Weak error for the Euler scheme approximation of diffusions with non-smooth coefficients[*]

Valentin Konakov[†]   Stéphane Menozzi[‡]

Abstract

We study the weak error associated with the Euler scheme of non degenerate diffusion processes with non smooth bounded coefficients. Namely, we consider the cases of Hölder continuous coefficients as well as piecewise smooth drifts with smooth diffusion matrices.

Keywords: diffusion processes; Euler scheme; parametrix; Hölder coefficients; Piecewise smooth bounded drifts.
AMS MSC 2010: Primary 60H10, Secondary 65C30.
Submitted to EJP on April 4, 2016, final version accepted on March 28, 2017.

1 Introduction

1.1 Setting

Let $T > 0$ be a fixed given deterministic final horizon and $x \in \mathbb{R}^d$ be an initial starting point. We consider the following multidimensional SDE:

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \ t \in [0, T], \tag{1.1}$$

where the coefficients $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ are bounded measurable in time and space and W is a Brownian motion on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. We assume that the diffusion matrix $a(t, x) := \sigma\sigma^*(t, x)$ is uniformly elliptic and at least Hölder continuous in time and space. We will consider two kinds of assumptions for the drift coefficient b: either Hölder continuous in time and space (as for the diffusion matrix), or piecewise smooth and having at most a finite set of spatial discontinuities. These 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.

Define now for a given $N \in \mathbb{N}^*$ the time step $h := T/N$ and set for all $i \in [\![1, N]\!]$, $t_i := ih$ where from now on the notation $[\![\cdot, \cdot]\!]$ is used to denote an interval of integers. Consider the continuous Euler scheme associated with (1.1) whose dynamics writes $X_0^h = x$ and for all $t \in [0, T]$:

$$X_t^h = x + \int_0^t b(\phi(u), X_{\phi(u)}^h)du + \int_0^t \sigma(\phi(u), X_{\phi(u)}^h)dW_u, \qquad (1.2)$$

where we set $\phi(u) = \inf\{(t_i)_{i \in [\![0, N-1]\!]} : t_i \leq u < t_{i+1}\}$.

A useful quantity to study, arising in many applicative fields from physics to finance, is the so-called weak error which for a suitable real valued test function f writes:

$$d(f, x, T, h) := \mathbb{E}[f(X_T^{h,0,x})] - \mathbb{E}[f(X_T^{0,x})], \qquad (1.3)$$

using the usual Markovian notations, i.e. $X_T^{h,0,x}, X_T^{0,x}$ respectively stand for the Euler scheme and the diffusion at time T which start at point x at 0.

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]. Under those conditions, the quantity $d(f, x, T, h)$ is of order h corresponding to the magnitude of the time step. In the non degenerate framework (under some uniform ellipticity or hypoellipticity conditions) it is even possible to take f to be a Dirac mass in the above expression (1.3). The associated convergence rate remains of order h for the Euler scheme, see [KM02] [BT96b] and $h^{1/2}$ in the more general case of Markov Chain approximations, see e.g [KM00] in which the Brownian increments appearing in (1.2) are replaced by i.i.d. sequences $(\xi_i)_{i \geq 1}$ that are not necessarily Gaussian. In the framework of Lipschitz coefficients we can also mention, in the scalar case, the recent work of Alfonsi *et al.* [AJKH14], who obtained bounds on the Wasserstein distances between the laws of the paths of the diffusion and its Euler scheme. Anyhow, the case of non smooth coefficients, Hölder continuous or less, has rarely been considered. Such cases might anyhow appear very naturally in many applications, when the drifts have for instance discontinuities at some given interfaces or when the diffusion coefficients are very irregular (random media).

In the framework of bounded non degenerate and Hölder continuous coefficients, let us mention the work of Mikulevičius and Platen [MP91] who obtained bounds for the weak error in (1.3) at rate $h^{\gamma/2}$ where $\gamma \in (0, 1)$ is the Hölder exponent of the coefficients b, σ in (1.1) provided $f \in C_b^{2+\gamma}(\mathbb{R}^d, \mathbb{R})$ (space of bounded functions with bounded derivatives up to order two and γ-Hölder continuous second derivatives). This regularity is essential in that work to apply Itô's formula. Our approach permits to establish that this bound holds true, up to an additional slowly varying factor in the exponent, for the difference of the densities itself, which again corresponds to the weak error (1.3) for a δ-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.* [KHLY15] who investigate the weak error for possibly discontinuous drifts and diffusion coefficients that are just continuous. We are able to extend some of their controls to densities. Indeed, in the quoted work, the authors investigate (1.3) for functions f that are at least continuous. We again have an additional slowly varying factor in the exponent which is due to our smoothing approach.

Our strategy is the following. Under the previous assumptions (stated after (1.1)), both processes $(X_t)_{t \in (0,T]}$ in (1.1) and $(X_{t_i}^h)_{i \in [\![1,N]\!]}$ in (1.2) have densities, see e.g. [KKM16] for the continuous process and Lemaire and Menozzi [LM10] for the scheme. Let us denote them respectively for $x \in \mathbb{R}^d$, $0 \le i < j \le N$, by $p(t_i, t_j, x, .)$ and $p^h(t_i, t_j, x, .)$ for the processes starting at time t_i from point x and considered at time t_j. To study the error $(p - p^h)(t_i, t_j, x, y)$ we introduce perturbed dynamics associated with (1.1) and (1.2) respectively. Namely, for a small parameter ε, we mollify suitably the coefficients, the mollification procedure is described in its whole generality in Section 2 and depends on the two considered sets of assumptions indicated above, and consider two additional processes with dynamics:

$$X_t^{(\varepsilon)} = x + \int_0^t b_\varepsilon(s, X_s^{(\varepsilon)})ds + \int_0^t \sigma_\varepsilon(s, X_s^{(\varepsilon)})dW_s, \tag{1.4}$$
$$X_0^{h,(\varepsilon)} = x, \ X_{t_{i+1}}^{h,(\varepsilon)} = X_{t_i}^{h,(\varepsilon)} + b_\varepsilon(t_i, X_{t_i}^{h,(\varepsilon)})h + \sigma_\varepsilon(t_i, X_{t_i}^{h,(\varepsilon)})(W_{t_{i+1}} - W_{t_i}),$$

where $b_\varepsilon, \sigma_\varepsilon$ are mollified versions of b, σ. It is clear that both processes $(X_t^{(\varepsilon)})_{t \in (0,T]}$ and $(X_{t_i}^{h,(\varepsilon)})_{i \in [\![1,N]\!]}$ have densities. The mollified coefficients indeed satisfy uniformly in the mollification parameter the previous assumptions. Let us denote those densities for $x \in \mathbb{R}^d, 0 \le t_i < t_j \le T$ by $p_\varepsilon(t_i, t_j, x, .)$, $p_\varepsilon^h(t_i, t_j, x, .)$ respectively.

The idea is now to decompose the global error as:

$$(p - p^h)(t_i, t_j, x, y) = (p - p_\varepsilon)(t_i, t_j, x, y) + (p_\varepsilon - p_\varepsilon^h)(t_i, t_j, x, y) + (p_\varepsilon^h - p^h)(t_i, t_j, x, y). \tag{1.5}$$

The key point is that the stability of the densities with respect to a perturbation has been thoroughly investigated for diffusions and Markov Chains in Konakov *et al.* [KKM16]. The results of that work allow to control the differences $p - p_\varepsilon$, $p_\varepsilon^h - p^h$. On the other hand, since the coefficients $b_\varepsilon, \sigma_\varepsilon$ of $(X_t^{(\varepsilon)})_{t \in [0,T]}$, $(X_{t_i}^{h,(\varepsilon)})_{i \in [\![0,N]\!]}$ are smooth, the central term $p_\varepsilon - p_\varepsilon^h$ in (1.5) can be investigated thanks to the work of Konakov and Mammen [KM02] giving the error expansion at order h on the densities for the weak error. The key point is that the coefficients in the expansion depend on the derivatives of $b_\varepsilon, \sigma_\varepsilon$ which explode when ε goes to zero. This last condition is natural in order to control $p - p_\varepsilon, p_\varepsilon^h - p^h$. Thus, two contributions need to be equilibrated to derive the global error bounds. This will be done through a careful analysis of the densities (heat kernel) of the processes with dynamics described in (1.1), (1.2), (1.4). The estimates required for the error analysis will lead us to refine some bounds previously established by Il'in *et al* [IKO62]. Let us indicate that this perturbative approach had also been considered by Kohatsu-Higa *et al.* [KHLY15] but for the weak error (1.3) involving at least a continuous function. Our approach, based on parametrix techniques, allows to handle directly the difference of the densities, and gives, up to an additional factor going to zero with the time step, the *expected* convergence rates.

1.2 Assumptions and main results

Let us introduce the following assumptions.

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

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |b(t,x)| \le K_1, \quad \sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |\sigma(t,x)| \le K_2.$$

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

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

We consider two types of *smoothness* assumptions for the coefficients b, σ in (1.1).

(H) (Hölder drift and diffusion coefficient). The drift b and the diffusion coefficient σ are time-space Hölder continuous in the following sense: for some $\gamma \in (0, 1]$, $\kappa < +\infty$, for all $(s, t) \in [0, T]^2, (x, y) \in (\mathbb{R}^d)^2$,

$$|\sigma(s, x) - \sigma(t, y)| + |b(s, x) - b(t, y)| \leq \kappa\{|s - t|^{\gamma/2} + |x - y|^{\gamma}\}.$$

Observe that the last condition also readily gives, thanks to the boundedness of σ, that the diffusion matrix $a = \sigma\sigma^*$ enjoys the same Hölder property.

(PS) (Piecewise smooth drift and Smooth diffusion coefficient). The drift b is piecewise smooth with bounded derivatives outside of the discontinuity sets. Precisely, $b \in C_b^{2,4}([0, T] \times (\mathbb{R}^d \backslash \mathcal{I}), \mathbb{R}^d)$ where the set of possible discontinuities $\mathcal{I}$ writes as $\mathcal{I} := \cup_{i=1}^m \mathcal{S}_i$, $m \in \mathbb{N}$. Here, for all $i \in [\![1, m]\!]$, $\mathcal{S}_i$ is a smooth bounded submanifold of $\mathbb{R}^d$ (at least C^4) of dimension lower or equal to $d - 1$, i.e. $\mathcal{S}_i := \{x \in \mathbb{R}^d : g_i(x) = 0\}$ for a corresponding smooth function g_i. We also assume that the $(\mathcal{S}_i)_{i \in [\![1, m]\!]}$ do not intersect: for all $1 \leq i < j \leq m$, $\mathcal{S}_j \cap \mathcal{S}_i = \emptyset$.

On the other hand we assume that the diffusion coefficient σ is **globally** $C_b^{2,4}([0, T] \times \mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)$.

We emphasize that, with the above definition, the discontinuity set of b only depends on the spatial variable. A time-dependent discontinuity set could *a priori* also be considered provided each of its components is the boundary of a smooth time-space domain. Namely, considering for $i \in [\![1, m]\!], t \in [0, T], \mathcal{S}_i(t) := \{x \in \mathbb{R}^d : g_i(t, x) = 0\}$, the smooth spatial submanifolds $\mathcal{S}_i(t)$ should as well evolve smoothly in time. We consider the case introduced in **(PS)** for simplicity.

From now on, we always assume conditions **(A1)-(A2)** to be in force. We say that assumption **($\mathbf{A}_H$)** (resp. **($\mathbf{A}_{PS}$)**) holds if additionally the coefficients satisfy **(H)** (resp. **(PS)**). We will write that **(A)** holds whenever **($\mathbf{A}_H$)** or **($\mathbf{A}_{PS}$)** is satisfied.

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. Other possible dependencies will be explicitly specified.

Theorem 1.1 (Error for the Euler scheme of a diffusion with Hölder coefficients). *Let $T > 0$ be fixed and consider a given time step $h := T/N$, for $N \in \mathbb{N}^*$. Set for $i \in \mathbb{N}$, $t_i := ih$. Under **($\mathbf{A}_H$)**, there exist $C \geq 1, c \in (0, 1]$ s.t. for all $0 \leq t_i < t_j \leq T$ s.t. $(t_j - t_i) \geq h^{1/(2-\gamma)}$ and $(x, y) \in (\mathbb{R}^d)^2$:*

$$p_c(t_j - t_i, y - x)^{-1}|(p - p^h)(t_i, t_j, x, y)| \leq \frac{C}{(t_j - t_i)^{(1-\gamma)\gamma/2}} h^{\frac{\gamma}{2} - C\psi(h)}, \tag{1.6}$$

where p, p^h respectively stand for the densities of the diffusion X and its Euler approximation X^h with time step h, for all $(t, z) \in \mathbb{R}_*^+ \times \mathbb{R}^d$, $p_c(t, z) := \frac{c^{d/2}}{(2\pi t)^{d/2}} \exp(-c\frac{|z|^2}{2t})$ and $\psi(h) = \frac{\log_3(h^{-1})}{\log_2(h^{-1})}$ where $\log_k$ denotes for $k \in \mathbb{N}^*$ the kth iterated logarithm. Let us observe that $\psi(h) \underset{h \to 0}{\longrightarrow} 0$. If we are now interested in the weak error in the sense of (1.3), for a function $f \in C^\beta(\mathbb{R}^d, \mathbb{R})$ (uniformly β-Hölder continuous functions), $\beta \in (0, 1]$:

$$|\mathbb{E}[f(X_{t_j}^{h, t_i, x})] - \mathbb{E}[f(X_{t_j}^{t_i, x})]| \leq C_f h^{\gamma/2}, \tag{1.7}$$

using again the usual Markovian notations, i.e. $X_{t_j}^{h,t_i,x}, X_{t_j}^{t_i,x}$ respectively stand for the Euler scheme and the diffusion at time t_j which start at point x at t_i.

Eventually, if we consider a smooth domain $A \subset \mathbb{R}^d$ (i.e. a connected open set at least C^2) with bounded boundary and non zero Lebesgue measure, we also get that for all $x \in \mathbb{R}^d$ s.t. $d(x, \partial A) \geq (t_j - t_i)^{1/2} h^{\gamma/2}$:

$$|\mathbb{E}[\mathbb{I}_{X_{t_j}^{h,t_i,x} \in A}] - \mathbb{E}[\mathbb{I}_{X_{t_j}^{t_i,x} \in A}]|$$

$$\leq C\Big\{ \frac{1}{\gamma d(x, \partial A)^\gamma} \mathbb{I}_{d(x,\partial A) \geq \exp(-\frac{1}{\gamma})} + |\ln(d(x, \partial A))|\mathbb{I}_{d(x,\partial A) < \exp(-\frac{1}{\gamma})} + 1 \Big\} h^{\gamma/2}, \tag{1.8}$$

where $d(., \partial A)$ stands for the distance to the boundary of A.

Remark 1.2. We point out that this result is to be compared with the one obtained by Mikulevičius and Platen [MP91] 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 $d(f, x, T, h) = \mathbb{E}[f(X_T^{h,0,x})] - \mathbb{E}[f(X_T^{0,x})]$ for a *smooth* function $f \in C_b^{2+\gamma}(\mathbb{R}^d, \mathbb{R})$ (space of bounded functions, with bounded derivatives up to order two and γ-Hölder continuous second derivatives). The above theorem establishes that $|d(f, x, T, h)| \leq Ch^{\gamma/2 - C\psi(h)}$ as soon as f is measurable and satisfies the growth condition

$$\exists c_0 < c/(2T), C_0 > 0, \forall x \in \mathbb{R}^d, \ |f(x)| \leq C_0 \exp(c_0|x|^2). \tag{1.9}$$

This control can be useful for specific and relevant applications, like for instance quantile estimation (that would involve functions of the form $f(x) = \mathbb{I}_{|x| \leq K}$ or $f(x) = \mathbb{I}_{|x| \leq K} \exp(c|x|)$) that appear in many applications: default probabilities in mathematical finance, fatigue of structures in random mechanics. We are able to find the expected convergence rate up to a vanishing contribution. The rate $h^{\gamma/2}$ again holds, without the additional term, as soon as $f \in C^\beta(\mathbb{R}^d, \mathbb{R}), \ \beta \in (0, 1]$. Some extensions to unbounded functions f satisfying the growth condition (1.9) are described in Remark 3.2 of Section 3.3.

The contribution in $\psi(h)$ appearing in (1.6), which slightly deteriorates the convergence, seems to be, with our approach, the price to pay to get rid of any smoothness on f. Observe anyhow that for indicator functions of smooth Borel sets, equation (1.8) provides a better result than (1.6) as soon as the initial distance to the boundary satisfies $d(x, \partial A) \geq (t_j - t_i)^{1/2} h^{\gamma/2}$ (see Section 3.3.2 for details). Observe that this control improves in that case what could be derived from [KHLY15] in which continuous test functions are considered.

Remark 1.3 (About the Convergence Rate)**.** We also emphasize that the convergence rate in $h^{\gamma/2}$ is closer to a rate associated with a *strong error*. It indeed 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, under $(\mathbf{A}_H)$, for all $i \in [\![0, N-1]\!]$:

$$\mathbb{E}[\sup_{u \in [t_i, t_{i+1}]} |b(u, X_u^h) - b(t_i, X_{t_i}^h)|] + \mathbb{E}[\sup_{u \in [t_i, t_{i+1}]} |\sigma(u, X_u^h) - \sigma(t_i, X_{t_i}^h)|]$$

$$\leq \kappa \left\{ h^{\gamma/2} + \mathbb{E}[\sup_{u \in [t_i, t_{i+1}]} |X_u^h - X_{t_i}^h|^\gamma] \right\}$$

$$\leq \kappa \left\{ h^{\gamma/2} + \mathbb{E}[\{\sup_{u \in [t_i, t_{i+1}]} |\sigma(t_i, X_{t_i}^h)(W_u - W_{t_i})| + K_1 h\}^2]^{\gamma/2} \right\} \leq ch^{\gamma/2}. \tag{1.10}$$

These terms typically appear in the error analysis when there is low regularity of the coefficients or of the value function $v(t, x) := \mathbb{E}[f(X_T^{t,x})]$. Under the previous assumptions,

if the function f belongs to $C_b^{2+\gamma}(\mathbb{R}^d, \mathbb{R})$, $\gamma \in (0,1)$ it is then well known, see e.g. Friedman [Fri64] or Ladyzhenskaya *et al.* [LSU68] that $v \in C_b^{1+\gamma/2, 2+\gamma}([0,T] \times \mathbb{R}^d, \mathbb{R})$. Also v satisfies the parabolic PDE $(\partial_t v + L_t v)(t,x) = 0$, $(t,x) \in [0,T] \times \mathbb{R}^d$, where L_t stands for the generator of (1.1) at time t, i.e. for all $\varphi \in C_0^2(\mathbb{R}^d, \mathbb{R}), x \in \mathbb{R}^d$,

$$L_t \varphi(x) = b(t,x) \cdot \nabla_x \varphi(x) + \frac{1}{2} \mathrm{Tr}(a(t,x) D_x^2 \varphi(x)).$$

Recalling that $t_0 = 0, t_N = T$, we decompose the error as:

$$d(f,x,T,h) := \mathbb{E}[f(X_T^{h,0,x})] - \mathbb{E}[f(X_T^{0,x})] = \sum_{i=0}^{N-1} \mathbb{E}[v(t_{i+1}, X_{t_{i+1}}^{h,0,x}) - v(t_i, X_{t_i}^{h,0,x})]$$

$$= \sum_{i=0}^{N-1} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} \Big\{ \partial_s v(s, X_s^{h,0,x}) + \nabla_x v(s, X_s^{h,0,x}) \cdot b(t_i, X_{t_i}^{h,0,x})$$

$$+ \frac{1}{2} \mathrm{Tr}(D_x^2 v(s, X_s^{h,0,x}) a(t_i, X_{t_i}^{h,0,x})) \Big\} ds \Big]$$

$$= \sum_{i=0}^{N-1} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} \Big\{ \partial_s v + L_s v \Big\}(s, X_s^{h,0,x}) ds \Big] \qquad (1.11)$$

$$+ \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} \Big\{ \nabla_x v(s, X_s^{h,0,x}) \cdot (b(t_i, X_{t_i}^{h,0,x}) - b(s, X_s^{h,0,x}))$$

$$+ \frac{1}{2} \mathrm{Tr}(D_x^2 v(s, X_s^{h,0,x})(a(t_i, X_{t_i}^{h,0,x}) - a(s, X_s^{h,0,x}))) \Big\} ds \Big]$$

$$= \sum_{i=0}^{N-1} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} \Big\{ \nabla_x v(s, X_s^{h,0,x}) \cdot (b(t_i, X_{t_i}^{h,0,x}) - b(s, X_s^{h,0,x}))$$

$$+ \frac{1}{2} \mathrm{Tr}\big(D_x^2 v(s, X_s^{h,0,x})(a(t_i, X_{t_i}^{h,0,x}) - a(s, X_s^{h,0,x}))\big) \Big\} ds \Big],$$

exploiting the PDE satisfied by v for the last equality. For a function f in $C_b^{2+\gamma}(\mathbb{R}^d, \mathbb{R})$, the spatial derivatives of v up to order two are globally bounded on $[0,T]$. Indeed, the classical Schauder estimates hold (see e.g. Theorem 5.2, p. 361 in [LSU68]). We are thus led to control in (1.11) quantities similar to those appearing in (1.10). The associated bound then precisely gives the convergence rate. The analysis extends if f is simply $C^\beta(\mathbb{R}^d, \mathbb{R})$, $\beta \in (0,1]$ and therefore possibly unbounded. In that case the second derivatives yield an integrable singularity in time for the second order partial derivatives. We refer to Proposition 3.4, which holds under the sole assumption $(\mathbf{A}_H)$ for multi-indices α, $|\alpha| \le 2$, and to the proof of Theorem 1.1 in Section 3.4.1. Extensions to locally β-Hölder functions f satisfying the growth condition (1.9) are discussed in Remark 3.2.

Remark 1.4. Even though we have considered $\gamma \in (0,1]$, our analysis should extend to the framework of Hölder spaces to $\gamma \in (1,2)$. On the other hand, Theorem 1.1 specifies the time-singularity in small time.

Remark 1.5. We feel that the bounds of Theorem 1.1 are *relevant* for functions which are **truly** Hölder continuous, that is for coefficients that would involve some simple transformations of the Weierstrass functions, see e.g. [Zyg36], or of an independent Brownian sample path in order that $(\mathbf{A}_H)$ is fulfilled. Indeed, for functions which are just *locally* Hölder continuous, like the mapping $x \mapsto 1 + |x|^\alpha \wedge K$, $\alpha \in (0,1]$, we think that it would be more appropriate to study some local regularizations, close to the neighborhoods of **real** Hölder continuity (0 and $K^{1/\alpha}$ for the indicated example) and to exploit that, outside of these neighborhoods, the **usual** sufficient smoothness is available. For such coefficients we think that the convergence rates might be definitely better.

Theorem 1.6 (Error for the Euler Scheme with Smooth Diffusion Coefficients and Piecewise Smooth Drift)**.** *Let $T > 0$ be fixed and ($\boldsymbol{A}_{PS}$) be in force. With the notations of Theorem 1.1 we have that:*

- *there exist $C \geq 1, c \in (0,1]$ s.t. for all $0 \leq t_i < t_j \leq T$ s.t. $(t_j - t_i) \geq h^{1/2}$ and $(x, y) \in (\mathbb{R}^d)^2$:*

$$p_c(t_j - t_i, y - x)^{-1}|(p - p^h)(t_i, t_j, x, y)| \leq C \left(\frac{h}{(t_j - t_i)^{1/2}} \right)^{1/(2d) - C\psi(h)}. \tag{1.12}$$

- *If $d(y, \mathcal{I})$ (distance of the final point y to the spatial discontinuity set $\mathcal{I}$) satisfies $d(y, \mathcal{I}) \geq h^{1/2 - \epsilon}$ for a fixed given $\epsilon \in (0, 1/2]$, then:*

$$p_c(t_j - t_i, y - x)^{-1}|(p - p^h)(t_i, t_j, x, y)| \leq C \Big[\left(\frac{h}{(t_j - t_i)^{1/2}} \right)^{1/(d+1) - C\psi(h)} + \frac{h^{1 - C\psi(h)}}{d(y, \mathcal{I})} \Big]. \tag{1.13}$$

- *In the special case $\sigma(t, x) = \sigma$, i.e. constant diffusion coefficient[1], the previous bound improves to:*

$$p_c(t_j - t_i, y - x)^{-1}|(p - p^h)(t_i, t_j, x, y)| \leq C \Big[\left(\frac{h}{(t_j - t_i)^{1/2}} \right)^{1/d - C\psi(h)} + \frac{h^{1 - C\psi(h)}}{d(y, \mathcal{I})} \Big]. \tag{1.14}$$

Remark 1.7. This result emphasizes that, as soon as the drift is irregular, a *true* diffusion coefficient deteriorates the convergence rate. This is clear since, in that case, the difference of the densities $p_\varepsilon - p_\varepsilon^h$ in (1.5) involves higher derivatives of densities of processes with mollified coefficients which are more explosive (see Section 3.4).

We also mention that the distance of the final point to the discontinuity set plays an important role. The global control (1.12) improves to (1.13) as soon as $h^{1 - 1/(2d)} \leq C(T)d(y, \mathcal{I})$.

Eventually, if the diffusion coefficient does not depend on space, we find, up to the additional term in $\psi(h)$, the usual convergence rate for the weak error if $d = 1$ as soon as $c_0(t_j - t_i)^{1/2} \leq d(y, \mathcal{I})$ for any given $c_0 > 0$.

However, our regularization approach clearly feels the dimension, when doing e.g. Hölder inequalities on neighborhoods of the discontinuity sets, and the convergence rates decrease with the dimension.

Let us carefully mention that considering the following weak error $d(f, x, t_i, t_j, h) := \mathbb{E}[f(X_{t_j}^{h, t_i, x})] - \mathbb{E}[f(X_{t_j}^{t_i, x})]$ for *smooth* functions f and not Dirac masses as we do, should improve the convergence rates and in particular allow to get rid of the terms in $\psi(h)$ through a careful investigation of the derivatives of the associated heat kernels. We refer to the estimates of Proposition 3.4 that could be refined when considering an additional integration w.r.t. to the final variable.

1.3 On some related applications

1.3.1 Some approximating dynamics for interest rates

A very popular model for interest rates in the financial literature is the Cox-Ingersoll-Ross process with dynamics:

$$dX_t = (a - kX_t)dt + \sigma|X_t|^{1/2}dW_t, \tag{1.15}$$

for given parameters $\sigma, k, a > 0$. From the numerical viewpoint, the behavior of the Euler scheme is not standard. For a given time-step h, the strong error was indeed proved to

---

[1]the case of an inhomogeneous diffusion coefficient independent of x, i.e. $\sigma(t, x) = \sigma(t)$ could also be handled provided the Gaussian part is simulated exactly in a modified Euler scheme.

be, as in the usual Lipschitz case, of order $h^{1/2}$ in Berkaoui *et al.* [BBD08] provided a is not *too small*. On the other hand, numerical experiments in Alfonsi [Alf05] emphasized very slow convergence, of order $(-\ln h)^{-1}$, for small values of a. This convergence order has been established by Gyöngy and Rásonyi [GR11].

Of course the dynamics in (1.15) does not enter our framework, since it is closer to the dynamics of a Bessel-like process whose density does not satisfy Gaussian bounds. However, we could introduce for positive parameters η, K, which are respectively meant to be small and large enough, the dynamics:

$$dX_t = (a - kX_t)dt + (\eta + \sigma|X_t|^{1/2} \wedge K)dW_t. \tag{1.16}$$

The diffusion coefficient $\tilde{\sigma}(x) = (\eta + \sigma|x|^{1/2} \wedge K)$ is then uniformly elliptic, 1/2 Hölder continuous and bounded. On the other hand the drift is not bounded but the analysis of Theorem 1.1 would still hold true thanks to the work of Konakov and Markova [KM15] that allows to get rid of the linear drift through a suitable transformation. We would then derive a convergence of order $h^{1/4-C\psi(h)}$ at least for the associated Euler scheme on the densities (see also Remark 1.5). Even though the marginals in (1.16) enjoy Gaussian bounds, see e.g. [DM10], the expected properties for an interest rate dynamics, mean reverting and positivity, should hold with some high probability. Also, the difference between the approximate dynamics in (1.16) and the original one in (1.15) might be investigated through stochastic analysis tools (occupation times).

1.3.2 Extension to some kinetic models

The results of Theorems 1.1 and 1.6 should extend without additional difficulties to the case of degenerate diffusions of the form:

$$\begin{aligned} dX_t^1 &= b(t, X_t)dt + \sigma(t, X_t)dW_t, \\ dX_t^2 &= X_t^1 dt, \end{aligned} \tag{1.17}$$

denoting $X_t = (X_t^1, X_t^2)$, under the same previous assumptions $(\mathbf{A}_H)$ or $(\mathbf{A}_{PS})$ on b, σ. The sensitivity analysis when we consider perturbations of the non-degenerate components, i.e. for a given $\varepsilon > 0$:

$$\begin{aligned} dX_t^{1,(\varepsilon)} &= b_\varepsilon(t, X_t^{(\varepsilon)})dt + \sigma_\varepsilon(t, X_t^{(\varepsilon)})dW_t, \\ dX_t^{2,(\varepsilon)} &= X_t^{1,(\varepsilon)}dt, \end{aligned} \tag{1.18}$$

has been performed by Kozhina [Koz16] following [KKM16]. The key point is that under **(A)**, the required parametrix expansions of the densities associated with the solutions of equation (1.17), (1.18) were established in [KMM10]. The analysis of the derivatives of the heat kernel, that would require to extend the results of Section 3 to the considered degenerate setting will concern further research.

The paper is organized as follows. We first introduce a suitable mollification procedure of the coefficients in Section 2 and derive from the stability results of Konakov *et al.* [KKM16] how the error of the mollifying procedure is then reflected on the densities. This allows to control the terms $p - p_\varepsilon$ and $p_\varepsilon^h - p^h$ in (1.5). We then give in Section 3 some pointwise bounds on the derivatives of the heat-kernels with mollified coefficients. From these controls and the previous error expansion obtained for the Euler scheme with smooth coefficients by Konakov and Mammen [KM02], we are able to control the remaining term $p_\varepsilon - p_\varepsilon^h$ in (1.5). We then establish our main estimates equilibrating the two errors. Eventually, Section 4 is dedicated to the proof of the controls stated in Section 3. These proofs are based on the parametrix expansions of the underlying densities following the Mc-Kean and Singer approach [MS67].

2 Mollification of the coefficients and stability results

For the error analysis, in order to apply the strategy described in the introduction, we first need to *regularize* in an appropriate manner the coefficients. The mollifying procedures differ under our two sets of assumptions.

2.1 Mollification under ($\mathbf{A}_H$) (Hölder continuous coefficients)

In this case both coefficients b, σ need to be globally regularized in time and space. We introduce the mollified coefficients defined for all $(t, x) \in [0, T] \times \mathbb{R}^d$ and $\varepsilon > 0$ by

$$b_{\varepsilon,S}(t,x) := b(t,\cdot) * \rho_\varepsilon(x), \sigma_{\varepsilon,S}(t,x) := \sigma(t,\cdot) * \rho_\varepsilon(x), \tag{2.1}$$

where $*$ stands for the spatial convolution and for $\varepsilon > 0$, ρ_ε is a spatial mollifier, i.e. for all $x \in \mathbb{R}^d$,

$$\rho_\varepsilon(x) := \varepsilon^{-d}\rho(x/\varepsilon),\ \rho \in C^\infty(\mathbb{R}^d, \mathbb{R}^+), \int_{\mathbb{R}^d} \rho(y)dy = 1, |\mathrm{supp}(\rho)| \subset K,$$

for some compact set $K \subset \mathbb{R}^d$. The subscript S in $b_{\varepsilon,S}, \sigma_{\varepsilon,S}$ appears to emphasize that the *spatial* convolution is considered. We will also need a mollification in time when the coefficients are inhomogeneous. Up to a symmetrization in time of the coefficients b, σ, i.e. we set for all $(t, x) \in [0, T] \times \mathbb{R}^d, b(-t, x) = b(t, x),\ \sigma(-t, x) = \sigma(t, x)$ we can define:

$$b_\varepsilon(t,x) = b_{\varepsilon,S}(.,x) \star \zeta_{\varepsilon^2}(t), \sigma_\varepsilon(t,x) = \sigma_{\varepsilon,S}(.,x) \star \zeta_{\varepsilon^2}(t), \tag{2.2}$$

where $\star$ stands for the time convolution and for $s \in \mathbb{R}$, $\zeta_{\varepsilon^2}(s) := \varepsilon^{-2}\zeta(s/\varepsilon^2)$, ζ being a scalar mollifier with compact support in $[-T, T]$. The *complete* regularization in the spatial and time variable reflects the usual parabolic scaling. This feature will be crucial to balance the singularities appearing in our analysis (see Propositions 3.4, 3.7 and their proofs below). We have the following controls.

Proposition 2.1 (First Controls on the Mollified Coefficients). *Assume that ($\mathbf{A}_H$) is in force. Then, there exists $C \geq 1$ s.t. for all $\varepsilon > 0$,*

$$\Delta_{\varepsilon,b} := \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |b(t,x) - b_\varepsilon(t,x)| \leq C\varepsilon^\gamma,$$

$$\Delta_{\varepsilon,\sigma} := \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\sigma(t,x) - \sigma_\varepsilon(t,x)| \leq C\varepsilon^\gamma, \tag{2.3}$$

$$\forall \eta \in (0, \gamma),\ \Delta_{\varepsilon,\sigma,\eta} := \Delta_{\varepsilon,\sigma} + \sup_{t\in[0,T]} |(\sigma - \sigma_\varepsilon)(t,.)|_\eta \leq C(\varepsilon^\gamma + \varepsilon^{\gamma-\eta}),$$

where for a given function $f : \mathbb{R}^d \to \mathbb{R}$, we denote for $\eta \in (0, 1)$, $|f|_\eta := \sup_{\substack{(x,\ y)\ \in\ (\mathbb{R}^d)^2, \\ x\ \neq\ y}} \frac{|f(x)-f(y)|}{|x-y|^\eta}$.

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

$$b(t,x) - b_{\varepsilon,S}(t,x) := \int_{\mathbb{R}^d} \{b(t,x) - b(t,y)\}\rho_\varepsilon(x-y)dy = \int_{\mathbb{R}^d} \{b(t,x) - b(t, x - z\varepsilon)\}\rho(z)dz.$$

From the Hölder continuity of b assumed in **(H)** and the above equation, we deduce that $b_{\varepsilon,S}$ satisfies **(H)** as well and that:

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |(b - b_{\varepsilon,S})(t,x)| \leq C_\rho\varepsilon^\gamma,\ C_\rho := \kappa \int_K |z|^\gamma\rho(z)dz. \tag{2.4}$$

The same analysis can be performed for the term $\sigma_{\varepsilon,S}$, so that $\sigma_{\varepsilon,S}$ satisfies **(H)** and $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |(\sigma - \sigma_{\varepsilon,S})(t,x)| \leq C_\rho\varepsilon^\gamma$. From **(H)**, we also have that $b_{\varepsilon,S}, \sigma_{\varepsilon,S}$ are both $\gamma/2$-Hölder continuous in time uniformly in $\varepsilon > 0$. Repeating the previous arguments replacing ρ_ε by ζ_{ε^2}, we deduce $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |(b_{\varepsilon,S} - b_\varepsilon)(t,x)| + \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |(\sigma_{\varepsilon,S} - \sigma_\varepsilon)(t,x)| \leq C_\zeta\varepsilon^\gamma$, which eventually yields:

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|(b-b_\varepsilon)(t,x)| + \sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|(\sigma-\sigma_\varepsilon)(t,x)| \le C\varepsilon^\gamma.$$

This gives the controls concerning the sup norms in (2.3).

Let us now turn to the Hölder norm. Observe first that, for all $t\in\mathbb{R}^+$, $(x,y)\in(\mathbb{R}^d)^2$:

$$\{\sigma(t,x)-\sigma_{\varepsilon,S}(t,x)\} - \{\sigma(t,y)-\sigma_{\varepsilon,S}(t,y)\}$$

$$= \int_{\mathbb{R}^d}\Big\{[\sigma(t,x)-\sigma(t,x-z\varepsilon)]-[\sigma(t,y)-\sigma(t,y-z\varepsilon)]\Big\}\rho(z)dz,$$

$$\{\sigma_\varepsilon(t,x)-\sigma_{\varepsilon,S}(t,x)\} - \{\sigma_\varepsilon(t,y)-\sigma_{\varepsilon,S}(t,y)\}$$

$$= \int_{\mathbb{R}}\Big\{[\sigma_{\varepsilon,S}(t-\varepsilon^2 u,x)-\sigma_{\varepsilon,S}(t,x)]-[\sigma_{\varepsilon,S}(t-\varepsilon^2 u,y)-\sigma_{\varepsilon,S}(t,y)]\Big\}\zeta(u)du.$$

It readily follows from the γ-Hölder continuity in space of σ (resp. the γ-Hölder continuity in space and the $\gamma/2$-Hölder continuity in time of $\sigma_{\varepsilon,S}$) that the following controls hold:

$$|[\sigma(t,x)-\sigma_\varepsilon(t,x)]-[\sigma(t,y)-\sigma_\varepsilon(t,y)]| \le C(|x-y|^\gamma \wedge \varepsilon^\gamma) \le C|x-y|^\eta\varepsilon^{\gamma-\eta},$$

$$|(\sigma-\sigma_\varepsilon)(t,.)|_\eta \le C\varepsilon^{\gamma-\eta},\ \eta\in(0,\gamma). \tag{2.5}$$

This completes the proof. $\qquad\square$

We will need as well some controls on the derivatives of the mollified coefficients.

Proposition 2.2 (Controls on the Derivatives of the Mollified Coefficients). *Under the assumptions of Proposition 2.1, we have that there exists $C\ge 1$ s.t. for all $\varepsilon\in(0,1)$ and for all multi-index α, $|\alpha|\in[\![1,4]\!]$:*

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|D_x^\alpha b_\varepsilon(t,x)| + \sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|D_x^\alpha\sigma_\varepsilon(t,x)| \le C\varepsilon^{-|\alpha|+\gamma},$$

$$\sup_{t\in[0,T]}|D_x^\alpha\sigma_\varepsilon(t,.)|_\gamma \le C\varepsilon^{-|\alpha|}. \tag{2.6}$$

Also, there exists a constant C s.t.:

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|\partial_t\sigma_\varepsilon(t,x)| \le C\varepsilon^{-2+\gamma},\ \sup_{t\in[0,T]}|\partial_t\sigma_\varepsilon(t,.)|_\eta \le C\varepsilon^{-2+\gamma-\eta},\ \forall\eta\in(0,\gamma]. \tag{2.7}$$

Proof. For all multi-index $\alpha,|\alpha|\in[\![1,4]\!]$ and $(t,x)\in[0,T]\times\mathbb{R}^d$ and all $\varepsilon>0$:

$$D_x^\alpha\sigma_{\varepsilon,S}(t,x) = \int_{\mathbb{R}^d}\sigma(t,z)D_x^\alpha\rho_\varepsilon(x-z)dz = \int_{\mathbb{R}^d}[\sigma(t,z)-\sigma(t,x)]D_x^\alpha\rho_\varepsilon(x-z)dz.$$

Indeed, setting for all $x\in\mathbb{R}^d$, $g_\varepsilon(x):=\int_{\mathbb{R}^d}\rho_\varepsilon(x-z)dz=1$ we have $D_x^\alpha g_\varepsilon(x):=\int_{\mathbb{R}^d}D_x^\alpha\rho_\varepsilon(x-z)dz=0$. Thus, since $|D_x^\alpha\rho_\varepsilon(x-z)|\le\varepsilon^{-(|\alpha|+d)}|D_w^\alpha\rho(w)||_{w=\frac{(x-z)}{\varepsilon}}$, we derive:

$$|D_x^\alpha\sigma_{\varepsilon,S}(t,x)| \le \int_{\mathbb{R}^d}|\sigma(t,z)-\sigma(t,x)|\varepsilon^{-(|\alpha|+d)}|D_w^\alpha\rho(w)|_{w=\frac{(x-z)}{\varepsilon}}dz$$

$$\le \kappa\varepsilon^{-|\alpha|+\gamma}\int_{\mathbb{R}^d}\left(\frac{|z-x|}{\varepsilon}\right)^\gamma\varepsilon^{-d}|D_w^\alpha\rho(w)|_{w=\frac{(x-z)}{\varepsilon}}dz \le c\varepsilon^{-|\alpha|+\gamma},$$

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

$$|D_x^\alpha\sigma_{\varepsilon,S}(t,x)-D_x^\alpha\sigma_{\varepsilon,S}(t,y)| \le \int_{\mathbb{R}^d}|\sigma(t,x-z)-\sigma(t,y-z)|\varepsilon^{-(|\alpha|+d)}|D_w^\alpha\rho(w)|_{w=\frac{z}{\varepsilon}}dz$$

$$\le C\kappa\varepsilon^{-|\alpha|}|x-y|^\gamma.$$

The same bounds hold for $b_{\varepsilon,S}$ as well. The previous controls readily imply (2.6) since the additional time convolution does not have any impact here.

Equation (2.7) is derived proceeding similarly for the time convolution, exploiting as well the $\gamma/2$-Hölder continuity in time of $\sigma_{\varepsilon,S}$. This completes the proof. $\qquad\square$

2.2 Mollification under ($\mathbf{A}_{PS}$) (piecewise smooth drift and smooth diffusion coefficient)

In this case we only need to regularize the drift in a neighborhood of the discontinuities. Let us denote by $m \in \mathbb{N}^*$, the finite number of spatial discontinuity sets and write $\mathcal{I} := \cup_{i=1}^m \mathcal{S}_i$, where we recall from ($\mathbf{A}_{PS}$) that each $\mathcal{S}_i := \{x \in \mathbb{R}^d : g_i(x) = 0\}$ is a smooth (at least C^4) bounded submanifold of $\mathbb{R}^d$ of dimension d_i lower or equal to $d-1$. For a given parameter $\varepsilon > 0$, define its neighborhood $\mathcal{V}_\varepsilon(\mathcal{I}) := \cup_{i=1}^m \mathcal{V}_\varepsilon(\mathcal{S}_i)$, where for $i \in [\![1,m]\!]$, $\mathcal{V}_\varepsilon(\mathcal{S}_i) := \{z \in \mathbb{R}^d : -\varepsilon \le d_S(z, \mathcal{S}_i) \le \varepsilon\}$. Here, $d_S(\cdot, \mathcal{S}_i)$ stands for the **signed** distance to $\mathcal{S}_i$. This function has the same smoothness as the boundary $\mathcal{S}_i$ (see e.g. Lemma 14.16 and its proof p. 355 in [GT98]). By convention, for $d_i \ge 1$, we choose $d_S(x, \mathcal{S}_i)$ to be positive for points x being in the bounded region with bounded boundary $\mathcal{S}_i$.

The fact is now that we set $b_\varepsilon(t,x) = b(t,x)$ on $\mathbb{R}^d \backslash \mathcal{V}_\varepsilon(\mathcal{I})$ and perform a smooth mollification on the neighborhood $\mathcal{V}_\varepsilon(\mathcal{I})$ of the discontinuity sets. A possible way to proceed is the following. Introduce for all $i \in [\![1,m]\!]$, $\partial \mathcal{V}_\varepsilon^{i,1} := \{x \in \mathbb{R}^d : d_S(x, \mathcal{S}_i) = -\varepsilon\}$, $\partial \mathcal{V}_\varepsilon^{i,2} := \{x \in \mathbb{R}^d : d_S(x, \mathcal{S}_i) = \varepsilon\}$. Denoting by $\left(\Pi_{\partial \mathcal{V}_\varepsilon^{i,j}}(x)\right)_{j \in \{1,2\}}$ the projection of x on the corresponding boundary $(\partial \mathcal{V}_\varepsilon^{i,j})_{j \in \{1,2\}}$ of $\mathcal{V}_\varepsilon(\mathcal{S}_i)$, which is again well defined on $\mathcal{V}_\varepsilon(\mathcal{S}_i)$, we set for all $(t,x) \in [0,T] \times \mathcal{V}_\varepsilon(\mathcal{S}_i)$:

$$
\begin{aligned}
b_\varepsilon(t,x) \quad := \quad & b\big(t, \Pi_{\partial \mathcal{V}_\varepsilon^{i,1}}(x)\big) \exp\big(\tfrac{1}{4}\big) \exp\left(-\frac{1}{4 - \frac{d_S(x, \partial \mathcal{V}_\varepsilon^{i,1})^2}{\varepsilon^2}}\right) \\
& + b\big(t, \Pi_{\partial \mathcal{V}_\varepsilon^{i,2}}(x)\big) \exp\big(\tfrac{1}{4}\big) \exp\left(-\frac{1}{4 - \frac{d_S(x, \partial \mathcal{V}_\varepsilon^{i,2})^2}{\varepsilon^2}}\right),
\end{aligned}
$$

where $d_S(x, \partial \mathcal{V}_\varepsilon^{i,j})$, $j \in \{1,2\}$ stands for the **signed** distance of x to the corresponding boundary $\partial \mathcal{V}_\varepsilon^{i,j}$ and is again a smooth function. Observing that for $x \in \partial \mathcal{V}_\varepsilon^{i,1}$ (resp. $x \in \partial \mathcal{V}_\varepsilon^{i,2}$) we indeed have $d_S(x, \partial \mathcal{V}_\varepsilon^{i,2})^2 = 4\varepsilon^2$ (resp. $d_S(x, \partial \mathcal{V}_\varepsilon^{i,1})^2 = 4\varepsilon^2$) we indeed have that for $(t,x) \in [0,T] \times \partial \mathcal{V}_\varepsilon(\mathcal{S}_i) = [0,T] \times \left(\partial \mathcal{V}_\varepsilon^{i,1} \cup \partial \mathcal{V}_\varepsilon^{i,2}\right)$, $b_\varepsilon(t,x) = b(t,x)$ and b_ε is smooth (as $\mathcal{S}_i$ on $\mathcal{V}_\varepsilon(\mathcal{S}_i)$). Thus b_ε is at least C^4 in the space variable.

Of course we have that $|(b - b_\varepsilon)(t,x)| \le C\mathbb{I}_{x \in \mathcal{V}_\varepsilon(\mathcal{I})}$ which is not necessarily small. Anyhow, for all $q > 1$, since the $(\mathcal{S}_i)_{i \in [\![1,m]\!]}$ are bounded, we derive as well:

$$
\begin{aligned}
\|b - b_\varepsilon\|_{L^q([0,T] \times \mathbb{R}^d)} &= \{\int_0^T dt \int_{\mathbb{R}^d} |(b - b_\varepsilon)(t,x)|^q dx\}^{1/q} \le C\{\int_0^T dt \int_{\mathcal{V}_\varepsilon(\mathcal{I})} dx\}^{1/q} \\
&\le C_{\mathcal{I}} \varepsilon^{1/q} T^{1/q}.
\end{aligned}
\tag{2.8}
$$

Recall indeed that since the $(\mathcal{S}_i)_{i \in [\![1,m]\!]}$ have zero Lebesgue measure and smooth boundary, for the *thickened* neighborhoods $\left(\mathcal{V}_\varepsilon(\mathcal{S}_i)\right)_{i \in [\![1,m]\!]}$, we have for all $i \in [\![1,m]\!]$, $|\mathcal{V}_\varepsilon(\mathcal{S}_i)| := \int_{\mathcal{V}_\varepsilon(\mathcal{S}_i)} dx \le C\varepsilon$. This is clear for a bounded portion of hyperplane. The smoothness of the boundary allows to locally map $\mathcal{V}_\varepsilon(\mathcal{S}_i)$ with a bounded neighborhood of a hyperplane if $\mathcal{S}_i$ has dimension $d-1$. For submanifolds of smaller dimension $d-i$, $i > 1$, the straightening of the boundary can be done in the corresponding dimension $d-i+1$ and the associated neighborhood would be smaller, namely $|\mathcal{V}_\varepsilon(\mathcal{S}_i)| \le C\varepsilon^{d-(d-i)} \le C\varepsilon^i$. We take the *worst* bound for simplicity.

Observe as well that the following control holds for the derivatives of the mollified coefficient. For all multi-index $\alpha, |\alpha| \le 4$, there exists $C \ge 1$ s.t. for all $(t,x) \in [0,T] \times \mathbb{R}^d$:

$$
|\partial_x^\alpha b_\varepsilon(t,x)| \le C\{\varepsilon^{-|\alpha|}\mathbb{I}_{x \in \mathcal{V}_\varepsilon(\mathcal{I})} + \mathbb{I}_{x \notin \mathcal{V}_\varepsilon(\mathcal{I})}\}.
\tag{2.9}
$$

Under the considered assumptions it is not necessary to mollify the diffusion coefficients. We thus set for all $(t, x) \in [0, T] \times \mathbb{R}^d$, $\sigma_\varepsilon(t, x) = \sigma(t, x)$, in order to keep homogeneous notations under our two running assumptions for the drift.

2.3 Stability results

Recall now that under $(\mathbf{A}_H)$ or $(\mathbf{A}_{PS})$ equation (1.1) admits a density (see e.g. [She91] under $(\mathbf{A}_H)$ or Proposition 1 in [KKM16] under $(\mathbf{A}_{PS})$), i.e. for all $0 \le s < t \le T$, $x \in \mathbb{R}^d$, $B \in \mathcal{B}(\mathbb{R}^d)$, $\mathbb{P}[X_t \in B | X_s = x] = \int_B p(s, t, x, y) dy$. The same holds for the Euler scheme in (1.2) (see e.g. Theorem 2.1 in [LM10]), for all $0 \le t_i < t_j \le T$, $x \in \mathbb{R}^d$, $\mathbb{P}[X_{t_j}^h \in B | X_{t_i}^h = x] = \int_B p^h(t_i, t_j, x, y) dy$. These properties remain valid for the respective perturbed diffusion and Euler scheme whose coefficients correspond to the procedures described in Section 2.1 and Section 2.2 depending on whether assumption $(\mathbf{A}_H)$ or $(\mathbf{A}_{PS})$ is in force. We denote the densities associated with the perturbed diffusion and discretization scheme by p_ε and p_ε^h respectively.

Let us now state the sensitivity result following from Theorems 1 and 2 in [KKM16].

Theorem 2.3 (Main Sensitivity Result). *Define for $q \in (d, +\infty]$ and $\eta \in (0, 1]$ the quantities:*

$$
\begin{aligned}
\Delta_{\varepsilon, b, q} &:= \sup_{t \in [0, T]} \|(b - b_\varepsilon)(t, \cdot)\|_{L^q(\mathbb{R}^d)}, \\
\Delta_{\varepsilon, \sigma, \eta} &:= \sup_{t \in [0, T]} \|\sigma(t, \cdot) - \sigma_\varepsilon(t, \cdot)\|_{L^\infty(\mathbb{R}^d)} + \sup_{t \in [0, T]} |\sigma(t, .) - \sigma_\varepsilon(t, .)|_\eta.
\end{aligned}
$$

Set $\Delta_{\varepsilon, \eta, q} := \Delta_{\varepsilon, b, q} + \Delta_{\varepsilon, \sigma, \eta}$. It holds under $(\mathbf{A})$ that there exists $C_{\eta, q} \ge 1$ s.t. for all $0 \le t_i < t_j \le T$ and $(x, y) \in (\mathbb{R}^d)^2$:

$$
p_c(t_j - t_i, y - x)^{-1} \left\{ |(p - p_\varepsilon)(t_i, t_j, x, y)| + |(p^h - p_\varepsilon^h)(t_i, t_j, x, y)| \right\} \le C_{\eta, q} \Delta_{\varepsilon, \eta, q}. \quad (2.10)
$$

Also, there exists $C \ge 1$ s.t.:

$$
C_{\eta, q} \le C \exp(C((\tfrac{\eta}{2} \wedge \alpha(q))^{-1} + 1)^{(\frac{\eta}{2} \wedge \alpha(q))^{-1} + 1}), \quad \alpha(q) := \frac{1}{2}(1 - \frac{d}{q}). \quad (2.11)
$$

Remark 2.4 (Constraint on q). The constraint $q > d$ in the above result is due to the fact that to establish (2.10) in the case $q < +\infty$, we are led to control quantities of the type

$$
Q := \int_0^{t_j - t_i} dt \int_{\mathbb{R}^d} p_c(t, x - w) |b - b_\varepsilon|(t, w) \frac{1}{((t_j - t_i) - t)^{1/2}} p_c((t_j - t_i) - t, y - w) dw,
$$

through Hölder's inequality. The constraint $q > d$ naturally appears in order to derive $Q \le \bar{C} \frac{\Gamma(\alpha(q))}{\Gamma(1 + \alpha(q))} T^{\alpha(q)} \|b - b_\varepsilon\|_{L^q(\mathbb{R}^d)}$ for a finite $\bar{C} := \bar{C}((\mathbf{A}), T) \ge 1$. We refer to the proof of Lemma 2 in [KKM16] for additional details.

We point out that under $(\mathbf{A}_H)$ the previous theorem will be used taking η *small* and $q = +\infty$, so that $\frac{\eta}{2} \wedge \alpha(q) = \frac{\eta}{2}$ and under $(\mathbf{A}_{PS})$ with $\eta = 1$ and $q \in (d, +\infty)$ so that $\frac{\eta}{2} \wedge \alpha(q) = \alpha(q)$.

Proof. Equation (2.10) readily follows from Theorems 1 and 2 in [KKM16]. The point is here to specify the control (2.11) on the constant appearing in (2.10). Lemma 3 in [KKM16], quantifies the explosive contributions for each term of the parametrix series giving the difference of the densities. It holds for both the diffusion and the Euler scheme, see Section 3.2 of [KKM16] for details, and yields:

$$
C_{\eta, q} \le \sum_{r \ge 1} (r + 1) \frac{\bar{C}^{r+1} \left[\Gamma(\frac{\eta}{2} \wedge \alpha(q)) \right]^r}{\Gamma(1 + r(\frac{\eta}{2} \wedge \alpha(q)))} T^{r(\frac{\eta}{2} \wedge \alpha(q))},
$$

for a constant $\bar{C} := \bar{C}((\mathbf{A}), T)$ which does not depend on η or q.

Introduce for $\theta \in (0, \frac{1}{2}]$ the quantity:

$$I(T, \theta) := \sum_{r \geq 1} (r+1) \frac{\bar{C}^{r+1} \left[\Gamma(\theta)\right]^r}{\Gamma(1 + r\theta)} T^{r\theta}.$$

One easily gets that for a given $T > 0$, there exists $\tilde{C} := \tilde{C}((\mathbf{A}), T) \geq 1$ independent of θ as well such that:

$$I(T, \theta) \leq C \sum_{r \geq 1} (r+1) \frac{\tilde{C}^{r+1} \left[\Gamma(\theta)\right]^r}{\Gamma(1 + r\theta)}.$$

Set now $r_0 := \lceil \frac{1}{\theta} \rceil$ and write by monotonicity of the Γ function (see e.g. formula 8.363 (8) in Gradstein and Ryzhik [GR14]):

$$
\begin{aligned}
I(T, \theta) &\leq C \sum_{k \geq 0} (k+1) r_0 \sum_{k r_0 \leq r < (k+1) r_0} \frac{\{\tilde{C} \Gamma(\theta)\}^r}{\Gamma(1+k)} \\
&\leq C \sum_{k \geq 0} \frac{(k+1) r_0}{\Gamma(k+1)} \sum_{k r_0 \leq r < (k+1) r_0} \{\tilde{C}(\theta^{-1} + \exp(-1))\}^r \\
&\leq C r_0^2 \sum_{k \geq 0} \frac{(k+1)}{\Gamma(k+1)} [\{\tilde{C}(\theta^{-1} + \exp(-1))\}^{r_0}]^{k+1} \leq C \exp(\tilde{C}(\theta^{-1} + 1)^{\theta^{-1}+1}).
\end{aligned}
$$

This gives (2.11) taking $\theta = \frac{\eta}{2} \wedge \alpha(q)$ and completes the proof. $\qquad\square$

From Theorem 2.3, we get the following key sensitivity results.

Lemma 2.5 (Sensitivity under **($\mathbf{A}_H$)**). *Under Assumption **($\mathbf{A}_H$)**, for $\eta \in (0, \gamma)$ there exists $c \leq 1$ s.t. for all $0 \leq t_i < t_j \leq T, (x, y) \in (\mathbb{R}^d)^2$:*

$$|(p - p_\varepsilon)(t_i, t_j, x, y)| + |(p^h - p_\varepsilon^h)(t_i, t_j, x, y)| \leq C_\eta \varepsilon^{\gamma - \eta} p_c(t_j - t_i, y - x), \qquad (2.12)$$

where $C_\eta := C_{\eta, \infty} \leq C \exp(C((\frac{\eta}{2})^{-1} + 1)^{(\frac{\eta}{2})^{-1}+1})$ for $C_{\eta, \infty}$ as in (2.11).

Proof. The lemma derives from Theorem 2.3 with $q = +\infty =$ and Proposition 2.1. The bound on C_η follows observing as well that for $\eta \in (0, \gamma)$, $\frac{\eta}{2} < \frac{\gamma}{2} \leq \frac{1}{2} = \alpha(\infty)$ so that $\frac{\eta}{2} \wedge \alpha(\infty) = \frac{\eta}{2}$. $\qquad\square$

Lemma 2.6 (Sensitivity under **($\mathbf{A}_{PS}$)**). *Under Assumption **($\mathbf{A}_{PS}$)**, for $q > d$ there exists $c \leq 1$ s.t. for all $0 \leq t_i < t_j \leq T, (x, y) \in (\mathbb{R}^d)^2$:*

$$|(p - p_\varepsilon)(t_i, t_j, x, y)| + |(p^h - p_\varepsilon^h)(t_i, t_j, x, y)| \leq C_q \varepsilon^{1/q} p_c(t_j - t_i, y - x), \qquad (2.13)$$

where $C_q := C_{1,q} \leq C \exp(C(\alpha(q)^{-1} + 1)^{\alpha(q)^{-1}+1})$ for $C_{1,q}$ as in (2.11).

Proof. Recall that under **($\mathbf{A}_{PS}$)**, since the diffusion coefficient is smooth, there is no need to regularize it and $\sigma = \sigma_\varepsilon$. Thus, $\Delta_{\varepsilon, \sigma, 1} = 0$. From this observation and equation (2.8), Theorem 2.3 then yields (2.13). The bound on C_q follows observing as well that for $q \in (d, +\infty), \frac{1}{2} \wedge \alpha(q) = \alpha(q)$. $\qquad\square$

Let us mention that the constants C_η, C_q in equations (2.12) and (2.13) respectively explode when η goes to 0 and q goes to d, which is precisely what we want in order to have the fastest convergence rate w.r.t. ε. On the other hand, the explosion rates that we have emphasized in (2.11) are crucial in order to equilibrate the global errors. This step is performed in Section 3.4 below.

3 Error analysis and derivation of the main results

3.1 Stream line to the proofs of the main results

This Section is devoted to the proof of Theorems 1.1 and 1.6.

Our main results are those controlling the difference of the densities, i.e. the estimates given in equations (1.6) under $(\mathbf{A}_H)$ and (1.13), (1.14) under $(\mathbf{A}_{PS})$.

To obtain these bounds, the strategy is the following. Let $0 \leq t_i < t_j \leq T$ and $(x,y) \in (\mathbb{R}^d)^2$ be given. One writes for $\varepsilon > 0$:

$$|p(t_i,t_j,x,y) - p^h(t_i,t_j,x,y)| \leq |p - p_\varepsilon|(t_i,t_j,x,y) + |p_\varepsilon - p_\varepsilon^h|(t_i,t_j,x,y) + |p_\varepsilon^h - p^h|(t_i,t_j,x,y).$$
(3.1)

Now, one derives from the sensitivity Lemma 2.5 that, under $(\mathbf{A}_H)$, for all $\eta \in (0,\gamma)$:

$$|p(t_i,t_j,x,y) - p^h(t_i,t_j,x,y)| \leq C_\eta \varepsilon^{\gamma - \eta} p_c(t_j - t_i, y - x) + |(p_\varepsilon - p_\varepsilon^h)|(t_i,t_j,x,y). \quad (3.2)$$

Similarly, Lemma 2.6 yields that, under $(\mathbf{A}_{PS})$, for all $q > d$:

$$|p(t_i,t_j,x,y) - p^h(t_i,t_j,x,y)| \leq C_q \varepsilon^{1/q} p_c(t_j - t_i, y - x) + |(p_\varepsilon - p_\varepsilon^h)|(t_i,t_j,x,y). \quad (3.3)$$

To investigate and minimize the contributions in the error it thus remains from equations (3.2) and (3.3) to precisely control the difference $|p_\varepsilon - p_\varepsilon^h|$ in (3.1). Let us now recall that, since the densities $p_\varepsilon, p_\varepsilon^h$ 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. We thus have that:

$$|(p_\varepsilon - p_\varepsilon^h)(t_i,t_j,x,y)| \leq C_{b_\varepsilon,\sigma_\varepsilon,t_j - t_i} h p_c(t_j - t_i, y - x), \quad (3.4)$$

where $C_{b_\varepsilon,\sigma_\varepsilon,t_j - t_i}$ depends on the derivatives of $b_\varepsilon, \sigma_\varepsilon$, it therefore explodes when ε goes to 0, and exhibits as well a singularity in $t_j - t_i$.

The delicate and crucial point is that we must here precisely quantify this explosion. A key ingredient, to proceed is the *parametrix* series representation for the densities of the diffusion and its Euler scheme. These aspects are recalled in Section 3.2 below.

Importantly, the parametrix expansion of the density of $X_{t_j}^{t_i,x}$ in (1.1), i.e. for the equation without mollified coefficients, also directly allows to derive, without any sensitivity analysis, exploiting the controls on the derivatives of the density $p(t_i,t_j,x,\cdot)$ of $X_{t_j}^{t_i,x}$ w.r.t. x up to order 2 under $(\mathbf{A}_H)$, the bounds in (1.7) and (1.8). The arguments follow from *cancellation* techniques that are also crucial to derive our main estimates. We first illustrate this approach in Section 3.3 which is dedicated to the proof of (1.7) and (1.8) (integrated weak error).

The main results corresponding to the controls of the difference of the densities are established in Section 3.4. As emphasized above, these results do rely on the sensitivity analysis. They also require a careful analysis of the explosions of the higher order derivatives of the involved heat kernels which need to be quantitatively controlled in terms of the corresponding regularization procedure. The main result in that direction is Proposition 3.4 below whose proof, which heavily exploits cancellation techniques, is postponed to Section 4. It yields a precise control of the constant $C_{b_\varepsilon,\sigma_\varepsilon}$ in (3.4). The main results of Theorems 1.1 and 1.6 are then derived in Section 3.4.1 and 3.4.2 respectively, balancing the errors appearing in (3.4), and (3.2) under $(\mathbf{A}_H)$ or (3.3) under $(\mathbf{A}_{PS})$.

3.2 Parametrix representation of densities

From Section 2 in [KKM16], we derive that under $(\mathbf{A})$ (i.e. the expansions below hold under both $(\mathbf{A}_H)$ and $(\mathbf{A}_{PS})$), for all $\varepsilon \geq 0$ (the expansion below even holds for the initial coefficients taking $\varepsilon = 0$), $0 \leq s < t \leq T$, $(x,y) \in (\mathbb{R}^d)^2$:

$$p_\varepsilon(s,t,x,y) := \sum_{r \in \mathbb{N}} \tilde{p}_\varepsilon \otimes H_\varepsilon^{(r)}(s,t,x,y), \quad (3.5)$$

where for $0 \leq u < t \leq T, (z, y) \in (\mathbb{R}^d)^2$:

$$H_\varepsilon(u, t, z, y) := (L_u^\varepsilon - \tilde{L}_u^{\varepsilon, y})\tilde{p}_\varepsilon(u, t, z, y), \tag{3.6}$$

and $L_u^\varepsilon, \tilde{L}_u^{\varepsilon, y}$ respectively stand for the generators at time u of the processes

$$X_t^{(\varepsilon)} = z + \int_u^t b_\varepsilon(v, X_v^{(\varepsilon)})dv + \int_u^t \sigma_\varepsilon(v, X_v^{(\varepsilon)})dW_v, \tilde{X}_t^{(\varepsilon), y} = z + \int_u^t \sigma_\varepsilon(v, y)dW_v, \tag{3.7}$$

i.e. for all $\varphi \in C^2(\mathbb{R}^d, \mathbb{R}), \ x \in \mathbb{R}^d$,

$$\begin{aligned} L_u^\varepsilon \varphi(x) &= \langle b_\varepsilon(u, x), \nabla_x \varphi(x) \rangle + \frac{1}{2}\mathrm{Tr}\Big(\sigma_\varepsilon \sigma_\varepsilon^*(u, x)D_x^2\varphi(x)\Big), \\ \tilde{L}_u^{\varepsilon, y}\varphi(x) &= \frac{1}{2}\mathrm{Tr}\Big(\sigma_\varepsilon \sigma_\varepsilon^*(u, y)D_x^2\varphi(x)\Big). \end{aligned}$$

Also $\tilde{p}_\varepsilon(u, t, z, y) := \tilde{p}_\varepsilon^y(u, t, z, w)|_{w=y}$ where $\tilde{p}_\varepsilon^y(u, t, z, .)$ stands for the density at time t of the process $\tilde{X}^{(\varepsilon), y}$ starting from z at time u. We denote in (3.5), $\tilde{p}_\varepsilon \otimes H_\varepsilon^{(0)}(s, t, x, y) = \tilde{p}_\varepsilon(s, t, x, y)$ and for all $r \geq 1$, $\tilde{p}_\varepsilon \otimes H_\varepsilon^{(r)}(s, t, x, y) = \int_s^t du \int_{\mathbb{R}^d} \tilde{p}_\varepsilon(s, u, x, z)H_\varepsilon^{(r)}(u, t, z, y)dz$ where for $r \geq 2$,

$$\begin{aligned} H_\varepsilon^{(r)}(u, t, z, y) &:= H_\varepsilon \otimes H_\varepsilon^{(r-1)}(u, t, z, y) \\ &:= \int_u^t dv \int_{\mathbb{R}^d} H_\varepsilon(u, v, z, w)H_\varepsilon^{(r-1)}(v, t, w, y)dw. \end{aligned}$$

More generally, the symbol $\otimes$ stands for the time-space convolution, i.e. for two real valued functions f, g defined on $[0, T]^2 \times (\mathbb{R}^d)^2$, $0 \leq s < t \leq T, f \otimes g(s, t, x, y) := \int_s^t du \int_{\mathbb{R}^d} f(s, u, x, z)g(u, t, z, y)dz$. We also recall that under (**A**$_{PS}$), since the diffusion coefficient is smooth we do not regularize it and denote in this case $\sigma_\varepsilon = \sigma$.

To investigate the contribution $p_\varepsilon - p_\varepsilon^h$ in (3.1) we will also use for $0 \leq t_i < t_j \leq T, (x, y) \in (\mathbb{R}^d)^2$ the function:

$$p_\varepsilon^d(t_i, t_j, x, y) := \sum_{r \in \mathbb{N}} \tilde{p}_\varepsilon \otimes_h H_\varepsilon^{(r)}(t_i, t_j, x, y), \tag{3.8}$$

where the quantities at hand are the same as above and the symbol $\otimes_h$ replacing the $\otimes$ in (3.5) denotes the *discrete* convolution, i.e. for all $r \geq 1$,

$$\tilde{p}_\varepsilon \otimes_h H_\varepsilon^{(r)}(t_i, t_j, x, y) = h \sum_{k=0}^{j-i-1} \int_{\mathbb{R}^d} \tilde{p}_\varepsilon(t_i, t_{i+k}, x, z)H_\varepsilon^{(r)}(t_{i+k}, t_j, z, y)dz.$$

Even though $p_\varepsilon^d(t_i, t_j, x, .)$ is not *a priori* a density, we will call it so with a slight abuse of terminology. An important control, under (**A**), for the terms in the parametrix series is the following:

$\forall 0 \leq s < t \leq T$,

$$|\tilde{p}_\varepsilon \otimes H_\varepsilon^{(r)}(s, t, x, y)| \leq \frac{((1 \vee T^{(1-\gamma)/2})c_1)^{r+1}\left[\Gamma(\frac{\gamma}{2})\right]^r}{\Gamma(1 + r\frac{\gamma}{2})}p_c(t - s, y - x)(t - s)^{\frac{r\gamma}{2}},$$

$\forall 0 \leq t_i < t_j \leq T$,

$$|\tilde{p}_\varepsilon \otimes_h H_\varepsilon^{(r)}(t_i, t_j, x, y)| \leq \frac{((1 \vee T^{(1-\gamma)/2})c_1)^{r+1}\left[\Gamma(\frac{\gamma}{2})\right]^r}{\Gamma(1 + r\frac{\gamma}{2})}p_c(t_j - t_i, y - x)(t_j - t_i)^{\frac{r\gamma}{2}},$$

$$\tag{3.9}$$

taking $\gamma = 1$ under (**A**$_{PS}$). We emphasize that those bounds are uniform w.r.t. $\varepsilon \geq 0$ and refer to [KM02] or Section 2 in [KKM16] for a proof.

From the same references (see also Lemma 3.6 in [KM00]), we have that the density of the Euler scheme also admits a similar *parametrix* representation. Introduce for $0 \le t_i < t_k \le T, (z,y) \in (\mathbb{R}^d)^2$, the schemes:

$$
\begin{aligned}
X_{t_k}^{h,(\varepsilon)} &= z + \sum_{l=i}^{k-1} \big(b_\varepsilon(t_l, X_{t_l}^{h,(\varepsilon)})h + \sigma_\varepsilon(t_l, X_{t_l}^{h,(\varepsilon)})(W_{t_{l+1}} - W_{t_l}) \big), \\
\tilde{X}_{t_k}^{h,(\varepsilon),y} &= z + \sum_{l=i}^{k-1} \sigma_\varepsilon(t_l, y)(W_{t_{l+1}} - W_{t_l}).
\end{aligned}
\tag{3.10}
$$

Viewed as Markov Chains, their generators write for all $\varphi \in C^2(\mathbb{R}^d, \mathbb{R})$, $x \in \mathbb{R}^d$:

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

Define now for $0 \le t_i < t_j \le T, (z,y) \in (\mathbb{R}^d)^2$ the *Markov chain* analogue of the parametrix kernel H in (3.6) by:

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

One gets the following parametrix representation for the density of the Euler scheme:

$$
p_\varepsilon^h(t_i, t_j, x, y) := \sum_{r=0}^{j-i} \tilde{p}_\varepsilon \otimes_h H_\varepsilon^{h,(r)}(t_i, t_j, x, y).
\tag{3.11}
$$

Again, the subscript ε is meant to explicitly express the dependence on the mollified coefficients. Also, the terms in the above series satisfy the controls of equation (3.9) uniformly in $\varepsilon \ge 0$.

3.3 Integrated weak error under (A$_H$)

We first prove the statements concerning the integrated weak error in (1.7) and (1.8). We insist that, in that case, no regularization of the coefficients is needed. We have the following result:

Proposition 3.1 (Controls of the Derivatives.). *Let $T > 0$ be fixed. Under (A$_H$), there exist constants $C \ge 1$, $c \in (0,1]$ s.t. for all $0 \le s < t \le T, (x,y) \in (\mathbb{R}^d)^2$ and all multi-index α, $|\alpha| \le 2$:*

$$
|D_x^\alpha p(s,t,x,y)| \le \frac{C}{(t-s)^{|\alpha|/2}} p_c(t-s, x-y).
\tag{3.12}
$$

As a consequence we also derive that for $t_j = jh \in [0,T]$ being fixed and setting for all $(t,x) \in [0,t_j] \times \mathbb{R}^d$, $v(t,x) := \mathbb{E}[f(X_{t_j}^{t,x})]$, as soon as f is bounded, we have that for all $(t,x) \in [0,t_j] \times \mathbb{R}^d$:

$$
|\nabla_x v(t,x)| \le \frac{C}{(t_j - t)^{1/2}}
\tag{3.13}
$$

and for $f \in C^\beta(\mathbb{R}^d, \mathbb{R})$, $\beta \in (0,1]$ (space of globally, and possibly unbounded, Hölder continuous functions), we have for a multi-index α, $|\alpha| \le 2$ and all $(t,x) \in [0,t_j] \times \mathbb{R}^d$:

$$
|D_x^\alpha v(t,x)| \le \frac{C}{(t_j - t)^{(|\alpha|-\beta)/2}}.
\tag{3.14}
$$

Proof. Equation (3.12) is a direct consequence of Proposition 3.4 below. This estimate readily gives (3.13). On the other hand, we get that for $f \in C^\beta(\mathbb{R}^d, \mathbb{R})$, $\beta \in (0,1]$, we have for a multi-index α, $|\alpha| \le 2, (t,x) \in [0,t_j] \times \mathbb{R}^d$:

$$
D_x^\alpha v(t,x) = \int_{\mathbb{R}^d} D_x^\alpha p(t,t_j,x,y) f(y)dy = \int_{\mathbb{R}^d} D_x^\alpha p(t,t_j,x,y)(f(y) - f(x))dy,
$$

recalling that $D_x^\alpha \int_{\mathbb{R}^d} p(t,t_j,x,y)dy = 0$ for the last identity. This is precisely what we call a *cancellation* technique. It allows here to exploit the spatial Hölder continuity of f to

get rid of the time singularity appearing in (3.12) when $|\alpha| = 2$, or to decrease the time singularity appearing in (3.13). Hence, from (3.12):

$$|D_x^\alpha v(t,x)| \le \frac{C|f|_\beta}{(t_j - t)^{(|\alpha|-\beta)/2}} \int_{\mathbb{R}^d} p_c(t_j - t, y - x)\left(\frac{|x-y|}{(t_j-t)^{1/2}}\right)^\beta dy.$$

Equation (3.14) readily follows. Similar operations will be recurrent in the proof of Proposition 3.4. □

3.3.1 Proof of (1.7): Hölder final test function

Set $t_{\beta,\gamma}^h := \sup\{(t_k)_{k\in[\![0,j]\!]} : t_k \le t_j - h^{\gamma/\beta} < t_{k+1}\}$ and $I_{\beta,\gamma}^h := t_{\beta,\gamma}^h/h$. In particular, if $\gamma \ge \beta$, $t_{\beta,\gamma}^h = t_{j-1}$ and if $\beta > \gamma$, $t_{\beta,\gamma}^h < t_{j-1}$.

Let v be the function defined in Proposition 3.1. It follows from Proposition 3.4 that $v \in C^{\beta/2,\beta}([0,t_j]\times\mathbb{R}^d,\mathbb{R}) \cap C^{1,2}([0,t_j)\times\mathbb{R}^d,\mathbb{R})$. An expansion similar to (1.11) yields:

$$|\mathbb{E}[f(X_{t_j}^{h,t_i,x}) - f(X_{t_j}^{t_i,x})]| \le |\mathbb{E}[f(X_{t_j}^{h,t_i,x}) - v(t_{\beta,\gamma}^h, X_{t_{\beta,\gamma}^h}^{h,t_i,x})]|$$

$$+C\sum_{k=i}^{I_{\beta,\gamma}^h-1}\int_{t_k}^{t_{k+1}} ds\,\mathbb{E}[\{|\nabla_x v(s, X_s^{h,t_i,x})| + |D_x^2 v(s, X_s^{h,t_i,x})|\}\{|s-t_k|^{\gamma/2} + |X_s^{h,t_i,x} - X_{t_k}^{h,t_i,x}|^\gamma\}]$$

$$=: (T_L + T_M)(h, t_i, t_j, x),$$
$$(3.15)$$

where T_L stands for the contribution associated with the last step(s) and T_M for the other *main* steps.

From equation (3.14) in Proposition 3.1, one readily gets:

$$T_M(h, t_i, t_j, x) \le Ch^{\gamma/2}\sum_{k=i}^{I_{\beta,\gamma}^h-1}\int_{t_k}^{t_{k+1}}\left(1 + \frac{1}{(t_j-s)^{1-\beta/2}}\right)ds \le Ch^{\gamma/2}. \quad (3.16)$$

The contribution T_L requires a more careful treatment. Let us write:

$$
\begin{aligned}
T_L(h, t_i, t_j, x) \le\;& \mathbb{E}[|f(X_{t_j}^{h,t_i,x}) - f(X_{t_j - h^{\gamma/\beta}}^{h,t_i,x})|]\\
&+ \mathbb{E}[|v(t_j, X_{t_j - h^{\gamma/\beta}}^{h,t_i,x}) - v(t_j - h^{\gamma/\beta}, X_{t_j - h^{\gamma/\beta}}^{h,t_i,x})|]\\
&+ |\mathbb{E}[v(t_j - h^{\gamma/\beta}, X_{t_j - h^{\gamma/\beta}}^{h,t_i,x}) - v(t_{\beta,\gamma}^h, X_{t_{\beta,\gamma}^h}^{h,t_i,x})]|\\
\le\;& \mathbb{E}[|f(X_{t_j}^{h,t_i,x}) - f(X_{t_j - h^{\gamma/\beta}}^{h,t_i,x})|]\\
&+ \mathbb{E}[|v(t_j, X_{t_j - h^{\gamma/\beta}}^{h,t_i,x}) - v(t_j - h^{\gamma/\beta}, X_{t_j - h^{\gamma/\beta}}^{h,t_i,x})|]\\
&+ \int_{t_{\beta,\gamma}^h}^{t_j - h^{\gamma/\beta}}\Big\{\mathbb{E}[|\nabla_x v(s, X_s^{h,t_i,x})||b(s, X_s^{h,t_i,x}) - b(\phi(s), X_{\phi(s)}^{h,t_i,x})|\\
&+ \frac{1}{2}|D_x^2 v(s, X_s^{h,t_i,x})||a(s, X_s^{h,t_i,x}) - a(\phi(s), X_{\phi(s)}^{h,t_i,x})|]\Big\}ds.
\end{aligned}
$$

Expanding as in (1.11) the term $|\mathbb{E}[v(t_j - h^{\gamma/\beta}, X_{t_j - h^{\gamma/\beta}}^{h,t_i,x}) - v(t_{\beta,\gamma}^h, X_{t_{\beta,\gamma}^h}^{h,t_i,x})]|$ with Itô's formula and using (1.10) we derive:

$$
\begin{aligned}
T_L(h, t_i, t_j, x) \le\;& C\mathbb{E}[|X_{t_j}^{h,t_i,x} - X_{t_j - h^{\gamma/\beta}}^{h,t_i,x}|^\beta] + Ch^{\gamma/2}\int_{t_{\beta,\gamma}^h}^{t_j - h^{\gamma/\beta}}\left(1 + \frac{1}{(t_j - s)^{1-\beta/2}}\right)ds\\
\le\;& Ch^{\gamma/2}.
\end{aligned}
\quad (3.17)
$$

Plugging (3.17) and (3.16) into (3.15) yields the required control in (1.7).

Remark 3.2 (Extensions to functions f with subquadratic exponential growth)**.** We stated (1.7) for $f \in C^\beta(\mathbb{R}^d, \mathbb{R})$ for simplicity. Observe anyhow that the above arguments can be adapted to derive the expected convergence rate as soon as f is locally β-Hölder and satisfies the growth condition:

$$\exists C_0 > 0, \forall x \in \mathbb{R}^d, \ |f(x)| \leq C_0 \exp(c_0 |x|^2), \ c_0 \leq \frac{c}{4T},$$

$$\forall (x,y) \in (\mathbb{R}^d)^2, \ |x - y| \leq 1, |f(x) - f(y)| \leq C_0 |x-y|^\beta \exp(c_0 |x|^2), \tag{3.18}$$

where c is as in equation (3.12). In that case, the controls of equations (3.13) and (3.14) would write in the following way. There exists a constant $C \geq 1$ s.t. for all $(t,x) \in [0, t_j) \times \mathbb{R}^d$:

$$
\begin{aligned}
|\nabla v(t,x)| &\leq \frac{C}{(t_j - t)^{1/2}} \int_{\mathbb{R}^d} p_c(t_j - t, y - x) \exp(c_0 |y|^2) dy, \\
&\leq \frac{C \exp(2c_0 |x|^2)}{(t_j - t)^{1/2}} \int_{\mathbb{R}^d} p_c(t_j - t, y - x) \exp(2c_0 |y - x|^2) dy \\
&\leq \frac{C \exp(2c_0 |x|^2)}{(t_j - t)^{1/2}}, \\
\forall \alpha, \ |\alpha| = 2, \ |D_x^\alpha v(t,x)| &\leq \frac{C \exp(2c_0 |x|^2)}{(t_j - t)^{1 - \beta/2}}.
\end{aligned}
\tag{3.19}
$$

Plugging (3.19) into (3.15) and (3.17) still yields, thanks to the condition on c_0 in (3.18) and (1.9), an integrable contribution.

3.3.2 Proof of (1.8): Indicator of a domain as test function

We have assumed A to be C^2 domain and ∂A bounded. Let us denote by $d_S(\cdot, \partial A)$ the **signed** distance to the boundary, i.e. $d(x, \partial A) > 0$ for $x \in A$ and $d(x, \partial A) \leq 0$ for $x \notin A$.

It is known (see e.g. Lemma 14.16 and its proof p. 355 in [GT98]) that for $\delta > 0$ small enough, on $V_\delta(A) := \{y \in \mathbb{R}^d : |d_S(y, \partial A)| \leq \delta\}$, the function $d_S(\cdot, \partial A)$ is C^2 and both the exterior and interior sphere conditions hold. The interior sphere condition writes that for $y \in A_\delta := V_\delta(A) \cap A := \{y \in \mathbb{R}^d : 0 < d_S(y, \partial A) \leq \delta\}$ (interior points of A whose distance to the boundary is lower or equal than δ), its orthogonal projection on the boundary $\Pi_{\partial A}(y)$ is also the unique point s.t. defining $B(y, d_S(y, \partial A)) := \{z \in \mathbb{R}^d : \|z - y\| \leq d_S(y, \partial A)\}$, $B(y, d_S(y, \partial A)) \cap \partial A = \Pi_{\partial A}(y)$. The exterior sphere condition writes similarly for the points $y \in V_\delta(A) \backslash \bar{A} := \{y \in \mathbb{R}^d : -\delta \leq d_S(y, \partial A) < 0\}$ (strictly exterior points of A whose distance to the boundary is lower or equal than δ).

For such a δ, let us now write for $0 \leq t_i < t_j \leq T$, $x \in \mathbb{R}^d$:

$$
\begin{aligned}
&\mathbb{E}[\mathbb{I}_{X_{t_j}^{h,t_i,x} \in A}] - \mathbb{E}[\mathbb{I}_{X_{t_j}^{t_i,x} \in A}] \\
&= \{\mathbb{E}[\mathbb{I}_{X_{t_j}^{h,t_i,x} \in A}] - \mathbb{E}[f_\delta(X_{t_j}^{h,t_i,x})]\} + \{\mathbb{E}[f_\delta(X_{t_j}^{h,t_i,x})] - \mathbb{E}[f_\delta(X_{t_j}^{t_i,x})]\} \\
&\quad + \{\mathbb{E}[f_\delta(X_{t_j}^{t_i,x})] - \mathbb{E}[\mathbb{I}_{X_{t_j}^{t_i,x} \in A}]\} =: \sum_{i=1}^{3} T_i^\delta,
\end{aligned}
\tag{3.20}
$$

where

$$
f_\delta(x) = \begin{cases}
1, \text{ if } x \in A, \\
\exp(1) \exp(-1/(1 - \frac{d_S(x, \partial A)^2}{\delta^2})), \text{ if } x \in V_\delta(A) \backslash A, \\
0 \text{ if } x \notin A \cup V_\delta(A).
\end{cases}
\tag{3.21}
$$

Namely, f_δ stands for a smooth approximation (at least C^2) of the mapping $x \mapsto \mathbb{I}_{x \in A}$.

Recalling again from the proof of Lemma 14.16 in [GT98] that for $x \in V_\delta(A)$, $\nabla_x d_S(x, \partial A) = n(\Pi_{\partial A}(x))$, where $n(\Pi_{\partial A}(x))$ stands for the inner unit normal associated with the projection on the boundary, we get for $x \in V_\delta(A) \backslash A$:

$$\nabla f_\delta(x) = -2 \frac{d_S(x, \partial A) n(\Pi_{\partial A}(x))}{\delta^2} \left(1 - \frac{d_S(x, \partial A)^2}{\delta^2} \right)^{-2} f_\delta(x). \qquad (3.22)$$

This yields in particular that $|\nabla f_\delta|_\infty = \sup_{x \in V_\delta(A)} |\nabla f_\delta(x)| \leq C\delta^{-1}$. This last bound in particular yields that there exists $C \geq 1$ s.t. for all $\eta \in (0, \gamma]$,

$$\sup_{x,y \in V_{2\delta}(A)} \frac{|f_\delta(x) - f_\delta(y)|}{|x - y|^\eta} \leq C\delta^{-\eta}. \qquad (3.23)$$

Indeed, from the control on $|\nabla f_\delta|_\infty$ and the smoothness of f_δ, we get for all $x, y \in V_{2\delta}(A)$, either $|x - y| \leq \delta$ and $|f_\delta(x) - f_\delta(y)| \leq C\delta^{-1}|x - y| \leq C\delta^{-\eta}|x - y|^\eta$, or $|x - y| \geq \delta$ and $|f_\delta(x) - f_\delta(y)| \leq C \leq C\delta^{-\eta}|x - y|^\eta$.

Now, the terms T_1^δ and T_3^δ in (3.20) can be handled similarly thanks to the Gaussian upper bound that is satisfied, under $(\mathbf{A}_H)$, by the density of both the diffusion and its Euler scheme, see Proposition 3.4 or again [She91], Theorem 2.1 in [LM10]. Precisely, with the notations of (3.20) and provided that $\delta \leq (t_j - t_i)^{1/2}$:

$$|T_1^\delta + T_3^\delta| \leq \mathbb{E}[\mathbb{I}_{X_{t_j}^{t_i,x} \in V_\delta(A)}] + \mathbb{E}[\mathbb{I}_{X_{t_j}^{h,t_i,x} \in V_\delta(A)}] \leq \frac{C\delta}{(t_j - t_i)^{1/2}} \exp\left(-c\frac{d(x, \partial A)^2}{t_j - t_i} \right), \quad (3.24)$$

where $d(x, \partial A) = |d_S(x, \partial A)|$ stands for the **nonnegative** distance to the boundary. Indeed, we have that locally, up to a change of coordinate, only one variable is orthogonal to the straightened image of the hypersurface ∂A. We can thus integrate the Gaussian bounds w.r.t. the other ones. This yields the above control.

Observe that to find the indicated convergence rate this imposes $\delta \leq (t_j - t_i)^{1/2} h^{\gamma/2}$ which specifies the admissible magnitude for the parameter δ. On the other hand, to analyze T_2^δ we recall from (3.23) that setting for all $(t, x) \in [0, t_j] \times \mathbb{R}^d$, $v_\delta(t, x) := \mathbb{E}[f_\delta(X_{t_j}^{t,x})]$ the terminal function f_δ is η-Hölder continuous, for all $\eta \in (0, \gamma]$, with Hölder modulus of continuity bounded by $\delta^{-\eta}$ on $V_{2\delta}(A)$. We will now establish, similarly to (3.14), that for all multi-index α, $|\alpha| \leq 2$, $(t, x) \in [0, t_j] \times \mathbb{R}^d$:

$$|D_x^\alpha v_\delta(t, x)| \leq \frac{C}{(\delta \vee d(x, \partial A))^\eta} \frac{1}{(t_j - t)^{(|\alpha| - \eta)/2}}. \qquad (3.25)$$

Recall indeed that

$$\begin{aligned} |D_x^\alpha v_\delta(t, x)| &= \left| \int_{\mathbb{R}^d} D_x^\alpha p(t, t_j, x, y)(f_\delta(y) - f_\delta(x)) dy \right| \\ &\leq \frac{C}{(t_j - t)^{|\alpha|/2}} \int_{\mathbb{R}^d} p_c(t_j - t, y - x)|f_\delta(y) - f_\delta(x)| dy, \end{aligned} \qquad (3.26)$$

exploiting Proposition 3.4 for the last inequality. Thus, from (3.21):

- if both $x, y \notin V_\delta(A)$, then $f_\delta(x) = \mathbb{I}_{x \in A}, f_\delta(y) = \mathbb{I}_{y \in A}$. If $x \in (A \cup V_\delta(A))^C, y \in A \backslash V_\delta(A)$, or by symmetry $y \in (A \cup V_\delta(A))^C, x \in A \backslash V_\delta(A)$, then $|x - y| \geq \delta \vee d(x, \partial A)$. If now $x, y \in (A \cup V_\delta(A))^C$ or $x, y \in A \backslash V_\delta(A)$ then $f_\delta(x) = f_\delta(y)$ yielding a trivial contribution in (3.26).

- if $x, y \in V_\delta(A)$, then the control of the Hölder modulus gives: $|f_\delta(x) - f_\delta(y)| \leq C\delta^{-\eta}|x - y|^\eta = C(\delta \vee d(x, \partial A))^{-\eta}|x - y|^\eta$.

- if $x \in V_\delta(A), y \notin V_\delta(A)$ (resp. $y \in V_\delta(A), x \notin V_\delta(A)$) we can exploit the Hölder continuity for $y \in V_{2\delta}(A)$ (resp. $x \in V_{2\delta}(A)$) and the fact that $|x - y| \geq \delta \vee d(x, \partial A)$ for $y \notin V_{2\delta}(A)$ (resp. $x \notin V_{2\delta}(A)$).

In all cases, we have established that $|f_\delta(x) - f_\delta(y)| \leq C(\delta \vee d(x, \partial A))^{-\eta}|x - y|^\eta$, which plugged into (3.26) yields the control (3.25). Recall now that, again from Proposition 3.4, we have $v_\delta \in C^{\eta/2,\eta}([0, t_j] \times \mathbb{R}^d) \cap C^{1,2}([0, t_j) \times \mathbb{R}^d)$. In particular, v_δ has the same Hölder continuity modulus as f_δ. We can as well assume w.l.o.g. that $\gamma/\eta \geq 1$ so that $h^{\gamma/\eta} \leq h \leq 1$.

Exploiting now (3.25) in an expansion similar to (1.11) and (3.15), we get:

$$
\begin{aligned}
|T_2^\delta| &\leq |\mathbb{E}[f_\delta(X_{t_j}^{h,t_i,x}) - f_\delta(X_{t_j-h^{\gamma/\eta}}^{h,t_i,x})| + |\mathbb{E}[v_\delta(t_j, X_{t_j-h^{\gamma/\eta}}^{h,t_i,x}) - v_\delta(t_j - h^{\gamma/\eta}, X_{t_j-h^{\gamma/\eta}}^{h,t_i,x})]| + \\
&\quad + |\mathbb{E}[v_\delta(t_j - h^{\gamma/\eta}, X_{t_j-h^{\gamma/\eta}}^{h,t_i,x}) - v_\delta(t_{j-1}, X_{t_{j-1}}^{h,t_i,x})]| \\
&\quad + C\sum_{k=i}^{j-2} \int_{t_k}^{t_{k+1}} ds \mathbb{E}[\{|\nabla_x v_\delta(s, X_s^{h,t_i,x})| + |D_x^2 v_\delta(s, X_s^{h,t_i,x})|\}\{|s - t_k|^{\gamma/2} \\
&\quad + |X_s^{h,t_i,x} - X_{t_k}^{h,t_i,x}|^\gamma\}] \\
&\leq Ch^{\gamma/2}\Big\{\mathbb{E}[(\delta \vee d(X_{t_j}^{h,t_i,x}, \partial A))^{-2\eta}]^{1/2} + 1 \\
&\quad + \int_{t_i}^{t_j-h^{\gamma/\eta}} (1 + \frac{1}{(t_j - s)^{1-\eta/2}}\mathbb{E}[\{\delta \vee d(X_s^{h,t_i,x}, \partial A)\}^{-2\eta}]^{1/2})ds\Big\}, \quad (3.27)
\end{aligned}
$$

where the term $|\mathbb{E}[v_\delta(t_j - h^{\gamma/\eta}, X_{t_j-h^{\gamma/\eta}}^{h,t_i,x}) - v_\delta(t_{j-1}, X_{t_{j-1}}^{h,t_i,x})]|$ is again expanded with Itô's formula which yields bounds similar to those appearing for the contributions associated with the indexes $k \in [\![i, j-2]\!]$.

Recalling as well that the Euler scheme satisfies the Aronson Gaussian bounds (see Proposition 3.4 and Theorem 2.1 in [LM10] for details) we obtain for all $s \in (t_i, t_j]$:

$$
\begin{aligned}
&\mathbb{E}[\{\delta \vee d(X_s^{h,t_i,x}, \partial A)\}^{-2\eta}] \\
&\leq C\Big\{(\delta \vee d(x, \partial A))^{-2\eta} + \int_{\frac{1}{2}d(x,\partial A) \geq d(y,\partial A)} \frac{\exp(-c\frac{|x-y|^2}{s-t_i})}{(\delta \vee d(y, \partial A))^{2\eta}} \frac{dy}{(s-t_i)^{d/2}}\Big\}.
\end{aligned}
$$

Since on $\{\frac{1}{2}d(x, \partial A) \geq d(y, \partial A)\}$ we have $|x - y| \geq |x - \Pi_{\partial A}(y)| - |\Pi_{\partial A}(y) - y| \geq \frac{d(x,\partial A)}{2} \geq d(y, \partial A)$, where $\Pi_{\partial A}(y)$ again denotes the projection of y on the boundary ∂A, we get:

$$
\begin{aligned}
&\mathbb{E}[\{\delta \vee d(X_s^{h,t_i,x}, \partial A)\}^{-2\eta}] \\
&\leq C\Big\{(\delta \vee d(x, \partial A))^{-2\eta} + \int_{\frac{1}{2}d(x,\partial A) \geq d(y,\partial A)} \frac{\exp(-\frac{c}{2}\frac{d(y,\partial A)^2}{s-t_i})\exp(-\frac{c}{2}\frac{|x-y|^2}{s-t_i})}{(\delta \vee d(y, \partial A))^{2\eta}} \frac{dy}{(s-t_i)^{d/2}}\Big\} \\
&\leq C\Big\{(\delta \vee d(x, \partial A))^{-2\eta} + 1\Big\}.
\end{aligned}
$$

Hence, since $d(x, \partial A) \geq (t_j - t_i)^{1/2}h^{\gamma/2} \geq \delta$, we get from (3.27)

$$
|T_2^\delta| \leq Ch^{\gamma/2}\Big\{\frac{1}{\eta d(x, \partial A)^\eta} + 1\Big\}. \quad (3.28)
$$

The point is now to find the $\eta \in (0, \gamma]$ maximizing $J_{x,A} : \eta \in (0, \gamma] \mapsto \eta d(x, \partial A)^\eta$ in order to minimize the associated contribution in $\frac{1}{\eta d(x,\partial A)^\eta}$ for T_2^δ. Two cases occur:

- $d(x, \partial A) \geq \exp(-\frac{1}{\gamma})$. In that case for $\eta \in (0, \gamma]$, $J'_{x,A}(\eta) = d(x, \partial A)^\eta(1 + \eta \ln(d(x, \partial A)) \geq 0$ and the maximum over the constraint set is attained for $\eta = \gamma$ and $J_{x,A}(\eta) = \gamma d(x, \partial A)^\gamma$.

- $0 < d(x, \partial A) < \exp(-\frac{1}{\gamma})$. The optimum is then attained for $\eta = -\frac{1}{\ln(d(x,\partial A))} \in (0,\gamma)$. This choice then yields: $J_{x,A}(\eta) := \frac{1}{|\ln(d(x,\partial A))|} \exp(\eta \ln(d(x,\partial A))) = \frac{1}{|\ln(d(x,\partial A))|} e^{-1}$.

This gives from (3.28) the global bound:

$$|T_2^\delta| \leq Ch^{\gamma/2}\Big(1 + \frac{1}{\gamma d(x,\partial A)^\gamma}\mathbb{I}_{d(x,\partial A)\geq\exp(-\frac{1}{\gamma})} + |\ln(d(x,\partial A))|\mathbb{I}_{d(x,\partial A)<\exp(-\frac{1}{\gamma})}\Big). \quad (3.29)$$

It is of course the last term above that becomes significant when the distance of the starting point comes closer to the boundary. The global error estimate deriving from (3.24), the previous computations on $J_{x,A}$ and (3.29) is then better, up to a multiplicative constant, than the one deriving from (1.6) as soon as $\frac{1}{J_{x,A}(\eta)} = |\ln(d(x,\partial A))|e \leq eh^{-C\psi(h)}$ which is equivalent to:

$$|\ln(d(x,\partial A))| \leq h^{-C\psi(h)} \iff d(x,\partial A) \geq \exp(-h^{-C\psi(h)}). \quad (3.30)$$

Since to apply the Aronson's estimates for T_1^δ, T_3^δ (see again eq. (3.20)) we had already assumed $d(x,\partial A) \geq (t_j - t_i)^{1/2}h^{\gamma/2} \geq h^{(1+\gamma)/2}$, we derive that the condition in (3.30) is always fulfilled. It can indeed be easily checked that $h^{(1+\gamma)/2} \geq \exp(-h^{-C\psi(h)})$ for h small enough. Equation (1.8) now follows from (3.24) and (3.29).

Remark 3.3 (Extension to piecewise smooth domains.)**.** Let us mention that results similar to (1.8) could also be derived for domains $A := \cap_{i=1}^n A_i$ that write as finite intersections of smooth domains $(A_i)_{i\in[\![1,n]\!]}$ with bounded boundaries, and therefore have piecewise smooth boundary. In that case, $d(x,\partial A) := \inf_{i\in\{1,\cdots,n\}} d(x,\partial A_i)$ is well defined, but the corresponding signed distance can fail to be smooth, precisely close to the resulting *corners*. Hence, f_δ cannot be directly defined as above. Namely, some additional mollification of the corresponding distance would be necessary as well.

3.4 Error expansion for the Euler scheme: Controls on the densities

From Theorem 1.1, Theorem 2.1 and their proofs in [KM02] we have with the notations of the previous paragraph:

$$(p_\varepsilon - p_\varepsilon^h)(t_i, t_j, x, y) = (p_\varepsilon - p_\varepsilon^d)(t_i, t_j, x, y) + h\int_0^1 (1-\tau)\Big\{p_\varepsilon^d \otimes_h (\tilde{L}_{\cdot,*}^\varepsilon - \tilde{L}^{*,\varepsilon})^2 p_\varepsilon^{\tau,h}(t_i, t_j, x, y)\Big\}d\tau, \quad (3.31)$$

where we denote for $0 \leq t_i < t_j \leq T, \tau \in [0,1]$:

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

$$\forall (k,z) \in (i,j]\!] \times \mathbb{R}^d, \ \tilde{p}_\varepsilon^\tau(t_i, t_k, x, z) := \int_{\mathbb{R}^d} \tilde{p}_\varepsilon^x(t_i, t_i + \tau h, x, w)\tilde{p}_\varepsilon^z(t_i + \tau h, t_k, w, z)dw.$$

Also, for $l \in \{1,2\}, t = t_{i+k}, \ k \in [\![0, j-i-1]\!], \ (\tilde{L}_{t,*}^\varepsilon)^l\phi(x,y) := (L_{t,\xi}^\varepsilon)^l\phi(x,y))|_{\xi=x}$, and $(\tilde{L}_t^{*,\varepsilon})^l\phi(x,y) := (\bar{L}_{t,\xi}^\varepsilon)^l\phi(x,y))|_{\xi=y}$ for

$$L_{t,\xi}^\varepsilon\phi(x,y) = \langle b_\varepsilon(t,\xi), \nabla_x\phi(x,y)\rangle + \frac{1}{2}\mathrm{Tr}(a_\varepsilon(t,\xi)D_x^2\phi(x,y)),$$

$$\bar{L}_{t,\xi}^\varepsilon\phi(x,y) = \frac{1}{2}\mathrm{Tr}(a_\varepsilon(t,\xi)D_x^2\phi(x,y)).$$

Observe that $L_t^\varepsilon\phi(x,y) = L_{t,*}^\varepsilon\phi(x,y)$, but more generally the operators do not coincide anymore when iterated. Also, we indicate that the operators involved slightly differ from [KM02] since we chose to use a Gaussian process without drift as proxy, see (3.7)

and (3.10). Another difference is the fact that we deal with inhomogeneous coefficients, and the notations $\tilde{L}_{\cdot,*}^{\varepsilon}, \tilde{L}^{*,\varepsilon}$ in (3.31) are used to emphasize the time dependence of the operators in the discrete convolution $\otimes_h$. Anyhow, reproducing the proof of [KM02] taking into account the indicated differences leads to the expression in (3.31).

We mention carefully that in order to analyze the contribution of the last term in the r.h.s. of (3.31) no smoothness in time of the coefficients is needed. On the other hand, such smoothness is clearly required to derive some convergence rates, since to control $p_\varepsilon - p_\varepsilon^d$ we need to investigate the difference between time integrals and Riemann sums (see Proposition 3.7 and its proof below).

The term $\int_0^1 (1-\tau)\{p_\varepsilon^d \otimes_h (\tilde{L}_{\cdot,*}^{\varepsilon} - \tilde{L}^{*,\varepsilon})^2 p_\varepsilon^{\tau,h}(t_i, t_j, x, y)\}d\tau$ involves derivatives of the coefficients and *heat kernels* up to order 4. The point is again that the derivatives of the coefficients and kernels explode with ε going to 0 (see equation (2.6)). It is precisely this aspect that deteriorates the convergence rate w.r.t. the usual *smooth* case. We carefully mention that if $\sigma(t, x) = \sigma$, the previous contributions involve lower derivatives of the heat kernel (up to order 2).

The key elements are now the following Propositions. The first one gives bounds for the derivatives of the densities involved in the parametrix series (3.5), (3.8). The second one controls the difference between the discrete and continuous convolutions in (3.31).

Proposition 3.4 (Controls for the derivatives of the densities). *Let α, $|\alpha| \leq 4$ be a multi-derivation index.*

*Under (**A_H**), there exist constants $C \geq 1, c \in (0,1]$ s.t. for all $0 \leq s < t \leq T$, $(x, y) \in (\mathbb{R}^d)^2$:*

$$|D_x^\alpha \bar{p}_\varepsilon(s, t, x, y)| \leq \frac{C}{(t-s)^{|\alpha|/2}} p_c(t-s, y-x), |\alpha| \leq 2,$$

$$|D_x^\alpha \bar{p}_\varepsilon(s, t, x, y)| \leq \frac{C}{(t-s)^{|\alpha|/2}} p_c(t-s, y-x)(1 + \varepsilon^{-|\alpha|+2}(t-s)^{\gamma/2}), |\alpha| \in [\![3,4]\!], \quad (3.32)$$

$$|D_y^\alpha \bar{p}_\varepsilon(s, t, x, y)| \leq \frac{C\varepsilon^{-|\alpha|+\gamma}}{(t-s)^{|\alpha|/2}} p_c(t-s, y-x), \ |\alpha| \geq 1.$$

*Under (**A_{PS}**), for all $q > d$, $\eta \in (0, \alpha(q))$, $\alpha(q) = \frac{1}{2}(1 - \frac{d}{q})$, there exist constants $C \geq 1, c \in (0,1]$ s.t. for all $0 \leq s < t \leq T$, $(x, y) \in (\mathbb{R}^d)^2$:*

$$|D_x \bar{p}_\varepsilon(s, t, x, y)| \leq \frac{C}{(t-s)^{1/2}} p_c(t-s, y-x),$$

$$|D_x^\alpha \bar{p}_\varepsilon(s, t, x, y)|$$

$$\leq \frac{C}{(t-s)^{|\alpha|/2}} p_c(t-s, y-x)(1 + \bar{C}_{\eta,q} \varepsilon^{-|\alpha|+2-\eta+(1/q)\mathbb{I}_{|\alpha| \geq 3}}(t-s)^{\eta/2}), |\alpha| \in [\![2,4]\!], \quad (3.33)$$

$$|D_y \bar{p}_\varepsilon(s, t, x, y)| \leq \frac{C}{(t-s)^{1/2}}(1 + \varepsilon^{-\eta} C_\eta(t-s)^{\eta/2}) p_c(t-s, y-x),$$

$$|D_y^\alpha \bar{p}_\varepsilon(s, t, x, y)| \leq \frac{C(1 + \bar{C}_{\eta,q} \varepsilon^{-|\alpha|+1-\eta}(t-s)^{\eta/2})}{(t-s)^{|\alpha|/2}} p_c(t-s, y-x), \ |\alpha| \in [\![2,4]\!],$$

where $\bar{C}_{\eta,q} = C_\eta \times C_q$ with C_q as in Lemma 2.6 and C_η as in Lemma 2.5.

In the above expressions $\bar{p}_\varepsilon$ can be any of the densities $p_\varepsilon, p_\varepsilon^d, p_\varepsilon^{\tau,h}$ uniformly in $\tau \in [0,1]$. For $p_\varepsilon^d, p_\varepsilon^{\tau,h}$, the time variables s, t are taken on the time grid.

Remark 3.5 (Spatial Hölder continuity and heat-kernel bounds). We point out that the previous controls (3.32) for $\bar{p}_\varepsilon = p_\varepsilon$ would also hold under the sole spatial Hölder continuity of the coefficients b, σ. This improves in some sense those of [IKO62] which require smoothness in time of the coefficients. We get here the same pointwise controls

for the derivatives of the non degenerate heat-kernel with spatial Hölder coefficients up to order 2, uniformly in $\varepsilon \in [0, 1]$.

Remark 3.6 (Constants in (3.33)). Even though we are currently considering $(\mathbf{A}_{PS})$, the associated small smoothing effect deriving from the regularization of the drift is the same as for the sensitivities of densities under $(\mathbf{A}_H)$, for which it was induced by the small Hölder parameter for the difference of the diffusion coefficient and its regularization. In both cases the constant C_η appears through the control of the corresponding parametrix series, see the proofs of Theorem 2.3, Lemma 2.5 and Proposition 3.4 below.

Proposition 3.7 (Bounds for the difference between continuous and discrete time convolutions). *Under* $(\mathbf{A}_H)$, *there exist* $C \geq 1, c \in (0, 1]$ *s.t. for all* $0 \leq t_i < t_j \leq T$, $(x, y) \in (\mathbb{R}^d)^2$, $\eta \in (0, \gamma)$:

$$|(p_\varepsilon - p_\varepsilon^d)(t_i, t_j, x, y)| \leq C_\eta h^{(\gamma - \eta)/2} p_c(t_j - t_i, y - x). \tag{3.34}$$

Under $(\mathbf{A}_{PS})$, *there exist* $C \geq 1, c \in (0, 1]$ *s.t. for all* $0 \leq t_i < t_j \leq T$, $(x, y) \in (\mathbb{R}^d)^2$, $q > d$, $\eta \in (0, \alpha(q))$:

$$|(p_\varepsilon - p_\varepsilon^d)(t_i, t_j, x, y)| \leq \bar{C}_{\eta,q} \left(\frac{h|\ln(h)|\varepsilon^{-(1+\eta)}}{(t_j - t_i)^{1/2}} + \frac{h^{1-\eta/2}\varepsilon^{-(1+\eta)}}{(t_j - t_i)^{1/2}} + h\varepsilon^{-2+1/q} \right) p_c(t_j - t_i, y - x), \tag{3.35}$$

with $\alpha(q), \bar{C}_{\eta,q}$ *as in Proposition 3.4.*

If now $d(y, \mathcal{V}_\varepsilon(\mathcal{I})) \geq 2\varepsilon$ *the previous bound improves to*

$$|p_\varepsilon - p_\varepsilon^d|(t_i, t_j, x, y) \leq \bar{C}_{\eta,q} \left(\frac{h|\ln(h)|\varepsilon^{-(1+\eta)}}{(t_j - t_i)^{1/2}} + \frac{h^{1-\eta/2}\varepsilon^{-(1+\eta)}}{(t_j - t_i)^{1/2}} + \frac{h^{1-\eta/2}}{d(y, \mathcal{V}_\varepsilon(\mathcal{I}))} \right) p_c(t_j - t_i, y - x). \tag{3.36}$$

If $d(y, \mathcal{V}_\varepsilon(\mathcal{I})) \geq 2\varepsilon$ *and additionally* $\sigma(t, x) = \sigma$, *i.e. constant diffusion term, then*

$$|p_\varepsilon - p_\varepsilon^d|(t_i, t_j, x, y) \leq \bar{C}_{\eta,q} \left(\frac{h|\ln(h)|\varepsilon^{-\eta}}{(t_j - t_i)^{1/2}} + \frac{h\varepsilon^{-(1+\eta)+1/q}}{(t_j - t_i)^{1/2}} + \frac{h^{1-\eta/2}}{d(y, \mathcal{V}_\varepsilon(\mathcal{I}))} \right) p_c(t_j - t_i, y - x). \tag{3.37}$$

We postpone the proof of Propositions 3.4 and 3.7 to Section 4 for clarity. It now remains to exploit Propositions 3.4, 3.7 and (3.31) to specifically control how the weak error for the densities depends on the explosive norms of the mollified coefficients.

3.4.1 Proof of the main results for Hölder coefficients (Theorem 1.1 under $(\mathbf{A}_H)$)

Observe from Proposition 3.4 that, for all $k \in [\![i, j-1]\!]$, $(z, y) \in (\mathbb{R}^d)^2$, $\tau \in [0, 1]$,

$$\left| \left(\widetilde{L}^\varepsilon_{t_k, *} - \widetilde{L}^{*, \varepsilon}_{t_k} \right) p_\varepsilon^{\tau, h}(t_k, t_j, z, y) \right| = \left| \langle b_\varepsilon(t_k, z), D_z p_\varepsilon^{\tau, h}(t_k, t_j, z, y) \rangle + $$

$$\frac{1}{2} \text{Tr} \left((a_\varepsilon(t_k, z) - a_\varepsilon(t_k, y)) D_z^2 p_\varepsilon^{\tau, h}(t_k, t_j, z, y) \right) \right| \leq \frac{C}{(t_j - t_k)^{1-\gamma/2}} p_c(t_j - t_k, y - z).$$

Iterating the frozen operator, we obtain that $\left(\widetilde{L}^\varepsilon_{t_k, *} - \widetilde{L}^{*, \varepsilon}_{t_k} \right)^2 p_\varepsilon^{\tau, h}(t_k, t_j, z, y)$ is a fourth order differential operator which is the sum of the following typical terms:

$$b_\varepsilon^l(t_k, z) b_\varepsilon^m(t_k, z) D_{z_l z_m}^2 p_\varepsilon^{\tau, h}(t_k, t_j, z, y)$$

$$=: \Psi_{l,m}^{\varepsilon, \tau, h}(t_k, t_j, z, y),$$

$$b_\varepsilon^l(t_k, z) \left((a_\varepsilon^{mq}(t_k, z) - a_\varepsilon^{mq}(t_k, y)) D_{z_l z_m z_q}^3 p_\varepsilon^{\tau, h}(t_k, t_j, z, y) \right)$$

$$=: \Psi_{l,m,q}^{\varepsilon, \tau, h}(t_k, t_j, z, y),$$

$$(a_\varepsilon^{lm}(t_k, z) - a_\varepsilon^{lm}(t_k, y))(a_\varepsilon^{qr}(t_k, z) - a_\varepsilon^{qr}(t_k, y))D_{z_l z_m z_q z_r}^4 p_\varepsilon^{\tau, h}(t_k, t_j, z, y)$$
$$=: \Psi_{l,m,q,r}^{\varepsilon, \tau, h}(t_k, t_j, z, y), \qquad (3.38)$$

for $l, m, q, r \in [\![1, d]\!]$. It is easy to see that the terms with fourth derivatives are the most singular. Hence, to evaluate $p_\varepsilon^d \otimes_h \left(\widetilde{L}_{\cdot,*}^{\varepsilon} - \widetilde{L}_{\cdot}^{*,\varepsilon} \right)^2 p_\varepsilon^{\tau,h}(t_i, t_j, x, y)$, it is enough to concentrate on:

$$p_\varepsilon^d \otimes_h \Psi_{l,m,q,r}^{\varepsilon, \tau, h}(t_i, t_j, x, y)$$
$$= h\big(a_\varepsilon^{lm}(t_i, x) - a_\varepsilon^{lm}(t_i, y)\big)\big(a_\varepsilon^{qr}(t_i, x) - a_\varepsilon^{qr}(t_i, y)\big)D_{z_l z_m z_q z_r}^4 p_\varepsilon^{\tau, h}(t_i, t_j, x, y)$$
$$+ h\sum_{k \in [\![i+1, \lceil \frac{i+j}{2} \rceil]\!]} \int_{\mathbb{R}^d} p_\varepsilon^d(t_i, t_k, x, z)\big(a_\varepsilon^{lm}(t_k, z) - a_\varepsilon^{lm}(t_k, y)\big)\big(a_\varepsilon^{qr}(t_k, z) - a_\varepsilon^{qr}(t_k, y)\big)$$
$$\times D_{z_l z_m z_q z_r}^4 p_\varepsilon^{\tau, h}(t_k, t_j, z, y)dz$$
$$+ h\sum_{k \in [\![\lceil \frac{i+j}{2} \rceil + 1, j-1]\!]} \int_{\mathbb{R}^d} p_\varepsilon^d(t_i, t_k, x, z)\big(a_\varepsilon^{lm}(t_k, z) - a_\varepsilon^{lm}(t_k, y)\big)\big(a_\varepsilon^{qr}(t_k, z) - a_\varepsilon^{qr}(t_k, y)\big)$$
$$\times D_{z_l z_m z_q z_r}^4 p_\varepsilon^{\tau, h}(t_k, t_j, z, y)dz$$
$$=: (T_1 + T_2 + T_3)(t_i, t_j, x, y). \qquad (3.39)$$

The tools to control the above terms are (3.32) in Proposition 3.4 and the Hölder continuity of the mollified coefficients under $(\mathbf{A}_H)$. We readily derive:

$$|T_1(t_i, t_j, x, y)| \leq \frac{Ch(1 + \varepsilon^{-2}(t_j - t_i)^{\gamma/2})}{(t_j - t_i)^{2-\gamma}} p_c(t_j - t_i, y - x). \qquad (3.40)$$

For the term T_2 in (3.39), integrating once by parts, we obtain from (3.32) and (2.6) that:

$$|T_2(t_i, t_j, x, y)| \leq \frac{C\varepsilon^{-1+\gamma}}{(t_j - t_i)^{1-\gamma}} p_c(t_j - t_i, y - x)\Big(1 + \varepsilon^{-1}(t_j - t_i)^{\gamma/2}\Big). \qquad (3.41)$$

The term T_3 in (3.39) can be handled using the same arguments and two integrations by parts in order to get rid of the time singularities. After integrations by parts, the most singular terms w.r.t. ε have the following form:

$$T_{31}(t_i, t_j, x, y) \;\; := \;\; h\sum_{k \in [\![\lceil \frac{i+j}{2} \rceil + 1, j-1]\!]} \int_{\mathbb{R}^d} p_\varepsilon^d(t_i, t_k, x, z)D_{z_l}a_\varepsilon^{lm}(t_k, z)D_{z_m}a_\varepsilon^{qr}(t_k, z)$$
$$\times D_{z_q z_r}^2 p_\varepsilon^{\tau, h}(t_k, t_j, z, y)dz,$$
$$T_{32}(t_i, t_j, x, y) \;\; := \;\; h\sum_{k \in [\![\lceil \frac{i+j}{2} \rceil + 1, j-1]\!]} \int_{\mathbb{R}^d} D_{z_l z_m}^2 p_\varepsilon^d(t_i, t_k, x, z) \big[(a_\varepsilon^{lm}(t_k, z) - a_\varepsilon^{lm}(t_k, y))$$
$$\times (a_\varepsilon^{qr}(t_k, z) - a_\varepsilon^{qr}(t_k, y))\big] D_{z_q z_r}^2 p_\varepsilon^{\tau, h}(t_k, t_j, z, y)dz. \qquad (3.42)$$

For T_{31}, we obtain from inequality (3.32) in Proposition 3.4 and (2.6) that:

$$\begin{aligned}
|T_{31}(t_i, t_j, x, y)| &\leq& C\varepsilon^{-2+2\gamma}p_c(t_j - t_i, y - x)\sum_{k \in [\![\lceil \frac{i+j}{2} \rceil + 1, j-1]\!]} \frac{h}{(t_j - t_k)} \\
&\leq& C\varepsilon^{-2+2\gamma}p_c(t_j - t_i, y - x)\int_{\frac{t_i + t_j}{2}}^{t_j - h} \frac{du}{t_j - u} \\
&\leq& C\varepsilon^{-2+2\gamma}p_c(t_j - t_i, y - x)|\ln h|. \qquad (3.43)
\end{aligned}$$

For T_{32}, Proposition 3.4 and the spatial Hölder continuity of a_ε yield:

$$|T_{32}(t_i, t_j, x, y)| \leq \frac{C\varepsilon^{-2+\gamma}}{(t_j - t_i)^{1-\gamma}} p_c(t_j - t_i, y - x). \tag{3.44}$$

An upper-bound for T_3 then follows summing (3.44) and (3.43). We then derive from (3.40), (3.41) and (3.39) that:

$$\left| p_\varepsilon^d \otimes_h \left(\widetilde{L}_{\cdot,*}^\varepsilon - \widetilde{L}_{\cdot}^{*,\varepsilon} \right)^2 p_\varepsilon^{\tau,h}(t_i, t_j, x, y) \right|$$

$$\leq C \left(\frac{\varepsilon^{-2+\gamma}}{(t_j - t_i)^{1-\gamma}} + \frac{h(1 + \varepsilon^{-2}(t_j - t_i)^{\gamma/2})}{(t_j - t_i)^{2-\gamma}} + \varepsilon^{-2+2\gamma} |\ln h| \right) p_c(t_j - t_i, y - x). \tag{3.45}$$

We thus eventually get from (3.31), (3.34) and (3.45):

$$\left| \left(p_\varepsilon - p_\varepsilon^h \right)(t_i, t_j, x, y) \right|$$

$$\leq C \left\{ C_\eta h^{(\gamma-\eta)/2} + h \left(\frac{h}{(t_j - t_i)^{2-\gamma}} + \frac{\varepsilon^{-2+\gamma}}{(t_j - t_i)^{1-\gamma}} \left(1 + \frac{\varepsilon^{-\gamma}}{(t_j - t_i)^{1-\gamma/2}} h \right) + \varepsilon^{-2+2\gamma} |\ln h| \right) \right\}$$

$$\times p_c(t_j - t_i, y - x). \tag{3.46}$$

Without loss of generality we assume now that $0 \leq t_j - t_i \leq T \leq 1$. We also suppose that:

$$\left(\frac{h}{(t_j - t_i)^{1-\gamma/2}} \right)^{1/\gamma} \leq \varepsilon. \tag{3.47}$$

We will check that (3.47) holds for the specific choice of the parameters ε, η which is performed below.

We derive from equations (3.46), (3.47), together with (3.1), (3.2) that:

$$\left| \left(p - p^h \right)(t_i, t_j, x, y) \right|$$

$$\leq C \left\{ C_\eta \left(\varepsilon^{\gamma-\eta} + h^{(\gamma-\eta)/2} \right) + \frac{h\varepsilon^{-2+\gamma}}{(t_j - t_i)^{1-\gamma}} + h\varepsilon^{-2+2\gamma} |\ln h| \right\} p_c(t_j - t_i, y - x). \tag{3.48}$$

Take now

$$C_\eta \varepsilon^{\gamma-\eta} = \frac{h\varepsilon^{-2+\gamma}}{(t_j - t_i)^{1-\gamma}} \iff \varepsilon = \left(\frac{h}{(t_j - t_i)^{1-\gamma}} \right)^{1/(2-\eta)} C_\eta^{-1/(2-\eta)}.$$

For such a choice of a mollifying parameter we have for $(t_j - t_i) \geq h^{1/(2-\gamma)}$:

$$\varepsilon^\gamma |\ln h| = \left(\frac{h}{(t_j - t_i)^{1-\gamma}} \right)^{\gamma/(2-\eta)} C_\eta^{-\gamma/(2-\eta)} |\ln h| \leq h^{\gamma/((2-\gamma)(2-\eta))} C_\eta^{-\gamma/(2-\eta)} |\ln h|.$$

Assume for a while that η can be taken so that:

$$h^{\gamma/((2-\gamma)(2-\eta))} C_\eta^{-\gamma/(2-\eta)} |\ln h| \leq h^{\gamma/((2-\gamma)2)} \impliedby \frac{2(2-\gamma)}{\gamma} \frac{\ln_2(h^{-1})}{\ln(h^{-1})} \leq \frac{\eta}{2}, \tag{3.49}$$

recalling as well that $C_\eta \geq 1$ for the last assertion. Then, for $(t_j - t_i) \geq h^{1/(2-\gamma)}$ if (3.49) holds:

$$\varepsilon^\gamma |\ln h| \leq h^{\gamma/((2-\gamma)2)} \leq (t_j - t_i)^{\gamma/2}.$$

Hence, from (3.48), if (3.49) holds:

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

$$\leq C \exp(C(2\eta^{-1}+1)^{2\eta^{-1}+1})\Big\{h^{(\gamma-\eta)/2} + \Big(\frac{h}{(t_j-t_i)^{1-\gamma}}\Big)^{\frac{\gamma}{2}-\eta\frac{1-\gamma/2}{2-\eta}}\Big\}p_c(t_j-t_i,y-x),\quad(3.50)$$

using the bounds of Lemma 2.5 for the last inequality. The point is now to carefully choose $\eta := \eta(h)$. Let us consider the specific sequence $\eta = \eta(h) := 2\frac{\log_3(h^{-1})}{\log_2(h^{-1})}$, where we recall that for $k \in \mathbb{N}$, $\log_k(x)$ stands for the kth iterated logarithm of x. Observe that this $\eta(h)$ satisfies the condition (3.49) for h small enough. Setting $\beta_h := h^{-\eta}$ and $\alpha_h := \exp\Big(C(2\eta^{-1}+1)^{2\eta^{-1}+1}\Big)$, we get that:

$$
\begin{aligned}
\log_2(\beta_h) &= \log(\eta\log(h^{-1})) = \log(2) + \log_4(h^{-1}) - \log_3(h^{-1}) + \log_2(h^{-1}),\\
\log_2(\alpha_h) &= \log(C(2\eta^{-1}+1)^{2\eta^{-1}+1}) = \log(C) + (2\eta^{-1}+1)\log(2\eta^{-1}+1)\\
&= \log(C) + (\frac{\log_2(h^{-1})}{\log_3(h^{-1})}+1)\log(2\eta^{-1}(1+\frac{\eta}{2}))\\
&= \log(C) + (\frac{\log_2(h^{-1})}{\log_3(h^{-1})}+1)\{\log(2\eta^{-1})+\log(1+\frac{\eta}{2})\}\\
&= \log(C) + (\frac{\log_2(h^{-1})}{\log_3(h^{-1})}+1)\{\log_3(h^{-1})-\log_4(h^{-1})+\log(1+\frac{\eta}{2})\}\\
&= \log_2(h^{-1}) - \frac{\log_2(h^{-1})\log_4(h^{-1})}{\log_3(h^{-1})} + \log_3(h^{-1}) - \log_4(h^{-1}) + R_h,\\
R_h &:= \log(C) + \log(1+\frac{\log_3(h^{-1})}{\log_2(h^{-1})})\left\{\frac{\log_2(h^{-1})}{\log_3(h^{-1})}+1\right\}.
\end{aligned}
$$

It is easily seen that there exists a finite constant $\bar{C} > 0$ s.t. for all h small enough, $R_h \leq \bar{C}$ and that $\log_2(\beta_h) \geq \log_2(\alpha_h) - \bar{C}$. By monotonicity of the exponential, recalling as well that $\eta \in (0,\gamma)$, we thus derive:

$$\big(\beta_h^{\frac{1}{2}} + \beta_h^{\frac{1-\gamma/2}{2-\eta}}\big)\alpha_h = \big(h^{-\frac{\eta}{2}} + h^{-\eta\frac{1-\gamma/2}{2-\eta}}\big)\exp(C(2\eta^{-1}+1)^{2\eta^{-1}+1}) \leq 2h^{-\eta(1/2+\exp(\bar{C}))}. \quad(3.51)$$

The previous choice of η yields that, since from Lemma 2.5 $C_\eta \leq C\alpha_h$, (3.47) is satisfied as well. Plugging (3.51) into (3.50) we complete the proof of equation (1.6) in Theorem 1.1.

3.4.2 Proof of the main results for piecewise smooth coefficients (Theorem 1.6 under ($\mathbf{A}_{PS}$))

Keeping the definitions of (3.38), the idea is to proceed as in the previous section from equations (3.31), and (3.39). To emphasize the specificity of Assumptions ($\mathbf{A}_{PS}$), due to the approximation of the piecewise smooth drift, we begin with the special case $\sigma(t,x) = \sigma$. In that framework, the only terms appearing in $(\tilde{L}^\varepsilon_{\cdot,*} - \tilde{L}^{*,\varepsilon}_{\cdot})^2 p^{\tau,h}_\cdot$ are the $\Psi^{\varepsilon,\tau,h}_{l,m}$ introduced in (3.38). From equation (3.33) in Proposition 3.4, using a direct control for the index $k = i$ and a global integration by part for $k > i$, associated with the bound of (2.9), we derive:

$$|[p^d_\varepsilon \otimes_h (\tilde{L}^\varepsilon_{\cdot,*} - \tilde{L}^{*,\varepsilon}_{\cdot})^2 p^{\tau,h}_\varepsilon](t_i,t_j,x,y)| \leq C\Big(\frac{\bar{C}_{\eta,q}\varepsilon^{-\eta}}{(t_j-t_i)}hp_c(t_j-t_i,y-x)$$

$$+ h\sum_{k\in[\![i+1,j-1]\!]}\Big|\int_{\mathbb{R}^d}\mathrm{div}_z\Big(p^d_\varepsilon(t_i,t_k,x,z)b_\varepsilon(t_k,z)\Big)\langle b_\varepsilon(t_k,z),\nabla_z p^{\tau,h}_\varepsilon(t_k,t_j,z,y)\rangle dz\Big|\Big)$$

$$\leq C\Big(\bar{C}_{\eta,q}\varepsilon^{-\eta}p_c(t_j-t_i,y-x)$$

$$+ h\sum_{k\in[\![i+1,j-1]\!]}\int_{\mathbb{R}^d}\Big(\frac{\bar{C}_{\eta,q}\varepsilon^{-\eta}}{(t_k-t_i)^{1/2}} + (1+\varepsilon^{-1}\mathbb{I}_{z\in\mathcal{V}_\varepsilon(\mathcal{I})})\Big)p_c(t_k-t_i,z-x)\frac{p_c(t_j-t_k,y-z)}{(t_j-t_k)^{1/2}}dz\Big).$$

The point is now to use the Hölder inequality to exploit that the set on which $\nabla_z b_\varepsilon$ gives an explosive bound is *small*. We get:

$$|[p_\varepsilon^d \otimes_h (\tilde{L}_{\cdot,*}^\varepsilon - \tilde{L}_{\cdot}^{*,\varepsilon})^2 p_\varepsilon^{\tau,h}](t_i, t_j, x, y)| \leq C\Big(\bar{C}_{\eta,q}\varepsilon^{-\eta} p_c(t_j - t_i, y - x)$$

$$+ h \sum_{k \in [\![i+1,j-1]\!]} \frac{1}{(t_j - t_k)^{1/2}} \Big\{ \frac{\bar{C}_{\eta,q}\varepsilon^{-\eta}}{(t_k - t_i)^{1/2}} p_c(t_j - t_i, y - x)$$

$$+ \varepsilon^{-1+1/q} \Big(\int_{\mathbb{R}^d} p_c(t_k - t_i, z - x)^{\bar{q}} p_c(t_j - t_k, y - z)^{\bar{q}} dz \Big)^{1/\bar{q}} \Big\} \Big),$$

denoting by $\bar{q} > 1$ the conjugate of q, $q^{-1} + \bar{q}^{-1} = 1$. Recall now that:

$$\Big(\int_{\mathbb{R}^d} p_c(t_k - t_i, z - x)^{\bar{q}} p_c(t_j - t_k, y - z)^{\bar{q}} dz \Big)^{1/\bar{q}}$$

$$= \Big(\frac{c(t_j - t_i)}{(2\pi)(t_k - t_i)(t_j - t_k)} \Big)^{d/(2q)} \bar{q}^{-d/(2\bar{q})} p_c(t_j - t_i, y - x).$$

This yields:

$$|[p_\varepsilon^d \otimes_h (\tilde{L}_{\cdot,*}^\varepsilon - \tilde{L}_{\cdot}^{*,\varepsilon})^2 p_\varepsilon^{\tau,h}](t_i, t_j, x, y)|$$

$$\leq C\Big(\bar{C}_{\eta,q}\varepsilon^{-\eta} + \frac{1}{\alpha(q)}\varepsilon^{-1+1/q}(t_j - t_i)^{\alpha(q)}\Big) p_c(t_j - t_i, y - x)$$

$$\leq C\bar{C}_{\eta,q}\varepsilon^{-1+1/q} p_c(t_j - t_i, y - x), \tag{3.52}$$

as soon as $\varepsilon^{1-\frac{1}{q}-\eta} \leq 1$ which holds true for η small enough (remember $q > d$).

Performing now in the general case, involving derivatives of the heat kernel up to order 4, an integration by part similar to the one described for (3.39) and using the Hölder inequality as above for the terms involving derivatives of b_ε, we derive from (3.33) in Proposition 3.4, that for all $q > d$, $\eta \in (0, \alpha(q))$:

$$|[p_\varepsilon^d \otimes_h (\tilde{L}_{\cdot,*}^\varepsilon - \tilde{L}_{\cdot}^{*,\varepsilon})^2 p_\varepsilon^{\tau,h}](t_i, t_j, x, y)| \leq C\Big\{1 + \bar{C}_{\eta,q}\varepsilon^{-(1+2\eta)}(1 + \varepsilon^\eta|\ln(h)|)\Big\} p_c(t_j - t_i, y - x). \tag{3.53}$$

We thus get in whole generality, from (3.3), (3.31), (3.53) and (3.35) in Proposition 3.7:

$$|(p - p^h)(t_i, t_j, x, y)| \leq C\Big[C_q\varepsilon^{1/q} + \bar{C}_{\eta,q}h\varepsilon^{-(1+\eta)}(1 + \varepsilon^{\eta/2}|\ln(h)|)$$

$$+ \bar{C}_{\eta,q}\Big(\frac{h|\ln(h)|\varepsilon^{-(1+\eta)}}{(t_j - t_i)^{1/2}} + \frac{h^{1-\eta/2}\varepsilon^{-(1+\eta)}}{(t_j - t_i)^{1/2}} + h\varepsilon^{-2+1/q}\Big)\Big] p_c(t_j - t_i, y - x).$$

If now $d(y, \mathcal{V}_\varepsilon(\mathcal{I})) \geq 2\varepsilon$, then, from (3.36) in Proposition 3.7:

$$|(p - p^h)(t_i, t_j, x, y)| \leq C\Big[C_q\varepsilon^{1/q} + \bar{C}_{\eta,q}h\varepsilon^{-(1+2\eta)}(1 + \varepsilon^\eta|\ln(h)|)$$

$$+ \bar{C}_{\eta,q}\Big(\frac{h|\ln(h)|\varepsilon^{-(1+\eta)}}{(t_j - t_i)^{1/2}} + \frac{h^{1-\eta/2}\varepsilon^{-(1+\eta)}}{(t_j - t_i)^{1/2}} + \frac{h^{1-\eta}}{d(y, \mathcal{V}_\varepsilon(\mathcal{I}))}\Big)\Big] p_c(t_j - t_i, y - x).$$

Eventually, if we additionally have that $\sigma(t, x) = \sigma$, (3.37) in Proposition 3.7 and (3.52) yield:

$$|(p - p^h)(t_i, t_j, x, y)| \leq C\Big[C_q\varepsilon^{1/q} + \bar{C}_{\eta,q}h\varepsilon^{-1+1/q}$$

$$+ \bar{C}_{\eta,q}\Big(\frac{h|\ln(h)|\varepsilon^{-\eta}}{(t_j - t_i)^{1/2}} + \frac{h\varepsilon^{-(1+\eta)+1/q}}{(t_j - t_i)^{1/2}} + \frac{h^{1-\eta}}{d(y, \mathcal{V}_\varepsilon(\mathcal{I}))}\Big)\Big] p_c(t_j - t_i, y - x).$$

We then set $C_q \varepsilon^{1/q} = \bar{C}_{\eta,q} \frac{h\varepsilon^{-2+1/q}}{(t_j-t_i)^{1/2}}$ in the general case, i.e. for b, σ depending both on the spatial variable and without any distance condition for the final point y. If $d(y, \mathcal{V}_\varepsilon(\mathcal{I})) \geq 2\varepsilon$, we take $C_q \varepsilon^{1/q} = \bar{C}_{\eta,q} \frac{h^{1-\eta/2}\varepsilon^{-(1+\eta)}}{(t_j-t_i)^{1/2}}$ for a general σ and $C_q \varepsilon^{1/q} = \bar{C}_{\eta,q} \frac{h\varepsilon^{-(1+\eta)+1/q}}{(t_j-t_i)^{1/2}}$ if $\sigma(t,x) = \sigma$. The results can be derived as in the previous section choosing $\eta := \eta(h) = \psi(h)$, $q := q(h)$ s.t. $\alpha(q) = 2\psi(h)$. For (1.12) and (1.13), we recall as well that if $d(y, \mathcal{I}) \geq h^{1/2-\epsilon}$ for a fixed given $\epsilon > 0$ for a general σ and $d(y, \mathcal{I}) \geq h^{1-\epsilon}$ for $\sigma(t,x) = \sigma$, the condition $d(y, \mathcal{V}_\varepsilon(\mathcal{I})) \geq 2\varepsilon$ is met.

4 Proof of the technical results from Section 3

4.1 Proof of Proposition 3.4

4.1.1 Proof under ($\mathbf{A}_H$)

Let us establish the result for p_ε. We start from the parametrix representation of p_ε obtained in (3.5). In all cases, we can readily derive from (3.7) (recall that $\tilde{X}^{\varepsilon,y}$ is a non degenerate Gaussian process) and (2.6) in Proposition 2.2 that for the *main* term in the expansion for all multi-index α, $|\alpha| \in [\![1,4]\!]$:

$$|D_x^\alpha \tilde{p}_\varepsilon(s,t,x,y)| \leq \frac{C}{(t-s)^{|\alpha|/2}} p_c(t-s,y-x), \quad |D_y^\alpha \tilde{p}_\varepsilon(s,t,x,y)| \leq \frac{C\varepsilon^{-|\alpha|+\gamma}}{(t-s)^{|\alpha|/2}} p_c(t-s,y-x).$$
(4.1)

Let us now concentrate on the remainder term:

$$R_\varepsilon(s,t,x,y) \quad := \quad \sum_{i \geq 1} \tilde{p}_\varepsilon \otimes H_\varepsilon^{(i)}(s,t,x,y) = \tilde{p}_\varepsilon \otimes \Phi_\varepsilon(s,t,x,y),$$

$$\Phi_\varepsilon(s,t,x,y) \quad := \quad \sum_{i \geq 1} H_\varepsilon^{(i)}(s,t,x,y).$$

We focus on the first two inequalities in (3.32), 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 (4.1) and (3.9), 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:

$$\begin{aligned} D_x^\alpha R_\varepsilon(s,t,x,y) \quad = \quad & \lim_{\tau \to 0} \int_{s+\tau}^{(t+s)/2} du \int_{\mathbb{R}^d} D_x^\alpha \tilde{p}_\varepsilon(s,u,x,z) \Phi_\varepsilon(u,t,z,y) dz + \\ & \int_{(t+s)/2}^{t} du \int_{\mathbb{R}^d} D_x^\alpha \tilde{p}_\varepsilon(s,u,x,z) \Phi_\varepsilon(u,t,z,y) dz \\ =: \quad & \lim_{\tau \to 0} D_x^\alpha R_\varepsilon^\tau(s,t,x,y) + D_x^\alpha R_\varepsilon^f(s,t,x,y). \end{aligned}$$
(4.2)

The contribution $D_x^\alpha R_\varepsilon^f(s,t,x,y)$ does not exhibit time singularities in the integral, since on the considered integration set $u - s \geq \frac{1}{2}(t-s)$. Let us now recall the usual control on the parametrix kernel under ($\mathbf{A}_H$), see e.g. Section 2 in [KKM16]. There exist c, c_1 s.t. for all $0 \leq u < t \leq T, (z,y) \in (\mathbb{R}^d)^2$:

$$|H_\varepsilon(u,t,z,y)| \leq \frac{c_1(1 \vee T^{(1-\gamma)/2})}{(t-u)^{1-\gamma/2}} p_c(t-u, z-y).$$
(4.3)

Inequality (4.3) for H_ε then yields for all $r \in \mathbb{N}^*, 0 \le s < t \le T, (x,y) \in (\mathbb{R}^d)^2$:

$$|H_\varepsilon^{(r)}(s,t,x,y)| \le ((1 \vee T^{(1-\gamma)/2})c_1)^r \prod_{i=1}^{r-1} B(\frac{\gamma}{2}, 1+(i-1)\frac{\gamma}{2}) p_c(t-s, y-x)(t-s)^{-1+\frac{r\gamma}{2}}, \quad (4.4)$$

with the convention $\prod_{i=1}^0 = 1$. We thus derive that for all $0 \le s < t \le T, (x,y) \in (\mathbb{R}^d)^2$:

$$|\Phi_\varepsilon(s,t,x,y)| \le \frac{C}{(t-s)^{1-\gamma/2}} p_c(t-s, y-x). \quad (4.5)$$

Thus, from inequalities (4.1) and (4.5):

$$|D_x^\alpha R_\varepsilon^f(s,t,x,y)| \le \frac{C}{(t-s)^{(|\alpha|-\gamma)/2}} p_c(t-s, y-x). \quad (4.6)$$

The delicate contribution is indeed $D_x^\alpha R_\varepsilon^\tau(s,t,x,y)$ 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 for $0 \le s < u \le T$, $\Sigma_\varepsilon(s,u,w) := \int_s^u \sigma_\varepsilon \sigma_\varepsilon^*(v,w) dv$,

$$\tilde{p}_\varepsilon^w(s,u,x,z) = \frac{\exp\left(-\frac{1}{2}\langle \Sigma_\varepsilon(s,u,w)^{-1}(z-x), z-x\rangle\right)}{(2\pi)^{d/2} \det(\Sigma_\varepsilon(s,u,w))^{1/2}},$$

$$D_{x_i x_j}^2 \tilde{p}_\varepsilon^w(s,u,x,z) = \left\{\left(\Sigma_\varepsilon^{-1}(s,u,w)(z-x)\right)_i \left(\Sigma_\varepsilon^{-1}(s,u,w)(z-x)\right)_j \right.$$
$$\left. -\delta_{ij}(\Sigma_\varepsilon^{-1}(s,u,w))_{ii}\right\} \tilde{p}_\varepsilon^w(s,u,x,z), \; \forall (i,j) \in [\![1,d]\!]^2, \quad (4.7)$$

where for $q \in \mathbb{R}^d$, we denote for $i \in [\![1,d]\!]$ by q_i its ith entry. Hence, for all multi-index α, $|\alpha| = 2$:

$$\int_{\mathbb{R}^d} D_x^\alpha \tilde{p}_\varepsilon^w(s,u,x,z) dz = 0. \quad (4.8)$$

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

$$D_x^\alpha R_\varepsilon^\tau(s,t,x,y) = \int_{s+\tau}^{(s+t)/2} du \int_{\mathbb{R}^d} (D_x^\alpha \tilde{p}_\varepsilon - c_\varepsilon^\alpha)(s,u,x,z) \Phi_\varepsilon(u,t,z,y) dz$$
$$+ \int_{s+\tau}^{(s+t)/2} du \int_{\mathbb{R}^d} c_\varepsilon^\alpha(s,u,x,z)(\Phi_\varepsilon(u,t,z,y) - \Phi_\varepsilon(u,t,x,y)) dz$$
$$:= (R_\varepsilon^{\tau,1} + R_\varepsilon^{\tau,2})(s,t,x,y), \quad (4.9)$$

exploiting the centering condition (4.8) to introduce the last term of the first equality. On the one hand, the terms $D_x^\alpha \tilde{p}_\varepsilon(s,u,x,z), c_\varepsilon^\alpha(s,u,x,z)$ 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:

$$|(D_x^\alpha \tilde{p}_\varepsilon - c_\varepsilon^\alpha)(s,u,x,z)| \le \frac{C|x-z|^\gamma}{(u-s)} p_c(u-s, z-x) \le \frac{C}{(u-s)^{1-\gamma/2}} p_c(u-s, z-x),$$

yielding an integrable singularity in time so that, from (4.5):

$$|R_\varepsilon^{\tau,1}(s,t,x,y)| \le \frac{C}{(t-s)^{1-\gamma}} p_c(t-s, y-x). \quad (4.10)$$

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

$$|\Phi_\varepsilon(u, t, x, y) - \Phi_\varepsilon(u, t, z, y)| \le C \frac{|x - z|^{\gamma/2}}{(t - u)^{1 - \gamma/4}} p_c(t - u, y - z). \qquad (4.11)$$

Then, we can derive from (4.1), (4.9) and (4.11):

$$|R_\varepsilon^{\tau,2}(s, t, x, y)| \le C^2 \int_{s+\tau}^{(s+t)/2} du \int_A \frac{|x - z|^{\gamma/2}}{(u - s)} p_c(u - s, z - x) \frac{1}{(t - u)^{1 - \gamma/4}} p_c(t - u, y - z) dz$$

$$+ \frac{C}{(t - s)^{\gamma/4}} \int_{s+\tau}^{(s+t)/2} du \int_{A^C} \frac{|x - z|^{\gamma/2}}{(u - s)} p_c(u - s, z - x)\{|\Phi_\varepsilon(u, t, z, y)| + |\Phi_\varepsilon(u, t, x, y)|\} dz. \qquad (4.12)$$

From (4.5), we finally get on the considered time set:

$$|R_\varepsilon^{\tau,2}(s, t, x, y)| \le C p_c(t - s, y - x) \int_{s+\tau}^{(s+t)/2} du \frac{1}{(u - s)^{1 - \gamma/4}} \frac{1}{(t - u)^{1 - \gamma/4}}$$

$$\le \frac{C}{(t - s)^{1 - \gamma/2}} p_c(t - s, y - x),$$

which together with (4.10), (4.9), (4.6) and (4.2) gives the statement. It remains to establish (4.11). From the definition of Φ_ε and the smoothing effect of the kernel H_ε in (4.4), it suffices to prove that on the set $\bar{A} := \{z \in \mathbb{R}^d : |x - z| \le c(u' - u)^{1/2}\}$:

$$|H_\varepsilon(u, u', x, w) - H_\varepsilon(u, u', z, w)| \le C \frac{|x - z|^{\gamma/2}}{(u' - u)^{1 - \gamma/4}} p_c(u' - u, w - z), \qquad (4.13)$$

for $u' \in (u, t]$, $u \in [s, (s + t)/2]$. Observe that $\bar{A} \subset A$. Indeed, recalling that we want to establish (4.11) on A if $z \notin \bar{A}$, we get from (4.4):

$$\int_u^t du' \int_{\bar{A}^c} |H_\varepsilon(u, u', x, w) - H_\varepsilon(u, u', z, w)||(\sum_{i \ge 2} H_\varepsilon^{(i)})(u', t, w, y)| dw$$

$$\le \int_u^t du' \int_{\bar{A}^c} \frac{C}{(u' - u)^{1 - \gamma/2}} (p_c(u' - u, w - x) + p_c(u' - u, w - z))$$

$$\times \frac{|x - z|^{\gamma/2}}{(u' - u)^{\gamma/4}} \frac{C}{(t - u')^{1 - \gamma}} p_c(t - u', y - w) dw \le C \frac{|x - z|^{\gamma/2}}{(t - u)^{1 - 5\gamma/4}} p_c(t - u, y - z)$$

$$\le C \frac{|x - z|^{\gamma/2}}{(t - u)^{1 - \gamma/4}} p_c(t - u, y - z),$$

exploiting that $z \in A$, $t - u \ge \frac{1}{2}(t - s)$, and the usual convexity inequality $\frac{|y-x|^2}{t-u} \ge \frac{|y-z|^2}{2(t-u)} - \frac{|z-x|^2}{t-u} \ge \frac{|y-z|^2}{2(t-u)} - 2c^2$ for the last but one inequality. On the other hand, on $\bar{A}$ we get (4.11) from (4.13) and (4.4).

Let us turn to the proof of (4.13). We concentrate on the second derivatives in H_ε which yield the most singular contributions:

$$\text{Tr}((a_\varepsilon(u, x) - a_\varepsilon(u, w)) D_x^2 \tilde{p}_\varepsilon(u, u', x, w)) - \text{Tr}((a_\varepsilon(u, z) - a_\varepsilon(u, w)) D_x^2 \tilde{p}_\varepsilon(u, u', z, w))$$

$$= \text{Tr}((a_\varepsilon(u, x) - a_\varepsilon(u, z)) D_x^2 \tilde{p}_\varepsilon(u, u', x, w))$$

$$- \text{Tr}((a_\varepsilon(u, z) - a_\varepsilon(u, w))(D_x^2 \tilde{p}_\varepsilon(u, u', z, w) - D_x^2 \tilde{p}_\varepsilon(u, u', x, w))) =: I + II. \qquad (4.14)$$

Then, from (4.1),

$$\begin{aligned}
|I| &\leq C\frac{|x-z|^\gamma}{(u-u')}p_c(u'-u,w-x) \leq \frac{C|x-z|^{\gamma/2}}{(u-u')^{1-\gamma/4}}p_c(u'-u,w-x) \\
&\leq \frac{C|x-z|^{\gamma/2}}{(u-u')^{1-\gamma/4}}p_c(u'-u,w-z),
\end{aligned} \tag{4.15}$$

using that $z \in \bar{A}$ for the second inequality, again combined with the convexity inequality $\frac{|x-w|^2}{u'-u} \geq \frac{|z-w|^2}{2(u'-u)} - \frac{|x-z|^2}{u'-u} \geq \frac{|z-w|^2}{2(u'-u)} - c^2$ for the last one. Now, from the explicit expression of the second order derivatives in (4.7), **(A2)** and usual computations we also derive:

$$|II| \leq \frac{C|z-w|^\gamma}{(u'-u)}\frac{|z-x|^{\gamma/2}}{(u'-u)^{\gamma/4}}p_c(u'-u,w-z) \leq \frac{C|z-x|^{\gamma/2}}{(u'-u)^{1-\gamma/4}}p_c(u'-u,w-z). \tag{4.16}$$

This gives (4.13) 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 (4.2), (4.6), we only have to consider the contribution $D_x^\alpha R_\varepsilon^\tau(s,t,x,y)$. Write:

$$\begin{aligned}
D_x^\alpha R_\varepsilon^\tau(s,t,x,y) &= D_x^{\alpha-l}\int_{s+\tau}^{(t+s)/2} du \int_{\mathbb{R}^d} D_x^l \tilde{p}_\varepsilon(s,u,x,z)\Phi_\varepsilon(u,t,z,y)dz \\
&= D_x^{\alpha-l}\int_{s+\tau}^{(t+s)/2} du \int_{\mathbb{R}^d} g^{l,\varepsilon}(s,u,x,z)\Phi_\varepsilon(u,t,z,y)dz, \tag{4.17}
\end{aligned}$$

where $g^{l,\varepsilon}(s,u,x,z) := D_x^l \tilde{p}_\varepsilon(s,u,x,z)$. Let us write introducing the cancellation term c_ε^l introduced after (4.8):

$$\begin{aligned}
D_x^\alpha R_\varepsilon^\tau(s,t,x,y) &= D_x^{\alpha-l}\int_{s+\tau}^{(s+t)/2} du \int_{\mathbb{R}^d} (g^{l,\varepsilon}-c_\varepsilon^l)(s,u,x,z)\Phi_\varepsilon(u,t,z,y)dz \\
&+ D_x^{\alpha-l}\int_{s+\tau}^{(s+t)/2} du \int_{\mathbb{R}^d} c_\varepsilon^l(s,u,x,z)(\Phi_\varepsilon(u,t,z,y)-\Phi_\varepsilon(u,t,x,y))dz \\
&= D_x^{\alpha-l}\int_{s+\tau}^{(s+t)/2} du \int_{\mathbb{R}^d} (g^{l,\varepsilon}-c_\varepsilon^l)(s,u,x,x+z)\Phi_\varepsilon(u,t,x+z,y)dz \\
&+ D_x^{\alpha-l}\int_{s+\tau}^{(s+t)/2} du \int_{\mathbb{R}^d} c_\varepsilon^l(s,u,x,x+z)(\Phi_\varepsilon(u,t,x+z,y)-\Phi_\varepsilon(u,t,x,y))dz. \tag{4.18}
\end{aligned}$$

The purpose of that change of variable, already performed in [KM02], is that we get integrable time singularities in the contributions $D_x^{\alpha-l}(g^{l,\varepsilon}-c_\varepsilon^l)(s,u,x,x+z)$. Anyhow, the mollified coefficients $b_\varepsilon, \sigma_\varepsilon$ have explosive derivatives. From the definition of $g^{l,\varepsilon}$ and (2.6) one easily gets that there exists c,C s.t. for all α, $|\alpha| \leq 4$:

$$\begin{aligned}
|D_x^{\alpha-l}(g^{l,\varepsilon}-c_\varepsilon^l)(s,u,x,x+z)| &\leq \frac{C\varepsilon^{-|\alpha-l|}}{(u-s)^{1-\gamma/2}}p_c(u-s,z), \\
|D_x^{\alpha-l}c_\varepsilon^l(s,u,x,x+z)| &\leq \frac{C\varepsilon^{-|\alpha-l|+\gamma}}{(u-s)}p_c(u-s,z).
\end{aligned} \tag{4.19}$$

From (4.18) and (4.19) it thus remains to control the terms $D_x^{\alpha-l}\Phi_\varepsilon(u,t,z+x,y)$ and $D_x^{\alpha-l}(\Phi_\varepsilon(u,t,x+z,y)-\Phi_\varepsilon(u,t,x,y))$ which are the most singular ones in $D_x^\alpha R_\varepsilon^\tau(s,t,x,y)$. To this end, we will establish by induction that the following control holds:

$$\exists c,C, \ \forall 0 \leq s < t \leq T, \ (x,y) \in (\mathbb{R}^d)^2, \ \forall \beta, \ |\beta| \leq 3, \ |D_x^\beta H_\varepsilon^{(i)}(s,t,x,y)| \leq$$

$$\frac{C^i \varepsilon^{-|\beta|}}{(t-s)^{|\beta|/2}}(t-s)^{-1+i\gamma/2}\prod_{j=1}^{i-1}B(\frac{\gamma}{2},1+(j-1)\frac{\gamma}{2})p_c(t-s,y-x), \qquad (4.20)$$

with the convention that $\prod_{j=1}^0 = 1$. Observe first that for $|\beta| = 0$ (no derivation), estimate (4.20) readily follows from (4.4). Let us now suppose $|\beta| > 0$. Observe from the definition of H_ε that (4.20) is satisfied for $i = 1$. Let us assume it holds for a given i and let us prove it for $i + 1$. Write again:

$$D_x^\beta H_\varepsilon^{(i+1)}(s,t,x,y) = \int_{(s+t)/2}^t du \int_{\mathbb{R}^d} D_x^\beta H_\varepsilon(s,u,x,z) H_\varepsilon^{(i)}(u,t,z,y)dz$$

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

The term $R_1^{i,\beta}$ is easily controlled by (4.20) for $\beta = 0$ and the induction hypothesis. Observe also that, from Proposition 2.2 one derives similarly to (4.19) that:

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

Together with the induction hypothesis and the Leibniz rule for differentiation, this allows to control $R_2^{i,\beta}$. The controls on $\{R_j^{i,\beta}\}_{j\in\{1,2\}}$ give (4.20) for $i+1$. We eventually derive (reminding that $|l| = 2$):

$$|D_x^{\alpha-l}\Phi_\varepsilon(u,t,x+z,y)| \le \frac{C}{(t-u)^{(|\alpha|-2)/2}}\frac{\varepsilon^{-|\alpha|+2}}{(t-u)^{1-\gamma/2}}p_c(t-u,y-(x+z)). \qquad (4.21)$$

The spatial Hölder continuity of the derivatives of the kernel Φ_ε 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 $|z| \le c(t-u)^{1/2}$:

$$|D_x^{\alpha-l}\Phi_\varepsilon(u,t,x+z,y) - D_x^{\alpha-l}\Phi_\varepsilon(u,t,x,y)|$$

$$\le \frac{C|z|^{\gamma/2}}{(t-u)^{(|\alpha|-2)/2}}\frac{\varepsilon^{-|\alpha|+2}}{(t-u)^{1-\gamma/4}}p_c(t-u,y-(x+z)),$$

which together with (4.21), (4.19), (4.18) gives (proceeding as above for $|z| \ge c(t-u)^{1/2}$):

$$|D_x^\alpha R_\varepsilon^\tau(s,t,x,y)| \le \frac{C\varepsilon^{-|\alpha|+2}(t-s)^{\gamma/2}}{(t-s)^{|\alpha|/2}}p_c(t-s,y-x).$$

The second equation of (3.32) follows for $\bar{p}_\varepsilon = p_\varepsilon$ from the above control and (4.6), (4.2). Observe that the control for the derivative w.r.t. y has additional singularity in ε. This is clear since we directly differentiate the frozen mollified coefficients. Now the statements readily hold for p_ε^d, since the *integration in time* played no role in the previous computations. For p_ε^h, the only point that should be totally justified is the smoothing property and Hölder continuity of the discrete Kernel $\Phi_\varepsilon^h(t_i,t_j,x,y) := \sum_{r=1}^{j-i} H_\varepsilon^{h,(r)}(t_i,t_j,x,y)$. The smoothing property, equivalent of (4.20), has been investigated in [LM10]. The spatial Hölder continuity can be derived as above.

4.1.2 Proof under (A$_{PS}$)

Let us now turn to the proof of the heat kernel bounds for p_ε under (A$_{PS}$), which almost follows the same lines. Observe first that the result for $|\alpha| = 1$ still follows from (4.1) and (3.9). The key point is again that the derivative of the Gaussian kernel yields an

integrable singularity. For $|\alpha| = 2$, we still separate the contribution $R_\varepsilon(s, t, x, y)$ as in (4.2) and again focus on $\lim_{\tau \to 0} D_x^\alpha R_\varepsilon^\tau(s, t, x, y)$ which is the only term yielding a potential singularity. With the notations of (4.9), it is sufficient to investigate $R_\varepsilon^{\tau,2}(s, t, x, y)$. Indeed, under $(\mathbf{A}_{PS})$, equation (4.10) actually holds with $\gamma = 1$. We recall that to control $R_\varepsilon^{\tau,2}(s, t, x, y)$, the key estimate was (4.11). We aim at proving the different control, for all $u \in [s, t)$, for all $\eta \in (0, 1]$:

$$|\Phi_\varepsilon(u, t, x, y) - \Phi_\varepsilon(u, t, z, y)| \leq C\varepsilon^{-\eta} \frac{|x - z|^\eta}{(t - u)^{3/4}} p_c(t - u, y - z), \qquad (4.22)$$

on $A := \{z \in \mathbb{R}^d : |x - z| \leq c[(t - s)^{1/2} \wedge \varepsilon]\}$. Then, we can derive from (4.1), (4.9) and (4.22):

$$
\begin{aligned}
|R_\varepsilon^{\tau,2}(s, t, x, y)| \leq{} & C\varepsilon^{-\eta} \int_{s+\tau}^{(s+t)/2} du \int_A \frac{|x - z|^\eta}{(u - s)} p_c(u - s, z - x) \frac{1}{(t - u)^{3/4}} p_c(t - u, y - z) dz \\
& + C((t - s)^{1/2} \wedge \varepsilon)^{-\eta} \int_{s+\tau}^{(s+t)/2} du \int_{A^C} \frac{|x - z|^\eta}{(u - s)} \\
& \qquad\qquad \times p_c(u - s, z - x)\{|\Phi_\varepsilon(u, t, z, y)| + |\Phi_\varepsilon(u, t, x, y)|\} dz.
\end{aligned}
$$

Since the drift b_ε is uniformly bounded, uniformly in $\varepsilon \in [0, 1]$, we have under $(\mathbf{A}_{PS})$ the following usual control on the parametrix kernel (see e.g. Section 2 in [KKM16]):

$$|H_\varepsilon(u, t, z, y)| \leq \frac{c_1}{(t - u)^{1/2}} p_c(t - u, y - z). \qquad (4.23)$$

Equation (4.23) for H_ε then yields

$$|H_\varepsilon^{(r)}(s, t, x, y)| \leq c_1^r \prod_{i=1}^{r-1} B\left(\frac{1}{2}, 1 + (i - 1)\frac{1}{2}\right) p_c(t - s, y - x)(t - s)^{-1 + \frac{r}{2}}, \qquad (4.24)$$

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

$$
\begin{aligned}
|R_\varepsilon^{\tau,2}(s, t, x, y)| &\leq C((t - s)^{1/2} \wedge \varepsilon)^{-\eta} p_c(t - s, y - x) \int_{s+\tau}^{(s+t)/2} du \frac{1}{(u - s)^{1 - \eta/2}} \frac{1}{(t - u)^{3/4}} \\
&\leq \frac{C((t - s)^{1/2} \wedge \varepsilon)^{-\eta}}{\eta(t - s)^{3/4 - \eta/2}} p_c(t - s, y - x).
\end{aligned}
$$

It remains to establish (4.22). From the definition of Φ_ε and the smoothing effect of the kernel H_ε in (4.24), it suffices to prove that on $\bar{A} := \{z \in \mathbb{R}^d : |x - z| \leq c[(u' - u)^{1/2} \wedge \varepsilon]\}$:

$$|H_\varepsilon(u, u', x, w) - H_\varepsilon(u, u', z, w)| \leq C\left\{\frac{|x - z|^\eta}{(u' - u)^{3/4}} ((u' - u)^{1/2} \wedge \varepsilon)^{-\eta}\right\} p_c(u' - u, w - z), \quad (4.25)$$

for $u' \in (u, t]$, $u \in [s, (s + t)/2]$. The contributions associated with $z \in \bar{A}^C$ can be handled as above. To establish the above control we focus on the first order terms involving the regularized coefficient with initial discontinuities. Indeed the second order contribution can be analyzed as in (4.14), (4.15), (4.16), taking $\gamma = 1$ in those expressions. In particular, the time singularity in $(u - u')^{3/4}$ in (4.25) precisely comes from those terms. Recalling that under $(\mathbf{A}_{PS})$ the driftless proxy does not depend on ε (since the diffusion is smooth, see (3.10) in which one has $\sigma_\varepsilon = \sigma$ under $(\mathbf{A}_{PS})$), we denote its density by $\tilde{p}$ and write:

$$\langle b_\varepsilon(u, x), D_x \tilde{p}(u, u', x, w)\rangle - \langle b_\varepsilon(u, z), D_x \tilde{p}(u, u', z, w)\rangle$$

$$= \langle b_\varepsilon(u,x) - b_\varepsilon(u,z), D_x \tilde{p}(u,u',x,w) \rangle$$
$$+ \langle b_\varepsilon(u,z), D_x \tilde{p}(u,u',x,w) - D_x \tilde{p}(u,u',z,w) \rangle := I + II.$$

On the one hand, from the mean value theorem and recalling that $|D_x b_\varepsilon|_\infty \leq C\varepsilon^{-1} \leq C((u'-u)^{1/2} \wedge \varepsilon)^{-1}$ we get:

$$
\begin{aligned}
|I| &\leq \frac{C}{(u'-u)^{1/2}} \Big\{ 2|b|_\infty \Big(\frac{|x-z|}{(u'-u)^{1/2} \wedge \varepsilon} \Big)^\eta \mathbb{I}_{|x-z|>(u'-u)^{1/2}\wedge\varepsilon} \\
&\qquad\qquad + \varepsilon^{-1}|x-z|\mathbb{I}_{|x-z|\leq(u'-u)^{1/2}\wedge\varepsilon} \Big\} p_c(u'-u, w-x) \\
&\leq C((u'-u)^{1/2} \wedge \varepsilon)^{-\eta} \frac{|x-z|^\eta}{(u'-u)^{1/2}} p_c(u'-u, w-x) \\
&\leq C((u'-u)^{1/2} \wedge \varepsilon)^{-\eta} \frac{|x-z|^\eta}{(u'-u)^{1/2}} p_c(u'-u, w-z),
\end{aligned}
$$

using again a convexity inequality for the last control, recalling that $z \in \bar{A}$. On the other hand still from the mean value Theorem and usual controls on the derivatives of the Gaussian density:

$$
\begin{aligned}
|II| &\leq \frac{C|x-z|}{(u'-u)} \int_0^1 p_c(u'-u, w - \{z + \lambda(x-z)\}) d\lambda \leq \frac{C|x-z|^\eta}{(u'-u)^{(1+\eta)/2}} p_c(u'-u, w-z) \\
&\leq \frac{C|x-z|^\eta}{(u'-u)^{1/2}((u'-u)^{1/2} \wedge \varepsilon)^\eta} p_c(u'-u, w-z).
\end{aligned}
$$

The above estimates give (4.25) and concludes the proof for $|\alpha| = 2$.

Let us turn to $|\alpha| \geq 3$. The idea is again to proceed as under $(\mathbf{A}_H)$, up to a suitable modification of the key estimate (4.20) which can now be localized and becomes for all $q > d$:

$$\exists c, C, \ \forall 0 \leq s < t \leq T, \ (x,y) \in (\mathbb{R}^d)^2, \ \forall \beta, \ |\beta| \leq 3,$$

$$|D_x^\beta H_\varepsilon^{(i)}(s,t,x,y)| \leq \frac{C^i(\varepsilon^{-|\beta|}\mathbb{I}_{x\in V_\varepsilon(\mathcal{I})} + \varepsilon^{-|\beta|+1/q})}{(t-s)^{|\beta|/2}} (t-s)^{-1+i\alpha(q)}$$

$$\times \prod_{j=1}^{i-1} B(\alpha(q), \alpha(q)j) p_c(t-s, y-x), \quad \alpha(q) = \frac{1}{2}\Big(1 - \frac{d}{q}\Big), \quad (4.26)$$

with $\prod_{j=1}^0 = 1$. We again proceed by induction. Observe first that for $|\beta| = 0$ (no derivation), estimate (4.26) readily follows from (4.23). Let us now suppose $|\beta| > 0$. Observe as well from the definition of H_ε that (4.26) is satisfied for $i = 1$. Let us assume it holds for a given i and let us prove it for $i + 1$. Write again:

$$D_x^\beta H_\varepsilon^{(i+1)}(s,t,x,y) = \int_{(s+t)/2}^t du \int_{\mathbb{R}^d} D_x^\beta H_\varepsilon(s,u,x,z) H_\varepsilon^{(i)}(u,t,z,y) dz$$

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

The term $R_1^{i,\beta}$ is easily controlled by (4.26) that holds from the induction hypothesis for $i_0 = 1$ (direct differentiation of H_ε) and $\beta = 0$ for the considered i (no differentiation of $H_\varepsilon^{(i)}$). Observe also that, similarly to (4.19), one has:

$$|D_x^\beta H_\varepsilon(s,u,x,x+z)| \leq \frac{C(\varepsilon^{-|\beta|}\mathbb{I}_{x\in V_\varepsilon(\mathcal{I})} + 1)}{(u-s)^{1/2}} p_c(u-s, z). \quad (4.27)$$

Now, from the Leibniz rule for differentiation, (4.27) and the induction hypothesis, we have:

$$|R_2^{i,\beta}(s,t,x,y)|$$

$$\leq C^{i+1}\prod_{j=1}^{i-1}B(\alpha(q),\alpha(q)j)\Big\{\sum_{\tilde{\beta},|\tilde{\beta}|\leq|\beta|}\left(\begin{array}{c}|\beta|\\|\tilde{\beta}|\end{array}\right)\int_s^{(s+t)/2}\frac{du(t-u)^{-1+i\alpha(q)}}{(u-s)^{1/2}(t-u)^{(|\beta|-|\tilde{\beta}|)/2}}$$

$$\times\int_{\mathbb{R}^d}p_c(u-s,z)(\varepsilon^{-|\tilde{\beta}|}\mathbb{I}_{x\in V_\varepsilon(\mathcal{I})}+1)(\varepsilon^{-(|\beta|-|\tilde{\beta}|)}\mathbb{I}_{x+z\in V_\varepsilon(\mathcal{I})}$$

$$+\varepsilon^{-(|\beta|-|\tilde{\beta}|)+1/q})p_c(t-u,y-x-z)dz\Big\}$$

$$\leq C^{i+1}\prod_{j=1}^{i-1}B(\alpha(q),\alpha(q)j)\Big\{\sum_{\tilde{\beta},|\tilde{\beta}|\leq|\beta|}\left(\begin{array}{c}|\beta|\\|\tilde{\beta}|\end{array}\right)\int_s^{(s+t)/2}\frac{du(t-u)^{-1+i\alpha(q)}}{(u-s)^{1/2}(t-u)^{(|\beta|-|\tilde{\beta}|)/2}}$$

$$\times[p_c(t-s,y-x)(\varepsilon^{-|\beta|}\mathbb{I}_{x\in V_\varepsilon(\mathcal{I})}+\varepsilon^{-|\beta|+1/q})$$

$$+\varepsilon^{-|\beta|+1/q}(\int_{\mathbb{R}^d}p_c(u-s,z)^{\bar{q}}p_c(t-u,y-x-z)^{\bar{q}}dz)^{1/\bar{q}}\Big\},\quad (4.28)$$

denoting by $\bar{q}>1$ the conjugate of q, $q^{-1}+\bar{q}^{-1}=1$ (see also Section 3.4.2 for similar arguments). Recall now that:

$$(\int_{\mathbb{R}^d}p_c(u-s,z)^{\bar{q}}p_c(t-u,y-x-z)^{\bar{q}}dz)^{1/\bar{q}}$$

$$=\Big(\frac{c(t-s)}{(2\pi)(u-s)(t-u)}\Big)^{d/(2q)}\bar{q}^{-d/(2\bar{q})}p_c(t-s,y-x)\leq C(u-s)^{-d/(2q)}p_c(t-s,y-x),$$

for $u\in[s,(s+t)/2]$. Hence,

$$|R_2^{i,\beta}(s,t,x,y)|\leq\frac{C^{i+1}}{(t-s)^{|\beta|/2}}\prod_{j=1}^{i-1}B(\alpha(q),\alpha(q)j)\Big\{\int_s^{(s+t)/2}\frac{du(t-u)^{-1+i\alpha(q)}}{(u-s)^{1/2(1+d/q)}}\Big\}$$

$$\times p_c(t-s,y-x)(\varepsilon^{-|\beta|}\mathbb{I}_{x\in V_\varepsilon(\mathcal{I})}+\varepsilon^{-|\beta|+1/q})$$

$$\leq\frac{C^{i+1}}{(t-s)^{|\beta|/2}}\prod_{j=1}^{i-1}B(\alpha(q),\alpha(q)j)(t-s)^{-1+(i+1)\alpha(q)}\int_0^{1/2}(1-u)^{-1+i\alpha(q)}u^{-1+\alpha(q)}du$$

$$\times p_c(t-s,y-x)(\varepsilon^{-|\beta|}\mathbb{I}_{x\in V_\varepsilon(\mathcal{I})}+\varepsilon^{-|\beta|+1/q}).$$

The controls on $\{R_j^{i,\beta}\}_{j\in\{1,2\}}$ give (4.26) for $i+1$.

Estimate (4.26) yields for every multi-index l, $|l|=2$:

$$|D_x^{\alpha-l}\Phi_\varepsilon(u,t,x+z,y)|\leq\frac{C_q}{(t-u)^{(|\alpha|-2)/2}}\frac{\mathbb{I}_{x+z\in V_\varepsilon(\mathcal{I})}\varepsilon^{-|\alpha|+2}+\varepsilon^{-|\alpha|+2+1/q}}{(t-u)^{1-\alpha(q)}}p_c(t-u,y-(x+z)).$$
$$(4.29)$$

The spatial Hölder continuity of the derivatives of the kernel Φ_ε 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 $|z|\leq c\{(t-u)^{1/2}\wedge\varepsilon\}$ for all $\eta\in(0,1]$:

$$|D_x^{\alpha-l}\Phi_\varepsilon(u,t,x+z,y)-D_x^{\alpha-l}\Phi_\varepsilon(u,t,x,y)|$$

$$\leq\frac{C_q\varepsilon^{-\eta}|z|^\eta}{(t-u)^{(|\alpha|-2)/2}}\frac{\varepsilon^{-|\alpha|+2}\mathbb{I}_{x+z\in V_\varepsilon(\mathcal{I})}+\varepsilon^{-|\alpha|+2+1/q}}{(t-u)^{1-\alpha(q)+\eta/2}}p_c(t-u,y-(x+z)).\quad (4.30)$$

Now, equation (4.18) still holds under $(\mathbf{A}_{PS})$, with $g^{l,\varepsilon}=g^l,c^l_\varepsilon=c^l$, i.e. the driftless proxy does not depend on ε. Also, the smoothness assumption on σ allows to improve

(4.19). Precisely, there exist c, C s.t. for all α, $|\alpha| \leq 4$:

$$|D_x^{\alpha-l}(g^l - c^l)(s, u, x, x+z)| \leq \frac{C}{(u-s)^{1/2}} p_c(u-s, z), \ |D_x^{\alpha-l}c^l(s, u, x, x+z)|$$

$$\leq \frac{C}{(u-s)} p_c(u-s, z),$$

which together with (4.29), (4.30), (4.18) and choosing $\alpha(q) > \eta$ gives (proceeding as above for $|z| \geq c\{(t-u)^{1/2} \wedge \varepsilon\}$):

$$|D_x^\alpha R_\varepsilon^\tau(s, t, x, y)| \leq \frac{C_{\eta,q}\varepsilon^{-(\eta+|\alpha|)+2+1/q}(t-s)^{\eta/2}}{(t-s)^{|\alpha|/2}} p_c(t-s, y-x).$$

The controls on the derivatives w.r.t. to the forward variables are derived similarly. We here simply illustrate on the first term $\tilde{p}_\varepsilon \otimes H^\varepsilon(s, t, x, y)$ of the parametrix series how the derivatives must be handled. The stated controls would follow from inductions similar to the previous ones. Write for a given multi-index β:

$$D_y^\beta\Big(\tilde{p}_\varepsilon \otimes H_\varepsilon(s, t, x, y)\Big)$$

$$= \int_s^{(s+t)/2} du \int_{\mathbb{R}^d} \tilde{p}(s, u, x, z) D_y^\beta\{\langle b_\varepsilon(u, z), D_z\tilde{p}(u, t, z, y)\rangle$$

$$+ \frac{1}{2}\mathrm{Tr}\{(a(u, z) - a(u, y))D_z^2\tilde{p}(u, t, z, y)\}\}dz +$$

$$\lim_{\tau\downarrow 0}\int_{(s+t)/2}^{t-\tau} du \int_{\mathbb{R}^d} D_y^\beta\Big(\tilde{p}(s, u, x, z)[\langle b_\varepsilon(u, z), D_z\tilde{p}(u, t, z, y)\rangle$$

$$+ \frac{1}{2}\mathrm{Tr}\{(a(u, z) - a(u, y))D_z^2\tilde{p}(u, t, z, y)\}]\Big)dz := (D_1^\beta + D_2^\beta)(s, t, x, y).$$

We readily get from the controls of (4.1) that:

$$|D_1^\beta(s, t, x, y)| \leq \frac{C}{(t-s)^{(|\beta|-1)/2}} p_c(t-s, y-x), \tag{4.31}$$

which is the expected control. Since a is smooth the terms involving the second derivatives w.r.t. z in D_2^β can be handled performing the change of variables $z' = z + y$ as above (see also [KM02] under the current smoothness assumption on the diffusion coefficient). Let us thus focus on the contribution:

$$D_{21}^\beta(s, t, x, y) := \lim_{\tau\downarrow 0}\int_{(s+t)/2}^{t-\tau} du \int_{\mathbb{R}^d} D_y^\beta\Big(\tilde{p}(s, u, x, z)\langle b_\varepsilon(u, z), D_z\tilde{p}(u, t, z, y)\rangle\Big)dz.$$

Consider first the case $|\beta| = 1$. Write:

$$D_{21}^{\beta,\tau}(s, t, x, y) := \int_{(s+t)/2}^{t-\tau} du \int_{\mathbb{R}^d} D_y^\beta\Big(\tilde{p}(s, u, x, z)\langle b_\varepsilon(u, z), D_z\tilde{p}(u, t, z, y)\rangle\Big)dz$$

$$= \int_{(s+t)/2}^{t-\tau} du \int_{\mathbb{R}^d} \tilde{p}(s, u, x, z)\langle b_\varepsilon(u, z), D_y^\beta D_z\tilde{p}(u, t, z, y)\rangle dz$$

$$= \int_{(s+t)/2}^{t-\tau} du \int_{\mathbb{R}^d} [\tilde{p}(s, u, x, z) - \tilde{p}(s, u, x, y)]\langle b_\varepsilon(u, z), D_y^\beta D_z\tilde{p}(u, t, z, y)\rangle dz$$

$$+ \int_{(s+t)/2}^{t-\tau} du \int_{\mathbb{R}^d} \tilde{p}(s, u, x, y)\langle b_\varepsilon(u, z) - b_\varepsilon(u, y), D_y^\beta D_z\tilde{p}(u, t, z, y)\rangle dz$$

$$=: [D_{211}^{\beta,\tau} + D_{212}^{\beta,\tau}](s, t, x, y), \tag{4.32}$$

recalling that for all $y \in \mathbb{R}^d$, $\int_{\mathbb{R}^d} D_z\tilde{p}(u, t, z, y)dz = 0$, so that $D_y^\beta \int_{\mathbb{R}^d} D_z\tilde{p}(u, t, z, y)dz = 0$, for the last but one equality. Still from the controls of (4.1), we readily get:

$$|D_{211}^{\beta;\tau}|$$

$$\leq \frac{C}{(t-s)^{1/2}} \int_{(s+t)/2}^{t-\tau} du \int_{\mathbb{R}^d} |z-y| \Big\{ \int_0^1 p_c(u-s, y-x+\lambda(z-y))d\lambda \mathbb{I}_{|z-y|\leq(t-s)^{1/2}}$$

$$+ (p_c(u-s, z-x) + p_c(u-s, y-x))\mathbb{I}_{|z-y|>(t-s)^{1/2}} \Big\} \frac{1}{(t-u)} p_c(t-u, y-z)dz$$

$$\leq C p_c(t-s, y-x).$$

On the other hand:

$$|D_{212}^{\beta;\tau}|(s,t,x,y)| \leq C p_c(t-s, y-x) \int_{(s+t)/2}^{t-\tau} du \int_{\mathbb{R}^d} \Big\{ \mathbb{I}_{|z-y|\leq\varepsilon} \frac{|z-y|}{\varepsilon} + \mathbb{I}_{|z-y|>\varepsilon} \frac{|z-y|^\eta}{\varepsilon^\eta} \Big\}$$

$$\times \frac{1}{(t-u)} p_c(t-u, y-z)dz \leq \frac{C}{\eta} \varepsilon^{-\eta} (t-s)^{\eta/2} p_c(t-s, y-x), \ \eta \in (0,1].$$

We therefore eventually derive from the above controls, (4.32) and (4.31) that for $|\beta| = 1$

$$D_y^\beta \Big(\tilde{p} \otimes H^\varepsilon(s,t,x,y) \Big) \leq C p_c(t-s, y-x)\{1 + \frac{C}{\eta} \varepsilon^{-\eta} (t-s)^{\eta/2}\}.$$

Take now $|\beta| \geq 2$, and let l be a multi-index s.t. $|l| = 1$ and $\beta - l \geq 0$. Set for all $0 \leq u < t \leq T, (z,y) \in \mathbb{R}^d$, $g_l(u,t,z,y) := D_y^l D_z \tilde{p}(u,t,z,y)$. Observe that there exists $C \geq 1$, $|g_l(u,t,z,y)| \leq C(t-u)^{-1} p_c(t-u, y-z)$ and also, similarly to (4.27), for all multi-index $\tilde{\beta}$, $|\tilde{\beta}| \leq 4$, $|D_y^{\tilde{\beta}} g_l(u,t,z+y,y)| \leq C(t-u)^{-1} p_c(t-u, z)$. Rewrite now from (4.32):

$$D_{21}^{\beta;\tau}(s,t,x,y) = D_y^{\beta-l} \int_{(s+t)/2}^{t-\tau} du \int_{\mathbb{R}^d} (\tilde{p}(s,u,x,z) - \tilde{p}(s,u,x,y))\langle b_\varepsilon(u,z), g_l(u,t,z,y)\rangle \Big) dz$$

$$+ D_y^{\beta-l} \int_{(s+t)/2}^{t-\tau} du \int_{\mathbb{R}^d} \tilde{p}(s,u,x,y)\langle b_\varepsilon(u,z) - b_\varepsilon(u,y), g_l(u,t,z,y)\rangle dz,$$

recalling that $\int_{\mathbb{R}^d} g_l(u,t,z,y)dz = 0$ for the last equality. Now,

$$|D_{21}^{\beta;\tau}(s,t,x,y)| =$$

$$\Big| D_y^{\beta-l} \int_{(s+t)/2}^{t-\tau} du \int_{\mathbb{R}^d} (\tilde{p}(s,u,x,z+y) - \tilde{p}(s,u,x,y))\langle b_\varepsilon(u,z+y), g_l(u,t,z+y,y)\rangle \Big) dz$$

$$+ D_y^{\beta-l} \int_{(s+t)/2}^{t-\tau} du \int_{\mathbb{R}^d} \tilde{p}(s,u,x,y)\langle b_\varepsilon(u,z+y) - b_\varepsilon(u,y), g_l(u,t,z+y,y)\rangle dz \Big|$$

$$\leq C \sum_{\substack{\beta_1,\beta_2,\beta_3, \\ \sum_{i=1}^3 |\beta_i| = |\beta|-1}} |\beta_1|, |\beta_2|, |\beta_3|)! \sum_{i=1}^d \int_{(s+t)/2}^{t-\tau} du \Big(\int_{\mathbb{R}^d} |D_y^{\beta_1} \tilde{p}(s,u,x,z+y) - D_y^{\beta_1} \tilde{p}(s,u,x,y)|$$

$$\times |D_y^{\beta_2} b_\varepsilon^i(u,z+y)| |D_y^{\beta_3} g_l^i(u,t,z+y,y)| dz$$

$$+ \int_{(s+t)/2}^{t-\tau} du \int_{\mathbb{R}^d} |D_y^{\beta_1} \tilde{p}(s,u,x,y)| |D_y^{\beta_2} b_\varepsilon^i(u,z+y) - D_y^{\beta_2} b_\varepsilon^i(u,y)| |D_y^{\beta_3} g_l^i(u,t,z+y,y)| dz \Big),$$

where $(|\beta_1|, |\beta_2|, |\beta_3|)! = \frac{(\sum_{i=1}^3 |\beta_i|)!}{\prod_{i=1}^3 (|\beta_i|!)}$ stands for the multinomial coefficients with entries $(|\beta|_i)_{i\in[\![1,3]\!]}$. Recall as well from (2.9) that we have the following control:

$$|D_y^{\beta_2} b_\varepsilon(u,z+y) - D_y^{\beta_2} b_\varepsilon(u,y)|$$

$$\leq C \Big((1 + \varepsilon^{-|\beta|+1} (\mathbb{I}_{y+z\in V_\varepsilon(\mathcal{I})} + \mathbb{I}_{y\in V_\varepsilon(\mathcal{I})}) \frac{|z|}{(t-s)^{1/2}} \mathbb{I}_{|z|>(t-s)^{1/2}}$$

$$+\big((1+\varepsilon^{-|\beta|}\mathbb{I}_{y\in V_{2\varepsilon}(\mathcal{I})})|z|\mathbb{I}_{|z|\le\varepsilon}$$

$$+\big(1+\varepsilon^{-|\beta|+1}(\mathbb{I}_{y+z\in V_\varepsilon(\mathcal{I})}+\mathbb{I}_{y\in V_\varepsilon(\mathcal{I})})(\frac{|z|}{\varepsilon})^\eta\mathbb{I}_{|z|>\varepsilon})\mathbb{I}_{|z|\le(t-s)^{1/2}}\big)$$

$$\le C\Big(\big(1+\varepsilon^{-|\beta|+1}(\mathbb{I}_{y+z\in V_\varepsilon(\mathcal{I})}+\mathbb{I}_{y\in V_\varepsilon(\mathcal{I})})\frac{|z|}{(t-s)^{1/2}}\mathbb{I}_{|z|>(t-s)^{1/2}}$$

$$+\big((1+\varepsilon^{-|\beta|+1}\mathbb{I}_{y\in V_{2\varepsilon}(\mathcal{I})})\mathbb{I}_{|z|\le\varepsilon}+\big(1+\varepsilon^{-|\beta|+1}(\mathbb{I}_{y+z\in V_\varepsilon(\mathcal{I})}$$

$$+\mathbb{I}_{y\in V_\varepsilon(\mathcal{I})})\mathbb{I}_{|z|>\varepsilon})\big)(\frac{|z|}{\varepsilon})^\eta\mathbb{I}_{|z|\le(t-s)^{1/2}}\Big).$$

Thus,

$$|D_{21}^{\beta;\tau}(s,t,x,y)|\le C\sum_{\substack{\beta_1,\beta_2,\beta_3,\\\sum_{i=1}^3|\beta_i|=|\beta|-1}}|\beta_1|,|\beta_2|,|\beta_3|)!$$

$$\times\int_{(s+t)/2}^{t-\tau}du\int_{\mathbb{R}^d}\int_0^1 d\lambda\Big\{\frac{|z|}{(u-s)^{(|\beta_1|+1)/2}}p_c(u-s,y-x+\lambda z)\mathbb{I}_{|z|\le(t-s)^{1/2}}$$

$$+\frac{|z|}{(t-s)^{(|\beta_1|+1)/2}}(p_c(u-s,y+z-x)+p_c(u-s,y-x))\mathbb{I}_{|z|>(t-s)^{1/2}}\Big\}$$

$$\times(1+\varepsilon^{-|\beta|+1}\mathbb{I}_{y+z\in V_\varepsilon(\mathcal{I})})\frac{1}{t-u}p_c(t-u,z)dz$$

$$+\int_{(s+t)/2}^{t-\tau}du\int_{\mathbb{R}^d}\frac{p_c(t-s,y-x)}{(t-s)^{|\beta_1|/2}}\Big\{(1+\varepsilon^{-|\beta|+1}(\mathbb{I}_{y+z\in V_\varepsilon(\mathcal{I})}+\mathbb{I}_{y\in V_\varepsilon(\mathcal{I})}))\frac{|z|}{(t-s)^{1/2}}\mathbb{I}_{|z|>(t-s)^{1/2}}$$

$$+\{(1+\varepsilon^{-|\beta|+1}\mathbb{I}_{y\in V_{2\varepsilon}(\mathcal{I})})\mathbb{I}_{|z|\le\varepsilon}+(1+\varepsilon^{-|\beta|+1}(\mathbb{I}_{y+z\in V_\varepsilon(\mathcal{I})}+\mathbb{I}_{y\in V_\varepsilon(\mathcal{I})}))\mathbb{I}_{|z|>\varepsilon}\}$$

$$\mathbb{I}_{|z|\le(t-s)^{1/2}}(\frac{|z|}{\varepsilon})^\eta\Big\}\frac{1}{t-u}p_c(t-u,z)dz$$

$$\le\frac{C}{(t-s)^{|\beta|}}p_c(t-s,y-x)(1+\varepsilon^{-|\beta|+1}(\frac{\varepsilon^{-\eta}}{\eta}+\frac{\varepsilon^{1/q}}{\alpha(q)})),$$

recalling that the contribution in $\frac{\varepsilon^{1/q}}{\alpha(q)}$ comes from the terms involving $\frac{\mathbb{I}_{y+z\in V_\varepsilon(\mathcal{I})}}{(t-u)^{1/2}}$ that can be handled using Hölder inequalities similarly to (4.28). This gives the stated control.

4.2 Proof of Proposition 3.7

Write similarly to the proof of Theorem 2.1 in [KM02]:

$$(p_\varepsilon-p_\varepsilon^d)(t_i,t_j,x,y)=(p_\varepsilon\otimes H_\varepsilon-p_\varepsilon\otimes_h H_\varepsilon)(t_i,t_j,x,y)+(p_\varepsilon-p_\varepsilon^d)\otimes_h H_\varepsilon(t_i,t_j,x,y)$$

$$=\sum_{r\ge0}(p_\varepsilon\otimes H_\varepsilon-p_\varepsilon\otimes_h H_\varepsilon)\otimes_h H_\varepsilon^{(r)}(t_i,t_j,x,y),\tag{4.33}$$

where we apply iteratively the first equality to get the second one. From (4.4) under $(\mathbf{A}_H)$ and (4.24) under $(\mathbf{A}_{PS})$, the key point is thus to control $p_\varepsilon\otimes H_\varepsilon-p_\varepsilon\otimes_h H_\varepsilon$. Write:

$$p_\varepsilon\otimes H_\varepsilon-p_\varepsilon\otimes_h H_\varepsilon)(t_i,t_j,x,y)$$

$$=\sum_{k=0}^{j-i-1}\int_{t_{i+k}}^{t_{i+k+1}}du\int_{\mathbb{R}^d}\{p_\varepsilon(t_i,u,x,z)H_\varepsilon(u,t_j,z,y)-p_\varepsilon(t_i,t_{i+k},x,z)H_\varepsilon(t_{i+k},t_j,z,y)\}dz$$

$$=\sum_{k=0}^{j-i-1}\Big\{\int_{t_{i+k}}^{t_{i+k+1}}du\int_{\mathbb{R}^d}\{[p_\varepsilon(t_i,u,x,z)-p_\varepsilon(t_i,t_{i+k},x,z)]H_\varepsilon(u,t_j,z,y)\}dz$$

$$+\int_{t_{i+k}}^{t_{i+k+1}}du\int_{\mathbb{R}^d}\{p_\varepsilon(t_i,t_{i+k},x,z)[H_\varepsilon(u,t_j,z,y)-H_\varepsilon(t_{i+k},t_j,z,y)]\}dz\Big\}$$

$$=:(D_\varepsilon^{d,1}+D_\varepsilon^{d,2})(t_i,t_j,x,y).\tag{4.34}$$

- Bounds for the term $D_\varepsilon^{d,1}$.

- Under $(\mathbf{A}_H)$, for $k = 0$, one readily gets from (4.3):

$$\int_{t_i}^{t_{i+1}} du \Big| \int_{\mathbb{R}^d} \{[p_\varepsilon(t_i, u, x, z) - p_\varepsilon(t_i, t_i, x, z)] H_\varepsilon(u, t_j, z, y)\} dz \Big|$$

$$\leq C p_c(t_j - t_i, y - x) \int_{t_i}^{t_{i+1}} \frac{du}{(t_j - u)^{1-\gamma/2}} \leq \frac{Ch}{(t_j - t_i)^{1-\gamma/2}} p_c(t_j - t_i, y - x)$$

$$\leq Ch^{\gamma/2} p_c(t_j - t_i, y - x). \qquad (4.35)$$

On the other hand, from the parametrix expansion of the density in (3.5), one gets that for all $\eta \in (0, \gamma)$ and $k \geq 1$, $u \in [t_{i+k}, t_{i+k+1}]$:

$$|p_\varepsilon(t_i, u, x, z) - p_\varepsilon(t_i, t_{i+k}, x, z)| \leq \frac{C}{\eta} \frac{(u - t_{i+k})^{(\gamma-\eta)/2}}{(u - t_i)^{\gamma/2}} p_c(u - t_i, z - x). \qquad (4.36)$$

Write indeed, recalling that $u \in [t_{i+k}, t_{i+k+1}]$:

$$p_\varepsilon(t_i, u, x, z) - p_\varepsilon(t_i, t_{i+k}, x, z) = \tilde{p}_\varepsilon(t_i, u, x, z) - \tilde{p}_\varepsilon(t_i, t_{i+k}, x, z)$$

$$+ \sum_{l \geq 1} \int_{t_i}^{t_{i+k}} ds \int_{\mathbb{R}^d} \tilde{p}_\varepsilon \otimes H_\varepsilon^{(l-1)}(t_i, s, x, w) \big(H_\varepsilon(s, u, w, z) - H_\varepsilon(s, t_{i+k}, w, z)\big) dw$$

$$+ \sum_{l \geq 1} \int_{t_{i+k}}^{u} ds \int_{\mathbb{R}^d} \tilde{p}_\varepsilon(t_i, s, x, w) H_\varepsilon^{(l)}(s, u, w, z) dw =: (T_{1,\varepsilon} + T_{2,\varepsilon} + T_{3,\varepsilon})(t_i, t_{i+k}, u, x, z),$$

$$(4.37)$$

with the convention $\tilde{p}_\varepsilon \otimes H_\varepsilon^{(0)} = \tilde{p}_\varepsilon$. Since $\tilde{p}_\varepsilon$ is a Gaussian non-degenerate kernel, and that for $k \geq 1$ we readily get from the mean value theorem:

$$|T_{1,\varepsilon}(t_i, t_{i+k}, u, x, z)| \leq \frac{C(u - t_{i+k})}{(u - t_i)} p_c(u - t_i, z - x) \leq \frac{C(u - t_{i+k})^{(\gamma-\eta)/2}}{(u - t_i)^{\gamma/2}} p_c(u - t_i, z - x). \qquad (4.38)$$

Also, from the definition of Φ_ε following (4.1) and the associated control (4.5), we get:

$$|T_{3,\varepsilon}(t_i, t_{i+k}, u, x, z)| \leq \int_{t_{i+k}}^{u} ds \int_{\mathbb{R}^d} \tilde{p}_\varepsilon(t_i, s, x, w) |\Phi_\varepsilon(s, u, w, z)| dw$$

$$\leq C p_c(u - t_i, z - x) \int_{t_{i+k}}^{u} \frac{ds}{(u - s)^{1-\gamma/2}} \qquad (4.39)$$

$$\leq C(u - t_{i+k})^{\gamma/2} p_c(u - t_i, z - x).$$

For $T_{2,\varepsilon}$, we again use some splitting in time. Write $T_{2,\varepsilon}(t_i, t_{i+k}, u, x, z) = (T_{21,\varepsilon} + T_{22,\varepsilon})(t_i, t_{i+k}, u, x, z)$ where:

$$\begin{aligned}
T_{21,\varepsilon}(t_i, t_{i+k}, u, x, z) &:= \int_{t_{i+k}-(u-t_{i+k})}^{t_{i+k}} ds \int_{\mathbb{R}^d} \tilde{p}_\varepsilon(t_i, s, x, w) \\
&\quad \times (\Phi_\varepsilon(s, u, w, z) - \Phi_\varepsilon(s, t_{i+k}, w, z)) dw, \\
|T_{21,\varepsilon}(t_i, t_{i+k}, u, x, z)| &\leq \int_{t_{i+k}-(u-t_{i+k})}^{t_{i+k}} ds \Big(\frac{p_c(u - t_i, z - x)}{(u - s)^{1-\gamma/2}} + \frac{p_c(t_{i+k} - t_i, z - x)}{(t_{i+k} - s)^{1-\gamma/2}} \Big) \\
&\leq C(u - t_{i+k})^{\gamma/2} p_c(u - t_i, z - x), \qquad (4.40)
\end{aligned}$$

recalling that, since $k \geq 1$, $(u - t_i) \geq t_{i+k} - t_i \geq \frac{1}{2}(u - t_i)$ for the last inequality. For

$$T_{22,\varepsilon}(t_i, t_{i+k}, u, x, z) := \sum_{l \geq 1} \int_{t_i}^{t_{i+k}-(u-t_{i+k})} ds \int_{\mathbb{R}^d} \tilde{p}_\varepsilon \otimes H_\varepsilon^{(l-1)}(t_i, s, x, w)$$

$$\times \big(H_\varepsilon(s, u, w, z) - H_\varepsilon(s, t_{i+k}, w, z)\big) dw,$$

we focus on the second order terms in the difference $\left(H_\varepsilon(s, u, w, z) - H_\varepsilon(s, t_{i+k}, w, z)\right)$. They are indeed the most singular. Note that on the considered time set $\frac{1}{2}(u - s) \leq (t_{i+k} - s)$. We thus get (with similar arguments than those used to handle $T_{1,\varepsilon}$):

$$\left| \text{Tr}\left((a(s, w) - a(s, z))(D_w^2 \tilde{p}_\varepsilon(s, u, w, z) - D_w^2 \tilde{p}_\varepsilon(s, t_{i+k}, w, z))\right) \right|$$
$$\leq C \frac{|z - w|^\gamma (u - t_{i+k})}{(u - s)^2} p_c(u - s, z - w)$$

$$\leq C \left(\frac{|z - w|}{(u - s)^{1/2}} \right)^\gamma \frac{(u - t_{i+k})^{(\gamma - \eta)/2}}{(u - s)^{1 - \eta/2}} p_c(u - s, z - w) \leq \frac{(u - t_{i+k})^{(\gamma - \eta)/2}}{(u - s)^{1 - \eta/2}} p_c(u - s, z - w).$$

The small loss on the time Hölder regularity index is here due to the fact we consider the forward time component and the η is needed to integrate in time. We obtain:

$$|T_{22,\varepsilon}(t_i, t_{i+k}, u, x, z)| \leq \frac{C(u - t_{i+k})^{(\gamma - \eta)/2}}{\eta} p_c(u - t_i, z - x).$$

Plugging this last bound and the controls of (4.40), (4.39), (4.38) into (4.37) yields (4.36). Now, from (4.35) and using (4.3) and (4.36) in (4.34), we get:

$$|D_\varepsilon^{d,1}|(t_i, t_j, x, y) \leq C_\eta h^{(\gamma - \eta)/2} p_c(t_j - t_i, y - x). \tag{4.41}$$

- Under $(\mathbf{A}_{PS})$, since we want to get higher convergence rates, we need to use the forward Kolmogorov equation in $D_\varepsilon^{d,1}$. Write for $k \in [\![1, j - i - 2]\!]$, $u \in [t_k, t_{k+1}]$:

$$p_\varepsilon(t_i, u, x, z) - p_\varepsilon(t_i, t_{i+k}, x, z) = (u - t_{i+k}) \int_0^1 \left(\partial_v p_\varepsilon(t_i, v, x, z)\right)_{v = t_{i+k} + \lambda(u - t_{i+k})} d\lambda$$

$$= (u - t_{i+k}) \int_0^1 \left((L_v^\varepsilon)^* p_\varepsilon(t_i, v, x, z)\right)_{v = t_{i+k} + \lambda(u - t_{i+k})} d\lambda.$$

$\diamond$ If now $\sigma(t, x) = \sigma$, the term H_ε in (4.34) only involves a first order derivative. We obtain:

$$|D_\varepsilon^{d,1}|(t_i, t_j, x, y)$$
$$\leq Ch\Big(\sum_{k=1}^{j-i-2} \int_0^1 d\lambda \int_{t_{i+k}}^{t_{i+k+1}} du \int_{\mathbb{R}^d} \Big[|\text{div}_z\left(b_\varepsilon(v, z)p_\varepsilon(t_i, v, x, z)\right)| \frac{p_c(t_j - u, y - z)}{(t_j - u)^{1/2}}$$
$$+ \frac{1}{2} |\langle a\nabla_z p_\varepsilon(t_i, v, x, z), \nabla_z \langle b_\varepsilon(u, z), \nabla_z \tilde{p}_\varepsilon(u, t_j, z, y)\rangle\rangle| \Big]_{v = t_{i+k} + \lambda(u - t_{i+k})} dz \Big)$$
$$+ C \int_{[t_{j-1}, t_j]} du \int_{\mathbb{R}^d} |p_\varepsilon(t_i, u, x, z) - p_\varepsilon(t_i, t_{j-1}, x, z)| \frac{1}{(t_j - u)^{1/2}} p_c(t_j - u, y - z) dz$$
$$+ C \frac{h}{(t_j - t_i)^{1/2}} p_c(t_j - t_i, y - x). \tag{4.42}$$

From the parametrix representation (3.5) of p_ε, it is again easily deduced similarly to (4.36) that for any $\eta \in (0, 1/2)$, $u \in [t_{j-1}, t_j]$:

$$|p_\varepsilon(t_i, u, x, z) - p_\varepsilon(t_i, t_{j-1}, x, z)| \leq \frac{C}{\eta} \frac{|u - t_{j-1}|^{(1-\eta)/2}}{(u - t_i)^{1/2}} p_c(u - t_i, z - x).$$

Plugging this estimate in (4.42) and using as well (2.9) and (3.33) yields for all $\eta \in (0, 1/2)$:

$$|D_\varepsilon^{d,1}|(t_i, t_j, x, y)$$
$$\leq C_\eta h\Big(\sum_{k=1}^{j-i-2} \int_0^1 d\lambda \int_{t_{i+k}}^{t_{i+k+1}} du \int_{\mathbb{R}^d} \Big[\left(\varepsilon^{-1} \mathbb{I}_{z \in \mathcal{V}_\varepsilon(\mathcal{I})} + \frac{\varepsilon^{-\eta}}{(v - t_i)^{1/2}}\right)$$

$$\times p_c(v - t_i, z - x)\frac{p_c(t_j - u, y - z)}{(t_j - u)^{1/2}} + \varepsilon^{-\eta}\frac{p_c(v - t_i, z - x)}{(v - t_i)^{1/2}}\Big(\frac{\varepsilon^{-1}}{(t_j - u)^{1/2}}\mathbb{I}_{z \in \mathcal{V}_\varepsilon(\mathcal{I})}$$

$$+ \frac{1}{(t_j - u)}\Big)p_c(t_j - u, y - z)\Big]_{v = t_{i+k} + \lambda(u - t_{i+k})}dz\Big) + \frac{Ch^{1-\eta/2}}{\eta(t_j - t_i)^{1/2}}p_c(t_j - t_i, y - x)$$

$$\leq C_\eta\Big(\frac{h\varepsilon^{-(1+\eta)+1/q}}{\alpha(q)(t_j - t_i)^{1/2-\alpha(q)}} + \frac{h|\ln(h)|\varepsilon^{-\eta}}{(t_j - t_i)^{1/2}} + \frac{h^{1-\eta/2}}{\eta(t_j - t_i)^{1/2}}\Big)p_c(t_j - t_i, y - x)$$

$$\leq \bar{C}_{\eta,q}\Big(\frac{h\varepsilon^{-(1+\eta)+1/q}}{(t_j - t_i)^{1/2-\alpha(q)}} + \frac{h|\ln(h)|\varepsilon^{-\eta}}{(t_j - t_i)^{1/2}} + \frac{h^{1-\eta/2}}{(t_j - t_i)^{1/2}}\Big)p_c(t_j - t_i, y - x), \qquad (4.43)$$

recalling as well for the previous computations that $u - v = (1 - \lambda)(u - t_{i+k}) \leq h$.

◇ For a general σ, an additional term appears in (4.42), which corresponds to the second order terms in H_ε for the indexes $k \in [\![1, j - i - 2]\!]$. The other contributions are controlled similarly. We have to bound:

$$Ch\sum_{k=1}^{j-i-2}\int_0^1 d\lambda \int_{t_{i+k}}^{t_{i+k+1}} du \int_{\mathbb{R}^d}\Big[\big|\sum_{l,m=1}^d D^2_{z_l z_m}\big(a_{lm}(t_i, z)p_\varepsilon(t_i, v, x, z)\big)$$

$$\mathrm{Tr}\big((a_\varepsilon(u, z) - a_\varepsilon(u, y))D^2_z\tilde{p}_\varepsilon(u, t_j, z, y)\big)\big|\Big]_{v = t_{i+k} + \lambda(u - t_{i+k})}dz$$

$$\leq \bar{C}_{\eta,q}h\sum_{k=1}^{j-i-2}\int_0^1 d\lambda \int_{t_{i+k}}^{t_{i+k+1}} du \int_{\mathbb{R}^d}\frac{\varepsilon^{-(1+\eta)}}{u - t_i}p_c(u - t_i, z - x)\frac{1}{(t_j - u)^{1/2}}p_x(t_j - u, y - z)dz$$

$$\leq \bar{C}_{\eta,q}\frac{h|\ln(h)|\varepsilon^{-(1+\eta)}}{(t_j - t_i)^{1/2}}.$$

This yields in the considered case:

$$|D^{d,1}_\varepsilon|(t_i, t_j, x, y) \leq \bar{C}_{\eta,q}\Big(\frac{h|\ln(h)|\varepsilon^{-(1+\eta)}}{(t_j - t_i)^{1/2}} + \frac{h^{1-\eta/2}}{(t_j - t_i)^{1/2}}\Big)p_c(t_j - t_i, y - x). \qquad (4.44)$$

• To control the term $D^{d,2}_\varepsilon$ appearing in (4.34), let us first introduce:

$$(\bar{D}^{d,21}_{\varepsilon,\sigma} + \bar{D}^{d,22}_{\varepsilon,\sigma})(t_i, t_{i+k}, u, t_j, x, y) :=$$

$$C\int_{\mathbb{R}^d} p_c(t_{i+k} - t_i, z - x)|a_\varepsilon(u, z) - a_\varepsilon(u, y) - (a_\varepsilon(t_{i+k}, z) - a_\varepsilon(t_{i+k}, y))|$$

$$\times \frac{p_c(t_j - t_{i+k}, y - z)}{t_j - t_{i+k}}dz$$

$$+ \Big|\int_{\mathbb{R}^d} p_\varepsilon(t_i, t_{i+k}, x, z)\mathrm{Tr}\big((a_\varepsilon(u, z) - a_\varepsilon(u, y))[D^2_z\tilde{p}_\varepsilon(u, t_j, z, y) - D^2_z\tilde{p}_\varepsilon(t_{i+k}, t_j, z, y)]\big)\Big|dz,$$

that correspond to the most singular contributions in $D^{d,2}_\varepsilon$ as far as the time singularity is concerned when the diffusion coefficient varies, i.e. $\sigma(t, x) \neq \sigma$.

- Under **(A$_H$)**. For $\bar{D}^{d,22}_{\varepsilon,\sigma}$ we can exploit the Hölder continuity in time of the Gaussian kernel $\tilde{p}_\varepsilon$ to derive, similarly to the computations performed above to investigate $T_{22,\varepsilon}$, that for all $\eta \in (0, \gamma)$, $k \in [\![0, j - i - 2]\!]$, $u \in [t_{i+k}, t_{i+k+1}]$:

$$|D^2_z\tilde{p}_\varepsilon(u, t_j, z, y) - D^2_z\tilde{p}_\varepsilon(t_{i+k}, t_j, z, y)| \leq C\frac{(u - t_{i+k})^{(\gamma-\eta)/2}}{(t_j - u)^{1+(\gamma-\eta)/2}}p_c(t_j - t_{i+k}, y - z).$$

From the spatial Hölder continuity of $a_\varepsilon(u, \cdot)$, we get:

$$\sum_{k=0}^{j-i-2}\int_{t_{i+k}}^{t_{i+k+1}} du|\bar{D}^{d,22}_{\varepsilon,\sigma}|(t_i, t_{i+k}, u, t_j, x, y) \leq \frac{C}{\eta}h^{(\gamma-\eta)/2}p_c(t_j - t_i, y - x). \qquad (4.45)$$

On the other hand, for $k = j - i - 1$, (4.3) readily yields:

$$
\int_{t_{j-1}}^{t_j} du |\bar{D}_{\varepsilon,\sigma}^{d,22}|(t_i, t_{i+k}, u, t_j, x, y) \leq C \int_{t_{j-1}}^{t_j} du \int_{\mathbb{R}^d} p_c(t_{j-1} - t_i, z - x)
$$
$$
\times \left(\frac{p_c(t_j - u, y - z)}{(t_j - u)^{1 - \gamma/2}} + \frac{p_c(t_j - t_{j-1}, y - z)}{(t_j - t_{j-1})^{1 - \gamma/2}} \right) dz \leq Ch^{\gamma/2} p_c(t_j - t_i, y - x).
$$
(4.46)

Also, using the uniform $\gamma/2$-Hölder continuity in time of a we get:

$$
|\bar{D}_{\varepsilon,\sigma}^{d,21}(t_i, t_{i+k}, u, t_j, x, y)|
$$
$$
\leq C \int_{\mathbb{R}^d} p_c(t_{i+k} - t_i, z - x) |u - t_{i+k}|^{\gamma/2} \frac{1}{t_j - t_{i+k}} p_c(t_j - t_{i+k}, y - z) dz
$$
$$
\leq Ch^{(\gamma - \eta)/2} p_c(t_j - t_i, y - x)(t_j - t_{i+k})^{-1 + \eta/2},
$$

for $\eta \in (0, \gamma)$, recalling $u \in [t_{i+k}, t_{i+k+1}]$ for the last inequality. The difference of the first order terms appearing in $D_\varepsilon^{d,2}$ in (4.34) yields similar controls. From the above bound, (4.45) and (4.46), we derive that under $(\mathbf{A}_H)$:

$$
|D_\varepsilon^{d,2}(t_i, t_j, x, y)| \leq C_\eta h^{(\gamma - \eta)/2} p_c(t_j - t_i, y - x).
$$
(4.47)

- Under $(\mathbf{A}_{PS})$, write:

$$
\sum_{k=0}^{j-i-1} \int_{t_{i+k}}^{t_{i+k+1}} du |\bar{D}_{\varepsilon,\sigma}^{d,22}|(t_i, t_{i+k}, u, t_j, x, y)
$$

$$
\leq C \Bigg(h \sum_{k=0}^{\lceil \frac{j-i-1}{2} \rceil} \int_{t_{i+k}}^{t_{i+k+1}} du \int_{\mathbb{R}^d} p_c(t_{i+k} - t_i, z - x) \frac{1}{(t_j - t_i)^{3/2}} p_c(t_j - u, y - z) dz
$$

$$
+ \sum_{k=\lceil \frac{j-i-1}{2} \rceil + 1}^{j-i-1} \int_{t_{i+k}}^{t_{i+k+1}} du \int_{\mathbb{R}^d} \sum_{l,m=1}^{d} \Big| D_{z_l z_m}^2 \Big(p_\varepsilon(t_i, t_{i+k}, x, z)(a_\varepsilon^{lm}(u, z) - a_\varepsilon^{lm}(u, y)) \Big) \Big|
$$

$$
\times |\tilde{p}_\varepsilon(u, t_j, z, y) - \tilde{p}_\varepsilon(t_{i+k}, t_j, z, y)| dz \Bigg)
$$

$$
\leq \bar{C}_{\eta,q} \Big(\frac{h}{(t_j - t_i)^{1/2}} + \frac{h^{1 - \eta/2} \varepsilon^{-(1 + \eta)}}{(t_j - t_i)^{1/2 - \eta/2}} \Big) p_c(t_j - t_i, y - x), \quad (4.48)
$$

where to derive the last inequality, we exploit (3.33), the time Hölder continuity of the Gaussian density $\tilde{p}_\varepsilon$ for $k \in \llbracket \lceil \frac{j-i-1}{2} \rceil + 1, j - i - 2 \rrbracket$ and direct computations for $k = j - i - 1$. Also, the smoothness in time (Lipschitz continuity) of the diffusion coefficients gives for $\eta \in (0, 1]$,

$$
|\bar{D}_{\varepsilon,\sigma}^{d,21}(t_i, t_{i+k}, u, t_j, x, y)| \leq Ch^{1 - \eta/2} p_c(t_j - t_i, y - x)(t_j - t_{i+k})^{-1 + \eta/2}.
$$
(4.49)

Let us now carefully mention that, under $(\mathbf{A}_{PS})$, because of the irregularity of the drift, it is very important as well to establish cautiously the bounds for the difference of the first order terms. Introduce:

$$
(\bar{D}_{\varepsilon,b}^{d,21} + \bar{D}_{\varepsilon,b}^{d,22})(t_i, t_{i+k}, u, t_j, x, y) :=
$$
$$
C \int_{\mathbb{R}^d} p_c(t_{i+k} - t_i, z - x) |b_\varepsilon(u, z) - b_\varepsilon(t_{i+k}, z)| \frac{1}{(t_j - t_{i+k})^{1/2}} p_c(t_j - t_{i+k}, y - z) dz
$$
$$
+ \Big| \int_{\mathbb{R}^d} p_\varepsilon(t_i, t_{i+k}, x, z) \langle b_\varepsilon(u, z), D_z \tilde{p}_\varepsilon(u, t_j, z, y) - D_z \tilde{p}_\varepsilon(t_{i+k}, t_j, z, y) \rangle \Big| dz. \quad (4.50)
$$

From the Lipschitz property in time of $b_\varepsilon(\cdot, z)$ we readily get:

$$\bar{D}_{\varepsilon,b}^{d,21}(t_i, t_{i+k}, u, t_j, x, y) \leq \frac{Ch}{(t_j - t_{i+k})^{1/2}} p_c(t_j - t_i, y - x). \tag{4.51}$$

Also, recalling that $\partial_u \tilde{p}_\varepsilon(u, t_j, z, y) + \frac{1}{2} \text{Tr}\big(a(u, y) D_z^2 \tilde{p}_\varepsilon(u, t_j, z, y)\big) = 0$, one readily gets:

$$\bar{D}_{\varepsilon,b}^{d,22}(t_i, t_{i+k}, u, t_j, x, y) \leq \frac{Ch}{(t_j - u)^{3/2}} p_c(t_j - t_i, y - x), \tag{4.52}$$

which once integrated in time gives the expected control for $k \in [\![0, \lceil \frac{j-i-1}{2} \rceil]\!]$. The indexes $k \in [\![\lceil \frac{j-i-1}{2} \rceil + 1, j - i - 1]\!]$ require a more careful treatment. Now, for such indexes and $u \in [t_{i+k}, t_{i+k+1}]$, using again the Kolmogorov equation satisfied by $\tilde{p}_\varepsilon$ and two spatial integration by parts in z, one obtains from (3.33) the following *global* control:

$$\bar{D}_{\varepsilon,b}^{d,22}(t_i, t_{i+k}, u, t_j, x, y)$$

$$\leq C(u - t_{i+k}) \int_0^1 d\lambda \sum_{l,m,q \in [\![1,d]\!]} \int_{\mathbb{R}^d} \Big[\big| D_{z_l z_m}^2 \big(p_\varepsilon(t_i, t_{i+k}, x, z) b_\varepsilon^l(u, z) \big) \big|$$

$$\times |D_{z_q} \tilde{p}_\varepsilon(v, t_j, z, y)| \Big]_{v = t_{i+k} + \lambda(u - t_{i+k})} dz$$

$$\leq Ch \int_0^1 d\lambda \int_{\mathbb{R}^d} \Big(\frac{\varepsilon^{-(1+\eta)}}{(t_{i+k} - t_i)} + \varepsilon^{-2} \mathbb{I}_{z \in \mathcal{V}_\varepsilon(\mathcal{I})} \Big) p_c(t_{i+k} - t_i, z - x)$$

$$\times \frac{p_c(t_j - v, y - z)}{(t_j - u)^{1/2}} \Big|_{v = t_{i+k} + \lambda(u - t_{i+k})} dz$$

$$\leq Ch p_c(t_j - t_i, y - x) \Big(\frac{\varepsilon^{-(1+\eta)}}{(t_j - t_i)(t_j - u)^{1/2}} + \frac{\varepsilon^{-2+1/q}}{(t_j - u)^{1/2 + d/(2q)}} \Big), q > d. \tag{4.53}$$

Plugging (4.53), (4.52) and (4.51) into (4.50) one derives:

$$\sum_{k=0}^{j-i-1} \int_{t_{i+k}}^{t_{i+k+1}} du \Big(\bar{D}_{\varepsilon,b}^{d,21}(t_i, t_{i+k}, u, t_j, x, y) + \bar{D}_{\varepsilon,b}^{d,22}(t_i, t_{i+k}, u, t_j, x, y) \Big)$$

$$\leq Ch \Big(\frac{\varepsilon^{-(1+\eta)}}{(t_j - t_i)^{1/2}} + \varepsilon^{-2+1/q} \Big) p_c(t_j - t_i, y - x). \tag{4.54}$$

We carefully, point out that, since $q > d$, this term will dominate the error associated with the time discretization when compared to (4.48).

We will now improve this bound using the (unsigned) distance of the final point to the neighborhood of the discontinuity sets $d(y, \mathcal{V}_\varepsilon(\mathcal{I}))$. We cannot hope to improve the control (4.52) for $k \in [\![0, \lceil \frac{j-i-1}{2} \rceil]\!]$ and therefore focus on the indexes $k \in [\![\lceil \frac{j-i-1}{2} \rceil + 1, j - i - 1]\!]$. For those indexes, performing one spatial integration by part in z from (4.50) yields:

$$\sum_{k=\lceil \frac{j-i-1}{2} \rceil + 1}^{j-i-1} \int_{t_{i+k}}^{t_{i+k+1}} du \, \bar{D}_{\varepsilon,b}^{d,22}(t_i, t_{i+k}, u, t_j, x, y)$$

$$\leq C \sum_{k=\lceil \frac{j-i-1}{2} \rceil + 1}^{j-i-1} \int_{t_{i+k}}^{t_{i+k+1}} du \Big(\int_{\mathbb{R}^d} \Big(\frac{\varepsilon^{-\eta}}{(t_{i+k} - t_i)^{1/2}} + \varepsilon^{-1} \mathbb{I}_{z \in \mathcal{V}_\varepsilon(\mathcal{I})} \Big) p_c(t_{i+k} - t_i, z - x)$$

$$\times |\tilde{p}_\varepsilon(u, t_j, z, y) - \tilde{p}_\varepsilon(t_{i+k}, t_j, z, y)| \Big) dz$$

$$\leq \frac{C}{\eta} \Big(\frac{h^{1-\eta/2} \varepsilon^{-\eta}}{(t_j - t_i)^{(1-\eta)/2}} p_c(t_j - t_i, y - x) + \bar{R}_\varepsilon^{d,22}(t_i, t_j, x, y) \Big),$$

using the Hölder continuity in time of $\tilde{p}_\varepsilon$ for $k \in [\![\lceil \frac{j-i-1}{2} \rceil + 1, j - i - 2]\!]$ and direct computations for $k = j - i - 1$ in the second inequality and where

$$\bar{R}_\varepsilon^{d,22}(t_i, t_j, x, y) := h^{1-\eta/2} \sum_{k=\lceil \frac{j-i-1}{2} \rceil + 1}^{j-i-2} \int_{t_{i+k}}^{t_{i+k+1}} \frac{du}{(t_j - u)^{1-\eta/2}} \varepsilon^{-1}$$

$$\times \int_{\mathbb{R}^d} p_c(u - t_i, z - x) \mathbb{I}_{z \in \mathcal{V}_\varepsilon(\mathcal{I})} p_c(t_j - u, y - z) dz$$

$$+ \varepsilon^{-1} \int_{t_{j-1}}^{t_j} du \int_{\mathbb{R}^d} p_c(u - t_i, z - x) \mathbb{I}_{z \in \mathcal{V}_\varepsilon(\mathcal{I})} \big(p_c(t_j - u, y - z) + p_c(t_j - t_{j-1}, y - z) \big) dz.$$

Since $|y - z| + |z - x| \geq |y - x|$ and $(u - t_i) \geq \frac{1}{2}(t_j - t_i)$ we get that up to a modification of c that for $k \in \llbracket \lceil \frac{j-i-1}{2} \rceil + 1, j - i - 1 \rrbracket$ and $u \in [t_{i+k}, t_{i+k+1}], s = u$ or $u \in [t_{j-1}, t_j], s = t_{j-1}$:

$$I_\varepsilon(t_i, t_{i+k}, u, s, t_j, x, y) := \varepsilon^{-1} \int_{\mathbb{R}^d} p_c(u - t_i, z - x) \mathbb{I}_{z \in \mathcal{V}_\varepsilon(\mathcal{I})} p_c(t_j - s, y - z) dz$$

$$\leq C \varepsilon^{-1} p_c(t_j - t_i, y - x) \int_{\mathbb{R}^d} \mathbb{I}_{z \in \mathcal{V}_\varepsilon(\mathcal{I})} p_c(t_j - s, y - z) dz. \quad (4.55)$$

Indeed, either $|z - x| \geq \frac{1}{2}|y - x|$ and in that case $p_c(u - t_i, z - x) \leq C p_c(t_j - t_i, y - x)$, or $|y - z| \geq \frac{1}{2}|y - x|$. In that case we use that $p_c(u - t_i, z - x) \leq C/(t_j - t_i)^{d/2}$ and write as well:

$$\exp\left(-\frac{c}{2} \frac{|y - z|^2}{t_j - s} \right) \leq \exp\left(-\frac{c}{16} \frac{|y - x|^2}{(t_j - s)} \right) \exp\left(-\frac{c}{4} \frac{|y - z|^2}{(t_j - s)} \right)$$

$$\leq \exp\left(-\frac{c}{16} \frac{|y - x|^2}{(t_j - t_i)} \right) \exp\left(-\frac{c}{4} \frac{|y - z|^2}{(t_j - s)} \right),$$

which also gives (4.55) modifying c.

Up to a change of coordinate, in order to straighten the boundary, we can write (following the arguments of Section 3.3.2 that led to (3.24)):

$$I_\varepsilon(t_i, t_{i+k}, u, s, t_j, x, y)$$

$$\leq C \varepsilon^{-1} p_c(t_j - t_i, y - x) \int_{-\varepsilon}^{\varepsilon} \exp\left(-\frac{|\bar{z} - d_S(y, \mathcal{V}_\varepsilon(\mathcal{I}))|^2}{2(t_j - s)} \right) \frac{d\bar{z}}{(t_j - s)^{1/2}}$$

$$\leq C \varepsilon^{-1} p_c(t_j - t_i, y - x) \int_{-\varepsilon}^{\varepsilon} \frac{d\bar{z}}{|\bar{z} - d_S(y, \mathcal{V}_\varepsilon(\mathcal{I}))|},$$

where $d_S(y, \mathcal{V}_\varepsilon(\mathcal{I}))$ stands for the **signed** distance[2] of y to the boundary of $\mathcal{V}_\varepsilon(\mathcal{I})$. Since we have assumed that for this part of the Proposition that $|d_S(y, \mathcal{V}_\varepsilon(\mathcal{I}))| \geq 2\varepsilon$ we get

$$|\bar{z} - d_S(y, \mathcal{V}_\varepsilon(\mathcal{I}))| \geq |d_S(y, \mathcal{V}_\varepsilon(\mathcal{I}))| - |\bar{z}| \geq |d_S(y, \mathcal{V}_\varepsilon(\mathcal{I}))| - \varepsilon \geq \frac{|d_S(y, \mathcal{V}_\varepsilon(\mathcal{I}))|}{2} =: \frac{d(y, \mathcal{V}_\varepsilon(\mathcal{I}))}{2},$$

where $d(y, \mathcal{V}_\varepsilon(\mathcal{I}))$ is the **unsigned** distance of y to the boundary of $\mathcal{V}_\varepsilon(\mathcal{I})$. We finally derive from the above computations (4.51) and (4.52):

$$\sum_{k=0}^{j-i-1} \int_{t_{i+k}}^{t_{i+k+1}} du (\bar{D}_{\varepsilon,b}^{d,21} + \bar{D}_{\varepsilon,b}^{d,22})(t_i, t_{i+k}, u, t_j, x, y)$$

$$\leq C_\eta h^{1-\eta/2} \left(\frac{\varepsilon^{-\eta}}{(t_j - t_i)^{1/2}} + \frac{1}{d(y, \mathcal{V}_\varepsilon(\mathcal{I}))} \right) p_c(t_j - t_i, y - x). \quad (4.56)$$

- Final derivation of the bounds.

---

[2] Since the discontinuity sets are bounded, we can for instance choose the distance to be positive for the points inside the bounded domain associated with the boundary. Anyhow, this choice plays no role here.

Recall first that:

$$
\begin{aligned}
|p_\varepsilon - p_\varepsilon^d|(t_i, t_j, x, y) \quad \leq \quad & |(p_\varepsilon \otimes H_\varepsilon - p_\varepsilon \otimes_h H_\varepsilon)|(t_i, t_j, x, y) \\
& + \sum_{r \geq 1} |(p_\varepsilon \otimes H_\varepsilon - p_\varepsilon \otimes_h H_\varepsilon)| \otimes_h |H_\varepsilon^{(r)}|(t_i, t_j, x, y).
\end{aligned}
$$

- Under $(\mathbf{A}_H)$, we first plug (4.47), (4.41) into (4.34). The bound (3.34) of the proposition then follows from the above inequality using (4.4).

- Under $(\mathbf{A}_{PS})$.

$\diamond$ For a general $\sigma(t, x)$ (which varies), we derive from (4.44), (4.48), (4.49), (4.54) and (4.33), (4.4):

$$
|p_\varepsilon - p_\varepsilon^d|(t_i, t_j, x, y) \leq \bar{C}_{\eta, q} \Big(\frac{h |\ln(h)| \varepsilon^{-(1+\eta)}}{(t_j - t_i)^{1/2}} + \frac{h^{1-\eta/2} \varepsilon^{-(1+\eta)}}{(t_j - t_i)^{1/2}} + h \varepsilon^{-2+1/q} \Big) p_c(t_j - t_i, y - x).
$$

Using (4.56) instead of (4.54) when $d(y, \mathcal{V}_\varepsilon(\mathcal{I})) \geq 2\varepsilon$ yields:

$$
\begin{aligned}
& |p_\varepsilon - p_\varepsilon^d|(t_i, t_j, x, y) \\
\leq \quad & \bar{C}_{\eta, q} \Big(\frac{h |\ln(h)| \varepsilon^{-(1+\eta)}}{(t_j - t_i)^{1/2}} + \frac{h^{1-\eta/2} \varepsilon^{-(1+\eta)}}{(t_j - t_i)^{1/2}} + \frac{h^{1-\eta/2}}{d(y, \mathcal{V}_\varepsilon(\mathcal{I}))} \Big) p_c(t_j - t_i, y - x).
\end{aligned}
$$

$\diamond$ For $\sigma(t, x) = \sigma$ (fixed diffusion coefficient), when $d(y, \mathcal{V}_\varepsilon(\mathcal{I})) \geq 2\varepsilon$, we derive from (4.43), (4.56) that:

$$
|p_\varepsilon - p_\varepsilon^d|(t_i, t_j, x, y) \quad \leq \quad \bar{C}_{\eta, q} \Big(\frac{h |\ln(h)| \varepsilon^{-\eta}}{(t_j - t_i)^{1/2}} + \frac{h \varepsilon^{-(1+\eta)+1/q}}{(t_j - t_i)^{1/2}} + \frac{h^{1-\eta/2}}{d(y, \mathcal{V}_\varepsilon(\mathcal{I}))} \Big) p_c(t_j - t_i, y - x).
$$

Observe that in this case the contribution $(\bar{D}_{\varepsilon, \sigma}^{d, 2j})_{j \in \{1, 2\}}$ vanish. The upper bound of (4.48) thus does not appear. This completes the proof.

Acknowledgments

References

[Alf05]   A. Alfonsi. On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods Appl.*, 11(4):355–384, 2005. MR-2186814

[AJKH14]   A. Alfonsi, B. Jourdain, and A. Kohatsu-Higa. Pathwise optimal transport bounds between a one-dimensional diffusion and its Euler scheme. *Ann. Appl. Probab.*, 24(3):1049–1080, 2014. MR-3199980

[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. MR-1367666

[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. MR-1401964

[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. MR-2642351

[BBD08]   Abdel Berkaoui, Mireille Bossy, and Awa Diop. Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. *ESAIM Probab. Stat.*, 12:1–11 (electronic), 2008. MR-2367990

[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. MR-2659772

[Fri64]   A. Friedman. *Partial Differential Equations of Parabolic Type*. Prentice-Hall, 1964. MR-0181836

[GT98]   D. Gilbarg and N.S. Trudinger. *Elliptic partial differential equations of second order, Second edition*. Springer Verlag, 1998. MR-0473443

[GR14]   I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. D. Zwillinger and V. Moll (eds.) Eighth edition, 2014. MR-3307944

[GR11]   I. Gyöngy and M. Rásonyi. A note on Euler approximations for SDEs with Hölder continuous diffusion coefficients. *Stochastic Process. Appl.*, 121(10):2189–2200, 2011. MR-2822773

[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. MR-0138888

[KHLY15]   A. Kohatsu-Higa, A. Lejay, and K. Yasuda. Weak Approximation Errors for Stochastic Differential Equations with Non-Regular Drifts. *hal-00840211*, 2015.

[KKM16]   V. Konakov, A. Kozhina, and S. Menozzi. Stability of Densities for Perturbed Diffusions and Markov Chains. *ESAIM Proba. and Stat.*, 21:88–112, 2017.

[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. MR-1777133

[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. MR-1931967

[KM15]   V. Konakov and A. Markova. Linear trend exclusion for models defined with stochastic differential and difference equations. *Automation and Remote Control*, 76–10:1771–1783, 2015. MR-3415457

[KMM10]   V. Konakov, S. Menozzi, and S. Molchanov. Explicit parametrix and local limit theorems for some degenerate diffusion processes. *Annales de l'Institut Henri Poincaré, Série B*, 46–4:908–923, 2010. MR-2744877

[Koz16]   A. Kozhina. Stability of densities for perturbed degenerate diffusions. *arXiv:1602.04770, To appear in Theory of Prob. and Appl.*, 2016.

[LSU68]   O.A. Ladyzenskaja, V.A. Solonnikov, and N.N. Ural'ceva. *Linear and quasi-linear equations of parabolic type*. Vol.23 Trans. Math. Monog., AMS, Providence, 1968. MR-0241822

[LM10]   V. Lemaire and S. Menozzi. On some non asymptotic bounds for the Euler scheme. *Electronic Journal of Probability*, 15:1645–1681, 2010. MR-2735377

[MS67]   H. P. McKean and I. M. Singer. Curvature and the eigenvalues of the Laplacian. *J. Differential Geometry*, 1:43–69, 1967. MR-0217739

[Men11]   S. Menozzi. Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electronic Communications in Probability*, 17:234–250, 2011. MR-2802040

[Mik12]   R. Mikulevicius. 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. MR-2926173

[MP91]   R. Mikulevičius and E. Platen. Rate of convergence of the Euler approximation for diffusion processes. *Math. Nachr.*, 151:233–239, 1991. MR-1121206

[MZ15]   R. Mikulevičius and C. Zhang. Weak Euler approximation for Itô diffusion and jump processes. *Stoch. Anal. Appl.*, 33(3):549–571, 2015. MR-3339318

[She91]   S. J. Sheu. Some estimates of the transition density of a nondegenerate diffusion Markov process. *Ann. Probab.*, 19–2:538–561, 1991. MR-1106275

[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. MR-1091544

[Zyg36]   A. Zygmund. *Trigonometric Series*. Cambridge University Press, 1936. MR-0236587

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm