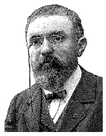
Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2010, Vol. 46, No. 4, 908–923
DOI: 10.1214/09-AIHP207
© Association des Publications de l'Institut Henri Poincaré, 2010



ANNALES
DE L'INSTITUT
HENRI
POINCARÉ
PROBABILITÉS
ET STATISTIQUES

www.imstat.org/aihp

Explicit parametrix and local limit theorems for some degenerate diffusion processes

Valentin Konakov[a], Stéphane Menozzi[b] and Stanislav Molchanov[c]

[a]*Central Economics Mathematical Institute, Academy of Sciences, Nahimovskii av. 47, 117418 Moscow, Russia.*
E-mail: valentin_konakov@yahoo.com
[b]*Laboratoire de Probabilités et Modèles Aléatoires, Université Paris VII, 175 Rue du Chevaleret, 75013 Paris, France.*
E-mail: menozzi@math.jussieu.fr
[c]*Department of Mathematics, University of North Carolina at Charlotte, Charlotte, USA.*
E-mail: smolchan@uncc.edu

Received 15 February 2009; revised 23 February 2009; accepted 2 July 2009

Abstract. For a class of degenerate diffusion processes of rank 2, i.e. when only Poisson brackets of order one are needed to span the whole space, we obtain a parametrix representation of McKean–Singer [*J. Differential Geom.* **1** (1967) 43–69] type for the density. We therefrom derive an explicit Gaussian upper bound and a partial lower bound that characterize the additional singularity induced by the degeneracy.

This particular representation then allows to give a local limit theorem with the usual convergence rate for an associated Markov chain approximation. The key point is that the "weak" degeneracy allows to exploit the techniques first introduced in Konakov and Molchanov [*Teor. Veroyatn. Mat. Statist.* **31** (1984) 51–64] and then developed in [*Probab. Theory Related Fields* **117** (2000) 551–587] that rely on Gaussian approximations.

Résumé. Pour une classe de processus de diffusion de rang deux, i.e. lorsque seuls des crochets de Poisson d'ordre un permettent d'engendrer l'espace, nous obtenons une représentation parametrix de type McMean–Singer [*J. Differential Geom.* **1** (1967) 43–69] de la densité. Nous en dérivons une borne supérieure Gaussienne explicite et une borne inférieure partielle qui caractérisent la singularité additionnelle induite par la dégénérescence.

Nous donnons ensuite un théorème limite local pour une approximation par chaîne de Markov associée. Le point crucial est que la faible dégénérescence permet d'exploiter les techniques initialement introduites par Konakov et Molchanov [*Teor. Veroyatn. Mat. Statist.* **31** (1984) 51–64] puis développées dans [*Probab. Theory Related Fields* **117** (2000) 551–587] et qui reposent sur des approximations Gaussiennes.

MSC: Primary 60J35; 60J60; secondary 35K65

Keywords: Degenerate diffusion processes; Parametrix; Markov chain approximation; Local limit theorems

1. Introduction

1.1. *Global overview*

Let us consider in $\mathbb{R}^d$, $d \geq 1$, the Markov diffusion process with generator

$$L = \frac{1}{2} \sum_{i,j \in [\![1,d]\!]^2} a_{ij}(x)\, \partial^2_{x_i x_j} + \sum_{i \in [\![1,d]\!]} b_i(x)\, \partial_{x_i}.$$

If the coefficients of L are smooth enough, say $C^1(\mathbb{R}^d)$, bounded, and the diffusion matrix $A(x) = (a_{ij}(x))$ is uniformly elliptic ($\forall \lambda \in \mathbb{R}^d$, $\langle A\lambda, \lambda\rangle \in [\delta|\lambda|^2, \delta^{-1}|\lambda|^2]$ for an appropriate $\delta > 0$) then the associated process $(X_t)_{t\ge 0}$ has a transition density $p(t, x, y)$ which is the fundamental solution of the parabolic problem $\partial_t p(\cdot) = L_x p(\cdot)$, $p(0, x, y) = \delta_y(x)$. Of course, one also has $\partial_t p(\cdot) = L_y^* p(\cdot)$, $p(0, x, y) = \delta_x(y)$.

Moreover, this density satisfies uniformly in $t \in]0, T]$ the following Gaussian bounds

$$\frac{M^{-1}}{t^{d/2}} \exp\left(-M\frac{|x-y|^2}{t}\right) \le p(t, x, y) \le \frac{M}{t^{d/2}} \exp\left(-\frac{|x-y|^2}{Mt}\right),$$

where the constant M depends on T, d, the ellipticity constant and the norms of the coefficients in $C^1(\mathbb{R}^d)$, see e.g. Aronson [1] or Stroock [24].

The above estimations express the following physically obvious fact: if the process starts from $x_0 \in \mathbb{R}^d$, then for small $t > 0$, in the neighborhood of x_0 it is “almost Gaussian” with the “frozen” diffusion tensor $A(x_0)$ and the drift $b(x_0)$.

The justification of this fact requires the solution of the perturbative integral equation for $p(\cdot)$ (so-called *Parametrix equation*), where the leading term of the perturbation theory for $p(\cdot)$ is *exactly* the Gaussian kernel $p_0(\cdot)$ corresponding to the “frozen” coefficients at x_0. For details concerning *Parametrix equations* we refer the reader to McKean and Singer [19], Friedman [11] or [14].

If the matrix $A(x)$ degenerates, but the coefficients a, b are still smooth, the diffusion process $(X_t)_{t\ge 0}$ with generator L exists (one can use the Itô calculus for the direct construction of the trajectories), but has generally speaking no density.

Consider now generators of the form $L = \sum_{i=1}^k A_i^2 + A_0$, $k < d$, where $(A_i)_{i\in[\![0,k]\!]}$ are first order operators (vector fields) on $\mathbb{R}^d$ (or more generally on smooth manifolds) with C^∞ coefficients. Sufficient conditions for the existence of the density can be formulated in terms of the structure of the Lie algebra of the vector fields on $\mathbb{R}^d$, with usual linear operations and the Poisson bracketing $[\cdot,\cdot]$. Namely, if the vector fields $A_1, \dots, A_k$, $[A_l, A_m]_{(l,m)\in[\![0,k]\!]^2}$, $[A_l, [A_m, A_n]]_{(l,m,n)\in[\![0,k]\!]^3}, \dots$ span $\mathbb{R}^d$ then the density exists. This result is due to Hörmander [13], see also Norris [22] for a Malliavin calculus based probabilistic proof. Operators having the previous property are said to be hypoelliptic. Also, in [13], Hörmander stressed that the seed of the idea of hypoellipticity goes back to Kolmogorov’s note [17].

A. Kolmogorov made the following important observation. Let $d = 2$. For the generator $L = \frac{1}{2}\partial^2_{xx} + ax\,\partial_y$, $a \ne 0$, the solution of the associated SDE writes $(X_t, Y_t) = (x_0 + W_t, y_0 + a(x_0 t + \int_0^t W_s\,\mathrm{d}s))$, where W is a standard one-dimensional Brownian motion. Thus (X_t, Y_t) has two-dimensional Gaussian distribution with mean $(x_0, y_0 + ax_0 t)$ and covariance matrix $C = \begin{pmatrix} t & \frac{at^2}{2} \\ \frac{at^2}{2} & \frac{a^2t^3}{3}\end{pmatrix}$. Note that the transition density for small t has higher singularity than the usual heat kernel. In Hörmander’s form $L = \frac{1}{2}A_1^2 + A_0$, $A_1 = \partial_x$, $A_0 = ax\,\partial_y$ so that $[A_1, A_0] = a\,\partial_y$ and thus, $A_1, [A_1, A_0]$ have together rank 2.

In this paper, using a parametrix approach derived from the work of McKean and Singer [19], we are able to derive a Gaussian upper bound, and a “partial” lower bound with the two previous time scales, and an associated local limit theorem in the following case.

1.2. *Statement of the problem*

We consider $\mathbb{R}^d \times \mathbb{R}^d$-valued diffusion processes that follow the dynamics

$$\begin{cases} X_t = x + \int_0^t b(X_s, Y_s)\,\mathrm{d}s + \int_0^t \sigma(X_s, Y_s)\,\mathrm{d}W_s, \\ Y_t = y + \int_0^t X_s\,\mathrm{d}s, \end{cases} \tag{1.1}$$

where $(W_t)_{t\ge 0}$ is a standard d-dimensional Brownian motion defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$ satisfying the usual assumptions. We assume that σ is uniformly elliptic and that b, σ are C^1, bounded, Lipschitz continuous functions so that there exists a unique strong solution to (1.1).

Such kind of processes appear in various applicative fields. For instance, in mathematical finance, when dealing with Asian options, X represents the dynamics of the underlying asset and its integral Y is involved in the option Pay-off. Typically, the price of such options write $\mathbb{E}_x[\psi(X_T, T^{-1}Y_T)]$, where for the *put* (resp. *call*) option the function $\psi(x, y) = (x - y)^+$ (resp. $(y - x)^+$), see [5] and [26].

The cross dependence of the dynamics of X in Y is also important when handling kinematic models or Hamiltonian systems. For a given Hamilton function of the form $H(x, y) = V(y) + \frac{|x|^2}{2}$, where V is a potential and $\frac{|x|^2}{2}$ the kinetic energy of a particle with unit mass, the associated stochastic Hamiltonian system would correspond to $b(X_s, Y_s) = -(\partial_y V(Y_s) + F(X_s, Y_s)X_s)$ in (1.1), where F is a friction term. When $F > 0$ natural questions arise concerning the asymptotic behavior of (X_t, Y_t), for instance, the geometric convergence to equilibrium for the Langevin equation is discussed in Mattingly and Stuart [20], numerical approximations of the invariant measures in Talay [25], the case of high degree potential V is investigated in Hérau and Nier [12]. Under the previous boundedness assumption on b, Eq. (1.1) describes frictionless Hamiltonian systems with "almost linear" potential.

Importantly, the two time-scales coming from Kolmogorov's example, and that we obtain for the density associated to (1.1), can be exploited to investigate small time asymptotics of the previous models. For instance, for the Asian option, a normalization is required in the pay-off to make both quantities scale-homogeneous.

As mentioned above, equation (1.1) provides one of the simplest forms of degenerated processes and the previous assumptions guarantee that Hörmander's theorem is satisfied taking only the first Poisson brackets between the vector fields associated to the drift and the diffusive part in (1.1). In a more general hypoelliptic setting, let us mention the work of Cattiaux [8,9] whose assumptions include the case (1.1), but who obtains less explicit controls, see his Proposition (1.12). Under the "strong" Hörmander condition that involves the Poisson brackets of the diffusive part of the process, small time asymptotics of the density are discussed in Ben Arous [3] or Ben Arous and Léandre [4]. Eventually, in whole generality two-sided bounds for the density of degenerate diffusions are investigated in Kusuoka and Stroock [18]. All these work strongly rely on Malliavin calculus techniques. We want to stress that the parametrix approach is not very well suited to study general degenerate processes. Anyhow, the counterpart is that it gives by construction more explicit controls. In the non-degenerate case, for α-Hölder continuous coefficients, it directly gives two-sided Gaussian estimates. The lower bound on the diagonal in small time derives from the series representation and the global lower bound is obtained thanks to a chaining argument as in [18]. Here, we still derive a lower bound in small time from the series, but a chaining argument needs to be developed.

Also, our controls remain valid if the coefficients in (1.1) are uniformly α-Hölder continuous, a case for which Hörmander's theorem breaks down, see Section 3, Remark 3.1 for details.

A natural question then concerns the Markov chain approximation of (1.1). For non-degenerated processes this aspect has been widely studied, see e.g. [15] for local limit theorems. In [7], using Malliavin calculus techniques, Bally and Talay obtain an expansion at order one w.r.t. the time step for the difference of the densities of the diffusion and a perturbed Euler scheme, i.e. the stochastic integrals are approximated by Gaussian variables and an artificial viscosity is added to ensure the discrete scheme has a density. This rate corresponds to the usual "weak error" bound. Since we follow the local limit theorem approach we can handle a wider class of random variables in the approximation but obtain a rate of order $1/2$ w.r.t. the time step. Of course, plugging Gaussian random variables in our approximation yields to rate h as in [7].

Importantly, as opposed to [7], we do not need to introduce an artificial viscosity to ensure the existence of the density for the underlying degenerate Markov chain. We develop analogously to the continuous case a parametrix approach to express the density of the Markov chain in term of the density of an auxiliary frozen random walk. The random walk is degenerated as well, but has a density after a sufficient number of time steps, see Section 4.4 for details. The local limit theorem is then derived from an accurate comparison of the parametrix expansions of the densities of the process and the chain. To motivate this result we can consider the case of the approximation of a "digital Asian call" i.e. of the quantity $\mathbb{P}[(T^{-1}Y_T - X_T)^+ > K]$ for a given $K \in \mathbb{R}^+$. Indeed, the local limit theorem associated to our scheme directly relates the densities of the discrete and continuous objects which is not the case if we only consider a discretization of the non degenerate component and a numerical estimation of the integral, since in that case the approximating couple can fail to have a density.

The paper is organized as follows. In Section 1.3, we give our assumptions and fix some notations. Then, since the form of the Markov chain approximation strongly relies on the proof of our results for the diffusion we choose to divide this paper into two parts. Sections 2 and 3 deal with the results for the diffusion and their proofs. Section 4 is dedicated to the Markov chain approximation of (1.1), the associated convergence results and the key points of the proofs. The complete proof of the local limit theorem can be found in the Appendix of [16].

1.3. *Assumptions and notations*

We suppose that the coefficients of Eq. (1.1) satisfy the following assumptions:

(UE) $\exists(\lambda_{\min}, \lambda_{\max}) \in (0,\infty)^2, \forall z \in \mathbb{R}^d, \forall (x,y) \in \mathbb{R}^{2d}, \lambda_{\min}|z|^2 \le \langle a(x,y)z, z\rangle \le \lambda_{\max}|z|^2$, denoting $a(x,y) = \sigma\sigma^*(x,y)$. From now on we suppose that σ is the unique symmetric matrix s.t. $\sigma\sigma = a$. We are interested in the density of the process and its approximation at a given time. Hence, from the uniqueness in law, the previous assumption concerning the factorization of a can be made without loss of generality.

(B) The coefficients b, σ in (1.1) are C^1, uniformly Lipschitz continuous and bounded.

Throughout the paper we consider the running diffusion (1.1) up to a fixed final time $T > 0$. We denote by C a generic positive constant that may change from line to line and only depends on T, and the parameters appearing in (UE), (B). We reserve the notation c for constants that only depend on parameters from (UE), (B). Other possible dependencies are explicitly indicated.

2. Explicit parametrix and associated controls for the density of the diffusion

The previous assumptions guarantee that Hörmander's theorem, see e.g. Nualart [23], holds true, and therefore that $\forall t > 0$, (X_t, Y_t) has a density w.r.t. the Lebesgue measure. Introduce the vector fields in $\mathbb{R}^{2d}$

$$A_0(x,y) = \begin{pmatrix} b_1(x,y) \\ \vdots \\ b_d(x,y) \\ x_1 \\ \vdots \\ x_d \end{pmatrix}, \qquad \forall j \in [\![1,d]\!] \quad A_j(x,y) = \begin{pmatrix} \sigma_{1j}(x,y) \\ \vdots \\ \sigma_{dj}(x,y) \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{2.1}$$

One directly derives the following proposition.

Proposition 2.1. *For all* $(x,y) \in \mathbb{R}^{2d}$,

$$\operatorname{Span}\bigl(A_1(x,y), \dots, A_d(x,y), \bigl[A_0(x,y), A_1(x,y)\bigr], \dots, \bigl[A_0(x,y), A_d(x,y)\bigr]\bigr) = \mathbb{R}^{2d},$$

where $\forall (i,j) \in [\![0,d]\!]^2$, $[A_i, A_j] = A_i \nabla A_j - A_j \nabla A_i$ *denotes the Poisson bracket.*

Fix $T > 0$ and $0 < t \le T$, $(x,y) \in \mathbb{R}^{2d}$. Since, we now know that (X_t, Y_t) has a transition density, i.e. $\mathbb{P}[X_t \in \mathrm{d}x', Y_t \in \mathrm{d}y' | X_0 = x, Y_0 = y] = p(t, (x,y), (x',y'))\,\mathrm{d}x'\,\mathrm{d}y'$, our aim is to develop a parametrix for (1.1) to obtain an explicit representation of this density.

Recall that we consider the following SDE

$$\begin{cases} \mathrm{d}X_s = b(X_s, Y_s)\,\mathrm{d}t + \sigma(X_s, Y_s)\mathrm{d}W_s, & X_0 = x, \\ \mathrm{d}Y_s = X_s\,\mathrm{d}s, & Y_0 = y. \end{cases} \tag{2.2}$$

For the parametrix development we need to introduce a "frozen" diffusion process, $(\widetilde{X}_s, \widetilde{Y}_s)_{s\in[0,t]}$ below. Namely for all $(x',y') \in \mathbb{R}^{2d}$, $t \in (0,T]$ define

$$\begin{cases} \mathrm{d}\widetilde{X}_s^{t,x',y'} = \sigma\bigl(x', y' - x'(t-s)\bigr)\,\mathrm{d}W_s + b\bigl(x',y'\bigr)\,\mathrm{d}s, & \widetilde{X}_0^{t,x',y'} = x, \\ \mathrm{d}\widetilde{Y}_s^{t,x',y'} = \widetilde{X}_s^{t,x',y'}\,\mathrm{d}s, & \widetilde{Y}_0^{t,x',y'} = y. \end{cases} \tag{2.3}$$

The key point is that the above process is Gaussian. The arguments in the second variable of the diffusion coefficient can seem awkward at first sight, it includes the transport of the frozen point x' with a time reversal. This particular choice is actually imposed by the natural metric of the frozen process, see Proposition 3.1, in order to allow the comparison of the singular parts of the generators.

The processes (X_s, Y_s) and $(\widetilde{X}_s^{t,x',y'}, \widetilde{Y}_s^{t,x',y'})$, $s \in [0,t]$, have the following generators: $\forall (x,y) \in \mathbb{R}^{2d}$, $\psi \in C^2(\mathbb{R}^{2d})$,

$$\begin{aligned} L\psi(x,y) &= \left(\frac{1}{2}\operatorname{Tr}\big(a(x,y)D_x^2\psi\big) + \big\langle b(x,y), \nabla_x\psi\big\rangle + \langle x, \nabla_y\psi\rangle\right)(x,y), \\ \widetilde{L}_s^{t,x',y'}\psi(x,y) &= \left(\frac{1}{2}\operatorname{Tr}\big(a\big(x', y' - x'(t-s)\big)D_x^2\psi\big) + \big\langle b\big(x',y'\big)\nabla_x\psi\big\rangle + \langle x, \nabla_y\psi\rangle\right)(x,y). \end{aligned} \tag{2.4}$$

From these operators we define for $0 < t \le T$, $((x,y),(x',y')) \in (\mathbb{R}^{2d})^2$ the kernel H by

$$H\big(t,(x,y),\big(x',y'\big)\big) = (L - \widetilde{L})\widetilde{p}\big(t,(x,y),\big(x',y'\big)\big),$$

where $\widetilde{p}(t,(x,y),(\cdot,\cdot)) := \widetilde{p}^{t,x',y'}(t,(x,y),(\cdot,\cdot))$, $\widetilde{L} := \widetilde{L}_0^{t,x',y'}$ respectively stand for the density of the process $(\widetilde{X}_t^{t,x',y'}, \widetilde{Y}_t^{t,x',y'})$ and the generator of $(\widetilde{X}_s^{t,x',y'}, \widetilde{Y}_s^{t,x',y'})_{s\in[0,t]}$ at time $s = 0$. We omit to explicitly emphasize the dependence in t, x', y' for notational convenience.

Remark 2.1. *Note carefully that in the above kernel H, because of the linear structure of the model the most singular terms, i.e. those involving derivatives w.r.t. y, vanish.*

The next proposition gives the expression of the density p in terms of an infinite sum involving iterated convolutions of the density $\widetilde{p}$ with the kernel H. Namely,

Proposition 2.2 (Parametrix expansion for (2.2)). *For all $0 \le t \le T$, $((x,y),(x',y')) \in (\mathbb{R}^{2d})^2$,*

$$p\big(t,(x,y),\big(x',y'\big)\big) = \sum_{r=0}^{+\infty} \widetilde{p} \otimes H^{(r)}\big(t,(x,y),\big(x',y'\big)\big), \tag{2.5}$$

where

$$\begin{aligned} &f \otimes g\big(t,(x,y),\big(x',y'\big)\big) \\ &\quad = \int_0^t \mathrm{d}u \int_{\mathbb{R}^{2d}} f\big(u,(x,y),(z,v)\big)g\big(t-u,(z,v),\big(x',y'\big)\big)\,\mathrm{d}z\,\mathrm{d}v, \end{aligned}$$

$\tilde{p} \otimes H^{(0)} = \tilde{p}$ and $H^{(r)} = H \otimes H^{(r-1)}$, $r > 0$ denotes the r-fold convolution of the kernel H.

The previous Proposition is a direct consequence of the usual parametrix recurrence relations. For the sake of completeness we provide its proof in Section 3.

Now, since for $0 < t \le T$ $(\tilde{X}_s, \tilde{Y}_s)_{s\in[0,t]}$, is a Gaussian process, $\widetilde{p}$ and its derivatives are well controlled. The previous expression is the starting point to derive the following theorem.

Theorem 2.1 (Gaussian bounds). *There exist constants $c, C > 0$ s.t. for all $0 < t \le T$, $((x,y),(x',y')) \in (\mathbb{R}^{2d})^2$, one has*:

$$p\big(t,(x,y),\big(x',y'\big)\big) \le C\widehat{p}_c\big(t,(x,y),\big(x',y'\big)\big), \tag{2.6}$$

where

$$\widehat{p}_c\big(t,(x,y),\big(x',y'\big)\big) := \frac{c^d 3^{d/2}}{(2\pi t^2)^d} \times \exp\left(-c\left[\frac{|x'-x|^2}{4t} + 3\frac{|y'-y-(x+x')t/2|^2}{t^3}\right]\right)$$

enjoys the semigroup property, i.e. $\forall 0 \le s < t \le T$,

$$\int_{\mathbb{R}^{2d}} \hat{p}_c\big(s,(x,y),(w,z)\big)\hat{p}_c\big(t-s,(w,z),\big(x',y'\big)\big)\,\mathrm{d}w\,\mathrm{d}z = \hat{p}_c\big(t,(x,y),\big(x',y'\big)\big).$$

Also, for a given $C_0 > 0$, $\exists t_0 := t_0(C_0, c, C)$ *s.t. for* $t \le t_0$, $[\frac{|x'-x|^2}{4t} + 3\frac{|y'-y-(x+x')t/2|^2}{t^3}] \le C_0$, $p(t,(x,y),(x',y')) \ge C^{-1}\widehat{p}_{c^{-1}}(t,(x,y),(x',y'))$.

Remark 2.2. *The lower bound, obtained in small time and compact sets, derives from the parametrix representation of Proposition* 2.2 *and the upper Gaussian control. It remains an open problem to find a well suited chaining argument to derive a global lower bound for this degenerate case.*

Remark 2.3. *Note that the above result would remain valid if we replaced the dynamics of* Y_t *in* (1.1) *by* $Y_t = y + \int_0^t F(X_s)\,\mathrm{d}s$ *for a* $C^{2+\alpha}$, $\alpha > 0$, *Lipschitz continuous mapping* $F:\mathbb{R}^d \to \mathbb{R}^d$ *s.t. the Gram matrix* $DFDF^*$ *is non-degenerated, i.e.* $\exists c_0 > 0, \forall (\xi, x) \in \mathbb{R}^d \times \mathbb{R}^d, |\langle DFDF^*(x)\xi,\xi\rangle| \ge c_0|\xi|^2$. *Indeed, in such a case,* $(\bar{X}_s, \bar{Y}_s)_{s\in[0,T]} := (F(X_s), Y_s)_{s\in[0,T]}$ *follows a dynamics of type* (1.1).

3. Proof of the main results: Diffusion process

3.1. *Proof of Proposition* 2.2: *Parametrix expansion*

Following Cattiaux [8] and Lemma 3.1 one derives that $p, \widetilde{p}$ have continuous densities with bounded derivatives. Hence, from the forward and backward Kolmogorov equations associated to $(X,Y), (\widetilde{X},\widetilde{Y})$ and denoting by L^* the adjoint of L, we have

$$\begin{aligned}
&p\big(t,(x,y),\big(x',y'\big)\big) - \widetilde{p}\big(t,(x,y),\big(x',y'\big)\big)\\
&\quad= \int_0^t \mathrm{d}u \frac{\partial}{\partial u}\int_{\mathbb{R}^{2d}} \mathrm{d}w\,\mathrm{d}z\, p\big(u,(x,y),(w,z)\big)\widetilde{p}\big(t-u,(w,z),\big(x',y'\big)\big)\\
&\quad= \int_0^t \mathrm{d}u \int_{\mathbb{R}^{2d}} \mathrm{d}w\,\mathrm{d}z\left[\frac{\partial p(u,(x,y),(w,z))}{\partial u}\widetilde{p}\big(t-u,(w,z),\big(x',y'\big)\big)\right.\\
&\qquad\left. + p\big(u,(x,y),(w,z)\big)\frac{\partial \widetilde{p}(t-u,(w,z),(x',y'))}{\partial u}\right]\\
&\quad= \int_0^t \mathrm{d}u \int_{\mathbb{R}^{2d}} \mathrm{d}w\,\mathrm{d}z\big[L^* p\big(u,(x,y),(w,z)\big)\widetilde{p}\big(t-u,(w,z),\big(x',y'\big)\big)\\
&\qquad - \widetilde{L}\widetilde{p}\big(t-u,(w,z),\big(x',y'\big)\big)p\big(u,(x,y),(w,z)\big)\big]\\
&\quad= \int_0^t \mathrm{d}u \int_{\mathbb{R}^{2d}} \mathrm{d}w\,\mathrm{d}z\, p\big(u,(x,y),(w,z)\big)(L-\widetilde{L})\widetilde{p}\big(t-u,(w,z),\big(x',y'\big)\big)\\
&\quad= p \otimes H\big(t,(x,y),\big(x',y'\big)\big).
\end{aligned}$$

A simple iteration completes the proof.

3.2. *Proof of Theorem* 2.1

3.2.1. *Proof of the upper bound*

The proof is divided into two parts. First an elementary control on the density of $(\widetilde{X},\widetilde{Y})$ is stated in Lemma 3.1. Then, this control is used to control the kernel H and the convolutions.

Step 1: *Gaussian controls for* $(\widetilde{X},\widetilde{Y})$.

Lemma 3.1. *There exist constants $c > 0, C > 0$, s.t. for all multi-index α, $|\alpha| \le 3$, $\forall 0 \le u < t \le T$, $\forall (w, z)$, $(x', y') \in \mathbb{R}^{2d}$,*

$$\begin{aligned}
&\bigl|\partial_w^\alpha \widetilde{p}\bigl(t-u,(w,z),(x',y')\bigr)\bigr| \\
&\quad \le C(t-u)^{-|\alpha|/2} \frac{c^d 3^{d/2}}{(2\pi t^2)^d} \exp\biggl(-c\biggl[\frac{|x'-w|^2}{4(t-u)} + 3\frac{|y'-z-(1/2)(w+x')(t-u)|^2}{(t-u)^3}\biggr]\biggr) \\
&\quad := C(t-u)^{-|\alpha|/2} \widehat{p}_c\bigl(t-u,(w,z),(x',y')\bigr), \\
&\widetilde{p}\bigl(t-u,(w,z),(x',y')\bigr) \ge 2C^{-1} \widehat{p}_{c^{-1}}\bigl(t-u,(w,z),(x',y')\bigr),
\end{aligned}$$

where $\hat{p}_c$ is as in Theorem 2.1 and enjoys the semigroup property.

The proof is postponed to Section 3.2.2.

Step 2: *Control of the kernel*. Recall that under (B), the coefficients a, b are uniformly Lipschitz continuous. Hence, it is easy to get from Lemma 3.1 and the previous definition of H that, up to a modification of $c > 0$ in $\widehat{p}_c$, $\exists C_1 > 0, \forall u \in [0, t)$,

$$\bigl|H\bigl(t-u,(w,z),(x',y')\bigr)\bigr| \le \frac{C_1}{\sqrt{t-u}} \widehat{p}_c\bigl(t-u,(w,z),(x',y')\bigr). \tag{3.1}$$

Lemma 3.1 also yields that $\exists C_2 > 0, \forall u \in (0, t]$, $\widetilde{p}(u,(x,y),(w,z)) \le C_2 \widehat{p}_c(u,(x,y),(w,z))$. Setting $C := C_1 \vee C_2$, we finally obtain

$$\begin{aligned}
\bigl|\widetilde{p} \otimes H\bigl(t,(x,y),(x',y')\bigr)\bigr| &\le \int_0^t \mathrm{d}u \int_{\mathbb{R}^{2d}} \widetilde{p}\bigl(u,(x,y),(w,z)\bigr)\bigl|H\bigl(t-u,(w,z),(x',y')\bigr)\bigr| \,\mathrm{d}w\,\mathrm{d}z, \\
&\le \int_0^t \mathrm{d}u \int_{\mathbb{R}^{2d}} C^2 \widehat{p}_c\bigl(u,(x,y),(w,z)\bigr) \frac{1}{\sqrt{t-u}} \widehat{p}_c\bigl(t-u,(w,z),(x',y')\bigr) \,\mathrm{d}w\,\mathrm{d}z \\
&\le C^2 t^{1/2} B\biggl(1,\frac{1}{2}\biggr) \widehat{p}_c\bigl(t,(x,y),(x',y')\bigr),
\end{aligned}$$

using the semigroup property of $\widehat{p}_c$ in the last inequality and where $B(m,n) = \int_0^1 u^{m-1}(1-u)^{n-1}\,\mathrm{d}u$ denotes the β-function. By induction in r,

$$\begin{aligned}
&\bigl|\widetilde{p} \otimes H^{(r)}\bigl(t,(x,y),(x',y')\bigr)\bigr| \\
&\quad \le C^{r+1} t^{r/2} B\biggl(1,\frac{1}{2}\biggr) B\biggl(\frac{3}{2},\frac{1}{2}\biggr) \times \cdots \times B\biggl(\frac{r+1}{2},\frac{1}{2}\biggr) \widehat{p}_c\bigl(t,(x,y),(x',y')\bigr), \quad r \in \mathbb{N}^*.
\end{aligned} \tag{3.2}$$

This implies that the series representing the density $p(t,(x,y),(x',y'))$,

$$p\bigl(t,(x,y),(x',y')\bigr) = \sum_{r=0}^{\infty} \widetilde{p} \otimes H^{(r)}\bigl(t,(x,y),(x',y')\bigr)$$

is absolutely convergent and the following estimate holds

$$\bigl|p\bigl(t,(x,y),(x',y')\bigr)\bigr| \le C \widehat{p}_c\bigl(t,(x,y),(x',y')\bigr).$$

Remark 3.1. *Note carefully that the above series still converges if the coefficients b, σ are only uniformly α-Hölder continuous. In such case Hörmander's theorem does not hold, but one can show by standard techniques, see e.g. Baldi* [2], *that $p(t,(x,y),(\cdot,\cdot)) := \sum_{r \in \mathbb{N}} \widetilde{p} \otimes H^{(r)}(t,(x,y),(\cdot,\cdot))$ is a probability density and derive with a Dynkin like argument, see e.g. Theorem* 2.3 *in* [10], *that it corresponds to the density of the weak solution of* (1.1).

3.2.2. *Proof of the partial lower bound*
From the previous proof and the Gaussian nature of $(\widetilde{X}_t, \widetilde{Y}_t)$, see Lemma 3.1, one gets

$$\begin{aligned} p\big(t,(x,y),(x',y')\big) &\geq \widetilde{p}\big(t,(x,y),(x',y')\big) - Ct^{1/2}\widehat{p}_c\big(t,(x,y),(x',y')\big) \\ &\geq 2C^{-1}\widehat{p}_{c^{-1}}\big(t,(x,y),(x',y')\big) - Ct^{1/2}\widehat{p}_c\big(t,(x,y),(x',y')\big) \\ &\geq C^{-1}\widehat{p}_{c^{-1}}\big(t,(x,y),(x',y')\big) \end{aligned}$$

for $\frac{|x'-x|^2}{4t} + 3\frac{|y'-y-(1/2)(x+x')t|^2}{t^3} \leq C_0$ and t small enough.

3.2.3. *Proof of the technical lemmas*

Proof of Lemma 3.1. . We prove the statement for $|\alpha| = 0$, i.e. without derivation. Indeed, since our computations only involve a finite number of derivations that introduce some polynomials in front of the exponential, the general bound can be derived similarly and the result holds taking the worst constants. Also, with respect to the statement of the lemma, we suppose w.l.o.g. $u = 0$ for notational convenience. We get from (2.3) with $x = w$, $y = z$ that for all $0 < t \leq T$,

$$\begin{aligned} \widetilde{Y}_t &= \left\{z + wt + b\big(x',y'\big)\frac{t^2}{2}\right\} + \int_0^t \left\{\int_0^v \sigma\big(x', y' - x'(t-s)\big)\,\mathrm{d}W_s\right\}\mathrm{d}v := m_{2,t} + A_t, \\ A_t &= \int_0^t (t-s)\sigma\big(x', y' - x'(t-s)\big)\,\mathrm{d}W_s := \int_0^t (t-s)\widetilde{\sigma}_s\,\mathrm{d}W_s, \end{aligned} \tag{3.3}$$

using Itô's formula for the last but one equality. Setting $\forall s \in [0,t], \widetilde{a}_s = \widetilde{\sigma}_s^2$, recall from (UE) that $\widetilde{\sigma}_s$ is symmetric, we finally obtain that the covariance matrix Σ_t of the vector $(\widetilde{X}_t, \widetilde{Y}_t)$ is equal to

$$\Sigma_t = \begin{pmatrix} \int_0^t \widetilde{a}_s\,\mathrm{d}s & \int_0^t (t-s)\widetilde{a}_s\,\mathrm{d}s \\ \int_0^t (t-s)\widetilde{a}_s\,\mathrm{d}s & \int_0^t (t-s)^2\widetilde{a}_s\,\mathrm{d}s \end{pmatrix}.$$

Note from (UE) that: $\exists c > 0, \forall s \in [0,T], \forall \xi \in \mathbb{R}^d, c^{-1}|\xi|^2 \geq \langle \widetilde{a}_s\xi, \xi\rangle \geq c|\xi|^2$. Hence, by the Cauchy Schwarz inequality

$$\forall Z \in \mathbb{R}^{2d} \quad 2/c\langle C_t Z, Z\rangle \geq \langle \Sigma_t Z, Z\rangle \geq c/2\langle C_t Z, Z\rangle, \quad C_t := \begin{pmatrix} tI_d & \frac{t^2}{2}I_d \\ \frac{t^2}{2}I_d & \frac{t^3}{3}I_d \end{pmatrix}, \tag{3.4}$$

where C_t is actually the covariance matrix of a d-dimensional standard Brownian motion and its integral.

The mean vector of $(\widetilde{X}_t, \widetilde{Y}_t)$ is equal to $(m_{1,t}, m_{2,t})$, with $m_{1,t} = w + b(x',y')t$ and $m_{2,t}$ as in (3.3). Note that $C_t = T_t\mathcal{A}T_t^*$, where

$$T_t^* = \begin{pmatrix} t^{1/2}I_d & \frac{t^{3/2}}{2}I_d \\ 0 & t^{3/2}I_d \end{pmatrix}, \qquad \mathcal{A} = \begin{pmatrix} I_d & 0 \\ 0 & \frac{1}{12}I_d \end{pmatrix}.$$

In the above decomposition, the matrix T_t can be seen as a "scale" matrix that gives the characteristic scale of each component, whereas $\mathcal{A}$ is a time independent "macroscopic" matrix.

Hence,

$$C_t^{-1} = \big(T_t^*\big)^{-1}\mathcal{A}^{-1}T_t^{-1} = \begin{pmatrix} t^{-1/2}I_d & -\frac{t^{-1/2}}{2}I_d \\ 0 & t^{-3/2}I_d \end{pmatrix}\begin{pmatrix} I_d & 0 \\ 0 & 12I_d \end{pmatrix}\begin{pmatrix} t^{-1/2}I_d & 0 \\ -\frac{t^{-1/2}}{2}I_d & t^{-3/2}I_d \end{pmatrix}.$$

Now, Σ_t can be factorized using the scale matrix T_t. Namely, $\Sigma_t = T_tD_tT_t^*$ where owing to (3.4),

$$D_t := \begin{pmatrix} t^{-1}\int_0^t \widetilde{a}_s\,\mathrm{d}s & t^{-2}\int_0^t \widetilde{a}_s(\frac{t}{2}-s)\,\mathrm{d}s \\ t^{-2}\int_0^t \widetilde{a}_s(\frac{t}{2}-s)\,\mathrm{d}s & t^{-3}\int_0^t \widetilde{a}_s(\frac{t}{2}-s)^2\,\mathrm{d}s \end{pmatrix}$$

is a macroscopic uniformly non-degenerated matrix w.r.t. $t \in (0,T]$. Thus, for a sufficiently small c, different than the one appearing in (3.4), we obtain that $\forall Z \in \mathbb{R}^{2d}$, $\mathcal{E} := -\langle \Sigma_t^{-1} Z, Z\rangle \le -c\langle \mathcal{A}^{-1}(T_t^{-1}Z), T_t^{-1}Z\rangle$. In particular, for $Z = (Z_1, Z_2)$, $Z_1 = x' - (w + b(x',y')t)$, $Z_2 = y' - (z + wt + b(x',y')\frac{t^2}{2})$, we get

$$T_t^{-1} Z = \begin{pmatrix} (x' - w - b(x',y')t)t^{-1/2} \\ (y' - z - \frac{1}{2}(x'+w)t)t^{-3/2} \end{pmatrix}.$$

We therefore derive

$$\frac{\mathcal{E}}{2} \le -\frac{c}{2t}\left|x' - w - b\bigl(x',y'\bigr)t\right|^2 - \frac{6c}{t^3}\left|y' - z - \frac{1}{2}\bigl(x'+w\bigr)t\right|^2.$$

From (B) (boundedness of b), we derive that there exist $c, C > 0$ s.t.

$$\frac{\mathcal{E}}{2} \le C - c\left[\frac{|x'-w|^2}{4t} + 3\frac{|y' - z - (1/2)(w+x')t)|^2}{t^3}\right].$$

Eventually

$$\begin{aligned} \widetilde{p}\bigl(t,(w,z),\bigl(x',y'\bigr)\bigr) &\le \frac{Cc^d 3^{d/2}}{(2\pi t^2)^d}\exp\left(-c\left[\frac{|x'-w|^2}{4t} + 3\frac{|y'-z-(1/2)(w+x')t)|^2}{t^3}\right]\right) \\ &:= C\widehat{p}_c\bigl(t,(w,z),\bigl(x',y'\bigr)\bigr). \end{aligned}$$

Note from [17] that $\widehat{p}_c$ enjoys the semigroup property. This gives the statement for $|\alpha| = 0$. The lower bound is derived similarly from the control $\forall Z \in \mathbb{R}^{2d}$, $\langle \Sigma_t^{-1} Z, Z\rangle \le c^{-1}\langle C_t^{-1}Z, Z\rangle$ achieved for c small enough. □

4. Markov Chain approximation and associated convergence results

4.1. *Global strategy*

Let us recall the strategy to derive a local limit theorem for the Markov chain approximation associated to a diffusion process. Suppose the underlying diffusion has a density with parametrix representation as in Proposition 2.2. If the "natural" Markov chain associated to the diffusion has a density, the main idea is to introduce a Markov chain with frozen coefficients that also has a density so that the density of the Markov chain can be written in parametrix form as well with a suitable discrete kernel.

The next step consists in comparing these two parametrix representations. To this end, two key steps are needed:

1. The comparison of the densities of the frozen Markov chain and frozen diffusion process.
2. The comparison of the kernels.

The first step relies on Edgeworth like expansions, see e.g. Bhattacharya and Rao [6], the second one on careful Taylor like expansions.

The local limit Theorem is then derived by controlling the iterated convolutions of differences of the kernels. This procedure has been applied successfully in [15] to derive a local limit theorem for the Markov chain approximation of uniformly elliptic inhomogeneous diffusions with bounded coefficients.

In our current framework new difficulties arise. First of all it is not obvious to derive that a "natural" Markov chain associated to (1.1) has a transition density. To guarantee such an existence a common trick in the literature consists in adding an artificial viscosity term in the discretization scheme, see e.g. [7]. Our strategy is here different. Namely, we manage to obtain a density for the natural frozen Markov chain deriving from (2.3) after a sufficient number of time steps. We therefore consider a "macro scale" frozen model corresponding to this number of time steps. We then obtain a good comparison between the densities of the "aggregated" chain at macro scale and the frozen diffusion process. This first step gives the structure of the random variables involved in the approximation in order to have the comparison of the densities. These variables have a density. From these variables, we then derive the Markov chain dynamics by letting the coefficients vary at macro scale.

A second difficulty is that the second component in (1.1) is unbounded. This yields to handle a supplementary term w.r.t. the analysis carried out in [15] and to a slightly different version of the local limit theorem. In the sequel we first give the dynamics of the Markov chains at macro scale and state the local limit theorem (Section 4.2). We give the lemma for the comparison of the densities (Section 4.3) and prove the existence of the density for the aggregated "frozen" Markov chain (Section 4.4). The whole proof of the local limit theorem is carried out in the appendix of [16].

4.2. *Models and results*

Now, fix $T > 0$, $N_0 \in \mathbb{N}^*$ and let $h_0 = T/N_0$ be the "micro" time discretization step. Let $n \in \mathbb{N}^*$ be large enough so that the natural "frozen" chain associated to (2.3) has a density, see Proposition 4.2, and define the "macro" scale time step $h = nh_0$ and set $N = N_0/n \in \mathbb{N}^*$ the total number of "macro" time steps over $[0, T]$.

For all $i \in [\![0, N]\!]$ set $t_i := ih$. For any $(x, y) \in \mathbb{R}^{2d}$, we define on the time grid $\{0, \dots, t_N\}$ an $\mathbb{R}^{2d}$ valued Markov chain $(Z^h_{t_i})_{i\in[\![0,N]\!]} = ((X^h_{t_i}, Y^h_{t_i})^*)_{i\in[\![0,N]\!]}$ whose dynamics is given by

$$
\begin{aligned}
&Z^h_0 = (x, y)^* \quad \text{and} \quad \forall i \in [\![0, N-1]\!],\\
&X^h_{t_{i+1}} = X^h_{t_i} + b\big(Z^h_{t_i}\big)h + \sigma\big(Z^h_{t_i}\big)\sqrt{h}\eta^1_{i+1},\\
&Y^h_{t_{i+1}} = Y^h_{t_i} + \left(X^h_{t_i} + \frac{\gamma_n}{2} b\big(Z^h_{t_i}\big)h + \sigma\big(Z^h_{t_i}\big)\sqrt{h}\eta^2_{i+1}\right)h,
\end{aligned}
\tag{4.1}
$$

where $\gamma_n := (1 + \frac{1}{n})$. The variables $(\vartheta_i)_{i\in(0,N]\!]} := (\eta^1_i, \eta^2_i)_{i\in(0,N]\!]}$ are i.i.d. $2d$-dimensional random variables s.t.

(A1) $\mathbb{E}[\vartheta_i] = 0$ and $\operatorname{Cov}(\vartheta_i) = \begin{pmatrix} \mathbf{I}_{d\times d} & \frac{1}{2}\gamma_n \mathbf{I}_{d\times d} \\ \frac{1}{2}\gamma_n \mathbf{I}_{d\times d} & \frac{1}{3}\gamma_n(1 + \frac{1}{2n})\mathbf{I}_{d\times d} \end{pmatrix}$.

(A2) The variable ϑ_i has density $q_n(\eta_1, \eta_2)$ and there exist a positive integer S' and a function $\psi : \mathbb{R}^{2d} \to \mathbb{R}$ with $\sup_{u\in\mathbb{R}^{2d}} \psi(u) < \infty$ and $\int \|u\|^S \psi(u)\,\mathrm{d}u < \infty$ for $S = 4dS' + 4$ such that

$$
\big|D^\nu_u q_n(u)\big| \le \psi(u)
$$

for all $|\nu| \in [\![0, 4]\!]$. The main result of the section, i.e. Theorem 4.1, is stated in terms of S'.

We finally need a "frozen" time inhomogeneous Markov chain. For $(x, y), (x', y') \in \mathbb{R}^{2d}$, $j \in (0, N]\!]$ we define $(\widetilde{Z}^h_{t_i})_{i\in[\![0,j]\!]} = ((\widetilde{X}^h_{t_i}, \widetilde{Y}^h_{t_i})^*)_{i\in[\![0,j]\!]}$ by

$$
\begin{aligned}
&\widetilde{Z}^h_0 = (x, y)^* \quad \text{and} \quad \forall i \in [\![0, j-1]\!],\\
&\widetilde{X}^h_{t_{i+1}} = \widetilde{X}^h_{t_i} + b\big(x', y'\big)h + \sigma\big(x', y' - x'(t_j - t_i)\big)\sqrt{h}\widetilde{\eta}^1_{i+1},\\
&\widetilde{Y}^h_{t_{i+1}} = \widetilde{Y}^h_{t_i} + \left\{\widetilde{X}^h_{t_i} + \frac{\gamma_n}{2} b\big(x', y'\big)h + \sigma\big(x', y' - x'(t_j - t_i)\big)\sqrt{h}\widetilde{\eta}^2_{i+1}\right\}h.
\end{aligned}
\tag{4.2}
$$

The i.i.d. variables $(\tilde{\eta}^1_i, \tilde{\eta}^2_i)_{i\in(0,j]\!]}$ have density $q_n(\cdot)$.

Remark 4.1. *Note that the models introduced in* (4.1) *and* (4.2) *can seem awkward at first sight. They actually derive from computations that yield the existence of the density for the natural frozen Markov chain associated to* (2.3) *after n "micro" time steps h_0, i.e. at the "macro" level with time step h. This is developed in Section* 4.4.

From now on, $p_h(t_{j'}, (x, y), (\cdot, \cdot))$ and $\widetilde{p}_h^{\,t_j,x',y'}(t_{j'}, (x, y), (\cdot, \cdot)) := \widetilde{p}_h(t_{j'}, (x, y), (\cdot, \cdot))$ denote the transition densities between times 0 and $t_{j'} \le t_j$ of the Markov chain (4.1), and "frozen" Markov chain (4.2) respectively. Introducing a discrete "analogue" to the generators we derive from the Markov property a relation similar to (2.5) between p_h and $\widetilde{p}_h$.

For a sufficiently smooth function f, define L_h and $\widetilde{L}_h$ by

$$\begin{aligned}
&L_h f\big(t_j,(x,y),\big(x',y'\big)\big)\\
&\quad = h^{-1}\biggl[\int p_h\big(h,(x,y),(u,v)\big) f\big(t_j-h,(u,v),\big(x',y'\big)\big)\,\mathrm{d}u\,\mathrm{d}v - f\big(t_j-h,(x,y),\big(x',y'\big)\big)\biggr],\\
&\widetilde{L}_h f\big(t_j,(x,y),\big(x',y'\big)\big)\\
&\quad = h^{-1}\biggl[\int \widetilde{p}_h^{t_j,x',y'}\big(h,(x,y),(u,v)\big) f\big(t_j-h,(u,v),\big(x',y'\big)\big)\,\mathrm{d}u\,\mathrm{d}v - f\big(t_j-h,(x,y),\big(x',y'\big)\big)\biggr].
\end{aligned}$$

Note that because of technical reasons, there is a shift in time in the above definitions, i.e. the time is t_j-h, instead of the "expected" t_j, in the right-hand side of the previous equations.

A discrete analogue H_h of the kernel H is defined as

$$H_h\big(t_j,(x,y),\big(x',y'\big)\big) = (L_h-\widetilde{L}_h)\widetilde{p}_h\big(t_j,(x,y),\big(x',y'\big)\big), \qquad 0<j\le N.$$

From the previous definition

$$\begin{aligned}
&H_h\big(t_j,(x,y),\big(x',y'\big)\big)\\
&\quad = h^{-1}\int \big[p_h-\widetilde{p}_h^{t_j,x',y'}\big]\big(h,(x,y),(u,v)\big)\widetilde{p}_h^{t_j,x',y'}\big(t_j-h,(u,v),\big(x',y'\big)\big)\,\mathrm{d}u\,\mathrm{d}v.
\end{aligned}$$

Analogously to Lemma 3.6 in [15] we obtain the following result.

Proposition 4.1 (Parametrix for Markov chain). *Assume* (A1), (A2), (UE), (B) *are in force. Then, for* $0<t_j\le T$,

$$p_h\big(t_j,(x,y),\big(x',y'\big)\big) = \sum_{r=0}^{j}\big(\widetilde{p}_h\otimes_h H_h^{(r)}\big)\big(t_j,(x,y),\big(x',y'\big)\big), \tag{4.3}$$

where the discrete time convolution type operator $\otimes_h$ *is defined by*

$$(g\otimes_h f)\big(t_j,(x,y),\big(x',y'\big)\big) = \sum_{i=0}^{j-1} h\int g\big(t_i,(x,y),(u,v)\big) f\big(t_j-t_i,(u,v),\big(x',y'\big)\big)\,\mathrm{d}u\,\mathrm{d}v,$$

$\widetilde{p}_h\otimes_h H_h^{(0)}=\tilde{p}_h$ *and* $H_h^{(r)}=H_h\otimes_h H_h^{(r-1)}$ *denotes the* r*-fold discrete convolution of the kernel* H_h. *W.r.t. the above definition, we use the convention that* $\widetilde{p}_h\otimes_h H_h^{(r)}(0,(x,y),(x',y'))=0$, $r\ge 1$.

Now (4.3) and (2.5) have the same form. Comparing these two expressions we obtain the following local limit theorem.

Theorem 4.1 (Local limit theorem for the densities). *Assume* (A1), (A2), (UE), (B) *hold true. Then*, $\exists c>0$,

$$\begin{aligned}
&\sup_{(x,y),(x',y')\in\mathbb{R}^{2d}}\biggl[\big(1+\big|x'\big|+|x|\big)\sup_{\delta\in[0,1]}\widehat{p}_c\big(T(1+\delta),(x,y),\big(x',y'\big)\big)+\chi_{\sqrt{T}}\biggl(x'-x,y'-y-T\biggl(\frac{x'+x}{2}\biggr)\biggr)\biggr]^{-1}\\
&\quad\times\big|p_h\big(T,(x,y),\big(x',y'\big)\big)-p\big(T,(x,y),\big(x',y'\big)\big)\big| = \mathrm{O}\big(h^{1/2}\big),
\end{aligned}$$

where $\widehat{p}_c$ *is as in Theorem* 2.1, p_h *denotes the density of the Markov chain* (4.1) *and* $\forall(\rho,u,v)\in\mathbb{R}^+\times\mathbb{R}^{2d}$,

$$\chi_\rho(u,v)=\rho^{-4d}\chi\big(u/\rho,v/\rho^3\big), \qquad \chi(u,v)=\big(1+\big(|u|^2+|v|^2\big)^{S'-1}\big)^{-1}.$$

Note from the above result that the bigger is S', the better is the control on the tails.

4.3. *Comparison of the discrete and continuous frozen densities*

The first step for the error analysis is achieved with the following

Lemma 4.1. *There exists $C>0$, s.t. for all $j\in(0,N]\!]$, $\rho^2:=t_j$,*

$$\big|(\widetilde{p}_h-\widetilde{p})\big(t_j,(x,y),\big(x',y'\big)\big)\big|\le Ch^{1/2}\rho^{-1}\zeta_\rho\bigg(x'-x,y'-y-\frac{x+x'}{2}\rho^2\bigg),\tag{4.4}$$

where $\zeta_\rho(u,v)=\rho^{-4d}\zeta(u/\rho,v/\rho^3)$, $\zeta(u,v)=\frac{1}{1+[|u|^2+|v|^2]^{(S-4)/2}}$, S being introduced in (A2).

Proof. Iterating (4.2) from 0 till t_j we get

$$\begin{aligned}\widetilde{X}^h_{t_j}&=x+b\big(x',y'\big)\rho^2+\rho\Bigg\{\frac{1}{j^{1/2}}\sum_{k=0}^{j-1}\sigma\big(x',y'-x'\big(\rho^2-t_k\big)\big)\widetilde{\eta}^1_{k+1}\Bigg\}\\ \widetilde{Y}^h_{t_j}&=y+x\rho^2+\frac{\rho^4}{2}b\big(x',y'\big)\bigg(1+\frac{1}{nj}\bigg)+\rho^3\Bigg\{\frac{1}{j^{1/2}}\sum_{k=0}^{j-1}\sigma\big(x',y'-x'\big(\rho^2-t_k\big)\big)\widetilde{\eta}^2_{k+1}\frac{1}{j}\\ &\quad+\frac{1}{j^{1/2}}\sum_{k=0}^{j-1}\sigma\big(x',y'-x'\big(\rho^2-t_k\big)\big)\widetilde{\eta}^1_{k+1}\bigg(1-\frac{k+1}{j}\bigg)\Bigg\}.\end{aligned}\tag{4.5}$$

Introduce

$$m_j=\begin{pmatrix}x+b(x',y')\rho^2\\ y+x\rho^2+\frac{\rho^4}{2}b(x',y')\gamma_{n,j}\end{pmatrix}:=\begin{pmatrix}m_j^1\\ m_j^2\end{pmatrix},\qquad\gamma_{n,j}:=1+\frac{1}{nj},$$

and

$$\Theta_j:=\begin{pmatrix}\{\frac{1}{j^{1/2}}\sum_{k=0}^{j-1}\sigma(x',y'-x'(\rho^2-t_k))\widetilde{\eta}^1_{k+1}\}\\ \{\frac{1}{j^{1/2}}\sum_{k=0}^{j-1}\sigma(x',y'-x'(\rho^2-t_k))\widetilde{\eta}^2_{k+1}\frac{1}{j}\\ \quad+\frac{1}{j^{1/2}}\sum_{k=0}^{j-1}\sigma(x',y'-x'(\rho^2-t_k))\widetilde{\eta}^1_{k+1}(1-\frac{k+1}{j})\}\end{pmatrix}.$$

The dynamics of (4.2) thus writes

$$\begin{pmatrix}\widetilde{X}^h_{t_j}\\ \widetilde{Y}^h_{t_j}\end{pmatrix}=m_j+\begin{pmatrix}\rho\mathbf{I}_{d\times d}&\mathbf{0}_{d\times d}\\ \mathbf{0}_{d\times d}&\rho^3\mathbf{I}_{d\times d}\end{pmatrix}\Theta_j.$$

Setting $\forall s\in[0,\rho^2]$, $\phi(s):=\inf\{t_i:=ih\colon t_i\le s<t_{i+1}\}$, $\widetilde{a}_s:=\sigma^2(x',y'-x'(\rho^2-s))$ we get $V_j:=\mathrm{Cov}(\Theta_j)=$

$$\begin{pmatrix}\frac{1}{t_j}\int_0^{t_j}\widetilde{a}_{\phi(s)}\,\mathrm{d}s&\frac{1}{t_j^2}\int_0^{t_j}\widetilde{a}_{\phi(s)}F_1^{j,h}(\phi(s))\,\mathrm{d}s\\ \frac{1}{t_j^2}\int_0^{t_j}\widetilde{a}_{\phi(s)}F_1^{j,h}(\phi(s))\,\mathrm{d}s&\frac{1}{t_j^3}\int_0^{t_j}\widetilde{a}_{\phi(s)}F_2^{j,h}(\phi(s))\,\mathrm{d}s\end{pmatrix},$$

where $F_1^{j,h}(\phi(s)):=[\frac{\gamma_nh}{2}+(t_j-(\phi(s)+h))]$, $F_2^{j,h}(\phi(s)):=[\frac{\gamma_nh^2}{3}(1+\frac{1}{2n})+\gamma_nh(t_j-(\phi(s)+h))+(t_j-(\phi(s)+h))^2]$. Now, similarly to the proof of Lemma 3.1, since $\exists c>0$ s.t. $\forall\xi\in\mathbb{R}^d$, $c^{-1}|\xi|^2\ge\langle\widetilde{a}_{\phi(s)}\xi,\xi\rangle\ge c|\xi|^2$, we derive from the Cauchy–Schwarz inequality that $\forall Z\in\mathbb{R}^{2d}$

$$2c^{-1}\langle A_jZ,Z\rangle\ge\langle V_jZ,Z\rangle\ge\frac{c}{2}\langle A_jZ,Z\rangle,\quad A_j:=\begin{pmatrix}\mathbf{I}_{d\times d}&\frac{1}{t_j^2}\int_0^{t_j}F_1^{j,h}(\phi(s))\,\mathrm{d}s\mathbf{I}_{d\times d}\\ \frac{1}{t_j^2}\int_0^{t_j}F_1^{j,h}(\phi(s))\,\mathrm{d}s\mathbf{I}_{d\times d}&\frac{1}{t_j^3}\int_0^{t_j}F_2^{j,h}(\phi(s))\,\mathrm{d}s\mathbf{I}_{d\times d}\end{pmatrix}.$$

Elementary computations also give $\exists c_1>0, \forall Z\in\mathbb{R}^{2d}, 2c_1^{-1}\langle AZ,Z\rangle \ge \langle A_j Z,Z\rangle \ge \frac{c_1}{2}\langle AZ,Z\rangle$, $A=\begin{pmatrix}\mathbf{I}_{d\times d} & \frac12\mathbf{I}_{d\times d}\\ \frac12\mathbf{I}_{d\times d} & \frac13\mathbf{I}_{d\times d}\end{pmatrix}$ for n large enough. Thus, for c small enough, and n large enough, $\forall Z\in\mathbb{R}^{2d}, c^{-1}|Z|^2\ge\langle V_j Z,Z\rangle\ge c|Z|^2$ and the covariance matrix V_j is uniformly invertible w.r.t. the parameter $j\in(0,N]\!]$. Denoting by g_n the density of the normalized sum $V_j^{-1/2}\Theta_j$ we derive

$$\widetilde{p}_h\big(t_j,(x,y),\big(x',y'\big)\big)=\frac{1}{\rho^{4d}\det(V_j^{1/2})}g_n\left(V_j^{-1/2}\begin{pmatrix}\frac{x'-m_j^1}{\rho}\\ \frac{y'-m_j^2}{\rho^3}\end{pmatrix}\right).$$

Applying the Edgeworth expansion for g_n (the key tool is the normal approximation of Bhattacharya and Rao, Theorem 19.3 in [6]) and exploiting arguments similar to those of the proof of Lemma 3.1, we obtain

$$\begin{aligned}&\left|\widetilde{p}_h\big(t_j,(x,y),\big(x',y'\big)\big)-\frac{1}{\rho^{4d}\det(V_j^{1/2})}g_G\left(V_j^{-1/2}\begin{pmatrix}\frac{x'-m_j^1}{\rho}\\ \frac{y'-m_j^2}{\rho^3}\end{pmatrix}\right)\right|\\ &\quad\le Ch^{1/2}\rho^{-1}\zeta_\rho\left(x'-x,y'-y-\frac{x+x'}{2}\rho^2\right),\end{aligned}\tag{4.6}$$

where g_G stands for the standard $2d$-dimensional Gaussian density. To conclude the proof, recall from the proof of Lemma 3.1 that

$$\widetilde{p}\big(t_j,(x,y),\big(x',y'\big)\big)=\frac{1}{\rho^{4d}\det(C_j^{1/2})}g_G\left(C_j^{-1/2}\begin{pmatrix}\frac{x'-m_{C,j}^1}{\rho}\\ \frac{y'-m_{C,j}^2}{\rho^3}\end{pmatrix}\right),\tag{4.7}$$

where

$$m_{C,j}=\begin{pmatrix}x+b(x',y')\rho^2\\ y+x\rho^2+\frac{\rho^4}{2}b(x',y')\end{pmatrix}:=\begin{pmatrix}m_{C,j}^1\\ m_{C,j}^2\end{pmatrix}$$

and

$$C_j=\begin{pmatrix}\frac{1}{t_j}\int_0^{t_j}\widetilde{a}_s\,\mathrm{d}s & \frac{1}{t_j^2}\int_0^{t_j}\widetilde{a}_s(t_j-s)\,\mathrm{d}s\\ \frac{1}{t_j^2}\int_0^{t_j}\widetilde{a}_s(t_j-s)\,\mathrm{d}s & \frac{1}{t_j^3}\int_0^{t_j}\widetilde{a}_s(t_j-s)^2\,\mathrm{d}s\end{pmatrix}.$$

The result eventually follows from (4.6), (4.7) and standard computations involving the mean value theorem. □

4.4. *Existence of the density for the aggregated frozen process*

Let $h_0>0$ be a given fixed time step. For $i\in\mathbb{N}$ set $t_i:=ih_0$. Fix $(x',y')\in\mathbb{R}^{2d}$, $t>0$. We consider the frozen model defined by $\widetilde{X}_0^{h_0}=x$, $\widetilde{Y}_0^{h_0}=y$ and for all $i\in\mathbb{N}$,

$$\begin{aligned}\widetilde{X}_{t_{i+1}}^{h_0}&=\widetilde{X}_{t_i}^{h_0}+b\big(x',y'\big)h_0+\sigma\big(x',y'-tx'\big)\sqrt{h_0}\widetilde{\xi}_{i+1},\\ \widetilde{Y}_{t_{i+1}}^{h_0}&=\widetilde{Y}_{t_i}^{h_0}+\widetilde{X}_{t_{i+1}}^{h_0}h_0=\widetilde{Y}_{t_i}^{h_0}+h_0\widetilde{X}_{t_i}^{h_0}+h_0^2b\big(x',y'\big)+h_0^{3/2}\sigma\big(x',y'-tx'\big)\widetilde{\xi}_{i+1},\end{aligned}\tag{4.8}$$

where $(\widetilde{\xi}_i)_{i\in\mathbb{N}^*}$ are i.i.d., centered with identity covariance. The aim of this section is to show that for i large enough $(\widetilde{X}_{t_i}^{h_0},\widetilde{Y}_{t_i}^{h_0})$ admits a density. We refer the reader to the work of Yurinski [27] or Molchanov and Varchenko [21] for related topics.

Conditionally to

$$\begin{pmatrix} \widetilde{X}^{h_0}_{t_i} = x^* \\ \widetilde{Y}^{h_0}_{t_i} = y^* \end{pmatrix}$$

and iterating the frozen model we get

$$\begin{aligned} \widetilde{X}^{h_0}_{t_{i+n}} &= x^* + (nh_0)b(x', y') + \sigma(x', y' - x't)\sqrt{nh_0}\widetilde{\xi}^{(1)}_{i,n}, \\ \widetilde{Y}^{h_0}_{t_{i+n}} &= y^* + (nh_0)x^* + \frac{\gamma_n}{2}(nh_0)^2 b(x', y') + (nh_0)^{3/2}\sigma(x', y' - tx')\widetilde{\xi}^{(2)}_{i,n}, \end{aligned} \tag{4.9}$$

where we recall $\gamma_n = (1 + \frac{1}{n})$ and

$$\begin{aligned} \widetilde{\xi}^{(1)}_{i,n} &= \frac{1}{\sqrt{n}}(\widetilde{\xi}_{i+1} + \widetilde{\xi}_{i+2} + \cdots + \widetilde{\xi}_{i+n}), \\ \widetilde{\xi}^{(2)}_{i,n} &= \frac{1}{\sqrt{n}}\left(\widetilde{\xi}_{i+1} + \left(1 - \frac{1}{n}\right)\widetilde{\xi}_{i+2} + \cdots + \left(1 - \frac{n-1}{n}\right)\widetilde{\xi}_{i+n}\right). \end{aligned}$$

We have

$$\begin{aligned} \operatorname{Var}(\widetilde{\xi}^{(2)}_{i,n}) &= \frac{(1 - (n-1)/n)^2 + \cdots + 1^2}{n} = \frac{2n^2 + 3n + 1}{6n^2} = \frac{1}{3}\gamma_n\left(1 + \frac{1}{2n}\right)\mathbf{I}_{d\times d}, \\ \operatorname{Cov}(\widetilde{\xi}^{(1)}_{i,n}, \widetilde{\xi}^{(2)}_{i,n}) &= \frac{(1 - (n-1)/n) + \cdots + 1}{n} = \frac{n+1}{2n} = \frac{\gamma_n}{2}\mathbf{I}_{d\times d}. \end{aligned}$$

Hence, the covariance matrix of the $2d$-dimensional vector $(\widetilde{\xi}^{(1)}_{i,n}, \widetilde{\xi}^{(2)}_{i,n})^*$ is non-degenerate for $n \geq 2$.

Estimating the characteristic function $\varphi_n(\tau_1, \tau_2)$ of the vector $(\widetilde{\xi}^{(1)}_{i,n}, \widetilde{\xi}^{(2)}_{i,n})^* \in \mathbb{R}^{2d}$ we derive the following proposition.

Proposition 4.2. *Let $\phi(\tau) := \mathbb{E}[\exp(\mathrm{i}\langle\widetilde{\xi}_1, \tau\rangle)]$, $\tau \in \mathbb{R}^d$ denote the characteristic function of the $(\widetilde{\xi}_i)_{i\in\mathbb{N}^*}$. If for all multi-index β, $|\beta| = S + 2d + 1$, $|D^\beta\phi(\tau)| \leq C(1 + |\tau|^{4+2d+1})^{-1}$, then for n large enough and for all multi-index α, $|\alpha| \leq 4$, one has*

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |(\tau_1, \tau_2)|^{|\alpha|} |D^{S+2d+1}\varphi_n(\tau_1, \tau_2)| \,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 < \infty.$$

In particular, by Fourier inversion the density

$$f_n(\theta_1, \theta_2) = \frac{1}{(2\pi)^{2d}}\int \exp(-\mathrm{i}\langle(\theta_1, \theta_2)^*, (\tau_1, \tau_2)^*\rangle)\varphi_n(\tau_1, \tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 \tag{4.10}$$

exists and there exists C s.t. for all multi-index ν, $|\nu| \leq 4$,

$$|D^\nu f_n(\theta_1, \theta_2)| \leq \frac{C}{1 + |(\theta_1, \theta_2)|^{S+2d+1}} := \psi_n(\theta_1, \theta_2).$$

Proof. Write

$$\begin{aligned} \varphi_n(\tau_1, \tau_2) &= \mathbb{E}[\exp\{\mathrm{i}\langle\tau_1, \widetilde{\xi}^{(1)}_{i,n}\rangle + \mathrm{i}\langle\tau_2, \widetilde{\xi}^{(2)}_{i,n}\rangle\}] \\ &= \prod_{j=0}^{n-1}\phi\left(\frac{\tau_1 + (1 - j/n)\tau_2}{\sqrt{n}}\right). \end{aligned} \tag{4.11}$$

We partition the space $\mathbb{R}^{2d}$ into the following disjoint sets

$$A_0 := \left\{(\tau_1, \tau_2) \in \mathbb{R}^{2d}\colon\ |\tau_1| \geq \left(1 - \frac{1}{n}\right)|\tau_2|\right\},$$
$$A_i := \left\{(\tau_1, \tau_2) \in \mathbb{R}^{2d}\colon\ \left(1 - \frac{i+1}{n}\right)|\tau_2| \leq |\tau_1| < \left(1 - \frac{i}{n}\right)|\tau_2|\right\}, \quad i \in [\![1, n-2]\!],$$
$$A_{n-1} := \left\{(\tau_1, \tau_2) \in \mathbb{R}^{2d}\colon\ |\tau_1| < \frac{1}{n}|\tau_2|\right\}.$$

If $(\tau_1, \tau_2) \in A_0$ then for $i \in [\![2, n-2]\!]$,

$$\left|\frac{\tau_1 + (1 - i/n)\tau_2}{\sqrt{n}}\right| \geq \frac{1}{\sqrt{n}}\left(|\tau_1| - \left(1 - \frac{i}{n}\right)|\tau_2|\right) \geq \frac{1}{\sqrt{n}}\left(\left(1 - \frac{1}{n}\right)|\tau_2| - \left(1 - \frac{i}{n}\right)|\tau_2|\right) = \frac{i-1}{n\sqrt{n}}|\tau_2|$$

and similarly $|\frac{\tau_1 + (1-i/n)\tau_2}{\sqrt{n}}| \geq \frac{i-1}{n\sqrt{n}}|\tau_1|$. Hence,

$$\left|\frac{\tau_1 + (1 - i/n)\tau_2}{\sqrt{n}}\right|^{2d+1} \geq \frac{((i-1)/2)^{2d+1}}{n^{3d+3/2}}\left|(\tau_1, \tau_2)\right|^{2d+1}. \tag{4.12}$$

If $(\tau_1, \tau_2) \in A_{i^*}$ for some i^*, $i^* \in [\![1, n-2]\!]$ and $l \in [\![2, n-1-i^*]\!]$ then elementary computations yield similarly

$$\left|\frac{\tau_1 + (1 - (i^* + l)/n)\tau_2}{\sqrt{n}}\right|^{2d+1} \geq \frac{((l-1)/2)^{2d+1}}{n^{3d+3/2}}\left|(\tau_1, \tau_2)\right|^{2d+1}, \tag{4.13}$$

and for $l \in [\![1, i^* - 1]\!]$,

$$\left|\frac{\tau_1 + (1 - (i^* - l)/n)\tau_2}{\sqrt{n}}\right|^{2d+1} \geq \frac{(l/2)^{2d+1}}{n^{3d+3/2}}\left|(\tau_1, \tau_2)\right|^{2d+1}. \tag{4.14}$$

If $(\tau_1, \tau_2) \in A_{n-1}$ then for $i \in [\![1, n-1]\!]$,

$$\left|\frac{\tau_1 + (1 - i/n)\tau_2}{\sqrt{n}}\right|^{2d+1} \geq \frac{(1/2)^{2d+1}}{n^{d+1/2}}\left(1 - \frac{i+1}{n}\right)^{2d+1}\left|(\tau_1, \tau_2)\right|^{2d+1}. \tag{4.15}$$

Use now the growth assumption on ϕ and the inequality $1 + \sum_{j=1}^N p_j \leq \prod_{j=1}^N (1 + p_j)$ where $p_j \geq 0$, to derive from (4.11)

$$\left|\varphi_n(\tau_1, \tau_2)\right| = \left|\prod_{j=0}^{n-1} \phi\left(\frac{\tau_1 + (1 - j/n)\tau_2}{\sqrt{n}}\right)\right| \leq \frac{C^n}{\prod_{j=0}^{n-1}(1 + |(\tau_1 + (1 - j/n)\tau_2)/\sqrt{n}|^{2d+1})}$$
$$\leq \frac{C^n}{1 + \sum_{j=0}^{n-1} |(\tau_1 + (1 - j/n)\tau_2)/\sqrt{n}|^{2d+1}}.$$

Now equations (4.12)–(4.15) yield that there exists n large enough s.t.

$$\left|\varphi_n(\tau_1, \tau_2)\right| \leq \frac{C(n)}{1 + |(\tau_1, \tau_2)|^{2d+1}},$$

where $C(n) \to_n +\infty$. Anyhow, for such a fixed n, one has $\varphi_n \in L^1(\mathbb{R}^{2d})$ which implies the existence of the density f_n of the vectors $(\widetilde{\xi}_{i,n}^{(1)}, \widetilde{\xi}_{i,n}^{(2)})^* \in \mathbb{R}^{2d}$. The properties concerning the growth and derivatives of f_n are derived from (4.10) and the growth and smoothness properties of ϕ. □

Hence we can set $(\eta_i^1, \eta_i^2) := (\widetilde{\xi}_{i,n}^{(1)}, \widetilde{\xi}_{i,n}^{(2)})$ where $(\widetilde{\xi}_{i,n}^{(1)}, \widetilde{\xi}_{i,n}^{(2)})$ are as in the above proposition. Introducing a "macro" scale time step $h = nh_0$, the discrete model (4.2) corresponds to the "aggregated" dynamics of (4.9). Set for all $(\theta_1, \theta_2) \in \mathbb{R}^{2d}$, $\psi(\theta_1, \theta_2) := \psi_n(\theta_1, \theta_2)$ defined in Proposition 4.2. With the notations of Section 4.2 one derives that $q_n(\theta_1, \theta_2) = f_n(\theta_1, \theta_2)$ satisfies (A2) with the above $\psi = \psi_n$.

References

[1] D. G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc.* **73** (1967) 890–896. MR0217444

[2] P. Baldi. Premières majorations de la densité d'une diffusion sur R^m, méthode de la parametrix. *Astérisques* **84–85** (1978) 43–53. MR0634964

[3] G. Ben Arous. Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus. *Ann. Sci. École Norm. Sup. (4)* **21** (1988) 307–331. MR0974408

[4] G. Ben Arous and R. Léandre. Décroissance exponentielle du noyau de la chaleur sur la diagonale, II. *Probab. Theory Related Fields* **90** (1991) 377–402. MR1133372

[5] E. Barucci, S. Polidoro and V. Vespri. Some results on partial differential equations and asian options. *Math. Models Methods Appl. Sci.* **3** (2001) 475–497. MR1830951

[6] R. Bhattacharya and R. Rao. *Normal Approximations and Asymptotic Expansions*. Wiley, New York, 1976. MR0436272

[7] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations, II. Convergence rate of the density. *Monte Carlo Methods Appl.* **2** (1996) 93–128. MR1401964

[8] P. Cattiaux. Calcul stochastique et opérateurs dégénérés du second ordre, I. Résolvantes, théorème de Hörmander et applications. *Bull. Sci. Math.* **114** (1990) 421–462. MR1077270

[9] P. Cattiaux. Calcul stochastique et opérateurs dégénérés du second ordre, II. Problème de Dirichlet. *Bull. Sci. Math.* **115** (1991) 81–122. MR1086940

[10] E. B. Dynkin. *Markov Processes*. Springer, Berlin, 1963. MR0193671

[11] A. Friedman. *Partial Differential Equations of Parabolic Type*. Prentice-Hall, Englewood Cliffs, 1964. MR0181836

[12] F. Hérau and F. Nier. Isotropic hypoellipticity and trend to equilibrium for the Fokker–Planck equation with a high-degree potential. *Arch. Ration. Mech. Anal.* **171** (2004) 151–218. MR2034753

[13] L. Hörmander. Hypoelliptic second order differential operators. *Acta. Math.* **119** (1967) 147–171. MR0222474

[14] V. D. Konakov and S. A Molchanov. On the convergence of Markov chains to diffusion processes. *Teor. Veroyatn. Mat. Statist. (in Russian)* **31** (1984) 51–64. English translation in *Theory Probab. Math. Statist.* **31** (1985) 59–73. MR0816126

[15] V. Konakov and E. Mammen. Local limit theorems for transition densities of Markov chains converging to diffusions. *Probab. Theory Related Fields* **117** (2000) 551–587. MR1777133

[16] V. Konakov, S. Menozzi and S. Molchanov. Explicit parametrix and local limit theorems for some degenerate diffusion processes, 2009. Available at http://hal.archives-ouvertes.fr/hal-00256588/fr/.

[17] A. N. Kolmogorov. Zufällige Bewegungen (zur Theorie der Brownschen Bewegung). *Ann. of Math. (2)* **35** (1934) 116–117. MR1503147

[18] S. Kusuoka and D Stroock. Applications of the Malliavin calculus, III. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **34** (1987) 391–442. MR0914028

[19] H. P. McKean and I. M. Singer. Curvature and the eigenvalues of the Laplacian. *J. Differential Geom.* **1** (1967) 43–69. MR0217739

[20] J Mattingly and A. Stuart. Geometric ergodicity of some hypo-elliptic diffusions for particle motions. Inhomogeneous random systems. *Markov Process. Related Fields* **8** (2004) 199–214. MR1924935

[21] S. A. Molchanov and A. N. Varchenko. Applications of the stationary phase method in limit theorems for Markov chains. *Dokl. Akad. Nauk SSSR* (Translated in *Soviet Math. Dokl.* (**18**) 265–269) **233** (1977) 11–14. MR0474506

[22] J. R. Norris. Simplified Malliavin calculus. *Séminaire de Probabilités*, XX 101–130. Springer, Berlin, 1986. MR0942019

[23] D. Nualart. *Malliavin Calculus and Related Topics*. Springer, New York, 1995. MR1344217

[24] D. W. Stroock. Diffusion semigroups corresponding to uniformly elliptic divergence form operators. *Séminaire de Probabilités*, XXII 316–347. Springer, Berlin, 1988. MR0960535

[25] D. Talay. Stochastic Hamiltonian systems: Exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Related Fields* **8** (2002) 163–198. MR1924934

[26] B. Lapeyre and E. Temam. Competitive Monte Carlo methods for the pricing of Asian Options. *Journal of Computational Finance* **5** (2001) 39–59.

[27] V. V. Yurinski. Estimates for the characteristic functions of certain degenerate multidimensional distributions. *Teor. Verojatn. Primen.* (Translated in *Theory Probab. Appl.* **22** 101–113) **17** (1972) 99–110. MR0297054