbb{E}\left[\int_0^1 \left\|\frac{1}{\sqrt{T}}\int_0^T w_s^T(u)H_s(\theta)\mathrm{d}M_s\right\|^2 \mathrm{d}u\right]$$
$$= \int_0^1 \mathbb{E}\left[\frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \|H_s(\theta)\|^2\lambda_s\mathrm{d}s\right]\mathrm{d}u$$
$$= \int_0^1 \left(\frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right]\mathrm{d}s\right)\mathrm{d}u$$
$$\leq \sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right] < \infty$$

by the martingale property and the Fubini theorem. This completes the proof. □

Proof of Theorem 4. Let us use Proposition 1 to check the asymptotic tightness of $\mathcal{M}_T(\cdot,\theta)$ in $L^2(0,1)$. First, let us confirm criterion (3) as follows:

$$\mathbb{E}\left[\left\|\frac{1}{\sqrt{T}}\int_0^T w_s^T H_s(\theta)\mathrm{d}M_s\right\|_{L^2(0,1)}^2\right]$$
$$= \int \left(\frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right]\mathrm{d}s\right) du$$
$$\to \mathrm{tr}C(\theta,\theta) < \infty.$$

The result of the limit operation above follows from the bounded convergence theorem because the pointwise convergence

$$\frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right]\mathrm{d}s$$
$$= \left(\frac{1-u}{Tu}\int_0^{Tu} + \frac{u}{T(1-u)}\int_{Tu}^T\right)\mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right]\mathrm{d}s$$
$$\to (1-u)\mathrm{tr}C(\theta,\theta) + u\mathrm{tr}C(\theta,\theta) = \mathrm{tr}C(\theta,\theta)$$

for all $u \in (0, 1)$ follows from assumption (12) and the fact that

$$\frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right]\mathrm{d}s \le \sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^2\lambda_s\right].$$

Next we argue the convergence of the inner product

$$\left\langle \frac{1}{\sqrt{T}}\int_0^T w_s^T H_s(\theta)\mathrm{d}M_s, h\right\rangle_{L^2(0,1)}$$

for $h \in L^2(0, 1)$ which also leads to (4). The preceding expression is equal to

$$\frac{1}{\sqrt{T}}\int_0^T \left\langle w_s^T H_s(\theta), h\right\rangle_{L^2(0,1)} \mathrm{d}M_s$$

by the Fubini theorem for stochastic integrals. We shall apply the central limit theorem for martingales. The predictable quadratic variation of the inner product is

$$\frac{1}{T}\int_0^T \left\langle w_s^T H_s(\theta), h\right\rangle_{L^2(0,1)}^2 \lambda_s \mathrm{d}s. \tag{17}$$

Define

$$V_T = \mathbb{E}\left[\frac{1}{T}\int_0^T \left\langle w_s^T H_s(\theta), h\right\rangle_{L^2(0,1)}^2 \lambda_s \mathrm{d}s\right],$$

then it holds that

$$\begin{aligned} V_T &= \mathbb{E}\left[\frac{1}{T}\int_0^T\int_0^1\int_0^1 w_s^T(u)w_s^T(v)h(u)^\top H_s(\theta)H_s(\theta)^\top h(v)\mathrm{d}u\mathrm{d}v\lambda_s\mathrm{d}s\right] \\ &= \int_0^1\int_0^1 \frac{1}{T}\int_0^T w_s^T(u)w_s^T(v)h(u)^\top \mathbb{E}\left[H_s(\theta)H_s(\theta)^\top\lambda_s\right]h(v)\mathrm{d}s\mathrm{d}u\mathrm{d}v. \end{aligned}$$

Thus, we see that

$$V_T \to \int_0^1\int_0^1 \frac{u\wedge v - uv}{\sqrt{u(1-u)v(1-v)}} h(u)^\top C(\theta,\theta)h(v)\mathrm{d}u\mathrm{d}v \tag{18}$$

as $T \to \infty$. Pointwise convergence for any u, v follows from Proposition 3. Because of the Schwartz inequality, it holds that

$$
\begin{aligned}
&\frac{1}{T}\int_0^T w_s^T(u)w_s^T(v)h(u)^\top \mathbb{E}\left[H_s(\theta)H_s(\theta)^\top \lambda_s\right]h(v)\mathrm{d}s \\
&\le \left(\frac{1}{T^2}\int_0^T \left(w_s^T(u)h(u)^\top \mathbb{E}\left[H_s(\theta)H_s(\theta)^\top \lambda_s\right]h(v)\right)^2 \mathrm{d}s \int_0^T (w_s^T(v))^2 \mathrm{d}s\right)^{1/2} \\
&\le \left(\frac{1}{T}\int_0^T (w_s^T(u))^2 \sup_{s\in[0,\infty)} \left(h(u)^\top \mathbb{E}\left[H_s(\theta)H_s(\theta)^\top \lambda_s\right]h(v)\right)^2 \mathrm{d}s\right)^{1/2} \\
&= \sup_{s\in[0,\infty)} \left|h(u)^\top \mathbb{E}\left[H_s(\theta)H_s(\theta)^\top \lambda_s\right]h(v)\right|.
\end{aligned}
$$

The right-hand side is integrable by the Schwartz inequality for the Euclidean inner product, which gives an upper bound for the right-hand side,

$$
\|h(u)\|\|h(v)\| \sup_{s\in[0,\infty)} \left\|\mathbb{E}\left[H_s(\theta)H_s(\theta)^\top \lambda_s\right]\right\|_{OP},
$$

and by Proposition 2. Therefore, the dominated convergence theorem yields (18). Though it is not obvious, it holds that (17) converges to the right-hand side of (18) in probability because of assumptions (6), (8) and (12). Finally, let us confirm the Lyapunov type condition

$$
\mathbb{E}\left[\frac{1}{T^{(2+\delta_0)/2}}\int_0^T \left\langle w_s^T H_s(\theta), h\right\rangle_{L^2(0,1)}^{2+\delta_0} \lambda_s ds\right] \to 0
$$

for some $\delta_0 > 0$. The Schwartz inequality and the Jensen inequality give an upper bound for the left-hand side

$$
\begin{aligned}
&\frac{1}{T^{(2+\delta_0)/2}}\mathbb{E}\left[\int_0^T \left\|w_s^T H_s(\theta)\right\|_{L^2(0,1)}^{2+\delta_0} \lambda_s \mathrm{d}s\right] \|h\|_{L^2(0,1)}^{2+\delta_0} \\
&\le \frac{1}{T^{(2+\delta_0)/2}}\mathbb{E}\left[\int_0^T\int_0^1 \|w_s^T(u)H_s(\theta)\|^{2+\delta_0} du\lambda_s \mathrm{d}s\right] \|h\|_{L^2(0,1)}^{2+\delta_0} \\
&= \frac{1}{T^{(2+\delta_0)/2}}\int_0^1\int_0^T |w_s^T(u)|^{2+\delta_0}\mathbb{E}\left[\|H_s(\theta)\|^{2+\delta_0}\lambda_s\right] \mathrm{d}s du \|h\|_{L^2(0,1)}^{2+\delta_0} \\
&\le \frac{1}{T^{\delta_0/2}}\int_0^1 \frac{u^{1+\delta_0}+(1-u)^{1+\delta_0}}{(u(1-u))^{\delta_0/2}}\mathrm{d}u \sup_{s\in[0,\infty)} \mathbb{E}\left[\|H_s(\theta)\|^{2+\delta_0}\lambda_s\right] \|h\|_{L^2(0,1)}^{2+\delta_0}.
\end{aligned}
$$

Setting $\delta_0 = 1$, the right-hand side converges to 0. Hence, the central limit theorem for martingales yields the conclusion. □

Proof of Proposition 3. It follows from

$$\frac{u\wedge v-uv}{\sqrt{uv(1-u)(1-v)}}=\frac{1}{T}\int_0^T w_s^T(u)w_s^T(v)\mathrm{d}s$$

that

$$\begin{aligned}
&\frac{1}{T}\int_0^T w_s^T(u)w_s^T(v)J_s(\theta)\mathrm{d}s-\frac{u\wedge v-uv}{\sqrt{uv(1-u)(1-v)}}C(\theta,\theta)\\
&=\frac{1}{T}\int_0^T w_s^T(u)w_s^T(v)(J_s(\theta)-C(\theta,\theta))\mathrm{d}s\\
&=\frac{u\wedge v}{T}\int_0^{T(u\wedge v)}\frac{(J_s(\theta)-C(\theta,\theta))}{\sqrt{uv(1-u)(1-v)}}\mathrm{d}s-\frac{u}{T}\int_0^{Tv}\frac{(J_s(\theta)-C(\theta,\theta))}{\sqrt{uv(1-u)(1-v)}}\mathrm{d}s\\
&\quad-\frac{v}{T}\int_0^{Tu}\frac{(J_s(\theta)-C(\theta,\theta))}{\sqrt{uv(1-u)(1-v)}}\mathrm{d}s+\frac{uv}{T}\int_0^{T}\frac{(J_s(\theta)-C(\theta,\theta))}{\sqrt{uv(1-u)(1-v)}}\mathrm{d}s.
\end{aligned}$$

All terms of the right-hand side converge to 0 by assumption (12). This completes the proof. □

Proof of Proposition 4. (i) By the assumptions, there exist constants C, $p\geq 1$ such that

$$\begin{aligned}
\sup_{s\in[0,\infty)}\mathbb{E}\left[\frac{\|\dot{S}(X_s,\theta_0)\|^4}{\sigma(X_s)^4}\right]&=\sup_{s\in[0,\infty)}\mathbb{E}\left[\frac{\left(\sum_{i=1}^d(\partial_i S(X_s,\theta_0))^2\right)^2}{\sigma(X_s)^4}\right]\\
&\leq\sup_{s\in[0,\infty)}\mathbb{E}\left[\frac{d\sum_{i=1}^d(\partial_i S(X_s,\theta_0))^4}{\sigma(X_s)^4}\right]\\
&\leq\sup_{s\in[0,\infty)}\mathbb{E}\left[d\sum_{i=1}^d\frac{\sup_{\theta\in N}|\partial_i S(X_s,\theta)|^4}{\inf_{x\in\mathbb{R}}\sigma(x)^4}\right]\\
&\leq\sup_{s\in[0,\infty)}\mathbb{E}\left[d\sum_{i=1}^d\frac{|C(1+|X_s|)^p}{\inf_{x\in\mathbb{R}}\sigma(x)^4}\right]\\
&=\frac{Cd^2}{\inf_{x\in\mathbb{R}}\sigma(x)^4}\sup_{s\in[0,\infty)}\mathbb{E}\left[|1+|X_s||^p\right]<\infty.
\end{aligned}$$

Hence, (15) holds. (16) follows from (15) and condition (II). This completes the proof.

(ii) The proof can be done in the same manner as for part (i). □

Proof of Lemma 2. It follows from the Itô formula that

$$\begin{aligned}
&\int_0^{Tu}\frac{\dot{S}(X_s,\theta)}{\sigma(X_s)}\mathrm{d}W_s\\
&=\int_{X_0}^{X_{Tu}}\frac{\dot{S}(x,\theta)}{\sigma(x)^2}\mathrm{d}x-\int_0^{Tu}\left(\frac{\dot{S}(X_s,\theta)S(X_s,\theta_0)}{\sigma(X_s)^2}+\frac{\dot{S}'(X_s,\theta)}{2}-\frac{\sigma'(X_s)\dot{S}(X_s,\theta)}{\sigma(X_s)}\right)\mathrm{d}s
\end{aligned}$$

and that

$$\int_{Tu}^{T} \frac{\dot{S}(X_s,\theta)}{\sigma(X_s)} \mathrm{d}W_s = \int_{X_{Tu}}^{X_T} \frac{\dot{S}(x,\theta)}{\sigma(x)^2}\mathrm{d}x - \int_{Tu}^{T}\left(\frac{\dot{S}(X_s,\theta)S(X_s,\theta_0)}{\sigma(X_s)^2} + \frac{\dot{S}'(X_s,\theta)}{2} - \frac{\sigma'(X_s)\dot{S}(X_s,\theta)}{\sigma(X_s)}\right)\mathrm{d}s.$$

Noting that

$$\mathbb{M}(u,\theta) = \frac{1}{\sqrt{Tu(1-u)}}\left((1-u)\int_0^{Tu}\frac{\dot{S}(X_s,\theta)}{\sigma(X_s)}\mathrm{d}W_s - u\int_{Tu}^{T}\frac{\dot{S}(X_s,\theta)}{\sigma(X_s)}\mathrm{d}W_s\right),$$

a Taylor expansion around θ_0 yields

$$\begin{aligned}
&\|\mathbb{M}(u,\hat{\theta}_T) - \mathbb{M}(u,\theta_0)\| \\
&\quad\le \frac{\|\hat{\theta}_T-\theta_0\|}{T\sqrt{u(1-u)}}\left\|(1-u)\int_{X_0}^{X_{Tu}}\frac{\ddot{S}(x,\tilde{\theta}_T)}{\sigma(x)^2}\mathrm{d}x - u\int_{X_{Tu}}^{X_T}\frac{\ddot{S}(x,\tilde{\theta}_T)}{\sigma(x)^2}\mathrm{d}x\right. \\
&\qquad -(1-u)\int_0^{Tu}\left(\frac{\ddot{S}(X_s,\tilde{\theta}_T)S(X_s,\theta_0)}{\sigma(X_s)^2} + \frac{\ddot{S}'(X_s,\check{\theta}_T)}{2} - \frac{\sigma'(X_s)\ddot{S}(X_s,\tilde{\theta}_T)}{\sigma(X_s)}\right)\mathrm{d}s \\
&\qquad \left. +u\int_{Tu}^{T}\left(\frac{\ddot{S}(X_s,\tilde{\theta}_T)S(X_s,\theta_0)}{\sigma(X_s)^2} + \frac{\ddot{S}'(X_s,\check{\theta}_T)}{2} - \frac{\sigma'(X_s)\ddot{S}(X_s,\tilde{\theta}_T)}{\sigma(X_s)}\right)\mathrm{d}s\right\|,
\end{aligned}$$

where $\tilde{\theta}_T$ and $\check{\theta}_T$ are elements between $\hat{\theta}_T$ and θ_0. The triangle inequality yields the bound

$$\begin{aligned}
&\|\mathbb{M}(u,\hat{\theta}_T) - \mathbb{M}(u,\theta_0)\| \\
&\quad\le \frac{\|\sqrt{T}(\hat{\theta}_T-\theta_0)\|}{T\sqrt{u(1-u)}}\left\|(1-u)\int_{X_0}^{X_{Tu}}\frac{\ddot{S}(x,\tilde{\theta}_T)}{\sigma(x)^2}\mathrm{d}x - u\int_{X_{Tu}}^{X_T}\frac{\ddot{S}(x,\tilde{\theta}_T)}{\sigma(x)^2}\mathrm{d}x\right\| \\
&\qquad +\frac{\|\sqrt{T}(\hat{\theta}_T-\theta_0)\|}{T\sqrt{u(1-u)}} \\
&\qquad \times\left\|-(1-u)\int_0^{Tu}\left(\frac{\ddot{S}(X_s,\tilde{\theta}_T)S(X_s,\theta_0)}{\sigma(X_s)^2} + \frac{\ddot{S}'(X_s,\check{\theta}_T)}{2} - \frac{\sigma'(X_s)\ddot{S}(X_s,\tilde{\theta}_T)}{\sigma(X_s)}\right)\mathrm{d}s\right. \\
&\qquad \left. +u\int_{Tu}^{T}\left(\frac{\ddot{S}(X_s,\tilde{\theta}_T)S(X_s,\theta_0)}{\sigma(X_s)^2} + \frac{\ddot{S}'(X_s,\check{\theta}_T)}{2} - \frac{\sigma'(X_s)\ddot{S}(X_s,\tilde{\theta}_T)}{\sigma(X_s)}\right)\mathrm{d}s\right\|. \qquad (19)
\end{aligned}$$

For the second factor of the first term of (19), the triangle inequality gives

$$
\begin{aligned}
&\left\|(1-u)\int_{X_0}^{X_{Tu}}\frac{\ddot{S}(x,\tilde{\theta}_T)}{\sigma(x)^2}\mathrm{d}x-u\int_{X_{Tu}}^{X_T}\frac{\ddot{S}(x,\tilde{\theta}_T)}{\sigma(x)^2}\mathrm{d}x\right\| \\
&\quad\le\left\|(1-u)\int_{X_0}^{X_{Tu}}\frac{\ddot{S}(x,\theta_0)}{\sigma(x)^2}\mathrm{d}x-u\int_{X_{Tu}}^{X_T}\frac{\ddot{S}(x,\theta_0)}{\sigma(x)^2}\mathrm{d}x\right\| \\
&\qquad+\left\|(1-u)\int_{X_0}^{X_{Tu}}\frac{(\ddot{S}(x,\tilde{\theta}_T)-\ddot{S}(x,\theta_0))}{\sigma(x)^2}\mathrm{d}x-u\int_{X_{Tu}}^{X_T}\frac{(\ddot{S}(x,\tilde{\theta}_T)-\ddot{S}(x,\theta_0))}{\sigma(x)^2}\mathrm{d}x\right\|.
\end{aligned}
\tag{20}
$$

By the Itô formula, the first term is equal to

$$
\begin{aligned}
&\Bigg\|(1-u)\Bigg[\int_0^{Tu}\frac{\ddot{S}(X_s,\theta_0)}{\sigma(X_s)}\mathrm{d}W_s \\
&\quad+\int_0^{Tu}\left(\frac{\ddot{S}(X_s,\theta_0)S(X_s,\theta_0)}{\sigma(X_s)^2}+\frac{\ddot{S}'(X_s,\theta_0)}{2}-\frac{\sigma'(X_s)\ddot{S}(X_s,\theta_0)}{\sigma(X_s)}\right)\mathrm{d}s\Bigg] \\
&\quad-u\Bigg[\int_{Tu}^{T}\frac{\ddot{S}(X_s,\theta_0)}{\sigma(X_s)}\mathrm{d}W_s \\
&\quad+\int_{Tu}^{T}\left(\frac{\ddot{S}(X_s,\theta_0)S(X_s,\theta_0)}{\sigma(X_s)^2}+\frac{\ddot{S}'(X_s,\theta_0)}{2}-\frac{\sigma'(X_s)\ddot{S}(X_s,\theta_0)}{\sigma(X_s)}\right)\mathrm{d}s\Bigg]\Bigg\|
\end{aligned}
$$

and the second term on the right-hand side of (20) is bounded above by

$$
\left\|(1-u)\int_{X_0}^{X_{Tu}}\frac{|\ddot{S}(x,\tilde{\theta}_T)-\ddot{S}(x,\theta_0)|}{\sigma(x)^2}\mathrm{d}x+u\int_{X_{Tu}}^{X_T}\frac{|\ddot{S}(x,\tilde{\theta}_T)-\ddot{S}(x,\theta_0)|}{\sigma(x)^2}\mathrm{d}x\right\|.
$$

Therefore, the right-hand side of (20) is bounded by

$$
\begin{aligned}
&\left\|\int_0^T(1\{s\le Tu\}-u)\frac{\ddot{S}(X_s,\theta_0)}{\sigma(X_s)}\mathrm{d}W_s\right\| \\
&\quad+\left\|\int_0^T(1\{s\le Tu\}-u)\left(\frac{\ddot{S}(X_s,\theta_0)S(X_s,\theta_0)}{\sigma(X_s)^2}+\frac{\ddot{S}'(X_s,\theta_0)}{2}-\frac{\sigma'(X_s)\ddot{S}(X_s,\theta_0)}{\sigma(X_s)}\right)\mathrm{d}s\right\| \\
&\quad+d\left|(1-u)\int_{X_0}^{X_{Tu}}\frac{K(x)}{\sigma(x)^2}\mathrm{d}x+u\int_{X_{Tu}}^{X_T}\frac{K(x)}{\sigma(x)^2}\mathrm{d}x\right|\|\tilde{\theta}_T-\theta_0\|
\end{aligned}
$$

because of condition (VI). The second term of (19) is equal to

$$
\begin{aligned}
&\left\|-\frac{1}{T}\int_0^T w_s^T(u)\left(\frac{\ddot{S}(X_s,\tilde{\theta}_T)S(X_s,\theta_0)}{\sigma(X_s)^2}+\frac{\ddot{S}'(X_s,\check{\theta}_T)}{2}-\frac{\sigma'(X_s)\ddot{S}(X_s,\tilde{\theta}_T)}{\sigma(X_s)}\right)\mathrm{d}s\right\| \\
&\|\sqrt{T}(\hat{\theta}_T-\theta_0)\|.
\end{aligned}
$$

Since $\|\sqrt{T}(\hat{\theta}_T - \theta_0)\| = O_p(1)$, let us confirm that the first factor converges in probability to 0. The triangle inequality yields an upper bound for the first factor of

$$\begin{aligned}&\left\| \frac{1}{T}\int_0^T w_s^T(u)\left(\frac{\ddot{S}(X_s,\theta_0)S(X_s,\theta_0)}{\sigma(X_s)^2} + \frac{\ddot{S}'(X_s,\theta_0)}{2} - \frac{\sigma'(X_s)\ddot{S}(X_s,\theta_0)}{\sigma(X_s)}\right)\mathrm{d}s\right\| \\ &\quad + \left\| \frac{1}{T}\int_0^T w_s^T(u)\left(\frac{(\ddot{S}(X_s,\tilde{\theta}_T) - \ddot{S}(X_s,\theta_0))S(X_s,\theta_0)}{\sigma(X_s)^2}\right.\right. \\ &\quad \left.\left. + \frac{\ddot{S}'(X_s,\check{\theta}_T) - \ddot{S}'(X_s,\theta_0)}{2} - \frac{\sigma'(X_s)(\ddot{S}(X_s,\tilde{\theta}_T) - \ddot{S}(X_s,\theta_0))}{\sigma(X_s)}\right)\mathrm{d}s\right\|.\end{aligned}$$

The first term will be considered in (22). The absolute value of each element in the norm of the second term is bounded above by

$$\begin{aligned}&\frac{1}{T}\int_0^T \left|w_s^T(u)\right|\left(\frac{|\partial_i\partial_j S(X_s,\tilde{\theta}_T) - \partial_i\partial_j S(X_s,\theta_0)||S(X_s,\theta_0)|}{\sigma(X_s)^2}\right. \\ &\qquad \left. + \frac{|\partial_i\partial_j S'(X_s,\check{\theta}_T) - \partial_i\partial_j S'(X_s,\theta_0)|}{2} + \frac{|\sigma'(X_s)||\partial_i\partial_j S(X_s,\tilde{\theta}_T) - \partial_i\partial_j S(X_s,\theta_0)|}{\sigma(X_s)}\right)\mathrm{d}s \\ &\le \frac{1}{T}\int_0^T \left|w_s^T(u)\right|\left(\frac{K(X_s)|S(X_s,\theta_0)|}{\sigma(X_s)^2} + \frac{K_d(X_s)}{2} + \frac{|\sigma'(X_s)|K(X_s)}{\sigma(X_s)}\right)\mathrm{d}s \\ &\|\hat{\theta}_T - \theta_0\|.\end{aligned}$$

The Schwartz inequality yields the following bound for the left-hand factor of the right-hand side:

$$\left(\frac{1}{T}\int_0^T \left(\frac{K(X_s)|S(X_s,\theta_0)|}{\sigma(X_s)^2} + \frac{K_d(X_s)}{2} + \frac{|\sigma'(X_s)|K(X_s)}{\sigma(X_s)}\right)^2 \mathrm{d}s\right)^{1/2}.$$

Its $L^2(0,1)$ norm is asymptotically tight in $\mathbb{R}$ because of ergodicity. Therefore, it suffices to prove that

$$\left\| \frac{1}{T}\int_0^T w_s^T(\cdot)\frac{\ddot{S}(X_s,\theta_0)}{\sigma(X_s)}\mathrm{d}W_s\right\|_{L^2((0,1))}^2 \to^p 0, \tag{21}$$

$$\begin{aligned}&\left\| \frac{1}{T}\int_0^T w_s^T(\cdot)\left(\frac{\ddot{S}(X_s,\theta_0)S(X_s,\theta_0)}{\sigma(X_s)^2} + \frac{\ddot{S}'(X_s,\theta_0)}{2}\right.\right. \\ &\qquad \left.\left. - \frac{\sigma'(X_s)\ddot{S}(X_s,\theta_0)}{\sigma(X_s)}\right)\mathrm{d}s\right\|_{L^2(0,1)}^2 \to^p 0,\end{aligned} \tag{22}$$

$$\left\| \frac{1-\cdot}{T^{3/2}\sqrt{\cdot(1-\cdot)}}\int_{X_0}^{X_{T\cdot}} \frac{K(x)}{\sigma(x)^2}\mathrm{d}x\right\|_{L^2(0,1)}^2 \quad \to^p 0, \tag{23}$$

and

$$\left\| \frac{\cdot}{T^{3/2}\sqrt{\cdot(1-\cdot)}} \int_{X_{T\cdot}}^{X_T} \frac{K(x)}{\sigma(x)^2} \mathrm{d}x \right\|_{L^2(0,1)}^2 \to^p 0. \tag{24}$$

for the limit in (21) follows from

$$\begin{aligned} &\frac{1}{T^2}\int_0^1 \mathbb{E}\left[\left|\int_0^T w_s^T(u)\frac{\partial_i\partial_j S(X_s,\theta_0)}{\sigma(X_s)}\mathrm{d}W_s\right|^2\right]\mathrm{d}u \\ &= \frac{1}{T}\int_0^1 \mathbb{E}\left[\frac{1}{T}\int_0^T \left(w_s^T(u)\right)^2 \frac{(\partial_i\partial_j S(X_s,\theta_0))^2}{\sigma(X_s)^2}\mathrm{d}s\right]\mathrm{d}u \\ &\le \frac{1}{T}\sup_{s\in[0,\infty)} \mathbb{E}\left[\frac{(\partial_i\partial_j S(X_s,\theta_0))^2}{\sigma(X_s)^2}\right] \to 0 \end{aligned}$$

for any i, j. To show (22), it is enough to prove the convergence of the expectation to 0. This follows from Proposition 5 and condition (I). For (23), since the Itô formula yields

$$\begin{aligned} &\int_{X_0}^{X_{Tu}} \frac{K(x)}{\sigma(x)^2}\mathrm{d}x \\ &= \int_0^{Tu} \frac{K(X_s)}{\sigma(X_s)}\mathrm{d}W_s + \int_0^{Tu}\left(\frac{K(X_s)S(X_s,\theta_0)}{\sigma(X_s)^2} + \frac{K'(X_s)}{2} - \frac{\sigma'(X_s)K(X_s)}{\sigma(X_s)}\right)\mathrm{d}s, \end{aligned}$$

it suffices to prove that

$$\mathbb{E}\left[\int_0^1 \left(\frac{1-u}{T^{3/2}\sqrt{u(1-u)}}\int_0^{Tu}\frac{K(X_s)}{\sigma(X_s)}\mathrm{d}W_s\right)^2 \mathrm{d}u\right] \to 0 \tag{25}$$

and that

$$\begin{aligned} \mathbb{E}\Bigg[\int_0^1 \frac{u(1-u)}{T}\Bigg(\frac{1}{Tu}\int_0^{Tu}\Bigg(&\frac{K(X_s)S(X_s,\theta_0)}{\sigma(X_s)^2} \\ &+\frac{K'(X_s)}{2} - \frac{\sigma'(X_s)K(X_s)}{\sigma(X_s)}\Bigg)\mathrm{d}s\Bigg)^2\mathrm{d}u\Bigg] \to 0. \end{aligned} \tag{26}$$

Limit (25) holds because the left-hand side is equal to

$$\begin{aligned} &\int_0^1 \frac{1-u}{T^3u}\mathbb{E}\left[\int_0^{Tu}\frac{(K(X_s))^2}{\sigma(X_s)^2}\mathrm{d}s\right]\mathrm{d}u \\ &\le \frac{1}{T^2}\int_0^1 (1-u)du \sup_{s\in[0,\infty)} \mathbb{E}\left[\frac{(K(X_s))^2}{\sigma(X_s)^2}\right] \to 0. \end{aligned}$$

Limit (26) holds because the Jensen inequality gives an upper bound for the left-hand side:

$$\mathbb{E}\left[\int_0^1 \frac{1-u}{T^2}\int_0^{Tu}\left(\frac{K(X_s)S(X_s,\theta_0)}{\sigma(X_s)^2}+\frac{K'(X_s)}{2}-\frac{\sigma'(X_s)K(X_s)}{\sigma(X_s)}\right)^2 \mathrm{d}s\mathrm{d}u\right]$$
$$\leq \frac{1}{T}\int_0^1 u(1-u)du \sup_{s\in[0,\infty)}\mathbb{E}\left[\left(\frac{K(X_s)S(X_s,\theta_0)}{\sigma(X_s)^2}+\frac{K'(X_s)}{2}-\frac{\sigma'(X_s)K(X_s)}{\sigma(X_s)}\right)^2\right]$$

which converges to 0. Limit (24) is also valid for the same reason as (23). This completes the proof of Lemma 2. □

Proof of Lemma 3. A Taylor expansion yields

$$\begin{aligned}\mathbb{Z}_T^p(u,\hat\theta_T) &= \frac{1}{\sqrt{T}}\int_0^T w_s^T(u)\frac{\dot S(X_s,\hat\theta_T)}{\sigma(X_s)^2}(S(X_s,\theta_0)-S(X_s,\hat\theta_T))\mathrm{d}s\\ &= \frac{1}{T}\int_0^T w_s^T(u)\frac{\dot S(X_s,\hat\theta_T)}{\sigma(X_s)^2}\dot S(X_s,\tilde\theta_T)^\top \mathrm{d}s\sqrt{T}(\hat\theta_T-\theta_0),\end{aligned}$$

where $\tilde\theta_T$ is a value between θ_0 and $\hat\theta_T$. Because it holds that $\sqrt{T}(\hat\theta_T-\theta_0)=O_P(1)$, it suffices to show the convergence to 0 in probability in $L^2(0,1)$ of the all elements in the matrix

$$\begin{aligned}&\frac{1}{T}\int_0^T w_s^T(\cdot)\frac{\dot S\left(X_s,\hat\theta_T\right)}{\sigma(X_s)^2}\dot S\left(X_s,\tilde\theta_T\right)^\top \mathrm{d}s\\ &=\frac{1}{T}\int_0^T w_s^T(\cdot)\frac{\dot S(X_s,\theta_0)}{\sigma(X_s)^2}\dot S\,(X_s,\theta_0)^\top\mathrm{d}s\\ &\quad+\frac{1}{T}\int_0^T w_s^T(\cdot)\frac{\dot S\left(X_s,\hat\theta_T\right)}{\sigma(X_s)^2}\left(\dot S\left(X_s,\tilde\theta_T\right)-\dot S(X_s,\theta_0)\right)^\top\mathrm{d}s\\ &\quad+\frac{1}{T}\int_0^T w_s^T(\cdot)\frac{\left(\dot S\left(X_s,\hat\theta_T\right)-\dot S(X_s,\theta_0)\right)}{\sigma(X_s)^2}\dot S(X_s,\theta_0)^\top\mathrm{d}s.\end{aligned} \tag{27}$$

It is sufficient to prove each term in the right-hand side converges to 0 in $L^2(0,1)$. The $L^2(0,1)$-norm of each element of the first term converges to 0 in mean square by Proposition 5. As for the second term, by the Schwartz inequality and the Taylor expansion, the absolute value of the (i,j)-element for any i,j is bounded above by

$$\left(\frac{1}{T^2}\int_0^T\left(w_s^T(u)\right)^2\mathrm{d}s\int_0^T\frac{\left(\partial_i S\left(X_s,\hat\theta_T\right)\partial_j S\left(X_s,\tilde\theta_T\right)-\partial_j S(X_s,\theta_0)\right)^2}{\sigma(X_s)^4}\mathrm{d}s\right)^{1/2}$$
$$=\left(\left(\tilde\theta_T-\theta_0\right)^\top\frac{1}{T}\int_0^T\frac{\left(\partial_i S\left(X_s,\hat\theta_T\right)\right)^2\partial_j\dot S\left(X_s,\grave\theta_T\right)\partial_j\dot S\left(X_s,\grave\theta_T\right)^\top}{\sigma(X_s)^4}\mathrm{d}s(\tilde\theta_T-\theta_0)\right)^{1/2}$$

for any $u \in (0, 1)$, where $\grave{\theta}_T$ is a value between $\tilde{\theta}_T$ and θ_0. Now it holds that

$$\frac{1}{T}\int_0^T \frac{\left(\partial_i S\left(X_s, \hat{\theta}_T\right)\right)^2 \partial_j \dot{S}\left(X_s, \grave{\theta}_T\right) \partial_j \dot{S}\left(X_s, \grave{\theta}_T\right)^\top}{\sigma(X_s)^4} \mathrm{d}s = O_p(1)$$

because the absolute value of the (i', j') element of its expectation is bounded above by

$$\begin{aligned}
&\sup_{s\in[0,\infty)} \mathbb{E}\left[\frac{\left(\partial_i S\left(X_s, \hat{\theta}_T\right)\right)^2 |\partial_i \partial_{i'} S\left(X_s, \grave{\theta}_T\right) \partial_j \partial_{j'} S\left(X_s, \grave{\theta}_T\right)|}{\sigma(X_s)^4}\right] \\
&\quad\leq \frac{1}{\inf_{x\in\mathbb{R}} \sigma(x)^4} \sup_{s\in[0,\infty)} \mathbb{E}\left[(\sup_{\theta\in\Theta} \left|\partial_i S(X_s, \theta)\right|)^2 \sup_{\theta\in\Theta} |\partial_i \partial_{i'} S(X_s, \theta)|\right. \\
&\quad\left. \sup_{\theta\in\Theta} |\partial_j \partial_{j'} S(X_s, \theta))|\right],
\end{aligned}$$

whereas $(\tilde{\theta}_T - \theta_0)$ converges to 0 in probability. Since the bound does not depend on u, the $L^2(0, 1)$-norm of the second term in (27) converges to 0 in probability because each element converges to 0 in probability. The $L^2(0, 1)$-norm of the third term in (27) also converges to 0 in probability for the same reason. This completes the proof. □

Proof of Proposition 5. It follows from the Schwartz inequality that

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{1}{T}\int_0^T w_s^T(u) f(X_s)\mathrm{d}s\right)^2\right] &\leq \mathbb{E}\left[\frac{1}{T^2}\int_0^T \left(w_s^T(u)\right)^2 \mathrm{d}s \int_0^T f(X_s)^2 \mathrm{d}s\right] \\
&\leq \sup_{s\in[0,\infty)} \mathbb{E}\left[f(X_s)^2\right].
\end{aligned}$$

The right-hand side is integrable with respect to u. Moreover, it holds that

$$\mathbb{E}\left[\left(\frac{1}{T}\int_0^T w_s^T(u) f(X_s)\mathrm{d}s\right)^2\right] \to 0$$

for any $u \in (0, 1)$ because

$$\frac{1}{T}\int_0^T w_s^T(u) f(X_s)\mathrm{d}s = u\frac{1}{Tu}\int_0^{Tu} f(X_s)\mathrm{d}u - u\frac{1}{T}\int_0^T f(X_s)\mathrm{d}u$$

converges to

$$u\int_I f(x)\mu_{\theta_0}(\mathrm{d}x) - u\int_I f(x)\mu_{\theta_0}(\mathrm{d}x) = 0$$

in mean square for any $u \in (0,1)$. Therefore, the Fubini theorem and the dominated convergence theorem yield the conclusion. This completes the proof. □

Proof of Lemma 5. (i) It follows from the Itô formula that

$$\frac{1}{T}\int_0^T w_s^T(u)\frac{\dot{S}(X_s,\theta)}{\sigma(X_s)}\mathrm{d}W_s|_{\theta=\hat{\theta}_T} - \frac{1}{T}\int_0^T w_s^T(u)\frac{\dot{S}(X_s,\theta_*)}{\sigma(X_s)}\mathrm{d}W_s$$
$$= \frac{1}{T\sqrt{u(1-u)}}\left((1-u)\int_{X_0}^{X_{Tu}}\frac{\dot{S}(x,\hat{\theta}_T)-\dot{S}(x,\theta_*)}{\sigma(x)^2}\mathrm{d}x\right.$$
$$\left.-u\int_{X_{Tu}}^{X_T}\frac{\dot{S}\left(x,\hat{\theta}_T\right)-\dot{S}(x,\theta_*)}{\sigma(x)^2}\mathrm{d}x\right)$$
$$-\frac{1}{T}\int_0^T w_s^T(u)\left(\frac{\dot{S}\left(X_s,\hat{\theta}_T\right)-\dot{S}(X_s,\theta_*)}{\sigma(X_s)}\left(\frac{S(X_s,\theta_{(s)})}{\sigma(X_s)}-\sigma'(X_s)\right)\right.$$
$$\left.+\frac{\dot{S}'\left(X_s,\hat{\theta}_T\right)-\dot{S}'(X_s,\theta_*)}{2}\right)\mathrm{d}s, \tag{28}$$

for any $u \in (0,1)$. By the Taylor expansion, the first term on the right-hand side of (28) is equal to

$$\left((1-u)\int_{X_0}^{X_{Tu}}\frac{\ddot{S}(x,\theta_*)}{\sigma(x)^2}\mathrm{d}x - u\int_{X_{Tu}}^{X_T}\frac{\ddot{S}(x,\theta_*)}{\sigma(x)^2}\mathrm{d}x\right)\frac{\left(\hat{\theta}_T-\theta_*\right)}{T\sqrt{u(1-u)}}$$
$$+(1-u)\int_{X_0}^{X_{Tu}}\frac{(\ddot{S}(x,\tilde{\theta}_T)-\ddot{S}(x,\theta_*))}{\sigma(x)^2}\mathrm{d}x\frac{\left(\hat{\theta}_T-\theta_*\right)}{T\sqrt{u(1-u)}}$$
$$-u\int_{X_{Tu}}^{X_T}\frac{(\ddot{S}(x,\tilde{\theta}_T)-\ddot{S}(x,\theta_*))}{\sigma(x)^2}\mathrm{d}x\frac{\left(\hat{\theta}_T-\theta_*\right)}{T\sqrt{u(1-u)}} \tag{29}$$

where $\tilde{\theta}_T$ lies between $\hat{\theta}_T$ and θ_*. The first term of (29) is

$$\int_0^T w_s^T(u)\frac{\ddot{S}(X_s,\theta_*)}{\sigma(X_s)^2}\mathrm{d}W_s\frac{\left(\hat{\theta}_T-\theta_*\right)}{T}$$
$$+\int_0^T w_s^T(u)\left(\frac{\ddot{S}(X_s,\theta_*)}{\sigma(X_s)}\left(\frac{S(X_s,\theta_{(s)})}{\sigma(X_s)}-\sigma'(X_s)\right)+\frac{\ddot{S}(X_s,\theta_*)}{2}\right)\mathrm{d}s\frac{\left(\hat{\theta}_T-\theta_*\right)}{T}$$

because of the Itô formula. Both terms converge in probability to 0 uniformly in $u \in (0,1)$ because $\hat{\theta}_T-\theta_*$ converges in probability to 0 and the expectation of the square of the remainders are $O(1)$ uniformly in $u \in (0,1)$. As for the second term of (29), it is enough to prove that

$$\frac{(1-u)}{T\sqrt{u(1-u)}}\int_{X_0}^{X_{Tu}}\frac{K(x)}{\sigma(x)^2}\mathrm{d}x \tag{30}$$

is $O_p(1)$ in $L^2(0,1)$ because

$$\begin{aligned}
&\frac{1-u}{T\sqrt{u(1-u)}}\int_{X_0}^{X_{Tu}}\frac{(\partial_i\partial_j S(x,\tilde{\theta}_T)-\partial_i\partial_j S(x,\theta_*))}{\sigma(x)^2}\mathrm{d}x\\
&\quad\le\frac{1-u}{T\sqrt{u(1-u)}}\int_{X_0}^{X_{Tu}}\frac{|\partial_i\partial_j S(x,\tilde{\theta}_T)-\partial_i\partial_j S(x,\theta_*)|}{\sigma(x)^2}\mathrm{d}x\\
&\quad\le\frac{(1-u)\|\tilde{\theta}_T-\theta_*\|}{T\sqrt{u(1-u)}}\int_{X_0}^{X_{Tu}}\frac{K(x)}{\sigma(x)^2}\mathrm{d}x
\end{aligned}$$

holds for any $u\in(0,1)$ and for all $i,j\in\{1\ldots,d\}$, and $\hat{\theta}_T$ converges in probability to θ_*. Since the Itô formula yields

$$\begin{aligned}
&\int_{X_0}^{X_{Tu}}\frac{K(x)}{\sigma(x)^2}\mathrm{d}x\\
&\quad=\int_0^{Tu}\frac{K(X_s)}{\sigma(X_s)}\mathrm{d}W_s+\int_0^{Tu}\left(\frac{K(X_s)S(X_s,\theta_{(s)})}{\sigma(X_s)^2}+\frac{K'(X_s)}{2}-\frac{\sigma'(X_s)K(X_s)}{\sigma(X_s)}\right)\mathrm{d}s,
\end{aligned}$$

it suffices to prove that

$$\begin{aligned}
&\mathbb{E}\left[\int_0^1\left(\frac{(1-u)}{T\sqrt{u(1-u)}}\int_0^{Tu}\frac{K(X_s)}{\sigma(X_s)}\mathrm{d}W_s\right)^2\mathrm{d}u\right]\\
&\quad=\int_0^1\frac{(1-u)}{T^2u}\mathbb{E}\left[\int_0^{Tu}\left(\frac{K(X_s)}{\sigma(X_s)}\right)^2\mathrm{d}s\right]\mathrm{d}u
\end{aligned}\tag{31}$$

converges to zero and that

$$\begin{aligned}
&\sup_{T\in[0,\infty)}\mathbb{E}\left[\int_0^1 u(1-u)\left(\frac{1}{Tu}\int_0^{Tu}\left(\frac{K(X_s)S(X_s,\theta_{(s)})}{\sigma(X_s)^2}\right.\right.\right.\\
&\quad\left.\left.\left.+\frac{K'(X_s)}{2}-\frac{\sigma'(X_s)K(X_s)}{\sigma(X_s)}\right)\mathrm{d}s\right)^2\mathrm{d}u\right]
\end{aligned}\tag{32}$$

is finite. (31) is bounded above by

$$\frac{1}{T}\int_0^1(1-u)du\sup_{s\in[0,\infty)}\mathbb{E}\left[\left(\frac{K(X_s)}{\sigma(X_s)}\right)^2\right]$$

and it converges to zero. (32) is bounded above by

$$\int_0^1 u(1-u)du \sup_{s\in[0,\infty)} \mathbb{E}\left[\left(\frac{K(X_s)S(X_s,\theta_{(s)})}{\sigma(X_s)^2}+\frac{K'(X_s)}{2}-\frac{\sigma'(X_s)K(X_s)}{\sigma(X_s)}\right)^2\right]$$

which is finite. For the third term of (29), it suffices to prove that

$$\frac{u}{T\sqrt{u(1-u)}}\int_{X_{Tu}}^{X_T}\frac{K(x)}{\sigma(x)^2}\mathrm{d}x$$

is $O_p(1)$ in $L^2(0,1)$, which we can see in the same way we see that (30) is $O_p(1)$ in $L^2(0,1)$. For the second term in the right-hand side of (28), because of the Schwartz inequality and the Taylor expansion, it suffices to prove that

$$\begin{aligned}
&\frac{1}{T}\int_0^T\left\|\frac{\dot{S}\left(X_s,\hat{\theta}_T\right)-\dot{S}(X_s,\theta_*)}{\sigma(X_s)}\left(\frac{S(X_s,\theta_{(s)})}{\sigma(X_s)}-\sigma'(X_s)\right)\right.\\
&\quad\left.+\frac{\dot{S}'\left(X_s,\hat{\theta}_T\right)-\dot{S}'(X_s,\theta_*)}{2}\right\|^2\mathrm{d}s\\
&\le\frac{\left\|\hat{\theta}_T-\theta_*\right\|^2}{T}\int_0^T\left\|\frac{\ddot{S}(X_s,\tilde{\theta}_T)}{\sigma(X_s)}\left(\frac{S(X_s,\theta_{(s)})}{\sigma(X_s)}-\sigma'(X_s)\right)+\frac{\ddot{S}'(X_s,\check{\theta}_T)}{2}\right\|^2\mathrm{d}s
\end{aligned}$$

converges in probability to 0, where $\tilde{\theta}_T$ and $\check{\theta}_T$ lie between $\hat{\theta}_T$ and θ_*. It follows from $\hat{\theta}_T\to^p\theta_*$, ergodicity and conditions (I–III). This completes the proof of part (i).

(ii) Since it holds that

$$\begin{aligned}
&\mathbb{Z}_T(u,\hat{\theta}_T)-\mathbb{Z}_T(u,\theta_*)\\
&\quad=\frac{1}{\sqrt{T}}\int_0^T w_s^T(u)\frac{(\dot{S}(X_s,\theta)-\dot{S}(X_s,\theta_*))}{\sigma(X_s)^2}\left(\mathrm{d}X_s-S\left(X_s,\hat{\theta}_T\right)\mathrm{d}s\right)|_{\theta=\hat{\theta}_T}\\
&\qquad+\frac{1}{\sqrt{T}}\int_0^T w_s^T(u)\frac{\dot{S}(X_s,\theta_*)\left(S(X_s,\hat{\theta}_T\right)-S(X_s,\theta_*))}{\sigma(X_s)^2}\mathrm{d}s,
\end{aligned}$$

it suffices to confirm that

$$\left\|\frac{1}{T}\int_0^T w_s^T(\cdot)\frac{(\dot{S}(X_s,\theta)-\dot{S}(X_s,\theta_*))}{\sigma(X_s)^2}\left(\mathrm{d}X_s-S\left(X_s,\hat{\theta}_T\right)\mathrm{d}s\right)|_{\theta=\hat{\theta}_T}\right\|_{L^2(0,1)}^2 \tag{33}$$

and

$$\left\|\frac{1}{T}\int_0^T w_s^T(\cdot)\frac{\dot{S}(X_s,\theta_*)\left(S\left(X_s,\hat{\theta}_T\right)-S(X_s,\theta_*)\right)}{\sigma(X_s)^2}\mathrm{d}s\right\|_{L^2(0,1)}^2 \tag{34}$$

converge in probability to 0. The expression in (33) is bounded above by

$$2\left\|\frac{1}{T}\int_0^T w_s^T(\cdot)\frac{(\dot{S}(X_s,\theta)-\dot{S}(X_s,\theta_*))}{\sigma(X_s)}\mathrm{d}W_s|_{\theta=\hat{\theta}_T}\right\|_{L^2(0,1)}^2$$
$$+2\left\|\frac{1}{T}\int_0^T w_s^T(\cdot)\frac{\left(\dot{S}\left(X_s,\hat{\theta}_T\right)-\dot{S}(X_s,\theta_*)\right)}{\sigma(X_s)^2}\left(S\left(X_s,\theta_{(s)}\right)-S(X_s,\hat{\theta}_T)\right)\mathrm{d}s\right\|_{L^2(0,1)}^2.$$

The convergence in probability to 0 of the first term is due to Lemma 5 (i). As for the second term, by a Taylor expansion, it is enough to see that

$$\frac{1}{T}\int_0^T\frac{(\sup_{\theta\in\Theta}|\partial_i\partial_j S(X_s,\theta)|)^2}{\sigma(X_s)^4}\left(\sup_{\theta\in\Theta}|S(X_s,\theta)|\right)^2\mathrm{d}s=O_p(1)\tag{35}$$

for all i and j, where $\tilde{\theta}_T$ lies between $\hat{\theta}_T$ and θ_*, since it holds that

$$\begin{aligned}
&\left(\hat{\theta}_T-\theta_*\right)^\top\frac{1}{T}\int_0^T w_s^T(u)\frac{\partial_i\dot{S}(X_s,\tilde{\theta}_T)}{\sigma(X_s)^2}(S(X_s,\theta_{(s)})-S(X_s,\theta_*))\mathrm{d}s\\
&\leq\|\hat{\theta}_T-\theta_*\|\left(\frac{2}{T}\int_0^T\frac{\sum_{j=1}^d(\partial_i\partial_j S(X_s,\tilde{\theta}_T))^2}{\sigma(X_s)^4}((S(X_s,\theta_{(s)}))^2+(S(X_s,\theta_*))^2)\mathrm{d}s\right)^{1/2}\\
&\leq\|\hat{\theta}_T-\theta_*\|\left(\frac{4\sum_{j=1}^d}{T}\int_0^T\frac{(\sup_{\theta\in\Theta}|\partial_i\partial_j S(X_s,\theta)|)^2}{\sigma(X_s)^4}(\sup_{\theta\in\Theta}|S(X_s,\theta)|)^2\mathrm{d}s\right)^{1/2}
\end{aligned}$$

and $\hat{\theta}_T\to^p\theta_*$. Equation (35) follows from ergodicity and conditions (I–III). The convergence in probability of (34) to 0 also follows from the same argument using a Taylor expansion and ergodicity. This completes the proof of part (ii).

(iii) The result follows from Propositions 2 and 4. This completes the whole proof. □

Proof of Theorem 5 (ii). In general, when M is a $d\times d$ non-negative definite matrix, it holds that

$$\begin{aligned}
2(v^\top M^{-1}v+w^\top M^{-1}w)&=(v+w)^\top M^{-1}(v+w)+(v-w)^\top M^{-1}(v-w)\\
&\geq(v-w)^\top M^{-1}(v-w)
\end{aligned}$$

for any d-dimensional vectors v and w. Since

$$\mathbb{Z}_T(u,\hat{\theta}_T)=\mathbb{Z}_T^p(u,\theta_*)+\mathbb{M}_T(u,\theta_*)+(\mathbb{Z}_T(u,\hat{\theta}_T)-\mathbb{Z}_T(u,\theta_*)),$$

the stated inequality yields

$$
\begin{aligned}
AD_T &= \int_0^1 \mathbb{Z}_T(u,\hat{\theta}_T)^\top \hat{C}_T \mathbb{Z}_T(u,\hat{\theta}_T)\mathrm{d}u \\
&\geq \frac{1}{4}\int_0^1 \mathbb{Z}_T^p(u,\theta_*)^\top \hat{C}_T \mathbb{Z}_T^p(u,\theta_*)\mathrm{d}u - \frac{1}{2}\int_0^1 \mathbb{M}_T(u,\theta_*)^\top \hat{C}_T \mathbb{M}_T(u,\theta_*)\mathrm{d}u \\
&\quad - \int_0^1 (\mathbb{Z}_T(u,\hat{\theta}_T) - \mathbb{Z}_T(u,\theta_*))^\top \hat{C}_T (\mathbb{Z}_T(u,\hat{\theta}_T) - \mathbb{Z}_T(u,\theta_*))\mathrm{d}u. \qquad (36)
\end{aligned}
$$

Define

$$
A_T(u) = \frac{1}{T}\int_0^T (1\{s \leq Tu\} - u)\frac{\dot{S}(X_s,\theta_*)}{\sigma(X_s)^2}(S(X_s,\theta_{(s)}) - S(X_s,\theta_*))\mathrm{d}s,
$$

and then

$$
\mathbb{Z}_T^p(u,\theta_*) = \left(\frac{T}{u(1-u)}\right)^{1/2} A_T(u) \geq T^{1/2} A_T(u).
$$

This shows that the first term of (36) is bounded below by

$$
\frac{T}{4}\int_0^1 A_T(u,\theta_*)^\top \hat{C}_T A_T(u,\theta_*)du. \qquad (37)
$$

For $u \leq u_*$, it follows from

$$
\begin{aligned}
A_T(u) = \frac{1-u}{T}&\int_0^{Tu} \frac{\dot{S}(X_s,\theta_*)}{\sigma(X_2)^2}(S(X_s,\theta_0) - S(X_s,\theta_*))\mathrm{d}s \\
-\frac{u}{T}&\int_{Tu}^{Tu_*} \frac{\dot{S}(X_s,\theta_*)}{\sigma(X_2)^2}(S(X_s,\theta_0) - S(X_s,\theta_*))\mathrm{d}s \\
-\frac{u}{T}&\int_{Tu_*}^{T} \frac{\dot{S}(X_s,\theta_*)}{\sigma(X_2)^2}(S(X_s,\theta_1) - S(X_s,\theta_*))\mathrm{d}s,
\end{aligned}
$$

that

$$
\begin{aligned}
\underset{T\to\infty}{\text{l.i.m.}}\, A_T(u) &= (u(1-u) - u(u_* - u))\Psi(\theta_*,\theta_0) - u(1-u_*)\Psi(\theta_*,\theta_1) \\
&= u(1-u_*)(\Psi(\theta_*,\theta_0) - \Psi(\theta_*,\theta_1)).
\end{aligned}
$$

For $u > u_*$, $\text{l.i.m.}_{T\to\infty} A_T(u) = u_*(1-u)(\Psi(\theta_*,\theta_0) - \Psi(\theta_*,\theta_1))$ for the same reason. Let us denote $\text{l.i.m.}_{T\to\infty} A_T(u)$ by $A_\infty(u)$ for all $u \in (0,1)$. Now, we prove that

$$
\mathbb{E}\left[\|A_T - A_\infty\|_{L^2(0,1)}^2\right] \to 0. \qquad (38)
$$

It holds that, for $\forall u \in (0,1)$,

$$\mathbb{E}\left[\|A_T(u) - A_\infty(u)\|^2\right] \le 2\mathbb{E}\left[\|A_T(u)\|^2\right] + 2\|A_\infty(u)\|^2$$

and the first term in the right-hand side is bounded above by

$$\begin{aligned}
&2\mathbb{E}\left[\frac{1}{T^2}\int_0^T (1\{s \le Tu\} - u)^2 \mathrm{d}s \int_0^T \left\|\frac{\dot{S}(X_s,\theta_*)}{\sigma(X_s)^2}(S(X_s,\theta_{(s)}) - S(X_s,\theta_*))\right\|^2 \mathrm{d}s\right]\\
&\le 4 \sup_{s\in[0,\infty)} \mathbb{E}\left[\frac{\|\dot{S}(X_s,\theta_*)\|^2}{\sigma(X_s)^4}\left((S(X_s,\theta_{(s)}))^2 + (S(X_s,\theta_*))^2\right)\right]\\
&\le 4 \sup_{s\in[0,\infty)} \mathbb{E}\left[\frac{\|\dot{S}(X_s,\theta_*)\|^2}{\sigma(X_s)^4}\left((S(X_s,\theta_0))^2 + (S(X_s,\theta_1))^2 + (S(X_s,\theta_*))^2\right)\right]
\end{aligned}$$

because of the Schwartz inequality and the bound is finite because of Proposition 4 (ii). Since the left-hand side of (38) is equal to

$$\int_0^1 \mathbb{E}\left[(A_T(u) - A_\infty(u))^2\right] \mathrm{d}u$$

and $(A_\infty(u))^2$ is integrable with respect to u, the dominated convergence theorem yields (38), and (38) gives $A_T \to^p A_\infty$ in $L^2(0,1)$. This result, the Slutsky theorem and the continuous mapping theorem yield

$$\int_0^1 A_T^\top(u)\hat{C}_T^{-1}A_T(u)\mathrm{d}u \to^p \int_0^1 A_\infty^\top(u)C_*^{-1}A_\infty(u)\mathrm{d}u,$$

where $C_* := u_* C_{\theta_0}(\theta_*,\theta_*) + (1-u_*)C_{\theta_1}(\theta_*,\theta_*)$, which is the limit in probability of $\hat{C}_T$. By simple calculations, the right-hand side of the limit is equal to

$$\frac{u_*^2(1-u_*)^2}{3}(\Psi(\theta_*,\theta_0) - \Psi(\theta_*,\theta_1))^\top C_*^{-1}(\Psi(\theta_*,\theta_0) - \Psi(\theta_*,\theta_1)).$$

Moreover, let us show that

$$\Psi(\theta_*,\theta_0) - \Psi(\theta_*,\theta_1) \ne 0\,. \tag{39}$$

Now $u_*\Psi(\theta_*,\theta_0) + (1-u_*)\Psi(\theta_*,\theta_1) = 0$ because of Lemma 1. If $\Psi(\theta_*,\theta_0) - \Psi(\theta_*,\theta_1)$ were zero, then $\Psi(\theta_*,\theta_1)$ and $\Psi(\theta_*,\theta_0)$ would be zero, but this contradicts condition (IV) and the assumption $\theta_0 \ne \theta_1$. Thus, (39) is valid. Hence, (37) and thus the first term of (36) converge to positive infinity and the convergence is faster than that of the third term of (36), which is $o_p(T)$ because of Lemma 4 (ii). Note that the second term of (36) is $O_p(1)$ because of Lemma 4 (iii). Therefore, it follows that AD_T converges to positive infinity in probability, so the test is consistent. This completes the proof. □

Acknowledgements The author thanks to Professor Shuhei Mano for comments on proofs and Professor Yoichi Nishiyama for many suggestions. The author also thanks to the referees and the associate editor for their feedbacks which improved this paper.

References

Brodsky, B. E., Darkhovsky, B. S. (2000). *Non-parametric statistical diagnosis. problems and methods. mathematics and its applications* (Vol. 509). Dordrecht: Kluwer Academic Publishers.

Chen, J., Gupta, A. K. (2012). *Parametric statistical change point analysis. With applications to genetics, medicine, and finance* (2nd ed.). New York: Springer.

Csörgő, M., Horváth, L. (1997). *Limit theorems in change-point analysis. Wiley series in probability and statistics*. Chichester: Wiley.

De Gregorio, A., Iacus, S. M. (2008). Least squares volatility change point estimation for partially observed diffusion processes. *Communications in Statistics-Theory and Methods*, *37*(15), 2342–2357.

Dehling, H., Franke, B., Kott, T., Kulperger, R. (2014). Change point testing for the drift parameters of a periodic mean reversion process. *Statistical Inference for Stochastic Processes*, *17*(1), 1–18.

Horváth, L. (1993). The maximum likelihood method for testing changes in the parameters of normal observations. *The Annals of Statistics*, *21*(2), 671–680.

Horváth, L., Parzen, E. (1994). Limit theorems for fisher-score change processes. In: E. Carlstein, H.-G. Müller, D. Siegmund (Eds.) *Change-point problems*, IMS Lecture Notes—Monograph Series, 23, 157–169.

Khmaladze, E. V. (1979). The use of ω^2 tests for testing parametric hypothesis. *Theory of Probability and Its Applications*, *24*(2), 283–301.

Kutoyants, Y. A. (2004). *Statistical inference for ergodic diffusion processes. Springer series in statistics*. London: Springer.

Lánska, V. (1979). Minimum contrast estimation in diffusion processes. *Journal of Applied Probability*, *16*(1), 65–75.

LaRiccia, V., Mason, D. M. (1986). Cramér-von Mises statistics based on the sample quantile function and estimated parameters. *Journal of Multivariate Analysis*, *18*(1), 93–106.

Lee, S., Nishiyama, Y., Yoshida, N. (2006). Test for parameter change in diffusion processes by cusum statistics based on one-step estimators. *Annals of the Institute of Statistical Mathematics*, *58*(2), 211–222.

Mason, D. M. (1984). Weak convergence of the weighted empirical quantile process in $L^2(0, 1)$. *The Annals of Probability*, *12*(1), 243–255.

Mihalache, S. (2012). Strong approximations and sequential change-point analysis for diffusion processes. *Statistics & Probability Letters*, *82*(3), 464–472.

Negri, I., Nishiyama, Y. (2012). Asymptotically distribution free test for parameter change in a diffusion process model. *Annals of the Institute of Statistical Mathematics*, *64*(5), 911–918.

Negri, I., Nishiyama, Y. (2014). Z-process method for change point problems. Quaderni del Dipartimento di Ingegneria dell'informazione e metodi matematici. Serie "Matematica e Statistica" n. 5/MS- 2014. Dalmine: Universita degli studi di Bergamo. Retrieved from http://hdl.handle.net/10446/30761

Nishiyama, Y. (2000). Entropy methods for martingales. CWI Tract, 128, Centrum voor Wiskunde en Informatica, Amsterdam.

Nishiyama, Y. (2011). *Statistical analysis by the theory of martingales*. (In Japanese). ISM Series, 1, Kindaikagakusha, Tokyo.

Prokhorov, Y. V. (1956). Convergence of random processes and limit theorems in probability. *Theory of Probability and Its Applications*, *1*(2), 157–214.

Song, J., Lee, S. (2009). Test for parameter change in discretely observed diffusion processes. *Statistical Inference for Stochastic Processes*, *12*(2), 165–183.

Suquet, Ch., Viano, M.-C. (1998). Change point detection in dependent sequences: invariance principles for some quadratic statistics. *Mathematical Methods of Statistics*, *7*, 157–191.

Tsukuda, K., Nishiyama, Y. (2014). On L^2 space approach to change point problems. *Journal of Statistical Planning and Inference*, *149*, 46–59.

van der Vaart, A. W. (1998). *Asymptotic statistics*. Cambridge: Cambridge University Press.

van der Vaart, A. W., Wellner, J. A. (1996). *Weak convergence and empirical processes: with applications to statistics*. New York: Springer.

van Zanten, H. (2003). On uniform laws of large numbers for ergodic diffusions and consistency of estimators. *Statistical Inferences for Stochastic Processes*, *6*, 199–213.