Probab. Theory Relat. Fields (2008) 142:1–20
DOI 10.1007/s00440-007-0097-7

Edgeworth expansion for ergodic diffusions

Masaaki Fukasawa

Received: 5 March 2006 / Revised: 20 June 2007 / Published online: 3 August 2007

Abstract The Edgeworth expansion for an additive functional of an ergodic diffusion is validated under fairly weak conditions. The validation procedure does not depend on the stationarity or the geometric mixing property, but exploits the strong Markov property of the process. In particular for an Itô-diffusion of dimension one, verifiable conditions for the validity of the expansion are given in terms of the coefficients of the corresponding stochastic differential equation. The maximum likelihood estimator for the CIR process is treated as example.

Mathematics Subject Classification (2000) 60F05 · 60J25 · 60J55 · 62E20

1 Introduction

The Edgeworth expansion (EE) is a refinement of the central limit theorem and has played an important role in statistics. Bhattacharya and Rao [5] and Hall [13] investigated the EE for independently and identically distributed (IID) sequences in detail. For weakly dependent sequences, Götze and Hipp [12] validated the EE by assuming the geometric mixing property. Their argument was extended to continuous-time processes by Kusuoka and Yoshida [16] and Yoshida [25]. Their methods based on mixing properties are, however, less direct for Markov processes in a sense than the so-called regenerative method; Malinovskii [17] and Jensen [14] validated the EE for a wider class of strongly dependent Markov chains by the regenerative method (see Bolthausen [6,7], Bertail and Clémençon [1]).

The present article validates the EE for continuous-time strong Markov processes by extending the argument of Malinovskii [17]. Let $X = (X_t)_{t\geq 0}$ be such a process

M. Fukasawa (✉)
Graduate School of Mathematical Sciences, The University of Tokyo,
3-8-1 Komaba, Meguro, Tokyo 153-8914, Japan
e-mail: fukasawa@ms.u-tokyo.ac.jp

and $Z = (Z_t)_{t\geq 0}$ be an n-dimensional additive functional of X. Denote by P_ν the law of X with initial distribution ν. We prove that for a given sufficiently smooth function $A : \mathbb{R}^n \to \mathbb{R}$,

$$\sup_{B\in\mathcal{B}(\mathbb{R})}\left| P_\nu\left[\sqrt{T}(A(Z_T/T) - A(\gamma))/\sigma \in B\right] - \int_B p_T(z)dz\right| = O(T^{-1}) \quad (1)$$

as $T \to \infty$ where $\mathcal{B}(\mathbb{R})$ is the Borel σ-field of $\mathbb{R}$,

$$p_T(z) = \phi(z)\left\{1 + T^{-1/2}\left\{\frac{A_1}{\sigma}z + \frac{A_2}{6\sigma^3}(z^3 - 3z)\right\}\right\}, \quad (2)$$

ϕ is the density of the standard normal distribution and $\gamma, \sigma, A_j, j = 1, 2$ are constants written explicitly in terms of moments of certain distributions depending on X and Z. As an application, we consider the case that X is the weak solution of the stochastic differential equation (SDE)

$$dX_t = b(X_t)dt + c(X_t)dB_t, \quad X_0 \sim \nu \quad (3)$$

where $B = (B_t)$ is a standard Brownian motion. Sufficient conditions for the validity of the EE (1) are presented in terms of the coefficients b and c. Note that we can also derive the validity of the asymptotic expansions of various estimators such as the maximum likelihood estimator in the context of parametric inference for b based on X by means of the delta method (see e.g., Hall [13], Sect. 2.7).

As mentioned above, Kusuoka and Yoshida [16] and Yoshida [25] established the EE for continuous-time processes with a mixing property. Although their framework itself is widely applicable to various contexts, the coefficients of their expansion are written in terms of the cumulants of $T^{-1/2}Z_T$, which hampers the practical use of their results. In contrast, the coefficients γ, σ and A_j in the formula (2) are independent of T; the effect of the initial distribution ν is specified in the definition of the coefficient A_1. Note also that our results can be applied to a wider class of one-dimensional diffusions including non-stationary diffusions, diffusions whose α-mixing coefficients decay slowly, and Itô-diffusions whose coefficients b and c (in (3)) are not smooth.

The paper is organized as follows. In Sect. 2, we introduce notation and state our basic result. Section 3 gives sufficient conditions for Itô-diffusions and includes the expansion formula of the maximum likelihood estimator of the CIR processes. The proof of the basic result and several remarks are in Sect. 4.

2 Notation and results

Here we give a rigorous formulation. We suppose that X is a càdlàg strong Markov process with state space E and infinite lifetime. We also suppose without loss of generality that X is canonical, i.e., $X_t(\omega) = \omega(t)$ for $t \geq 0$, $\omega \in \mathcal{D}$ where $\mathcal{D}$ is the set of the càdlàg mappings from $[0, \infty)$ to E. Recall that P_ν is the law of X with $X_0 \sim \nu$, where ν is a given distribution on E. Let P_x stand for P_{δ_x} for $x \in E$. Throughout this paper, the filtration is supposed to be the canonical one of X.

Definition 1 Let T be a given finite stopping time. The shift operator θ_T is a mapping $\theta_T : \mathcal{D} \to \mathcal{D}$ defined as

$$\theta_T(\omega)(t) = \omega(t + T(\omega)), \qquad t \in [0, \infty), \quad \omega \in \mathcal{D}.$$

For example, we have $X_t \circ \theta_T = X_{t+T}$ for deterministic $t \geq 0$, and for the hitting time H_x defined as $H_x = \inf\{t \geq 0; X_t = x\}$, $x \in E$, we have

$$H_x \circ \theta_T = \inf\{t \geq 0; X_{t+T} = x\} = \inf\{t \geq T; X_t = x\} - T.$$

Moreover, it holds $H_{x,y} := H_y + H_x \circ \theta_{H_y} = \inf\{t \geq H_y; X_t = x\}$ for any $x, y \in E$ if H_y is a.s. finite.

Definition 2 An adapted process Z is said to be an additive functional if it holds

$$Z_{T+S\circ\theta_T} = Z_T + Z_S \circ \theta_T$$

for any finite stopping times S and T.

An example of additive functionals is $Z_t = u(X_t) - u(X_0)$ where u is a given measurable function $u\colon E \to \mathbb{R}$. Another example is $Z_t = \int_0^t u(X_t)dt$.

Now, fix $n \in \mathbb{N}$ and n-dimensional additive functional Z. Let us introduce an assumption which we will refer as (A1-s, x, y) for $s \in \mathbb{N}$, $x, y \in E$, $x \neq y$.

(A1-s, x, y): *Assume*

$$P_\nu[H_{x,y}^2] + P_\nu[|Z_{H_{x,y}}|^2] + P_x[H_{x,y}^s] + P_x[|Z_{H_{x,y}}|^s] + P_x\left[\int_0^{H_{x,y}} |Z_t|^2 dt\right]$$

to be finite.

Since $H_{x,y}$ corresponds to the so-called regeneration time in the context of the regenerative method, this condition is related to the ergodicity of X. Here we used another point y in order to define the regeneration time which is a.s. positive. Hereafter we at least assume that there exist x, y such that (A1-2, x, y) holds. Then, we can put $\alpha_{x,y} = P_x[H_{x,y}] > 0$, $\beta_{x,y} = P_x[Z_{H_{x,y}}] \in \mathbb{R}^n$ and $\bar{Z}_{x,y} = Z_{H_{x,y}} - H_{x,y}\beta_{x,y}/\alpha_{x,y}$. Assume without loss of generality that there exists $0 \leq n' \leq n$ such that $\bar{Z}_{x,y}$ has the form $\bar{Z}_{x,y} = (G_{x,y}, N_{x,y})$, where $G_{x,y}$ is an n'-dimensional random vector which has no P_x-a.s. zero components and $N_{x,y}$ is an $(n - n')$-dimensional random vector with $N_{x,y} = \mathbf{0}$, P_x-a.s.. Then, let $\mathcal{Z}_{x,y}$ stand for the $(n' + 1)$-dim random vector $(G_{x,y}, H_{x,y})$. We will assume

(A2-x, y): *There exists* $p \geq 1$ *such that*

$$\int_{\mathbb{R}^{n'+1}} |P_x[\exp\{u \cdot \mathcal{Z}_{x,y}\}]|^p du < \infty.$$

(A2-x, y) requires the law of $\mathcal{Z}_{x,y}$ to be smooth, which is related to the reason why we eliminated the P_x-degenerate component $N_{x,y}$ when we defined $\mathcal{Z}_{x,y}$.

Now, we would like to describe the coefficients of the EE. Denote by $\Sigma^0_{x,y}$ the covariance matrix of $G_{x,y}$. Put $\gamma = (\gamma_j)_{1\leq j\leq n} = \beta_{x,y}/\alpha_{x,y}$,

$$\Gamma = (\gamma_{i,j})_{1\leq i,j,\leq n'} = \Sigma^0_{x,y}/\alpha_{x,y} \quad \text{and} \quad \rho = (\rho_j)_{1\leq j\leq n'} = P_x[G_{x,y}H_{x,y}].$$

Here we dropped the suffix 'x, y' in the definitions of γ, Γ and ρ for short. Note that γ and Γ actually do not depend on x, y because they are nothing but the asymptotic mean and variance of Z_T respectively in the light of (1), while ρ does. For a given function $A : \mathbb{R}^n \to \mathbb{R}$, assume that A is four times continuously differentiable in a neighborhood of γ. Put for $1 \leq i, j, \leq n$,

$$a_j = \partial_j A(\gamma), \quad a_{i,j} = \partial_i\partial_j A(\gamma), \quad \sigma = \sqrt{\sum_{i,j=1}^{n'} a_i a_j \gamma_{i,j}} \tag{4}$$

and $a = (a_j) \in \mathbb{R}^n$.

Theorem 1 *Assume (A1-*$(n'+2)\vee 4, x, y$*), (A2-*x, y*) and* $\sigma > 0$*. Then, there exists a unique stationary distribution* μ *of* X *and* $P_\mu[|\bar{Z}_{x,y}|] < \infty$*. Moreover, it holds*

$$\sup_{B\in\mathcal{B}(\mathbb{R})} \left| P_\nu\left[\sqrt{T}(A(Z_T/T) - A(\gamma))/\sigma \in B\right] - \int_B p_T(z)dz \right| = O(T^{-1})$$

as $T \to \infty$ *where* $\mathcal{B}(\mathbb{R})$ *is the Borel* σ*-field of* $\mathbb{R}$*,*

$$p_T(z) = \phi(z)\left\{1 + T^{-1/2}\left\{\frac{A_1}{\sigma}z + \frac{A_2}{6\sigma^3}(z^3 - 3z)\right\}\right\},$$

ϕ *is the density of the standard normal distribution,*

$$\begin{aligned} A_1 &= a\cdot\left(P_\nu[\bar{Z}_{x,y}] - P_\mu[\bar{Z}_{x,y}]\right) + \frac{1}{2}\sum_{i,j=1}^{n'} a_{i,j}\gamma_{i,j}, \\ A_2 &= \sum_{i,j,k=1}^{n'} a_i a_j a_k \gamma_{i,j,k} + 3\sum_{i,j,k,l=1}^{n'} a_i a_j a_{k,l}\gamma_{i,k}\gamma_{j,l}, \\ \gamma_{i,j,k} &= (\kappa_{i,j,k} - \rho_i\gamma_{j,k} - \rho_j\gamma_{k,i} - \rho_k\gamma_{i,j})/\alpha_{x,y} \end{aligned} \tag{5}$$

and $\kappa_{i,j,k}$, $1 \leq i, j, k \leq n'$ *are the third cumulants of* $G_{x,y}$*.*

Proof See Sect. 4. □

3 Itô-diffusions

3.1 Diffusions on the whole line

This section presents several sufficient conditions for the assumptions of Theorem 1 to hold. Here we suppose that $E = \mathbb{R}$ and that P_ν is the law of the weak solution of the stochastic differential equation (3). Assume that b is locally integrable on $\mathbb{R}$ and that both c and $1/c$ are locally bounded on $\mathbb{R}$. Define the scale function s as

$$s(u) = \int_0^u \exp\left\{-2\int_0^v \frac{b(w)}{c(w)^2}dw\right\} dv,$$

and the speed measure m as

$$m(du) = m(u)du = \frac{1}{c(u)^2} \exp\left\{2\int_0^u \frac{b(w)}{c(w)^2}dw\right\} du.$$

Assume $s(\mathbb{R}) = \mathbb{R}$ and $m(\mathbb{R}) < \infty$ in order to assure the ergodicity (see e.g., Skorokhod [23], Theorem 16). It holds then $P_x[H_y] < \infty$ for all $x, y \in \mathbb{R}$ (see e.g., Gikhman and Skorokhod [11], Sect. 18). Let $Z^{(1)}$ be an n'-dimensional additive functional of the form

$$Z_T^{(1)} = \int_0^T f(X_t)dt \tag{6}$$

where $f = (f_1, \ldots, f_{n'})$ is a Borel function on $\mathbb{R}$ to $\mathbb{R}^{n'}$. We will consider an n-dimensional additive functional $Z = (Z^{(1)}, Z^{(2)})$ where $n' \leq n$, $Z_\cdot^{(2)} = u(X_\cdot) - u(X_0)$ for a Borel function $u : \mathbb{R} \to \mathbb{R}^{n-n'}$. The reason why we focus on additive functionals of this form is that many of statistics are smooth functions of some of them in the context of parametric inference for b. Note that $\mu :\ = m/m(\mathbb{R})$ is the stationary distribution of X and that $\lim_{T\to\infty} Z_T/T = \gamma = (\mu[f], \mathbf{0}), \mathbf{0} \in \mathbb{R}^{n-n'}$ if f is μ-integrable (see e.g., Skorokhod [23], Theorems 16 and 18).

Let $\tilde{c}(z) = c(s^{-1}(z))s'(s^{-1}(z))$ and

$$\psi_\omega(t) = \int_0^t \frac{ds}{\tilde{c}(\omega(s))^2}, \quad \psi_\omega^-(a) = \inf\{t \geq 0; \psi_\omega(t) > a\}$$

for $\omega \in \mathcal{D}$. The following argument relies on the fact that $P_u = Q_{s(u)}^{\hat{X}}$ where Q_v is the law of a Brownian motion starting at v and $\hat{X} : \mathcal{D} \to \mathcal{D}$ is defined as

$$\hat{X}(\omega) = s^{-1}(\omega(\psi_\omega^-(\cdot))).$$

See Karatzas and Shreve [15], Sect. 5.5 for details.

Lemma 1 **(Borisov [8])** *Consider the case $n' = 1$ in (6). Let $\xi < \eta \in \mathbb{R}$ and $\delta > 0$, $K > k > 0$ be constants. Let Ψ be the characteristic function of $Z^{(1)}_{H_\eta}$ under Q_ξ. Then, for all $d > 0$ there exists $B > 0$ which depending only d, δ, k, and K such that*

$$|\Psi(u)| \le \exp\left\{-B\sqrt{|u|}\right\} \tag{7}$$

for all u with $|u| > d$ and all f with the following property:

1. $|f(\zeta)| \le K$ *for all* $\zeta \in [2x - y, y]$.
2. *There exists an interval* $\Delta \subset [x, y]$ *such that* $|\Delta| = \delta$ *and* $|f| \ge k$ *on* Δ.
3. $Q_\xi[|Z^{(1)}_{H_\eta}| < \infty] = 1$.

Proof See the proof of Borisov [8], Lemma 3. □

Lemma 2 *Assume that there exist two distinct points $x, y \in \mathbb{R}$ such that the following conditions hold for the intervals $\Delta_0 = (x \wedge y, x \vee y)$ and $\Delta_1 = s^{-1}([s(x) - |s(\Delta_0)|, s(x) + |s(\Delta_0)|])$.*

1. f *is bounded on* Δ_1,
2. f *is continuous on* Δ_0,
3. $\{1, f_1, \dots, f_{n'}\}$ *is linearly independent as a subset of* $C(\Delta_0)$,
4. $P_x[|Z^{(1)}_{H_y}| < \infty] = 1$.

Let Ψ be the characteristic function of the $(n' + 1)$-dimensional random vector $(Z^{(1)}_{H_y}, H_y)$ under P_x. Then, for all $d > 0$, there exists $B > 0$ such that the inequality (7) holds for all $u \in \mathbb{R}^{n'+1}$ with $|u| > d$.

Proof The assertion essentially results from Lemma 1 which treats the case that X is a Brownian motion and $n' = 1$. In fact, the law of $(Z^{(1)}_{H_y}, H_y)$ under P_x equals to that of

$$\left(\int_0^{\hat{H}_y} f(\hat{X}_t)dt, \hat{H}_y\right) = \left(\int_0^{H_{s(y)}} \frac{f(s^{-1}(X_a))}{\tilde{c}(X_a)^2}da, \int_0^{H_{s(y)}} \frac{1}{\tilde{c}(X_a)^2}da\right)$$

under $Q_{s(x)}$ where $\hat{H}_y = \inf\{t \ge 0; \hat{X}_t = y\}$. Let the inner product $F(v, \xi) = v \cdot (f(s^{-1}(\xi))/\tilde{c}(\xi), 1)$ for $v \in \mathbb{R}^{n'+1}$ with $|v| = 1$. It is trivial that $K := \sup_{|v|=1, \xi \in s(\Delta_1)} |F(v, \xi)| < \infty$ and that there exists $\Delta_v \subset s(\Delta_0)$ such that $k(v) := \inf_{\xi \in \Delta_v} |F(v, \xi)| > 0$. Lemma 1 implies that for all $d > 0$, there exists $B_v > 0$ such that $|\Psi(tv)| \le \exp\{-B_v\sqrt{t}\}$ for all $t > d$. Let $k(u, v) = \inf_{\xi \in \Delta_v} |F(u, \xi)|$. Since the open covering $\bigcup_v \{u; k(u, v) > k(v)/2\}$ of the compact set $\{v; |v| = 1\}$ has a finite subcovering, the result follows. □

Now, put $N = (n' + 2) \vee 4$ and

$$r = -\limsup_{|z| \to \infty} \frac{zb(z)}{c(z)^2} \in [-\infty, \infty].$$

Theorem 2 *Assume that for constants $C, k, p \in \mathbb{R}$ and an interval I,*

1. $\{|f(z)| \vee 1\} z^2 c(z)^{-2} \le C|1+|z||^p$, $|u(z)| \le C|1+|z||^k$ *for all* $z \in \mathbb{R}$,
2. $2r+1 > \{Np\} \vee p \vee \{2k+p\}$,
3. $\{1, f_1, \ldots, f_{n'}\}$ *is continuous and linearly independent on* I,
4. ν *has the* $2\{k \vee p \vee 0\}$*-th moments.*

Then, (*A1-*N, x, y) *and* (*A2-*x, y) *hold for all* $x, y \in I$.

Proof Let us prove first that (A1-s, x, y) holds for all $x, y \in \mathbb{R}$ and $s \le N$. Put $g = |f| \vee 1$. By a similar argument to Gikhman and Skorokhod [11], Chap. 3, Sect. 15, Theorem 2 and Sect. 18, Lemma 1, we have for any positive function ψ that

$$P_x\left[\int_0^{H_y} \psi(X_t) g(X_t) dt\right] = 2\int_x^y (s(y)-s(u))\psi(u)g(u)m(du) + 2(s(y)-s(x))\int_{-\infty}^x \psi(u)g(u)m(du) \tag{8}$$

in the case $x < y$ and

$$P_x\left[\int_0^{H_y} \psi(X_t) g(X_t) dt\right] = 2(s(x)-s(y))\int_x^\infty \psi(u)g(u)m(du) + 2\int_y^x (s(u)-s(y))\psi(u)g(u)m(du) \tag{9}$$

in the case $x > y$. It is straightforward to show that that if ψ is of polynomial growth of order q and $2r+1 > q+p$, then the right hand sides of (8) and (9) are of polynomial growth of order $q+p$ as $x \to -\infty$, $x \to \infty$ respectively. In the light of Kac's moment formula [9], the iterative use of this fact leads to that

$$x \mapsto P_x\left[\left\{\int_0^{H_y} g(X_t) dt\right\}^s\right]$$

is of polynomial growth of order sp as long as $sp < 2r+1$.

Fix $x, y \in I$ arbitrarily and then define $\mathcal{Z}_{x,y}$ as before. It is now straightforward to see that (A1-N, x, y) holds by the strong Markov property. (A2-x, y) results from Lemma 2. □

Remark 1 It is possible to describe σ, A_1 and A_2 more explicitly in terms of f, s, m, ν and derivatives of A by calculating (5) with the aid of Kac's moment formula.

Remark 2 The results of Yoshida [25] can be applied to (3) in order to obtain the EE if the corresponding α-mixing coefficient decays faster than the reciprocal of any polynomial. However, it seems to have been unknown whether the EE can be validated in the case, for example, that $b(z) \sim -r/z$ as $|z| \to \infty$, $r \in (0, \infty)$ and $c(z) \equiv 1$, for which Veretennikov [24] presents a polynomial lower bound of the corresponding α-mixing coefficient.

Remark 3 Consider the following parametric model

$$dX_t = b(\theta, X_t)dt + c(X_t)dB_t, \quad X_0 \sim \nu^\theta.$$

Sakamoto and Yoshida [21,22] showed that M-estimators $\hat{\theta}_T$ such as the maximum likelihood estimator of θ admits the stochastic expansion of the form $\sqrt{T}(\hat{\theta}_T - \theta) = \sqrt{T}(A(Z_T/T) - \theta) + O_p(T^{-1})$ under a certain smoothness condition in θ. Thus, the asymptotic expansion of $\hat{\theta}_T$ is validated in the view of Theorem 2 with the aid of the delta method which enables us to control the remainder $O_p(T^{-1})$ under an appropriate moment condition (see e.g., Hall [13], Sect. 2.7). It is now straightforward to prove the parametric bootstrap method to be valid in e.g., constructing second-order correct confidence intervals of those estimators (see Mykland [18] which proved the validity in the sense of a weak topology).

3.2 Diffusions on the half line

In the preceding subsection, we treated Itô-diffusions whose speed measures have the whole line as their supports. The results can be extended to diffusions on the half line $E = \mathbb{R}_+ = (0, \infty)$. Assume in (3) that b is locally integrable on $\mathbb{R}_+$, that both c and $1/c$ are locally bounded on $\mathbb{R}_+$ and that the support of ν is $\mathbb{R}_+$. Let $Z = (Z^{(1)}, Z^{(2)})$ as before where f and u are considered to be functions on $\mathbb{R}_+$. Put

$$\mathcal{G}_1^q = \left\{ f : \mathbb{R}_+ \to \mathbb{R};\ \text{locally bounded Borel},\ \limsup_{z \to 0} z^q |f(z)| < \infty \right\},$$

$$\mathcal{G}_2^p = \left\{ f : \mathbb{R}_+ \to \mathbb{R};\ \text{locally bounded Borel},\ \limsup_{z \to \infty} z^{-p} |f(z)| < \infty \right\}$$

for $p, q \geq 0$ and $\mathcal{G}_2 = \bigcup_{p \geq 0} \mathcal{G}_2^p$.

Theorem 3 *Assume* $|1/c|^2 \vee |u|^3 \vee |f| \in \mathcal{G}_1^1 \cap \mathcal{G}_2$ *and that there exist* $\kappa \in [0, 1)$, $q > 1$ *and an interval* $I \subset \mathbb{R}_+$ *such that*

1. $\limsup_{z \to \infty} z^\kappa b(z) c(z)^{-2} < 0$.
2. $\liminf_{z \to 0} 2z b(z) c(z)^{-2} > q$.
3. $\nu[|\psi|] < \infty$ *for any* $\psi \in \mathcal{G}_1^q \cap \mathcal{G}_2$.
4. $\{1, f_1, \ldots, f_{n'}\}$ *is continuous and linearly independent on* I,

Then, (*A1*-s, x, y) *and* (*A2*-x, y) *hold for all* $x, y \in I$ *and* $s \in \mathbb{N}$.

Proof By using counterparts of (8) and (9) we can prove that if $|f| \in \mathcal{G}_1^q \cap \mathcal{G}_2$, then $P_x[|Z_{H_y}^{(1)}|]$ belongs to $\mathcal{G}_1^{q-1} \cap \mathcal{G}_2$ as a function of x for all $y \in \mathbb{R}_+$. It is then straightforward to see that for all $s \in \mathbb{N}$ and $y \in \mathbb{R}_+$, $x \mapsto P_x[|Z_{H_y}^{(1)}|^s]$ belongs to $\mathcal{G}_1^q \cap \mathcal{G}_2$ by Kac's moment formula. The rest of the proof is clear. □

Example 1 Theorem 3 can be applied to the CIR processes

$$dX_t = -(aX_t - b)dt + \sigma\sqrt{X_t}dB_t, \quad X_0 \sim \nu$$

and the corresponding maximum likelihood estimators $\hat{a}_T$ and $\hat{b}_T$ as long as $2b > \sigma^2$. For example, in the case that b and σ are known, we have

$$P_\nu[\sqrt{IT}(\hat{a}_T - a) \in B] = \int_B \phi(z)\left\{1 + \frac{\sigma(z^3 - z)}{2\sqrt{abT}}\right\}dz + O(T^{-1})$$

uniformly in $B \in \mathcal{B}(\mathbb{R})$ where $I = b/a\sigma^2$.

4 Proof of Theorem 1 and remarks

4.1 Outline of the proof

This subsection gives the outline of the proof of Theorem 1. Detailed calculation are included in the next subsection for interested readers. Let $Z_t^\dagger = Z_t - t\gamma$ for $t \geq 0$ and $A^\dagger(u) = A(u+\gamma) - A(\gamma)$ for $u \in \mathbb{R}^n$. Observe that $A(Z_T/T) - A(\gamma) = A^\dagger(Z^\dagger/T)$ and that the constants $a_j, a_{i,j}, \sigma, \rho_j, \gamma_{i,j}, \kappa_{i,j,k}$ and A_j remain unchanged even if we replace A and Z with $A^\dagger$ and $Z^\dagger$ respectively in their definitions, so that we can assume $\gamma = 0$ and $A(0) = 0$ without loss of generality. The points x, y in (A1-s, x, y) and (A2-x, y) are now fixed, hence we abbreviate $\alpha_{x,y}$ to α and drop the suffix 'x, y' also when introducing other notation. Let $\pi : \mathbb{R}^n \to \mathbb{R}^{n'}$ be the projection onto the first n'-dim subspace and $\iota : \mathbb{R}^{n'} \to \mathbb{R}^n$ be the natural inclusion, that is $\iota(u) = (u, 0, \dots, 0)$ for $u \in \mathbb{R}^{n'}$. Note that $\sigma > 0$ implies $n' \geq 1$.

Step 1 [The regenerative method]: Here we use a variant of the argument used by Bolthausen [6,7] and Malinovskii [17] in order to apply the classical theory of the EE for the IID case. Define a sequence of finite stopping times $\{\tau_j\}$ as

$$\tau_0 = 0, \qquad \tau_j = \tau_{j-1} + H_{x,y} \circ \theta_{\tau_{j-1}}, \quad j \geq 1$$

and put $L_j = \tau_{j+1} - \tau_j$, $\hat{Z}_j = Z_{\tau_{j+1}} - Z_{\tau_j}$. Observe that $(\hat{Z}_j, L_j)$, $j \geq 0$ are IID under P_x by the strong Markov property and the definition of the additive functional. Note that this fact is essential to the following regenerative argument. Define

$$U_m = \frac{1}{\sqrt{m}}\sum_{j=1}^m \pi(\hat{Z}_j), \quad V_m = \frac{1}{\sqrt{m}}\sum_{j=1}^m (L_j - \alpha), \quad a_m = \frac{T - \alpha m}{\sqrt{m}}$$

for $m \geq 1$ and $U_0 = V_0 = a_0 = 0$. Notice that the $(n'+1)$-dim random vector (U_m, V_m) is a sum of IID sequence and that (U_m, V_m) and $(\hat{Z}_0, L_0)$ are independent under P_ν. Put $M_T = \max\left\{M \geq 0; \sum_{m=0}^{M} L_m \leq T\right\}$. It is straightforward to show $P_\nu[|T - \alpha M_T| \geq \delta T] = O(T^{-1})$ for an arbitrarily fixed $\delta \in (0, 1/2)$ by using Chebyshev's inequality and (A1-2, x, y). Now, observe that

$$Z_T = \hat{Z}_0 + \sqrt{M_T}\iota(U_{M_T}) + R_{M_T},$$

where $R_m = R_m(L_0, V_m)$, $R_m(l, \eta) = Z_T - Z_{\hat{T}_m(l,\eta)}$ and $\hat{T}_m(l, \eta) = (l + \sqrt{m}\eta + \alpha m) \wedge T$ for $m \geq 0$. Let $\psi_1(f, \xi, r)$ and $\psi_2(l, \eta, t)$ are the indicator functions of the sets $\{(f, \xi, r) \in \mathbb{R}^n \times \mathbb{R}^{n'} \times \mathbb{R}^n; \sqrt{T}A((f + \sqrt{m}\iota(\xi) + r)/T)/\sigma \in B\}$ and $\{(l, \eta, t) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}; 0 \leq \sqrt{m}(a_m - \eta) - l < t\}$, respectively. Then, $\{M_T = m\} = \{\psi_2(L_0, V_m, L_{m+1}) = 1\}$, so that we have

$$\begin{aligned} &P_\nu[\sqrt{T}A(Z_T/T)/\sigma \in B] \\ &= \sum_{m;|T-\alpha m|<\delta T} P_\nu[\psi_1(\hat{Z}_0, U_m, R_m)\psi_2(L_0, V_m, L_{m+1})] + O(T^{-1}) \end{aligned}$$

uniformly in $B \in \mathcal{B}(\mathbb{R})$. Hereafter, we always drop "uniformly in $B \in \mathcal{B}(\mathbb{R})$" for short whenever stating an identity with $O(T^{-1})$ and $B \in \mathcal{B}(\mathbb{R})$. By the strong Markov property (see Sect. 4.2 for details), we have

$$\begin{aligned} &P_\nu\left[\psi_1(\hat{Z}_0, U_m, R_m)\psi_2(L_0, V_m, L_{m+1})\right] \\ &= \int \psi_1(f, \xi, r)\psi_2(l, \eta, t) \\ &\quad \times P_x^{(\hat{R}_m(l,\eta),\tau_1)}(dr, dt)p_m(\xi, \eta)d\xi d\eta P_\nu^{(\hat{Z}_0, L_0)}(df, dl), \end{aligned} \tag{10}$$

where $\hat{R}_m(l, \eta) = Z_{T-\hat{T}_m(l,\eta)}$ and p_m is a bounded density of the $(n'+1)$-dim random vector (U_m, V_m). Note that the existence of p_m is assured for sufficiently large m by (A2-x, y) (see Bhattacharya and Rao [5], Theorem 19.1). We can now apply Bhattacharya and Rao [5], Theorem 19.2 under (A1-$(n'+2) \vee 4$, x, y) to obtain

$$\sup_{\zeta \in \mathbb{R}^{n'+1}} \left(1 + |\zeta|^{(n'+2)\vee 4}\right) \left|p_m(\zeta) - \phi_\Sigma(\zeta)\left(1 + m^{-1/2}p^\Sigma(\zeta)\right)\right| = O(m^{-1}),$$

where Σ is the covariance matrix of $\mathcal{Z}_{x,y}$ under P_x, ϕ_Σ is the normal density with mean 0, covariance Σ, and p^Σ is a polynomial. Note that (A2-x, y) results in particular to the regularity of the matrix Σ. We have then

$$
\begin{aligned}
&P_\nu\left[\sqrt{T}A(Z_T/T)/\sigma \in B\right] \\
&= \sum_{m;|T-\alpha m|<\delta T} \int \psi_1(f,\xi,r)\psi_2(l,\eta,t)\phi_\Sigma(\xi,\eta)\left(1+m^{-1/2}p^\Sigma(\xi,\eta)\right) \\
&\quad \times P_x^{(\hat{R}_m(l,\eta),\tau_1)}(dr,dt)P_\nu^{(\hat{Z}_0,L_0)}(df,dl)d\xi d\eta + O(T^{-1}) \qquad (11)
\end{aligned}
$$

by estimating the error terms suitably (see Sect. 4.2 for details).

Step 2 [Calculation of the sum over m]: This part also is a variant of the argument of Malinovskii [17]. Here we deal with the sum over m by changing variable and using Taylor's expansion. Let $\psi_T^{f,r}$ be the indicator function of the set

$$
\left\{u \in \mathbb{R}^{n'};\ \frac{\sqrt{T}}{\sigma}A\left(\frac{\iota(u)}{\sqrt{T}}+\frac{f+r}{T}\right)\in B\right\}
$$

for $f,r \in \mathbb{R}^n$ and $\phi^{\Sigma,m} = \phi_\Sigma\{1+m^{-1/2}p^\Sigma\}$, $\phi_\xi^{\Sigma,m} = (\partial_j\phi^{\Sigma,m})_{j=1}^{n'}$, $\phi_\eta^{\Sigma,m} = \partial_{n'+1}\phi^{\Sigma,m}$. Changing variables: $\xi = u\sqrt{T/m}$, $\eta = (v-\alpha m - l)/\sqrt{m}$, and using Taylor's expansion, the summand of (11) turns to

$$
\begin{aligned}
&\frac{1}{\sqrt{m}}\left(\frac{T}{m}\right)^{n'/2}\int \psi_T^{f,r}(u)1_{\{0\le T-v<t\}}(v) \\
&\times\left\{\phi^{\Sigma,m}(\sqrt{\alpha}u,a_m)+\phi_\xi^{\Sigma,m}(\sqrt{\alpha}u,a_m)\cdot\theta_\xi^m+\phi_\eta^{\Sigma,m}(\sqrt{\alpha}u,a_m)\theta_\eta^m + A^m\right\} \\
&\times P_x^{(R(v),\tau_1)}(dr,dt)P_\nu^{(\hat{Z}_0,L_0)}(df,dl)dudv,
\end{aligned}
$$

where $R(\eta) = Z_{T-\eta}$, $\theta_\xi^m = (\sqrt{T/m}-\sqrt{\alpha})u$, $\theta_\eta^m = (v-l-T)/\sqrt{m}$, and A^m is a negligible remainder term in the sense that

$$
\begin{aligned}
&\sum_{|T-\alpha m|<\delta T}\frac{1}{\sqrt{m}}\left(\frac{T}{m}\right)^{n'/2}\int |A^m|1_{\{0\le T-v<t\}} \\
&\quad\times P_x^{(R(v),\tau_1)}(dr,dt)P_\nu^{(\hat{Z}_0,L_0)}(df,dl)dudv = O(T^{-1}). \qquad (12)
\end{aligned}
$$

See Sect. 4.2 for the proof.

It suffices then to cope with the expectations of the following terms with respect to $P(dr, dt, df, dl, dv) := P_x^{(R(v),\tau_1)}(dr, dt) P_v^{(\hat{Z}_0, L_0)}(df, dl) dv$.

$$T_1 = \sum_{|T-\alpha m|<\delta T} \frac{1}{\sqrt{m}} \left(\frac{T}{m}\right)^{n'/2} \int \psi_T^{f,r}(u) 1_{\{0 \le T-v<t\}} \phi^{\Sigma,m}(\sqrt{\alpha}u, a_m) du,$$

$$T_2 = \sum_{|T-\alpha m|<\delta T} \frac{1}{\sqrt{m}} \left(\frac{T}{m}\right)^{n'/2} \int \psi_T^{f,r}(u) 1_{\{0 \le T-v<t\}} \phi_\xi^{\Sigma,m}(\sqrt{\alpha}u, a_m) du \cdot \theta_\xi^m,$$

$$T_3 = \sum_{|T-\alpha m|<\delta T} \frac{1}{\sqrt{m}} \left(\frac{T}{m}\right)^{n'/2} \int \psi_T^{f,r}(u) 1_{\{0 \le T-v<t\}} \phi_\eta^{\Sigma,m}(\sqrt{\alpha}u, a_m) du \theta_\eta^m.$$

In the sequel, we write $S_1 \equiv S_2$ to mean

$$\int (S_1(r, t, f, l, v) - S_2(r, t, f, l, v)) P(dr, dt, df, dl, dv) = O(T^{-1}).$$

Following Malinovskii [17] (see also Sect. 4.2), we obtain

$$\begin{aligned} T_1 \equiv \int \psi_T^{f,r}(u) \int_{-\infty}^{\infty} \alpha^{n'/2-1} \phi_\Sigma(\sqrt{\alpha}u, \lambda) \\ \times \left\{1 + \sqrt{\frac{\alpha}{T}} \left(p^\Sigma(\sqrt{\alpha}u, \lambda) + \frac{\lambda(n'-1)}{2\alpha}\right)\right\} d\lambda du 1_{\{0 \le T-v<t\}}. \end{aligned} \tag{13}$$

In the same manner, we can prove $T_3 \equiv 0$ and

$$T_2 \equiv \frac{\alpha^{n'/2-1}}{2\sqrt{T}} \int \int_{-\infty}^{\infty} \psi_T^{f,r}(u) u \cdot \partial_\xi \phi_\Sigma(\sqrt{\alpha}u, \lambda) \lambda d\lambda du 1_{\{0 \le T-v<t\}}.$$

Further, integrating in λ, we have

$$T_1 + T_2 \equiv J \int \psi_T^{f,r}(u) \phi_\Gamma(u) \left\{1 - \frac{\rho \cdot q^\Gamma(u)}{\alpha\sqrt{T}} + \frac{1}{6\sqrt{T}} \sum_{i,j,k=1}^{n'} \gamma_{i,j,k} p_{i,j,k}^\Gamma(u)\right\} du, \tag{14}$$

where ϕ_Γ is the normal density with mean 0, covariance matrix Γ,

$$q^\Gamma(u) = -\phi_\Gamma(u)^{-1} \partial_u \phi_\Gamma(u), \quad p_{i,j,k}^\Gamma(u) = -\phi_\Gamma(u)^{-1} \partial_i \partial_j \partial_k \phi_\Gamma(u)$$

and $J = J(f, r, t, v, T) = \alpha^{-1} 1_{\{0 \le T-v<t\}}$.

Step 3 [Transformation by A and calculation of the coefficients]: This is the last step of the proof and here we derive (5). We exploit Lemma 2.1 of Bhattacharya and Ghosh [4] to treat the transformation by A. We, however, cannot directly apply it because of the existence of f and r, which derive from the first and last blocks in the decomposition of Z. Let H be a compact neighborhood of $0 \in \mathbb{R}^n$ on which A is smooth. Put $A^h(u) = A(\iota(u) + h) - A(h)$ for $h \in H$. Define a_j^h, $a_{i,j}^h$ and σ^h for $1 \le i, j \le n'$ as (4) with $\gamma = 0$ and A^h instead of A. By the argument of Lemma 2.1 of Bhattacharya and Ghosh [4] (see also Theorem 2.2 of Hall [13]), we have

$$\begin{aligned}
&\int_S \phi_\Gamma(u)\left\{1 - \frac{\rho \cdot q^\Gamma(u)}{\alpha\sqrt{T}} + \frac{1}{6\sqrt{T}}\sum_{i,j,k=1}^{n'} \gamma_{i,j,k} p_{i,j,k}^\Gamma(u)\right\} du \\
&\quad = \int_{\hat{B}} \phi(z)\left\{1 + \frac{1}{\sqrt{T}}\left(\frac{A_1^h}{\sigma^h} z + \frac{A_2^h}{6\{\sigma^h\}^3}(z^3 - 3z)\right)\right\} dz + O(T^{-1})
\end{aligned}$$

uniformly in $\hat{B} \in \mathcal{B}(\mathbb{R})$ and $h \in H$, where $S = \{u; \sqrt{T}A^h(u/\sqrt{T})/\sigma^h \in \hat{B}\}$, A_2^h is defined as A_2 in (5) with $a_j^h, a_{i,j}^h$ instead of $a_j, a_{i,j}$ and

$$A_1^h = -\frac{1}{\alpha}\sum_{j=1}^{n'} a_j^h \rho_j + \frac{1}{2}\sum_{i,j=1}^{n'} a_{i,j}^h \gamma_{i,j}.$$

Now, since

$$\int 1_{\left\{\frac{f+r}{T} \notin H,\ 0 \le T - v < t\right\}} P_\nu^{\hat{Z}_0}(df) P_x^{(R(v),\tau_1)}(dr, dt) dv = O(T^{-1}),$$

we can let $\hat{B} = B^h :\ = \sigma B/\sigma^h - \sqrt{T}A(h)/\sigma^h$ and $h = (f + r)/T$ to have

$$\begin{aligned}
&J\int \psi_T^{f,r}(u)\phi_\Gamma(u)\left\{1 - \frac{\rho \cdot q^\Gamma(u)}{\alpha\sqrt{T}} + \frac{1}{6\sqrt{T}}\sum_{i,j,k=1}^{n'} \gamma_{i,j,k} p_{i,j,k}^\Gamma(u)\right\} du \\
&\quad \equiv J\int_{B^h} \phi(z)\left\{1 + \frac{1}{\sqrt{T}}\left(\frac{A_1^h}{\sigma^h} z + \frac{A_2^h}{6\{\sigma^h\}^3}(z^3 - 3z)\right)\right\} dz \\
&\quad \equiv J\int_{B} \phi(z)\left\{1 + \frac{1}{\sqrt{T}}\left(\frac{\hat{A}_1}{\sigma} z + \frac{A_2}{6\sigma^3}(z^3 - 3z)\right)\right\} dz,
\end{aligned}$$

where $\hat{A}_1 = a \cdot (f + r - \iota(\rho)/\alpha) + \sum_{i,j} a_{i,j}\gamma_{i,j}/2$. Taking the expectation with respect to P, we have (1) with the coefficient

$$\int \hat{A}_1 \alpha^{-1} 1_{\{0 \le T - v < t\}} P_x^{(R(v),\tau_1)}(dr, dt) P_v^{(\hat{Z}_0, L_0)}(df, dl) dv$$
$$= a \cdot \left\{ P_\nu[\hat{Z}_0] + \alpha^{-1} P_x \left[\int_0^{\tau_1} Z_s ds \right] - \alpha^{-1}\iota(\rho) \right\} + \frac{1}{2} \sum_{i,j=1}^{n'} a_{i,j}\gamma_{i,j}$$

instead of A_1. Now, the proof is completed by showing the following lemma.

Lemma 3 *There exists a unique stationary distribution μ of X, and it holds*

$$P_\mu[|Z_{\tau_1}|] < \infty, \quad \alpha P_\mu[Z_{\tau_1}] = \iota(\rho) - P_x \left[\int_0^{\tau_1} Z_t dt \right].$$

Proof As is well known, μ defined as

$$\mu(B) = \frac{1}{\alpha} P_x \left[\int_0^{\tau_1} 1_B(X_t) dt \right], \ B \in \mathcal{B}(\mathbb{R})$$

is the unique stationary distribution of X. Hence

$$\alpha P_\mu[|Z_{\tau_1}|] = \alpha \int P_u[|Z_{\tau_1}|]\mu(du) = P_x \left[\int_0^{\tau_1} P_{X_t}[|Z_{\tau_1}|] dt \right]$$
$$= P_x \left[\int_0^{\tau_1} P_x[|Z_{t+\tau_1\circ\theta_t} - Z_t| |\mathcal{B}_t] dt \right]$$
$$= P_x \left[\int_0^{\tau_1} |Z_{t+\tau_1\circ\theta_t} - Z_t| dt \right].$$

Since $t + \tau_1 \circ \theta_t = \inf\{s > t; X_s = x, \text{ there exists } u \in [t, s) \text{ such that } X_u = y\}$, there are two possibilities for $t < \tau_1$, that are $t + \tau_1 \circ \theta_t = \tau_1$ and $t + \tau_1 \circ \theta_t = \tau_2$. The first case is that X hits y after the time t and then hits x at τ_1. In the second case, X has already hit y at some time before t and never hits y until it hits x at τ_1. From now on, we assume $x < y$ without loss of generality. Define

$$q = \sup \left\{ t \ge 0; t < \tau_1, \sup_{s \ge t} X_{s\wedge\tau_1} \ge y \right\}.$$

Note that q is an $\mathcal{B}_{\tau_1}$-measurable random variable. In fact, for any u,

$$
\begin{aligned}
\{q < a\} \cap \{\tau_1 \le u\} &= \left\{ \sup_{s \ge a} X_{s\wedge\tau_1} < y \right\} \cap \{\tau_1 \le u\} \\
&= \left\{ \sup_{s \ge a} X_{s\wedge u\wedge\tau_1} < y \right\} \cap \{\tau_1 \le u\} \in \mathcal{B}_u.
\end{aligned}
$$

Using this q, we have

$$
\begin{aligned}
P_x\left[\int_0^{\tau_1} |Z_{t+\tau_1\circ\theta_t} - Z_t| dt\right] &= P_x\left[\int_0^{q} |Z_{\tau_1} - Z_t| dt\right] + P_x\left[\int_q^{\tau_1} |Z_{\tau_2} - Z_t| dt\right] \\
&\le P_x[|Z_{\tau_1}|\tau_1] + P_x\left[\int_0^{\tau_1} |Z_t| dt\right] + \alpha P_x[|Z_{\tau_1}|],
\end{aligned}
$$

hence $P_\nu[|Z_{\tau_1}|] < \infty$. By repeating the same process with Z_t instead of $|Z_t|$, we have

$$
\alpha P_\mu[Z_{\tau_1}] = P_x[\tau_1 Z_{\tau_1}] - P_x\left[\int_0^{\tau_1} Z_t dt\right] + P_x[(\tau_1 - q)(Z_{\tau_2} - Z_{\tau_1})].
$$

The last term vanishes by the strong Markov property since $\gamma = 0$. □

4.2 Detailed calculation

Derivation of (10). It suffices to show the following lemma.

Lemma 4 *Let $S_m = (\hat{Z}_0, U_m, L_0, V_m)$. Then, $P_\nu^{S_m}$-a.s. (f, ξ, l, η), it holds*

$$
\begin{aligned}
&P_\nu\left[\Psi\left(S_m, R_m, L_{m+1}\right)\psi_2\left(L_0, V_m, L_{m+1}\right) | S_m = (f, \xi, l, \eta)\right] \\
&= P_x\left[\Psi\left(f, \xi, l, \eta, \hat{R}_m(l, \eta), \tau_1\right)\psi_2(l, \eta, \tau_1)\right]
\end{aligned}
$$

for any bounded measurable function Ψ.

Proof Denote by $(\mathcal{B}_t)$ the canonical filtration of X. By Galmarino's test (see e.g., Revuz and Yor [20], 1.4.21),

$$
\hat{T}\left(L_0, V_m\right) = T \wedge \left(L_0 + \sqrt{m} V_m + \alpha m\right) = T \wedge \sum_{j=0}^{m} L_j
$$

is a $(\mathcal{B}_t)$-stoppiping time and S_m is $\mathcal{B}_{\hat{T}(L_0,V_m)}$-measurable. Putting $(f,\xi,l,\eta)=S_m(\omega)$ for P_ν-a.s. ω, we have

$$
\begin{aligned}
&P_\nu\left[\Psi(S_m,R_m,L_{m+1})\psi_2(L_0,V_m,L_{m+1})|S_m\right](\omega)\\
&\quad=P_\nu\left[P_\nu\left[\Psi(S_m,R_m,L_{m+1})\psi_2(L_0,V_m,L_{m+1})|\mathcal{B}_{\hat{T}(L_0,V_m)}\right]|S_m\right](\omega)\\
&\quad=P_\nu\left[P_\nu\left[\Psi\left(f,\xi,l,\eta,R_m(l,\eta),L_{m+1}\right)\psi(l,\eta,L_{m+1})|\mathcal{B}_{\hat{T}(l,\eta)}\right]|S_m=(f,\xi,l,\eta)\right]\\
&\quad=P_\nu\left[P_x\left[\Psi\left(f,\xi,l,\eta,\hat{R}_m(l,\eta),\tau_1\right)\psi_2(l,\eta,\tau_1)\right]|S_m=(f,\xi,l,\eta)\right]\\
&\quad=P_x\left[\Psi\left(f,\xi,l,\eta,\hat{R}_m(l,\eta),\tau_1\right)\psi_2(l,\eta,\tau_1)\right].
\end{aligned}
$$

We used the optional sampling theorem in the second equality. □

Derivation of (11). Observe that

$$
\begin{aligned}
&\sum_{m;|T-\alpha m|<\delta T}\int\frac{\psi_2(l,\eta,t)}{m(1+|\eta|^2)}P_x^{(\hat{R}_m(l,\eta),\tau_1)}(dr,dt)P_\nu^{(\hat{Z}_0,L_0)}(df,dl)d\eta\\
&\quad\leq\sum_{m;|T-\alpha m|<\delta T}m^{-3/2}\int\left(1+\left|\frac{\eta-\alpha m-l}{\sqrt{m}}\right|^2\right)^{-1}1_{\{0\leq T-\eta\leq t\}}\\
&\qquad\times P_x^{(R(\eta),\tau_1)}(dr,dt)P_\nu^{(\hat{Z}_0,L_0)}(df,dl)d\eta
\end{aligned}
$$

where $R(\eta)=Z_{T-\eta}$, and that

$$
\begin{aligned}
&\sum_{m;|T-\alpha m|<\delta T}m^{-3/2}\left(1+\left|\frac{\eta-\alpha m-l}{\sqrt{m}}\right|^2\right)^{-1}\\
&\quad\leq((1-\delta)T/\alpha)^{-3/2}\left(3+\int_{-\infty}^{\infty}\left(1+\left|\frac{\eta-\alpha u-l}{\sqrt{(1+\delta)T/\alpha}}\right|^2\right)^{-1}du\right)\\
&\quad=((1-\delta)T/\alpha)^{-3/2}\left(3+\pi\sqrt{(1+\delta)T/\alpha^3}\right)=O(T^{-1}).
\end{aligned}
$$

Derivation of (12). We can put $A^m=A^m_{2,0}+A^m_{1,1}+A^m_{0,2}$ where

$$
A^m_{2,0}=\sum_{i,j=1}^{n'}\int_0^1(1-s)\partial_i\partial_j\phi^{\Sigma,m}\left(\sqrt{\alpha}u+s\theta^m_\xi,a_m+s\theta^m_\eta\right)ds\left\{\theta^m_\xi\right\}_i\left\{\theta^m_\xi\right\}_j,
$$

$$A_{1,1}^m = \sum_{j=1}^{n'} \int_0^1 2(1-s)\partial_j \partial_{n'+1} \phi^{\Sigma,m} \left(\sqrt{\alpha}u + s\theta_\xi^m, a_m + s\theta_\eta^m\right) ds \left\{\theta_\xi^m\right\}_j \theta_\eta^m,$$

$$A_{0,2}^m = \int_0^1 (1-s)\partial_{n'+1}^2 \phi^{\Sigma,m} \left(\sqrt{\alpha}u + s\theta_\xi^m, a_m + s\theta_\eta^m\right) ds \left\{\theta_\eta^m\right\}^2.$$

We shall show that these $A_{i,j}$'s are negligible up to $O(T^{-1})$. Notice that

$$0 \le \frac{\sqrt{m}}{a_m}\left(\sqrt{\frac{T}{\alpha m}} - 1\right) = \alpha^{-1}\left(\sqrt{\frac{T}{\alpha m}} + 1\right)^{-1} < \frac{1}{\alpha(1 + (1+\delta)^{-1/2})},$$

so that $|\theta_\xi^m| \le m^{-1/2} C |a_m u|$ for m with $|T - \alpha m| < \delta T$ and some constant C. In the following, we use ϵ and C as generic positive constants independent of T and m. It also holds $|\sqrt{\alpha}u + s\theta_\xi^m|^2 \ge \alpha u^2/(1+\delta)$ for all $s \in [0, 1]$, so that

$$\int |A_{2,0}^m| du \le \frac{C a_m^2}{m} \int_0^1 \exp(-\epsilon |a_m + s\theta_\eta^m|^2) ds.$$

Since

$$\begin{aligned}
&\sum_{m;|T-\alpha m|<\delta T} m^{-3/2} a_m^2 \exp(-\epsilon |a_m + s\theta_\eta^m|^2) \\
&\le C T^{-3/2} \sum_{m=-\infty}^{\infty} \frac{|T-\alpha m|^2}{T} \exp\left\{-\frac{\epsilon\alpha(T - \alpha m - s(T - v + l))^2}{T(1+\delta)}\right\} \\
&\le C T^{-3/2} \left\{1 + |T - v + l|^2/\sqrt{T} + \int_{-\infty}^{\infty} \frac{|T - \alpha z - s(T - v + l)|^2}{T} \right. \\
&\qquad \left. \times \exp\left\{-\frac{\epsilon\alpha(T - \alpha z - s(T - v + l))^2}{T(1+\delta)}\right\} dz \right\} \\
&\le C T^{-1}(1 + |T - v + l|^2/T)
\end{aligned} \tag{15}$$

uniformly in $s \in [0, 1]$, we conclude

$$\sum_{m;|T-\alpha m|<\delta T} \frac{1}{\sqrt{m}} \left(\frac{T}{m}\right)^{n'/2} \int |A_{2,0}^m| du 1_{\{0 \le T - v < t\}} \times dP_x^{(R(v),\tau_1)} dP_\nu^{(\hat{Z}_0, L_0)} dv = O(T^{-1}).$$

In the same manner, we obtain

$$\sum_{|T-\alpha m|<\delta T} \frac{1}{\sqrt{m}} \left(\frac{T}{m}\right)^{n'/2} \int |A_{0,2}| du 1_{\{0 \le T-v<t\}}$$
$$\le CT^{-1}|T-v+l|^2 1_{\{0\le T-v<t\}}$$

and

$$\sum_{|T-\alpha m|<\delta T} \frac{1}{\sqrt{m}} \left(\frac{T}{m}\right)^{n'/2} \int |A_{1,1}| du 1_{\{0 \le T-v<t\}}$$
$$\le CT^{-1}|T-v+l|(1+|T-v+l|/\sqrt{T}) 1_{\{0\le T-v<t\}},$$

hence $A_{1,1}$ and $A_{0,2}$ also are negligible.
Derivation of (13). Using Taylor's expansion, we have

$$\int_{a_{m+1}}^{a_m} \phi^{\Sigma,m}\left(\sqrt{\alpha}u, \lambda\right) - \frac{\lambda}{2\alpha\sqrt{m}} \phi_\Sigma\left(\sqrt{\alpha}u, \lambda\right) - \frac{\alpha}{\sqrt{m}} \frac{\partial \phi_\Sigma}{\partial \eta}\left(\sqrt{\alpha}u, \lambda\right) d\lambda$$
$$= \frac{\alpha}{\sqrt{m}} \phi^{\Sigma,m}\left(\sqrt{\alpha}u, a_m\right) + B_m(u),$$

where B_m is a remainder term which satisfies $|B_m(u)| \le Cm^{-3/2}\exp(-\epsilon(|u|^2+a_m^2))$ or $\le Cm^{-3/2}\exp(-\epsilon(|u|^2+a_{m+1}^2))$, depending on the signature of $T-\alpha m$. Moreover, by observing

$$\left(\frac{T}{\alpha m}\right)^{k/2} - 1 = \frac{k a_m}{2\alpha\sqrt{m}} + a_m^2 b_{k,m}^T, \quad \sup_{m;|T-\alpha m|<\delta T} |b_{k,m}^T| = O(T^{-1})$$

for each k and by a similar estimate to (15), we have

$$\sum_{m;|T-\alpha m|<\delta T} \left(\frac{\alpha^{(n'+1)/2}}{\sqrt{T}} - \frac{1}{\sqrt{m}}\left(\frac{T}{m}\right)^{n'/2}\right) \int_{a_{m+1}}^{a_m} \exp\left(-\epsilon\lambda^2\right) d\lambda = O(T^{-1}),$$

for any $\epsilon > 0$ and

$$\sum_{|T-\alpha m|<\delta T} \left(\left(\frac{T}{m}\right)^{n'/2} - \alpha^{n'/2}\right) \int\int_{a_{m+1}}^{a_m} \phi_\Sigma\left(\sqrt{\alpha}u, \lambda\right) d\lambda du$$
$$- \int\int_{a_{m+1}}^{a_m} \frac{\lambda n' \alpha^{n'/2-1}}{2\sqrt{m}} \phi_\Sigma\left(\sqrt{\alpha}u, \lambda\right) d\lambda du = O(T^{-1}).$$

(13) is obtained by combining these estimates.

4.3 Remarks on Theorem 1

Here are several remarks on possible extensions of Theorem 1 and notes on the regenerative method.

Remark 4 The higher order expansions are validated in the same manner by assuming the existence of higher order moments.

Remark 5 Theorem 1 remains valid even if we replace $H_{x,y}$ with $H_{x,y,z} := \inf\{t \geq H_y \wedge H_z; X_t = x\}$ for $y < x < z$, or $H_x^\delta = \inf\{t \geq \delta; X_t = x\}$ for $\delta > 0$, of which the proof is the same as that of Theorem 1. The point $x \in E$ plays a role of an atom in the context of the regenerative method. Therefore, we do not need to use the splitting method of Nummelin [19]. The reason why we need another point $y \in E$ (another two points $y, z \in E$ when we use $H_{x,y,z}$, or $\delta > 0$ when we use H_x^δ) is that it usually holds that $\inf\{t > 0; X_t = x\} = 0$, P_x-a.s..

Remark 6 We saw in Sect. 3 that Theorem 1 is applicable to a large class of one-dimensional diffusions. It is also possible to verify the conditions of Theorem 1 under appropriate assumptions when considering continuous-time Markov chains. On the other hand, unfortunately, $H_{x,y}$ will be infinite in cases of general multi-dimensional diffusions or diffusions with jumps. It is still possible to treat those cases by a similar regeneration-based argument, with the aid of an extension of the splitting method used in Sect. 7.4 of Nummelin [19]. Roughly speaking, the path of such a continuous-time Markov process can be decomposed into a discrete-time stationary 1-dependent sequence. Then, it is possible to prove e.g., the central limit theorem in the same manner as Nummelin [19]. However, for the EE by this method, it remains for further research to find verifiable conditions in terms of the coefficients of the corresponding SDE. Since the regeneration-based argument directly relies on the Markov property, it will require milder conditions for the validation of the EE than the mixing argument does, where the Markov property is only utilized via mixing properties.

Remark 7 The regenerative argument will be crucial in order to prove the validity of a specific bootstrap method for diffusions, which is a variant of the Regenerative Block Bootstrap (RBB) designed for discrete-time Markov processes by Bertail and Clémençon [2,3]. The author intends to treat the RBB for diffusions in an accompanying paper [10].

Acknowledgments The author is grateful to Prof. Nakahiro Yoshida for the the support and numerous helpful discussions. The author thanks Dr. Mariusz Bialecki, Dr. Shin Kanaya, Dr. Hideaki Noda, the referees and the associate editor for their helpful advise on the manuscript.

References

1. Bertail, P., Clémençon, S.: Edgeworth expansions of suitably normalized sample mean statistics for atomic Markov chains. Probab. Theory Relat. Fields **130**, 388–414 (2004)
2. Bertail, P., Clémençon, S.: Regenerative block bootstrap for Markov chains. Bernoulli **12**(4), 689–712 (2006)
3. Bertail, P., Clémençon, S.: Second-order properties of regeneration-based bootstrap for atomic Markov chains. Test **16**, 109–122 (2007)

4. Bhattacharya, R.N., Ghosh, J.K.: On the validity of the formal Edgeworth expansion. Ann. Stat. **6**, 434–451 (1976)
5. Bhattacharya, R.N., Rao, R.R.: Normal Approximation and Asymptotic Expansions. Wiley, New York (1976)
6. Bolthausen, E.: The Berry–Esseen theorem for functionals of discrete Markov chains. Z. Wahr. **54**, 59–73 (1980)
7. Bolthausen, E.: The Berry–Esseén theorem for strongly mixing Harris recurrent Markov chains. Z. Wahr. **60**, 283–289 (1982)
8. Borisov, I.S.: Estimate of the rate of convergence of distributions of additive functionals of a sequence of sums of independent random variables. Siberian Math. J. **19**(3), 371–383 (1978)
9. Fitzsimmons, P.J., Pitman, J.: Kac's moment formula and Feynman–Kac formula for additive functionals of a Markov process. Stoc. Proc. Appl. **79**, 117–134 (1999)
10. Fukasawa, M.: Regenerative block bootstrap for ergodic diffusions (in press)
11. Gikhman, I.I., Skorokhod, A.V.: Stochastic Differential Equations. Springer, Berlin (1972)
12. Götze, F., Hipp, C.: Asymptotic expansions for sums of weakly dependent random vectors. Z. Wahr. **64**, 211–239 (1983)
13. Hall, P.: The bootstrap and Edgeworth expansion. Springer, New York (1992)
14. Jensen, J.L.: Asymptotic expansions for strongly mixing Harris recurrent Markov chains. Scand. J. Stat. **16**, 47–63 (1989)
15. Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus. Springer, New York (1991)
16. Kusuoka, S., Yoshida, N.: Malliavin calculus, geometric mixing, and expansion of diffusion functionals. Probab. Theory Relat. Fields **116**, 457–484 (2000)
17. Malinovskii, V.K.: Limit theorems for Harris Markov chains, 1. Theory Probab. Appl. **31**(2), 269–285 (1987)
18. Mykland, P.A.: Asymptotic expansions and bootstrapping distributions for dependent variables: a martingale approach. Ann. Stat. **20**, 623–654 (1992)
19. Nummelin, E.: General Irreducible Markov Chains and Non-negative Operators. Cambridge University Press, London (1984)
20. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion. Springer, Berlin (1999)
21. Sakamoto, Y., Yoshida, N.: Asymptotic expansion of M-estimator over Wiener space. Stat. Inference Stoch. Process. **1**, 85–103 (1998)
22. Sakamoto, Y., Yoshida, N.: Asymtotic expansion formulas for functionals of ϵ-Markov processes with a mixing property. Ann. Inst. Stat. Math. **56**(3), 545–597 (2004)
23. Skorokhod, A.V.: Asymptotic Methods in the Theory of Stochastic Differential Equations. American Mathematical Society, USA (1989)
24. Veretennikov, A.Yu.: On lower bounds for mixing coefficients of Markov diffusions. In: Stoyanov, Y.M., Kabanov, J.M., Lipster, R. (eds.) From Stochastic Calculus to Mathematical Finance., pp. 623–633. Springer, Berlin (2006)
25. Yoshida, N.: Partial mixing and Edgeworth expansion. Probab. Theory Relat. Fields **129**, 559–624 (2004)