



**HAL**  
open science

# STABILITY OF DENSITIES FOR PERTURBED DIFFUSIONS AND MARKOV CHAINS

V Konakov, A Kozhina, S Menozzi

► **To cite this version:**

V Konakov, A Kozhina, S Menozzi. STABILITY OF DENSITIES FOR PERTURBED DIFFUSIONS AND MARKOV CHAINS. 2015. hal-01169685v1

**HAL Id: hal-01169685**

**<https://hal.science/hal-01169685v1>**

Preprint submitted on 30 Jun 2015 (v1), last revised 25 Nov 2016 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# STABILITY OF DENSITIES FOR PERTURBED DIFFUSIONS AND MARKOV CHAINS

V. KONAKOV, A. KOZHINA, AND S. MENOZZI

ABSTRACT. We are interested in studying the sensitivity of diffusion processes or their approximations by Markov Chains with respect to a perturbation of the coefficients. As an important application, we give a first order expansion for the difference of the densities of a diffusion with Hölder coefficients and its approximation by the Euler scheme.

## 1. INTRODUCTION

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

$$(1.1) \quad dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad t \in [0, T],$$

where  $b : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $\sigma : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d$  are bounded coefficients that are measurable in time and Hölder continuous in space 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 [BP09], [Men11] from which the uniqueness to the martingale problem for the associated generator can be derived under the current assumptions.

We now introduce a perturbed version of (1.1) with dynamics:

$$(1.2) \quad dX_t^{(n)} = b_n(t, X_t^{(n)})dt + \sigma_n(t, X_t^{(n)})dW_t, \quad t \in [0, T],$$

where  $b_n : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $\sigma_n : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d$  satisfy at least the same assumptions as  $b, \sigma$  and are in some sense meant to be *close* to  $b, \sigma$  for large values of  $n \in \mathbb{N}$ .

It is known that, under the previous assumptions, the density of the processes  $(X_t)_{t \geq 0}, (X_t^{(n)})_{t \geq 0}$  exists and satisfies some Gaussian bounds, see e.g. Aronson [Aro59], Sheu [She91] or [DM10] for extensions to some degenerate cases.

The first goal of this work is to investigate how the closeness of  $(b_n, \sigma_n)$  and  $(b, \sigma)$  is reflected on the respective densities of the associated processes. Important applications include, for instance, the case of mollification by spatial convolution, i.e. when  $b_n(t, x) := b(t, \cdot) \star \rho_n(x), \sigma_n(t, x) := \sigma(t, \cdot) \star \rho_n(x)$ ,  $x \in \mathbb{R}^d$  where  $\rho_n$  is a smooth mollifying kernel, i.e.  $\rho_n : \mathbb{R}^d \rightarrow \mathbb{R}^+, \rho_n \in C^\infty(\mathbb{R}^d, \mathbb{R}^+), \int_{\mathbb{R}^d} \rho_n(x) dx = 1, |\text{supp}(\rho_n)| \xrightarrow{n} 0$ . This specific kind of perturbation will be used to derive our

---

*Date:* June 30, 2015.

*2000 Mathematics Subject Classification.* Primary 60H10; Secondary 65C30.

*Key words and phrases.* Diffusion Processes, Parametrix, Hölder Coefficients, Euler Scheme, Weak Error.

second main result which precisely quantifies the error between the densities of a non-degenerate diffusion of type (1.1) with Hölder coefficients and its Euler scheme.

Other possible applications can be found in mathematical finance. If the dynamics of (1.1) models the evolution of the price of a financial asset, it is often very useful to know how a perturbation of the volatility  $\sigma$  impacts the density, and therefore the associated option prices. More generally, this situation can appear in every applicative field for which the diffusion coefficient might be misspecified.

The previously mentioned Gaussian bounds on the density are derived through the so-called *parametrix* expansion which will be the crux of our approach. Roughly speaking, it consists in approximating the process by a proxy which has a known density, here a Gaussian one, and then in investigating the difference through the Kolmogorov equations. Various approaches to the parametrix expansion exist, see e.g. Il'in *et al.* [IKO62], Friedman [Fri64] and McKean and Singer [MS67]. The latter approach will be the one used in this work since it appears to be the most adapted to handle coefficients with no *a priori* smoothness in time and can also be directly extended to the discrete case for Markov chain approximations of equations (1.1) and (1.2). Let us mention in this setting the works of Konakov *et al.* ([KM00], [KM02]).

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

$$(1.3) \quad \begin{aligned} Y_{t_{k+1}} &= Y_{t_k} + b(t_k, Y_{t_k})h + \sigma(t_k, Y_{t_k})\sqrt{h}\xi_{k+1}, Y_0 = x, \\ Y_{t_{k+1}}^{(n)} &= Y_{t_k}^{(n)} + b_n(t_k, Y_{t_k}^{(n)})h + \sigma_n(t_k, Y_{t_k}^{(n)})\sqrt{h}\xi_{k+1}, Y_0^{(n)} = x, \end{aligned}$$

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 i.i.d. random variables satisfying some integrability conditions. It is exactly the combination of the two stability Theorems, for diffusions and Markov chains, that will give the error bound on the densities for the Euler scheme of a non-degenerate diffusion in the weak setting of Hölder coefficients. There is a huge literature concerning the weak error for smooth and/or non-degenerate coefficients, from the seminal paper of Talay and Tubaro [TT90], to the extensions to the hypoelliptic framework [BT96a], [BT96b] or for Markov Chain approximations [KM00], [KM02], but the Hölder case has rarely been considered.

In this framework, let us mention the work of Mikulevičius and Platen [MP91] who obtained bounds for the weak error, namely for the quantity  $d(f, x, T, h) := \mathbb{E}[f(X_T)] - \mathbb{E}[f(X_T^h)]$  where  $X_T^h$  stands for the terminal value at time  $T$  of the Euler scheme  $X^h$  with time step  $h = T/N$ ,  $N \in \mathbb{N}^*$  associated with  $X$ . They managed to get a bound at rate  $h^{\gamma/2}$  where  $\gamma \in (0, 1)$  is the Hölder exponent of the coefficients  $b, \sigma$  in (1.1) provided  $f \in C^{2+\gamma}$ . This regularity is essential in that work to apply Itô's formula. Our approach permits to establish that this bound holds true for the difference of the densities itself, which corresponds to the weak error for a  $\delta$ -function. We also mention the recent work of Mikulevičius *et al.* [Mik12], [MZ15], concerning some extensions of [MP91] to jump-driven SDEs with Hölder coefficients.

Finally, concerning numerical schemes for diffusions with non-regular coefficients, we refer to the recent work of Kohatsu-Higa *et al.* [KHL15] who investigate the weak error for possibly discontinuous drifts and diffusion coefficients that are just continuous.

**1.2. Assumptions and Main Results.** Let us introduce the following assumptions.

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

$$\begin{aligned} \sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |b(t, x)| + \sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |b_n(t, x)| &\leq K_1, \\ \sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |\sigma(t, x)| + \sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |\sigma_n(t, x)| &\leq K_2. \end{aligned}$$

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

$$\begin{aligned} \Lambda^{-1}|\xi|^2 &\leq \langle a(t, x)\xi, \xi \rangle \leq \Lambda|\xi|^2 \\ \Lambda^{-1}|\xi|^2 &\leq \langle a_n(t, x)\xi, \xi \rangle \leq \Lambda|\xi|^2. \end{aligned}$$

**(A3) (Hölder continuity in space).** For some  $\gamma \in (0, 1]$ ,  $\kappa < \infty$ ,  $N \in \mathbb{N}$  for all  $t \in \mathbb{R}_+$ ,  $n \geq N$ ,

$$\begin{aligned} |b(t, x) - b(t, y)| + |b_n(t, x) - b_n(t, y)| &\leq \kappa |x - y|^\gamma; \\ |\sigma(t, x) - \sigma(t, y)| + |\sigma_n(t, x) - \sigma_n(t, y)| &\leq \kappa |x - y|^\gamma. \end{aligned}$$

Observe that the last condition also readily gives, thanks to the boundedness of  $\sigma, \sigma_n$  that  $a, a_n$  are also uniformly  $\gamma$ -Hölder continuous.

**(A4) (Closeness of  $(b_n, \sigma_n)$  and  $(b, \sigma)$ ).** Set for  $n \in \mathbb{N}$ :

$$\Delta_{n,b} := \sup_{(t,x) \in [0,T] \times \mathbb{R}^d} \{|b(t, x) - b_n(t, x)|\}.$$

Since  $\sigma, \sigma_n$  are both  $\gamma$ -Hölder continuous, see **(A3)** we also define

$$\Delta_{n,\sigma} := \sup_{u \in [0,T]} |\sigma(u, \cdot) - \sigma_n(u, \cdot)|_\gamma,$$

where  $|\cdot|_\gamma$  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 [Kry96]) i.e. :

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

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

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

Setting  $\Delta_n := \Delta_{n,\sigma} + \Delta_{n,b}$  we assume  $\Delta_n \xrightarrow[n]{} 0$ .

We say that assumption **(A)** holds when conditions **(A1)**-**(A4)** are in force. 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

We are now in position to state our main results.

*Theorem 1* (Stability Control for diffusions). Fix  $T > 0$ . Under **(A)** there exists  $C \geq 1, c \in (0, 1]$  s.t. for all  $0 \leq s < t \leq T$ ,  $(x, y) \in (\mathbb{R}^d)^2$ :

$$(1.4) \quad p_c(t - s, y - x)^{-1} |(p - p_n)(s, t, x, y)| \leq C\Delta_n,$$

where  $p(s, t, x, \cdot), p_n(s, t, x, \cdot)$  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  $(u, z) \in \mathbb{R}^{+*} \times \mathbb{R}^d$ ,  $p_c(u, z) := \frac{c^{d/2}}{(2\pi u)^{d/2}} \exp(-c \frac{|z|^2}{2u})$ .

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

(**I<sub>G</sub>**) 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) correspond to the Euler discretization of equations (1.1) and (1.2).

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

$$(1.5) \quad |D^\nu f_\xi(z)| \leq C Q_M(z),$$

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

*Theorem 2* (Stability Control for Markov Chains). Fix  $T > 0$ . For  $h = T/N$ ,  $N \in \mathbb{N}^*$ , we set for  $i \in \mathbb{N}$ ,  $t_i := ih$ . Under (**A**), assuming that either (**I<sub>G</sub>**) or (**I<sub>P,M</sub>**) holds for  $M > 2d + 4$ , there exist  $C := C \geq 1, c \in (0, 1]$  s.t. for all  $0 \leq t_i < t_j \leq T$ ,  $(x, y) \in (\mathbb{R}^d)^2$ :

$$(1.6) \quad \chi_c(t_j - t_i, y - x)^{-1} |(p^h - p_n^h)(t_i, t_j, x, y)| \leq C \Delta_n,$$

where  $p^h(t_i, t_j, x, \cdot), p_n^h(t_i, t_j, x, \cdot)$  respectively stand for the transition densities at time  $t_j$  of equations in (1.3) starting from  $x$  at time  $t_i$ . Also:

- If (**I<sub>G</sub>**) holds:

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

with  $p_c$  as in Theorem 1.

- If (**I<sub>P,M</sub>**) holds:

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

As an application of the previous results we also derive the following important Theorem.

*Theorem 3* (Error for the Euler scheme of a diffusion with Hölder coefficients). Let  $T > 0$  be fixed and consider a given time step  $h := T/N$ ,  $N \in \mathbb{N}^*$ . Set for  $i \in \mathbb{N}$ ,  $t_i := ih$ . Under (**A**), if we additionally assume that the coefficients  $a, b$  are  $\gamma/2$ -Hölder continuous in time, there exist  $C \geq 1, c \in (0, 1]$  s.t. for all  $0 \leq t_i < t_j \leq T$ ,  $(x, y) \in (\mathbb{R}^d)^2$ :

$$(1.7) \quad p_c(t - s, y - x)^{-1} |(p - p^h)(t_i, t_j, x, y)| \leq \frac{C}{(t_j - t_i)^{(1-\gamma/4)\gamma/2}} h^{\frac{\gamma}{2}},$$

where  $p, p^h$  respectively stand for the densities of the diffusion  $X$  and its Euler approximation  $X^h$  with time step  $h$  and  $p_c$  is as in Theorem 1.

Observe here that the rate  $h^{\gamma/2}$  is quite *natural*. It corresponds to the typical magnitude of the quantity  $\mathbb{E}[|W_h|^\gamma] \leq c_\gamma h^{\gamma/2}$ , which reflects the variations, on one

time-step of length  $h$ , of the Euler scheme with Hölder coefficients. Indeed, for all  $i \in \llbracket 0, N-1 \rrbracket$ :

$$\begin{aligned} \mathbb{E}\left[\sup_{u \in [t_i, t_{i+1}]} |\sigma(u, X_u^h) - \sigma(t_i, X_{t_i}^h)|\right] &\leq \kappa \left\{ h^{\gamma/2} + \mathbb{E}\left[\sup_{u \in [t_i, t_{i+1}]} |X_u^h - X_{t_i}^h|^\gamma\right] \right\} \\ &\leq \kappa \left\{ h^{\gamma/2} + \mathbb{E}\left[\left\{\sup_{u \in [t_i, t_{i+1}]} |\sigma(t_i, X_{t_i}^h)(W_u - W_{t_i})| + K_1 h\right\}^2\right]^{\gamma/2} \right\} \leq ch^{\gamma/2}, \end{aligned}$$

supposing that the Hölder constant in time is as well  $\kappa$ . We also refer to the Lemma in Section 6 of [MP91].

*Remark 1.* We point out that the previous result extends to the densities of the processes involved the results from Mikulevičius and Platen [MP91], obtained for the weak error. The framework they considered is similar to ours, and their main results consists in controlling at rate  $h^{\gamma/2}$  the *weak* error  $d(f, x, T, h) := \mathbb{E}_x[f(X_T)] - \mathbb{E}_x[f(X_T^h)]$  for a *smooth* function  $f \in C^{2+\gamma}$ . The above theorem establishes that  $|d(f, x, T, h)| \leq Ch^{\gamma/2}$  as soon as  $f$  is measurable and satisfies the following growth condition:

$$\exists c_0 < c, C_0, \forall x \in \mathbb{R}^d, |f(x)| \leq C_0(1 + \exp(c_0|x|^2)),$$

where the  $c$  is the one of  $p_c$  in (1.7). This point can be particularly relevant if we think for instance about quantile estimation (that would involve functions of the form  $f(x) = \mathbb{1}_{|x| \leq K}$  or  $f(x) = \mathbb{1}_{|x| \leq K} \exp(c|x|)$ ) that appear in many applications: default probabilities in mathematical finance, fatigue of structures in random mechanics.

*Remark 2.* Even though we have considered  $\gamma \in (0, 1]$ , our analysis should be extendable in the framework of Hölder spaces to  $\gamma \in (1, 2]$ . On the other hand, Theorem 3 specifies the time-singularity in small time. The result is again coherent with the one obtained for *smooth* coefficients, see Gobet and Labart [GL08] for an explicit bound, or [KM02] from which it can be derived that the time singularity behaves as  $(t_j - t_i)^{-1/2}$ . This bound is obtained from (1.7) taking  $\gamma = 2$  which is the threshold that, as already observed in [MP91], gives the usual convergence rate for the weak error.

The paper is organized as follows. We recall in Section 2 some basic facts about parametrix expansions for the densities of diffusions and Markov Chains. We then detail in Section 3 how to perform a stability analysis of the parametrix expansions in order to derive the results of Theorems 1 and 2. Section 4 is eventually devoted to the convergence analysis of the weak error for the Euler scheme of a non-degenerate diffusion with Hölder coefficients. The proof combines two elements. We first apply our previous sensitivity results of Theorems 1 and 2 to quantify the difference of densities between the initial diffusion or scheme and a perturbation associated with a mollification of coefficients. The next step consists in adapting the weak error expansion for the densities in [KM02] for diffusions/schemes having coefficients with explosive derivatives. The final result follows equilibrating the two error sources.

## 2. DERIVATION OF FORMAL SERIES EXPANSION FOR DENSITIES

**2.1. Parametrix Representation of the Density for Diffusions.** 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 solution of (1.1) starting from  $x$  at time  $s$ .

Assume that  $(X_t^{s,x})_{t \geq s}$  has for all  $t > s$  a smooth density  $p(s, t, x, \cdot)$  (which is the case if additionally to **(A)** the coefficients are smooth see e.g. Friedman [Fri64]). We would like to estimate this density at a given point  $y \in \mathbb{R}^d$ . To this end, we introduce the following Gaussian inhomogeneous process with spatial variable frozen at  $y$ . For all  $(s, x) \in [0, T] \times \mathbb{R}^d$ ,  $t \geq s$  we set:

$$\tilde{X}_t^y = x + \int_s^t b(u, y) du + \int_s^t \sigma(u, y) dW_u.$$

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

$$(2.1) \quad \begin{cases} \partial_u \tilde{p}^y(u, t, z, y) + \tilde{L}_u^y \tilde{p}^y(u, t, z, y) = 0, & s \leq u < t, z \in \mathbb{R}^d, \\ \tilde{p}^y(u, t, \cdot, y) \xrightarrow{u \uparrow t} \delta_y(\cdot), \end{cases}$$

where for all  $\varphi \in C_0^2(\mathbb{R}^d, \mathbb{R})$ ,  $z \in \mathbb{R}^d$ :

$$\tilde{L}_u^y \varphi(z) = \frac{1}{2} \text{Tr} (\sigma \sigma^*(u, y) D_z^2 \varphi(z)) + \langle b(u, y), D_z \varphi(z) \rangle,$$

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 [Dyn65]). For a given starting point  $x \in \mathbb{R}^d$  at time  $s$ ,

$$(2.2) \quad \begin{cases} \partial_u p(s, u, x, z) = L_u^* p(s, u, x, z) = 0, & s < u \leq t, z \in \mathbb{R}^d, \\ p(s, u, x, \cdot) \xrightarrow{u \downarrow s} \delta_x(\cdot), \end{cases}$$

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:

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

Equations (2.1), (2.2) yield the formal expansion below which is initially due to McKean and Singer [MS67].

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

using the Dirac convergence for the first equality, equations (2.2) and (2.1) for the second 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

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

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:

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

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:

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

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 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,

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

and

$$(2.7) \quad p_n(s, t, x, y) = \tilde{p}_n(s, t, x, y) + \sum_{r=1}^{\infty} \tilde{p}_n \otimes H_n^{(r)}(s, t, x, y).$$

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

The convergence of the series (2.6) and (2.7) is in some sense *standard* (see e.g. [Men11] or Friedman [Fri64]) under **(A)**. We recall for the sake of completeness the key steps for (2.6).

Let us recall that there exist  $c_1 \geq 1$ ,  $c \in (0, 1]$  s.t. for all  $T > 0$  and all multi-index  $\alpha$ ,  $|\alpha| \leq 4$ ,

$$(2.8) \quad |D_z^\alpha \tilde{p}(u, t, z, y)| \leq \frac{c_1 \exp(c_1 K_1^2(t-u))}{(t-u)^{|\alpha|/2}} p_c(t-u, y-z),$$

where we used that  $|z - y - \int_u^t b(v, y) dv|^2 \geq \frac{|z-y|^2}{2} - K_1^2(t-u)^2$  with  $K_1$  being the bound for the drift in Assumption **(A1)** and

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

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

$$(2.9) \quad |H(u, t, z, y)| \leq \frac{c_1 \exp(c_1 K_1^2(t-u))}{(t-u)^{1-\gamma/2}} p_c(t-u, z-y).$$

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:

$$(2.10) \quad \begin{aligned} & |\tilde{p} \otimes H^{(r)}(s, t, x, y)| \\ & \leq \exp(c_1 K_1^2(t-s)) c_1^{r+1} \prod_{i=1}^r B\left(\frac{\gamma}{2}, 1 + (i-1)\frac{\gamma}{2}\right) p_c(t-s, y-x)(t-s)^{\frac{r\gamma}{2}} \\ & = \exp(c_1 K_1^2(t-s)) \frac{c_1^{r+1} [\Gamma(\frac{\gamma}{2})]^r}{\Gamma(1 + r\frac{\gamma}{2})} p_c(t-s, y-x)(t-s)^{\frac{r\gamma}{2}}. \end{aligned}$$

These bounds readily yield the convergence of the series as well as a Gaussian upper-bound. Namely

$$(2.11) \quad p(s, t, x, y) \leq c_1 \exp(c_1[(t-s)^{\gamma/2} + K_1^2(t-s)]) p_c(t-s, y-x).$$

An important application of the *stability* of the perturbation consists in considering coefficients  $b_n := b \star \zeta_n$ ,  $\sigma := \sigma \star \zeta_n$  in (2.7), where  $\zeta_n$  is a mollifier in time and space. For mollified coefficients, the existence and smoothness of the density  $p_n$  for the associated process  $X^{(n)}$  in (1.2) can be derived from [IKO62]. Observe carefully that the previous Gaussian bounds also hold for  $p_n$  uniformly in  $n$  and independently of the mollifying procedure. This therefore gives that

$$(2.12) \quad p_n(s, t, x, y) \xrightarrow[n]{} p(s, t, x, y),$$

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

$$\mathbb{E}_{s,x}[f(X_t^{(n)})] = \int_{\mathbb{R}^d} f(y) p_n(s, t, x, y) dy \xrightarrow[n]{} \int_{\mathbb{R}^d} f(y) p(s, t, x, y) dy.$$

In particular, taking  $f = 1$  gives that  $\int_{\mathbb{R}^d} p(s, t, x, y) dy = 1$ .

We now specify how we can derive that under **(A)** the sum  $p(s, t, x, y)$  of the parametrix series can actually be identified with a density. Recall indeed, that the uniform convergence in (2.12) gives that  $p(s, t, x, \cdot)$  is non negative.

It can on the other hand be proved by Kolmogorov's tightness criterion, see e.g. [Bil99], that the sequence of measures  $(\mu^n)_{n \geq 0}$  associated with  $(X_t^{(n)})_{t \geq 0}$  on the space of continuous functions is tight. By Prokhorov's Theorem it is thus weakly relatively compact. Now, for a given  $t > s \geq 0$ ,  $x \in \mathbb{R}^d$ , we therefore deduce that for any continuous bounded continuous  $f$  and any converging subsequence

$(\mu_t^{(n_k),s,x})_{k \geq 0}$ , denoting for all  $n \in \mathbb{N}$  by  $\mu_t^{(n),s,x}$  the law of  $X_t^{(n),s,x}$ , we have for every continuous bounded function  $f$ :

$$\mu_t^{(n_k),s,x}(f) := \mathbb{E}_{s,x}[f(X_t^{(n_k)})] \xrightarrow[k]{\rightarrow} \int_{\mathbb{R}^d} f(y)p(s,x,t,y)dy.$$

This proves that the weak limit of  $(\mu_t^{(n),s,x})_{n \geq 0}$  (and even  $(\mu^n)_{n \geq 0}$ ) exists and is unique.

It can now be proved following the steps of Oleinik *et al.* [IKO62] that the density  $p(s,t,x,\cdot)$  is actually the unique fundamental solution of

$$\begin{aligned} (\partial_s + L_s)p(s,t,x,y) &= 0, \quad 0 \leq s < t, (x,y) \in (\mathbb{R}^d)^2, \\ p(s,t,x,y)dy &\xrightarrow[t \downarrow s]{} \delta_x(dy). \end{aligned}$$

We anyhow carefully mention that the freezing in the quoted paper is performed w.r.t. the starting time and space variables. This approach leads to an extra Hölder continuity in time assumption, see also Friedman [Fri64], which we do not need here since we follow the McKean and Singer approach to the parametrix expansion [MS67].

We can thus now conclude that the processes  $X, X^n$  in (1.1), (1.2) have transition densities given by the sum of the series (2.6), (2.7). □

**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  $(Y_j^{t_i,x})_{j \geq i}$  the Markov chain with dynamics (1.3) starting from  $x$  at time  $t_i$ . Observe first that if the innovations  $(\xi_k)_{k \geq 1}$  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  $\varphi \in C_0^2(\mathbb{R}^d, \mathbb{R})$ ,  $x \in \mathbb{R}^d$ :

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

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  $(t_i, x) \in [0, T] \times \mathbb{R}^d$ ,  $t_j \geq t_i$  we set:

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

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

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

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:

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

Analogously to Lemma 3.6 in [KM00], 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 1.

*Proposition 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

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

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. to the coefficients through the difference of the series.

**3.1. Stability for Diffusions: Proof of Theorem 1.** Let us consider the difference between two parametrix expansions:

$$\begin{aligned} & |p(s, t, x, y) - p_n(s, t, x, y)| = \\ & = \left| \sum_{r=0}^{\infty} \tilde{p} \otimes H^{(r)}(s, t, x, y) - \sum_{r=0}^{\infty} \tilde{p}_n \otimes H_n^{(r)}(s, t, x, y) \right| \leq \\ (3.1) \quad & \leq |(\tilde{p} - \tilde{p}_n)(s, t, x, y)| + \left| \sum_{r=1}^{\infty} \tilde{p} \otimes H^{(r)}(s, t, x, y) - \sum_{r=1}^{\infty} \tilde{p}_n \otimes H_n^{(r)}(s, t, x, y) \right|. \end{aligned}$$

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 1* (Difference of the first terms and their derivatives). 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  $\alpha$ ,  $|\alpha| \leq 4$ ,

$$|D_x^\alpha \tilde{p}(s, t, x, y) - D_x^\alpha \tilde{p}_n(s, t, x, y)| \leq \frac{c_1}{(t-s)^{|\alpha|/2}} \exp(c_1 K_1^2(t-s)) \Delta_n p_c(t-s, y-x).$$

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

$$\begin{aligned} & b(s, t, y) := \int_s^t b(u, y) du, \quad b_n(s, t, y) := \int_s^t b_n(u, y) du, \\ (3.2) \quad & \Sigma(s, t, y) := \int_s^t a(u, y) du, \quad \Sigma_n(s, t, y) := \int_s^t a_n(u, y) du. \end{aligned}$$

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

$$\Sigma(s, t, y) = ( \Sigma^1 \mid \Sigma^2 \mid \cdots \mid \Sigma^d ) (s, t, y).$$

Recall now that  $b(s, t, y), b_n(s, t, y)$  are also viewed as  $d$ -dimensional column vectors, we rewrite:

$$\begin{aligned}\tilde{p}(s, t, x, y) &= f_{x,y}(\Theta(s, t, y)), \quad \Theta(s, t, y) = (b^*, (\Sigma^1)^*, \dots, (\Sigma^d)^*)^*(s, t, y), \\ \tilde{p}_n(s, t, x, y) &= f_{x,y}(\Theta_n(s, t, y)), \quad \Theta_n(s, t, y) = (b_n^*, (\Sigma_n^1)^*, \dots, (\Sigma_n^d)^*)^*(s, t, y),\end{aligned}$$

with

$$(3.3) \quad \begin{aligned}f_{x,y} : \mathbb{R}^{d(d+1)} &\rightarrow \mathbb{R} \\ \Gamma &\mapsto f_{x,y}(\Gamma) = \frac{1}{(2\pi)^{d/2} \det(\Gamma^{2:(d+1)})^{1/2}} \\ &\quad \times \exp\left(-\frac{1}{2} \langle (\Gamma^{2:(d+1)})^{-1} (y - x - \Gamma^1), y - x - \Gamma^1 \rangle\right),\end{aligned}$$

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

$$\Gamma^{2:(d+1)} := (\Gamma^2 \mid \Gamma^3 \mid \dots \mid \Gamma^{d+1}),$$

the  $d \times d$  matrix formed with the entries  $(\Gamma^i)_{i \in [2, d+1]}$ , each entry being viewed as a column.

The multidimensional Taylor expansion now gives:

$$(3.4) \quad \begin{aligned} |(\tilde{p} - \tilde{p}_n)(s, t, x, y)| &= |f_{x,y}(\Theta(s, t, y)) - f_{x,y}(\Theta_n(s, t, y))| \\ &= \left| \sum_{|\nu|=1} D_\nu f_{x,y}(\Theta(s, t, y)) \{(\Theta_n - \Theta)(s, t, y)\}^\nu \right. \\ &\quad \left. + 2 \sum_{|\nu|=2} \frac{\{(\Theta_n - \Theta)(s, t, y)\}^\nu}{\nu!} \int_0^1 (1 - \delta) D^\nu f_{x,y}([\Theta + \delta(\Theta_n - \Theta)](s, t, y)) d\delta \right|,\end{aligned}$$

where for a multi-index  $\nu := (\nu_1, \dots, \nu_{d(d+1)}) \in \mathbb{N}^{d(d+1)}$ , we denote by  $|\nu| := \sum_{i=1}^{d(d+1)} \nu_i$  the length of the multi-index,  $\nu! = \prod_{i=1}^{d(d+1)} \nu_i!$  and for  $h \in \mathbb{R}^{d(d+1)}$ ,  $h^\nu := \prod_{i=1}^{d(d+1)} h_i^{\nu_i}$  (with the convention that  $0^0 = 1$ ). With these notations, from (3.2), (3.3), (3.4) and Assumption **(A4)** we get:

$$(3.5) \quad \begin{aligned} &|f_{x,y}(\Theta(s, t, y)) - f_{x,y}(\Theta_n(s, t, y))| \\ &\leq c \left\{ \sum_{|\nu|=1} |D^\nu f_{x,y}(\Theta(s, t, y))| \Delta_n(t - s) \right. \\ &\quad \left. + \Delta_n^2(t - s)^2 \max_{\delta \in [0, 1]} \sum_{|\nu|=2} |D^\nu f_{x,y}([\Theta + \delta(\Theta_n - \Theta)](s, t, y))| \right\}.\end{aligned}$$

For our computations we need to separate the coordinates associated with drift contributions and those that correspond to an entry of the diffusion matrix. To this end, for a multi-index  $\nu \in \mathbb{N}^{d(d+1)}$  we write  $\nu = (\nu_1, \nu_{2:(d+1)})$  where  $\nu_1 := (\nu_1^1, \dots, \nu_1^d) \in \mathbb{N}^d$  is associated with the drift, and  $\nu_{2:(d+1)} := (\nu_{2:(d+1)}^1, \dots, \nu_{2:(d+1)}^{d^2}) \in \mathbb{N}^{d^2}$  with the diffusion matrix.

For  $\nu = (\nu_1, \nu_{2:(d+1)})$ , since  $f_{x,y}$  in (3.3) is a Gaussian density in the parameters  $x, y$ , the following properties hold:

- If  $|\nu_1| = 1$ ,  $\forall \Gamma \in \mathbb{R}^{d(d+1)}$ ,  $D^{\nu_1} f_{x,y}(\Gamma) = \prod_{i=1}^d \left( \frac{\partial}{\partial x_i} \right)^{\nu_1^i} f_{x,y}(\Gamma)$ ,
- If  $|\nu_{2:(d+1)}| = 1$ , we recall from Cramer ad Leadbetter [CL04] (see eq. (2.10.3) therein), that for all  $\Gamma \in \mathbb{R}^{d(d+1)}$ :

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

where  $\lfloor \cdot \rfloor$  stands for the integer part.

- If  $|\nu_1| = 1$ ,  $|\nu_{2:(d+1)}| = 1$  one combines the previous controls to derive that for all  $\Gamma \in \mathbb{R}^{d(d+1)}$ :

$$D^{\nu_1, \nu_{2:(d+1)}} f_{x,y}(\Gamma) = \prod_{i=1}^d \left( \frac{\partial}{\partial x_i} \right)^{\nu_1^i} \frac{1}{2} \left( \prod_{i=1}^{d^2} \left( \frac{\partial^2}{\partial x_{\lfloor \frac{i-1}{d} \rfloor + 1} \partial x_{i - \lfloor \frac{i-1}{d} \rfloor d}} \right)^{\nu_{2:(d+1)}^i} f_{x,y}(\Gamma) \right).$$

This gives that the associated singularity is homogeneous to the one induced by a third spatial derivative of the Gaussian density. We thus have:

$$|D^{\nu_1, \nu_{2:(d+1)}} f_{x,y}(\Gamma)| \leq \bar{c}_1 \prod_{i=1}^d |\{(\Gamma^{2:(d+1)})^{-1}(x - y - \Gamma_1)\}_i|^{\nu_1^i} \\ \prod_{i=1}^{d^2} |\{(\Gamma^{2:(d+1)})^{-1}(x - y - \Gamma_1)\}_{\lfloor \frac{i-1}{d} \rfloor + 1} \{(\Gamma^{2:(d+1)})^{-1}(x - y - \Gamma_1)\}_{i - \lfloor \frac{i-1}{d} \rfloor d}|^{\nu_{2:(d+1)}^i} \\ \times f_{x,y}(\Gamma).$$

Hence, taking from (3.5), for all  $\delta \in [0, 1]$ ,  $\Gamma_{n,\delta}(s, t, y) = [\Theta + \delta(\Theta_n - \Theta)](s, t, y)$  yields, thanks to the non-degeneracy conditions:

$$|D^{\nu_1, \nu_{2:(d+1)}} f_{x,y}(\Gamma_{n,\delta}(s, t, y))| \leq \frac{\bar{c}_1}{(t-s)^{3/2}} f_{x,y}(\Gamma_{n,\delta}(s, t, y)) \\ \leq \frac{\bar{c}_1}{(t-s)^{3/2}} \exp(\bar{c}_1 K_1^2(t-s)) p_{\bar{c}}(t-s, y-x),$$

for some  $\bar{c}_1 \geq 1$ ,  $\bar{c} \in (0, 1]$ . Observe now that similar computations actually give that:

$$(3.6) \quad \forall \nu, |\nu| = i, i \in \{1, 2\}, |D^\nu f_{x,y}(\Gamma_{n,\delta}(s, t, y))| \leq \frac{\bar{c}_1}{(t-s)^i} \exp(\bar{c}_1 K_1^2(t-s)) p_{\bar{c}}(t-s, y-x),$$

where in the above equation, for  $|\nu| = i$ ,  $i \in \{1, 2\}$ , we consider the *worst* singularity which is in small time associated with the multi-indexes  $\nu$  s.t.  $|\nu| = |\nu_{2:(d+1)}|$ . If  $t-s \geq 1$ , the exponential factor in (3.6) would absorb the remaining contributions in time.

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

$$|\tilde{p}(s, t, x, y) - \tilde{p}_n(s, t, x, y)| \leq \bar{c}_1 \Delta_n \exp(\bar{c}_1 K_1^2(t-s)) p_{\bar{c}}(s, t, x, y).$$

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.  $\square$

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 1 and 2 below will yield by induction the result stated in Theorem 1.

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

$$(3.7) \quad \begin{aligned} & |\tilde{p} \otimes H^{(1)}(s, t, x, y) - \tilde{p}_n \otimes H_n^{(1)}(s, t, x, y)| \\ & \leq 2c_1^2 \exp(c_1 K_1^2(t-s)) \Delta_n p_c(s, t, x, y) B(1, \frac{\gamma}{2})(t-s)^{\frac{\gamma}{2}}, \end{aligned}$$

where  $c_1, c$  are as in Lemma 1.

*Proof.* Let us write:

$$(3.8) \quad \begin{aligned} & |\tilde{p} \otimes H^{(1)}(s, t, x, y) - \tilde{p}_n \otimes H_n^{(1)}(s, t, x, y)| \leq \\ & |(\tilde{p} - \tilde{p}_n) \otimes H(s, t, x, y)| + |\tilde{p}_n(s, u, x, z) \otimes (H - H_n)(s, t, x, y)| := I + II. \end{aligned}$$

From Lemma 1 and (2.9) we readily get:

$$(3.9) \quad |(\tilde{p} - \tilde{p}_n) \otimes H(s, t, x, y)| \leq c_1^2 \exp(c_1 K_1^2(t-s)) \Delta_n p_c(t-s, y-x) B(1, \frac{\gamma}{2})(t-s)^{\frac{\gamma}{2}}.$$

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

$$(3.10) \quad (H - H_n)(u, t, z, y) \leq \Delta_n \frac{c_1 \exp(c_1 K_1^2(t-u))}{(t-u)^{1-\frac{\gamma}{2}}} p_c(t-u, y-z).$$

Equations (3.10) and (2.8) give that  $II$  can be handled as  $I$  which yields the result. It therefore remains to prove (3.10). Let us write with the notations of (3.3):

$$\begin{aligned} (H - H_n)(u, t, z, y) & := \left[ \frac{1}{2} \text{Tr} \left( (a(u, z) - a(u, y)) D_z^2 f_{z,y}(\Theta(u, t, y)) \right) \right. \\ & \quad \left. + \langle b(u, z) - b(u, y), D_z f_{z,y}(\Theta(u, t, y)) \rangle \right] \\ & \quad - \left[ \frac{1}{2} \text{Tr} \left( (a_n(u, z) - a_n(u, y)) D_z^2 f_{z,y}(\Theta_n(u, t, y)) \right) \right. \\ & \quad \left. + \langle b_n(u, z) - b_n(u, y), D_z f_{z,y}(\Theta_n(u, t, y)) \rangle \right]. \end{aligned}$$

Thus,

$$\begin{aligned} & (H - H_n)(u, t, z, y) \\ & = \frac{1}{2} \left[ \text{Tr} \left( (a(u, z) - a(u, y)) \{ D_z^2 f_{z,y}(\Theta(u, t, y)) - D_z^2 f_{z,y}(\Theta_n(u, t, y)) \} \right) \right. \\ & \quad \left. - \text{Tr} \left( [(a_n(u, z) - a_n(u, y)) - (a(u, z) - a(u, y))] D_z^2 f_{z,y}(\Theta_n(u, t, y)) \right) \right] \\ & \quad + \left[ \langle b(u, z) - b(u, y), \{ D_z f_{z,y}(\Theta(u, t, y)) - D_z f_{z,y}(\Theta_n(u, t, y)) \} \rangle \right. \\ & \quad \left. - \langle [(b_n(u, z) - b_n(u, y)) - (b(u, z) - b(u, y))], D_z f_{z,y}(\Theta_n(u, t, y)) \rangle \right]. \end{aligned}$$

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

$$\begin{aligned} |D_z^i f_{z,y}(\Theta(u, t, y))| + |D_z^i f_{z,y}(\Theta_n(u, t, y))| &\leq \frac{\tilde{c}_1 \exp(\tilde{c}_1 K_1^2(t-u))}{(t-u)^{i/2}} p_{\tilde{c}}(t-u, y-z), \\ |D_z^i f_{z,y}(\Theta(u, t, y)) - D_z^i f_{z,y}(\Theta_n(u, t, y))| &\leq \frac{\tilde{c}_1 \Delta_n \exp(\tilde{c}_1 K_1^2(t-s))}{(t-u)^{i/2}} p_{\tilde{c}}(t-u, y-z). \end{aligned}$$

Thus, provided that  $c_1, c$  have been chosen large and small enough respectively in Lemma 1, the Hölder continuity assumption in **(A4)** gives:

$$|(H - H_n)(u, t, z, y)| \leq \frac{c_1 \Delta_n \exp(c_1 K_1^2(t-u))}{(t-u)^{1-\gamma/2}} p_c(t-u, y-z).$$

This establishes (3.10).  $\square$

The following Lemma associated with Lemmas 1 and 2 allows to complete the proof of Theorem 1.

*Lemma 3* (Difference of the iterated kernels). For all  $0 \leq s < t \leq T$ ,  $(x, y) \in (\mathbb{R}^d)^2$  and for all  $r \in \mathbb{N}$ :

$$\begin{aligned} &|(\tilde{p} \otimes H^{(r)} - \tilde{p}_n \otimes H_n^{(r)})(s, t, x, y)| \\ (3.11) \quad &\leq (r+1) \exp(c_1 K_1^2(t-s)) \Delta_n \frac{c_1^{r+1} [\Gamma(\frac{\gamma}{2})]^r}{\Gamma(1+r\frac{\gamma}{2})} p_c(t-s, y-x)(t-s)^{\frac{r\gamma}{2}}. \end{aligned}$$

where  $c, c_1$  are as in Lemma 1.

*Proof.* Observe that Lemmas 1 and 2 respectively give (3.11) 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  $r \geq 1$ ,  $\eta_r(s, t, x, y) := |(\tilde{p} \otimes H^{(r)} - \tilde{p}_n \otimes H_n^{(r)})(s, t, x, y)|$ . Write

$$\begin{aligned} \eta_{r+1}(s, t, x, y) &\leq |[\tilde{p} \otimes H^{(r)} - \tilde{p}_n \otimes H_n^{(r)}] \otimes H(s, t, x, y)| \\ &\quad + |\tilde{p}_n \otimes H_n^{(r)} \otimes (H - H_n)(s, t, x, y)| \\ &\leq \eta_r \otimes |H|(s, t, x, y) + |\tilde{p}_n \otimes H_n^{(r)}| \otimes |(H - H_n)|(s, t, x, y). \end{aligned}$$

Recall now that under **(A)**, the terms  $|H|(s, t, x, y)$  and  $|\tilde{p}_n \otimes H_n^{(r)}|$  satisfy respectively and uniformly in  $n$  the controls of equations (2.9), (2.10). Thus thanks to (3.10) and the induction hypothesis we get the result.  $\square$

Theorem 1 now simply follows from the controls of Lemma 3, the parametrix expansions (2.6) and (2.7) of the densities  $p, p_n$  and the asymptotics of the  $\Gamma$  function.

**3.2. Stability for Markov Chains.** In this Section we prove Theorem 2. 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.

Recall first from Section 2.2 that we have the following representations for the density  $p^h, p_n^h$  of the Markov chains  $Y, Y^{(n)}$  in (1.3). For all  $0 \leq t_i < t_j \leq T$ ,  $(x, y) \in$

$(\mathbb{R}^d)^2$ :

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

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

3.2.1. *Comparison of the frozen densities.* The first key point for the analysis with Markov chains is the following Lemma.

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

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

$$|D_x^\alpha \tilde{p}^h(t_i, t_j, x, y) - D_x^\alpha \tilde{p}_n^h(t_i, t_j, x, y)| \leq \frac{\Delta_n}{(t_j - t_i)^{|\alpha|/2}} \psi_{c, c_1}(t_j - t_i, y - x),$$

where

- Under  $(\mathbf{I}_G)$ :

$$\psi_{c, c_1}(t_j - t_i, y - x) := c_1 \exp(c_1 K_1^2(t_j - t_i)) p_c(t_j - t_i, y - x),$$

- Under  $(\mathbf{I}_{P, M})$ :

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

*Proof.* Note first that under  $(\mathbf{I}_G)$  the statement has already been proved in Lemma 1. We thus assume that  $(\mathbf{I}_{P, M})$  holds. Introduce first the random vectors with zero mean:

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

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

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

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

$$D_x^\alpha \tilde{p}_n^h(t_k, t_j, x, y) = \frac{1}{(t_j - t_k)^{(d+|\alpha|)/2}} (-1)^{|\alpha|} D_z^\alpha q_{j-k,n}(z) \Big|_{z = \frac{y-x-h \sum_{l=k}^{j-1} b_n(t_l, y)}{(t_j - t_k)^{1/2}}}.$$

(3.12)

From the Edgeworth expansion of Theorem 19.3 in Bhattacharya and Rao [BR76], for  $q_{j-k}, q_{j-k,n}$ , one readily derives under  $(\mathbf{A})$ , for  $|\alpha| = 0$  that there exists  $c_1$  s.t. for all  $0 \leq t_k < t_j \leq T$ ,  $(x, y) \in (\mathbb{R}^d)^2$ ,

$$(3.13) \quad \tilde{p}^h(t_k, t_j, x, y) + \tilde{p}_n^h(t_k, t_j, x, y) \leq \frac{c_1}{(t_j - t_k)^{d/2}} \frac{1}{\left(1 + \frac{|x-y|}{(t_j - t_k)^{1/2}}\right)^m},$$

for all integer  $m < M - d$ , where we recall that  $M$  stands for the initial decay of the density  $f_\xi$  of the innovations bounded by  $Q_M$  (see equation (1.5)).

We can as well derive similarly to the proof of Theorem 19.3 in [BR76], see also Lemma 3.7 in [KM00], that for all  $\alpha, |\alpha| \leq 4$ :

$$(3.14) \quad |D_x^\alpha \tilde{p}^h(t_k, t_j, x, y)| + |D_x^\alpha \tilde{p}_n^h(t_k, t_j, x, y)| \leq \frac{c_1}{(t_j - t_k)^{(d+|\alpha|)/2}} \frac{1}{\left(1 + \frac{|x-y|}{(t_j - t_k)^{1/2}}\right)^m},$$

for all  $m < M - d - 4$ . Note indeed that differentiating in  $D_x^\alpha$  the density and the terms of the Edgeworth expansion corresponds to a multiplication of the Fourier transforms involved by  $\zeta^\alpha$ ,  $\zeta$  standing for the Fourier variable. Hence, from our smoothness assumptions in  $(\mathbf{I}_{P,M})$ , after obvious modifications, the estimates of Theorem 9.11 and Lemma 14.3 from [BR76] 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 (3.13) up to the derivation order. On the other hand, the bound in (3.13) is itself deteriorated w.r.t. the initial concentration condition in (1.5). The key point is that the techniques of Theorem 19.3 in [BR76] actually provide concentration bounds for inhomogeneous sums of random variables with concentration as in (1.5) in terms of the moments of the innovations. To explain the bound in (3.13) let us observe that the  $m^{\text{th}}$  moment of  $\xi$  is finite for  $m < M - d$ .

Equations (3.13) and (3.14) give the first part of the lemma. Still from the proof of Theorem 19.3 in [BR76], one gets, under  $(\mathbf{A})$ , that there exists  $C > 0$  s.t. for all multi-indexes  $\bar{\alpha}$ ,  $|\bar{\alpha}| \leq 4$ ,  $\bar{\beta}$ ,  $|\bar{\beta}| \leq m \leq M - d - 5$  for all  $j > k$ :

$$(3.15) \quad \int_{\mathbb{R}^d} |\zeta^{\bar{\alpha}}| \left\{ |D_\zeta^{\bar{\beta}} \hat{q}_{j-k}(\zeta)| + |D_\zeta^{\bar{\beta}} \hat{q}_{j-k,n}(\zeta)| \right\} d\zeta \leq C,$$

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

To investigate the quantity  $|D_x^\alpha \tilde{p}^h(t_k, t_j, x, y) - D_x^\alpha \tilde{p}_n^h(t_k, t_j, x, y)|$  thanks to (3.12) define now for all  $\alpha$ ,  $|\alpha| \leq 4$ ,  $\beta$ ,  $|\beta| \leq m \leq M - d - 5$ :

$$(3.16) \quad \begin{aligned} \forall z \in \mathbb{R}^d, \quad \Theta_{j-k,n}(z) &:= z^\beta D_z^\alpha (q_{j-k}(z) - q_{j-k,n}(z)), \\ \forall \zeta \in \mathbb{R}^d, \quad \hat{\Theta}_{j-k,n}(\zeta) &:= (-i)^{|\alpha|+|\beta|} D_\zeta^\beta (\zeta^\alpha \{\hat{q}_{j-k}(\zeta) - \hat{q}_{j-k,n}(\zeta)\}). \end{aligned}$$

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

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

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

$$\begin{aligned}
& (i^{\bar{\beta}})^{-1} \zeta^{\bar{\alpha}} (D_{\zeta}^{\bar{\beta}} \hat{q}_{j-k}(\zeta) - D_{\zeta}^{\bar{\beta}} \hat{q}_{j-k,n}(\zeta)) = \\
& \zeta^{\bar{\alpha}} \mathbb{E} \left[ (\tilde{Z}_{k,j,1}^y + \tilde{Z}_{k,j,2}^y)^{\bar{\beta}} \exp[i\zeta \cdot \tilde{Z}_{k,j,1}^y] \exp[i\zeta \cdot \tilde{Z}_{k,j,2}^y] \right] - \\
& \mathbb{E} \left[ (\tilde{Z}_{k,j,1}^{y,(n)} + \tilde{Z}_{k,j,2}^{y,(n)})^{\bar{\beta}} \exp[i\zeta \cdot \tilde{Z}_{k,j,1}^{y,(n)}] \exp[i\zeta \cdot \tilde{Z}_{k,j,2}^{y,(n)}] \right] = \\
& \zeta^{\bar{\alpha}} \left\{ \sum_{l, |l| \leq |\bar{\beta}|} C_{\bar{\beta}}^l \mathbb{E} \left[ (\tilde{Z}_{k,j,1}^y)^l \exp[i\zeta \cdot \tilde{Z}_{k,j,1}^y] \right] \mathbb{E} \left[ (\tilde{Z}_{k,j,2}^y)^{\bar{\beta}-l} \exp[i\zeta \cdot \tilde{Z}_{k,j,2}^y] \right] - \right. \\
& \left. \sum_{l, |l| \leq |\bar{\beta}|} C_{\bar{\beta}}^l \mathbb{E} \left[ (\tilde{Z}_{k,j,1}^{y,(n)})^l \exp[i\zeta \cdot \tilde{Z}_{k,j,1}^{y,(n)}] \right] \mathbb{E} \left[ (\tilde{Z}_{k,j,2}^{y,(n)})^{\bar{\beta}-l} \exp[i\zeta \cdot \tilde{Z}_{k,j,2}^{y,(n)}] \right] \right\} = \\
& \zeta^{\bar{\alpha}} \left\{ \sum_{l, |l| \leq |\bar{\beta}|} C_{\bar{\beta}}^l \left\{ \mathbb{E} \left[ (\tilde{Z}_{k,j,1}^y)^l \exp[i\zeta \cdot \tilde{Z}_{k,j,1}^y] \right] - \mathbb{E} \left[ (\tilde{Z}_{k,j,1}^{y,(n)})^l \exp[i\zeta \cdot \tilde{Z}_{k,j,1}^{y,(n)}] \right] \right\} \right. \\
& \times \mathbb{E} \left[ (\tilde{Z}_{k,j,2}^y)^{\bar{\beta}-l} \exp[i\zeta \cdot \tilde{Z}_{k,j,2}^y] \right] + \mathbb{E} \left[ (\tilde{Z}_{k,j,1}^{y,(n)})^l \exp[i\zeta \cdot \tilde{Z}_{k,j,1}^{y,(n)}] \right] \\
& \left. \times \left[ \mathbb{E} \left[ (\tilde{Z}_{k,j,2}^y)^{\bar{\beta}-l} \exp[i\zeta \cdot \tilde{Z}_{k,j,2}^y] \right] - \mathbb{E} \left[ (\tilde{Z}_{k,j,2}^{y,(n)})^{\bar{\beta}-l} \exp[i\zeta \cdot \tilde{Z}_{k,j,2}^{y,(n)}] \right] \right] \right\},
\end{aligned}$$

where in the above expression we considered the binomial expansion for multi-indexes denoting by  $C_{\bar{\beta}}^l := \frac{\bar{\beta}!}{(\bar{\beta}-l)!l!}$  with the corresponding definitions for factorials (see the proof of Lemma 1). Introduce now, for a multi-index  $l, |l| \in \llbracket 0, |\bar{\beta}| \rrbracket$ , the functions:

$$\begin{aligned}
\Psi_1^{\bar{\alpha}, \bar{\beta}-l}(\zeta) & := \zeta^{\bar{\alpha}} \mathbb{E} \left[ (\tilde{Z}_{k,j,2}^y)^{\bar{\beta}-l} \exp[i\zeta \cdot \tilde{Z}_{k,j,2}^y] \right], \\
\Psi_2^{\bar{\alpha}, l}(\zeta) & := \zeta^{\bar{\alpha}} \mathbb{E} \left[ (\tilde{Z}_{k,j,1}^{y,(n)})^l \exp[i\zeta \cdot \tilde{Z}_{k,j,1}^{y,(n)}] \right],
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{E}_{1,l}(\zeta) & := \left[ \mathbb{E} \left[ (\tilde{Z}_{k,j,1}^y)^l \exp[i\zeta \cdot \tilde{Z}_{k,j,1}^y] \right] - \mathbb{E} \left[ (\tilde{Z}_{k,j,1}^{y,(n)})^l \exp[i\zeta \cdot \tilde{Z}_{k,j,1}^{y,(n)}] \right] \right], \\
\mathcal{E}_{2, \bar{\beta}-l}(\zeta) & := \left[ \mathbb{E} \left[ (\tilde{Z}_{k,j,2}^y)^{\bar{\beta}-l} \exp[i\zeta \cdot \tilde{Z}_{k,j,2}^y] \right] - \mathbb{E} \left[ (\tilde{Z}_{k,j,2}^{y,(n)})^{\bar{\beta}-l} \exp[i\zeta \cdot \tilde{Z}_{k,j,2}^{y,(n)}] \right] \right].
\end{aligned}$$

Thus, we can rewrite from the previous computations:

$$(3.17) \quad (i^{\bar{\beta}})^{-1} \zeta^{\bar{\alpha}} (D_{\zeta}^{\bar{\beta}} \hat{q}_{j-k}(\zeta) - D_{\zeta}^{\bar{\beta}} \hat{q}_{j-k,n}(\zeta)) = \sum_{l, |l| \leq |\bar{\beta}|} C_{\bar{\beta}}^l \left\{ (\mathcal{E}_{1,l} \Psi_1^{\bar{\alpha}, \bar{\beta}-l})(\zeta) + (\mathcal{E}_{2, \bar{\beta}-l} \Psi_2^{\bar{\alpha}, l})(\zeta) \right\}.$$

Recall from (3.15) that we already have integrability for the contributions  $\Psi_1^{\bar{\alpha}, \bar{\beta}-l}(\zeta)$  and  $\Psi_2^{\bar{\alpha}, l}(\zeta)$ . Let us thus start 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. Write:

$$\begin{aligned}
& |\mathcal{E}_{1,l}(\zeta)| \leq \\
& \mathbb{E} \left[ |(\tilde{Z}_{k,j,1}^y)^l - (\tilde{Z}_{k,j,1}^{y,n})^l| \right] + \mathbb{E} \left[ |(\tilde{Z}_{k,j,1}^{y,n})^l| \left| \exp(i\zeta \cdot \tilde{Z}_{k,j,1}^y) - \exp(i\zeta \cdot \tilde{Z}_{k,j,1}^{y,n}) \right| \right] \\
& \leq C \left\{ \mathbb{E} \left[ |\tilde{Z}_{k,j,1}^y - \tilde{Z}_{k,j,1}^{y,n}| \left( |\tilde{Z}_{k,j,1}^y|^{l-1} + |\tilde{Z}_{k,j,1}^{y,n}|^{l-1} \right) \right] + \mathbb{E} \left[ |\tilde{Z}_{k,j,1}^{y,n}|^l |\zeta| |\tilde{Z}_{k,j,1}^y - \tilde{Z}_{k,j,1}^{y,n}| \right] \right\}.
\end{aligned}$$

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

$$(3.18) \quad \begin{aligned} & |\mathcal{E}_{1,l}(\zeta)| \leq \\ & C \left\{ \mathbb{E}[|\tilde{Z}_{k,j,1}^y - \tilde{Z}_{k,j,1}^{y,n}|^{|l|}]^{1/|l|} \left\{ \mathbb{E}[|\tilde{Z}_{k,j,1}^y|^{|l|}]^{(|l|-1)/|l|} + \mathbb{E}[|\tilde{Z}_{k,j,1}^{y,n}|^{|l|}]^{(|l|-1)/|l|} \right\} + \right. \\ & \left. |\zeta| \mathbb{E}[|\tilde{Z}_{k,j,1}^{y,n}|^{|l|+1}]^{1/(|l|+1)} \mathbb{E}[|\tilde{Z}_{k,j,1}^y - \tilde{Z}_{k,j,1}^{y,n}|^{|l|+1}]^{1/(|l|+1)} \right\}. \end{aligned}$$

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  $r \leq m + 1$ ,

$$(3.19) \quad \mathbb{E}[|\tilde{Z}_{k,j,1}^y - \tilde{Z}_{k,j,1}^{y,n}|^r]^{1/r} \leq c\Delta_n, \quad \mathbb{E}[|\tilde{Z}_{k,j,1}^y|^r]^{1/r} + \mathbb{E}[|\tilde{Z}_{k,j,1}^{y,n}|^r]^{1/r} \leq c.$$

Let us establish the point for the difference, the other bounds can be derived similarly. Define for all  $i \in \llbracket k, j \rrbracket$ ,  $\tilde{M}_i := \sqrt{h} \sum_{r=k}^{i-1} (\sigma - \sigma_n)(t_r, y) \xi_{r+1}$ . The process  $(\tilde{M}_i)_{i \in \llbracket k, j \rrbracket}$  is a square integrable martingale (in discrete time, w.r.t.  $\mathcal{F}_i := \sigma(\xi_r, r \leq i)$ ). Its quadratic variation writes  $[\tilde{M}]_i = h \sum_{r=k}^{i-1} |(\sigma - \sigma_n)(t_r, y)|^2 |\xi_{r+1}|^2$  and the Burkholder-Davies-Gundy inequalities, see e.g. Shiryaev [Shi96], give for all  $r \leq M - d - 4$ :

$$(3.20) \quad \mathbb{E}[\sup_{i \in \llbracket k, j \rrbracket} |\tilde{M}_i|^r] \leq c_r \mathbb{E}[[\tilde{M}]_j^{r/2}] = c_r h^{r/2} \mathbb{E}[\left(\sum_{i=k}^{j-1} |(\sigma - \sigma_n)(t_i, y)|^2 |\xi_{i+1}|^2\right)^{r/2}].$$

If  $r = 2$  one readily gets:

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

Let us thus assume  $r > 2$  and derive from (3.20)

$$\begin{aligned} \mathbb{E}[|\tilde{Z}_{k,j,1}^y - \tilde{Z}_{k,j,1}^{y,n}|^r] & \leq \frac{c_r}{(t_j - t_k)^{r/2}} \mathbb{E}[\sup_{i \in \llbracket k, j \rrbracket} |\tilde{M}_i|^r] \\ & \leq \frac{c_r h^{r/2}}{(t_j - t_k)^{r/2}} \mathbb{E}\left[\left(\sum_{i=k}^{j-1} |(\sigma - \sigma_n)(t_i, y)|^r |\xi_{i+1}|^r\right) \left(\sum_{i=k}^{j-1} 1\right)^{r/2(1-2/r)}\right], \end{aligned}$$

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

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

Since we have assumed  $r \leq m + 1 \leq M - d - 4$ , this gives the first control in (3.19). The other one readily follows replacing  $\sigma - \sigma_n$  by  $\sigma$  or  $\sigma_n$ .

From equations (3.18), (3.19) and similar controls for  $\mathcal{E}_{2,\bar{\beta}-l}(\zeta)$  we finally derive:

$$|\mathcal{E}_{1,l}(\zeta)| + |\mathcal{E}_{1,\bar{\beta}-l}(\zeta)| \leq C_1 \Delta_n (1 + |\zeta|).$$

As a result we have from (3.16) and (3.17):

$$\begin{aligned} & |D_\zeta^\beta(\zeta^\alpha(\hat{q}_{j-k}(\zeta) - D_\zeta^\beta \hat{q}_{j-k,n}(\zeta)))| \\ & \leq C\Delta_n \left\{ \sum_{\substack{\bar{\beta}, |\bar{\beta}| \leq |\beta| \\ \bar{\alpha} = \alpha - (\beta - \bar{\beta})}} \sum_{l, |l| \leq |\bar{\beta}|} (|\Psi_1^{\bar{\alpha}, \bar{\beta}-l}(\zeta)| + |\Psi_2^{\bar{\alpha}, l}(\zeta)|)(1 + |\zeta|) \right\}. \end{aligned}$$

We finally derive from (3.16) and (3.15) (which thanks to the smoothness assumption on  $Q_M$  in  $(\mathbf{I}_{P,M})$  holds as well for a multi-index  $\bar{\alpha}$ ,  $|\bar{\alpha}| = 5$ ):

$$(3.21) \quad |\Theta_{j-k,n}(z)| \leq \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} |\hat{\Theta}_{j-k,n}(\zeta)| d\zeta \leq c\Delta_n.$$

From (3.12) this concludes the proof for  $j > k + 1$  and  $b = b_n = 0$ . To handle the drift, write now setting  $z := \frac{y-x}{(t_j-t_k)^{1/2}}$ ,  $b_{j-k} := \frac{h \sum_{i=k}^{j-1} b(t_i, y)}{(t_j-t_k)^{1/2}}$ ,  $b_{j-k,n} := \frac{h \sum_{i=k}^{j-1} b_n(t_i, y)}{(t_j-t_k)^{1/2}}$ :

$$\begin{aligned} & D_z^\alpha q_{j-k}(z-b) - D_z^\alpha q_{j-k,n}(z-b_n) = \{D_z^\alpha q_{j-k} - D_z^\alpha q_{j-k,n}\}(z-b) \\ & + \{D_z^\alpha q_{j-k,n}(z-b) - D_z^\alpha q_{j-k,n}(z-b_n)\} =: Q_{j-k,n}^1(z, b) + Q_{j-k,n}^2(z, b, b_n). \end{aligned}$$

From (3.16) and (3.21) we get  $|Q_{j-k,n}^1(z, b)| \leq c \frac{\Delta_n}{(1+|z|)^m}$ . Now,

$$\begin{aligned} |Q_{j-k,n}^2(z, b, b_n)| & \leq \sup_{\theta \in [0,1]} |DD_z^\alpha q_{j-k,n}(z - (\theta b + (1-\theta)b_n))| |b - b_n| \\ & \leq \frac{c\Delta_n(t_j - t_k)^{1/2}}{(1+|z|)^m}, \end{aligned}$$

exploiting again (3.15) for a multi-index  $\bar{\alpha}$ ,  $|\bar{\alpha}| = 5$ . This concludes the proof if  $j > k + 1$ . 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 actually have the following Lemma.

*Lemma 5 (Difference of the Discrete Kernels.).* There exists  $c_1, c$  s.t. for  $0 \leq t_k < t_j \leq T$ ,  $(z, y) \in (\mathbb{R}^d)^2$ :

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

with

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

where  $\psi_{c,c_1}$  is defined according to the assumptions on the innovations in Lemma 4.

*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  $k \in \llbracket 0, N \rrbracket$ ,  $(x, w) \in (\mathbb{R}^d)^2$  the one step transitions:

$$(3.22) \quad T^h(t_k, x, w) := b(t_k, x)h + h^{1/2}\sigma(t_k, x)w, \quad T_n^h(t_k, x, w) := b_n(t_k, x)h + h^{1/2}\sigma_n(t_k, x)w.$$

From the definition of  $H^h, H_n^h$ , recalling that  $f_\xi$  stands for the density of the innovation, the difference of the kernels writes:

$$\begin{aligned}
& (H^h - H_n^h)(t_k, t_j, z, y) \\
= & h^{-1} \int_{\mathbb{R}^d} dw f_\xi(w) \left[ \left\{ \tilde{p}^h(t_{k+1}, t_j, z + T^h(t_k, z, w), y) - \tilde{p}^h(t_{k+1}, t_j, z + T^h(t_k, y, w), y) \right\} \right. \\
& \left. - \left\{ \tilde{p}_n^h(t_{k+1}, t_j, z + T_n^h(t_k, z, w), y) - \tilde{p}_n^h(t_{k+1}, t_j, z + T_n^h(t_k, y, w), y) \right\} \right].
\end{aligned} \tag{3.23}$$

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:

$$\begin{aligned}
\varphi_\lambda^h : \mathbb{R}^d \times \mathbb{R}^d & \longrightarrow \mathbb{R} \\
(T_1, T_2) & \longmapsto \text{Tr} \left( D_z^2 \tilde{p}^h(t_{k+1}, t_j, z + \lambda T_1, y) [T_2 T_2^*] \right), \\
\varphi_{\lambda, n}^h : \mathbb{R}^d \times \mathbb{R}^d & \longrightarrow \mathbb{R} \\
(T_1, T_2) & \longmapsto \text{Tr} \left( D_z^2 \tilde{p}_n^h(t_{k+1}, t_j, z + \lambda T_1, y) [T_2 T_2^*] \right).
\end{aligned} \tag{3.24}$$

Recalling as well that  $\xi$  is centered we get:

$$\begin{aligned}
\Delta H^{h, n}(t_k, t_j, z, y) & := (H^h - H_n^h)(t_k, t_j, z, y) \\
& = \left[ \left\{ D_z \tilde{p}^h(t_{k+1}, t_j, z, y) \cdot (b(t_k, z) - b(t_k, y)) \right\} \right. \\
& \quad \left. - \left\{ D_z \tilde{p}_n^h(t_{k+1}, t_j, z, y) \cdot (b_n(t_k, z) - b_n(t_k, y)) \right\} \right] \\
& \quad + h^{-1} \int_{\mathbb{R}^d} dw f_\xi(w) \int_0^1 d\lambda (1 - \lambda) \\
& \quad \times \left[ \left\{ \varphi_\lambda^h(T^h(t_k, z, w), T^h(t_k, z, w)) - \varphi_\lambda^h(T^h(t_k, y, w), T^h(t_k, y, w)) \right\} \right. \\
& \quad \left. - \left\{ \varphi_{\lambda, n}^h(T_n^h(t_k, z, w), T_n^h(t_k, z, w)) - \varphi_{\lambda, n}^h(T_n^h(t_k, y, w), T_n^h(t_k, y, w)) \right\} \right] \\
& =: (\Delta_1 H^{h, n} + \Delta_2 H^{h, n})(t_k, t_j, z, y),
\end{aligned} \tag{3.25}$$

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

$$\begin{aligned}
& \Delta_1 H^{h, n}(t_k, t_j, z, y) \\
= & \left[ \left\{ D_z \tilde{p}^h(t_{k+1}, t_j, z, y) \cdot [(b(t_k, z) - b(t_k, y)) - (b_n(t_k, z) - b_n(t_k, y))] \right\} \right. \\
& \quad \left. - \left\{ (D_z \tilde{p}_n^h - D_z \tilde{p}^h)(t_{k+1}, t_j, z, y) \cdot (b_n(t_k, z) - b_n(t_k, y)) \right\} \right].
\end{aligned}$$

From Assumptions **(A3)**, **(A4)** and Lemma 4 we derive:

$$|\Delta_1 H^{h, n}(t_k, t_j, z, y)| \leq \frac{\Delta_n \{1 + \kappa |z - y|^\gamma\}}{(t_j - t_k)^{1/2}} \psi_{c, c_1}(t_j - t_k, y - z). \tag{3.26}$$

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

$$\begin{aligned} \Delta\varphi_\lambda^{h,n}(t_k, z, y, w) &:= \\ &\left\{ \varphi_\lambda^h(T^h(t_k, z, w), T^h(t_k, z, w)) - \varphi_\lambda^h(T^h(t_k, y, w), T^h(t_k, y, w)) \right\} \\ &- \left\{ \varphi_{\lambda,n}^h(T_n^h(t_k, z, w), T_n^h(t_k, z, w)) - \varphi_{\lambda,n}^h(T_n^h(t_k, y, w), T_n^h(t_k, y, w)) \right\}. \end{aligned}$$

Let us then decompose:

$$\begin{aligned} &\Delta\varphi_\lambda^{h,n}(t_k, z, y, w) \\ &:= \left[ \left\{ \varphi_\lambda^h(T^h(t_k, z, w), T^h(t_k, z, w)) - \varphi_\lambda^h(T^h(t_k, z, w), T^h(t_k, y, w)) \right\} \right. \\ &\quad \left. - \left\{ \varphi_{\lambda,n}^h(T_n^h(t_k, z, w), T_n^h(t_k, z, w)) - \varphi_{\lambda,n}^h(T_n^h(t_k, z, w), T_n^h(t_k, y, w)) \right\} \right] \\ &\quad + \left[ \left\{ \varphi_\lambda^h(T^h(t_k, z, w), T^h(t_k, y, w)) - \varphi_\lambda^h(T^h(t_k, y, w), T^h(t_k, y, w)) \right\} \right. \\ &\quad \left. + \left\{ \varphi_{\lambda,n}^h(T_n^h(t_k, y, w), T_n^h(t_k, y, w)) - \varphi_{\lambda,n}^h(T_n^h(t_k, z, w), T_n^h(t_k, y, w)) \right\} \right] \\ (3.27) \qquad \qquad \qquad &:= (\Delta_1\varphi_\lambda^{h,n} + \Delta_2\varphi_\lambda^{h,n})(t_k, z, y, w), \end{aligned}$$

and write from (3.25):

$$\begin{aligned} \Delta_2 H^{h,n}(t_k, t_j, z, y) &= h^{-1} \int_{\mathbb{R}^d} dw f_\xi(w) \int_0^1 d\lambda (1-\lambda) (\Delta_1\varphi_\lambda^{h,n} + \Delta_2\varphi_\lambda^{h,n})(t_k, z, y, w) \\ (3.28) \qquad \qquad \qquad &=: (\Delta_{21}H^{h,n} + \Delta_{22}H^{h,n})(t_k, t_j, z, y), \end{aligned}$$

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

**Term**  $\Delta_{21}H^{h,n}$ . We first write from (3.27):

$$\begin{aligned}
& \Delta_1 \varphi_\lambda^{h,n}(t_k, z, y, w) \\
= & \left[ \left\{ \varphi_\lambda^h(T^h(t_k, z, w), T^h(t_k, z, w)) - \varphi_\lambda^h(T^h(t_k, z, w), T^h(t_k, y, w)) \right\} - \right. \\
& \left. \left\{ \varphi_\lambda^h(T^h(t_k, z, w), T_n^h(t_k, z, w)) - \varphi_\lambda^h(T^h(t_k, z, w), T_n^h(t_k, y, w)) \right\} \right] \\
& + \left[ \left\{ \varphi_\lambda^h(T^h(t_k, z, w), T_n^h(t_k, z, w)) - \varphi_\lambda^h(T^h(t_k, z, w), T_n^h(t_k, y, w)) \right\} \right. \\
& \left. - \left\{ \varphi_\lambda^h(T_n^h(t_k, z, w), T_n^h(t_k, z, w)) - \varphi_\lambda^h(T_n^h(t_k, z, w), T_n^h(t_k, y, w)) \right\} \right] \\
& - \left[ \left\{ \varphi_{\lambda,n}^h(T_n^h(t_k, z, w), T_n^h(t_k, z, w)) - \varphi_{\lambda,n}^h(T_n^h(t_k, z, w), T_n^h(t_k, y, w)) \right\} \right. \\
& \left. - \left\{ \varphi_\lambda^h(T_n^h(t_k, z, w), T_n^h(t_k, z, w)) - \varphi_\lambda^h(T_n^h(t_k, z, w), T_n^h(t_k, y, w)) \right\} \right] \\
(3.29) \quad & =: \sum_{i=1}^3 \Delta_{1i} \varphi_\lambda^{h,n}(t_k, z, y, w).
\end{aligned}$$

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

$$\begin{aligned}
D(t_k, z, y, w) & := T^h(t_k, z, w)T^h(t_k, z, w)^* - T^h(t_k, y, w)T^h(t_k, y, w)^*, \\
D_n(t_k, z, y, w) & := T_n^h(t_k, z, w)T_n^h(t_k, z, w)^* - T_n^h(t_k, y, w)T_n^h(t_k, y, w)^*,
\end{aligned}$$

we have from **(A3)**, **(A4)**:

$$\begin{aligned}
(|D| + |D_n|)(t_k, z, y, w) & \leq \bar{c}(h^2 + h^{3/2}|w| + h(1 \wedge |z - y|)^\gamma |w|^2), \\
(3.30) \quad |D - D_n|(t_k, z, y, w) & \leq \bar{c}\Delta_n(h^2 + h^{3/2}|w| + h(1 \wedge |z - y|)^\gamma |w|^2).
\end{aligned}$$

From the definition of  $\varphi_\lambda^h$  in (3.24), equation (3.27) and the control (3.30), we get:

$$\begin{aligned}
(3.31) \quad & |\Delta_{11} \varphi_\lambda^{h,n}|(t_k, z, y, w) \\
& \leq \bar{c}\Delta_n \frac{\psi_{c,c_1}(t_j - t_k, y - (z + \lambda T^h(t_k, z, w)))}{(t_j - t_k)} (h^2 + h^{3/2}|w| + h(1 \wedge |z - y|)^\gamma |w|^2).
\end{aligned}$$

We would similarly get from Lemma 4 and (3.30):

$$\begin{aligned}
& |\Delta_{13}\varphi_\lambda^{h,n}|(t_k, z, y, w) \\
\leq & \bar{c}\Delta_n \frac{\psi_{c,c_1}(t_j - t_k, y - (z + \lambda T_n^h(t_k, z, w)))}{(t_j - t_k)} (h^2 + h^{3/2}|w| + h(1 \wedge |z - y|)^\gamma |w|^2), \\
& |\Delta_{12}\varphi_\lambda^{h,n}|(t_k, z, y, w) \\
\leq & \frac{\psi_{c,c_1}(t_j - t_k, y - (z + \theta\lambda T^h(t_k, z, w) + (1 - \theta)\lambda T_n^h(t_k, z, w)))}{(t_j - t_k)^{3/2}} \\
& \times |(T^h - T_n^h)(t_k, z, w)| |D_n| \\
\leq & \bar{c}\Delta_n \frac{\psi_{c,c_1}(t_j - t_k, y - (z + \theta\lambda T^h(t_k, z, w) + (1 - \theta)\lambda T_n^h(t_k, z, w)))}{(t_j - t_k)^{3/2}} \\
(3.32) \quad & \times (h^2 + h^{3/2}|w| + h(1 \wedge |z - y|)^\gamma |w|^2)(h + h^{1/2}|w|),
\end{aligned}$$

for some  $\theta \in (0, 1)$ . The point is now to get rid of the transitions appearing in the function  $\psi_{c,c_1}$ . We separate here the two assumptions at hand.

- Under  $(\mathbf{I}_G)$ , it suffices to remark that by a convexity inequality, for all  $\Theta \in \mathbb{R}^d$ :

$$\begin{aligned}
& \psi_{c,c_1}(t_j - t_k, y - z - \Theta) \\
\leq & c_1 \exp(c_1 K_1^2(t_j - t_k)) \frac{c^{d/2}}{(2\pi(t_j - t_k))^{d/2}} \exp\left(-\frac{c}{2} \frac{|z - y|^2}{t_j - t_k}\right) \exp\left(c \frac{|\Theta|^2}{t_j - t_k}\right).
\end{aligned}$$

Now, if  $\Theta$  is one of the above transitions or linear combination of transitions, we get from (3.22):

$$\begin{aligned}
(3.33) \quad & \psi_{c,c_1}(t_j - t_k, y - z - \Theta) \\
\leq & c_1 \exp(c_1 K_1^2(t_j - t_k)) \frac{c^{d/2}}{(2\pi(t_j - t_k))^{d/2}} \exp\left(-\frac{c}{2} \frac{|z - y|^2}{t_j - t_k}\right) \exp(c K_2^2 |w|^2),
\end{aligned}$$

up to a modification of  $c_1$  observing that  $h/(t_j - t_k) \leq 1$ . Since  $c$  can be chosen small enough in the previous controls, up to deteriorating the concentration properties in Lemma 4, the last term can be integrated by the standard Gaussian density  $f_\xi$  appearing in (3.28). We thus derive, from (3.33), (3.31), (3.32) and the definition in (3.29), up to modifications of  $c, c_1$ :

$$\begin{aligned}
& |\Delta_1 \varphi_\lambda^{h,n}|(t_k, z, y, w) \leq \\
\Delta_n h \bar{c} \psi_{c,c_1}(t_j - t_k, z - y) \exp(c|w|^2) & \left\{ 1 + \frac{|w|}{(t_j - t_k)^{1/2}} + \frac{|z - y|^\gamma |w|^2}{t_j - t_k} \right\},
\end{aligned}$$

which plugged into (3.28) yields up to modifications of  $c, c_1$ :

$$(3.34) \quad |\Delta H_{21}^{h,n}(t_k, t_j, z, y)| \leq \frac{\Delta_n \psi_{c,c_1}(t_j - t_k, z - y)}{(t_j - t_k)^{1-\gamma/2}}.$$

- Under  $(\mathbf{I}_{P,M})$ , in the off diagonal regime  $|z - y| \geq c(t_j - t_k)^{1/2}$  we have to discuss according to the position of  $w$  w.r.t.  $y - z$ . With the notations of  $(\mathbf{A3})$ , 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 (3.31), (3.32),

$$\begin{aligned} & (|\Delta_{11}\varphi_\lambda^{h,n}| + |\Delta_{13}\varphi_\lambda^{h,n}|)(t_k, z, y, w) \\ & \leq \bar{c}\Delta_n \frac{\psi_{c,c_1}(t_j - t_k, y - z)}{(t_j - t_k)} (h^2 + h^{3/2}|w| + h(1 \wedge |z - y|)^\gamma |w|^2), \\ & \qquad \qquad \qquad |\Delta_{12}\varphi_\lambda^{h,n}|(t_k, z, y, w) \\ & \leq c\Delta_n \frac{\psi_{c,c_1}(t_j - t_k, y - z)}{(t_j - t_k)^{3/2}} (h^2 + h^{3/2}|w| + h(1 \wedge |z - y|)^\gamma |w|^2)(h + h^{1/2}|w|). \end{aligned}$$

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

$$\begin{aligned} (3.35) \quad f_\xi(w) & \leq c \frac{1}{(1 + |w|)^M} \leq c \frac{1}{(1 + \frac{|z-y|}{h^{1/2}})^{M-(d+4)}} \frac{1}{(1 + |w|)^{d+4}} \\ & \leq c \frac{1}{(1 + \frac{|z-y|}{(t_j - t_k)^{1/2}})^{M-(d+4)}} \frac{1}{(1 + |w|)^{d+4}}, \end{aligned}$$

where the last splitting is performed in order to integrate the contribution in  $|w|^3$  coming from the upper bound for  $|\Delta_{12}\varphi_\lambda^{h,n}|$  in (3.32). Plugging the above controls in (3.28) yields:

$$(3.36) \quad |\Delta H_{21}^{h,n}(t_k, t_j, z, y)| \leq \frac{\Delta_n \Phi_{c,c_1}(t_j - t_k, z - y)}{(t_j - t_k)^{1-\gamma/2}}.$$

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

**Term  $\Delta_{22}H^{h,n}$ .** This term can be handled with the same arguments as  $\Delta_{21}H^{h,n}$ . For the sake of completeness we anyhow specify how the different contributions appear. Namely, with the notations of (3.27) and (3.28):

$$\begin{aligned} & \Delta_2 \varphi_\lambda^{h,n}(t_k, z, y, w) = \\ & \int_0^1 d\lambda \left\{ D_{T_1} \varphi_\lambda^h(T^h(t_k, y, w) + \lambda(T^h(t_k, z, w) - T^h(t_k, y, w)), T^h(t_k, y, w)) \right. \\ & \qquad \qquad \qquad \cdot (T^h(t_k, z, w) - T^h(t_k, y, w)) - \\ & \qquad \qquad \qquad D_{T_1} \varphi_{\lambda,n}^h(T_n^h(t_k, y, w) + \lambda(T_n^h(t_k, z, w) - T_n^h(t_k, y, w)), T_n^h(t_k, y, w)) \\ & \qquad \qquad \qquad \left. \cdot (T_n^h(t_k, z, w) - T_n^h(t_k, y, w)) \right\} \\ & = \left\{ \int_0^1 d\lambda \left\{ D_{T_1} \varphi_\lambda^h(T^h(t_k, y, w) + \lambda(T^h(t_k, z, w) - T^h(t_k, y, w)), T^h(t_k, y, w)) \right. \right. \\ & \qquad \qquad \qquad \left. \cdot [(T^h(t_k, z, w) - T^h(t_k, y, w)) - (T_n^h(t_k, z, w) - T_n^h(t_k, y, w))] \right\} \\ & - \left\{ \int_0^1 d\lambda \left[ D_{T_1} \varphi_{\lambda,n}^h(T_n^h(t_k, y, w) + \lambda(T_n^h(t_k, z, w) - T_n^h(t_k, y, w)), T_n^h(t_k, y, w)) \right. \right. \\ & \qquad \qquad \qquad \left. \left. - D_{T_1} \varphi_\lambda^h(T^h(t_k, y, w) + \lambda(T^h(t_k, z, w) - T^h(t_k, y, w)), T^h(t_k, y, w)) \right] \right. \\ & \qquad \qquad \qquad \left. \cdot (T_n^h(t_k, z, w) - T_n^h(t_k, y, w)) \right\} \\ & =: (\Delta_{21}\varphi_\lambda^{h,n} + \Delta_{22}\varphi_\lambda^{h,n})(t_k, z, y, w). \end{aligned}$$

In  $\Delta_{21}\varphi_\lambda^{h,n}$  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:

$$\begin{aligned} & |T^h(t_k, y, w)[T^h(t_k, y, w)]^*| \\ & \times |(T^h(t_k, z, w) - T^h(t_k, y, w)) - (T_n^h(t_k, z, w) - T_n^h(t_k, y, w))| \\ & \leq \bar{c}(h^2 + h^{3/2}|w| + h|w|^2)\Delta_n(h + h^{1/2}(1 \wedge |z - y|)^\gamma|w|), \end{aligned}$$

where the last inequality is obtained similarly to (3.30) using as well **(A4)**. The same kind of controls can be established for  $\Delta_{22}\varphi_\lambda^{h,n}$ . 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  $(\mathbf{I}_{P,M})$ , i.e. we need to integrate a term in  $|w|^4$  (see also equation (3.35) in which we performed the splitting of  $f_\xi$  on the *off-diagonal* region to integrate a contribution in  $|w|^3$ ).

We can thus claim that

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

Plugging the above control and (3.36) (or (3.34) under  $(\mathbf{I}_G)$ ) into (3.28) we derive:

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

which together with (3.26) and the decomposition (3.25) completes the proof.  $\square$

From Lemmas 4 and 5 the proof of Theorem 2 is achieved following the steps of Lemmas 2 and 3. The point is that we get want to justify the following inequality:

$$(3.37) \quad \begin{aligned} & |(\tilde{p}^h \otimes_h H^{h,(r)} - \tilde{p}_n^h \otimes_h H_n^{h,(r)})(t_i, t_j, x, y)| \\ & \leq (r+1)\Delta_n \frac{(1 \vee T)c_1^{r+1} [\Gamma(\frac{\gamma}{2})]^r}{\Gamma(1+r\frac{\gamma}{2})} \frac{c^{d/2}}{(t_j - t_i)^{d/2}} Q_{M-(d+5+\gamma)} \left( \frac{y-x}{(t_j - t_i)^{1/2}/c} \right) (t_j - t_i)^{\frac{r\gamma}{2}}. \end{aligned}$$

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 [Kol00] to investigate convolutions of “stable like” densities. Set  $m := M - (d + 5 + \gamma)$  and denote for all  $0 \leq i < j \leq N$ ,  $x \in \mathbb{R}^d$  by  $q_m(t_j - t_i, x) := \frac{c^{d/2}}{(t_j - t_i)^{d/2}} Q_{M-(d+5+\gamma)} \left( \frac{x}{(t_j - t_i)^{1/2}/c} \right)$  the density with polynomial decay appearing in Lemmas 4 and 5. Let us consider for fixed  $i < k < j$ ,  $(x, y) \in (\mathbb{R}^d)^2$  the convolution:

$$(3.38) \quad I_{t_k}^1(t_i, t_j, x, y) := \int_{\mathbb{R}^d} dz q_m(t_k - t_i, z - x) q_m(t_j - t_k, y - z).$$

- If  $|x - y| \leq c(t_j - t_i)^{1/2}$  (diagonal regime for the parabolic scaling), it is easily seen that one of the two densities in the integral (3.38) is homogeneous to  $q_m(t_j - t_i, y - x)$ . Namely, if  $(t_k - t_i) \geq (t_j - t_i)/2$ ,  $q_m(t_k - t_i, z - x) \leq \frac{c^{d/2} c_m}{(t_k - t_i)^{d/2}} \leq \frac{(2c)^{d/2} c_m}{(t_j - t_i)^{d/2}} \leq \tilde{c} q_m(t_j - t_i, y - x)$ . Thus,

$$I_{t_k}^1(t_i, t_j, x, y) \leq \tilde{c} q_m(t_j - t_i, y - x) \int_{\mathbb{R}^d} dz q_m(t_j - t_k, y - z) = \tilde{c} q_m(t_j - t_i, y - x).$$

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.

- If  $|x - y| > c(t_j - t_i)^{1/2}$  (off-diagonal regime), we introduce  $A_1 := \{z \in \mathbb{R}^d : |x - z| \geq \frac{1}{2}|x - y|\}$ ,  $A_2 := \{z \in \mathbb{R}^d : |z - y| \geq \frac{1}{2}|x - y|\}$ . 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  $|z - y| \geq \frac{c}{2}(t_j - t_i)^{1/2} \geq \frac{c}{2}(t_j - t_k)^{1/2}$  so that the density  $q_m(t_j - t_k, y - z)$  is itself in the off-diagonal regime. Write:

$$\begin{aligned} & \int_{A_2} dz q_m(t_k - t_i, z - x) q_m(t_j - t_k, y - z) \\ & \leq \int_{A_2} dz q_m(t_k - t_i, z - x) \frac{c_m(t_k - t_i)^{(m-d)/2}}{|z - y|^m} \\ & \leq \frac{c_m 2^m (t_j - t_i)^{(m-d)/2}}{|x - y|^m} \int_{A_2} dz q_m(t_k - t_i, z - x) \leq \bar{c} q_m(t_i, t_j, x, y), \end{aligned}$$

recalling that, under  $(\mathbf{I}_{P,M})$ ,  $m > d$  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  $0 \leq i < k < j$ ,  $(x, y) \in (\mathbb{R}^d)^2$  :

$$I_{t_k}^1(t_i, t_j, x, y) \leq \bar{c} q_m(t_j - t_i, y - x).$$

From the controls of Lemma 5 and following the strategy of Lemma 3, we will be led to consider convolutions of the previous type involving  $\Gamma$  functions. The above strategy thus yields (3.37) by induction.

#### 4. EULER SCHEME WITH HÖLDER COEFFICIENTS

This section is devoted to the proof of Theorem 3.

**4.1. Strategy.** Fix a given time step  $h := T/N$ ,  $N \in \mathbb{N}^*$  and define  $t_j := jh$ ,  $j \in \mathbb{N}$ . For a given  $i \in \llbracket 0, N \rrbracket$ , let  $(X_t^{t_i, x})_{t \in [t_i, T]}$  stand for the diffusion (1.1) starting from  $x$  at time  $t_i$ , with coefficients  $b, \sigma$  satisfying  $(\mathbf{A})$  and the additional  $\gamma/2$ -Hölder continuity in time. Denote by  $(X_{t_j}^{t_i, x, h})_{j \geq i}$  the associated Euler scheme which we will from now on view as a Markov Chain with dynamics of the form (1.3). It has been established in Section 2 that for all  $(i, j) \in \llbracket 0, N \rrbracket^2$ ,  $j > i$ , the respective densities of  $X_{t_j}^{t_i, x}$ ,  $X_{t_j}^{t_i, x, h}$ , denoted by  $p(t_i, t_j, x, \cdot)$ ,  $p^h(t_i, t_j, x, \cdot)$ , exist.

To investigate the weak error  $p(t_i, t_j, x, \cdot) - p^h(t_i, t_j, x, \cdot)$ , the idea is now to introduce *perturbed* processes  $(X_t^{t_i, x, (n)})_{t \geq t_i}$ ,  $(X_{t_j}^{t_i, x, h, (n)})_{t_j \geq t_i}$  with dynamics (1.2) and (1.3) respectively, which also have densities, denoted by  $p_n(t_i, t_j, x, \cdot)$ ,  $p_n^h(t_i, t_j, x, \cdot)$  at time  $t_j$  and s.t. for all  $y \in \mathbb{R}^d$ , the difference:

$$p_n(t_i, t_j, x, y) - p_n^h(t_i, t_j, x, y)$$

can be explicitly investigated. Now, if the coefficients of  $(X_t^{t_i, x, (n)})_{t \geq t_i}$ ,  $(X_{t_j}^{t_i, x, h, (n)})_{t_j \geq t_i}$  satisfy  $(\mathbf{A})$ , the  $\gamma/2$ -Hölder continuity in time, and are *smooth*, Theorem 1.1 in [KM02] extends to the inhomogeneous case and precisely gives an error expansion for the quantity  $p_n - p_n^h$  in terms of increasing powers of  $h$ . This is why a natural choice consists in considering a mollification procedure of the coefficients as perturbation.

We thus decompose the global error as:

$$\begin{aligned} & (p - p^h)(t_i, t_j, x, y) \\ & = (p - p_n)(t_i, t_j, x, y) + (p_n - p_n^h)(t_i, t_j, x, y) + (p_n^h - p^h)(t_i, t_j, x, y). \end{aligned}$$

The idea is then to control the differences  $p - p_n, p^h - p_n^h$  thanks to Theorems 1 and 2 respectively and to exploit the error expansion in [KM02] for  $p_n - p_n^h$ . We point out that two error sources will need to be equilibrated. On the one hand, if the mollified coefficients tend to the initial ones, then  $p_n^h - p^h$  and  $p_n - p$  go to zero. On the other hand, the constants in the error expansion  $p_n - p_n^h$  will depend on the derivatives of the mollified coefficients, which explode when the mollifying parameter goes to zero.

**4.2. Mollified Coefficients.** Let us now introduce the mollified coefficients defined for all  $(t, x) \in [0, T] \times \mathbb{R}^d$  by  $b_n(t, x) := b(t, \cdot) * \rho_n(x), \sigma_n(t, x) := \sigma(t, \cdot) * \rho_n(x)$  where  $\rho_n$  is a spatial mollifier, i.e. for all  $n \in \mathbb{N}, x \in \mathbb{R}^d$ ,

$$\rho_n(x) := n^d \rho(nx), \quad \rho \in C^\infty(\mathbb{R}^d, \mathbb{R}^+), \quad \int_{\mathbb{R}^d} \rho(y) dy = 1, \quad |\text{supp}(\rho)| \subset K,$$

for some compact set  $K \subset \mathbb{R}^d$ .

Write now for all  $(t, x) \in [0, T] \times \mathbb{R}^d$ :

$$\begin{aligned} b(t, x) - b_n(t, x) &:= \int_{\mathbb{R}^d} \{b(t, x) - b(t, y)\} \rho_n(x - y) dy \\ &= \int_{\mathbb{R}^d} \{b(t, x) - b(t, x - \frac{z}{n})\} \rho(z) dz. \end{aligned}$$

From the Hölder continuity of  $b$  assumed in **(A3)** and the above equation, we deduce that  $b_n$  satisfies **(A3)** as well and that:

$$\Delta_{n,b} \leq Cn^{-\gamma}, \quad C := \kappa \int_K |z|^\gamma \rho(z) dz.$$

The same analysis can be performed for  $\sigma_n$ , so that  $|\sigma(t, \cdot) - \sigma_n(t, \cdot)|_\infty \leq Cn^{-\gamma}$ . We anyhow need to control the Hölder norm as well. Precisely, for all  $t \in \mathbb{R}^+, (x, y) \in (\mathbb{R}^d)^2$ :

$$\begin{aligned} &[\sigma(t, x) - \sigma_n(t, x)] - [\sigma(t, y) - \sigma_n(t, y)] \\ &= \int_{\mathbb{R}^d} [\sigma(t, x) - \sigma(t, x - \frac{z}{n})] - [\sigma(t, y) - \sigma(t, y - \frac{z}{n})] \rho(z) dz. \end{aligned}$$

It readily follows from the  $\gamma$ -Hölder continuity of  $\sigma$  that one has the following controls:

$$(4.1) \quad |[\sigma(t, x) - \sigma_n(t, x)] - [\sigma(t, y) - \sigma_n(t, y)]| \leq 2(\kappa \vee C)(|x - y|^\gamma \wedge n^{-\gamma}).$$

**4.2.1. Theorems 1 and 2 for Mollified coefficients.** From (4.1) one easily derives, observing that one of the two terms is greater or equal than the other one, that:

$$(4.2) \quad |\sigma - \sigma_n|_{\varepsilon_n} \leq Cn^{-(\gamma - \varepsilon_n)}, \quad \forall \varepsilon_n \in (0, 1), \quad \forall n \in \mathbb{N}^*,$$

and where  $\varepsilon_n$  is meant to tend to 0 with  $n$ . We now want to exploit the results from Theorems 1 and 2 taking  $\Delta_n := Cn^{-(\gamma - \varepsilon_n)}$ . This choice induces to consider  $\varepsilon_n$  as Hölder index in **(A3)**. The difficulty, w.r.t. the previous proofs is to thoroughly specify how the *small* Hölder regularity, here in  $\varepsilon_n$ , affects the constants in the expansions. From Section 3 (especially Lemma 3), we have to investigate the convergence of the series:

$$S_T(\varepsilon_n) := \sum_{r \geq 0} (r+1) c_1^{r+1} \frac{\Gamma(\frac{\varepsilon_n}{2})^r}{\Gamma(1 + r\frac{\varepsilon_n}{2})} T^{\frac{r\varepsilon_n}{2}}.$$

We can assume w.l.o.g. that  $T \leq 1$ , so that  $S_T(\varepsilon_n) \leq S_1(\varepsilon_n)$ . Set  $r_0 := \lceil \frac{2}{\varepsilon_n} \rceil$  and write:

$$\begin{aligned}
S_1(\varepsilon_n) &\leq c \sum_{k \in \mathbb{N}} (k+1)r_0 \sum_{kr_0 \leq r < (k+1)r_0} \frac{\{c_1 \Gamma(\frac{\varepsilon_n}{2})\}^r}{\Gamma(1+k)} \\
&\leq c \sum_{k \in \mathbb{N}} \frac{(k+1)r_0}{\Gamma(k+1)} \sum_{kr_0 \leq r < (k+1)r_0} \{c_1(2\varepsilon_n^{-1} + \exp(-1))\}^r \\
&\leq cr_0^2 \sum_{k \in \mathbb{N}} \frac{(k+1)}{\Gamma(k+1)} [c_1(2\varepsilon_n^{-1} + \exp(-1))^{r_0}]^{k+1} \\
(4.3) \quad &\leq C \exp(c(2\varepsilon_n^{-1} + 1)^{2\varepsilon_n^{-1}+1}).
\end{aligned}$$

With the notations introduced in Section 4.1, we derive from (4.3), following the proofs of Theorems 1 and 2 (under  $(\mathbf{I}_G)$ ) that there exist constants  $C, c$  s.t. for all  $(j, x, y) \in (i, N] \times (\mathbb{R}^d)^2$ :

$$\begin{aligned}
(4.4) \quad &|p - p_n|(t_i, t_j, x, y) + |p^h - p_n^h|(t_i, t_j, x, y) \\
&\leq C n^{-\gamma + \varepsilon_n} \exp(c(2\varepsilon_n^{-1} + 1)^{2\varepsilon_n^{-1}+1}) p_c(t_i, t_j, x, y).
\end{aligned}$$

**4.2.2. Auxiliary Results for Mollifiers.** Finally, let us give some useful controls on the derivatives of the mollified coefficients that will be needed for the analysis below. These results will be crucial to analyze the weak error between the diffusion with mollified coefficients and the Euler scheme.

For all multi-index  $\alpha$ ,  $|\alpha| \in \llbracket 1, 4 \rrbracket$  and  $(t, x) \in [0, T] \times \mathbb{R}^d$  write:

$$\begin{aligned}
D_x^\alpha \sigma_n(t, x) &= \int_{\mathbb{R}^d} \sigma(t, z) D_x^\alpha \rho_n(x - z) dz \\
&= \int_{\mathbb{R}^d} [\sigma(t, z) - \sigma(t, x)] D_x^\alpha \rho_n(x - z) dz.
\end{aligned}$$

Indeed, setting for all  $x \in \mathbb{R}^d$ ,  $g_n(x) := \int_{\mathbb{R}^d} \rho_n(x - z) dz = 1$  we have  $D_x^\alpha g_n(x) := \int_{\mathbb{R}^d} D_x^\alpha \rho_n(x - z) dz = 0$ . Thus, since  $|D_x^\alpha \rho_n(x - z)| \leq cn^{|\alpha|+d} |D_w^\alpha \rho(w)|_{w=n(x-z)}$ , we derive:

$$\begin{aligned}
|D_x^\alpha \sigma_n(t, x)| &\leq c \int_{\mathbb{R}^d} |\sigma(t, z) - \sigma(t, x)| n^{d+|\alpha|} |D_w^\alpha \rho(w)|_{w=n(x-z)} dz \\
&\leq c \kappa n^{|\alpha|-\gamma} \int_{\mathbb{R}^d} (n|z - x|)^\gamma n^d |D_w^\alpha \rho(w)|_{w=n(x-z)} dz \leq cn^{|\alpha|-\gamma},
\end{aligned}$$

exploiting the Hölder continuity assumption **(A3)** for  $\sigma$  in the last but one inequality and the assumptions on  $\rho$  for the last one. Similarly, we derive for all  $(t, x, y) \in [0, T] \times (\mathbb{R}^d)^2$ :

$$\begin{aligned}
|D_x^\alpha \sigma_n(t, x) - D_x^\alpha \sigma_n(t, y)| &\leq \int_{\mathbb{R}^d} |\sigma(t, x - z) - \sigma(t, y - z)| n^{d+|\alpha|} |D_x^\alpha \rho(w)|_{w=nz} dz \\
&\leq \kappa n^{|\alpha|} |x - y|^\gamma.
\end{aligned}$$

The same bounds hold for  $b_n$  as well. We can summarize those controls as follows:

$$(4.5) \quad |D_x^\alpha b_n|_\infty + |D_x^\alpha \sigma_n|_\infty \leq cn^{|\alpha|-\gamma}, |D_x^\alpha \sigma_n|_\gamma \leq cn^{|\alpha|}.$$

**4.3. Error Analysis Through Stability.** With the notations of Section 4.1 write now for all  $0 \leq i < j \leq N$ ,  $(x, y) \in (\mathbb{R}^d)^2$  from (4.4):

$$\begin{aligned} & |p(t_i, t_j, x, y) - p^h(t_i, t_j, x, y)| \leq \\ & |p - p_n|(t_i, t_j, x, y) + |(p_n - p_n^h)|(t_i, t_j, x, y) + |p_n^h - p^h|(t_i, t_j, x, y) \\ & \leq Cn^{-\gamma+\varepsilon_n} \exp(c(2\varepsilon_n^{-1} + 1)^{2\varepsilon_n^{-1}+1}) p_c(t_j - t_i, y - x) + |(p_n - p_n^h)|(t_i, t_j, x, y). \end{aligned} \quad (4.6)$$

Our goal is now to choose a suitable sequence  $(\varepsilon_n)_{n \geq 1}$  to make the contribution  $n^{\varepsilon_n} \exp(c(2\varepsilon_n^{-1} + 1)^{2\varepsilon_n^{-1}+1})$  go rather slowly to infinity, so that we can hope to get an equilibrium with the error coming from the Euler discretization in (4.6) and analyzed below.

Let us consider a specific sequence  $\varepsilon_n := 2^{\frac{\log_3(n)}{\log_2(n)}}$ , where we denote for  $k \in \mathbb{N}$ , by  $\log_k(x)$  the  $k^{\text{th}}$  iterated logarithm of  $x$ . Setting  $b_n := n^{\varepsilon_n}$  and  $a_n := \exp\left(c(2\varepsilon_n^{-1} + 1)^{2\varepsilon_n^{-1}+1}\right)$ , we get that:

$$\begin{aligned} \log_2(b_n) &= \log(\varepsilon_n \log(n)) = \log(2) + \log_4(n) - \log_3(n) + \log_2(n), \\ \log_2(a_n) &= \log(c(2\varepsilon_n^{-1} + 1)^{2\varepsilon_n^{-1}+1}) = \log(c) + (2\varepsilon_n^{-1} + 1) \log(2\varepsilon_n^{-1} + 1) \\ &= \log(c) + \left(\frac{\log_2(n)}{\log_3(n)} + 1\right) \log(2\varepsilon_n^{-1}(1 + \frac{\varepsilon_n}{2})) \\ &= \log(c) + \left(\frac{\log_2(n)}{\log_3(n)} + 1\right) \{\log(2\varepsilon_n^{-1}) + \log(1 + \frac{\varepsilon_n}{2})\} \\ &= \log(c) + \left(\frac{\log_2(n)}{\log_3(n)} + 1\right) \{\log_3(n) - \log_4(n) + \log(1 + \frac{\varepsilon_n}{2})\} \\ &= \log_2(n) - \frac{\log_2(n) \log_4(n)}{\log_3(n)} + \log_3(n) - \log_4(n) + R_n, \\ R_n &:= \log(c) + \log(1 + \frac{\log_3(n)}{\log_2(n)}) \left\{ \frac{\log_2(n)}{\log_3(n)} + 1 \right\}. \end{aligned}$$

It is easily seen that there exists a finite constant  $\bar{c} > 0$  s.t. for all  $n$  large enough,  $R_n \leq \bar{c}$  and that  $\log_2(b_n) \geq \log_2(a_n) - \bar{c}$ . By monotonicity of the exponential we thus derive that there exists  $c := c(\bar{c}) > 0$  s.t. for all  $n \geq 0$ :

$$(4.7) \quad b_n a_n = n^{\varepsilon_n} \exp(c(2\varepsilon_n^{-1} + 1)^{2\varepsilon_n^{-1}+1}) \leq c(b_n)^2 = cn^{2\varepsilon_n}.$$

To analyze the last contribution in equation (4.6), we recall that since the densities  $p_n, p_n^h$  are now respectively associated with a diffusion process and its Euler scheme with *smooth* coefficients, they can be compared thanks to the results in [KM02] adapted to the current inhomogeneous setting. The only delicate, but crucial, point is that we must here specify the dependence on the derivatives of the coefficients, which here explode when  $n$  goes to infinity. Precisely, from Theorem 1.1, Theorem 2.1 and their proofs in [KM02] we have:

$$(4.8) \quad \begin{aligned} & (p_n - p_n^h)(t_i, t_j, x, y) = \\ & (p_n - p_n^d)(t_i, t_j, x, y) + \frac{h}{2} \left\{ p_n^d \otimes_h (\tilde{L}_{\cdot, *}^n - \tilde{L}_{\cdot, *}^{*,n})^2 p_n^h(t_i, t_j, x, y) \right\} \\ & \quad \quad \quad + h^2 R_n(t_i, t_j, x, y), \end{aligned}$$

where we denote for  $k \in \{1, 2\}$ ,  $t \in (t_i, t_j)$ ,  $(\tilde{L}_{t,*}^n)^k \phi(x, y) := (L_{t,\xi}^n)^k \phi(x, y)|_{\xi=x}$ ,  $(\tilde{L}_t^{*,n})^k \phi(x, y) := (L_{t,\xi}^n)^k \phi(x, y)|_{\xi=y}$  for

$$L_{t,\xi}^n \phi(x, y) = \langle b_n(t, \xi), D_x \phi(x, y) \rangle + \frac{1}{2} \text{Tr}(a_n(t, \xi) D_x^2 \phi(x, y)).$$

Observe that  $L_t^n \phi(x) = L_{t,*}^n \phi(x)$ , but more generally the operators do not coincide anymore when iterated. The notations  $\tilde{L}_{t,*}^n, \tilde{L}_t^{*,n}$  in (4.8) are used to emphasize the time dependence of the operators in the discrete convolution  $\otimes_h$ . Also  $p_n^d(t_i, t_j, x, y) := \sum_{r \in \mathbb{N}} \tilde{p}_n \otimes_h H_n^{(r)}(t_i, t_j, x, y)$  is the expression obtained from the parametrix expansion of the diffusion density (2.5) replacing the continuous time by the discrete time convolution. Again, the subscript  $n$  is meant to explicitly express the dependence on the mollified coefficients. Also, even though  $p_n^d(t_i, t_j, x, \cdot)$  is not *a priori* a density, we will call it so with a slight abuse of terminology. We mention carefully that for the second contribution in the previous expansion the smoothness in time of the coefficients is not needed. On the other hand, it is clearly required to derive some convergence rates, since we investigate the difference between integrals and Riemann sums to control  $p_n - p_n^d$  (see Proposition 4 and its proof below).

The contribution  $h^2 R_n(t_i, t_j, x, y)$  can be seen as a *remainder*. It involves iterated differences of generators up to order 3. When coefficients are smooth, this term is negligible w.r.t. to the other ones in (4.8). Observe anyhow that, since it involves derivatives of the heat kernel up to order 6 (whereas the first term only up to order 4), and that the derivatives of our coefficients explode with  $n$  (see equation (4.5)), this term will actually dominate in the previous expansion.

The key elements are now the following Propositions. The first one gives bounds for the derivatives of the densities involved in the parametrix series. The second one controls the difference between the discrete and continuous convolutions in (4.8).

*Proposition 3* (Controls for the derivatives of the densities). Let  $\alpha$ ,  $|\alpha| \leq 6$  be a multi-derivation index. There exist constants  $C, c$  s.t. for all  $0 \leq s < t \leq T$ ,  $(x, y) \in (\mathbb{R}^d)^2$ :

$$(4.9) \quad \begin{aligned} |D_x^\alpha \bar{p}_n(s, t, x, y)| &\leq \frac{C}{(t-s)^{|\alpha|/2}} p_c(t-s, y-x), |\alpha| \leq 2, \\ |D_x^\alpha \bar{p}_n(s, t, x, y)| &\leq \frac{C}{(t-s)^{|\alpha|/2}} p_c(t-s, y-x) (1 + n^{|\alpha|-2} (t-s)^{\gamma/4}), |\alpha| \geq 3, \\ |D_y^\alpha \bar{p}_n(s, t, x, y)| &\leq \frac{C n^{|\alpha|-\gamma}}{(t-s)^{|\alpha|/2}} p_c(t-s, y-x). \end{aligned}$$

where in the above expression  $\bar{p}_n$  can be any of the densities  $p_n, p_n^d, p_n^h$ . For  $p_n^d, p_n^h$ ,  $s, t$  are taken on the time grid.

*Remark 3.* We point out that the previous controls for  $\bar{p}_n = p_n$  improve in some sense those of [IKO62], since we do not exploit any smoothness in time of the coefficients and we get the same pointwise controls for the derivatives of the non degenerate heat-kernel with Hölder coefficients in space up to order 2.

*Proposition 4* (Bounds for the difference between continuous and discrete time convolutions). There exists  $C, c$  s.t. for all  $0 \leq t_i < t_j \leq T$ ,  $(x, y) \in (\mathbb{R}^d)^2$ :

$$(4.10) \quad |(p_n - p_n^d)(t_i, t_j, x, y)| \leq C \left\{ h^{(\gamma-\varepsilon_n)/2} + \frac{n^{2-\gamma} h}{(t_j - t_i)^{1-\gamma/2}} \right\} p_c(t_j - t_i, y - x).$$

We postpone the proof of Propositions 3 and 4 for clarity. It now remains to exploit Propositions 3, 4 and (4.8) to specifically control how the weak error for the densities depends on the explosive norms of the mollified coefficients.

Observe from Proposition 3 that, for all  $k \in \llbracket 1, j-1 \rrbracket$ ,  $(z, y) \in (\mathbb{R}^d)^2$ :

$$|(\tilde{L}_{t_k, *}^n - \tilde{L}_{t_k}^{*,n})p_n^h(t_k, t_j, z, y)| \leq \frac{C}{(t_j - t_k)^{1-\gamma/2}} p_c(t_j - t_k, y - z).$$

We analyze the contribution  $p_n^d \otimes_h (\tilde{L}_{\cdot, *}^n - \tilde{L}_{\cdot}^{*,n})^2 p_n^h(t_i, t_j, x, y)$  in (4.8) thanks to Proposition 3 as follows:

$$\begin{aligned} & |[p_n^d \otimes_h (\tilde{L}_{\cdot, *}^n - \tilde{L}_{\cdot}^{*,n})^2 p_n^h](t_i, t_j, x, y)| \leq Ch \\ & \left\{ \sum_{k \in \llbracket i+1, \lceil (i+j)/2 \rceil \rrbracket} \int_{\mathbb{R}^d} \frac{n^{1-\gamma}}{(t_k - t_i)^{1/2}} p_c(t_k - t_i, z - x) \left[ \frac{n(t_j - t_k)^{\gamma/4}}{(t_j - t_k)^{3/2}} \right] p_c(t_j - t_k, y - z) dz \right. \\ & \left. + \sum_{k \in \llbracket \lceil (i+j)/2 \rceil + 1, j-1 \rrbracket} \int_{\mathbb{R}^d} \frac{n^{2-\gamma}}{(t_k - t_i)} p_c(t_k - t_i, z - x) \frac{1}{(t_j - t_k)^{1-\gamma/2}} p_c(t_j - t_k, y - z) dz \right\}, \end{aligned}$$

where we perform one integration by part w.r.t.  $z$  for the first integral and two for the second one (taking once the adjoints). We thus get:

$$(4.11) \quad |[p_n^d \otimes_h (\tilde{L}_{\cdot, *}^n - \tilde{L}_{\cdot}^{*,n})^2 p_n^h](t_i, t_j, x, y)| \leq \frac{C}{(t_j - t_i)^{1-\gamma/4}} n^{2-\gamma} p_c(t_j - t_i, y - x).$$

For the remainder  $R_n(t_i, t_j, x, y)$  in (4.8), since it involves differential operators of order six, the idea would be to proceed as above doing 1 integration by parts for  $k \in \llbracket i+1, \lceil (i+j)/2 \rceil \rrbracket$  and 4 if  $k \in \llbracket \lceil (i+j)/2 \rceil + 1, j-1 \rrbracket$ . In any case this would yield an explosive constant in  $n^{4-\gamma}$ . Indeed, for  $k \in \llbracket i+1, \lceil (i+j)/2 \rceil \rrbracket$ , the most explosive term corresponds to the product of the constants associated with a first derivative of the densities w.r.t the frozen coefficient (of order  $n^{1-\gamma}$  from Proposition 3) and the 5<sup>th</sup> derivatives of one of the densities w.r.t. to the backward (or unfrozen) variable in Proposition 3 (of order  $n^3$ ). For  $k \in \llbracket \lceil (i+j)/2 \rceil + 1, j-1 \rrbracket$ , taking twice the adjoints yields to consider derivatives up to order 4 of the coefficients or the densities in the frozen coefficients. This still yields a bound in  $n^{4-\gamma}$  (see again Proposition 3 and (4.5)).

We thus finally derive from (4.8), (4.10) and (4.11):

$$\begin{aligned} & |(p_n - p_n^h)(t_i, t_j, x, y)| \\ & \leq C \left\{ h^{(\gamma-\varepsilon_n)/2} n^{\varepsilon_n} + \frac{hn^{2-\gamma}}{(t_j - t_i)^{1-\gamma/4}} + \frac{h^2 n^{4-\gamma}}{(t_j - t_i)^{2-\gamma/4}} \right\} p_c(t_j - t_i, y - x), \end{aligned}$$

which together with (4.6) and (4.7) yields:

$$\begin{aligned} & |(p - p^h)(t_i, t_j, x, y)| \\ & \leq C \left\{ h^{(\gamma-\varepsilon_n)/2} n^{\varepsilon_n} + \frac{1}{n^{\gamma-2\varepsilon_n}} + \frac{hn^{2-\gamma}}{(t_j - t_i)^{1-\gamma/4}} + \frac{h^2 n^{4-\gamma}}{(t_j - t_i)^{2-\gamma/4}} \right\} p_c(t_j - t_i, y - x). \end{aligned}$$

Equilibrating the two error sources, taking  $\frac{1}{n^{\gamma-2\varepsilon_n}} = \frac{h^2 n^{4-\gamma}}{(t_j - t_i)^{2-\gamma/4}} \geq \frac{hn^{2-\gamma}}{(t_j - t_i)^{1-\gamma/4}}$ , leads to

$$|(p - p^h)(t_i, t_j, x, y)| \leq C \left( \frac{h}{(t_j - t_i)^{1-\gamma/4}} \right)^{\frac{\gamma-2\varepsilon_n}{2-\varepsilon_n}} p_c(t_j - t_i, y - x).$$

We complete the proof of Theorem 3 letting  $n$  go to infinity in the previous expression (recall that  $\varepsilon_n \xrightarrow[n]{\rightarrow} 0$ ).

#### 4.4. Proof of Proposition the Technical Results.

##### 4.4.1. Proof of Proposition 3.

*Proof.* Let us establish the result for  $p_n$ . We start from the parametrix representation  $p_n$  discussed in Section 2. Namely, for all  $0 \leq s < t \leq T$ ,  $(x, y) \in (\mathbb{R}^d)^2$ :

$$p_n(s, t, x, y) = \tilde{p}_n(s, t, x, y) + \sum_{i \geq 1} \tilde{p}_n \otimes H_n^{(i)}(s, t, x, y).$$

In all cases, we can readily derive from (2.8) and (4.5) that for the *main* term in the expansion:

$$\begin{aligned} |D_x^\alpha \tilde{p}_n(s, t, x, y)| &\leq \frac{C}{(t-s)^{|\alpha|/2}} p_c(t-s, y-x), \\ |D_y^\alpha \tilde{p}_n(s, t, x, y)| &\leq \frac{Cn^{|\alpha|-\gamma}}{(t-s)^{|\alpha|/2}} p_c(t-s, y-x). \end{aligned}$$

Let us now concentrate on the remainder term:

$$\begin{aligned} R_n(s, t, x, y) &:= \sum_{i \geq 1} \tilde{p}_n \otimes H_n^{(i)}(s, t, x, y) = \tilde{p}_n \otimes \Phi_n(s, t, x, y), \\ \Phi_n(s, t, x, y) &:= \sum_{i \geq 1} H_n^{(i)}(s, t, x, y). \end{aligned}$$

We focus on the first two inequalities in (4.9), the last one can be proved similarly. The ideas are close to those in [IKO62], but we need to adapt them since they considered the “forward” version of the parametrix expansions. The key point is that, for Hölder coefficients we have bounded controls for the derivatives of the remainder in the backward variable up to order two. It is first easily seen for the first derivatives, since the first order derivation gives an integrable singularity in time in the previous expansions. Indeed, from (2.8) and (2.10), one readily gets the statement if  $|\alpha| = 1$ . The case  $|\alpha| \geq 2$  is much more subtle and needs to be discussed thoroughly. Write indeed:

$$\begin{aligned} D_x^\alpha R_n(s, t, x, y) &= \lim_{\varepsilon \rightarrow 0} \int_{s+\varepsilon}^{(t+s)/2} du \int_{\mathbb{R}^d} D_x^\alpha \tilde{p}_n(s, u, x, z) \Phi_n(u, t, z, y) dy + \\ &\quad \int_{(t+s)/2}^t du \int_{\mathbb{R}^d} D_x^\alpha \tilde{p}_n(s, u, x, z) \Phi_n(u, t, z, y) dy \\ (4.12) \quad &:= \lim_{\varepsilon \rightarrow 0} D_x^\alpha R_n^\varepsilon(s, t, x, y) + D_x^\alpha R_n^f(s, t, x, y). \end{aligned}$$

The contribution  $D_x^\alpha R_n^f(s, t, x, y)$  does not exhibit time singularities in the integral, since on the considered integration set  $u-s \geq \frac{1}{2}(t-s)$ . Thus, still from (2.8), (2.10):

$$(4.13) \quad |D_x^\alpha R_n^f(s, t, x, y)| \leq \frac{C}{(t-s)^{|\alpha|/2}} p_c(t-s, y-x).$$

The delicate contribution is indeed  $D_x^\alpha R_n^\varepsilon(s, t, x, y)$  for which we need to be more careful. If  $|\alpha| = 2$  we exploit some cancellation properties of the derivatives of the

Gaussian kernels. Recall now that for an arbitrary  $w \in \mathbb{R}^d$ , setting

$$\begin{aligned} \tilde{p}_n^w(s, u, x, z) &= \frac{1}{(2\pi)^{d/2} \det(\Sigma_n(s, u, w))^{1/2}} \\ &\quad \times \exp\left(-\frac{1}{2} \langle \Sigma_n(s, u, w)^{-1}(z - x - b_n(s, u, w)), z - x - b_n(s, u, w) \rangle\right), \\ D_{x_i x_j}^2 \tilde{p}_n^w(s, u, x, z) &= \left\{ (\Sigma_n^{-1}(s, u, w)(z - x - b_n(s, u, w)))_i (\Sigma_n^{-1}(s, u, w)(z - x - b_n(s, u, w)))_j \right. \\ &\quad \left. - \delta_{ij} (\Sigma_n^{-1}(s, u, w))_{ii} \right\} \tilde{p}_n^w(s, u, x, z), \quad \forall (i, j) \in \llbracket 1, d \rrbracket^2, \end{aligned} \quad (4.14)$$

so that

$$(4.15) \quad \int_{\mathbb{R}^d} D_{x_i x_j}^2 \tilde{p}_n^w(s, u, x, z) dz = 0.$$

Introducing the centering function  $c_n^\alpha(s, u, x, z) := (D_x^\alpha \tilde{p}_n^w(s, u, x, z))|_{w=x}$ , we rewrite:

$$\begin{aligned} D_x^\alpha R_n^\varepsilon(s, t, x, y) &= \int_{s+\varepsilon}^{(s+t)/2} du \int_{\mathbb{R}^d} (D_x^\alpha \tilde{p}_n - c_n^\alpha)(s, u, x, z) \Phi_n(u, t, z, y) dz \\ &\quad + \int_{s+\varepsilon}^{(s+t)/2} du \int_{\mathbb{R}^d} c_n^\alpha(s, u, x, z) (\Phi_n(u, t, z, y) - \Phi_n(u, t, x, y)) dz \\ (4.16) \quad &:= (R_n^{\varepsilon,1} + R_n^{\varepsilon,2})(s, t, x, y), \end{aligned}$$

exploiting the centering condition (4.15) to introduce the last term of the first equality. On the one hand, the terms  $D_x^\alpha \tilde{p}_n(s, u, x, z)$ ,  $c_n^\alpha(s, u, x, z)$  only differ in their frozen coefficients (respectively at point  $z$  and  $x$ ). Exploiting the Hölder property in space of the mollified coefficients, it is then easily seen that:

$$\begin{aligned} |(D_x^\alpha \tilde{p}_n - c_n^\alpha)(s, u, x, z)| &\leq \frac{C|x-z|^\gamma}{(u-s)} p_c(u-s, z-x) \\ &\leq \frac{C}{(u-s)^{1-\gamma/2}} p_c(u-s, z-x), \end{aligned}$$

yielding an integrable singularity in time so that:

$$(4.17) \quad |R_n^{\varepsilon,1}(s, t, x, y)| \leq \frac{C}{(t-s)^{1-\gamma/2}} p_c(t-s, y-x).$$

Let us now control the other contribution. The key idea is now to exploit the *smoothing property* of the kernel  $\Phi_n$ . Assume indeed that for  $A := \{z \in \mathbb{R}^d : |x-z| \leq c(t-s)^{1/2}\}$  (recall as well that  $u \in [\frac{s+t}{2}, t]$ ) one has:

$$(4.18) \quad |\Phi_n(u, t, x, y) - \Phi_n(u, t, z, y)| \leq C \frac{|x-z|^{\gamma/2}}{(t-u)^{1-\beta}} p_c(t-u, y-z), \quad \beta > 0.$$

Then, we can derive from (2.8), (4.16) and (4.18):

$$\begin{aligned} &|R_n^{\varepsilon,2}(s, t, x, y)| \\ &\leq C^2 \int_{s+\varepsilon}^{(s+t)/2} du \int_A \frac{|x-z|^{\gamma/2}}{(u-s)} p_c(u-s, z-x) \frac{1}{(t-u)^{1-\beta}} p_c(t-u, y-z) dz \\ &+ \frac{C}{(t-s)^{\gamma/2}} \int_{s+\varepsilon}^{(s+t)/2} du \int_{A^c} \frac{|x-z|^\gamma}{(u-s)} p_c(u-s, z-x) \{|\Phi_n(u, t, z, y)| + |\Phi_n(u, t, x, y)|\} dz. \end{aligned}$$

Since equation (2.9) for  $H_n$  yields

$$(4.19) \quad |H_n^{(r)}(s, t, x, y)| \leq \exp(c_1 K_1^2(t-s)) c_1^r \prod_{i=1}^{r-1} B\left(\frac{\gamma}{2}, 1 + (i-1)\frac{\gamma}{2}\right) p_c(t-s, y-x)(t-s)^{-1+\frac{r\gamma}{2}},$$

with the convention  $\prod_{i=1}^0 = 1$ , we derive  $|\Phi_n(u, t, z, y)| \leq \frac{C}{(t-u)^{1-\gamma/2}} p_c(t-u, y-z)$  and  $|\Phi_n(u, t, x, y)| \leq \frac{C}{(t-u)^{1-\gamma/2}} p_c(t-u, y-x)$ . We finally get on the considered time set:

$$|R_n^{\varepsilon, 2}(s, t, x, y)| \leq \frac{C}{t-s} p_c(t-s, y-x),$$

which together with (4.17), (4.16), (4.13) and (4.12) gives the statement. It remains to establish (4.18). From the definition of  $\Phi_n$  and the smoothing effect of the kernel  $H_n$  in (4.19), it suffices to prove that on the set  $\bar{A} := \{z \in \mathbb{R}^d : |z-x| \leq c(u'-u)^{1/2}\}$ :

$$(4.20) \quad |H_n(u, u', x, w) - H_n(u, u', z, w)| \leq C \frac{|x-z|^{\gamma/2}}{(u'-u)^{1-\beta}} p_c(u'-u, w-z), \quad \beta > 0,$$

for  $u' \in (u, t]$ ,  $u \in [s, (s+t)/2]$ . Observe that  $\bar{A} \subset A$ . Indeed, recalling that we want to establish (4.18) on  $A$  if  $z \notin \bar{A}$ , we get from (4.19):

$$\begin{aligned} & \int_u^t du' \int_{\bar{A}^c} |H_n(u, u', x, w) - H_n(u, u', z, w)| \left| \sum_{i \geq 1} H_n^{(i)}(u', t, w, y) \right| dw \\ & \leq \int_u^t du' \int_{\bar{A}^c} \frac{C}{(u'-u)^{1-\gamma/2}} (p_c(u'-u, w-x) + p_c(u'-u, w-z)) \\ & \times \frac{|x-z|^{\gamma/2}}{(u'-u)^{\gamma/4}} \frac{C}{(t-u')^{1-\gamma/2}} p_c(t-u', y-w) dw \leq C \frac{|x-z|^{\gamma/2}}{(t-u)^{1-\gamma/4}} p_c(t-u, y-z), \end{aligned}$$

exploiting that  $z \in A$ ,  $t-u \geq \frac{1}{2}(t-s)$ , and the usual convexity inequality  $\frac{|y-x|^2}{t-u} \geq \frac{|y-z|^2}{2(t-u)} - \frac{|z-x|^2}{t-u}$  for the last inequality. On the other hand, on  $\bar{A}$  we get (4.18) from (4.20) and (4.19).

Let us turn to the proof of (4.20). We concentrate on the second derivatives in  $H_n$  which yield the most singular contributions:

$$\begin{aligned} & \text{Tr}((a_n(u, x) - a_n(u, w)) D_x^2 \tilde{p}_n(u, u', x, w)) - \text{Tr}((a_n(u, z) - a_n(u, w)) D_x^2 \tilde{p}_n(u, u', z, w)) \\ & = \text{Tr}((a_n(u, x) - a_n(u, z)) D_x^2 \tilde{p}_n(u, u', x, w)) \\ & \quad - \text{Tr}((a_n(u, z) - a_n(u, w)) (D_x^2 \tilde{p}_n(u, u', z, w) - D_x^2 \tilde{p}_n(u, u', x, w))) \\ & =: I + II. \end{aligned}$$

Then, from (2.8),

$$\begin{aligned} |I| & \leq C \frac{|x-z|^\gamma}{(u-u')^\gamma} p_c(u'-u, w-x) \leq \frac{C|x-z|^{\gamma/2}}{(u-u')^{1-\gamma/4}} p_c(u'-u, w-x) \\ & \leq \frac{C|x-z|^{\gamma/2}}{(u-u')^{1-\gamma/4}} p_c(u'-u, w-z), \end{aligned}$$

using that  $z \in \bar{A}$  for the second inequality, again combined with the convexity inequality  $\frac{|x-w|^2}{u'-u} \geq \frac{|z-w|^2}{2(u'-u)} - \frac{|x-z|^2}{u'-u}$  for the last one. Now, from the explicit expression of the second order derivatives in (4.14), **(A2)** and usual computations we

also derive:

$$|II| \leq \frac{C|z-w|^\gamma}{(u'-u)} \frac{|z-x|^{\gamma/2}}{(u'-u)^{\gamma/4}} p_c(u'-u, z-w) \leq \frac{C|z-x|^{\gamma/2}}{(u'-u)^{1-\gamma/4}} p_c(u'-u, w-z).$$

This gives (4.20) and completes the proof for  $|\alpha| \leq 2$ .

Let us now turn to  $|\alpha| \geq 3$ . In those cases the singularities induced by the derivatives are not integrable in short time, even if we exploit cancellations. We are thus led to perform integration by parts, deteriorating the bounds since these operations make the derivatives of the mollified coefficients appear.

Recalling  $\alpha \in \mathbb{N}^d$ , denote by  $l$  a multi-index s.t.  $|l| = 2$  and  $\alpha - l \geq 0$  (where the inequality is to be understood componentwise). From equations (4.12), (4.13), we only have to consider the contribution  $D_x^\alpha R_n^\varepsilon(s, t, x, y)$ . Write:

$$\begin{aligned} & D_x^\alpha R_n^\varepsilon(s, t, x, y) \\ &= D_x^{\alpha-l} \int_{s+\varepsilon}^{(t+s)/2} du \int_{\mathbb{R}^d} D_x^l \tilde{p}_n(s, u, x, z) \Phi_n(u, t, z, y) dz \\ (4.21) \quad &= D_x^{\alpha-l} \int_{s+\varepsilon}^{(t+s)/2} du \int_{\mathbb{R}^d} g^{l,n}(s, u, x, z) \Phi_n(u, t, z, y) dz, \end{aligned}$$

where  $g^{l,n}(s, u, x, z) := D_x^l \tilde{p}_n(s, u, x, z)$ . Let us write introducing the cancellation term  $c_n^l$ :

$$\begin{aligned} D_x^\alpha R_n^\varepsilon(s, t, x, y) &= D_x^{\alpha-l} \int_{s+\varepsilon}^{(s+t)/2} du \int_{\mathbb{R}^d} (g^{l,n} - c_n^l)(s, u, x, z) \Phi_n(u, t, z, y) dz \\ &\quad + D_x^{\alpha-l} \int_{s+\varepsilon}^{(s+t)/2} du \int_{\mathbb{R}^d} c_n^l(s, u, x, z) (\Phi_n(u, t, z, y) - \Phi_n(u, t, x, y)) dz \\ &= D_x^{\alpha-l} \int_{s+\varepsilon}^{(s+t)/2} du \int_{\mathbb{R}^d} (g^{l,n} - c_n^l)(s, u, x, x+z) \Phi_n(u, t, x+z, y) dz \\ (4.22) \quad &\quad + D_x^{\alpha-l} \int_{s+\varepsilon}^{(s+t)/2} du \int_{\mathbb{R}^d} c_n^l(s, u, x, x+z) (\Phi_n(u, t, x+z, y) - \Phi_n(u, t, x, y)) dz. \end{aligned}$$

The purpose of that change of variable, already performed in [KM02], is that we get integrable time singularities in the contributions  $D_x^{\alpha-l}(g^{l,n} - c_n^l)(s, u, x, x+z)$ . Anyhow, the mollified coefficients  $b_n, \sigma_n$  have explosive derivatives. From the definition of  $g^{l,n}$  and (4.5) one easily gets that there exists  $c, C$  s.t. for all  $\alpha$ ,  $|\alpha| \leq 6$ :

$$\begin{aligned} (4.23) \quad |D_x^{\alpha-l}(g^{l,n} - c_n^l)(s, u, x, x+z)| &\leq \frac{Cn^{|\alpha-l|}}{(u-s)^{1-\gamma/2}} p_c(u-s, z), \\ |D_x^{\alpha-l} c_n^l(s, u, x, x+z)| &\leq \frac{Cn^{|\alpha-l|-\gamma}}{(u-s)} p_c(u-s, z). \end{aligned}$$

From (4.22) and (4.23) it thus remains to control the terms  $D_x^{\alpha-l} \Phi_n(u, t, z+x, y), D_x^{\alpha-l} (\Phi_n(u, t, x+z, y) - \Phi_n(u, t, x, y))$  which are the most singular ones in  $D_x^\alpha R_n^\varepsilon(s, t, x, y)$ . To this end, we will establish by induction that the following

control holds:

$$(4.24) \quad \exists c, C, \forall 0 \leq s < t \leq T, (x, y) \in (\mathbb{R}^d)^2, \forall \beta, |\beta| \leq 5, |D_x^\beta H_n^{(i)}(s, t, x, y)| \leq \frac{C^i n^{|\beta|}}{(t-s)^{|\beta|/2}} (t-s)^{-1+i\gamma/2} \prod_{j=1}^{i-1} B(\gamma/2, j\gamma/2) p_c(t-s, y-x).$$

Observe first that for  $|\beta| = 0$  (no derivation), estimate (4.24) readily follows from (2.9). Let us now suppose  $|\beta| > 0$ . Observe from the definition of  $H_n$  that (4.24) is satisfied for  $i = 1$ . Let us assume it holds for a given  $i$  and let us prove it for  $i + 1$ . Write again:

$$D_x^\beta H_n^{(i+1)}(s, t, x, y) = \int_{(s+t)/2}^t du \int_{\mathbb{R}^d} D_x^\beta H_n(s, u, x, z) H_n^{(i)}(u, t, z, y) dz + D_x^\beta \int_s^{(s+t)/2} du \int_{\mathbb{R}^d} H_n(s, u, x, x+z) H_n^{(i)}(u, t, x+z, y) dz =: (R_1^{i,\beta} + R_2^{i,\beta})(s, t, x, y).$$

The term  $R_1^{i,\beta}$  is easily controlled by (4.24) for  $\beta = 0$  and the induction hypothesis. Observe also that, similarly to (4.23), one has:

$$|D_x^\beta H_n(s, u, x, x+z)| \leq \frac{C n^{|\beta|}}{(u-s)^{1-\gamma/2}} p_c(u-s, z).$$

Together with the induction hypothesis and Leibniz's rule for differentiation, this allows to control  $R_2^{i,\beta}$ . The controls on  $\{R_j^{i,\beta}\}_{j \in \{1,2\}}$  give (4.24) for  $i + 1$ . We eventually derive (reminding that  $|l| = 2$ ):

$$(4.25) \quad |D_x^{\alpha-l} \Phi_n(u, t, x+z, y)| \leq \frac{C}{(t-u)^{(|\alpha|-2)/2}} \frac{n^{|\alpha|-2}}{(t-u)^{1-\gamma/2}} p_c(t-u, y-(x+z)).$$

The spatial Hölder continuity of the derivatives of the kernel  $\Phi_n$  could be checked following the previous steps performed respectively to get the spatial Hölder continuity of the kernel and the controls on its derivatives. One gets, on  $|z| \leq c(t-u)^{1/2}$ :

$$\begin{aligned} & |D_x^{\alpha-l} \Phi_n(u, t, x+z, y) - D_x^{\alpha-l} \Phi_n(u, t, x, y)| \\ & \leq \frac{C |z|^{\gamma/2}}{(t-u)^{(|\alpha|-2)/2}} \frac{n^{|\alpha|-2}}{(t-u)^{1-\gamma/4}} p_c(t-u, y-(x+z)), \end{aligned}$$

which together with (4.25), (4.23), (4.22) gives (proceeding as above for  $|z| \geq c(t-u)^{1/2}$ ):

$$|D_x^\alpha R_n^\varepsilon(s, t, x, y)| \leq \frac{C n^{|\alpha|-2} (t-s)^{\gamma/4}}{(t-s)^{|\alpha|/2}} p_c(t-s, y-x).$$

The second equation of (4.9) follows for  $\bar{p}_n = p_n$  from the above control and (4.13), (4.12). Observe that the control for the derivative w.r.t.  $y$  has additional singularity in  $n$ . This is clear since we directly differentiate the frozen mollified coefficients. Now the statements readily hold for  $p_n^d$ , since the *integration in time* played no role in the previous computations. For  $p_n^h$ , the only point that should be totally justified is the smoothing property and Hölder continuity of the discrete Kernel  $\Phi_n^h(t_i, t_j, x, y) := \sum_{r=1}^{j-i} H_n^{h,(r)}(t_i, t_j, x, y)$ . The smoothing property, equivalent of (4.24), has been investigated in [LM10]. The spatial Hölder continuity can be derived as above.  $\square$

4.4.2. *Proof of Proposition 4.* Write similarly to the proof of Theorem 2.1 in [KM02]:

$$(4.26) \quad \begin{aligned} (p_n - p_n^d)(t_i, t_j, x, y) &= (p_n \otimes H_n - p_n \otimes_h H_n)(t_i, t_j, x, y) + (p_n - p_n^d) \otimes_h H_n(t_i, t_j, x, y) \\ &= \sum_{r \geq 0} (p_n \otimes H_n - p_n \otimes_h H_n) \otimes_h H_n^{(r)}(t_i, t_j, x, y), \end{aligned}$$

where we apply iteratively the first equality to get the second one. From (4.19), the key point is thus to control  $p_n \otimes H_n - p_n \otimes_h H_n$ . Write:

$$(4.27) \quad \begin{aligned} &(p_n \otimes H_n - p_n \otimes_h H_n)(t_i, t_j, x, y) \\ &= \sum_{k=0}^{j-i-1} \int_{t_i+kh}^{t_i+(k+1)h} du \int_{\mathbb{R}^d} dz \{ p_n(t_i, u, x, z) H_n(u, t_j, z, y) \\ &\quad - p_n(t_i, t_{i+k}, x, z) H_n(t_{i+k}, t_j, z, y) \} \\ &= \sum_{k=0}^{j-i-1} \left\{ \int_{t_i+kh}^{t_i+(k+1)h} du \int_{\mathbb{R}^d} dz \{ [p_n(t_i, u, x, z) - p_n(t_i, t_{i+k}, x, z)] H_n(u, t_j, z, y) \} \right. \\ &\quad \left. + \int_{t_i+kh}^{t_i+(k+1)h} du \int_{\mathbb{R}^d} dz \{ p_n(t_i, t_{i+k}, x, z) [H_n(u, t_j, z, y) - H_n(t_{i+k}, t_j, z, y)] \} \right\} \\ (4.28) \quad &=: (D_n^{d,1} + D_n^{d,2})(t_i, t_j, x, y). \end{aligned}$$

For the term  $D_n^{d,1}$  we first write:

$$\begin{aligned} p_n(t_i, u, x, z) - p_n(t_i, t_{i+k}, x, z) &= (u - t_{i+k}) \int_0^1 \partial_u p_n(t_i, t_{i+k} + \lambda(u - t_{i+k}), x, z) d\lambda \\ &= (u - t_{i+k}) \int_0^1 L_{u+\lambda(u-t_{i+k})}^* p_n(t_i, u + \lambda(t_{i+k} - u), x, z) d\lambda. \end{aligned}$$

Reproducing the integration by parts strategy that led to (4.11), we then derive:

$$(4.29) \quad |D_n^{d,1}|(t_i, t_j, x, y) \leq C \frac{hn^{2-\gamma}}{(t_j - t_i)^{1-\gamma/2}} p_c(t_j - t_i, y - x).$$

On the other hand introduce:

$$\begin{aligned} &(\bar{D}_n^{d,21} + \bar{D}_n^{d,22})(t_i, t_j, t_{i+k}, u, x, y) := \\ &C \int_{\mathbb{R}^d} p_c(t_{i+k} - t_i, z - x) |a_n(u, z) - a_n(u, y) - (a_n(t_{i+k}, z) - a_n(t_{i+k}, y))| \\ &\quad \times \frac{1}{t_j - t_{i+k}} p_c(t_j - t_{i+k}, y - z) dz \\ &+ \left| \int_{\mathbb{R}^d} p_n(t_i, t_{i+k}, x, z) (a_n(u, z) - a_n(u, y)) [D_z^2 \tilde{p}_n(u, t_j, z, y) - D_z^2 \tilde{p}_n(t_{i+k}, t_j, z, y)] \right|, \end{aligned}$$

that correspond to the most singular contributions in  $D_n^{d,2}$ . For  $\bar{D}_n^{d,22}$  we can again perform Taylor expansion in time, use the Kolmogorov equations and integrate by parts as above to derive:

$$(4.30) \quad \sum_{k=0}^{j-i-1} \int_{t_i+kh}^{t_i+(k+1)h} du |\bar{D}_n^{d,22}|(t_i, t_j, t_{i+k}, u, x, y) \leq C \frac{hn^{2-\gamma}}{(t_j - t_i)^{1-\gamma/2}} p_c(t_j - t_i, y - x).$$

On the other hand, using the  $\gamma/2$ -Hölder continuity in time, we get:

$$\begin{aligned} & |\bar{D}_n^{d,21}(t_i, t_j, t_{i+k}, u, x, y)| \\ & \leq C \int_{\mathbb{R}^d} p_c(t_{i+k} - t_i, z - x) |u - t_{i+k}|^{\gamma/2} \frac{1}{t_j - t_{i+k}} p_c(t_j - t_{i+k}, y - z) dz \\ & \leq Ch^{(\gamma-\varepsilon_n)/2} p_c(t_j - t_i, y - x) (t_j - t_{i+k})^{-1+\varepsilon_n/2}. \end{aligned}$$

Plugging now the above control, (4.30), (4.29) in (4.28) we derive the result from (4.19) and (4.26).

#### ACKNOWLEDGMENTS

The article was prepared within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

#### REFERENCES

- [Aro59] D. G. Aronson. The fundamental solution of a linear parabolic equation containing a small parameter. *Ill. Journ. Math.*, 3:580–619, 1959.
- [Bil99] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [BP09] 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*, pages 47–53, 2009.
- [BR76] R. Bhattacharya and R. Rao. *Normal approximations and asymptotic expansions*. Wiley and sons, 1976.
- [BT96a] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations: I. Convergence rate of the distribution function. *Prob. Th. Rel. Fields*, 104-1:43–60, 1996.
- [BT96b] 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.*, 2:93–128, 1996.
- [CL04] 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.
- [DM10] F. Delarue and S. Menozzi. Density estimates for a random noise propagating through a chain of differential equations. *Journal of Functional Analysis*, 259-6:1577–1630, 2010.
- [Dyn65] E. B Dynkin. *Markov Processes*. Springer Verlag, 1965.
- [Fri64] A. Friedman. *Partial Differential Equations of Parabolic Type*. Prentice-Hall, 1964.
- [GL08] E. Gobet and C. Labart. Short time asymptotics of the density of the Euler scheme. *Electronic Communications in Probability*, 23, 2008.
- [IKO62] A. M. Il'in, A. S. Kalashnikov, and O. A. Oleinik. Second-order linear equations of parabolic type. *Uspehi Mat. Nauk*, 17-3(105):3–146, 1962.
- [KHLY15] A. Kohatsu-Higa, A. Lejay, and K. Yasuda. Weak Approximation Errors for Stochastic Differential Equations with Non-Regular Drifts. *hal-00840211*, 2015.
- [KM00] V. Konakov and E. Mammen. Local limit theorems for transition densities of Markov chains converging to diffusions. *Prob. Th. Rel. Fields*, 117:551–587, 2000.
- [KM02] V. Konakov and E. Mammen. Edgeworth type expansions for Euler schemes for stochastic differential equations. *Monte Carlo Methods Appl.*, 8-3:271–285, 2002.
- [Kol00] V. Kolokoltsov. Symmetric stable laws and stable-like jump diffusions. *Proc. London Math. Soc.*, 80:725–768, 2000.
- [Kry96] N. V. Krylov. *Lectures on elliptic and parabolic equations in Hölder spaces*. Graduate Studies in Mathematics 12. AMS, 1996.
- [LM10] V. Lemaire and S. Menozzi. On some non asymptotic bounds for the Euler scheme. *Electronic Journal of Probability*, 15:1645–1681, 2010.

- [Men11] S. Menozzi. Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electronic Communications in Probability*, 17:234–250, 2011.
- [Mik12] R. Mikulevičius. On the rate of convergence of simple and jump-adapted weak Euler schemes for Lévy driven SDEs. *Stochastic Process. Appl.*, 122(7):2730–2757, 2012.
- [MP91] R. Mikulevičius and E. Platen. Rate of convergence of the Euler approximation for diffusion processes. *Math. Nachr.*, 151:233–239, 1991.
- [MS67] H. P. McKean and I. M. Singer. Curvature and the eigenvalues of the Laplacian. *J. Differential Geometry*, 1:43–69, 1967.
- [MZ15] R. Mikulevičius and C. Zhang. Weak Euler approximation for Itô diffusion and jump processes. *Stoch. Anal. Appl.*, 33(3):549–571, 2015.
- [She91] S. J. Sheu. Some estimates of the transition density of a nondegenerate diffusion Markov process. *Ann. Probab.*, 19–2:538–561, 1991.
- [Shi96] A.N. Shiryaev. *Probability, Second Edition*. Graduate Texts in Mathematics, 95. Springer-Verlag, New York., 1996.
- [TT90] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch. Anal. and App.*, 8-4:94–120, 1990.

HIGHER SCHOOL OF ECONOMICS, SHABOLOVKA 31, MOSCOW, RUSSIAN FEDERATION.  
*E-mail address:* [VKonakov@hse.ru](mailto:VKonakov@hse.ru)

HIGHER SCHOOL OF ECONOMICS, SHABOLOVKA 31, MOSCOW, RUSSIAN FEDERATION.  
*E-mail address:* [aakozhina@hse.ru](mailto:aakozhina@hse.ru)

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.

*E-mail address:* [stephane.menozzi@univ-evry.fr](mailto:stephane.menozzi@univ-evry.fr)