STABILITY OF DENSITIES FOR PERTURBED DIFFUSIONS AND MARKOV CHAINS [*],[**]

Valentin Konakov[1], A. Kozhina[2] and Stéphane Menozzi[3]

Abstract. We study the sensitivity of the densities of non degenerate diffusion processes and related Markov Chains with respect to a perturbation of the coefficients. Natural applications of these results appear in models with misspecified coefficients or for the investigation of the weak error of the Euler scheme with irregular coefficients.

1. Introduction

1.1. Setting

For a fixed given deterministic final horizon $T > 0$, let us consider the following multidimensional SDE:

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t, \ t \in [0, T], \tag{1.1}$$

where $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ are bounded coefficients that are measurable in time and Hölder continuous in space (this last condition will be possibly relaxed for the drift term b) and W is a Brownian motion on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. Also, $a(t, x) := \sigma\sigma^*(t, x)$ is assumed to be uniformly elliptic. In particular those assumptions guarantee that (1.1) admits a unique weak solution, see *e.g.* Bass and Perkins [3, 20] from which the uniqueness to the martingale problem for the associated generator can be derived under the current assumptions.

We now introduce, for a given parameter $\varepsilon > 0$, a perturbed version of (1.1) with dynamics:

$$\mathrm{d}X_t^{(\varepsilon)} = b_\varepsilon(t, X_t^{(\varepsilon)})\mathrm{d}t + \sigma_\varepsilon(t, X_t^{(\varepsilon)})\mathrm{d}W_t, \ t \in [0, T], \tag{1.2}$$

[1] Higher School of Economics, Shabolovka 31, Moscow, Russian Federation. vkonakov@hse.ru

[2] Higher School of Economics, Shabolovka 31, Moscow, Russian Federation and RTG 1953, Institute of Applied Mathematics, Heidelberg University, Germany. aakozhina@hse.ru

[3] Higher School of Economics, Shabolovka 31, Moscow, Russian Federation and LaMME, UMR CNRS 8070, Université d'Evry Val d'Essonne, 23 Boulevard de France, 91037 Evry, France. stephane.menozzi@univ-evry.fr

1 where $b_\varepsilon : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma_\varepsilon : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ satisfy at least the same assumptions as b, σ and are in
2 some sense meant to be *close* to b, σ when ε is small.

3   It is known that, under the previous assumptions, the density of the processes $(X_t)_{t \geq 0}, (X_t^{(\varepsilon)})_{t \geq 0}$ exists in
4 positive time and satisfies some Gaussian bounds, see *e.g.* Aronson [1,23] or [8] for extensions to some degenerate
5 cases.

6   The goal of this work is to investigate how the closeness of $(b_\varepsilon, \sigma_\varepsilon)$ and (b, σ) is reflected on the respective
7 densities of the associated processes. Important applications can for instance be found in mathematical finance.
8 If the dynamics of (1.1) models the evolution of the (log-)price of a financial asset, it is often very useful to
9 know how a perturbation of the volatility σ impacts the density, and therefore the associated option prices (see
10 also Corielli *et al.* [6] or Benhamou *et al.* [2] for related problems).

11   In the framework of parameter estimation it can be useful, having at hand estimators $(b_\varepsilon, \sigma_\varepsilon)$ of the true
12 parameters (b, σ) and some controls for the differences $|b - b_\varepsilon|, |\sigma - \sigma_\varepsilon|$ in a suitable sense, to quantify the
13 difference $p_\varepsilon - p$ of the densities corresponding respectively to the dynamics with the estimated parameters and
14 the one of the model.

15   Another important application includes the case of mollification by spatial convolution. This specific kind of
16 perturbation is useful to investigate the error between the densities of a non-degenerate diffusion of type (1.1)
17 with Hölder coefficients (or with piecewise smooth bounded drift) and its Euler scheme. In this framework, some
18 explicit convergence results can be found in [16].

19   More generally, this situation can appear in every applicative field for which the diffusion coefficient might
20 be misspecified.

21   The previously mentioned Gaussian bounds on the density are derived through the so-called *parametrix*
22 expansion which will be the crux of our approach. Roughly speaking, it consists in approximating the process
23 by a proxy which has a known density, here a Gaussian one, and then in investigating the difference through
24 the Kolmogorov equations. Various approaches to the parametrix expansion exist, see *e.g.* Il'in *et al.* [13],
25 Friedman [10] and McKean and Singer [22]. The latter approach will be the one used in this work since it
26 appears to be the most adapted to handle coefficients with no *a priori* smoothness in time and can also be
27 directly extended to the discrete case for Markov chain approximations of equations (1.1) and (1.2). Let us
28 mention in this setting the works of Konakov *et al.*, see [14,15].

29   Our stability results will also apply to two Markov chains with respective dynamics:

$$Y_{t_{k+1}} = Y_{t_k} + b(t_k, Y_{t_k})h + \sigma(t_k, Y_{t_k})\sqrt{h}\xi_{k+1}, Y_0 = x,$$

$$Y_{t_{k+1}}^{(\varepsilon)} = Y_{t_k}^{(\varepsilon)} + b_\varepsilon(t_k, Y_{t_k}^{(\varepsilon)})h + \sigma_\varepsilon(t_k, Y_{t_k}^{(\varepsilon)})\sqrt{h}\xi_{k+1}, Y_0^{(\varepsilon)} = x, \tag{1.3}$$

32 where $h > 0$ is a given time step, for which we denote for all $k \geq 0$, $t_k := kh$ and the $(\xi_k)_{k \geq 1}$ are centered
33 i.i.d. random variables satisfying some integrability conditions. Again, the key tool will be the parametrix
34 representation for the densities of the chains and the Gaussian local limit theorem.

1.2. Assumptions and main results

36   Let us introduce the following assumptions. Below, the parameter $\varepsilon > 0$ is fixed and the constants appearing
37 in the assumptions do not depend on ε.

38 (**A1**) (**Boundedness of the coefficients**). The components of the vector-valued functions $b(t,x), b_\varepsilon(t,x)$
39   and the matrix-functions $\sigma(t,x), \sigma_\varepsilon(t,x)$ are bounded measurable. Specifically, there exist constants
40        $K_1, K_2 > 0$ s.t.

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |b(t,x)| + \sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |b_\varepsilon(t,x)| \leq K_1,$$

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |\sigma(t,x)| + \sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |\sigma_\varepsilon(t,x)| \leq K_2.$$

(**A2**) (**Uniform Ellipticity**). The matrices $a := \sigma\sigma^*, a_\varepsilon := \sigma_\varepsilon\sigma_\varepsilon^*$ are uniformly elliptic, *i.e.* there exists $\Lambda \geq 1$, $\forall(t, x, \xi) \in [0, T] \times (\mathbb{R}^d)^2$,

$$\Lambda^{-1}|\xi|^2 \leq \langle a(t, x)\xi, \xi\rangle \leq \Lambda|\xi|^2, \Lambda^{-1}|\xi|^2 \leq \langle a_\varepsilon(t, x)\xi, \xi\rangle \leq \Lambda|\xi|^2.$$

(**A3**) (**Hölder continuity in space**). For some $\gamma \in (0, 1]$, $\kappa < \infty$, for all $t \in [0, T]$,

$$|\sigma(t, x) - \sigma(t, y)| + |\sigma_\varepsilon(t, x) - \sigma_\varepsilon(t, y)| \leq \kappa |x - y|^\gamma.$$

Observe that the last condition also readily gives, thanks to the boundedness of $\sigma, \sigma_\varepsilon$ that a, a_ε are also uniformly γ-Hölder continuous.

For a given $\varepsilon > 0$, we say that assumption (A) holds when conditions (A1)$-$(A3) are in force. Let us now introduce, under (A), the quantities that will bound the difference of the densities in our main results below. Set for $\varepsilon > 0$:

$$\Delta_{\varepsilon,b,\infty} := \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |b(t, x) - b_\varepsilon(t, x)|, \ \forall q \in (1, +\infty), \ \Delta_{\varepsilon,b,q} := \sup_{t\in[0,T]} \|b(t, \ldots) - b_\varepsilon(t, \ldots)\|_{L^q(\mathbb{R}^d)}.$$

Since $\sigma, \sigma_\varepsilon$ are both γ-Hölder continuous, see (A3) we also define

$$\Delta_{\varepsilon,\sigma,\gamma} := \sup_{u\in[0,T]} |\sigma(u, \ldots) - \sigma_\varepsilon(u, \ldots)|_\gamma, \tag{1.4}$$

where for $\gamma \in (0, 1]$, $|.|_\gamma$ stands for the usual Hölder norm in space on $C_b^\gamma(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)$ (space of Hölder continuous bounded functions, see *e.g.* Krylov [19]) *i.e.*:

$$|f|_\gamma := \sup_{x\in\mathbb{R}^d} |f(x)| + [f]_\gamma, \ [f]_\gamma := \sup_{x\neq y,(x,y)\in(\mathbb{R}^d)^2} \frac{|f(x) - f(y)|}{|x - y|^\gamma}.$$

The previous control in particular implies for all $(u, x, y) \in [0, T] \times (\mathbb{R}^d)^2$:

$$|a(u, x) - a(u, y) - a_\varepsilon(u, x) + a_\varepsilon(u, y)| \leq 2(K_2 + \kappa)\Delta_{\varepsilon,\sigma,\gamma}|x - y|^\gamma.$$

We eventually set for $q \in (1, +\infty]$,
$$\Delta_{\varepsilon,\gamma,q} := \Delta_{\varepsilon,\sigma,\gamma} + \Delta_{\varepsilon,b,q}, \tag{1.5}$$

which will be the key quantity governing the error in our results.

We will denote, from now on, by C a constant depending on the parameters appearing in (A) and T. We reserve the notation c for constants that only depend on (A) but not on T. The values of C, c may change from line to line and do not depend on the considered parameter ε. Also, for given integers $i, j \in \mathbb{N}$ s.t. $i < j$, we will denote by $[\![i, j]\!]$ the set $\{i, i + 1, \ldots, j\}$.

We are now in position to state our main results.

Theorem 1.1 (Stability control for diffusions). *Fix $\varepsilon > 0$ and a final deterministic time horizon $T > 0$. Under* (A) *and for $q > d$, there exist constants $C := C(q) \geq 1$ and $c \in (0, 1]$ s.t. for all $0 \leq s < t \leq T, (x, y) \in (\mathbb{R}^d)^2$:*

$$p_c(t - s, y - x)^{-1}|(p - p_\varepsilon)(s, t, x, y)| \leq C\Delta_{\varepsilon,\gamma,q}, \tag{1.6}$$

where $p(s, t, x, \ldots), p_\varepsilon(s, t, x, \ldots)$ respectively stand for the transition densities at time t of equations (1.1), (1.2) *starting from x at time s. Also, we denote for a given $c > 0$ and for all $u > 0, z \in \mathbb{R}^d$, $p_c(u, z) := \frac{c^{d/2}}{(2\pi u)^{d/2}} \exp(-c\frac{|z|^2}{2u})$.*

Remark 1.2 (About the constants)**.** We mention that the constant $C := C(q)$ in (1.6) explodes when $q \downarrow d$ and is decreasing in q (see Lems. 3.3 and 3.4 below for details). In particular, it can be chosen uniformly as soon as $q \geq q_0 > d$.

Before stating our results for Markov Chains we introduce two kinds of innovations in (1.3). Namely:

(I_G).   The i.i.d. random variables $(\xi_k)_{k \geq 1}$ are Gaussian, with law $\mathcal{N}(0, I_d)$. In that case the dynamics in (1.3) corresponds to the Euler discretization of equations (1.1) and (1.2).

($I_{P,M}$). For a given integer $M > 2d + 5 + \gamma$, the innovations $(\xi_k)_{k \geq 1}$ are centered and have C^5 density f_ξ which has, together with its derivatives up to order 5, at least polynomial decay of order M. Namely, for all $z \in \mathbb{R}^d$ and multi-index $\nu, |\nu| \leq 5$:

$$|D^\nu f_\xi(z)| \leq C Q_M(z), \tag{1.7}$$

where we denote for all $r > d$, $z \in \mathbb{R}^d$, $Q_r(z) := c_r \frac{1}{(1+|z|)^r}$, $\int_{\mathbb{R}^d} \mathrm{d}z Q_r(z) = 1$.

Theorem 1.3 (Stability control for Markov chains)**.** *Fix $\varepsilon > 0$ and a final deterministic time horizon $T > 0$. For $h = T/N$, $N \in \mathbb{N}^* := \mathbb{N} \setminus \{0\}$, we set for $i \in \mathbb{N}$, $t_i := ih$. Under* (A)*, assuming that either* (I_G) *or* ($I_{P,M}$) *holds, and for $q > d$ there exist $C := C(q) \geq 1, c \in (0,1]$ s.t. for all $0 \leq t_i < t_j \leq T, (x,y) \in (\mathbb{R}^d)^2$:*

$$\chi_c(t_j - t_i, y - x)^{-1} |(p^h - p_\varepsilon^h)(t_i, t_j, x, y)| \leq C \Delta_{\varepsilon, \gamma, q}, \tag{1.8}$$

where $p^h(t_i, t_j, x, \dots), p_\varepsilon^h(t_i, t_j, x, \dots)$ respectively stand for the transition densities at time t_j of the Markov Chains Y and $Y^{(\varepsilon)}$ in (1.3) *starting from x at time t_i. Also:*

– *If* (I_G) *holds:*

$$\chi_c(t_j - t_i, y - x) := p_c(t_j - t_i, y - x),$$

with p_c as in Theorem 1.1.

– *If* ($I_{P,M}$) *holds:*

$$\chi_c(t_j - t_i, y - x) := \frac{c^d}{(t_j - t_i)^{d/2}} Q_{M-(d+5+\gamma)} \left(\frac{|y-x|}{(t_j - t_i)^{1/2}/c} \right).$$

The behavior of $C := C(q)$ is similar to what is described in Remark 1.2.

1.3. On some related applications

1.3.1. Model sensitivity for option prices

Assume for instance that the (log)-price of a financial asset is given by the dynamics in (1.1). Under suitable assumptions the price of an option on that asset writes at time t and when $X_t = x$ as $\mathbb{E}[f(\exp(X_T^{t,x}))]$ up to an additional discounting factor. In the previous expression f is the pay-off function. For a rather large class of pay-offs, say measurable functions with polynomial growth, including irregular ones, Theorem 1.1 allows to specifically quantify how a perturbation of the coefficients impacts the option prices. Precisely for a given $\varepsilon > 0$, under (A):

$$|\mathcal{E}_\varepsilon(t, T, x, f)| := |\mathbb{E}\left[f\left(\exp(X_T^{t,x}) \right) \right] - \mathbb{E}\left[f(\exp(X_T^{(\varepsilon), t, x})) \right]| \leq C \Delta_{\varepsilon, \gamma, q} \int_{\mathbb{R}^d} f(\exp(y)) p_c(T - t, x, y) \mathrm{d}y.$$

This previous control can be as well exploited to investigate perturbations of a model which provides some closed formulas, *e.g.* a perturbation of the Black and Scholes model that would include a stochastic volatility taking for instance $\sigma_\varepsilon(x) = \sigma + \varepsilon \psi(x)$ for some bounded γ-Hölder continuous function ψ and ε small enough. In that case, assuming that the drift is known and unperturbed, we have $\Delta_{\varepsilon, \gamma, \infty} = |\sigma_\varepsilon - \sigma|_\gamma = \varepsilon |\psi|_\gamma$.

In connection with this application, we can quote the work of Corielli *et al.* [6] who give estimates on option prices through parametrix expansions truncating the series. Some of their results, see *e.g.* their Theorem 3.1, can be related to a perturbation analysis since they obtain an approximation of an option price for a local volatility model in terms of the Black-Scholes price and a correction term corresponding to the first order term in the parametrix series. A more probabilistic approach to similar problems can be found in Benhamou *et al.* [2]. However, none of the indicated works indeed deals with the global perturbation analysis we perform here.

1.3.2. Weak error analysis

It is well known that if the coefficients b, σ in (1.1) are smooth and a satisfies the non-degeneracy condition (A2), then the weak error on the densities for the approximation by the Euler scheme is well controlled. Precisely, for a given time step $h > 0$, let us set for $i \in \mathbb{N}, t_i := ih$. Introduce now, for a given starting time t_i, the Euler scheme $X_{t_i}^h = x, \ \forall j > i, \ X_{t_{j+1}}^h = X_{t_j}^h + b(t_j, X_{t_j}^h)h + \sigma(t_j, X_{t_j}^h)(W_{t_{j+1}} - W_{t_j})$ and denote by $p^h(t_i, t_j, x, \ldots)$ its density at time t_j. The dynamics of the Euler scheme clearly enters the scheme (1.3). It can be derived from Konakov and Mammen [15] (see also Bally and Talay [5] for an extension to the hypoelliptic setting) that:

$$|p - p^h|(t_i, t_j, x, y) \leq C h p_c(t_j - t_i, y - x).$$

If the coefficients in (1.1) are not smooth, it is then possible to use a mollification procedure, taking for $x \in \mathbb{R}^d$, $b_\varepsilon(t, x) := b(t, \ldots) \star \rho_\varepsilon(x), \sigma_\varepsilon(t, x) := \sigma(t, \ldots) \star \rho_\varepsilon(x)$ with $\rho_\varepsilon(x) := \varepsilon^{-d}\rho(x/\varepsilon)$ and $\rho \in C^\infty(\mathbb{R}^d, \mathbb{R}^+)$, $\int_{\mathbb{R}^d} \rho(x)\mathrm{d}x = 1, \mathrm{supp}(\rho) \subset K$ for some compact set K of $\mathbb{R}^d$. For the mollifying kernel ρ_ε, one then easily checks that for γ-Hölder continuous in space coefficients b, σ there exists C s.t.

$$\sup_{t \in [0,T]} |b(t, \ldots) - b_\varepsilon(t, \ldots)|_\infty \leq C\varepsilon^\gamma, \ \sup_{t \in [0,T]} |\sigma(t, \ldots) - \sigma_\varepsilon(t, \ldots)|_\eta \leq C\varepsilon^{\gamma-\eta}, \ \eta \in (0, \gamma). \tag{1.9}$$

The important aspect is that we lose a bit with respect to the sup norm when investigating the Hölder norm. We then have by Theorems 1.1 and 1.3 and their proof, that, for γ-Hölder continuous in space coefficients b, σ and taking $q = \infty$, there exist c, C s.t. for all $0 \leq s < t \leq T, 0 \leq t_i < t_j \leq T, \ (x, y) \in (\mathbb{R}^d)^2$:

$$|(p - p_\varepsilon)(s, t, x, y)| \leq C C_\eta \varepsilon^{\gamma-\eta} p_c(t - s, y - x), \ |(p^h - p_\varepsilon^h)(t_i, t_j, x, y)| \leq C C_\eta \varepsilon^{\gamma-\eta} p_c(t_j - t_i, y - x),$$

where the constant C_η explodes when η tends to 0.

To investigate the global weak error $(p - p^h)(t_i, t_j, x, y) = \{(p - p_\varepsilon) + (p_\varepsilon - p_\varepsilon^h) + (p_\varepsilon^h - p^h)\}(t_i, t_j, x, y)$, it therefore remains to analyze the contribution $(p_\varepsilon - p_\varepsilon^h)(t_i, t_j, x, y)$. The results of [15] indeed apply but yield $|(p_\varepsilon - p_\varepsilon^h)(t_i, t_j, x, y)| \leq C_\varepsilon h p_c(t_j - t_i, y - x)$ where C_ε is explosive when ε goes to zero. The global error thus writes:

$$|(p - p^h)(t_i, t_j, x, y)| \leq C\{C_\eta \varepsilon^{\gamma-\eta} + C_\varepsilon h\} p_c(t_j - t_i, y - x),$$

and a balance is needed to derive a global error bound. This is precisely the analysis we perform in [16]. In this work, we extend to densities (up to a slowly growing factor) the results previously obtained by Mikulevičius and Platen [21] on the weak error, *i.e.* they showed $|\mathbb{E}[f(X_T) - f(X_T^h)]| \leq C h^{\gamma/2}$ provided $f \in C_b^{2+\gamma}(\mathbb{R}^d, \mathbb{R})$. Precisely, we obtain through a suitable analysis of the constants C_η, C_ε, which respectively depend on behavior of the parametrix series and of the derivatives of the heat kernel with mollified coefficients, that $|p - p^h|(t_i, t_j, x, y) \leq C h^{\gamma/2 - \psi(h)} p_c(t_j - t_i, y - x)$ for a function $\psi(h)$ going to 0 as $h \to 0$ (which is induced by the previous loss of η in (1.9)). In the quoted work, we also obtain some error bounds for piecewise smooth drifts having a countable set of discontinuities. This part explicitly requires the stability result of Theorems 1.1, Theorems 1.3 for $q < +\infty$. The idea being that the difference between the piece-wise smooth drift and its smooth approximation (actually the mollification procedure is only required around the points of discontinuity), is well controlled in L^q norm, $q < +\infty$.

1.3.3. Extension to some Kinetic models

The results of Theorems 1.1 and 1.3 should extend without additional difficulties to the case of degenerate diffusions of the form:

$$\begin{aligned} \mathrm{d}X_t^1 &= b(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t, \\ \mathrm{d}X_t^2 &= X_t^1 \mathrm{d}t, \end{aligned} \tag{1.10}$$

1 denoting $X_t = (X_t^1, X_t^2)$, under the same previous assumptions on b, σ when we consider perturbations of the
2 non-degenerate components, *i.e.* for a given $\varepsilon > 0$, $X_t^{(\varepsilon)} = (X_t^{1,(\varepsilon)}, X_t^{2,(\varepsilon)})$ where:

$$\mathrm{d}X_t^{1,(\varepsilon)} = b_\varepsilon(t, X_t^{(\varepsilon)})\mathrm{d}t + \sigma_\varepsilon(t, X_t^{(\varepsilon)})\mathrm{d}W_t,$$
$$\mathrm{d}X_t^{2,(\varepsilon)} = X_t^{1,(\varepsilon)}\mathrm{d}t. \tag{1.11}$$

3 Indeed, under (A), the required parametrix expansions of the densities associated with the solutions of
4 equation (1.10), (1.11) have been established in [17].
5

6 *1.3.4. A posteriori Controls in Parameter Estimation*

7   Let us consider to illustrate this application a parametrized family of diffusions of the form:

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \sigma(\vartheta, t, X_t)\mathrm{d}W_t, \tag{1.12}$$

8 where $\vartheta \in \Theta \subset \mathbb{R}^d$, the coefficients b, σ are smooth, bounded and the non-degeneracy condition (A2) holds.
9 A natural *practical* problem consists in estimating the *true* parameter ϑ from an observed discrete sample
10 $(X_{t_i^n})_{i \in [\![0,n]\!]}$ where the $\{(t_i)_{i \in [\![0,n]\!]}\}$ form a partition of the observation interval, *i.e.* if $T = 1$, $0 = t_0^n < t_1^n < \cdots <$
11 $t_n^n = 1$.
12   Introducing the contrast:

13 $$U^n(\vartheta) := \frac{1}{n}\sum_{i=1}^n \left[\log\left(\det(a(\vartheta, t_{i-1}^n, X_{t_{i-1}^n}))\right) + \langle a^{-1}(\vartheta, t_{i-1}^n, X_{t_{i-1}^n})X_i^n, X_i^n\rangle\right], \ \forall i \in [\![1, n]\!], X_i^n := \frac{X_{t_i^n} - X_{t_{i-1}^n}}{\sqrt{t_i^n - t_{i-1}^n}},$$

14 and denoting by $\hat{\vartheta}_n$ the corresponding minimizer, it was shown by Genon$-$Catalot and Jacod [12] that under $\mathbb{P}^\vartheta$,
15 $\sqrt{n}(\hat{\vartheta}_n - \vartheta)$ converges in law towards a *mixed normal variable S* which is, conditionally to $\mathcal{F}_1 := \sigma\{(X_s)_{s\in[0,1]}\}$,
16 centered and Gaussian. For a precise expression of the covariance which depends on the whole path of $(X_t)_{t\in[0,1]}$
17 we refer to Theorem 3 and its proof in [12].
18   This means that, when n is large, conditionally to $\mathcal{F}_1$, we have on a subset $\bar{\Omega} \subset \Omega$ which has *high probability*,
19 that $|\hat{\vartheta}_n - \vartheta| \leq \frac{C}{\sqrt{n}}$ for a certain threshold C. Setting $\varepsilon_n = n^{-1/2}$, $\sigma_{\varepsilon_n}(t, x) := \sigma(\hat{\vartheta}_n, t, x)$ and with a slight abuse
20 of notation $\sigma(t, x) := \sigma(\vartheta, t, x)$, one gets that, on $\bar{\Omega}$:

21 $$|\sigma(t, x) - \sigma_{\varepsilon_n}(t, x) - (\sigma(t, y) - \sigma_{\varepsilon_n}(t, y))| \leq |x - y| \wedge Cn^{-1/2} \Rightarrow |\sigma - \sigma_\varepsilon|_\eta \leq (Cn^{-1/2})^{1-\eta}, \ \eta \in (0, 1].$$

22 We can then invoke our Theorem 1.1 to compare the densities of the diffusions with the estimated parameter
23 and the exact one in (1.12).
24   The paper is organized as follows. We recall in Section 2 some basic facts about parametrix expansions for
25 the densities of diffusions and Markov Chains. We then detail in Section 3 how to perform a stability analysis
26 of the parametrix expansions in order to derive the results of Theorems 1.1 and 1.3.

27   2. DERIVATION OF FORMAL SERIES EXPANSION FOR DENSITIES

28 **2.1. Parametrix representation of the density for diffusions**

29   In the following, for given $(s, x) \in \mathbb{R}^+ \times \mathbb{R}^d$, we use the standard Markov notation $(X_t^{s,x})_{t\geq s}$ to denote the
30 solution of (1.1) starting from x at time s.
31   Assume that $(X_t^{s,x})_{t\geq s}$ has for all $t > s$ a smooth density $p(s, t, x, \dots)$ (which is the case if additionally to
32 (A) the coefficients are smooth see *e.g.* Friedman [10]). We would like to estimate this density at a given point
33 $y \in \mathbb{R}^d$. To this end, we introduce the following Gaussian inhomogeneous process with spatial variable frozen at
34 y. For all $(s, x) \in [0, T] \times \mathbb{R}^d$, $t \geq s$ we set:

35 $$\tilde{X}_t^y = x + \int_s^t \sigma(u, y)\mathrm{d}W_u.$$

Its density $\tilde{p}^y$ readily satisfies the Kolmogorov backward equation:

$$\begin{cases} \partial_u \tilde{p}^y(u,t,z,y) + \tilde{L}_u^y \tilde{p}^y(u,t,z,y) = 0, \ s \leq u < t, z \in \mathbb{R}^d, \\ \tilde{p}^y(u,t,\ldots,y) \underset{u\uparrow t}{\to} \delta_y(.), \end{cases} \qquad (2.1)$$

where for all $\varphi \in C_0^2(\mathbb{R}^d, \mathbb{R})$ (twice continuously differentiable functions with compact support) and $z \in \mathbb{R}^d$:

$$\tilde{L}_u^y \varphi(z) = \frac{1}{2} \mathrm{Tr}\left(\sigma\sigma^*(u,y) D_z^2 \varphi(z)\right),$$

stands for the generator of $\tilde{X}^y$ at time u.

On the other hand, since we have assumed the density of X to be smooth, it must satisfy the Kolmogorov forward equation (see *e.g.* Dynkin [9]). For a given starting point $x \in \mathbb{R}^d$ at time s,

$$\begin{cases} \partial_u p(s,u,x,z) - L_u^* p(s,u,x,z) = 0, \ s < u \leq t, z \in \mathbb{R}^d, \\ p(s,u,x,\ldots) \underset{u\downarrow s}{\to} \delta_x(.), \end{cases} \qquad (2.2)$$

where L_u^* stands for the *formal* adjoint (which is again well defined if the coefficients in (1.1) are smooth) of the generator of (1.1) which for all $\varphi \in C_0^2(\mathbb{R}^d, \mathbb{R}), z \in \mathbb{R}^d$ writes:

$$L_u \varphi(z) = \frac{1}{2} \mathrm{Tr}\left(\sigma\sigma^*(u,z) D_z^2 \varphi(z)\right) + \langle b(u,z), D_z \varphi(z) \rangle.$$

Equations (2.1), (2.2) yield the formal expansion below which is initially due to McKean and Singer [22].

$$\begin{aligned}
(p - \tilde{p}^y)(s,t,x,y) &= \int_s^t \mathrm{d}u \partial_u \int_{\mathbb{R}^d} \mathrm{d}z\, p(s,u,x,z) \tilde{p}^y(u,t,z,y) \\
&= \int_s^t \mathrm{d}u \int_{\mathbb{R}^d} \mathrm{d}z \left(\partial_u p(s,u,x,z)\tilde{p}^y(u,t,z,y) + p(s,u,x,z)\partial_u \tilde{p}^y(u,t,z,y) \right) \\
&= \int_s^t \mathrm{d}u \int_{\mathbb{R}^d} \mathrm{d}z \left(L_u^* p(s,u,x,z)\tilde{p}^y(u,t,z,y) - p(s,u,x,z)\tilde{L}_u^y \tilde{p}^y(u,t,z,y) \right) \\
&= \int_s^t \mathrm{d}u \int_{\mathbb{R}^d} \mathrm{d}z\, p(s,u,x,z)(L_u - \tilde{L}_u^y)\tilde{p}^y(u,t,z,y), \qquad (2.3)
\end{aligned}$$

using the Dirac convergence for the first equality, equations (2.2) and (2.1) for the third one. We eventually take the adjoint for the last equality. Note carefully that the differentiation under the integral is also here formal since we would need to justify that it can actually be performed using some growth properties of the density and its derivatives which we *a priori* do not know.

Let us now introduce the notation

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

for the time-space convolution and let us define $\tilde{p}(s,t,x,y) := \tilde{p}^y(s,t,x,y)$, that is in $\tilde{p}(s,t,x,y)$ we consider the density of the frozen process at the final point and observe it at *that specific* point. We now introduce the *parametrix* kernel:

$$H(s,t,x,y) := (L_s - \tilde{L}_s)\tilde{p}(s,t,x,y) := (L_s - \tilde{L}_s^y)\tilde{p}^y(s,t,x,y). \qquad (2.4)$$

With those notations equation (2.3) rewrites:

$$(p - \tilde{p})(s,t,x,y) = p \otimes H(s,t,x,y).$$

From this expression, the idea then consists in iterating this procedure for $p(s, u, x, z)$ in (2.3) introducing the density of a process with frozen characteristics in z which is here the integration variable. This yields to iterated convolutions of the kernel and leads to the formal expansion:

$$p(s, t, x, y) = \sum_{r=0}^{\infty} \tilde{p} \otimes H^{(r)}(s, t, x, y), \tag{2.5}$$

where $\tilde{p} \otimes H^{(0)} = \tilde{p}, H^{(r)} = H \otimes H^{(r-1)}, \ r \geq 1$. Obtaining estimates on p from the formal expression (2.5) requires to have good controls on the right-hand side. The remarkable property of this formal expansion is now that the right-hand-side of (2.5) only involves controls on Gaussian densities which in particular will provide, associated with our assumption (A) a *smoothing* in time property for the kernel H.

Proposition 2.1. *Under the sole assumption* (A), *for $t > s$, the density of $X_t^{x,s}$ solving* (1.1) *exists and can be written as in* (2.5).

Proof. The proof can already be derived from a sensitivity argument. We first introduce two parametrix series of the form (2.5). Namely,

$$p(s, t, x, y) := \widetilde{p}(s, t, x, y) + \sum_{r=1}^{\infty} \tilde{p} \otimes H^{(r)}(s, t, x, y) \tag{2.6}$$

and

$$p_\varepsilon(s, t, x, y) := \widetilde{p}_\varepsilon(s, t, x, y) + \sum_{r=1}^{\infty} \tilde{p}_\varepsilon \otimes H_\varepsilon^{(r)}(s, t, x, y). \tag{2.7}$$

Let us point out that, at this stage, p and p_ε are defined as sums of series. The purpose is then to identify those sums with the densities of the processes $X_t^{s,x}, X_t^{(\varepsilon),s,x}$ at point y.

The convergence of the series (2.6) and (2.7) is in some sense *standard* (see *e.g.* [20] or Friedman [10]) under (A). We recall for the sake of completeness the key steps for (2.6).

From direct computations, there exist $c_1 \geq 1, c \in (0, 1]$ s.t. for all $T > 0$ and all multi-index $\alpha, |\alpha| \leq 8$,

$$\forall 0 \leq u < t \leq T, \ (z, y) \in (\mathbb{R}^d)^2, \ |D_z^\alpha \widetilde{p}(u, t, z, y)| \leq \frac{c_1}{(t-u)^{|\alpha|/2}} p_c(t-u, y-z), \tag{2.8}$$

where

$$p_c(t-u, y-z) = \frac{c^{d/2}}{(2\pi(t-u))^{d/2}} \exp\left(-\frac{c}{2} \frac{|y-z|^2}{t-u}\right),$$

stands for the usual Gaussian density in $\mathbb{R}^d$ with 0 mean and covariance $(t-u)c^{-1}I_d$. From (2.8), the boundedness of the drift and the Hölder continuity in space of the diffusion matrix we readily get that there exists $c_1 \geq 1, \ c \in (0, 1]$,

$$|H(u, t, z, y)| \leq \frac{c_1(1 \vee T^{(1-\gamma)/2})}{(t-u)^{1-\gamma/2}} p_c(t-u, z-y). \tag{2.9}$$

Now the key point is that the control (2.9) yields an integrable singularity giving a smoothing effect in time once integrated in space in the time-space convolutions appearing in (2.6) and (2.7). It follows by induction that:

$$|\widetilde{p} \otimes H^{(r)}(s, t, x, y)| \leq ((1 \vee T^{(1-\gamma)/2})c_1)^{r+1} \prod_{i=1}^{r} B\left(\frac{\gamma}{2}, 1 + (i-1)\frac{\gamma}{2}\right) p_c(t-s, y-x)(t-s)^{\frac{r\gamma}{2}}$$

$$= \frac{((1 \vee T^{(1-\gamma)/2})c_1)^{r+1} \left[\Gamma(\frac{\gamma}{2})\right]^r}{\Gamma(1 + r\frac{\gamma}{2})} p_c(t-s, y-x)(t-s)^{\frac{r\gamma}{2}}, \tag{2.10}$$

where for $a, b > 0, B(a,b) = \int_0^1 t^{-1+a}(1-t)^{-1+b}dt$ stands for the β-function, and using as well the identity $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ for the last equality. These bounds readily yield the convergence of the series as well as a Gaussian upper-bound. Namely

$$p(s,t,x,y) \le c_1 \exp((1 \vee T^{(1-\gamma)/2})c_1[(t-s)^{\gamma/2}])p_c(t-s, y-x). \tag{2.11}$$

An important application of the *stability* of the perturbation consists in considering coefficients $b_\varepsilon := b \star \zeta_\varepsilon, \sigma_\varepsilon := \sigma \star \zeta_\varepsilon$ in (2.7), where ζ_ε is a mollifier in time and space. For mollified coefficients which satisfy the non-degeneracy assumption (A2), the existence and smoothness of the density of $X^{(\varepsilon)}$ is a standard property (see *e.g.* Friedman [10,11]). It is also well known, see [14], that this density p_ε admits the series representation given in (2.7). Observe now carefully that the previous Gaussian bounds also hold for p_ε uniformly in ε and independently of the mollifying procedure. This therefore gives that

$$p_\varepsilon(s,t,x,y) \xrightarrow[\varepsilon \to 0]{} p(s,t,x,y), \tag{2.12}$$

boundedly and uniformly. Thus, for every continuous bounded function f we derive from the bounded convergence theorem and (2.11) that for all $0 \le s < t \le T, \ x \in \mathbb{R}^d$:

$$\mathbb{E}[f(X_t^{(\varepsilon),s,x})] = \int_{\mathbb{R}^d} f(y)p_\varepsilon(s,t,x,y)dy \xrightarrow[\varepsilon \to 0]{} \int_{\mathbb{R}^d} f(y)p(s,t,x,y)dy. \tag{2.13}$$

Taking $f = 1$ gives that $\int_{\mathbb{R}^d} p(s,t,x,y)dy = 1$ and the uniform convergence in (2.12) gives that $p(s,t,x,\ldots)$ is non negative. We therefore derive that $p(s,t,x,\cdot)$ is a probability density on $\mathbb{R}^d$.

On the other hand, under (A), we can derive from Theorem 11.3.4 of [25] that $(X_s^\varepsilon)_{s \in [0,T]} \xRightarrow[\varepsilon \to 0]{(\text{law})} (X_s)_{s \in [0,T]}$. This gives that for any bounded continuous function f:

$$\mathbb{E}\left[f\left(X_t^{(\varepsilon),s,x}\right)\right] \xrightarrow[\varepsilon \to 0]{} \mathbb{E}\left[f\left(X_t^{s,x}\right)\right].$$

This convergence and (2.13) then yield that the random variable $X_t^{s,x}$ admits $p(s,t,x,\cdot)$ as density.

We can thus now conclude that the processes $X, X^{(\varepsilon)}$ in (1.1), (1.2) have transition densities given by the sum of the series (2.6), (2.7). $\qquad\square$

2.2. Parametrix for Markov chains

One of the main advantages of the formal expansion in (2.5) is that it has a direct discrete counterpart in the Markov chain setting. Indeed, denote by $(Y_{t_j}^{t_i,x})_{j \ge i}$ the Markov chain with dynamics (1.3) starting from x at time t_i. Observe first that if the innovations $(\xi_k)_{k \ge 1}$ have a density then so does the chain at time t_k.

Let us now introduce its generator at time t_i, *i.e.* for all $\varphi \in C_0^2(\mathbb{R}^d, \mathbb{R}), \ x \in \mathbb{R}^d$:

$$L_{t_i}^h \varphi(x) := h^{-1} \mathbb{E}\left[\varphi(Y_{t_{i+1}}^{t_i,x}) - \varphi(x)\right].$$

In order to give a representation of the density of $p^h(t_i, t_j, x, y)$ of $Y_{t_j}^{t_i,x}$ at point y for $j > i$, we introduce similarly to the continuous case, the Markov chain (or inhomogeneous random walk) with coefficients frozen in space at y. For given $(t_i, x) \in [0,T] \times \mathbb{R}^d, \ t_j \ge t_i$ we set:

$$\tilde{Y}_{t_j}^{t_i,x,y} := x + h^{1/2}\sum_{k=i}^{j-1} \sigma(t_k, y)\xi_{k+1},$$

and denote its density by $\tilde{p}^{h,y}(t_i, t_j, x, \ldots)$. Its generator at time t_i writes for all $\varphi \in C_0^2(\mathbb{R}^d, \mathbb{R}), x \in \mathbb{R}^d$:

$$\tilde{L}_{t_i}^{h,y}\varphi(x) = h^{-1}\mathbb{E}\left[\varphi(\tilde{Y}_{t_{i+1}}^{t_i,x,y}) - \varphi(x)\right].$$

Using the notation $\tilde{p}^h(t_i, t_j, x, y) := \tilde{p}^{h,y}(t_i, t_j, x, y)$, we introduce now for $0 \leq i < j \leq N$ the *parametrix* kernel:

$$H^h(t_i, t_j, x, y) := \left(L_{t_i}^h - \tilde{L}_{t_i}^{h,y} \right) \tilde{p}^h \left(t_i + h, t_j, x, y \right).$$

Analogously to Lemma 3.6 in [14], which follows from a direct algebraic manipulation, we derive the following representation for the density which can be viewed as the Markov chain analogue of Proposition 2.1.

Proposition 2.2 (Parametrix expansion for the Markov chain). *Assume* (A) *is in force. Then, for* $0 \leq t_i < t_j \leq T$,

$$p^h(t_i, t_j, x, y) = \sum_{r=0}^{j-i} \tilde{p}^h \otimes_h H^{h,(r)}(t_i, t_j, x, y),$$

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

$$f \otimes_h g(t_i, t_j, x, y) = \sum_{k=0}^{j-i-1} h \int_{\mathbb{R}^d} f(t_i, t_{i+k}, x, z) g(t_{i+k}, t_j, z, y) \mathrm{d}z.$$

Also $g \otimes_h H^{h,(0)} = g$ *and for all* $r \geq 1, H^{h,(r)} := H^h \otimes_h H^{h,(r-1)}$ *denotes the r-fold discrete convolution of the kernel* H^h.

3. STABILITY OF PARAMETRIX SERIES

We will now investigate more specifically the sensitivity of the density w.r.t. the coefficients through the difference of the series. For a given fixed parameter ε, under (A) the densities $p(s, t, x, \ldots), p_\varepsilon(s, t, x, \cdot)$ at time t of the processes in (1.1), (1.2) starting from x at time s both admit a parametrix expansion of the previous type.

3.1. Stability for diffusions: Proof of Theorem 1.1

Let us consider the difference between two parametrix expansions:

$$|p(s, t, x, y) - p_\varepsilon(s, t, x, y)| = \left| \sum_{r=0}^{\infty} \tilde{p} \otimes H^{(r)}(s, t, x, y) - \sum_{r=0}^{\infty} \tilde{p}_\varepsilon \otimes H_\varepsilon^{(r)}(s, t, x, y) \right|$$

$$\leq |(\tilde{p} - \tilde{p}_\varepsilon)(s, t, x, y)| + \left| \sum_{r=1}^{\infty} \tilde{p} \otimes H^{(r)}(s, t, x, y) - \sum_{r=1}^{\infty} \tilde{p}_\varepsilon \otimes H_\varepsilon^{(r)}(s, t, x, y) \right|. \quad (3.1)$$

The strategy to study the above difference, using some well known properties of the Gaussian kernels and their derivatives recalled in (2.8), consists in first studying the difference of the *main* terms.

We have the following Lemma.

Lemma 3.1 (Difference of the first terms and their derivatives). *Under (A), there exist* $c_1 \geq 1, \ c \in (0,1]$ *s.t. for all* $0 \leq s < t, \ (x,y) \in (\mathbb{R}^d)^2$ *and all multi-index* $\alpha, \ |\alpha| \leq 4$,

$$|D_x^\alpha \tilde{p}(s, t, x, y) - D_x^\alpha \tilde{p}_\varepsilon(s, t, x, y)| \leq \frac{c_1}{(t-s)^{|\alpha|/2}} \Delta_{\varepsilon, \sigma, \gamma} p_c(t-s, y-x).$$

Proof. Let us first consider $|\alpha| = 0$ and introduce some notations. Set:

$$\Sigma(s, t, y) := \int_s^t a(u, y) \mathrm{d}u, \ \Sigma_\varepsilon(s, t, y) := \int_s^t a_\varepsilon(u, y) \mathrm{d}u. \quad (3.2)$$

Let us now identify the columns of the matrices $\Sigma(s,t,y), \Sigma_\varepsilon(s,t,y)$ with d-dimensional column vectors, *i.e.* for $\Sigma(s,t,y)$:

$$\Sigma(s,t,y) = \left(\Sigma^1 \big| \Sigma^2 \big| \ldots \big| \Sigma^d \right)(s,t,y).$$

We now rewrite $\tilde{p}(s,t,x,y)$ and $\tilde{p}_\varepsilon(s,t,x,y)$ in terms of vectors $\Theta(s,t,y), \Theta_\varepsilon(s,t,y) \in \mathbb{R}^{d^2}$:

$$\tilde{p}(s,t,x,y) = f_{x,y}(\Theta(s,t,y)), \ \Theta(s,t,y) = ((\Sigma^1)^*, \ldots, (\Sigma^d)^*)^*(s,t,y),$$
$$\tilde{p}_\varepsilon(s,t,x,y) = f_{x,y}(\Theta_\varepsilon(s,t,y)), \ \Theta_\varepsilon(s,t,y) = ((\Sigma_\varepsilon^1)^*, \ldots, (\Sigma_\varepsilon^d)^*)^*(s,t,y),$$

where $(\cdot)^*$ stands for the transpose and

$$f_{x,y} : \mathbb{R}^{d^2} \to \mathbb{R}$$
$$\Gamma \mapsto f_{x,y}(\Gamma) = \frac{1}{(2\pi)^{d/2}\det(\Gamma^{1:d})^{1/2}} \exp\left(-\frac{1}{2}\langle (\Gamma^{1:d})^{-1}(y-x), y-x \rangle\right), \tag{3.3}$$

where $\Gamma := \begin{pmatrix} \Gamma^1 \\ \Gamma^2 \\ \vdots \\ \Gamma^d \end{pmatrix}$ and each $(\Gamma^i)_{i\in[\![1,d]\!]}$ belongs to $\mathbb{R}^d$. Also, we have denoted:

$$\Gamma^{1:d} := \left(\Gamma^1 \big| \Gamma^2 \big| \ldots \big| \Gamma^d \right),$$

the $d \times d$ matrix formed with the entries $(\Gamma_i)_{i\in[\![1,d]\!]}$, each entry being viewed as a column.

The multidimensional Taylor expansion now gives:

$$(\tilde{p} - \tilde{p}_\varepsilon)(s,t,x,y)| = |f_{x,y}(\Theta(s,t,y)) - f_{x,y}(\Theta_\varepsilon(s,t,y))|$$
$$= \Bigg| \sum_{|\nu|=1} D_\nu f_{x,y}(\Theta(s,t,y))\{(\Theta_\varepsilon - \Theta)(s,t,y)\}^\nu$$
$$+ 2\sum_{|\nu|=2} \frac{\{(\Theta_\varepsilon - \Theta)(s,t,y)\}^\nu}{\nu!} \int_0^1 (1-\delta)D^\nu f_{x,y}([\Theta + \delta(\Theta_\varepsilon - \Theta)](s,t,y))\mathrm{d}\delta \Bigg|, \tag{3.4}$$

where for a multi-index $\nu := (\nu_1, \ldots, \nu_{d^2}) \in \mathbb{N}^{d^2}$, we denote by $|\nu| := \sum_{i=1}^{d^2} \nu_i$ the length of the multi-index, $\nu! = \prod_{i=1}^{d^2} \nu_i!$ and for $h \in \mathbb{R}^{d^2}$, $h^\nu := \prod_{i=1}^{d^2} h_i^{\nu_i}$ (with the convention that $0^0 = 1$). With these notations, from (3.2)−(3.4) we get:

$$|f_{x,y}(\Theta(s,t,y)) - f_{x,y}(\Theta_\varepsilon(s,t,y))| \leq c \Bigg\{ \sum_{|\nu|=1} |D^\nu f_{x,y}(\Theta(s,t,y))|\Delta_{\varepsilon,\sigma,\gamma}(t-s)$$
$$+ \Delta_{\varepsilon,\sigma,\gamma}^2(t-s)^2 \max_{\delta\in[0,1]} \sum_{|\nu|=2} |D^\nu f_{x,y}([\Theta + \delta(\Theta_\varepsilon - \Theta)](s,t,y))| \Bigg\}, \tag{3.5}$$

recalling that $\Delta_{\varepsilon,\sigma,\gamma}$ has previously been defined in (1.4)[4].

---

[4]Actually only the sup norm term in $\Delta_{\varepsilon,\sigma,\gamma}$ could be considered for this part of the anlalysis.

Since $f_{x,y}$ in (3.3) is a Gaussian density in the parameters x, y, we recall from Cramer and Leadbetter [7] (see Eq. (2.10.3) therein), that for all $\Gamma \in \mathbb{R}^{d^2}$ and any multi index ν, $|\nu| \leq 2$:

$$D^\nu f_{x,y}(\Gamma) = \frac{1}{2^{|\nu|}} \left(\prod_{i=1}^{d^2} \left(\frac{\partial^2}{\partial_{x_{\lfloor \frac{i-1}{d} \rfloor + 1}} \partial_{x_{i - \lfloor \frac{i-1}{d} \rfloor d}}} \right)^{\nu_i} f_{x,y}(\Gamma) \right),$$

where $\lfloor \cdot \rfloor$ stands for the integer part. Hence, taking from (3.5), for all $\delta \in [0,1]$, $\Gamma_{\varepsilon,\delta}(s,t,y) := [\Theta + \delta(\Theta_\varepsilon - \Theta)](s,t,y)$ yields, thanks to the non-degeneracy conditions (see Eq. (2.8)):

$$|D^\nu f_{x,y}(\Gamma_{\varepsilon,\delta}(s,t,y))| \leq \frac{\bar{c}_1}{(t-s)^{|\nu|}} f_{x,y}\big(\bar{c}\Gamma_{\varepsilon,\delta}(s,t,y)\big) \leq \frac{\bar{c}_1}{(t-s)^{|\nu|}} p_{\bar{c}}(t-s, y-x), \tag{3.6}$$

for some $\bar{c}_1 \geq 1, \bar{c} \in (0,1]$.

Thus, from (3.3), (3.4), equations (3.5) and (3.6) give:

$$|\tilde{p}(s,t,x,y) - \tilde{p}_\varepsilon(s,t,x,y)| \leq \bar{c}_1 \Delta_{\varepsilon,\sigma,\gamma} p_{\bar{c}}(t-s, y-x).$$

Up to a modification of $\bar{c}_1, \bar{c}$ or c_1, c in (2.8) we can assume that the statement of the lemma and (2.8) hold with the same constants c_1, c. The bounds for the derivatives are established similarly using the controls of (2.8). This concludes the proof. $\qquad \square$

Remark 3.2. Observe from equation (3.4) that the previous Lemma still holds with $\Delta_{\varepsilon,\sigma,\gamma}$ replaced by $\Delta_{\varepsilon,\sigma,\infty} := \sup_{t \in [0,T]} |\sigma(t,\ldots) - \sigma_\varepsilon(t,\ldots)|_\infty$. The Hölder norm is required to control the differences of the parametrix kernels.

The previous lemma quantifies how close are the main parts of the expansions. To proceed we need to consider the difference between the one-step convolutions. Combining the estimates of Lemmas 3.1 and 3.3 below will yield by induction the result stated in Theorem 1.1.

Lemma 3.3 (Control of the one-step convolution). *For all $0 \leq s < t \leq T$, $(x,y) \in (\mathbb{R}^d)^2$ and for $q \in (d, +\infty]$:*

$$|\tilde{p} \otimes H^{(1)}(s,t,x,y) - \tilde{p}_\varepsilon \otimes H_\varepsilon^{(1)}(s,t,x,y)|$$

$$\leq c_1^2 p_c(t-s, y-x) \Bigg\{ 2(1 \vee T^{(1-\gamma)/2})^2 [\Delta_{\varepsilon,\sigma,\gamma} + \mathbb{I}_{q=+\infty} \Delta_{\varepsilon,b,+\infty}] B\left(1, \frac{\gamma}{2}\right)(t-s)^{\frac{\gamma}{2}}$$

$$+ \mathbb{I}_{q \in (d,+\infty)} \Delta_{\varepsilon,b,q} B\left(\frac{1}{2} + \alpha(q), \alpha(q)\right)(t-s)^{\alpha(q)} \Bigg\}, \tag{3.7}$$

where c_1, c are as in Lemma 3.1 and for $q \in (d, +\infty]$ we set $\alpha(q) = \frac{1}{2}(1 - \frac{d}{q})$. The above control then yields for a fixed $q \in (d, +\infty]$:

$$|\tilde{p} \otimes H^{(1)}(s,t,x,y) - \tilde{p}_\varepsilon \otimes H_\varepsilon^{(1)}(s,t,x,y)|$$

$$\leq 2\bar{C}^2 \Delta_{\varepsilon,\gamma,q} p_c(t-s, y-x)(t-s)^{\frac{\gamma}{2} \wedge \alpha(q)} \left(B\left(1, \frac{\gamma}{2}\right) \vee B\left(\frac{1}{2} + \alpha(q), \alpha(q)\right) \right), \quad \bar{C} = c_1(1 \vee T^{(1-\gamma)/2}), \tag{3.8}$$

which will be useful for the iteration (see Lem. 3.4).

Proof. Let us write:

$$|\tilde{p} \otimes H^{(1)}(s,t,x,y) - \tilde{p}_\varepsilon \otimes H_\varepsilon^{(1)}(s,t,x,y)| \leq |\tilde{p} - \tilde{p}_\varepsilon| \otimes |H|(s,t,x,y) + \tilde{p}_\varepsilon \otimes |H - H_\varepsilon|(s,t,x,y) =: I + II. \tag{3.9}$$

From Lemma 3.1 and (2.9) we readily get for all $q \in (d, +\infty]$:

$$|\tilde{p} - \tilde{p}_\varepsilon| \otimes |H|(s,t,x,y) \le ((1 \vee T^{(1-\gamma)/2})c_1)^2 \Delta_{\varepsilon,\gamma,q} p_c(t-s, y-x) B\left(1, \frac{\gamma}{2}\right)(t-s)^{\frac{\gamma}{2}}. \qquad (3.10)$$

Now we will establish that for all $0 \le u < t \le T$, $(z,y) \in (\mathbb{R}^d)^2$ and $q = +\infty$:

$$|(H - H_\varepsilon)(u,t,z,y)| \le \Delta_{\varepsilon,\gamma,\infty} \frac{(1 \vee T^{(1-\gamma)/2})c_1}{(t-u)^{1-\frac{\gamma}{2}}} p_c(t-u, y-z). \qquad (3.11)$$

Equations (3.11) and (2.8) give that II can be handled as I which yields the result for $q = +\infty$. It therefore remains to prove (3.11). Let us write with the notations of (3.3):

$$(H - H_\varepsilon)(u,t,z,y) := \left[\frac{1}{2}\mathrm{Tr}\Big((a(u,z) - a(u,y))D_z^2 f_{z,y}\big(\Theta(u,t,y)\big)\Big) + \langle b(u,z), D_z f_{z,y}\big(\Theta(u,t,y)\big)\rangle\right]$$
$$- \left[\frac{1}{2}\mathrm{Tr}\Big((a_\varepsilon(u,z) - a_\varepsilon(u,y))D_z^2 f_{z,y}\big(\Theta_\varepsilon(u,t,y)\big)\Big) + \langle b_\varepsilon(u,z), D_z f_{z,y}\big(\Theta_\varepsilon(u,t,y)\big)\rangle\right].$$

Thus,

$$(H - H_\varepsilon)(u,t,z,y) = \frac{1}{2}\left[\mathrm{Tr}\Big((a(u,z) - a(u,y))\{D_z^2 f_{z,y}\big(\Theta(u,t,y)\big) - D_z^2 f_{z,y}\big(\Theta_\varepsilon(u,t,y)\big)\}\Big)\right.$$
$$\left. - \mathrm{Tr}\Big([(a_\varepsilon(u,z) - a_\varepsilon(u,y) - (a(u,z) - a(u,y))]D_z^2 f_{z,y}\big(\Theta_\varepsilon(u,t,y)\big)\Big)\right]$$
$$+ \left[\langle b(u,z), \{D_z f_{z,y}\big(\Theta(u,t,y)\big) - D_z f_{z,y}\big(\Theta_\varepsilon(u,t,y)\big)\}\rangle - \langle(b_\varepsilon(u,z) - b(u,z)), D_z f_{z,y}\big(\Theta_\varepsilon(u,t,y)\big)\rangle\right]. \quad (3.12)$$

Observe now that, similarly to (3.6) one has for all $i \in \{1,2\}$:

$$|D_z^i f_{z,y}\big(\Theta(u,t,y)\big)| + |D_z^i f_{z,y}\big(\Theta_\varepsilon(u,t,y)\big)| \le \frac{\tilde{c}_1}{(t-u)^{i/2}} p_{\tilde{c}}(t-u, y-z),$$

$$|D_z^i f_{z,y}\big(\Theta(u,t,y)\big) - D_z^i f_{z,y}\big(\Theta_\varepsilon(u,t,y)\big)| \le \frac{\tilde{c}_1 \Delta_{\varepsilon,\sigma,\gamma}}{(t-u)^{i/2}} p_{\tilde{c}}(t-u, y-z).$$

Also,

$$|(a_\varepsilon(u,z) - a_\varepsilon(u,y) - (a(u,z) - a(u,y))| \le c\Delta_{\varepsilon,\sigma,\gamma}|z-y|^\gamma, |b_\varepsilon(u,z) - b(u,z)| \le c\Delta_{\varepsilon,b,\infty}.$$

Thus, provided that c_1, c have been chosen large and small enough respectively in Lemma 3.1, the definition in (1.5) gives:

$$|(H - H_\varepsilon)(u,t,z,y)| \le \frac{(1 \vee T^{(1-\gamma)/2})c_1 \Delta_{\varepsilon,\gamma,\infty}}{(t-u)^{1-\gamma/2}} p_c(t-u, y-z).$$

This establishes (3.11) for $q = +\infty$. For $q \in (d, +\infty)$ we have to use Hölder's inequality in the time-space convolution involving the difference of the drifts (last term in (3.12)). Set:

$$D(s,t,x,y) := \int_s^t \mathrm{d}u \int_{\mathbb{R}^d} \tilde{p}_\varepsilon(s,u,x,z)\langle(b_\varepsilon(u,z) - b(u,z)), D_z f_{z,y}\big(\Theta_\varepsilon(u,t,y)\big)\rangle \mathrm{d}z.$$

Denoting by $\bar{q}$ the conjugate of q, *i.e.* $q, \bar{q} > 1, q^{-1} + \bar{q}^{-1} = 1$, we get from (2.8) and for $q > d$ that:

$$|D(s,t,x,y)| \leq c_1^2 \int_s^t \frac{\mathrm{d}u}{(t-u)^{1/2}} \|b(u,\dots) - b_\varepsilon(u,\dots)\|_{L^q(\mathbb{R}^d)} \left\{ \int_{\mathbb{R}^d} [p_c(u-s,z-x)p_c(t-u,y-z)]^{\bar{q}} \mathrm{d}z \right\}^{1/\bar{q}}$$

$$\leq c_1^2 \Delta_{\varepsilon,b,q} \int_s^t \frac{c^d}{(2\pi)^{d\left(1-\frac{1}{\bar{q}}\right)}(c\bar{q})^{d/\bar{q}}} \left\{ \int_{\mathbb{R}^d} p_{c\bar{q}}(u-s,z-x)p_{c\bar{q}}(t-u,y-z)\mathrm{d}z \right\}^{1/\bar{q}} \frac{\mathrm{d}u}{(u-s)^{\frac{d}{2}\left(1-\frac{1}{\bar{q}}\right)}(t-u)^{\frac{1}{2}+\frac{d}{2}(1-\frac{1}{\bar{q}})}}$$

$$\leq c_1^2 \left(\frac{c(t-s)}{2\pi} \right)^{\frac{d}{2}\left(1-\frac{1}{\bar{q}}\right)} \bar{q}^{-\frac{d}{2\bar{q}}} \Delta_{\varepsilon,b,q} p_c(t-s,y-x) \int_s^t \frac{\mathrm{d}u}{(u-s)^{\frac{d}{2}\left(1-\frac{1}{\bar{q}}\right)}(t-u)^{\frac{1}{2}+\frac{d}{2}\left(1-\frac{1}{\bar{q}}\right)}}.$$

Now, the constraint $d < q < +\infty$ precisely gives that $1 < \bar{q} < d/(d-1) \Rightarrow \frac{1}{2} + \frac{d}{2}(1 - \frac{1}{\bar{q}}) < 1$ so that the last integral is well defined. We therefore derive:

$$|D(s,t,x,y)| \leq c_1^2 (t-s)^{\frac{1}{2}-\frac{d}{2}(1-\frac{1}{\bar{q}})} \Delta_{\varepsilon,b,q} p_c(t-s,y-x) B\left(1 - \frac{d}{2}\left(1-\frac{1}{\bar{q}}\right), \frac{1}{2} - \frac{d}{2}\left(1-\frac{1}{\bar{q}}\right) \right).$$

In the case $d < q < +\infty$, recalling that $\alpha(q) = \frac{1}{2}(1 - \frac{d}{q})$, we eventually get :

$$\tilde{p}_\varepsilon \otimes |H - H_\varepsilon|(s,t,x,y) \leq c_1^2 p_c(t-s,y-x) \left\{ \Delta_{\varepsilon,b,q}(t-s)^{\alpha(q)} B\left(\frac{1}{2} + \alpha(q), \alpha(q) \right) \right.$$

$$\left. + 2\Delta_{\varepsilon,\sigma,\gamma}(1 \vee T^{(1-\gamma)/2})(t-s)^{\gamma/2} B(1,\gamma/2) \right\}. \tag{3.13}$$

The statement now follows in whole generality from (3.9), (3.10), equations (3.11), (2.8) for $q = \infty$ and (3.13) for $d < q < +\infty$. □

The following Lemma associated with Lemmas 3.1 and 3.3 allows to complete the proof of Theorem 1.1.

Lemma 3.4 (Difference of the iterated kernels)**.** *For all* $0 \leq s < t \leq T$, $(x,y) \in (\mathbb{R}^d)^2$ *and for all* $q \in (d, +\infty]$, $r \in \mathbb{N}$:

$$|(\tilde{p} \otimes H^{(r)} - \tilde{p}_\varepsilon \otimes H_\varepsilon^{(r)})(s,t,x,y)| \leq (r+1)\Delta_{\varepsilon,\gamma,q} \frac{\bar{C}^{r+1} \left[\Gamma(\frac{\gamma}{2} \wedge \alpha(q)) \right]^r}{\Gamma(1 + r(\frac{\gamma}{2} \wedge \alpha(q)))} p_c(t-s,y-x)(t-s)^{r(\frac{\gamma}{2} \wedge \alpha(q))}. \tag{3.14}$$

where c is as in Lemma 3.1 and $\bar{C}$ as in equation (3.8) of Lemma 3.3.

Proof. Observe that Lemmas 3.1 and 3.3 respectively give (3.14) for $r = 0$ and $r = 1$. Let us assume that it holds for a given $r \in \mathbb{N}^*$ and let us prove it for $r + 1$.

Let us denote for all $r \geq 1$, $\eta_r(s,t,x,y) := |(\tilde{p} \otimes H^{(r)} - \tilde{p}_\varepsilon \otimes H_\varepsilon^{(r)})(s,t,x,y)|$. Write

$$\eta_{r+1}(s,t,x,y) \leq |[\tilde{p} \otimes H^{(r)} - \tilde{p}_\varepsilon \otimes H_\varepsilon^{(r)}] \otimes H(s,t,x,y)| + |\tilde{p}_\varepsilon \otimes H_\varepsilon^{(r)} \otimes (H - H_\varepsilon)(s,t,x,y)|$$

$$\leq \eta_r \otimes |H|(s,t,x,y) + |\tilde{p}_\varepsilon \otimes H_\varepsilon^{(r)}| \otimes |(H - H_\varepsilon)|(s,t,x,y).$$

Recall now that under (A), the terms $|H|(s,t,x,y)$ and $|\tilde{p}_\varepsilon \otimes H_\varepsilon^{(r)}|$ satisfy respectively and uniformly in ε the controls of equations (2.9), (2.10). The result then follows from the proof of Lemma 3.3 (see Eq. (3.11) for $q = \infty$ and (3.13) for $q \in (d, +\infty)$) and the induction hypothesis. □

Theorem 1.1 now simply follows from the controls of Lemma 3.4, the parametrix expansions (2.6) and (2.7) of the densities p, p_ε and the asymptotics of the gamma function.

3.2. Stability for Markov chains

In this Section we prove Theorem 1.3. The strategy is rather similar to the one of Section 3.1 thanks to the series representation of the densities of the chains given in Proposition 2.2.

Recall first from Section 2.2 that we have the following representations for the density p^h, p_ε^h of the Markov chains $Y, Y^{(\varepsilon)}$ in (1.3). For all $0 \le t_i < t_j \le T$, $(x,y) \in (\mathbb{R}^d)^2$:

$$p^h(t_i, t_j, x, y) = \sum_{r=0}^{j-i} \tilde{p}^h \otimes_h H^{h,(r)}(t_i, t_j, x, y),$$

$$p_\varepsilon^h(t_i, t_j, x, y) = \sum_{r=0}^{j-i} \tilde{p}_\varepsilon^h \otimes_h H_\varepsilon^{h,(r)}(t_i, t_j, x, y).$$

3.2.1. Comparison of the frozen densities

The first key point for the analysis with Markov chains is the following Lemma.

Lemma 3.5 (Controls and comparison of the densities and their derivatives). *There exist c, c_1 s.t. for all $0 \le t_i < t_j \le T$, $(x,y) \in (\mathbb{R}^d)^2$ and for all multi-index α, $|\alpha| \le 4$:*

$$|D_x^\alpha \tilde{p}^h(t_i, t_j, x, y)| + |D_x^\alpha \tilde{p}_\varepsilon^h(t_i, t_j, x, y)| \le \frac{1}{(t_j - t_i)^{|\alpha|/2}} \psi_{c,c_1}(t_j - t_i, y - x),$$

$$|D_x^\alpha \tilde{p}^h(t_i, t_j, x, y) - D_x^\alpha \tilde{p}_\varepsilon^h(t_i, t_j, x, y)| \le \frac{\Delta_{\varepsilon,\sigma,\gamma}}{(t_j - t_i)^{|\alpha|/2}} \psi_{c,c_1}(t_j - t_i, y - x),$$

where

– Under (I_G)*:*

$$\psi_{c,c_1}(t_j - t_i, y - x) := c_1 p_c(t_j - t_i, y - x),$$

– Under $(\mathrm{I}_{P,M})$*:*

$$\psi_{c,c_1}(t_j - t_i, y - x) := \frac{c_1 c^d}{(t_j - t_i)^{d/2}} Q_{M-d-5}\left(\frac{|y - x|}{(t_j - t_i)^{1/2}/c}\right).$$

Proof. Note first that under (I_G) the statement has already been proved in Lemma 3.1. We thus assume that $(\mathrm{I}_{P,M})$ holds. Introduce first the random vectors with zero mean:

$$\tilde{Z}_{k,j}^y := \frac{1}{(t_j - t_k)^{1/2}} \sum_{l=k}^{j-1} \sigma(t_l, y)\sqrt{h}\xi_{l+1}, \quad \tilde{Z}_{k,j}^{y,(\varepsilon)} := \frac{1}{(t_j - t_k)^{1/2}} \sum_{l=k}^{j-1} \sigma_\varepsilon(t_l, y)\sqrt{h}\xi_{l+1}.$$

Denoting by $q_{j-k}, q_{j-k,\varepsilon}$ their respective densities, one has:

$$D_x^\alpha \tilde{p}^h(t_k, t_j, x, y) = \frac{1}{(t_j - t_k)^{(d+|\alpha|)/2}} (-1)^{|\alpha|} D_z^\alpha q_{j-k}(z)\big|_{z = \frac{y-x}{(t_j - t_k)^{1/2}}},$$

$$D_x^\alpha \tilde{p}_\varepsilon^h(t_k, t_j, x, y) = \frac{1}{(t_j - t_k)^{(d+|\alpha|)/2}} (-1)^{|\alpha|} D_z^\alpha q_{j-k,\varepsilon}(z)\big|_{z = \frac{y-x}{(t_j - t_k)^{1/2}}}. \tag{3.15}$$

From the Edgeworth expansion of Theorem 19.3 in Bhattacharya and Rao [4], for $q_{j-k}, q_{j-k,\varepsilon}$, one readily derives under (A), for $|\alpha| = 0$ that there exists c_1 s.t. for all $0 \le t_k < t_j \le T$, $(x,y) \in (\mathbb{R}^d)^2$,

$$\tilde{p}^h(t_k, t_j, x, y) + \tilde{p}^h_\varepsilon(t_k, t_j, x, y) \le \frac{c_1}{(t_j - t_k)^{d/2}} \frac{1}{\left(1 + \frac{|x-y|}{(t_j - t_k)^{1/2}}\right)^m}, \tag{3.16}$$

for all integer $m < M - d$, where we recall that M stands for the initial decay of the density f_ξ of the innovations bounded by Q_M (see Eq. (1.7)).

We can as well derive similarly to the proof of Theorem 19.3 in [4], see also Lemma 3.7 in [14], that for all $\alpha, |\alpha| \le 4$:

$$|D_x^\alpha \tilde{p}^h(t_k, t_j, x, y)| + |D_x^\alpha \tilde{p}^h_\varepsilon(t_k, t_j, x, y)| \le \frac{c_1}{(t_j - t_k)^{(d+|\alpha|)/2}} \frac{1}{\left(1 + \frac{|x-y|}{(t_j - t_k)^{1/2}}\right)^m}, \tag{3.17}$$

for all $m < M - d - 4$. Note indeed that differentiating in D_x^α the density and the terms of the Edgeworth expansion corresponds to a multiplication of the Fourier transforms involved by ζ^α, ζ standing for the Fourier variable. Hence, from our smoothness assumptions in $(\mathrm{I}_{P,M})$, after obvious modifications, the estimates of Theorem 9.11 and Lemma 14.3 from [4] apply for these derivatives. With these bounds, one then simply has to copy the proof of Theorem 19.3. Roughly speaking, taking derivatives deteriorates the concentration of the initial control in (3.16) up to the derivation order. On the other hand, the bound in (3.16) is itself deteriorated w.r.t. the initial concentration condition in (1.7). The key point is that the techniques of Theorem 19.3 in [4] actually provide concentration bounds for inhomogeneous sums of random variables with concentration as in (1.7) in terms of the moments of the innovations. To explain the bound in (3.16) let us observe that the mth moment of ξ is finite for $m < M - d$.

Equations (3.16) and (3.17) give the first part of the lemma. Still from the proof of Theorem 19.3 in [4], one gets, under (A), that there exists $C > 0$ s.t. for all multi-indexes $\bar{\alpha}$, $|\bar{\alpha}| \le 4, \bar{\beta}$, $|\bar{\beta}| \le m \le M - d - 5$ for all $j > k$:

$$\int_{\mathbb{R}^d} |\zeta^{\bar{\alpha}}| \left\{ \left| D_\zeta^{\bar{\beta}} \hat{q}_{j-k}(\zeta) \right| + \left| D_\zeta^{\bar{\beta}} \hat{q}_{j-k,\varepsilon}(\zeta) \right| \right\} \mathrm{d}\zeta \le C, \tag{3.18}$$

where $\hat{q}_{j-k}(\zeta), \hat{q}_{j-k,\varepsilon}(\zeta)$ stand for the respective characteristic functions of the random variables $\tilde{Z}_{k,j}^y$, $\tilde{Z}_{k,j}^{y,(\varepsilon)}$ at point ζ.

To investigate the quantity $|D_x^\alpha \tilde{p}^h(t_k, t_j, x, y) - D_x^\alpha \tilde{p}^h_\varepsilon(t_k, t_j, x, y)|$ thanks to (3.15) define now for all $\alpha, |\alpha| \le 4, \beta, |\beta| \le m \le M - d - 5$:

$$\forall z \in \mathbb{R}^d, \; \Theta_{j-k,\varepsilon}(z) := z^\beta D_z^\alpha \left(q_{j-k}(z) - q_{j-k,\varepsilon}(z) \right),$$
$$\forall \zeta \in \mathbb{R}^d, \; \widehat{\Theta}_{j-k,\varepsilon}(\zeta) := (-i)^{|\alpha|+|\beta|} D_\zeta^\beta \left(\zeta^\alpha \left\{ \hat{q}_{j-k}(\zeta) - \hat{q}_{j-k,\varepsilon}(\zeta) \right\} \right). \tag{3.19}$$

Let us now estimate the difference between the characteristic functions. From the Leibniz formula, we are led to investigate for all multi-indexes $\bar{\beta}, \bar{\alpha}$, $|\bar{\beta}| \le |\beta|, |\bar{\alpha}| \le |\alpha|$ quantities of the form:

$$(i^{\bar{\beta}})^{-1} \zeta^{\bar{\alpha}} (D_\zeta^{\bar{\beta}} \hat{q}_{j-k}(\zeta) - D_\zeta^{\bar{\beta}} \hat{q}_{j-k,\varepsilon}(\zeta)) = \zeta^{\bar{\alpha}} \mathbb{E}\left[(\tilde{Z}_{k,j}^y)^{\bar{\beta}} \exp[i\zeta \cdot \tilde{Z}_{k,j}^y] - \left(\tilde{Z}_{k,j}^{y,(\varepsilon)} \right)^{\bar{\beta}} \exp\left[i\zeta \cdot \tilde{Z}_{k,j}^{y,(\varepsilon)} \right] \right].$$

Assume first that $j > k + 1$. In that case, set now $\tilde{Z}^y_{k,j,1} := \tilde{Z}^y_{k,\lceil(j+k)/2\rceil}$, $\tilde{Z}^y_{k,j,2} := \tilde{Z}^y_{k,j} - \tilde{Z}^y_{k,j,1}$. Denoting similarly $\tilde{Z}^{y,(\varepsilon)}_{k,j,1} := \tilde{Z}^{y,(\varepsilon)}_{k,\lceil(j+k)/2\rceil}$, $\tilde{Z}^{y,(\varepsilon)}_{k,j,2} := \tilde{Z}^{y,(\varepsilon)}_{k,j} - \tilde{Z}^{y,(\varepsilon)}_{k,j,1}$ for the perturbed process, we get:

$$(i^{\bar{\beta}})^{-1}\zeta^{\bar{\alpha}}(D^{\bar{\beta}}_\zeta \hat{q}_{j-k}(\zeta) - D^{\bar{\beta}}_\zeta \hat{q}_{j-k,\varepsilon}(\zeta))$$

$$= \zeta^{\bar{\alpha}}\left\{ \mathbb{E}\left[(\tilde{Z}^y_{k,j,1} + \tilde{Z}^y_{k,j,2})^{\bar{\beta}}\exp[i\zeta \cdot \tilde{Z}^y_{k,j,1}]\exp[i\zeta \cdot \tilde{Z}^y_{k,j,2}]\right] - \mathbb{E}\left[(\tilde{Z}^{y,(\varepsilon)}_{k,j,1} + \tilde{Z}^{y,(\varepsilon)}_{k,j,2})^{\bar{\beta}}\exp[i\zeta \cdot \tilde{Z}^{y,(\varepsilon)}_{k,j,1}]\exp\left[i\zeta \cdot \tilde{Z}^{y,(\varepsilon)}_{k,j,2}\right]\right]\right\}$$

$$= \zeta^{\bar{\alpha}}\left\{ \sum_{l,|l|\le|\bar{\beta}|} C^l_{\bar{\beta}}\mathbb{E}\left[(\tilde{Z}^y_{k,j,1})^l\exp[i\zeta \cdot \tilde{Z}^y_{k,j,1}]\right]\mathbb{E}\left[(\tilde{Z}^y_{k,j,2})^{\bar{\beta}-l}\exp[i\zeta \cdot \tilde{Z}^y_{k,j,2}]\right]\right.$$

$$\left. - \sum_{l,|l|\le|\bar{\beta}|} C^l_{\bar{\beta}}\mathbb{E}\left[(\tilde{Z}^{y,(\varepsilon)}_{k,j,1})^l\exp[i\zeta \cdot \tilde{Z}^{y,(\varepsilon)}_{k,j,1}]\right]\mathbb{E}\left[(\tilde{Z}^{y,(\varepsilon)}_{k,j,2})^{\bar{\beta}-l}\exp[i\zeta \cdot \tilde{Z}^{y,(\varepsilon)}_{k,j,2}]\right]\right\}$$

$$= \zeta^{\bar{\alpha}}\left\{ \sum_{l,|l|\le|\bar{\beta}|} C^l_{\bar{\beta}}\left\{\left[\mathbb{E}\left[(\tilde{Z}^y_{k,j,1})^l\exp[i\zeta \cdot \tilde{Z}^y_{k,j,1}]\right] - \mathbb{E}\left[(\tilde{Z}^{y,(\varepsilon)}_{k,j,1})^l\exp[i\zeta \cdot \tilde{Z}^{y,(\varepsilon)}_{k,j,1}]\right]\right]\mathbb{E}\left[(\tilde{Z}^y_{k,j,2})^{\bar{\beta}-l}\exp[i\zeta \cdot \tilde{Z}^y_{k,j,2}]\right]\right.\right.$$

$$\left.\left. + \mathbb{E}\left[(\tilde{Z}^{y,(\varepsilon)}_{k,j,1})^l\exp[i\zeta \cdot \tilde{Z}^{y,(\varepsilon)}_{k,j,1}]\right]\left[\mathbb{E}\left[(\tilde{Z}^y_{k,j,2})^{\bar{\beta}-l}\exp[i\zeta \cdot \tilde{Z}^y_{k,j,2}]\right] - \mathbb{E}\left[(\tilde{Z}^{y,(\varepsilon)}_{k,j,2})^{\bar{\beta}-l}\exp[i\zeta \cdot \tilde{Z}^{y,(\varepsilon)}_{k,j,2}]\right]\right]\right\}\right\},$$

where in the above expression we considered the binomial expansion for multi-indexes denoting by $C^l_{\bar{\beta}} := \frac{\bar{\beta}!}{(\bar{\beta}-l)!l!}$ with the corresponding definitions for factorials (see the proof of Lem. 3.1). Introduce now, for a multi-index $l, |l| \in [\![0, |\bar{\beta}|]\!]$, the functions:

$$\Psi^{\bar{\alpha},\bar{\beta}-l}_1(\zeta) := \zeta^{\bar{\alpha}}\mathbb{E}\left[(\tilde{Z}^y_{k,j,2})^{\bar{\beta}-l}\exp\left[i\zeta \cdot \tilde{Z}^y_{k,j,2}\right]\right], \quad \Psi^{\bar{\alpha},l}_2(\zeta) := \zeta^{\bar{\alpha}}\mathbb{E}\left[\left(\tilde{Z}^{y,(\varepsilon)}_{k,j,1}\right)^l\exp\left[i\zeta \cdot \tilde{Z}^{y,(\varepsilon)}_{k,j,1}\right]\right],$$

and

$$\mathcal{E}_{1,l}(\zeta) := \left[\mathbb{E}\left[(\tilde{Z}^y_{k,j,1})^l\exp[i\zeta \cdot \tilde{Z}^y_{k,j,1}]\right] - \mathbb{E}\left[(\tilde{Z}^{y,(\varepsilon)}_{k,j,1})^l\exp\left[i\zeta \cdot \tilde{Z}^{y,(\varepsilon)}_{k,j,1}\right]\right]\right],$$

$$\mathcal{E}_{2,\bar{\beta}-l}(\zeta) := \left[\mathbb{E}\left[(\tilde{Z}^y_{k,j,2})^{\bar{\beta}-l}\exp[i\zeta \cdot \tilde{Z}^y_{k,j,2}]\right] - \mathbb{E}\left[(\tilde{Z}^{y,(\varepsilon)}_{k,j,2})^{\bar{\beta}-l}\exp\left[i\zeta \cdot \tilde{Z}^{y,(\varepsilon)}_{k,j,2}\right]\right]\right].$$

Thus, we can rewrite from the previous computations:

$$(i^{\bar{\beta}})^{-1}\zeta^{\bar{\alpha}}(D^{\bar{\beta}}_\zeta \hat{q}_{j-k}(\zeta) - D^{\bar{\beta}}_\zeta \hat{q}_{j-k,\varepsilon}(\zeta)) = \sum_{l,|l|\le|\bar{\beta}|} C^l_{\bar{\beta}}\left\{(\mathcal{E}_{1,l}\Psi^{\bar{\alpha},\bar{\beta}-l}_1)(\zeta) + \left(\mathcal{E}_{2,\bar{\beta}-l}\Psi^{\bar{\alpha},l}_2\right)(\zeta)\right\}. \tag{3.20}$$

Recall from (3.18) that we already have integrability for the contributions $\Psi^{\bar{\alpha},\bar{\beta}-l}_1(\zeta)$ and $\Psi^{\bar{\alpha},l}_2(\zeta)$. Let us thus proceed with the control of $\mathcal{E}_{1,l}(\zeta), \mathcal{E}_{2,\bar{\beta}-l}(\zeta)$. We only give details for $\mathcal{E}_{1,l}(\zeta)$, the contribution $\mathcal{E}_{2,\bar{\beta}-l}$ can be handled similarly. We also consider $|l| \ge 2$, since the cases $|l| < 2$ can be handled more directly. Write:

$$|\mathcal{E}_{1,l}(\zeta)| \le \mathbb{E}\left[\left|(\tilde{Z}^y_{k,j,1})^l - (\tilde{Z}^{y,(\varepsilon)}_{k,j,1})^l\right|\right] + \mathbb{E}\left[\left|(\tilde{Z}^{y,(\varepsilon)}_{k,j,1})^l||\exp(i\zeta \cdot \tilde{Z}^y_{k,j,1}) - \exp\left(i\zeta \cdot \tilde{Z}^{y,(\varepsilon)}_{k,j,1}\right)\right|\right]$$

$$\le C\left\{\mathbb{E}\left[|\tilde{Z}^y_{k,j,1} - \tilde{Z}^{y,(\varepsilon)}_{k,j,1}|\left(|\tilde{Z}^y_{k,j,1}|^{|l|-1} + |\tilde{Z}^{y,(\varepsilon)}_{k,j,1}|^{|l|-1}\right)\right] + \mathbb{E}\left[\left|\tilde{Z}^{y,(\varepsilon)}_{k,j,1}\right|^{|l|}|\zeta||\tilde{Z}^y_{k,j,1} - \tilde{Z}^{y,(\varepsilon)}_{k,j,1}|\right]\right\}.$$

Apply now Hölder's inequality with $p_1 = |l|, q_1 = |l|/(|l|-1)$ for the first term and $p_2 = (|l|+1)/|l|, q_2 = |l|+1$ for the second one so that all the contribution appear with the same power (in order to equilibrate the constraints concerning the intregrability conditions). One gets:

$$|\mathcal{E}_{1,l}(\zeta)| \le C\left\{\mathbb{E}[|\tilde{Z}^y_{k,j,1} - \tilde{Z}^{y,(\varepsilon)}_{k,j,1}|^{|l|}]^{1/|l|}\left\{\mathbb{E}[|\tilde{Z}^y_{k,j,1}|^{|l|}]^{(|l|-1)/|l|} + \mathbb{E}[|\tilde{Z}^{y,(\varepsilon)}_{k,j,1}|^{|l|}]^{(|l|-1)/|l|}\right\}\right.$$

$$\left. + |\zeta|\mathbb{E}[|\tilde{Z}^{y,(\varepsilon)}_{k,j,1}|^{|l|+1}]^{|l|/(|l|+1)}\mathbb{E}\left[|\tilde{Z}^y_{k,j,1} - \tilde{Z}^{y,(\varepsilon)}_{k,j,1}|^{|l|+1}\right]^{1/(|l|+1)}\right\}. \tag{3.21}$$

The point is now to prove, since we have assumed $m \leq M - d - 5 \iff m + 1 \leq M - d - 4$, that there exists c s.t. for all $r \leq m + 1$,

$$\mathbb{E}[|\tilde{Z}_{k,j,1}^{y} - \tilde{Z}_{k,j,1}^{y,(\varepsilon)}|^{r}]^{1/r} \leq c\Delta_{\varepsilon,\sigma,\gamma}, \ \mathbb{E}[|\tilde{Z}_{k,j,1}^{y}|^{r}]^{1/r} + \mathbb{E}[|\tilde{Z}_{k,j,1}^{y,(\varepsilon)}|^{r}]^{1/r} \leq c. \tag{3.22}$$

Let us establish the point for the difference, the other bounds can be derived similarly. Define for all $i \in [\![k,j]\!]$, $\tilde{M}_i := \sqrt{h} \sum_{r=k}^{i-1} (\sigma - \sigma_\varepsilon)(t_r, y)\xi_{r+1}$. The process $(\tilde{M}_i)_{i \in [\![k,j]\!]}$ is a square integrable martingale (in discrete time, w.r.t. $\mathcal{F}_i := \Sigma(\xi_r, r \leq i)$, Σ-field generated by the innovation up to the current time). Its quadratic variation writes $[\tilde{M}]_i = h \sum_{r=k}^{i-1} |(\sigma - \sigma_\varepsilon)(t_r, y)|^2 |\xi_{r+1}|^2$ and the Burkholder-Davies-Gundy inequalities, see *e.g.* Shiryaev [24], give for all $r \leq M - d - 4$:

$$\mathbb{E}[\sup_{i \in [\![k,j]\!]} |\tilde{M}_i|^r] \leq c_r \mathbb{E}[[\tilde{M}]_j^{r/2}] = c_r h^{r/2} \mathbb{E}\left[\left(\sum_{i=k}^{j-1} |(\sigma - \sigma_\varepsilon)(t_i, y)|^2 |\xi_{i+1}|^2\right)^{r/2}\right]. \tag{3.23}$$

If $r = 2$ one readily gets:

$$\mathbb{E}[|\tilde{Z}_{k,j,1}^{y} - \tilde{Z}_{k,j,1}^{y,(\varepsilon)}|^2] \leq \frac{c_2}{(t_j - t_k)} \mathbb{E}[\sup_{i \in [\![k,j]\!]} |\tilde{M}_i|^2] \leq \frac{c_2 h}{(t_j - t_k)} \Delta_{\varepsilon,\sigma,\gamma}^2 \sum_{i=k}^{j-1} \mathbb{E}[|\xi_{i+1}|^2] \leq \bar{c}_2 \Delta_{\varepsilon,\sigma,\gamma}^2.$$

Let us thus assume $r > 2$ and derive from (3.23)

$$\mathbb{E}[|\tilde{Z}_{k,j,1}^{y} - \tilde{Z}_{k,j,1}^{y,(\varepsilon)}|^r] \leq \frac{c_r}{(t_j - t_k)^{r/2}} \mathbb{E}[\sup_{i \in [\![k,j]\!]} |\tilde{M}_i|^r]$$

$$\leq \frac{c_r h^{r/2}}{(t_j - t_k)^{r/2}} \mathbb{E}\left[\left(\sum_{i=k}^{j-1} |(\sigma - \sigma_\varepsilon)(t_i, y)|^r |\xi_{i+1}|^r\right)\left(\sum_{i=k}^{j-1} 1\right)^{r/2(1-2/r)}\right],$$

applying Hölder's inequality for the counting measure with $p = r/2, q = r/(r-2)$ for the last inequality. This finally gives:

$$\mathbb{E}[|\tilde{Z}_{k,j,1}^{y} - \tilde{Z}_{k,j,1}^{y,(\varepsilon)}|^r] \leq \frac{c_r h^{r/2}}{(t_j - t_k)^{r/2}} (j - k)^{r/2-1} \Delta_{\varepsilon,\sigma,\gamma}^r \sum_{i=k}^{j-1} \mathbb{E}[|\xi_{i+1}|^r] \leq \bar{c}_r \Delta_{\varepsilon,\sigma,\gamma}^r.$$

Since we have assumed $r \leq m + 1 \leq M - d - 4$, this gives the first control in (3.22). The other one readily follows replacing $\sigma - \sigma_\varepsilon$ by σ or σ_ε.

From equations (3.21), (3.22) and similar controls for $\mathcal{E}_{2,\bar{\beta}-l}(\zeta)$ we finally derive:

$$|\mathcal{E}_{1,l}(\zeta)| + |\mathcal{E}_{2,\bar{\beta}-l}(\zeta)| \leq C_1 \Delta_{\varepsilon,\sigma,\gamma}(1 + |\zeta|).$$

As a result we have from (3.19) and (3.20):

$$|D_\zeta^\beta(\zeta^\alpha(\hat{q}_{j-k}(\zeta) - D_\zeta^\beta \hat{q}_{j-k,\varepsilon}(\zeta)))| \leq C\Delta_{\varepsilon,\sigma,\gamma} \left\{ \sum_{\substack{\bar{\beta}, |\bar{\beta}| \leq |\beta| \\ \bar{\alpha} = \alpha - (\beta - \bar{\beta}).}} \sum_{l, |l| \leq |\bar{\beta}|} (|\Psi_1^{\bar{\alpha},\bar{\beta}-l}(\zeta)| + |\Psi_2^{\bar{\alpha}l}(\zeta)|)(1 + |\zeta|) \right\}.$$

We finally derive from (3.19) and (3.18) (which thanks to the smoothness assumption on Q_M in $(\mathrm{I}_{P,M})$ holds as well for a multi-index $\bar{\alpha}, |\bar{\alpha}| = 5$):

$$|\Theta_{j-k,\varepsilon}(z)| \leq \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} |\hat{\Theta}_{j-k,\varepsilon}(\zeta)| \mathrm{d}\zeta \leq c\Delta_{\varepsilon,\sigma,\gamma}. \tag{3.24}$$

From (3.15) this concludes the proof for $j > k + 1$. If $j = k + 1$ the previous arguments can be simplified and lead to the same results. $\square$

3.2.2. Comparison of the parametrix kernels

This step is crucial and actually the key to the result for the Markov chains. We focus for simplicity on the case $q = +\infty$, for which pointwise controls for the differences between the drift coefficients are available, and which already emphasizes all the difficulties. The case $q \in (d, +\infty)$ for the drifts could be handled as in Lemma 3.3, using similar Hölder inequalities.

We actually have the following Lemma.

Lemma 3.6 (Control of the One-Step convolution for the chain). *There exists c_1, c s.t. for all $0 \leq t_k < t_j \leq T, (z, y) \in (\mathbb{R}^d)^2$:*

$$|(H^h - H^h_\varepsilon)(t_k, t_j, z, y)| \leq \frac{\Delta_{\varepsilon,\gamma,\infty}}{(t_j - t_k)^{1-\gamma/2}} \Phi_{c,c_1}(t_j - t_k, z - y),$$

with

– $\Phi_{c,c_1}(t_j - t_k, z - y) = \psi_{c,c_1}(t_j - t_k, z - y)$ *under* (I_G).
– $\Phi_{c,c_1}(t_j - t_k, z - y) = \psi_{c,c_1}(t_j - t_k, z - y) \left(1 + \frac{|z-y|}{(t_j - t_k)^{1/2}}\right)^\gamma$, *under* $(\mathrm{I}_{P,M})$,

where ψ_{c,c_1} is defined according to the assumptions on the innovations in Lemma 3.5.

Proof. The case $k = j + 1$ involves directly differences of densities and could be treated more directly than the case $k > j + 1$. We thus focus on the latter. Introduce for $k \in [\![0, N]\!], (x, w) \in (\mathbb{R}^d)^2$ the one step transitions:

$$T^h(t_k, x, w) := b(t_k, x)h + h^{1/2}\sigma(t_k, x)w, \ T^h_\varepsilon(t_k, x, w) := b_\varepsilon(t_k, x)h + h^{1/2}\sigma_\varepsilon(t_k, x)w,$$
$$T^h_0(t_k, x, w) := h^{1/2}\sigma(t_k, x)w, \ T^h_{\varepsilon,0}(t_k, x, w) := h^{1/2}\sigma_\varepsilon(t_k, x)w. \tag{3.25}$$

From the definition of H^h, H^h_ε, recalling that f_ξ stands for the density of the innovation, the difference of the kernels writes:

$$(H^h - H^h_\varepsilon)(t_k, t_j, z, y) = h^{-1} \int_{\mathbb{R}^d} dw f_\xi(w) \Bigg[\Big\{ \tilde{p}^h(t_{k+1}, t_j, z + T^h(t_k, z, w), y) - \tilde{p}^h(t_{k+1}, t_j, z + T^h_0(t_k, y, w), y) \Big\}$$

$$- \Big\{ \tilde{p}^h_\varepsilon(t_{k+1}, t_j, z + T^h_\varepsilon(t_k, z, w), y) - \tilde{p}^h_\varepsilon(t_{k+1}, t_j, z + T^h_{0,\varepsilon}(t_k, y, w), y) \Big\} \Bigg]. \tag{3.26}$$

Let us now perform a Taylor expansion at order 2 with integral rest. To this end, let us first introduce for $\lambda \in [0, 1]$ the mappings:

$$\begin{aligned} \varphi^h_\lambda : \mathbb{R}^d \times \mathbb{R}^d &\longrightarrow \mathbb{R} \\ (T_1, T_2) &\longmapsto \mathrm{Tr}\Big(D^2_z \tilde{p}^h(t_{k+1}, t_j, z + \lambda T_1, y)[T_2 T_2^*] \Big), \end{aligned}$$
$$\begin{aligned} \varphi^h_{\lambda,\varepsilon} : \mathbb{R}^d \times \mathbb{R}^d &\longrightarrow \mathbb{R} \\ (T_1, T_2) &\longmapsto \mathrm{Tr}\Big(D^2_z \tilde{p}^h_\varepsilon(t_{k+1}, t_j, z + \lambda T_1, y)[T_2 T_2^*] \Big), \end{aligned} \tag{3.27}$$

where T_2 is viewed as a column vector and T_2^* denotes its transpose. Recalling as well that ξ is centered we get:

$$
\begin{aligned}
\Delta H^{h,\varepsilon}(t_k, t_j, z, y) &:= (H^h - H^h_\varepsilon)(t_k, t_j, z, y) \\
&= \left[\left\langle D_z \tilde{p}^h(t_{k+1}, t_j, z, y), b(t_k, z) \right\rangle - \left\langle D_z \tilde{p}^h_\varepsilon(t_{k+1}, t_j, z, y), b_\varepsilon(t_k, z) \right\rangle \right] \\
&\quad + h^{-1} \int_{\mathbb{R}^d} \mathrm{d}w f_\xi(w) \int_0^1 \mathrm{d}\lambda (1 - \lambda) \\
&\quad \times \left[\left\{ \varphi_\lambda^h(T^h(t_k, z, w), T^h(t_k, z, w)) - \varphi_\lambda^h(T_0^h(t_k, y, w), T_0^h(t_k, y, w)) \right\} \right. \\
&\quad \left. - \left\{ \varphi_{\lambda,\varepsilon}^h(T_\varepsilon^h(t_k, z, w), T_\varepsilon^h(t_k, z, w)) - \varphi_{\lambda,\varepsilon}^h(T_{0,\varepsilon}^h(t_k, y, w), T_{0,\varepsilon}^h(t_k, y, w)) \right\} \right] \\
&=: (\Delta_1 H^{h,\varepsilon} + \Delta_2 H^{h,\varepsilon})(t_k, t_j, z, y), \tag{3.28}
\end{aligned}
$$

where for $i \in \{1, 2\}$, $\Delta_i H^{h,\varepsilon}$ is associated with the terms of order i. The idea is now to make $\Delta_{\varepsilon,\gamma,\infty}$ appear explicitly. The term $\Delta_1 H^{h,\varepsilon}$ is the easiest to handle. We can indeed readily write:

$$
\Delta_1 H^{h,\varepsilon}(t_k, t_j, z, y) = \left[\left\langle D_z \tilde{p}^h(t_{k+1}, t_j, z, y), [b(t_k, z) - b_\varepsilon(t_k, z)] \right\rangle - \left\langle (D_z \tilde{p}^h_\varepsilon - D_z \tilde{p}^h)(t_{k+1}, t_j, z, y), b_\varepsilon(t_k, z) \right\rangle \right].
$$

From Assumption (A3), equation (1.5) and Lemma 3.5 we derive for $q = +\infty$:

$$
|\Delta_1 H^{h,\varepsilon}(t_k, t_j, z, y)| \leq \frac{C \Delta_{\varepsilon,\gamma,\infty}}{(t_j - t_k)^{1/2}} \psi_{c,c_1}(t_j - t_k, y - z). \tag{3.29}
$$

The term $\Delta_2 H^{h,\varepsilon}$ is trickier to handle. Define to this end:

$$
\begin{aligned}
\Delta \varphi_\lambda^{h,\varepsilon}(t_k, z, y, w) &:= \left\{ \varphi_\lambda^h(T^h(t_k, z, w), T^h(t_k, z, w)) - \varphi_\lambda^h(T_0^h(t_k, y, w), T_0^h(t_k, y, w)) \right\} \\
&\quad - \left\{ \varphi_{\lambda,\varepsilon}^h(T_\varepsilon^h(t_k, z, w), T_\varepsilon^h(t_k, z, w)) - \varphi_{\lambda,\varepsilon}^h(T_{0,\varepsilon}^h(t_k, y, w), T_{0,\varepsilon}^h(t_k, y, w)) \right\}.
\end{aligned}
$$

Let us then decompose:

$$
\begin{aligned}
\Delta \varphi_\lambda^{h,\varepsilon}(t_k, z, y, w) &:= \left[\left\{ \varphi_\lambda^h(T^h(t_k, z, w), T^h(t_k, z, w)) - \varphi_\lambda^h(T^h(t_k, z, w), T_0^h(t_k, y, w)) \right\} \right. \\
&\quad \left. - \left\{ \varphi_{\lambda,\varepsilon}^h(T_\varepsilon^h(t_k, z, w), T_\varepsilon^h(t_k, z, w)) - \varphi_{\lambda,\varepsilon}^h(T_\varepsilon^h(t_k, z, w), T_{0,\varepsilon}^h(t_k, y, w)) \right\} \right] \\
&\quad + \left[\left\{ \varphi_\lambda^h(T^h(t_k, z, w), T_0^h(t_k, y, w)) - \varphi_\lambda^h(T_0^h(t_k, y, w), T_0^h(t_k, y, w)) \right\} \right. \\
&\quad \left. - \left\{ \varphi_{\lambda,\varepsilon}^h(T_\varepsilon^h(t_k, y, w), T_{0,\varepsilon}^h(t_k, y, w)) - \varphi_{\lambda,\varepsilon}^h(T_{0,\varepsilon}^h(t_k, z, w), T_{0,\varepsilon}^h(t_k, y, w)) \right\} \right] \\
&=: (\Delta_1 \varphi_\lambda^{h,\varepsilon} + \Delta_2 \varphi_\lambda^{h,\varepsilon})(t_k, z, y, w), \tag{3.30}
\end{aligned}
$$

and write from (3.28):

$$
\begin{aligned}
\Delta_2 H^{h,\varepsilon}(t_k, t_j, z, y) &= h^{-1} \int_{\mathbb{R}^d} \mathrm{d}w f_\xi(w) \int_0^1 \mathrm{d}\lambda (1 - \lambda)(\Delta_1 \varphi_\lambda^{h,\varepsilon} + \Delta_2 \varphi_\lambda^{h,\varepsilon})(t_k, z, y, w) \\
&=: (\Delta_{21} H^{h,\varepsilon} + \Delta_{22} H^{h,\varepsilon})(t_k, t_j, z, y), \tag{3.31}
\end{aligned}
$$

for the associated contributions in $\Delta_2 H^{h,\varepsilon}$. Again, we have to consider these two terms separately.

Term $\Delta_{21} H^{h,\varepsilon}$. We first write from (3.30):

$$\Delta_1 \varphi_\lambda^{h,\varepsilon}(t_k, z, y, w) = \left[\left\{\varphi_\lambda^h(T^h(t_k, z, w), T^h(t_k, z, w)) - \varphi_\lambda^h(T^h(t_k, z, w), T_0^h(t_k, y, w))\right\}\right.$$

$$- \left\{\varphi_\lambda^h(T^h(t_k, z, w), T_\varepsilon^h(t_k, z, w)) - \varphi_\lambda^h(T^h(t_k, z, w), T_{0,\varepsilon}^h(t_k, y, w))\right\}\right]$$

$$+ \left[\left\{\varphi_\lambda^h(T^h(t_k, z, w), T_\varepsilon^h(t_k, z, w)) - \varphi_\lambda^h(T^h(t_k, z, w), T_{0,\varepsilon}^h(t_k, y, w))\right\}\right.$$

$$- \left\{\varphi_\lambda^h(T_\varepsilon^h(t_k, z, w), T_\varepsilon^h(t_k, z, w)) - \varphi_\lambda^h(T_\varepsilon^h(t_k, z, w), T_{0,\varepsilon}^h(t_k, y, w))\right\}\right]$$

$$- \left[\left\{\varphi_{\lambda,\varepsilon}^h(T_\varepsilon^h(t_k, z, w), T_\varepsilon^h(t_k, z, w)) - \varphi_{\lambda,\varepsilon}^h(T_\varepsilon^h(t_k, z, w), T_{0,\varepsilon}^h(t_k, y, w))\right\}\right.$$

$$- \left\{\varphi_\lambda^h(T_\varepsilon^h(t_k, z, w), T_\varepsilon^h(t_k, z, w)) - \varphi_\lambda^h(T_\varepsilon^h(t_k, z, w), T_{0,\varepsilon}^h(t_k, y, w))\right\}\right]$$

$$=: \sum_{i=1}^{3} \Delta_{1i} \varphi_\lambda^{h,\varepsilon}(t_k, z, y, w). \tag{3.32}$$

We now state some useful controls for the analysis. Namely, setting:

$$D(t_k, z, y, w) := T^h(t_k, z, w) T^h(t_k, z, w)^* - T_0^h(t_k, y, w) T_0^h(t_k, y, w)^*,$$

$$D_\varepsilon(t_k, z, y, w) := T_\varepsilon^h(t_k, z, w) T_\varepsilon^h(t_k, z, w)^* - T_{0,\varepsilon}^h(t_k, y, w) T_{0,\varepsilon}^h(t_k, y, w)^*,$$

we have from (A3) and equation (1.5) for $q = +\infty$:

$$(|D| + |D_\varepsilon|)(t_k, z, y, w) \le \bar{c}(h^2 + h^{3/2}|w| + h(1 \wedge |z - y|)^\gamma |w|^2),$$

$$|D - D_\varepsilon|(t_k, z, y, w) \le \bar{c}\Delta_{\varepsilon,\gamma,\infty}(h^2 + h^{3/2}|w| + h(1 \wedge |z - y|)^\gamma |w|^2). \tag{3.33}$$

From the definition of φ_λ^h in (3.27), equation (3.32), the control (3.33) and Lemma 3.5, we get:

$$|\Delta_{11} \varphi_\lambda^{h,\varepsilon}|(t_k, z, y, w) \le \bar{c}\Delta_{\varepsilon,\gamma,\infty} \frac{\psi_{c,c_1}(t_j - t_k, y - (z + \lambda T^h(t_k, z, w)))}{(t_j - t_k)}(h^2 + h^{3/2}|w| + h(1 \wedge |z - y|)^\gamma |w|^2). \tag{3.34}$$

We would similarly get from Lemma 3.5 and (3.33):

$$|\Delta_{13} \varphi_\lambda^{h,\varepsilon}|(t_k, z, y, w) \le \bar{c}\Delta_{\varepsilon,\gamma,\infty} \frac{\psi_{c,c_1}(t_j - t_k, y - (z + \lambda T_\varepsilon^h(t_k, z, w)))}{(t_j - t_k)}(h^2 + h^{3/2}|w| + h(1 \wedge |z - y|)^\gamma |w|^2),$$

$$|\Delta_{12} \varphi_\lambda^{h,\varepsilon}|(t_k, z, y, w) \le \frac{\psi_{c,c_1}(t_j - t_k, y - (z + \theta\lambda T^h(t_k, z, w) + (1 - \theta)\lambda T_\varepsilon^h(t_k, z, w)))}{(t_j - t_k)^{3/2}}$$

$$\times |(T^h - T_\varepsilon^h)(t_k, z, w)||D_\varepsilon|(t_k, z, y, w)$$

$$\le \bar{c}\Delta_{\varepsilon,\gamma,\infty} \frac{\psi_{c,c_1}(t_j - t_k, y - (z + \theta\lambda T^h(t_k, z, w) + (1 - \theta)\lambda T_\varepsilon^h(t_k, z, w)))}{(t_j - t_k)^{3/2}}$$

$$\times (h^2 + h^{3/2}|w| + h(1 \wedge |z - y|)^\gamma |w|^2)(h + h^{1/2}|w|), \tag{3.35}$$

for some $\theta \in (0, 1)$, using as well (3.25) and (1.5) for the last inequality. The point is now to get rid of the transitions appearing in the function ψ_{c,c_1}. We separate here the two assumptions at hand.

1   – Under (I_G), it suffices to remark that by the convexity inequality $|z - y - \Theta|^2 \geq \frac{1}{2}|z - y|^2 - |\Theta|^2$, for all
2    $\Theta \in \mathbb{R}^d$:

$$\psi_{c,c_1}(t_j - t_k, y - z - \Theta) \leq c_1 \frac{c^{d/2}}{(2\pi(t_j - t_k))^{d/2}} \exp\left(-\frac{c}{4}\frac{|z - y|^2}{t_j - t_k}\right) \exp\left(\frac{c}{2}\frac{|\Theta|^2}{t_j - t_k}\right).$$

4   Now, if Θ is one of the above transitions or linear combination of transitions, we get from (3.25):

$$\psi_{c,c_1}(t_j - t_k, y - z - \Theta) \leq c_1 \frac{(c/2)^{d/2}}{(2\pi(t_j - t_k))^{d/2}} \exp\left(-\frac{c}{4}\frac{|z - y|^2}{t_j - t_k}\right) \exp\left(\frac{c}{2}K_2^2|w|^2\right), \tag{3.36}$$

6   up to a modification of c_1 observing that $h/(t_j - t_k) \leq 1$ and with K_2 as in (A_1). Since c can be cho-
7   sen small enough in the previous controls, up to deteriorating the concentration properties in Lemma 3.5,
8   the last term can be integrated by the standard Gaussian density f_ξ appearing in (3.31). We thus derive,
9   from (3.36), (3.34), (3.35) and the definition in (3.32), up to modifications of c, c_1:

$$|\Delta_1\varphi_\lambda^{h,\varepsilon}|(t_k, z, y, w) \leq \Delta_{\varepsilon,\gamma,\infty}h\bar{c}\psi_{c,c_1}(t_j - t_k, z - y)\exp(c|w|^2)\left\{1 + \frac{|w|}{(t_j - t_k)^{1/2}} + \frac{|z - y|^\gamma|w|^2}{t_j - t_k}\right\},$$

11   which plugged into (3.31) yields up to modifications of $\bar{c}, c, c_1$:

$$|\Delta_{21}H^{h,\varepsilon}(t_k, t_j, z, y)| \leq \bar{c}\frac{\Delta_{\varepsilon,\gamma,\infty}(1 \vee T^{(1-\gamma)/2})\psi_{c,c_1}(t_j - t_k, z - y)}{(t_j - t_k)^{1-\gamma/2}}. \tag{3.37}$$

12   – Under ($I_{P,M}$), we only detail the computations for the off diagonal regime $|z - y| \geq c(t_j - t_k)^{1/2}$ which is the
13   most delicate to handle. In this case, we have to discuss according to the position of w w.r.t. $y - z$. With the
14   notations of (A2), introduce $\mathcal{D} := \{\bar{w} \in \mathbb{R}^d : \{\Lambda h\}^{1/2}|\bar{w}| \leq |z - y|/2\}$. If $w \in \mathcal{D}$, then, still from (3.34), (3.35),
15

16    $(|\Delta_{11}\varphi_\lambda^{h,\varepsilon}| + |\Delta_{13}\varphi_\lambda^{h,\varepsilon}|)(t_k, z, y, w)$

$$17 \qquad \leq \bar{c}\Delta_{\varepsilon,,\gamma,\infty}\frac{\psi_{c,c_1}(t_j - t_k, y - z)}{(t_j - t_k)}(h^2 + h^{3/2}|w| + h(1 \wedge |z - y|)^\gamma|w|^2), |\Delta_{12}\varphi_\lambda^{h,\varepsilon}|(t_k, z, y, w)$$

$$18 \qquad\qquad \leq c\Delta_{\varepsilon,\gamma,\infty}\frac{\psi_{c,c_1}(t_j - t_k, y - z)}{(t_j - t_k)^{3/2}}(h^2 + h^{3/2}|w| + h(1 \wedge |z - y|)^\gamma|w|^2)(h + h^{1/2}|w|).$$
19

20   On the other hand, when $w \notin \mathcal{D}$ we use f_ξ to make the off-diagonal bound of $\psi_{c,c_1}(t_j - t_k, y - z)$ appear.
21   Namely, we can write:

$$f_\xi(w) \leq c\frac{1}{(1 + |w|)^M} \leq c\frac{1}{\left(1 + \frac{|z-y|}{h^{1/2}}\right)^{M-(d+4)}}\frac{1}{(1 + |w|)^{d+4}}$$

$$\leq c\frac{1}{\left(1 + \frac{|z-y|}{(t_j-t_k)^{1/2}}\right)^{M-(d+4)}}\frac{1}{(1 + |w|)^{d+4}}, \tag{3.38}$$

24   where the last splitting is performed in order to integrate the contribution in $|w|^3$ coming from the upper
25   bound for $|\Delta_{12}\varphi_\lambda^{h,\varepsilon}|$ in (3.35). Plugging the above controls in (3.31) yields:

$$|\Delta H_{21}^{h,\varepsilon}(t_k, t_j, z, y)| \leq \frac{\Delta_{\varepsilon,\gamma,\infty}\Phi_{c,c_1}(t_j - t_k, z - y)}{(t_j - t_k)^{1-\gamma/2}}. \tag{3.39}$$

26   We emphasize that in the case of innovations with polynomial decays, the control on the difference of the
27   kernels again induces a loss of concentration of order γ in order to equilibrate the time singularity.

Term $\Delta_{22}H^{h,\varepsilon}$. This term can be handled with the same arguments as $\Delta_{21}H^{h,\varepsilon}$. For the sake of completeness we anyhow specify how the different contributions appear. Namely, with the notations of (3.30) and (3.31):

$$\Delta_2\varphi_\lambda^{h,\varepsilon}(t_k,z,y,w)$$

$$= \int_0^1 \mathrm{d}\mu\Big\{\Big\langle D_{T_1}\varphi_\lambda^h(T_0^h(t_k,y,w)+\mu(T^h(t_k,z,w)-T_0^h(t_k,y,w)),T_0^h(t_k,y,w)),T^h(t_k,z,w)-T_0^h(t_k,y,w)\Big\rangle$$

$$\Big\langle -D_{T_1}\varphi_{\lambda,\varepsilon}^h(T_{0,\varepsilon}^h(t_k,y,w)+\mu(T_\varepsilon^h(t_k,z,w)-T_{0,\varepsilon}^h(t_k,y,w)),T_{0,\varepsilon}^h(t_k,y,w)),T_\varepsilon^h(t_k,z,w)-T_{0,\varepsilon}^h(t_k,y,w)\Big\rangle\Big\}$$

$$= \Big\{\int_0^1 \mathrm{d}\mu\Big\{\Big\langle D_{T_1}\varphi_\lambda^h(T_0^h(t_k,y,w)+\mu(T^h(t_k,z,w)-T_0^h(t_k,y,w)),T_0^h(t_k,y,w)),$$

$$\Big[(T^h(t_k,z,w)-T_0^h(t_k,y,w))-(T_\varepsilon^h(t_k,z,w)-T_{0,\varepsilon}^h(t_k,y,w))\Big]\Big\rangle\Big\}$$

$$-\Big\{\int_0^1 \mathrm{d}\mu\Big[\Big\langle D_{T_1}\varphi_{\lambda,\varepsilon}^h(T_{0,\varepsilon}^h(t_k,y,w)+\mu(T_\varepsilon^h(t_k,z,w)-T_{0,\varepsilon}^h(t_k,y,w)),T_{0,\varepsilon}^h(t_k,y,w))$$

$$-D_{T_1}\varphi_\lambda^h(T_0^h(t_k,y,w)+\mu(T^h(t_k,z,w)-T_0^h(t_k,y,w)),T_0^h(t_k,y,w))\Big],T_\varepsilon^h(t_k,z,w)-T_{0,\varepsilon}^h(t_k,y,w)\Big\rangle\Big\}$$

$$=: (\Delta_{21}\varphi_\lambda^{h,\varepsilon}+\Delta_{22}\varphi_\lambda^{h,\varepsilon})(t_k,z,y,w).$$

In $\Delta_{21}\varphi_\lambda^{h,\varepsilon}$ we have sensitivities of order 3 for the density, giving time singularities in $(t_j-t_k)^{-3/2}$, which are again equilibrated by the the multiplicative factor:

$$|T_0^h(t_k,y,w)[T_0^h(t_k,y,w)]^*|\times|(T^h(t_k,z,w)-T_0^h(t_k,y,w))-(T_\varepsilon^h(t_k,z,w)-T_{0,\varepsilon}^h(t_k,y,w))|$$

$$\leq \bar{c}(h^2+h^{3/2}|w|+h|w|^2)\Delta_{\varepsilon,\gamma,\infty}(h+h^{1/2}(1\wedge|z-y|)^\gamma|w|),$$

where the last inequality is obtained similarly to (3.33) using as well (1.5). The same kind of controls can be established for $\Delta_{22}\varphi_\lambda^{h,\varepsilon}$. Anyhow, the analysis of this term leads to investigate the difference of third order derivatives, which finally yields contributions involving derivatives of order four. This is what induces the final concentration loss under $(\mathrm{I}_{P,M})$, *i.e.* we need to integrate a term in $|w|^4$ (see also Eq. (3.38) in which we performed the splitting of f_ξ on the *off-diagonal* region to integrate a contribution in $|w|^3$).

We can thus claim that

$$|\Delta H_{22}^{h,\varepsilon}(t_k,t_j,z,y)|\leq\frac{\Delta_{\varepsilon,\gamma,\infty}\Phi_{c,c_1}(t_j-t_k,z-y)}{(t_j-t_k)^{1-\gamma/2}}.$$

Plugging the above control and (3.39) (or (3.37) under (I_G)) into (3.31) we derive:

$$|\Delta H_2^{h,\varepsilon}(t_k,t_j,z,y)|\leq\frac{\Delta_{\varepsilon,\gamma,\infty}\Phi_{c,c_1}(t_j-t_k,z-y)}{(t_j-t_k)^{1-\gamma/2}},$$

which together with (3.29) and the decomposition (3.28) completes the proof. $\square$

From Lemmas 3.5 and 3.6 the proof of Theorem 1.3 is achieved, under (I_G), following the steps of Lemmas 3.3 and 3.4, using the Hölder inequalities for the differences of the drift terms for $q\in(d,+\infty)$.

The point is that we want to justify the following inequality under $(\mathrm{I}_{P,M})$ and $q=+\infty$:

$$|(\tilde{p}^h\otimes_h H^{h,(r)}-\tilde{p}_\varepsilon^h\otimes_h H_\varepsilon^{h,(r)})(t_i,t_j,x,y)|$$

$$\leq (r+1)\Delta_{\varepsilon,\gamma,\infty}\frac{\{(1\vee T^{(1-\gamma)/2})c_1\}^{r+1}\left[\Gamma(\frac{\gamma}{2})\right]^r}{\Gamma(1+r\frac{\gamma}{2})}\frac{c^d}{(t_j-t_i)^{d/2}}Q_{M-(d+5+\gamma)}\left(\frac{y-x}{(t_j-t_i)^{1/2}/c}\right)(t_j-t_i)^{\frac{r\gamma}{2}}. \quad (3.40)$$

The only delicate point, w.r.t. the analysis performed for diffusions, consists in controlling the convolutions of the densities with polynomial decay. To this end, we can adapt a technique used by Kolokoltsov [18] to investigate

convolutions of "stable like" densities. Set $m := M - (d + 5 + \gamma)$ and denote for all $0 \le i < j \le N,\ x \in \mathbb{R}^d$ by

$q_m(t_j - t_i, x) := \frac{c^d}{(t_j - t_i)^{d/2}} Q_{M-(d+5+\gamma)} \left(\frac{x}{(t_j - t_i)^{1/2}/c} \right)$ the density with polynomial decay appearing in Lemmas 3.5

and 3.6. Let us consider for fixed $i < k < j, (x, y) \in (\mathbb{R}^d)^2$ the convolution:

$$I^1_{t_k}(t_i, t_j, x, y) := \int_{\mathbb{R}^d} \mathrm{d}z\, q_m(t_k - t_i, z - x) q_m(t_j - t_k, y - z). \tag{3.41}$$

– If $|x - y| \le c(t_j - t_i)^{1/2}$ (diagonal regime for the parabolic scaling), it is easily seen that one of the two densities in the integral (3.41) is homogeneous to $q_m(t_j - t_i, y - x)$. Namely, if $(t_k - t_i) \ge (t_j - t_i)/2$,

$q_m(t_k - t_i, z - x) \le \frac{c^d c_m}{(t_k - t_i)^{d/2}} \le \frac{2^{d/2} c^d c_m}{(t_j - t_i)^{d/2}} \le \tilde{c} q_m(t_j - t_i, y - x).$ Thus,

$$I^1_{t_k}(t_i, t_j, x, y) \le \tilde{c} q_m(t_j - t_i, y - x) \int_{\mathbb{R}^d} \mathrm{d}z\, q_m(t_j - t_k, y - z) = \tilde{c} q_m(t_j - t_i, y - x).$$

If $(t_k - t_i) < (t_j - t_i)/2$, the same operation can be performed taking $q_m(t_j - t_k, y - z)$ out of the integral, observing again that in that case $q_m(t_j - t_k, y - z) \le \tilde{c} q_m(t_j - t_i, y - x)$.

– If $|x - y| > c(t_j - t_i)^{1/2}$ (off-diagonal regime), we introduce $A_1 := \{z \in \mathbb{R}^d : |x - z| \ge \frac{1}{2}|x - y|\}, A_2 := \{z \in \mathbb{R}^d : |z - y| \ge \frac{1}{2}|x - y|\}$. Every $z \in \mathbb{R}^d$ belongs at least to one of the $\{A_i\}_{i \in \{1,2\}}$. Let us assume w.l.o.g. that $z \in A_2$. Then $|z - y| \ge \frac{c}{2}(t_j - t_i)^{1/2} \ge \frac{c}{2}(t_j - t_k)^{1/2}$ so that the density $q_m(t_j - t_k, y - z)$ is itself in the off-diagonal regime. Write:

$$\int_{A_2} \mathrm{d}z\, q_m(t_k - t_i, z - x) q_m(t_j - t_k, y - z) \le \int_{A_2} \mathrm{d}z\, q_m(t_k - t_i, z - x) \frac{c_m(t_k - t_i)^{(m-d)/2}}{|z - y|^m}$$

$$\le \frac{c_m 2^m (t_j - t_i)^{(m-d)/2}}{|x - y|^m} \int_{A_2} \mathrm{d}z\, q_m(t_k - t_i, z - x) \le \bar{c} q_m(t_j - t_i, y - x),$$

recalling that, under $(I_{P,M})$, $m > d$ for the last but one inequality. The same operation could be performed on A_1.

We have thus established that, there exist $\bar{c} > 1$ s.t. for all $0 \le i < k < j, (x, y) \in (\mathbb{R}^d)^2$:

$$I^1_{t_k}(t_i, t_j, x, y) \le \bar{c} q_m(t_j - t_i, y - x).$$

From the controls of Lemma 3.6 and following the strategy of Lemma 3.4, we will be led to consider convolutions of the previous type involving Γ functions. The above strategy thus yields (3.40) by induction.

References

[1] D.G. Aronson, The fundamental solution of a linear parabolic equation containing a small parameter. *Ill. J. Math.* **3** (1959) 580–619.

[2] E. Benhamou, E. Gobet and M. Miri, Expansion formulas for European options in a local volatility model. *Inter. J. Theor. Appl. Finance* **13** (2010) 602–634.

[3] R.F. Bass and E.A. Perkins, A new technique for proving uniqueness for martingale problems. *From Probability to Geometry (I): Volume in Honor of the 60th Birthday of Jean–Michel Bismut* (2009) 47–53.

[4] R. Bhattacharya and R. Rao, *Normal approximations and asymptotic expansions.* Wiley and sons (1976).

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

[6] F. Corielli, P. Foschi and Pascucci, Parametrix approximation of diffusion transition densities. *SIAM J. Financial Math.* **1** (2010) 833–867.

[7] H. Cramér and M.R. Leadbetter, *Stationary and related stochastic processes.* Dover Publications, Inc., Mineola, NY (2004). Sample function properties and their applications, Reprint of the 1967 original.

[8] F. Delarue and S. Menozzi, Density estimates for a random noise propagating through a chain of differential equations. *J. Funct. Anal.* **259** (2010) 1577–1630.

[9] E.B. Dynkin, *Markov Processes.* Springer Verlag (1965).

[10] A. Friedman, *Partial Differential Equations of Parabolic Type.* Prentice-Hall (1964).

[11] A. Friedman, *Stochastic differential equations.* Chapmann-Hall (1975).

[12] V. Genon−Catalot and J. Jacod, On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Annal. l'I.H.P. Probab. Stat.* **29** (1993) 119–151.

[13] A.M. Il'in, A.S. Kalashnikov and O.A. Oleinik, Second-order linear equations of parabolic type. *Uspehi Mat. Nauk* **17** (1962) 3–146.

[14] V. Konakov and E. Mammen, Local limit theorems for transition densities of Markov chains converging to diffusions. *Prob. Th. Rel. Fields* **117** (2000) 551–587.

[15] V. Konakov and E. Mammen, Edgeworth type expansions for Euler schemes for stochastic differential equations. *Monte Carlo Methods Appl.* **8** (2002) 271–285.

[16] V. Konakov and S. Menozzi, Weak error for the Euler scheme of a diffusion with non-regular coefficients. *Preprint ArXiV:1604.00771* (2015).

[17] V. Konakov, S. Menozzi and S. Molchanov, Explicit parametrix and local limit theorems for some degenerate diffusion processes. *Annales de l'Institut Henri Poincaré, Série B* **46** (2010) 908–923.

[18] V. Kolokoltsov, Symmetric stable laws and stable-like jump diffusions. *Proc. London Math. Soc.* **80** (2000) 725–768.

[19] N.V. Krylov, *Lectures on elliptic and parabolic equations in Hölder spaces.* Graduate Studies in Mathematics, Vol. 12. AMS (1996).

[20] S. Menozzi, Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electr. Commun. Prob.* **17** (2011) 234–250.

[21] R. Mikulevičius and E. Platen, Rate of convergence of the Euler approximation for diffusion processes. *Math. Nachr.* **151** (1991) 233–239.

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

[23] S.J. Sheu, Some estimates of the transition density of a nondegenerate diffusion Markov process. *Ann. Probab.* **19** (1991) 538–561.

[24] A.N. Shiryaev, *Probability, 2nd Edition.* Graduate Texts in Mathematics, Vol. 95. Springer-Verlag, New York (1996).

[25] D.W. Stroock and S.R.S. Varadhan, *Multidimensional diffusion processes.* Springer-Verlag Berlin Heidelberg New York (1979).