Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

Moment inequalities for mixing long-span high-frequency data and strongly consistent estimation of OU integrated diffusion process

Shanchao Yang[1], Jiaying Xie[1*], Shuyi Luo[1], Zhiyong Li[1] and Xin Yang[2*]

*Correspondence:
jiayingxiee@qq.com;
xinyang_emily@sina.cn
[1]School of Mathematics and Statistics, Guangxi Normal University, Guilin, 541004, China
[2]School of Mathematical Sciences, Guilin University of Aerospace Technology, Guilin, 541004, China

Abstract

Mixing is not much used in the high-frequency literature so far. However, mixing is a common weakly dependent property of continuous and discrete stochastic processes, such as Gaussian, Ornstein–Uhlenberck (OU), Vasicek, CIR, CKLS, logistic diffusion, generalized logistic diffusion, and double-well diffusion processes. So, long-span high-frequency data typically have weak dependence, and using mixing to study them is also an alternative approach. In this paper, we give some moment inequalities for long-span high-frequency data with ϕ-mixing, ρ-mixing, and α-mixing. These inequalities are effective tools for studying asymptotic properties. Applying these inequalities, we investigate the strong consistency of parameter estimation for the OU-integrated diffusion process. We also derive the mean square error of the estimation of the OU process and the optimal interval for the drift parameter estimator.

Keywords: Long-span high-frequency data; Mixing property; Moment inequality; OU-integrated diffusion process; Strong consistency

1 Introduction

Let $X_{i\Delta_n}$ $(i = 0, 1, 2, \dots, n)$ be the observation data of the continuous-time stochastic process $\{X_t, t \geq 0\}$ at time points $t_{i\Delta_n} = i\Delta_n$ $(i = 0, 1, 2, \dots, n)$ over an interval $[0, T]$ with $\Delta_n > 0$ and $T = n\Delta_n$. These data are called high-frequency if $\Delta_n \to 0$ as $n \to \infty$ and low-frequency if $\Delta_n = c$.

High-frequency data are commonly used in many fields, especially in finance. For example, in studying the asymptotic properties of the estimation of diffusion processes, it is often necessary to assume the basic condition $\Delta_n \to 0$, i.e., high-frequency samples. For details, one can refer to Andersen and Bollerslev [1], Barndorff-Nielsen and Shephard [6, 7], Christensen and Podolskij [13], Bandi and Russell [5], Fan and Wang [17], Fan et al. [16], Li et al. [31], Li and Guo [30], Chang et al. [11], and Yang et al. [50]. In these studies, the observation time intervals of high-frequency data have both fixed and increasing intervals. In the case of increasing time interval $[0, T]$ with $T \to \infty$, the high-frequency data is called long-span high-frequency data, which typically has weak dependence and is usually described as mixing dependence.

Assume that $\{X_t, t \geq 0\}$ is a continuous-time stochastic process, $\mathcal{F}_a^b$ represents a σ-algebraic field generated by $(X_t : a \leq t \leq b)$. For $\tau > 0$, let

$$\alpha(\tau) = \sup_{s\geq 0} \sup_{A\in\mathcal{F}_0^s, B\in\mathcal{F}_{s+\tau}^\infty} \left|P(A\cap B) - P(A)P(B)\right|, \tag{1.1}$$

$$\phi(\tau) = \sup_{s\geq 0} \sup_{A\in\mathcal{F}_0^s, B\in\mathcal{F}_{s+\tau}^\infty, P(A)>0} \left|P(B|A) - P(B)\right|, \tag{1.2}$$

$$\rho(\tau) = \sup_{s\geq 0} \sup_{X\in L^2(\mathcal{F}_0^s), Y\in L^2(\mathcal{F}_{s+\tau}^\infty)} \frac{|\mathrm{Cov}(X,Y)|}{\sqrt{E(X-EX)^2E(Y-EY)^2}}. \tag{1.3}$$

If $\alpha(\tau) \to 0, \phi(\tau) \to 0, \rho(\tau) \to 0$ as $\tau \to \infty$, then the process is called to be α-mixing, ϕ-mixing, and ρ-mixing, respectively. The long-span high-frequency data $X_{i\Delta_n}$ ($i = 0, 1, 2, \ldots, n$) are said to be α-mixing (ϕ-mixing, or ρ-mixing) if the corresponding process $\{X_t, t \geq 0\}$ is α-mixing (ϕ-mixing, or ρ-mixing).

Mixing is not much used in the high-frequency literature so far. One reason might be that it seems very difficult to establish the mixing properties of the models of interest. However, the mixing properties of many stochastic processes have been studied. Kolmogorov and Rozanov [29] first proved that ρ-mixing and α-mixing are equivalent for the stationary Gaussian process, and the process is ρ-mixing under appropriate conditions on the spectral density. Gorodetskii [22] discussed that linear processes are α-mixing under certain conditions. Later, Withers [44] improved the conclusions and gave sufficient conditions that are easier to verify. From that, it is not difficult to know that the stationary and reversible ARMA processes with normal white noise are α-mixing. The stationary GARCH process and the stationary Markov chain are both α-mixing processes (Carrasco and Chen [10]; Fan and Yao [18]), and the vector autoregressive (VAR) process, multivariate ARCH process and multivariate GARCH process are also α-mixing processes (Hafner and Preminger [23]; Boussama et al. [9]; Wong et al. [45]). Recently, Chen et al. [12] also gave some sufficient conditions for diffusion processes to be β-mixing, ρ-mixing and α-mixing, which provide us with an effective method to verify the mixing properties of some interesting diffusion processes, such as Ornstein–Uhlenberck (OU), Vasicek, Cox–Ingersoll–Ross (CIR), Chan–Karolyi–Longstaff–Sanders (CKLS), logistic diffusion, generalized logistic diffusion, double-well diffusion processes (Sect. 3). Therefore, mixing property can provide a selection method to study long-span high-frequency data of these interesting models.

In addition, although diffusion processes are semi-martingale and have Markov properties, integrated diffusion processes (see (4.2) below) no longer have these properties (Ditlevsen and Sørensen [15]). However, if diffusion processes are mixing, then the integrated diffusion processes also have the same mixing properties. So, mixing provides a new method for studying integrated diffusion processes, as we did in Sect. 4.

For mixing low-frequency data, moment and maximal inequalities are very useful for statistics to prove the asymptotic theory. These inequalities have been established before Billingsley [8], Yokoyama [51], Peligrad [34–36], Roussas and Ioannides [37], Shao [38, 39], Shao and Yu [40], Yang [47–49], Zhang [52], Wei et al. [43], and Xing et al. [46]. However, there is currently no literature on moment inequalities for mixing long-span high-frequency data. This article will provide such inequalities and apply them to study the uniformly strong consistency of parameter estimation of the OU-integrated process.

In Sect. 2, we give some moment inequalities for mixing long-span high-frequency data. To show that some long-span high-frequency data have mixing properties, in Sect. 3, we summarize some conclusions about the mixing of continuous-time stochastic processes from the existing literature, and verify the mixing properties of some interesting diffusion processes. As a simple application of the moment inequalities, we study the strong consistency of parameter estimates for the OU-integrated diffusion process in Sect. 4 and discuss the optimal sampling interval for the estimates. The last section is the conclusion of this paper.

2 Inequalities for mixing long-span high-frequency data

In this section, we give some moment inequalities for mixing long-span high-frequency data with $\Delta_n \to 0$ and $n\Delta_n \to \infty$ as $n \to \infty$. Let

$$\begin{aligned} &\tau_n = [1/\Delta_n] + 1, \qquad \lambda_n = \left[n/(2\tau_n)\right] + 1, \\ &\xi_j = \sum_{i=((j-1)\tau_n)\wedge n+1}^{(j\tau_n)\wedge n} X_{i\Delta_n}, \quad j = 1, 2, \dots, 2\lambda_n, \end{aligned}$$

where $[x]$ denotes the integer part of x, $a \wedge b = \min\{a, b\}$. If $(j-1)\tau_n \geq n$, we redefine $\xi_j = 0$. Clearly,

$$2(\lambda_n - 1)\tau_n \leq n < 2\lambda_n\tau_n,$$

and

$$\sum_{i=1}^{n} X_{i\Delta_n} = \sum_{j=1}^{2\lambda_n} \xi_j.$$

Theorem 2.1 *Suppose that $\{X_t, t \geq 0\}$ is a ϕ-mixing stochastic process with $EX_t = 0$ and $E|X_t|^r < \infty$ where $r \geq 2$. Let $\Delta_n \to 0$ and $n\Delta_n \to \infty$ as $n \to \infty$.*

(1) *If*

$$\sum_{k=0}^{\infty} \phi^{1/2}\left(2^k\right) < \infty, \tag{2.1}$$

then there exists a positive constant $C = C(r, \phi)$ independent of n such that

$$E\left|\sum_{i=1}^{n} X_{i\Delta_n}\right|^r \leq C\left\{E \max_{1\leq j\leq 2\lambda_n} |\xi_j|^r + \left(\lambda_n \max_{1\leq j\leq 2\lambda_n} E|\xi_j|^2\right)^{r/2}\right\}. \tag{2.2}$$

(2) *If*

$$\sum_{k=1}^{\infty} \phi^{1/2}(k) < \infty, \tag{2.3}$$

then there exists a positive constant $C = C(r,\phi)$ independent of n such that

$$E\left|\sum_{i=1}^{n} X_{i\Delta_n}\right|^r \leq C\left\{E\max_{1\leq j\leq 2\lambda_n}|\xi_j|^r + \left(\sum_{j=1}^{2\lambda_n} E|\xi_j|^2\right)^{r/2}\right\}. \tag{2.4}$$

Remark 2.1 Obviously, the second inequality (2.4) implies the first inequality (2.2), but the condition (2.1) with logarithmic decay mixing coefficient is weaker than the condition (2.3) with polynomial decay mixing coefficient. So, the first inequality is suitable for processes with longer dependence, while the second inequality is suitable for processes with shorter dependence.

There are also various inequalities for the cases of ρ-mixing and α-mixing, as shown in Theorem 2.2–2.5 below, which are suitable for different types of dependency processes.

The idea to prove the theorem is to transition from moment inequalities for mixing low-frequency data to moment inequalities for mixing long-span high-frequency data. So, we first give the following moment inequalities for ϕ-mixing low-frequency data.

Let $S_n = \sum_{i=1}^n X_i$, where $\{X_i; i \geq 1\}$ is a sequence of random variables.

Lemma 2.1 (*Shao* [38]) *Let $\{X_i; i \geq 1\}$ be a sequence of ϕ-mixing random variables, r,η be positive real numbers satisfying $r > 1$ and $0 < \eta < 1/(1+4^r)$. If there exist $A_n > 0$ and an integer $p \geq 1$ such that*

$$\phi(p) + \max_{p\leq m\leq n} P\big(|S_n - S_m| > A_n\big) < \eta, \quad \forall n \geq p, \tag{2.5}$$

then, for any $n \geq p$,

$$E\max_{1\leq i\leq n}|S_i|^r \leq \big(1-\eta-4^r\eta\big)^{-1}\Big\{(8A_n)^r + 2(4p)^r E\max_{1\leq i\leq n}|X_i|^r\Big\}. \tag{2.6}$$

Lemma 2.2 *Let $\{X_i; i \geq 1\}$ be a sequence of ϕ-mixing random variables with $E|X_i|^r < \infty$ where $r \geq 2$. If there exists a sequence of real numbers $C_n > 0$ such that*

$$E\left(\sum_{i=a+1}^{a+m} X_i\right)^2 \leq C_n, \quad \forall 1 \leq m \leq n, a \geq 0, \tag{2.7}$$

then there exists a positive constant $C = C(r,\phi)$ independent of n such that

$$E\max_{1\leq i\leq n}|S_i|^r \leq C\Big\{E\max_{1\leq i\leq n}|X_i|^r + C_n^{r/2}\Big\}. \tag{2.8}$$

Proof Let $A_n^2 = 4(1+4^r)C_n$. For any $n \geq m \geq p \geq 1$, we have

$$P\big(|S_n - S_m| > A_n\big) \leq A_n^{-2}E|S_n - S_m|^2 \leq A_n^{-2}C_n = \frac{1}{4(1+4^r)}.$$

Since $\phi(p) \to 0$ as $p \to \infty$, there exists $p > 1$ such that $\phi(p) < \frac{1}{4(1+4^r)}$. Thus,

$$\phi(p) + \max_{p\leq m\leq n} P\big(|S_n - S_m| > A_n\big) < \frac{1}{2(1+4^r)} =: \eta, \quad \forall n \geq p.$$

Note that $\eta < 1/(1+4^r)$. By Lemma 2.1, we have, for any $n \geq p$,

$$E \max_{1\leq i\leq n} |S_i|^r \leq \left(1-\eta-4^r\eta\right)^{-1}\left\{\left(16\sqrt{\left(1+4^r\right)}\right)^r C_n^{r/2} + 2(4p)^r E \max_{1\leq i\leq n} |X_i|^r\right\}$$
$$\leq C\left\{C_n^{r/2} + E \max_{1\leq i\leq n} |X_i|^r\right\}.$$

When $n < p$, it is obvious that

$$E \max_{1\leq i\leq n} |S_i|^r \leq p^r E \max_{1\leq i\leq n} |X_i|^r.$$

Combining the above two equations leads to the conclusion. This completes the proof. □

Lemma 2.3 (*Ibragimov* [24], *Lemma* 1.1) *Let* $\{X_i; i \geq 1\}$ *be a sequence of* ϕ*-mixing random variables,* $\mathcal{F}_k^l = \sigma(X_i, k \leq i \leq l)$. *Suppose that* X *and* Y *are* $\mathcal{F}_1^k$ *measurable and* $\mathcal{F}_{k+n}^\infty$ *measurable, respectively, random variables with* $E|X|^p < \infty$ *and* $E|Y|^q < \infty$, *where* $p > 1, q > 1$ *and* $1/p + 1/q = 1$. *Then*

$$\left|E(XY) - (EX)(EY)\right| \leq 2\phi^{1/p}(n)\left(E|X|^p\right)^{1/p}\left(E|Y|^q\right)^{1/q}.$$

Lemma 2.4 *Let* $\{X_i; i \geq 1\}$ *be a sequence of* ϕ*-mixing random variables with* $EX_i = 0$ *and* $E|X_i|^r < \infty$ *where* $r \geq 2$. *If* $\sum_{k=0}^\infty \phi^{1/2}(2^k) < \infty$, *then there exists a positive constant* $C = C(r, \phi)$ *independent of* n *such that*

$$E \max_{1\leq i\leq n} |S_i|^r \leq C\left\{E \max_{1\leq i\leq n} |X_i|^r + \left(n \max_{1\leq i\leq n} E|X_i|^2\right)^{r/2}\right\}. \tag{2.9}$$

Proof Denote $\|X\|_r = (E|X|^r)^{1/r}$ and

$$S_a(m) = \sum_{i=a+1}^{a+m} X_i, \qquad \sigma_m = \sup_{a\geq 1}\left\|S_a(m)\right\|_2, \qquad \sigma_1 = \sup_{i\geq 1}\|X_i\|_2.$$

Obviously

$$S_a(2m) = S_a(m) + S_{a+m}\left(\left[m^{1/3}\right]\right) + S_{a+m+[m^{1/3}]}(m) - S_{a+2m}\left(\left[m^{1/3}\right]\right).$$

By Minkowski's inequality, we have

$$\left\|S_a(2m)\right\|_2 \leq \left\|S_a(m) + S_{a+m+[m^{1/3}]}(m)\right\|_2 + \left\|S_{a+m}\left(\left[m^{1/3}\right]\right)\right\|_2 + \left\|S_{a+2m}\left(\left[m^{1/3}\right]\right)\right\|_2$$
$$\leq \left\|S_a(m) + S_{a+m+[m^{1/3}]}(m)\right\|_2 + 2\left[m^{1/3}\right]\sigma_1.$$

From Lemma 2.3, we have

$$E\left(S_a(m) + S_{a+m+[m^{1/3}]}(m)\right)^2$$
$$= ES_a^2(m) + ES_{a+m+[m^{1/3}]}^2(m) + 2E\left(S_a(m)S_{a+m+[m^{1/3}]}(m)\right)$$
$$= 2\sigma_m^2 + 2\phi^{1/2}\left(\left[m^{1/3}\right]\right)\left\|S_a(m)\right\|_2\left\|S_{a+m+[m^{1/3}]}(m)\right\|_2$$

$$\leq 2\left(1+\phi^{1/2}\left(\left[m^{1/3}\right]\right)\right)\sigma_m^2.$$

Therefore,

$$\sigma_{2m} \leq 2^{1/2}\left(1+\phi^{1/2}\left(\left[m^{1/3}\right]\right)\right)^{1/2}\sigma_m + 2\left[m^{1/3}\right]\sigma_1.$$

Let $m = 2^{k-1}$ for any integer $k \geq 1$, we have

$$\sigma_{2^k} \leq 2^{1/2}\left(1+\phi^{1/2}\left(\left[2^{(k-1)/3}\right]\right)\right)^{1/2}\sigma_{2^{k-1}} + 2\left[2^{(k-1)/3}\right]\sigma_1.$$

Using the above formula to iterate repeatedly, we get

$$\begin{aligned}
\sigma_{2^k} &\leq 2^{1/2}\left(1+\phi^{1/2}\left(\left[2^{(k-1)/3}\right]\right)\right)^{1/2}\sigma_{2^{k-1}} + 2\left[2^{(k-1)/3}\right]\sigma_1 \\
&\leq 2^{2/2}\left(1+\phi^{1/2}\left(\left[2^{(k-2)/3}\right]\right)\right)^{1/2}\left(1+\phi^{1/2}\left(\left[2^{(k-1)/3}\right]\right)\right)^{1/2}\sigma_{2^{k-2}} \\
&\quad + 2\times 2^{1/2}\left(1+\phi^{1/2}\left(\left[2^{(k-1)/3}\right]\right)\right)^{1/2}\left[2^{(k-2)/3}\right]\sigma_1 + 2\left[2^{(k-1)/3}\right]\sigma_1 \\
&\quad \cdots \\
&\quad \leq 2\sigma_1 \sum_{j=1}^{k} 2^{(j-1)/2}\left[2^{(k-j)/3}\right]\prod_{i=1}^{j-1}\left(1+\phi^{1/2}\left(\left[2^{(k-i)/3}\right]\right)\right)^{1/2} \\
&\leq 2^{k/3+1/2}\sigma_1 \sum_{j=1}^{k} 2^{j/6}\prod_{i=1}^{k-1}\left(1+\phi^{1/2}\left(\left[2^{(k-i)/3}\right]\right)\right)^{1/2} \\
&\leq C2^{k/2}\sigma_1\left\{\prod_{i=1}^{k-1}\left(1+\phi^{1/2}\left(\left[2^{(k-i)/3}\right]\right)\right)\right\}^{1/2}.
\end{aligned}$$

Since $\log(1+x) < x$ for any $x > 0$, so

$$\begin{aligned}
\log\left(\prod_{i=1}^{k-1}\left(1+\phi^{1/2}\left(\left[2^{(k-i)/3}\right]\right)\right)\right) &= \sum_{i=1}^{k-1}\log\left(1+\phi^{1/2}\left(\left[2^{(k-i)/3}\right]\right)\right) \\
&\leq \sum_{i=1}^{k-1}\phi^{1/2}\left(\left[2^{(k-i)/3}\right]\right) \\
&\leq \sum_{j=1}^{k}\phi^{1/2}\left(\left[2^{j/3}\right]\right).
\end{aligned}$$

For the integer $[2^{j/3}]$, there exists an integer $s \geq 1$ such that $2^{s-1} \leq [2^{j/3}] < 2^s$. Obviously, $2^{s-1} \leq 2^{j/3} < 2^s$. Thus, $s-1 \leq j/3 < s$, i.e., $3s-3 \leq j < 3s$. Therefore, there are only three values of j that meet the condition $2^{s-1} \leq [2^{j/3}] < 2^s$. By the monotonicity of $\phi(n)$, we have

$$\sum_{j=1}^{k}\phi^{1/2}\left(\left[2^{j/3}\right]\right) \leq 3\sum_{i=0}^{\infty}\phi^{1/2}\left(2^i\right) < \infty.$$

Thereby, $\prod_{i=1}^{k-1}(1+\phi^{1/2}([2^{(k-i)/3}])) \leq C < \infty$. Hence, $\sigma_{2^k} \leq C2^{k/2}\sigma_1$, i.e.

$$ES_{2^k}^2 \leq C2^k \sup_{i\geq 1} EX_i^2.$$

For any $n \geq 1$, there exists an integer $k > 0$ such that $2^{k-1} \leq n < 2^k$. Let $X_i = 0$ for $i > n$. Then, we have

$$ES_n^2 = ES_{2^k}^2 \leq C2^k \max_{1\leq i\leq n} EX_i^2 \leq 2Cn \max_{1\leq i\leq n} EX_i^2.$$

It follows the desired conclusion by Lemma 2.2. This completes the proof. □

Lemma 2.5 *Let $\{X_i; i \geq 1\}$ be a sequence of ϕ-mixing random variables with $EX_i = 0$ and $E|X_i|^r < \infty$ where $r \geq 2$. If $\sum_{k=1}^{\infty} \phi^{1/2}(k) < \infty$, then there exists a positive constant $C = C(r, \phi)$ independent of n such that*

$$E \max_{1\leq i\leq n} |S_i|^r \leq C\left\{E \max_{1\leq i\leq n} |X_i|^r + \left(\sum_{i=1}^{n} EX_i^2\right)^{r/2}\right\}. \tag{2.10}$$

Proof From Lemma 2.3, we have

$$\begin{aligned}
E\left(\sum_{i=1}^{n} X_i\right)^2 &= \sum_{i=1}^{n} EX_i^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} E(X_iX_j) \\
&\leq \sum_{i=1}^{n} EX_i^2 + C\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \phi^{1/2}(j-i)\left(EX_i^2\right)^{1/2}\left(EX_j^2\right)^{1/2} \\
&= \sum_{i=1}^{n} EX_i^2 + C\sum_{i=1}^{n-1}\sum_{k=1}^{n-i} \phi^{1/2}(k)\left(EX_i^2\right)^{1/2}\left(EX_{i+k}^2\right)^{1/2} \\
&\leq \sum_{i=1}^{n} EX_i^2 + C\sum_{i=1}^{n-1}\sum_{k=1}^{n-i} \phi^{1/2}(k)\left(EX_i^2 + EX_{i+k}^2\right) \\
&\leq \sum_{i=1}^{n} EX_i^2 + C\sum_{k=1}^{n} \phi^{1/2}(k)\sum_{i=1}^{n} EX_i^2 + C\sum_{k=1}^{n-1}\sum_{i=1}^{n-k} \phi^{1/2}(k)EX_{i+k}^2 \\
&\leq \left(1 + C\sum_{k=1}^{\infty} \phi^{1/2}(k)\right)\sum_{i=1}^{n} EX_i^2,
\end{aligned}$$

This implies that the condition (2.7) in Lemma 2.2 holds, which leads to the desired conclusion. This completes the proof. □

Proof of Theorem 2.1 Let

$$Y_j = \xi_{2j-1}, Z_j = \xi_{2j}, \quad j = 1, 2, \ldots, \lambda_n.$$

Obviously,

$$\sum_{i=1}^{n} X_{i\Delta_n} = \sum_{j=1}^{\lambda_n} Y_j + \sum_{j=1}^{\lambda_n} Z_j.$$

As the subscript time interval $\tau_n\Delta_n$ between random variables Y_j and Y_{j+1} satisfies $\tau_n\Delta_n \geq 1$, $\{Y_1, Y_2, \ldots, Y_{\lambda_n}\}$ are low-frequency ϕ-mixing random variables. Thus, using Lemma 2.4, we have

$$E\left|\sum_{j=1}^{\lambda_n} Y_j\right|^r \le C\left\{E\max_{1\le j\le\lambda_n}|Y_j|^r + \left(\lambda_n\max_{1\le j\le\lambda_n}E|Y_j|^2\right)^{r/2}\right\}$$
$$\le C\left\{E\max_{1\le j\le 2\lambda_n}|\xi_j|^r + \left(\lambda_n\max_{1\le j\le 2\lambda_n}E|\xi_j|^2\right)^{r/2}\right\}.$$

Similarly,

$$E\left|\sum_{j=1}^{\lambda_n} Z_j\right|^r \le C\left\{E\max_{1\le j\le 2\lambda_n}|\xi_j|^r + \left(\lambda_n\max_{1\le j\le 2\lambda_n}E|\xi_j|^2\right)^{r/2}\right\}.$$

Therefore, conclusion (2.2) holds. Conclusion (2.4) is easily obtained by using Lemma 2.5 and the similar procedure as above. This completes the proof. □

Theorem 2.2 *Suppose $\{X_t, t \ge 0\}$ is a ρ-mixing stochastic process with $EX_t = 0$ and $E|X_t|^r < \infty$ where $r > 1$. Let $\Delta_n \to 0$ and $n\Delta_n \to \infty$ as $n \to \infty$.*

(1) *If $r \ge 2$ and*

$$\sum_{k=0}^{\infty}\rho^{2/r}\left(2^k\right) < \infty, \tag{2.11}$$

then there exists a positive constant $C = C(r,\rho)$ independent of n such that

$$E\left|\sum_{i=1}^{n} X_{i\Delta_n}\right|^r \le C\left\{\lambda_n\max_{1\le j\le 2\lambda_n}E|\xi_j|^r + \left(\lambda_n\max_{1\le j\le 2\lambda_n}E|\xi_j|^2\right)^{r/2}\right\}. \tag{2.12}$$

(2) *If*

$$\rho(\tau) = O\left(\tau^{-\theta}\right), \quad \theta > 0, \tag{2.13}$$

then for any given $\varepsilon > 0$, there exists a positive constant $C = C(r,\rho(\cdot),\theta,\varepsilon)$ independent of n such that

$$E\left|\sum_{i=1}^{n} X_{i\Delta_n}\right|^r \le C\lambda_n^{\varepsilon}\sum_{j=1}^{2\lambda_n}E|\xi_j|^r, \quad 1 < r \le 2, \tag{2.14}$$

and

$$E\left|\sum_{i=1}^{n} X_{i\Delta_n}\right|^r \le C\lambda_n^{\varepsilon}\left\{\sum_{j=1}^{2\lambda_n}E|\xi_j|^r + \left(\sum_{j=1}^{2\lambda_n}E|\xi_j|^2\right)^{r/2}\right\}. \tag{2.15}$$

Proof It is easy to obtain (2.12) using the proof process of Theorem 2.1 and Theorem 1.1 in Shao [39], while (2.14) and (2.15) are obtained using Theorem 1 in Yang [47]. This completes the proof. □

Theorem 2.3 *Suppose* $\{X_t, t \geq 0\}$ *is an* α*-mixing stochastic process with* $EX_t = 0$ *and* $E|X_t|^{r+\delta} < \infty$ *where* $r > 2, \delta > 0, 2 < \nu \leq r + \delta$. *Let* $\Delta_n \to 0$ *and* $n\Delta_n \to \infty$ *as* $n \to \infty$. *If*

$$\alpha(\tau) = O(\tau^{-\theta}), \quad \theta > 0, \tag{2.16}$$

then for any given $\varepsilon > 0$, *there exists a positive constant* $K = K(\varepsilon, r, \delta, \nu, \theta, C) < \infty$ *such that*

$$E\left|\sum_{i=1}^{n} X_{i\Delta_n}\right|^r \leq K\left\{(\lambda_n C_n)^{r/2} \max_{1\leq j\leq 2\lambda_n} \|\xi_j\|_\nu^r + \lambda_n^{(r-\delta\theta/(r+\delta))\vee(1+\varepsilon)} \max_{1\leq j\leq 2\lambda_n} \|\xi_j\|_{r+\delta}^r\right\}, \tag{2.17}$$

where $C_n = (\sum_{i=0}^{\lambda_n}(i+1)^{2/(\nu-2)}\alpha(i))^{(\nu-2)/\nu}$.

In particular, if $\theta > \nu/(\nu - 2)$ *and* $\theta \geq (r - 1)(r + \delta)/\delta$, *then for any given* $\varepsilon > 0$,

$$E\left|\sum_{i=1}^{n} X_{i\Delta_n}\right|^r \leq K\left\{\lambda_n^{r/2} \max_{1\leq j\leq 2\lambda_n} \|\xi_j\|_\nu^r + \lambda_n^{1+\varepsilon} \max_{1\leq j\leq 2\lambda_n} \|\xi_j\|_{r+\delta}^r\right\}; \tag{2.18}$$

If $\theta \geq r(r + \delta)/(2\delta)$, *then*

$$E\left|\sum_{i=1}^{n} X_{i\Delta_n}\right|^r \leq K\lambda_n^{r/2} \max_{1\leq j\leq 2\lambda_n} \|\xi_j\|_{r+\delta}^r. \tag{2.19}$$

Proof The conclusion is derived from Theorem 4.1 in Shao & Yu [40]. This completes the proof. □

Theorem 2.4 *Suppose* $\{X_t, t \geq 0\}$ *is an* α*-mixing stochastic process with* $EX_t = 0$ *and* $E|X_t|^{r+\delta} < \infty$ *where* $r > 2, \delta > 0, 2 < \nu \leq r + \delta$. *Let* $\Delta_n \to 0$ *and* $n\Delta_n \to \infty$ *as* $n \to \infty$. *If*

$$\alpha(\tau) = O(\tau^{-\theta}), \quad \theta > 0, \tag{2.20}$$

and θ *satisfies*

$$\theta > \max\{\nu/(\nu - 2), (r - 1)(r + \delta)/\delta\}, \tag{2.21}$$

then for any given $\varepsilon > 0$, *there exists a positive constant independent of* n $K = K(\varepsilon, r, \delta, \nu, \theta, C) < \infty$ *such that*

$$E\left|\sum_{i=1}^{n} X_{i\Delta_n}\right|^r \leq K\left\{\lambda_n^\varepsilon \sum_{j=1}^{2\lambda_n} E|\xi_j|^r + \sum_{j=1}^{2\lambda_n} \|\xi_j\|_{r+\delta}^r + \left(\sum_{j=1}^{2\lambda_n} \|\xi_j\|_\nu^2\right)^{r/2}\right\}. \tag{2.22}$$

Proof The conclusion is obtained from Theorem 2.1 in Yang [49]. This completes the proof. □

Theorem 2.5 *Suppose* $\{X_t, t \geq 0\}$ *is an* α*-mixing stochastic process with* $EX_t = 0$ *and* $E|X_t|^{r+\delta} < \infty$ *where* $r > 2, \delta > 0$. *Let* $\Delta_n \to 0$ *and* $n\Delta_n \to \infty$ *as* $n \to \infty$. *If the condition* (2.20) *holds and* θ *satisfies*

$$\theta > r(r + \delta)/(2\delta), \tag{2.23}$$

then for any given $\varepsilon > 0$, there exists a positive constant independent of n $K = K(\varepsilon, r, \delta, \theta, C) < \infty$ such that

$$E\left|\sum_{i=1}^{n} X_{i\Delta_n}\right|^r \le K\left\{\lambda_n^{\varepsilon}\sum_{j=1}^{2\lambda_n} E|\xi_j|^r + \left(\sum_{j=1}^{2\lambda_n}\|\xi_j\|_{r+\delta}^2\right)^{r/2}\right\}. \tag{2.24}$$

Proof The conclusion is obtained from Theorem 2.2 in Yang [49]. This completes the proof. □

From Theorem 2.4 and Theorem 2.5 the following corollary is immediately obtained.

Corollary 2.1 *Suppose $\{X_t, t \ge 0\}$ is an α-mixing stochastic process with $EX_t = 0$, $E|X_t|^{r+\delta_0} < \infty$ and $\alpha(\tau) = O(e^{-\theta\tau})$ where $r > 2, \delta_0 > 0, \theta > 0$. Let $\Delta_n \to 0$ and $n\Delta_n \to \infty$ as $n \to \infty$. Then for any given $\varepsilon > 0$ and $\delta \in (0, \delta_0]$, there exists a positive constant independent of n $K = K(\varepsilon, r, \delta, \theta, C) < \infty$ such that*

$$E\left|\sum_{i=1}^{n} X_{i\Delta_n}\right|^r \le K\left\{\lambda_n^{\varepsilon}\sum_{j=1}^{2\lambda_n} E|\xi_j|^r + \sum_{j=1}^{2\lambda_n}\|\xi_j\|_{r+\delta}^r + \left(\sum_{j=1}^{2\lambda_n}\|\xi_j\|_{2+\delta}^2\right)^{r/2}\right\}, \tag{2.25}$$

$$E\left|\sum_{i=1}^{n} X_{i\Delta_n}\right|^r \le K\left\{\lambda_n^{\varepsilon}\sum_{j=1}^{2\lambda_n} E|\xi_j|^r + \left(\sum_{j=1}^{2\lambda_n}\|\xi_j\|_{r+\delta}^2\right)^{r/2}\right\}. \tag{2.26}$$

Remark 2.2 (1) The inequalities given in Theorems 2.1–2.5 and Corollary 2.1 use the moments of ξ_j as the upper bounds. ξ_j is a sum of τ_n random variables in which the time subscript interval between any two variables $X_{i\Delta_n}$ and $X_{k\Delta_n}$ is less than 2. Therefore, the mixing (i.e., asymptotic independence) property cannot be used to calculate the moments of ξ_j. In this sense, using the moments of ξ_j as the upper bound control terms for the moment inequalities of mixing high-frequency random variables is appropriate. Moreover, in the application, to calculate the moments of ξ_j, we can no longer use mixing properties but can only use other properties of random processes, as shown in the proofs of 4.2 and Theorem 4.3 later.

(2) If $\rho(\tau) = O((\log\tau)^{-r/2}(\log\log\tau)^{-r})$ for $r \ge 2$, then $\sum_{k=0}^{\infty}\rho^{2/r}(2^k) < \infty$. It implies that condition (2.11) in Theorem 2.2 only requires the ρ-mixing coefficient to have logarithmic decay, while condition (2.13) requires the mixing coefficient to have polynomial decay. In practice, the mixing coefficient tends to zero at different speeds, see Kolmogorov and Rozanov [29], Chen et al. [12], and the next section. Hence, it is reasonable to assume whether the mixing coefficients are short-range- or long-range-dependent.

(3) For each mixing process, we provide multiple inequalities. It is clear that $\sum_{j=1}^{2\lambda_n} E|\xi_j|^r$ and $(\sum_{j=1}^{2\lambda_n} E|\xi_j|^2)^{r/2}$ are superior to $\lambda_n \max_{1\le j\le 2\lambda_n} E|\xi_j|^r$ and $(\lambda_n \max_{1\le j\le 2\lambda_n} E|\xi_j|^2)^{r/2}$, respectively, for nonstationary processes. Therefore, the upper bound of the inequality obtained under the condition that the mixing coefficient approaches zero at a faster rate is superior to the upper bound obtained under the condition that the mixing coefficient approaches zero at a slower rate.

3 Mixing property of random process

Since the concept of mixing was proposed, many scholars have studied the mixing properties of stochastic processes. They mainly discussed the sufficient conditions for stochastic processes to have mixing property and the decay rate of mixing coefficient. Since high-frequency data can be regarded as discretizations of a continuous stochastic process, we are interested in the mixing property of a continuous-time stochastic process. Therefore, we summarize some meaningful conclusions about the mixing of continuous stochastic processes, which can be applied to long-span high-frequency data.

3.1 Mixing property of stationary Gaussian process

In the cases of continuous time and discrete time, Kolmogorov and Rozanov [29] proved that the α-mixing of the stationary Gaussian process is equivalent to ρ-mixing and derived some sufficient conditions for ρ-mixing. Later, Ibragimov [25] derived the necessary conditions of α-mixing for the discrete stationary Gaussian process and further discussed some sufficient conditions (Ibragimov [26]). The following conclusions are from Kolmogorov and Rozanov [29].

Theorem 3.1 *Suppose that X_t is a continuous stationary Gaussian process and $f(\lambda)$ is the spectral density of the process. Then, the following several statements hold:*

(1) $\alpha(\tau) \le \rho(\tau) \le 2\pi\alpha(\tau)$.

(2)

$$\rho(\tau) = \inf_{\varphi} \operatorname{ess\,sup}_{\lambda} \left\{ \left| f(\lambda) - e^{\mathrm{i}\lambda\tau}\varphi(\lambda) \right| \frac{1}{f(\lambda)} \right\},$$

where $\inf_\varphi$ is taken over all functions $\varphi(z)$, which are extended analytically into the lower semi-plane.

(3) *If $f(\lambda)$ is positive and uniformly continuous and for sufficiently large λ satisfies the inequality*

$$\frac{m}{\lambda^k} \le f(\lambda) \le \frac{M}{\lambda^{k-1}}$$

for some positive m, M, and integral $k > 0$, then X_t is ρ-mixing.

(4) *If there exists an analytic function $\varphi_0(z)$ such that $|f/\varphi_0| \ge \varepsilon > 0$, and the derivative $(f/\varphi_0)^{(k)}$ is bounded uniformly, then X_t is ρ-mixing with polynomial decay $\rho(\tau) = O(\tau^{-k})$.*

(5) *If X_t is a Markov process, then X_t is ρ-mixing with exponential decay.*

Remark 3.1 ess sup is the essential supremum of g defined by $\operatorname{ess\,sup}_x g(x) = \inf\{a \in \mathrm{R} : \mu(\{x : g(x) > a\}) = 0\}$, where μ is a measure.

Conclusion (1) implies that α-mixing and ρ-mixing are equivalent for stationary Gaussian process, and both of these mixing coefficients have the same decay rate. So, conclusions (3)–(5) are also valid for α-mixing. Conclusion (2) gives the expression of ρ-mixing coefficient determined by spectral density. Conclusion (3) gives a sufficient condition for ρ-mixing, while conclusion (4) provides a sufficient condition for ρ-mixing with polynomial decay. We know from (1) and (5) that a stationary Gaussian–Markov process is ρ-mixing and α-mixing with exponential decay.

3.2 Mixing property of time-homogeneous diffusion process

Suppose that X_t is the strong solution of the time-homogeneous stochastic differential equation (SDE)

$$dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dB_t \tag{3.1}$$

with left boundary l and right boundary r, either of which can be infinite. The function $\mu(x)$ is the drift, $\sigma(x)$ is the diffusion function, and B_t is a standard Brownian motion.

Let $s(z) = \exp\{-\int_{z_0}^{z} \frac{2\mu(x)}{\sigma^2(x)}\,dx\}$ be the scale density function ($z_0 \in (l,r)$), $S(u) = \int_{z_0}^{u} s(z)\,dz$ the scale function, and $m(x) = (\sigma^2(x)s(x))^{-1}$ the speed density function. From Corollary 4.2 and Remark 4.3 in Chen et al. [12], we have the following conclusion.

Theorem 3.2 *Suppose that the following conditions are satisfied.*

A.1 $\mu(x)$ and $\sigma(x)$ are continuous on (l,r) with $\sigma(x)$ strictly positive on this interval.

A.2 $S(l) = -\infty$ and $S(r) = +\infty$.

A.3 $\limsup_{x\nearrow r}(\frac{\mu(x)}{\sigma(x)} - \frac{\sigma'(x)}{2}) < 0$ and $\liminf_{x\searrow l}(\frac{\mu(x)}{\sigma(x)} - \frac{\sigma'(x)}{2}) > 0$.

Then X_t is ρ-mixing and α-mixing with exponential decay, and $\int_l^r m(x)\,dx < \infty$.

The strong solution of the SDE (3.1) has the Markov property by Theorem 5.6 in Klebaner [28]. Under the conditions of Theorem 3.2, X_t has an invariant distribution and its invariant density $\pi(x) = m(x)/\int_l^r m(x)\,dx$. If the initial distribution is the invariant distribution, then X_t is stationary (Arnold [2]).

Below, we will verify mixing properties for some interesting diffusion processes based on this theorem.

3.2.1 OU diffusion process

The OU diffusion process X_t is the strong solution of the SDE

$$dX_t = \mu X_t\,dt + \sigma\,dB_t, \tag{3.2}$$

with $l = -\infty$ and $r = \infty$, where $\mu < 0$ and $\sigma > 0$.

For this model, $\mu(x) = \mu x$ and $\sigma(x) = \sigma$ are linear functions, which implies that A.1 holds. As $s(z) = \exp\{-\mu(z^2 - z_0^2)/\sigma^2\}$ and $\lim_{|z|\to\infty} s(z) = +\infty$, we have $S(l) = -\infty$ and $S(r) = +\infty$, so then A.2 holds. Obviously, $\mu(x)/\sigma(x) - \sigma'(x)/2 = \mu x/\sigma$, it follows A.3. Thus, the OU diffusion process is ρ-mixing and α-mixing with exponential decay, and its invariant distribution is $N(0, \sigma^2/(-2\mu))$.

3.2.2 Vasicek diffusion process

The Vasicek diffusion process X_t is the strong solution of the SDE

$$dX_t = (\mu_1 X_t + \mu_0)\,dt + \sigma\,dB_t, \tag{3.3}$$

with $l = -\infty$ and $r = \infty$, where $\mu_1 < 0$, $-\infty < \mu_0 < \infty$ and $\sigma > 0$.

For this model, $\mu(x) = \mu_1 x + \mu_0$ and $\sigma(x) = \sigma$ are linear functions, which implies that A.1 holds. It is easy to get that

$$s(z) = \exp\left\{-\frac{\mu_1(z^2 - z_0^2)}{\sigma^2} - \frac{2\mu_0(z - z_0)}{\sigma^2}\right\},$$

which implies $\lim_{|z|\to\infty} s(z) = +\infty$, so A.2 holds. Note that $\mu(x)/\sigma(x) - \sigma'(x)/2 = (\mu_1 x + \mu_0)/\sigma$ and $\mu_1 < 0$, we know that A.3 holds. Therefore, the Vasicek diffusion process is ρ-mixing and α-mixing with exponential decay, and its invariant distribution is $N(-\mu_0/\mu_1, (\sigma/\sqrt{-2\mu_1})^2)$.

3.2.3 CIR diffusion process

The CIR diffusion process X_t is the strong solution of the SDE

$$dX_t = (\mu_1 X_t + \mu_0)\,dt + \sigma\sqrt{X_t}\,dB_t, \tag{3.4}$$

with $l = 0$ and $r = \infty$, where $\mu_1 < 0$, $\mu_0 > 0$ and $\sigma > 0$. We suppose that $4\mu_0 > \sigma^2$.

For this model, $\mu(x) = \mu_1 x + \mu_0$ and $\sigma(x) = \sigma\sqrt{x}$, so A.1 holds. And

$$s(z) = \exp\bigl\{-2\sigma^{-2}\mu_1(z - z_0) - 2\sigma^{-2}\mu_0 \ln(z/z_0)\bigr\}.$$

Hence $\lim_{z\to+\infty} s(z) = +\infty$ and $\lim_{z\to 0} s(z) = +\infty$, it implies the condition A.2 is satisfied. Moreover,

$$\mu(x)/\sigma(x) - \sigma'(x)/2 = \frac{4\mu_1 x + 4\mu_0 - \sigma^2}{4\sigma\sqrt{x}},$$

which follows the condition A.3 for $4\mu_0 > \sigma^2$. Therefore, the CIR diffusion process is ρ-mixing and α-mixing with exponential decay, and its invariant density is

$$\pi(x) = \frac{(-2\mu_1/\sigma^2)^{2\mu_0/\sigma^2}}{\Gamma(2\mu_0/\sigma^2)} x^{2\mu_0/\sigma^2 - 1} e^{-(-2\mu_1/\sigma^2)x},$$

which is the density of gamma distribution.

3.2.4 Generalized CIR diffusion process

The generalized CIR diffusion process X_t is the strong solution of the SDE

$$dX_t = \beta(\tau - X_t)\,dt + \sqrt{\sigma^2 + \lambda(X_t - \mu)^2}\,dB_t, \tag{3.5}$$

with $l = -\infty$ and $r = \infty$, where $\beta > 0, \tau \geq 0, \sigma > 0, \lambda > 0, -\infty < \mu < \infty$ (Nicolau [33]).

For the case, $\mu(x) = \beta(\tau - x)$ and $\sigma(x) = \sqrt{\sigma^2 + \lambda(x - \mu)^2}$. We have that $s(z) = e^{-g(z)+g(z_0)}$, where

$$g(z) = \frac{2\beta(\tau - \mu)}{\sigma\sqrt{\lambda}} \arctan\left(\frac{\sqrt{\lambda}(z - \mu)}{\sigma}\right) - \frac{\beta}{\lambda} \ln\bigl(\sigma^2 + \lambda(z - \mu)^2\bigr).$$

Hence $\lim_{z\to\pm\infty} s(z) = \lim_{z\to\pm\infty} e^{-g(z)+g(z_0)} = +\infty$, it follows condition A.2. Note that

$$\frac{\mu(x)}{\sigma(x)} - \frac{\sigma'(x)}{2} = \frac{2\beta(\tau - \mu) + (\lambda + 2\beta)(\mu - x)}{2\sqrt{\sigma^2 + \lambda(\mu - x)^2}},$$

we have that

$$\limsup_{x\to+\infty}\left(\frac{\mu(x)}{\sigma(x)} - \frac{\sigma'(x)}{2}\right) = -\frac{\lambda + 2\beta}{2\sqrt{\lambda}} < 0,$$

$$\liminf_{x\to-\infty}\left(\frac{\mu(x)}{\sigma(x)}-\frac{\sigma'(x)}{2}\right)=\frac{\lambda+2\beta}{2\sqrt{\lambda}}>0.$$

It implies condition A.3. Therefore, the diffusion process is ρ-mixing and α-mixing with exponential decay. Its invariant density is

$$\pi(x)\propto\frac{e^{g(x)}}{\sigma^2+\lambda(x-\mu)^2}.$$

3.2.5 CKLS diffusion process

The CKLS diffusion process X_t is the strong solution of the SDE

$$dX_t=(\mu_1X_t+\mu_0)\,dt+\sigma X_t^\gamma\,dB_t \tag{3.6}$$

with $l=0$ and $r=\infty$, where $\mu_1<0$, $\mu_0>0$, $\sigma>0$ and $\gamma>0$.

For this process, $\mu(x)=\mu_1x+\mu_0$ and $\sigma(x)=\sigma x^\gamma$, so A.1 holds. To verify A.2 and A.3, we will discuss several situations of γ.

(1) $0<\gamma<1/2$. Let

$$g(z)=\frac{2}{\sigma^2}\left(\frac{\mu_1}{2(1-\gamma)}z^{2(1-\gamma)}+\frac{\mu_0}{1-2\gamma}z^{1-2\gamma}\right)$$

Then, $s(z)=\exp\{-g(z)+g(z_0)\}$. $\lim_{z\to+\infty}s(z)=+\infty$, but $\lim_{z\to0}s(z)=\exp\{g(z_0)\}\neq+\infty$, which implies that $S(0)\neq+\infty$. So, A.2 is not satisfied for this case.

(2) $\gamma=1/2$. It is the CIR diffusion process that has discussed before.

(3) $1/2<\gamma<1$. At that time, $s(z)=\exp\{-g(z)+g(z_0)\}$. So, $\lim_{z\to+\infty}s(z)=+\infty$ and $\lim_{z\to0}s(z)=+\infty$. And

$$\frac{\mu(x)}{\sigma(x)}-\frac{\sigma'(x)}{2}=\frac{\mu_1x^{1-\gamma}+\mu_0x^{-\gamma}}{\sigma}-\frac{\sigma\gamma x^{\gamma-1}}{2}=\begin{cases}<0 & \text{as } x\to+\infty,\\ >0 & \text{as } x\to0.\end{cases} \tag{3.7}$$

Therefore, the conditions of Theorem 3.2 are satisfied for the case.

(4) $\gamma=1$. For this case, $s(z)=z^{-2\mu_1/\sigma^2}e^{\frac{2\mu_0}{\sigma^2}z^{-1}}z_0^{2\mu_1/\sigma^2}e^{-\frac{2\mu_0}{\sigma^2}z_0^{-1}}$. Then, $\lim_{z\to+\infty}s(z)=+\infty$ and $\lim_{z\to0}s(z)=+\infty$. And

$$\frac{\mu(x)}{\sigma(x)}-\frac{\sigma'(x)}{2}=\frac{\mu_1+\mu_0x^{-1}}{\sigma}-\frac{\sigma}{2}=\begin{cases}<0 & \text{as } x\to+\infty,\\ >0 & \text{as } x\to0.\end{cases}$$

Thus, the conditions of Theorem 3.2 are satisfied for the case.

(5) $\gamma>1$. At that time, $s(z)=\exp\{-g(z)+g(z_0)\}$. So $\lim_{z\to+\infty}s(z)=\exp\{g(z_0)\}>0$ and $\lim_{z\to0}s(z)=+\infty$, and (3.7) holds. Hence, the conditions of Theorem 3.2 are satisfied for the case.

As discussed above, the CKLS diffusion process is ρ-mixing and α-mixing with exponential decay for $\gamma\geq1/2$.

3.2.6 Logistic diffusion process

The Logistic diffusion process X_t is the strong solution of the SDE

$$dX_t=\alpha X_t(1-\beta X_t)\,dt+\sigma X_t\,dB_t \tag{3.8}$$

with $l = 0$ and $r = \infty$, where $\alpha > 0$, $\beta > 0$, $\sigma > 0$ and $\sigma^2 < 2\alpha$. The diffusion process is useful for modeling the population systems under environmental noise (Bahar and Mao [3]; Mao [32]).

For this process, $\mu(x) = \alpha x(1-\beta x)$ and $\sigma(x) = \sigma x$. After calculation, we have

$$s(z) = \exp\left\{-2\alpha\left[(\ln z - \beta z) - (\ln z_0 - \beta z_0)\right]/\sigma^2\right\} \propto z^{-2\alpha/\sigma^2} e^{2\alpha\beta z/\sigma^2}.$$

Hence $\lim_{z\to+\infty} s(z) = +\infty$ and $\lim_{z\to 0} s(z) = +\infty$, it implies that $S(l) = -\infty$ and $S(r) = +\infty$.

$$\frac{\mu(x)}{\sigma(x)} - \frac{\sigma'(x)}{2} = \frac{\alpha(1-\beta x)}{\sigma} - \frac{\sigma}{2} = \begin{cases} \frac{2\alpha-\sigma^2}{2\sigma} > 0 & \text{as } x \to 0, \\ < 0 & \text{as } x \to +\infty. \end{cases}$$

Hence, the conditions of Theorem 3.2 are satisfied. So, the Logistic diffusion process is ρ-mixing and α-mixing with exponential decay. Its invariant density is

$$\pi(x) \propto x^{(2\alpha-\sigma^2)/\sigma^2-1} e^{-2\alpha\beta x/\sigma^2},$$

which is the density of gamma distribution.

3.2.7 Double-well diffusion process

The double-well diffusion X_t is the strong solution of the SDE

$$dX_t = \alpha X_t\left(\gamma^2 - X_t^2\right)dt + \sigma\, dB_t \tag{3.9}$$

with $l = -\infty$ and $r = \infty$, where $\alpha > 0$, $-\infty < \gamma < \infty$, $\sigma > 0$. This diffusion process is ergodic, and its invariant measure is the bimodal distribution with modes at $x = \pm\gamma$ and with density

$$\pi(x) \propto \exp\left\{-\frac{\alpha}{4\sigma^2}x^2\left(x^2 - 2\gamma^2\right)\right\}.$$

It is a widely used benchmark for nonlinear inference problems. The parameter α governs the rate at which sample trajectories are pushed toward either mode. If σ is small in comparison to α, mode-switching occurs relatively rarely.

For this process, $\mu(x) = \alpha x(\gamma^2 - x^2)$ and $\sigma(x) = \sigma$, so A.1 holds, and

$$s(z) = \exp\left\{-\frac{\alpha}{4\sigma^2}\left[\left(2\gamma^2 z^2 - z^4\right) - \left(2\gamma^2 z_0^2 - z_0^4\right)\right]\right\}.$$

Hence $\lim_{|z|\to+\infty} s(z) = +\infty$, it implies A.2 is satisfied, and

$$\frac{\mu(x)}{\sigma(x)} - \frac{\sigma'(x)}{2} = \frac{\alpha x(\gamma^2 - x^2)}{\sigma} = \begin{cases} +\infty & \text{as } x \to -\infty, \\ -\infty & \text{as } x \to +\infty. \end{cases}$$

It follows A.3. Thus, the double-well diffusion process is ρ-mixing and α-mixing with exponential decay.

3.2.8 Generalized logistic diffusion process

The generalized logistic diffusion process X_t is the strong solution of the SDE

$$\begin{aligned} dX_t = {} & \bigl\{(\theta_1 - \theta_2)\cosh(X_t/2) - (\theta_1 + \theta_2)\sinh(X_t/2)\bigr\}\cosh(X_t/2)\,dt \\ & + 2\cosh(X_t/2)\,dB_t \end{aligned} \tag{3.10}$$

with $l = -\infty$ and $r = \infty$, where $\sinh(x) = (e^x - e^{-x})/2, \cosh(x) = (e^x + e^{-x})/2, \theta_1 > 0$ and $\theta_2 > 0$. This diffusion is ergodic, and its invariant measure is the generalized logistic distribution with density

$$\pi(x) = B(\theta_1 + 1, \theta_2 + 1)e^{(\theta_1+1)x}\bigl(1 + e^x\bigr)^{-(\theta_1+\theta_2+2)},$$

here $B(a, b)$ denotes the beta function. It is used in many areas of application, e.g., mathematical finance and turbulence (Kessler and Sørensen [27]).

After simple calculation, it can be concluded that

$$s(z) \propto e^{-(\theta_1-\theta_2)z/2}\bigl(e^{z/2} + e^{-z/2}\bigr)^{\theta_1+\theta_2},$$

which follows that $\lim_{|z|\to+\infty} s(z) = +\infty$. So $S(l) = -\infty$ and $S(r) = +\infty$. Moreover,

$$\begin{aligned} \frac{\mu(x)}{\sigma(x)} - \frac{\sigma'(x)}{2} &= \frac{(\theta_1 - \theta_2)\cosh(x/2) - (\theta_1 + \theta_2 + \sigma)\sinh(x/2)}{2\sigma} \\ &= \frac{-(2\theta_2 + \sigma)e^{x/2} + (2\theta_1 + \sigma)e^{-x/2}}{4\sigma} \\ &= \begin{cases} +\infty & \text{as } x \to -\infty, \\ -\infty & \text{as } x \to +\infty. \end{cases} \end{aligned}$$

Hence, the conditions of Theorem 3.2 are satisfied. Thus, the generalized logistic diffusion process is ρ-mixing and α-mixing with exponential decay.

4 Contrast estimation of the Ornstein–Uhlenberck (OU) integrated diffusion process

As an application of the moment inequalities in Sect. 2, we discuss the strong consistency of parameter estimates for the following OU-integrated diffusion process

$$\begin{cases} Y_t = \int_0^t X_s\,ds, \\ dX_t = \mu X_t\,dt + \sigma\,dB_t, \end{cases} \tag{4.1}$$

where $\mu < 0$ and $\sigma > 0$ are unknown parameters. We assume the initial condition $X_0 \sim N(0, -\sigma^2/2\mu)$, which is the invariant distribution of the diffusion process, to be independent of B_t. Generally, the integrated diffusion process

$$\begin{cases} dY_t = X_t\,dt, \\ dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dB_t, \end{cases} \tag{4.2}$$

where $\mu(x)$ and $\sigma(x)$ are the drift and diffusion coefficients.

Many scholars have studied the integrated process. Gloter [19] studied the asymptotic representation of the integrated diffusion process and showed the consistency and asymptotic mixed normality of the minimum contrast estimate of the diffusion coefficient. Gloter [20] proved limit theorems for functionals associated with the observations of the integrated diffusion process, applied these results to obtain a contrast function, and showed the associated minimum contrast estimators are consistent and asymptotically Gaussian with different rates for drift and diffusion coefficient parameters. Applying these results to the OU-integrated diffusion process, the consistency and asymptotic normality of parameter estimation are obtained. Ditlevsen and Sørensen [15] studied the statistical inference problem of the integrated diffusion process with some weight function, obtained an estimation function based on the optimal prediction, and proved that the estimates are consistent and asymptotically normal. The method is applied to inference based on integrated data from the OU process and from the CIR model, for both of which an explicit optimal estimating function is found. Nicolau [33] studied the Nadaraya–Watson kernel estimates of the drift and diffusion coefficients and proved that the estimates are weakly consistent and asymptotically normal. Yang et al. [50] improved the asymptotic property of the nonparametric kernel estimate of Nicolau [33] by generalizing weak consistency to strong consistency under weaker conditions. Gloter and Gobet [21] proved the local asymptotic mixed normality property for the statistical model given by the observation of local means of a diffusion process. Using discrete observations of the integrated diffusion process, Comte et al. [14] established a nonparametric adaptive estimation based on penalized least squares methods for both the drift function and the diffusion coefficient of the unobserved diffusion, which is a stationary and β-mixing diffusion. Wang and Lin [41] proposed a local linear estimation of the diffusion coefficient. Wang at al. [42] proposed a re-weighted estimator of the diffusion coefficient in the second-order diffusion model and showed the consistency and the asymptotic normality of the estimator under appropriate conditions.

In the literature mentioned above, there is relatively little discussion on the strong consistency of estimation. Yang et al. [50] only studied the strong consistency of the nonparametric kernel estimates of the drift and diffusion coefficients for the model (4.2). In this section, we will provide sufficient conditions for strong consistency of parameter estimates for the model (4.1).

4.1 Contrast estimation of the OU-integrated diffusion process

We introduce the notation

$$\overline{X}_i = \Delta_n^{-1} \int_{(i-1)\Delta_n}^{i\Delta_n} X_s \, ds = \Delta_n^{-1} (Y_{i\Delta_n} - Y_{(i-1)\Delta_n}), \quad i \geq 1. \tag{4.3}$$

According to Gloter [20], we obtain the contrast function for the OU-integrated diffusion process

$$\mathcal{L}_n(\theta) = \sum_{i=1}^{n} \left(\frac{3}{2\Delta_n} \frac{(\overline{X}_{i+1} - \overline{X}_i - \mu \Delta_n \overline{X}_i)^2}{\sigma^2} + \frac{3\mu}{4\sigma^2} (\overline{X}_{i+1} - \overline{X}_i)^2 + \log\bigl(\sigma^2\bigr) \right). \tag{4.4}$$

The contrast estimator $\widehat{\theta}_n = \operatorname{arginf}_{\theta \in \Theta} \mathcal{L}_n(\theta)$, where $\theta = (\mu, \sigma^2)$.

The equations are obtained by differentiating the contrast function

$$\begin{cases} \sum_{i=1}^{n}\left(\frac{-3}{\Delta_n}\frac{(\overline{X}_{i+1}-\overline{X}_i-\mu\Delta_n\overline{X}_i)\Delta_n\overline{X}_i}{\sigma^2}+\frac{3}{4\sigma^2}(\overline{X}_{i+1}-\overline{X}_i)^2\right)=0, \\ \sum_{i=1}^{n}\left(-\frac{3}{2\Delta_n}\frac{(\overline{X}_{i+1}-\overline{X}_i-\mu\Delta_n\overline{X}_i)^2}{\sigma^4}-\frac{3\mu}{4\sigma^4}(\overline{X}_{i+1}-\overline{X}_i)^2+\frac{1}{\sigma^2}\right)=0. \end{cases}$$

This is equivalent to

$$\begin{cases} 4\sum_{i=1}^{n}(\overline{X}_{i+1}-\overline{X}_i-\mu\Delta_n\overline{X}_i)\overline{X}_i=\sum_{i=1}^{n}(\overline{X}_{i+1}-\overline{X}_i)^2, \\ \sum_{i=1}^{n}\left(\frac{3}{2\Delta_n}(\overline{X}_{i+1}-\overline{X}_i-\mu\Delta_n\overline{X}_i)^2+\frac{3\mu}{4}\sum_{i=1}^{n}(\overline{X}_{i+1}-\overline{X}_i)^2=n\sigma^2. \right. \end{cases}$$

Hence, the contrast estimators of μ and σ are

$$\widehat{\mu}_n=\frac{\sum_{i=1}^{n}(\overline{X}_{i+1}-\overline{X}_i)\overline{X}_i}{\Delta_n\sum_{i=1}^{n}\overline{X}_i^2}-\frac{\sum_{i=1}^{n}(\overline{X}_{i+1}-\overline{X}_i)^2}{4\Delta_n\sum_{i=1}^{n}\overline{X}_i^2}, \tag{4.5}$$

$$\begin{aligned} \widehat{\sigma}_n^2={}&\frac{3}{2n\Delta_n}\sum_{i=1}^{n}(\overline{X}_{i+1}-\overline{X}_i)^2-\frac{3\widehat{\mu}_n}{n}\sum_{i=1}^{n}(\overline{X}_{i+1}-\overline{X}_i)\overline{X}_i \\ &+\frac{3\widehat{\mu}_n^2\Delta_n}{2n}\sum_{i=1}^{n}\overline{X}_i^2+\frac{3\widehat{\mu}_n}{4n}\sum_{i=1}^{n}(\overline{X}_{i+1}-\overline{X}_i)^2. \end{aligned} \tag{4.6}$$

From the Itô formula and $X_t\sim N(0,-\sigma^2/2\mu)$, we can obtain

$$E(\overline{X}_i^2)=-\frac{\sigma^2}{2\mu}+O(\Delta), \tag{4.7}$$

$$E(\overline{X}_{i+1}-\overline{X}_i)^2=\frac{2}{3}\sigma^2\Delta_n+O(\Delta_n^2), \tag{4.8}$$

$$E\left[(\overline{X}_{i+1}-\overline{X}_i)\overline{X}_i\right]=-\frac{1}{3}\sigma^2\Delta_n+O(\Delta_n^2). \tag{4.9}$$

Therefore,

$$E\widehat{\mu}_n=\mu+O(\Delta_n), \tag{4.10}$$

That is, $\widehat{\mu}_n$ is an asymptotically unbiased estimator of μ. In (4.6), the first term on the right is the asymptotically unbiased term of σ^2, and the rest of the terms converge to zero. So, the estimator of σ^2 can be written as

$$\widehat{\sigma}_n^2=\frac{3}{2n\Delta_n}\sum_{i=1}^{n}(\overline{X}_{i+1}-\overline{X}_i)^2. \tag{4.11}$$

4.2 Mean square error and optimal interval

Theorem 4.1 *The mean squared errors of $\widehat{\mu}_n$ and $\widehat{\sigma}_n^2$ are*

$$\mathrm{MSE}(\widehat{\mu}_n)=\frac{2|\mu|}{n\Delta_n}+\frac{25}{144}\mu^4\Delta_n^2+o\left(\frac{1}{n\Delta_n}+\Delta_n^2\right), \tag{4.12}$$

$$\mathrm{MSE}(\widehat{\sigma}_n^2)=\frac{9\sigma^4}{4n}+\frac{9}{16}\mu^2\sigma^4\Delta_n^2+o\left(\frac{1}{n}+\Delta_n^2\right). \tag{4.13}$$

To prove the theorem, we need the following lemma.

Lemma 4.1 *Suppose X_t is the diffusion process in (4.2) and $\mathcal{F}_t = \sigma(X_s, s \le t)$, then*

$$
\begin{aligned}
&E\big(\overline{X}_i^2 | \mathcal{F}_{(i-1)\Delta_n}\big) \\
&\quad = X_{(i-1)\Delta_n}^2 + X_{(i-1)\Delta_n}\mu(X_{(i-1)\Delta_n})\Delta_n + \frac{1}{3}\sigma^2(X_{(i-1)\Delta_n})\Delta_n + O\big(\Delta_n^2\big), \qquad (4.14)\\
&E\big((\overline{X}_{i+1} - \overline{X}_i)^2 | \mathcal{F}_{(i-1)\Delta_n}\big) \\
&\quad = \frac{2}{3}\sigma^2(X_{(i-1)\Delta_n})\Delta_n + \mu^2(X_{(i-1)\Delta_n})\Delta_n^2 + f(X_{(i-1)\Delta_n})\Delta_n^2 + O\big(\Delta_n^3\big), \qquad (4.15)
\end{aligned}
$$

where $f(x) = \sigma^2(x)\mu'(x) + \frac{4}{3}\{\mu(x)\sigma(x)\sigma'(x) + \frac{1}{2}\sigma^2(x)(\sigma'(x))^2 + \frac{1}{2}\sigma^3(x)\sigma''(x)\}$.

Further, if X_t is stationary, then

$$
\begin{aligned}
&E\big\{(\overline{X}_{i+1} - \overline{X}_i)\overline{X}_i\big\} \\
&\quad = -\frac{1}{3}\sigma^2(X_{i\Delta_n})\Delta_n - \frac{1}{2}\mu^2(X_{i\Delta_n})\Delta_n^2 - \frac{1}{2}f(X_{i\Delta_n})\Delta_n^2 + O\big(\Delta_n^3\big). \qquad (4.16)
\end{aligned}
$$

Proof From Ito's formula, (4.14) and (4.15) can be derived through some complicated calculations.

$$
\begin{aligned}
&E\big((\overline{X}_{i+1} - \overline{X}_i)^2 | \mathcal{F}_{(i-1)\Delta_n}\big) \\
&\quad = E\big((\overline{X}_{i+1} - \overline{X}_i)\overline{X}_{i+1} | \mathcal{F}_{(i-1)\Delta_n}\big) - E\big((\overline{X}_{i+1} - \overline{X}_i)\overline{X}_i | \mathcal{F}_{(i-1)\Delta_n}\big) \\
&\quad = E\big(\overline{X}_{i+1}^2 | \mathcal{F}_{(i-1)\Delta_n}\big) - E\big(\overline{X}_{i+1}\overline{X}_i | \mathcal{F}_{(i-1)\Delta_n}\big) - E\big((\overline{X}_{i+1} - \overline{X}_i)\overline{X}_i | \mathcal{F}_{(i-1)\Delta_n}\big) \\
&\quad = E\big(\overline{X}_{i+1}^2 | \mathcal{F}_{(i-1)\Delta_n}\big) - E\big(\overline{X}_i^2 | \mathcal{F}_{(i-1)\Delta_n}\big) - 2E\big((\overline{X}_{i+1} - \overline{X}_i)\overline{X}_i | \mathcal{F}_{(i-1)\Delta_n}\big)
\end{aligned}
$$

According to the stationary of the process, we have $E\overline{X}_{i+1}^2 = E\overline{X}_i^2$. Thus,

$$
E(\overline{X}_{i+1} - \overline{X}_i)^2 = -2E\big\{(\overline{X}_{i+1} - \overline{X}_i)\overline{X}_i\big\}.
$$

From this equation and (4.15), we get (4.16). This completes the proof. □

Proof of Theorem 4.1 Since the OU process is stationary and $X_t \sim N(0, -\sigma^2/2\mu)$, it follows from Lemma 4.1 that

$$
\begin{aligned}
&E\big[(\overline{X}_{i+1} - \overline{X}_i)^2\big] \\
&\quad = E\bigg[\frac{2}{3}\sigma^2\Delta_n + X_{i\Delta}^2\mu^2\Delta_n^2 + \mu\sigma^2\Delta_n^2 + O\big(\Delta_n^3\big)\bigg] \\
&\quad = \frac{2}{3}\sigma^2\Delta_n + (\sigma/\sqrt{-2\mu})^2\mu^2\Delta_n^2 + \mu\sigma^2\Delta_n^2 + O\big(\Delta_n^3\big), \\
&\quad = \frac{2}{3}\sigma^2\Delta_n + \frac{1}{2}\mu\sigma^2\Delta_n^2 + O\big(\Delta_n^3\big), \\
&E\big\{(\overline{X}_{i+1} - \overline{X}_i)\overline{X}_i\big\} \\
&\quad = E\bigg[-\frac{1}{3}\sigma^2\Delta_n - \frac{1}{2}X_{i\Delta}^2\mu^2\Delta_n^2 - \frac{1}{2}\mu\sigma^2\Delta_n^2 + O\big(\Delta_n^3\big)\bigg]
\end{aligned}
$$

$$
\begin{aligned}
&= -\frac{1}{3}\sigma^2\Delta_n - \frac{1}{4}\mu\sigma^2\Delta_n^2 + O(\Delta_n^3), \\
&E(\overline{X}_i^2) \\
&= EX^2_{(i-1)\Delta_n} + \mu\Delta_n EX^2_{(i-1)\Delta_n} + \frac{1}{3}\sigma^2\Delta_n + O(\Delta_n^2) \\
&= (\sigma/\sqrt{-2\mu})^2 + (\sigma/\sqrt{-2\mu})^2\mu\Delta_n + \frac{1}{3}\sigma^2\Delta_n + O(\Delta_n^2) \\
&= -\frac{\sigma^2}{2\mu} - \frac{1}{6}\sigma^2\Delta_n + O(\Delta_n^2).
\end{aligned}
$$

So,

$$
\begin{aligned}
E\widehat{\mu}_n &= \frac{-\frac{1}{3}\sigma^2\Delta_n - \frac{1}{4}\mu\sigma^2\Delta_n^2 - \frac{1}{4}\{\frac{2}{3}\sigma^2\Delta_n + \frac{1}{2}\mu\sigma^2\Delta_n^2\}}{\{-\frac{\sigma^2}{2\mu} - \frac{1}{6}\sigma^2\Delta_n\}\Delta_n} + O(\Delta_n^2) \\
&= \frac{-\frac{1}{3} - \frac{1}{4}\mu\Delta_n - \frac{1}{4}\{\frac{2}{3} + \frac{1}{2}\mu\Delta_n\}}{\{-\frac{1}{2\mu} - \frac{1}{6}\Delta_n\}} + O(\Delta_n^2) \\
&= \frac{\mu + \frac{3}{4}\mu^2\Delta_n}{1 + \frac{1}{3}\mu\Delta_n} + O(\Delta_n^2).
\end{aligned}
$$

Using the Taylor expansion to expand the function $\frac{1}{x}$ at $x = 1$, we get $\frac{1}{x} = 1 - (x-1) + O(x-1)^2$. This yields

$$
\begin{aligned}
E\widehat{\mu}_n &= \left\{\mu + \frac{3}{4}\mu^2\Delta_n\right\}\left\{1 - \frac{1}{3}\mu\Delta_n + O(\Delta_n^2)\right\} + O(\Delta_n^2) \\
&= \mu + \frac{3}{4}\mu^2\Delta_n - \frac{1}{3}\mu^2\Delta_n + O(\Delta_n^2) \\
&= \mu + \frac{5}{12}\mu^2\Delta_n + O(\Delta_n^2).
\end{aligned}
$$

Thus, the asymptotic biased term of $\widehat{\mu}_n$ is

$$
\mathrm{Bias}(\widehat{\mu}_n) = \frac{5}{12}\mu^2\Delta_n + O(\Delta_n^2).
$$

By Gloter [20], $\sqrt{n\Delta_n}(\widehat{\mu}_n - \mu) \xrightarrow{d} N(0, 2|\mu|)$. Thus, the asymptotic variance is

$$
\mathrm{Var}(\widehat{\mu}_n) = \frac{2|\mu|}{n\Delta_n} + o\left(\frac{1}{n\Delta_n}\right).
$$

Therefore, we obtain the mean square error (4.12).

On the other hand, by Gloter [20], $\sqrt{n}(\widehat{\sigma}_n^2 - \sigma^2) \xrightarrow{d} N(0, 9\sigma^4/4)$. It follows $\mathrm{Var}(\widehat{\sigma}_n^2) = \frac{9\sigma^4}{4n} + o(1/n)$. From the Itô formula, it is easy to get that

$$
E(\widehat{\sigma}_n^2) = \sigma^2 + \frac{3}{4}\mu\sigma^2\Delta_n + O(\Delta_n^2),
$$

Therefore, we obtain the mean square error (4.13). This completes the proof. □

Table 1 Simulation estimate values

n	$\Delta_{\mu,\mathrm{opt}}$	$\widehat{\mu}_n$	$\widehat{\sigma}_n$
500	0.2336	−0.9561 (0.1237)	0.9286 (0.0345)
1000	0.1854	−0.9614 (0.0986)	0.9450 (0.0226)
5000	0.1084	−0.9687 (0.0552)	0.9650 (0.0106)
10000	0.0861	−0.9705 (0.0459)	0.9720 (0.0069)

From (4.12), we get the optimal interval for $\widehat{\mu}_n$ as

$$\Delta_{\mu,\mathrm{opt}} = \left(-\frac{144}{25\mu^3}\right)^{1/3} n^{-1/3}. \tag{4.17}$$

$\mathrm{MSE}(\widehat{\sigma}_n^2)$ is a monotonically decreasing function with respect to Δ_n. The smaller Δ_n, the smaller the mean square error. Therefore, there is no optimal interval for $\widehat{\sigma}_n^2$.

Now we use numerical simulations to demonstrate the performance of the optimal interval. Consider the Euler discrete form of the OU diffusion process

$$X_{i\Delta_n} = X_{(i-1)\Delta_n} + \mu X_{(i-1)\Delta_n}\Delta_n + \sigma\sqrt{\Delta_n}\varepsilon_i,$$

where $\varepsilon_i \sim N(0,1)$. Given $\mu = -1, \sigma^2 = 1$.

In practice, since μ and σ^2 are unknown, the optimal interval $\Delta_{\mu,\mathrm{opt}}$ cannot be obtained. Here, we use a simulation method to estimate the optimal interval. Based on the expression of the optimal interval, we select $\Delta_n = n^{-1/3}$ and $n = 10000$ and generate samples to obtain the estimates of μ and σ as follows

$$\widehat{\mu}_n = -0.9667, \qquad \widehat{\sigma}_n = 0.9819,$$

where $\widehat{\sigma}_n = \sqrt{\widehat{\sigma}_n^2}$. As a result, the optimal interval is

$$\Delta_{\mu,\mathrm{opt}} = \left(-\frac{144}{25\widehat{\mu}_n^3}\right)^{1/3} n^{-1/3} = 1.8543n^{-1/3}.$$

Let $\Delta_n = \Delta_{\mu,\mathrm{opt}}$. For different sample sizes n, the samples $\{X_{i\Delta_n}, i = 1,2,3\cdots n\}$ can be generated using the Euler discretization model. To obtain the samples $\overline{X}_{i\Delta_n}$, each time interval $[(i-1)\Delta_n, i\Delta_n]$ is equally divided into some small intervals. Then, we can generate the sample X_t using the Euler discretization model again and get the approximation of the integral $\int_{(i-1)\Delta_n}^{i\Delta_n} X_s\,ds$. Finally, we obtain the estimates $\widehat{\mu}_n$ and $\widehat{\sigma}_n^2$, and the simulation results in Table 1 by repeating simulation, where the numerical values in parentheses are standard deviations. The results show that as the sample size n gradually increases, the estimated values $\widehat{\mu}_n$ and $\widehat{\sigma}_n^2$ are closer to the true values, and the standard deviations gradually decrease. It implies that the optimal interval is effective.

4.3 Strong consistency of estimation

Gloter [20] gave weak consistent and asymptotically normal properties for $\widehat{\mu}_n$ and $\widehat{\sigma}_n^2$. Let us now discuss the strong consistency of the estimators.

Theorem 4.2 *Suppose there exist real numbers $a \in (0,1)$ such that $\Delta_n \to 0$ and $n^{1-a}\Delta_n \to \infty$. Then*

$$\widehat{\sigma}_n^2 \xrightarrow{a.s.} \sigma^2. \tag{4.18}$$

Theorem 4.3 *Suppose there exist real numbers $b \in (0,1)$ such that $\Delta_n \to 0$ and $n^{1-b}\Delta_n^2 \to \infty$. Then*

$$\widehat{\mu}_n \xrightarrow{a.s.} \mu. \tag{4.19}$$

The conditions of Theorem 4.3 are stronger than those of Theorem 4.2. The optimal interval $\Delta_{\mu,\mathrm{opt}}$ of $\widehat{\mu}_n$ satisfies the conditions of Theorem 4.3. The proof of the theorem requires the following Lévy continuous modulus.

Lemma 4.2 (Lévy modulus of continuity of diffusions)

$$P\left(\limsup_{\Delta_n\to 0} \frac{k_n}{(\Delta_n \log(1/\Delta_n))^{1/2}} = k_0\right) = 1,$$

where k_0 is a constant,

$$k_n = \max_{1\le i\le n} \sup_{(i-1)\Delta_n \le s\le i\Delta_n} |X_s - X_{i\Delta_n}|,$$

or

$$k_n = \max_{1\le i\le n} \sup_{(i-1)\Delta_n \le s\le i\Delta_n} |X_s - X_{(i-1)\Delta_n}|.$$

Proof The conclusion of the lemma is obtained from Theorem 7.2.5 of Arnold ([2], P121), see also Bandi and Phillips ([4], (7.7)). This completes the proof. □

We can easily generalize the Lévy continuous modulus to integral diffusion processes as follows.

Lemma 4.3 *Denoting $\beta_n = (\Delta_n \log(1/\Delta_n)^{1/2})$, we have*

$$\max_{1\le i\le n} |X_{i\Delta_n} - X_{(i-1)\Delta_n}| = O_{a.s.}(\beta_n),$$

$$\max_{1\le i\le n} |\overline{X}_i - X_{i\Delta_n}| = O_{a.s.}(\beta_n),$$

$$\max_{1\le i\le n} |\overline{X}_i - X_{(i-1)\Delta_n}| = O_{a.s.}(\beta_n),$$

$$\max_{1\le i\le n} |\overline{X}_i - \overline{X}_i| = O_{a.s.}(\beta_n).$$

Proof of Theorem 4.2 Since $E\widehat{\sigma}_n^2 = \sigma^2 + O(\Delta_n)$, we only need to prove that $\widehat{\sigma}_n^2 - E\widehat{\sigma}_n^2 \xrightarrow{a.s} 0$. Let

$$Z_{i\Delta_n} = (\overline{X}_{i+1} - \overline{X}_i)^2 - E(\overline{X}_{i+1} - \overline{X}_i)^2.$$

Then $\widehat{\sigma}_n^2 - E\widehat{\sigma}_n^2 = \frac{3}{2n\Delta_n}\sum_{i=1}^n Z_{i\Delta_n}$. In Sect. 3, we have verified that the OU process X_t is a geometrically decaying ρ-mixing process. It implies that $\{\overline{X}_i, 1 \le i \le n\}$ are ρ-mixing with geometrical decay. By the moment inequality (2.12) of Theorem 2.2, for any given $\varepsilon > 0$ and $r \ge 2$, we have

$$\begin{aligned} P\left(\left|\widehat{\sigma}_n^2 - E\widehat{\sigma}_n^2\right| > \varepsilon\right) &\le C(n\Delta_n)^{-r} E\left|\sum_{i=1}^n Z_{i\Delta_n}\right|^r \\ &\le C(n\Delta_n)^{-r}\left\{\lambda_n \max_{1\le j\le 2\lambda_n} E|\xi_j|^r + \left(\lambda_n \max_{1\le j\le 2\lambda_n} E|\xi_j|^2\right)^{r/2}\right\}, \end{aligned}$$

where $\xi_j = \sum_{i=(j-1)\tau_n\wedge n+1}^{j\tau_n\wedge n} Z_{i\Delta_n}$. By the Lévy continuous modulus (Lemma 4.3), we know that $|Z_{i\Delta_n}| \le C\Delta_n \log(1/\Delta_n)$. Thus,

$$E|\xi_j|^r \le \tau_n^{r-1}\sum_{i=(j-1)\tau_n\wedge n+1}^{j\tau_n\wedge n} E|Z_{i\Delta_n}|^r \le C\tau_n^r\left(\Delta_n\log(1/\Delta_n)\right)^r \le C\log^r(1/\Delta_n).$$

Therefore,

$$\begin{aligned} P\left(\left|\widehat{\sigma}_n^2 - E\widehat{\sigma}_n^2\right| > \varepsilon\right) &\le C(n\Delta_n)^{-r}\left\{\lambda_n\log^r(1/\Delta_n) + \left(\lambda_n\log^2(1/\Delta_n)\right)^{r/2}\right\} \\ &\le C(n\Delta_n)^{-r}\left\{n\Delta_n\log^r(1/\Delta_n) + \left(n\Delta_n\log^2(1/\Delta_n)\right)^{r/2}\right\} \\ &\le C(n\Delta_n)^{-r}\left(n\Delta_n\log^2(1/\Delta_n)\right)^{r/2} \\ &\le C(n\Delta_n)^{-r/2}\log^r(1/\Delta_n). \end{aligned}$$

Since $n^{1-a}\Delta_n \to \infty$, so $1/\Delta_n \le Cn^{1-a}$. Thereby,

$$(n\Delta_n)^{-r/2}\log^r(1/\Delta_n) \le Cn^{-ar/2}\log^r n.$$

Taking $r > 2/a$, then we have $\sum_{n=1}^\infty P(|\widehat{\sigma}_n^2 - E\widehat{\sigma}_n^2| > \varepsilon) < \infty$. Thereby, we have $\widehat{\sigma}_n^2 - E\widehat{\sigma}_n^2 \xrightarrow{a.s} 0$ by the Borel–Cantelli Lemma. This completes the proof. □

Proof of Theorem 4.3 We introduce the notations

$$\begin{aligned} A_{1n} &= n^{-1}\sum_{i=1}^n \overline{X}_i^2, \\ A_{2n} &= (n\Delta_n)^{-1}\sum_{i=1}^n(\overline{X}_{i+1} - \overline{X}_i)\overline{X}_i, \\ A_{3n} &= (n\Delta_n)^{-1}\sum_{i=1}^n(\overline{X}_{i+1} - \overline{X}_i)^2. \end{aligned}$$

Then, $\widehat{\mu}_n$ can be written as

$$\widehat{\mu}_n = \frac{A_{2n}}{A_{1n}} - \frac{A_{3n}}{4A_{1n}}.$$

By Theorem 4.2, we have $A_{3n} \xrightarrow{a.s.} \frac{2\sigma^2}{3}$. Moreover, $E(A_{1n}) = \frac{\sigma^2}{2|\mu|} + O(\Delta_n), EA_{2n} = -\frac{1}{3}\sigma^2 + O(\Delta_n)$. Therefore, to prove $\widehat{\mu}_n \xrightarrow{a.s.} \mu$, we only need to prove the following two facts

$$A_{1n} - EA_{1n} \xrightarrow{a.s.} 0, \qquad A_{2n} - EA_{2n} \xrightarrow{a.s.} 0.$$

(1) To prove that $A_{1n} - EA_{1n} \xrightarrow{a.s.} 0$. Let $Z_{i\Delta_n}(1) = \overline{X}_i^2 - E\overline{X}_i^2$. Then $A_{1n} - EA_{1n} = n^{-1}\sum_{i=1}^n Z_{i\Delta_n}(1)$. By the moment inequality (2.12) of Theorem 2.2, for any given $\varepsilon > 0$ and $r \geq 2$, we have

$$\begin{aligned} P\big(|A_{1n} - EA_{1n}| > \varepsilon\big) &\leq Cn^{-r}E\left|\sum_{i=1}^n Z_{i\Delta_n}(1)\right|^r \\ &\leq Cn^{-r}\left\{\lambda_n \max_{1\leq j\leq 2\lambda_n} E|\xi_j|^r + \left(\lambda_n \max_{1\leq j\leq 2\lambda_n} E|\xi_j|^2\right)^{r/2}\right\}, \end{aligned}$$

where $\xi_j = \sum_{i=(j-1)\tau_n\wedge n+1}^{j\tau_n\wedge n} Z_{i\Delta_n}(1)$. By the integral Cauchy inequality, for any $r > 1$, we have

$$\begin{aligned} E|\overline{X}_i|^r &\leq \Delta_n^{-r}E\left\{\int_{(i-1)\Delta_n}^{i\Delta_n} |X_s|\,ds\right\}^r \\ &\leq \Delta_n^{-r}E\left\{\left(\int_{(i-1)\Delta_n}^{i\Delta_n} |X_s|^r\,ds\right)^{1/r}\left(\int_{(i-1)\Delta_n}^{i\Delta_n} ds\right)^{(r-1)/r}\right\}^r \\ &= \Delta_n^{-1}E\left(\int_{(i-1)\Delta_n}^{i\Delta_n} |X_s|^r\,ds\right) \\ &= \Delta_n^{-1}\int_{(i-1)\Delta_n}^{i\Delta_n} E|X_s|^r\,ds \\ &= E|X_0|^r. \end{aligned}$$

It follows that $E|Z_{i\Delta_n}(1)|^r \leq CE|\overline{X}_i|^{2r} \leq C < \infty$. So,

$$E|\xi_j|^r \leq \tau_n^{r-1}\sum_{i=(j-1)\tau_n\wedge n+1}^{j\tau_n\wedge n} E\big|Z_{i\Delta_n}(1)\big|^r \leq C\tau_n^r \leq C\Delta_n^{-r}.$$

Thus,

$$\begin{aligned} P\big(|A_{1n} - EA_{1n}| > \varepsilon\big) &\leq Cn^{-r}\big\{\lambda_n\Delta_n^{-r} + \big(\lambda_n\Delta_n^{-2}\big)^{r/2}\big\} \\ &\leq Cn^{-r}\big\{n\Delta_n^{-r+1} + \big(n\Delta_n^{-1}\big)^{r/2}\big\} \\ &\leq Cn^{-r}\big(n\Delta_n^{-1}\big)^{r/2} \\ &\leq C(n\Delta_n)^{-r/2} \\ &\leq Cn^{-br/2}. \end{aligned}$$

Taking $r > 2/b$, then we have $\sum_{n=1}^\infty P(|A_{1n} - EA_{1n}| > \varepsilon) < \infty$. Thereby, $A_{1n} - EA_{1n} \xrightarrow{a.s.} 0$.

(2) To prove that $A_{2n} - EA_{2n} \xrightarrow{a.s.} 0$. Let

$$Z_{i\Delta_n}(2) = (\overline{X}_{i+1} - \overline{X}_i)\overline{X}_i - E\big\{(\overline{X}_{i+1} - \overline{X}_i)\overline{X}_i\big\}.$$

Then $A_{2n}-EA_{2n}=(n\Delta_n)^{-1}\sum_{i=1}^{n}Z_{i\Delta_n}(2)$. By the moment inequality (2.12) of Theorem 2.2, for any given $\varepsilon>0$ and $r\geq 2$, we have

$$\begin{aligned}P\bigl(|A_{2n}-EA_{2n}|>\varepsilon\bigr)&\leq C(n\Delta_n)^{-r}E\left|\sum_{i=1}^{n}Z_{i\Delta_n}(2)\right|^r\\&\leq C(n\Delta_n)^{-r}\left\{\lambda_n\max_{1\leq j\leq 2\lambda_n}E|\xi_j|^r+\left(\lambda_n\max_{1\leq j\leq 2\lambda_n}E|\xi_j|^2\right)^{r/2}\right\},\end{aligned}$$

where $\xi_j=\sum_{i=(j-1)\tau_n\wedge n+1}^{j\tau_n\wedge n}Z_{i\Delta_n}(2)$. By the Lévy continuous modulus, we have

$$E\bigl|Z_{i\Delta_n}(2)\bigr|^r\leq CE\bigl|(\overline{X}_{i+1}-\overline{X}_i)\overline{X}_i\bigr|^r\leq C\bigl(\Delta_n\log(1/\Delta_n)\bigr)^{r/2},$$

and

$$\begin{aligned}E|\xi_j|^r&\leq\tau_n^{r-1}\sum_{i=(j-1)\tau_n\wedge n+1}^{j\tau_n\wedge n}E\bigl|Z_{i\Delta_n}(2)\bigr|^r\\&\leq C\tau_n^r\bigl(\Delta_n\log(1/\Delta_n)\bigr)^{r/2}\\&\leq C\bigl(\Delta_n^{-1}\log(1/\Delta_n)\bigr)^{r/2}.\end{aligned}$$

Hence,

$$\begin{aligned}P\bigl(|A_{2n}-EA_{2n}|>\varepsilon\bigr)&\leq C(n\Delta_n)^{-r}\bigl\{\lambda_n\bigl(\Delta_n^{-1}\log(1/\Delta_n)\bigr)^{r/2}+\bigl(\lambda_n\Delta_n^{-1}\log(1/\Delta_n)\bigr)^{r/2}\bigr\}\\&\leq C(n\Delta_n)^{-r}\bigl\{n\Delta_n\bigl(\Delta_n^{-1}\log(1/\Delta_n)\bigr)^{r/2}+\bigl(n\log(1/\Delta_n)\bigr)^{r/2}\bigr\}\\&\leq C(n\Delta_n)^{-r}\bigl(n\log(1/\Delta_n)\bigr)^{r/2}\\&\leq C\bigl(n\Delta_n^2\bigr)^{-r/2}\bigl(\log(1/\Delta_n)\bigr)^{r/2}.\end{aligned}$$

Since $n^{1-b}\Delta_n^2\to\infty$, so $1/\Delta_n\leq Cn^{(1-b)/2}$. It follows that

$$\bigl(n\Delta_n^2\bigr)^{-r/2}\bigl(\log(1/\Delta_n)\bigr)^{r/2}\leq Cn^{-br/2}\log^{r/2}n.$$

Taking $r>2/b$, we have $\sum_{n=1}^{\infty}P(|A_{2n}-EA_{2n}|>\varepsilon)<\infty$. Thereby, $A_{2n}-EA_{2n}\xrightarrow{a.s.}0$. This completes the proof. □

5 Conclusion

This paper provides some moment inequalities for mixing long-span high-frequency data and verifies some interesting diffusion processes with mixing properties. These results indicate that mixing is feasible for studying long-span high-frequency data of some interesting models.

Acknowledgements

The authors are grateful to the referees and the editor for their valuable comments, which improved the structure and the presentation of the paper.

Funding

This research was financially supported by the Guangxi Natural Science Foundation (No. 2022GXNSFAA035516) and the Natural Science Foundation of China (No. 11861017, 11461009).

Availability of data and materials
Not applicable.

Declarations

Competing interests
The authors declare no competing interests.

Author contributions
All authors carried out the mathematical studies. All authors read and approved the final manuscript.

Received: 14 July 2023 Accepted: 20 November 2023 Published online: 28 November 2023

References

1. Andersen, T.G., Bollerslev, T.: Answering the skeptics: yes, standard volatility models do provide accurate forecasts. Int. Econ. Rev. **39**(4), 885–905 (1998)
2. Arnold, L.: Stochastic Differential Equations: Theory and Applications. Wiley, New York (1974)
3. Bahar, A., Mao, X.: Stochastic delay Lotka–Volterra model. J. Math. Anal. Appl. **292**, 364–380 (2004)
4. Bandi, F.M., Phillips, P.C.B.: Fully nonparametric estimation of scalar diffusion models. Econometrica **71**(1), 241–283 (2003)
5. Bandi, F.M., Russell, J.R.: Microstructure noise, realized variance, and optimal sampling. Rev. Econ. Stud. **75**(2), 339–369 (2008)
6. Barndorff-Nielsen, O.E., Shephard, N.: Econometric analysis of realized volatility and its use in estimating stochastic volatility models. J. R. Stat. Soc. **64**(2), 253–280 (2002)
7. Barndorff-Nielsen, O.E., Shephard, N.: Estimating quadratic variation using realized variance. J. Appl. Econom. **17**, 457–477 (2002)
8. Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1968)
9. Boussama, F., Fuchs, F., Stelzer, R.: Stationary and geometric ergodicity of BEKK multivariate GARCH models. Stoch. Process. Appl. **121**, 2331–2360 (2011)
10. Carrasco, M., Chen, X.: Mixing and moment properties of various GARCH and stochastic volatility models. Econom. Theory **18**, 17–39 (2002)
11. Chang, J., Hu, Q., Liu, C., Tang, C.: Optimal covariance matrix estimation for high-dimensional noise in high-frequency data. J. Econom. (2022). https://doi.org/10.1016/j.jeconom.2022.06.010
12. Chen, X., Hansen, L.P., Carrasco, M.: Nonlinearity and temporal dependence. J. Econom. **155**(2), 155–169 (2010)
13. Christensen, K., Podolskij, M.: Asymptotic theory for range-based estimation of integrated variance of a continuous semi-martingale. Technical Reports (2005)
14. Comte, F., Genon-Catalot, V., Rozenholc, Y.: Nonparametric adaptive estimation for integrated diffusions. Stoch. Process. Appl. **119**, 811–834 (2009)
15. Ditlevsen, S., Sørensen, M.: Inference for observations of integrated diffusion processes. Scand. J. Stat. **31**(3), 417–429 (2004)
16. Fan, J., Li, Y., Yu, K.: Vast volatility matrix estimation using high-frequency data for portfolio selection. J. Am. Stat. Assoc. **107**(497), 412–428 (2010)
17. Fan, J., Wang, Y.: Multi-scale jump and volatility analysis for high-frequency financial data. J. Am. Stat. Assoc. **102**, 1349–1362 (2007)
18. Fan, J., Yao, Q.: Nonlinear Time Series: Nonparametric and Parametric Methods. Springer, Berlin (2003)
19. Gloter, A.: Discrete sampling of an integrated diffusion process and parameter estimation of the diffusion coefficient. ESAIM Probab. Stat. **4**, 204–227 (2000)
20. Gloter, A.: Parameter estimation for a discretely observed integrated diffusion process. Scand. J. Stat. **33**, 83–104 (2006)
21. Gloter, A., Gobet, E.: LAMN property for hidden processes: the case of integrated diffusions. Ann. Inst. Henri Poincaré Probab. Stat. **44**(1), 104–128 (2008)
22. Gorodetskii, V.V.: On the strong mixing property for linear sequences. Theory Probab. Appl. **22**, 411–413 (1977)
23. Hafner, C.M., Preminger, A.: On asymptotic theory for multivariate GARCH models. J. Multivar. Anal. **100**, 2044–2054 (2009)
24. Ibragimov, I.A.: Some limit theorems for stationary processes. Theory Probab. Appl. **7**(4), 349–382 (1962)
25. Ibragimov, I.A.: On the spectrum of stationary Gaussian sequences satisfying the strong mixing condition. I. Necessary conditions. Theory Probab. Appl. **10**(1), 85–106 (1965)
26. Ibragimov, I.A.: On the spectrum of stationary Gaussian sequences satisfying the strong mixing condition. II. Sufficient conditions mixing rate. Theory Probab. Appl. **15**(1), 23–36 (1970)
27. Kessler, M., Sørensen, M.: Estimating equations based on eigenfunctions for a discretely observed diffusion process. Bernoulli **5**(2), 299–314 (1999)
28. Klebaner, C.F.: Introduction to Stochastic Calculus with Applications. Imperial College Press, London (2012)
29. Kolmogorov, A.N., Rozanov, Y.A.: On strong mixing conditions for stationary Gaussian processes. Theory Probab. Appl. **5**, 204–208 (1960)
30. Li, C., Guo, E.: Estimation of the integrated volatility using noisy high-frequency data with jumps and endogeneity. Commun. Stat., Theory Methods **47**(3), 521–531 (2018)
31. Li, Y., Xie, S., Zheng, X.: Efficient estimation of integrated volatility incorporating trading information. J. Econom. **195**, 33–50 (2016)
32. Mao, X.: Stationary distribution of stochastic population systems. Syst. Control Lett. **60**, 398–405 (2011)
33. Nicolau, J.: Nonparametric estimation of second-order stochastic differential equations. Econom. Theory **23**, 880–898 (2007)

34. Peligrad, M.: Invariance principles for mixing sequencs of random variales. Ann. Probab. **10**, 968–981 (1982)
35. Peligrad, M.: Convergence rates of the strong law for stationary mixing sequences. Z. Wahrscheinlichkeitstheor. Verw. Geb. **70**, 307–314 (1985)
36. Peligrad, M.: On the central limit theorem for ρ-mixing sequences of random variables. Ann. Probab. **15**, 1387–1394 (1987)
37. Roussas, G.G., Ioannides, D.: Moment inequalities for mixing sequences of random variables. Stoch. Anal. Appl. **5**(1), 60–120 (1987)
38. Shao, Q.M.: A moment inequality and its applications. Acta Math. Sin. **31**, 736–747 (1988). (in Chinese)
39. Shao, Q.M.: Maximal inequalities for partial sums of ρ-mixing sequences. Ann. Probab. **23**, 948–965 (1995)
40. Shao, Q.M., Yu, H.: Weak convergence for weighed emprirical processes of dependent sequences. Ann. Probab. **24**, 2098–2127 (1996)
41. Wang, H., Lin, Z.: Local liner estimation of second-order diffusion models. Commun. Stat., Theory Methods **40**, 394–407 (2011)
42. Wang, Y., Zhang, L., Tang, M.: Re-weighted functional estimation of second-order diffusion processes. Metrika **75**(8), 1129–1151 (2012)
43. Wei, X., Yang, S., Yu, K., Yang, X., Xing, G.: Bahadur representation of linear kernel quantile estimator of VaR under α-mixing assumptions. J. Stat. Plan. Inference **140**, 1620–1634 (2010)
44. Withers, C.S.: Conditions for linear processes to be strong-mixing. Z. Wahrscheinlichkeitstheor. Verw. Geb. **57**, 477–480 (1981)
45. Wong, K.C., Li, Z., Tewari, A.: Lasso guarantees for β-mixing heavy-tailed time series. Ann. Stat. **48**, 1124–1142 (2020)
46. Xing, G., Kang, Q., Yang, S., Chen, Z.: Maximal moment inequality for partial sums of ρ-mixing sequences and its applications. J. Math. Inequal. **15**(2), 827–844 (2021)
47. Yang, S.: Moment inequality for mixing sequences and nonparametric estimation. Acta Math. Sin. **40**, 271–279 (1997). (in Chinese)
48. Yang, S.: Moment bounds for strong mixing sequences and their application. J. Math. Res. Exposition **20**(3), 349–359 (2000)
49. Yang, S.: Maximal moment inequality for partial sums of strong mixing sequences and application. Acta Math. Sin. Engl. Ser. **23**(6), 1013–1024 (2007)
50. Yang, S., Zhang, S., Xing, G., Yang, X.: Strong consistency of nonparametric kernel estimators for integrated diffusion process. Commun. Stat., Theory Methods (2022). https://doi.org/10.1080/03610926.2022.2148540
51. Yokoyama, R.: Moment bounds for stationary mixing sequences. Z. Wahrscheinlichkeitstheor. Verw. Geb. **52**, 45–57 (1980)
52. Zhang, L.: Rosenthal type inequalities for B-valued shrong mixing random fields and their applications. Sci. China Ser. A **41**(7), 736–745 (1998)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.