

Local asymptotic mixed normality of
approximate maximum likelihood estimator of
drift parameters in diffusion model*

Snježana Lubura Strunjak* and Miljenko Huzak*

*University of Zagreb

September 5, 2017

*This work has been fully supported by Croatian Science Foundation under the project
3526

Abstract

We assume that the diffusion X satisfies a stochastic differential equation of the form: $dX_t = \mu(X_t, \theta)dt + \sigma_0\nu(X_t)dW_t$, with unknown drift parameter θ and known diffusion coefficient parameter σ_0 . We prove that approximate maximum likelihood estimator of drift parameter $\bar{\theta}_n$ obtained from discrete observations $(X_{i\Delta_n}, 0 \leq i \leq n)$ along fixed time interval $[0, T]$, and when $\Delta_n = \frac{T}{n}$ tends to zero, is locally asymptotic mixed normal, with covariance matrix which depends on MLE $\hat{\theta}$ obtained from continuous observations $(X_t, 0 \leq t \leq T)$ along fixed time interval $[0, T]$, and on path $(X_t, 0 \leq t \leq T)$.

Keywords: asymptotic mixed normality, diffusion processes, discrete observation, parameter estimation

1 Introduction

Let $W = (W_t, t \geq 0)$ be a one-dimensional standard Brownian motion defined on filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. We suppose that filtration $(\mathcal{F}_t)_{t \geq 0}$ is the smallest one such that satisfies the usual conditions and to which W is adapted. Let $X = (X_t, t \geq 0)$ be a one-dimensional diffusion which satisfies Itô's stochastic differential equation (SDE) of the form (for details see Revuz, D. and Yor, M. (1999))

$$dX_t = \mu(X_t, \theta)dt + \sigma_0\nu(X_t)dW_t, \quad X_0 = x_0, \quad t \geq 0, \quad (1.1)$$

where ν and μ are real functions and x_0 is a given deterministic initial value of X . For example, if both functions ν and $\mu(\cdot, \theta)$ are locally Lipschitz, then diffusion given by (1.1) has strong solution, see (Rogers, L. C. G. and Williams, D., 1987, Theorem V.12.1).

Let θ_0 and σ_0 be true values of drift parameter and diffusion coefficient parameter respectively. We assume that $\sigma_0 > 0$ is known. Let $T > 0$ be a fixed real number and $0 =: t_0 < t_1 < \dots < t_n := T, n \in \mathbb{N}$ be a deterministic subdivision of segment $[0, T]$ such that $t_i = i\Delta_n, i = 0, \dots, n, \Delta_n = \frac{T}{n}$. Given a discrete observation $(X_{t_i}, 0 \leq i \leq n)$ of the trajectory $(X_t, t \in [0, T])$, the problem is to estimate the unknown drift parameter θ of X . We assume that θ belongs to the parameter space Θ which is a relatively compact, open and convex set in Euclidean space \mathbb{R}^d .

The maximum likelihood estimator (MLE) of drift parameters based on continuous observation $(X_t, t \in [0, T])$ has good properties, such as consistency and asymptotic efficiency when T goes to infinity in case of ergodic diffusions (see Brown, B. M. and Hewitt, J. I. (1975), and Liptser, R. S. and Shirayayev, A. N. (1977)). The same holds for MLE based on discrete observations (see Dacunha-Castelle, D. and Florens-Zmirou, D. (1986)) but in most cases it can not be explicitly calculated. Hence, other methods of estimations have to be considered. In this paper we use method based on a Gaussian approximation of the transition density (see Kessler, M. (1997)) to construct contrast function used for getting approximate maximum likelihood estimator (AMLE) of drift parameters based on discrete observations. We consider the simplest case based on approximation of diffusion X by a solution to difference equation obtained from (1.1) by Euler approximation of Riemann and Itô integrals (see Huzak, M. (2001), LeBreton, A. (1976)). In such a way obtained AMLE is consistent and

asymptotically normal in case of ergodic diffusions when $n\Delta_n$ goes to infinity ($T \rightarrow +\infty$) such that $n\Delta_n^3$ goes to zero (see Florens-Zmirou, D. (1989)). In case when T is fixed and Δ_n goes to zero it was proved that measurable AMLE exists, and it converges to MLE based on continuous observations over $[0, T]$ with rate $\sqrt{\Delta_n}$ in probability (see LeBreton, A. (1976) for linear case, and Huzak, M. (2017) for general case).

The existence, measurability and some asymptotic properties of the AMLE are proved in LeBreton, A. (1976), under assumption that observations can be taken only up to some maximal observational time T , when Δ_n tends to zero and under the assumption that the diffusion satisfies a linear SDE with completely determined and constant diffusion coefficient function. The similar results are proved in Huzak, M. (2017) in more general case (without assumption on linearity of SDE).

In this paper we prove that the difference between AMLE and MLE based on continuous observations over fixed observational time interval $[0, T]$, normed with $\sqrt{\Delta_n}$ is asymptotically mixed normal when Δ_n goes to zero, and that covariance matrix of the limit random vector depends on MLE and the continuous trajectory. This result can be applied for estimations of AMLE's standard errors, and hence asymptotic confidence intervals for drift parameters in ergodic diffusions case, which include a correction due to discretization, for example, in a sense of Example 3 from Sect. 6.

The paper is organized in the following way. In the next section we introduce notation and definitions we need. The main results are presented in Sect. 4, followed by proofs in Sect. 5. In the last section we will present simulation study throughout two examples.

2 Preliminaries

We denote by $\langle \cdot, \cdot \rangle$ the scalar product in the d -dimensional Euclidean space \mathbb{R}^d and by $\| \cdot \|$ the induced norm. The drift parameter space $\Theta \subseteq \mathbb{R}^d$ is relatively compact, which means that its closure $Cl(\Theta)$ is a compact set in \mathbb{R}^d . We will denote by E an open interval such that $E \subseteq \mathbb{R}$. We will say that U is relatively compact in E if $U \subseteq E$ and $Cl(U)$ is compact set in E . If $(x, \theta) \mapsto f(x, \theta)$ is a real function then we will denote by $D_\theta^m f$ the m -th partial derivatives with respect to θ of the function f , $m \in \mathbb{N}$, provided that they exist. If $\theta \mapsto f(\theta)$ is a real-valued function defined on an open subset of \mathbb{R}^d , then we will denote by $Df(\theta), D^2f(\theta)$ its first and second derivatives with respect to θ , provided that they exist. Let K be a relatively compact set in \mathbb{R}^d . We say that a partial derivative $D_\theta^m f$ of a real function $f : E \times Cl(K) \rightarrow \mathbb{R}$ exists on $E \times Cl(K)$ if there exists an open set $U^K \subseteq \mathbb{R}^d$, such that $E \times Cl(K) \subseteq E \times U^K$, and an extension \tilde{f} of f , defined on $E \times U^K$, such that $D_\theta^m \tilde{f}$ exists.

For fixed $T > 0$ let $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ be a given filtered probability space. Let $(\tilde{\Omega}, \tilde{\mathcal{F}}_T, \tilde{\mathbb{P}} = (\tilde{\mathcal{F}}_t)_{0 \leq t \leq T}, \tilde{\mathbb{P}})$ be an *extension* of $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ (for details see Jacod, J. (1997)). The extension is called *very good* if all martingales on the space $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ are also martingales on $(\tilde{\Omega}, \tilde{\mathcal{F}}_T, \tilde{\mathbb{P}} = (\tilde{\mathcal{F}}_t)_{0 \leq t \leq T}, \tilde{\mathbb{P}})$. Let A be some Polish space. In our case A will be the Skorokhod space $A = D([0, T], \mathbb{R}^d)$ or d -dimensional Euclidean space $A = \mathbb{R}^d$. Let (Z_n) be a sequence of A -valued random vectors, all defined on $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$, and let Z be an A -valued random vector defined on the extension $(\tilde{\Omega}, \tilde{\mathcal{F}}_T, \tilde{\mathbb{P}} =$

$(\tilde{\mathcal{F}}_t)_{0 \leq t \leq T, \tilde{\mathbb{P}}}$. We will say that (Z_n) converges stably in law to Z , and write $Z_n \xrightarrow{st} Z$, if

$$\lim_{n \rightarrow \infty} \mathbb{E}[Y f(Z_n)] = \tilde{\mathbb{E}}[Y f(Z)],$$

for all bounded continuous functions $f: A \rightarrow \mathbb{R}$ and all bounded random variable Y on $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$. This kind of convergence is introduced by Rényi (see Rényi, A. (1963)) and studied by Aldous and Eagleson (see Aldous, D. J. and Eagleson, G. K. (1978)). Nice properties of this kind of convergence can also be found in Jacod, J. and Protter, P. (2012) and Jacod, J. (1997).

We will say that an \mathbb{R}^d -valued random vector Y has *mixed normal distribution* with \mathcal{F}_T -measurable random covariance matrix $C = (C^{jk})_{j,k=1,\dots,d}$, and we write $Y \sim MN(0, C)$ if $\mathbb{E}[e^{i\langle t, Y \rangle} | \mathcal{F}_T] = e^{-\frac{1}{2} \sum_{j,k=1,\dots,d} t_j t_k C^{jk}}$ (see Aldous, D. J. and Eagleson, G. K. (1978)). If $Y \sim MN(0, C)$, then Y has the same distribution as $\sqrt{C}Z$, where \sqrt{C} is square symmetric root of C and $Z \sim N(0, I)$ is standard normal random vector independent of \mathcal{F}_T . An \mathbb{R}^d -valued process $(Y_t)_{0 \leq t \leq T}$ is *centered Gaussian process* if for all $0 \leq s_1 < s_2 < \dots < s_k \leq T$ random matrix $(Y_{s_1}, \dots, Y_{s_k}) \in \mathbb{R}^{dk}$ has multivariate normal distribution and $\mathbb{E}[Y_t] = 0$, $t \in [0, T]$.

For $n \in \mathbb{N}$, let $t_i := i\Delta_n, i = 0, \dots, n$ be an equidistant subdivision of segment $[0, T]$, $\Delta_n = \frac{T}{n}$. Let $A_n^t := \max\{j: t_j \leq t\}$ and $\mathcal{F}_{n,i} := \mathcal{F}_{t_i}, i = 0, \dots, n, n \in \mathbb{N}$. For proving local mixed normality of the AMLE we use the following theorem, which is a version of Theorem 3-2 from Jacod, J. (1997), but with notation adjusted for our article.

Theorem 2.1. *Let W be a one-dimensional Brownian motion on $(\Omega, \mathcal{F}_T, \mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$, and let χ_i^n be \mathcal{F}_{t_i} -measurable square integrable \mathbb{R}^d -valued random vectors. Let $C = (C^{jk})$ be continuous adapted process defined on $(\Omega, \mathcal{F}_T, \mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ such that C_t is positive semidefinite symmetric $d \times d$ matrix for all $t \in [0, T]$. Assume that next conditions hold:*

$$\sup_{0 \leq t \leq T} \left\| \sum_{i=1}^{A_n^t} \mathbb{E}[\chi_i^n | \mathcal{F}_{n,i-1}] \right\| \xrightarrow{\mathbb{P}} 0, \quad (2.1)$$

$$\sum_{i=1}^{A_n^t} (\mathbb{E}[\chi_i^{n,j} \chi_i^{n,k} | \mathcal{F}_{n,i-1}] - \mathbb{E}[\chi_i^{n,j} | \mathcal{F}_{n,i-1}] \mathbb{E}[\chi_i^{n,k} | \mathcal{F}_{n,i-1}]) \xrightarrow{\mathbb{P}} C_t^{j,k},$$

$$\forall t \in [0, T], j, k = 1, \dots, d, \quad (2.2)$$

$$\sum_{i=1}^{A_n^t} \mathbb{E}[\chi_i^n (W_{t_i} - W_{t_{i-1}}) | \mathcal{F}_{n,i-1}] \xrightarrow{\mathbb{P}} 0, \quad \forall t \in [0, T], \quad (2.3)$$

$$\sum_{i=1}^n \mathbb{E}[\|\chi_i^n\|^2 1_{\{\|\chi_i^n\| > \epsilon\}} | \mathcal{F}_{n,i-1}] \xrightarrow{\mathbb{P}} 0, \quad \forall \epsilon > 0, \quad (2.4)$$

$$\sum_{i=1}^{A_n^t} \mathbb{E}[\chi_i^n (N_{t_i} - N_{t_{i-1}}) | \mathcal{F}_{n,i-1}] \xrightarrow{\mathbb{P}} 0, \quad \forall t \in [0, T], \quad (2.5)$$

where N is bounded \mathcal{F}_t -martingale orthogonal to W .

Then we have

$$\sum_{i=1}^{A_n^t} \chi_i^n \xrightarrow{st} Y \quad \text{on } D([0, T], \mathbb{R}^d),$$

where Y is a continuous process defined on a very good filtered extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}} = (\tilde{\mathcal{F}}_t)_{0 \leq t \leq T}, \tilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}_T, \mathbb{P} = (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ and which, conditionally on σ -field \mathcal{F}_T , is centered Gaussian \mathbb{R}^d -valued process with independent increments satisfying $\tilde{\mathbb{E}}[Y_t^j Y_t^k | \mathcal{F}_T] = C_t^{jk}, t \in [0, T], j, k = 1, \dots, d$.

Let $(\gamma_n, n \in \mathbb{N})$ be a sequence of positive numbers and let $Y = (Y_n, n \in \mathbb{N})$ be a sequence of random vectors defined on some probability space. We will say that Y is of order $O_{\mathbb{P}}(\gamma_n), n \in \mathbb{N}$, and we will write $Y_n = O_{\mathbb{P}}(\gamma_n)$, if $(\frac{Y_n}{\gamma_n}, n \in \mathbb{N})$ is bounded in probability, which means that

$$\lim_{A \rightarrow +\infty} \overline{\lim}_n \mathbb{P}\{\gamma_n^{-1} \|Y_n\| > A\} = 0.$$

We will denote $L^1(\mathbb{P})$ -norm by $\|\cdot\|_1 := \mathbb{E}[|\cdot|]$ and $L^2(\mathbb{P})$ -norm by $\|\cdot\|_2 := \sqrt{\mathbb{E}[(\cdot)^2]}$.

3 Estimation method

Let us now introduce our AMLE of d -dimensional parameter $\theta \in \Theta$. Let X be a strong solution of SDE (1.1) for true parameter value $\theta_0 \in \Theta$ and given Brownian motion W . If we discretize the SDE over interval $[t_{i-1}, t_i]$ by using the Euler discretization scheme, we get stochastic difference equation of the form:

$$Z_{t_i} - Z_{t_{i-1}} = \mu(Z_{t_{i-1}}, \theta)(t_i - t_{i-1}) + \sigma_0 \nu(Z_{t_{i-1}})(W_{t_i} - W_{t_{i-1}}), \quad Z_0 = x_0, \quad (3.1)$$

for $1 \leq i \leq n$. If it exists, the solution of (3.1) is a discrete-time process $Z = (Z_{t_0}, Z_{t_1}, \dots, Z_{t_n})$ which is an approximation of X on $[0, T]$. The log likelihood function, shortly LLF, of the process Z , up to the constant not depending on the parameter θ is

$$-\frac{n}{2} \ln(\sigma_0^2) - \frac{1}{2} \sum_{i=1}^n \frac{(z_i - z_{i-1} - \mu(z_{i-1}, \theta)(t_i - t_{i-1}))^2}{\sigma_0^2 \nu^2(z_{i-1})(t_i - t_{i-1})}. \quad (3.2)$$

The contrast function is obtained from (3.2) by substituting vector (z_0, z_1, \dots, z_n) with the discrete observations $(X_{t_0}, X_{t_1}, \dots, X_{t_n})$ of the process X and neglecting the part which does not depend on parameter θ :

$$L_n(\theta) = \sum_{i=1}^n \left(\frac{(X_{t_i} - X_{t_{i-1}}) \mu(X_{t_{i-1}}, \theta)}{\sigma_0^2 \nu^2(X_{t_{i-1}})} - \frac{1}{2} \frac{\mu^2(X_{t_{i-1}}, \theta)(t_i - t_{i-1})}{\sigma_0^2 \nu^2(X_{t_{i-1}})} \right). \quad (3.3)$$

Here we assumed that $\nu \neq 0$ on the state space of X . A point of global maximum $\bar{\theta}_n$ of the function (3.3), if it exists, is an estimator of the parameter θ which we call approximate maximum likelihood estimator, shortly AMLE. Further in the text, if it is written point of maximum, it means point of global maximum.

4 Main results

We assume that drift parameter space Θ is a relatively compact, open and convex set in $\mathbb{R}^d, d \geq 1$, and state space E is an open interval in \mathbb{R} .

Let the following assumptions be satisfied:

- (A1) For all $\theta \in \Theta$, there exists strong solution X of SDE (1.1) on time interval $[0, +\infty)$. This solution X has continuous paths with values in E and if X' is any other solution of SDE (1.1), with the same Brownian motion W , then the law of X' is identical to the law of X .
- (A2) For all $\theta \in Cl(\Theta)$, function $\mu(\cdot, \theta): E \rightarrow \mathbb{R}$ is continuously differentiable on E . Functions $(x, \theta) \mapsto \mu(x, \theta)$, $(x, \theta) \mapsto \frac{\partial^3}{\partial x^2 \partial \theta_j} \mu(x, \theta)$, $(x, \theta) \mapsto \frac{\partial^2}{\partial x^2} \mu(x, \theta)$, $(x, \theta) \mapsto \frac{\partial}{\partial x} \mu(x, \theta)$, $(x, \theta) \mapsto \frac{\partial}{\partial \theta_i} \frac{\partial}{\partial x} \frac{\partial}{\partial \theta_j} \mu(x, \theta)$ are continuous on $E \times Cl(\Theta)$, for all $i, j = 1, \dots, d$. Function $x \mapsto \nu(x)$ is two-times continuously differentiable on E , and $\nu(x) > 0$ for all $x \in E$.
- (A3) For all $m \leq d+3$, there exist partial derivatives $D_\theta^m \mu(x, \theta)$ and $\frac{\partial}{\partial x} D_\theta^m \mu(x, \theta)$ on $E \times Cl(\Theta)$. Furthermore, functions $(x, \theta) \mapsto D_\theta^m \mu(x, \theta)$, $(x, \theta) \mapsto \frac{\partial}{\partial x} D_\theta^m \mu(x, \theta)$, $m \leq d+3$, are continuous on $E \times Cl(\Theta)$.

Let

$$L_T(\theta) = \int_0^T \frac{\mu(X_s, \theta)}{\sigma_0^2 \nu^2(X_s)} dX_s - \frac{1}{2} \int_0^T \frac{\mu^2(X_s, \theta)}{\sigma_0^2 \nu^2(X_s)} ds \quad (4.1)$$

be a continuous-time log-likelihood function, shortly LLF (see Feigin, P. D. (1976)).

- (A4) For all $\omega \in \Omega$, LLF $\theta \mapsto L_T(\theta) \equiv L_T(\omega, \theta)$ has a unique point of global maximum $\hat{\theta}(\omega) \in \Theta$, and $D^2 L_T(\hat{\theta}) < 0$ which means that Hessian $D^2 L_T(\hat{\theta})$ is a negatively definite matrix.

For example, the general growth diffusion process (see Huzak, M. (1998)) satisfies assumptions (A1-3) obviously, and (A4) on an event which probability tends to unity. Generally, these assumptions (and in the same sense) satisfies any ergodic diffusion such that (H1B-5B) from Huzak, M. (2017) hold. More precisely, these diffusions have a property that a.s. there exists a time $\tau > 0$ such that for all times $T \geq \tau$, (A4) holds. Diffusions that satisfy (A4) with certainty are those that satisfy (A1-3) and have drift functions that are linear combinations of linearly independent functions defined on state space with coefficients equal to drift parameters such that the values of (vector) parameters are in $\Theta = \mathbb{R}^d$ (see Brown, B. M. and Hewitt, J. I. (1975)). For such diffusions we will say briefly that are linear in drift parameters. To prove that the main results of this paper hold for diffusions linear in drift parameters we do not need the assumption of relatively compactness of drift parameter space (see Remark 5.2).

For each $\theta \in \Theta$ let $\Sigma(\theta)$ be $d \times d$ random matrix which jk component is defined by

$$\Sigma(\theta)^{jk} = \frac{1}{2} \int_0^T \nu^4(X_s) \frac{\partial}{\partial x} \frac{\partial}{\partial \theta_j} \mu(X_s, \theta) \frac{\partial}{\partial x} \frac{\partial}{\partial \theta_k} \mu(X_s, \theta) \frac{ds}{\nu^2(X_s)}.$$

Under conditions (A1)-(A4) in Huzak, M. (2017) is proved that there exists a sequence $(\hat{\theta}_n, n \in \mathbb{N}) \subseteq \Theta$ of \mathcal{F}_T -measurable random vectors such that the

following hold:

$$\lim_n \mathbb{P}\{DL_n(\hat{\theta}_n) = 0\} = 1, \quad (4.2)$$

$$\hat{\theta}_n \xrightarrow{\mathbb{P}} \hat{\theta}, n \rightarrow \infty, \quad (4.3)$$

if $(\tilde{\theta}_n, n \in \mathbb{N})$ is any other sequence of random vectors which satisfies

$$(4.2) \text{ and } (4.3), \text{ then } \lim_n \mathbb{P}\{\tilde{\theta}_n = \hat{\theta}_n\} = 1, \quad (4.4)$$

$$\left(\frac{1}{\sqrt{\Delta_n}} \|\hat{\theta}_n - \hat{\theta}\|, n \in \mathbb{N}\right) \text{ is bounded in probability.} \quad (4.5)$$

(4.2)-(4.5) are consequences of

$$\sup_{\theta \in \Theta} \|D^r L_n(\theta) - D^r L_T(\theta)\| = O_{\mathbb{P}}(\sqrt{\Delta_n}), \quad n \in \mathbb{N}, r = 1, 2, \quad (4.6)$$

and so called the general theorem on approximate maximum likelihood estimation (see Huzak, M. (2001)). In the proof of this theorem $\hat{\theta}_n$ is constructed to be a point of global maximum of L_n in Θ on an event with probability tending to unity, and hence we can take $\bar{\theta}_n = \hat{\theta}_n$ on this event. It is sufficient to assume that $(\bar{\theta}_n, n \in \mathbb{N})$ satisfies (4.2)-(4.5), and all of our results are stated for that kind of sequence.

Theorem 4.1. *Assume that (A1)-(A4) hold. Assume that $(\bar{\theta}_n, n \in \mathbb{N})$ satisfies (4.2)-(4.5). Then*

$$\frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}) \xrightarrow{st} MN(0, (D^2 L_T(\hat{\theta}))^{-1} \Sigma(\hat{\theta}) (D^2 L_T(\hat{\theta}))^{-1}).$$

For real symmetric $d \times d$ matrix C let us denote by $\lambda_k(C)$ its k -th in order eigenvalue in the ordered sequence of eigenvalues $\lambda_1(C) \leq \lambda_2(C) \leq \dots \leq \lambda_d(C)$. Let us additionally assume:

(A5) For all $\omega \in \Omega$, matrix $\Sigma(\omega, \hat{\theta}(\omega))$ is regular. There exists a constant $\lambda_{\Sigma} > 0$ such that $\lambda_1(\Sigma(\omega, \hat{\theta}(\omega))) \geq \lambda_{\Sigma}$, for all $\omega \in \Omega$.

For $n \in \mathbb{N}$ let us define random matrix $\Sigma_n(\theta)$ by defining its jk - component in the following way:

$$\Sigma_n(\theta)^{jk} = \sum_{i=1}^n \frac{1}{2} \nu^4(X_{t_{i-1}}) \partial_x \left(\frac{\partial \theta_j \mu(X_{t_{i-1}}, \theta)}{\nu^2(X_{t_{i-1}})} \right) \partial_x \left(\frac{\partial \theta_k \mu(X_{t_{i-1}}, \theta)}{\nu^2(X_{t_{i-1}})} \right) \Delta_n,$$

where $\partial \theta_j := \frac{\partial}{\partial \theta_j}, j = 1, \dots, d$. It is obvious, from the definition, that $\Sigma_n(\theta)^{jk}$ can be considered as a discretization of $\Sigma(\theta)^{jk}$.

Corollary 4.1. *Assume that (A1)-(A5) hold. Assume that $(\bar{\theta}_n, n \in \mathbb{N})$ satisfies (4.2)-(4.5). Then*

$$\left(\sqrt{\Sigma_n(\bar{\theta}_n)} \right)^{-1} D^2 L_n(\bar{\theta}_n) \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}) 1_{\{\Sigma_n(\bar{\theta}_n) \text{ is regular matrix}\}} \xrightarrow{st} N(0, I).$$

Remark 4.1. In the case when the subdivision of the segment $[0, T]$ is not equidistant, all of the results will still hold under the assumptions that $\lim_n \Delta_n = 0$ and $\lim_{n \rightarrow \infty} \max_{i=1, \dots, n} |\frac{h_i - \Delta_n}{\Delta_n}| = 0$, where $h_i := t_i - t_{i-1}$, $i = 1, \dots, n$, $\Delta_n := \max_{1 \leq i \leq n} h_i$, and by replacing $\Sigma_n(\theta)$ with $\tilde{\Sigma}_n(\theta)$ such that

$$\tilde{\Sigma}_n(\theta)^{jk} = \sum_{i=1}^n \frac{1}{2} \nu^4(X_{t_{i-1}}) \partial_x \left(\frac{\partial \theta_j \mu(X_{t_{i-1}}, \theta)}{\nu^2(X_{t_{i-1}})} \right) \partial_x \left(\frac{\partial \theta_k \mu(X_{t_{i-1}}, \theta)}{\nu^2(X_{t_{i-1}})} \right) h_i.$$

5 Proofs

We will prove Theorem 4.1 in number of steps. Some of the lemmas will be proved in Appendix.

Let $\hat{\theta}$ be a continuous-time MLE, and $\bar{\theta}_n$ be an AMLE which satisfies (4.2)-(4.5). Using mean value theorem for the function DL_n which is, due to (A2), continuously differentiable on Θ , for all $\omega \in \Omega$, we get

$$DL_n(\theta) = DL_n(\hat{\theta}) + A_{\hat{\theta}, \theta}(\theta - \hat{\theta}),$$

for all $\theta \in \Theta$, where $A_{\hat{\theta}, \theta} := \int_0^1 D^2 L_n(\hat{\theta} + v(\theta - \hat{\theta})) dv$, and integration is by components. If we put $\theta = \bar{\theta}_n$, then on the event $\{DL_n(\bar{\theta}_n) = 0\}$ we have

$$0 = DL_n(\bar{\theta}_n) = DL_n(\hat{\theta}) + A_{\hat{\theta}, \bar{\theta}_n}(\bar{\theta}_n - \hat{\theta}),$$

where $A_{\hat{\theta}, \bar{\theta}_n} := \int_0^1 D^2 L_n(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta})) dv$. Since $DL_T(\hat{\theta}) = 0$, we can write

$$\begin{aligned} A_{\hat{\theta}, \bar{\theta}_n}(\bar{\theta}_n - \hat{\theta}) &= -DL_n(\hat{\theta}) = DL_T(\hat{\theta}) - DL_n(\hat{\theta}) \Rightarrow \\ A_{\hat{\theta}, \bar{\theta}_n} \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}) &= \frac{1}{\sqrt{\Delta_n}}(DL_T(\hat{\theta}) - DL_n(\hat{\theta})), \end{aligned} \quad (5.1)$$

on the event $\{DL_n(\bar{\theta}_n) = 0\}$.

First, in Theorem 5.1 we will prove that the right hand side of equation (5.1) converges stably in law, for fixed parameter $\theta \in \Theta$, and then in Theorem 5.2 we will prove the same result but for MLE $\hat{\theta}$. After that we will prove Theorem 4.1.

Let $\theta \in \Theta$ be fixed, arbitrary parameter value. Let us denote $X_i := X_{t_i}$, $i = 0, \dots, n$, and $\partial \theta_j := \frac{\partial}{\partial \theta_j}$, $j = 1, \dots, d$. Define functions

$$g_j(x) = g_j(x, \theta) := \frac{\partial \theta_j \mu(x, \theta)}{\nu^2(x)}, j = 1, \dots, d,$$

and denote by g'_j and g''_j its first and second derivative with respect to x . Define functions

$$f_j(x) = f_j(x, \theta) := \frac{\mu(x, \theta) \partial \theta_j \mu(x, \theta)}{\nu^2(x)}, j = 1, \dots, d,$$

and denote by f'_j and f''_j its first and second derivative with respect to x .

(A6) Functions $\mu, \nu, \frac{1}{\nu}, g_j, g'_j, g''_j, f_j, f'_j, f''_j, \frac{\partial}{\partial x} \mu, \frac{\partial^2}{\partial x^2} \mu, \nu', \nu''$ are bounded.

Remark 5.1. The assumption (A6) is only here for technical reasons. First, we will prove our results under assumption (A6), and then we will prove our results without that assumption.

Lemma 5.1. *Assume that (A1)-(A4), (A6) hold. Then, for arbitrary fixed $\theta \in \Theta$ we have*

$$\begin{aligned} & \frac{1}{\sqrt{\Delta_n}}(DL_T(\theta) - DL_n(\theta)) = Z_n(\theta) + \\ & + \frac{1}{\sqrt{\Delta_n}} \begin{bmatrix} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \nu(X_s) \int_{t_{i-1}}^s g'_1(X_u) \nu(X_u) dW_u dW_s \\ \vdots \\ \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \nu(X_s) \int_{t_{i-1}}^s g'_d(X_u) \nu(X_u) dW_u dW_s \end{bmatrix}, \end{aligned}$$

where $Z_n(\theta) \xrightarrow{\mathbb{P}} 0, n \rightarrow +\infty$.

Proof. For details see the Appendix. \square

Theorem 5.1. *Assume that (A1)-(A4) hold. Then, for arbitrary fixed parameter $\theta \in \Theta$ we have*

$$\frac{1}{\sqrt{\Delta_n}}(DL_T(\theta) - DL_n(\theta)) \xrightarrow{st} Y(\theta), \quad (5.2)$$

where $Y(\theta) \sim MN(0, \Sigma(\theta))$.

Proof. First, let us assume that (A6) holds.

Let us denote

$$R_t^j(\theta) := g'_j(X_t) \nu(X_t) = g'_j(X_t, \theta) \nu(X_t),$$

hence,

$$\Sigma(\theta)^{jk} = \frac{1}{2} \int_0^T \nu^2(X_s) R_s^j(\theta) R_s^k(\theta) ds, \quad j, k = 1, \dots, d.$$

Let us define continuous adapted process $C = (C_t^{jk})_{0 \leq t \leq T}$ by

$$C_t^{jk} = \frac{1}{2} \int_0^t \nu^2(X_s) R_s^j(\theta) R_s^k(\theta) ds, \quad j, k = 1, \dots, d, t \in [0, T],$$

and define \mathcal{F}_{t_i} -measurable random vectors $\chi_i^n, i = 1, \dots, n$ by

$$\chi_i^n = \begin{bmatrix} \chi_i^{n,1} \\ \vdots \\ \chi_i^{n,d} \end{bmatrix} = \frac{1}{\sqrt{\Delta_n}} \begin{bmatrix} \int_{t_{i-1}}^{t_i} \nu(X_s) \int_{t_{i-1}}^s R_u^1(\theta) dW_u dW_s \\ \vdots \\ \int_{t_{i-1}}^{t_i} \nu(X_s) \int_{t_{i-1}}^s R_u^d(\theta) dW_u dW_s \end{bmatrix}.$$

We will prove that for C and χ_i^n Theorem 2.1 holds. Random vectors χ_i^n are square integrable due to (A6). Because of its definition, C_t is symmetric positive semidefinite random matrix, for all $t \in [0, T]$. Using notation $\mathcal{F}_{n,i} := \mathcal{F}_{t_i}, i = 1, \dots, n$, we have

$$\mathbb{E}[\chi_i^{n,j} | \mathcal{F}_{n,i-1}] = 0, \quad \forall j = 1, \dots, d, \forall i = 1, \dots, n,$$

hence (2.1) is trivially satisfied.

Let $\epsilon > 0$. Then, there exists constant $k_9 > 0$, such that

$$\left\| \sum_{i=1}^n \mathbb{E}[\|\chi_i^n\|^2 1_{\{\|\chi_i^n\| > \epsilon\}} | \mathcal{F}_{n,i-1}] \right\|_1 \leq \frac{d}{\epsilon^2} \sum_{i=1}^n \mathbb{E}[\sum_{j=1}^d (\chi_i^{n,j})^4] \leq k_9 \Delta_n T,$$

hence (2.4) is satisfied. The last inequality and existence of k_9 follows in a way similar to the proof of the existence of constant k_2 (see the proof of Lemma 5.1 in Appendix).

Let $N = (N_t)_{0 \leq t \leq T}$ be any bounded \mathcal{F}_t -martingale orthogonal to W . Since $(\mathcal{F}_t)_{0 \leq t \leq T}$ is generated by Brownian motion W , it follows from martingale representation theorem (see Jacod, J. and Shiryaev, A. N. (2003)[Theorem III.4.33]) that any \mathcal{F}_t -martingale can be represented as the sum of constant term and a stochastic integral with respect to W . Since N_t is bounded \mathcal{F}_t -martingale orthogonal to W , it follows that N_t is equal to constant, so (2.5) is satisfied.

Let us recall that θ_0 is the true value of (vector) drift parameter. For $j = 1, \dots, d$, we have

$$\begin{aligned} & \sum_{i=1}^{A_n^t} \mathbb{E}[\chi_i^{n,j} (W_{t_i} - W_{t_{i-1}}) | \mathcal{F}_{n,i-1}] = \\ &= \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{A_n^t} \mathbb{E}\left[\int_{t_{i-1}}^{t_i} \nu(X_s) \int_{t_{i-1}}^s R_u^j(\theta) dW_u ds | \mathcal{F}_{n,i-1} \right] = \\ &= \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{A_n^t} \mathbb{E}\left[\int_{t_{i-1}}^{t_i} \left(\int_{t_{i-1}}^s (\nu'(X_u) \mu(X_u, \theta_0) + \frac{1}{2} \nu''(X_u) \sigma_0^2 \nu^2(X_u)) du \right) \right. \\ & \quad \left. \cdot \left(\int_{t_{i-1}}^s R_u^j(\theta) dW_u \right) ds | \mathcal{F}_{n,i-1} \right] + \end{aligned} \tag{5.3}$$

$$+ \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{A_n^t} \mathbb{E}\left[\int_{t_{i-1}}^{t_i} \left(\int_{t_{i-1}}^s \nu'(X_u) \sigma_0 \nu(X_u) dW_u \right) \cdot \left(\int_{t_{i-1}}^s R_u^j(\theta) dW_u \right) ds | \mathcal{F}_{n,i-1} \right]. \tag{5.4}$$

If $\nu \equiv c > 0$ is a constant function, then (2.3) is trivially satisfied, because

$$\begin{aligned} & \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{A_n^t} \mathbb{E}\left[\int_{t_{i-1}}^{t_i} \nu(X_s) \int_{t_{i-1}}^s R_u^j(\theta) dW_u ds | \mathcal{F}_{n,i-1} \right] = \\ &= \frac{c}{\sqrt{\Delta_n}} \sum_{i=1}^{A_n^t} \mathbb{E}\left[\int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s R_u^j(\theta) dW_u ds | \mathcal{F}_{n,i-1} \right] = 0. \end{aligned}$$

So, let us assume that ν is not a constant function. There exist constants k_{10}, k_{11}

such that

$$\begin{aligned}
& \left\| \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{A_n^t} \mathbb{E} \left[\int_{t_{i-1}}^{t_i} \left(\int_{t_{i-1}}^s (\nu'(X_u) \mu(X_u, \theta_0) + \frac{1}{2} \nu''(X_u) \sigma_0^2 \nu^2(X_u)) du \right) \cdot \right. \right. \\
& \quad \left. \left. \cdot \left(\int_{t_{i-1}}^s R_u^j(\theta) dW_u \right) ds \middle| \mathcal{F}_{n,i-1} \right] \right\|_1 \leq k_{10} \sqrt{\Delta_n} T, \\
& \left| \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{A_n^t} \mathbb{E} \left[\int_{t_{i-1}}^{t_i} \left(\int_{t_{i-1}}^s \nu'(X_u) \sigma_0 \nu(X_u) dW_u \right) \cdot \left(\int_{t_{i-1}}^s R_u^j(\theta) dW_u \right) ds \middle| \mathcal{F}_{n,i-1} \right] \right| = \\
& = \left| \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{A_n^t} \mathbb{E} \left[\int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s \nu'(X_u) \sigma_0 \nu(X_u) R_u^j(\theta) duds \middle| \mathcal{F}_{n,i-1} \right] \right| \leq \\
& \leq k_{11} \sqrt{\Delta_n} T,
\end{aligned}$$

so (5.3) converges in L^1 -norm to zero and (5.4) converges almost surely to zero, hence (2.3) is satisfied. For example, for proving the existence of constant k_{10} let us define $M := \sup_{0 \leq s \leq T} \left| \int_0^s R_u^j(\theta) dW_u \right|$. We have

$$\begin{aligned}
& \mathbb{E} \left[\left| \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{A_n^t} \mathbb{E} \left[\int_{t_{i-1}}^{t_i} \left(\int_{t_{i-1}}^s (\nu'(X_u) \mu(X_u, \theta_0) + \frac{1}{2} \nu''(X_u) \sigma_0^2 \nu^2(X_u)) du \right) \cdot \right. \right. \right. \\
& \quad \left. \left. \left. \cdot \left(\int_{t_{i-1}}^s R_u^j(\theta) dW_u \right) ds \middle| \mathcal{F}_{n,i-1} \right] \right| \right] \leq \\
& \leq \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{A_n^t} \mathbb{E} \left[\left| \int_{t_{i-1}}^{t_i} \left(\int_{t_{i-1}}^s (\nu'(X_u) \mu(X_u, \theta_0) + \frac{1}{2} \nu''(X_u) \sigma_0^2 \nu^2(X_u)) du \right) \cdot \right. \right. \\
& \quad \left. \left. \cdot \left(\int_{t_{i-1}}^s R_u^j(\theta) dW_u \right) ds \right| \right] \leq \\
& \leq \frac{const_1}{\sqrt{\Delta_n}} \sum_{i=1}^{A_n^t} \int_{t_{i-1}}^{t_i} (s - t_{i-1}) \mathbb{E}[M] ds \leq const_2 \sqrt{\Delta_n} T \mathbb{E}[M].
\end{aligned}$$

From Doob's inequality we conclude that $\mathbb{E}[M] < +\infty$, and we define $k_{10} := const_2 \mathbb{E}[M]$.

It is only left to prove that (2.2) is satisfied. Let $1 \leq j \leq k \leq d$. Then

$$\begin{aligned}
& \sum_{i=1}^{A_n^t} (\mathbb{E}[\chi_i^{n,j} \chi_i^{n,k} | \mathcal{F}_{n,i-1}] - \mathbb{E}[\chi_i^{n,j} | \mathcal{F}_{n,i-1}] \mathbb{E}[\chi_i^{n,k} | \mathcal{F}_{n,i-1}]) = \\
& = \frac{1}{\Delta_n} \sum_{i=1}^{A_n^t} \mathbb{E} \left[\int_{t_{i-1}}^{t_i} \nu^2(X_s) \int_{t_{i-1}}^s J_u^{(i,j)} R_u^k(\theta) dW_u ds \middle| \mathcal{F}_{n,i-1} \right] + \quad (5.5)
\end{aligned}$$

$$+ \frac{1}{\Delta_n} \sum_{i=1}^{A_n^t} \mathbb{E} \left[\int_{t_{i-1}}^{t_i} \nu^2(X_s) \int_{t_{i-1}}^s J_u^{(i,k)} R_u^j(\theta) dW_u ds \middle| \mathcal{F}_{n,i-1} \right] + \quad (5.6)$$

$$+ \frac{1}{\Delta_n} \sum_{i=1}^{A_n^t} \mathbb{E} \left[\int_{t_{i-1}}^{t_i} \nu^2(X_s) \int_{t_{i-1}}^s R_u^j(\theta) R_u^k(\theta) duds \middle| \mathcal{F}_{n,i-1} \right], \quad (5.7)$$

where $J_s^{(i,j)} := \int_{t_{i-1}}^s R_u^j(\theta) dW_u$. If ν is constant function, then (5.5) and (5.6) are equal to zero, and the proof that (5.7) converges in probability to C_t^{jk} , $\forall t \in [0, T]$ would stay the same as in the rest of the proof. So, let us assume that ν is not a constant function. It can be shown that there exist constants k_{12}, k_{13} , such that

$$\begin{aligned} & \left\| \frac{1}{\Delta_n} \sum_{i=1}^{A_n^t} \mathbb{E} \left[\int_{t_{i-1}}^{t_i} \nu^2(X_s) \int_{t_{i-1}}^s J_u^{(i,j)} R_u^k(\theta) dW_u ds \middle| \mathcal{F}_{n,i-1} \right] \right\|_2 \leq k_{12} \sqrt{\Delta_n} (T + \sqrt{T}), \\ & \left\| \frac{1}{\Delta_n} \sum_{i=1}^{A_n^t} \mathbb{E} \left[\int_{t_{i-1}}^{t_i} \nu^2(X_s) \int_{t_{i-1}}^s J_u^{(i,k)} R_u^j(\theta) dW_u ds \middle| \mathcal{F}_{n,i-1} \right] \right\|_2 \leq k_{13} \sqrt{\Delta_n} (T + \sqrt{T}), \end{aligned}$$

so (5.5) and (5.6) converge in probability to zero. It is left to prove that (5.7) converges in probability to C_t^{jk} . There exists constant k_{14} such that

$$\begin{aligned} & \frac{1}{\Delta_n} \sum_{i=1}^{A_n^t} \mathbb{E} \left[\int_{t_{i-1}}^{t_i} \nu^2(X_s) \int_{t_{i-1}}^s R_u^j(\theta) R_u^k(\theta) duds \middle| \mathcal{F}_{n,i-1} \right] = \\ & = \frac{1}{\Delta_n} \sum_{i=1}^{A_n^t} \left(\mathbb{E} \left[\int_{t_{i-1}}^{t_i} \nu^2(X_s) \int_{t_{i-1}}^s R_u^j(\theta) R_u^k(\theta) duds \middle| \mathcal{F}_{n,i-1} \right] - \right. \\ & \left. - \int_{t_{i-1}}^{t_i} \nu^2(X_s) \int_{t_{i-1}}^s R_u^j(\theta) R_u^k(\theta) duds \right) + \end{aligned} \quad (5.8)$$

$$+ \frac{1}{\Delta_n} \sum_{i=1}^{A_n^t} \int_{t_{i-1}}^{t_i} \nu^2(X_s) \int_{t_{i-1}}^s R_u^j(\theta) R_u^k(\theta) duds, \quad (5.9)$$

and

$$\begin{aligned} & \left\| \frac{1}{\Delta_n} \sum_{i=1}^{A_n^t} \left(\mathbb{E} \left[\int_{t_{i-1}}^{t_i} \nu^2(X_s) \int_{t_{i-1}}^s R_u^j(\theta) R_u^k(\theta) duds \middle| \mathcal{F}_{n,i-1} \right] - \right. \right. \\ & \left. \left. \int_{t_{i-1}}^{t_i} \nu^2(X_s) \int_{t_{i-1}}^s R_u^j(\theta) R_u^k(\theta) duds \right) \right\|_2 \leq k_{14} \sqrt{\Delta_n} \sqrt{T}, \end{aligned}$$

hence (5.8) converges in probability to zero.

We will show that (5.9) converges in probability to C_t^{jk} . Let

$$l(u, s)(\omega) := (\nu^2(X_s) R_u^j(\theta) R_u^k(\theta))(\omega)$$

be a function defined on $[0, T] \times [0, T] \times \Omega$. For fixed $\omega \in \Omega$ function $l(u, s)(\omega)$ is bounded, continuous function on $[0, T] \times [0, T]$, which means that there exist $u_i^*(\omega), s_i^*(\omega) \in [t_{i-1}, t_i]$ such that $u_i^*(\omega) \leq s_i^*(\omega)$ and

$$\int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s l(u, s) duds = \frac{\Delta_n^2}{2} l(u_i^*, s_i^*), \quad i \in \{1, \dots, n\}.$$

Hence,

$$\begin{aligned} & \frac{1}{\Delta_n} \sum_{i=1}^{A_n^t} \int_{t_{i-1}}^{t_i} \nu^2(X_s) \int_{t_{i-1}}^s R_u^j(\theta) R_u^k(\theta) du ds - C_t^{jk} = \\ & = \sum_{i=1}^{A_n^t} \frac{\Delta_n}{2} l(u_i^*, u_i^*) - \frac{1}{2} \int_0^t \nu^2(X_s) R_s^j(\theta) R_s^k(\theta) ds + \end{aligned} \quad (5.10)$$

$$+ \sum_{i=1}^{A_n^t} \frac{\Delta_n}{2} (l(u_i^*, s_i^*) - l(u_i^*, u_i^*)). \quad (5.11)$$

Continuity of integrand function assures that there exist some $t^*(\omega) \in [t_{A_n^t}, t]$, such that $\frac{1}{2} \int_{t_{A_n^t}}^t \nu^2(X_s) R_s^j(\theta) R_s^k(\theta) ds = \frac{1}{2} (t - t_{A_n^t}) \nu^2(X_{t^*}) R_{t^*}^j(\theta) R_{t^*}^k(\theta)$, therefore for (5.10) we have

$$\begin{aligned} & \sum_{i=1}^{A_n^t} \frac{\Delta_n}{2} l(u_i^*, u_i^*) - \frac{1}{2} \int_0^t \nu^2(X_s) R_s^j(\theta) R_s^k(\theta) ds = \\ & = \sum_{i=1}^{A_n^t} \frac{\Delta_n}{2} \nu^2(X_{u_i^*}) R_{u_i^*}^j(\theta) R_{u_i^*}^k(\theta) + \frac{1}{2} (t - t_{A_n^t}) \nu^2(X_{t^*}) R_{t^*}^j(\theta) R_{t^*}^k(\theta) - \\ & - \frac{1}{2} \int_0^t \nu^2(X_s) R_s^j(\theta) R_s^k(\theta) ds - \frac{1}{2} \int_{t_{A_n^t}}^t \nu^2(X_s) R_s^j(\theta) R_s^k(\theta) ds \xrightarrow{a.s.} 0, \end{aligned}$$

hence (5.10) converges in probability to zero. For (5.11) there exists constant k_{15} such that we have

$$\begin{aligned} & \left| \sum_{i=1}^{A_n^t} \frac{\Delta_n}{2} (l(u_i^*, s_i^*) - l(u_i^*, u_i^*)) \right| = \left| \sum_{i=1}^{A_n^t} \frac{\Delta_n}{2} (\nu^2(X_{s_i^*}) - \nu^2(X_{u_i^*})) R_{u_i^*}^j(\theta) R_{u_i^*}^k(\theta) \right| \leq \\ & \leq k_{15} \sum_{i=1}^{A_n^t} \frac{\Delta_n}{2} |\nu^2(X_{s_i^*}) - \nu^2(X_{u_i^*})|. \end{aligned}$$

Let $\epsilon > 0$. For fixed $\omega \in \Omega$ the function $t \mapsto \nu^2(X_t(\omega))$ is continuous on $[0, T]$, so it is uniformly continuous. It means that

$$(\exists \delta > 0)(\forall s, t \in [0, T])(|s - t| < \delta) \Rightarrow |\nu^2(X_s(\omega)) - \nu^2(X_t(\omega))| < \epsilon.$$

We omitted writing ω , because it is fixed. $\lim_{n \rightarrow \infty} \Delta_n = 0$, so there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ holds $\Delta_n < \delta$. Then, for all $n \geq n_0$ we have

$$\sum_{i=1}^{A_n^t} \frac{\Delta_n}{2} |\nu^2(X_{s_i^*}) - \nu^2(X_{u_i^*})| \leq \epsilon \frac{T}{2},$$

which means that (5.11) converges almost surely to zero 0, so it converges in probability to zero.

Therefore, conditions of Theorem 2.1 are satisfied provided that (A6) holds. If we denote by π_T projection function $\pi_T: D([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}^d$, defined by

$\pi_T((X_s, s \in [0, T])) := X_T$, then by (Billingsley, P., 1999, Theorem 12.5) projection π_T is continuous function. Notice that

$$\frac{1}{\sqrt{\Delta_n}}(DL_T(\theta) - DL_n(\theta)) = \pi_T\left(\left(\sum_{i=1}^{A_n^t} \chi_i^n, t \in [0, T]\right) + Z_n(\theta)\right),$$

hence by definition of stable convergence in law, our theorem holds under assumption (A6).

In general case (i.e. without assumption (A6)), let $(E_M, M \in \mathbb{N})$ be a sequence of open and relatively compact subsets of E such that $x_0 \in E_1$, $Cl(E_M) \subseteq E_{M+1}$, $\forall M \in \mathbb{N}$ and $\cup_{M=1}^{\infty} E_M = E$. For that sequence and the solution $(X_t, t \geq 0)$ of our SDE (1.1) define for $M \in \mathbb{N}$

$$T_M := \inf\{t \geq 0: X_t \in E_M^c\},$$

where $\inf \emptyset = +\infty$. Since X exists on $[0, \infty)$ and it is continuous process, $(T_M, M \in \mathbb{N})$ is an increasing sequence of stopping times such that $T_M \uparrow +\infty$ a.s., when $M \rightarrow +\infty$. Let $(\Phi_M, M \in \mathbb{N})$ be a sequence of $C^\infty(E)$ - functions such that $\Phi_M = 1$ on $Cl(E_M)$, and $\Phi_m = 0$ on $Cl(E_{M+1})^c$. Let us define the functions $\mu_M(x, \theta) := \Phi_M(x)\mu(x, \theta)$, $(x, \theta) \in E \times \Theta$ and let ν_M be continuous functions on E such that $\nu_M(x) = \nu(x)$ for $x \in Cl(E_M)$ and $\nu_M(x) = const$ for $x \in E \setminus Cl(E_{M+1})$ (for example, ν_M can be defined as $\nu_M(x) = \Phi_M(x)\nu(x) + (1 - \Phi_M(x))c_M$, where $c_M := \min_{x \in Cl(E_{M+1})} \nu(x)$). The functions μ_M, ν_M are bounded and satisfy assumptions (A2) and (A3). Let, for a fixed $M \in \mathbb{N}$, the process $X^M = (X_t^M, 0 \leq t \leq T)$ be a unique strong solution of the SDE (see Protter, P. (2004))

$$dX_t^M = \mu_M(X_t^M, \theta_0)dt + \sigma_0 \nu_M(X_t^M)dW_t, \quad X_0^M = x_0, x_0 \in E.$$

Moreover, for almost all $\omega \in \Omega$ and all $t \in [0, T_M(\omega)]$, it holds $X_t(\omega) = (X^M)_t(\omega)$. Let $Y(\theta)$ be a random vector such that $Y(\theta) = \sqrt{\Sigma(\theta)}Z$, where Z is standard normal random vector independent of \mathcal{F}_T . Let Y_M be a random vector such that $Y_M(\theta) = \sqrt{\Sigma_M(\theta)}Z$, where $\Sigma_M(\theta)$ is a random matrix $\Sigma(\theta)$ from first part of this proof which we apply on process X^M and functions ν_M and μ_M . Let $\frac{1}{\sqrt{\Delta_n}}(DL_{T,M}(\theta) - DL_{n,M}(\theta))$ be the left hand side of (5.2) for the process X^M , and the functions ν_M and μ_M . Then, from first part of the proof $\frac{1}{\sqrt{\Delta_n}}(DL_{T,M}(\theta) - DL_{n,M}(\theta)) \xrightarrow{st} Y_M(\theta)$. With notation $Z_{n,M}(\theta) := \frac{1}{\sqrt{\Delta_n}}(DL_{T,M}(\theta) - DL_{n,M}(\theta))$ we can write $Z_{n,M}(\theta) \xrightarrow{st} Y_M(\theta)$. Let us denote by $Z_n(\theta) := \frac{1}{\sqrt{\Delta_n}}(DL_T(\theta) - DL_n(\theta))$. We want to show that $Z_n(\theta) \xrightarrow{st} Y(\theta)$, for fixed $\theta \in \Theta$.

Let $f: \mathbb{R}^d \rightarrow \mathbb{R}$ be bounded continuous function, and let U be bounded \mathcal{F}_T -measurable random variable. Let $B, H > 0$ be constants such that $|f| \leq H$ and $|U| \leq B$. Now, we have:

$$\begin{aligned} & |\mathbb{E}[f(Z_n(\theta))U] - \tilde{\mathbb{E}}[f(Y(\theta))U]| \leq \\ & \leq |\mathbb{E}[f(Z_{n,M}(\theta))U1_{\{T_M > T\}}] - \tilde{\mathbb{E}}[f(Y_M(\theta))U1_{\{T_M > T\}}]| + 2BH\mathbb{P}(T_M \leq T). \end{aligned}$$

$U1_{\{T_M > T\}}$ is bounded, \mathcal{F}_T -measurable random variable, and $Z_{n,M}(\theta) \xrightarrow{st} Y_M(\theta)$, hence

$$\overline{\lim}_n |\mathbb{E}[f(Z_n(\theta))U] - \tilde{\mathbb{E}}[f(Y(\theta))U]| \leq 2BH\mathbb{P}(T_M \leq T),$$

and by letting $M \rightarrow \infty$ we have

$$\overline{\lim}_n |\mathbb{E}[f(Z_n(\theta))U] - \tilde{\mathbb{E}}[f(Y(\theta))U]| = 0,$$

which implies

$$\lim_n |\mathbb{E}[f(Z_n(\theta))U] - \tilde{\mathbb{E}}[f(Y(\theta))U]| = 0,$$

which proves our statement. \square

Theorem 5.2. *Assume that (A1)-(A4) hold. Then*

$$\frac{1}{\sqrt{\Delta_n}}(DL_T(\hat{\theta}) - DL_n(\hat{\theta})) \xrightarrow{st} Y(\hat{\theta}),$$

where $Y(\hat{\theta}) \sim MN(0, \Sigma(\hat{\theta}))$.

Remark 5.2. From the proofs of Lemma 5.1 and Theorem 5.1 the following holds. If X is a diffusion linear in drift parameters then the statement of Lemma 5.1 holds directly with θ replaced with $\hat{\theta}$. In that case the double-Itô integral term from the lemma does not depend on θ . Hence the statement of Theorem 5.2 can be proved directly by following the proof of Theorem 5.1 and so without need for the additional assumption about relatively compactness of Θ . Consequently $Y(\theta)$ does not depend on θ .

Proof. Let $Z_n(\theta) \equiv Z_n(\omega, \theta) := \frac{1}{\sqrt{\Delta_n}}(DL_T(\omega, \theta) - DL_n(\omega, \theta))$. Since $\hat{\theta}$ is \mathcal{F}_T -measurable random vector, see (Huzak, M., 2001, Lema 4.1.), and the functions $(\omega, \theta) \mapsto Z_n \equiv Z_n(\omega, \theta)$ are $\mathcal{F}_T \otimes \mathcal{B}(\Theta)$ -measurable (see the beginning of the proof of Lemma 4.2. in Huzak, M. (2001)), for each n , then $\omega \mapsto Z_n(\hat{\theta}) \equiv Z_n(\omega, \hat{\theta}(\omega))$ are \mathcal{F}_T -measurable functions. We want to prove that

$$Z_n(\hat{\theta}) \xrightarrow{st} Y(\hat{\theta}).$$

First, let us assume that all of the functions which appear in the proof are bounded. Let $t \in \mathbb{R}^d$ be arbitrary, fixed, vector, and let U be bounded \mathcal{F}_T -measurable random variable. Let $B > 0$ be a constant such that $|U| \leq B$. For each $n \in \mathbb{N}$, let us define the functions F_n in the following way:

$$F_n(\omega, \theta) := \cos(\langle t, Z_n(\omega, \theta) \rangle) + i \sin(\langle t, Z_n(\omega, \theta) \rangle) - e^{-\frac{1}{2} \sum_{i,j=1}^d t_i t_j \Sigma(\omega, \theta)^{ij}}.$$

For all $\omega \in \Omega$, the functions $\theta \mapsto F_n(\theta) \equiv F_n(\omega, \theta)$ are continuously differentiable on Θ , and hence there exist nonnegative random variable G_n and constant $R \geq 0$ (not depending on n) such that $\mathbb{E}[G_n] \leq R$ and

$$|F_n(\theta_1) - F_n(\theta_2)| \leq G_n \|\theta_1 - \theta_2\|, \quad \forall n \in \mathbb{N}, \forall \theta_1, \theta_2 \in \Theta. \quad (5.12)$$

Namely, since the rationals from $\Theta \cap \mathbb{Q}^d$ are dense in Θ and for $i = 1, \dots, d$, $|\partial \theta_i F_n(\theta)|$ are continuous on compact $Cl(\Theta)$,

$$G_n := \sup_{\theta \in \Theta} \sum_{i=1}^d |\partial \theta_i F_n(\theta)| = \sup_{\theta \in \Theta \cap \mathbb{Q}^d} \sum_{i=1}^d |\partial \theta_i F_n(\theta)|$$

is bounded and measurable, i.e. random variable. By Proposition 1 in Aldous, D. J. and Eagleson, G. K. (1978) it is enough to prove that

$$\lim_n |\mathbb{E}[F_n(\hat{\theta})U]| = 0.$$

Let $\epsilon > 0$ be arbitrary, fixed real number. Define $\delta := \frac{\epsilon}{2RB}$. For all $\theta \in \Theta$, let $K(\theta, \delta)$ be a open ball in \mathbb{R}^d with center θ and radius δ . Since $Cl(\Theta)$ is compact set and $\Theta \subseteq Cl(\Theta) \subseteq \bigcup_{\theta \in \Theta} K(\theta, \delta)$, there exist finite number of balls $K(\theta_l, \delta), l = 1, \dots, N$ such that $\Theta \subseteq \bigcup_{l=1}^N K(\theta_l, \delta)$ and $\theta_l \in \Theta, l = 1, \dots, N$. Let us define a finite partition $\{K_1, \dots, K_N\}$ of Θ in the following way:

$$\begin{aligned} K_1 &:= K(\theta_1, \delta) \cap \Theta \\ K_2 &:= K(\theta_2, \delta) \cap \Theta \cap K_1^c \\ &\vdots \\ K_N &:= K(\theta_N, \delta) \cap \Theta \cap K_1^c \cap K_2^c \cap \dots \cap K_{N-1}^c. \end{aligned}$$

This partition does not depend on subdivision of the segment $[0, T]$. Since $1 = 1_{\{\hat{\theta} \in \Theta\}} = \sum_{l=1}^N 1_{\{\hat{\theta} \in K_l\}}$, we have

$$\begin{aligned} |\mathbb{E}[F_n(\hat{\theta})U]| &= \left| \sum_{l=1}^N \mathbb{E}[F_n(\hat{\theta})U 1_{\{\hat{\theta} \in K_l\}}] \right| = \\ &= \left| \sum_{l=1}^N \mathbb{E}[(F_n(\hat{\theta}) - F_n(\theta_l))U 1_{\{\hat{\theta} \in K_l\}}] + \sum_{l=1}^N \mathbb{E}[F_n(\theta_l)U 1_{\{\hat{\theta} \in K_l\}}] \right| \leq \\ &\leq \left| \sum_{l=1}^N \mathbb{E}[(F_n(\hat{\theta}) - F_n(\theta_l))U 1_{\{\hat{\theta} \in K_l\}}] \right| + \left| \sum_{l=1}^N \mathbb{E}[F_n(\theta_l)U 1_{\{\hat{\theta} \in K_l\}}] \right|. \end{aligned} \tag{5.13}$$

Events $\{\hat{\theta} \in K_1\}, \dots, \{\hat{\theta} \in K_N\}$ make complete system of events for Θ . On the event $\{\hat{\theta} \in K_l\}$ holds $\|\hat{\theta} - \theta_l\| < \delta$, so from (5.12) it follows

$$\begin{aligned} \left| \sum_{l=1}^N \mathbb{E}[(F_n(\hat{\theta}) - F_n(\theta_l))U 1_{\{\hat{\theta} \in K_l\}}] \right| &\leq \sum_{l=1}^N \mathbb{E}[\|F_n(\hat{\theta}) - F_n(\theta_l)\| |U| 1_{\{\hat{\theta} \in K_l\}}] \leq \\ &\leq B \sum_{l=1}^N \mathbb{E}[G_n \|\hat{\theta} - \theta_l\| 1_{\{\hat{\theta} \in K_l\}}] < B\delta \sum_{l=1}^N \mathbb{E}[G_n 1_{\{\hat{\theta} \in K_l\}}] = B\delta \mathbb{E}[G_n] \leq BR\delta = \frac{\epsilon}{2}. \end{aligned}$$

Since $\hat{\theta}$ is \mathcal{F}_T -measurable random vector, it implies that $U 1_{\{\hat{\theta} \in K_l\}}$ is bounded, \mathcal{F}_T -measurable random variable, $l = 1, \dots, N$. Theorem 5.1 and (Aldous, D. J. and Eagleson, G. K., 1978, Proposition 1) implies

$$\lim_n \mathbb{E}[F_n(\theta_l)U 1_{\{\hat{\theta} \in K_l\}}] = 0, \quad l = 1, \dots, N,$$

which yields

$$\lim_n \left| \sum_{l=1}^N \mathbb{E}[F_n(\theta_l)U 1_{\{\hat{\theta} \in K_l\}}] \right| = 0.$$

Let $n_0 = n_0(\epsilon) \in \mathbb{N}$ be natural number such that for all $n \geq n_0$

$$\left| \sum_{l=1}^N \mathbb{E}[F_n(\theta_l) U 1_{\{\hat{\theta} \in K_l\}}] \right| < \frac{\epsilon}{2}.$$

For (5.13) we have

$$\begin{aligned} |\mathbb{E}[F_n(\hat{\theta})U]| &\leq \left| \sum_{l=1}^N \mathbb{E}[(F_n(\hat{\theta}) - F_n(\theta_l))U 1_{\{\hat{\theta} \in K_l\}}] \right| + \left| \sum_{l=1}^N \mathbb{E}[F_n(\theta_l)U 1_{\{\hat{\theta} \in K_l\}}] \right| \leq \\ &\frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \quad \forall n \geq n_0, \end{aligned}$$

so the theorem holds for bounded functions. In the general case, let us define stopping times T_M as in the proof of Theorem 5.1. Let $\hat{\theta}_M$ be a solution of the equation $DL_{T,M}(\theta) = 0$. Since $\hat{\theta}$ is a unique solution of $DL_T(\theta) = 0$, and $DL_{T,M} = DL_T$ on $\{T_M > T\} \times \Theta$, it implies that $\hat{\theta}_M$ and $\hat{\theta}$ coincide on $\{T_M > T\}$. In the first part of the proof we proved $\frac{1}{\sqrt{\Delta_n}}(DL_{T,M}(\hat{\theta}_M) - DL_{n,M}(\hat{\theta}_M)) \xrightarrow{st} Y_M(\hat{\theta}_M)$, shortly $Z_{n,M}(\hat{\theta}_M) \xrightarrow{st} Y_M(\hat{\theta}_M)$. Let $f: \mathbb{R}^d \rightarrow \mathbb{R}$ be bounded continuous function and let $H > 0$ be a real number such that $|f(x)| \leq H, \forall x \in \mathbb{R}^d$.

$$\begin{aligned} &|\mathbb{E}[f(Z_n(\hat{\theta}))U] - \tilde{\mathbb{E}}[f(Y(\hat{\theta}))U]| \leq \\ &\leq |\mathbb{E}[f(Z_{n,M}(\hat{\theta}_M))U 1_{\{T_M > T\}}] - \tilde{\mathbb{E}}[f(Y_M(\hat{\theta}_M))U 1_{\{T_M > T\}}]| + 2BH\mathbb{P}(T_M \leq T). \end{aligned}$$

Since $U 1_{\{T_M > T\}}$ is bounded, \mathcal{F}_T -measurable random variable and $Z_{n,M}(\hat{\theta}_M) \xrightarrow{st} Y_M(\hat{\theta}_M)$, we have

$$\overline{\lim}_n |\mathbb{E}[f(Z_n(\hat{\theta}))U] - \tilde{\mathbb{E}}[f(Y(\hat{\theta}))U]| \leq 2BH\mathbb{P}(T_M \leq T),$$

and by letting $M \rightarrow \infty$, we get

$$\overline{\lim}_n |\mathbb{E}[f(Z_n(\hat{\theta}))U] - \tilde{\mathbb{E}}[f(Y(\hat{\theta}))U]| = 0,$$

which implies

$$\lim_n |\mathbb{E}[f(Z_n(\hat{\theta}))U] - \tilde{\mathbb{E}}[f(Y(\hat{\theta}))U]| = 0.$$

□

Lemma 5.2. *Assume that (A1)-(A4) hold. Assume that $(\bar{\theta}_n, n \in \mathbb{N})$ satisfies (4.2)-(4.5). Then*

$$A_{\hat{\theta}, \bar{\theta}_n} \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}) \xrightarrow{st} Y(\hat{\theta}),$$

where $Y(\hat{\theta}) \sim MN(0, \Sigma(\hat{\theta}))$.

Proof. First, let ν^{-1} be a bounded function and let μ, ν and all of the functions which appear in the proof be bounded. Since (4.2) holds, it implies that

$$1_{\{DL_n(\bar{\theta}_n)=0\}} \xrightarrow{\mathbb{P}} 1 \text{ and } 1_{\{DL_n(\bar{\theta}_n) \neq 0\}} \xrightarrow{\mathbb{P}} 0. \text{ For } A_{\hat{\theta}, \bar{\theta}_n} \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}) \text{ holds}$$

$$\begin{aligned} &A_{\hat{\theta}, \bar{\theta}_n} \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}) = \\ &\frac{1}{\sqrt{\Delta_n}} ((DL_T(\hat{\theta}) - DL_n(\hat{\theta})) 1_{\{DL_n(\bar{\theta}_n)=0\}} + A_{\hat{\theta}, \bar{\theta}_n} \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}) 1_{\{DL_n(\bar{\theta}_n) \neq 0\}}). \end{aligned}$$

Theorem 5.2 and (Jacod, J. and Protter, P., 2012, (2.2.5)) imply $(\frac{1}{\sqrt{\Delta_n}}((DL_T(\hat{\theta}) - DL_n(\hat{\theta})), 1_{\{DL_n(\bar{\theta}_n)=0\}}) \xrightarrow{st} (Y(\hat{\theta}), 1)$, hence from (Aldous, D. J. and Eagleson, G. K., 1978, Theorem 1) follows that $\frac{1}{\sqrt{\Delta_n}}((DL_T(\hat{\theta}) - DL_n(\hat{\theta})), 1_{\{DL_n(\bar{\theta}_n)=0\}}) \xrightarrow{st} Y(\hat{\theta})$. It holds

$$\|A_{\hat{\theta}, \bar{\theta}_n} \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}) 1_{\{DL_n(\bar{\theta}_n) \neq 0\}}\| \leq \|A_{\hat{\theta}, \bar{\theta}_n}\| \frac{1}{\sqrt{\Delta_n}} \|\bar{\theta}_n - \hat{\theta}\| 1_{\{DL_n(\bar{\theta}_n) \neq 0\}}.$$

Sequences $\|A_{\hat{\theta}, \bar{\theta}_n}\|$ and $\frac{1}{\sqrt{\Delta_n}} \|\bar{\theta}_n - \hat{\theta}\|$ are bounded in probability and $1_{\{DL_n(\bar{\theta}_n) \neq 0\}} \xrightarrow{\mathbb{P}} 0$, so we can conclude that $A_{\hat{\theta}, \bar{\theta}_n} \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}) 1_{\{DL_n(\bar{\theta}_n) \neq 0\}} \xrightarrow{\mathbb{P}} 0$, hence by (Jacod, J., 1997, (2.2)) this theorem holds for bounded functions. For general case, we define a sequence of stopping times T_M (as in the proof of Theorem 5.2). Let $\hat{\theta}_M$ be defined in the same way as in Theorem 5.2. Let $\bar{\theta}_{n,M}$ be a sequence which satisfies (4.2)-(4.5) for the functions μ_M, ν_M and the process X^M . Since (4.4) holds, we have $\lim_n \mathbb{P}(\bar{\theta}_{n,M} = \bar{\theta}_n | T_M > T) = 1$. Let f be bounded real function, and U be a bounded random variable. Let $B, H > 0$ be constants such that $|U| \leq B$ and $|f| \leq H$. Then, with definition $A_{\theta_1, \theta_2, M} := \int_0^1 D^2 L_{n,M}(\theta_1 + v(\theta_2 - \theta_1)) dv$, we have

$$\begin{aligned} & |\mathbb{E}[f(A_{\hat{\theta}, \bar{\theta}_n} \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}))U] - \tilde{\mathbb{E}}[f(Y(\hat{\theta}))U]| \leq \\ & \leq |\mathbb{E}[f(A_{\hat{\theta}, \bar{\theta}_{n,M}} \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}))U 1_{\{T_M > T\}} 1_{\{\bar{\theta}_{n,M} = \bar{\theta}_n\}}] - \\ & - \tilde{\mathbb{E}}[f(Y_M(\hat{\theta}_M))U 1_{\{T_M > T\}}]| + 2BH\mathbb{P}(T_M \leq T) + BH\mathbb{P}(\bar{\theta}_{n,M} \neq \bar{\theta}_n, T_M > T), \end{aligned}$$

and the proof is over once we let first $n \rightarrow +\infty$, and then $M \rightarrow +\infty$. \square

Lemma 5.3. *Assume that (A1)-(A4) hold. Assume that $(\bar{\theta}_n, n \in \mathbb{N})$ satisfies (4.2)-(4.5). Then*

$$\int_0^1 D^2 L_T(\hat{\theta}) - D^2 L_T(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta})) dv \xrightarrow{\mathbb{P}} 0. \quad (5.14)$$

Proof. It is sufficient to prove that $\sup_{v \in [0,1]} \|D^2 L_T(\hat{\theta}) - D^2 L_T(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta}))\| \xrightarrow{\mathbb{P}} 0$. We will prove that $\sup_{v \in [0,1]} F(\bar{\theta}_n, v) \xrightarrow{\mathbb{P}} 0$, for the functions

$$F(\theta, v) \equiv F(\omega, \theta, v) := \|D^2 L_T(\omega, \hat{\theta}(\omega)) - D^2 L_T(\omega, \hat{\theta}(\omega) + v(\theta - \hat{\theta}(\omega)))\|.$$

For details see the Appendix. \square

Lemma 5.4. *Assume that (A1)-(A4) hold. Assume that $(\bar{\theta}_n, n \in \mathbb{N})$ satisfies (4.2)-(4.5). Then*

$$D^2 L_T(\hat{\theta}) \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}) \xrightarrow{st} MN(0, \Sigma(\hat{\theta})). \quad (5.15)$$

Proof. We will prove that the sequence $D^2 L_T(\hat{\theta}) \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta})$ can be written as the sum of the sequence which converges in probability to zero and the sequence which converges stably in law to $MN(0, \Sigma(\hat{\theta}))$. For details see the Appendix. \square

Lemma 5.5. *Assume that (A1)-(A4) hold. Assume that $(\bar{\theta}_n, n \in \mathbb{N})$ satisfies (4.2)-(4.5). Then*

$$D^2 L_n(\bar{\theta}_n) \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}) \xrightarrow{st} Y(\hat{\theta}).$$

Proof.

$$\begin{aligned} & D^2 L_n(\bar{\theta}_n) \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}) = \\ & = (D^2 L_n(\bar{\theta}_n) - D^2 L_T(\bar{\theta}_n)) \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}) + \\ & + (D^2 L_T(\bar{\theta}_n) - D^2 L_T(\hat{\theta})) \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}) + \\ & + D^2 L_T(\hat{\theta}) \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}). \end{aligned}$$

(4.6) implies

$$D^2 L_n(\bar{\theta}_n) - D^2 L_T(\bar{\theta}_n) \xrightarrow{\mathbb{P}} 0,$$

and since the sequence $\frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta})$ is bounded in probability, it follows that

$$(D^2 L_n(\bar{\theta}_n) - D^2 L_T(\bar{\theta}_n)) \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}) \xrightarrow{\mathbb{P}} 0.$$

From Lemma 5.3, when $v = 1$, we get

$$D^2 L_T(\bar{\theta}_n) - D^2 L_T(\hat{\theta}) \xrightarrow{\mathbb{P}} 0,$$

hence

$$(D^2 L_T(\bar{\theta}_n) - D^2 L_T(\hat{\theta})) \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}) \xrightarrow{\mathbb{P}} 0.$$

Lemma 5.4 and (Jacod, J., 1997, (2.2)) imply the conclusion. \square

Lemma 5.6. *Assume that (A1)-(A5) hold. Assume that $(\bar{\theta}_n, n \in \mathbb{N})$ satisfies (4.2)-(4.5). Then*

$$(\sqrt{\Sigma(\hat{\theta})})^{-1} D^2 L_n(\bar{\theta}_n) \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}) \xrightarrow{st} N(0, I).$$

Proof. This Lemma is consequence of Lemma 5.5 and (Aldous, D. J. and Eagleson, G. K., 1978, Theorem 1). \square

Lemma 5.7. *Assume that (A1)-(A4) hold. Assume that $(\bar{\theta}_n, n \in \mathbb{N})$ satisfies (4.2)-(4.5). Then*

$$\Sigma_n(\bar{\theta}_n) \xrightarrow{\mathbb{P}} \Sigma(\hat{\theta}).$$

Proof. We will prove that each component $\Sigma_n(\bar{\theta}_n)^{jk}$ can be written as the sum of sequence which converges in probability to zero and the random variable $\Sigma(\hat{\theta})^{jk}$. For details see the Appendix. \square

Lemma 5.8. *Assume that (A1)-(A5) hold. Assume that $(\bar{\theta}_n, n \in \mathbb{N})$ satisfies (4.2)-(4.5). Then*

$$\lim_{n \rightarrow \infty} \mathbb{P}(\Sigma_n(\bar{\theta}_n) \text{ is regular matrix}) = 1.$$

Proof. For details see the Appendix. \square

5.1 Proofs of Theorem 4.1 and Corollary 4.1

Proof of Theorem 4.1. Since $D^2L_T(\hat{\theta})$ is regular symmetric matrix, there exists $(D^2L_T(\hat{\theta}))^{-1}$, and it is symmetric matrix. Since $D^2L_T(\hat{\theta})$ is \mathcal{F}_T -measurable random matrix, and multiplication with inverse is continuous mapping, (Aldous, D. J. and Eagleson, G. K., 1978, Theorem 1) and Lemma 5.4 imply

$$\frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}) \xrightarrow{st} (D^2L_T(\hat{\theta}))^{-1}MN(0, \Sigma(\hat{\theta})),$$

where $(D^2L_T(\hat{\theta}))^{-1}MN(0, \Sigma(\hat{\theta})) \sim MN(0, (D^2L_T(\hat{\theta}))^{-1}\Sigma(\hat{\theta})(D^2L_T(\hat{\theta}))^{-1})$. \square

Proof of Corollary 4.1. Using continuous mapping theorem, from Lemma 5.7 and Lemma 5.8 it follows that

$$\left(\sqrt{\Sigma_n(\bar{\theta}_n)}\right)^{-1} 1_{\{\Sigma_n(\bar{\theta}_n) \text{ is regular matrix}\}} \xrightarrow{\mathbb{P}} \left(\sqrt{\Sigma(\hat{\theta})}\right)^{-1}, \quad n \rightarrow \infty.$$

For $\left(\sqrt{\Sigma_n(\bar{\theta}_n)}\right)^{-1} D^2L_n(\bar{\theta}_n) \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}) 1_{\{\Sigma_n(\bar{\theta}_n) \text{ is regular matrix}\}}$ one may write

$$\begin{aligned} & \left(\sqrt{\Sigma_n(\bar{\theta}_n)}\right)^{-1} D^2L_n(\bar{\theta}_n) \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}) 1_{\{\Sigma_n(\bar{\theta}_n) \text{ is regular matrix}\}} = \\ & = \left(\left(\sqrt{\Sigma_n(\bar{\theta}_n)}\right)^{-1} 1_{\{\Sigma_n(\bar{\theta}_n) \text{ is regular matrix}\}}\right)^{-} \\ & - \left(\sqrt{\Sigma(\hat{\theta})}\right)^{-1} D^2L_n(\bar{\theta}_n) \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}) + \\ & + \left(\sqrt{\Sigma(\hat{\theta})}\right)^{-1} D^2L_n(\bar{\theta}_n) \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}). \end{aligned}$$

Since $D^2L_n(\bar{\theta}_n) \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}), n \in \mathbb{N}$ converges stably in law (Lemma 5.5), it also converges in distribution, which implies that this sequence is bounded in probability. Hence, it follows that

$$\begin{aligned} & \left(\left(\sqrt{\Sigma_n(\bar{\theta}_n)}\right)^{-1} 1_{\{\Sigma_n(\bar{\theta}_n) \text{ is regular matrix}\}}\right)^{-} \\ & \left(\sqrt{\Sigma(\hat{\theta})}\right)^{-1} D^2L_n(\bar{\theta}_n) \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}) \xrightarrow{\mathbb{P}} 0, \end{aligned}$$

and Lemma 5.6 and (Jacod, J., 1997, (2.2)) yields the result. \square

6 Simulations

Example 1. SDE

$$dX_t = aX_t dt + dW_t, \quad X_0 = 0,$$

is given, where $a \in \mathbb{R}$ is unknown parameter, and process is observed over fixed time interval $[0, T]$. For $n \in \mathbb{N}$, define $h = \frac{T}{n}$ and equidistant subdivision of $[0, T]$ with $t_i = ih, i = 1, \dots, n$. Then MLE, based on the continuous-time sample $(X_t)_{t \in [0, T]}$, \hat{a} is given by (see Liptser, R. S. and Shirayayev, A. N. (1977))

$$\hat{a} = \frac{X_T^2 - T}{2 \int_0^T X_t^2 dt}.$$

Let $(X_1, \dots, X_n) := (X_{t_1}, \dots, X_{t_n})$ be discrete random sample from the model. For the simulation study, $\int_0^T X_t^2 dt$ will be approximated by $h \sum_{i=1}^n X_{i-1}^2$, hence \hat{a} will be estimated by the formula

$$\hat{a} = \frac{X_T^2 - T}{2h \sum_{i=1}^n X_{i-1}^2}. \quad (6.1)$$

If we define the functions $\mu(x, a) = ax$, $\nu(x) = 1$, and $\sigma_0 = 1$, then using (3.3) we get

$$L_n(a) = \sum_{i=1}^n \left((X_i - X_{i-1})X_{i-1}a - \frac{1}{2}X_{i-1}^2 ha^2 \right),$$

from which we calculate AMLE \bar{a}_n , which is given by formula

$$\bar{a}_n = \frac{\sum_{i=1}^n (X_i - X_{i-1})X_{i-1}}{h \sum_{i=1}^n X_{i-1}^2}.$$

Therefore, all of the assumptions of Corollary 4.1 are satisfied, and we have

$$-\frac{\sqrt{2}}{\sqrt{n}} \sum_{i=1}^n X_{i-1}^2 (\bar{a}_n - \hat{a}) \xrightarrow{st} N(0, 1). \quad (6.2)$$

Let $-z_{\frac{\alpha}{2}}$ be $\frac{\alpha}{2}$ -quantile of standard normal distribution. We simulated M realizations of discrete random sample $(X_{t_1}, \dots, X_{t_n})$ with parameter $a = 2$ over equidistant points $t_i = i\Delta_n$, where $n = 2^k$, $T = 1$ and $\Delta_n = \frac{T}{n}$. Then we calculated how many times values of the variable

$$-\frac{\sqrt{2}}{\sqrt{n}} \sum_{i=1}^n X_{i-1}^2 (\bar{a}_n - \hat{a}) \quad (6.3)$$

are in the interval $[-z_{\frac{\alpha}{2}}, z_{\frac{\alpha}{2}}]$ and we present that number as a percentage. Results are presented in Tables 1, 2 and 3.

Simulations showed (see Tables 1, 2 and 3) that increasing the number n causes increase in percentage of values of variable (6.3) which are in the interval $[-z_{\frac{\alpha}{2}}, z_{\frac{\alpha}{2}}]$, in the sense that percentage becomes value near $(1 - \alpha) \cdot 100\%$, what was expected.

Example 2. Gompertz model is given by SDE

$$dX_t = (a - b \ln X_t)X_t dt + \sigma_0 X_t dW_t, \quad X(0) = 1,$$

where $a, b \in \mathbb{R}$ are unknown parameters, and $\sigma_0^2 = 0.84$. Process is observed over time-interval $[0, T]$. In this example, we define the functions $\mu(x, a, b) = (a - b \ln x)x$ and $\nu(x) = x$. MLE \hat{a}, \hat{b} of the parameters a, b are calculated using (Brown, B. M. and Hewitt, J. I., 1975, Equation 20) and given by

$$\begin{aligned} \hat{b} &= \frac{T \int_0^T \frac{\ln X_s}{X_s} dX_s - \int_0^T \frac{1}{X_s} dX_s \cdot \int_0^T \ln X_s ds}{\left(\int_0^T \ln X_s ds \right)^2 - T \int_0^T (\ln X_s)^2 ds} \\ \hat{a} &= \frac{\hat{b} \int_0^T \ln X_s ds + \int_0^T \frac{1}{X_s} dX_s}{T}. \end{aligned} \quad (6.4)$$

If $(X_1, \dots, X_n) := (X_{t_1}, \dots, X_{t_n})$ is discrete random sample from the model, then using (3.3) we get

$$L_n(a, b) = \sum_{i=1}^n \left(\frac{(X_i - X_{i-1})(a - b \ln X_{i-1})}{0.84 X_{i-1}} - \frac{1}{2} \frac{(a - b \ln X_{i-1})^2 h}{0.84} \right).$$

L_n is quadratic function of a and b , which means that there exists a point of maximum of that function. Hence, AMLE \bar{a}_n and \bar{b}_n of the parameters a and b can be calculated using formulas

$$\begin{aligned} \bar{b}_n &= \frac{\sum_{i=1}^n \frac{X_i - X_{i-1}}{X_{i-1}} \cdot \sum_{i=1}^n \ln X_{i-1} - n \sum_{i=1}^n \frac{X_i - X_{i-1}}{X_{i-1}} \ln X_{i-1}}{T \sum_{i=1}^n (\ln X_{i-1})^2 - h (\sum_{i=1}^n \ln X_{i-1})^2} \\ \bar{a}_n &= \frac{\bar{b}_n h \sum_{i=1}^n \ln X_{i-1} + \sum_{i=1}^n \frac{X_i - X_{i-1}}{X_{i-1}}}{T}. \end{aligned} \quad (6.5)$$

In this case MLE given by (6.4) can not be calculated exactly, so we will use formulas (6.5) as their good approximation. For that purpose we used different (much larger) number of points for calculating MLE then for AMLE. In this way obtained estimate of MLE is not equal to estimate of AMLE. If we define the functions $\mu(x, a, b) = a - b \ln x$, $\nu(x) = x$, then all of the assumptions of Corollary 4.1 are satisfied, and we have

$$\begin{aligned} \langle \left(\sqrt{\Sigma_n(\bar{\theta}_n)} \right)^{-1} D^2 L_n(\bar{\theta}_n) \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}), \left(\sqrt{\Sigma_n(\bar{\theta}_n)} \right)^{-1} D^2 L_n(\bar{\theta}_n) \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}) \rangle \\ \xrightarrow{st} \chi_2^2, \end{aligned}$$

where $\theta = (a, b)$, and matrices $\Sigma_n(\bar{\theta}_n)$ and $D^2 L_n(\bar{\theta}_n)$ are given by formulas

$$\begin{aligned} \Sigma_n(\bar{\theta}_n) &= \begin{bmatrix} \frac{T}{2} & \frac{T - h \sum_{i=1}^n \ln X_{i-1}}{2} \\ \frac{T - h \sum_{i=1}^n \ln X_{i-1}}{2} & \frac{T - 2h \sum_{i=1}^n \ln X_{i-1} + h \sum_{i=1}^n (\ln X_{i-1})^2}{2} \end{bmatrix}, \\ D^2 L_n(\bar{\theta}_n) &= \frac{1}{\sigma_0^2} \begin{bmatrix} -T & h \sum_{i=1}^n \ln X_{i-1} \\ h \sum_{i=1}^n \ln X_{i-1} & -h \sum_{i=1}^n (\ln X_{i-1})^2 \end{bmatrix}. \end{aligned}$$

We simulated M realizations of discrete random sample $(X_{t_1}, \dots, X_{t_n})$ with parameters $a = 9.72$, $b = 3.7$, over equidistant points $t_i = i\Delta_n$, where $n = 2^k$, $T = 1$ and $\Delta_n = \frac{T}{n}$. We simulated in the way that $n' = 2^l > 2^k$ and we use all of these points to estimate MLE $\hat{\theta}$ by using (6.5). Then we take sub-sample of length $n = 2^k$ and calculate AMLE $\bar{\theta}_n$ by (6.5). Then, we calculate percentage of values $\langle (\sqrt{\Sigma_n(\bar{\theta}_n)})^{-1} D^2 L_n(\bar{\theta}_n) \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}), (\sqrt{\Sigma_n(\bar{\theta}_n)})^{-1} D^2 L_n(\bar{\theta}_n) \frac{1}{\sqrt{\Delta_n}} (\bar{\theta}_n - \hat{\theta}) \rangle$ which are in the interval $[0, \chi_{1-\alpha}^2]$, where $\chi_{1-\alpha}^2$ is $(1 - \alpha)$ -quantile of χ^2 distribution with 2 degrees of freedom. Results are presented in Tables 4, 5, 6, 7 and 8. Simulations showed (see Tables 6, 7, 8) that increase in number k causes increase of percentage, in the sense that percentage, for large enough k , is value close to $1 - \alpha$, for given α , what we expected.

Example 3. Let us consider ergodic recurrent diffusion X which satisfies SDE (1.1). Let $\bar{\theta}_{n,T}$ and $\hat{\theta}_T$ be AMLE and MLE of drift parameters over interval $[0, T]$. Let us assume that $\lim_{T \rightarrow +\infty} \hat{\theta}_T = \theta_0$ a.s., $\lim_{T \rightarrow +\infty} \frac{1}{T} D^2 L_T(\hat{\theta}_T) =$

$-M_0$ a.s., $\sqrt{T}(\hat{\theta}_T - \theta_0) \Rightarrow N(0, M_0^{-1})$, $T \rightarrow +\infty$ and $\lim_{T \rightarrow +\infty} \frac{1}{T} \Sigma(\hat{a}_T) = \Sigma_0$ a.s. for some $d \times d$ positive definite matrices M_0 and Σ_0 . Since the difference $\sqrt{T}(\bar{\theta}_{n,T} - \theta_0)$ can be written in the form

$$\sqrt{T}(\bar{\theta}_{n,T} - \theta_0) = \sqrt{\Delta_n T} \frac{\bar{\theta}_{n,T} - \hat{\theta}_T}{\sqrt{\Delta_n}} + \sqrt{T}(\hat{\theta}_T - \theta_0),$$

for large enough T and n one can say that the difference $\sqrt{T}(\bar{\theta}_{n,T} - \theta_0)$ asymptotically behaves as $\sqrt{\Delta_n} (\frac{1}{T} D^2 L_T(\hat{\theta}_T))^{-1} (\frac{1}{T} \Sigma(\hat{\theta}_T))^{\frac{1}{2}} Z_T + (M_0^{-1})^{\frac{1}{2}} Z$. But if T is large enough, then we can say that the difference $\sqrt{T}(\bar{\theta}_{n,T} - \theta_0)$ asymptotically behaves as $\sqrt{\Delta_n} (-M_0)^{-1} (\Sigma_0)^{\frac{1}{2}} Z_T + (M_0^{-1})^{\frac{1}{2}} Z$, where Z_T is standard normal random vector independent of \mathcal{F}_T , and Z is standard normal random vector. For example, if we look at the variance at one dimensional case

$$\begin{aligned} \text{Var}(\sqrt{\Delta_n} (\frac{1}{T} D^2 L_T(\hat{\theta}_T))^{-1} (\frac{1}{T} \Sigma(\hat{\theta}_T))^{\frac{1}{2}} Z_T + \sqrt{T}(\hat{\theta}_T - a_0)) = \\ = \text{Var}(\sqrt{\Delta_n} (\frac{1}{T} D^2 L_T(\hat{\theta}_T))^{-1} (\frac{1}{T} \Sigma(\hat{\theta}_T))^{\frac{1}{2}} Z_T) + \text{Var}(\sqrt{T}(\hat{\theta}_T - a_0)), \end{aligned}$$

then for simulation study we can use this variance to explain standard error and to get asymptotic confidence intervals for true parameter value. This is only the hint how to use the result presented in this paper in future simulation study in the case of ergodic diffusions.

APPENDIX

Proof of Lemma 5.1. For the j -th component of $\frac{1}{\sqrt{\Delta_n}} (DL_T(\theta) - DL_n(\theta))$, $j = 1, \dots, d$, using Itô's formula, we get

$$\begin{aligned} \frac{1}{\sqrt{\Delta_n}} (DL_T(\theta) - DL_n(\theta))_j = \\ = \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \frac{\mu(X_s, \theta_0)}{\sigma_0^2} \left(\int_{t_{i-1}}^s (g'_j(X_u) \mu(X_u, \theta_0) + \frac{1}{2} g''_j(X_u) \sigma_0^2 \nu^2(X_u)) du \right) ds - \end{aligned} \quad (6.6)$$

$$- \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \frac{1}{\sigma_0^2} \left(\int_{t_{i-1}}^s (f'_j(X_u) \mu(X_u, \theta_0) + \frac{1}{2} f''_j(X_u) \sigma_0^2 \nu^2(X_u)) du \right) ds + \quad (6.7)$$

$$+ \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \frac{\mu(X_s, \theta_0)}{\sigma_0} \left(\int_{t_{i-1}}^s g'_j(X_u) \nu(X_u) dW_u \right) ds - \quad (6.8)$$

$$- \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \frac{1}{\sigma_0} \left(\int_{t_{i-1}}^s f'_j(X_u) \nu(X_u) dW_u \right) ds + \quad (6.9)$$

$$+ \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \frac{\nu(X_s)}{\sigma_0} \left(\int_{t_{i-1}}^s (g'_j(X_u) \mu(X_u, \theta_0) + \frac{1}{2} g''_j(X_u) \sigma_0^2 \nu^2(X_u)) du \right) dW_s + \quad (6.10)$$

$$+ \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \nu(X_s) \left(\int_{t_{i-1}}^s g'_j(X_u) \nu(X_u) dW_u \right) dW_s. \quad (6.11)$$

If we define $Z_{n,j}(\theta) := (6.6) + (6.7) + (6.8) + (6.9) + (6.10)$, then

$$Z_n(\theta) = \begin{bmatrix} Z_{n,1}(\theta) \\ \vdots \\ Z_{n,d}(\theta) \end{bmatrix}.$$

If, for some $j = 1, \dots, d$, g_j is a constant function, then we define $Z_{n,j}(\theta)$ as $Z_{n,j}(\theta) := (6.7) + (6.9)$. If, for some $j = 1, \dots, d$, f_j is a constant function, then we define $Z_{n,j}(\theta)$ as $Z_{n,j}(\theta) := (6.6) + (6.8) + (6.10) + (6.11)$. So, the conclusion and the proof of this Lemma stay the same in those special cases. It is sufficient to prove that each component $Z_{n,j}(\theta), j = 1, \dots, d$, of the vector $Z_n(\theta)$ converges in probability to zero. Using (A6) it is easy to prove that there exists constant $k_1 > 0$, such that

$$\begin{aligned} & \left| \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \frac{\mu(X_s, \theta_0)}{\sigma_0^2} \left(\int_{t_{i-1}}^s (g'_j(X_u) \mu(X_u, \theta_0) + \frac{1}{2} g''_j(X_u) \sigma_0^2 \nu^2(X_u)) du \right) ds - \right. \\ & \left. - \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \frac{1}{\sigma_0^2} \left(\int_{t_{i-1}}^s (f'_j(X_u) \mu(X_u, \theta_0) + \frac{1}{2} f''_j(X_u) \sigma_0^2 \nu^2(X_u)) du \right) ds \right| \leq \\ & k_1 \sqrt{\Delta_n} T, \end{aligned}$$

so we conclude that (6.6) + (6.7) converges in probability to zero.

Let us denote with μ' and μ'' first and second derivative of the function $x \mapsto \mu(x, \theta_0)$ with respect to x . For (6.8), we have

$$\begin{aligned} & \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \frac{\mu(X_s, \theta_0)}{\sigma_0} \left(\int_{t_{i-1}}^s g'_j(X_u) \nu(X_u) dW_u \right) ds = \\ & = \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s \frac{\mu(X_{i-1}, \theta_0)}{\sigma_0} g'_j(X_u) \nu(X_u) dW_u ds + \end{aligned} \quad (6.12)$$

$$\begin{aligned} & + \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \left(\int_{t_{i-1}}^s (\mu'(X_u, \theta_0) \mu(X_u, \theta_0) + \frac{1}{2} \mu''(X_u, \theta_0) \sigma_0^2 \nu^2(X_u)) du \right) \cdot \\ & \cdot \left(\int_{t_{i-1}}^s \frac{1}{\sigma_0} g'_j(X_u) \nu(X_u) dW_u \right) ds + \end{aligned} \quad (6.13)$$

$$\begin{aligned} & + \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s \left(\int_{t_{i-1}}^u \mu'(X_l, \theta_0) \nu(X_l) dW_l \right) \cdot (g'_j(X_u) \nu(X_u)) dW_u ds + \\ & \end{aligned} \quad (6.14)$$

$$\begin{aligned} & + \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s \left(\int_{t_{i-1}}^u g'_j(X_l) \nu(X_l) dW_l \right) \cdot (\mu'(X_u, \theta_0) \nu(X_u)) dW_u ds + \\ & \end{aligned} \quad (6.15)$$

$$\begin{aligned} & + \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s \mu'(X_u, \theta_0) \nu^2(X_u) g'_j(X_u) du ds, \\ & \end{aligned} \quad (6.16)$$

where (6.14)-(6.16) we get by applying Itô's formula for

$$\int_{t_{i-1}}^s \mu'(X_u, \theta_0) \nu(X_u) dW_u \int_{t_{i-1}}^s g'_j(X_u) \nu(X_u) dW_u.$$

In the case when $\mu(\cdot, \theta_0)$ is a constant function, (6.8) would have the same form as (6.12). It can be shown that there exist constants k_2, k_3, k_4, k_5, k_6 such that

$$\begin{aligned}
& \left\| \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s \frac{\mu(X_{i-1}, \theta_0)}{\sigma_0} g'_j(X_u) \nu(X_u) dW_u ds \right\|_2 \leq k_2 \sqrt{\Delta_n} \sqrt{T}, \\
& \left\| \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \left(\int_{t_{i-1}}^s (\mu'(X_u, \theta_0) \mu(X_u, \theta_0) + \frac{1}{2} \mu''(X_u, \theta_0) \sigma_0^2 \nu^2(X_u)) du \right) \right. \\
& \quad \cdot \left. \left(\int_{t_{i-1}}^s \frac{1}{\sigma_0} g'_j(X_u) \nu(X_u) dW_u \right) ds \right\|_2 \leq k_3 \sqrt{\Delta_n} T \\
& \left\| \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s \left(\int_{t_{i-1}}^u \mu'(X_l, \theta_0) \nu(X_l) dW_l \right) \cdot (g'_j(X_u) \nu(X_u)) dW_u ds \right\|_2 \leq \\
& \leq k_4 \Delta_n \sqrt{T} \\
& \left\| \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s \left(\int_{t_{i-1}}^u g'_j(X_l) \nu(X_l) dW_l \right) \cdot (\mu'(X_u, \theta_0) \nu(X_u)) dW_u ds \right\|_2 \leq \\
& \leq k_5 \Delta_n \sqrt{T} \\
& \left| \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s \mu'(X_u, \theta_0) \nu^2(X_u) g'_j(X_u) dud s \right| \leq k_6 \sqrt{\Delta_n} T,
\end{aligned}$$

so (6.12), (6.13), (6.14) and (6.15) converge in L^2 -norm to zero, and (6.16) converges almost surely to zero. From this we conclude that (6.8) converges in probability to zero.

For example, let us show how we can get constant k_2 . The other constants in this article can be calculated in similar way. Let $h_i := t_i - t_{i-1}$. For (6.12) we have

$$\begin{aligned}
& \mathbb{E} \left[\left(\frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s \frac{\mu(X_{i-1}, \theta_0)}{\sigma_0} g'_j(X_u) \nu(X_u) dW_u ds \right)^2 \right] = \\
& = \frac{1}{\Delta_n} \sum_{i=1}^n \mathbb{E} \left[\left(\int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s \frac{\mu(X_{i-1}, \theta_0)}{\sigma_0} g'_j(X_u) \nu(X_u) dW_u ds \right)^2 \right] + \\
& + \frac{2}{\Delta_n} \sum_{1 \leq i < k \leq n} \mathbb{E} \left[\left(\int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s \frac{\mu(X_{i-1}, \theta_0)}{\sigma_0} g'_j(X_u) \nu(X_u) dW_u ds \right) \cdot \right. \\
& \quad \cdot \left. \left(\int_{t_{k-1}}^{t_k} \int_{t_{k-1}}^s \frac{\mu(X_{k-1}, \theta_0)}{\sigma_0} g'_j(X_u) \nu(X_u) dW_u ds \right) \right] = \\
& = \frac{1}{\Delta_n} \sum_{i=1}^n \mathbb{E} \left[\left(\int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s \frac{\mu(X_{i-1}, \theta_0)}{\sigma_0} g'_j(X_u) \nu(X_u) dW_u ds \right)^2 \right] \leq \\
& \leq \frac{1}{\Delta_n} \sum_{i=1}^n h_i \int_{t_{i-1}}^{t_i} \mathbb{E} \left[\left(\int_{t_{i-1}}^s \frac{\mu(X_{i-1}, \theta_0)}{\sigma_0} g'_j(X_u) \nu(X_u) dW_u \right)^2 \right] ds = \\
& = \frac{1}{\Delta_n} \sum_{i=1}^n h_i \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s \mathbb{E} \left[\left(\frac{\mu(X_{i-1}, \theta_0)}{\sigma_0} g'_j(X_u) \nu(X_u) \right)^2 \right] dud s \leq \\
& \leq \frac{k_2^2}{\Delta_n} \sum_{i=1}^n h_i^3 \leq k_2^2 \Delta_n T.
\end{aligned}$$

For (6.9) there exists constant k_7 such that

$$\left\| \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \frac{1}{\sigma_0} \left(\int_{t_{i-1}}^s f'_j(X_u) \nu(X_u) dW_u \right) ds \right\|_2 \leq k_7 \sqrt{\Delta_n} \sqrt{T},$$

from which follows that (6.9) converges in probability to zero. For (6.10) there exists constant k_8 such that

$$\begin{aligned} & \left\| \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \frac{\nu(X_s)}{\sigma_0} \left(\int_{t_{i-1}}^s (g'_j(X_u) \mu(X_u, \theta_0) + \frac{1}{2} g''_j(X_u) \sigma_0^2 \nu^2(X_u)) du \right) dW_s \right\|_2 \leq \\ & \leq k_8 \sqrt{\Delta_n} \sqrt{T}, \end{aligned}$$

from which we conclude that (6.10) converges in probability to zero. We can conclude that $Z_n(\theta)$ converges in probability to zero. \square

Proof of Lemma 5.3. Since

$$\begin{aligned} & \left\{ \sup_{v \in [0,1]} |(D^2 L_T(\hat{\theta}) - D^2 L_T(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta})))_{ij}| < \epsilon \right\} \subseteq \\ & \subseteq \left\{ \left| \int_0^1 (D^2 L_T(\hat{\theta}) - D^2 L_T(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta})))_{ij} dv \right| < \epsilon \right\}, \end{aligned}$$

for $i, j = 1, \dots, d$ it is sufficient to prove that

$$\sup_{v \in [0,1]} \|D^2 L_T(\hat{\theta}) - D^2 L_T(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta}))\| \xrightarrow{\mathbb{P}} 0.$$

Let us define the functions

$$F(\theta, v) \equiv F(\omega, \theta, v) := \|D^2 L_T(\omega, \hat{\theta}(\omega)) - D^2 L_T(\omega, \hat{\theta}(\omega) + v(\theta - \hat{\theta}(\omega)))\|.$$

Function F is uniformly continuous on $Cl(\Theta) \times [0, 1]$, for each $\omega \in \Omega$ and the functions $\sup_{v \in [0,1]} F(\theta, v)$ are continuous for each $\omega \in \Omega$.

We want to prove that

$$\sup_{v \in [0,1]} F(\bar{\theta}_n, v) \xrightarrow{\mathbb{P}} 0.$$

If it is not true then there exist $\epsilon, \delta > 0$ such that for all $k \in \mathbb{N}$ exists $n_k \geq k$ such that

$$\mathbb{P}(\omega : \sup_{v \in [0,1]} F(\omega, \bar{\theta}_{n_k}(\omega), v) \geq \epsilon) \geq \delta. \quad (6.17)$$

(4.3) implies $\bar{\theta}_{n_k} \xrightarrow{\mathbb{P}} \hat{\theta}_T$, when $k \rightarrow \infty$, hence there exist subsequence $(\bar{\theta}_{n_{k_j}})$ of the sequence $(\bar{\theta}_{n_k})$ such that

$$\bar{\theta}_{n_{k_j}} \xrightarrow{a.s.} \hat{\theta}, \quad j \rightarrow \infty.$$

Let us denote by Ω_0 the event such that $\mathbb{P}(\Omega_0) = 1$ and for all $\omega \in \Omega_0$ holds $\bar{\theta}_{n_{k_j}}(\omega) \rightarrow \hat{\theta}(\omega), j \rightarrow \infty$. Let $\omega \in \Omega_0$. Then, because of continuity of the function $\sup_{v \in [0,1]} F(\omega, \cdot, v)$, follows

$$\sup_{v \in [0,1]} F(\omega, \bar{\theta}_{n_{k_j}}(\omega), v) \rightarrow \sup_{v \in [0,1]} F(\omega, \hat{\theta}(\omega), v) = 0, \quad j \rightarrow \infty,$$

from which we conclude that $\sup_{v \in [0,1]} F(\bar{\theta}_{n_{k_j}}, v)$ converges almost surely to 0, when $j \rightarrow \infty$, hence it converges in probability to 0, which is contradiction with (6.17). \square

Proof of Lemma 5.4. For the sequence $D^2 L_T(\hat{\theta}) \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta})$ holds

$$\begin{aligned} D^2 L_T(\hat{\theta}) \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}) &= \\ &= \int_0^1 (D^2 L_T(\hat{\theta}) - D^2 L_T(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta}))) dv \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}) + \\ &+ \int_0^1 (D^2 L_T(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta})) - D^2 L_n(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta}))) dv \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}) + \\ &+ A_{\hat{\theta}, \bar{\theta}_n} \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}). \end{aligned}$$

Sequence $\frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta})$ is bounded in probability, so by Lemma 5.3 follows

$$\int_0^1 (D^2 L_T(\hat{\theta}) - D^2 L_T(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta}))) dv \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}) \xrightarrow{\mathbb{P}} 0.$$

(4.6) yields $\sup_{v \in [0,1]} \|(D^2 L_T(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta})) - D^2 L_n(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta}))\| \xrightarrow{\mathbb{P}} 0$. Hence, we can conclude that

$$\int_0^1 (D^2 L_T(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta})) - D^2 L_n(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta}))) dv \xrightarrow{\mathbb{P}} 0,$$

which implies

$$\int_0^1 (D^2 L_T(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta})) - D^2 L_n(\hat{\theta} + v(\bar{\theta}_n - \hat{\theta}))) dv \frac{1}{\sqrt{\Delta_n}}(\bar{\theta}_n - \hat{\theta}) \xrightarrow{\mathbb{P}} 0.$$

Lemma 5.2 and (Jacod, J., 1997, (2.2)) yields the conclusion. \square

Proof of Lemma 5.7. It is sufficient to prove that Lemma holds for each jk component of matrices. Since

$$\begin{aligned} \Sigma_n(\bar{\theta}_n)^{jk} &= \Sigma_n(\bar{\theta}_n)^{jk} - \Sigma(\bar{\theta}_n)^{jk} + \\ &+ \Sigma(\bar{\theta}_n)^{jk} - \Sigma(\hat{\theta})^{jk} + \\ &+ \Sigma(\hat{\theta})^{jk}, \end{aligned}$$

it is sufficient to prove that $\Sigma_n(\bar{\theta}_n)^{jk} - \Sigma(\bar{\theta}_n)^{jk} \xrightarrow{\mathbb{P}} 0$ and $\Sigma(\bar{\theta}_n)^{jk} - \Sigma(\hat{\theta})^{jk} \xrightarrow{\mathbb{P}} 0$. Let $\omega \in \Omega$ be fixed. Then the functions $\bar{f}_{jk} : [0, T] \rightarrow \mathbb{R}$, defined by $\bar{f}_{jk}(t) := f_{jk}(X_t(\omega), \bar{\theta}_n(\omega))$, are uniformly continuous. Let $\epsilon > 0$ be arbitrary fixed number. Then, there exists $\delta > 0$, such that $|s - t| < \delta \Rightarrow |f_{jk}(X_s, \bar{\theta}_n) - f_{jk}(X_t, \bar{\theta}_n)| < \frac{\epsilon}{T}$. Let $n_0 \in \mathbb{N}$, such that for all $n \geq n_0$ holds $\Delta_n < \delta$. Then

$$\begin{aligned} |\Sigma_n(\bar{\theta}_n)^{jk} - \Sigma(\bar{\theta}_n)^{jk}| &= \sum_{i=1}^n \int_{t_{i-1}}^{t_i} (f_{jk}(X_{t_{i-1}}, \bar{\theta}_n) - f_{jk}(X_s, \bar{\theta}_n)) ds < \\ &< \sum_{i=1}^n \frac{\epsilon}{T} \Delta_n = \epsilon, \end{aligned}$$

so $\Sigma_n(\bar{\theta}_n)^{jk} - \Sigma(\bar{\theta}_n)^{jk} \xrightarrow{a.s.} 0$, which implies $\Sigma_n(\bar{\theta}_n)^{jk} - \Sigma(\bar{\theta}_n)^{jk} \xrightarrow{\mathbb{P}} 0$. For fixed $\omega \in \Omega$ and $s \in [0, T]$ let $\hat{f}_{jk,s}: \Theta \rightarrow \mathbb{R}$ be the function defined by $\hat{f}_{jk,s}(\theta) = f_{jk}(X_s(\omega), \theta)$. (A3) yields that $\hat{f}_{jk,s} \in C^1(\Theta)$, for all $\omega \in \Omega$, hence

$$|\hat{f}_{jk,s}(\bar{\theta}_n) - \hat{f}_{jk,s}(\hat{\theta})| \leq \sum_{l=1}^d \sup_{\theta \in \Theta} |\partial_{\theta_l} \hat{f}_{jk,s}(\theta)| \|\bar{\theta}_n - \hat{\theta}\|.$$

If we assume that all of the functions are bounded, then there exists constant $M > 0$ such that

$$|\hat{f}_{jk,s}(\bar{\theta}_n) - \hat{f}_{jk,s}(\hat{\theta})| \leq M \|\bar{\theta}_n - \hat{\theta}\|.$$

Let $\epsilon > 0$ be arbitrary, fixed number. Let $\delta > 0$. Then, due to (4.3), there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ holds

$$\mathbb{P}(\|\bar{\theta}_n - \hat{\theta}\| \geq \frac{\delta}{MT}) < \epsilon.$$

For $n \geq n_0$ let us define the set $E_n := \{\|\bar{\theta}_n - \hat{\theta}\| < \frac{\delta}{MT}\}$. Let $\omega \in E_n$. Then

$$|\hat{f}_{jk,s}(\bar{\theta}_n) - \hat{f}_{jk,s}(\hat{\theta})| = |f_{jk}(X_s, \bar{\theta}_n) - f_{jk}(X_s, \hat{\theta})| < M \frac{\delta}{MT} = \frac{\delta}{T},$$

for all $s \in [0, T]$, which implies

$$|\int_0^T (f_{jk}(X_s, \bar{\theta}_n) - f_{jk}(X_s, \hat{\theta})) ds| < \delta,$$

hence

$$\mathbb{P}(|\int_0^T (f_{jk}(X_s, \bar{\theta}_n) - f_{jk}(X_s, \hat{\theta})) ds| \geq \delta) \leq \mathbb{P}(E_n^c) < \epsilon,$$

which proves that $\Sigma(\bar{\theta}_n)^{jk} - \Sigma(\hat{\theta})^{jk} \xrightarrow{\mathbb{P}} 0$. For general case convergence in probability can be proved by defining the stopping times T_M , as seen before. \square

Proof of Lemma 5.8. Let $\epsilon > 0$ such that $\lambda_\Sigma - \epsilon > 0$. Lemma 5.7 implies $\lambda_1(\Sigma_n(\bar{\theta}_n)) - \lambda_1(\Sigma(\hat{\theta})) \xrightarrow{\mathbb{P}} 0$. Since $\Sigma_n(\bar{\theta}_n)$ is symmetric real matrix, for all $\omega \in \Omega$, it follows that $\mathbb{P}(\Sigma_n(\bar{\theta}_n) \text{ is regular matrix}) = \mathbb{P}(\lambda_1(\Sigma_n(\bar{\theta}_n)) > 0)$. Since

$$\{|\lambda_1(\Sigma_n(\bar{\theta}_n)) - \lambda_1(\Sigma(\hat{\theta}))| < \epsilon\} \subseteq \{\lambda_\Sigma - \epsilon < \lambda_1(\Sigma_n(\bar{\theta}_n))\} \subseteq \{0 < \lambda_1(\Sigma_n(\bar{\theta}_n))\},$$

we can conclude that $\mathbb{P}(\lambda_1(\Sigma_n(\bar{\theta}_n)) > 0) \xrightarrow{\mathbb{P}} 1$, which proves our Lemma. \square

Address

Snježana Lubura Strunjak, Department of Mathematics, University of Zagreb,
Faculty of Science, Bijenička cesta 30, 10000 Zagreb, Croatia,
snjezana.lubura.strunjak@math.hr

Acknowledgements

The authors would like to thank Professor Jean Jacod for many helpful comments and suggestions.

References

- Aldous, D. J. and Eagleson, G. K. (1978) On mixing and stability of limit theorems, *The Annals of Probability*, Vol. 6, No. 2, 325-331
- Billingsley, P. (1999) *Convergence of probability measures*, 2nd edn. New York: Wiley and Sons.
- Brown, B. M. and Hewitt, J. I. (1975). Asymptotic likelihood theory for diffusion processes, *J. Appl. Prob.*, 12 228-238.
- Dacunha-Castelle, D. and Florens-Zmirou, D. (1986). Estimation of the coefficients of a diffusion from discrete observations, *Stochastics*, 19 263-284.
- Feigin, P. D. (1976). Maximum likelihood estimation for continuous-time stochastic processes. *Adv. Appl. Prob.*, 8 712-736.
- Florens-Zmirou, D. (1989). Approximate discrete time schemes for statistics of diffusion processes, *Statistics : A Journal of Theoretical and Applied Statistics*, 20 547-557.
- Huzak, M. (2001). A general theorem on approximate maximum likelihood estimation, *Glasnik matematički*, 36(56), 139-153. (hrcak.srce.hr/file/7900)
- Huzak, M. (1998). Parameter estimation of diffusion models, *Glasnik matematički*, Vol.3 No.1, 129-134. (hrcak.srce.hr/1797)
- Huzak, M. (2017). Estimating a class of diffusions from discrete observations via approximate maximum likelihood method, *Statistics* (accepted for publication).
- Huzak, M. (1997). *Selection of diffusion growth process and parameter estimation from discrete observation*, Ph.D. thesis, University of Zagreb, (in Croatian).
- Jacod, J. (1997). On continuous conditional Gaussian martingales and stable convergence in law, *Séminaire de probabilités (Strasbourg)*, tome 31, p. 232-246.
- Jacod, J. and Protter, P. (2012). *Discretization of processes*, Berlin: Springer-Verlag.
- Jacod, J. and Shiryaev, A. N. (2003). *Limit theorems for stochastic processes*, 2nd edn. New York: Springer-Verlag.
- Kessler, M. (1997). Estimation of an Ergodic Diffusion from Discrete Observations, *Scand. J. Statist.*, 24 211-229.
- LeBreton, A. (1976). On continuous and discrete sampling for parameter estimation in diffusion type processes, *Math. Prog. Study*, 5 124-144.
- Liptser, R. S. and Shiryaev, A. N. (1977). *Statistics of random processes II, Applications*, 1st edn. New York: Springer-Verlag.
- Protter, P. (2004). *Stochastic integration and differential equations*, 2nd edn. Berlin: Springer-Verlag.

Rényi, A. (1963). On stable sequences of events, *The Indian Journal of Statistics, Series A (1961-2002)*, Vol. 25, No. 3, pp. 293-302.

Revuz, D. and Yor, M. (1999). *Continuous martingales and Brownian motion*, 3rd edn. Berlin: Springer-Verlag.

Rogers, L. C. G. and Williams, D. (1987). *Diffusion, Markov Processes, and Martingales, Vol. 2: Itô Calculus*, Wiley, Chichester.

Table 1: Example 1., $M=100$, $\alpha = 0.05$

k	2	4	6	8	10	12	14	16
%	0.51	0.55	0.82	0.84	0.90	0.97	0.95	0.96

Table 2: Example 1., $M=100$, $\alpha = 0.01$

k	2	4	6	8	10	12	14	16
%	0.52	0.72	0.83	0.91	0.95	0.98	1	0.97

Table 3: Example 1., $M=1000$, $\alpha = 0.05$

k	2	4	6	8	10	12	14	16
%	0.541	0.670	0.825	0.918	0.976	0.989	0.987	0.988

Table 4: Example 2., $M=100$, $\alpha=0.025$

l	4	5	6	8	10	10	10	10	12	13
k	2	2	2	2	2	4	6	8	6	6
%	0.26	0.26	0.17	0.29	0.28	0.65	0.93	0.99	0.9	0.9
l	13	13	14	14	14	14	15	15	15	16
k	7	8	7	8	9	10	9	10	11	8
%	0.96	0.97	0.91	0.93	0.98	0.99	0.98	0.98	0.97	0.98

Table 5: Example 2., $M=100$, $\alpha=0.05$

l	4	5	6	8	10	10	10	10	12	13	13
k	2	2	2	2	2	4	6	8	6	6	7
%	0.21	0.23	0.14	0.22	0.25	0.53	0.90	0.99	0.84	0.88	0.93
l	13	14	14	14	14	15	15	15	16	16	16
k	8	7	8	9	10	9	10	11	8	10	12
%	0.90	0.91	0.92	0.96	0.98	0.95	0.98	0.95	0.94	0.93	0.94

Table 6: Example 2., $M=100$, $\alpha=0.05$, $l=12$

k	4	5	6	7	8	9	10	11
%	0.66	0.75	0.83	0.92	0.94	0.94	0.95	1

Table 7: Example 2., $M=100$, $\alpha=0.05$, $l=14$

k	4	5	6	7	8	9	10	11	12	13
%	0.67	0.75	0.85	0.85	0.95	0.94	0.96	0.99	1	1

Table 8: Example 2., $M=100$, $\alpha=0.05$, $l=16$

k	4	5	6	7	8	9	10	11	12	13	14	15
%	0.60	0.67	0.86	0.85	0.94	0.97	0.94	0.94	0.94	0.95	0.98	0.99