Probab. Theory Relat. Fields 117, 551–587 (2000)
Digital Object Identifier (DOI) 10.1007/s004400000063

Valentin Konakov · Enno Mammen

Local limit theorems for transition densities of Markov chains converging to diffusions*

Received: 28 August 1998 / Revised version: 6 September 1999 /
Published online: 14 June 2000 –

Abstract. We consider triangular arrays of Markov chains that converge weakly to a diffusion process. Local limit theorems for transition densities are proved.

1. Introduction and results

In this paper we study triangular arrays of Markov chains $X_n(k)$ $(n \geq 1, 0 \leq k \leq n)$ that converge weakly to a diffusion process (for $n \to \infty$). Our main result is that the transition densities converges with rate $\mathrm{O}(n^{-1/2})$ to the transition density in the diffusion model.

Weak convergence of the distribution of scaled discrete time Markov processes to diffusions has been extensively studied in the literature. First general results have been received by A.V. Skorohod in the early 60's [see Skorohod (1965)]. In chapter 6 of this monography he proves a general weak convergence theorem where the limiting process consists of two components: a diffusion and a jump process. The continuous diffusion is described by a stochastic Itô integral, the jump component is a stochastic integral with respect to a Poissonian random measure. Later, Strook and Varadhan (1979, chapter 11) developped an elegant "martingale problem" approach for continuous Markov processes. Probably, there the most general results on weak convergence to a continuous diffusion can be found. The results in Skorohod (1965) and in Strook and Varadhan (1979) are obtained by probabilistic methods. In this paper, for the treatment of the convergence of transition densities we will use an analytical approach. We will apply the parametrix method for parabolic PDEs and a modification of this method for discrete time Markov chains.

V. Konakov: Central Economics Mathematical Institute, Academy of Sciences Nahimovskii av. 47, 117418 Moscow, Russia

E. Mammen: Institut für Angewandte Mathematik, Ruprecht-Karls-Universität Heidelberg Im Neuenheimer Feld 294, 69120 Heidelberg, Germany. e-mail: mammen@statlab.uni-heidelberg.de

* This research was supported by grant 436 - RUS 113/467/1 from the Deutsche Forschungsgemeinschaft and by grants 98-01-04108 and 98-01-00524 from the Russian Foundation of Fundamental Researches.

Mathematics Subject Classification (1991): Primary 62G07; Secondary 60G60

Key words and phrases: Markov chains – Diffusion processes – Transition densities

Standard reference for the parametrix method are the books by A. Friedman (1964) and O.A. Ladyženskaja, V.A. Solonnikov and Ural´ceva (1968) on parabolic PDEs. But for our purposes a slightly different version of the parametrix method used in McKean and Singer (1967) is more appropriate. This method permits to obtain a tractable representation of transition densities of diffusions that is based on Gaussian densities, see Lemma 3.1. A similar representation is derived for the discrete time Markov chain X_n, see Lemma 3.6. The parametrix method was used e.g. in Kuznetsov (1998) to obtain bounds for Poisson kernels. But as far as we know for Markov chains the parametrix method was not systematically developed before. A description of the parametrix method is given in Subsection 3.1. Local limit theorems for Markov chains were given in Kasymdzganova (1981) and Konakov and Molchanov (1984). Kasymdzganova (1981) considered the case of a homogeneous random walk on the lattice $\mathbb{Z}^p$[with no drift, e.g. $m \equiv 0$]. Local limit theorems for homogeneous Markov chains with continuous state space were given in Konakov and Molchanov (1984).

We now give a more detailed description of the Markov chains and their diffusion limit. For each $n \geq 1$ we consider Markov chains $X_n(k)$ where the time k runs from 0 to n. The Markov chain X_n is assumed to take values in $\mathbb{R}^p$. The dynamics of the chain X_n is described by

$$\begin{aligned} X_n(k+1) = X_n(k) &+ \Delta_n(k+1) m\{s_n(k), X_n(k)\} \\ &+ \Delta_n(k+1)^{1/2} \varepsilon_n(k+1). \end{aligned} \tag{1.1}$$

Here $\Delta_n(k) > 0$ are real numbers with

$$\sum_{k=1}^{n} \Delta_n(k) = 1.$$

The numbers $s_n(k)$ are defined as $s_n(0) = 0$ and

$$s_n(k) = \sum_{i=1}^{k} \Delta_n(i) \quad \text{for} \quad k \geq 1.$$

Furthermore, m is a function $m : [0, 1] \times \mathbb{R}^p \to \mathbb{R}^p$. We make the Markov assumption that the conditional distribution of the innovation $\varepsilon_n(k+1)$ given the past $X_n(k), X_n(k-1), \ldots$ depends only on the last value $X_n(k)$. Given $X_n(i) = x(i)$ for $i = 0, \ldots, k$ the variable $\varepsilon_n(k+1)$ has a conditional density $q\{s_n(k), x(k), \bullet\}$. The conditional covariance matrix of $\varepsilon_n(k+1)$ is denoted by $\Sigma\{s_n(k), x(k)\}$. Here q is a function mapping $[0, 1] \times \mathbb{R}^p \times \mathbb{R}^p$ into $\mathbb{R}_+$. Furthermore, Σ is a function mapping $[0, 1] \times \mathbb{R}^p$ into the set of positive definite $p \times p$ matrices. The conditional density of $X_n(n)$, given $X_n(0) = x$, is denoted by $p_n(x, \bullet)$. Study of the transition densities $p_n(x, y)$ is the topic of this paper. Conditions on $\Delta_n(k)$, m, $q\{s_n(k), x(k), \bullet\}$ and $\Sigma\{s_n(k), x(k)\}$ will be given below.

By time change the Markov chain X_n defines a process Y_n on $[0, 1]$. More precisely, put $\kappa_n(t) = \sup\{k : s_n(k) \leq t, 0 \leq k \leq n\}$. This defines a monotone time transform $\kappa_n : [0, 1] \to \{0, \ldots, n\}$. Using this time transform we get the following process:

$$Y_n(t) = X_n\{\kappa_n(t)\}.$$

Under our assumptions, see below, the process Y_n converges to a diffusion $Y(t)$. This follows for instance from Theorem 1, p. 82 in Skorohod (1987). The diffusion is defined by $Y(0) = x$ and

$$dY(t) = m\{t, Y(t)\}dt + \Lambda\{t, Y(t)\}dW(t),$$

where W is a p dimensional Brownian motion. The matrix $\Lambda(t, z)$ is the unique symmetric matrix defined by $\Lambda(t, z)\Lambda(t, z)^T = \Sigma(t, z)$. The conditional density of $Y(1)$, given $Y(0) = x$, is denoted by $p(x, \bullet)$. Note that the conditional density of $Y_n(1)$, given $Y_n(0) = x$, is denoted by $p_n(x, \bullet)$.

For our result we use the following conditions.

(A1) For $t \in [0, 1]$ and $x \in \mathbb{R}^p$ let $q\{t, x, \bullet\}$ be a density in $\mathbb{R}^p$ with

$$\int q\{t, x, z\}z\, dz = 0 \quad \text{for all } t \in [0, 1], x \in \mathbb{R}^p,$$

$$\int q\{t, x, z\}z_i z_j\, dz = \sigma_{ij}(t, x) \qquad \text{for all } t \in [0, 1], x \in \mathbb{R}^p \text{ and } i, j = 1, \ldots, p.$$

The matrix with elements $\sigma_{ij}(t, x)$ is denoted by $\Sigma(t, x)$.

(A2) There exist a positive integer S' and a function $\psi : \mathbb{R}^p \to \mathbb{R}$ with $\sup_{x\in\mathbb{R}^p} |\psi(x)| < \infty$ and $\int_{\mathbb{R}^p} \|x\|^S |\psi(x)|\, dx < \infty$ for $S = 2pS' + 4$ such that

$$|D_z^\nu q\{t, x, z\}| \le \psi(z) \quad \text{for all } t \in [0, 1], x, z \in \mathbb{R}^p, \text{ and } |\nu| = 0, \ldots, 4,$$

$$|D_x^\nu q\{t, x, z\}| \le \psi(z) \quad \text{for all } t \in [0, 1], x, z \in \mathbb{R}^p \text{ and } |\nu| = 0, \ldots, 2.$$

[For the case that $S' = 1$ Theorem 2.1 can be shown under the weaker assumption that (A2) holds for a function ψ with $\sup_{x\in\mathbb{R}^p} |\psi(x)| < \infty$ and $\int_{\mathbb{R}^p} \|x\|^k |\psi(x)|\, dx < \infty$ for an integer $k > p + 4$.] Furthermore, for all $x \in \mathbb{R}^p$ we assume that $\int \left|q\{t, x, z\} - q\{t', x, z\}\right| dz \to 0$ for $\left|t - t'\right| \to 0$. [Under our assumptions this follows e.g. if $q\{t, x, z\}$ is continuous in t for fixed x and z.]

(A3) There exist positive constants c and C such that

$$c \le \theta^T \Sigma(t, x)\theta \le C$$

for all θ, $\|\theta\| = 1$, t and x.

(A4) There exists a constant B with

$$B^{-1} < \frac{\Delta_n(k)}{\Delta_n(l)} < B$$

for $n \ge 1$ and $1 \le k, l \le n$. [Then it follows that $\Delta_{max} = \max_{1\le j\le n} \Delta_n(j) = O(n^{-1})$.]

(A5) The functions $m(t, x)$ and $\Sigma(t, x)$ and their first derivatives with respect to x and with respect to t 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). Furthermore, $\partial^2/(\partial x_j \partial x_k)\Sigma(t, x)$ exists for $1 \le j, k \le p$ and is Holder continuous with respect to x (with positive exponent δ and constant that do not depend on t).

The following theorem contains our main result. It gives a bound for the rate of convergence of p_n to p.

Theorem 1.1. *Assume (A1) – (A5). Then the following estimate holds:*

$$\sup_{x,y \in R^p} \left(1 + \|y - x\|^{2(S'-1)}\right) |p_n(x, y) - p(x, y)| = O(n^{-1/2}),$$

where S' is defined in Assumption (A2). The norm $\| \dots \|$ is the usual Euclidean norm.

2. Examples and extensions

(i) *Approximation by diffusions that depend on n.* The result can be extended to the case that q, m and Σ depend on n. For this purpose conditions (A2), . . ., (A5) have to be replaced by assumptions that hold uniformly in n. Then the limiting fixed diffusion has to to replaced by a sequence of approximating diffusions depending on n.

(ii) *Unbounded drift function.* Our result can be extended to the case of an unbounded drift function m that is of the form $b(t)x+a(t, x)$ where a fullfills the conditions stated for m and where $b(t)$ is a matrix that depends continuously on t.

(iii) *Unbounded one step transition density.* Our results can be extended to unbounded transition densities if the transition density for a finite number of steps is bounded, see e.g. (vii).

(iv) *Functionals of Markov chains.* Our theorem implies that the density of $(Y_n(t_1), \dots, Y_n(t_k))$ converges to the density of $(Y(t_1), \dots, Y(t_k))$ in L_1 norm for any tuple $0 \le t_1 < \dots < t_k \le 1$. We conjecture that with the approach of Davydov (1980, 1981) these results can be used to show that the density of $H(Y_n(\bullet))$ converges to the density of $H(Y(\bullet))$ for a wide range of functionals H.

(v) *Conditional Markov chains.* In particular, our result can be used to show that the conditional density of $(Y_n(t_1), \dots, Y_n(t_k))$ given $Y_n(1))$ converges to the conditional density of $(Y(t_1), \dots, Y(t_k))$ given $Y(1))$ (in L_1 norm), where tuple $t_1, \dots, t_k$ is a tuple with $0 \le t_1 < \dots < t_k < 1$.

(vi) *Euler approximations.* The case where q is a normal density corresponds to Euler approximations that are the simplest strong Taylor approximations used as numerical solutions to stochastic differential equations, see Kloeden and Platen (1992). Rates of convergence for the distribution functions and densities of transition probabilities can be found in Bally and Talay (1996a, b).

A detailed discussion of new literature on numerical methods for SDE's can be found in Platen (1999).

(vii) *Transport processes.* Let us consider a symmetric and positive definite $p \times p$ matrix $S(x)$ and vector $m(x) = (m_1(x), \dots, m_p(x))^T$ where $x \in I\!R^p$. For $a > 0$ we consider independent variables $R_{a,1}, R_{a,2}, \dots, U_{a,1}, U_{a,2}, \dots$ where $R_{a,i}$ have density $a^{-1}\exp(-r/a)$ and where $U_{a,i}$ are uniformly distributed on the unit sphere in $I\!R^p$. We define the following chain (transport processes, see e.g. Pinsky, 1991):

$$\begin{aligned} X_a(0) &= x, \\ X_a(i+1) &= X_a(i) + a^2 m(X_a(i)) + S(X_a(i))R_{a,i}U_{a,i}, \\ &\qquad \text{for} \quad 0 \le i \le [1/a^2]. \end{aligned}$$

This process has no bounded one step transition density and it does not fulfill the conditions of our theorem for this reason. However it is easy to show that for a finite numbers k of steps the transition density of $X_a(i+k)$ given $X_a(i)$ is bounded, so that we can apply our theorem to the process $i \to X_a(ik)$. This shows that the density $X_a([1/a^2])$ converges to the density of the diffusion Y at time point $t = 1$ for $a \to 0$ where

$$\begin{aligned} Y(0) &= x, \\ dY(t) &= cS(Y(t))\,dW(t) + m(Y(t))\,dt, \quad \text{for} \quad 0 \le t \le 1, \end{aligned}$$

where c is an appropriate constant. The speed of convergence is of order $O(a)$.

(viii) *Lattice distributions.* Our approach can be extended to obtain local limit theorems for a general class of nonhomogeneous random walks on a lattice $Z\!\!Z^p$. A treatment of the homogeneous case can be found in Konakov and Mammen (1999). This generalizes the results of Konovalov (1981) and Kasymdzganova (1981).

(ix) *Edgeworth expansions.* We conjecture that Edgeworth expansions for Markov chains can be proved by our approach. This will be treated elsewhere.

(x) *Statistical applications.* This research was partially motivated by recent new approaches in time series analysis. In a series of papers [see e.g. Robinson (1983), Tjøstheim (1994), Franke, Kreiss and Mammen (1996)] it has been proposed to use nonparametric approaches to model time series. In particular nonparametric autoregression models have been considered:

$$X(k+1) = m(X(k)) + \sigma(X(k))\varepsilon(k+1), \tag{2.1}$$

where the innovations $\varepsilon(1), \varepsilon(2), \dots$ are typically assumed to be i.i.d. mean zero variables. For the functions m and σ nonparametric smoothness assumptions are made and nonparametric smoothing methods are proposed for their estimation. For a discussion of different nonparametric statistical problems in these models we refer to the references above. Under regularity conditions on m, σ and the distribution of $\varepsilon(i)$, solutions of (2.1) are stationary processes.

In Dahlhaus (1997) models are proposed for time series that are not stationary, however locally stationary. In particular he considers autoregressive processes with time varying coefficients:

$$X_n(k+1) = a(\frac{k}{n})X_n(k) + \varepsilon(k+1). \tag{2.2}$$

Our model (1.1) is strongly related to the models (2.1)–(2.2). As in (2.2), we observe a function on a finer grid for $n \to \infty$. The main result of this paper may be applied to discuss statistical nonparametric estimation problems of the transition density and the shift function m under different smoothness and structural assumptions. Our result reduces the discussion of some of such problems in model (1.1) to the analysis of corresponding problems in diffusion models. For the discussion of some nonparametric estimation problems in diffusion models see Kutoyants (1997a, b).

3. Proofs

The proof of Theorem 1.1 is organized as follows. In the next two subsections we will state series expansions for the transition densities of the limiting diffusion and for the Markov chain. The series only depend on transition densities of “frozen” processes. The “frozen” diffusion is a Gaussian process that has a Gaussian density as transition density. For the “frozen” Markov chain we get transition densities that are densities of sums of independent variables. The difference between these densities and the Gaussian densities can be treated by Edgeworth expansions. This is done in Subsection 3.3. These are the main steps of the proof of Theorem 1.1. The remaining steps of the proof of Theorem 1.1 are given in Subsection 3.4. Longer proofs of some lemmas are given in Subsection 3.5.

3.1. The parametrix method

In this subsection we will state an infinite series expansion of the transition density p of the limiting diffusion process Y, see Lemma 3.1. We will give a similar expansion for the Markov chain in the next subsection, see Lemma 3.6. Our proof of Theorem 1.1 will be based on the comparison of these two series. The series for the transition densities will be derived by the parametrix method. We will give a description of the parametrix method below.

For the statement of the expansion of p in Lemma 3.1 we have to introduce additional diffusion processes. For $0 < s < 1$ and $x, y \in \mathbb{R}^p$ we define diffusions $\tilde{Y} = \tilde{Y}_{s,x,y}$ that are defined for $s \le t \le 1$ by

$$\tilde{Y}(s) = x$$

and

$$d\tilde{Y}(t) = m\{t, y\}dt + \Lambda\{t, y\}dW(t).$$

The processes $\tilde{Y}$ are called “frozen” diffusions. We define $\tilde{p}(s, t, x, y)$ as the conditional density of $\tilde{Y}(t)[= \tilde{Y}_{s,x,y}(t)]$ at the point y, given $\tilde{Y}(s) = x$. Note that the variable y acts here twice: as the argument of the density and as a defining quantity of the process $\tilde{Y} = \tilde{Y}_{s,x,y}$. Furthermore, we denote by $\tilde{p}_j^y(x, z)$ the conditional density of $\tilde{Y}(s_n(j+1))[= \tilde{Y}_{s_n(j),x,y}(s_n(j+1))]$ at the point z, given $\tilde{Y}(s_n(j)) = x$. The process $\tilde{Y}$ is a simple Gaussian process. Its transition densities $\tilde{p}$ are given explicitely. By definition, we have that

$$\begin{aligned}\tilde{p}(s,t,x,y) &= (2\pi)^{-p/2}(\det \Sigma(s,t,y))^{-1/2} \\ &\quad \times \exp[-\frac{1}{2}\{y - x - m(s,t,y)\}' \Sigma(s,t,y)^{-1}\{y - x - m(s,t,y)\}],\end{aligned} \tag{3.1}$$

where

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

Let us introduce the following differential operators L and $\tilde{L}$:

$$Lf(s,t,x,y) = m(s,x)^T \frac{\partial f(s,t,x,y)}{\partial x} + \frac{1}{2}\,\mathrm{tr}[\Lambda(s,x)^T \frac{\partial^2 f(s,t,x,y)}{(\partial x)^2}\Lambda(s,x)]$$

and

$$\tilde{L}f(s,t,x,y) = m(s,y)^T \frac{\partial f(s,t,x,y)}{\partial x} + \frac{1}{2}\,\mathrm{tr}[\Lambda(s,y)^T \frac{\partial^2 f(s,t,x,y)}{(\partial x)^2}\Lambda(s,y)].$$

Note that L and $\tilde{L}$ corresponds to the infinitesimal operators of Y or of the frozen process $\tilde{Y}_{s,x,y}$, respectively, i.e.

$$Lf(s,t,x,y) = \lim_{h\to 0} h^{-1}\{E\,[f(s,t,Y(s+h),y)|Y(s) = x] - f(s,t,x,y)\}, \tag{3.2}$$

$$\tilde{L}f(s,t,x,y) = \lim_{t\to 0} t^{-1}\{E[f(s,t,\tilde{Y}_{s,x,y}(s+h),y)] - f(s,t,x,y)\}. \tag{3.3}$$

We put

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

Then

$$\begin{aligned}H(s,t,x,y) &= \frac{1}{2}\sum_{i,j=1}^{p}(\sigma_{ij}(s,x) - \sigma_{ij}(s,y))\frac{\partial^2 \tilde{p}(s,t,x,y)}{\partial x_i \partial x_j} \\ &\quad + \sum_{i=1}^{p}(m_i(s,x) - m_i(s,y))\frac{\partial \tilde{p}(s,t,x,y)}{\partial x_i}.\end{aligned} \tag{3.4}$$

Now we define the following convolution type binary operation $\otimes$:

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

We write $g \otimes H^{(0)}$ for g and for $r = 1, 2, \ldots$ we denote the r-fold "convolution" $(g \otimes H^{(r-1)}) \otimes H$ by $g \otimes H^{(r)}$. With these notations we can state our expansion for p.

Lemma 3.1. *For* $0 \le s < t \le 1$ *the following formula holds:*

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

A proof of Lemma 3.1 can be found in McKean and Singer (1967). It is based on application of the parametrix method. We give now a description of this approach. It is well-known [see e.g. Dynkin (1965)] that the transition density p is the fundamental solution of Kolmogorov's backward differential equation:

$$\frac{\partial p}{\partial s} + Lp = 0.$$

Moreover under Condition (A5) p satisfies Kolmogorov's forward differential equation (or the Fokker-Plank equation)

$$-\frac{\partial p}{\partial t} + L^* p = 0,$$

where

$$L^* f(s,t,x,y) = -\sum_{i=1}^{p} \frac{\partial [m_i(t,y) f(s,t,x,y)]}{\partial y_i} + \frac{1}{2} \sum_{i,j=1}^{p} \frac{\partial^2 [\sigma_{i,j}(t,y) f(s,t,x,y)]}{\partial y_i \, \partial y_j}.$$

We use now

$$\tilde{p}(t,t,x,y) = \delta(x-y), \tag{3.5}$$

$$p(t,t,x,y) = \delta(x-y), \tag{3.6}$$

$$\int_{I\!R^p} \tilde{p}(u,t,z,y) L^* p(s,u,x,z) dz = \int_{I\!R^p} p(s,u,x,z) L \tilde{p}(u,t,z,y) dz, \tag{3.7}$$

where δ is the Dirac delta function. Equations (3.5) and (3.6) are the initial conditions of the processes Y and $\tilde{Y}$, respectively. Equation (3.7) follows by partial integration (or this is just the property that, as is well known, L and L^* correspond to adjoint operators.) With (3.5)–(3.7) we can write the basic parametrix equation

$$\begin{aligned} p(s,t,x,y) &- \tilde{p}(s,t,x,y) \\ &= \int_s^t du \frac{\partial}{\partial u} \left[\int_{I\!R^d} p(s,u,x,z) \tilde{p}(u,t,z,y) dz \right] \\ &= \int_s^t du \int_{I\!R^d} \left[\tilde{p}(u,t,z,y) L^* p(s,u,x,z) - p(s,u,x,z) \tilde{L} \tilde{p}(u,t,z,y) \right] dz \\ &= \int_s^t du \int_{I\!R^d} p(s,u,x,z) \left[L - \tilde{L} \right] \tilde{p}(u,t,z,y) dz. \end{aligned} \tag{3.8}$$

This equation can be rewritten as

$$p = \tilde{p} + p \otimes H. \tag{3.9}$$

Iterative application of (3.9) gives

$$p = \sum_{i=0}^{r} \tilde{p} \otimes H^{(i)} + p \otimes H^{(r+1)}. \tag{3.10}$$

Lemma 3.1 follows by taking the limit $r \to \infty$ in (3.10).

We will make use of the bounds on H and $\tilde{p} \otimes H^{(r)}$ that are stated in the following lemma. Proofs of these bounds can be found again in McKean and Singer (1967). For a more detailed proof of Lemma 3.2 see also Ladyženskaja, Solonnikov and Ural´ceva (1968).

Lemma 3.2. *There exist constants C and C_1 (that do not depend on x and y) such that the following inequalities hold:*

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

and

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

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

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

In the proof of Theorem 1.1 we will need bounds on the derivatives of H, $\tilde{p}$ and $\tilde{p} \otimes H^{(r)}$. These are stated in the next three lemmas.

Lemma 3.3. *There exist constants C and C_1 such that the following estimate holds*

$$\left| \frac{\partial H(s,t,x,y)}{\partial s} \right| \le C_1 \rho^{-3} \phi_{C,\rho}(y-x),$$

where ρ and $\phi_{C,\rho}$ are defined as in Lemma 3.2.

Proof of Lemma 3.3. By Assumption (A5), $\sigma_{ij}(s,x)$ and $m_i(s,x)$ have partial derivatives with respect to s that are Lipschitz continuous with respect to x. Using (3.4), one sees that for the statement of the lemma it suffices to show for some constants C_1' and C_2' that

$$\left| \frac{\partial^2 \tilde{p}(s,t,x,y)}{\partial x_i \partial x_j} \right| \le C_1' \rho^{-2} \phi_{C_2',\rho}(y-x),$$

$$\left| \frac{\partial}{\partial s} \frac{\partial^2 \tilde{p}(s,t,x,y)}{\partial x_i \partial x_j} \right| \le C_1' \rho^{-4} \phi_{C_2',\rho}(y-x).$$

These claims follow from Assumption (A5) by taking partial derivatives of $\tilde{p}$, see (3.1). □

Lemma 3.4. *There exist constants C_1 and C such that the following estimates hold for $1 \le k \le p$*

$$\left|\frac{\partial}{\partial y_k} H(s,t,x,y) + \frac{\partial}{\partial x_k} H(s,t,x,y)\right| \le C_1 \rho^{-1} \phi_{C,\rho}(y-x), \qquad (3.11)$$

$$\left|\frac{\partial}{\partial s} H(s,t,x,y) + \frac{\partial}{\partial t} H(s,t,x,y)\right| \le C_1 \rho^{-1} \phi_{C,\rho}(y-x), \qquad (3.12)$$

where ρ and $\phi_{C,\rho}$ are defined as in Lemma 3.2.

Proof of Lemma 3.4. The statements of the lemma can be seen from the definition of $H(s,t,x,y)$, well-known properties of Gaussian densities and (A5). □

Lemma 3.5. *There exist constants C_1 and C such that the following estimate holds for $r \ge 0$*

$$\left|\frac{\partial \tilde{p} \otimes H^{(r)}(s,t,x,y)}{\partial t}\right| \le C_1^{r+1} \frac{\rho^{r-2}}{\Gamma(1+\frac{r}{2})} \phi_{C,\rho}(y-x), \qquad (3.13)$$

where ρ and $\phi_{C,\rho}$ are defined as in Lemma 3.2.

The proof of Lemma 3.5 is deferred to subsection 3.5.

3.2. Application of the parametrix method to Markov chains

In this subsection we derive a finite series expansion of the transition density $p_n(s,t,x,y)$ of the Markov chain, see Lemma 3.6. Here, $p_n(s,t,x,\bullet)$ denotes the conditional density of $Y_n(t)$, given $Y_n(s) = x$ (in particular, $p_n(0,1,x,y) = p_n(x,y)$). We proceed similarly as in the last subsection. Again we apply the parametrix method and for this purpose we introduce additional "frozen" Markov chains. These are defined as follows. For all $0 \le j \le n$ and $x, y \in I\!R^p$ we define the Markov chains $\tilde{X}_n = \tilde{X}_{n,j,x,y}$. For fixed j, x and y, the chain is defined for i with $j \le i \le n$. The dynamics of the chain is described by

$$\tilde{X}_n(j) = x$$

and

$$\tilde{X}_n(i+1) = \tilde{X}_n(i) + \Delta_n(i+1) m\{s_n(i), y\} + \Delta_n(i+1)^{1/2} \tilde{\varepsilon}_n(i+1).$$

The stochastic structure of the $I\!R^p$ valued innovations $\tilde{\varepsilon}_n(i)$ is described as follows. Given $\tilde{X}_n(l) = x(l)$ for $l = j, \dots, i$ the variable $\tilde{\varepsilon}_n(i+1)$ has a conditional density $q\{s_n(i), y, \bullet\}$. Note that the conditional distribution of $\tilde{X}_n(i+1) - \tilde{X}_n(i)$ does not depend on the past $\tilde{X}_n(l)$ for $l = j, \dots, i$. Let us call $\tilde{X}_n$ the Markov chain *frozen at* y. We put $\tilde{Y}_n(t) = \tilde{X}_n\{\kappa_n(t)\}$ and we write $\tilde{p}_n(s_n(j), s_n(k), x, y)$ for the conditional density of $\tilde{X}_n(k)[= \tilde{X}_{n,j,x,y}(k)]$ at the point y, given $\tilde{X}_n(j) = x$. Note that, as in the case of a "frozen" diffusion the variable y acts here twice: as the argument of the density and as a defining quantity of the process $\tilde{X}_n = \tilde{X}_{n,j,x,y}$. Let us introduce the following infinitesimal operators L_n and $\tilde{L}_n$:

$$\begin{aligned} &L_n\ f(s_n(j), s_n(k), x, y) \\ &\quad = \frac{\int p_{n,j}(x,z) f(s_n(j+1), s_n(k), z, y)dz - f(s_n(j+1), s_n(k), x, y)}{\Delta_n(j+1)}, \\ &\tilde{L}_n\ f(s_n(j), s_n(k), x, y) \\ &\quad = \frac{\int \tilde{p}^y_{n,j}(x,z) f(s_n(j+1), s_n(k), z, y)dz - f(s_n(j+1), s_n(k), x, y)}{\Delta_n(j+1)}, \end{aligned}$$

where we write $p_{n,j}(x,z) = p_n(s_n(j), s_n(j+1), x, z)$ and where $\tilde{p}^y_{n,j}(x, \bullet)$ denotes the conditional density of $\tilde{X}_n(j+1)[= \tilde{X}_{n,j,x,y}(j+1)]$ given $\tilde{X}_n(j) = x$. Note that L_n and $\tilde{L}_n$ are defined in analogy with the definition of L and $\tilde{L}$, see (3.2)-(3.3). We remark that for some technical reasons on the right hand side of the definitions of $L_n\ f$ and $\tilde{L}_n\ f$ the terms $f(s_n(j+1), \ldots)$ appear instead of $f(s_n(j), \ldots)$. The reasons will become apparent in the development of the proof of Theorem 1.1. For $k > j$ we put in analogy with the definition H

$$H_n = \{L_n - \tilde{L}_n\}\tilde{p}_n.$$

In the following we use the following convolution type binary operation $\otimes_n$:

$$\begin{aligned} &(g \otimes_n f)(s_n(j), s_n(k), x, y) \\ &\quad = \sum_{i=j}^{k-1} \Delta_n(i+1) \int_{I\!R^p} g(s_n(j), s_n(i), x, z) f(s_n(i), s_n(k), z, y)dz, \end{aligned}$$

where $0 \le j < k \le n$. In this definition the convention is used that $\sum_{i=j}^{k-1} \ldots = 0$ if $j \ge k$. We write $g \otimes_n H_n^{(0)}$ for g and for $r = 1, 2, \ldots, n$, we denote the r fold "convolution" $(g \otimes_n H_n^{(r-1)}) \otimes_n H_n$ by $g \otimes_n H_n^{(r)}$. Note that $\otimes_n$ is a discretized version of $\otimes$.

The next lemma gives the "parametrix" expansion of p_n.

Lemma 3.6. *For $0 \le j < k \le n$ the following formula holds:*

$$p_n(s_n(j), s_n(k), x, y) = \sum_{r=0}^{k-j} (\tilde{p}_n \otimes_n H_n^{(r)})(s_n(j), s_n(k), x, y),$$

where in the calculation of $\tilde{p}_n \otimes_n H_n^{(r)}$ we define

$$p_n(s_n(j), s_n(j), x, y) = \tilde{p}_n(s_n(k), s_n(k), x, y) = \delta(x - y).$$

Here δ denotes the Dirac function.

Proof of Lemma 3.6. Note that by definition:

$$H_n(s_n(j), s_n(k), x, y) \\ = \frac{\int [p_{n,j}(x,z) - \tilde{p}^y_{n,j}(x,z)]\tilde{p}_n(s_n(j+1), s_n(k), z, y)dz}{\Delta_n(j+1)}. \tag{3.14}$$

Using the Markov property we get the following identity:

$$\begin{aligned} &p_n(s_n(j), s_n(k), x, y) - \tilde{p}_n(s_n(j), s_n(k), x, y) \\ &= \sum_{i=j}^{k-1} \Delta_n(i+1) \int p_n(s_n(j), s_n(i), x, z) \\ &\quad \times \int \frac{[p_{n,i}(z,z') - \tilde{p}^y_{n,i}(z,z')]\tilde{p}_n(s_n(i+1), s_n(k), z', y)}{\Delta_n(i+1)}\, dz'\, dz \\ &= \sum_{i=j}^{k-1} \Delta_n(i+1) \int p_n(s_n(j), s_n(i), x, z) H_n(s_n(i), s_n(k), z, y) dz \\ &= (p_n \otimes_n H_n)(s_n(j), s_n(k), x, y). \end{aligned}$$

The lemma follows by iterative application of this identity. □

3.3. *Bounds on $\tilde{p}_n - \tilde{p}$ based on Edgeworth expansions*

In this subsection we will develop some tools that are helpful for the comparison of the expansion of p (see Lemma 3.1) and the expansion of p_n (see Lemma 3.6). These expansions are simple expressions in $\tilde{p}$ or $\tilde{p}_n$, respectively. Recall that $\tilde{p}$ is a Gaussian density, see (3.1), and that $\tilde{p}_n$ is the density of a sum of independent variables. The densities $\tilde{p}$ and $\tilde{p}_n$ can be compared by application of Edgeworth expansions. This is done in Lemma 3.8. This is the essential step for the comparison of the expansions of p and p_n. The other lemmas of this subsection give bounds for several quantities. In the next lemma bounds will be given for derivatives of $\tilde{p}_n$. The proof of this lemma also makes essential use of Edgeworth expansions. In Lemma 3.9 we give an approximation for $H_n = (L_n - \tilde{L}_n)\tilde{p}_n$. We show that this term can be approximated by $K_n + M_n$, where $K_n = (L - \tilde{L})\tilde{p}_n$ and where M_n is defined in Lemma 3.9. Bounds on H_n, K_n, M_n and $\tilde{p}_n \otimes_n H_n^{(r)}$ are given in Lemmas 3.10 and 3.11. These bounds will be used in the proof of Lemma 3.12 to show that in the expansion of p_n the terms $\tilde{p}_n \otimes_n H_n^{(r)}$ can be replaced by $\tilde{p}_n \otimes_n (M_n + K_n)^{(r)}$. Finally, in Lemma 3.13 we use our Gaussian approximation $\tilde{p}$ for the transition density $\tilde{p}_n$ of the Markov chain and we show that in the expansion of p_n, the density $\tilde{p}_n$ can be replaced by $\tilde{p}$.

Lemma 3.7. *The following bound holds:*

$$|D^\nu_u \tilde{p}_n(s_n(j), s_n(k), x, y)| \le C\rho^{-|\nu|}\, \xi_\rho(y-x) \tag{3.15}$$

for all $j < k$, *for all* x *and* y *and for all* ν *with* $0 \le |\nu| \le 2$. *Here,* $\rho = [s_n(k) - s_n(j)]^{1/2}$ *[for simplicity the indices* n, j *and* k *are suppressed in the notation],* $\xi_\rho(\bullet) = \rho^{-p}\xi(\bullet/\rho)$ *and*

$$\xi(z) = \frac{[1 + \|z\|^{S-2}]^{-1}}{\int [1 + \|z'\|^{S-2}]^{-1}\, dz'}.$$

The constant S *has been defined in Assumption (A2).*

Proof of Lemma 3.7. We first note that $\tilde{p}_n(s_n(j), s_n(k), x, \bullet)$ is the density of the vector

$$x + \mu_{j,k} + \sum_{i=j}^{k-1} \eta_i,$$

where $\mu_{j,k} = \sum_{i=j}^{k-1} \Delta_n(i+1) m\{s_n(i), y\}$ is deterministic, where $\eta_i = [\Delta_n(i+1)]^{1/2}\tilde{\varepsilon}_n(i+1)$, $[i = j, \dots, k-1]$, and where, as above in the definition of the "frozen" Markov chain $\tilde{Y}_n$, $\tilde{\varepsilon}_n(i+1)$ is a sequence of independent variables with densities $q(s_n(i), y, \bullet)$. Let $f_n(\bullet)$ be the density of the normalized sum $V_{j,k}^{-1/2} \sum_{i=j}^{k-1} \eta_i$ where

$$V_{j,k} = \sum_{i=j}^{k-1} \Delta_n(i+1)\Sigma(s_n(i), y).$$

It follows from (A3) that for some constants $c_1, \dots, c_4 > 0$ the following inequalities hold for all θ with $\|\theta\| = 1$ and all $j < k$

$$c_1\rho^{-1} \le \theta^T V_{j,k}^{-1/2}\theta \le c_2\rho^{-1} \tag{3.16}$$

and

$$c_3\rho^{-p} \le \det V_{j,k}^{-1/2} \le c_4\rho^{-p}. \tag{3.17}$$

Clearly, we have

$$\tilde{p}_n(s_n(j), s_n(k), x, \bullet) = \det V_{j,k}^{-1/2} f_n\{V_{j,k}^{-1/2}(\bullet - x - \mu_{j,k})\}.$$

We now argue that an Edgeworth expansion holds for f_n. Because of (3.16) and (3.17) this implies the following expansion for $\tilde{p}_n(s_n(j), s_n(k), x, \bullet)$.

$$\begin{aligned}
&\tilde{p}_n(s_n(j), s_n(k), x, \bullet) \\
&= \det V_{j,k}^{-1/2}\Big[\sum_{r=0}^{S-3}(k-j)^{-r/2}P_r(-\phi : \{\bar{\chi}_{\beta,r}\})(V_{j,k}^{-1/2}[\bullet - x - \mu_{j,k}]) \\
&\quad + O([k-j]^{-(S-2)/2}[1 + \|V_{j,k}^{-1/2}(\bullet - x - \mu_{j,k})\|^S]^{-1})\Big]
\end{aligned} \tag{3.18}$$

with standard notations, see Bhattacharya and Rao (1976), p. 53. In particular, P_r denotes a product of a standard normal density with a polynomial that has coefficients depending only on cumulants of order $\le r + 2$.

We now argue that expansion (3.18) follows from Theorem 19.3 in Bhattacharya and Rao (1976). For this claim we have to show that for the sum $\sum_{i=j}^{k-1} \eta_i$ the conditions (19.27), (19.29) and (19.30) of Theorem 19.3 are verified. The matrix B_n on page 194 of Bhattacharya and Rao (1976) corresponds to $B_{j,k} = (k-j)^{1/2} \cdot V_{j,k}$. Now because of (3.16) we have for some constants C_1 and C_2

$$\begin{aligned}
\mathbf{E}\|B_{j,k}\eta_i\|^S &= \mathbf{E}\|(k-j)^{1/2}V_{j,k}\eta_i\|^S \\
&\leq C\Delta_{max}^{-S/2}\mathbf{E}\|\rho V_{j,k}\eta_i\|^S \\
&\leq C\Delta_{max}^{-S/2}\|\rho V_{j,k}\|^S\mathbf{E}\|\eta_i\|^S \\
&\leq C_1\Delta_{max}^{-S/2}\|\mathbf{E}\eta_i\|^S \\
&\leq C_1\int\|z\|^S q(s_n(i), y, z)dz \\
&\leq C_2\int\|z\|^S\psi(z)dz \\
&< \infty,
\end{aligned}$$

see (A2). This shows (19.27).

Next, we show (19.29) with $p = 2$. We write $\varphi_X(u) = \mathbf{E}\exp(iu^TX)$ for the characteristic function of a random variable X. Using the substitution $s = u\cdot(k-j)^{1/2}V_{j,k}^{-1/2}$ we get for a constant C_3

$$\begin{aligned}
\gamma = &\sup_{i=j,\ldots,k-2}\int|\mathbf{E}\exp\{iu^T(k-j)^{1/2}V_{j,k}^{-1/2}\eta_i\}| \\
&\cdot|\mathbf{E}\exp\{iu^T(k-j)^{1/2}V_{j,k}^{-1/2}\eta_{i+1}\}|du \\
&\leq \sup_{i=j,\ldots,k-2}[\det((k-j)^{1/2}V_{j,k}^{-1/2})]^{-1}\sup_{i=j,\ldots,k-2}\int|\varphi_{\eta_i}(s)||\varphi_{\eta_{i+1}}(s)|ds \\
&\leq C_3\Delta_{max}^{p/2}\sup_{i=j,\ldots,k-2}\int|\varphi_{\eta_i}(s)|\,|\varphi_{\eta_{i+1}}(s)|ds \\
&\leq C_3\sup_{i=j,\ldots,k-2}\int|\varphi_{\tilde{\varepsilon}_n(i+1)}(s)||\varphi_{\tilde{\varepsilon}_n(i+2)}(s)|ds \\
&\leq C_3\sup_{i=j,\ldots,k-2}\|\varphi_{\tilde{\varepsilon}_n(i+1)}\|_2\|\varphi_{\tilde{\varepsilon}_n(i+2)}\|_2 \\
&= C_3\sup_{i=j,\ldots,k-2}\|q(s_n(i), y)\|_2\|q(s_n(i+1), y\|_2 \\
&\leq C_3\|\psi\|_2^2 < \infty.
\end{aligned}$$

Hence, (19.29) holds with p = 2.

To check (19.30) we prove that for any $b > 0$

$$\sup_{t\in[0,1],\|u\|>b}|\varphi_{\tilde{\varepsilon}(t)}(u)| < 1 \tag{3.19}$$

where $\tilde{\varepsilon}(t)$ is a random variable with density $q(t, y, \bullet)$. Suppose that this does not hold. Then there exists $b_0 > 0$ with

$$\sup_{t\in[0,1],\|u\|>b_0}|\varphi_{\tilde{\varepsilon}(t)}(u)| = 1.$$

Hence, there exist sequences $\{u_k\}$ and t_k with $|\varphi_{\tilde{\varepsilon}(t_k)}(u_k)| \to 1$. For constant B_0 it must hold that $b_0 \le \|u_k\| \le B_0$ because, under our conditions, $|\varphi_{\tilde{\varepsilon}(t)}(u)| \to 0$, $\|u\| \to \infty$ uniformly in $t \in [0,1]$. Hence there exist $u_0 \ge b_0 \ne 0$ and t_0 such that for subsequences $u_{k_j} \overset{j\to\infty}{\longrightarrow} u_0$ and $t_{k_j} \overset{j\to\infty}{\longrightarrow} t_0$. This implies $\varphi_{\tilde{\varepsilon}(t_0)}(u_0) = 1$. Define $e = \frac{u_0}{\|u_0\|}$ and consider

$$\mathbf{E}\exp\{ise^T\tilde{\varepsilon}(t_0)\} = \mathbf{E}\exp\{ise^T\tilde{\varepsilon}(t_0)\} = \varphi_{Pr\tilde{\varepsilon}(t_0)}(s).$$

$\varphi_{Pr\tilde{\varepsilon}(t_0)}(s)$ is the characteristic function of the projection of $\tilde{\varepsilon}(t_0)$ onto e. We have $\varphi_{Pr\tilde{\varepsilon}(t_0)}(0) = \varphi_{Pr\tilde{\varepsilon}(t_0)}(\|u_0\|) = 1$. This implies that the distribution of this projection is purely discrete. Clearly this is impossible because it contradicts our assumption that $\varepsilon(t_0)$ has a density. Therefore (3.19) holds.

Our setting is slightly different from that of Theorem 19.3 in Bhattacharya and Rao (1976). We consider triangular arrays of independent random vectors instead of a sequence of independent random vectors. But the same proof applies because in our setting the conditions (19.27), (19.29) and (19.30) hold uniformly.

We now argue that, for C large enough it holds that

$$\tilde{p}_n(s_n(j), s_n(k), x, \bullet) \le C\xi_\rho(\bullet - x).$$

For seeing this note that for all c there exists a constant C' with

$$\exp(-c\|z\|^2) \le C'\frac{1}{1+\|z\|^S}.$$

This shows the lemma for $|\nu| = 0$.

For $|\nu| = 1, 2$ one again proceeds simlarly as in the proof of Theorem 19.3 in Bhattacharya and Rao (1976) to obtain Edgeworth expansions for $\mathrm{D}_u^\nu\tilde{p}_n(s_n(j), s_n(k), u, y)$. Note that differentiation D^ν of the density and of the terms of the Edgeworth expansion corresponds to multiplication of their Fourier transforms with t^ν. Hence, after obvious modifications the estimates of Theorem 9.11 and Lemma 14.3 from Bhattacharya and Rao (1976) apply for these derivatives. Then with these bounds one simply has to copy the proof of Theorem 19.3. Proceeding as above one gets (3.15). $\square$

Lemma 3.8. *The following bound holds with a constant C*

$$|\tilde{p}_n(s_n(j), s_n(k), x, y) - \tilde{p}(s_n(j), s_n(k), x, y)| \le C\Delta_{max}^{1/2}\rho^{-1}\zeta_\rho(y-x) \qquad (3.20)$$

for all $j < k$, x and y. Here again ρ denotes the term $\rho = [s_n(k) - s_n(j)]^{1/2}$. We write $\zeta_\rho(\bullet) = \rho^{-p}\zeta(\bullet/\rho)$ where

$$\zeta(z) = \frac{[1+\|z\|^{S-4}]^{-1}}{\int[1+\|z'\|^{S-4}]^{-1}\,dz'}.$$

Proof of Lemma 3.8. It follows from the proof of Lemma 3.7 (see (3.18)) and from Condition (A4) that

$$|\tilde{p}_n(s_n(j), s_n(k), x, y) - \hat{p}_n(s_n(j), s_n(k), x, y)| \\ \leq C\Delta_{max}^{1/2}\rho^{-1}\zeta_\rho(y-x), \tag{3.21}$$

where with $V_{j,k}$ and $\mu_{j,k}$ as in the proof of Lemma 3.7

$$\hat{p}_n(s_n(j), s_n(k), x, y) = \det V_{j,k}^{-1/2}(2\pi)^{-p/2} \\ \times \exp\{-\frac{1}{2}(y-x-\mu_{j,k})^T V_{j,k}^{-1}(y-x-\mu_{j,k})\}.$$

Note that by (A5) we easily get

$$\left\|\mu_{j,k} - m(s_n(j), s_n(k), y)\right\| \leq C\Delta_{\max}\rho^2,$$
$$\left\|V_{j,k} - \Sigma(s_n(j), s_n(k), y)\right\| \leq C\Delta_{\max}\rho^2.$$

This implies

$$|\hat{p}_n(s_n(j), s_n(k), x, y) - \tilde{p}(s_n(j), s_n(k), x, y)| \leq C\Delta_{max}\zeta_\rho(y-x) \\ \leq C\Delta_{max}^{1/2}\rho^{-1}\zeta_\rho(y-x). \tag{3.22}$$

The lemma follows from (3.22) and (3.21). □

In the next lemma we compare the infinitesimal operators L_n and $\tilde{L}_n$ with the differential operators L and $\tilde{L}$. We give an approximation for the error if, in the definition of H_n, the terms L_n and $\tilde{L}_n$ are replaced by L or $\tilde{L}$, respectively.

Lemma 3.9. *The following bound holds with a constant C*

$$|H_n(s_n(j), s_n(k), x, y) - K_n(s_n(j), s_n(k), x, y) - M_n(s_n(j), s_n(k), x, y)| \\ \leq C\Delta_{max}^{1/2}\rho^{-1}\zeta_\rho(y-x) \tag{3.23}$$

with ζ_ρ as in Lemma 3.8 for all $j < k, x$ and y. Here again ρ denotes the term $\rho = [s_n(k) - s_n(j)]^{1/2}$. For $j < k-1$ the function K_n is defined as

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

Furthermore, for $j < k-1$ we define

$$M_n(s_n(j), s_n(k), x, y) = 3\Delta_n(j+1)^{1/2}\sum_{|\nu|=3}\sum_{|\mu|=1}\int_{\mathbb{R}^p}\int_0^1 D_y^\mu q(s_n(j), y, \theta)(x-y)^\mu \\ \frac{\theta^\nu}{\nu!}D_x^\nu \tilde{p}_n(s_n(j+1), s_n(k), x+\delta\theta\Delta_n(j+1)^{1/2}, y)(1-\delta)^2 d\delta d\theta.$$

For $j = k-1$ we define

$$K_n(s_n(j), s_n(k), x, y) = M_n(s_n(j), s_n(k), x, y) = 0.$$

The proof of Lemma 3.9 is based on some lengthy elementary calculations. It is deferred to Subsection 3.5.

Lemma 3.10. *The following bound holds with a constant C*

$$|K_n(s_n(j), s_n(k), x, y)| \le C\rho^{-1}\, \zeta_\rho(y-x), \tag{3.24}$$
$$|H_n(s_n(j), s_n(k), x, y)| \le C\rho^{-1}\, \zeta_\rho(y-x), \tag{3.25}$$
$$|M_n(s_n(j), s_n(k), x, y)| \le C\rho^{-1}\, \zeta_\rho(y-x), \tag{3.26}$$

with ζ_ρ as in Lemma 3.8 for all $j < k$, x and y. Here again, $\rho = [s_n(k) - s_n(j)]^{1/2}$.

Proof of Lemma 3.10. Note first that (3.25) follows from (3.24) with Lemma 3.9 and (A4). Claim (3.26) follows from the fact that $\Delta_{max} \le c\rho$ for a constant c and from simple estimates. It remains to show (3.24). We have that

$$\begin{aligned} |K_n(s_n(j), s_n(k), x, y)| \ &\le |f'(u)^T[m(s_n(j), x) - m(s_n(j), y)]| \\ &\quad + \frac{1}{2}\mathrm{tr}\big\{[\Lambda(s_n(j), x) - \Lambda(s_n(j), y)] \\ &\quad \times f''(x)[\Lambda(s_n(j), x) + \Lambda(s_n(j), y)]\big\}, \end{aligned} \tag{3.27}$$

where $f(x) = \tilde{p}_n(s_n(j+1), s_n(k), x, y)$. It follows from (A2) and (A3) that for C' large enough

$$\|m(s_n(j), x) - m(s_n(j), y)\| \le C'\rho\Big[\frac{\|y-x\|}{\rho} + 1\Big] \tag{3.28}$$

and

$$\|\Lambda(s_n(j), x) - \Lambda(s_n(j), y)\| \le C'\rho\Big[\frac{\|y-x\|}{\rho} + 1\Big]. \tag{3.29}$$

Now the lemma follows from Lemma 3.7, (3.27) – (3.29) and (A4). □

Lemma 3.11. *There exists a constant C_1 (that does not depend on x and y) such that the following inequality holds:*

$$|\tilde{p}_n \otimes_n H_n^{(r)}(s_n(j), s_n(k), x, y)| \le \frac{C_1^{r+1}\rho^r}{\Gamma(1+\frac{r}{2})}\chi_\rho(y-x)$$

for $0 < j < k \le n$, where

$$\chi(z) = \frac{[1 + \|z\|^{2S'-2}]^{-1}}{\int [1 + \|z'\|^{2S'-2}]^{-1}\, dz'}$$

and $\rho = [s_n(k) - s_n(j)]^{1/2}$.

Proof of Lemma 3.11. With the help of Lemmas 3.10 and 3.7 [note that ξ/ζ is bounded] we get

$$
\begin{aligned}
&|\tilde{p}_n \otimes_n H_n(s_n(j), s_n(k), x, z)| \\
&\quad \le \sum_{i=j}^{k-1} \Delta_n(j+1) \int_{IR^p} \tilde{p}_n(s_n(j), s_n(i), x, z')|H_n(s_n(i), s_n(k), z', z)|\, dz' \\
&\quad \le C^2 \sum_{i=j}^{k-1} \Delta_n(i+1)[s_n(k) - s_n(i)]^{-1/2} \zeta^{2,j,k}(z-x),
\end{aligned}
$$

where we put

$$
\begin{aligned}
\zeta^{l,j,k}(x) \ &= \max\{\zeta_{\rho_1} * \ldots * \zeta_{\rho_l}(x) : \rho_1 \ge 0, \ldots, \rho_l \ge 0, \\
&\qquad \rho_1^2 + \ldots + \rho_l^2 = \rho^2\}.
\end{aligned} \tag{3.30}
$$

Here ζ_0 denotes the δ-function. We use now that $\sum_{i=j}^{k-1} \Delta_n(i+1)[s_n(k) - s_n(i)]^{-1/2} \le \int_{s_n(j)}^{s_n(k)} [s_n(k) - v]^{-1/2}\, dv = \rho B(1, \frac{1}{2})$, where $B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\, dt$ is the beta function. We get

$$
|\tilde{p}_n \otimes_n H_n(s_n(j), s_n(k), x, z)| \le C^2 \rho B(1, \frac{1}{2}) \zeta^{2,j,k}(z-x). \tag{3.31}
$$

Using (3.31) and (3.25) we get

$$
\begin{aligned}
&\left|\tilde{p}_n \otimes_n H_n^{(2)}(s_n(j), s_n(k), x, z)\right| \\
&\quad \le \sum_{i=j}^{k-1} \Delta_n(i+1) \int_{IR^p} |\tilde{p}_n \otimes_n H_n(s_n(j), s_n(i), x, z')||H_n(s_n(i), s_n(k), z', z)| dz' \\
&\quad \le C^3 \rho^2 \zeta^{3,j,k}(z-x) B(1, \frac{1}{2}) B(\frac{3}{2}, \frac{1}{2}),
\end{aligned}
$$

where it has been used that $\int_{s_n(j)}^{s_n(k)} [s_n(k) - v]^{1/2} [v - s_n(j)]^{-1/2}\, dv = \rho^2 B(\frac{3}{2}, \frac{1}{2})$. Using iteratively similar bounds we get

$$
\begin{aligned}
&\left|\tilde{p}_n \otimes_n H_n^{(r)}(s_n(j), s_n(k), x, z)\right| \\
&\quad \le C^{r+1} \rho^r \zeta^{r+1,j,k}(z-x) B(1, \frac{1}{2}) B(\frac{3}{2}, \frac{1}{2}) \times \ldots \times B(\frac{r+1}{2}, \frac{1}{2}). \\
&\quad \le C^{r+1} \Gamma(\frac{1}{2})^r \rho^r \zeta^{r+1,j,k}(z-x) \frac{1}{\Gamma(\frac{r}{2}+1)}.
\end{aligned} \tag{3.32}
$$

For the statement of the lemma it suffices to show that $[1 + \|x/\rho\|^{2S'-2}]\rho^p \zeta^{r+1,j,k}(x)$ is bounded by $(C')^{r+1}$ for a constantC'. For this purpose note that due to our choice of S', see Assumption (A2), with constants C_1, C_2

$$
\zeta(x) = \frac{C_1}{1 + \|x\|^{2pS'}} \le C_2 \prod_{i=1}^{p} \lambda(x_i),
$$

where $\lambda(x) = [1 + x^{2S'}]^{-1}\{\int [1 + u^{2S'}]^{-1}\, du\}^{-1}$. This shows that for $\rho_1, ..., \rho_{r+1}$, $\rho_1^2 + ... + \rho_{r+1}^2 = \rho^2$,

$$\zeta_{\rho_1} * ... * \zeta_{\rho_{r+1}}(x) \le C_2^{r+1} \prod_{i=1}^{p} \eta(x_i), \tag{3.33}$$

where $\eta(u) = \lambda_{\rho_1} * ... * \lambda_{\rho_{r+1}}(u)$. Let us denote the Fourier transform of a function γ by $\hat{\gamma}(t) = \int \exp(itu)\gamma(u)du$. Furthermore, here $\|\bullet\|_1$ is the usual L_1−norm in R^1. We will show that

$$\left\|\hat{\eta}^*\right\|_1 \le C_3^{r+1}\rho^{-1}, \tag{3.34}$$

where $\eta^*(u) = [1 + (u/\rho)^{2S'-2}]\eta(u)$ and where C_3 is a constant that does not depend on the special choice of $\rho_1, ..., \rho_{r+1}$. [Note that the function η^* is in $L_1(R^1)$, and that for this reason its Fourier transform is well defined.]

From (3.34) we get by the Fourier Inversion Theorem

$$|\eta(u)| \le \frac{C^{r+1}\rho^{-1}}{1 + (u/\rho)^{2S'-2}}.$$

Because of

$$\prod_{i=1}^{p} \frac{1}{1 + x_i^{2S'-2}} \le \frac{C_4}{1 + \|x\|^{2S'-2}}$$

[with some constant C_4] we therefore get from (3.33) that

$$\zeta^{r+1,j,k}(x) \le C_5^{r+1}\rho^{-p}\frac{1}{1 + \|x/\rho\|^{2S'-2}}$$

with some constant C_5, i.e. (3.34) holds and the lemma is proved.

It remains to show claim (3.34).

Proof of (3.34). Note first that

$$\left\|\hat{\eta}^*\right\|_1 \le \left\|\hat{\eta}\right\|_1 + \frac{1}{\rho^{2S'-2}}\left\|\hat{\eta}^{(2S'-2)}\right\|_1 \tag{3.35}$$

where $\hat{\eta}^{(2S'-2)}$ means the derivative of order $2S' - 2$ of the Fourier transform $\hat{\eta}$ of η. We now show that

$$\left\|\hat{\eta}\right\|_1 \le (r+1)^{1/2}\rho^{-1}\left\|\hat{\lambda}\right\|_1. \tag{3.36}$$

For the proof of claim (3.36) note first that there exists an i_* with $\rho_{i_*}^2 \ge \rho^2/(r+1)$. We get the following inequality:

$$\begin{aligned}\int |\hat{\eta}(t)|dt &\le \int |\hat{\lambda}(t\rho_1)| \cdot \ldots \cdot |\hat{\lambda}(t\rho_l)|dt \\ &\le \int |\hat{\lambda}(t\rho_{i_*})|dt \\ &\le (r+1)^{1/2}\rho^{-1}\int |\hat{\lambda}(t)|dt.\end{aligned}$$

Note now that $\int |\hat{\lambda}(t)|dt$ is bounded, see Lemma 3.10. This shows (3.36).

To estimate $\|\hat{\eta}^{(2S'-2)}\|_1$ note first that $|\hat{\lambda}_{\rho_i}(t)| = |\hat{\lambda}(\rho_i t)| \leq 1$ and

$$|\hat{\lambda}^{(k)}_{\rho_i}(t)| \leq \rho_i^k \int |u|^k \lambda(u)du < \infty$$

for $k = 1, \ldots, 2S' - 2$. Furthermore, for $\sum_{i=1}^{r+1} = 2S' - 2$ we have with some constants C_6 and C_7

$$\begin{aligned}\int |\hat{\lambda}^{(k_1)}_{\rho_1}(t)| \cdot \ldots \cdot |\hat{\lambda}^{(k_{r+1})}_{\rho_{r+1}}(t)|dt &\leq C_6^r \prod_{i \neq i_*} \rho_i^{k_i} \int |\hat{\lambda}^{(k_{i*})}_{\rho_{i*}}(t)|\, dt \\ &\leq C_6^r \rho^{2S'-2-k_{i*}} \int \rho_{i_*}^{k_{i*}} |\hat{\lambda}^{(k_{i*})}(u)|\, \frac{dt}{\rho_{i_*}} \\ &\leq C_7^{r+1} \rho^{2S'-3} \|\hat{\lambda}^{(k_{i*})}\|_1. \end{aligned} \tag{3.37}$$

Using Leibnitz formula for $\eta(u) = \lambda_{\rho_1} * \ldots * \lambda_{\rho_{r+1}}(u)$ we get the following estimate from (3.37) with a constant C_8

$$\left\| \hat{\eta}^{(2S'-2)} \right\|_1 \leq C_8 \rho^{2S'-3} \left(\left\| \hat{\lambda}^{(1)} \right\|_1 + \ldots + \left\| \hat{\lambda}^{(2S'-2)} \right\|_1 \right). \tag{3.38}$$

It is well known that $\left\| \hat{\lambda}^{(q)} \right\|_1$ is uniformly bounded for $q = 0, 1, \ldots, 2S' - 2$, see e.g. Lemma 1 in Gel'fand and Shilov (1958), p. 236. Claim (3.34) now follows from (3.35)–(3.38). □

Lemma 3.12. *For $0 \leq j < k \leq n$ the following formula holds:*

$$p_n(s_n(j), s_n(k), x, y) = \sum_{r=0}^{k-j} (\tilde{p}_n \otimes_n (M_n + K_n)^{(r)})(s_n(j), s_n(k), x, y) + R,$$

where

$$|R| \leq C \Delta_{max}^{\frac{1}{2}} \chi_\rho(y - x)$$

for some constant C. The function χ has been defined in Lemma 3.11. Here again $\rho = [s_n(k) - s_n(j)]^{1/2}$.

Proof of Lemma 3.12. By Lemma 3.6 we have that

$$p_n(s_n(j), s_n(k), x, y) = \sum_{r=0}^{k-j} (\tilde{p}_n \otimes_n H_n^{(r)})(s_n(j), s_n(k), x, y).$$

For $r = 0$ we have that

$$(\tilde{p}_n \otimes_n H_n^{(r)})(s_n(j), s_n(k), x, y) = (\tilde{p}_n \otimes_n (M_n + K_n)^{(r)})(s_n(j), s_n(k), x, y),$$

by definition. For $r = 1$ we have by Lemmas 3.7 and 3.9 that

$$(\tilde{p}_n \otimes_n H_n^{(r)})(s_n(j), s_n(k), x, y) = (\tilde{p}_n \otimes_n (M_n + K_n)^{(r)})(s_n(j), s_n(k), x, y) + R_1,$$

where

$$|R_1| \le \sum_{i=j}^{k-1} \Delta_n(i+1) \int_{\mathbb{R}^p} \tilde{p}_n(s_n(j), s_n(i), x, z) \times |H_n - M_n - K_n|(s_n(i), s_n(k), z, y)dz$$

$$\le C^2 \zeta^{2,j,k}(y-x)\Delta_{max}^{1/2} \sum_{i=j}^{k-1} \Delta_n(i+1)\rho^{-1}, \tag{3.39}$$

where the function $\zeta^{l,j,k}$ was defined in (3.30). For the proof of (3.39) we use Lemma 3.9. We now apply that

$$\sum_{i=j}^{k-1} \Delta_n(i+1)[s_n(k) - s_n(i)]^{-1/2} \le \int_{s_n(j)}^{s_n(k)} [s_n(k) - v]^{-1/2}\, dv = \rho\ B(1, 1/2).$$

Therefore we get from (3.39) that

$$|R_1| \le C^2 \zeta^{2,j,k}(y-x)\Delta_{max}^{\frac{1}{2}} \rho\ B(1, 1/2).$$

With similar arguments we get

$$(\tilde{p}_n \otimes_n H_n^{(2)})(s_n(j), s_n(k), x, y) = (\tilde{p}_n \otimes_n (M_n + K_n)^{(2)})(s_n(j), s_n(k), x, y) + R_2,$$

where

$$|R_2| \le 2C^3 \zeta^{3,j,k}(y-x)\Delta_{max}^{\frac{1}{2}} \rho^2\ B(1, 1/2)B(3/2, 1/2).$$

For arbitrary r it holds that

$$(\tilde{p}_n \otimes_n H_n^{(r)})(s_n(j), s_n(k), x, y) = (\tilde{p}_n \otimes_n (M_n + K_n)^{(r)})(s_n(j), s_n(k), x, y) + R_r,$$

where

$$|R_r| \le C_1^{r+1} \zeta^{r+1,j,k}(y-x)\Delta_{max}^{\frac{1}{2}} \rho^r\ \frac{\Gamma(1/2)^r}{\Gamma([r+3]/2)}.$$

In the proof of Lemma 3.11 we have shown that

$$\zeta^{r+1,j,k}(y-x) \le C^{r+1}\rho^{-p} \frac{1}{1 + \|(y-x)/\rho\|^{2S'-2}}.$$

This gives

$$p_n(s_n(j), s_n(k), x, y) = \sum_{r=0}^{k-j} \tilde{p}_n \otimes_n (M_n + K_n)^{(r)}(s_n(j), s_n(k), x, y) + R,$$

where

$$\left[1+\|(y-x)/\rho\|^{2S'-2}\right]|R| \le \sum_{r=1}^{\infty}\left[1+\|(y-x)/\rho\|^{2S'-2}\right]|R_r|$$
$$\le \Delta_{max}^{\frac{1}{2}}\rho^{-p}\sum_{r=1}^{\infty}\rho^r C_2^r \frac{\Gamma(1/2)^r}{\Gamma([r+3]/2)}.$$

Because this is bounded by $C\Delta_{max}^{\frac{1}{2}}\rho^{-p}$ for some constant C, this shows the statement of the lemma. □

We now show that in the expansion of Lemma 3.12 for p_n the densities $\tilde{p}_n$ can be replaced by the Gaussian densities $\tilde{p}$.

Lemma 3.13. *For* $0 \le j < k \le n$ *the following formula holds:*

$$p_n(s_n(j), s_n(k), x, y) = \sum_{r=0}^{k-j}(\tilde{p}\otimes_n (M_n+K_n)^{(r)})(s_n(j), s_n(k), x, y) + R,$$

where

$$|R| \le C\Delta_{max}^{\frac{1}{2}}\chi_\rho(y-x)$$

for some constant C. *The function* χ *has been defined in Lemma 3.11. Here again* $\rho = [s_n(k) - s_n(j)]^{1/2}$.

Proof of Lemma 3.13. The lemma follows from Lemma 3.12 and

$$\sum_{r=0}^{k-j}([\tilde{p}-\tilde{p}_n]\otimes_n (M_n+K_n)^{(r)})(s_n(j), s_n(k), x, y) \le C'\Delta_{max}^{\frac{1}{2}}\chi_\rho(y-x) \tag{3.40}$$

for some constant C'.

It remains to show (3.40) Mimicking the proof of Lemma 3.11 with Lemma 3.8 instead of Lemma 3.7 we get

$$\left|\left(\hat{p}_n - \tilde{p}_n\right)\otimes_n (M_n+K_n)^{(r)}(s_n(j), s_n(k), x, y)\right|$$
$$\le C^{r+1}\rho^r\Delta_{\max}^{1/2}B(1/2,1/2)B(1,1/2)...B(r/2,1/2)\zeta^{r+1,j,k}(y-x).$$

The lemma follows by application of this bound. □

We now come to the proof of our theorem.

3.4. *Proof of Theorem 1.1*

The main steps of the proof of Theorem 1.1 have been given in Subsections 3.1 – 3.3. We will now prove some technical bounds for the infinite series expansion of transition denities for diffusions. We will show that in this expansion the “convolution” $\otimes$ can be replaced by the “convolution” $\otimes_n$ and that the kernel H can be replaced by the the kernel $M_n + K_n$, see below. This will be done by some careful estimates. First we argue that this will imply the theorem.

From Lemmas 3.1 and 3.2 we get

$$p(s,t,x,y) = \sum_{r=0}^{n} (\tilde{p} \otimes H^{(r)})(s,t,x,y) + O(\Delta_{\max}^{1/2} \exp\left(-\frac{C\,\|y-x\|^2}{t-s}\right)). \tag{3.41}$$

Furthermore, Lemma 3.13 implies that

$$\begin{aligned} p_n(0,1,x,y) = {} & \sum_{r=0}^{n} (\tilde{p} \otimes_n (M_n + K_n)^{(r)})(0,1,x,y) \\ & + O(\Delta_{\max}^{1/2} \frac{1}{1+\|y-x\|^{2S'-2}}). \end{aligned} \tag{3.42}$$

Because of (3.41) and (3.42) for the statement of the theorem it remains to show that

$$\begin{aligned} & \left| \sum_{r=0}^{n} \left(\tilde{p} \otimes H^{(r)}(0,1,x,y) - \tilde{p} \otimes_n (M_n + K_n)^{(r)}(0,1,x,y) \right) \right| \\ & = O(\Delta_{\max}^{1/2} \frac{1}{1+\|y-x\|^{2S'-2}}). \end{aligned} \tag{3.43}$$

For the proof of (3.43) note that

$$\begin{aligned} & \left| \sum_{r=0}^{n} \left[\tilde{p} \otimes H^{(r)}(0,1,x,y) - \tilde{p} \otimes_n (M_n + K_n)^{(r)}(0,1,x,y) \right] \right| \\ & \le \left| \sum_{r=0}^{n} \left[\tilde{p} \otimes H^{(r)}(0,1,x,y) - \tilde{p} \otimes_n H^{(r)}(0,1,x,y) \right] \right| \\ & \quad + \left| \sum_{r=0}^{n} \left[\tilde{p} \otimes_n H^{(r)}(0,1,x,y) - \tilde{p} \otimes_n (M_n + H)^{(r)}(0,1,x,y) \right] \right| \\ & \quad + \left| \sum_{r=0}^{n} \left[\tilde{p} \otimes_n (M_n + H)^{(r)}(0,1,x,y) - \tilde{p} \otimes_n (M_n + K_n)^{(r)}(0,1,x,y) \right] \right| \\ & = T_1 + T_2 + T_3. \end{aligned} \tag{3.44}$$

For T_1, T_2 and T_3 we will show the following estimates

$$T_k = O(\Delta_{\max}^{1/2} \frac{1}{1+\|y-x\|^{2S'-2}}), \tag{3.45}$$

where $k = 1, \ldots, 3$. This shows (3.43). It remains to show (3.45).

Proof of (3.45) *for* $k = 1$. We conjecture that under additional smoothness assumptions on m and Σ it can be shown with the methods of Bally and Talay (1996b) that T_1 is of order $O(\Delta_{\max})$. We have

$$\begin{aligned}
T_1 \leq \sum_{r=1}^{n} \Bigg| & \int_0^1 ds_r \int \Psi_r(0, s_r, x, z) H(s_r, 1, z, y) dz \\
& - \sum_{j=0}^{n-1} \Delta_n(j+1) \int \Psi_r(0, s_n(j), x, z) H(s_n(j), 1, z, y) dz \Bigg| \\
& + \sum_{r=2}^{n} \Bigg| \sum_{j=0}^{n-1} \Delta_n(j+1) \int \left(\Psi_r(0, s_n(j), x, z) - \Psi_r^{\Delta}(0, s_n(j), x, z) \right) \\
& \qquad \times H(s_n(j), 1, z, y) dv \Bigg|
\end{aligned}$$

where

$$\begin{aligned}
\Psi_1(s, t, x, y) &= \tilde{p}(s, t, x, y), \\
\Psi_r(s, t, x, y) &= \tilde{p} \otimes H^{(r-1)}(s, t, x, y), \\
\Psi_1^{\Delta}(0, s_n(j), x, z) &= \tilde{p}(0, s_n(i), x, z), \\
\Psi_r^{\Delta}(0, s_n(j), x, z) &= \sum_{i=0}^{j-1} \Delta_n(i+1) \int \Psi_{r-1}^{\Delta}(0, s_n(i), x, z') \\
& \quad \times H(s_n(i), s_n(j), z', z) dz',
\end{aligned}$$

for $r \geq 2$.

Denote $A_r(0, 0, x, v) = 0$ and

$$\begin{aligned}
A_r(0, s_n(k), x, z) = & \int_0^{s_n(k)} ds_r \int \Psi_r(0, s_r, x, z') H(s_r, s_n(k), z', z) dz' \\
& - \sum_{j=0}^{k-1} \Delta_n(j+1) \int \Psi_r(0, s_n(j), x, z') \\
& \times H(s_n(j), s_n(k), z', z) dz'.
\end{aligned}$$

Then we can rewrite our inequality in the form

$$T_1 \leq \sum_{r=1}^{n} |A_r(0, 1, x, y)| + \sum_{r=2}^{n} \left| \left(\left(\Psi_r - \Psi_r^{\Delta} \right) \otimes_n H \right) (0, 1, x, y) \right|. \tag{3.46}$$

Note that for $r \geq 2$

$$\begin{aligned}
& \Psi_r(0, s_n(j), x, z) - \Psi_r^{\Delta}(0, s_n(j), x, z) \\
& \quad = A_{r-1}(0, s_n(j), x, z) + \left(\left(\Psi_{r-1} - \Psi_{r-1}^{\Delta} \right) \otimes_n H \right) (0, s_n(j), x, z).
\end{aligned} \tag{3.47}$$

We now apply Lemma 3.5 to estimate $A_r(0, s_n(j), x, z)$. Let us consider the function

$$\Lambda_r(t) = \int \Psi_r(0, t, x, z')H(t, s, z', z)dz'.$$

Let $t, t + \Delta t \in (0, s)$. We have by Lemmas 3.3, 3.5 and 3.2 for $\Delta t \geq 0$ with constants $C_1, C_2, \ldots$

$$\begin{aligned}
&|\Lambda_r(t + \Delta t) - \Lambda_r(t)| \\
&\quad = \Delta t \,| \int \int_0^1 \frac{\partial}{\partial t}[\Psi_r(0, t + h\Delta t, x, z')H(t + h\Delta t, s, z', z)]dhdz' \,| \\
&\quad = \Delta t \left| \int_0^1 dh \left[\int H(t + h\Delta t, s, z', z) \frac{\partial}{\partial t} \Psi_r(0, t + h\Delta t, x, z') \right.\right. \\
&\qquad\quad \left.\left. + \Psi_r(0, t + h\Delta t, x, z') \frac{\partial}{\partial t} H(t + h\Delta t, s, z', z)dz' \right] \right| \\
&\quad \leq \Delta t \int_0^1 dh \left\{ \int C_1^r \frac{(t + h\Delta t)^{\frac{r-1}{2} - 1 - \frac{p}{2}}}{\Gamma(1 + \frac{r-1}{2})} \exp\left(-\frac{C_2 \left|z' - x\right|^2}{t + h\Delta t} \right) \right. \\
&\qquad\quad \times (s - t - h\Delta t)^{-\frac{p}{2} - \frac{1}{2}} \exp\left(-\frac{C_3 \left|z - z'\right|^2}{s - t - h\Delta t} \right) \\
&\qquad\quad + C_4^r \frac{(t + h\Delta t)^{\frac{r-1}{2} - \frac{p}{2}}}{\Gamma(1 + \frac{r-1}{2})} \exp\left(-\frac{C_5 \left|z' - x\right|^2}{t + h\Delta t} \right) \\
&\qquad\quad \left. \times C_6 (s - t - h\Delta t)^{-\frac{p}{2} - \frac{3}{2}} \exp\left(-\frac{C_7 \left|z - z'\right|^2}{s - t - h\Delta t} \right) dz' \right\} \\
&\quad \leq \frac{C_8^r \Delta t}{\Gamma(1 + \frac{r-1}{2})} s^{-p/2} \exp\left(-\frac{C_9 \left|z - x\right|^2}{s} \right) \\
&\qquad\quad \times \int_0^1 dh \left((s - t - h\Delta t)^{-\frac{3}{2}} + (t + h\Delta t)^{-\frac{3}{2}} \right).
\end{aligned}$$

This gives

$$\begin{aligned}
&|\Lambda_r(t + \Delta t) - \Lambda_r(t)| \\
&\leq \frac{C_8^r}{\Gamma(1 + \frac{r-1}{2})} s^{-p/2} \exp\left(-\frac{C_9 \left|z - x\right|^2}{s} \right) \left(\frac{\Delta t}{t^{3/2}} + \frac{\Delta t}{(s - t - \Delta t)^{3/2}} \right)
\end{aligned}$$

and hence (with $s = s_n(k)$)

$$\begin{aligned}
&\left| \int_{s_n(j)}^{s_n(j+1)} \Lambda_r(t)dt - \Delta_n(j + 1)\Lambda_r(s_n(j)) \right| \\
&\leq \int_{s_n(j)}^{s_n(j+1)} \max_{t \in [s_n(j), s_n(j+1)]} |\Lambda_r(t) - \Lambda_r(s_n(j))|dt
\end{aligned}$$

$$\leq \frac{C_8^r}{\Gamma(1+\frac{r-1}{2})} s_n^{-p/2}(k) \exp\left(-\frac{C_9\,|z-x|^2}{s_n(k)}\right) \times \left(\frac{\Delta_n^2(j+1)}{s_n^{3/2}(j)} + \frac{\Delta_n^2(j+1)}{(s_n(k)-s_n(j+1))^{3/2}}\right).$$

Suppose now that $s_n(k) \geq 2\Delta_{\max}^{1/2}$. We put

$$B = \left[0,\, \Delta_{\max}^{1/2}\right] \cup [s_n(k) - \Delta_{\max}, s_n(k)]\,,$$
$$B_n = \left\{ j : 0 \leq s_n(j) \leq \Delta_{\max}^{1/2} \text{ or } s_n(k) - \Delta_{\max} \leq s_n(j) \leq s_n(k) \right\}.$$

Then

$$\begin{aligned} &|A_r(0, s_n(k), x, z)| \\ &= \left| \int_0^{s_n(k)} \Lambda_r(t)dt - \sum_{j=0}^{k-1} \Delta_n(j+1)\Lambda_r(s_n(j)) \right| \\ &\leq \int_B |\Lambda_r(t)|dt + \sum_{j\in B_n} \Delta_n(j+1)\,|\Lambda_r(s_n(j))| \\ &\quad + \frac{C_8^r}{\Gamma(1+\frac{r-1}{2})} s_n^{-p/2}(k) \exp\left(-\frac{C_9\,|z-x|^2}{s_n(k)}\right)(S_1+S_2+S_3+S_4)\,, \end{aligned} \tag{3.48}$$

where

$$S_1 = \sum_{\{j:\Delta_{\max}^{1/3} \leq s_n(j) \leq s_n(k)\}} \frac{\Delta_n^2(j+1)}{s_n^{3/2}(j)},$$
$$S_2 = \sum_{\{j:\Delta_{\max}^{1/2} \leq s_n(j) \leq \Delta_{\max}^{1/3}\}} \frac{\Delta_n^2(j+1)}{s_n^{3/2}(j)},$$
$$S_3 = \sum_{\{j:0 \leq s_n(j+1) \leq s_n(k) - \Delta_{\max}^{1/3}\}} \frac{\Delta_n^2(j+1)}{(s_n(k)-s_n(j+1))^{3/2}},$$
$$S_4 = \sum_{\{j:s_n(k)-\Delta_{\max}^{1/3} \leq s_n(j+1) \leq s_n(k)-\Delta_{\max}\}} \frac{\Delta_n^2(j+1)}{(s_n(k)-s_n(j+1))^{3/2}}.$$

We have

$$S_1 \leq \Delta_{\max}^{-1/2}\Delta_{\max} s_n(k) = \Delta_{\max}^{1/2} s_n(k), \tag{3.49}$$
$$S_2 \leq \Delta_{\max}^{-3/4}\Delta_{\max}\Delta_{\max}^{1/3} = o(\Delta_{\max}^{1/2}), \tag{3.50}$$
$$S_3 \leq \Delta_{\max}^{-1/2}\Delta_{\max} s_n(k) = \Delta_{\max}^{1/2} s_n(k), \tag{3.51}$$
$$\begin{aligned} S_4 &= \Delta_{\max} \sum_{\{j:s_n(k)-\Delta_{\max}^{1/3} \leq s_n(j+1) \leq s_n(k)-\Delta_{\max}\}} \frac{\Delta_n(j+1)}{(s_n(k)-s_n(j+1))^{3/2}} \\ &\leq C\Delta_{\max} \int_{s_n(k)-\Delta_{\max}^{1/3}}^{s_n(k)-\Delta_{\max}} (s_n(k)-v)^{-3/2}\,dv \leq C_1\Delta_{\max}^{1/2}. \end{aligned} \tag{3.52}$$

From the estimates of Lemma 3.2 we obtain (remind that now $s_n(k) \geq 2\Delta_{\max}^{1/2}$)

$$\int_B |\Lambda_r(t)|dt \leq \frac{C_1^r}{\Gamma(1+\frac{r-1}{2})} s_n^{-p/2}(k) \exp\left(-\frac{C_2\,|z-x|^2}{s_n(k)}\right)$$
$$\times\left(\int_0^{\Delta_{\max}^{1/2}} (s_n(k)-t)^{-1/2} t^{(r-1)/2}\,dt + \int_{s_n(k)-\Delta_{\max}}^{s_n(k)} (s_n(k)-t)^{-1/2}\,dt\right)$$
$$\leq \frac{C_1^r}{\Gamma(1+\frac{r-1}{2})} s_n^{-p/2}(k) \exp\left(-\frac{C_2\,|z-x|^2}{s_n(k)}\right) \Delta_{\max}^{1/2} s_n(k)^{0\wedge(r-3/2)}, \tag{3.53}$$

$$\sum_{j\in B_n} \Delta_n(j+1)\,|\Lambda_r(s_n(j))|$$
$$\leq \frac{C_1^r}{\Gamma(1+\frac{r-1}{2})} s_n^{-p/2}(k) \exp\left(-\frac{C_2\,|z-x|^2}{s_n(k)}\right)$$
$$\times \sum_{j\in B_n} \Delta_n(j+1) \frac{s_n^{(r-1)/2}(j)}{(s_n(k)-s_n(j))^{1/2}}$$
$$\leq \frac{C_1^r}{\Gamma(1+\frac{r-1}{2})} s_n^{-p/2}(k) \exp\left(-\frac{C_2\,|z-x|^2}{s_n(k)}\right) \sum_{j\in B_n} \Delta_n(j+1)$$
$$\leq \frac{C_3^r}{\Gamma(1+\frac{r-1}{2})} s_n^{-p/2}(k) \exp\left(-\frac{C_2\,|z-x|^2}{s_n(k)}\right) \Delta_{\max}^{1/2} s_n(k)^{0\wedge(r-3/2)}. \tag{3.54}$$

We now get from (3.48)–(3.54) for $r \geq 1$

$$|A_r(0, s_n(k), x, z)|$$
$$\leq \frac{C_3^r}{\Gamma(1+\frac{r-1}{2})} s_n^{-p/2}(k) \exp\left(-\frac{C_2\,|z-x|^2}{s_n(k)}\right) \Delta_{\max}^{1/2} s_n(k)^{-1/2}. \tag{3.55}$$

It follows from the inequalities of Lemma 3.2 that the same estimate (3.55) holds for $s_n(k) \leq 2\Delta_{\max}^{1/2}$. Now, iterative application of (3.46) and (3.47) gives

$$\sum_{r=2}^{n} \left|\left(\left(\Psi_r - \Psi_r^{\Delta}\right) \otimes_n H\right)(0,1,x,y)\right| \leq \sum_{r=1}^{\infty}\sum_{l=1}^{\infty} \left|\left(A_r \otimes_n H^{(l)}\right)(0,1,x,y)\right|. \tag{3.56}$$

From (3.55) just as in Lemma 3.11 we obtain

$$\sum_{r=1}^{\infty}\sum_{l=1}^{\infty} \left|\left(A_r \otimes_n H^{(l)}\right)(0,1,x,y)\right|$$
$$\leq \left(\sum_{r=1}^{\infty} \frac{C_3^r}{\Gamma(1+\frac{r}{2})}\right)\left(\sum_{l=1}^{\infty} \frac{C_4^l}{\Gamma(1+\frac{l}{2})}\right) \exp\left(-C_5(y-x)^2\right) \Delta_{\max}^{1/2}. \tag{3.57}$$

The desired estimate for T_1 follows from (3.46), (3.55), (3.56) and (3.57).

Proof of (3.45) *for* $k = 2$. For $r = 1$ we have

$$\begin{aligned}
&\tilde{p} \otimes_n H^{(r)}(0, s_n(k), x, y) - \tilde{p} \otimes_n (M_n + H)^{(r)}(0, s_n(k), x, y) \\
&= \tilde{p} \otimes_n M_n^{(r)}(0, s_n(k), x, y) \\
&= \sum_{j=0}^{k-1} \Delta_n (j+1)^{3/2} \sum_{|\mu|=1} \sum_{|\nu|=3} a_{\mu,\nu}(j),
\end{aligned}$$

where

$$\begin{aligned}
a_{\mu,\nu}(j) = 3 \int_{I\!R^p} \int_{I\!R^p} \int_0^1 & \tilde{p}(0, s_n(j), x, z) D_y^{\mu} q(s_n(j), y, \theta)(y-z)^{\mu} \\
& \frac{\theta^{\nu}}{\nu!} D_z^{\nu} \tilde{p}_n(s_n(j+1), s_n(k), z + \delta\theta\Delta_n(j+1)^{1/2}, y)(1-\delta)^2 \, d\delta \, d\theta \, dz.
\end{aligned}$$

We consider the index sets $J_1 = \{j \le k : s_n(j) \le s_n(k)/2\}$ and $J_2 = \{j \le k : s_n(j) > s_n(k)/2\}$. For $j \in J_1$ we get the following bound for $a_{\mu,\nu}(j)$ with constants C_1, C_2 and with $\kappa^2 = s_n(k)$, $\lambda^2 = s_n(k) - s_n(j)$

$$\begin{aligned}
|a_{\mu,\nu}(j)| &\le C_1 \int \tilde{p}(0, s_n(j), x, z) \lambda^{-2} \zeta_{\lambda}(y-z) dz \\
&\le C_2 \lambda^{-2} \zeta_{\kappa}(y-x).
\end{aligned}$$

This gives with a constant C_3

$$\begin{aligned}
&\left| \sum_{j \in J_1} \Delta_n (j+1)^{3/2} \sum_{|\mu|=1} \sum_{|\nu|=3} a_{\mu,\nu}(j) \right| \\
&\le C_3 \sum_{j \in J_1} \Delta_n (j+1)^{3/2} [s_n(k) - s_n(j)]^{-1} \zeta_{\kappa}(y-x) \\
&\le C_3 \Delta_{\max}^{1/2} \zeta_{\kappa}(y-x) \int_0^{s_n(k)/2} [s_n(k) - u]^{-1} \, du \\
&\le C_3 \Delta_{\max}^{1/2} \zeta_{\kappa}(y-x) \, [\ln(s_n(k)) - \ln(s_n(k)/2)] \\
&\le C_3 \ln(2) \Delta_{\max}^{1/2} \zeta^{2,0,k}(y-x).
\end{aligned}$$

We now consider $a_{\mu,\nu}(j)$ for $j \in J_2$. Denote the index l with $\mu_l = 1$ by $l(\mu)$. We first consider the case that $\nu_{l(\mu)} < 3$. Then there exists an $l^* \ne l(\mu)$ with $\nu_{l^*} \ge 1$. Define $\nu_l^* = \nu_l$ for $l \ne l^*$ and $\nu_l^* = \nu_l - 1$ for $l = l^*$. By integration by parts we get

$$\begin{aligned}
a_{\mu,\nu}(j) \ = 3 \int_{I\!R^p} \int_{I\!R^p} \int_0^1 & \frac{\partial}{\partial z_{l^*}} \tilde{p}(0, s_n(j), x, z) D_y^{\mu} q(s_n(j), y, \theta)(y-z)^{\mu} \\
& \frac{\theta^{\nu}}{\nu!} D_z^{\nu^*} \tilde{p}_n(s_n(j+1), s_n(k), z + \delta\theta\Delta_n(j+1)^{1/2}, y)(1-\delta)^2 \, d\delta \, d\theta \, dz.
\end{aligned}$$

Using this equation we get the following bound for $a_{\mu,\nu}(j)$ [with $\nu_{l(\mu)} < 3$]

$$|a_{\mu,\nu}(j)| \le C_4 \int \frac{\partial}{\partial z_{l^*}} \tilde{p}(0, s_n(j), x, z) \lambda^{-1} \zeta_{\lambda}(y-z) dz,$$

where C_4 is a constant and where again $\lambda^2 = s_n(k) - s_n(j)$.

By calculating $\partial/\partial z_{l^*} \tilde{p}(0, s_n(j), x, z)$ using the explicite definition (3.1) one can show that

$$|a_{\mu,\nu}(j)| \leq C_5 \iota^{-1} \lambda^{-1} \zeta^{2,0,k}(y - x),$$

where C_5 is a constant and where $\iota^2 = s_n(j)$ and again $\lambda^2 = s_n(k) - s_n(j)$. For a definition of $\zeta^{2,0,k}$ see (3.30). For $a_{\mu,\nu}(j)$ with $\nu_{l(\mu)} = 3$ note that after partial integration $a_{\mu,\nu}(j)$ is of the form

$$\int f(z) z g^{(3)}(z) dz.$$

By integration by parts one gets under conditions on the tails of f and g that

$$\begin{aligned} \int f(z) z g^{(3)}(z) dz &= \int f(z)[(zg(z))^{(3)} - 3g^{(2)}(z)] dz \\ &= -\int f'(z)[(zg(z))^{(2)} - 3g'(z)] dz. \end{aligned}$$

By application of this equality one can show that for a constant C_6

$$|a_{\mu,\nu}(j)| \leq C_6 \left[\iota^{-2} + \iota^{-1} \lambda^{-1} \right] \zeta^{2,0,k}(y - x).$$

Application of these bounds gives for $j \in J_2$ with some constant C_7

$$\begin{aligned} &\left| \sum_{j \in J_2} \Delta_n (j+1)^{3/2} \sum_{|\mu|=1} \sum_{|\nu|=3} a_{\mu,\nu}(j) \right| \\ &\qquad \leq C_7 \Delta_{max}^{1/2} s_n(k)^{-1/2} \zeta^{2,0,k}(y - x). \end{aligned}$$

This gives that for $r = 1$ it holds with some constant C_8

$$\begin{aligned} &\left| \tilde{p} \otimes_n H^{(r)}(0, s_n(k), x, y) - \tilde{p} \otimes_n (M_n + H)^{(r)}(0, s_n(k), x, y) \right| \\ &\quad \leq C_8 \Delta_{\max}^{1/2} s_n(k)^{-1/2} \zeta^{2,0,k}(y - x). \end{aligned}$$

We now claim that for $r \geq 1$ it holds that

$$\begin{aligned} &\left| \tilde{p} \otimes_n H^{(r)}(0, s_n(k), x, y) - \tilde{p} \otimes_n (M_n + H)^{(r)}(0, s_n(k), x, y) \right| \\ &\quad \leq \frac{C_8^r}{\Gamma([r+2]/2)} \Delta_{\max}^{1/2} s_n(k)^{(r-2)/2} \zeta^{r+1,0,k}(y - x). \end{aligned} \tag{3.58}$$

This claim can be proved similarly as for the case $r = 1$. An essential tool is Lemma 3.4. The first statement of this lemma implies the following bound

$$\frac{\partial}{\partial w_l} (\tilde{p} \otimes_n H^{(s)})(0, s_n(k), x, z) \leq \frac{C_9^{s+1} \kappa^{s-1} \zeta^{s+1,0,k}(z - x)}{\Gamma([s+1]/2)}$$

for $s < r$. This inequality can be shown by iterative application of integration by parts. With the help of this inequality and with Lemma 3.10 claim (3.58) follows with similar arguments as in the proof of Lemma 3.12.

Proof of (3.45) *for* $k = 3$. First note that our conditions imply that (formal) differentiation with respect to x up to second order is possible in both sides of (3.18). After calculations similar to the ones presented in the proofs of Lemmas 3.7 and 3.8 this gives

$$\begin{aligned}&|H(s_n(i), s_n(k), x, y) - K_n(s_n(i), s_n(k), x, y)|\\&\leq C\Delta_{\max}^{1/2}(s_n(k) - s_n(i))^{-1/2}\zeta_\rho(y-u).\end{aligned}\tag{3.59}$$

Proceeding as in the proof of Lemma 3.11 we get with a constant C [in the following arguments we will suppose that C is sufficiently large]

$$\begin{aligned}&|\tilde{p}\otimes_n [(H+M_n)-(K_n+M_n)](0, s_n(k), x, y)|\\&\leq \sum_{j=0}^{k-1}\Delta_n(j+1)\int \tilde{p}(0, s_n(j), x, z)(H-K_n)(s_n(j), s_n(k), z, y)dz\\&\leq C^2\Delta_{\max}^{1/2}\sum_{j=0}^{k-1}\Delta_n(j+1)(s_n(k)-s_n(j))^{-1/2}\zeta^{2,0,k}(y-x)\\&\leq C^2\Delta_{\max}^{1/2}s_n^{1/2}(k)B(1,1/2)\zeta^{2,0,k}(y-x).\end{aligned}\tag{3.60}$$

Now

$$\begin{aligned}&\tilde{p}\otimes_n (H+M_n)\otimes_n (H+M_n)(0, s_n(k), x, y)\\&\quad -\tilde{p}\otimes_n (K_n+M_n)\otimes_n (K_n+M_n)(0, s_n(k), x, y)\\&\qquad = (\tilde{p}\otimes_n H - \tilde{p}\otimes_n K_n)\otimes_n (K_n+M_n)(0, s_n(k), x, y)\\&\qquad\quad +\tilde{p}\otimes_n (H+M_n)\otimes_n (H-K_n)(0, s_n(k), x, y)\\&\qquad = I + II.\end{aligned}\tag{3.61}$$

From (3.60) and (3.24) we get

$$\begin{aligned}|I| &\leq C^3\Delta_{\max}^{1/2}B(1,1/2)\sum_{j=0}^{k-1}\Delta_n(j+1)s_n^{1/2}(j)(s_n(k)-s_n(j))^{-1/2}\zeta^{3,0,k}(y-x)\\&\leq C^3\Delta_{\max}^{1/2}B(1,1/2)B(3/2,1/2)s_n(k)\zeta^{3,0,k}(y-x).\end{aligned}\tag{3.62}$$

Proceeding as in the proof of Lemma 3.11 and using Lemma 3.2 instead of Lemma 3.10 we have analogously to (3.32)

$$|II| \leq C^3\Delta_{\max}^{1/2}\Gamma^2(1/2)s_n(k)\zeta^{3,0,k}(y-x).\tag{3.63}$$

From (3.61), (3.62) and (3.63) we get

$$\begin{aligned}&|\tilde{p}\otimes_n (H+M_n)\otimes_n (H+M_n)(0, s_n(k), x, y)\\&\quad -\tilde{p}\otimes_n (K_n+M_n)\otimes_n (K_n+M_n)(0, s_n(k), x, y)|\\&\qquad \leq (2C)^3\Delta_{\max}^{1/2}B(1,1/2)B(3/2,1/2)s_n(k)\zeta^{3,0,k}(y-x).\end{aligned}\tag{3.64}$$

Iterative application of analogous arguments gives

$$\begin{aligned}
&\tilde{p} \otimes_n (H + M_n)^{(r)}(0, s_n(k), x, y) - \tilde{p} \otimes_n (K_n + M_n)^{(r)}(0, s_n(k), x, y) \\
&\quad = \left(\tilde{p} \otimes_n (H + M_n)^{(r-1)} - \tilde{p} \otimes_n (K_n + M_n)^{(r-1)}\right) \\
&\qquad \otimes_n (K_n + M_n)(0, s_n(k), x, y) \\
&\qquad + \tilde{p} \otimes_n (H + M_n)^{(r-1)} \otimes_n (H - K_n)(0, s_n(k), x, y),
\end{aligned} \tag{3.65}$$

where

$$\begin{aligned}
&\left|\tilde{p} \otimes_n (H + M_n)^{(r-1)} \otimes_n (H - K_n)(0, s_n(k), x, y)\right| \\
&\quad \le 2C^{r+2}\Delta_{\max}^{1/2}\Gamma^{r+1}(1/2)s_n^{(r+1)/2}(k)\zeta^{r+2,0,k}(y - x)/\,\Gamma((r + 2)/2)
\end{aligned} \tag{3.66}$$

and

$$\begin{aligned}
&\left|\left(\tilde{p} \otimes_n (H + M_n)^{(r-1)} - \tilde{p} \otimes_n (K_n + M_n)^{(r-1)}\right) \otimes_n (M_n + K_n)(0, s_n(k), x, y)\right| \\
&\quad \le 2^r C^{r+1}\Delta_{\max}^{1/2} B(1, 1/2)\ldots B((r + 1)/2, 1/2)s_n^{r/2}(k)\zeta^{r+1,0,k}(y - x).
\end{aligned} \tag{3.67}$$

Claim (3.45) follows from (3.65) – (3.67). □

3.5. *Additional proofs*

Proof of Lemma 3.5. It suffices to prove (3.13) for $r = 1$ and the following recursion formula for $r \ge 1$

$$\begin{aligned}
\frac{\partial}{\partial t}\tilde{p} \otimes H^{(r)}(s, t, x, y) \;&= \int_s^t du \int \frac{\partial}{\partial u}\left[\tilde{p} \otimes H^{(r-1)}(s, u, x, z)\right] \cdot H(u, t, z, y)dz \\
&\quad + \delta_{1,r} H(s, t, x, y) + R_r(s, t, x, y),
\end{aligned} \tag{3.68}$$

where $\delta_{1,r}$ is Kronecker's delta [i.e. $\delta_{1,1} = 1$ and $\delta_{1,r} = 0$ for $r \neq 1$] and where for some constants C_1' and C_2'

$$|R_r(s, t, x, y)| \;\le\; \frac{[C_1']^r}{\Gamma(1 + \frac{r}{2})}\rho^r \phi_{C_2',\rho}(y - x). \tag{3.69}$$

These claims imply the statement of the lemma: iterating (3.68) we get (3.13). We now prove (3.68). From (3.12) we have for $u \in (s, t)$ and $r \ge 1$

$$\begin{aligned}
&\frac{\partial}{\partial t}\left(\int \tilde{p} \otimes H^{(r-1)}(s, u, x, z) \cdot H(u, t, z, y)dz\right) \\
&\quad = \int \tilde{p} \otimes H^{(r-1)}(s, u, x, z) \cdot \frac{\partial}{\partial t}H(u, t, z, y)dz \\
&\quad = -\int \tilde{p} \otimes H^{(r-1)}(s, u, x, z) \cdot \frac{\partial}{\partial u}H(u, t, z, y)dz + R_r(s, u, t, x, y),
\end{aligned} \tag{3.70}$$

where

$$|R_r(s, u, t, x, y)| \le \frac{C_1^r (u - s)^{\frac{r-1}{2}} \cdot (t - u)^{-1/2}}{\Gamma(\frac{1+r}{2})}\phi_{C_2,\rho}(y - x).$$

Note now that

$$
\begin{aligned}
&\frac{\partial}{\partial u}\int \tilde p \otimes H^{(r-1)}(s,u,x,z)\cdot H(u,t,z,y)dz \\
&\quad = \int \frac{\partial}{\partial u}\left[\tilde p \otimes H^{(r-1)}(s,u,x,z)\right]\cdot H(u,t,z,y)dz \\
&\quad\quad + \int \tilde p \otimes H^{(r-1)}(s,u,x,z)\cdot \frac{\partial}{\partial u}H(u,t,z,y)dz. \qquad (3.71)
\end{aligned}
$$

Comparing (3.70) and (3.71) we get

$$
\begin{aligned}
&\frac{\partial}{\partial t}\int \tilde p \otimes H^{(r-1)}(s,u,x,z)\cdot H(u,t,z,y)dz \\
&\quad + \frac{\partial}{\partial u}\int \tilde p \otimes H^{(r-1)}(s,u,x,z)\cdot H(u,t,z,y)dz \\
&\quad = \int \frac{\partial}{\partial u}\left[\tilde p \otimes H^{(r-1)}(s,u,x,z)\right]\cdot H(u,t,z,y)dz + R_r(s,u,t,x,y).
\end{aligned}
\qquad (3.72)
$$

Integrating (3.70) in u we have from (3.72)

$$
\begin{aligned}
&\int_s^t du \frac{\partial}{\partial t}\left(\int \tilde p \otimes H^{(r-1)}(s,u,x,z)\cdot H(u,t,z,y)dz\right) \\
&\quad = \int_s^t du \int \frac{\partial}{\partial u}\left[\tilde p \otimes H^{(r-1)}(s,u,x,z)\right]\cdot H(u,t,z,y)dz \\
&\quad\quad - \int \tilde p \otimes H^{(r-1)}(s,u,x,z)H(u,t,z,y)dz \,\Big|_{u=s}^{u=t} + R_r(s,t,x,y), \qquad (3.73)
\end{aligned}
$$

where $R_r(s,t,x,y)$ satisfies (3.69). Now (3.68) immediately follows from (3.73) if we take into account that

$$
\int \tilde p \otimes H^{(r-1)}(s,u,x,z)H(u,t,z,y)dz \,\big|_{u=s} = \delta_{1,r}H(s,t,x,y)
$$

and

$$
\begin{aligned}
\frac{\partial}{\partial t}\tilde p \otimes H^{(r)}(s,t,x,y) &= \int_s^t du \frac{\partial}{\partial t}\left(\int \tilde p \otimes H^{(r-1)}(s,u,x,z)\cdot H(u,t,z,y)dz\right) \\
&\quad + \int \tilde p \otimes H^{(r-1)}(s,u,x,z)\cdot H(u,t,z,y)dz \,|_{u=t} .
\end{aligned}
$$

For the statement of the lemma it remains to show that (3.13) holds for $r = 1$. This follows from (3.68) and from arguments that are very similar to those in Ladyženskaja, Solonnikov and Ural´ceva (1968) [p. 378, formula (13.5) with $n = p, s = 0, r = 1, \alpha = 1$.] So we omit the details. This completes the proof. □

Proof of Lemma 3.9. For $j = k - 1$ note that $H_n(s_n(j), s_n(k), x, y) = 0$. So it remains to consider the case $j < k - 1$. First note that [see (3.14)]

$$
H_n(s_n(j), s_n(k), x, y) = H_n^1(s_n(j), s_n(k), x, y) - H_n^2(s_n(j), s_n(k), x, y), \qquad (3.74)
$$

where

$$H_n^1(s_n(j), s_n(k), x, y)$$
$$= \Delta_n(j+1)^{-1} \int p_{n,j}(x, z)$$
$$[\tilde{p}_n(s_n(j+1), s_n(k), z, y) - \tilde{p}_n(s_n(j+1), s_n(k), x, y)]dz \qquad (3.75)$$

and

$$H_n^2(s_n(j), s_n(k), x, y)$$
$$= \Delta_n(j+1)^{-1} \int \tilde{p}_{n,j}^{y}(x, z)$$
$$[\tilde{p}_n(s_n(j+1), s_n(k), z, y) - \tilde{p}_n(s_n(j+1), s_n(k), x, y)]dz. \qquad (3.76)$$

On the right hand side of (3.75) we now use the substitution $\theta = \Delta_n(j+1)^{-1/2}(z-x) - \Delta_n(j+1)^{1/2} m\{s_n(j), x\}$. With the notation $\lambda(z) = \tilde{p}_n(s_n(j+1), s_n(k), z, y)$ and $h(\theta) = m(s_n(j), x)\Delta_n(j+1) + \theta\Delta_n(j+1)^{1/2}$ this gives

$$H_n^1(s_n(j), s_n(k), x, y) = \Delta_n(j+1)^{-1} \int q(s_n(j), x, \theta)[\lambda\{x + h(\theta)\} - \lambda(x)]d\theta.$$

Remind that $q(s_n(j), x, \bullet)$ denotes the conditional density of $\varepsilon_n(j+1)$. We now use the expansion

$$\lambda\{x + h(\theta)\} - \lambda(x) = \sum_{1 \le |\nu| \le 2} \frac{h(\theta)^\nu}{\nu!}(\mathrm{D}^\nu\lambda)(x)$$
$$+3 \sum_{|\nu|=3} \frac{h(\theta)^\nu}{\nu!} \int_0^1 (1-\delta)^2 (\mathrm{D}^\nu\lambda)\{x + \delta h(\theta)\}\, d\delta.$$

Using now that $\varepsilon_n(j)$ has conditional mean 0 we get that

$$H_n^1(s_n(j), s_n(k), x, y)$$
$$= \lambda'(x)^T m(s_n(j), x) + \frac{1}{2}\mathrm{tr}[\Sigma\{s_n(j), x\}\, \lambda''(x)]$$
$$+\Delta_n(j+1) \sum_{|\nu|=2} \frac{m(s_n(j), x)^\nu}{\nu!}(\mathrm{D}^\nu\lambda)(x) + 3 \sum_{|\nu|=3} \Delta_n(j+1)^{-1}$$
$$\times \int \int_0^1 q(s_n(j), x, \theta)\frac{h(\theta)^\nu}{\nu!}(1-\delta)^2(\mathrm{D}^\nu\lambda)\{x + \delta h(\theta)\} d\delta\, d\theta. \qquad (3.77)$$

Note that the first two terms on the right hand side of (3.77) are equal to $L\tilde{p}_n(s_n(j+1), s_n(k), x, y)$.

We now treat the term $H_n^2(s_n(j), s_n(k), x, y)$. On the right hand side of (3.76) we use the substitution $\theta = \Delta_n(j+1)^{-1/2}(z-x) - \Delta_n(j+1)^{1/2} m\{s_n(j), y\}$.

With the notation $\tilde{h}(\theta) = m(s_n(j), y)\Delta_n(j+1) + \theta\Delta_n(j+1)^{1/2}$ this gives

$$\begin{aligned}
&H_n^2(s_n(j), s_n(k), x, y) \\
&\quad = \tilde{L}\tilde{p}_n(s_n(j), s_n(k), x, y) + \Delta_n(j+1)\sum_{|\nu|=2}\frac{m(s_n(j), y)^\nu}{\nu!}(\mathrm{D}^\nu\lambda)(x) \\
&\qquad +3\sum_{|\nu|=3}\Delta_n(j+1)^{-1}\int\int_0^1 q(s_n(j), y, \theta)\frac{\tilde{h}(\theta)^\nu}{\nu!}(1-\delta)^2(\mathrm{D}^\nu\lambda)\{x+\delta\tilde{h}(\theta)\}d\delta\, d\theta.
\end{aligned} \tag{3.78}$$

It remains to show that there exists a constant C with

$$\Delta_n(j+1)|m(s_n(j), x)^\nu - m(s_n(j), y)^\nu|\,|(\mathrm{D}^\nu\lambda)(x)| \le C\Delta_{max}\rho^{-1}\zeta_\rho(y-x) \tag{3.79}$$

for ν with $|\nu| = 2$ and

$$\begin{aligned}
&\Bigg|\Delta_n(j+1)^{-1}\sum_{|\nu|=3}\frac{3}{\nu!}\int\int_0^1 \Big[q(s_n(j), x, \theta)h(\theta)^\nu(\mathrm{D}^\nu\lambda)\{x+\delta h(\theta)\} \\
&\qquad - q(s_n(j), y, \theta)\tilde{h}(\theta)^\nu(\mathrm{D}^\nu\lambda)\{x+\delta\tilde{h}(\theta)\}\Big](1-\delta)^2\, d\delta\, d\theta \\
&\qquad - M_n(s_n(j), s_n(k), x, y)\Bigg| \le C\Delta_{\max}^{1/2}\rho^{-1}\zeta_\rho(y-x).
\end{aligned} \tag{3.80}$$

Proof of (3.79). Because of assumption (A3) we have that for a constant C it holds that $|m(s_n(j), x)^\nu - m_n(s_n(j), y)^\nu| \le C\|x-y\|$. Claim (3.79) follows from Lemma 3.7, monotonicity of $\zeta(x)$ and (A4).

Proof of (3.80). Note that for $|\nu| = 3$

$$\begin{aligned}
\max\{|\tilde{h}(\theta)^\nu|, |h(\theta)^\nu|\} &\le C\Delta_n^{\frac{3}{2}}(j+1)\,(1+\|\theta\|)^3, \\
|\tilde{h}(\theta)^\nu - h(\theta)^\nu| &\le C\Delta_n^2(j+1)\,(1+\|\theta\|)^2\,\|x-y\|.
\end{aligned}$$

So the left hand side of (3.80) does not exceed the following sum

$$\begin{aligned}
&C\Delta_n^{\frac{1}{2}}(j+1)\sum_{|\nu|=3}\int \|x-y\|^2\psi(\theta)\,(1+\|\theta\|)^3\,|(D^\nu\lambda)\{x+\delta h(\theta)\}|d\theta \\
&+C\Delta_n(j+1)\sum_{|\nu|=3}\int \|x-y\|\psi(\theta)\,(1+\|\theta\|)^2\,|(D^\nu\lambda)\{x+\delta h(\theta)\}|d\theta \\
&+C\Delta_n^{\frac{1}{2}}(j+1)\sum_{|\nu|=3}\int \psi(\theta)\,(1+\|\theta\|)^3\,|(D^\nu\lambda)\{x+\delta h(\theta)\} \\
&-(D^\nu\lambda)\{x+\delta\tilde{h}(\theta)\}|d\theta.
\end{aligned} \tag{3.81}$$

We now use the following simple estimate. For an $\varepsilon > 0$ suppose that $\|v\| \le \varepsilon$. Then

$$\frac{1}{1+\|u+v\|^s} \le \frac{1}{1+[\|u\|-\varepsilon]^s} \le \frac{1}{1+[\frac{\|u\|}{2}]^s} \le \frac{2^s}{1+\|u\|^s}$$

for $\|u\| \ge 2\varepsilon$ and

$$\frac{1}{1+\|u+v\|^s} \le 1 \le \frac{(2\varepsilon)^s+1}{1+\|u\|^s}$$

for $\|u\| < 2\varepsilon$. Hence,

$$\frac{1}{1+\|u+v\|^s} \le \frac{C(s,\varepsilon)}{1+\|u\|^s} \tag{3.82}$$

with $C(s,\varepsilon) = \max\{2^s, (2\varepsilon)^s + 1\}$ for all u.

From assumptions (A2), (A4), (3.16), (3.17) and (3.18) it follows that for $|\nu| = 3$

$$|(D^\nu\lambda)\{x+\delta h(\theta)\}| \le c\rho^{-p-3}\Big[1+\|\frac{y-x-\delta m(s_n(j),x)\Delta_n(j+1)-\theta\delta\Delta_n(j+1)^{\frac{1}{2}}}{\rho}\|^s\Big]^{-1}.$$

Similarly we get that

$$|(D^\nu\lambda)\{x+\delta\tilde{h}(\theta)\}| \le c\rho^{-p-3}\Big[1+\|\frac{y-x-\delta m(s_n(j),y)\Delta_n(j+1)-\theta\delta\Delta_n(j+1)^{\frac{1}{2}}}{\rho}\|^s\Big]^{-1}.$$

Applying (3.82) with $v = [\delta m(s_n(j),z)\Delta_n(j+1)+\theta\delta\Delta_n(j+1)^{\frac{1}{2}}]/\rho$, $z = x$ or y, and $\varepsilon = C\Delta_n(j+1)^{\frac{1}{2}} + \|\theta\|$ we get [note that $\|v\| \le \varepsilon$] for $|\nu| = 3$ with a constant $C(s)$ depending on s

$$\max\{|(D^\nu\lambda)\{x+\delta h(\theta)\}|, |(D^\nu\lambda)\{x+\delta\tilde{h}(\theta)\}|\} \le c\rho^{-p-3}\frac{C(s)(1+\|\theta\|^s)}{1+\|\frac{y-x}{\rho}\|^s}. \tag{3.83}$$

Note now that for ν with $|\nu| = 4$ and for κ with $|\kappa| \le 1$ we have [because of $|\delta h(\theta) + \kappa\delta(h(\theta)-\tilde{h}(\theta))| \le C\Delta_n(j+1) + \|\theta\|\Delta_n(j+1)^{\frac{1}{2}}$]

$$|(D^\nu\lambda)\{x+\delta h(\theta)+\kappa\delta(h(\theta)-\tilde{h}(\theta))\}| \le c\rho^{-p-4}\frac{C(s)(1+\|\theta\|^s)}{1+\|\frac{y-x}{\rho}\|^s}. \tag{3.84}$$

Furthermore we get for the difference in the integrand of the third term in (3.81) that

$$|(D^\nu\lambda)\{x+\delta h(\theta)\} - (D^\nu\lambda)\{x+\delta\tilde{h}(\theta)\}| \le c\rho^{-p-4}\Delta_n(j+1)\|x-y\|\frac{C(s)(1+\|\theta\|^s)}{1+\|\frac{y-x}{\rho}\|^s}. \tag{3.85}$$

Substituting (3.83), (3.85) into (3.81) and taking $s = S - 3$ (see (A2)) we get that the left hand side of (3.80) does not exceed

$$C\Delta_{\max}^{\frac{1}{2}}\rho^{-1}\zeta_\rho(y-x). \qquad \square$$

Acknowledgements. We would like to thank three referees for a careful reading of the paper and for remarks that helped to essentially improve the exposition of the material in the paper.

References

Bally, V., Talay, D.: The law of the Euler scheme for stochastic differential equations: I. Convergence rate of the distribution function. Probab. Theory Rel. Fields **104**, 43–60 (1996a)

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 (1996b)

Bhattacharya, R., Rao, R.: Normal approximations and asymptotic expansions. John Wiley & Sons, New York (1976)

Dahlhaus, R.: Fitting time series models to nonstationary processes. Ann. Statist. **25**, 1–37 (1997)

Davydov, Y.A.: On strong convergence of distributions of functionals of random processes. I. Theory Probab. Appl. **25**, 772–789 (1980)

Davydov, Y.A.: On strong convergence of distributions of functionals of random processes. II. Theory Probab. Appl. **26**, 258–278 (1981)

Franke J., Kreiss J.- P., Mammen, E.: Bootstrap of kernel smoothing in nonlinear time series. Manuscript. (1996)

Friedman, A.: Partial differential equations of parabolic type. Prentice-Hall, Englewood Cliffs, New Jersey (1964)

Gel'fand, I.M., Shilov, G.E.: *Generalized functions, vol.2.* Fizmatgiz, Moscow (in russian) (1958)

Kasymdzganova, M.: Local limit theorems for homogenieus Markov chains. Ph.D. Thesis. Moscow Statw Univ. (in russian) (1981)

Kloeden, P.E., Platen, E.: *Numerical solution of stochastic differential equations.* Springer, Berlin, Heidelberg (1992)

Konakov, V.D., Mammen, E.: Local approximations of Markov random walks by diffusions. Manuscript (1999)

Konakov, V.D., Molchanov, S.A.: On the convergence of Markov chains to diffusion processes. Teor. Veroyatn. Mat. Stat.**31**, 51–64, (in russian) [English translation in Theory Probab. Math. Stat. (1985) **31**, 59–73] (1984)

Konovalov, A.G.: A local limit theorem on convergence to a diffusion process. Theory Probab. Appl., **26**, 581–587 (1981)

Kutoyants, Y.A.: Efficiency of the empirical distribution for ergodic diffusion. Bernoulli **3**, 445–456 (1997a)

Kutoyants, Y.A.: Some problems of nonparametric estimation by observations of ergodic diffusion process. Statist. Probab. Letters **32**, 311–320 (1997b)

Kuznetsov, S.E.: Polar boundary sets for superdiffusions and removable lateral singularities for nonlinear parabolic PDEs. Comm. Pure Appl. Math. **51**, 303–340 (1998)

Ladyženskaja, O.A., Solonnikov, V.A. and and Ural´ceva: *Linear and quasi-linear equations of parabolic type.* Amer. Math. Soc., Providence, Rhode Island (1968)

McKean, H.P., Singer, I.M.: Curvature and the eigenvalues of the Laplacian. J. Differential Geometry **1**, 43–69 (1967)

Pinsky, M.A. *Lectures on random evolution.* World Scientific, Singapore, New Jersey (1991)

Platen, E.: An introduction to numerical methods for stochastic differential equations. Research Paper 6, UTS, Sidney (1999)

Robinson, P.M.: Nonparametric estimators for time series. J. Time Ser. Anal. **4**, 185–207 (1983)

Stroock, D.W., Varadhan, S. R.: Multidimensional diffusion processes. Springer, Berlin, Heidelberg, New York (1979)

Skorohod, A.V.: Studies in the theory of random processes. Addison-Wesley, Reading, Massachussetts.[English translation of Skorohod, A. V. (1961). Issledovaniya po teorii sluchaynykh protsessov. Kiev University Press] (1965)
Skorohod, A.V. Asymptotic methods for Stochastic differential equations. (in russian). Kiev Naukova dumka (1987)
Tjøstheim, D. Non-linear time series: a selective review. Scand. J. Statist. **21**, 97–130 (1994)