Analytical approximations of non-linear SDEs of McKean-Vlasov type

Emmanuel Gobet, Stefano Pagliarani

Analytical approximations of non-linear SDEs of McKean-Vlasov type

Emmanuel Gobet * Stefano Pagliarani †

November 11, 2016

Abstract

We provide analytical approximations for the law of the solutions to a certain class of scalar McKean-Vlasov stochastic differential equations (MKV-SDEs) with random initial datum. "Propagation of chaos" results ([Szn91]) connect this class of SDEs with the macroscopic limiting behavior of a particle, evolving within a mean-field interaction particle system, as the total number of particles tends to infinity. Here we assume the mean-field interaction only acting on the drift of each particle, this giving rise to a MKV-SDE where the drift coefficient depends on the law of the unknown solution. By perturbing the non-linear forward Kolmogorov equation associated to the MKV-SDE, we perform a two-steps approximating procedure that decouples the McKean-Vlasov interaction from the standard dependence on the state-variables. The first step yields an expansion for the marginal distribution at a given time, whereas the second yields an expansion for the transition density. Both the approximating series turn out to be asymptotically convergent in the limit of short times and small noise, the convergence order for the latter expansion being higher than for the former. The resulting approximation formulas are expressed in semi-closed form and can be then regarded as a viable alternative to the numerical simulation of the large-particle system, which can be computationally very expensive. Moreover, these results pave the way for further extensions of this approach to more general dynamics and to high-dimensional settings.

Keywords: Nonlinear SDE in the McKean sense, small time/noise approximations, interacting random processes, nonlinear parabolic equations

2010 Mathematics Subject Classification: 65C30, 34E10, 35Q83, 65K35, 35K55

*Centre de Mathématiques Appliquées, Ecole Polytechnique and CNRS, Route de Saclay, Université Paris-Saclay, 91128 Palaiseau Cedex, France. Email: `emmanuel.gobet@polytechnique.edu`. Corresponding author. The author's research is part of the Chair *Financial Risks* of the *Risk Foundation*, the *Finance for Energy Market Research Centre* and the ANR project *CAESARS* (ANR-15-CE05-0024).

†Dipartimento di Scienze Economiche, Aziendali, Matematiche e Statistiche, Università di Trieste, Via Tigor 22, 34124 Trieste, Italy. Email: `stefano.pagliarani@deams.units.it`. The author's research was supported by the Chair *Financial Risks* of the *Risk Foundation*, under the aegis of *Louis Bachelier Finance and Sustainable Growth* laboratory.

1 Introduction

Model. Consider the non-linear diffusion

$$\begin{cases} \mathrm{d}X_t = \mathbb{E}[b(x, X_t)]|_{x=X_t}\mathrm{d}t + \sigma \mathrm{d}W_t, \qquad t > 0 \\ X_0 = Y. \end{cases} \tag{1.1}$$

Here, W is a scalar Brownian motion and Y is a square integrable random variable, independent of W. Throughout the paper, we assume that there exist two positive constants $M, \bar{\sigma} > 0$ such that the following standing assumptions hold:

[**Hyp-**b.0] $b : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ is a globally Lipschitz function, and is bounded by $M\sigma^2$;

[**Hyp-**σ] The diffusion coefficient σ is such that $0 < \sigma \leq \bar{\sigma}$.

For high order expansions, [**Hyp-**b.0] will be reinforced by adding the following further assumption, for a given $N \in \mathbb{N}$, $N \geq 1$.

[**Hyp-**$b.N$] For any $y \in \mathbb{R}$, the function $b(\cdot, y) \in C^N(\mathbb{R})$ with all the derivatives $\partial_1^n b(\cdot, \cdot)$ up to order N being measurable and bounded by $M\sigma^2$. Moreover, $\partial_1 b(\cdot, \cdot)$ is continuous.

Such non-linear SDEs, where the coefficient b of the equation depends not only on the state of the solution at time t, but also on its whole distribution, is a particular case of a class of SDEs known as *McKean-type non-linear diffusions*. It is well known that, under [**Hyp-**b.0], Eq. (1.1) admits a unique strong solution (see for instance [Szn91]). The extra assumption [**Hyp-**σ], along with other additional regularity and boundedness assumptions on b, will be used to derive expansions for the density of the distribution of X_t. In particular, the need for the constant $\bar{\sigma}$ will be clarified in the sequel. Loosely speaking, it will allow to prove sharp error estimates not only for small times, but also for small σ.

Background results and main contributions. So far, the study of numerical approximations of SDEs of McKean-type has been mainly conducted under the point of view of time discretization and simulation through an interacting particles system. References are numerous and we refer to [Mél96, BT97, AK02, TV03, Tra08] among others. Recently, an alternative method using cubature formula has been investigated in [CG15]. Our approach is quite different and relies on analytical expansions; to the best of our knowledge this is fully novel in this context. We emphasize that during the last decade, there has been an increasing gain of interest in the study of SDEs of McKean-type, with new applications ranging from modeling economic interactions and mean-field games [CDL13, CD15], to financial portfolio [BK10, JR15] and neuroscience [DIRT15]. The first main contribution of the paper is a semi-closed N-th order approximation $\bar{\tilde{P}}_{N,t}$ for the density P_t of X_t, for which we are able to prove an asymptotic error bound (Theorem 2.9) that can be roughly summarized as

$$\left\| P_t - \bar{\tilde{P}}_{N,t} \right\|_{L^1(\mathbb{R})} = O\big(\sigma^2 t\big)^{\frac{N+1}{2}} \qquad \text{as } \sigma^2 t \to 0^+.$$

The second main contribution is a family of semi-closed N-th order approximations $\bar{p}_N^{\bar{x}}(s, \xi; t, x)$ for the transition density $p(s, \xi; t, x)$ of X (seen as a time-inhomogeneous standard SDE), the latter depending on

the previous approximation $\bar{\bar{P}}_{N,t}$ in a way that will be specified in Section 2. In this case we are able to prove an asymptotic result (Theorem 2.16) that roughly reads as

$$\left|\left(p-\bar{p}_N^{\bar{x}}\right)(s,\xi;t,x)\right| = e^{-\frac{(x-\xi)^2}{4\sigma^2(t-s)}} O\left(\sigma^2 t\right)^{\frac{N+1}{2}} \qquad \text{as } \sigma^2 t \to 0^+,$$

uniformly w.r.t. $x, \xi \in \mathbb{R}$.

We emphasize that, even though such results are carried out here for a scalar Mc-Kean SDE as in (1.1), our approach can be generalized to multi-dimensional settings allowing for Mc-Kean interactions not only in the drift but also in the diffusion coefficient. These extensions, as well as numerical tests to illustrate the accuracy of different approximation formulas, will be handled in a further work.

Organization of the paper. In the rest of this section we introduce extra notations, which will be used throughout the whole paper. Section 2 is then devoted to present our approximation strategy and state the main results (approximation formulas and error analysis). Section 3 gathers the proofs about the expansion of the marginal distribution of the diffusion process. The proofs about approximations of the transition density are given in Section 4.

Notation 1.1. For any random variable U, we denote by U' an independent copy of U, and by $\mathbb{E}'$ the expectation w.r.t. the distribution of U' only, i.e. $\mathbb{E}'[\varphi(U',Z)] = \mathbb{E}[\varphi(U,z)]|_{z=Z}$ for any random variable Z independent on U. With this notation we can rewrite (1.1) as

$$\begin{cases} \mathrm{d}X_t = \mathbb{E}'[b(X_t, X_t')]\mathrm{d}t + \sigma \mathrm{d}W_t, \qquad t > 0 \\ X_0 = Y. \end{cases} \tag{1.2}$$

Notation 1.2. For any measure μ on $(\mathbb{R}, \mathcal{B})$, and any function $f \in L^1(\mathbb{R}, \mu)$ we define the average

$$\mu[f] := \int_{\mathbb{R}} f(x)\mu(\mathrm{d}x).$$

Moreover, for any functions $P \in L^1(\mathbb{R}, \mathrm{Leb})$ and f such that $(fP) \in L^1(\mathbb{R}, \mathrm{Leb})$ we set

$$P[f] := \int_{\mathbb{R}} f(x)P(x)\mathrm{d}x.$$

In particular, if μ is absolutely continuous w.r.t. the Lebesgue measure with density P, i.e. $\mu(\mathrm{d}x) = P(x)\mathrm{d}x$, then we have $\mu[f] = P[f]$.

Notation 1.3. We will denote by μ_{X_t} and μ_Y the laws of the random variable X_t and Y respectively. Sometimes, to shorten notation, we well use μ_t instead of μ_{X_t} when the dependence on X is clear from the context. Moreover, under the standing assumptions **[Hyp-b.0]** and **[Hyp-σ]**, μ_t is absolutely continuous with respect to the Lebesgue measure at any time $t > 0$ (see the discussion below), and we will denote by $P_t(\cdot)$ its probability density, i.e. $\mu_t(\mathrm{d}x) = P_t(x)\mathrm{d}x$ for any $t > 0$.

Remark 1.4. Under assumptions **[Hyp-b.0]** and **[Hyp-σ]**, the function $(t,x) \mapsto \mu_t[b(x,\cdot)]$ is bounded and continuous, and the function $x \mapsto \mu_t[b(x,\cdot)]$ is Lipschitz continuous uniformly w.r.t. $t \geq 0$. Therefore, it is well known (see [Fri64, Chapter 1, p. 23]) that the parabolic operator $(\partial_s + \mathcal{A}_s^b)$, with $\mathcal{A}_s^b$ given by

$$\mathcal{A}_s^b = \frac{\sigma^2}{2}\partial_{\xi\xi} + \mu_s[b(\xi,\cdot)]\partial_\xi, \qquad s \geq 0,$$

has a fundamental solution $p^{(\mu)}$, i.e. a continuous function $p^{(\mu)}(s,\xi;t,x)$ defined for any $\xi, x \in \mathbb{R}$ and $0 \leq s < t$, such that, for any $(t,x) \in]0,\infty[\times\mathbb{R}$ the function $p^{(\mu)}(\cdot,\cdot;t,x)$ solves the backward Cauchy problem

$$\begin{cases} (\partial_s + \mathcal{A}_s^b)u = 0 \quad \text{on } [0,t[\times\mathbb{R}, \\ u(t,\cdot) = \delta_x. \end{cases} \tag{1.3}$$

Moreover, if we additionally assume [**Hyp-**$b.2$] to be in force, then the function $(t,x) \mapsto \partial_x \mu_t[b(x,\cdot)]$ is bounded and continuous, and the function $x \mapsto \partial_x \mu_t[b(x,\cdot)]$ is Lipschitz continuous uniformly w.r.t. $t \geq 0$. Thus $p^{(\mu)}(s,\xi;t,x)$ also coincides with the fundamental solution of the linear parabolic operator $(-\partial_t + \mathcal{A}_t)$ (see [Fri64, Chapter 1, p. 28]), where $\mathcal{A}_t$ is the so called *formal adjoint* of $\mathcal{A}_t^b$, acting as

$$\mathcal{A}_t u(x) = \frac{\sigma^2}{2}\partial_{xx}u(x) - \partial_x\big(u(x)\mu_t[b(x,\cdot)]\big), \qquad t \geq 0. \tag{1.4}$$

In particular, for any $(s,\xi) \in [0,\infty[\times\mathbb{R}$ the function $p^{(\mu)}(s,\xi;\cdot,\cdot)$ solves the forward Cauchy problem

$$\begin{cases} (-\partial_t + \mathcal{A}_t)u = 0 \quad \text{on }]s,\infty[\times\mathbb{R}, \\ u(s,\cdot) = \delta_\xi. \end{cases}$$

Note that, once existence of the solution X is ensured, the McKean-Vlasov SDE (1.1) can also be regarded, a fortiori, as an ordinary SDE with random initial condition Y and *unknown* variable drift-coefficient $\mu_t[b(x,\cdot)]$, i.e.

$$\begin{cases} \mathrm{d}X_t = \mu_t[b(X_t,\cdot)]\mathrm{d}t + \sigma \mathrm{d}W_t, \qquad t > 0 \\ X_0 = Y. \end{cases} \tag{1.5}$$

Therefore, a simple application of Feynman-Kac representation formulas shows that the Markovian process X, solution of (1.5), has a transition density kernel that coincides with the fundamental solution $p^{(\mu)}$ in Remark 1.4. Precisely, $p^{(\mu)}(s,\cdot;t,\cdot)$ is the density of the marginal X_t of the process conditioned to X_s. Note that the superscript (μ) emphasizes the fact that the transition kernel does depend on the distribution of X; in particular it depends on the $\mu_u(.)$ for any $0 < u \leq t$, and on the initial distribution μ_Y. This fact represents a key difference with respect to standard SDEs.

Now, by Chapman-Kolmogorov equation, we can conclude that the distribution μ_t has a density P_t given by

$$P_t(x) = \int_{\mathbb{R}} p^{(\mu)}(0,\xi;t,x)\mu_Y(\mathrm{d}\xi), \qquad t > 0, \quad x \in \mathbb{R}.$$

In particular, by changing the notation of $p^{(\mu)}$ into $p^{(P)}$, the density P can be regarded as the solution of the fixed-point functional equation

$$P_t(x) = \int_{\mathbb{R}} p^{(P)}(0,\xi;t,x)\mu_Y(\mathrm{d}\xi), \qquad t > 0, \quad x \in \mathbb{R}. \tag{1.6}$$

Hereafter, to simplify the notation, we will suppress the suffix P, or μ, and we will use $p(s,\xi;t,x)$ to denote the transition density of X. However, the reader should always bear in mind that the transition density p depends on the law μ (with density P) of the process X.

Notation

- $\mathbb{N} = \{0, 1, \dots, n, \dots\}$ denotes the set of non-negative integers.
- $\mathbb{N}^* = \mathbb{N}\backslash\{0\}$ denotes the set of positive integers.

2 Analytical Approximations

In order to expand the density $P_t(\cdot)$ of the solution X_t to the MKV SDE (1.5) we propose a *two-steps* (or *decoupling*) perturbation scheme. Loosely speaking, the main idea is as follows: we obtain a first approximating expansion $\bar{\bar{P}}_{N,t}$ of the marginal P_t, then we further approximate the transition density of the solution to the SDE

$$\mathrm{d}Y_t = \bar{\bar{P}}_{N,t}[b(Y_t, \cdot)]\mathrm{d}t + \sigma \mathrm{d}W_t, \qquad t > 0, \tag{2.1}$$

and finally, we obtain an new approximation of P_t by reintegrating w.r.t. μ_Y (see Eq. (1.6)).

Besides providing with a first approximation for the marginal distributions μ_t of the solution X_t, the first step is relevant because it allows to separate (or decouple) the two kinds of interactions in (1.5): the McKean-Vlasov interaction through the law of the solution, and that through the realization of the solution. This first step should be regarded as the main element of novelty in this paper. Once the decoupling of the two interactions is done, the problem boils down to approximating the transition density of a standard SDE of the type (2.1), where the coefficient only depends on state and time. Thus the second approximation, the one for the transition density, follows by adapting some previous PDE techniques that allow to expand the transition density of the solution to a standard SDE (see [LPP15]). Note that the latter techniques admit a stochastic counterpart that allows to obtain similar results (see [BG12] for a review). It is important to mention that, however certainly more standard, this second step returns a higher order approximation compared to the first one. The improvement is relevant because it has a major impact for low values of N, which are the only cases when the approximation can be easily computed in practice.

2.1 Expansion of the marginal distributions

We carry out an approximating expansion for the marginal distributions (marginal densities) $P_t(\cdot)$ of X_t.

2.1.1 Approximation strategy

We introduce an interpolation parameter ε. For any $\varepsilon \in [0, 1]$, let us consider the family of McKean SDEs given by

$$\begin{cases} \mathrm{d}\tilde{X}_t^\varepsilon = \varepsilon \mathbb{E}'\big[b\big(Y + \varepsilon(\tilde{X}_t^\varepsilon - Y), \tilde{X}_t'^\varepsilon\big)\big]\mathrm{d}t + \sigma \mathrm{d}W_t, \qquad t > 0, \\ \tilde{X}_0^\varepsilon = Y. \end{cases} \tag{2.2}$$

Due to the presence of $Y = \tilde{X}_0^\varepsilon$ in the b-term, this is a path-dependent McKean SDE: Lemma B.1 justifies the existence and uniqueness of a strong solution to the above equation under **[Hyp-b.0]**. Note that, if $\varepsilon = 1$, then (2.2) reduces to the original McKean equation (1.2). For any $t > 0$, denote by $\tilde{P}_t^\varepsilon(\cdot)$ the density of $\tilde{X}_t^\varepsilon$, which exists for the same reason as for P_t (see Remark 1.4 and following discussion). The density

$\tilde{P}^\varepsilon$ can be given an interpretation within the PDE framework. In fact, even though the process $\tilde{X}^\varepsilon$ in (2.2) is not Markovian, it becomes Markovian when conditioned to Y. Therefore $\tilde{P}^\varepsilon$ can be written as

$$\tilde{P}^\varepsilon_t(x) = \int_{\mathbb{R}} \tilde{p}^\varepsilon(0,\xi;t,x)\mu_Y(\mathrm{d}\xi), \qquad t>0, \quad x\in\mathbb{R},$$

where $\tilde{p}^\varepsilon = \tilde{p}^\varepsilon(0,\xi;t,\cdot)$ is the density of the marginal $\tilde{X}^\varepsilon_t|_{Y=\xi}$, or the fundamental solution of the linear parabolic operator $\big(-\partial_t + \tilde{\mathcal{A}}^{\varepsilon,\xi}\big)$, with $\tilde{\mathcal{A}}^{\varepsilon,\xi}$ acting as

$$\tilde{\mathcal{A}}^{\varepsilon,\xi}u(t,x) = \frac{\sigma^2}{2}\partial_{xx}u(t,x) - \varepsilon\partial_x\Big(u(t,x)\tilde{P}^\varepsilon_t\big[b\big(\xi+\varepsilon(x-\xi),\cdot\big)\big]\Big). \tag{2.3}$$

In particular, the function $\tilde{p}^\varepsilon(0,\xi;\cdot,\cdot)$ satisfies

$$\begin{cases} \big(-\partial_t + \tilde{\mathcal{A}}^{\varepsilon,\xi}\big)\tilde{p}^\varepsilon(0,\xi;\cdot,\cdot) = 0 & \text{on }]0,\infty[\times\mathbb{R}, \\ \tilde{p}^\varepsilon(0,\xi;0,\cdot) = \delta_\xi. \end{cases} \tag{2.4}$$

For a given $N\in\mathbb{N}^*$, consider the N-th order approximations

$$\tilde{p}^\varepsilon(0,\xi;t,x)|_{\varepsilon=1} \approx \bar{\tilde{p}}_N(0,\xi;t,x) := \sum_{n=0}^N \frac{1}{n!}\tilde{p}_n(0,\xi;t,x), \quad 0\le s<t, \quad x,\xi\in\mathbb{R}, \tag{2.5}$$

$$\begin{aligned} \tilde{P}^\varepsilon_t(x)|_{\varepsilon=1} \approx \bar{\tilde{P}}_{N,t}(x) &:= \int_{\mathbb{R}} \bar{\tilde{p}}_N(0,\xi;t,x)\mu_Y(\mathrm{d}\xi) = \sum_{n=0}^N \frac{1}{n!}\tilde{P}_{n,t}(x), \\ &\text{with} \quad \tilde{P}_{n,t}(x) = \int_{\mathbb{R}} \tilde{p}_n(0,\xi;t,x)\mu_Y(\mathrm{d}\xi), \qquad t>0,\ x\in\mathbb{R}, \end{aligned} \tag{2.6}$$

where the rigorous definition of each function $\tilde{p}_n$ will be given later. Had $\tilde{p}^\varepsilon$ and $\tilde{P}^\varepsilon$ had a Taylor expansion in ε, we would have taken naturally

$$\tilde{p}_n(s,\xi;t,x) = \frac{\mathrm{d}^n}{\mathrm{d}\varepsilon^n}\tilde{p}^\varepsilon(s,\xi;t,x)\Big|_{\varepsilon=0}, \qquad \tilde{P}_{n,t}(x) = \frac{\mathrm{d}^n}{\mathrm{d}\varepsilon^n}\tilde{P}^\varepsilon_t(x)\Big|_{\varepsilon=0}, \qquad 0\le s<t, \quad x,\xi\in\mathbb{R}. \tag{2.7}$$

In what follows we will not prove, strictly mathematically speaking, the existence of Taylor expansions in ε, but this Taylor expansion principle will guide us through the definitions of $\tilde{p}_n$ (and therefore $\tilde{P}_{n,t}$ owing to (2.6)).

2.1.2 0-th order approximation

Here we will give an explicit representation of the leading term $\tilde{P}_{0,t} := \tilde{P}^0_t$ appearing in the expansion (2.6) of the marginal density $\tilde{P}^\varepsilon_t$.

All the definitions of this subsection are well posed under assumption [**Hyp-**$b.0$] and [**Hyp-**σ]. By setting $\varepsilon=0$ in (2.2) we obtain

$$\begin{cases} \mathrm{d}\tilde{X}^0_t = \sigma\mathrm{d}W_t, \qquad t>0, \\ \tilde{X}^0_0 = Y, \end{cases}$$

which clearly yields

$$\tilde{P}_{0,t}(x) := \int_{\mathbb{R}} \tilde{p}_0(0,\xi;t,x)\mu_Y(\mathrm{d}\xi), \qquad\qquad t>0, \quad x\in\mathbb{R}, \tag{2.8}$$

$$\tilde{p}_0(0,\xi;t,x) := \Gamma_\sigma\big(t, x-\xi\big), \qquad t>0, \quad \xi, x \in \mathbb{R}, \tag{2.9}$$

where $\Gamma_\sigma(\cdot,\cdot)$ is the Gaussian density with variance proportional to σ^2

$$\Gamma_\sigma(t,x) := \frac{1}{\sigma\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2\sigma^2 t}\right), \qquad x \in \mathbb{R},\ t>0.$$

In order to maintain the parallel with the PDE's setting, note that, by setting $\varepsilon = 0$ in (2.3), we have that the kernel $\tilde{p}_0$ represents the fundamental solution of the operator $\big(-\partial_t + \tilde{\mathcal{A}}_0\big)$, with

$$\tilde{\mathcal{A}}_0 = \frac{\sigma^2}{2}\partial_{xx}, \tag{2.10}$$

i.e. for any $\xi \in \mathbb{R}$ the function $\tilde{p}_0(0,\xi;\cdot,\cdot)$ solves the forward Cauchy problem

$$\begin{cases} (-\partial_t + \tilde{\mathcal{A}}_0)u = 0 \quad \text{on }]0,\infty[\times\mathbb{R}, \\ u(0,\cdot) = \delta_\xi. \end{cases} \tag{2.11}$$

2.1.3 Higher orders approximations

In order to achieve higher orders expansions for the marginal density $P_t = \tilde{P}_t^\varepsilon|_{\varepsilon=1}$, we follow the strategy explained previously. Hereafter throughout this subsection we fix $N \in \mathbb{N}^*$ and we assume the assumptions [**Hyp-**b**.**$N-1$] and [**Hyp-**σ] to be in force.

Formal derivation. We start by freely assuming that all quantities are smooth in ε and that all subsequent PDEs are well posed. It will enable us to formally represent the terms $\tilde{p}_n(s,\xi;t,x)$ appearing in (2.5) as the solutions of some nested PDEs. By formally differentiating both sides of (2.4) we obtain

$$\begin{aligned}
\big(-\partial_t + \tilde{\mathcal{A}}_0\big)\partial_\varepsilon^n \tilde{p}^\varepsilon(0,\xi;t,x) &= \partial_x \partial_\varepsilon^n \Big(\varepsilon\, \tilde{p}^\varepsilon(0;\xi;t,x)\tilde{P}_t^\varepsilon\big[b\big(\xi+\varepsilon(x-\xi),\cdot\big)\big]\Big) \\
&= \varepsilon \partial_x \partial_\varepsilon^n \Big(\tilde{p}^\varepsilon(0;\xi;t,x)\tilde{P}_t^\varepsilon\big[b\big(\xi+\varepsilon(x-\xi),\cdot\big)\big]\Big) \\
&\quad + n\partial_x \sum_{h=0}^{n-1}\sum_{i=0}^{h} \binom{n-1}{h}\binom{h}{i} \big(\partial_\varepsilon^{n-1-h}\tilde{p}^\varepsilon(0;\xi;t,x)\big) \\
&\qquad\qquad \cdot (x-\xi)^i \big(\partial_\varepsilon^{h-i}\tilde{P}_t^\varepsilon\big)\big[\partial_1^i b\big(\xi+\varepsilon(x-\xi),\cdot\big)\big],
\end{aligned} \tag{2.12}$$

along with the terminal condition $\partial_\varepsilon^n \tilde{p}^\varepsilon(0,\xi;0,x) = 0$. Now, if we were to define $\tilde{p}_n(0,\xi;t,x)$ as in (2.7), by shifting the index h, reorganizing the binomial coefficients, and setting now $\varepsilon = 0$ in (2.12) we would obtain

$$\begin{cases} \big(-\partial_t + \tilde{\mathcal{A}}_0\big)\tilde{p}_n(0,\xi;\cdot,\cdot) = \sum_{h=1}^n \sum_{i=0}^{h-1} h\binom{n}{h}\binom{h-1}{i}\mathcal{B}^\xi_{h-1,i}(\cdot,\cdot)\,\tilde{p}_{n-h}(0;\xi;\cdot,\cdot), & \text{on }]0,\infty[\times\mathbb{R}, \\ \tilde{p}_n(0,\xi;0,\cdot) = 0, & \text{on } \mathbb{R}, \end{cases} \tag{2.13}$$

with $\mathcal{B}^\xi_{k,i} = \mathcal{B}^\xi_{k,i}(\cdot,\cdot)$ being the differential operator acting as

$$\big(\mathcal{B}^\xi_{k,i}(t,x)u\big)(t,x) = \tilde{P}_{k-i,t}\big[\partial_1^i b(\xi,\cdot)\big]\partial_x\big((x-\xi)^i u(t,x)\big), \qquad 0 \le i \le k. \tag{2.14}$$

Rigorous definition. Although all the previous computations were only meant to be heuristic, the Cauchy problems (2.13) lead us through giving rigorous definitions for the expansion terms $\tilde{p}_n$ entering in the definition (2.6) of the approximate marginal distribution $\bar{\tilde{P}}_{N,t}$. Precisely, by applying Duhamel's principle we can give integral definitions that are coherent with aforementioned Cauchy problems. A fortiori, in Section 3.2 (Remark 3.6) it will be shown that $\tilde{p}_n$ actually solves (2.13), with the terminal condition meant in the distributional sense.

Definition 2.1. For any $n \in \mathbb{N}^*$ with $n \leq N$, set

$$\tilde{P}_{n,t}(x) := \int_{\mathbb{R}} \tilde{p}_n(0,\xi;t,x)\mu_Y(\mathrm{d}\xi), \qquad t>0, \quad x \in \mathbb{R}, \tag{2.15}$$

where, for any $\xi \in \mathbb{R}$, the functions $\big(\tilde{p}_n(0,\xi;\cdot,\cdot)\big)_{1\leq n\leq N}$ are defined, recursively, as

$$\tilde{p}_n(0,\xi;t,x) := -\sum_{h=1}^{n}\sum_{i=0}^{h-1} h\binom{n}{h}\binom{h-1}{i}\int_0^t\int_{\mathbb{R}} \Gamma_\sigma(t-s;x-y)\mathcal{B}^{\xi}_{h-1,i}(s,y)\,\tilde{p}_{n-h}(0;\xi;s,y)\mathrm{d}y\mathrm{d}s \tag{2.16}$$

with $\tilde{p}_0$ as defined in (2.9).

As it is stated in Theorem 2.2 below, Definition 2.1 is well posed under assumptions [**Hyp-**$b.N-1$] and [**Hyp-**σ], and the function $\tilde{P}_n$ can be given a representation in terms of differential operators acting on the leading term $\tilde{P}_0$. Such representation can be manipulated, see Corollary 2.5, in order to achieve a fully explicit characterization.

Theorem 2.2. *Under assumptions* [**Hyp-**$b.N-1$] *and* [**Hyp-**σ], *for any $n \in \mathbb{N}^*$ with $n \leq N$, we have:*

(i) the functions $\tilde{P}_n$ and $\tilde{p}_n$ as in (2.15)-(2.16) *are well defined;*

(ii) it holds:

$$\tilde{p}_n(0,\xi;t,x) = -\tilde{\mathcal{L}}^{\xi}_n(t,x)\,\tilde{p}_0(0,\xi;t,x), \qquad t>0, \quad x,\xi\in\mathbb{R} \tag{2.17}$$

where $\tilde{\mathcal{L}}^{\xi}_n = \tilde{\mathcal{L}}^{\xi}_n(\cdot,\cdot)$ is the differential operator defined as

$$\tilde{\mathcal{L}}^{\xi}_n(t,x) := n!\sum_{h=1}^{n}\int_0^t \mathrm{d}s_1\int_0^{s_1}\mathrm{d}s_2\cdots\int_0^{s_{h-1}}\mathrm{d}s_h \sum_{i\in I_{n,h}}\tilde{\mathcal{G}}^{\xi}_{i_1}(s_1,t,x)\cdots\tilde{\mathcal{G}}^{\xi}_{i_h}(s_h,t,x),$$

where[1]

$$I_{n,h} = \{i=(i_1,\dots,i_h)\in(\mathbb{N}^*)^h \mid i_1+\cdots+i_h = n\}, \qquad 1\leq h\leq n, \tag{2.18}$$

and the operator $\tilde{\mathcal{G}}^{\xi}_k = \tilde{\mathcal{G}}^{\xi}_k(\cdot,\cdot,\cdot)$ is defined as, for $k\geq 1$,

$$\tilde{\mathcal{G}}^{\xi}_k(s,t,x) := \frac{1}{(k-1)!}\sum_{j=0}^{k-1}\binom{k-1}{j}\mathcal{B}^{\xi}_{k-1,j}\big(s,\tilde{\mathcal{M}}(s,t,x)\big), \qquad \tilde{\mathcal{M}}(s,t,x) := x+\sigma^2(t-s)\,\partial_x, \tag{2.19}$$

In (2.19), $\mathcal{B}^{\xi}_{k-1,j}\big(s,\tilde{\mathcal{M}}(s,t,x)\big)$ *is a slight abuse of notation for the operator $\mathcal{B}^{\xi}_{k-1,j}$ acting as in* (2.14) *composed with $\tilde{\mathcal{M}}$, i.e. $\mathcal{B}^{\xi}_{k-1,j}\big(s,\tilde{\mathcal{M}}(s,t,x)\big)u = \tilde{P}_{k-1-j,s}\big[\partial_1^j b(\xi,\cdot)\big]\partial_x\big((x+\sigma^2(t-s)\,\partial_x-\xi)^j u\big)$.*

[1] For instance, for $n=3$ we have $I_{3,3}=\{(1,1,1)\}$, $I_{3,2}=\{(1,2),(2,1)\}$ and $I_{3,1}=\{(3)\}$.

We are now going to re-write the operator $\tilde{\mathcal{L}}_n^\xi(t,x)$ in a different way that is more explicit, though less intuitive. Such representation is useful in the practical implementation of the expansion. We state the next result in terms of the following functions and operators.

Notation 2.3. For any $l,j,i,m \in \mathbb{N}^h$ and $\xi \in \mathbb{R}$, we let the function $\mathbf{F}^\xi_{\mathbf{0},l,j,i} = \mathbf{F}^\xi_{\mathbf{0},l,j,i}(\cdot)$ be defined as

$$\mathbf{F}^\xi_{\mathbf{0},l,j,i}(t) := \int_0^t \mathrm{d}s_1 \cdots \int_0^{s_{h-1}} \mathrm{d}s_h \prod_{q=1,\cdots,h} \tilde{P}_{i_q-j_q,s_q}\big[\partial_1^{j_q} b(\xi,\cdot)\big]\big(\sigma^2(t-s_q)\big)^{l_q}, \qquad t \geq 0, \tag{2.20}$$

the operator $\mathcal{H}^\xi_{j,l} = \mathcal{H}^\xi_{j,l}(\cdot)$ be defined as

$$\begin{aligned}\mathcal{H}^\xi_{j,l}(x) := \Big((x-\xi)^{j_1-l_1-1}\big(j_1-l_1+(x-\xi)\partial_x\big)\partial_x^{l_1}\Big)\Big((x-\xi)^{j_2-l_2-1}\big(j_2-l_2+(x-\xi)\partial_x\big)\partial_x^{l_2}\Big)\cdots \\ \cdots\Big((x-\xi)^{j_h-l_h-1}\big(j_h-l_h+(x-\xi)\partial_x\big)\partial_x^{l_h}\Big), \qquad x \in \mathbb{R},\end{aligned} \tag{2.21}$$

and the constant $c_{\mathbf{0},m,l,j,i}$ be defined as

$$c_{\mathbf{0},m,l,j,i} := \prod_{q=1,\cdots,h} (2m_q-1)!!\binom{i_q}{j_q}\binom{j_q}{2m_q}\binom{j_q-2m_q}{l_q}.$$

Remark 2.4. For any $l,j \in \mathbb{N}^h$ we have

$$\mathcal{H}^\xi_{j,l}(x) = \mathcal{H}^\xi_{j_1,l_1}(x)\mathcal{H}^\xi_{j_2,l_2}(x)\cdots\mathcal{H}^\xi_{j_h,l_h}(x), \qquad \xi,x \in \mathbb{R}.$$

Corollary 2.5. *Under assumptions* [**Hyp-**$b.N-1$] *and* [**Hyp-**σ]*, for any $n \in \mathbb{N}^*$ with $n \leq N$, the function $\tilde{p}_n(0,\xi;t,x)$ is as in* (2.17) *where*

$$\tilde{\mathcal{L}}_n^\xi(t,x) := n!\sum_{h=1}^n \sum_{i\in I_{n,h}} \Big(\prod_{q=1,\cdots,h}\frac{1}{(i_q-1)!}\Big) \sum_{\substack{m,l,j\in\mathbb{N}^h \\ l_q\leq j_q\leq i_q-1 \\ m_q\leq\lfloor\frac{j_q-l_q}{2}\rfloor}} c_{\mathbf{0},m,l,j,i-\mathbf{1}}\,\mathbf{F}^\xi_{\mathbf{0},l+m,j,i-\mathbf{1}}(t)\,\mathcal{H}^\xi_{j-2m,l}(x), \tag{2.22}$$

with $I_{n,h}$ as in (2.18).

Proof. Consider the operator $\mathcal{O}$ acting as $(\mathcal{O}f)(x) = \big(a(x-c)+b\partial_x\big)f(x)$. Then, one can prove (proof made with Mathematica) that

$$(\mathcal{O}^j f)(x) = \sum_{l=0}^j \sum_{m=0}^{\lfloor\frac{j-l}{2}\rfloor} (2m-1)!!\binom{j}{2m}\binom{j-2m}{l}(x-c)^{j-l-2m}a^{j-l-m}b^{l+m}\partial_x^l f(x).$$

Therefore, the operators $\tilde{\mathcal{G}}_k^\xi$ in (2.19) can be represented in a more explicit, though less compact, fashion as follows:

$$\begin{aligned}\tilde{\mathcal{G}}_k^\xi = \tilde{\mathcal{G}}_k^\xi(s,t,x) = \frac{1}{(k-1)!}\sum_{j=0}^{k-1}\sum_{l=0}^{j}\sum_{m=0}^{\lfloor\frac{j-l}{2}\rfloor}(2m-1)!!\binom{k-1}{j}\binom{j}{2m}\binom{j-2m}{l}\tilde{P}_{k-1-j,s}\big[\partial_1^j b(\xi,\cdot)\big] \\ \big(\sigma^2(t-s)\big)^{l+m}(x-\xi)^{j-2m-l-1}\big(j-2m-l+(x-\xi)\partial_x\big)\partial_x^l.\end{aligned}$$

Finally Corollary 2.5 stems from Theorem 2.2. □

Remark 2.6. It is important to observe that the representation (2.17)-(2.22) for $\tilde{P}_{n,t}(x)$ is fully explicit up to computing the coefficient functions $\mathbf{F}^{\xi}_{\mathbf{0},l,j,i}(t)$ defined in (2.20). In fact, the operator $\mathcal{H}^{\xi}_{j,l}(x)$ can be easily computed even at high orders, possibly by means of a symbolic computation software.

Example 2.7. For $n = 1$ we have

$$\tilde{\mathcal{L}}^{\xi}_1(t,x) = \mathbf{F}^{\xi}_{0,0,0,0}(t)\mathcal{H}^{\xi}_{0,0}(x), \qquad \mathbf{F}^{\xi}_{0,0,0,0}(t) = \int_0^t \tilde{P}_{0,r}\big[b(\xi,\cdot)\big]\mathrm{d}r, \qquad \mathcal{H}^{\xi}_{0,0}(x) = \partial_x.$$

Example 2.8. For $n = 2$ we have

$$\begin{aligned}\tilde{\mathcal{L}}^{\xi}_2(t,x) = 2\Big(&\mathbf{F}^{\xi}_{0,0,0,1}(t)\mathcal{H}^{\xi}_{0,0}(x) + \mathbf{F}^{\xi}_{0,0,1,1}(t)\mathcal{H}^{\xi}_{1,0}(x) + \mathbf{F}^{\xi}_{0,1,1,1}(t)\mathcal{H}^{\xi}_{1,1}(x)\\ &+ \mathbf{F}^{\xi}_{(0,0),(0,0),(0,0),(0,0)}(t)\mathcal{H}^{\xi}_{(0,0),(0,0)}(x)\Big),\end{aligned}$$

with

$$\begin{aligned}\mathbf{F}^{\xi}_{0,0,0,1}(t) &= \int_0^t \tilde{P}_{1,s_1}\big[b(\bar{x},\cdot)\big]\mathrm{d}s_1, \qquad \mathbf{F}^{\xi}_{0,0,1,1}(t) = \int_0^t \tilde{P}_{0,s_1}\big[\partial_1 b(\bar{x},\cdot)\big]\mathrm{d}s_1,\\ \mathbf{F}^{\xi}_{0,1,1,1}(t) &= \int_0^t \tilde{P}_{0,s_1}\big[\partial_1 b(\bar{x},\cdot)\big]\sigma^2(t-s_1)\,\mathrm{d}s_1,\\ \mathbf{F}^{\xi}_{(0,0),(0,0),(0,0),(0,0)}(t) &= \int_0^t \tilde{P}_{0,s_1}\big[b(\xi,\cdot)\big]\,\mathrm{d}s_1 \int_0^{s_1} \tilde{P}_{0,s_2}\big[b(\xi,\cdot)\big]\,\mathrm{d}s_2,\end{aligned}$$

and

$$\mathcal{H}^{\xi}_{0,0}(x) = \partial_x, \qquad \mathcal{H}^{\xi}_{1,0}(x) = \big(1 + (x-\xi)\partial_x\big), \qquad \mathcal{H}^{\xi}_{1,1}(x) = \mathcal{H}^{\xi}_{(0,0),(0,0)}(x) = \partial_x^2.$$

2.1.4 Error estimates

In this subsection we provide some rigorous error bounds for the N-th order approximation $\bar{\tilde{P}}_{N,t}$ of the true marginal density P_t, as it is defined in (2.6).

We are in the position to prove the following result.

Theorem 2.9. *Let $\bar{N} \in \mathbb{N}^*$, $T > 0$, and assume* [**Hyp-σ**], [**Hyp-b.0**] *and* [**Hyp-b.$\bar{N}+1$**] *to be in force. Then, for any $N \in \mathbb{N}$ with $N \le \bar{N}+1$, and for any $c > 1$, we have*

$$\big|\big(p - \bar{\tilde{p}}_N\big)(0,\xi;t,x)\big| \le C\big(\sigma^2 t\big)^{\frac{N+1}{2}}\Gamma_{c\sigma}\big(t, x-\xi\big), \qquad \xi, x \in \mathbb{R},\ 0 < t \le T, \tag{2.23}$$

where $C > 0$ only depends on $\bar{N}, T, c$, and on the constants M, $\bar{\sigma}$ of assumptions [**Hyp-σ**], [**Hyp-b.$\bar{N}+1$**]. *In particular, we have*

$$\big\|P_t - \bar{\tilde{P}}_{N,t}\big\|_{L^1(\mathbb{R})} \le C\big(\sigma^2 t\big)^{\frac{N+1}{2}}, \qquad 0 < t \le T.$$

2.2 Expansion of the transition density

We carry out an approximating expansion for the transition densities $p(s,\xi;t,x)$ of X.

2.2.1 Approximation strategy

We consider the family of standard Markovian non-homogeneous SDEs, indexed by $\bar{x} \in \mathbb{R}$ and $\varepsilon \in [0,1]$, given by

$$\mathrm{d}X_t^{\varepsilon,\bar{x}} = \tilde{P}_t^{\varepsilon}\big[b\big(\bar{x} + \varepsilon(X_t^{\varepsilon,\bar{x}} - \bar{x}), \cdot\big)\big]\mathrm{d}t + \sigma \mathrm{d}W_t, \qquad t > 0. \tag{2.24}$$

Recall that $\tilde{P}_t^{\varepsilon}$ represents the marginal distribution of the process $\tilde{X}_t^{\varepsilon}$ in (2.2). The initial point $X_0^{\varepsilon,\bar{x}}$ is deterministic; its value is unimportant since we are interested only in the transition density of $X^{\varepsilon,\bar{x}}$. Again, one can observe that, if $\varepsilon = 1$, then (2.24) with initial point sampled according to μ_Y reduces to the original McKean SDE (1.5). We now denote by $p^{\varepsilon,\bar{x}}(s,\xi;t,x)$, $s < t$, the transition density of $X^{\varepsilon,\bar{x}}$. To ease the notation we preferred here to use $p^{\varepsilon,\bar{x}}$ instead of $p^{(\tilde{P}^{\varepsilon}),\varepsilon,\bar{x}}$; however, the reader should bear in mind that the drift coefficient in (2.24), and thus also the kernel $p^{\varepsilon,\bar{x}}$, do depend on the density $\tilde{P}^{\varepsilon}$. From the PDEs perspective, the kernel $p^{\varepsilon,\bar{x}}$ can be interpreted as the fundamental solution of the linear parabolic operator $\big(-\partial_t + \mathcal{A}^{\varepsilon,\bar{x}}\big)$, with $\mathcal{A}^{\varepsilon,\bar{x}}$ acting as

$$\mathcal{A}^{\varepsilon,\bar{x}}u(t,x) = \frac{\sigma^2}{2}\partial_{xx}u(t,x) - \partial_x\Big(u(t,x)\tilde{P}_t^{\varepsilon}\big[b\big(\bar{x} + \varepsilon(x-\bar{x}), \cdot\big)\big]\Big).$$

In particular, the function $p^{\varepsilon,\bar{x}}(s,\xi;\cdot,\cdot)$ satisfies

$$\begin{cases} \big(-\partial_t + \mathcal{A}^{\varepsilon,\bar{x}}\big)p^{\varepsilon,\bar{x}}(s,\xi;\cdot,\cdot) = 0 & \text{on }]s,\infty[\times\mathbb{R}, \\ p^{\varepsilon,\bar{x}}(s,\xi;s,\cdot) = \delta_{\xi}. \end{cases} \tag{2.25}$$

Although (2.2) and (2.24) with μ_Y-random initialization coincide at $\varepsilon = 1$, they differ at $\varepsilon \neq 1$ because of the different scalings in ε and because of different form of the interpolated drift (involving Y or a fixed point $\bar{x}$). This difference is instrumental for our decoupling approach.

For a fixed $\bar{x} \in \mathbb{R}$ and given $N \in \mathbb{N}^*$, consider the N-th order approximation,

$$p^{\varepsilon,\bar{x}}(s,\xi;t,x)|_{\varepsilon=1} \approx \bar{p}_N^{\bar{x}}(s,\xi;t,x) := \sum_{n=0}^{N}\frac{1}{n!}p_n^{\bar{x}}(s,\xi;t,x), \qquad 0 \le s < t, \quad x,\xi \in \mathbb{R}, \tag{2.26}$$

where the rigorous definition of each function $p_n^{\bar{x}}$ will be provided later. Had $p^{\varepsilon,\bar{x}}$ had a Taylor series expansion in ε, we would take naturally

$$p_n^{\bar{x}}(s,\xi;t,x) = \frac{\mathrm{d}^n}{\mathrm{d}\varepsilon^n}p^{\varepsilon,\bar{x}}(s,\xi;t,x)\Big|_{\varepsilon=0}; \tag{2.27}$$

this principle will serve as a guide to define $p_n^{\bar{x}}$.

Remark 2.10. Instead of considering the parametrization in (2.24), one might prefer to fix $N \in \mathbb{N}$ and consider the dynamics

$$\mathrm{d}X_t^{\varepsilon,\bar{x}} = \tilde{\bar{P}}_{N,t}\big[b\big(\bar{x} + \varepsilon(X_t^{\varepsilon,\bar{x}} - \bar{x}), \cdot\big)\big]\mathrm{d}t + \sigma \mathrm{d}W_t, \qquad t > 0. \tag{2.28}$$

Now, an expansion for the transition density would be readily available be simply employing perturbation methods for standard SDEs and linear PDEs, like those introduced by the authors in [BG12], [LPP15]. Although this way of proceeding would certainly return an approximation with the same order of asymptotic convergence as the expansion in (2.26)-(2.27), the expansion stemming from (2.28) would contain some extra terms that lengthen the approximation formulas without improving the order of convergence.

Now, reintegrating w.r.t. μ_Y, by setting $\bar{x} = \xi$ or $\bar{x} = x$, and in view of (1.6), we obtain two *N-th order approximations* for the marginal density $P_t(x)$ of X_t, namely

$$P^{\text{start}}_{N,t}(x) := \int_{\mathbb{R}} \bar{p}^{\xi}_N(0,\xi;t,x)\mu_Y(\mathrm{d}\xi), \qquad P^{\text{end}}_{N,t}(x) := \int_{\mathbb{R}} \bar{p}^{x}_N(0,\xi;t,x)\mu_Y(\mathrm{d}\xi), \qquad t>0,\ x\in\mathbb{R}.$$

In general, we will show that it is a sensible choice any $\bar{x}$ that lies on the segment connecting the initial point ξ and the terminal point x, which yields the λ-approximation

$$P^{\lambda}_{N,t}(x) := \int_{\mathbb{R}} \bar{p}^{\bar{x}}_N(0,\xi;t,x)\big|_{\bar{x}=\lambda x+(1-\lambda)\xi}\,\mu_Y(\mathrm{d}\xi), \qquad \lambda\in[0,1]. \tag{2.29}$$

Note that $P^{\lambda}_{N,t}$ reduces to $P^{\text{start}}_{N,t}$ and $P^{\text{start}}_{N,t}$ for $\lambda = 0$ and $\lambda = 1$, respectively. Theoretically, all the above approximations have the same asymptotic accuracy, uniform in λ, in the limit $\sigma^2 t \to 0$ (see Theorem 2.16). Their actual accuracy for different choices of λ will be compared in a further work.

2.2.2 0-th order approximation

Here we will give an explicit representation of the leading term $p^{\bar{x}}_0(s,\xi;t,x) := p^{0,\bar{x}}(s,\xi;t,x)$ appearing in the expansion (2.26)-(2.27) of the transition density $p^{\varepsilon,\bar{x}}(s,\xi;t,x)$.

All the definitions of this subsection are well posed under assumptions [**Hyp-**$b.0$] and [**Hyp-**σ]. By setting $\varepsilon = 0$ in (2.24) we obtain

$$\mathrm{d}X^{0,\bar{x}}_t = \tilde{P}_{0,t}[b(\bar{x},\cdot)]\mathrm{d}t + \sigma\,\mathrm{d}W_t, \qquad t>0,$$

which yields

$$p^{\bar{x}}_0(s,\xi;t,x) := \Gamma_\sigma\big(t-s, x-\xi-\mathbf{m}^{\bar{x}}(s,t)\big), \qquad \mathbf{m}^{\bar{x}}(s,t) := \int_s^t \tilde{P}_{0,r}[b(\bar{x},\cdot)]\mathrm{d}r, \qquad 0\le s<t, \quad \xi,x\in\mathbb{R}. \tag{2.30}$$

Note that, owing to the boundedness of b in assumption [**Hyp-**$b.0$], the mean $\mathbf{m}^{\bar{x}}(s,t)$ in (2.30) is well defined. Furthermore, owing to the continuity of $b(\bar{x},\cdot)$, again by [**Hyp-**$b.0$], the function $t\to\tilde{P}_{0,t}[b(\bar{x},\cdot)]$ is continuous and bounded and thus $p^{\bar{x}}_0$ corresponds to the fundamental solution of the parabolic linear operator $\big(-\partial_t + \mathcal{A}^{\bar{x}}_0\big)$, where

$$\mathcal{A}^{\bar{x}}_0 u(t,x) = \frac{\sigma^2}{2}\partial_{xx}u(t,x) - \tilde{P}_{0,t}[b(\bar{x},\cdot)]\partial_x u(t,x),$$

i.e. for any $(s,\xi)\in[0,\infty[\times\mathbb{R}$ the function $p^{\bar{x}}_0(s,\xi;\cdot,\cdot)$ solves the forward Cauchy problem

$$\begin{cases} (-\partial_t + \mathcal{A}^{\bar{x}}_0)u = 0 \quad \text{on }]s,\infty[\times\mathbb{R}, \\ u(s,\cdot) = \delta_\xi. \end{cases} \tag{2.31}$$

2.2.3 Higher orders approximations

In order to achieve higher orders expansions for the transition kernel $p(s,\xi;t,x) = p^{\varepsilon,\bar{x}}(s,\xi;t,x)|_{\varepsilon=1}$, we follow the strategy explained previously. Hereafter throughout this subsection we fix $N\in\mathbb{N}^*$ and we assume assumptions [**Hyp-**$b.N$] and [**Hyp-**σ] to be in force.

Formal derivation. In analogy to what was done in Section 2.1.3, we start by freely assuming that all quantities are smooth in ε and that all subsequent PDEs are well posed. This will allow us to formally represent the terms $p_n^{\bar{x}}(s,\xi;t,x)$ appearing in (2.26) as the solutions of some nested PDEs.

By formally differentiating both the left and the right-hand sides of (2.25) we obtain

$$\begin{aligned}\big(-\partial_t+\mathcal{A}_0\big)\partial_\varepsilon^n p^{\varepsilon,\bar{x}}(s,\xi;t,x) &= \partial_x\partial_\varepsilon^n\Big(p^{\varepsilon,\bar{x}}(s;\xi;t,x)\tilde{P}_t^\varepsilon\big[b\big(\bar{x}+\varepsilon(x-\bar{x}),\cdot\big)\big]\Big)\\ &= \partial_x\sum_{h=0}^n\sum_{i=0}^h\binom{n}{h}\binom{h}{i}\big(\partial_\varepsilon^{n-h}p^{\varepsilon,\bar{x}}(s;\xi;t,x)\big)\\ &\qquad\cdot(x-\bar{x})^i\big(\partial_\varepsilon^{h-i}\tilde{P}_t^\varepsilon\big)\big[\partial_1^i b\big(\bar{x}+\varepsilon(x-\bar{x}),\cdot\big)\big],\end{aligned}$$

along with the terminal condition $\partial_\varepsilon^n p^{\varepsilon,\bar{x}}(s,\xi;s,x)=0$. Once again, setting $\varepsilon=0$ yields

$$\begin{cases}\big(-\partial_t+\mathcal{A}_0^{\bar{x}}\big)p_n^{\bar{x}}(s,\xi;\cdot,\cdot)=\sum_{h=1}^n\sum_{i=0}^h\binom{n}{h}\binom{h}{i}\mathcal{B}_{h,i}^{\bar{x}}(\cdot,\cdot)\,p_{n-h}^{\bar{x}}(s;\xi;\cdot,\cdot), & \text{on }]s,\infty[\times\mathbb{R},\\ p_n^{\bar{x}}(s,\xi;s,\cdot)=0, & \text{on } \mathbb{R}.\end{cases}\tag{2.32}$$

with the operators $\mathcal{B}_{h,i}^{\bar{x}}$ acting as in (2.14).

Rigorous definition. Proceeding as we did in Section 2.1.3, we use the previous heuristic computations, in particular the Cauchy problems (2.32), in order to give rigorous definitions for the correcting terms $p_n^{\bar{x}}$. Precisely, by applying Duhamel's principle we can give integral definitions that are coherent with aforementioned Cauchy problems. A fortiori, in Section 4.2, it will be shown that $p_n^{\bar{x}}$ actually solves (2.32), with the terminal condition meant in the distributional sense.

Definition 2.11. For any $n\in\mathbb{N}^*$ with $n\le N$, and $\bar{x},\xi\in\mathbb{R}$, the functions $\big(p_n^{\bar{x}}(s,\xi;\cdot,\cdot)\big)_{n\le N}$ are recursively defined as

$$p_n^{\bar{x}}(s,\xi;t,x)=-\sum_{h=1}^n\sum_{i=0}^h\binom{n}{h}\binom{h}{i}\int_s^t\int_{\mathbb{R}}\Gamma_\sigma\big(t-r;x-y-\mathbf{m}^{\bar{x}}(r,t)\big)\mathcal{B}_{h,i}^{\bar{x}}(r,y)\,p_{n-h}^{\bar{x}}(s;\xi;r,y)\mathrm{d}y\mathrm{d}r,\tag{2.33}$$

for any $0\le s<t$ and $x\in\mathbb{R}$.

As it is stated in Theorem 2.12 below, Definition 2.11 is well posed under assumptions [**Hyp-**$b.N$] and [**Hyp-**σ], and the function $p_n^{\bar{x}}$ can be given a representation in terms of differential operators acting on the leading term $p_0^{\bar{x}}$. Such representation can be manipulated, see Corollary 2.14 in order to achieve a fully explicit characterization.

Theorem 2.12. *Under assumptions* [**Hyp-**$b.N$] *and* [**Hyp-**σ]*, for any $n\in\mathbb{N}^*$ with $n\le N$ and $\bar{x}\in\mathbb{R}$, we have:*

(i) the functions $p_n^{\bar{x}}$ as in (2.33) *are well defined;*

(ii) it holds:

$$p_n^{\bar{x}}(s,\xi;t,x)=-\mathcal{L}_n^{\bar{x}}(s,t,x)\,p_0^{\bar{x}}(s,\xi;t,x),\qquad 0\le s<t,\quad x,\xi\in\mathbb{R},\tag{2.34}$$

where $\mathcal{L}_n^{\bar{x}}=\mathcal{L}_n^{\bar{x}}(\cdot,\cdot,\cdot)$ is the differential operator defined as

$$\mathcal{L}_n^{\bar{x}}(s,t,x):=n!\sum_{h=1}^n\int_s^t\mathrm{d}s_1\int_s^{s_1}\mathrm{d}s_2\cdots\int_s^{s_{h-1}}\mathrm{d}s_h\sum_{i\in I_{n,h}}\mathcal{G}_{i_1}^{\bar{x}}(s_1,t,x)\cdots\mathcal{G}_{i_h}^{\bar{x}}(s_h,t,x),$$

where the set $I_{n,h}$ is defined in (2.18)*, and the operator $\mathcal{G}_k^{\bar{x}} = \mathcal{G}_k^{\bar{x}}(\cdot,\cdot,\cdot)$ is defined as*

$$\mathcal{G}_k^{\bar{x}}(s,t,x) := \frac{1}{k!}\sum_{j=0}^{k}\binom{k}{j}\mathcal{B}_{k,j}^{\bar{x}}\big(s,t,\mathcal{M}^{\bar{x}}(s,t,x)\big), \quad \mathcal{M}^{\bar{x}}(s,t,x) := x - \mathbf{m}^{\bar{x}}(s,t) + \sigma^2(t-s)\,\partial_x, \tag{2.35}$$

with $\mathbf{m}^{\bar{x}}(s,t)$ as in (2.30)*. In* (2.35)*, $\mathcal{B}_{k,j}^{\bar{x}}\big(s,t,\mathcal{M}^{\bar{x}}(s,t,x)\big)$ is a slight abuse of notation for the operator $\mathcal{B}_{k,j}^{\bar{x}}$ acting as in* (2.14) *composed with $\mathcal{M}^{\bar{x}}$, i.e. $\mathcal{B}_{k,j}^{\bar{x}}\big(s,t,\mathcal{M}^{\bar{x}}(s,t,x)\big)u = \tilde{P}_{k-j,s}\big[\partial_1^j b(\bar{x},\cdot)\big]\partial_x\big((x - \mathbf{m}^{\bar{x}}(s,t) + \sigma^2(t-s)\,\partial_x - \bar{x})^j u\big)$.*

In analogy to Corollary 2.5, we are now going to re-write the operator $\mathcal{L}_n^{\bar{x}}(s,t,x)$ in a more explicit way, which is useful in order to implement the expansion. We generalize Notation 2.3.

Notation 2.13. For any $\gamma,l,j,i,m \in \mathbb{N}^h$ and $\bar{x} \in \mathbb{R}$, we let the function $\mathbf{F}_{\gamma,l,j,i}^{\bar{x}} = \mathbf{F}_{\gamma,l,j,i}^{\bar{x}}(\cdot,\cdot)$ be defined as

$$\mathbf{F}_{\gamma,l,j,i}^{\bar{x}}(s,t) := \int_s^t \mathrm{d}s_1 \cdots \int_s^{s_{h-1}} \mathrm{d}s_h \prod_{q=1,\cdots,h} \tilde{P}_{i_q-j_q,s_q}\big[\partial_1^{j_q} b(\bar{x},\cdot)\big]\,\big(-\mathbf{m}^{\bar{x}}(s_q,t)\big)^{\gamma_q}\,\big(\sigma^2(t-s_q)\big)^{l_q}, \qquad 0 \le s < t,$$

with $\mathbf{m}^{\bar{x}}(\cdot,\cdot)$ as in (2.30), and the constant $c_{\gamma,m,l,j,i}$ be defined as

$$c_{\gamma,m,l,j,i} := \prod_{q=1,\cdots,h}(2m_q-1)!!\binom{i_q}{j_q}\binom{j_q}{2m_q}\binom{j_q-2m_q}{l_q}\binom{j_q-2m_q-l_q}{\gamma_q}.$$

Corollary 2.14. *Under assumptions* [**Hyp-**$b.N$] *and* [**Hyp-**σ]*, for any $n \in \mathbb{N}^*$ with $n \le N$, the function $p_n^{\bar{x}}(s,\xi;t,x)$ is as in* (2.34)*, where*

$$\mathcal{L}_n^{\bar{x}}(s,t,x) := n!\sum_{h=1}^{n}\sum_{i\in I_{n,h}}\sum_{\substack{\gamma,m,l,j\in\mathbb{N}^h\\ l_q\le j_q\le i_q\\ m_q\le\lfloor\frac{j_q-l_q}{2}\rfloor\\ \gamma_q\le j_q-2m_q-l_q}}\Big(\prod_{q=1,\cdots,h}\frac{1}{i_q!}\Big)c_{\gamma,m,l,j,i}\,\mathbf{F}_{\gamma,l+m,j,i}^{\bar{x}}(s,t)\,\mathcal{H}_{j-2m-\gamma,l}^{\bar{x}}(x)$$

where $I_{n,h}$ and $\mathcal{H}_{j,l}^{\bar{x}}(x)$ are respectively defined in (2.18) *and* (2.21)*.*

Proof. Analogous to that of Corollary 2.5. □

Example 2.15. For $n = 1$ we have

$$\mathcal{L}_1^{\bar{x}}(s,t,x) := \mathbf{F}_{0,1,1,1}^{\bar{x}}(s,t)\,\mathcal{H}_{1,1}^{\bar{x}}(x) + \mathbf{F}_{0,0,1,1}^{\bar{x}}(s,t)\,\mathcal{H}_{1,0}^{\bar{x}}(x) + \big(\mathbf{F}_{0,0,0,1}^{\bar{x}}(s,t) + \mathbf{F}_{1,0,1,1}^{\bar{x}}(s,t)\big)\mathcal{H}_{0,0}^{\bar{x}}(x),$$

with

$$\mathbf{F}_{0,0,0,1}^{\bar{x}}(s,t) = \int_s^t \tilde{P}_{1,s_1}\big[b(\bar{x},\cdot)\big]\mathrm{d}s_1, \qquad \mathbf{F}_{0,1,1,1}^{\bar{x}}(s,t) = \int_s^t \tilde{P}_{0,s_1}\big[\partial_1 b(\bar{x},\cdot)\big]\sigma^2(t-s_1)\,\mathrm{d}s_1,$$
$$\mathbf{F}_{1,0,1,1}^{\bar{x}}(s,t) = -\int_s^t \tilde{P}_{0,s_1}\big[\partial_1 b(\bar{x},\cdot)\big]\mathbf{m}^{\bar{x}}(s_1,t)\,\mathrm{d}s_1, \quad \mathbf{F}_{0,0,1,1}^{\bar{x}}(s,t) = \int_s^t \tilde{P}_{0,s_1}\big[\partial_1 b(\bar{x},\cdot)\big]\mathrm{d}s_1,$$

and

$$\mathcal{H}_{0,0}^{\bar{x}}(x) = \partial_x, \qquad \mathcal{H}_{1,0}^{\bar{x}}(x) = \big(1 + (x-\bar{x})\partial_x\big), \qquad \mathcal{H}_{1,1}^{\bar{x}}(x) = \partial_x^2.$$

2.2.4 Error estimates

In this subsection we provide some rigorous error bounds for the N-th order approximation $\bar{p}_N^{\bar{x}}(s,\xi;t,x)$ with $\bar{x} = \lambda x + (1-\lambda)\xi$ and $\lambda \in [0,1]$, as defined in (2.26), of the transition density $p(s,\xi;t,x)$ of X; such bounds in turn imply analogous error bounds for the approximation $P_{N,t}^{\lambda}(x)$, as defined in (2.29), of the density $P_t(x)$ of X_t. We are in the position to prove the following result.

Theorem 2.16. *Let $\bar{N} \in \mathbb{N}^*$, $T > 0$, and assume assumptions* [**Hyp-σ**], [**Hyp-b.0**] *and* [**Hyp-b.$\bar{N}+1$**] *to be in force. Then, for any $N \in \mathbb{N}$ with $N \le \bar{N}$, $\lambda \in [0,1]$, and for any $c > 1$, we have*

$$\Big|\big(p - \bar{p}_N^{\bar{x}}\big)(s,\xi;t,x)\big|_{\bar{x}=\lambda x+(1-\lambda)\xi}\Big| \le C\big(\sigma^2 t\big)^{\frac{N+1}{2}} \big(\sigma^2(t-s)\big)^{\frac{1}{2}} \Gamma_{c\sigma}\big(t-s, x-\xi\big) \tag{2.36}$$

for any $\xi, x \in \mathbb{R}$ and $0 \le s < t \le T$, where $C > 0$ only depends on $\bar{N}, T, c$, and on the constants M and $\bar{\sigma}$ of assumptions [**Hyp-σ**], [**Hyp-b.$\bar{N}+1$**]. *In particular, we have*

$$\big\|\big(P_t - P_{N,t}^{\lambda}\big)\big\|_{L^1(\mathbb{R})} \le C\big(\sigma^2 t\big)^{\frac{N+2}{2}}, \qquad 0 < t \le T.$$

Note that the factor $\big(\sigma^2(t-s)\big)^{\frac{1}{2}}$ brings extra accuracy when the transition densities are computed on a small time interval.

3 Proof of Theorems 2.2 and 2.9

3.1 Proof of Theorem 2.2

Hereafter, throughout this subsection, we assume assumptions [**Hyp-b.$N-1$**] and [**Hyp-σ**] to be in force for a fixed $N \in \mathbb{N}^*$. We start by stating the following a priori estimates on the functions $\partial_x^k \tilde{p}_n(0,\xi;t,x)$ and $\tilde{P}_{n,t}[\partial_1^j b(x,\cdot)]$, which will be employed recursively in order to prove Theorem 2.2, in particular to prove wellposedness of Definitions 2.1.

Lemma 3.1. *Assume Corollary 2.5 to hold. Then, for any $c > 1$, and for any $n, k, j \in \mathbb{N}$ with $n \le N$ and $j \le N-1$, we have*

$$\big|\partial_x^k \tilde{p}_n(0,\xi;t,x)\big| \le C_c\big(\sigma^2 t\big)^{\frac{n-k}{2}} \Gamma_{c\sigma}\big(t, x-\xi\big), \tag{3.1}$$

$$\big|\tilde{P}_{n,t}[\partial_1^j b(x,\cdot)]\big| \le C_c\, \sigma^2 \big(\sigma^2 t\big)^{\frac{n}{2}}, \tag{3.2}$$

for any $0 < t \le T$ and $x, \xi, \in \mathbb{R}$, where $C_c > 0$ depends at most on N, k, T and on the constants M and $\bar{\sigma}$ of assumptions [**Hyp-σ**] *and* [**Hyp-b.$N-1$**]

Proof. We proceed by induction on n. First consider the case $n = 0$. Then, (3.1) follows by definition (2.9) and by Lemma A.2. Consequently, (3.2) stems from definition (2.8), since by assumption $|\partial_1^j b| \le M\sigma^2$.

Fix now $\bar{n} \in \mathbb{N}$ with $\bar{n} < N$, assume (3.1) and (3.2) to hold true for any $n \le \bar{n}$, and prove it true for $n = \bar{n}+1$. By Corollary 2.5 we have

$$\big|\partial_x^k \tilde{p}_{\bar{n}+1}(0,\xi;t,x)\big| \le C_c \sum_{h=1}^{n} \sum_{i \in I_{\bar{n}+1,h}} \sum_{\substack{m,l,j \in \mathbb{N}^h \\ l_q \le j_q \le i_q - 1 \\ m_q \le \lfloor \frac{j_q - l_q}{2} \rfloor}} \big|\mathbf{F}_{\mathbf{0},l+m,j,i-\mathbf{1}}^{\xi}(t)\big|\, \big|\partial_x^k \mathcal{H}_{j-2m,l}^{\xi}(x) \tilde{p}_0(0,\xi;t,x)\big|, \tag{3.3}$$

and by definition (2.20), we obtain

$$\big|\mathbf{F}^{\xi}_{\mathbf{0},l+m,j,i-\mathbf{1}}(t)\big| \le C_c \int_0^t \mathrm{d}s_1 \cdots \int_0^{s_{h-1}} \mathrm{d}s_h \prod_{q=1,\cdots,h} \big|\tilde{P}_{i_q-1-j_q,s_q}\big[\partial_1^{j_q} b(\xi,\cdot)\big]\big| \big(\sigma^2(t-s_q)\big)^{l_q+m_q}$$

(by inductive hypothesis, since $i_q - 1 - j_q \le \bar{n}$)

$$\le C_c \sigma^{2(\sum_{q=1}^h \frac{i_q+1-j_q}{2}+l_q+m_q)} \int_0^t \mathrm{d}s_1 \cdots \int_0^{s_{h-1}} \mathrm{d}s_h \prod_{q=1,\cdots,h} s_q^{(i_q-1-j_q)/2}(t-s_q)^{l_q+m_q}$$

(by solving the time integrals)

$$\le C_c(\sigma^2 t)^{\frac{\sum_{q=1}^h (i_q+1-j_q+2l_q+2m_q)}{2}}. \tag{3.4}$$

On the other hand, by employing the Gaussian estimates of Lemmas A.1 and A.2, along with definition (2.21), we obtain

$$\big|\partial_x^k \mathcal{H}^{\xi}_{j-2m,l}(x)\tilde{p}_0(0,\xi;t,x)\big| \le C_c(\sigma^2 t)^{\frac{-k+\sum_{q=1}^h (j_q-2m_q-2l_q-1)}{2}} \Gamma_{c\sigma}\big(t, x-\xi\big).$$

Therefore, by (3.3)-(3.4) and since $\sum_{q=1}^h i_q = \bar{n}+1$, we get (3.1) with $n = \bar{n}+1$. By integrating (3.1), and by using (2.15) and that $|\partial_1^j b| \le M\sigma^2$, we also obtain (3.2). □

The rest of the section is devoted to the proof of Theorem 2.2, which is based on the a priori estimates in Lemma 3.1 combined with following lemmas. Hereafter, throughout the rest of this section, we denote by $\mathcal{S}(\mathbb{R}^n)$ the Schwartz space of rapidly decreasing functions on $\mathbb{R}^n$.

Lemma 3.2. *For any $t_0 < s < t$, $x, y, \xi \in \mathbb{R}$, and $k \in \mathbb{N}^*$ with $k \le N$, we have*

$$\frac{1}{(k-1)!}\sum_{j=0}^{k-1}\binom{k-1}{j}\int_{\mathbb{R}} \Gamma_\sigma(t-s;x-y)\mathcal{B}^{\xi}_{k-1,j}(s,y)f(y)\mathrm{d}y = \tilde{\mathcal{G}}^{\xi}_k(s,t,x)\int_{\mathbb{R}} \Gamma_\sigma(t-s;x-\eta)f(\eta)\mathrm{d}\eta, \tag{3.5}$$

$$\frac{1}{(k-1)!}\sum_{j=0}^{k-1}\binom{k-1}{j}\int_{\mathbb{R}} f(x)\mathcal{B}^{\xi}_{k-1,j}(s,x)\Gamma_\sigma(s-t_0;x-y)\mathrm{d}x = \bar{\tilde{\mathcal{G}}}^{\xi}_k(t_0,s,y)\int_{\mathbb{R}} \Gamma_\sigma(s-t_0;\eta-y)f(\eta)\mathrm{d}\eta, \tag{3.6}$$

for any $f \in \mathcal{S}(\mathbb{R})$, where $\bar{\tilde{\mathcal{G}}}^{\xi}_k = \bar{\tilde{\mathcal{G}}}^{\xi}_k(t_0,s,y)$ is the differential operator acting as

$$\bar{\tilde{\mathcal{G}}}^{\xi}_k(t_0,s,y)u(y) = -\frac{1}{(k-1)!}\sum_{j=0}^{k-1}\binom{k-1}{j}\tilde{P}_{k-1-j,s}\big[\partial_1^j b(\xi,\cdot)\big]\big(\tilde{\mathcal{M}}(t_0,s,y)-\xi\big)^j \partial_y u(y), \tag{3.7}$$

and the operator $\tilde{\mathcal{M}} = \tilde{\mathcal{M}}(t_0,s,y)$ is as defined in (2.19)*. Moreover, the following relation holds:*

$$\tilde{\mathcal{G}}^{\xi}_k(s,t,x)\Gamma_\sigma(t-t_0;x-y) = \bar{\tilde{\mathcal{G}}}^{\xi}_k(t_0,s,y)\Gamma_\sigma(t-t_0;x-y). \tag{3.8}$$

Proof. We start by observing that

$$\partial_x \Gamma_\sigma(t-s;x-y) = -\partial_y \Gamma_\sigma(t-s;x-y), \tag{3.9}$$

and, for any $j \in \mathbb{N}^*$,

$$y^j\,\Gamma_\sigma(t-s;x-y) = \tilde{\mathcal{M}}^j(s,t,x)\Gamma_\sigma(t-s;x-y), \qquad x^j\,\Gamma_\sigma(t-s;x-y) = \tilde{\mathcal{M}}^j(s,t,y)\Gamma_\sigma(t-s;x-y). \tag{3.10}$$

Actually the above identity is clear for $j = 1$ thanks to (3.9) and the case $j > 1$ is obtained by simple iteration. Therefore, for any $\xi \in \mathbb{R}$ and $j \in \mathbb{N}^*$, by (3.10) we also have

$$(x-\xi)^j\Gamma_\sigma(t-s;x-y) = \big(\tilde{\mathcal{M}}(s,t,y)-\xi\big)^j\,\Gamma_\sigma(t-s;x-y), \tag{3.11}$$

$$(y-\xi)^j\Gamma_\sigma(t-s;x-y) = \big(\tilde{\mathcal{M}}(s,t,x)-\xi\big)^j\,\Gamma_\sigma(t-s;x-y), \tag{3.12}$$

for any $0 \le i \le k$. We now prove (3.5). By (2.14), for any $j \le k-1$ we get

$$\int_{\mathbb{R}} \Gamma_\sigma(t-s;x-y)\mathcal{B}^\xi_{k-1,j}(s,y)f(y)\mathrm{d}y = \tilde{P}_{k-1-j,s}\big[\partial_1^j b(\xi,\cdot)\big]\int_{\mathbb{R}} \Gamma_\sigma(t-s;x-y)\,\partial_y\big((y-\xi)^j f(y)\big)\mathrm{d}y$$

(integrating by parts, and applying (3.9))

$$= \tilde{P}_{k-1-j,s}\big[\partial_1^j b(\xi,\cdot)\big]\partial_x\int_{\mathbb{R}} \Gamma_\sigma(t-s;x-y)(y-\xi)^j f(y)\mathrm{d}y$$

(by (3.12))

$$= \tilde{P}_{k-1-j,s}\big[\partial_1^j b(\xi,\cdot)\big]\partial_x\Big(\big(\tilde{\mathcal{M}}(s,t,x)-\xi\big)^j\int_{\mathbb{R}} \Gamma_\sigma(t-s;x-y)f(y)\mathrm{d}y\Big)$$

$$= \mathcal{B}^\xi_{k-1,j}\big(s,\tilde{\mathcal{M}}(s,t,x)\big)\int_{\mathbb{R}} \Gamma_\sigma(t-s;x-y)f(y)\mathrm{d}y.$$

Now (3.5) stems from (2.19). We proceed analogously to prove (3.6). For any $j \le k-1$ we have

$$\int_{\mathbb{R}} f(x)\mathcal{B}^\xi_{k-1,j}(s,x)\Gamma_\sigma(s-t_0;x-y)\mathrm{d}x = \tilde{P}_{k-1-j,s}\big[\partial_1^j b(\xi,\cdot)\big]\int_{\mathbb{R}} f(x)\,\partial_x\big((x-\xi)^j\Gamma_\sigma(s-t_0;x-y)\big)\mathrm{d}x$$

(by (3.11))

$$= \tilde{P}_{k-1-j,s}\big[\partial_1^j b(\xi,\cdot)\big]\big(\tilde{\mathcal{M}}(t_0,s,y)-\xi\big)^j\int_{\mathbb{R}} f(x)\partial_x\Gamma_\sigma(s-t_0;x-y)\mathrm{d}x$$

$$= -\tilde{P}_{k-1-j,s}\big[\partial_1^j b(\xi,\cdot)\big]\big(\tilde{\mathcal{M}}(t_0,s,y)-\xi\big)^j\partial_y\int_{\mathbb{R}} f(x)\Gamma_\sigma(s-t_0;x-y)\mathrm{d}x,$$

where we used (3.9) in the last equality. Now (3.6) stems from (3.7). Eventually, identity (3.8) follows by combining (3.5) with $f(\eta) = \Gamma_\sigma(s-t_0;\eta-y)$ and (3.6) with $f(\eta) = \Gamma_\sigma(t-s;x-\eta)$, together with the Chapman-Kolmogorov identity

$$\int_{\mathbb{R}} \Gamma_\sigma(s-t_0,\eta-y)\Gamma_\sigma(t-s,x-\eta)\mathrm{d}\eta = \Gamma_\sigma(t-t_0,x-y), \qquad t_0 < s < t, \quad x,y \in \mathbb{R}.$$

□

In the next statement we denote by $\mathcal{C} : \mathcal{S}(\mathbb{R}^2)\times\mathcal{S}(\mathbb{R}^2) \to \mathcal{S}(\mathbb{R}^2)$ the operator acting as

$$\mathcal{C}(f,g)(\xi,x) = f(\cdot,x)[g(\xi,\cdot)], \qquad \xi,x \in \mathbb{R}.$$

Lemma 3.3. *Let $\big(\varphi_{s,t}\big)_{0\le s<t}$ be a family of functions in $\mathcal{S}(\mathbb{R}^2)$, such that*

$$\varphi_{t_0,t} = \mathcal{C}(\varphi_{s,t}, \varphi_{t_0,s}), \qquad 0 \le t_0 < s < t. \tag{3.13}$$

Let also $\big(\mathcal{P}_{h,s,t,z}\big)_{h\in\mathbb{N}^*,0\le s<t,z\in\mathbb{R}}$, $\big(\bar{\mathcal{P}}_{h,s,t,z}\big)_{h\in\mathbb{N}^*,0\le s<t,z\in\mathbb{R}}$ be two families of operators from $\mathcal{S}(\mathbb{R}^2)$ onto itself that can be represented as finite sums of the form

$$\begin{aligned}\mathcal{P}_{h,s,t,z}f(\xi,x) &= \sum_{i\ge 0,j\ge 1} a_{i,j}^{(h,z)}(s,t)x^i\partial_x^j f(\xi,x), \\ \bar{\mathcal{P}}_{h,s,t,z}f(\xi,x) &= \sum_{i\ge 0,j\ge 1} \bar{a}_{i,j}^{(h,z)}(s,t)\xi^i\partial_\xi^j f(\xi,x), \qquad \xi,x\in\mathbb{R},\end{aligned} \tag{3.14}$$

where $a_{i,j}^{(h,z)}(\cdot,t)$ and $\bar{a}_{i,j}^{(h,z)}(\cdot,t)$ are bounded measurable functions on $[0,t]$, and such that

$$\mathcal{P}_{h,s,t,z}\varphi_{t_0,t} = \bar{\mathcal{P}}_{h,t_0,s,z}\varphi_{t_0,t}, \qquad 0 \le t_0 < s < t. \tag{3.15}$$

Then, $\big(f_{n,t_0,t,z}\big)_{n\in\mathbb{N},0\le t_0<t,z\in\mathbb{R}}$ given by

$$f_{0,t_0,t,z} := \varphi_{t_0,t}, \tag{3.16}$$

$$f_{n,t_0,t,z} := n!\sum_{h=1}^{n}\int_{t_0}^{t}\mathrm{d}s_1\cdots\int_{t_0}^{s_{h-1}}\mathrm{d}s_h\sum_{i\in I_{n,h}}\mathcal{P}_{i_1,s_1,t,z}\cdots\mathcal{P}_{i_h,s_h,t,z}\,\varphi_{t_0,t}, \qquad n\in\mathbb{N}^*, \tag{3.17}$$

is well defined as a family of functions in $\mathcal{S}(\mathbb{R}^2)$. Here, the set $I_{n,h}$ is as defined in (2.18)*. Moreover, for any $n\in\mathbb{N}^*$, $z\in\mathbb{R}$ and $0\le t_0<t$ we have*

$$f_{n,t_0,t,z} = \sum_{h=1}^{n}\frac{n!}{(n-h)!}\int_{t_0}^{t}\mathcal{P}_{h,s,t,z}\,\mathcal{C}(\varphi_{s,t}, f_{n-h,t_0,s,z})\mathrm{d}s. \tag{3.18}$$

Proof. The first part of the statement easily follows from (3.14) (note that we mainly use that $a_{i,j}^{(h,z)}$ are bounded in time). Now note that for any $f,g\in\mathcal{S}(\mathbb{R}^2)$, we have the following commutation properties:

$$\mathcal{C}(f,\bar{\mathcal{P}}_{h,s,t,z}\,g) = \bar{\mathcal{P}}_{h,s,t,z}\mathcal{C}(f,g), \qquad \mathcal{P}_{h_1,s_1,t_1,z}\bar{\mathcal{P}}_{h_2,s_2,t_2,z}f = \bar{\mathcal{P}}_{h_2,s_2,t_2,z}\mathcal{P}_{h_1,s_1,t_1,z}f. \tag{3.19}$$

In order to prove (3.18) we first need to prove that, for any $h\in\mathbb{N}^*$, $0\le t_0<s<t$ and $z\in\mathbb{R}$, we have

$$\mathcal{P}_{i_1,s_1,t,z}\cdots\mathcal{P}_{i_h,s_h,t,z}\,\varphi_{t_0,t} = \mathcal{C}\big(\varphi_{s,t},\mathcal{P}_{i_1,s_1,s,z}\cdots\mathcal{P}_{i_h,s_h,s,z}\,\varphi_{t_0,s}\big), \tag{3.20}$$

for any $i\in(\mathbb{N}^*)^h$ and $t_0<s_1<\cdots<s_h<s$. We proceed by induction on h. If $h=1$, by using (3.15), (3.13), (3.19) and then again (3.15), exactly in this order, one has

$$\mathcal{P}_{i_1,s_1,t,z}\varphi_{t_0,t} = \bar{\mathcal{P}}_{i_1,t_0,s_1,z}\varphi_{t_0,t} = \bar{\mathcal{P}}_{i_1,t_0,s_1,z}\mathcal{C}(\varphi_{s,t},\varphi_{t_0,s}) = \mathcal{C}(\varphi_{s,t},\bar{\mathcal{P}}_{i_1,t_0,s_1,z}\varphi_{t_0,s}) = \mathcal{C}(\varphi_{s,t},\mathcal{P}_{i_1,s_1,s,z}\varphi_{t_0,s}).$$

We now assume (3.20) to hold for $h\in\mathbb{N}^*$ and we prove it true for $h+1$. We get

$$\begin{aligned}&\mathcal{C}\big(\varphi_{s,t},\mathcal{P}_{i_1,s_1,s,z}\cdots\mathcal{P}_{i_h,s_h,s,z}\mathcal{P}_{i_{h+1},s_{h+1},s,z}\,\varphi_{t_0,s}\big)\\ &= \mathcal{C}\big(\varphi_{s,t},\mathcal{P}_{i_1,s_1,s,z}\cdots\mathcal{P}_{i_h,s_h,s,z}\bar{\mathcal{P}}_{i_{h+1},t_0,s_{h+1},z}\,\varphi_{t_0,s}\big) \qquad \text{(by (3.15))}\\ &= \bar{\mathcal{P}}_{i_{h+1},t_0,s_{h+1},z}\mathcal{C}\big(\varphi_{s,t},\mathcal{P}_{i_1,s_1,s,z}\cdots\mathcal{P}_{i_h,s_h,s,z}\,\varphi_{t_0,s}\big) \qquad \text{(by (3.19))}\end{aligned}$$

$$
\begin{aligned}
&= \bar{\mathcal{P}}_{i_{h+1},t_0,s_{h+1},z}\mathcal{P}_{i_1,s_1,t,z}\cdots\mathcal{P}_{i_h,s_h,t,z}\,\varphi_{t_0,t} && \text{(by inductive hypothesis)}\\
&= \mathcal{P}_{i_1,s_1,t,z}\cdots\mathcal{P}_{i_h,s_h,t,z}\bar{\mathcal{P}}_{i_{h+1},t_0,s_{h+1},z}\,\varphi_{t_0,t} && \text{(by (3.19))}\\
&= \mathcal{P}_{i_1,s_1,t,z}\cdots\mathcal{P}_{i_h,s_h,t,z}\mathcal{P}_{i_{h+1},s_{h+1},t,z}\,\varphi_{t_0,t} && \text{(by (3.15))},
\end{aligned}
$$

which is (3.20) for $h+1$.

We are now ready to conclude the proof of (3.18). For $n=1$, (3.18) directly stems from (3.13) For $n\geq 2$, by definition (3.17) we have

$$
\frac{f_{n,t_0,t,z}}{n!} = \sum_{l=1}^{n}\int_{t_0}^{t}\mathrm{d}r_1\int_{t_0}^{r_1}\mathrm{d}r_2\cdots\int_{t_0}^{r_{l-1}}\mathrm{d}r_l\sum_{i\in I_{n,l}}\mathcal{P}_{i_1,r_1,t,z}\mathcal{P}_{i_2,r_2,t,z}\cdots\mathcal{P}_{i_l,r_l,t,z}\,\varphi_{t_0,t}
$$

(by (2.18))

$$
\begin{aligned}
&= \int_{t_0}^{t}\mathrm{d}r_1\mathcal{P}_{n,r_1,t,z}\varphi_{t_0,t}\\
&\quad+\sum_{l=2}^{n}\sum_{h=1}^{n+1-l}\int_{t_0}^{t}\mathrm{d}r_1\int_{t_0}^{r_1}\mathrm{d}r_2\cdots\int_{t_0}^{r_{l-1}}\mathrm{d}r_l\sum_{i\in I_{n-h,l-1}}\mathcal{P}_{h,r_1,t,z}\mathcal{P}_{i_1,r_2,t,z}\cdots\mathcal{P}_{i_{l-1},r_l,t,z}\,\varphi_{t_0,t}
\end{aligned}
$$

(by replacing the integration variables: $(\mathrm{d}r_1,\cdots,\mathrm{d}r_l)\to(\mathrm{d}s,\cdots,\mathrm{d}s_{l-1})$)

$$
\begin{aligned}
&= \int_{t_0}^{t}\mathrm{d}s\mathcal{P}_{n,s,t,z}\varphi_{t_0,t}\\
&\quad+\sum_{l=2}^{n}\sum_{h=1}^{n+1-l}\int_{t_0}^{t}\mathrm{d}s\int_{t_0}^{s}\mathrm{d}s_1\cdots\int_{t_0}^{s_{l-2}}\mathrm{d}s_{l-1}\sum_{i\in I_{n-h,l-1}}\mathcal{P}_{h,s,t,z}\mathcal{P}_{i_1,s_1,t,z}\cdots\mathcal{P}_{i_{l-1},s_{l-1},t,z}\,\varphi_{t_0,t}
\end{aligned}
$$

(by setting $j=l-1$)

$$
\begin{aligned}
&= \int_{t_0}^{t}\mathrm{d}s\mathcal{P}_{n,s,t,z}\varphi_{t_0,t}\\
&\quad+\sum_{j=1}^{n-1}\sum_{h=1}^{n-j}\int_{t_0}^{t}\mathrm{d}s\int_{t_0}^{s}\mathrm{d}s_1\cdots\int_{t_0}^{s_{j-1}}\mathrm{d}s_j\sum_{i\in I_{n-h,j}}\mathcal{P}_{h,s,t,z}\mathcal{P}_{i_1,s_1,t,z}\cdots\mathcal{P}_{i_j,s_j,t,z}\,\varphi_{t_0,t}
\end{aligned}
$$

(by exchanging summation and integration, using again that $a_{i,j}^{(h,z)}$ are bounded)

$$
\begin{aligned}
&= \int_{t_0}^{t}\mathrm{d}s\mathcal{P}_{n,s,t,z}\varphi_{t_0,t}\\
&\quad+\sum_{h=1}^{n-1}\int_{t_0}^{t}\mathrm{d}s\,\mathcal{P}_{h,s,t,z}\underbrace{\sum_{j=1}^{n-h}\int_{t_0}^{s}\mathrm{d}s_1\cdots\int_{t_0}^{s_{j-1}}\mathrm{d}s_j\sum_{i\in I_{n-h,j}}\mathcal{P}_{i_1,s_1,t,z}\cdots\mathcal{P}_{i_j,s_j,t,z}\,\varphi_{t_0,t}}_{:=\Psi_{h,s,t_0,t,z}}.
\end{aligned}
$$

Observe that, under the assumptions on the operator $\mathcal{P}_{h,s,t,z}$, $\Psi_{h,s,t_0,t,z}$ is a function in $\mathcal{S}(\mathbb{R}^2)$ uniformly in s, in the sense that derivatives of a given order are rapidly decreasing uniformly in s. The resulting function $\mathcal{P}_{h,s,t,z}\Psi_{h,s,t_0,t,z}$ is also in $\mathcal{S}(\mathbb{R}^2)$, uniformly in s. We now give another representation of this function: by (3.20), we have for a given $s\in(t_0,t)$

$$
\Psi_{h,s,t_0,t,z} = \sum_{j=1}^{n-h}\int_{t_0}^{s}\mathrm{d}s_1\cdots\int_{t_0}^{s_{j-1}}\mathrm{d}s_j\sum_{i\in I_{n-h,j}}\mathcal{C}\big(\varphi_{s,t},\mathcal{P}_{i_1,s_1,s,z}\cdots\mathcal{P}_{i_j,s_j,s,z}\,\varphi_{t_0,s}\big)
$$

(by Fubini's Theorem)

$$= \mathcal{C}\Big(\varphi_{s,t}, \sum_{j=1}^{n-h} \int_{t_0}^{s} \mathrm{d}s_1 \cdots \int_{t_0}^{s_{j-1}} \mathrm{d}s_j \sum_{i \in I_{n-h,j}} \mathcal{P}_{i_1,s_1,s,z} \cdots \mathcal{P}_{i_j,s_j,s,z}\, \varphi_{t_0,s}\Big)$$

(finally, by (3.16)-(3.17))

$$= \frac{1}{(n-h)!} \mathcal{C}\Big(\varphi_{s,t}, f_{n-h,t_0,s,z}\Big).$$

Therefore, we obtain

$$\frac{f_{n,t_0,t,z}}{n!} = \int_{t_0}^{t} \mathrm{d}s \mathcal{P}_{n,s,t,z} \varphi_{t_0,t} + \frac{1}{(n-h)!} \sum_{h=1}^{n-1} \int_{t_0}^{t} \mathrm{d}s\, \mathcal{P}_{h,s,t,z} \mathcal{C}\big(\varphi_{s,t}, f_{n-h,t_0,s,z}\big),$$

which proves (3.18) and concludes the proof. □

We are now ready to prove Theorems 2.2, which will be proved by induction on n.

Proof of Theorem 2.2. The result will follow by applying Lemma 3.3 with:

$$\varphi_{s,t}(\xi, x) = \Gamma_\sigma(t-s, x-\xi), \qquad \xi, x \in \mathbb{R}.$$

We first prove the statement for $n = 1$. Set

$$\mathcal{P}_{1,s,t,z} = \tilde{\mathcal{G}}_1^z(s,t,x) = \tilde{P}_{0,s}[b(z,\cdot)]\partial_x, \qquad \bar{\mathcal{P}}_{1,s,t,z} = \bar{\tilde{\mathcal{G}}}_1^z(s,t,\xi) = -\tilde{P}_{0,s}[b(z,\cdot)]\partial_\xi,$$

which are operators from $\mathcal{S}(\mathbb{R}^2)$ onto itself of the form (3.14). Moreover, by definition (2.8)-(2.9), assumption [**Hyp-***b*.$N-1$], and by estimate (3.2) with $n = 0$, the function $s \mapsto \tilde{P}_{0,s}[b(z,\cdot)]$ are measurable and bounded on $[0,t]$. Finally, (3.8) implies (3.15) for $h = 1$. Therefore, by Lemma 3.3 we have that $(f_{1,t_0,t,z})_{0 \le t_0 < t, z \in \mathbb{R}}$ as in (3.17) is well defined as a family of functions in $\mathcal{S}(\mathbb{R}^2)$, and

$$f_{1,t_0,t,z}(\xi, x) = \int_{t_0}^{t} \tilde{\mathcal{G}}_1^z(s,t,x) \int_{\mathbb{R}} \Gamma_\sigma(t-s, x-y) \varphi_{t_0,s}(\xi, y) \mathrm{d}y \mathrm{d}s$$

(by (3.5))

$$= \int_{0}^{t} \int_{\mathbb{R}} \Gamma_\sigma(t-s, x-y) \mathcal{B}_{0,0}^z(s,y) \varphi_{t_0,s}(\xi, y) \mathrm{d}y \mathrm{d}s.$$

This proves that the function $\tilde{p}_1(0,\xi;t,x) := -f_{1,0,t,\xi}(\xi, x)$ is well defined, satisfies (2.16), and also

$$\tilde{p}_1(0,\xi;t,x) = -\tilde{\mathcal{L}}_1^\xi(t,x)\, \tilde{p}_0(0,\xi;t,x).$$

Eventually, the wellposedness of definition (2.15) for $\tilde{P}_{1,t}$ stems directly from estimate (3.1) with $n = 1$, which is integrable w.r.t. $\mu_Y(d\xi)$.

Set now $\bar{n} \in \mathbb{N}^*$, $\bar{n} < N$, assume the statement true for any $n \le \bar{n}$, and prove it true for $n = \bar{n} + 1$. Set

$$\mathcal{P}_{h,s,t,z} = \tilde{\mathcal{G}}_h^z(s,t,x), \qquad \bar{\mathcal{P}}_{h,s,t,z} = \bar{\tilde{\mathcal{G}}}_h^z(s,t,\xi), \qquad h \le \bar{n}+1,$$

with $\tilde{\mathcal{G}}_h^z$ and $\bar{\tilde{\mathcal{G}}}_h^z$ as in (2.19) and (3.7), respectively. In particular, by inductive hypothesis, the functions $\tilde{P}_{n,t}$, $n \le \bar{n}$, are well defined and satisfy the a priori estimates (3.2). Therefore, it follows by (2.19) and (3.7)

that $\mathcal{P}_{h,s,t,z}$ and $\bar{\mathcal{P}}_{h,s,t,z}$ are well defined as operators from $\mathcal{S}(\mathbb{R}^2)$ onto itself, and they admit a representation of the form (3.14) where $s \mapsto a_{i,j}^{(h,z)}(s,t)$ and $s \mapsto \bar{a}_{i,j}^{(h,z)}(s,t)$ are bounded and measurable (by assumption [**Hyp-**$b.N-1$] and estimate (3.2)) functions on $[0,t]$. Moreover, (3.8) implies the symmetry property (3.15). Therefore, by Lemma 3.3 we have that $(f_{n,t_0,t,z})_{n\le\bar{n}+1,0\le t_0<t,z\in\mathbb{R}}$ as in (3.17) is well defined as a family of functions in $\mathcal{S}(\mathbb{R}^2)$, and that in particular

$$f_{\bar{n}+1,t_0,t,z}(\xi,x) = \sum_{h=1}^{\bar{n}+1} \frac{(\bar{n}+1)!}{(\bar{n}+1-h)!} \int_{t_0}^{t} \tilde{\mathcal{G}}_h^z(s,t,x) \int_{\mathbb{R}} \Gamma_\sigma(t-s,x-y) f_{\bar{n}+1-h,t_0,s,z}(\xi,y)\mathrm{d}y\mathrm{d}s$$

(by (3.5))

$$= \sum_{h=1}^{\bar{n}+1} h\binom{\bar{n}+1}{h} \sum_{j=0}^{h-1} \binom{h-1}{j} \int_0^t \int_{\mathbb{R}} \Gamma_\sigma(t-s,x-y) \mathcal{B}_{h-1,j}^z(s,y) f_{\bar{n}+1-h,t_0,s,z}(\xi,y)\mathrm{d}y\mathrm{d}s.$$

This proves that the function $\tilde{p}_{\bar{n}+1}(0,\xi;t,x) := -f_{\bar{n}+1,0,t,\xi}(\xi,x)$ is well defined, satisfies (2.16), and also

$$\tilde{p}_{\bar{n}+1}(0,\xi;t,x) = -\tilde{\mathcal{L}}_{\bar{n}+1}^{\xi}(t,x)\, \tilde{p}_0(0,\xi;t,x).$$

Eventually, the wellposedness of definition (2.15) for $\tilde{P}_{\bar{n}+1,t}$ stems directly from estimate (3.1) with $n = \bar{n}+1$, as before. This concludes the proof. □

3.2 Proof of Theorem 2.9

Throughout this section we fix a time $T > 0$ and we consider assumptions [**Hyp-**σ], [**Hyp-**$b.0$] and [**Hyp-**$b.\bar{N}+1$] to be in force for a fixed $\bar{N} \in \mathbb{N}^*$.

Notation 3.4. Throughout the rest of this section, unless explicitly stated, we will denote by $(C_c)_{c>0}$ any family of positive constants that depend at most on $\bar{N}, T$ and on the constants $M, \bar{\sigma}$ of assumptions [**Hyp-**σ], [**Hyp-**$b.0$] and [**Hyp-**$b.\bar{N}+1$]. Note that, in particular, C_c is independent of σ.

In light of Remark 1.4, we have the following classical Gaussian upper bounds for $p(s,\xi;t,x)$ and $\partial_\xi p(s,\xi;t,x)$ (see [Fri64, Chapter 1, p. 28]) that will be used here below.

Lemma 3.5. *For any $m,n \in \mathbb{N}$ with $m \le \bar{N}+1$ and $n \le 1$ and for any $c > 1$, we have*

$$|x-\xi|^m \left|\partial_\xi^n p(s,\xi;t,x)\right| \le C_c\big(\sigma^2(t-s)\big)^{\frac{m-n}{2}} \Gamma_{c\sigma}\big(t-s,x-\xi\big),$$

for any $\xi, x \in \mathbb{R}$ and $0 \le s < t \le T$.

Remark 3.6. Under assumptions [**Hyp-**$b.\bar{N}+1$] and [**Hyp-**σ], the functions $\tilde{p}_n(0,\xi;\cdot,\cdot)$, $1 \le n \le \bar{N}+1$, are solutions of the Cauchy problems (2.13) , meaning that the PDE is satisfied on the internal domain $]0,\infty[\times\mathbb{R}$, and that the initial condition is fulfilled in the distributional sense, i.e.

$$\lim_{\substack{(t,\xi)\to(0,\bar{\xi})\\ t>0}} \int_{\mathbb{R}} \tilde{p}_n(0,\xi;t,x)\varphi(t,x)\mathrm{d}x = 0, \qquad \bar{\xi} \in \mathbb{R}, \qquad \varphi \in C_b([0,\bar{T}[\times\mathbb{R}), \quad \bar{T} \in]0,\infty[. \tag{3.21}$$

In fact, the initial condition above is a straightforward consequence of the upper bound (3.1), whereas the fact that $\tilde{p}_n(0,\xi;\cdot,\cdot)$ solves the PDE in (2.13) follows by differentiating (2.16), and by using again estimate (3.1) and integrating by parts to deal with the time-integral in (2.16), which is singular near 0 and t.

Before to continue, let the family of operators $\big(\bar{\tilde{\mathcal{A}}}^{\xi}_{n,s}\big)_{n\leq \bar{N}+1, 0\leq s\leq T, \xi\in\mathbb{R}}$ be defined as

$$\bar{\tilde{\mathcal{A}}}^{\xi}_{n,s} := \tilde{\mathcal{A}}_0 + \sum_{h=1}^{n} \tilde{\mathcal{A}}^{\xi}_{h,s}, \qquad \text{with} \qquad \tilde{\mathcal{A}}^{\xi}_{h,s} := -\frac{1}{(h-1)!}\sum_{i=0}^{h-1}\binom{h-1}{i}\mathcal{B}^{\xi}_{h-1,i}(s,\cdot), \tag{3.22}$$

where the family of operators $\big(\mathcal{B}^{\xi}_{k,i}\big)_{0\leq i\leq k,\xi\in\mathbb{R}}$ is as defined in (2.14). We also recall to the reader the definitions of $\bar{\tilde{p}}_N$, $\bar{\tilde{P}}_N$ given in (2.5)-(2.6).

Lemma 3.7. *For any $N\in\mathbb{N}$ with $N\leq \bar{N}+1$, the following identity holds:*

$$\big(p-\bar{\tilde{p}}_N\big)(0,\xi;t,x) = \sum_{n=0}^{N}\frac{1}{(N-n)!}\int_0^t\int_{\mathbb{R}} p(s,y;t,x)\big(\mathcal{A}_s(y)-\bar{\tilde{\mathcal{A}}}^{\xi}_{n,s}(y)\big)\tilde{p}_{N-n}(0,\xi;s,y)\mathrm{d}y\mathrm{d}s, \tag{3.23}$$

for any $t>0$ and $x,\xi\in\mathbb{R}$.

Proof. By induction on N. We first prove the statement for $N=0$. We set

$$\psi_0(s) := -\int_{\mathbb{R}} p(s,y;t,x)\tilde{p}_0(0,\xi;s,y)\mathrm{d}y, \qquad 0<s<t, \quad \xi,x\in\mathbb{R}.$$

By the continuity of $p(\cdot,\cdot;t,x)$ and $\tilde{p}_0(0,\xi;\cdot,\cdot)$ along with the terminal condition in (1.3) and the initial condition in (2.11), one readily has

$$\lim_{s\to 0^+}\psi_0(s) = -p(0,\xi;t,x), \qquad \lim_{s\to t^-}\psi_0(s) = -\tilde{p}_0(0,\xi;t,x).$$

On the other hand, for any $s\in]0,t[$ we obtain

$$\partial_s\psi_0(s) = -\int_{\mathbb{R}}\Big(\big(\partial_s p(s,y;t,x)\big)\tilde{p}_0(0,\xi;s,y) + p(s,y;t,x)\big(\partial_s\tilde{p}_0(0,\xi;s,y)\big)\Big)\mathrm{d}y$$

(by the PDEs in (1.3) and (2.11))

$$= \int_{\mathbb{R}}\Big(\big(\mathcal{A}^b_s(y)p(s,y;t,x)\big)\tilde{p}_0(0,\xi;s,y) - p(s,y;t,x)\big(\tilde{\mathcal{A}}_0(y)\tilde{p}_0(0,\xi;s,y)\big)\Big)\mathrm{d}y$$

(as $\mathcal{A}_s$ is the adjoint of $\mathcal{A}^b_s$)

$$= \int_{\mathbb{R}} p(s,y;t,x)\big(\mathcal{A}_s(y)-\tilde{\mathcal{A}}_0(y)\big)\tilde{p}_0(0,\xi;s,y)\mathrm{d}y.$$

Now, by employing the estimates in Lemmas 3.1 and 3.5, it is straightforward to see that $\partial_s\psi_0$ is integrable on the interval $[0,t]$. Therefore, by Newton-Leibniz axiom along with definition (2.5) we get

$$(p-\bar{\tilde{p}}_0)(0,\xi;t,x) = \int_0^t\int_{\mathbb{R}} p(s,y;t,x)\big(\mathcal{A}_s(y)-\tilde{\mathcal{A}}_0(y)\big)\tilde{p}_0(0,\xi;s,y)\mathrm{d}y\mathrm{d}s,$$

which, by (3.22), yields (3.23) for $N=0$.

Fix now $\widetilde{N}\leq\bar{N}$, assume that (3.23) holds true for $N=\widetilde{N}$, and prove it true for $N=\widetilde{N}+1$. We set

$$\psi_{\widetilde{N}+1}(s) := \int_{\mathbb{R}} p(s,y;t,x)\tilde{p}_{\widetilde{N}+1}(0,\xi;s,y)\mathrm{d}y, \qquad 0<s<t, \quad \xi,x\in\mathbb{R}.$$

By the continuity of $p(\cdot,\cdot;t,x)$ and $\tilde{p}_{\widetilde{N}+1}(0,\xi;\cdot,\cdot)$ along with the terminal condition in (1.3) and the initial condition (3.21), one readily has

$$\lim_{s\to 0^+}\psi_{\widetilde{N}+1}(s) = 0, \qquad \lim_{s\to t^-}\psi_{\widetilde{N}+1}(s) = \tilde{p}_{\widetilde{N}+1}(0,\xi;t,x).$$

On the other hand, for any $s \in]0,t[$ we get

$$\partial_s \psi_{\widetilde{N}+1}(s) = \int_{\mathbb{R}} \Big(\big(\partial_s p(s,y;t,x)\big)\tilde{p}_{\widetilde{N}+1}(0,\xi;s,y) + p(s,y;t,x)\big(\partial_s \tilde{p}_{\widetilde{N}+1}(0,\xi;s,y)\big)\Big)\mathrm{d}y$$

(by the PDE in (1.3) and since $\mathcal{A}_s^b$ is the adjoint of $\mathcal{A}_s$)

$$= \int_{\mathbb{R}} p(s,y;t,x)\big(\partial_s - \tilde{\mathcal{A}}_0(y) + \tilde{\mathcal{A}}_0(y) - \mathcal{A}_s(y)\big)\tilde{p}_{\widetilde{N}+1}(0,\xi;s,y)\mathrm{d}y$$

(by the PDE in (2.13) (see Remark 3.6) and by definition (3.22))

$$\begin{aligned} &= \int_{\mathbb{R}} p(s,y;t,x)\big(\tilde{\mathcal{A}}_0(y) - \mathcal{A}_s(y)\big)\tilde{p}_{\widetilde{N}+1}(0,\xi;s,y)\mathrm{d}y \\ &\quad + \sum_{n=1}^{\widetilde{N}+1} \frac{(\widetilde{N}+1)!}{(\widetilde{N}+1-n)!} \int_{\mathbb{R}} p(s,y;t,x)\,\tilde{\mathcal{A}}^{\xi}_{n,s}(y)\,\tilde{p}_{\widetilde{N}+1-n}(0,\xi;s,y)\mathrm{d}y. \end{aligned}$$

Now, by employing again the estimates in Lemmas 3.1 and 3.5, it is straightforward to see that $\partial_s \psi_{\widetilde{N}+1}$ is integrable on the interval $[0,t]$. Therefore, it holds

$$\begin{aligned} \frac{1}{(\widetilde{N}+1)!}\tilde{p}_{\widetilde{N}+1}(0,\xi;t,x) &= \frac{1}{(\widetilde{N}+1)!}\int_0^t \int_{\mathbb{R}} p(s,y;t,x)\big(\tilde{\mathcal{A}}_0(y) - \mathcal{A}_s(y)\big)\tilde{p}_{\widetilde{N}+1}(0,\xi;s,y)\mathrm{d}y\mathrm{d}s \\ &\quad + \sum_{n=1}^{\widetilde{N}+1} \frac{1}{(\widetilde{N}+1-n)!}\int_0^t \int_{\mathbb{R}} p(s,y;t,x)\,\tilde{\mathcal{A}}^{\xi}_{n,s}(y)\,\tilde{p}_{\widetilde{N}+1-n}(0,\xi;s,y)\mathrm{d}y\mathrm{d}s. \end{aligned} \tag{3.24}$$

Now, by definition (2.5) we finally obtain

$$\big(p - \bar{\tilde{p}}_{\widetilde{N}+1}\big)(0,\xi;t,x) = \big(p - \bar{\tilde{p}}_{\widetilde{N}} - \frac{1}{(\widetilde{N}+1)!}\tilde{p}_{\widetilde{N}+1}\big)(0,\xi;t,x)$$

(by inductive hypothesis, and shifting the index n, and by (3.24))

$$\begin{aligned} &= \sum_{n=1}^{\widetilde{N}+1} \frac{1}{(\widetilde{N}+1-n)!}\int_0^t \int_{\mathbb{R}} p(s,y;t,x)\big(\mathcal{A}_s(y) - \bar{\tilde{\mathcal{A}}}^{\xi}_{n-1}(y)\big)\tilde{p}_{\widetilde{N}+1-n}(0,\xi;s,y)\mathrm{d}y\mathrm{d}s \\ &\quad + \frac{1}{(\widetilde{N}+1)!}\int_0^t \int_{\mathbb{R}} p(s,y;t,x)\big(\mathcal{A}_s(y) - \tilde{\mathcal{A}}_0(y)\big)\tilde{p}_{\widetilde{N}+1}(0,\xi;s,y)\mathrm{d}y\mathrm{d}s \\ &\quad - \sum_{n=1}^{\widetilde{N}+1} \frac{1}{(\widetilde{N}+1-n)!}\int_0^t \int_{\mathbb{R}} p(s,y;t,x)\,\tilde{\mathcal{A}}^{\xi}_{n,s}(y)\,\tilde{p}_{\widetilde{N}+1-n}(0,\xi;s,y)\mathrm{d}y\mathrm{d}s, \end{aligned}$$

which, by (3.22), yields (3.23) with $N = \widetilde{N}+1$ and concludes the proof. □

Lemma 3.8. *For any $n \in \mathbb{N}$ with $n \le \bar{N}+1$ we have*

$$\big(\mathcal{A}_s(y) - \bar{\tilde{\mathcal{A}}}^{\xi}_{n,s}(y)\big) = \partial_y\Big(P_s\big[T^{(\xi)}_{n-1}b(y,\cdot) - b(y,\cdot)\big] + \sum_{i=0}^{n-1}\frac{1}{i!}\Big(\bar{\tilde{P}}_{n-1-i,s}\big[\partial_1^i b(\xi,\cdot)\big] - P_s\big[\partial_1^i b(\xi,\cdot)\big]\Big)(y-\xi)^i\Big), \tag{3.25}$$

where $T^{(\xi)}_k b(y,\cdot)$ represents the k-th order Taylor expansion of $b(y,\cdot)$ centered at ξ, i.e.

$$T^{(\xi)}_{-1}b(y,\cdot) \equiv 0, \qquad T^{(\xi)}_k b(y,\cdot) = \sum_{i=0}^{k}\frac{\partial_1^i b(\xi,\cdot)}{i!}(y-\xi)^i, \qquad k \ge 0. \tag{3.26}$$

Proof. We prove the statement by induction on N. For $N = 0$, it directly stems from (3.22), and from the definition of $\mathcal{A}_s$ and $\tilde{\mathcal{A}}_0$ in (1.4) and (2.10), respectively.

Fix now $\tilde{n} \leq \bar{N}$, assume that (3.25) holds true for any $n = \tilde{n}$, and we prove it true for $n = \tilde{n} + 1$. By (3.22) and by induction hypothesis, we get

$$\begin{aligned}\big(\mathcal{A}_s(y) - \bar{\tilde{\mathcal{A}}}^{\xi}_{\tilde{n}+1,s}(y)\big) &= \partial_y\bigg(P_s\big[T^{(\xi)}_{\tilde{n}-1}b(y,\cdot) - b(y,\cdot)\big] + \sum_{i=0}^{\tilde{n}-1}\frac{1}{i!}\Big(\bar{\tilde{P}}_{\tilde{n}-1-i,s}[\partial_1^i b(\xi,\cdot)] - P_s[\partial_1^i b(\xi,\cdot)]\Big)(y-\xi)^i\bigg)\\ &\quad + \frac{1}{\tilde{n}!}\sum_{i=0}^{\tilde{n}}\binom{\tilde{n}}{i}\mathcal{B}^{\xi}_{\tilde{n},i}(s,y)\end{aligned}$$

(by (2.14))

$$\begin{aligned}&= \partial_y\bigg(P_s\big[T^{(\xi)}_{\tilde{n}-1}b(y,\cdot) - b(y,\cdot)\big] + \sum_{i=0}^{\tilde{n}-1}\frac{1}{i!}\Big(\bar{\tilde{P}}_{\tilde{n}-1-i,s}[\partial_1^i b(\xi,\cdot)] - P_s[\partial_1^i b(\xi,\cdot)]\Big)(y-\xi)^i\bigg)\\ &\quad + \partial_y\bigg(\sum_{i=0}^{\tilde{n}}\frac{1}{(\tilde{n}-i)!i!}\tilde{P}_{\tilde{n}-i,s}\big[\partial_1^i b(\xi,\cdot)\big](y-\xi)^i\bigg)\end{aligned}$$

(by definition of $\bar{\tilde{P}}_{\tilde{n}-i}$ in (2.6))

$$\begin{aligned}&= \partial_y\bigg(P_s\big[T^{(\xi)}_{\tilde{n}-1}b(y,\cdot) - b(y,\cdot)\big] + \sum_{i=0}^{\tilde{n}-1}\frac{1}{i!}\Big(\bar{\tilde{P}}_{\tilde{n}-i,s}[\partial_1^i b(\xi,\cdot)] - P_s[\partial_1^i b(\xi,\cdot)]\Big)(y-\xi)^i\bigg)\\ &\quad + \partial_y\bigg(\frac{1}{\tilde{n}!}\tilde{P}_{0,s}\big[\partial_1^{\tilde{n}} b(\xi,\cdot)\big](y-\xi)^{\tilde{n}}\bigg)\\ &= \partial_y\bigg(P_s\big[T^{(\xi)}_{\tilde{n}-1}b(y,\cdot) - b(y,\cdot)\big] + \sum_{i=0}^{\tilde{n}-1}\frac{1}{i!}\Big(\bar{\tilde{P}}_{\tilde{n}-i,s}[\partial_1^i b(\xi,\cdot)] - P_s[\partial_1^i b(\xi,\cdot)]\Big)(y-\xi)^i\bigg)\\ &\quad + \partial_y\bigg(\frac{1}{\tilde{n}!}\big(\tilde{P}_{0,s}\big[\partial_1^{\tilde{n}} b(\xi,\cdot)\big] - P_s\big[\partial_1^{\tilde{n}} b(\xi,\cdot)\big] + P_s\big[\partial_1^{\tilde{n}} b(\xi,\cdot)\big]\big)(y-\xi)^{\tilde{n}}\bigg)\\ &= \partial_y\bigg(P_s\big[T^{(\xi)}_{\tilde{n}}b(y,\cdot) - b(y,\cdot)\big] + \sum_{i=0}^{\tilde{n}}\frac{1}{i!}\Big(\bar{\tilde{P}}_{\tilde{n}-i,s}[\partial_1^i b(\xi,\cdot)] - P_s[\partial_1^i b(\xi,\cdot)]\Big)(y-\xi)^i\bigg),\end{aligned}$$

which is (3.25) with $n = \tilde{n} + 1$. □

We are now in the position to prove Theorem 2.9.

Proof of Theorem 2.9. We proceed by induction on N. For $N = 0$, by (3.23)-(3.25) we have

$$\big(p - \bar{\tilde{p}}_0\big)(0,\xi;t,x) = -\int_0^t\int_{\mathbb{R}} p(s,y;t,x)\partial_y\big(P_s[b(y,\cdot)]\tilde{p}_0(0,\xi;s,y)\big)\mathrm{d}y\mathrm{d}s$$

(integrating by parts)

$$= -\int_0^t\int_{\mathbb{R}}\big(\partial_y p(s,y;t,x)\big)P_s[b(y,\cdot)]\,\tilde{p}_0(0,\xi;s,y)\mathrm{d}y\mathrm{d}s.$$

By applying Lemma 3.5, Lemma 3.1 and Lemma A.1 on $\partial_y p(s,y;t,x)$, $P_t[b(y,\cdot)]$ and $\tilde{p}_0(0,\xi;s,y)$, respectively, we get

$$\big|\big(p - \bar{\tilde{p}}_0\big)(0,\xi;t,x)\big| \leq C_c\sigma\int_0^t (t-s)^{-\frac{1}{2}}\int_{\mathbb{R}}\Gamma_{c\sigma}\big(t-s,x-y\big)\Gamma_{c\sigma}(s,y-\xi)\mathrm{d}y\mathrm{d}s$$

(by Chapman-Kolmogorov identity)

$$\leq C_c\sigma\int_0^t (t-s)^{-\frac{1}{2}}\mathrm{d}s\,\Gamma_{c\sigma}(t,x-\xi) = C_c\,\sigma t^{\frac{1}{2}}\,\Gamma_{c\sigma}(t,x-\xi),$$

which is (2.23) for $N=1$.

Fix now $\widetilde{N}\leq\bar{N}$, assume that (2.23) holds true for any $N\leq\widetilde{N}$, and prove it true for $N=\widetilde{N}+1$. By (3.23)-(3.25) we obtain

$$\big(p-\bar{\tilde{p}}_{\widetilde{N}+1}\big)(0,\xi;t,x) = \sum_{n=0}^{\widetilde{N}+1}\frac{1}{(\widetilde{N}+1-n)!}I_n,$$

where

$$\begin{aligned}I_n &= \int_0^t\int_{\mathbb{R}} p(s,y;t,x)\partial_y\big(P_s\big[T^{(\xi)}_{n-1}b(y,\cdot)-b(y,\cdot)\big]\tilde{p}_{\widetilde{N}+1-n}(0,\xi;s,y)\big)\mathrm{d}y\mathrm{d}s\\ &\quad+\sum_{i=0}^{n-1}\frac{1}{i!}\int_0^t\int_{\mathbb{R}} p(s,y;t,x)\partial_y\Big(\big(\bar{\tilde{P}}_{n-1-i,s}[\partial_1^i b(\xi,\cdot)]-P_s[\partial_1^i b(\xi,\cdot)]\big)(y-\xi)^i\tilde{p}_{\widetilde{N}+1-n}(0,\xi;s,y)\Big)\mathrm{d}y\mathrm{d}s,\end{aligned}$$

(integrating by parts)

$$\begin{aligned}&= \int_0^t\int_{\mathbb{R}}\big(\partial_y p(s,y;t,x)\big)P_s\big[T^{(\xi)}_{n-1}b(y,\cdot)-b(y,\cdot)\big]\tilde{p}_{\widetilde{N}+1-n}(0,\xi;s,y)\mathrm{d}y\mathrm{d}s\\ &\quad+\sum_{i=0}^{n-1}\frac{1}{i!}\int_0^t\int_{\mathbb{R}}\big(\partial_y p(s,y;t,x)\big)\big(\bar{\tilde{P}}_{n-1-i,s}[\partial_1^i b(\xi,\cdot)]-P_s[\partial_1^i b(\xi,\cdot)]\big)(y-\xi)^i\tilde{p}_{\widetilde{N}+1-n}(0,\xi;s,y)\mathrm{d}y\mathrm{d}s.\end{aligned}$$

Now, by Taylor Theorem with Lagrange remainder along with assumption [**Hyp-**$b.\bar{N}+1$], we get

$$\big|P_s\big[T^{(\xi)}_{n-1}b(y,\cdot)-b(y,\cdot)\big]\big|\leq M\sigma^2|y-\xi|^n,$$

whereas, by induction hypothesis, we have

$$\big|\bar{\tilde{P}}_{n-1-i,s}[\partial_1^i b(\xi,\cdot)]-P_s[\partial_1^i b(\xi,\cdot)]\big|\leq C_c\,\sigma^2\big(\sigma^2 s\big)^{\frac{n-i}{2}}$$

as a corollary of (2.23) with $N=n-1-i$. Therefore, one has

$$\begin{aligned}|I_n| &\leq M\sigma^2\int_0^t\int_{\mathbb{R}}\big|\partial_y p(s,y;t,x)\big|\,|y-\xi|^n\big|\tilde{p}_{\widetilde{N}+1-n}(0,\xi;s,y)\big|\mathrm{d}y\mathrm{d}s\\ &\quad+C_c\,\sigma^2\sum_{i=0}^{n-1}\int_0^t\big(\sigma^2 s\big)^{\frac{n-i}{2}}\int_{\mathbb{R}}\big|\partial_y p(s,y;t,x)\big|\,|y-\xi|^i\big|\tilde{p}_{\widetilde{N}+1-n}(0,\xi;s,y)\big|\mathrm{d}y\mathrm{d}s.\end{aligned}$$

Eventually, by applying Lemma 3.5 on $\big|\partial_y p(s,y;t,x)\big|$, and by applying Lemma 3.1 with Lemma A.1 on $|y-\xi|^i\big|\tilde{p}_{\widetilde{N}+1-n}(0,\xi;s,y)\big|$, we obtain

$$|I_n|\leq C_c\,\sigma^{\widetilde{N}+2}\int_0^t(t-s)^{-\frac{1}{2}}s^{\frac{\widetilde{N}+1}{2}}\int_{\mathbb{R}}\Gamma_{c\sigma}\big(t-s,x-y\big)\Gamma_{c\sigma}(s,y-\xi)\mathrm{d}y\mathrm{d}s$$

(by Chapman-Kolmogorov identity)

$$\leq C_c\,\sigma^{\widetilde{N}+2}\int_0^t(t-s)^{-\frac{1}{2}}s^{\frac{\widetilde{N}+1}{2}}\mathrm{d}s\,\Gamma_{c\sigma}(t,x-\xi) = C_c\,\big(\sigma^2 t\big)^{\frac{\widetilde{N}+2}{2}}\,\Gamma_{c\sigma}(t,x-\xi),$$

which yields (2.23) with $N=\widetilde{N}+1$. □

4 Proof of Theorems 2.12 and 2.16

4.1 Proof of Theorem 2.12

Hereafter, throughout this subsection, we assume the hypotheses [**Hyp-**$b.N$] and [**Hyp-**σ] to be in force for a fixed $N \in \mathbb{N}^*$. Recall that $\mathcal{S}(\mathbb{R}^n)$ denotes the Schwartz space of rapidly decreasing functions on $\mathbb{R}^n$.

The proof of Theorem 2.12, which is based on Lemma 3.3 combined with the following

Lemma 4.1. *For any $t_0 < s < t$, $x, y, \bar{x} \in \mathbb{R}$, and $k \in \mathbb{N}^*$ with $k \leq N$, we have*

$$\begin{aligned}
&\frac{1}{k!}\sum_{j=0}^{k}\binom{k}{j}\int_{\mathbb{R}}\Gamma_\sigma\big(t-s;x-y-\mathbf{m}^{\bar{x}}(s,t)\big)\mathcal{B}^{\bar{x}}_{k,j}(s,y)f(y)\mathrm{d}y\\
&=\mathcal{G}^{\bar{x}}_k(s,t,x)\int_{\mathbb{R}}\Gamma_\sigma\big(t-s;x-y-\mathbf{m}^{\bar{x}}(s,t)\big)f(y)\mathrm{d}y,
\end{aligned}\tag{4.1}$$

and

$$\begin{aligned}
&\frac{1}{k!}\sum_{j=0}^{k}\binom{k}{j}\int_{\mathbb{R}}f(x)\mathcal{B}^{\bar{x}}_{k,j}(s,x)\Gamma_\sigma\big(s-t_0;x-y-\mathbf{m}^{\bar{x}}(t_0,s)\big)\mathrm{d}x\\
&=\bar{\mathcal{G}}^{\bar{x}}_k(t_0,s,y)\int_{\mathbb{R}}\Gamma_\sigma\big(s-t_0;x-y-\mathbf{m}^{\bar{x}}(t_0,s)\big)f(x)\mathrm{d}x,
\end{aligned}\tag{4.2}$$

for any $f \in \mathcal{S}(\mathbb{R})$, where $\bar{\mathcal{G}}^{\bar{x}}_k = \bar{\mathcal{G}}^{\bar{x}}_k(t_0,s,y)$ is the differential operator acting as

$$\bar{\mathcal{G}}^{\bar{x}}_k(t_0,s,y)u(y) = -\frac{1}{k!}\sum_{j=0}^{k}\binom{k}{j}\tilde{P}_{k-j,s}\big[\partial_1^j b(\bar{x},\cdot)\big]\big(\bar{\mathcal{M}}^{\bar{x}}(t_0,s,y)-\bar{x}\big)^j\partial_y u(y),\tag{4.3}$$

and the operator $\bar{\mathcal{M}}^{\bar{x}} = \bar{\mathcal{M}}^{\bar{x}}(t_0,s,y)$ acts as

$$\bar{\mathcal{M}}^{\bar{x}}(t_0,s,y) = y + \mathbf{m}^{\bar{x}}(t_0,s) + \sigma^2(s-t_0)\,\partial_y.$$

Moreover, the following relation holds:

$$\mathcal{G}^{\bar{x}}_k(s,t,x)\Gamma_\sigma(t-t_0;x-y) = \bar{\mathcal{G}}^{\bar{x}}_k(t_0,s,y)\Gamma_\sigma(t-t_0;x-y).\tag{4.4}$$

Proof. Similarly to (3.9) and (3.10), we have

$$\begin{aligned}
\partial_x\Gamma_\sigma\big(t-s;x-y-\mathbf{m}^{\bar{x}}(s,t)\big) &= -\partial_y\Gamma_\sigma\big(t-s;x-y-\mathbf{m}^{\bar{x}}(s,t)\big),\\
y\,\Gamma_\sigma\big(t-s;x-y-\mathbf{m}^{\bar{x}}(s,t)\big) &= \mathcal{M}^{\bar{x}}(s,t,x)\Gamma_\sigma\big(t-s;x-y-\mathbf{m}^{\bar{x}}(s,t)\big),\\
x\,\Gamma_\sigma\big(t-s;x-y-\mathbf{m}^{\bar{x}}(s,t)\big) &= \bar{\mathcal{M}}^{\bar{x}}(s,t,y)\Gamma_\sigma\big(t-s;x-y-\mathbf{m}^{\bar{x}}(s,t)\big).
\end{aligned}$$

Now the proof of (4.1)-(4.2) is completely analogous to that of (3.5)-(3.6), and thus we omit the details for brevity. Eventually, identity (4.4) is a consequence of (4.1)-(4.2) combined with the Chapman-Kolmogorov equation:

$$\int_{\mathbb{R}}\Gamma_\sigma\big(t-s,x-\eta-\mathbf{m}^{\bar{x}}(s,t)\big)\Gamma_\sigma\big(s-t_0,\eta-y-\mathbf{m}^{\bar{x}}(t_0,s)\big)\mathrm{d}\eta = \Gamma_\sigma\big(t-t_0,x-y-\mathbf{m}^{\bar{x}}(t_0,t)\big),$$

for any $t_0 < s < t$ and $x, y, \bar{x} \in \mathbb{R}$. □

We are now ready to prove Theorem 2.12.

Proof of Theorem 2.12. The proof is a straightforward application of Lemma 3.3 with the kernel

$$\varphi_{s,t}(\xi,x)=\Gamma_\sigma\big(t-s;x-\xi-\mathbf{m}^{\bar x}(s,t)\big),\qquad \xi,x\in\mathbb{R},$$

and the operators

$$\mathcal{P}_{h,s,t,\bar x}=\mathcal{G}_h^{\bar x}(s,t,x),\qquad \bar{\mathcal{P}}_{h,s,t,\bar x}=\bar{\mathcal{G}}_h^{\bar x}(s,t,\xi),\qquad h\in\mathbb{N}^*,\ h\le N,$$

with $\mathcal{G}_h^{\bar x}$ and $\bar{\mathcal{G}}_h^{\bar x}$ as in (2.35) and (4.3), respectively. In particular, under assumption [**Hyp-***b.N*], the functions $(\tilde P_{n,t})_{n\le N}$ appearing in (2.35)-(4.3) are well defined and satisfy the estimates (3.2). It follows that $\mathcal{P}_{h,s,t,\bar x}$ and $\bar{\mathcal{P}}_{h,s,t,\bar x}$ are well defined as operators from $\mathcal{S}(\mathbb{R}^2)$ onto itself, and they admit a representation of the form (3.14) where $s\mapsto a_{i,j}^{(h,\bar x)}(s,t)$ and $s\mapsto \bar a_{i,j}^{(h,\bar x)}(s,t)$ are bounded and measurable (by assumption [**Hyp-***b.N*]) functions on $[0,t]$. Moreover, (4.4) implies the symmetry property (3.15). Therefore, by Lemma 3.3 we have that $(f_{n,t_0,t,\bar x})_{n\le N,0\le t_0<t,\bar x\in\mathbb{R}}$ as in (3.17) is well defined as a family of functions in $\mathcal{S}(\mathbb{R}^2)$, and in particular

$$f_{n,t_0,t,\bar x}(\xi,x)=\sum_{h=1}^n\frac{n!}{(n-h)!}\int_{t_0}^t\mathcal{G}_h^{\bar x}(s,t,x)\int_{\mathbb{R}}\Gamma_\sigma\big(t-s,x-y-\mathbf{m}^{\bar x}(s,t)\big)f_{n-h,t_0,s,\bar x}(\xi,y)\mathrm{d}y\mathrm{d}s$$

(by (4.1))

$$=\sum_{h=1}^n\binom{n}{h}\sum_{j=0}^h\binom{h}{j}\int_0^t\int_{\mathbb{R}}\Gamma_\sigma\big(t-s,x-y-\mathbf{m}^{\bar x}(s,t)\big)\mathcal{B}_{h,j}^{\bar x}\big(s,y\big)f_{n-h,t_0,s,\bar x}(\xi,y)\mathrm{d}y\mathrm{d}s.$$

This proves that the function $p_n^{\bar x}(s,\xi;t,x):=-f_{n,s,t,\bar x}(\xi,x)$ is well defined, satisfies (2.33), and also

$$p_n^{\bar x}(s,\xi;t,x)=-\mathcal{L}_n^{\bar x}(s,t,x)\,p_0^{\bar x}(s,\xi;t,x).$$

This concludes the proof.

□

4.2 Proof of Theorem 2.16

Throughout this section we fix a time $T>0$ and we consider assumptions [**Hyp-**σ], [**Hyp-***b*.0] and [**Hyp-***b*.$\bar N+1$] to be in force for a fixed $\bar N\in\mathbb{N}^*$.

Notation 4.2. Throughout the rest of this section, unless explicitly stated, we will denote by $(C_c)_{c>0}$ any family of positive constants that depend at most on $\bar N, T$ and on the constants $M,\bar\sigma$ of assumptions [**Hyp-**σ], [**Hyp-***b*.0] and [**Hyp-***b*.$\bar N+1$]. Note that, in particular, C_c is independent of σ.

We start with the following upper bounds for the functions $p_n^{\bar x}(s,\xi;t,x)$.

Lemma 4.3. *For any $c>1$, $k\in\mathbb{N}$, and $n\in\mathbb{N}^*$ with $n\le\bar N$, we have*

$$\big|\partial_x^k p_0^{\bar x}(s,\xi;t,x)\big|\le C_c\big(\sigma^2(t-s)\big)^{-\frac{k}{2}}\Gamma_{c\sigma}\big(t-s,x-\xi\big),\tag{4.5}$$

$$\big|\partial_x^k p_n^{\bar x}(s,\xi;t,x)\big|\le C_c\big(\sigma^2(t-s)\big)^{\frac{n+1-k}{2}}\bigg(\sum_{j=0}^n|x-\bar x|^j\big(\sigma^2(t-s)\big)^{-\frac{j}{2}}\bigg)\Gamma_{c\sigma}\big(t-s,x-\xi\big),\tag{4.6}$$

for any $0<t\le T$ and $x,\xi,\bar x\in\mathbb{R}$.

Proof. Note that (2.30) and (3.2) yield $|\mathbf{m}^{\bar{x}}(s,t)| \le C_c\,\sigma^2(t-s)$. Therefore, (4.5) follows by applying Lemma A.2 and A.3. Now, for $n \in \mathbb{N}^*$ with $n \le \bar{N}$, by Corollary 2.14 we have

$$\big|\partial_x^k p_n^{\bar{x}}(s,\xi;t,x)\big| \le C_c \sum_{h=1}^{n} \sum_{i\in I_{n,h}} \sum_{\substack{\gamma,m,l,j\in\mathbb{N}^h \\ l_q\le j_q\le i_q \\ m_q\le \lfloor \frac{j_q-l_q}{2}\rfloor \\ \gamma_q\le j_q-2m_q-l_q}} \big|\mathbf{F}^{\bar{x}}_{\gamma,l+m,j,i}(s,t)\big|\,\big|\partial_x^k \mathcal{H}^{\bar{x}}_{j-2m-\gamma,l}(x) p_0^{\bar{x}}(s,\xi;t,x)\big|.$$

Now, proceeding as we did before to prove (3.4), one easily gets

$$\begin{aligned}\big|\mathbf{F}^{\bar{x}}_{\gamma,l+m,j,i}(s,t)\big| \le C_c\big(\sigma^2(t-s)\big)^{\frac{\sum_{q=1}^h (i_q+2-j_q+2l_q+2m_q+2\gamma_q)}{2}} &\le C_c\big(\sigma^2(t-s)\big)^{\frac{n+2h+\sum_{q=1}^h (2l_q+2m_q+2\gamma_q-j_q)}{2}} \\ &\le C_c\big(\sigma^2(t-s)\big)^{\frac{n+1+h+\sum_{q=1}^h (2l_q+2m_q+\gamma_q-j_q)}{2}}. \quad (4.7)\end{aligned}$$

On the other hand, definition (2.21) yields

$$\big|\partial_x^k \mathcal{H}^{\bar{x}}_{j-2m-\gamma,l}(x) p_0^{\bar{x}}(s,\xi;t,x)\big| \le C_c \sum_{\substack{0\le\rho_1\le n,\,0\le\rho_2\le 2n \\ \rho_1-\rho_2=-h-k+\sum_{q=1}^h (j_q-2m_q-\gamma_q-2l_q)}} |x-\bar{x}|^{\rho_1}\big|\partial_x^{\rho_2} p_0^{\bar{x}}(s,\xi;t,x)\big|$$

(by Lemma A.2)

$$\le C_c \sum_{\substack{0\le\rho_1\le n,\,0\le\rho_2\le 2n \\ \rho_1-\rho_2=-h-k+\sum_{q=1}^h (j_q-2m_q-\gamma_q-2l_q)}} |x-\bar{x}|^{\rho_1}\big(\sigma^2(t-s)\big)^{-\frac{\rho_2}{2}}\Gamma_{c\sigma}\big(t-s,x-\xi\big),$$

which in turn, combined with (4.7), yields

$$\begin{aligned}&\big|\mathbf{F}^{\bar{x}}_{\gamma,l+m,j,i}(s,t)\big|\,\big|\partial_x^k \mathcal{H}^{\bar{x}}_{j-2m-\gamma,l}(x) p_0^{\bar{x}}(s,\xi;t,x)\big| \\ &\le C_c\big(\sigma^2(t-s)\big)^{\frac{n+1}{2}}\Bigg(\sum_{\rho_1=0}^{n}|x-\bar{x}|^{\rho_1}\big(\sigma^2(t-s)\big)^{-\frac{\rho_1+k}{2}}\Bigg)\Gamma_{c\sigma}\big(t-s,x-\xi\big),\end{aligned}$$

and this concludes the proof. □

Remark 4.4. Under assumptions [**Hyp-**$b.\bar{N}+1$] and [**Hyp-**σ], the functions $p_n^{\bar{x}}(s,\xi;\cdot,\cdot)$, $1 \le n \le \bar{N}$, are solutions of the Cauchy problems (2.32) , meaning that the PDE is satisfied on the internal domain $]s,\infty[\times\mathbb{R}$, and that the initial condition is fulfilled in the distributional sense, i.e.

$$\lim_{\substack{(t,\xi)\to(s,\bar{\xi}) \\ t>0}} \int_{\mathbb{R}} p_n^{\bar{x}}(s,\xi;t,x)\varphi(t,x)\mathrm{d}x = 0, \qquad \bar{\xi}\in\mathbb{R}, \qquad \varphi \in C_b([s,\bar{T}[\times\mathbb{R}), \quad \bar{T}\in]s,\infty[.$$

In fact, the initial condition above is a straightforward consequence of the upper bound (4.6), whereas the fact that $p_n^{\bar{x}}(s,\xi;\cdot,\cdot)$ solves the PDE in (2.32) follows by differentiating (2.33), after using again estimate (4.6) and integrating by parts to deal with the time-integral in (2.33), which is singular near 0 and t.

In order to continue, let us introduce the family of operators $\big(\bar{\mathcal{A}}^{\bar{x}}_{n,s}\big)_{n\le\bar{N},0\le s\le T,\bar{x}\in\mathbb{R}}$

$$\mathcal{A}^{\bar{x}}_{n,s} := \mathcal{A}^{\bar{x}}_0 + \sum_{h=1}^{n}\mathcal{A}^{\bar{x}}_{h,s}, \qquad \text{with} \qquad \mathcal{A}^{\bar{x}}_{h,s} := -\frac{1}{h!}\sum_{i=0}^{h}\binom{h}{i}\mathcal{B}^{\bar{x}}_{h,i}(s,\cdot),$$

where the family of operators $\big(\mathcal{B}^{\bar{x}}_{h,i}\big)_{0\le i\le h,\bar{x}\in\mathbb{R}}$ is as defined in (2.14). We also recall to the reader the definition of $\bar{p}^{\bar{x}}_N$ given in (2.26).

The proof of Theorem 2.16 is preceded by the following two Lemmas.

Lemma 4.5. *For any $N\in\mathbb{N}$ with $N\le\bar{N}$, and $\bar{x}\in\mathbb{R}$, the following identity holds:*

$$\big(p-\bar{p}^{\bar{x}}_N\big)(s,\xi;t,x)=\sum_{n=0}^{N}\frac{1}{(N-n)!}\int_s^t\int_{\mathbb{R}}p(r,y;t,x)\big(\mathcal{A}_r(y)-\bar{\mathcal{A}}^{\bar{x}}_{n,r}(y)\big)p^{\bar{x}}_{N-n}(s,\xi;r,y)\mathrm{d}y\mathrm{d}r, \tag{4.8}$$

for any $0\le s<t$ and $x,\xi\in\mathbb{R}$.

Lemma 4.6. *For any $n\in\mathbb{N}$ with $n\le\bar{N}$ we have*

$$\big(\mathcal{A}_r(y)-\bar{\mathcal{A}}^{\bar{x}}_{n,r}(y)\big)=\partial_y\bigg(P_r\big[T^{(\bar{x})}_n b(y,\cdot)-b(y,\cdot)\big]+\sum_{i=0}^{n}\frac{1}{i!}\Big(\bar{\tilde{P}}_{n-i,r}\big[\partial_1^i b(\bar{x},\cdot)\big]-P_r\big[\partial_1^i b(\bar{x},\cdot)\big]\Big)(y-\bar{x})^i\bigg), \tag{4.9}$$

where $T^{(\bar{x})}_k b(y,\cdot)$ represents the k-th order Taylor expansion of $b(y,\cdot)$ centered at $\bar{x}$, see (3.26).

The proofs of Lemmas 4.5 and 4.6 are totally analogous to those of Lemmas 3.7 and 3.8 respectively, and they are based on Remark 4.4 along with the upper bounds in Lemma 4.3, and the Cauchy problems (1.3)-(2.31). For brevity, we leave the detailed proofs to the reader.

We are now in the position to prove Theorem 4.2.

Proof of Theorem 2.16. Let $N\in\mathbb{N}$ with $N\le\bar{N}$. By (4.8)-(4.9) we get

$$\big(p-\bar{p}^{\bar{x}}_N\big)(s,\xi;t,x)=\sum_{n=0}^{N}\frac{1}{(N-n)!}I^{\bar{x}}_n, \tag{4.10}$$

where

$$\begin{aligned}I^{\bar{x}}_n&=\int_s^t\int_{\mathbb{R}}p(r,y;t,x)\partial_y\big(P_r\big[T^{(\bar{x})}_n b(y,\cdot)-b(y,\cdot)\big]p^{\bar{x}}_{N-n}(s,\xi;r,y)\big)\mathrm{d}y\mathrm{d}r\\&\quad+\sum_{i=0}^{n}\frac{1}{i!}\int_s^t\int_{\mathbb{R}}p(r,y;t,x)\partial_y\Big(\big(\bar{\tilde{P}}_{n-i,r}[\partial_1^i b(\bar{x},\cdot)]-P_r[\partial_1^i b(\bar{x},\cdot)]\big)(y-\bar{x})^i p^{\bar{x}}_{N-n}(s,\xi;r,y)\Big)\mathrm{d}y\mathrm{d}r,\end{aligned}$$

(integrating by parts)

$$\begin{aligned}&=\int_s^t\int_{\mathbb{R}}\big(\partial_y p(r,y;t,x)\big)P_r\big[T^{(\bar{x})}_n b(y,\cdot)-b(y,\cdot)\big]p^{\bar{x}}_{N-n}(s,\xi;r,y)\mathrm{d}y\mathrm{d}r\\&\quad+\sum_{i=0}^{n}\frac{1}{i!}\int_s^t\int_{\mathbb{R}}\big(\partial_y p(r,y;t,x)\big)\big(\bar{\tilde{P}}_{n-i,r}[\partial_1^i b(\bar{x},\cdot)]-P_r[\partial_1^i b(\bar{x},\cdot)]\big)(y-\bar{x})^i p^{\bar{x}}_{N-n}(s,\xi;r,y)\mathrm{d}y\mathrm{d}r.\end{aligned}$$

Now, by Taylor Theorem with Lagrange remainder along with assumption [**Hyp-**$b.\bar{N}+1$], we get

$$\big|P_r\big[T^{(\bar{x})}_n b(y,\cdot)-b(y,\cdot)\big]\big|\le M\sigma^2|y-\bar{x}|^{n+1},$$

whereas, by Theorem 2.9, we have

$$\big|\bar{\tilde{P}}_{n-i,r}[\partial_1^i b(\bar{x},\cdot)]-P_r[\partial_1^i b(\bar{x},\cdot)]\big|\le C_c\,\sigma^2\big(\sigma^2 r\big)^{\frac{n-i+1}{2}}$$

as a corollary of (2.23) with $N = n - i$. Therefore, one has

$$\big|I_n^{\bar{x}}\big| \le C_c\,\sigma^2 \sum_{i=0}^{n+1} \int_s^t \big(\sigma^2 r\big)^{\frac{n-i+1}{2}} \int_{\mathbb{R}} \big|\partial_y p(r,y;t,x)\big|\,|y-\bar{x}|^i \big|p_{N-n}^{\bar{x}}(s,\xi;r,y)\big| \mathrm{d}y \mathrm{d}r. \tag{4.11}$$

Consider now the case $0 \le n < N$. By applying Lemma 4.3 we obtain

$$|y-\bar{x}|^i \big|p_{N-n}^{\bar{x}}(s,\xi;r,y)\big| \le C_c\big(\sigma^2(r-s)\big)^{\frac{1}{2}} |y-\bar{x}|^i \Big(|y-\bar{x}|^{N-n} + \big(\sigma^2(r-s)\big)^{\frac{N-n}{2}}\Big) \Gamma_{\sqrt{c}\sigma}\big(r-s, y-\xi\big),$$

which combined with the identity

$$\big|y - \big(\lambda x + (1-\lambda)\xi\big)\big| \le \lambda|y-x| + (1-\lambda)|y-\xi| \le \max\big(|y-x|, |y-\xi|\big), \qquad \lambda \in [0,1]$$

yields

$$\begin{aligned}
&|y-\bar{x}|^i \big|p_{N-n}^{\bar{x}}(s,\xi;r,y)\big|\Big|_{\bar{x}=\lambda x+(1-\lambda)\xi} \\
&\qquad \le C_c\big(\sigma^2(r-s)\big)^{\frac{1}{2}} |y-\xi|^i \Big(|y-\xi|^{N-n} + \big(\sigma^2(r-s)\big)^{\frac{N-n}{2}}\Big) \Gamma_{\sqrt{c}\sigma}\big(r-s, y-\xi\big) \\
&\qquad\quad + C_c\big(\sigma^2(r-s)\big)^{\frac{1}{2}} |y-x|^i \Big(|y-x|^{N-n} + \big(\sigma^2(r-s)\big)^{\frac{N-n}{2}}\Big) \Gamma_{\sqrt{c}\sigma}\big(r-s, y-\xi\big)
\end{aligned}$$

(by applying Lemma A.1)

$$\begin{aligned}
&\le C_c\big(\sigma^2(r-s)\big)^{\frac{1}{2}} \big(\sigma^2(r-s)\big)^{\frac{N-n+i}{2}} \Gamma_{c\sigma}\big(r-s, y-\xi\big) \\
&\quad + C_c\big(\sigma^2(r-s)\big)^{\frac{1}{2}} |y-x|^i \Big(|y-x|^{N-n} + \big(\sigma^2(r-s)\big)^{\frac{N-n}{2}}\Big) \Gamma_{c\sigma}\big(r-s, y-\xi\big).
\end{aligned} \tag{4.12}$$

Now, replacing (4.12) into (4.11), and applying Lemma 3.5 along with Chapman-Kolmogorov equation, we finally obtain

$$\begin{aligned}
\Big|I_n^{\bar{x}}\big|_{\bar{x}=\lambda x+(1-\lambda)\xi}\Big| &\le C_c\,\sigma^{N+3} \int_s^t (t-r)^{-\frac{1}{2}} \Big(\sum_{i=0}^{n+1} r^{\frac{n-i+1}{2}} (r-s)^{\frac{N-n+1+i}{2}}\Big) \mathrm{d}r\, \Gamma_{c\sigma}\big(t-s, x-\xi\big) \\
&\quad + C_c\,\sigma^{N+3} \int_s^t (r-s)^{\frac{1}{2}} \Big(|t-r|^{\frac{N-n}{2}} + (r-s)^{\frac{N-n}{2}}\Big) \Big(\sum_{i=0}^{n+1} r^{\frac{n-i+1}{2}} (t-r)^{\frac{i-1}{2}}\Big) \mathrm{d}r\, \Gamma_{c\sigma}\big(t-s, x-\xi\big)
\end{aligned}$$

(by using $0 \le s < t \le T$)

$$\le C_c\,\big(\sigma^2 t\big)^{\frac{N}{2}} \big(\sigma^2(t-s)\big)^{\frac{3}{2}}\, \Gamma_{c\sigma}\big(t-s, x-\xi\big). \tag{4.13}$$

Similar computations give a bound for I_N:

$$\begin{aligned}
\Big|I_N^{\bar{x}}\big|_{\bar{x}=\lambda x+(1-\lambda)\xi}\Big| &\le C_c\,\sigma^{N+2} \int_s^t (t-r)^{-\frac{1}{2}} \Big(\sum_{i=0}^{N+1} r^{\frac{N-i+1}{2}} (r-s)^{\frac{i}{2}}\Big) \mathrm{d}r\, \Gamma_{c\sigma}\big(t-s, x-\xi\big) \\
&\quad + C_c\,\sigma^{N+2} \int_s^t \Big(\sum_{i=0}^{N+1} r^{\frac{N-i+1}{2}} (r-s)^{\frac{i-1}{2}}\Big) \mathrm{d}r\, \Gamma_{c\sigma}\big(t-s, x-\xi\big)
\end{aligned}$$

(by using $0 \le s < t \le T$)

$$\le C_c\,\big(\sigma^2 t\big)^{\frac{N+1}{2}} \big(\sigma^2(t-s)\big)^{\frac{1}{2}}\, \Gamma_{c\sigma}\big(t-s, x-\xi\big). \tag{4.14}$$

Combining (4.13)-(4.14) with (4.10) yields (2.36). □

A Gaussian Estimates

We recall here some standard estimate for the Gaussian density

$$\Gamma_\sigma(t,x) := \frac{1}{\sigma\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2\sigma^2 t}\right), \qquad x \in \mathbb{R},\ t > 0. \tag{A.1}$$

In the following, $\sigma > 0$ is fixed.

Lemma A.1. *For any $n \in \mathbb{N}$ and $c > 1$ we have*

$$\left(\frac{|x|}{\sigma\sqrt{t}}\right)^n \Gamma_\sigma(t,x) \le c\left(\frac{c^2 n}{(c^2-1)e}\right)^{\frac{n}{2}} \Gamma_{c\sigma}(t,x), \qquad x \in \mathbb{R},\ t > 0.$$

Proof. Set $z = \frac{|x|}{\sigma\sqrt{t}}$. For any $c > 1$ we have

$$\left(\frac{|x|}{\sigma\sqrt{t}}\right)^n \Gamma_\sigma(t,x) = \frac{z^n}{\sigma\sqrt{2\pi t}} \exp\left(-\frac{z^2}{2}\right) = c\, G(z)\Gamma_{c\sigma}(t,x),$$

with

$$G(z) = z^n \exp\left(-\frac{z^2}{2}\left(1 - \frac{1}{c^2}\right)\right), \qquad z \ge 0.$$

The statement now follows by observing that G attains a global maximum at $z_n^* = \sqrt{\frac{c^2 n}{c^2-1}}$ and that

$$G(z_n^*) = e^{-\frac{n}{2}} \left(\frac{c^2 n}{c^2-1}\right)^{n/2}.$$

□

Lemma A.2. *For any $n \in \mathbb{N}$ and $c > 1$ we have*

$$|\partial_x^n \Gamma_\sigma(t,x)| \le C\left(\sigma\sqrt{t}\right)^{-n} \Gamma_{c\sigma}(t,x), \qquad x \in \mathbb{R},\ t > 0,$$

where C is a positive constant only dependent on n and c.

Proof. Let us define the n-th order Hermite polynomial as $\mathbf{H}_n(z) := e^{\frac{1}{2}z^2}\partial_z^n e^{-\frac{1}{2}z^2}$. Then, by definition (A.1) we have

$$\partial_x^n \Gamma_\sigma(t,x) = (\sigma^2 t)^{-\frac{n}{2}} \mathbf{H}_n\left(\frac{x}{\sigma\sqrt{t}}\right) \Gamma_\sigma(t,x),$$

and thus the statement easily stems from Lemma A.1. □

Lemma A.3. *For any $c > 1$ we have*

$$\Gamma_\sigma(t, x+\mu) \le c \exp\left(\frac{\mu^2}{2\sigma^2 t(c^2-1)}\right) \Gamma_{c\sigma}(t,x), \qquad x, \mu \in \mathbb{R},\ t > 0.$$

Proof. This is straightforward using the inequality $x^2 \le (x+\mu)^2(1+\delta) + \mu^2(1+1/\delta)$ (available for any $\delta > 0$), with $1 + \delta = c^2$. □

B Path-dependent McKean SDE

Lemma B.1. *Assume that $F = (F(t,x,x') : 0 \le t \le T, x, x' \in C^0([0,T],\mathbb{R}^d))$ is a family of bounded continuous functional valued in $\mathbb{R}^d$, globally Lipschitz in the paths x and x', i.e. there exists a finite constant L_F such that*

$$|F(t,x,x') - F(t,y,y')| \le L_F \left(\sup_{0\le s\le t} |x_s - x'_s| + \sup_{0\le s\le t} |y_s - y'_s| \right), \qquad \forall t \in [0,T],$$

for any continuous functions $x, y, x', y' : C^0([0,T],\mathbb{R}^d) \mapsto \mathbb{R}^d$. Let $\sigma \in \mathbb{R}$; consider the equation

$$\begin{cases} \mathrm{d}Z_t = \mathbb{E}'\big[F(t,Z,Z')\big]\mathrm{d}t + \sigma \mathrm{d}W_t, \qquad t > 0, \\ Z_0 = Y, \end{cases} \tag{B.1}$$

where Y is a square integrable random variable, independent of the Brownian motion W. Then there exists an unique solution to (B.1)*, continuous in time and square integrable uniformly in time.*

Proof. This is a sort of extension of [Pro04, Theorem 7, Chapter V] to the McKean-Vlasov case. We closely follow the proof of [Szn91] by constructing the solution using a fixed point argument. For this, we introduce the Wasserstein distance $D_T(\cdot,\cdot)$, defined on the set $\mathcal{M}(\mathcal{C}_T)$ of the probability measures on the continuous trajectories $\mathcal{C}_T := C^0([0,T],\mathbb{R}^d)$:

$$D_T(\nu_1,\nu_2) := \inf_{\nu \in \mathcal{M}(\mathcal{C}_T\times\mathcal{C}_T) \text{ with marginals } \nu_1 \text{ and } \nu_2} \Big\{ \int (\sup_{s\le T} |\omega_{1,s} - \omega_{2,s}| \wedge 1) \nu(\mathrm{d}\omega_1, \mathrm{d}\omega_2) \Big\}.$$

This defines a complete metric on $\mathcal{M}(\mathcal{C}_T)$, whose topology is that of weak convergence. For ν in $\mathcal{M}(\mathcal{C}_T)$, let us denote by $\Phi(\nu)$ the distribution of the solution defined as

$$Z_t = Y + \int_0^t \left(\int_{\mathcal{C}_s} F(s,Z,\omega')\nu(d\omega') \right) \mathrm{d}s + \sigma W_t.$$

Note that this path-dependent stochastic differential equation has a unique strong solution [Pro04, Theorem 7, Chapter V] since $x \in C^0([0,T],\mathbb{R}^d) \mapsto \int_{\mathcal{C}_s} F(s,x,\omega)\nu(\mathrm{d}\omega)$ is Lipschitz, uniformly in $s \in [0,T]$. Given two probability measures ν_1 and ν_2 in $\mathcal{M}(\mathcal{C}_T)$, compare the two SDEs corresponding to Z_1 and Z_2: for any coupling ν on $\mathcal{C}_t$ with marginal distributions ν_1 and ν_2, we have

$$\begin{aligned} \sup_{s\le t} |Z_{1,s} - Z_{2,s}| &\le \int_0^t \Big| \int_{\mathcal{C}_t} F(s,Z_1,\omega_1)\nu_1(\mathrm{d}\omega_1) - \int_{\mathcal{C}_t} F(s,Z_2,\omega_2)\nu_2(\mathrm{d}\omega_2) \Big| \mathrm{d}s \\ &\le \int_0^t \Big[L_F \sup_{r\le s} |Z_{1,r} - Z_{2,r}| + \int_{\mathcal{C}_t\times\mathcal{C}_t} \big([L_F \sup_{r\le s} |\omega_{1,r} - \omega_{2,r}|] \wedge [2|F|_\infty] \big) \nu(\mathrm{d}\omega_1, \mathrm{d}\omega_2) \Big] \mathrm{d}s. \end{aligned}$$

Set $K = \max(L_F, 2|F|_\infty)$; then, taking the infimum over the couplings ν and using the Gronwall lemma, we get

$$\sup_{s\le t} |Z_{1,s} - Z_{2,s}| \le K \int_0^t \Big[\sup_{r\le s} |Z_{1,r} - Z_{2,r}| + D_s(\nu_1,\nu_2) \Big] \mathrm{d}s \le K e^{KT} \int_0^t D_s(\nu_1,\nu_2) \mathrm{d}s.$$

From this we deduce a precise control of $D_t(\Phi(\nu_1),\Phi(\nu_2)) \le K e^{KT} \int_0^t D_s(\nu_1,\nu_2)\mathrm{d}s$. We can easily conclude to the result: we take $\nu \in \mathcal{M}(\mathcal{C}_T)$ and iterating this procedure, we get $D_T(\Phi^k(\nu),\Phi^{k+1}(\nu)) \le \frac{(KTe^{KT})^k}{k!} D_T(\nu,\Phi(\nu))$, which allows us to prove easily that $(\Phi^k(\nu))_k$ is a Cauchy sequence, converging to the fixed point of Φ. This implies existence and uniqueness as announced. □

References

[AK02] F. Antonelli and A. Kohatsu-Higa. Rate of convergence of a particle method to the solution of the McKean-Vlasov equation. *Ann. Appl. Probab.*, 12(2):423–476, 2002.

[BG12] R. Bompis and E. Gobet. Asymptotic and non asymptotic approximations for option valuation. In T. Gerstner and P. Kloeden, editors, *Recent Developments in Computational Finance: Foundations, Algorithms and Applications*, chapter 4, pages 159–241. World Scientific Publishing Company, 2012.

[BK10] V.S. Borkar and K.S. Kumar. McKean-Vlasov limit in portfolio optimization. *Stoch. Anal. Appl.*, 28(5):884–906, 2010.

[BT97] M. Bossy and D. Talay. A stochastic particle method for the McKean-Vlasov and the Burgers equation. *Math. Comp.*, 66(217):157–192, 1997.

[CD15] R. Carmona and F. Delarue. Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics. *Ann. Probab.*, 43(5):2647–2700, 2015.

[CDL13] R. Carmona, F. Delarue, and A. Lachapelle. Control of McKean-Vlasov dynamics versus mean field games. *Math. Financ. Econ.*, 7(2):131–166, 2013.

[CG15] P.E. Chaudru de Raynal and C.A. Garcia Trillos. A cubature based algorithm to solve decoupled McKean-Vlasov forward-backward stochastic differential equations. *Stochastic Process. Appl.*, 125(6):2206–2255, 2015.

[DIRT15] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré. Global solvability of a networked integrate-and-fire model of McKean-Vlasov type. *Ann. Appl. Probab.*, 25(4):2096–2133, 2015.

[Fri64] A. Friedman. *Partial differential equations of parabolic type.* Prentice-Hall Inc., Englewood Cliffs, N.J., 1964.

[JR15] B. Jourdain and J. Reygner. Capital distribution and portfolio performance in the mean-field Atlas model. *Ann. Finance*, 11(2):151–198, 2015.

[LPP15] M. Lorig, S. Pagliarani, and A. Pascucci. Analytical expansions for parabolic equations. *SIAM J. Appl. Math.*, 75(2):468–491, 2015.

[Mél96] S. Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, volume 1627 of *Lecture Notes in Math.*, pages 42–95. Springer, Berlin, 1996.

[Pro04] P. Protter. *Stochastic integration and differential equations.* Springer Verlag, second edition, 2004.

[Szn91] Alain-Sol Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.

[Tra08] V.C. Tran. A wavelet particle approximation for McKean-Vlasov and 2D-Navier-Stokes statistical solutions. *Stochastic Process. Appl.*, 118(2):284–318, 2008.

[TV03] D. Talay and O. Vaillant. A stochastic particle method with random weights for the computation of statistical solutions of McKean-Vlasov equations. *Ann. Appl. Probab.*, 13(1):140–180, 2003.