Bernoulli **20**(2), 2014, 623–644
DOI: 10.3150/12-BEJ500

Statistical convergence of Markov experiments to diffusion limits

VALENTIN KONAKOV[1], ENNO MAMMEN[2] and JEANNETTE WOERNER[3]

[1]*Higher School of Economics, Pokrovskii Boulevard 11, 103012 Moscow, Russia. E-mail: VKonakov@hse.ru*
[2]*Department of Economics, University of Mannheim, L7,3-5, 68229 Mannheim, Germany. E-mail: emammen@rumms.uni-mannheim.de*
[3]*Technische Universität Dortmund, Fakultät für Mathematik, Vogelpothsweg 87, 44227 Dortmund, Germany. E-mail: jeannette.woerner@math.uni-dortmund.de*

Assume that one observes the kth, $2k$th, $\dots, nk$th value of a Markov chain $X_{1,h}, \dots, X_{nk,h}$. That means we assume that a high frequency Markov chain runs in the background on a very fine time grid but that it is only observed on a coarser grid. This asymptotics reflects a set up occurring in the high frequency statistical analysis for financial data where diffusion approximations are used only for coarser time scales. In this paper, we show that under appropriate conditions the L_1-distance between the joint distribution of the Markov chain and the distribution of the discretized diffusion limit converges to zero. The result implies that the LeCam deficiency distance between the statistical Markov experiment and its diffusion limit converges to zero. This result can be applied to Euler approximations for the joint distribution of diffusions observed at points $\Delta, 2\Delta, \dots, n\Delta$. The joint distribution can be approximated by generating Euler approximations at the points $\Delta k^{-1}, 2\Delta k^{-1}, \dots, n\Delta$. Our result implies that under our regularity conditions the Euler approximation is consistent for $n \to \infty$ if $nk^{-2} \to 0$.

Keywords: deficiency distance; diffusion processes; Euler approximations; high frequency time series; Markov chains

1. Introduction

In this paper, we consider approximations of the joint distribution of a partially observed Markov chain by the law of a discretely observed diffusion. More precisely we consider a Markov chain $X_{1,h}, \dots, X_{nk,h}$ with values at nk time points. This time points are equal to $h, 2h, \dots, nkh$ where h is a time interval that converges to zero. We assume that this process is only observed at each kth point, that is, at the time points $kh, 2kh, \dots, nkh$. That means we assume that a high frequency Markov chain runs in the background on a very fine time grid but that it is only observed on a coarser grid. This asymptotics reflects a set up occurring in the high frequency statistical analysis for financial data where diffusion approximations are used for coarser time scales. For the finest scale, discrete pattern in the price processes become transparent that could not be modeled by diffusions. The joint distribution of the observed values of the Markov chain is denoted by P_h. We assume that this joint distribution can be approximated by the distribution of $(Y_1^*, \dots, Y_n^*)$ where $Y_1^*, \dots, Y_n^*$ are the values of a diffusion Y on the equidistant grid $kh, 2kh, \dots, nkh$, that is, $Y(ikh) = Y_i^*$. The joint distribution of $(Y_1^*, \dots, Y_n^*)$ is denoted by Q_h.

1350-7265

In this paper, we show that

$$\|P_h - Q_h\|_1 \to 0$$

under some regularity conditions if

$$\frac{n}{k} \to 0.$$

This result can be applied to the asymptotic study of Markov experiments $(P_{h,\theta} : \theta \in \Theta)$ where Θ is a finite or infinite-dimensional parameter set. Suppose that for this family of Markov chains our assumptions apply uniformly for $\theta \in \Theta$. Then one gets that $\sup_{\theta\in\Theta} \|P_{h,\theta} - Q_{h,\theta}\|_1 \to 0$ where $Q_{h,\theta}$ is the distribution of the discretized limiting diffusion. This implies that the Markov experiment $(P_{h,\theta} : \theta \in \Theta)$ and the diffusion experiment $(Q_{h,\theta} : \theta \in \Theta)$ are asymptotically equivalent in the sense of Le Cam's statistical theory of asymptotic equivalence of experiments. Asymptotic equivalence of nonparametric experiments has been discussed in a series of papers starting with [2] and [14]. Work of statistical experiments that converge to diffusions include [3,6,7,10,13,16]. Recently, Reiss [15] provided asymptotic equivalence of a stochastic volatility model with microstructure noise to a Gaussian shift experiment and a regression model whereas Buchmann and Müller [4] considered the relation between GARCH and COGARCH in the framework of statistical equivalence. Our result justifies approximating diffusion models for high frequency financial processes that are observed on a coarser grid. We also outline that the Markov experiment and its diffusion approximation differ in first order if n/k does not converge to zero. Then skewness properties of the Markov chain do not vanish in first order. For a related paper see [8]. They consider estimation of the intensity of a discretely observed compound Poisson process with symmetric Bernoulli jumps. For this model, they discuss limit experiments under different assumptions on the limit of the difference between neighbored time points.

We only discuss Markov chains with continuous state space. The distribution of Markov chains with discrete state space cannot be approximated by the distribution of continuous diffusions. For asymptotic equivalence of the experiments $(P_{h,\theta} : \theta \in \Theta)$ and $(Q_{h,\theta} : \theta \in \Theta)$, one has to show that there exist Markov kernels K_n and L_n with $\sup_{\theta\in\Theta} \|K_n P_{h,\theta} - Q_{h,\theta}\|_1 \to 0$ and $\sup_{\theta\in\Theta} \|P_{h,\theta} - L_n Q_{h,\theta}\|_1 \to 0$. We expect that such results could be shown by using expansions for transition densities of Markov random walks. The approach of this paper is based on expansions developed in [12]. The latter paper only considers Markov chains with continuous state space. To treat Markov random walks, their approach has to be carried over to the case of discrete state spaces.

2. The main result

We consider a Markov chain $X_{l,h}$ in $\mathbb{R}$ that runs on very fine time grid and has the following form

$$X_{l+1,h} = X_{l,h} + m(X_{l,h})h + \sqrt{h}\xi_{l+1,h}, \qquad X_{0,h} = x_0 \in \mathbb{R}, \qquad l = 0, \ldots, nk - 1. \tag{1}$$

The innovation sequence $(\xi_{l,h})_{l=1,\ldots,nk}$ is assumed to satisfy the Markov assumption: the conditional distribution of $\xi_{l+1,h}$ given the past $X_{l,h} = x_l, \ldots, X_{0,h} = x_0$ depends only on the last value $X_{l,h} = x_l$ and has a conditional density $q(x_l, \cdot)$. The conditional variance corresponding to this density is denoted by $\sigma^2(x_l)$ and the conditional νth order moment by $\mu_\nu(x_l)$. The transition densities of $(X_{r,h})$ given $(X_{l,h})$ are denoted by $p_h(rh - lh, x_l, \cdot)$.

In the following, C denotes a finite strictly positive constant whose meaning may vary from line to line. We make the following assumptions.

(A1) It holds that $\int_{\mathbb{R}} yq(x, y)\,\mathrm{d}y = 0$ for $x \in \mathbb{R}$.

(A2) There exist positive constants $\sigma_\star$ and $\sigma^\star$ such that the variance $\sigma^2(x) = \int_{\mathbb{R}} y^2 q(x, y)\,\mathrm{d}y$ satisfies

$$\sigma_\star \le \sigma^2(x) \le \sigma^\star$$

for all $x \in \mathbb{R}$.

(A3) There exist a positive integer $S' > 1$ and a real nonnegative function $\psi(y)$, $y \in \mathbb{R}$ satisfying $\sup_{y\in\mathbb{R}} \psi(y) < \infty$ and $\int_{\mathbb{R}} |y|^S \psi(y)\,\mathrm{d}y < \infty$ with $S = 2S' + 4$ such that

$$\left|D_y^\nu q(x, y)\right| \le \psi(y), \qquad x, y \in \mathbb{R}, 0 \le \nu \le 4.$$

Moreover, for all $x, y \in R$, $j \ge 1$

$$\left|D_x^\nu q^{(j)}(x, y)\right| \le Cj^{-1/2}\psi\left(j^{-1/2}y\right), \qquad 0 \le \nu \le 3$$

for a constant $C < \infty$. Here, $q^{(j)}(x, y)$ denotes the usual j-fold convolution of q for fixed x as a function of y:

$$q^{(j)}(x, y) = \int q^{(j-1)}(x, u)q(x, y - u)\,\mathrm{d}u,$$

$q^{(1)}(x, y) = q(x, y)$.

Note that the last condition is very weak. It is motivated by (A2) and the classical local limit theorem.

(A4) The functions $m(x)$ and $\sigma(x)$ and their derivatives up to the order six are continuous and bounded. Furthermore, $D_x^6\sigma(x)$ is Hölder continuous of order $0 < \alpha < 1$.

(A5) There exists $\varkappa < \frac{1}{5}$ and a constant $C > 0$ such that

$$C^{-1}k^{-\varkappa} < hk < C.$$

The Markov chain $X_{l,h}$, see (1), is an approximation to the following stochastic differential equation in $\mathbb{R}$:

$$\mathrm{d}Y_s = m(Y_s)\,\mathrm{d}s + \sigma(Y_s)\,\mathrm{d}W_s, \qquad Y_0 = x_0 \in \mathbb{R}, \qquad s \in [0, T], \tag{2}$$

where $(W_s)_{s\ge 0}$ is the standard Wiener process. The conditional density of Y_t, given $Y_s = x$ is denoted by $p(t - s, x, \cdot)$. We also write $Y(s)$ for Y_s. The joint distribution of Y on the equidistant grid $kh, 2kh, \dots, nkh$ is denoted by Q_h.

Our main result is stated in the following theorem.

Theorem 1. *Assume* (A1)–(A5) *and* $nk^{-1} \to 0$. *Then it holds that* $\|P_h - Q_h\|_1 \to 0$.

Remark 1. Theorem 1 can be generalized to higher dimensions and to the nonhomogenous case. We only treat the univariate homogenous case for simplicity. In our proof, we make use of the representation (4) from [5] that is only available for the univariate case. For multivariate reducible diffusions, one can apply the Hermite expansion given in [1].

Remark 2. The assumptions of Theorem 1 allow to apply second order expansions for the transition densities of Markov chains that have been developed in [12]. In the proof of Theorem 1, we make only use of first order expansions. For this reason, the assumptions could be weakened. For example, we expect that one needs only four derivatives in (A4) instead of six. We do not pursue this here because we will need the second order expansions for getting the results in the following theorem.

Theorem 2. *Assume* (A1)–(A5), $nh^{1+\delta} \to 0$ *and* $nk^{-2} \to 0$, *where* $\delta > 0$ *is chosen such that the statement of Theorem 4 holds for this choice. Suppose that the third conditional moment* $\mu_3(x)$ *of innovations of the Markov chain fulfills* $\mu_3(x) \equiv 0$. *Then it holds that* $\|P_h - Q_h\|_1 \to 0$.

Remark 3. This result can be applied to Euler approximations of diffusions and to Markov chains with symmetric innovations. For Euler schemes that approximate the joint density of a diffusion at points $\Delta, 2\Delta, \ldots, n\Delta$ it means that one has to generate Euler approximations of the diffusions at points $\Delta k^{-1}, 2\Delta k^{-1}, \ldots, n\Delta$ where $k \to \infty$ is chosen such that $nk^{-2} \to 0$ and $n(\Delta/k)^{1+\delta} \to 0$. The joint distribution of the Euler values at the points $\Delta, 2\Delta, \ldots, n\Delta$ is then the approximation of the joint distribution of the diffusion at these points. Under the regularity assumptions of Theorem 2, the Euler approximation is consistent. A more detailed discussion of the necessity of the above assumptions on k will be given elsewhere.

We now show that our assumption on the growth of k in Theorem 1 is sharp. For this purpose, we consider a simple model of Markov chains that converge to a Gaussian process and we show that for this case $\|P_h - Q_h\|_1$ does not converge to zero if the condition on the growth of k in Theorem 1 is not met.

Theorem 3. *Assume* (A1)–(A5) *for Markov chains with* $m(x) \equiv 1$ *and innovation density* $q(x, \cdot) = q(\cdot)$ *not depending on* x. *We assume that* $nk^{-1} \to c$ *for a constant* $c \neq 0$. *Furthermore, suppose, that* $\mu_3(x) = \mu_3 \neq 0$ *and that* $kh \to 0$. *Then* $\|P_h - Q_h\|_1$ *does not converge to zero.*

3. Proofs

The proof of Theorem 1 will be divided into several lemmas. For the proof, we will make use of the results in [12] where Edgeworth type expansions of p_h were given for nonhomogenous Markov chains in $\mathbb{R}^d$ for $d \geq 1$. We now restate their main result for one-dimensional homogenous Markov chains. To formulate their result, we need some additional notation.

We will use the following differential operators L and $\widetilde{L}$:

$$
\begin{aligned}
Lf(t,x,y) &= \frac{1}{2}\sigma^2(x)\frac{\partial^2 f(t,x,y)}{(\partial x)^2} + m(x)\frac{\partial f(t,x,y)}{\partial x},\\
\widetilde{L}f(t,x,y) &= \frac{1}{2}\sigma^2(y)\frac{\partial^2 f(t,x,y)}{(\partial x)^2} + m(y)\frac{\partial f(t,x,y)}{\partial x}.
\end{aligned}
\tag{3}
$$

We also need the following convolution type binary operation $\otimes$:

$$f \otimes g(t,x,y) = \int_0^t \mathrm{d}u \int_R f(u,x,z)g(t-u,z,y)\,\mathrm{d}z.$$

We now introduce the following differential operators

$$
\begin{aligned}
\mathcal{F}_1[f](t,x,y) &= \frac{\mu_3(x)}{6} D_x^3 f(t,x,y),\\
\mathcal{F}_2[f](t,x,y) &= \frac{\mu_4(x) - 3\sigma^4(x)}{24} D_x^4 f(t,x,y).
\end{aligned}
$$

The Gaussian transition densities $\widetilde{p}(t,x,y)$ are defined as

$$\widetilde{p}(t,x,y) = (2\pi)^{-1/2}\sigma(y)^{-1}t^{-1/2}\exp\left(-\frac{1}{2t}\bigl(y-x-tm(y)\bigr)^2\sigma(y)^{-2}\right).$$

We are now in the position to state the Edgeworth type expansion for Markov chain transition densities from [12].

Theorem 4 ([12]). *Assume* (A1)–(A5). *Then there exists a constant* $\delta > 0$ *such that the following expansion holds*:

$$
\begin{aligned}
&\sup_{x,y\in R} (kh)^{1/2}\left(1 + \left|\frac{y-x}{\sqrt{kh}}\right|^{S'}\right)\\
&\quad\times \left|p_h(kh,x,y) - p(kh,x,y) - h^{1/2}\pi_1(kh,x,y) - h\pi_2(kh,x,y)\right| = \mathrm{O}\bigl(h^{1+\delta}\bigr),
\end{aligned}
$$

where S' *is defined in Assumption* (A3) *and where*

$$
\begin{aligned}
\pi_1(t-s,x,y) &= \bigl(p \otimes \mathcal{F}_1[p]\bigr)(t-s,x,y),\\
\pi_2(t-s,x,y) &= \bigl(p \otimes \mathcal{F}_2[p]\bigr)(t-s,x,y) + p \otimes \mathcal{F}_1\bigl[p \otimes \mathcal{F}_1[p]\bigr](t-s,x,y)\\
&\quad + \tfrac{1}{2} p \otimes \bigl(L_\star^2 - L^2\bigr)p(t-s,x,y).
\end{aligned}
$$

Here the operator $L_\star$ *is defined as* $\widetilde{L}$, *but with the coefficients "frozen" at the point* x, *that is*,

$$L_\star^2 f(t,x,y) = \frac{1}{4}\sigma^4(x)\frac{\partial^4 f(t,x,y)}{(\partial x)^4} + \sigma^2(x)m(x)\frac{\partial^3 f(t,x,y)}{(\partial x)^3} + m(x)^2\frac{\partial^2 f(t,x,y)}{(\partial x)^2}.$$

We will apply this theorem for transition densities over the interval $(ikh, (i+1)kh]$. The expansion of the theorem holds uniformly over $0 \le i \le n-1$.

We denote now the signed measure on $\mathbb{R}^n$ defined by the products of $p + h^{1/2}\pi_1$ as Q_h^1 and the signed measure defined by the products of $p + h^{1/2}\pi_1 + h\pi_2$ as Q_h^2.

Proof of Theorem 1. Theorem 1 immediately follows from the following two lemmas. □

In all lemmas of this section, we make the assumptions of Theorem 1.

Lemma 1. *It holds that*:

$$\| Q_h^1 - Q_h \|_1 = \mathrm{o}(1) \qquad \textit{for } n \to \infty.$$

Lemma 2. *It holds that*:

$$\| P_h - Q_h^1 \|_1 = \mathrm{o}(1) \qquad \textit{for } n \to \infty.$$

The hard part of these two lemmas is the proof of Lemma 1. For the proof of the two lemmas, we will use a series of lemmas that are stated and proved now. We will come back to the proofs of Lemmas 1 and 2 afterwards.

In our proofs, we make use of the following representation of transition densities. For the transition density $p(t-s, x, \xi)$ of the diffusion (2), the following formula holds, see formula (3.2) in [5]

$$\begin{aligned} p(t-s,x,y) &= \widehat{p}(t-s,x,y) \\ &\quad \times E\exp\biggl[(t-s)\int_0^1 g\bigl[z_\delta\bigl(S(x),S(y)\bigr)+\sqrt{(t-s)}B_\delta\bigr]\,\mathrm{d}\delta\biggr], \end{aligned} \tag{4}$$

where for $0 \le \delta \le 1$ B_δ is a Brownian bridge. Furthermore, for $u \ge 0$ we put $g(u) = -\frac{1}{2}(C^2(u) + C'(u))$ and $z_\delta(x,y) = (1-\delta)x + \delta y$ with

$$\widehat{p}(t-s,x,y) = \frac{1}{\sqrt{2\pi(t-s)}\sigma(y)}\exp\biggl[-\frac{(S(y)-S(x))^2}{2(t-s)} + H(y) - H(x)\biggr], \tag{5}$$

$$\begin{aligned} S(x) &= \int_0^x \frac{\mathrm{d}u}{\sigma(u)}, \\ H(x) &= \int_0^{S(x)} C(u)\,\mathrm{d}u \qquad \text{with } C(u) = \frac{m(u)}{\sigma(u)} - \frac{1}{2}\sigma'(u) \end{aligned} \tag{6}$$

for $x, y, s, t \in \mathbb{R}$.

Note that under our assumptions g is bounded, $|g(x)| \le M$, and, hence, for $t - s \le kh$

$$E\exp\biggl[(t-s)\int_0^1 g\bigl[z_\delta\bigl(S(x),S(\xi)\bigr)+\sqrt{(t-s)}B_\delta\bigr]\,\mathrm{d}\delta\biggr] \le \exp[Mkh] \le C^* \tag{7}$$

for some constant $C^* > 0$ because of (A5). For the proof of Lemma 1 we make use of the following lemmas. These lemmas make use of some further technical lemmas, given in Section 4 that bound $\delta_1(x,y) = \sqrt{h}\pi_1(kh,x,y)/p(kh,x,y)$, $\delta_2(x,y) = h\pi_2(kh,x,y)/p(kh,x,y)$ and partial derivatives of the transition densities.

Lemma 3. *Put $\Delta_i = \delta_1(Y((i-1)kh), Y(ikh))$. Then we have for all $p \geq 1$ that under Q_h*

$$\sup_{1 \leq i \leq n} E_{Q_h} |\Delta_i|^p \leq C_p k^{-p/2}$$

for some constants C_p depending on p.

Proof of Lemma 3. This lemma directly follows from Lemma 9 and the representation (4). Using these results, the moments of Δ_i can be easily bounded by Gaussian moments. □

Lemma 3 implies that for all $\rho > 0$ under Q_h

$$\sup_{1 \leq i \leq n} |\Delta_i| = \mathrm{O}_p\bigl(k^{-1/2} n^{\rho}\bigr).$$

This bound would suffice for our purposes but for completeness we state the following sharper bound that follows (from our Lemma 9 and) from Theorem 1 in [9], where bounds for moments for the modulus of continuity of diffusions are given.

Lemma 4. *We have that under Q_h that*

$$\sup_{1 \leq i \leq n} |\Delta_i| = \mathrm{O}_p\bigl(k^{-1/2} (\log n)^{3/2}\bigr).$$

We now state a result on the order of sums of Δ_i's.

Lemma 5. *Under Q_h it holds that*

$$\sum_{i=1}^{n} \Delta_i = \mathrm{O}_p\left(\sqrt{\frac{n}{k}}\right).$$

Proof. We have that $E_{Q_h}[\Delta_i \Delta_j] = 0$ for $i \neq j$ because the definition of Δ_i implies that $E_{Q_h}[\Delta_i | Y_j : j \leq i-1] = 0$. Thus, it holds

$$E_{Q_h}\left[\left(\sum_{i=1}^{n} \Delta_i\right)^2\right] = \sum_{i=1}^{n} E_{Q_h}\bigl[(\Delta_i)^2\bigr] = \mathrm{O}\bigl(nk^{-1}\bigr).$$

This follows from Lemma 3. □

Put $A_n = \{(Y(kh), \dots, Y(nkh)) : \sup_{1 \leq i \leq n} |\Delta_i| \leq \tau_n \sqrt{\frac{n}{k}}, |\sum_{i=1}^{n} \Delta_i| \leq \tau_n \sqrt{\frac{n}{k}}\}$, where $\tau_n \to \infty$ with $\tau_n \sqrt{\frac{n}{k}} \to 0$.

Then we get from Lemmas 4–5 that

$$Q_h(A_n) \to 1. \tag{8}$$

For the proof of Lemma 1, we need the following additional simple lemma.

Lemma 6. *Consider the set* $B_n = \{x \in \mathbb{R}^n : |x_i| \le \tau_n\sqrt{\frac{n}{k}}$ *for* $i = 1, \ldots, n; |\sum_{i=1}^n x_i| \le \tau_n\sqrt{\frac{n}{k}}\} \subset \mathbb{R}^n$, *where* $\tau_n \to \infty$ *with* $\tau_n\sqrt{\frac{n}{k}} \to 0$. *Then it holds that*

$$\sup_{x\in B_n}\left|1 - \prod_{i=1}^n (1 + x_i)\right| \to 0 \qquad \text{for } n \to \infty.$$

The lemma implies that

$$\max_{1\le j\le n}\sup_{x\in B_{nj}}\left|1 - \prod_{i=1}^j (1 + x_i)\right| \to 0 \qquad \text{for } n \to \infty,$$

where $B_{nj} = \{x \in \mathbb{R}^j : |x_i| \le \tau_n\sqrt{\frac{n}{k}}$ for $i = 1, \ldots, j; |\sum_{i=1}^j x_i| \le \tau_n\sqrt{\frac{n}{k}}\} \subset \mathbb{R}^n$. This follows by putting $x_i = 0$ for $i = j + 1, \ldots, n$.

The next lemma states that the expansion (8) also holds under the measure Q_h^1.

Lemma 7. *It holds that*

$$|Q_h^1|(A_n) \to 1.$$

Here, $|Q_h^1|$ *means the total variation measure of* Q_h^1.

Proof. By application of Lemma 6, we get that

$$\begin{aligned}\left||Q_h^1|(A_n) - Q_h(A_n)\right| &\le \int I(A_n)\left|\mathrm{d}Q_h^1 - \mathrm{d}Q_h\right| \\ &= \int \left|1 - \prod_{i=1}^n (1 + \Delta_i)\right| I(A_n)\,\mathrm{d}Q_h \\ &= \mathrm{o}(1).\end{aligned}$$

This implies the statement of the lemma because of (8). □

We now prove Lemma 1.

Proof of Lemma 1. We have that

$$\begin{aligned}\|Q_h^1 - Q_h\|_1 &= E_{Q_h}\left[\left|1 - \prod_{i=1}^{n}(1+\Delta_i)\right|\right] \\ &= E_{Q_h}\left[\left|1 - \prod_{i=1}^{n}(1+\Delta_i)\right| I(A_n)\right] + \mathrm{o}(1),\end{aligned}$$

because of (8) and Lemma 7. Now the lemma follows from Lemma 6. □

It remains to prove Lemma 2.

Proof of Lemma 2. We can write $P_h = P_{h,1} \times \cdots \times P_{h,n}$ and $Q_h^1 = Q_{h,1}^1 \times \cdots \times Q_{h,n}^1$ where $P_{h,j}$, $Q_{h,j}^1$ are suitably defined (signed) Markov kernels. By using a telescope argument, we get with constants $C^*, C^{**} > 0$ that for n large enough

$$\begin{aligned}\|P_h - Q_h^1\|_1 &= \int \big| p_h(kh, x, z_1) \times \cdots \times p_h(kh, z_{n-1}, z_n) \\ &\qquad - \big(p + h^{1/2}\pi_1\big)(kh, x, z_1) \times \cdots \times \big(p + h^{1/2}\pi_1\big)(kh, z_{n-1}, z_n)\big| \, \mathrm{d}z_1 \cdots \mathrm{d}z_n \\ &\le \int \big|\big(p_h - p - h^{1/2}\pi_1\big)(kh, x, z_1)\big| \\ &\qquad \times p_h(kh, z_1, z_2) \times \cdots \times p_h(kh, z_{n-1}, z_n) \, \mathrm{d}z_1 \cdots \mathrm{d}z_n \\ &\quad + \int \big|p + h^{1/2}\pi_1\big|(kh, x, z_1) \big|\big(p_h - p - h^{1/2}\pi_1\big)(kh, z_1, z_2)\big| \\ &\qquad \times p_h(kh, z_2, z_3) \times \cdots \times p_h(kh, z_{n-1}, z_n) \, \mathrm{d}z_1 \cdots \mathrm{d}z_n \\ &\quad + \cdots + \int \big|p + h^{1/2}\pi_1\big|(kh, x, z_1) \times \cdots \times \big|p + h^{1/2}\pi_1\big|(kh, z_{n-2}, z_{n-1}) \\ &\qquad \times \big|\big(p_h - p - h^{1/2}\pi_1\big)(kh, z_{n-1}, z_n)\big| \, \mathrm{d}z_1 \cdots \mathrm{d}z_n \\ &\le \frac{C^*}{k}\left(1 + \sum_{j=1}^{n-1} E_{Q_h}\left[\prod_{i=1}^{j}\big(|1+\Delta_i|\big)\right]\right) \\ &\le \frac{C^*}{k}\left(1 + \sum_{j=1}^{n} \|Q_{h,j}^1\|_1\right) \\ &\le n\frac{C^*}{k}\big(1 + \mathrm{o}(1)\big) \le \frac{C^{**}n}{k} = \mathrm{o}(1), \qquad n \to \infty,\end{aligned}$$

where $\|Q^1_{h,j}\|_1 = \int |p + h^{1/2}\pi_1|(kh, x, z_1) \times \cdots \times |p + h^{1/2}\pi_1|(kh, x, z_j)\,\mathrm{d}z_1 \cdot \cdots \cdot dz_j$. We used that

$$\sup_x \int \bigl|\bigl(p_h - p - h^{1/2}\pi_1\bigr)(kh, x, z)\bigr|\,\mathrm{d}z \le \frac{C^*}{k} \tag{9}$$

for some constant $C^* > 0$ and that uniformly in $1 \le j \le n$, $\|Q^1_{h,j}\|_1 \le \|Q_{h,j}\|_1 + \|Q^1_{h,j} - Q_{h,j}\|_1 = 1 + \|Q^1_{h,j} - Q_{h,j}\|_1 = 1 + \mathrm{o}(1)$. For the last equality, we used Lemma 1.

From Theorem 4, we get that the left-hand side of the inequality can be bounded by:

$$\sup_x \int \bigl|h\pi_2(kh, x, z)\bigr|\,\mathrm{d}z + \mathrm{O}\bigl(h^{1+\delta}\bigr) \sup_x \int (kh)^{-1/2}\left(1 + \left|\frac{z-x}{\sqrt{kh}}\right|^{S'}\right)^{-1} \mathrm{d}z. \tag{10}$$

According to Assumption (A5) hk is bounded. Thus, the second term in (10) is of order $\mathrm{O}(h^{1+\delta}) = \mathrm{O}((hk)^{1+\delta}k^{-1-\delta}) = \mathrm{O}(k^{-1-\delta}) = \mathrm{O}(k^{-1})$.

For the first term, we have the following bound from Lemma 11:

$$\mathrm{O}\bigl(k^{-1}\bigr) \sup_x \int p(kh, x, z)\left[1 + \left(\frac{|z-x|}{\sqrt{kh}}\right)^7\right] \mathrm{d}z.$$

Now, the second factor of this bound is of order $\mathrm{O}(1)$ because of (4). Thus, the bound is of order $\mathrm{O}(k^{-1})$. This shows claim (9) and concludes the proof of the lemma. □

Proof of Theorem 2. It is enough to prove that

$$\bigl\|Q_h - Q_h^2\bigr\|_1 \to 0, \qquad n \to \infty \tag{11}$$

and

$$\bigl\|P_h - Q_h^2\bigr\|_1 \le Cnh^{1+\delta}. \tag{12}$$

Claim (12) can be shown with arguments similar to the ones used in the proof of Lemma 2. Instead of the bound 10, one now uses the expansion of Theorem 4.

The proof of (11) is close to the proof of Lemma 1. With $\Delta_i^{(2)} = \delta_2(Y((i-1)kh), Y(ikh))$, we obtain as it was done before

$$\sup_{1\le i\le n} E\bigl|\Delta_i^{(2)}\bigr|^p \le C_p k^{-p}, \tag{13}$$

$$\sup_{1\le i\le n} \bigl|\Delta_i^{(2)}\bigr| = \mathrm{O}_p\bigl(k^{-1}(\log n)^{7/2}\bigr), \tag{14}$$

$$\sum_{i=1}^n \Delta_i^{(2)} = \mathrm{O}_p\left(\frac{\sqrt{n}}{k}\right) \tag{15}$$

and the assertion of Theorem 2 follows with the same arguments as used in the proof of Theorem 1. □

Proof of Theorem 3. Without loss of generality, we assume that $\int x^2 q(x)\,dx = 1$. Suppose that $\|P_h - Q_h\|_1$ does converge to zero. This implies that the loglikelihood $\log(dP_h/dQ_h)$ converges to zero in Q_h-probability. Thus, we have that

$$\sum_{i=1}^{n} \log\bigl(1 + \Delta_i + \Delta_i^{(2)}\bigr) \xrightarrow{Q_h} 0.$$

Note that the bounds (13)–(15) remain valid under the assumptions of Theorem 3. We now apply Lemma 4 and (14). With a Taylor expansion of the logarithm, we get from the last expression that

$$\sum_{i=1}^{n} \Delta_i - \frac{1}{2}\Delta_i^2 + \Delta_i^{(2)} \xrightarrow{Q_h} 0.$$

Because of (15) this shows that

$$\sum_{i=1}^{n} \Delta_i - \frac{1}{2}\Delta_i^2 \xrightarrow{Q_h} 0. \tag{16}$$

We will show that under Q_h

$$\sum_{i=1}^{n} \Delta_i \xrightarrow{d} N\bigl(0, \sigma^2\bigr) \tag{17}$$

with $\sigma^2 = 22c\mu_3^2 > 0$ where c is the limit of n/k. Note that (17) contradicts (16) because these two limit statements would imply that

$$\frac{1}{2}\sum_{i=1}^{n} \Delta_i^2 \xrightarrow{d} N\bigl(0, \sigma^2\bigr).$$

This is not possible because non negative random variables cannot converge in distribution to a normal limit with strictly positive variance. Thus for the statement of the theorem, it remains to prove (17).

For the proof of (17), we will use a martingale central limit theorem for the martingale $\sum_{j=1}^{i} \Delta_j$ with σ-field $\mathcal{F}_{h,i} = \sigma(Y(0), Y(kh), \ldots, Y(ikh))$. According to Theorem 3.2 and Corollary 3.1 in [11], we have for (17) to check that

$$\sum_{i=1}^{n} E\bigl[\Delta_i^2 | \mathcal{F}_{h,i-1}\bigr] \to \sigma^2, \qquad \text{in probablity}, \tag{18}$$

$$\max_{1\le i\le n} \Delta_i^2 \to 0, \qquad \text{in probablity}, \tag{19}$$

$$E\Bigl[\max_{1\le i\le n} \Delta_i^2\Bigr] = \mathrm{O}(1). \tag{20}$$

Claims (19)–(20) follow directly from Lemmas 3–4. Here, for the proof of (20) one can use the simple bound $\max_{1\le i\le n}\Delta_i^2 \le \sum_{i=1}^n \Delta_i^2$. Thus for the statement of the theorem it only remains to prove (18).

For the limiting diffusion Y_s we get that $\mathrm{d}Y_s = \mathrm{d}s + \mathrm{d}W_s$. For this case, it holds that $p(s,t,x,y) = \widehat{p}(s,t,x,y) = (2\pi(t-s))^{-1/2}\exp(-(y-x-(t-s))^2(2(t-s))^{-1})$. We now give an estimate for

$$6h^{-1/2}\delta_1(x,y)p(kh,x,y) = \mu_3\int_0^{kh}\mathrm{d}u\int p(u,x,\xi)\frac{\partial^3}{\partial\xi^3}p(kh-u,\xi,y)\,\mathrm{d}\xi. \tag{21}$$

Calculations close to the proof of (33) give the following estimate with a constant $C>0$:

$$\left|\frac{\partial^3 p(t,x,y)}{\partial x^3} + \frac{\partial^3 p(t,x,y)}{\partial y^3}\right| \le \frac{Cp(t,x,y)}{\sqrt{t}}\left(1+\left|\frac{y-x}{\sqrt{t}}\right|^3\right).$$

Using this estimate in (21), we obtain

$$6h^{-1/2}\delta_1(x,y)p(kh,x,y) = I(x,y) + \mathit{II}(x,y),$$

where

$$\begin{aligned}
\mathit{II}(x,y) &= -\mu_3\int_0^{kh}\mathrm{d}u\int p(u,x,\xi)\frac{\partial^3}{\partial y^3}p(kh-u,\xi,y)\,\mathrm{d}\xi \\
&= -\mu_3\frac{\partial^3}{\partial y^3}\int_0^{kh}\mathrm{d}u\int p(u,x,\xi)p(kh-u,\xi,y)\,\mathrm{d}\xi \\
&= -\mu_3 kh\frac{\partial^3}{\partial y^3}p(kh,x,y) \qquad\qquad (22)\\
&= -\mu_3\frac{p(kh,x,y)}{\sqrt{kh}}\left(\left(\sqrt{kh}-\frac{y-x}{\sqrt{kh}}\right)^3 - 2\left(\sqrt{kh}-\frac{y-x}{\sqrt{kh}}\right)^2 - \left(\sqrt{kh}-\frac{y-x}{\sqrt{kh}}\right)\right) \\
&= -\mu_3\frac{p(kh,x,y)}{\sqrt{kh}}Q_3\left(\sqrt{kh}-\frac{y-x}{\sqrt{kh}}\right)
\end{aligned}$$

with $Q_3(z) = z^3 - 2z^2 - z$. For the term $I(x,y)$, we have the following bound with a (new) constant $C>0$:

$$\left|I(x,y)\right| \le C\int_0^{kh}\mathrm{d}u\int \frac{p(u,x,\xi)p(kh-u,\xi,y)}{\sqrt{kh-u}}\left(1+\left|\frac{y-x}{\sqrt{kh-u}}\right|^3\right)\mathrm{d}\xi.$$

Using the same substitution as in the proof of Lemma 9, we obtain the following estimate

$$\left|I(x,y)\right| \le C\sqrt{kh}\,p(kh,x,y)P_3\left(\left|\frac{y-x}{\sqrt{kh}}\right|\right) \tag{23}$$

for a polynomial $P_3(z)$ of degree 3 with positive coefficients. Put now

$$\Delta_{1,i} = \frac{h^{1/2} I(Y((i-1)kh), Y(ikh))}{6p(kh, Y((i-1)kh), Y(ikh))},$$

$$\Delta_{2,i} = \frac{h^{1/2} II(Y((i-1)kh), Y(ikh))}{6p(kh, Y((i-1)kh), Y(ikh))}.$$

We will show that

$$\sum_{i=1}^{n} E\big[\Delta_{1,i}^2 | \mathcal{F}_{h,i-1}\big] \to 0, \qquad \text{in probablity,} \tag{24}$$

$$\sum_{i=1}^{n} E\big[\Delta_{2,i}^2 | \mathcal{F}_{h,i-1}\big] \to \sigma^2, \qquad \text{in probablity.} \tag{25}$$

Because of $\Delta_i = \Delta_{1,i} + \Delta_{2,i}$ this shows (18).

We get from (23) with a new constant $C > 0$ that

$$E\big[\Delta_{1,i}^2 | \mathcal{F}_{h,i-1}\big] \leq Ckh^2 E\left[P_3\left(\left| \frac{Y(ikh) - Y((i-1)kh)}{\sqrt{kh}} \right| \right)^2 \middle| \mathcal{F}_{h,i-1} \right]. \tag{26}$$

Conditionally given $\mathcal{F}_{h,i-1}$, $(Y(ikh) - Y((i-1)kh))/\sqrt{kh}$ has a normal distribution with mean $\sqrt{kh}$ and variance 1. Because of $kh \to 0$ (by assumption), we get that the expectation on the right hand side of (26) is uniformly bounded, for $1 \leq i \leq n, n \geq 1$. Furthermore, we have that $nkh^2 = (n/k)(kh)^2 \to 0$. Thus, (26) implies (24).

It remains to check (25). For the proof of this claim, we apply the explicit expression (22) and we get that

$$\begin{aligned}
\sum_{i=1}^{n} E\big[\Delta_{2,i}^2 | \mathcal{F}_{h,i-1}\big] &= \frac{1}{k}\mu_3^2 \sum_{i=1}^{n} E\left[Q_3\left(\sqrt{kh} - \frac{Y(ikh) - Y((i-1)kh)}{\sqrt{kh}} \right)^2 \middle| \mathcal{F}_{h,i-1} \right] \\
&= \frac{n}{k}\mu_3^2 \frac{1}{\sqrt{2\pi}} \int (z^6 + 2z^4 + z^2) \mathrm{e}^{-z^2/2}\, \mathrm{d}z \\
&= 22\frac{n\mu_3^2}{k}.
\end{aligned}$$

Now, because of $n/k \to c$, we get that the right-hand side of this equation converges to σ^2. This concludes the proof. □

4. Some technical lemmas

This section collects some technical lemmas that were used in the proofs of the last section. In all lemmas of this section, we make the assumptions of Theorem 1.

Lemma 8. *For all $c > 0$ there exists a constant $C > 0$ such that the following estimates hold for $0 \le t - s \le c$*

$$\left|\frac{\partial}{\partial x} p(t-s, x, y)\right| \le C \frac{p(t-s, x, y)}{\sqrt{t-s}} \left(\sqrt{t-s} + \frac{|y-x|}{\sqrt{t-s}}\right), \tag{27}$$

$$\left|\frac{\partial}{\partial y} p(t-s, x, y)\right| \le C \frac{p(t-s, x, y)}{\sqrt{t-s}} \left(\sqrt{t-s} + \frac{|y-x|}{\sqrt{t-s}}\right), \tag{28}$$

$$\left|\frac{\partial^2}{\partial y^2} p(t-s, x, y)\right| \le C \frac{p(t-s, x, y)}{t-s} \left(1 + \sqrt{t-s} + \frac{|y-x|}{\sqrt{t-s}}\right)^2, \tag{29}$$

$$\left|\frac{\partial^2}{\partial x^2} p(t-s, x, y)\right| \le C \frac{p(t-s, x, y)}{t-s} \left(1 + \sqrt{t-s} + \frac{|y-x|}{\sqrt{t-s}}\right)^2, \tag{30}$$

$$\left|\frac{\partial^3}{\partial x^3} p(t-s, x, y)\right| \le C \frac{p(t-s, x, y)}{(t-s)^{3/2}} \left(1 + \sqrt{t-s} + \frac{|y-x|}{\sqrt{t-s}}\right)^3, \tag{31}$$

$$\left|\frac{\partial^4}{\partial x^2 \, \partial y^2} p(t-s, x, y)\right| \le C \frac{p(t-s, x, y)}{(t-s)^2} \left(1 + \sqrt{t-s} + \frac{|y-x|}{\sqrt{t-s}}\right)^4, \tag{32}$$

$$\begin{aligned} \left|\left(\frac{\partial^2}{\partial x \, \partial y} - \frac{\partial^2}{\partial x^2}\right) p(t-s, x, y)\right| &\le C \frac{p(t-s, x, y)}{\sqrt{t-s}} \\ &\quad \times \left(1 + \frac{|y-x|}{\sqrt{t-s}} + \left|\frac{y-x}{\sqrt{t-s}}\right|^2 + \left|\frac{y-x}{\sqrt{t-s}}\right|^3\right). \end{aligned} \tag{33}$$

Proof. We prove the second, the third and the last inequality. The remaining inequalities can be proved exactly in the same way. From (5), we obtain

$$\begin{aligned} \frac{\partial}{\partial y} \widehat{p}(t-s, x, y) &= -\frac{\sigma'(y)}{\sqrt{2\pi(t-s)}\sigma^2(y)} \exp\left[-\frac{(S(y)-S(x))^2}{2(t-s)} + H(y) - H(x)\right] \\ &\quad + \frac{1}{\sqrt{2\pi(t-s)}\sigma(y)} \exp\left[-\frac{(S(y)-S(x))^2}{2(t-s)} + H(y) - H(x)\right] \\ &\quad \times \left(H'(y) - \frac{(S(y)-S(x))}{(t-s)\sigma(y)}\right) \\ &= \widehat{p}(t-s, x, y)\left[-\frac{\sigma'(y)}{\sigma(y)} + H'(y) - \frac{S(y)-S(x)}{(t-s)\sigma(y)}\right], \end{aligned} \tag{34}$$

$$\begin{aligned} \frac{\partial^2}{\partial y^2} \widehat{p}(t-s, x, y) &= \frac{\partial}{\partial y} \widehat{p}(t-s, x, y)\left[-\frac{\sigma'(y)}{\sigma(y)} + H'(y) - \frac{S(y)-S(x)}{(t-s)\sigma(y)}\right] \\ &\quad + \widehat{p}(t-s, x, y)\left[\frac{(\sigma'(y))^2 - \sigma(y)\sigma''(y)}{\sigma^2(y)} + H''(y)\right. \end{aligned}$$

$$
\begin{aligned}
&\qquad\qquad\qquad - \frac{1-\sigma'(y)(S(y)-S(x))}{(t-s)\sigma^2(y)}\Bigg] \\
&= \widehat{p}(t-s,x,y)\bigg[-\frac{\sigma'(y)}{\sigma(y)} + H'(y) - \frac{S(y)-S(x)}{(t-s)\sigma(y)}\bigg]^2 \\
&\quad + \widehat{p}(t-s,x,y)\bigg[\frac{(\sigma'(y))^2-\sigma(y)\sigma''(y)}{\sigma^2(y)} + H''(y) \\
&\qquad\qquad - \frac{1}{(t-s)\sigma^2(y)} + \frac{(S(y)-S(x))}{(t-s)}\frac{\sigma'(y)}{\sigma^2(y)}\bigg].
\end{aligned} \tag{35}
$$

It follows from (34) and (35) and our assumptions that

$$
\left|\frac{\partial}{\partial y}\widehat{p}(t-s,x,y)\right| \le C\frac{\widehat{p}(t-s,x,y)}{\sqrt{t-s}}\left(\sqrt{t-s} + \frac{|S(y)-S(x)|}{\sqrt{t-s}}\right), \tag{36}
$$

$$
\left|\frac{\partial^2}{\partial y^2}\widehat{p}(t-s,x,y)\right| \le C\frac{\widehat{p}(t-s,x,y)}{t-s}\left(1+\sqrt{t-s} + \frac{|S(y)-S(x)|}{\sqrt{t-s}}\right)^2. \tag{37}
$$

It is easy to see that

$$
\begin{aligned}
&\left|\frac{\partial}{\partial y}E\exp\bigg[(t-s)\int_0^1 g\big[z_\delta\big(S(x),S(y)\big)+\sqrt{(t-s)}B_\delta\big]\,\mathrm{d}\delta\bigg]\right| \\
&\quad \le C(t-s)E\exp\bigg[(t-s)\int_0^1 g\big[z_\delta\big(S(x),S(y)\big)+\sqrt{(t-s)}B_\delta\big]\,\mathrm{d}\delta\bigg],
\end{aligned} \tag{38}
$$

$$
\begin{aligned}
&\left|\frac{\partial^2}{\partial y^2}E\exp\bigg[(t-s)\int_0^1 g\big[z_\delta\big(S(x),S(y)\big)+\sqrt{(t-s)}B_\delta\big]\,\mathrm{d}\delta\bigg]\right| \\
&\quad \le C(t-s)^2E\exp\bigg[(t-s)\int_0^1 g\big[z_\delta\big(S(x),S(y)\big)+\sqrt{(t-s)}B_\delta\big]\,\mathrm{d}\delta\bigg].
\end{aligned} \tag{39}
$$

The second and the third inequality of the statement of the lemma now follow from our assumptions and from (4), (7), (36)–(39).

It remains to show (33). For a proof of this claim, note that

$$
\begin{aligned}
&\left(\frac{\partial^2}{\partial x\,\partial y} - \frac{\partial^2}{\partial x^2}\right)\widehat{p}(t-s,x,y) \\
&\quad = \widehat{p}(t-s,x,y)\Bigg[\left(\frac{S(y)-S(x)}{(t-s)\sigma(x)} - H'(x)\right) \\
&\qquad\qquad \times\left(-\frac{\sigma'(y)}{\sigma(y)} + H'(y) - H'(x) + \frac{(S(y)-S(x))(\sigma^{-1}(x)-\sigma^{-1}(y))}{t-s}\right) \\
&\qquad\qquad + \left(-H''(x) - \frac{\sigma'(x)}{\sigma^2(x)}\frac{(S(y)-S(x))}{t-s} - \frac{1}{\sigma(x)}\frac{\sigma^{-1}(x)-\sigma^{-1}(y)}{t-s}\right)\Bigg].
\end{aligned}
$$

Claim (33) follows from our assumptions and (4). □

Put

$$\delta_1(x, y) = \sqrt{h}\frac{\pi_1(kh, x, y)}{p(kh, x, y)}.$$

We will also make use of the following bound.

Lemma 9. *There exists a constant C such that for $x, y \in \mathbb{R}$*

$$\left|\delta_1(x, y)\right| \le \frac{C}{\sqrt{k}}\left(1 + \frac{|y-x|}{\sqrt{kh}}\right)^3.$$

Proof. Note that by definition of π_1:

$$\begin{aligned} 6h^{-1/2}\delta_1(x, y)p(kh, x, y) &= \int_0^{kh} \mathrm{d}u \int p(u, x, \xi)\mu_3(\xi)\frac{\partial^3}{\partial \xi^3} p(kh-u, \xi, y)\,\mathrm{d}\xi \\ &= \int_0^{kh/2} \mathrm{d}u \cdots + \int_{kh/2}^{kh} \mathrm{d}u \cdots \\ &\doteqdot \Im_1 + \Im_2. \end{aligned} \tag{40}$$

We now apply the estimates of Lemma 8 to obtain the upper bounds for $\Im_1$ and $\Im_2$ in (40). For $u \in [\frac{kh}{2}, kh]$, we apply two times integrations by parts. From our assumptions on $\mu_3(\xi)$ and from (7), (28) and (29) we obtain that

$$\begin{aligned} |\Im_2| &= \left|\int_{kh/2}^{kh} \mathrm{d}u \int \frac{\partial^2}{\partial \xi^2}\bigl[p(u, x, \xi)\mu_3(\xi)\bigr]\frac{\partial}{\partial \xi} p(kh-u, \xi, y)\,\mathrm{d}\xi\right| \\ &\le \int_{kh/2}^{kh} \mathrm{d}u \int \left|\frac{\partial^2}{\partial \xi^2}\bigl[p(u, x, \xi)\mu_3(\xi)\bigr]\right|\left|\frac{\partial}{\partial \xi} p(kh-u, \xi, y)\right| \mathrm{d}\xi \\ &\le C\int_{kh/2}^{kh} \mathrm{d}u \int \frac{p(u, x, \xi)}{u}\left(1 + \sqrt{u} + \frac{|S(\xi) - S(x)|}{\sqrt{u}}\right)^2 \\ &\qquad \times \frac{p(kh-u, \xi, y)}{\sqrt{kh-u}}\left(\sqrt{kh-u} + \frac{|S(y) - S(\xi)|}{\sqrt{kh-u}}\right)\mathrm{d}\xi \\ &\le \frac{C}{kh}\exp[2Mkh]\int_{kh/2}^{kh} \frac{\mathrm{d}u}{\sqrt{kh-u}} \int \widehat{p}(u, x, \xi)\widehat{p}(kh-u, \xi, y)\left(1 + \sqrt{u} + \frac{|S(\xi) - S(x)|}{\sqrt{u}}\right)^2 \\ &\qquad \times \left(\sqrt{kh-u} + \frac{|S(y) - S(\xi)|}{\sqrt{kh-u}}\right)\mathrm{d}\xi. \end{aligned} \tag{41}$$

For $u \in [0, \frac{kh}{2}]$ we get from (7), (28) and (29) again by applying integration by parts:

$$
\begin{aligned}
|\Im_1| &= \int_0^{kh/2} \mathrm{d}u \int \left| \frac{\partial}{\partial \xi}\big[p(u,x,\xi)\mu_3(\xi)\big]\right| \left| \frac{\partial^2}{\partial \xi^2} p(kh-u,\xi,y)\right| \mathrm{d}\xi \\
&\le C \int_0^{kh/2} \mathrm{d}u \int \frac{p(u,x,\xi)}{\sqrt{u}} \left(\sqrt{u} + \frac{|S(\xi)-S(x)|}{\sqrt{u}}\right) \\
&\qquad \times \frac{p(kh-u,\xi,y)}{kh-u} \left(1 + \sqrt{kh-u} + \frac{|S(y)-S(\xi)|}{\sqrt{kh-u}}\right)^2 \mathrm{d}\xi \qquad (42) \\
&\le \frac{C}{kh} \exp[2Mkh] \int_0^{kh/2} \frac{\mathrm{d}u}{\sqrt{u}} \int \widehat{p}(u,x,\xi)\widehat{p}(kh-u,\xi,y) \left(\sqrt{u} + \frac{|S(\xi)-S(x)|}{\sqrt{u}}\right) \\
&\qquad \times \left(1 + \sqrt{kh-u} + \frac{|S(y)-S(\xi)|}{\sqrt{kh-u}}\right)^2 \mathrm{d}\xi.
\end{aligned}
$$

We now use the following substitution:

$$
\begin{aligned}
u' &= kh - u, \\
z(\xi) &= \left(\frac{kh}{kh-u'}\right)^{1/2} \frac{(S(\xi)-S(y))}{\sqrt{u'}} \\
&\quad + \left(\frac{u'}{kh-u'}\right)^{1/2} \frac{(S(y)-S(x))}{\sqrt{kh}}.
\end{aligned}
$$

Note that

$$
\mathrm{d}\xi = \big(kh-u'\big)^{1/2}\big(u'\big)^{1/2}(kh)^{-1/2}\sigma(\xi)\,\mathrm{d}z, \qquad (43)
$$

$$
\begin{aligned}
z^2 + \frac{(S(y)-S(x))^2}{kh} &= \frac{kh}{(kh-u')} \frac{(S(\xi)-S(y))^2}{(u')} \\
&\quad + \frac{(u')}{(kh-u')} \frac{(S(y)-S(x))^2}{kh} \\
&\quad + 2\frac{(S(\xi)-S(y))(S(y)-S(x))}{kh-u'} + \frac{(S(y)-S(x))^2}{kh} \\
&= \frac{(S(\xi)-S(y))^2}{u'} + \frac{(S(\xi)-S(y))^2}{kh-u'} \qquad (44) \\
&\quad + 2\frac{(S(\xi)-S(y))(S(y)-S(x))}{kh-u'} + \frac{(S(y)-S(x))^2}{kh-u'} \\
&= \frac{(S(\xi)-S(y))^2}{u'} + \frac{(S(\xi)-S(x))^2}{kh-u'}.
\end{aligned}
$$

From (43) and (44), we get that

$$
\begin{aligned}
|\Im_1| \le{}& \frac{C}{kh}\exp[2Mkh]\exp\bigl[H(y)-H(x)\bigr]\\
&\times\int_{kh/2}^{kh}\frac{\mathrm{d}u'}{\sqrt{kh-u'}}\int\frac{1}{\sqrt{2\pi(kh-u')}\sigma(\xi)}\frac{1}{\sqrt{2\pi(u')}\sigma(y)}\\
&\qquad\times\exp\biggl[-\frac{(S(\xi)-S(x))^2}{2(kh-u')}-\frac{(S(y)-S(\xi))^2}{2(u')}\biggr]\\
&\qquad\times\biggl(\sqrt{kh-u'}+\frac{|S(\xi)-S(x)|}{\sqrt{kh-u'}}\biggr)\\
&\qquad\times\biggl(1+\sqrt{u'}+\frac{|S(y)-S(\xi)|}{\sqrt{u'}}\biggr)^2\,\mathrm{d}\xi\\
\le{}& \frac{C}{kh}\exp[2Mkh]\frac{\exp[H(y)-H(x)]}{\sqrt{2\pi kh}\sigma(y)}\exp\biggl(-\frac{(S(y)-S(x))^2}{2kh}\biggr)\\
&\times\int_{kh/2}^{kh}\frac{\mathrm{d}u'}{\sqrt{kh-u'}}\int\frac{\mathrm{d}z}{\sqrt{2\pi}}\exp\biggl(-\frac{z^2}{2}\biggr)\\
&\times\biggl(1+\sqrt{u'}+\sqrt{\frac{u'}{kh}}\frac{|S(y)-S(x)|}{\sqrt{kh}}+|z|\sqrt{\frac{kh-u'}{kh}}\biggr)^2\\
&\times\biggl(\sqrt{kh-u'}+\sqrt{\frac{kh-u'}{kh}}\frac{|S(y)-S(x)|}{\sqrt{kh}}+|z|\sqrt{\frac{u'}{kh}}\biggr)\\
\le{}& \frac{C}{kh}\exp[2Mkh]\widehat{p}(kh,x,y)\int_{kh/2}^{kh}\frac{\mathrm{d}u'}{\sqrt{kh-u'}}\int\frac{\mathrm{d}z}{\sqrt{2\pi}}\exp\biggl(-\frac{z^2}{2}\biggr)\\
&\times\biggl(1+\sqrt{kh}+\frac{|S(y)-S(x)|}{\sqrt{kh}}+|z|\biggr)^2\biggl(\sqrt{kh}+\frac{|S(y)-S(x)|}{\sqrt{kh}}+|z|\biggr)\\
\le{}& \frac{C}{\sqrt{kh}}\exp[2Mkh]\widehat{p}(kh,x,y)\int\frac{\mathrm{d}z}{\sqrt{2\pi}}\exp\biggl(-\frac{z^2}{2}\biggr)\biggl(1+\sqrt{kh}+\frac{|S(y)-S(x)|}{\sqrt{kh}}+|z|\biggr)^3\\
\le{}& \frac{C}{\sqrt{kh}}p(kh,x,y)\biggl(1+\frac{|S(y)-S(x)|}{\sqrt{kh}}\biggr)^3.
\end{aligned}
\tag{45}
$$

By similar calculations, we obtain that

$$
|\Im_2| \le \frac{C}{\sqrt{kh}}p(kh,x,y)\biggl(1+\frac{|S(y)-S(x)|}{\sqrt{kh}}\biggr)^3. \tag{46}
$$

The lemma now follows from our assumptions on σ, (40), (45) and (46). □

We will also make use of the following bound.

Lemma 10. *For any polynomials $P_l(x)$ and $P_m(x)$ of degrees l and m, there exists a constant C, depending only on l, m and the coefficients of the polynomials, such that uniformly for $w \in [0, kh/4)$ the following inequalities hold*

$$\int_0^{kh/4} u^{-1/2} \int p(u, x, z) P_l\left(\left|\frac{z - x}{\sqrt{u}}\right|\right) p(kh - w - u, z, y) P_m\left(\left|\frac{y - z}{\sqrt{kh - w - u}}\right|\right) \mathrm{d}z \, \mathrm{d}u$$
$$\leq C\sqrt{kh}\, p(kh - w, x, y)\left(1 + \left|\frac{y - x}{\sqrt{kh - w}}\right|^{l+m}\right).$$

Proof. These bounds can be easily shown by using the representation (4) and calculations of similar convolution integrals for Gaussian densities. □

Put

$$\delta_2(x, y) = h\frac{\pi_2(kh, x, y)}{p(kh, x, y)}.$$

We now state a bound for $\delta_2(x, y)$.

Lemma 11. *There exists a constant C such that for $x, y \in \mathbb{R}$*

$$\left|\delta_2(x, y)\right| \leq \frac{C}{k}\left[1 + \left(\frac{|y - x|}{\sqrt{kh}}\right)^7\right].$$

Proof. Note that the function $\pi_2(kh, x, y)$ can be written as

$$\pi_2(kh, x, y) = \Im_3 + \Im_4,$$

where

$$\Im_3 = \sum_{i=1}^4 \int_0^{kh} \mathrm{d}u \int p(u, x, \xi) f_i(\xi) \frac{\partial^i}{\partial \xi^i} p(kh - u, \xi, y) \, \mathrm{d}\xi,$$

with $f_4(\xi) = \mu_4(\xi) - 3\sigma^4(\xi)$ and $f_i(\xi)$, $i = 1, 2, 3$, depending on the coefficients of the operator L and their derivatives up to the order 2. Furthermore, the term $\Im_4$ is defined as

$$\Im_4 = \frac{1}{36} \int_{u+w \leq kh; u,w \geq 0} p\bigl(u, x, \xi^*\bigr)\mu_3\bigl(\xi^*\bigr) \frac{\partial^3}{(\partial \xi^*)^3} p\bigl(kh - u - w, \xi^*, \xi\bigr)\mu_3(\xi)$$
$$\times \frac{\partial^3}{(\partial \xi)^3} p(w, \xi, y) \, \mathrm{d}\xi \, \mathrm{d}\xi^* \, \mathrm{d}u \, \mathrm{d}w.$$

Applying the same arguments as in the proof of Lemma 9, we get

$$\left|\sum_{i=1}^{3}\int_0^{kh}\mathrm{d}u\int p(u,x,\xi)f_i(\xi)\frac{\partial^i}{\partial\xi^i}p(kh-u,\xi,y)\,\mathrm{d}\xi\right|\le\frac{C}{\sqrt{kh}}p(kh,x,y)\left(1+\frac{|y-x|}{\sqrt{kh}}\right)^3.$$

For $i=4$, we have to estimate the integral

$$\int_0^{kh}\mathrm{d}u\int p(u,x,\xi)f_4(\xi)\frac{\partial^4}{\partial\xi^4}p(kh-u,\xi,y)\,\mathrm{d}\xi.$$

With calculations very similar to the ones used in the proof of Lemma 9 we get

$$|\Im_3|\le\frac{C}{kh}p(kh,x,y)\left[1+\left(\frac{|y-x|}{\sqrt{kh}}\right)^4\right]. \tag{47}$$

It remains to bound $\Im_4$. We write

$$\Im_4=\Im_{4a}+\Im_{4b}+\Im_{4c},$$

where

$$\begin{aligned}
\Im_{4a}&=\frac{1}{36}\int_{I_a}\cdots\,\mathrm{d}\xi\,\mathrm{d}\xi^*\,\mathrm{d}u\,\mathrm{d}w,\\
\Im_{4b}&=\frac{1}{36}\int_{I_b}\cdots\,\mathrm{d}\xi\,\mathrm{d}\xi^*\,\mathrm{d}u\,\mathrm{d}w,\\
\Im_{4c}&=\frac{1}{36}\int_{I_c}\cdots\,\mathrm{d}\xi\,\mathrm{d}\xi^*\,\mathrm{d}u\,\mathrm{d}w,\\
I_a&=\bigl\{(u,w,\xi,\xi^*):u,w,\xi,\xi^*\in\mathbb{R};u+w\le kh;0\le u;kh/4\le w\bigr\},\\
I_b&=\bigl\{(u,w,\xi,\xi^*):u,w,\xi,\xi^*\in\mathbb{R};u+w\le kh;kh/4\le u;0\le w<kh/4\bigr\},\\
I_c&=\bigl\{(u,w,\xi,\xi^*):u,w,\xi,\xi^*\in\mathbb{R};u+w\le kh;0\le u<kh/4;0\le w<kh/4\bigr\}.
\end{aligned}$$

We now show that for some constant $C>0$

$$|\Im_{4c}|\le\frac{C}{kh}p(kh,x,y)\left[1+\left(\frac{|y-x|}{\sqrt{kh}}\right)^4\right]. \tag{48}$$

For this estimate one applies the following bound that follows by partial integration:

$$\begin{aligned}
|\Im_{4c}|\le\frac{1}{36}\biggl|\int_{I_c}&\frac{\partial}{\partial\xi^*}\bigl[p\bigl(u,x,\xi^*\bigr)\mu_3\bigl(\xi^*\bigr)\bigr]\frac{\partial^4}{(\partial\xi^*)^2(\partial\xi)^2}p\bigl(kh-u-w,\xi^*,\xi\bigr)\mu_3(\xi)\\
&\times\frac{\partial}{\partial\xi}p(w,\xi,y)\,\mathrm{d}\xi\,\mathrm{d}\xi^*\,\mathrm{d}u\,\mathrm{d}w\biggr|.
\end{aligned}$$

The integrand can be bounded with the help of (27), (28) and (32). Because of the bounds of Lemma 10 this implies (48).

To bound $\Im_{4a}$ we use that:

$$\begin{aligned}
36|\Im_{4a}| \le{}& \left|\int_{I_a} \frac{\partial}{\partial\xi^*}\big[p\big(u,x,\xi^*\big)\mu_3(\xi)\big]\left[\frac{\partial^2}{(\partial\xi^*)^2} - \frac{\partial^2}{(\partial\xi^*)(\partial\xi)}\right] p\big(kh-u-w,\xi^*,\xi\big)\mu_3(\xi)\right.\\
&\quad\left.\times \frac{\partial^3}{(\partial\xi)^3} p(w,\xi,y)\,\mathrm{d}\xi\,\mathrm{d}\xi^*\,\mathrm{d}u\,\mathrm{d}w\right|\\
&+\left|\int_{I_a} \frac{\partial}{\partial\xi^*}\big[p\big(u,x,\xi^*\big)\mu_3(\xi)\big]\frac{\partial}{\partial\xi^*} p\big(kh-u-w,\xi^*,\xi\big)\frac{\partial}{\partial\xi}\mu_3(\xi)\right.\\
&\quad\left.\times \frac{\partial^3}{(\partial\xi)^3} p(w,\xi,y)\,\mathrm{d}\xi\,\mathrm{d}\xi^*\,\mathrm{d}u\,\mathrm{d}w\right|\\
&+\left|\int_{I_a} \frac{\partial}{\partial\xi^*}\big[p\big(u,x,\xi^*\big)\mu_3(\xi)\big]\frac{\partial^2}{(\partial\xi^*)(\partial\xi)}\big[p\big(kh-u-w,\xi^*,\xi\big)\mu_3(\xi)\big]\right.\\
&\quad\left.\times \frac{\partial^3}{(\partial\xi)^3} p(w,\xi,y)\,\mathrm{d}\xi\,\mathrm{d}\xi^*\,\mathrm{d}u\,\mathrm{d}w\right|\\
={}& \Im_{4aa} + \Im_{4ab} + \Im_{4ac}.
\end{aligned}$$

These terms can be easily bounded by using the bounds of Lemma 8. Because of the bounds of Lemma 10, this implies

$$|\Im_{4a}| \le \frac{C}{kh} p(kh,x,y)\left[1 + \left(\frac{|y-x|}{\sqrt{kh}}\right)^7\right]. \tag{49}$$

To get a bound for $\Im_{4ac}$ we use that by partial integration:

$$\begin{aligned}
\Im_{4ac} ={}& \left|\int_{I_a} \frac{\partial}{\partial\xi^*}\big[p\big(u,x,\xi^*\big)\mu_3(\xi)\big]\frac{\partial}{\partial\xi^*}\big[p\big(kh-u-w,\xi^*,\xi\big)\mu_3(\xi)\big]\right.\\
&\quad\left.\times \frac{\partial^4}{(\partial\xi)^4} p(w,\xi,y)\,\mathrm{d}\xi\,\mathrm{d}\xi^*\,\mathrm{d}u\,\mathrm{d}w\right|.
\end{aligned}$$

Similarly one shows that

$$|\Im_{4b}| \le \frac{C}{kh} p(kh,x,y)\left[1 + \left(\frac{|y-x|}{\sqrt{kh}}\right)^7\right]. \tag{50}$$

The statement of Lemma 11 follows now from (47), (49), (50) and (48). □

Acknowledgements

This study was carried out within "The National Research University Higher School of Economics" Academic Fund Program in 2012–2013, research Grant 11-01-0083. Support by Grant 436RUS113/467/81-2 from the Deutsche Forschungsgemeinschaft is also acknowledged. The research of Enno Mammen and Jeannette Woerner was supported by the German Science Foundation (DFG) in the framework of the German–Swiss Research Group FOR 916 "Statistical Regularization and Qualitative Constraints".

References

[1] Aït-Sahalia, Y. (2008). Closed-form likelihood expansions for multivariate diffusions. *Ann. Statist.* **36** 906–937. MR2396819

[2] Brown, L.D. and Low, M.G. (1996). Asymptotic equivalence of nonparametric regression and white noise. *Ann. Statist.* **24** 2384–2398. MR1425958

[3] Brown, L.D., Wang, Y. and Zhao, L.H. (2003). On the statistical equivalence at suitable frequencies of GARCH and stochastic volatility models with the corresponding diffusion model. *Statist. Sinica* **13** 993–1013. Statistical applications in financial econometrics. MR2026059

[4] Buchmann, B. and Müller, G. (2012). Limit experiments of GARCH. *Bernoulli* **18** 64–99. MR2888699

[5] Dacunha-Castelle, D. and Florens-Zmirou, D. (1986). Estimation of the coefficients of a diffusion from discrete observations. *Stochastics* **19** 263–284. MR0872464

[6] Dalalyan, A. and Reiß, M. (2006). Asymptotic statistical equivalence for scalar ergodic diffusions. *Probab. Theory Related Fields* **134** 248–282. MR2222384

[7] Dalalyan, A. and Reiß, M. (2007). Asymptotic statistical equivalence for ergodic diffusions: The multidimensional case. *Probab. Theory Related Fields* **137** 25–47. MR2278451

[8] Duval, C. and Hoffmann, M. (2011). Statistical inference across time scales. *Electron. J. Stat.* **5** 2004–2030. MR2870155

[9] Fischer, M. and Nappo, G. (2010). On the moments of the modulus of continuity of Itô processes. *Stoch. Anal. Appl.* **28** 103–122. MR2597982

[10] Genon-Catalot, V., Laredo, C. and Nussbaum, M. (2002). Asymptotic equivalence of estimating a Poisson intensity and a positive diffusion drift. *Ann. Statist.* **30** 731–753. Dedicated to the memory of Lucien Le Cam. MR1922540

[11] Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application. Probability and Mathematical Statistics*. New York: Academic Press [Harcourt Brace Jovanovich Publishers]. MR0624435

[12] Konakov, V. and Mammen, E. (2009). Small time Edgeworth-type expansions for weakly convergent nonhomogeneous Markov chains. *Probab. Theory Related Fields* **143** 137–176. MR2449126

[13] Milstein, G. and Nussbaum, M. (1998). Diffusion approximation for nonparametric autoregression. *Probab. Theory Related Fields* **112** 535–543. MR1664703

[14] Nussbaum, M. (1996). Asymptotic equivalence of density estimation and Gaussian white noise. *Ann. Statist.* **24** 2399–2430. MR1425959

[15] Reiß, M. (2011). Asymptotic equivalence for inference on the volatility from noisy observations. *Ann. Statist.* **39** 772–802. MR2816338

[16] Wang, Y. (2002). Asymptotic nonequivalence of Garch models and diffusions. *Ann. Statist.* **30** 754–783. Dedicated to the memory of Lucien Le Cam. MR1922541

Received January 2012 and revised November 2012