Probab. Theory Relat. Fields (2009) 143:137–176
DOI 10.1007/s00440-007-0123-9

Small time Edgeworth-type expansions for weakly convergent nonhomogeneous Markov chains

Valentin Konakov · Enno Mammen

Received: 14 December 2006 / Revised: 29 October 2007 / Published online: 6 December 2007

Abstract We consider triangular arrays of Markov chains that converge weakly to a diffusion process. Second order Edgeworth type expansions for transition densities are proved. The paper differs from recent results in two respects. We allow nonhomogeneous diffusion limits and we treat transition densities with time lag converging to zero. Small time asymptotics are motivated by statistical applications and by resulting approximations for the joint density of diffusion values at an increasing grid of points.

Keywords Markov chains · Diffusion processes · Transition densities · Edgeworth expansions

Mathematics Subject Classification (2000) Primary: 62G07; Secondary: 60G60

1 Introduction

Recently, there was some activity on Edgeworth-type expansions for dependent data. In most approaches higher order expansions have been derived by application of

This research was supported by grant 436RUS113/467/81-2 from the Deutsche Forschungsgemeinschaft and by grants 05-01-04004 and 04-01-00700 from the Russian Foundation of Fundamental Researches. The first author worked on the paper during a visit at the Laboratory of Probability Theory and Random Models of the University Paris VI in 2006. He is grateful for the hospitality during his stay. We would like to thank Stephane Menozzi, two referees and the associate editor for helpful comments.

V. Konakov
Central Economics Mathematical Institute, Academy of Sciences,
Nahimovskii av. 47, 117418 Moscow, Russia
e-mail: kv24@mail.ru

E. Mammen (✉)
Department of Economics, University of Mannheim, L 7, 3-5, 68131 Mannheim, Germany
e-mail: emammen@rumms.uni-mannheim.de

classical Edgeworth expansions for independent data. The approaches differ in their main idea how the dependence structure can be reduced to the case of independent data. For sums of independent random variables and for functionals of such sums the theory of Edgeworth expansions is classical and well understood in a very general setting (see [5,11]). For models with dependent variables three approaches have been developed where the expansion is derived from models with sums of independent random variables. In the first approach mixing properties are used to approximate the Markov chain by a sum of independent random variables and it is shown that their Edgeworth expansion carries over to the Markov chain up to a certain accuracy. The mixing approach was first used by Götze and Hipp [12] and it was further applied to continuous time processes in [23,31]. Under appropriate conditions Markov chains can be split at regeneration times into a sequence of i.i.d. variables. This fact has been used in [6,7] to get Berry–Esseen bounds for Markov chains. For the statement of Edgeworth expansions the regenerative method has been used in [3,9,17,25]. The higher order Edgeworth expansions have been used to show higher order accuracy of different bootstrap schemes, see [4,10,26].

Both approaches, the mixing method and the regenerative method only have been used for Markov chains with a Gaussian limit. In this paper we study Markov chains that converge weakly to a diffusion limit. For the treatment of this case we make use of the parametrix method. In this approach the transition density is represented as a nested sum of functionals of densities of sums of independent variables. Plugging Edgeworth expansions into this representation will result in an expansion for the transition density. Thus as in the mixing method and in the regenerative method the expansion is reduced to models with sums of independent random variables.

The parametrix method permits to obtain tractable representations of transition densities of diffusions and of Markov chains. For diffusions the parametrix expansion is based on Gaussian densities, see Lemma 2 below, and standard references for the parametrix method are the books of [8] and Ladyzenskaja, Solonnikov and Ural'ceva (1968) on parabolic PDE (see also [27]). For a short exposition of the parametrix method, see Sect. 3 and [19]. Similar representations hold for discrete time Markov chains $X_{k,h}$, see Lemma 4 below. The parametrix method for Markov chains was developed in [19] and it is exposed in Sect. 3.2. In [20] the approach was used to state Edgeworth-type expansions for Euler schemes for stochastic differential equations. Related treatments of Euler schemes can be found in [1,2,13–16,28].

In this paper we study triangular arrays of Markov chains $X_{k,h}$ $(k \geq 0)$ that converge weakly to a diffusion process Y_s $(s \geq 0)$ for $n \longrightarrow \infty$. We consider the Markov chains for the time interval $(0 \leq k \leq n)$. The corresponding time interval of the diffusion is $(0 \leq s \leq T)$. The term $h = T/n$ denotes the discretization step. We allow that T depends on n. In particular, we consider the case that $T \to 0$ for $n \to \infty$. Furthermore, we allow nonhomogeneous diffusion limits.

Weak convergence of the distribution of scaled discrete time Markov processes to diffusions has been extensively studied in the literature (see [29,30]). Local limit theorems for Markov chains were given in [18–20]. In [19] it was shown that the transition density of a Markov chain converges with rate $O(n^{-1/2})$ to the transition density in the diffusion model. For the proof there an analytical approach was chosen that made essential use of the parametrix method.

The main result of this paper will give Edgeworth type expansions for the transition densities of the Markov chains $X_{k,h}$ $(0 \le k \le n)$. The first order term of the expansion is the transition density of the diffusion process Y_s $(0 \le s \le T)$. The order of the expansion is $o(h^{-1-\delta})$ with $\delta > 0$. Related results were shown in [21]. The work of this paper generalizes the results in [21] in two directions. The time horizon T is allowed to converge to 0 and also cases are treated with nonhomogeneous diffusion limit. Small time asymptotics is done for two reasons. First of all it allows approximations for the joint density of values of the Markov chain at an increasing grid of points. Secondly, it is motivated by statistical applications. In statistics, diffusion models are used as an approximation to the truth. They can be motivated by a high frequency Markov chain that is assumed to run in the background on a very fine time grid and is only observed on a coarser grid. If the number of time steps between two observed values of the process converges to infinity this allows diffusion approximations (under appropriate conditions). 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. For the finest scale discrete pattern in the price processes become transparent and do not allow diffusion approximations. The statistical implications of our result will be discussed elsewhere. The mathematical treatment of nonhomogeneous diffusion limits with time horizon T going to zero contributes some additional qualitatively new problems. In this case some additional terms appear that explode for $T \to 0$ and for this reason these terms need a qualitatively different treatment as in the case with fixed T. The nonhomogeneity adds an additional term in the Edgeworth expansion. See also below for more details.

The paper is organized as follows. In the next section we will present our model for the Markov chain and state our main result that gives an Edgeworth-type expansion for Markov chains. Connections with previously known results are also discussed in Sect. 2. In Sect. 3.1 we will give a short introduction into the parametrix method for diffusions. In Sect. 3.2 we will recall the parametrix approach developed in [19] for Markov chains. Technical discussions, auxiliary results and proofs are given in Sects. 4 and 5.

2 The main result: an Edgeworth-type expansion for Markov chains converging to diffusions

We consider a family of Markov processes in $\mathbb{R}^d$ that have the following form

$$X_{k+1,h} = X_{k,h} + m\left(kh, X_{k,h}\right) h + \sqrt{h}\xi_{k+1,h}, \quad X_{0,h} = x \in \mathbb{R}^d, \quad k = 0, \ldots, n-1. \tag{1}$$

The innovation sequence $\left(\xi_{i,h}\right)_{i=1,\ldots,n}$ is assumed to satisfy the Markov assumption: the conditional distribution of $\xi_{k+1,h}$ given the past $X_{k,h} = x_k, \ldots, X_{0,h} = x_0$ depends only on the last value $X_{k,h} = x_k$ and has a conditional density $q\,(kh, x_k, \cdot)$. The conditional covariance matrix corresponding to this density is denoted by $\sigma\,(kh, x_k)$ and the conditional νth cumulant by $\chi_\nu(kh, x_k)$. The transition densities of $\left(X_{i,h}\right)_{i=1,\ldots,n}$ are denoted by $p_h(0, kh, x, \cdot)$. The time horizon $T = T(n) \leq 1$ is allowed to depend on n and $h = T/n$ is the discretization step.

We make the following assumptions.

(A1) It holds that $\int_{\mathbb{R}^d} yq(t, x, y)\, dy = 0$ for $0 \leq t \leq 1,\ x \in \mathbb{R}^d$.

(A2) There exist positive constants $\sigma_\star$ and $\sigma^\star$ such that the covariance matrix $\sigma(t, x) = \int_{\mathbb{R}^d} yy^T q(t, x, y)\, dy$ satisfies

$$\sigma_\star \leq \theta^T \sigma(t, x)\, \theta \leq \sigma^\star$$

for all $\|\theta\| = 1$ and $t \in [0, 1]$ and $x \in \mathbb{R}^d$.

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

$$\left| D_y^\nu q(t, x, y) \right| \leq \psi(y), \quad t \in [0, 1],\ x, y \in \mathbb{R}^d\ |\nu| = 0, 1, 2, 3, 4$$

and

$$\left| D_x^\nu q(t, x, y) \right| \leq \psi(y), \quad t \in [0, 1],\ x, y \in \mathbb{R}^d\ |\nu| = 0, 1, 2.$$

Moreover, for all $x, y \in R^d$, $h > 0, 0 \leq t, t + jh \leq 1$, $j \geq j_0$, with a bound j_0 that does not depend on x, t,

$$\left| D_x^\nu q^{(j)}(t, x, y) \right| \leq C j^{-d/2} \psi\left(j^{-1/2} y \right), \quad |\nu| = 0, 1, 2, 3$$

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

$$q^{(j)}(t, x, y) = \int q^{(j-1)}(t, x, u) q(t + (j-1)h, x, y - u) du,$$

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

Note that the last condition is motivated by (A2) and the classical local limit theorem. Note also that for $1 \leq j \leq j_0$

$$\int \|y\|^S q^{(j)}(t, x, y) dy \leq C(j_0, S).$$

(B1) The functions $m(t, x)$ and $\sigma(t, x)$ and their first and second derivatives w.r.t. t and their derivatives up to the order six w.r.t. x are continuous and bounded uniformly in t and x. All these functions are Lipschitz continuous with respect to x with a Lipschitz constant that does not depend on t. The functions $\chi_\nu(t, x)$, $|\nu| = 3, 4$, are Lipschitz continuous with respect to t with a Lipschitz constant that does not depend on x. A sufficient condition for this is the following inequality

$$\int_{\mathbb{R}^d} (1 + \|z\|^4) \left|q\,(t, x, z) - q\left(t', x, z\right)\right| dz \le C \left|t - t'\right|, 0 \le t, t' \le 1, x \in \mathbb{R}^d,$$

with a constant that does not depend on $x \in \mathbb{R}^d$. Furthermore, $D_x^\nu \sigma\,(t, x)$ exist for $|\nu| \le 6$ and are Holder continuous w.r.t. x with a positive exponent and a constant that does not depend on t.

(B2) There exists $\varkappa < \frac{1}{5}$ such that $\liminf_{n\to\infty} T(n)n^{\varkappa} > 0$.

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

$$dY_s = m\,(s, Y_s)\,ds + \Lambda\,(s, Y_s)\,dW_s, \quad Y_0 = x \in \mathbb{R}^d, \quad s \in [0, T],$$

where $(W_s)_{s\ge 0}$ is the standard Wiener process and Λ is a symmetric positive definite $d \times d$ matrix such that $\Lambda\,(s, y)\,\Lambda\,(s, y)^T = \sigma\,(s, y)$. The conditional density of Y_t, given $Y_0 = x$ is denoted by $p\,(0, t, x, \cdot)$. We will use the following differential operators L and $\widetilde{L}$:

$$Lf(s, t, x, y) = \frac{1}{2} \sum_{i,j=1}^{d} \sigma_{ij}(s, x) \frac{\partial^2 f(s, t, x, y)}{\partial x_i \partial x_j} + \sum_{i=1}^{d} m_i(s, x) \frac{\partial f(s, t, x, y)}{\partial x_i},$$

$$\tilde{L} f(s, t, x, y) = \frac{1}{2} \sum_{i,j=1}^{d} \sigma_{ij}(s, y) \frac{\partial^2 f(s, t, x, y)}{\partial x_i \partial x_j} + \sum_{i=1}^{d} m_i(s, y) \frac{\partial f(s, t, x, y)}{\partial x_i}. \quad (2)$$

To formulate our main result we need also the following operators

$$L' f(s, t, x, y) = \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial \sigma_{ij}(s, x)}{\partial s} \frac{\partial^2 f\,(s, t, x, y)}{\partial x_i \partial x_j} + \sum_{i=1}^{d} \frac{\partial m_i(s, x)}{\partial s} \frac{\partial f\,(s, t, x, y)}{\partial x_i},$$

$$\widetilde{L}' f(s, t, v, z) = \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial \sigma_{ij}(s, y)}{\partial s} \frac{\partial^2 f\,(s, t, x, y)}{\partial x_i \partial x_j} + \sum_{i=1}^{d} \frac{\partial m_i(s, y)}{\partial s} \frac{\partial f\,(s, t, x, y)}{\partial x_i}. \quad (3)$$

and the convolution type binary operation $\otimes$:

$$f \otimes g\,(s, t, x, y) = \int_s^t du \int_{R^d} f\,(s, u, x, z)\,g\,(u, t, z, y)\,dz.$$

Konakov and Mammen [19] obtained a nonuniform rate of convergence for the difference $p_h(0, T, x, \cdot) - p(0, T, x, \cdot)$ as $n \to \infty$ in the case $T \asymp 1$. Edgeworth type expansions for the case $T \asymp 1$ and homogenous diffusions were obtained in [21]. The goal of the present paper is to obtain an Edgeworth type expansion for nonhomogeneous case which remains valid for the both cases $T \asymp 1$ or $T = o(1)$. The following theorem contains our main result. It gives Edgeworth type expansions for p_h. For the statement of the theorem we introduce the following differential operators

$$\mathcal{F}_1[f](s, t, x, y) = \sum_{|\nu|=3} \frac{\chi_\nu(s, x)}{\nu!} D_x^\nu f(s, t, x, y),$$
$$\mathcal{F}_2[f](s, t, x, y) = \sum_{|\nu|=4} \frac{\chi_\nu(s, y)}{\nu!} D_x^\nu f(s, t, x, y).$$

Furthermore, we introduce two terms corresponding to the classical Edgeworth expansion (see [5])

$$\widetilde{\pi}_1(s, t, x, y) = (t - s) \sum_{|\nu|=3} \frac{\overline{\chi}_\nu(s, t, y)}{\nu!} D_x^\nu \widetilde{p}(s, t, x, y), \tag{4}$$

$$\widetilde{\pi}_2(s, t, x, y) = (t - s) \sum_{|\nu|=4} \frac{\overline{\chi}_\nu(s, t, y)}{\nu!} D_x^\nu \widetilde{p}(s, t, x, y) + \frac{1}{2}(t - s)^2 \left\{ \sum_{|\nu|=3} \frac{\overline{\chi}_\nu(s, t, y)}{\nu!} D_x^\nu \right\}^2 \widetilde{p}(s, t, x, y), \tag{5}$$

where

$$\overline{\chi}_\nu(s, t, y) = \frac{1}{t - s} \int_s^t \chi_\nu(u, y) du$$

and $\chi_\nu(t, x)$ is the νth cumulant of the density of the innovations $q(t, x, \cdot)$. The gaussian transition densities $\widetilde{p}(s, t, x, y)$ are defined in (6). Note, that in the homogenous case $\chi_\nu(u, y) \equiv \chi_\nu(y)$ and $\overline{\chi}_\nu(s, t, y) \equiv \chi_\nu(y)$, where $\chi_\nu(y)$ is the νth cumulant of the density $q(y, \cdot)$.

Theorem 1 *Assume* (A1)–(A3), (B1)–(B2). *Then there exists a constant* $\delta > 0$ *such that the following expansion holds*:

$$\sup_{x, y \in R^d} T^{d/2} \left(1 + \left\| \frac{y - x}{\sqrt{T}} \right\|^{S'} \right) \times \Big| p_h(0, T, x, y) - p(0, T, x, y) - h^{1/2} \pi_1(0, T, x, y) - h \pi_2(0, T, x, y) \Big| = O(h^{1+\delta}),$$

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

$$\pi_1(0,T,x,y) = (p \otimes \mathcal{F}_1[p])(0,T,x,y),$$
$$\pi_2(0,T,x,y) = (p \otimes \mathcal{F}_2[p])(0,T,x,y) + p \otimes \mathcal{F}_1[p \otimes \mathcal{F}_1[p]](0,T,x,y) + \frac{1}{2} p \otimes (L_\star^2 - L^2) p(0,T,x,y) - \frac{1}{2} p \otimes (L' - \widetilde{L}') p(0,T,x,y).$$

Here $p(s,t,x,y)$ is the transition density of the limiting diffusion Y_s and the operator $L_\star$ is defined as $\widetilde{L}$, but with the coefficients "frozen" at the point x. The norm $\|\cdot\|$ is the usual Euclidean norm.

Remark 1 The terms of the Edgeworth expansion have subgaussian tails and are of order $n^{-1/2}$ or n^{-1}, respectively:

$$\left|h^{1/2}\pi_1(0,T,x,y)\right| \le C_1 n^{-1/2} T^{-d/2} \exp\left[-C_2 \left\|\frac{y-x}{\sqrt{T}}\right\|^2\right],$$
$$|h\pi_2(0,T,x,y)| \le C_1 n^{-1} T^{-d/2} \exp\left[-C_2 \left\|\frac{y-x}{\sqrt{T}}\right\|^2\right],$$

with some positive constants C_1 and C_2.

Remark 2 If the innovation density $q(t,x,\cdot)$ and the conditional mean $m(t,x)$ do not depend on x then we are in the classical case of independent nonidentically distributed random vectors. We now show that then the Edgeworth expansion of Theorem 1 coincides with the first two terms of the classical Edgeworth expansion $h^{1/2}\widetilde{\pi}_1(0,T,x,y) + h\widetilde{\pi}_2(0,T,x,y)$. Note first that in this case $L_\star = L$, $L' = \widetilde{L}'$ and $p(s,t,x,y) = \widetilde{p}(s,t,x,y)$ where $\widetilde{p}$ is defined in (6) with $\sigma(s,t,y) = \sigma(s,t) = \int_s^t \sigma(u)du$ and $m(s,t,y) = m(s,t) = \int_s^t m(u)du$. This gives

$$\begin{aligned}
\pi_1(0,T,x,y) &= \int_0^T ds \int \widetilde{p}(0,s,x,v) \sum_{|\nu|=3} \frac{\chi_\nu(s)}{\nu!} D_v^\nu \widetilde{p}(s,T,v,y)dv \\
&= -\sum_{|\nu|=3} \int_0^T \frac{\chi_\nu(s)}{\nu!} ds D_y^\nu \int \widetilde{p}(0,s,x,v)\widetilde{p}(s,T,v,y)dv \\
&= -\sum_{|\nu|=3} \frac{T}{\nu!} \overline{\chi}_\nu(0,T) D_y^\nu \widetilde{p}(0,T,x,y) \\
&= \sum_{|\nu|=3} \frac{T}{\nu!} \overline{\chi}_\nu(0,T) D_x^\nu \widetilde{p}(0,T,x,y) = \widetilde{\pi}_1(0,T,x,y),
\end{aligned}$$

$$\widetilde{p} \otimes \mathcal{F}_1[\widetilde{p}](s,T,z,y) = \int_s^T du \int \widetilde{p}(s,u,z,w) \sum_{|\nu|=3} \frac{\chi_\nu(u)}{\nu!} D_w^\nu \widetilde{p}(u,T,w,y) dw$$
$$= -\sum_{|\nu|=3} \int_s^T \frac{\chi_\nu(u)}{\nu!} du \; D_y^\nu \widetilde{p}(s,T,z,y)$$
$$= (T-s) \sum_{|\nu|=3} \frac{\overline{\chi}_\nu(s,T)}{\nu!} D_z^\nu \widetilde{p}(s,T,z,y),$$

$$\mathcal{F}_1[\widetilde{p} \otimes \mathcal{F}_1[\widetilde{p}]](s,T,z,y) = (T-s) \sum_{|\nu|=3} \frac{\chi_\nu(s)}{\nu!} D_z^\nu \times \left[\sum_{|\nu'|=3} \frac{\overline{\chi}_{\nu'}(s,T)}{\nu'!} D_z^{\nu'} \widetilde{p}(s,T,z,y) \right]$$
$$= (T-s) \sum_{|\nu|=3, |\nu'|=3} \frac{\chi_\nu(s)}{\nu!} \frac{\overline{\chi}_{\nu'}(s,T)}{\nu'!} D_z^{\nu+\nu'} \widetilde{p}(s,T,z,y),$$

$$\widetilde{p} \otimes \mathcal{F}_2[\widetilde{p}](0,T,x,y) + \widetilde{p} \otimes \mathcal{F}_1[\widetilde{p} \otimes \mathcal{F}_1[\widetilde{p}]](0,T,x,y)$$
$$= T \sum_{|\nu|=4} \frac{\overline{\chi}_\nu(0,T)}{\nu!} D_x^\nu \widetilde{p}(0,T,x,y)$$
$$+ \int_0^T ds \int \widetilde{p}(0,s,x,z)(T-s) \sum_{|\nu|=3, |\nu'|=3} \frac{\chi_\nu(s)}{\nu!} \frac{\overline{\chi}_{\nu'}(s,T)}{\nu'!} D_z^{\nu+\nu'} \widetilde{p}(s,T,z,y) dz$$
$$= T \sum_{|\nu|=4} \frac{\overline{\chi}_\nu(0,T)}{\nu!} D_x^\nu \widetilde{p}(0,T,x,y)$$
$$+ \sum_{|\nu|=3, |\nu'|=3} \frac{1}{\nu!} \frac{1}{\nu'!} \int_0^T \chi_\nu(s) \left(\int_s^T \chi_{\nu'}(u) du \right) ds D_x^{\nu+\nu'} \widetilde{p}(s,T,x,y).$$

For $\nu = \nu'$ we have

$$\int_0^T \chi_\nu(s) \left(\int_s^T \chi_{\nu'}(u) du \right) ds = \frac{1}{2} \int_0^T \int_0^T \chi_\nu(s) \chi_\nu(u) ds du = \frac{T^2}{2} \overline{\chi}_\nu(0,T) \overline{\chi}_\nu(0,T).$$

For $\nu \neq \nu'$ we get

$$\int_0^T \chi_\nu(s) \left(\int_s^T \chi_{\nu'}(u) du \right) ds + \int_0^T \chi_{\nu'}(s) \left(\int_s^T \chi_\nu(u) du \right) ds$$

$$
\begin{aligned}
&= \int_0^T \int_s^T [\chi_\nu(s)\chi_{\nu'}(u) + \chi_{\nu'}(s)\chi_\nu(u)]\, dsdu \\
&= \frac{1}{2} \int_0^T \int_0^T [\chi_\nu(s)\chi_{\nu'}(u) + \chi_{\nu'}(s)\chi_\nu(u)]\, dsdu \\
&= \frac{T^2}{2} \overline{\chi}_\nu(0,T)\overline{\chi}_{\nu'}(0,T) + \frac{T^2}{2} \overline{\chi}_{\nu'}(0,T)\overline{\chi}_\nu(0,T).
\end{aligned}
$$

From these equations we obtain

$$
\begin{aligned}
&\widetilde{p} \otimes \mathcal{F}_2[\widetilde{p}](0,T,x,y) + \widetilde{p} \otimes \mathcal{F}_1[\widetilde{p} \otimes \mathcal{F}_1[\widetilde{p}](0,T,x,y) \\
&= T \sum_{|\nu|=4} \frac{\overline{\chi}_\nu(0,T)}{\nu!} D_x^\nu \widetilde{p}(0,T,x,y) + \frac{T^2}{2} \left\{ \sum_{|\nu|=3} \frac{\overline{\chi}_\nu(0,T)}{\nu!} D_x^\nu \right\}^2 \widetilde{p}(0,T,x,y) \\
&= \widetilde{\pi}_2(0,T,x,y).
\end{aligned}
$$

This shows the claim that we get for this case the first two terms of the classical Edgeworth expansion.

Remark 3 If $\chi_\nu(t,x) = 0$ for $|\nu| = 3$ and for $t \in [0,T] \times R^d$ then it holds that $\mathcal{F}_1 \equiv 0$. The Theorem 1 holds with

$$
\begin{aligned}
\pi_1(0,T,x,y) &= 0, \\
\pi_2(0,T,x,y) &= (p \otimes \mathcal{F}_2[p])(0,T,x,y) + \frac{1}{2} p \otimes (L_\star^2 - L^2)p(0,T,x,y) \\
&\quad - \frac{1}{2} p \otimes (L' - \widetilde{L}')p(0,T,x,y).
\end{aligned}
$$

If in addition $\chi_\nu(t,x) = 0$ for $|\nu| = 4$ then the first four moments of the innovations coincide with the first four moments of a normal distribution with zero mean and covariance matrix $\sigma(t,x)$. In this case we have $\mathcal{F}_2 = 0$ and we have

$$
\begin{aligned}
\pi_1(0,T,x,y) &= 0, \\
\pi_2(0,T,x,y) &= \frac{1}{2} p \otimes (L_\star^2 - L^2)p(0,T,x,y) - \frac{1}{2} p \otimes (L' - \widetilde{L}')p(0,T,x,y)
\end{aligned}
$$

and the first two terms of the Edgeworth expansion do not depend on the innovation density. In particular, it holds that $\chi_\nu(t,x) = 0$ for $|\nu| = 3, 4$ for Markov chains that are defined by Euler approximations to diffusions. Thus, an Edgeworth expansion for the Euler scheme holds with the same π_1 and π_2 as just defined. For the homogenous case we have that $L' = \widetilde{L}' = 0$ and we obtain for the Euler scheme in this case

$$
\begin{aligned}
\pi_1(0,T,x,y) &= 0, \\
\pi_2(0,T,x,y) &= \frac{1}{2} p \otimes (L_\star^2 - L^2)p(0,T,x,y).
\end{aligned}
$$

This result for $T = [0, 1]$ under Hormander's condition on a diffusion matrix was obtained by Bally and Talay [1,2].

Remark 4 We now shortly discuss an application of our result to statistics. Assume that one observes a Markov process $X_{1,h}, \dots, X_{nk,h}$ at time points $k, 2k, \dots, nk$. 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. 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 process 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 on the equidistant grid $kh, 2kh, \dots, nkh$. The joint distribution of $(Y_1, \dots, Y_n)$ is denoted by Q_h. According to our theorem the one-dimensional marginal distributions of P_h can be approximated by the one-dimensional marginal distributions of Q_h. Under appropriate conditions the L_1-norm of this difference is of order $k^{-1/2}$. This implies that the L_1-norm of the difference between the joint distributions P_h and Q_h is of order $nk^{-1/2}$. That means the diffusion approximation is only accurate if $k \gg n^2$, i.e., only if the grid of observed points is very coarse in comparison to the grid on which the Markov process lives. Only in this case it can be guaranteed that a statistical inference that is based on the diffusion model is accurate. Or put it in another way, data that come from the Markov model could not be asymptotically statistically distinguished from diffusion observations. Our results help to analyze what may go wrong if $k \gg n^2$ does not hold. The (signed) transition densities $p + h^{1/2}\pi_1 + h\pi_2$ given in the statement of Theorem 1 define a joined (signed) measure R_h. According to Theorem 1, the marginal distributions of R_h approximate the one-dimensional marginal distributions of P_h with order $o(k^{-1-\delta})$. One may conjecture that under some regularity assumptions the exact order is $k^{-3/2}$. This implies that $\| P_h - R_h \|_1$ is of order $nk^{-3/2}$. Thus, this approximation is appropriate as long as $k \gg n^{2/3}$. This is a much more acceptable assumption. Now, one can check which statistical procedures behave differently under the models Q_h and R_h. These procedures may lead to erroneous conclusions for the Markov data.

3 The parametrix method

3.1 The parametrix method for diffusions

We now give a short overview on the parametrix method for diffusions. For any $s \in [0, T]$, $x, y \in \mathbb{R}^d$ we consider the following family of "frozen" diffusion processes

$$d\widetilde{Y}_t = m\,(t, y)\,dt + \Lambda\,(t, y)\,dW_t, \quad \widetilde{Y}_s = x, \quad s \le t \le T.$$

Let $\widetilde{p}^{\,y}\,(s, t, x, \cdot)$ be the conditional density of $\widetilde{Y}_t$, given $\widetilde{Y}_s = x$. In the sequel for any z we will denote $\widetilde{p}\,(s, t, x, z) = \widetilde{p}^{\,z}\,(s, t, x, z)$, where the variable z acts here twice: as the argument of the density and as defining quantity of the process $\widetilde{Y}_t$.

The transition densities $\widetilde{p}$ can be computed explicitly

$$\widetilde{p}\,(s,t,x,y) = (2\pi)^{-d/2}\,(\det \sigma\,(s,t,y))^{-1/2} \times \exp\left(-\frac{1}{2}\,(y-x-m\,(s,t,y))^T\,\sigma^{-1}\,(s,t,y)\,(y-x-m\,(s,t,y))\right), \tag{6}$$

where

$$\sigma\,(s,t,y) = \int_s^t \sigma\,(u,y)\,du, \qquad m\,(s,t,y) = \int_s^t m\,(u,y)\,du.$$

Note that the following differential operators L and $\tilde{L}$ correspond to the infinitesimal operators of Y or of the frozen process $\widetilde{Y}$, respectively, i.e.,

$$Lf(s,t,x,y) = \lim_{h\to 0} h^{-1}\{E[f(s,t,Y(s+h),y)\mid Y(s)=x] - f(s,t,x,y)\},$$
$$\tilde{L}f(s,t,x,y) = \lim_{h\to 0} h^{-1}\{E[f(s,t,\tilde{Y}(s+h),y)\mid \widetilde{Y}(s)=x] - f(s,t,x,y)\}.$$

We put

$$H = (L-\tilde{L})\tilde{p}.$$

Then

$$H\,(s,t,x,y) = \frac{1}{2}\sum_{i,j=1}^{d}\left(\sigma_{ij}\,(s,x) - \sigma_{ij}\,(s,y)\right)\frac{\partial^2 \widetilde{p}\,(s,t,x,y)}{\partial x_i\,\partial x_j} + \sum_{i,j=1}^{d}(m_i\,(s,x) - m_i\,(s,y))\,\frac{\partial \widetilde{p}\,(s,t,x,y)}{\partial x_i}.$$

In the following lemmas the k-fold convolution of H is denoted by $H^{(k)}$. The following results have been proved in [19].

Lemma 2 *Let* $0 \le s < t \le T$. *It holds*

$$p(s,t,x,y) = \sum_{r=0}^{\infty} \widetilde{p} \otimes H^{(r)}(s,t,x,y).$$

Lemma 3 *Let* $0 \le s < t \le T$. *There are constants* C *and* C_1 *such that*

$$|H(s,t,x,y)| \le C_1 \rho^{-1} \phi_{C,\rho}(y-x)$$

and

$$\left|\widetilde{p} \otimes H^{(r)}(s,t,x,y)\right| \le C_1^{r+1} \frac{\rho^r}{\Gamma(1+\frac{r}{2})} \phi_{C,\rho}(y-x),$$

where $\rho^2 = t - s$, $\phi_{C,\rho}(u) = \rho^{-d}\phi_C(u/\rho)$ *and*

$$\phi_C(u) = \frac{\exp(-C\,\|u\|^2)}{\int \exp(-C\,\|v\|^2\,dv)}.$$

3.2 The parametrix method for Markov chains

We now give a short overview on the parametrix method for Markov chains. This theory was developed in [19]. For any $0 \le jh \le T$, $x, y \in \mathbb{R}^d$ we consider an additional family of "frozen" Markov chains defined for $jh \le ih \le T$ as

$$\widetilde{X}_{i+1,h} = \widetilde{X}_{i,h} + m\,(ih, y)\,h + \sqrt{h}\widetilde{\xi}_{i+1,h}, \quad \widetilde{X}_{j,h} = x \in \mathbb{R}^d, \quad j \le i \le n, \tag{7}$$

where $\widetilde{\xi}_{j+1,h}, \ldots, \widetilde{\xi}_{n,h}$ is an innovation sequence such that the conditional density of $\widetilde{\xi}_{i+1,h}$ given the past $\widetilde{X}_{i,h} = x_i, \ldots, \widetilde{X}_{0,h} = x_0$ equals to $q\,(ih, y, \cdot)$. Let us introduce the infinitesimal operators corresponding to Markov chains (1) and (7) respectively,

$$L_h f\,(jh, kh, x, y) = h^{-1}\bigg(\int p_h\,(jh, (j+1)\,h, x, z)\, f\,((j+1)\,h, kh, z, y)\,dz - f\,((j+1)\,h, kh, x, y)\bigg)$$

and

$$\widetilde{L}_h f\,(jh, kh, x, y) = h^{-1}\bigg(\int \widetilde{p}_h^{\,y}\,(jh, (j+1)\,h, x, z)\, f\,((j+1)\,h, kh, z, y)\,dz - f\,((j+1)\,h, kh, x, y)\bigg),$$

where $\widetilde{p}_h^{\,y}\left(jh, j'h, x, \cdot\right)$ denotes the conditional density of $\widetilde{X}_{j',h}$ given $\widetilde{X}_{j,h} = x$. Similarly as above, for brevity for any z we write $\widetilde{p}_h\left(jh, j'h, x, z\right) = \widetilde{p}_h^{\,z}\left(jh, j'h, x, z\right)$, where the variable z acts here twice: as the argument of the density and as defining quantity of the process $\widetilde{X}_{i,h}$. For technical convenience the terms $f\,((j+1)\,h, kh, z, y)$ on the right hand side of $L_h f$ and $\widetilde{L}_h f$ appear instead of $f\,(jh, kh, z, y)$.

In analogy with the definition of H we put, for $k > j$,

$$H_h\,(jh, kh, x, y) = \left(L_h - \widetilde{L}_h\right)\widetilde{p}_h\,(jh, kh, x, y)\,.$$

We also shall use the convolution type binary operation $\otimes_h$ which is a discrete version of $\otimes$:

$$g \otimes_h f(jh, kh, x, y) = \sum_{i=j}^{k-1} h \int_{\mathbb{R}^d} g(jh, ih, x, z) f(ih, kh, z, y)\, dz,$$

where $0 \le j < k \le n$. We write $g \otimes_h H_h^{(0)} = g$ and $g \otimes_h H_h^{(r)} = \left(g \otimes_h H_h^{(r-1)}\right) \otimes_h H_h$ for $r \ge 1$. For the higher order convolutions we use the convention $\sum_{i=j}^{l} = 0$ for $l < j$. One can show the following analog of the "parametrix" expansion for p_h (see [19]).

Lemma 4 *Let* $0 \le jh < kh \le T$. *It holds*

$$p_h(jh, kh, x, y) = \sum_{r=0}^{k-j} \widetilde{p}_h \otimes_h H_h^{(r)}(jh, kh, x, y),$$

where

$$\widetilde{p}_h(jh, jh, x, y) = p_h(kh, kh, x, y) = \delta(y - x)$$

and δ *is the Dirac delta symbol.*

4 Some technical tools

4.1 Plugged in Edgeworth expansions for independent observations

In this Section we will develop some tools that are helpful for the comparison of the expansion of p (see Lemma 2) and the expansion of p_h (see Lemma 4). These expansions are simple expressions in $\tilde{p}$ or $\tilde{p}_h$, respectively. Recall that $\tilde{p}$ is a Gaussian density, see (6), and that $\tilde{p}_h$ is the density of a sum of independent variables. The densities $\tilde{p}$ and $\tilde{p}_h$ can be compared by application of the classical Edgeworth expansions. This is done in Lemma 5 and this is the essential step for the comparison of the expansions of p and p_h. Lemma 7 contains technical tools that will be used below. The proof of Lemma 5 can be found in the extended version of this paper, see [22]. Lemma 7 contains bounds on derivatives of $\tilde{p}_h$ that will be used at several places in the proof of Theorem 1. Its proof makes use of Lemma 6 that is a generalization of a result in [18] (Lemma 4 on page 68). Lemma 5 is a higher order extension of the results from Sect. 3.3 in [19].

Denote

$$\mu_{j,k}(y) = h \sum_{i=j}^{k-1} m(ih, y),\ V_{j,k}(y) = h \sum_{i=j}^{k-1} \sigma(ih, y). \tag{8}$$

Lemma 5 *The following bound holds with a constant C for $\nu = (\nu_1, \dots \nu_p)^T$ with $0 \le |\nu| \le 6$*

$$\Big| D_z^\nu \widetilde{p}_h(jh, kh, x, y) - D_z^\nu \widetilde{p}(jh, kh, x, y) - \sqrt{h} D_z^\nu \widetilde{\pi}_1(jh, kh, x, y) - h D_z^\nu \widetilde{\pi}_2(jh, kh, x, y) \Big| \le C h^{3/2} \rho^{-3} \zeta_\rho^{S-|\nu|}(y-x)$$

for all $j < k$, x and y. Here D_z^ν denotes the partial differential operator of order ν with respect to $z = V_{j,k}^{-1/2}(y)(y - x - \mu_{j,k}(y))$. The quantity ρ denotes again the term $\rho = [h(k-j)]^{1/2}$ and the functions $\widetilde{\pi}_1$ and $\widetilde{\pi}_2$ are defined in (4) and (5). We write $\zeta_\rho^k(\cdot) = \rho^{-d} \zeta^k(\cdot/\rho)$ where

$$\zeta^k(z) = \frac{[1 + \|z\|^k]^{-1}}{\int [1 + \|z'\|^k]^{-1} dz'}.$$

Lemma 6 *Let $L(d)$ be the set of symmetric matrices, and for $0 < \lambda^- < \lambda^+ < \infty$ let $D_{\lambda^+,\lambda^-} \subset L(d)$ be the open subset of $L(d)$ that contains all $\Lambda \in L(d)$ with $\lambda^- I < \Lambda < \lambda^+ I$. For $\Lambda \in L(d)$ define $A = A(\Lambda)$ as the symmetric solution of the equation $A^2 = \Lambda$. Then for any $k, l, i, j \le d$ and $\Lambda \in D_{\lambda^+,\lambda^-}$ we have that with a constant C_m depending on m*

$$\left| \frac{\partial^m a_{ij}(\Lambda)}{(\partial \lambda_{kl})^m} \right| \le C_m (\lambda^-)^{-(2m-1)/2}. \tag{9}$$

Here $a_{ij}(\Lambda)$ are the elements of $A = A(\Lambda)$.

Proof of Lemma 6. For $m = 1$ the lemma was proved in [18] (see Lemma 4). Suppose now that (9) holds for $m \le l$. From the equality $AA = \Lambda$ we obtain for $m = l + 1$

$$d^{l+1}(AA) = (d^{l+1}A)A + \binom{l+1}{1}(d^l A) dA + \cdots + \binom{l+1}{l} dA(d^l A) + A(d^{l+1}A) = 0,$$

where d denotes elementwise differentiation of a matrix with respect to a fixed element of Λ. This implies

$$(d^{l+1}A)A + A(d^{l+1}A) = -\binom{l+1}{1}(d^l A) dA - \cdots - \binom{l+1}{l} dA(d^l A). \tag{10}$$

Denote the symmetric matrix in the right hand side of (10) by $\widetilde{\Lambda}$. Then equality (10) determines a linear operator ℓ mapping $d^{l+1}A$ to $\widetilde{\Lambda}$. In the linear space of symmetric $d \times d$ matrices we introduce the scalar product $\langle X, Y \rangle = \text{trace}(XY)$. The operator ℓ determines a quadratic form

$$\langle \ell X, X\rangle = \mathrm{trace}[(XA + AX)X] = 2\mathrm{trace}[XAX] \geq 2\sqrt{\lambda^-}\mathrm{trace}[XX] = 2\sqrt{\lambda^-}\langle X, X\rangle,$$

where in the inequality we have used that $A - \sqrt{\lambda^-}I$ positive definite implies that $X(A - \sqrt{\lambda^-}I)X = XAX - \sqrt{\lambda^-}XX$ is positive definite. Similarly, we get $\langle \ell X, X\rangle \leq 2\sqrt{\lambda^+}\langle X, X\rangle$. Hence,

$$2\sqrt{\lambda^-} \leq \|\ell\| = \sup_{X\neq 0} \frac{\|\ell X\|}{\|X\|} \leq 2\sqrt{\lambda^+}$$

and

$$\frac{1}{2\sqrt{\lambda^+}} \leq \|\ell^{-1}\| \leq \frac{1}{2\sqrt{\lambda^-}}.$$

We obtain

$$\|d^{l+1}A\| \leq \frac{1}{2\sqrt{\lambda^-}}\|\widetilde{\Lambda}\|.$$

Using the induction hypothesis we get from (10)

$$\|d^{l+1}A\| \leq C_{l+1}(\lambda^-)^{(2l+1)/2}.$$

This completes the proof. □

From Lemmas 5 and 6 we get the following corollary. The statement of the next lemma is an extension of Lemma 3.7 in Konakov and Mammen (2000) where the result has been shown for $0 \leq |b| \leq 2$, $a = 0$. (In formula (3.15) of this lemma there is a typo. Differentiation with respect to u should be replaced by differentiation with respect to x.)

Lemma 7 *The following bound holds*:

$$\left|D_y^a D_x^b \widetilde{p}_h(jh, kh, x, y)\right| \leq C\rho^{-|a|-|b|}\zeta_\rho^{S-|a|}(y - x)$$

for all $j < k$, *for all* x *and* y *and for all* a, b *with* $0 \leq |a| + |b| \leq 6$. *Here,* $\rho = [(k - j)h]^{1/2}$. *The constant* S *has been defined in Assumption* (A3).

Proof of Lemma 7. For two matrices A and B with elements a_{ij} or b_{kl}, respectively where $a_{ij}(B)$ are smooth functions of b_{kl} we write $\left|\frac{\partial A}{\partial B}\right| \leq C$ if $\left|\frac{\partial a_{ij}}{\partial b_{kl}}\right| \leq C$ for all $1 \leq i, j \leq d$, $1 \leq k, l \leq d$. To obtain the assertion of the lemma we have to estimate the derivatives $D_y^a D_x^b z$, where $z = V_{j,k}^{-1/2}(y)(y - x - \mu_{j,k}(y))$. Note that $z = z(x, y)$, where $V_{j,k}^{-1/2} = V_{j,k}^{-1/2}(y)$ and $\mu_{j,k} = \mu_{j,k}(y)$. For $l = 1, \ldots, 6$ it follows from condition (B1) and (8) that

$$\left|\frac{\partial^l \mu_{j,k}(y)}{(\partial y)^l}\right| \leq C\rho^2, \left|\frac{\partial^l V_{j,k}(y)}{(\partial y)^l}\right| \leq C\rho^2. \tag{11}$$

It follows from Lemma 6 that

$$\left|\frac{\partial^l V_{j,k}^{1/2}}{(\partial V_{j,k})^l}\right| \le C\rho^{-(2l-1)/2}. \tag{12}$$

From inequalities (3.16) in [19] and from the representation of an inverse matrix in terms of cofactors divided by the determinant we obtain that

$$\left|\frac{\partial^l V_{j,k}^{-1/2}}{(\partial V_{j,k}^{1/2})^l}\right| \le C\rho^{-(l+1)}. \tag{13}$$

From (11) to (13) and from the chain rule (or from de Bruno's formula) we get

$$\left|\frac{\partial^l V_{j,k}^{-1/2}(y)}{(\partial y)^l}\right| \le C\rho^{-1}. \tag{14}$$

This gives for $l \ge 1$

$$\left|\frac{\partial^l z}{(\partial y)^l}\right| \le C\rho^{-1}.$$

This bound, $\left|\frac{\partial z}{\partial x}\right| \le C\rho^{-1}$ and $\frac{\partial^l z}{(\partial x)^l} = 0$ (for $l \ge 2$) can be plugged into the formula of Lemma 5. This implies the assertion of Lemma 7. □

4.2 Bounds on operator kernels used in the parametrix expansions

In this Section we will present bounds for operator kernels appearing in the expansions based on the parametrix method. In Lemma 8 we compare the infinitesimal operators L_h and $\tilde{L}_h$ with the differential operators L and $\tilde{L}$. We give an approximation for the error if, in the definition of $H_h = (L_h - \tilde{L}_h)\tilde{p}_h$, the terms L_h and $\tilde{L}_h$ are replaced by L or $\tilde{L}$, respectively. We show that this term can be approximated by $K_h + M_h$, where $K_h = (L - \tilde{L})\tilde{p}_h$ and where M_h is defined in Remark 5 after Lemma 8. The bounds obtained in Lemma 9 will be used in the proof of our theorem to show that in the expansion of p_h the terms $\widetilde{p}_h \otimes_h H_h^{(r)}$ can be replaced by $\widetilde{p}_h \otimes_h (K_h + M_h)^{(r)}$.

Lemma 8 *The following bound holds with a constant C*

$$\begin{aligned}&\left|H_h(jh, kh, x, y) - K_h'(jh, kh, x, y) - M_h'(jh, kh, x, y) - R_h(jh, kh, x, y)\right|\\&\quad \le Ch^{3/2}\rho^{-1}\zeta_\rho^S(y - x)\end{aligned}$$

with ζ_ρ^S as in Lemma 5 for all $j < k$, x and y. For $j<k-1$ we define

$$\begin{aligned}
K_h'(jh, kh, x, y) &= (L-\widetilde{L})\lambda(x), M_h'(jh, kh, x, y) \\
&= M_{h,1}(jh, kh, x, y)+M_{h,2}(jh, kh, x, y) + M_{h,3}'(jh, kh, x, y), \\
M_{h,1}(jh, kh, x, y) &= h^{1/2} \sum_{|\nu|=3} \frac{D_x^\nu \lambda(x)}{\nu!}(\chi_\nu(jh, x) - \chi_\nu(jh, y)), \\
M_{h,2}(jh, kh, x, y) &= h \sum_{|\nu|=4} \frac{D_x^\nu \lambda(x)}{\nu!}(\chi_\nu(jh, x) - \chi_\nu(jh, y)), \\
M_{h,3}'(jh, kh, x, y) &= \frac{h}{2}(L_\star^2 - \widetilde{L}^2)\lambda(x), \\
R_h(jh, kh, x, y) &= h^{3/2} \sum_{|\nu|=4} \frac{D_x^\nu \lambda(x)}{\nu!} \sum_{r=1}^{d} \nu_r [m_r(jh, x)\mu_{\nu-e_r}(jh, x) \\
&\quad - m_r(jh, y)\mu_{\nu-e_r}(jh, y)] \\
&\quad + 5 \sum_{|\nu|=5} \frac{1}{\nu!} \sum_{k=1}^{d} (m_k(jh, x) \\
&\quad - m_k(jh, y)) \Bigg\{ \nu_k \int q(jh, x, \theta)\widetilde{h}^{\nu-e_k}(\theta) \\
&\quad \times \left[\int_0^1 (1-u)^4 D^\nu \lambda(x+u\widetilde{h}(\theta))du \right] d\theta \\
&\quad + \int q(jh, x, \theta)\widetilde{h}^\nu(\theta) \\
&\quad \times \left[\int_0^1 (1-u)^4 u D^{\nu+e_k} \lambda(x+u\widetilde{h}(\theta))du \right] d\theta \Bigg\} \\
&\quad + h^2 \sum_{|\nu|=4} \frac{D_x^\nu \lambda(x)}{\nu!} \sum_{|\nu'|=2} \nu! N(\nu, \nu')[m^{\nu'}(jh, x)\mu_{\nu-\nu'}(jh, x) \\
&\quad - m^{\nu'}(jh, y)\mu_{\nu-\nu'}(jh, y)].
\end{aligned}$$

Here $L_\star$ is defined as $\widetilde{L}$ but with the coefficients "frozen" at the point x, e_r denotes a d-dimensional vector with the rth element equal to 1 and with all other elements equal to 0. Furthermore, for $|\nu| = 4$, $\left|\nu'\right| = 2$ we define

$$N(\nu, \nu') = 2^{\chi[\nu'!=1]+\chi[(\nu-\nu')!=1]-2},$$

where $\chi(\cdot)$ is the indicator function. We put $m(x)^\nu = m_1(x)^{\nu_1} \times \cdots \times m_d(x)^{\nu_d}$ and we define $\mu_\nu(t, x) = \int z^\nu q(t, x, z)dz$. We put $m(x)^\nu = 0$, $\mu_\nu(t, x) = 0$, and $\nu! = 0$, if at *least one of the coordinates of* $\nu = (\nu_1, \ldots, \nu_d)$ *is negative. We use also the following definitions*

$$\lambda(x) = \widetilde{p}_h((j+1)h, kh, x, y),$$
$$\widetilde{h}(\theta) = m(jh, y)h + \theta h^{1/2}.$$

Here again ρ *denotes the term* $\rho = [h(k-j)]^{1/2}$ *. For* $j = k-1$ *we define*

$$K_h'(jh, kh, x, y) = R_h(jh, kh, x, y) = M_{h,2}(jh, kh, x, y)$$
$$= M_{h,3}'(jh, kh, x, y) = 0$$

and

$$M_{h,1}(jh, kh, x, y) = h^{-(d+2)/2}\Big[q\left\{jh, x, h^{-1/2}(y - x - m[jh, x]h)\right\} - q\left\{jh, y, h^{-1/2}(y - x - m[jh, y]h)\right\}\Big].$$

Proof of Lemma 8. As in the proof of Lemma 3.9 in [19] we have

$$H_h(jh, kh, x, y) = H_h^1(jh, kh, x, y) - H_h^2(jh, kh, x, y),$$

where

$$H_h^1(jh, kh, x, y) = h^{-1}\int q(jh, x, \theta)[\lambda(x + h(\theta)) - \lambda(x)]d\theta, \tag{15}$$

$$H_h^2(jh, kh, x, y) = h^{-1}\int q(jh, y, \theta)[\lambda(x + \widetilde{h}(\theta)) - \lambda(x)]d\theta, \tag{16}$$

$$h(\theta) = m(jh, x)h + \theta h^{1/2}, \widetilde{h}(\theta) = m(jh, y)h + \theta h^{1/2}.$$

For $[\lambda(x+h(\theta)) - \lambda(x)]$ and $[\lambda(x+\widetilde{h}(\theta)) - \lambda(x)]$ in (15), (16) we use now the Taylor expansion up to order 5 with remaining term in integral form. To pass from moments to cumulants we use the well known relations (see, e.g., relation (6.11) on page 46 in [5]). After long but simple calculations we come to the conclusion of the lemma. □

Remark 5 We show now that the function $K_h'(jh, kh, x, y) + M_{h,3}'(jh, kh, x, y)$ in Lemma 8 is equal to $K_h(jh, kh, x, y) + \frac{h}{2}(L_\star^2 - 2L\widetilde{L} + \widetilde{L}^2)\lambda(x) + M_{h,3}''(jh, kh, x, y)$ where

$$M_{h,3}''(jh, kh, x, y) = -h^2 \sum_{|\mu|=2} \frac{m^\mu(jh, y)}{\mu!}(L - \widetilde{L})D^\mu \lambda(x) - 3\sum_{|\mu|=3}\int_0^1 (1-\delta)^2 d\delta \times \int q(jh, y, \theta)\frac{\widetilde{h}(\theta)^\mu}{\mu!}(L - \widetilde{L})D^\mu\lambda(x + \delta\widetilde{h}(\theta))d\theta. \tag{17}$$

Thus in Lemma 8 we can replace $K'_h(jh, kh, x, y) + M'_h(jh, kh, x, y)$ by $K_h(jh, kh, x, y) + M_h(jh, kh, x, y)$ where $K_h(jh, kh, x, y) = (L - \widetilde{L})\widetilde{p}_h(jh, kh, x, y)$, $M_h(jh, kh, x, y) = \frac{h}{2}(L_\star^2 - 2L\widetilde{L} + \widetilde{L}^2)\lambda(x) + M''_h$, $M''_h = M_{h,1}(jh, kh, x, y) + M_{h,2}(jh, kh, x, y) + M''_{h,3}(jh, kh, x, y)$ and

$$\max\left\{\left|M'_h(jh, kh, x, y)\right|, |M_h(jh, kh, x, y)|\right\} \le C\rho^{-1}\zeta_\rho(y - x),$$

$\rho^2 = kh - jh$. To show this we note that

$$\widetilde{p}_h(jh, kh, x, y) = \int q(jh, y, \theta)\lambda(x + \widetilde{h}(\theta))d\theta,$$

where $\widetilde{h}(\theta) = m(jh, y)h + h^{1/2}\theta$. From the Taylor expansion we get

$$\begin{aligned}\widetilde{p}_h(jh, kh, x, y) &= \lambda(x) + h\widetilde{L}\lambda(x) + h^2 \sum_{|\mu|=2} \frac{m^\mu(jh, y)}{\mu!} D^\mu \lambda(x) \\ &\quad + 3 \sum_{|\mu|=3} \int_0^1 (1 - \delta)^2 d\delta \int q(jh, y, \theta) \frac{\widetilde{h}(\theta)^\mu}{\mu!} D^\mu \lambda(x + \delta\widetilde{h}(\theta))d\theta\end{aligned}$$

and, hence,

$$\begin{aligned}K'_h(jh, kh, x, y) &= K_h(jh, kh, x, y) + (L - \widetilde{L})[\lambda(x) - \widetilde{p}_h(jh, kh, x, y)] \\ &= K_h(jh, kh, x, y) + h(\widetilde{L}^2 - L\widetilde{L})\lambda(x) + M''_{h,3}(jh, kh, x, y). \qquad (18)\end{aligned}$$

From

$$\begin{aligned}h(\widetilde{L}^2 - L\widetilde{L})\lambda(x) + M'_{h,3}(jh, kh, x, y) &= h(\widetilde{L}^2 - L\widetilde{L})\lambda(x) + \frac{h}{2}(L_\star^2 - \widetilde{L}^2)\lambda(x) \\ &= \frac{h}{2}(L_\star^2 - 2L\widetilde{L} + \widetilde{L}^2)\lambda(x)\end{aligned}$$

and from the definitions of the operators L, $\widetilde{L}$ and $L_\star$ and from the Lipschitz conditions on the coefficients $m(t, x)$ and $\sigma(t, x)$ we obtain that

$$\left|\frac{h}{2}(L_\star^2 - 2L\widetilde{L} + \widetilde{L}^2)\lambda(x)\right| \le Ch\rho^{-3}\zeta_\rho^S(y - x) \qquad (19)$$

with ζ^S defined as in Lemma 5 and with the constant S introduced in (A3). Analogously, we have

$$\left| h^2 \sum_{|\mu|=2} \frac{m^\mu(jh,y)}{\mu!}(L-\widetilde{L})D^\mu\lambda(x)\right| \le Ch^2\rho^{-3}\zeta_\rho^S(y-x), \tag{20}$$

$$\left| 3\sum_{|\mu|=3} \int_0^1 (1-\delta)^2 d\delta \int q(jh,y,\theta)\frac{\widetilde{h}(\theta)^\mu}{\mu!}(L-\widetilde{L})D^\mu\lambda(x+\delta\widetilde{h}(\theta))d\theta\right|$$
$$\le Ch^{3/2}\rho^{-4}\zeta_\rho^S(y-x). \tag{21}$$

Now (18)–(21) imply the assertion of this remark.

Lemma 9 *The following bound holds*:

$$\left|\sum_{r=0}^n \widetilde{p}_h \otimes_h (K_h+M_h+R_h)^{(r)}(0,T,x,y) - \sum_{r=0}^n \widetilde{p}_h \otimes_h (K_h+M_h)^{(r)}(0,T,x,y)\right|$$
$$\le C(\varepsilon)hn^{-1/2+\varepsilon}\zeta_{\sqrt{T}}^S(y-x), \tag{22}$$

where $\lim_{\varepsilon\downarrow 0} C(\varepsilon) = +\infty$.

The proof of Lemma 9 can be found in the extended version of this paper, see [22].

Lemma 10 *Let* $A(s,t,x,y)$, $B(s,t,x,y)$, $C(s,t,x,y)$ *be some functions with absolute value less than* $C(t-s)^{-1/2}\zeta_{\sqrt{t-s}}^\kappa(y-x)$ *for a constant* C *and an integer* $\kappa \ge S'd$. *Then*

$$\sum_{r=0}^\infty A\otimes_h (B+C)^{(r)}(ih,jh,x,y) - \sum_{r=0}^\infty A\otimes_h B^{(r)}(ih,jh,x,y)$$
$$= \sum_{r=1}^\infty \left[A\otimes_h B^\infty\right]\otimes_h \left[C\otimes_h B^\infty\right]^{(r)}(ih,jh,x,y),$$

where $B^\infty = \sum_{r=0}^\infty B^{(r)}$.

Proof of Lemma 10. Under the conditions of the lemma all series are absolutely convergent. The assertion of this lemma is a consequence of the linearity of the operation $\otimes_h$ and of the possibility to permutate the terms in absolutely convergent series. □

5 Proof of Theorem 1

We now come to the proof of Theorem 1. Main tools for the proof have been given in Sects. 3.1, 3.2, 4.1 and 4.2. From Lemmas 2 and 3 we get that

$$p(0, T, x, y) = \sum_{r=0}^{n} \tilde{p} \otimes H^{(r)}(0, T, x, y) + o(h^2 T)\phi_{C,\sqrt{T}}(y - x).$$

With Lemma 4 this gives

$$p(0, T, x, y) - p_h(0, T, x, y) = T_1 + \cdots + T_7 + o(h^2 T)\phi_{C,\sqrt{T}}(y - x), \quad (23)$$

where

$$T_1 = \sum_{r=0}^{n} \tilde{p} \otimes H^{(r)}(0, T, x, y) - \sum_{r=0}^{n} \tilde{p} \otimes_h H^{(r)}(0, T, x, y),$$

$$T_2 = \sum_{r=0}^{n} \tilde{p} \otimes_h H^{(r)}(0, T, x, y) - \sum_{r=0}^{n} \tilde{p} \otimes_h (H + M_h'' + \sqrt{h} N_1)^{(r)}(0, T, x, y),$$

$$T_3 = \sum_{r=0}^{n} \tilde{p} \otimes_h (H + M_h'' + \sqrt{h} N_1)^{(r)}(0, T, x, y) - \sum_{r=0}^{n} \tilde{p} \otimes_h (H + M_h + \sqrt{h} N_1)^{(r)}(0, T, x, y),$$

$$T_4 = \sum_{r=0}^{n} \tilde{p} \otimes_h (H + M_h + \sqrt{h} N_1)^{(r)}(0, T, x, y) - \sum_{r=0}^{n} \tilde{p} \otimes_h (K_h + M_h)^{(r)}(0, T, x, y),$$

$$T_5 = \sum_{r=0}^{n} \tilde{p} \otimes_h (K_h + M_h)^{(r)}(0, T, x, y) - \sum_{r=0}^{n} \tilde{p}_h \otimes_h (K_h + M_h)^{(r)}(0, T, x, y),$$

$$T_6 = \sum_{r=0}^{n} \tilde{p}_h \otimes_h (K_h + M_h)^{(r)}(0, T, x, y) - \sum_{r=0}^{n} \tilde{p}_h \otimes_h (K_h + M_h + R_h)^{(r)}(0, T, x, y),$$

$$T_7 = \sum_{r=0}^{n} \tilde{p}_h \otimes_h (K_h + M_h + R_h)^{(r)}(0, T, x, y) - \sum_{r=0}^{n} \tilde{p}_h \otimes_h H_h^{(r)}(0, T, x, y).$$

Here we put $N_1(s, t, x, y) = (L - \widetilde{L})\widetilde{\pi}_1(s, t, x, y)$.

We now discuss the asymptotic behavior of the terms $T_1, \ldots, T_7$.

Asymptotic treatment of the term T_1.

We start from the recurrence relations for $r = 1, 2, 3, \ldots$

$$\begin{aligned}&\left(\widetilde{p} \otimes H^{(r)}\right)(0, jh, x, y) - \left(\widetilde{p} \otimes_h H^{(r)}\right)(0, jh, x, y)\\&= \left[\left(\widetilde{p} \otimes H^{(r-1)}\right) \otimes H - \left(\widetilde{p} \otimes H^{(r-1)}\right) \otimes_h H\right](0, jh, x, y)\\&\quad + \left[\left(\widetilde{p} \otimes H^{(r-1)}\right) - \left(\widetilde{p} \otimes_h H^{(r-1)}\right)\right] \otimes_h H\,(0, jh, x, y). \qquad (24)\end{aligned}$$

By summing up the identities in (24) from $r = 1$ to ∞ and by using the linearity of the operations $\otimes$ and $\otimes_h$ we get

$$\begin{aligned}(p - p^d)(0, jh, x, y) &= (p \otimes H - p \otimes_h H)(0, jh, x, y)\\&\quad + (p - p^d) \otimes_h H\,(0, jh, x, y), \qquad (25)\end{aligned}$$

where we put

$$p^d(ih, i'h, x, y) = \sum_{r=0}^{\infty} (\widetilde{p} \otimes_h H^{(r)})(ih, i'h, x, y). \qquad (26)$$

By iterative application of (25) we obtain

$$\begin{aligned}(p - p^d)(0, jh, x, y) &= (p \otimes H - p \otimes_h H)(0, jh, x, y)\\&\quad + (p \otimes H - p \otimes_h H) \otimes_h \Xi\,(0, jh, x, y), \qquad (27)\end{aligned}$$

where $\Xi(ih, i'h, z, z') = H(ih, i'h, z, z') + H \otimes_h H(ih, i'h, z, z') + \cdots = \sum_{r=1}^{\infty} H^{(r)}(ih, i'h, z, z')$.

By application of a Taylor expansion we get

$$\begin{aligned}(p \otimes H - p \otimes_h H)(0, jh, x, z) &= \sum_{i=0}^{j-1} \int_{ih}^{(i+1)h} du \int_{R^d} [\lambda(u) - \lambda(ih)]\, dv\\&= \sum_{i=0}^{j-1} \int_{ih}^{(i+1)h} (u - ih) du \int_{R^d} \lambda'(ih) dv\\&\quad + \sum_{i=0}^{j-1} \int_{ih}^{(i+1)h} (u - ih)^2\\&\quad \times \int_0^1 (1 - \delta) \int_{R^d} \lambda''(s)\,|_{s=s_i}\, dv d\delta du, \qquad (28)\end{aligned}$$

where $\lambda(u) = p(0, u, x, v) H(u, jh, v, z)$, $s_i = s_i(u, i, \delta, h) = ih + \delta(u - ih)$.

Note that

$$\int_{R^d} \lambda'(ih)dv = \int_{R^d} \frac{\partial}{\partial s} p(0,s,x,v) \mid_{s=ih} H(ih,jh,v,z)dv$$
$$+ \int_{R^d} p(0,ih,x,v)\frac{\partial}{\partial s} H(s,jh,v,z) \mid_{s=ih} dv$$
$$= \int_{R^d} L^t p(0,ih,x,v) \ \ (L-\widetilde{L})\widetilde{p}(ih,jh,v,z)dv$$
$$- \int_{R^d} p(0,ih,x,v)[(L-\widetilde{L})\widetilde{L}\widetilde{p}(ih,jh,v,z)$$
$$- H_1(ih,jh,v,z)]dv$$
$$= \int_{R^d} p(0,ih,x,v)H_1(ih,jh,v,z)dv$$
$$+ \int_{R^d} p(0,ih,x,v)(L^2 - 2L\widetilde{L} + \widetilde{L}^2)\widetilde{p}(ih,jh,v,z)dv, \qquad (29)$$

where $H_1(s,t,v,z)$ is defined below in (35). We get from (29)

$$\sum_{i=0}^{j-1} \int_{ih}^{(i+1)h} (u-ih)du \int_{R^d} \lambda'(ih)dv = \frac{h}{2}(p \otimes_h H_1)(0,jh,x,z)$$
$$+ \frac{h}{2}(p \otimes_h A_0)(0,jh,x,z), \qquad (30)$$

where $A_0(s,jh,v,z) = (L^2 - 2L\widetilde{L} + \widetilde{L}^2)\widetilde{p}(s,jh,v,z)$. Direct calculations show that

$$A_0(s,jh,v,z) = \frac{1}{4}\sum_{p,q,r,l=1}^{d} (\sigma_{pq}(s,v) - \sigma_{pq}(s,z))(\sigma_{rl}(s,v)$$
$$- \sigma_{rl}(s,z))\frac{\partial^4 \widetilde{p}(s,jh,v,z)}{\partial v_p \partial v_q \partial v_r \partial v_l}$$
$$+ \sum_{p,q,r=1}^{d} (\sigma_{pq}(s,v) - \sigma_{pq}(s,z))(m_r(s,v) - m_r(s,z))\frac{\partial^3 \widetilde{p}(s,jh,v,z)}{\partial v_p \partial v_q \partial v_r}$$
$$+ \frac{1}{2}\sum_{p,q,r,l=1}^{d} \sigma_{pq}(s,v)\frac{\partial \sigma_{rl}(s,v)}{\partial v_p}\frac{\partial^3 \widetilde{p}(s,jh,v,z)}{\partial v_q \partial v_r \partial v_l} + (\leq 2), \qquad (31)$$

where we denote by (≤ 2) the sum of terms that contain derivatives of $\widetilde{p}(s, jh, v, z)$ of order less or equal 2. Note that for a constant $C < \infty$ and for any $0 < \varepsilon < \frac{1}{2}$

$$\left| \frac{h}{2}(p \otimes_h H_1)(0, jh, x, z) \right| \leq Ch\phi_{C,\sqrt{jh}}(z - x)\,,$$
$$\left| \frac{h}{2}(p \otimes_h A_0)(0, jh, x, z) \right| \leq C(\varepsilon)h^{1/2}j^{-(1/2-\varepsilon)}\phi_{C,\sqrt{jh}}(z - x)\,. \tag{32}$$

The first inequality (32) follows from (B1) and well known estimates for the diffusion density p and for the kernel H_1. The second inequality (32) follows from (B1), (31) and the following estimate

$$\begin{aligned}
&\frac{h}{2}\sum_{i=0}^{j-1} h \left| \int_{R^d} p(0, ih, x, v) \frac{\partial^3 \widetilde{p}(ih, jh, v, z)}{\partial v_q \partial v_r \partial v_l} dv \right| \\
&\leq \frac{h^3}{2} \left| \frac{\partial^3 \widetilde{p}(0, jh, x, z)}{\partial v_q \partial v_r \partial v_l} \right| + \frac{h}{2}\sum_{i=1}^{j-1} h \left| \int_{R^d} \frac{\partial p(0, ih, x, v)}{\partial v_q} \frac{\partial^2 \widetilde{p}(ih, jh, v, z)}{\partial v_r \partial v_l} dv \right| \\
&\leq \frac{h^3}{2} \left| \frac{\partial^3 \widetilde{p}(0, jh, x, z)}{\partial v_q \partial v_r \partial v_l} \right| + Ch^{1/2} j^{-(1/2-\varepsilon)} B\left(\frac{1}{2}, \varepsilon\right) \phi_{C,\sqrt{jh}}(z - x)\,.
\end{aligned} \tag{33}$$

Now we estimate the second summand in the right hand side of (28). Clearly,

$$\begin{aligned}
\lambda''(s) &= \frac{\partial^2}{\partial s^2} p(0, s, x, v) H(s, jh, v, z) + 2\frac{\partial}{\partial s} p(0, s, x, v) \frac{\partial}{\partial s} H(s, jh, v, z) \\
&\quad + p(0, s, x, v) \frac{\partial^2}{\partial s^2} H(s, jh, v, z).
\end{aligned}$$

Using forward and backward Kolmogorov equations we get from this equation after long but simple calculations

$$\begin{aligned}
&\sum_{i=0}^{j-1} \int_{ih}^{(i+1)h} (u - ih)^2 \int_0^1 (1 - \delta) \int_{R^d} \lambda''(s) \mid_{s=s_i} dv d\delta du \\
&= \sum_{i=0}^{j-1} \int_{ih}^{(i+1)h} (u - ih)^2 \int_0^1 (1 - \delta) \sum_{k=1}^{4} \int_{R^d} p(0, s, x, v) A_k(s, jh, v, z) \mid_{s=s_i} dv d\delta du,
\end{aligned} \tag{34}$$

where

$$
\begin{aligned}
A_1(s, jh, v, z) &= (L^3 - 3L^2\widetilde{L} + 3L\widetilde{L}^2 - \widetilde{L}^3)\widetilde{p}(s, jh, v, z),\\
A_2(s, jh, v, z) &= (L_1 H + 2LH_1)(s, jh, v, z),\\
A_3(s, jh, v, z) &= [(L - \widetilde{L})\widetilde{L}_1 + 2(L_1 - \widetilde{L}_1)\widetilde{L}]\widetilde{p}(s, jh, v, z),\\
A_4(s, jh, v, z) &= H_2(s, jh, v, z),\\
H_l(s, t, v, z) &= (L_l - \widetilde{L}_l)\widetilde{p}(s, t, v, z)\\
&= \frac{1}{2}\sum_{i,j=1}^{d}\left(\frac{\partial^l \sigma_{ij}(s, v)}{\partial s^l} - \frac{\partial^l \sigma_{ij}(s, z)}{\partial s^l}\right)\frac{\partial^2 \widetilde{p}(s, t, v, z)}{\partial v_i \partial v_j}\\
&\quad + \sum_{i=1}^{d}\left(\frac{\partial^l m_i(s, v)}{\partial s^l} - \frac{\partial^l m_i(s, z)}{\partial s^l}\right)\frac{\partial \widetilde{p}(s, t, v, z)}{\partial v_i}, \quad l = 1, 2.
\end{aligned}
\tag{35}
$$

Using integration by parts and the definition of A_2, A_3 and A_4 it is easy to get that for any $0 < \varepsilon < 1/2$ and for $k = 2, 3, 4$

$$
\begin{aligned}
&\left|\sum_{i=0}^{j-1}\int_{ih}^{(i+1)h}(u - ih)^2\int_0^1(1-\delta)\int_{R^d} p(0, s, x, v)A_k(s, jh, v, z)\,|_{s=s_i}\, dv d\delta du\right|\\
&\quad \le C(\varepsilon)h^{3/2-\varepsilon}\phi_{C,\sqrt{jh}}(z - x).
\end{aligned}
\tag{36}
$$

For $k = 1$ we shall prove the following estimate for any $0 < \varepsilon < \frac{1}{2}$

$$
\begin{aligned}
&\left|\sum_{i=0}^{j-1}\int_{ih}^{(i+1)h}(u - ih)^2\int_0^1(1-\delta)\int_{R^d} p(0, s, x, v)A_1(s, jh, v, z)\,|_{s=s_i}\, dv d\delta du\right|\\
&\quad \le C(\varepsilon)hj^{-(1/2-\varepsilon)}\phi_{C,\sqrt{jh}}(z - x).
\end{aligned}
\tag{37}
$$

Note that the function $A_1(s, jh, v, z)$ can be written as the following sum

$$
\begin{aligned}
A_1(s, jh, v, z) = \frac{1}{8}\sum_{i,j,p,q,l,r=1}^{d}&(\sigma_{ij}(s, v) - \sigma_{ij}(s, z))(\sigma_{pq}(s, v) - \sigma_{pq}(s, z))(\sigma_{lr}(s, v)\\
&- \sigma_{lr}(s, z))\frac{\partial^6 \widetilde{p}(s, jh, v, z)}{\partial v_i \partial v_j \partial v_p \partial v_q \partial v_l \partial v_r}\\
&+ \frac{3}{4}\sum_{i,j,p,q,l=1}^{d}(\sigma_{ij}(s, v) - \sigma_{ij}(s, z))(\sigma_{pq}(s, v)\\
&- \sigma_{pq}(s, z))(m_l(s, v) - m_l(s, z))\frac{\partial^5 \widetilde{p}(s, jh, v, z)}{\partial v_i \partial v_j \partial v_p \partial v_q \partial v_l}
\end{aligned}
$$

$$+\frac{3}{4}\sigma_{ij}(s,v)\frac{\partial\sigma_{pq}(s,v)}{\partial v_i}(\sigma_{lr}(s,v)$$
$$-\sigma_{lr}(s,z))\frac{\partial^5\widetilde{p}(s,jh,v,z)}{\partial v_j\partial v_p\partial v_q\partial v_l\partial v_r}+(\leq 4), \qquad (38)$$

where we denote by (≤ 4) the sum of terms that contain derivatives of $\widetilde{p}(s, jh, v, z)$ of order less or equal 4. By (B1) and (38) it is clear that (up to a constant) the estimate for the left hand side of (36) for $k = 1$ will be the same as for the following sum (with fixed p, q, r, l)

$$\left|\sum_{i=0}^{j-1}\int_{ih}^{(i+1)h}(u-ih)^2\int_0^1(1-\delta)\int_{R^d}p(0,s,x,v)\frac{\partial^4\widetilde{p}(s,jh,v,z)}{\partial v_p\partial v_q\partial v_l\partial v_r}\,|_{s=s_i}\,dvd\delta du\right|.$$

After integration by parts w.r.t. v_p and after the substitution $hw = (u - ih)$ in each integral we obtain

$$\left|\sum_{i=0}^{j-1}\int_{ih}^{(i+1)h}(u-ih)^2\int_0^1(1-\delta)\int_{R^d}p(0,s,x,v)\frac{\partial^4\widetilde{p}(s,jh,v,z)}{\partial v_p\partial v_q\partial v_l\partial v_r}\,|_{s=s_i}\,dvd\delta du\right|$$
$$=\left|\sum_{i=0}^{j-1}\int_{ih}^{(i+1)h}(u-ih)^2\int_0^1(1-\delta)\int_{R^d}\frac{\partial p(0,s,x,v)}{\partial v_p}\frac{\partial^3\widetilde{p}(s,jh,v,z)}{\partial v_q\partial v_l\partial v_r}\,|_{s=s_i}\,dvd\delta du\right|$$
$$\leq Ch^2\phi_{C,\sqrt{jh}}(z-x)\int_0^1 w^2\int_0^1(1-\delta)$$
$$\times\sum_{i=0}^{j-1}h\frac{1}{\sqrt{ih+\delta hw}}\frac{1}{[(j-i)h-\delta hw)]^{3/2}}d\delta dw$$
$$\leq Ch^{3/2-\varepsilon}\phi_{C,\sqrt{jh}}(z-x)\int_0^1 w^2\int_0^1(1-\delta)^{1/2-\varepsilon}$$
$$\times\sum_{i=0}^{j-1}h\frac{1}{\sqrt{ih+\delta hw}}\frac{1}{[(j-\delta w)h-ih)]^{1-\varepsilon}}d\delta dw$$
$$\leq Ch^{3/2-\varepsilon}\phi_{C,\sqrt{jh}}(z-x)\int_0^1 w^2dw\int_0^1(1-\delta)^{1/2-\varepsilon}d\delta\int_0^{(j-1)h}\frac{dt}{\sqrt{t}[(j-1)h-t]^{1-\varepsilon}}$$
$$\leq Chj^{-(1/2-\varepsilon)}B\left(\frac{1}{2},\varepsilon\right)\phi_{C,\sqrt{jh}}(z-x), \qquad (39)$$

where $B(p, q)$ is the Beta function and $\phi_{C,\rho}(z - x)$ is defined as in Lemma 3. As we mentioned above, (37) follows now from (39). By (B2), (28), (30), (31), (33), (36) and (37) we obtain for $0 < \varepsilon < \frac{1}{2}$, $j \geq 1$

$$|(p \otimes H - p \otimes_h H)(0, T, x, z)| \leq C(\varepsilon) h^{1/2} n^{-(1/2-\varepsilon)} \phi_{C,\sqrt{T}}(z - x). \quad (40)$$

We use now the following estimate for $\Xi(ih, i'h, z, z')$

$$\left|\Xi(ih, i'h, z, z')\right| \leq C \frac{1}{\sqrt{i'h - ih}} \phi_{C,\sqrt{i'h-ih}}\left(z' - z\right), \quad (41)$$

see (5.7) in [20]. From (B2), (28), (30), (39), (40) and (41) we obtain

$$\begin{aligned}(p - p^d)(0, T, x, y) &= \frac{h}{2}(p \otimes_h H_1)(0, T, x, y) + \frac{h}{2}(p \otimes_h A_0)(0, T, x, y) \\ &+ \frac{h}{2}(p \otimes_h H_1 \otimes_h \Xi)(0, T, x, y) + \frac{h}{2}(p \otimes_h A_0 \otimes_h \Xi)(0, T, x, y) + R(0, T, x, y),\end{aligned} \quad (42)$$

where for any $0 < \varepsilon < 1/2$ it holds that $|R(0, T, x, y)| \leq C(\varepsilon)(h^{3/2-\varepsilon} + hn^{-(1/2-\varepsilon)})$ $\phi_{C,\sqrt{T}}(y - x) = \phi_{C,\sqrt{T}}(y - x)\, o(h^{1+\delta})$. This representation implies that

$$\begin{aligned}T_1 = &\frac{h}{2}[p \otimes_h (L^2 - 2L\widetilde{L} + \widetilde{L}^2)\widetilde{p} \otimes_h \Xi_0)](0, T, x, y) \\ &+ \frac{h}{2}[p \otimes_h (L' - \widetilde{L}')\widetilde{p} \otimes_h \Xi_0](0, T, x, y) + R_T(0, T, x, y),\end{aligned} \quad (43)$$

where for any $0 < \varepsilon < 1/2$

$$|R_T(0, T, x, y)| \leq C(\varepsilon) h n^{-1/2+\varepsilon} \phi_{C,\sqrt{T}}(y - x) \leq C(\varepsilon) h^{1+\delta} \phi_{C,\sqrt{T}}(y - x) \quad (44)$$

for $\delta > 0$ small enough and where $\Xi_0(s, t, x, y) = \sum_{r=0}^{\infty} H^{(r)}(s, t, x, y)$. Here the summand $H^{(0)}(s, t, x, y)$ is introduced to shorten the notation. By definition we suppose that $g \otimes_h H^{(0)}(s, t, x, y) = g(s, t, x, y)$ for a function g. Note, that in the homogenous case $\sigma_{ij}(s, x) = \sigma_{ij}(x)$, $m_i(s, x) = m_i(x)$ and thus the second summand in (43) is equal to 0.

Asymptotic treatment of the term T_2. We will show that

$$\begin{aligned}\Bigg| T_2 - 3\sum_{r=0}^{\infty} \tilde{p} \otimes_h H^{(r)}(0, T, x, y) + \sum_{r=0}^{\infty} \tilde{p} \otimes_h (H + M_{h,1} + \sqrt{h} N_1)^{(r)}(0, T, x, y) \\ + \sum_{r=0}^{\infty} \tilde{p} \otimes_h (H + M_{h,2})^{(r)}(0, T, x, y) + \sum_{r=0}^{\infty} \tilde{p} \otimes_h (H + M''_{h,3})^{(r)}(0, T, x, y)\Bigg| \\ \leq C h n^{-\delta} \zeta_{\sqrt{T}}(y - x)\end{aligned} \quad (45)$$

with some positive $\delta > 0$. The proof of this estimate can be found in the extended version of this paper, see [22].
Asymptotic treatment of the term T_3. We will show that

$$\left| T_3 - \left[\sum_{r=0}^{\infty} \tilde{p} \otimes_h (H+A)^{(r)}(0,T,x,y) - \sum_{r=0}^{\infty} \tilde{p} \otimes_h H^{(r)}(0,T,x,y) \right] \right| \\ \leq Chn^{-\delta} \zeta_{\sqrt{T}}(y-x), \tag{46}$$

where $A = M_h'' - M_h = -\frac{h}{2}(L_\star^2 - 2L\widetilde{L} + \widetilde{L}^2)\lambda(x)$. Write

$$\begin{aligned} C_r &= \tilde{p} \otimes_h (H + M_h'' + \sqrt{h}N_1)^{(r)}(0,T,x,y) \\ &\quad - \tilde{p} \otimes_h (H + M_h + \sqrt{h}N_1)^{(r)}(0,T,x,y) \\ &\quad - [\tilde{p} \otimes_h (H+A)^{(r)} - \tilde{p} \otimes_h H^{(r)}](0,T,x,y). \end{aligned}$$

We use the following recurrence relation

$$\begin{aligned} C_r &= C_{r-1} \otimes_h H + \Big[\tilde{p} \otimes_h (H + M_h'' + \sqrt{h}N_1)^{(r-1)} \\ &\quad - \tilde{p} \otimes_h (H + M_h + \sqrt{h}N_1)^{(r-1)} \Big] \otimes_h (M_h'' + \sqrt{h}N_1) \\ &\quad + \Big[\tilde{p} \otimes_h (H + M_h + \sqrt{h}N_1)^{(r-1)} - \tilde{p} \otimes_h (H+A)^{(r-1)} \Big] \otimes_h A \\ &= I + II + III. \end{aligned} \tag{47}$$

With the notation

$$D_{r-1} = \tilde{p} \otimes_h (H + M_h + \sqrt{h}N_1)^{(r-1)} - \tilde{p} \otimes_h (H+A)^{(r-1)}$$

we get

$$D_{r-1} = D_{r-2} \otimes_h (H + M_h + \sqrt{h}N_1) + \tilde{p}_h \otimes_h (H+A)^{(r-2)} \otimes_h (M_h - A + \sqrt{h}N_1).$$

Iterative application gives

$$\begin{aligned} III &= D_{r-1} \otimes_h A \\ &= \sum_{l=0}^{r-2} \widetilde{p}_{4,l} \otimes_h (M_h - A + \sqrt{h}N_1) \otimes_h (H + M_h + \sqrt{h}N_1)^{(r-l-2)} \otimes_h A(0,T,x,y), \end{aligned}$$

where $\widetilde{p}_{4,l} = \widetilde{p} \otimes_h (H+A)^{(l)}$. This sum can be estimated in exactly the same way as the sum in (86) in [22]. This gives for $r = 2, 3, \ldots$

$$|III| \leq C(\varepsilon) h^{3/2-2\varepsilon} \frac{C^r}{\Gamma(\frac{r-1}{2})} T^{3\varepsilon + \frac{r-4}{2}} \zeta_{\sqrt{T}}(v-x). \tag{48}$$

To estimate II we write

$$E_{r-1} = \tilde{p} \otimes_h (H + M_h'' + \sqrt{h}N_1)^{(r-1)} - \tilde{p} \otimes_h (H + M_h + \sqrt{h}N_1)^{(r-1)}.$$

For $r = 2$ we have $E_1 = \tilde{p} \otimes_h A$ and we get

$$|E_1| \le Ch^{1-\varepsilon}(kh)^{\varepsilon-1/2} B\left(\frac{1}{2}, \varepsilon\right) \zeta^S_{\sqrt{kh}}(y-x).$$

For $r \ge 3$ we use the recurrence relation

$$\begin{aligned} E_{r-1} &= E_{r-2} \otimes_h (H + M_h'' + \sqrt{h}N_1) + [\tilde{p} \otimes_h (H + M_h + \sqrt{h}N_1)^{(r-2)}] \otimes_h A \\ &= I' + II'. \end{aligned}$$

The terms I' and II' have a similar structure as the corresponding terms in (81) in [22] and they can be estimated similarly. This gives the following estimates for $r = 2, 3, \ldots$

$$\begin{aligned} |E_{r-1}| &\le C^r h^{1-\varepsilon} B\left(\frac{1}{2}, \varepsilon\right) \times \cdots \times B\left(\frac{1}{2}, \varepsilon + \frac{r-2}{2}\right) (kh)^{\varepsilon + \frac{r-3}{2}} \zeta^S_{\sqrt{kh}}(y-x), \\ |II| &= \left| E_{r-1} \otimes_h (M_h'' + \sqrt{h}N_1)(0, T, x, y) \right| \\ &\le C(\varepsilon) h^{3/2-2\varepsilon} \frac{C^r}{\Gamma(\frac{r-1}{2})} T^{3\varepsilon + \frac{r-4}{2}} \zeta_{\sqrt{T}}(v-x). \end{aligned}$$

The claim (46) follows from (47), (48) and the last two inequalities.
Asymptotic treatment of the term T_4. It suffices to show that

$$T_4 = \sum_{r=1}^{\infty} \tilde{p} \otimes_h H^{(r)}(0, T, x, y) - \sum_{r=1}^{\infty} \tilde{p} \otimes_h [H + hN_2]^{(r)}(0, T, x, y) + R_h^\star(x, y)$$

with $N_2(s, t, x, y) = (L - \widetilde{L})\widetilde{\pi}_2(s, t, x, y)$, $\left|R_h^\star(x, y)\right| \le Chn^{-\delta} \zeta^S_{\sqrt{T}}(y-x)$ for $\delta > 0$ small enough, and a constant C depending on δ. The proof of this claim is elementary but rather long. For this reason we omit it. Details can be found in [22].
Asymptotic treatment of the term T_5. We will show that,

$$\begin{aligned} T_5 = &-\sqrt{h} \sum_{r=0}^{\infty} \widetilde{\pi}_1 \otimes_h (H + M_{h,1} + \sqrt{h}N_1)^{(r)}(0, T, x, y) \\ &- h \sum_{r=0}^{\infty} \widetilde{\pi}_2 \otimes_h H^{(r)}(0, T, x, y) + R_h(x, y), \end{aligned} \tag{49}$$

where $|R_h(x, y)| \le Chn^{-\gamma} \zeta^{S-2}_{\sqrt{T}}(y-x)$ for some $\gamma > 0$. Note that with $S_h(s, t, x, y) = \sum_{r=1}^{n} (K_h + M_h)^{(r)}(s, t, x, y)$ the term T_5 can be rewritten as

$$T_5 = (\widetilde{p} - \widetilde{p}_h)(0, T, x, y) + (\widetilde{p} - \widetilde{p}_h) \otimes_h S_h(0, T, x, y).$$

We start by showing that for $\varkappa < \delta < \frac{1-\varkappa}{4}$ uniformly for $x, y \in R$

$$\left| h \sum_{1 \le j \le n^\delta} \int (\widetilde{p}_h - \widetilde{p})(0, jh, x, u) S_h(jh, T, u, y) du \right| \\ \le O(hn^{-1/2(1-\varkappa-4\delta)}) \zeta^{S-2}_{\sqrt{T}}(y-x) \tag{50}$$

for δ small enough. For the proof of (50) we will show that uniformly for $1 \le j \le n^\delta$ and for $x, y \in R^d$

$$\int \widetilde{p}_h(0, jh, x, u) S_h(jh, T, u, y) du = S_h(jh, T, x, y) \\ + O[h^{1/2} T^{-1/2} n^{-1/2+\delta} + h^{1/2} T^{-1} + n^{\delta/2} h^{1/2}] \zeta^{S-2}_{\sqrt{T}}(y-x), \tag{51}$$

$$\int \widetilde{p}(0, jh, x, u) S_h(jh, T, u, y) du = S_h(jh, T, x, y) \\ + O[h^{1/2} T^{-1/2} n^{-1/2+\delta} + h^{1/2} T^{-1} + n^{\delta/2} h^{1/2}] \zeta^{S-2}_{\sqrt{T}}(y-x). \tag{52}$$

Claim (50) immediately follows from (51) to (52). For the proof we will make use of the fact that for all $1 \le j \le n^\delta$ and for all $x, y \in R^d$ and $|\nu| = 1$

$$\left| D^\nu_x S_h(jh, T, x, y) \right| \le C(T - jh)^{-1} \zeta^{S-2}_{\sqrt{T-jh}}(y-x). \tag{53}$$

Claim (53) can be shown with the same arguments as in the proof of (5.7) in [20]. Note that the function Ξ in that paper has a similar structure as S_h. For $1 \le j \le n^\delta$ the bound (53) immediately implies for a constant C'

$$\left| D^\nu_x S_h(jh, T, x, y) \right| \le C' T^{-1} \zeta^{S-2}_{\sqrt{T}}(y-x). \tag{54}$$

We have $\widetilde{p}_h(0, jh, x, u) = h^{-d/2} q^{(j)}[0, u, h^{-1/2}(u - x - h \sum_{i=0}^{j-1} m(ih, u))]$. Denote the determinant of the Jacobian matrix of $u - h \sum_{i=0}^{j-1} m(ih, u)$ by Δ_h. From the condition (A3) and (54) we get that for $1 \le j \le n^\delta$

$$\int \widetilde{p}_h(0, jh, x, u) S_h(jh, T, u, y) du \\ = \int h^{-d/2} q^{(j)} \left[0, u, h^{-1/2} \left(u - x - h \sum_{i=0}^{j-1} m(ih, u) \right) \right] S_h(jh, T, u, y) du \\ = \int q^{(j)} \left(0, x + h^{1/2} w + h \sum_{i=0}^{j-1} m(ih, u(w)), w \right) \left| \Delta_h^{-1} \right| S_h \left(jh, T, x + h^{1/2} w \right. \\ \left. + h \sum_{i=0}^{j-1} m(ih, u(w)), y \right) dw$$

$$
\begin{aligned}
&= \int \Big[q^{(j)}(0,x,w) + O(j^{-d/2}h^{1/2})(\|w\|+1)\psi(j^{-1/2}w) \Big] \\
&\quad \times [1+O(jh)][S_h(jh,T,x,y) \\
&\quad + O(h^{1/2}T^{-1})\zeta^{S-2}_{\sqrt{T}}(y-x)(1+h^{(S-2)/2}\,\|w\|^{S-2})(\|w\|+1)]dw \\
&= S_h(jh,T,x,y) + O[h^{1/2}T^{-1/2}n^{-1/2+\delta} + h^{1/2}T^{-1} + h^{1/2}n^{\delta/2}]\zeta^{S-2}_{\sqrt{T}}(y-x)
\end{aligned}
$$

with $u = u(w)$ in $\sum_{i=0}^{j-1} m(ih,u)$ defined by the Inverse Function Theorem from the substitution $w = h^{-1/2}(u - x - h\sum_{i=0}^{j-1} m(ih,u))$. This proves (51). Claim (52) follows by similar arguments. From (50) we get that for $\delta < \frac{1-\varkappa}{4}$ (with $\varkappa$ defined as in (B2))

$$
\begin{aligned}
T_5 &= (\widetilde{p} - \widetilde{p}_h)(0,T,x,y) \\
&\quad + h \sum_{n^\delta < j < n} \int (\widetilde{p} - \widetilde{p}_h)(0,jh,x,u)S_h(jT,u,y)du + R_h(x,y)
\end{aligned}
$$

with $|R_h(x,y)| \le O(hn^{-1/2(1-\varkappa-4\delta)})\zeta^{S-2}_{\sqrt{T}}(y-x)$. We now make use of the expansion of $\widetilde{p}_h - \widetilde{p}$ given in Lemma 5. We have with $\rho = (jh)^{1/2} \ge h^{1/2}n^{\delta/2}$

$$
\begin{aligned}
&\left| h \sum_{j=n^\delta}^{n} h^{3/2}\rho^{-3} \int \zeta^S_\rho(u-x)S_h(jh,T,u,y)du \right| \\
&\quad \le Ch^2T^{-\delta'}n^{-\delta''} \sum_{j=n^\delta}^{n} \rho^{-2+2\delta'} \int \Big| \zeta^S_\rho(u-x)S_h(jh,T,u,y) \Big| du, \qquad (55)
\end{aligned}
$$

where $\delta' < \frac{1}{2}\delta(1-\delta)^{-1}$, $2\,\delta'' = \delta + 2\delta\delta' - 2\delta'$. Now we get that

$$
h \sum_{j=n^\delta}^{n} \rho^{-2+2\delta'} \int \Big| \zeta^S_\rho(u-x)S_h(jh,T,u,y) \Big| du \le CB(\delta',1/2)T^{\delta'-1/2}\zeta^{S-2}_{\sqrt{T}}(y-x) \qquad (56)
$$

for a constant C. This shows that for $\delta' > 0$ small enough

$$
\begin{aligned}
T_5 &= -\Big[\sqrt{h}\widetilde{\pi}_1 + h\widetilde{\pi}_2\Big](0,T,x,y) \\
&\quad - h \sum_{n^\delta < j < n} \int \Big[\sqrt{h}\widetilde{\pi}_1 + h\widetilde{\pi}_2\Big](0,jh,x,u)S_h(jh,T,u,y)du + R'_h(x,y)
\end{aligned}
$$

with $\left|R_h'(x,y)\right| \le O(hn^{-(\delta''-\varkappa/2)})\zeta_{\sqrt{T}}^{S-2}(y-x)$ with a constant in $O(\cdot)$ depending on δ'. It follows from (50), (55) and (56) that

$$T_5 = -\sum_{r=0}^{\infty}\left[\sqrt{h}\widetilde{\pi}_1 + h\widetilde{\pi}_2\right] \otimes_h (K_h + M_h)^{(r)}(0,T,x,y) + R_h''(x,y), \qquad (57)$$

where $\left|R_h''(x,y)\right| \le O(hn^{-(\delta''-\varkappa/2)})\zeta_{\sqrt{T}}^{S-2}(y-x)$. Now we apply Lemma 10 with $A = \sqrt{h}\widetilde{\pi}_1$, $B = H + M_{h,1} + \sqrt{h}N_1$, $C = (K_h - H - \sqrt{h}N_1) + (M_h - M_{h,1})$ to

$$\begin{aligned} &-\sum_{r=0}^{\infty}\sqrt{h}\widetilde{\pi}_1 \otimes_h (K_h + M_h)^{(r)}(0,T,x,y) \\ &\quad+\sum_{r=0}^{\infty}\sqrt{h}\widetilde{\pi}_1 \otimes_h (H + M_{h,1} + \sqrt{h}N_1)^{(r)}(0,T,x,y) \end{aligned} \qquad (58)$$

and with $A = h\widetilde{\pi}_2$, $B = H$, $C = (K_h - H) + M_h$ to

$$-\sum_{r=0}^{\infty} h\widetilde{\pi}_2 \otimes_h (K_h + M_h)^{(r)}(0,T,x,y) + \sum_{r=0}^{\infty} h\widetilde{\pi}_2 \otimes_h H^{(r)}(0,T,x,y). \qquad (59)$$

The estimate (49) follows from (56), (58), (59), Lemmas 10 and 5.
Asymptotic treatment of the term T_6. By application of Lemma 9 we get that

$$|T_6| \le C(\varepsilon)hn^{-1/2+\varepsilon}\zeta_{\sqrt{T}}^{S}(y-x).$$

Asymptotic treatment of the term T_7. From the recurrence relation for $r = 2,3,\ldots$

$$\begin{aligned} &\widetilde{p}_h \otimes_h (K_h + M_h + R_h)^{(r)}(0,T,x,y) - \widetilde{p}_h \otimes_h {H_h}^{(r)}(0,T,x,y) \\ &\quad= \left[\widetilde{p}_h \otimes_h (K_h + M_h + R_h)^{(r-1)} - \widetilde{p}_h \otimes_h {H_h}^{(r-1)}\right] \otimes_h H_h(0,T,x,y) \\ &\qquad+ \left[\widetilde{p}_h \otimes_h (K_h + M_h + R_h)^{(r-1)} \otimes_h (K_h + M_h + R_h - H_h)\right](0,T,x,y) \end{aligned}$$

and from Lemma 8 with $r = 1$ we get by similar arguments as in the proof of Lemma 9 that

$$|T_7| \le Ch^{3/2}T^{-1/2}\zeta_{\sqrt{T}}^{S}(y-x) = Chn^{-1/2}\zeta_{\sqrt{T}}^{S}(y-x).$$

Plugging in the asymptotic expansions of $T_1, \ldots, T_7$. We now plug the asymptotic expansions of $T_1, \ldots, T_7$ into (23). Using Lemma 10, Theorem 2.1 in [20] we get

$$\begin{aligned}
& p_h(0, T, x, y) - p(0, T, x, y) \\
&= \sqrt{h}\left[\widetilde{\pi}_1 + p^d \otimes_h \Re_1\right] \otimes_h \Xi(0, T, x, y) \\
&\quad + h\left\{\left[\widetilde{\pi}_2 + \widetilde{\pi}_1 \otimes_h \Xi \otimes_h \Re_1 + p^d \otimes_h \Re_2 + p^d \otimes_h \Re_3\right] \otimes_h \Xi(0, T, x, y)\right. \\
&\quad + p^d \otimes_h (\Re_1 \otimes_h \Xi)^{(2)}(0, T, x, y) \\
&\quad \left. + \frac{1}{2} p \otimes_h (L_\star^2 - L^2) p^d(0, T, x, y) - \frac{1}{2} p \otimes_h (L' - \widetilde{L}') p^d(0, T, x, y)\right\} \\
&\quad + O(h^{1+\delta} \zeta_{\sqrt{T}}(y - x)),
\end{aligned} \tag{60}$$

where

$$\begin{aligned}
p^d(ih, i'h, x, y) &= \sum_{r=0}^{\infty} \widetilde{p} \otimes_h H^{(r)}(ih, i'h, x, y), \\
\Re_1(s, t, x, y) &= N_1(s, t, x, y) + M_1(s, t, x, y) - \widetilde{M}_1(s, t, x, y), \\
\Re_2(s, t, x, y) &= N_2(s, t, x, y) + \Pi_1(s, t, x, y) - \widetilde{\Pi}_1(s, t, x, y), \\
\Re_3(s, t, x, y) &= \sum_{|\nu|=4} \frac{(\chi_\nu(s, x) - \chi_\nu(s, y))}{\nu!} D_x^\nu \widetilde{p}(s, t, x, y), \\
M_1(s, t, x, y) &= \sum_{|\nu|=3} \frac{\chi_\nu(s, x)}{\nu!} D_x^\nu \widetilde{p}(s, t, x, y), \\
\widetilde{M}_1(s, t, x, y) &= \sum_{|\nu|=3} \frac{\chi_\nu(s, y)}{\nu!} D_x^\nu \widetilde{p}(s, t, x, y), \\
\Pi_1(s, t, x, y) &= \sum_{|\nu|=3} \frac{\chi_\nu(s, x)}{\nu!} D_x^\nu \widetilde{\pi}_1(s, t, x, y), \\
\widetilde{\Pi}_1(s, t, x, y) &= \sum_{|\nu|=3} \frac{\chi_\nu(s, y)}{\nu!} D_x^\nu \widetilde{\pi}_1(s, t, x, y).
\end{aligned}$$

Note that for the homogenous case and $T = [0, 1]$ (60) coincides with formula (53) on page 623 in [21].

Asymptotic replacement of p^d by p. It follows from (26), (39) and (40) that

$$\left|(p^d - p)(ih, jh, x, z)\right| \leq C(\varepsilon) h^{1-\varepsilon} (jh - ih)^{\varepsilon - 1/2} \phi_{C, \sqrt{(j-i)h}}(z - x) \tag{61}$$

for any $0 < \varepsilon < 1/2$. Using (61) and making an integration by parts we can replace p^d by p in (60). For example the operator $L_\star^2 - L^2$ is an operator of order three.

Applying integration by parts we get for $|\nu| = 3$

$$\left|\sum_{i=1}^{n-1} h \int D_z^\nu p(0, ih, x, z)(p^d - p)(ih, T, z, y)dz\right|$$
$$\le C(\varepsilon)h^{1-\varepsilon} \sum_{i=1}^{n-1} h \frac{1}{(ih)^{3/2}} \frac{1}{(T-ih)^{1/2-\varepsilon}} \phi_{C,\sqrt{T}}(y-x)$$
$$\le C(\varepsilon)h^{1/2-2\varepsilon} T^{2\varepsilon-1/2} B(\varepsilon, \varepsilon + \frac{1}{2})\phi_{C,\sqrt{T}}(y-x).$$

By (B2) we have $0 < \varkappa < 1 - 4\varepsilon$. This implies

$$\left|\frac{h}{2} p \otimes_h (L_\star^2 - L^2)(p^d - p)(0, T, x, y)\right| \le C(\varepsilon)hT^{1/2} n^{-(1/2-2\varepsilon-\varkappa/2)} \phi_{C,\sqrt{T}}(y-x)$$
$$\le C(\varepsilon)h^{1+\delta}\phi_{C,\sqrt{T}}(y-x)$$

for some $0 < \delta < 1/2$. The other terms in (60) containing p^d can be estimated analogously. Thus we get the following representation

$$\begin{aligned}
&p_h(0, T, x, y) - p(0, T, x, y) \\
&\quad = \sqrt{h}\,[\widetilde{\pi}_1 + p \otimes_h \Re_1] \otimes_h \Xi(0, T, x, y) \\
&\qquad + h \Big\{ [\widetilde{\pi}_2 + \widetilde{\pi}_1 \otimes_h \Xi \otimes_h \Re_1 + p \otimes_h \Re_2 + p \otimes_h \Re_3] \otimes_h \Xi(0, T, x, y) \\
&\qquad + p \otimes_h (\Re_1 \otimes_h \Xi)^{(2)}\,(0, T, x, y) \\
&\qquad + \frac{1}{2} p \otimes_h (L_\star^2 - L^2)p(0, T, x, y) - \frac{1}{2} p \otimes_h (L' - \widetilde{L}')p(0, T, x, y) \Big\} \\
&\qquad + O(h^{1+\delta}\zeta_{\sqrt{T}}(y-x)).
\end{aligned}$$

In the further analysis we make use of the following binary operation $\otimes_h'$. This operator generalizes the binary operation $\otimes$ introduced in [21]. For $s \in [0, t-h]$ and $t \in \{h, 2h, \dots, T\}$ the operation $\otimes_h'$ is defined as follows

$$f \otimes_h' g(s, t, x, y) = \sum_{s \le jh \le t-h} h \int f(s, jh, x, z) g(jh, t, z, y)dz.$$

Note that for $s \in \{0, h, 2h, \dots, T\}$ the two operations $\otimes_h'$ and $\otimes_h$ coincide.
Asymptotic replacement of $(p \otimes_h \Re_i) \otimes_h \Xi(0, T, x, y)$ *by* $p \otimes (\Re_i \otimes_h' \Xi)(0, T, x, y) = (p \otimes \Re_i) \otimes_h \Xi(0, T, x, y), i = 1, 2, 3, [p \otimes_h (\Re_1 \otimes_h \Xi)] \otimes_h (\Re_1 \otimes_h \Xi)\,(0, T, x, y)$ *by* $p \otimes [(\Re_i \otimes_h' \Xi) \otimes_h' (\Re_i \otimes_h' \Xi)](0, T, x, y)$, $p \otimes_h (L_\star^2 - L^2)p(0, T, x, y)$ *by* $p \otimes (L_\star^2 - L^2)p(0, T, x, y)$ *and* $p \otimes_h (L' - \widetilde{L}')p(0, T, x, y)$ *by* $p \otimes (L' - \widetilde{L}')p(0, T, x, y)$.

These replacements follow from the definitions of $\Re_i, i = 1, 2, 3$, and can be proved by the same method as in the treatment of T_1. For more details see [22]. With these replacements we come to the following representation

$$\begin{aligned}
&p_h(0,T,x,y)-p(0,T,x,y)\\
&\quad=\sqrt{h}\left[\widetilde{\pi}_1\otimes_h'\,\Xi(0,T,x,y)+p\otimes(\Re_1\otimes_h'\,\Xi)(0,T,x,y)\right]\\
&\qquad+h\left[\widetilde{\pi}_2\otimes_h'\,\Xi(0,T,x,y)+p\otimes(\Re_2\otimes_h'\,\Xi)(0,T,x,y)\right.\\
&\qquad+\left.\ p\otimes_h(\Re_3\otimes_h'\,\Xi)(0,T,x,y)\right]\\
&\qquad+h\left[\widetilde{\pi}_1\otimes_h'\,\Xi+p\otimes\left(\Re_1\otimes_h'\,\Xi\right)\right]\otimes_h'(\Re_1\otimes_h'\,\Xi)(0,T,x,y)\\
&\qquad+\frac{h}{2}p\otimes(L_\star^2-L^2)p(0,T,x,y)-\frac{h}{2}p\otimes(L'-\widetilde{L}')p(0,T,x,y)\\
&\qquad+O(h^{1+\delta}\zeta_{\sqrt{T}}(y-x)).
\end{aligned}\tag{62}$$

We now further simplify our expansion of p_h-p. We start by showing the following expansion

$$\begin{aligned}
&p_h(0,T,x,y)-p(0,T,x,y)\\
&\quad=\sqrt{h}(p\otimes\mathcal{F}_1[p_\Delta])(0,T,x,y)+h\,(p\otimes\mathcal{F}_2[p_\Delta])\,(0,T,x,y)\\
&\qquad+h\,(p\otimes\mathcal{F}_1[p\otimes\mathcal{F}_1[p_\Delta]])\,(0,T,x,y)\\
&\qquad+\frac{h}{2}p\otimes(L_\star^2-L^2)p(0,T,x,y)-\frac{h}{2}p\otimes(L'-\widetilde{L}')p(0,T,x,y)\\
&\qquad+O(h^{1+\delta}\zeta_{\sqrt{T}}(y-x)),
\end{aligned}\tag{63}$$

where for $s\in[0,t-h]$, $t\in\{h,2h,\ldots,T\}$

$$\begin{aligned}
p_\Delta(s,t,z,y)&=(\widetilde{p}\otimes_h'\,\Xi)(s,t,z,y)\\
&=\widetilde{p}(s,t,z,y)+\sum_{s\le jh\le t-h}h\int\widetilde{p}(s,jh,z,v)\Xi'(jh,t,v,y)dv.
\end{aligned}$$

Here $\Xi'=H+H\otimes_h' H+H\otimes_h' H\otimes_h' H+\cdots$. We now treat the term $p\otimes\widetilde{L}\widetilde{\pi}_1(s,t,x,y)$.

$$\begin{aligned}
p\otimes\widetilde{L}\widetilde{\pi}_1(s,t,x,y)&=\int_s^t d\tau\int p(s,\tau,x,v)(t-\tau)\sum_{|\nu|=3}\frac{\overline{\chi}_\nu(\tau,t,y)}{\nu!}D_v^\nu(\widetilde{L}_v\widetilde{p}(\tau,t,v,y))dv\\
&=-\sum_{|\nu|=3}\frac{1}{\nu!}\int dv\left[\int_s^t p(s,\tau,x,v)\left(\int_\tau^t\chi_\nu(u,y)du\right)\right.\\
&\qquad\left.\times\frac{\partial}{\partial\tau}D_v^\nu\widetilde{p}(\tau,t,v,y)d\tau\right]
\end{aligned}$$

$$= - \sum_{|\nu|=3} \frac{1}{\nu!} \int dv \int_s^{\frac{s+t}{2}} \cdots - \sum_{|\nu|=3} \frac{1}{\nu!} \int dv \int_{\frac{s+t}{2}}^{t} \cdots$$
$$= I + II. \tag{64}$$

By integrating by parts w.r.t. the time variable we obtain for I

$$\begin{aligned} I = &- \sum_{|\nu|=3} \frac{1}{\nu!} \int dv \Bigg[p(s,\tau,x,v) \left(\int_\tau^t \chi_\nu(u,y)du \right) D_v^\nu \widetilde{p}(\tau,t,v,y) \,\Big|_{\tau=s}^{\tau=(s+t)/2} \\ &- \int_s^{\frac{s+t}{2}} D_v^\nu \widetilde{p}(\tau,t,v,y) \left(\frac{\partial p(s,\tau,x,v)}{\partial \tau} \int_\tau^t \chi_\nu(u,y)du \right. \\ &- \left. p(s,\tau,x,v)\chi_\nu(\tau,y) \right) d\tau \Bigg] \\ = &- \sum_{|\nu|=3} \frac{1}{\nu!} \int dv \Bigg[p\left(s, \frac{s+t}{2}, x, v\right) \left(\int_{\frac{s+t}{2}}^t \chi_\nu(u,y)du \right) D_v^\nu \widetilde{p}(\frac{s+t}{2},t,v,y) \\ &+ \sum_{|\nu|=3} \frac{1}{\nu!} \left(\int_s^t \chi_\nu(u,y)du \right) D_v^\nu \widetilde{p}(s,t,x,y) \Bigg] \\ &+ \sum_{|\nu|=3} \frac{1}{\nu!} \int_s^{\frac{s+t}{2}} d\tau \left(\int_\tau^t \chi_\nu(u,y)du \right) \int L^t p(s,\tau,x,v) D_v^\nu \widetilde{p}(\tau,t,v,y)dv \\ &- \sum_{|\nu|=3} \frac{1}{\nu!} \int_s^{\frac{s+t}{2}} \chi_\nu(\tau,y)d\tau \int p(s,\tau,x,v) D_v^\nu \widetilde{p}(\tau,t,v,y)dv. \end{aligned} \tag{65}$$

For the second term we get

$$\begin{aligned} II = &\sum_{|\nu|=3} \frac{1}{\nu!} \int p\left(s, \frac{s+t}{2}, x, v\right) \left(\int_{\frac{s+t}{2}}^t \chi_\nu(u,y)du \right) D_v^\nu \widetilde{p}(\frac{s+t}{2},t,v,y)dv \\ &+ \sum_{|\nu|=3} \frac{1}{\nu!} \int_{\frac{s+t}{2}}^t d\tau \left(\int_\tau^t \chi_\nu(u,y)du \right) \int L^t p(s,\tau,x,v) D_v^\nu \widetilde{p}(\tau,t,v,y)dv \end{aligned}$$

$$-\sum_{|\nu|=3}\frac{1}{\nu!}\int\limits_{\frac{s+t}{2}}^{t}\chi_\nu(\tau,y)d\tau\int p(s,\tau,x,v)D_v^\nu\widetilde{p}(\tau,t,v,y)dv. \tag{66}$$

From (64) to (66) we have

$$p\otimes\widetilde{L}\widetilde{\pi}_1(s,t,x,y)=\widetilde{\pi}_1(s,t,x,y)+p\otimes L\widetilde{\pi}_1(s,t,x,y)-p\otimes\widetilde{M}_1(s,t,x,y).$$

This shows that

$$\begin{aligned}\widetilde{\pi}_1(s,t,x,y)+p\otimes\Re_1(s,t,x,y)&=\widetilde{\pi}_1(s,t,x,y)\\&\quad+p\otimes L\widetilde{\pi}_1(s,t,x,y)-p\otimes\widetilde{L}\widetilde{\pi}_1(s,t,x,y)\\&\quad+p\otimes M_1(s,t,x,y)-p\otimes\widetilde{M}_1(s,t,x,y)\\&=p\otimes M_1(s,t,x,y).\end{aligned} \tag{67}$$

It follows from (67) and the definitions of the operations $\otimes$ and $\otimes_h'$ that

$$\begin{aligned}&\sqrt{h}\left[\widetilde{\pi}_1\otimes_h'\Xi(s,t,x,y)+(p\otimes\Re_1)\otimes_h'\Xi(s,t,x,y)\right]\\&\quad=\sqrt{h}(\widetilde{\pi}_1+p\otimes\Re_1)\otimes_h'\Xi(s,t,x,y)\\&\quad=\sqrt{h}(p\otimes M_1)\otimes_h'\Xi(s,t,x,y)\\&\quad=\sqrt{h}\sum_{0\le jh\le t-h}h\int(p\otimes M_1)(s,jh,x,z)\Xi(jh,t,z,y)dz\\&\quad=\sqrt{h}\sum_{0\le jh\le t-h}h\int\left[\int\limits_s^{jh}du\int p(s,u,x,v)M_1(u,jh,v,z)dv\right]\Xi(jh,t,z,y)dz\\&\quad=\sqrt{h}\sum_{0\le jh\le t-h}h\int\left[\int\limits_s^{t}du\,\chi[s,jh]\int p(s,u,x,v)M_1(u,jh,v,z)dv\right]\\&\qquad\times\Xi(jh,t,z,y)dz\\&\quad=\sqrt{h}\int\limits_s^t du\int p(s,u,x,v)\sum_{|\nu|=3}\frac{\chi_\nu(u,v)}{\nu!}\\&\qquad\times D_v^\nu\left[\sum_{0\le jh\le t-h}h\chi[s,jh]\int\widetilde{p}(u,jh,v,z)\Xi(jh,t,z,y)dz\right]dv\\&\quad=\sqrt{h}\int\limits_s^t du\int p(s,u,x,v)\times\sum_{|\nu|=3}\frac{\chi_\nu(u,v)}{\nu!}D_v^\nu p_\Delta(u,t,v,y)dv\\&\quad=\sqrt{h}(p\otimes\mathcal{F}_1)[p_\Delta](s,t,x,y).\end{aligned} \tag{68}$$

Here, $\chi[s, jh]$ denotes the indicator of the interval $[s, jh]$. Using similar arguments as in the proof of (68) one can show that

$$h\left[\widetilde{\pi}_2 \otimes_h' \Xi(s,t,x,y) + (p \otimes \Re_2) \otimes_h' \Xi(s,t,x,y) + p \otimes_h (\Re_3 \otimes_h' \Xi)(s,t,x,y)\right]$$
$$= h(p \otimes \mathcal{F}_2)[p_\Delta](s,t,x,y) + hp \otimes \Pi_1 \otimes_h' \Xi(s,t,x,y). \tag{69}$$

For the first two terms in the right hand side of (62) we obtain from (68) and (69)

$$\begin{aligned}
&\sqrt{h}\left[\widetilde{\pi}_1 \otimes_h' \Xi(0,T,x,y) + (p \otimes \Re_1) \otimes_h' \Xi(0,T,x,y)\right] \\
&\quad + h\left[\widetilde{\pi}_2 \otimes_h' \Xi(0,T,x,y) + p \otimes (\Re_2 \otimes_h' \Xi)(0,T,x,y)\right. \\
&\quad \left. + \; p \otimes_h (\Re_3 \otimes_h' \Xi)(0,T,x,y)\right] \\
&= \sqrt{h}(p \otimes \mathcal{F}_1)[p_\Delta](0,T,x,y) + h(p \otimes \mathcal{F}_2)[p_\Delta](s,t,x,y) \\
&\quad + hp \otimes \Pi_1 \otimes_h' \Xi(s,t,x,y).
\end{aligned} \tag{70}$$

Using (68) we get

$$\begin{aligned}
&h\left[\widetilde{\pi}_1 \otimes_h' \Xi + p \otimes \left(\Re_1 \otimes_h' \Xi\right)\right] \otimes_h' (\Re_1 \otimes_h' \Xi)(0,T,x,y) \\
&\quad = h(p \otimes \mathcal{F}_1[p_\Delta]) \otimes_h' (\Re_1 \otimes_h' \Xi)(0,T,x,y) \\
&\quad = hp \otimes \mathcal{F}_1\left[p_\Delta \otimes_h' (\Re_1 \otimes_h' \Xi)\right](0,T,x,y).
\end{aligned}$$

Note that

$$\begin{aligned}
hp \otimes \Pi_1 \otimes_h' \Xi(s,t,x,y) &= h\int_s^t du \int p(s,u,x,v) \\
&\qquad \times \sum_{|\nu|=3} \frac{\chi_\nu(u,v)}{\nu!} D_v^\nu[\widetilde{\pi}_1 \otimes_h' \Xi](u,t,v,y) \\
&= hp \otimes \mathcal{F}_1[\widetilde{\pi}_1 \otimes_h' \Xi](s,t,x,y).
\end{aligned}$$

For the proof of (63) it remains to show that

$$\begin{aligned}
&hp \otimes \mathcal{F}_1[\widetilde{\pi}_1 \otimes_h' \Xi + p_\Delta \otimes_h' (\Re_1 \otimes_h' \Xi)](0,T,x,y) \\
&\quad = h\left(p \otimes \mathcal{F}_1[p \otimes \mathcal{F}_1[p_\Delta]]\right)(0,T,x,y) + O(h^{1+\delta}\zeta_{\sqrt{T}}(y-x)).
\end{aligned} \tag{71}$$

We will show that

$$hp \otimes \mathcal{F}_1[(p - p_\Delta) \otimes_h' (\Re_1 \otimes_h' \Xi)](0,T,x,y) = O(h^{1+\delta}\zeta_{\sqrt{T}}(y-x)), \tag{72}$$

$$\begin{aligned}
&hp \otimes \mathcal{F}_1[p \otimes_h' (\Re_1 \otimes_h' \Xi)](0,T,x,y) - hp \otimes \mathcal{F}_1[p \otimes (\Re_1 \otimes_h' \Xi)](0,T,x,y) \\
&\quad = O(h^{1+\delta}\zeta_{\sqrt{T}}(y-x)).
\end{aligned} \tag{73}$$

Claim (71) follows from (72), (73) and (68). The estimate (73) can be shown similarly as in the proof of (112) in [22]. An additional singularity arising from the derivatives in the operator $\mathcal{F}_1[\cdot]$ can be treated by using the additional factor h in (73). For the estimate (72) see [22].

The proof of Theorem 1 is now completed by showing that

$$hp \otimes \mathcal{F}_2[p_\Delta](0, T, x, y) - hp \otimes \mathcal{F}_2[p](0, T, x, y) = O(h^{1+\gamma} \phi_{C,\sqrt{T}}(y - x)).$$
$$hp \otimes \mathcal{F}_1[p \otimes \mathcal{F}_1[p - p_\Delta]] = O(h^{1+\delta} \phi_{C,\sqrt{T}}(y - x)).$$

For a proof of these claims see [22].

References

1. Bally, V., Talay, D.: The law of the Euler scheme for stochastic differential equations: I. Convergence rate of the distribution function. Probab. Theory Relat. Fields **104**, 43–60 (1996)
2. Bally, V., Talay, D.: The law of the Euler scheme for stochastic differential equations: II. Convergence rate of the density. Monte Carlo Methods Appl. **2**, 93–128 (1996)
3. Bertail, P., Clémençon, S.: Edgeworth expansions of suitably normalized sample mean statistics for atomic Markov chains. Probab. Theory Relat. Fields **130**, 388–414 (2004)
4. Bertail, P., Clémençon, S.: Regenerative block bootstrap for Markov chains. Bernoulli **12**, 689–712 (2006)
5. Bhattacharya, R., Rao, R.: Normal Approximations and Asymptotic Expansions. Wiley, New York (1976)
6. Bolthausen, E.: The Berry–Esseen theorem for functionals of discrete Markov chains. Z. Wahrsch. Verw. Geb. **54**, 59–73 (1980)
7. Bolthausen, E.: The Berry–Esseen theorem for strongly mixing Harris recurrent Markov chains. Z. Wahrsch. Verw. Geb. **60**, 283–289 (1982)
8. Friedman, A.: Partial differential equations of parabolic type. Prentice-Hall, Englewood Cliffs (1964)
9. Fukasawa, M.: Edgeworth expansion for ergodic diffusions. Probab. Theory Relat. Fields (in print) (2007a)
10. Fukasawa, M.: Regenerative block bootstrap for ergodic diffusions (preprint) (2007b)
11. Götze, F.: Edgeworth expansions in functional limit theorems. Ann. Probab. **17**(4), 1602–1634 (1989)
12. Götze, F., Hipp, C.: Asymptotic expansions for sums of weakly dependent random vectors. Z. Wahrsch. Verw. Geb. **64**, 211–239 (1983)
13. Guyon, J.: Euler scheme and tempered distributions. Stoch. Proc. Appl. **116**, 877–904 (2006)
14. Jacod, J.: The Euler scheme for Levy driven stochastic differential equations: limit theorems. Ann. Probab. **32**, 1830–1872 (2004)
15. Jacod, J., Protter, P.: Asymptotic error distributions for the Euler method for stochastic differential equations. Ann. Probab. **26**, 267–307 (1998)
16. Jacod, J., Kurtz, T., Meleard, S., Protter, P.: The approximate Euler method for Levy driven stochastic differential equations. Ann. de l'I.H.P. **41**, 523–558 (2005)
17. Jensen, J.L.: Asymptotic expansions for strongly mixing Harris recurrent Markov chains. Scand. J. Statist. **16**, 47–63 (1989)
18. Konakov, V., Molchanov, S.: On the convergence of Markov chains to diffusion processes. Teoria veroyatnostei i matematiceskaya statistika, **31**, 51–64 (1984) (in russian) [English translatiion in Theory Probab. Math. Stat. **31**, 59–73 (1985)]
19. Konakov, V., Mammen, E.: Local limit theorems for transition densities of Markov chains converging to diffusions. Probab. Theory Relat. Fields. **117**, 551–587 (2000)
20. Konakov, V., Mammen, E.: Edgeworth type expansions for Euler schemes for stochastic differential equations. Monte Carlo Methods Appl. **8**, 271–286 (2002)
21. Konakov, V., Mammen, E.: Edgeworth-type expansions for transition densities of Markov chains converging to diffusions. Bernoulli **11**(4), 591–641 (2005)
22. Konakov, V., Mammen, E.: Small time Edgeworth-type expansions for weakly convergent nonhomogeneous Markov chains. Extended version (preprint). ArXiv: 0705.3139 (available under http://fr.arxiv.org/abs/0705.3139) (2007)

23. Kusuoka, S., Yoshida, N.: Malliavin calculus, geometric mixing, and expansion of diffusion functionals. Probab. Theory Relat. Fields **116**, 457–484 (2000)
24. Ladyzenskaya, O.A., Solonnikov, V.A., Ural'ceva, N.: Linear and quasi-linear equations of parabolic type. Amer. Math. Soc., Providence, Rhode Island (1968)
25. Malinovskii, V.K.: Limit theorems for Harris Markov chains, 1. Theory Probab. Appl. **31**, 269–285 (1987)
26. Mykland, P.A.: Aymptotic expansions and bootstrapping distributions for dependent variables: A martingale approach. Ann. Statist. **20**, 623–654 (1992)
27. McKean, H.P., Singer, I.M.: Curvature and the eigenvalues of the Laplacian. J. Diff. Geom. **1**, 43–69 (1967)
28. Protter, P., Talay, D.: The Euler scheme for Levy driven stochastic differential equations. Ann. Probab. **25**, 393–323 (1997)
29. Skorohod, A.V.: Studies in the Theory of Random Processes. Addison-Wesley, Reading [English translation of Skorohod A. V. (1961). Issledovaniya po teorii sluchainykh processov. Kiev University Press] (1965)
30. Stroock, D.W., Varadhan, S.R.: Multidimensional Diffusion Processes. Springer, Berlin (1979)
31. Yoshida, N.: Partial mixing and Edgeworth expansion. Probab. Theory Relat. Fields **129**, 559–624 (2004)