National Research University Higher School of Economics
International Laboratory of Stochastic Analysis and Its Applications

Anna Kozhina

Parametrix Method and its Applications in Probability Theory

Summary of the PhD thesis
for the purpose of obtaining
Philosophy Doctor in Mathematics HSE

Academic supervisor:
Valentin Konakov Doctor of Science, professor

Moscow – 2018

Modelling of many natural phenomena is still a challenging task. Data and observations which we receive from the real word usually contain a lot of inaccuracies and noisy factors. Using deterministic models only often makes predictions inefficient and imprecise. Thus, researchers in many fields are forced to apply models with additional randomness inside.

A possible way to model the uncertainty is to describe dynamics of the process in terms of Stochastic Differential Equations (SDEs further). We are interested in studying Brownian SDEs of the following form

$$Z_t = z + \int_0^t b(s, Z_s)ds + \int_0^t \sigma(s, Z_s)dW_s, \tag{1}$$

where $(W_{s\geq 0})$ is an $\mathbb{R}^k$-valued Brownian motion on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, Z_t is $\mathbb{R}^m$ valued, with $m \in \mathbb{N}$ possibly different of k. The coefficients b, σ are $\mathbb{R}^m$ and $\mathbb{R}^m \otimes \mathbb{R}^k$ valued respectively and such that a unique weak solution to (1) exists.

Equation (1) appears in many applied fields varied from physics to finance. Let us mention Hamiltonian mechanics [Tal02], financial mathematics [JYC10] and biologic simple epidemic model ([Bai17]; [BY89]).

Except from some very specific cases, the SDE (1) cannot be solved explicitly and it therefore seems natural to investigate some related approximation procedures. The Euler - Maruyama method (usually simply called the Euler method), introduced in the current SDE framework in [Mar55], is still one of the simplest effective computational methods. Let us fix a finite time horizon $T > 0$. For a given integer N, representing the number of time steps to be considered along the time interval $[0, T]$, introducing the time step $h = T/N$ and for all $t \in [0, T]$:

$$Z_t^h = z + \int_0^t b(\phi(s), Z^h_{\phi(s)})ds + \int_0^t \sigma(\phi(s), Z^h_{\phi(s)})dW_s, \tag{2}$$

Studying the accuracy of the approximation of the scheme proposed in (2) for the initial SDE (1) two main types of errors are usually considered. The first one to be investigated (see e.g. [Mar55], Gikhman and Skorokhod [GS67], [GS82]) was the so-called *strong error*. Namely, for all $p \in [1, +\infty)$, with the usual Markovian notations for the processes Z^h_s, Z_s, started from z at the moment 0 it holds:

$$\mathcal{E}_S(T, z, h, p) := \Big(\mathbb{E}_z[\sup_{s\in[0,T]} |Z_s^{h,0,z} - Z_s^{0,z}|^p]\Big)^{1/p}. \tag{3}$$

When the coefficients in (1) are Lipschitz continuous in space and at least 1/2-Hölder continuous in time, it is easily seen from usual stochastic analysis techniques, namely Itô' s formula Burkholder-Davis-Gundy inequalities and the Gronwall, Lemma that:

$$\exists C_p(T, b, \sigma),\ \mathcal{E}_S(T, z, h, p) \leq C_p(T, b, \sigma)h^{1/2}.$$

On the other hand, in many applications, such as the pricing and hedging of financial derivatives, only so called weak error, introduced in (1) and (2), is of interest. For a *suitable* test function f (we remain here a bit vague about the function space to which f belongs to), one introduces:

$$\mathcal{E}_W(T,z,h,f) := \mathbb{E}_z[f(Z_T^{h,0,z})] - \mathbb{E}_z[f(Z_T^{0,z})]. \tag{4}$$

There are two sets of assumptions which guarantee that the convergence rate for $\mathcal{E}_W(T,z,h,f)$ is actually of order h. Namely, if

(i) b,σ,f are smooth and without any specific non-degeneracy assumptions

or

(ii) b,σ enjoy some structure property (i.e. the generator associated with (1) is elliptic or hypoelliptic) and some smoothness, and for f that enjoys suitable growth conditions (and that can even be a Dirac mass)

then

$$|\mathcal{E}_W(T,z,h,f)| = |\mathbb{E}_z[f(Z_T^h)] - \mathbb{E}_z[f(Z_T)]| \le C(T,f,\sigma,b)h. \tag{5}$$

In the both cases the main tool for the analysis is the correspondence between $\mathbb{E}_z[f(Z_T^h)]$ and the solution of a second order parabolic PDE. This correspondence is provided by the Feynman-Kac representation formula. Precisely, under the above assumptions we have that, with the usual Markovian notations, $v(t,z) := \mathbb{E}[f(Z_T^{t,z})]$ solves

$$\begin{cases} (\partial_t + L_t)v(t,z) = 0, \ (t,z) \in [0,T[\times\mathbb{R}^m, \\ v(T,z) = f(z), \ z \in \mathbb{R}^m, \end{cases} \tag{6}$$

where

$$L_t v(t,z) = \langle b(t,z), \nabla_z v(t,z)\rangle + \frac{1}{2}\mathrm{Tr}\Big(a(t,z)D_z^2 v(t,z)\Big), a(t,z) := \sigma\sigma^*(t,z),$$

is the generator associated with (1). Assuming some smoothness on v, one can write

$$
\begin{aligned}
\mathcal{E}_W(T,z,h,f) &= \mathbb{E}[f(Z_T^{h,0,z})] - \mathbb{E}[f(Z_T^{0,z})] = \sum_{i=0}^{N-1} \mathbb{E}[v(t_{i+1}, Z_{t_{i+1}}^{h,0,z}) - v(t_i, Z_{t_i}^{h,0,z})] \qquad (7)\\
&= \sum_{i=0}^{N-1} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} \Big\{\partial_s v(s, Z_s^{h,0,z}) + \nabla_z v(s, Z_s^{h,0,z}) b(t_i, Z_{t_i}^{h,0,z})\\
&+ \frac{1}{2}\mathrm{Tr}(D_z^2 v(s, Z_s^{h,0,z}) a(t_i, Z_{t_i}^{h,0,z}))\Big\} ds\Big] = \sum_{i=0}^{N-1} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} \Big\{\partial_s v + L_s v\Big\}(Z_s^{h,0,z}) ds\Big]\\
&+ \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} \Big\{\nabla_z v(s, Z_s^{h,0,z}) \cdot (b(t_i, Z_{t_i}^{h,0,z}) - b(s, Z_s^{h,0,z}))\\
&+ \frac{1}{2}\mathrm{Tr}(D_z^2 v(s, Z_s^{h,0,z})(a(t_i, Z_{t_i}^{h,0,z}) - a(s, Z_s^{h,0,z})))\Big\} ds\Big]\\
&= \sum_{i=0}^{N-1} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} \Big\{\nabla_z v_\varepsilon(s, Z_s^{h,0,z}) \cdot (b(t_i, Z_{t_i}^{h,0,z}) - b(s, Z_s^{h,0,z}))\\
&+ \frac{1}{2}\mathrm{Tr}(D_z^2 v(s, Z_s^{h,0,z})(a(t_i, Z_{t_i}^{h,0,z}) - a(s, Z_s^{h,0,z})))\Big\} ds\Big], \qquad (8)
\end{aligned}
$$

exploiting the PDE satisfied by v for the last equality and Itô formula for the third equality. For a function$f$ in $C_b^{2+\beta}(\mathbb{R}^k, \mathbb{R}), \forall \beta \in (0,1]$ the spatial derivatives of v up to order two are globally bounded on $[0,T]$. Through Taylor like expansions, whenever (i) or (ii) holds, one can control (8), deriving that each contribution in (8) has the order h^2. This leads to the error of order h achieved after summing from 0 to $N-1$.

In case (i), which was considered among the others in the seminal paper by Talay and Tubaro [TT90], the smoothness of v is simply derived via stochastic flow techniques. In case (ii) let us mention that in the hypoelliptic setting (see Section 4.1.1 for additional details on hypoellipticity), Bally and Talay [BT96a], [BT96b] established (5) for bounded Borel functions f and Dirac masses. It respectively bases on the controls of Kusuoka and Stroock [KS84], [KS85] for the derivatives of the density of the diffusion process. We carefully mention that, for this method, which anyhow allows to consider a broad class of potential degeneracies, to apply, the coefficients are assumed to be smooth. The estimates on the tangent processes and Malliavin matrices in the works by Kusuoka and Stroock indeed require such a smoothness. In the uniformly elliptic case yet another approach has been developed by Konakov and Mammen [KM00], [KM02] which is based on parametrix expansions.

Parametrix expansions, which roughly consists in approximating the density of a process with variable coefficients by the density of the corresponding dynamics with constant coefficients, have been a successful tool in many fields. In particular, when a good *proxy* is available (which is, for instance, the case the coefficients b ad σ in (1) are non-degenerate and bounded), parametrix allowed to derive the controls required for the analysis of the weak error under rather mild assumptions. We can mention the work of Il'in *et al.* [IKO62], who derived Gaussian heat kernel for the density of (1) for bounded Hölder coefficients when $\sigma\sigma^*$ is non-degenerate. Similar bounds have been successfully exploited by Konakov and Menozzi [KM17] to derive, in the non-degenerate Hölder continuous setting, that for $b, \sigma \in C^{\gamma/2,\gamma}([0,T], \mathbb{R}^k)$, $\gamma \in (0,1]$ and $f \in C^{\beta}(\mathbb{R}^k, \mathbb{R}), \beta \in (0,1]$:

$$|\mathcal{E}_W(T, z, h, f)| = |\mathbb{E}_z[f(Z_s^h)] - \mathbb{E}_z[f(Z_s)]| \leq C(T, f, \sigma, b)h^{\gamma/2}, \tag{9}$$

improving the previous result by Mikulevičius and Platen [MP91] who also obtained the bound (9) for a function $f \in C^{2+\gamma}(\mathbb{R}^k, \mathbb{R})$. This additional smoothness was due to the fact that they based their analysis on the associated Schauder estimates (which could already be found in [IKO62]). Going directly to the heat-kernel allows to notably alleviate the smoothness assumptions on the final condition, which might be useful for applications.

Intuitively, the above convergence rate can be explained by the fact that, in the low regularity setting, the terms of order greater than one in the telescopic sum (7) cannot be expanded much further. Namely, we can only exploit the γ-Hölder continuity of the coefficients which leads to an error controlled by the increments

$$\mathbb{E}[|b(s, Z_s^h) - b(\phi(s), Z_{\phi(s)}^h)|] + \mathbb{E}[|a(s, Z_s^h) - a(\phi(s), Z_{\phi(s)}^h)|] \leq C(b, \sigma)h^{\gamma/2}.$$

In other words, the convergence rate is closer to the one associated with the *strong error* in (3).

For many applications, e.g. for neuro-sciences or diffusions in random media, it is far important to handle rougher coefficients, for instance piecewise smooth drifts in (1). In that case, the previously mentioned heat-kernel bounds do not hold. Motivated by the investigation of the related weak error for Dirac masses test functions, we have developed, with V. Konakov and S. Menozzi, a sensitivity analysis of the density of (1) (when suitable good Gaussian bounds exist) with respect to a perturbation of the coefficients. This is the first main result of the Thesis which led to the publication [KKM17] and is thoroughly developed in Chapter 3.

Namely, let us introduce the SDE of the form:

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \ t \in [0, T], \tag{10}$$

where $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d, \ \sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ are bounded coefficients that are measurable in time and Hölder continuous in space (this last condition will be relaxed for the drift term b). Also, $a(t,x) := \sigma\sigma^*(t,x)$ is assumed to be uniformly elliptic. In particular, those assumptions guarantee that (10) admits a unique weak solution, see e.g. Bass and Perkins [BP09], [Men11], from which the uniqueness to the martingale problem for the associated generator can be derived under the current assumptions.

We now introduce, for a given parameter $\varepsilon > 0$, a perturbed version of (10) with dynamics:

$$dX_t^{(\varepsilon)} = b_\varepsilon(t, X_t^{(\varepsilon)})dt + \sigma_\varepsilon(t, X_t^{(\varepsilon)})dW_t, \ t \in [0, T], \tag{11}$$

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

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

In the Chapter 3 we investigate, applying the parametrix technique, how the closeness of $(b_\varepsilon, \sigma_\varepsilon)$ and (b, σ) is reflected on the respective densities of the associated processes. Our stability results will also apply to two Markov chains with respective dynamics:

$$\begin{aligned} Y_{t_{k+1}} &= Y_{t_k} + b(t_k, Y_{t_k})h + \sigma(t_k, Y_{t_k})\sqrt{h}\xi_{k+1}, Y_0 = x, \\ Y^{(\varepsilon)}_{t_{k+1}} &= Y^{(\varepsilon)}_{t_k} + b_\varepsilon(t_k, Y^{(\varepsilon)}_{t_k})h + \sigma_\varepsilon(t_k, Y^{(\varepsilon)}_{t_k})\sqrt{h}\xi_{k+1}, Y^{(\varepsilon)}_0 = x, \end{aligned} \tag{12}$$

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

Let us specify the following assumptions **(A)** which we use in Chapter 3. The parameter $\varepsilon > 0$ below is fixed and the constants appearing in the assumptions **do not depend** on ε.

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

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

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

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

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

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

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

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

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

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

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

where for $\gamma \in (0,1]$, $||\cdot||_\gamma$ stands for the usual Hölder norm in space on $C_b^\gamma(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^d)$ (space of Hölder continuous bounded functions, see e.g. Krylov [Kry96]) i.e. :

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

We eventually set for $q \in (1,+\infty]$,

$$\Delta_{\varepsilon,\gamma,q} := \Delta_{\varepsilon,\sigma,\gamma} + \Delta_{\varepsilon,b,q}.$$

Theorem (3.1.1)**.** *Fix $\varepsilon > 0$ and a final deterministic time horizon $T > 0$. Under assumptions **(A)**, specified before, for $q > d$, there exist $C := C(q) \geq 1, c := c(q) \in (0,1]$ s.t. for all $0 \leq s < t \leq T, (x,y) \in (\mathbb{R}^d)^2$:*

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

where $p(s,t,x,.), p_\varepsilon(s,t,x,.)$ respectively stand for the transition densities at time t of equations (10)*,* (11) *starting from x at time s. Also, we denote for a given $c > 0$ and for all $(u,z) \in \mathbb{R}^+ \times \mathbb{R}^d$, $p_c(u,z) := \frac{c^{d/2}}{(2\pi u)^{d/2}} \exp(-c\frac{|z|^2}{2u})$. If $q = \infty$, the constants C, c do not depend on q.*

This and the next theorem will be restated and discussed in Section 3.1.1.

Before stating our results for Markov Chains we introduce two kinds of innovations in (12). Namely:

(**IG**) The i.i.d. random variables $(\xi_k)_{k\geq1}$ are Gaussian, with law $\mathcal{N}(0,I_d)$. In that case the dynamics in (12) correspond to the Euler discretization of equations (10) and (11).

(**IP**) For a given integer $M > 2d + 5 + \gamma$, the innovations $(\xi_k)_{k\geq1}$ are centered and have C^5 density f_ξ which has, together with its derivatives up to order 5, at most polynomial decay of order M. Namely, for all $z \in \mathbb{R}^d$ and multi-index $\nu, |\nu| \leq 5$:

$$|D^\nu f_\xi(z)| \leq CQ_M(z),$$

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

Theorem (3.1.2)**.** *Fix $\varepsilon > 0$ and a final deterministic time horizon $T > 0$. For $h = T/N,\ N \in \mathbb{N}^*$, we set for $i \in \mathbb{N},\ t_i := ih$. Under **(A)**, assuming that either(**IG**) or (**IP**) holds, and for $q > d$ there exist $C := C(q) \geq 1, c := c(q) \in (0,1]$ s.t. for all $0 \leq t_i < t_j \leq T, (x,y) \in (\mathbb{R}^d)^2$:*

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

where $p^h(t_i, t_j, x, .), p^h_\varepsilon(t_i, t_j, x, .)$ respectively stand for the transition densities at time t_j of the Markov Chains Y and $Y^{(\varepsilon)}$ in (12) *starting from x at time t_i. Also:*

- *If (**IG**) holds:*

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

with p_c as in Theorem 3.1.1.

- *If (**IP**) holds:*

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

Again, if $q = +\infty$ the constants C, c do not depend on q.

Continuing the research, V. Konakov and S. Menozzi applied results mentioned above to study the weak error of the Euler scheme approximations in the paper [KM17]. To investigate the weak error for rough drifts, the idea in [KM17] is to mollify the drifts. The difference between the density of the initial diffusion and the one with mollified coefficients is precisely controlled by the previous result. The same occurs for the Euler scheme case. It therefore remains to control the difference between the densities of the mollified diffusion and the scheme which can be addressed from previous results of [KM02] provided, that high order derivatives (which explode with the mollifying parameter) are sharply controlled.

Motivated by the extension of the previous study, we continue with the weak error controls for the case of rough coefficients to Kolmogorov's degenerate SDEs in Chapter 4. Namely, we specify the model in (1) writing $Z_t = (X_t, Y_t)$ with:

$$\begin{cases} dX_t = b(X_t, Y_t)dt + \sigma(X_t, Y_t)dW_t, \\ dY_t = X_t dt, t \in [0, T], \end{cases} \tag{13}$$

where $b : \mathbb{R}^{2d} \to \mathbb{R}^d,\ \sigma : \mathbb{R}^{2d} \to \mathbb{R}^d \otimes \mathbb{R}^d$ are bounded coefficients that are Hölder continuous in space (this condition will be relaxed for the drift term b) and W is a Brownian motion on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. In (13), $T > 0$ is a fixed deterministic final time. Also, $a(x, y) := \sigma\sigma^*(x, y)$ is assumed to be uniformly elliptic.

We point out that those assumptions (specified below) are actually sufficient to guarantee weak uniqueness for the solution of equation (13), see Remark 4.2.1.

Such equations were first introduced in the seminal paper [Kol34] by Kolmogorov. In that work, he found the explicit expression of the density when the coefficients are constants. The parametrix approach in that framework has then been applied by various authors, Weber [Web51], Sonin [Son67] and the more recent [KMM10] under various kinds of assumptions. Adapting the techniques introduced in the last quoted work, which deals with Lipschitz coefficients, it is now possible to consider the Hölder setting for the degenerate Kolmogorov diffusions of type (13). The sensitivity analysis naturally extends to this framework. These aspects are detailed in Chapter 4 (see as well the published article [Koz16]).

Precisely, let us introduce the Euler scheme for the SDE (13) first. For a fixed N and $T > 0$ we define a time grid $\{0, t_1, \ldots, t_N\}$ with a given step $h := T/N$, i.e. $t_i = ih$, for $i = 0, \ldots, N$ and the scheme

$$\begin{cases} X_t^h = x + \int_0^t b(X_{\phi(s)}^h, Y_{\phi(s)}^h)ds + \int_0^t \sigma(X_{\phi(s)}^h, Y_{\phi(s)}^h)dW_s, \\ Y_t^h = y + \int_0^t X_s^h ds. \end{cases} \tag{14}$$

where $\phi(t) = t_i \ \forall t \in [t_i, t_{i+1})$. Observe that the above scheme is in fact well defined even though the non degenerate component of the scheme itself appears in the integral. On every time-step the increments of $(X_t^h, Y_t^h)_{t\in[t_i,t_{i+1}]}, \ i \geq 0$ are actually Gaussian. They indeed correspond to a suitable rescaling of the Brownian increment and its integral on the considered time step, see also Remark 4.2.3.

Let us also denote for a given $c > 0$ and for all $(x, y), (x', y') \in \mathbb{R}^{2d}$ the Kolmogorov-type density

$$p_{c,K}(t, (x, y), (x', y')) := \frac{c^d 3^{d/2}}{(2\pi t^2)^d} \exp\left(-c\left[\frac{|x' - x|^2}{4t} + 3\frac{|y' - y - (x + x')t/2|^2}{t^3}\right]\right). \tag{15}$$

The subscript K in the notation $p_{c,K}(t, (x, y), (x', y'))$ stands for Kolmogorov-like equations.

We would like to emphasize that in Chapter 4 we are considering time-homogeneous coefficients b, σ under the following assumptions:

(**AD1**) **(Boundedness of the coefficients)**.

The components of the vector-valued function $b(x, y)$ and the matrix-valued function $\sigma(x, y)$ are bounded measurable. Specifically, there exists a constant K s.t.

$$\sup_{(x,y)\in\mathbb{R}^{2d}} |b(x, y)| + \sup_{(x,y)\in\mathbb{R}^{2d}} |\sigma(x, y)| \leq K.$$

(**AD2**) **(Uniform Ellipticity)**.

The matrix $a := \sigma\sigma^*$ is uniformly elliptic, i.e. there exists $\Lambda \geq 1$, such that $\forall (x, y, \xi) \in (\mathbb{R}^d)^3$,

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

(**AD3**) **(Hölder continuity in space)**.

For some $\gamma \in (0,1]$, κ, we have

$$|b(x,y)-b(x',y')|+|\sigma(x,y)-\sigma(x',y')| \quad \leq \quad \kappa\left(|x-x'|^{\gamma}+|y-y'|^{\gamma/3}\right).$$

We say that assumption **(AD)** holds when conditions **(AD1)-(AD3)** are in force.

Under the above mentioned assumptions, we now introduce perturbed versions of (13) and (14). Namely, for $b_\varepsilon : \mathbb{R}^{2d} \to \mathbb{R}^d$, $\sigma_\varepsilon : \mathbb{R}^{2d} \to \mathbb{R}^d \otimes \mathbb{R}^d$ satisfying at least the same assumptions as b, σ and being meant to be *close* to b, σ for small values of $\varepsilon > 0$ one denote:

$$\begin{cases} dX_t^{(\varepsilon)} = b_\varepsilon(X_t^{(\varepsilon)}, Y_t^{(\varepsilon)})dt + \sigma(X_t^{(\varepsilon)}, Y_t^{(\varepsilon)})dW_t, \\ dY_t^{(\varepsilon)} = X_t^{(\varepsilon)}dt, t \in [0,T], \end{cases} \tag{16}$$

and similarly:

$$\begin{cases} X_t^{\varepsilon,h} = x + \int_0^t b_\varepsilon(X_{\phi(s)}^{\varepsilon,h}, Y_{\phi(s)}^{\varepsilon,h})ds + \int_0^t \sigma_\varepsilon(X_{\phi(s)}^{\varepsilon,h}, Y_{\phi(s)}^{\varepsilon,h})dW_s, \\ Y_t^{\varepsilon,h} = y + \int_0^t X_s^{\varepsilon,h} ds. \end{cases} \tag{17}$$

for $t \in [0, t_j), 0 < j \leq N$, where $\phi(t) = t_i \ \forall t \in [t_i, t_{i+1})$.

Considering a specific kind of Hölder continuity associated with the intrinsic scales of the system and the time-homogeneous case we set for $\varepsilon > 0$:

$$\forall q \in (1, +\infty],\ \Delta_{\varepsilon,b,q}^d := |b(.,.) - b_\varepsilon(.,.)|_{L^q(\mathbb{R}^{2d})}.$$

We also define

$$\Delta_{\varepsilon,\sigma,\gamma}^d := |\sigma(.,.) - \sigma_\varepsilon(.,.)|_{d,\gamma},$$

where for $\gamma \in (0,1]$, $|.|_{d,\gamma}$ stands for the Hölder norm in space on $C_{b,\mathbf{d}}^\gamma(\mathbb{R}^d \otimes \mathbb{R}^d)$, which denotes the space of Hölder continuous bounded functions with respect to the distance $\mathbf{d}$ defined as follows:

$$\forall (x,y),\ (x',y') \in (\mathbb{R}^d)^2,\ \mathbf{d}\big((x,y),(x',y')\big) := |x-x'| + |y'-y|^{1/3}.$$

Namely, a measurable function f is in $C_{b,\mathbf{d}}^\gamma(\mathbb{R}^d \otimes \mathbb{R}^d)$ if

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

We eventually set $\forall q \in (1,+\infty]$,

$$\Delta_{\varepsilon,\gamma,q}^d := \Delta_{\varepsilon,\sigma,\gamma}^d + \Delta_{\varepsilon,b,q}^d,$$

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

Theorem (4.3.1)**.** *Fix $T > 0$. Under **AD**, for $q \in (4d, +\infty]$, there exist $C := C(q) \geq 1, c \in (0,1]$ s.t. for all $0 < t \leq T, ((x,y),(x',y')) \in (\mathbb{R}^{2d})^2$:*

$$|(p - p_\varepsilon)(t,(x,y),(x',y'))| \leq C\Delta^d_{\varepsilon,\gamma,q} p_{c,K}(t,(x,y),(x',y')),$$

where $p(t,(x,y),(.,.)), p_\varepsilon(t,(x,y),(.,.))$ respectively stand for the transition densities at time t of equations (13)*,* (16) *starting from (x,y) at time* 0*.*

Theorem (4.3.5)**.** *Fix $T > 0$ and let us define a time-grid $\Lambda_h := \{(t_i)_{i\in[\![1,N]\!]}\}, N \in \mathbb{N}^*$. Under **AD**, there exist $C \geq 1, c \in (0,1]$ s.t. for all $0 < t_j \leq T, ((x,y),(x',y')) \in (\mathbb{R}^{2d})^2$:*

$$|p^\varepsilon_h - p_h|(t_j,(x,y),(x',y')) \leq C\Delta^d_{\varepsilon,\gamma,q} p_{c,K}(t_j,(x,y),(x',y')),$$

where $p^\varepsilon_h(t,(x,y),(.,.)), p_h(t,(x,y),(.,.))$ respectively stand for the transition densities at time t of equations (14)*,* (17) *starting from (x,y) at time* 0*.*

These two theorems will be restated and discussed in Section 4.3.1.

The sensitivity analysis will then be applied, in the flavour of [KM17] to investigate the weak error associated to a specific Euler scheme which had already been considered in [LM10] for equations of type (13). However, to perform the analysis we need to change assumptions **(AD)** slightly. Precisely, we have to assume more about Hölder properties of coefficients than in **(AD)**.

Instead of (**AD3**), we assume for some $\gamma \in (0,1]$, $0 < \kappa < \infty$ it holds:

$$|b(x,y) - b(x',y')| + |\sigma(x,y) - \sigma(x',y')| \quad \leq \quad \kappa \left(|x - x'|^\gamma + |y - y'|^{\gamma/2}\right).$$

and denote that as ($\hat{\textbf{A}}$**D3**). Thus, we say that assumption **($\hat{\textbf{A}}$D)** holds when conditions **(AD1)**,**(AD2)**, **($\hat{\textbf{A}}$D3)** are in force.

Theorem. *Fix $T > 0$. Under assumptions **($\hat{A}$D)** for any test function $f \in C^{\beta,\beta/2}(\mathbb{R}^{2d})$ ($\beta-$Hölder in the first variable and $\beta/2-$Hölder in the second variable functions) for $\beta \in (0,1]$, there exists $C > 0$, such that:*

$$|\mathbb{E}_{(x,y)}[f(X^h_T, Y^h_T)] - \mathbb{E}_{(x,y)}[f(X_T, Y_T)]| \leq Ch^{\gamma/2}(1 + |x|^{\gamma/2}).$$

where $\gamma \in (0,1]$ stands for the Hölder index of $\gamma, \gamma/2$ (γ for the variable x, $\gamma/2$ for y) Hölder continuous time-homogeneous functions b, σ.

The theorem will be restated in Section 4.4.

We also would like to present our control for the direct difference of transition densities $p(t,(x,y),(x',y'))$ and $p_h(t,(x,y),(x',y'))$. The result below is in clear contrast with the one of Theorem 4.4.1 for the weak error, i.e. when additionally one considers an integration of a Hölder function w.r.t. the final (or forward variable). We finally can reach a global error of order $h^\beta, \beta < \gamma - 1/2$ which is *close* to the expected one in $h^{\gamma/2}$ when γ goes to 1.

Theorem (4.5.1)**.** *Fix a final time horizon $T > 0$ and a time step $h = T/N, N \in \mathbb{N}^*$ for the Euler scheme. Under assumptions* ***($\hat{\mathbf{A}}$D)****, for $\gamma \in (1/2, 1]$ and $\beta \in (0, \gamma - \frac{1}{2})$, for all t in the time grid $\Lambda_h := \{(t_i)_{i \in [\![1,N]\!]}\}$ and $(x, y), (x', y') \in \mathbb{R}^{2d}$ there exist $C := (T, b, a, \beta), c > 0$ such that :*

$$\begin{aligned} &|p(t, (x, y), (x', y')) - p_h(t, (x, y), (x', y')| \\ &\leq Ch^\beta (1 + (|x| \wedge |x'|))^{1+\gamma}) \sup_{s \in [t-h, t]} p_{c,K}(s, (x, y), (x', y')), \end{aligned} \tag{18}$$

where $p_{c,K}(s, (x, y), (x', y'))$ stands for the Kolmogorov-type Gaussian density (15) *at time s.*

The theorem will be discussed in Section 4.5.

References

[Aro59] D. G Aronson. The fundamental solution of a linear parabolic equation containing a small parameter. *Illinois Journal of Mathematics*, 3:580–619, 1959.

[Bai17] N.T.G. Bailey. *The mathematical theory of infectious diseases and its applications.* 2nd ed. Hafner Press, New York, 2017.

[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.

[BT96a] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations: I. Convergence rate of the distribution function. *Probability Theory and Related Fields*, 104-1:43–60, 1996.

[BT96b] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations, II. Convergence rate of the density. *Monte Carlo Methods and Applications*, 2:93–128, 1996.

[BY89] N. Becker and P. Yip. Analysis of variations in an infection rate. *Australian Journal of Statistics*, 31(1), 1989.

[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.

[Fri18] N. Frikha. On the weak approximation of a skew diffusion by an Euler-type scheme. *Bernoulli*, 24(3):1653–1691, 2018.

[GS67] I. Gihman and A. Skorohod. *Stochastic Differential Equations.* Naukova dumka, Kiev., 1967.

[GS82] I. Gihman and A. Skorohod. *Stochastic Differential Equations and Applications*. Naukova dumka, Kiev., 1982.

[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.

[JYC10] M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical Methods for Financial Markets*. Springer Finance, London, 2010.

[KKM17] V. Konakov, A. Kozhina, and S Menozzi. Stability of densities for perturbed Diffusions and Markov Chains. *ESAIM: Probability and Statistics*, 21:88–112, 2017.

[KM00] V. Konakov and E. Mammen. Local limit theorems for transition densities of Markov chains converging to diffusions. *Probability Theory and Related Fields*, 117:551–587, 2000.

[KM02] V. Konakov and E. Mammen. Edgeworth type expansions for Euler schemes for stochastic differential equations. *Monte Carlo Methods and Applications*, 8–3:271–285, 2002.

[KM10] V. Konakov and S. Menozzi. Weak error for stable driven stochastic differential equations: Expansion of the densities. *Journal of Theoretical Probability*, 24-2:554–578, 2010.

[KM17] V. Konakov and S Menozzi. Weak Error for the Euler Scheme Approximation of Diffusions with non-smooth coefficients. *Electronic Journal of Probability*, 22:1–47, 2017.

[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.

[Kol34] A. N. Kolmogorov. Zufällige Bewegungen (zur Theorie der Brownschen Bewegung). *Annals of Mathematics*, 2-35:116–117, 1934.

[Koz16] A. Kozhina. Stability of densities for perturbed degenerate diffusions. *Teoriya Veroyatnostei i ee Primeneniya*, 3:570–579, 2016.

[Kry96] N. V. Krylov. *Lectures on elliptic and parabolic equations in Hölder spaces*. Graduate Studies in Mathematics 12. AMS, 1996.

[KS84] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. I. *Stochastic analysis (Katata/Kyoto, 1982)*, North-Holland Math. Library, 32:271–306, 1984.

[KS85] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. II. *Journal of the Faculty of Science, the University of Tokyo*, 32:1–76, 1985.

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

[Mar55] G. Maruyama. Continuous markov processes and stochastic equations. *Rendiconti del Circolo Matematico di Palermo*, 4:48, 1955.

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

[MP91] R. Mikulevičius and E. Platen. Rate of convergence of the Euler approximation for diffusion processes. *Mathematische Nachrichten*, 151:233–239, 1991.

[Son67] I. M. Sonin. A class of degenerate diffusion processes (russian). *Teoriya Veroyatnostei i ee Primeneniya*, 12:540–547, 1967.

[Tal02] D. Talay. Stochastic Hamiltonian dissipative systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Processes and Related Fields*, 8–2:163–198, 2002.

[TT90] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Analysis and Applications*, 8-4:94–120, 1990.

[Web51] M. Weber. The fundamental solution of a degenerate partial differential equation of parabolic type. *Transactions of the American Mathematical Society*, 71:24–37, 1951.