Alma Mater Studiorum · Università di Bologna

DOTTORATO DI RICERCA IN MATEMATICA
Ciclo: XXXIII

Settore concorsuale: 01/A3
Settore Scientifico Disciplinare: MAT/06

The parametrix method for SPDEs and conditional transition densities

Presentata da: Antonello Pesce

Coordinatore Didattico:
Chiar.mo Prof.
Fabrizio Caselli

Supervisore:
Chiar.mo Prof.
Andrea Pascucci

Esame finale anno 2021

Introduction

This thesis is made up of two, quite distinct parts. The first part fits in the stream of literature on the theory of SPDEs in Hölder spaces; the second one on the other hand concerns the study of density estimates for diffusion processes with unbounded drift. Although the two themes may seem far from each other, we shall see that a large class of SPDEs is still heavily connected with the theory of diffusions, as they describe the evolution of the conditioned distributions of some underlying partially observable, finite dimensional processes. On the other hand, some key techniques and strategies used throughout the study are also shared between the two parts; the most prominent example is the parametrix method, which we could present both in its analytic and in its more probabilistic interpretations in a single work.

Part I: The parametrix method for evolution SPDEs

Interest in Stochastic Partial Differential Equations (SPDEs) of evolution type began to arise in the early seventies, driven by the demand from modern applications and advances in natural sciences. Indeed this kind of equations, which naturally generalize both ordinary stochastic equations (SDEs) and deterministic evolution PDEs, are well suited to model any kind of stochastic influence in nature or man-made complex systems. The most notable examples are turbolent flows in fluid dynamics, diffusion and waves propagation in random media, as well as population growth, among many others (see, for instance [4], [10], [30], [12], [21]).

Another relevant source of SPDEs is provided by the study of stochastic flows defined by ordinary SDEs. The following result is due to Krylov ([33], [40]).

Example 0.0.1 (Backward diffusion equation)**.** *Consider the diffusion defined by the SDE*

$$dX_s^{t,x} = b(X_s^{t,x})ds + \sigma(X_s^{t,x})dB_s, \quad s > t,\ X_t^{t,x} = x \in \mathbb{R},$$

where B_s is a Brownian motion. Then, as a function of t and x, the process $X_s^{t,x}$ is a solution to *the backward SPDE*

$$\begin{aligned}
-dX_s^{t,x} &= \left(\sigma^2(X_s^{t,x})\partial_x^2 X_s^{t,x} + b(X_s^{t,x})\partial_x X_s^{t,x}\right) dt + b(X_s^{t,x})\partial_x X_s^{t,x} \star dB_t, \quad t < s,\\
X_s^{s,x} &= s \in \mathbb{R},
\end{aligned}$$

where $\star dB_t$ *denotes the backward Itô integral.*

As we shall see, the backward diffusion equation also comes in play in the study of *conditional distributions* of finite dimensional processes. Here stochastic (ultra)parabolic type equations appear naturally in the form of *filtering equations.*

Example 0.0.2 (Filtering equation)**.** *The filtering problem, in his most simple formulation, consists in estimating a certain 'signal', by observing it when it is mixed with a noise. Suppose that the signal* X *is modelled by a diffusion*

$$dX_s = b(X_s)ds + \sigma(X_s)dB_s, \quad s > 0,\ X_0 = x \in \mathbb{R},$$

and we are given an observation with dynamics

$$dY_s = h(X_s)ds + dW_s, \tag{0.0.1}$$

where W_s *is a different Brownian motion than* B_s*: for instance* Y_s *may describe the position of a moving object on the basis of a GPS observation,* W_s *the measurement error, and the signal* X_s *the true coordinates of the object. If* B_s *and* W_s *are independent, then the function of the paths* $\{Y_s,\ 0 \leqslant s \leqslant T\}$ *which best approximates, in the least squared sense, a quantity* $f(X_T)$*, is given by*

$$E\left[f(X_T) \mid \sigma(Y_s,\ 0 \leqslant s \leqslant T)\right] = \int_{\mathbb{R}} \mathbf{\Gamma}(0,x,T,\xi)f(\xi)d\xi,$$

where $\mathbf{\Gamma}$ *is the normalized stochastic fundamental solution of the SPDE*

$$dp_s(\xi) = \left(\frac{1}{2}\partial_{\xi\xi}\left(\sigma^2(\xi)p_s(\xi)\right) - \partial_\xi\left(b(\xi)p_s(\xi)\right)\right) ds + h(\xi)p_s(\xi)dY_s, \quad s \geqslant 0,$$

which generalizes the classic Fokker-Plank equation of X_s*.*

Filtering models may encompass much more general situations as well. For instance, while in (0.0.2) the variance of the noise W_s is equal to s, in applications it is much more likely that the scattering of the observed process is dependent on its position; the noises could be possibly correlated, and the dynamics of the signal could be affected by the observations Y_s as well. As a matter of fact, filtering models provide a large and relevant class of evolution SPDEs that can be written in the form

$$du_s(\xi) = \mathcal{L}u_s(\xi)ds + \sum_{k=1}^{d_1} \mathcal{G}^k u_s(\xi)dW_s^k, \quad s \geqslant 0,\ \xi \in \mathbb{R}^d, \tag{0.0.2}$$

where $\mathcal{L}$ is a (possibly degenerate) second order operator, $(\mathcal{G}^k)_{k=1,\cdots d_1}$ are first order operators and $W_s = (W_s^1, \cdots, W_s^{d_1})$ is a multi-dimensional Brownian motion on some probability space; importantly, since the coefficients of these equations may depend on the observation process Y, they are generally assumed to be random and only measurable in the time variable. Equations (0.0.3) in Hölder classes will be the main subject of this study.

The Cauchy problem for evolution SPDEs has been studied by several authors. Under coercivity conditions analogous to uniform ellipticity for PDEs, there exists a complete theory in Sobolev spaces (see e.g. [62] and the references therein) and in the spaces of Bessel potentials ([34], [35]). Classical solutions in Hölder classes were first considered in [61], [64] and more recent results were proved in [11] and [50], but in these cases the authors only considered equations with non-random coefficients and with no derivatives of the unknown function in the stochastic term. As we will explain in detail in Section 1.2 these restrictions are ultimately needed to recover a Duhamel principle, which does not hold in the stochastic framework in general. On the other hand, because of what we said above, it is worth to consider equations with random coefficients; moreover the filtering equation can include derivatives of the unknown function in the operators $\mathcal{G}$ even in very simple models.

In the last decades, the use of analytical or PDE techniques in the study of SPDEs has become widespread. For instance, the results in [11], [50], [19], [73] are based on classical methods of deterministic PDEs, such as the Duhamel principle and a priori Schauder estimates; the L^p estimates in [15] are proved by adapting the classical Moser's iterative argument; [66] provides short-time asymptotics of random heat kernels. A further remarkable example is given by the recent series of papers by Krylov [36, 37, 38] where the Hörmander's theorem for SPDEs is proved; see also the very recent results in [60] for backward SPDEs.

In this thesis we aim at extending another classical tool that, to the best of our knowledge, has not yet been considered in the study of SPDEs, the well-known *parametrix method* for the construction of the fundamental solution of PDEs with Hölder continuous coefficients.

In Chapter 1 we begin investigating the possibility to use a parametrix based method to prove existence and estimates of the fundamental solution to a parabolic SPDE. More precisely we consider

$$\begin{aligned} \mathcal{L}u_s(\xi) &= \frac{1}{2}a_s(\xi)\partial_{\xi\xi}u_s(\xi) + b_s(\xi)\partial_\xi u_s(\xi) + c_s(\xi)u_s(\xi), \qquad (0.0.3)\\ \mathcal{G}^k u_s(\xi) &= \sigma_s^k(\xi)\partial_\xi u_s(\xi), \end{aligned}$$

under the coercivity assumption

$$a_s - \sigma_s\sigma_s^* \geqslant \lambda > 0, \quad s > 0.$$

The lack of the Duhamel principle and the roughness of the coefficients we already mentioned constitute the two main obstacles that one faces when trying to apply the parametrix method to SPDEs. Specifically the Duhamel principle is the core of the usual parametrix iterative procedure. We propose to use the Itô-Wentzell formula to make a random change of variables and transform the SPDE to a PDE with random coefficients; the latter admits a Duhamel principle and we use it to extend the parametrix method to parabolic PDEs with measurable coefficients in the time variable. Importantly, this approach allows the operators in (0.0.4) to sport random (*stochastic Hölder*) coefficients, which compensate the extra-regularity required in the coefficients of $\mathcal{G}^k$ to allow the change of variables.

In Chapter 2 we examine a stochastic version of the degenerate Fokker-Plank equation

$$\mathbf{B}u_s(\xi,\nu) = a_s(\xi,\nu)\partial_{\nu\nu}u_s(\xi,\nu), \quad \mathbf{B} = \partial_s + \nu\partial_\xi,$$

which is characterized by a linear (unbounded) drift and it is a standard example of equation satisfying the *weak Hörmander condition.* As far as we are aware this is a novely in the context of SPDEs. Here we anticipate that, compared to the uniformly parabolic case, as well as the deterministic degenerate case, two main new difficulties arise in the analysis: the Itô-Wentzell transform drastically affects the drift $\mathbf{B}$ which will no longer have polynomial coefficients after the change of variables; moreover, again the roughness in time prevents the use of the so called *intrinsic Hölder spaces*, which would be more natural in the study of the singular kernels that come into play in the parametrix procedure (see [58], [16]).

Finally, in Chapter 3 we will show how the results on the Fokker-Plank SPDE allow to directly derive the filtering equations for a system of SDEs of Langevin type, both in their forward and backward formulations, without resorting to the general results from filtering theory. Here we follow the approaches recently proposed by Krylov and Zatezalo [42] and Veretennikov [70]. Again, as far as we are aware, this kind of problem was never considered in the literature, possibly because the known results for hypoelliptic SPDEs don't apply in this case.

Many of the results presented in Chapters one, two and three are taken from our works [57], [55], [56] with A. Pascucci.

Part II: Brownian SDEs with unbounded measurable drift

Consider the following diffusion

$$dX_s = b(s,X_s)ds + \sigma(s,X_s)dW_s, \quad s \geqslant 0,\ X_0 = x \in \mathbb{R}^d. \tag{0.0.4}$$

When both coefficients b,σ are *bounded* and Hölder continuous and σ is separated from zero (non-degeneracy condition), it is well known that there exists a unique weak solution to (0.0.5) which admits a density (see for instance [67], [25], [9]), that is for all $A \in \mathcal{B}(\mathbb{R}^d)$ (Borel σ-field of $\mathbb{R}^d$),

$$P\left(X_s \in A | X_0 = x\right) = \int_A \Gamma(0, x; s, \xi) d\xi.$$

Furthermore, it can be proved by the parametrix method that the transition density $\Gamma(0, x, s, \xi)$ enjoys the following two sided Gaussian estimates on a compact set in time:

$$C^{-1}\Gamma_{\mu^{-1}}(s, x-\xi) \leqslant \Gamma(0, x; s, \xi) \leqslant C\Gamma_{\mu}(s, x-\xi) \tag{0.0.5}$$

as well as the following gradient estimate

$$|\nabla_x^j \Gamma(0, x; s, \xi)| \leqslant C s^{-\frac{j}{2}} \Gamma_{\mu}(s, x-\xi), \quad j = 1, 2,$$

where

$$\Gamma_{\mu}(t, x) := t^{-\frac{d}{2}} \exp\left(-\mu |x|^2/t \right), \quad \mu \in (0, 1], t > 0, \tag{0.0.6}$$

and the constants $\mu \in (0, 1]$, $C \geqslant 1$ only depend on the regularity of the coefficients, the non-degeneracy constants of the diffusion coefficients, the dimension d, and for the constant C, on the maximal time considered (see [24] and [1], [2]). Such methods have been successfully applied to derive upper bounds up to the second order derivative for more general cases, such as operators satisfying a strong Hörmander condition (see [5]), Kolmogorov operators with linear drift (see [58] and [16]), as well as the SPDEs in Chapters one and two. A different approach consists in viewing a logarithmic transformation of Γ as the value function of a certain stochastic control problem, as proposed by Fleming and Sheu in [22]: this idea allows then to get the desired density estimates by choosing appropriate controls and eventually an upper bound for the *logarithmic gradient* (see [63]).

When the drift is unbounded and non-linear fewer results are available. In fact, in this case it is no longer expected that the two sided estimates as given in (0.0.6) hold.

Example 0.0.3. *The following Ornstein-Uhlenbeck (OU)-process*

$$dX_s = X_s ds + dW_s, \quad X_0 = x,$$

has, with the notations of (0.0.7)*, the non-spatial homeogenous density*

$$\Gamma_{\mathrm{OU}}(0, x; s, \xi) = (\pi(\varepsilon^{2s} - 1))^{-d/2} \Gamma_{s/(\varepsilon^{2s}-1)}(t, \varepsilon^s x - \xi).$$

In [14], Delarue and Menozzi derive two sided density bounds for a class of degenerate operators with *unbounded and Lipschitz drift*, satisfying a weak Hörmander condition, by combining the two previous approaches: parametrix and logarithmic transform. Indeed, when the drift is unbounded it becomes difficult to get good controls for the iterated kernels in the parametrix expansion. In our non-degenerate parabolic setting those bounds still hold provided the drift is *globally* Lipschitz continuous in space. Then, they read as:

$$C^{-1}\Gamma_{\mu^{-1}}(s,\gamma_s(x)-\xi) \leqslant \Gamma(0,x;s,\xi) \leqslant C\Gamma_{\mu}(s,\gamma_s(x)-\xi), \tag{0.0.7}$$

where γ stands for the deterministic flow associated with the drift, that is

$$\dot{\gamma}_s(x) = b(s,\gamma_s(x)), \quad s \geqslant 0,\, \gamma_0(x) = x$$

and $C,\mu > 0$ both depend on the maximal time considered. This means that the diffusion starting from x, oscillates around $\gamma_s(x)$ at time s with fluctuations of order $s^{-\frac{1}{2}}$. Notice that if b is bounded, then (0.0.8) reduces to (0.0.6) since

$$s^{-\frac{1}{2}}|x-\xi| - \|b\|_\infty s^{\frac{1}{2}} \leqslant s^{-\frac{1}{2}}|\gamma_s(x)-\xi| \leqslant s^{-\frac{1}{2}}|x-\xi| + \|b\|_\infty s^{\frac{1}{2}}.$$

Hence, taking or not into consideration the flow does not give much additional information. The above control also clearly emphasizes why C might depend on some maximal time interval considered. In the case where b is bounded but not necessarily *smooth*, the above bounds remain valid for *any* regularizing flow.

Diffusion with dynamics (0.0.5) and unbounded drifts appear in many applicative fields. We can mention for instance the work [26] which concerned issues related to statistics of diffusions and also [52] for the numerical approximation of ergodic diffusions.

In these frameworks, estimates on the density and its derivatives are naturally required. Some gradient estimates of the density were established in [26]. The approach developed therein relies on the Malliavin calculus and thus required some extra regularity on the drift. Also, since the deterministic flow was not taken into consideration, an additional penalizing exponential term in the right hand side of the bounds appeared. Similar features appeared in the work [13] which established the existence of fundamental solutions for a strictly sublinear Hölder continuous drift.

In Chapter 2 we obtain some estimates for the derivatives in the non-degenerate direction for a Kolmogorov equation with Lipschitz drift that appears after the change of variables in the SPDE: these controls reflect both the singularities associated with the differentiation, as in equation (0.0.6) above, and also reflect the key importance of the flow for unbounded drifts as it appears in the two-sided heat kernel estimate (0.0.8). The analysis here builds on the work of Delarue and Menozzi: specifically, we remark that the density lower bound allows to recover, a posteriori, the good controls on the iterated kernels which allow to pursuit the usual parametrix procedure.

In Chapter 4 we concentrate only on the non-degenerate case but we develop a new approach to the derivation of these estimates, based on a circular argument: the point of this work is to be completely self-contained, to provide estimates both in the forward and in the backward variables under minimal regularity assumptions, and to be sufficiently robust to be generalized, as soon as some suitable two sided bounds hold. We can actually address various frameworks. We manage to obtain two-sided heat kernel bounds for a Hölder continuous in space diffusion coefficient σ in (0.0.5) and a drift b which is uniformly bounded in time at the origin and has linear growth in space. Importantly, when the drift b is itself not smooth, the heat kernel bounds can be stated in the form (0.0.8) for any flow associated with a mollification of b. In particular, if the drift is continuous in space they actually hold for *any* Peano flow. These conditions are also sufficient to obtain gradient bounds with respect to the *backward* variable x. To derive controls for the second order derivatives with respect to x, an additional spatial Hölder continuity assumption naturally appears for the drift. Eventually, imposing some additional spatial smoothness on the diffusion coefficient, we also succeed in establishing a gradient bound with respect to the *forward* variable y.

To the best of our knowledge, ours are among the first results for derivatives of heat kernels with unbounded drifts.

The results presented in Chapter 4 are part of our work ([49]) (with S. Menozzi and X. Zhang).

Some general notations

In the following analysis, the main settings will substantially change from chapter to chapter, making it difficult to keep a consistent notation throughout. Nonetheless we try to follow as much as possible some general guidelines.

The time variables are denoted with t or s, where t usually stands for the *initial time* and s stands for the *final time* when it matters; the spatial variables in $\mathbb{R}^N$ are denoted with $x = (x_1, \cdots x_N)$ and $\xi = (\xi_1, \cdots \xi_N)$, with x usually standing for the *initial point* and ξ for the *final point*; similarly the spatial points in $\mathbb{R}^{N+1}$ are denoted with $z = (x, v)$ and $\zeta = (\xi, \nu)$, and share the analogous convention. Moreover, as a general rule, when a quantity depends on both an *initial state* and a *final state*, the variables which describe the initial state are always appended first, regardless of whether they may act as the *pole* or not: in particular this is the case when denoting deterministic or stochastic flows, conditioned or unconditioned densities, deterministic or stochastic fundamental solutions.

We use the notation ∇, ∇^2 to denote respectively the gradient and Hessian matrix with respect to the spatial variable; by extension, we denote by ∇^j the j^{th} order derivative. When required by the context, we may use the notation ∂_i for the i^{th} partial derivative and $\partial_{ij} = \partial_i \partial_j$, as well as the multi-index notation $\partial^{\beta} = \partial_1^{\beta_1} \cdots \partial_N^{\beta_N}$, with $\beta \in \mathbb{N}_0^N$.

Throughout the work, the summation convention over repeated indices is enforced regardless of whether they stand at the same level or at different ones; the letter C usually stands for a positive constant, only dependent on the quantities in the assumptions; in the context of the proofs, its value may update from line to line. Other possible dependences are explicitly indicated when needed.

Next, we introduce the general functional setting to be used throughout the study. Let $k \in \mathbb{N}$, $\alpha \in (0,1)$ and $0 \leqslant t < T$. Denote by $m\mathcal{B}_{t,T}$ (resp. $b\mathcal{B}_{t,T}$) the space of all real-valued (resp. bounded) Borel measurable functions $f = f(s,x)$ on $[t,T] \times \mathbb{R}^d$ and

- $C^0_{t,T}$ (resp. $C^0_{t,T}$) is the space of functions $f \in m\mathcal{B}_{t,T}$ (resp. $f \in b\mathcal{B}_{t,T}$) that are continuous in x;

- $C^{\alpha}_{t,T}$ (resp. $bC^{\alpha}_{t,T}$) is the space of functions $f \in m\mathcal{B}_{t,T}$ (resp. $f \in b\mathcal{B}_{t,T}$) that are α-Hölder continuous in x uniformly with respect to $s \in [t,T]$, that is
$$\sup_{\substack{s \in [t,T] \\ x \neq \xi}} \frac{|f(s,x) - f(s,\xi)|}{|x-\xi|^{\alpha}} < \infty;$$
We also denote by $C^{0,1}_{t,T}$ the space of functions $f \in m\mathcal{B}_{t,T}$ that are Lipschitz continuous in x uniformly with respect to $s \in [t,T]$;

- $C^{k}_{t,T}$ (resp. $bC^{k}_{t,T}$) is the space of functions $f \in m\mathcal{B}_{t,T}$ (resp. $f \in b\mathcal{B}_{t,T}$) that are k-times differentiable with respect to x with derivatives in $C^0_{t,T}$ (resp. $bC^0_{t,T}$);

- $C^{k+\alpha}_{t,T}$ (resp. $bC^{k+\alpha}_{t,T}$) is the space of functions $f \in m\mathcal{B}_{t,T}$ that are k-times differentiable in with respect to x with derivatives in $C^{\alpha}_{t,T}$ (resp. $bC^{\alpha}_{t,T}$)

As a general rule, a random field $u = u(s,\xi,\omega)$ on $[0,\infty) \times \mathbb{R}^d \times \Omega$ is denoted by $u_s(\xi)$ and we shall systematically omit the explicit dependence on $\omega \in \Omega$; we keep the notation $u(s,\xi)$ for deterministic functions on $[0,\infty) \times \mathbb{R}^d$.

Contents

Chapter 1

The parametrix method for parabolic SPDEs

1.1 Introduction

In this chapter, we introduce the fundamental tools and ideas that constitute the starting point of our analysis: here we set the stochastic Hölder spaces which will be used throughout the dissertation and the notion of stochastic fundamental solution; we recall the Itô-Wentzell formula and establish the time-dependent parametrix for a parabolic SPDE.

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space with an increasing filtration $(\mathcal{F}_t)_{t\geqslant 0}$ of complete with respect to $(\mathcal{F}, P)$ σ-fields $\mathcal{F}_t \subseteq \mathcal{F}$. Let $d_1 \in \mathbb{N}$ and let W^k, $k = 1, \cdots, d_1$, be one-dimensional independent Wiener processes with respect to $(\mathcal{F}_t)_{t\geqslant 0}$. We consider the parabolic SPDE

$$du_s(\xi) = (\mathcal{L}_s u_s(\xi) + f_s(\xi))\, ds + \left(\mathcal{G}_s^k u_s(\xi) + g_s^k(\xi)\right) dW_s^k, \qquad \xi \in \mathbb{R}^d \tag{1.1.1}$$

where $\mathcal{L}_s$ is the second-order operator

$$\mathcal{L}_s u_s(\xi) = \frac{1}{2} a_s^{ij}(\xi)\partial_{ij} u_s(\xi) + b_s^j(\xi)\partial_j u_s(\xi) + c_s(\xi) u_s(\xi)$$

and $\mathcal{G}_s^k$ is the first-order operator

$$\mathcal{G}_s^k u_s(\xi) = \sigma_s^{ik}(\xi)\partial_i u_s(\xi).$$

The coefficients a_s, b_s, σ_s^k, c_s and f_s are intended to be random and not smooth.

In the remaining part of this section we introduce the functional setting, set the assumptions and state the main result on the SPDE (1.1.1), Theorem 1.1.5; for illustrative purposes, the particular case of the stochastic heat equation is discussed in Section 1.2. In Section 1.3 we recall the Itô-Wentzell formula and provide some estimates for the related flow of diffeomorphisms. In Section 1.4 we present the parametrix method. Since the complete proofs are rather technical and to a large extent similar to the classical case, we only provide the details on those aspects that require significant modifications: in particular, in Section 1.4.3 we present a proof of the Gaussian lower bound for the fundamental solution which requires some non trivial adaptation of an original argument by Aronson (cf. [20]).

1.1.1 Functional setting and main results

Let $k \in \mathbb{N} \cup \{0\}$, $\alpha \in [0,1)$ and $\mathcal{P}_{t,T}$ be the predictable σ-algebra on $[t,T] \times \Omega$. We denote by $\mathbf{C}^{k+\alpha}_{t,T}$ the family of functions $f = f_s(x,\omega)$ on $[t,T] \times \mathbb{R}^d \times \Omega$ such that:

i) $(s,x) \mapsto f_s(x,\omega) \in C^{k+\alpha}_{t,T}$ P-a.s.;

ii) $(s,\omega) \mapsto f_s(x,\omega)$ is $\mathcal{P}_{t,T}$-measurable for any $x \in \mathbb{R}^d$.

Moreover, $\mathbf{bC}^{k+\alpha}_{t,T}$ is the space of functions $f \in \mathbf{C}^{k+\alpha}_{t,T}$ such that

$$\sum_{|\beta| \leqslant k} \sup_{\substack{s \in [t,T] \\ x \in \mathbb{R}^d}} |\partial^\beta f_s(x)| < \infty \qquad P\text{-a.s.}$$

We say that $f = f_s(x)$ is *non-rapidly increasing uniformly on* $(t,T] \times \mathbb{R}^d$ if, for any $\delta > 0$, $e^{-\delta|x|^2} |f_s(x)|$ is a bounded function on $(t,T] \times \mathbb{R}^d$, P-a.s.; in case f does not depend on s, we simply say that f is non-rapidly increasing on $\mathbb{R}^d$.

Definition 1.1.1. *A stochastic fundamental solution* $\mathbf{\Gamma} = \mathbf{\Gamma}(t,x,s,\xi)$ *for the SPDE* (1.1.1) *is a function defined for* $0 \leqslant t < s \leqslant T$ *and* $x, \xi \in \mathbb{R}^d$*, such that for any* $(t,x) \in [0,T) \times \mathbb{R}^d$ *we have:*

i) $\mathbf{\Gamma}(t,x;\cdot,\cdot) \in \mathbf{C}^2_{t_0,T}(\mathbb{R}^d)$ *and with probability one satisfies*

$$\mathbf{\Gamma}(t,x;s,\xi) = \mathbf{\Gamma}(t,x;t_0,\xi) + \int_{t_0}^s \mathcal{L}_\tau \mathbf{\Gamma}(t,x;\tau,\xi) d\tau + \int_{t_0}^s \mathcal{G}^k_\tau \mathbf{\Gamma}(t,x;\tau,\xi) dW^k_\tau \tag{1.1.2}$$

for $t < t_0 \leqslant s \leqslant T$ *and* $\xi \in \mathbb{R}^d$*;*

ii) for any continuous and non-rapidly increasing function φ *on* $\mathbb{R}^d$ *and* $x_0 \in \mathbb{R}^d$

$$\lim_{\substack{(s,\xi) \to (t,x_0) \\ s > t}} \int_{\mathbb{R}^d} \varphi(x) \mathbf{\Gamma}(t,x;s,\xi) dx = \varphi(x_0), \qquad P\text{-a.s.}$$

Next we state the standing assumptions on the coefficients of the SPDE (1.1.1).

Assumption 1.1.2 (Regularity). *For some $\alpha \in (0,1)$ and for every $i,j = 1,\dots,d$ and $k = 1,\dots,d_1$, we have: $a^{ij} \in \mathbf{bC}^{\alpha}_{0,T}$, $\sigma^{ik} \in \mathbf{bC}^{3+\alpha}_{0,T}$ and b^j, $c \in \mathbf{bC}_{0,T}$.*

Assumption 1.1.3 (Coercivity). *Let*

$$\alpha_t(x) := \Big(a^{ij}_t(x) - \sigma^{ik}_t(x)\sigma^{jk}_t(x)\Big)_{i,j=1,\dots,d}.$$

There exists a positive random variable $\mathbf{m}$ *such that*

$$\langle \alpha_t(x)\xi,\xi\rangle \geqslant \mathbf{m}|\xi|^2, \qquad t \in [0,T],\ x,\xi \in \mathbb{R}^d,\ P\text{-a.s.}$$

We now introduce a random change of coordinates that will play a central role in the following analysis. We fix $(t,x) \in [0,T) \times \mathbb{R}^d$ and consider the stochastic ordinary differential equation

$$x_s = x - \int_t^s \sigma^k_\tau(x_\tau)dW^k_\tau, \qquad s \in [t,T]. \tag{1.1.3}$$

It is well-known (see, for instance, Theor. 4.6.5 in [44])) that, under Assumption 1.1.2, equation (1.1.3) admits a solution $\gamma^{\mathrm{IW}} = \gamma^{\mathrm{IW}}_{t,s}(x,\omega)$ that is a stochastic flow of diffeomorphisms: precisely, $\gamma^{\mathrm{IW}}_{t,s} \in \mathbf{C}^{3+\alpha'}_{t,T}$, for any $\alpha' < \alpha$, the matrix $\nabla\gamma^{\mathrm{IW}}_{t,s}(x)$ satisfies

$$\nabla\gamma^{\mathrm{IW}}_{t,s}(x) = I_d - \int_t^s \nabla\sigma^k_\tau(\gamma^{\mathrm{IW}}_{t,\tau}(x))\nabla\gamma^{\mathrm{IW}}_{t,\tau}(x)dW^k_\tau, \tag{1.1.4}$$

and, for any $i,j = 1,\dots,d$, $\partial^2_{ij}\gamma^{\mathrm{IW}}_{t,s}(x)$ satisfies

$$\begin{aligned}\partial^2_{ij}(\gamma^{\mathrm{IW}}_{t,s}(x))_h = -\int_t^s \Big[&(\nabla\sigma^k_\tau(\gamma^{\mathrm{IW}}_{t,\tau}(x))\partial^2_{ij}\gamma^{\mathrm{IW}}_{t,\tau}(x))_h \\ &+ \Big((\nabla\gamma^{\mathrm{IW}}_{t,\tau}(x))^*\nabla^2\sigma^{hk}_\tau(\gamma^{\mathrm{IW}}_{t,\tau}(x))\nabla\gamma^{\mathrm{IW}}_{t,\tau}(x)\Big)_{ij}\Big]dW^k_\tau\end{aligned} \tag{1.1.5}$$

with probability one.

Since we are going to use γ^{IW} as a global change of variables, we need some control over the stochastic integrals in (1.1.4) and (1.1.5) for x varying in $\mathbb{R}^d$: this issue is addressed in Section 1.3 (see, in particular, Proposition 1.3.2) under the following additional condition. For any suitably regular function $f = f(w) : \mathbb{R}^N \longrightarrow \mathbb{R}$, $\varepsilon > 0$ and multi-index $\beta \in \mathbb{N}^N_0$, we set

$$\{f\}_{\varepsilon,\beta} := \sup_{w\in\mathbb{R}^N}(1+|w|^2)^{\varepsilon}|\partial^\beta_w f(w)|. \tag{1.1.6}$$

Assumption 1.1.4. *There exist $\varepsilon > 0$ and a random variable $M \in L^{\bar p}(\Omega)$, with $\bar p > \max\left\{2, d, \frac{d}{2\varepsilon}\right\}$, such that, with probability one*

$$\sup_{t\in[0,T]}\{\sigma^k_t\}_{\varepsilon,\beta} \leqslant M, \qquad |\beta| = 1,2,3, \quad k = 1,\cdots,d_1.$$

Assumption 1.1.4 requires that $\sigma_t^k(x)$ flattens as $x \to \infty$: in particular, this condition is clearly satisfied if σ depends only on t or, more generally, if the spatial gradients of σ has compact support.

In order to state the main result of this section we introduce the following notation: let $\mathcal{C} = \left(\mathcal{C}^{ij}\right)_{1\leqslant i,j\leqslant d}$ be a constant, symmetric and positive definite matrix. We denote by

$$\Gamma^{\text{heat}}(\mathcal{C}, x) = \frac{1}{\sqrt{(2\pi)^d \det \mathcal{C}}} e^{-\frac{1}{2}\langle \mathcal{C}^{-1}x, x\rangle}, \qquad x \in \mathbb{R}^d, \tag{1.1.7}$$

the d-dimensional Gaussian kernel with covariance matrix $\mathcal{C}$. Clearly Γ^{heat} is a smooth function and satisfies

$$\partial_t \Gamma^{\text{heat}}(t\mathcal{C}, x) = \frac{1}{2}\text{tr}\left(\mathcal{C}\nabla^2\Gamma^{\text{heat}}(t\mathcal{C}, x)\right), \qquad t > 0,\ x \in \mathbb{R}^d.$$

The main result of the chapter is the following

Theorem 1.1.5. *Let Assumptions 1.1.2, 1.1.3 and 1.1.4 be in force. Then there exists a fundamental solution* $\mathbf{\Gamma}$ *for the SPDE* (1.1.1)*. Moreover, there exist two positive random variables* μ_1 *and* μ_2 *such that, with probability one we have*

$$\begin{aligned}
\mathbf{\Gamma}(t, x; s, \xi) &\geqslant \mu_2^{-1}\Gamma^{\text{heat}}\left(\mu_1^{-1}\mathcal{I}_{s-t}, \gamma_{t,s}^{\text{IW},-1}(\xi) - x\right), && (1.1.8)\\
\mathbf{\Gamma}(t, x; s, \xi) &\leqslant \mu_2\Gamma^{\text{heat}}\left(\mu_1\mathcal{I}_{s-t}, \gamma_{t,s}^{\text{IW},-1}(\xi) - x\right),\\
\left|\nabla_\xi\mathbf{\Gamma}(t, x; s, \xi)\right| &\leqslant \frac{\mu_2}{\sqrt{s-t}}\Gamma^{\text{heat}}\left(\mu_1\mathcal{I}_{s-t}, \gamma_{t,s}^{\text{IW},-1}(\xi) - x\right), && (1.1.9)\\
\left|\nabla_\xi^2\mathbf{\Gamma}(t, x; s, \xi)\right| &\leqslant \frac{\mu_2}{s-t}\Gamma^{\text{heat}}\left(\mu_1\mathcal{I}_{s-t}, \gamma_{t,s}^{\text{IW},-1}(\xi) - x\right),
\end{aligned}$$

for every $0 \leqslant t < s \leqslant T$ *and* $x, \xi \in \mathbb{R}^d$*, where* $\gamma_{t,s}^{\text{IW},-1}$ *is the inverse of the Itô-Wentzell stochastic flow* $\zeta \mapsto \gamma_{t,s}^{\text{IW}}(\zeta)$ *and* I_t *denotes the diagonal* $d \times d$ *matrix* $\text{diag}(t, \cdots, t)$.

The proof of Theorem 1.1.5 is postponed to Section 1.4.4.

Corollary 1.1.6. *Let* u_0 *be a* $\mathcal{F}_0\otimes\mathcal{B}$*-measurable function on* $\Omega\times\mathbb{R}^d$ *such that* $u_0(\omega, \cdot)$ *is continuous and non-rapidly increasing on* $\mathbb{R}^d$ *for a.e.* $\omega \in \Omega$*. Let* $f \in \mathbf{C}_{0,T}^{\bar{\alpha}}$*, for some* $\bar{\alpha} \in (0, 1)$*, be non-rapidly increasing uniformly on* $[0, T] \times \mathbb{R}^d$*. Then*

$$u_s(\xi) = \int_{\mathbb{R}^d} u_0(x)\mathbf{\Gamma}(0, x; s, \xi)dx + \int_0^s\int_{\mathbb{R}^d} f_\tau(x)\mathbf{\Gamma}(\tau, x; s, \xi)dxd\tau$$

is a classical solution of (1.1.1) *with initial value* u_0*, in the sense that* $u \in \mathbf{C}_{0,T}^2$ *and with probability one satisfies*

$$u_s(\xi) = u_0(\xi) + \int_0^s \left(\mathcal{L}_\tau u_\tau(\xi) + f_\tau(\xi)\right)d\tau + \int_0^s \mathcal{G}_\tau^k u_\tau(\xi)dW_\tau^k, \qquad s \in [0, T],\ \xi \in \mathbb{R}^d.$$

Such a solution is unique in the class of functions with quadratic exponential growth: precisely, u is the unique solution such that there exists a positive random variable C such that $|u_s(\xi)|\, e^{-C|\xi|^2}$ is bounded on $[0,T] \times \mathbb{R}^d$, P-a.s.

1.2 Stochastic heat equation and Duhamel principle

For illustrative purposes, in this section we consider the prototype case of the stochastic heat equation. We focus our attention on the Duhamel principle that is the crucial ingredient in the parametrix method for the construction of the fundamental solution. More generally, the Duhamel principle is a powerful tool for studying the existence and regularity properties of parabolic PDEs. In the framework of SPDEs of the form (1.1.1), it is still possible to have a Duhamel representation when the coefficients a^{ij} are deterministic and $\mathcal{G}_s^k$ are replaced by an operator of order zero: this case has been considered in [64] and [50] where the Cauchy problem for parabolic SPDEs is studied. For the general SPDE (1.1.1) however, as also noticed by other authors (see, for instance, Sowers [65], Sect.3), measurability issues arise that do not appear in the deterministic case.

To be more specific, let us consider the stochastic heat equation

$$du_s(\xi) = \frac{a^2}{2}\partial_{\xi\xi}u_s(\xi)ds + \left(\sigma\partial_\xi u_s(\xi) + g_s(\xi)\right)dW_s. \tag{1.2.1}$$

Under the *coercivity condition* $\alpha^2 := a^2 - \sigma^2 > 0$, the Gaussian kernel

$$p\,(t,x;s,\xi) := \frac{1}{\sqrt{2\pi\alpha^2(s-t)}}\exp\left(-\frac{(\xi+\sigma(W_s-W_t)-x)^2}{2\alpha^2(s-t)}\right), \qquad s > t \geqslant 0,\ x,\xi \in \mathbb{R}, \tag{1.2.2}$$

is well defined, and if $\sigma = 0$ or $g \equiv 0$ then the function

$$u_s(\xi) := \int_{\mathbb{R}} u_0(x)p\,(0,x;s,\xi)\,dx + \int_0^s \int_{\mathbb{R}} g_\tau(x)p\,(\tau,x;s,\xi)\,dx\,dW_\tau \tag{1.2.3}$$

is a classical solution to (1.2.1), for any suitable initial value u_0. This follows directly from the Itô formula and the fact that the change of variable

$$\hat{u}_s(\xi) := u_s(\xi - \sigma W_s)$$

transforms the homogeneous version of (1.2.1) into the *deterministic* heat equation

$$d\hat{u}_s(\xi) = \frac{\alpha^2}{2}\partial_{\xi\xi}\hat{u}_s(\xi)ds.$$

The difficulty in considering the case when σ and g are *both* not null, comes from the fact that the integrand $g_\tau(x)p\,(\tau,x;s,\xi)$ in (1.2.3) becomes measurable with respect to the *future* σ-algebra $\mathcal{F}_s$ in the filtered space: thus in general the last integral in (1.2.3) is not well-defined in the framework of classical Itô-based stochastic calculus. For this reason, in the context of SPDEs, the Duhamel principle has been used only under rather specific assumptions.

We observe that a naive application of the parametrix method for SPDE (1.1.1) would consist precisely of a successive application of the Duhamel formula (1.2.3) with g and $\sigma = \sigma_s(\xi)$ that are not null and not even constant. Hence, the lack of a general Duhamel formula seems to preclude a direct use of the whole parametrix approach.

Incidentally formula (1.2.2) shows that, even for SPDEs with constant coefficients, the stochastic fundamental solution p has distinctive properties compared to the Gaussian deterministic heat kernel. In particular, the asymptotic behaviour near the pole of p is affected by the presence of the Brownian motion: this fact was studied also in [65] in the more general framework of Riemannian manifolds and is coherent with the Gaussian lower and upper bounds (1.1.8).

1.3 Itô-Wentzell change of coordinates

In this section we consider the random change of coordinates (1.1.3) and use the Itô-Wentzell formula to transform the SPDE (1.1.1) into a PDE with random coefficients. For simplicity, we only consider the case $t = 0$ and set $\gamma_s(\xi) \equiv \gamma^{\mathrm{IW}}_{0,s}(\xi)$. We define the operation "hat" which transforms any function $u_s(\xi)$ into

$$\hat{u}_s(\xi) = u_s(\gamma_s(\xi)) \tag{1.3.1}$$

and recall the classical Itô-Wentzell formula (see, for instance, Theor. 1.17 in [62] or Theor. 3.3.1 in [44]).

Theorem 1.3.1 (Itô-Wentzell)**.** *Let $u \in \mathbf{C}^2_{0,T}$, $h \in \mathbf{C}^0_{0,T}$ and $g^k \in \mathbf{C}^1_{0,T}$ be such that*

$$du_s(\xi) = h_s(\xi)ds + g^k_s(\xi)dW^k_s. \tag{1.3.2}$$

Then we have

$$d\hat{u}_s(\xi) = \left(\hat{h}_s(\xi) + \frac{1}{2}\widehat{\sigma^{ik}_t\sigma^{jk}_s}(\xi)\widehat{\partial_{ij}u_s}(\xi) - \widehat{\partial_i g^k_s}(x)\hat{\sigma}^{ik}_s(\xi)\right)ds + \left(\hat{g}^k_s(\xi) - \widehat{\mathcal{G}^k_s u_s}(\xi)\right)dW^k_s. \tag{1.3.3}$$

In order to apply Itô-Wentzell formula to our SPDE, we prove the following crucial estimate for the gradient of $\gamma_s(\xi)$.

Proposition 1.3.2. *Let*

$$Y_s := (\nabla\gamma_s)^{-1}.$$

We have $\nabla\gamma$, $Y \in \mathbf{bC}^1_{0,T}$ and there exists a positive random variable $\widetilde{\mathbf{m}}$ such that

$$|Y^*_s(\xi)x|^2 \geqslant \widetilde{\mathbf{m}}|x|^2, \qquad s \in [0,T],\ x,\xi \in \mathbb{R}^d,\ P\text{-}a.s. \tag{1.3.4}$$

The proof of Proposition 1.3.2 is based on the following preliminary lemma:

Lemma 1.3.3. *Let Z be a continuous random field defined on $[t,T]\times\mathbb{R}^d$. Assume that for some $\varepsilon>0$ and $p>\left(d\vee\frac{d}{2\varepsilon}\right)$ there exists a constant $C>0$ such that*

$$E\left[\sup_{s\in[t,T]}|Z_s(x)|^p\right]\leqslant C(1+|x|^2)^{-\varepsilon p},\tag{1.3.5}$$

$$E\left[\sup_{s\in[t,T]}|\nabla Z_s(x)|^p\right]\leqslant C(1+|x|^2)^{-\varepsilon p},\tag{1.3.6}$$

for every $x\in\mathbb{R}^d$. Then Z has a modification in $\mathbf{bC}_{t,T}^{1-\frac{d}{p}}$.

Proof. By the classical Sobolev embedding theorem, for every $f\in W^{1,p}(\mathbb{R}^d)$, with $p>d$, we have

$$|f(x)|+\frac{|f(x)-f(y)|}{|x-y|^{1-\frac{d}{p}}}\leqslant N\|f\|_{W^{1,p}(\mathbb{R}^d)},\qquad\text{a.e. } x,y\in\mathbb{R}^d,$$

where N is a constant dependent only on p and d. Hence the statement directly follows from the following estimate

$$\sup_{s\in[t,T]}\|Z_s\|_{W^{1,p}(\mathbb{R}^d)}<\infty\quad P\text{-a.e.}$$

and the continuity of Z. To this end, we check that

$$E\left[\sup_{s\in[t,T]}\|Z_s\|_{W^{1,p}(\mathbb{R}^d)}\right]<\infty.$$

By (1.3.5) and since $p>\frac{d}{2\varepsilon}$, we have

$$E\left[\sup_{s\in[t,T]}\|Z_s\|^p_{L^p(\mathbb{R}^d)}\right]\leqslant E\left[\int_{\mathbb{R}^d}\sup_{s\in[t,T]}|Z_s(x)|^p dx\right]\leqslant\int_{\mathbb{R}^d}C(1+|x|^2)^{-\varepsilon p}dx<\infty,$$

and analogously by (1.3.6) we have

$$E\left[\sup_{s\in[t,T]}\|\nabla Z_s\|^p_{L^p(\mathbb{R}^d)}\right]\leqslant\int_{\mathbb{R}^d}C(1+|x|^2)^{-\varepsilon p}dx<\infty.$$

□

Proof of Proposition 1.3.2. Let

$$Z_s(\xi):=\nabla\gamma_s(\xi)-I=\int_0^s\nabla\sigma_\tau^k(\gamma_\tau(\xi))\nabla\gamma_\tau(\xi)dW_\tau^k.\tag{1.3.7}$$

We show that the matrix-valued random field $Z_t(x)$ satisfies estimates (1.3.5) and (1.3.6) of Lemma 1.3.3 for every p such that $\left(2 \vee d \vee \frac{d}{2\varepsilon}\right) < p < \bar{p}$, with ε and $\bar{p}$ as in Assumption 1.1.4. Indeed, by the well-known L^p-estimates for $\gamma_s(\xi)$ (see [44], Chapter 4), for any $0 \leqslant s \leqslant T$ and $x \in \mathbb{R}^d$ we have

$$E\left[(1 + |\gamma_s(\xi)|^2)^p\right] \leqslant N_1(1 + |\xi|^2)^p, \qquad p \in \mathbb{R}, \tag{1.3.8}$$

$$E\left[\left|\nabla^j \gamma_s(\xi)\right|^p\right] \leqslant N_2, \qquad p \geqslant 2,\ 1 \leqslant j \leqslant 3, \tag{1.3.9}$$

where the constants N_1 and N_2 depend only on p and d. We have

$$E\left[\sup_{s\in[0,T]} \left|Z_s^{ij}(\xi)\right|^p\right] \leqslant C \sum_{k=1}^{d_1} \sum_{h=1}^{d} E\left[\sup_{s\in[0,T]} \left|\int_0^s \partial_h \sigma_\tau^{ik}(\gamma_\tau(\xi)) \partial_j \gamma_\tau^h(\xi) dW_\tau^k\right|^p\right]$$

(by Burkolder inequality)

$$\leqslant C_p' \sum_{k=1}^{d_1} \sum_{h=1}^{d} E\left[\left(\int_0^T \left(\partial_h \sigma_\tau^{ik}(\gamma_\tau(\xi)) \partial_j \gamma_\tau^h(\xi)\right)^2 d\tau\right)^{\frac{p}{2}}\right]$$

(by Hölder inequality with conjugate exponents $\frac{p}{2}$ and $\frac{p}{p-2}$)

$$\leqslant C_p' T^{\frac{p-2}{2}} \sum_{k=1}^{d_1} \sum_{h=1}^{d} \int_0^T E\left[\left|\partial_h \sigma_\tau^{ik}(\gamma_\tau(\xi)) \partial_j X_\tau^h(\xi)\right|^p\right] d\tau$$

(by Hölder inequality with conjugate exponents r and $q < \frac{\bar{p}}{p}$)

$$\leqslant C_p' T^{\frac{p-2}{2}} \sum_{k=1}^{d_1} \sum_{h=1}^{d} \int_0^T E\left[\left|\partial_h \sigma_\tau^{ik}(\gamma_\tau(\xi))\right|^{pq}\right]^{\frac{1}{q}} E\left[\left|\partial_j \gamma_\tau^h(\xi)\right|^{pr}\right]^{\frac{1}{r}} ds$$

(by the flattening Assumption 1.1.4 and estimate (1.3.9))

$$\leqslant C_p'' T^{\frac{p-2}{2}} N_2^{\frac{1}{r}} \int_0^T E\left[M^{pq}(1 + |\gamma_\tau(\xi)|^2)^{-\varepsilon pq}\right]^{\frac{1}{q}} ds$$

(by Hölder inequality with conjugate exponents $\bar{r}$ and $\bar{q} := \frac{\bar{p}}{pq} > 1$)

$$\leqslant C_p'' T^{\frac{p-2}{2}} N_2^{\frac{1}{r}} \int_0^T E\left[M^{\bar{p}}\right]^{\frac{p}{\bar{p}}} E\left[(1 + |\gamma_\tau(\xi)|^2)^{-\varepsilon pq\bar{r}}\right]^{\frac{1}{q\bar{r}}} ds$$

(by estimate (1.3.8))

$$\leqslant C_p'' T^{\frac{p}{2}} N_1^{\frac{1}{q\bar{r}}} N_2^{\frac{1}{r}} \|M\|_{\bar{p}}^p (1 + |\xi|^2)^{-\varepsilon p}.$$

This proves (1.3.5). Estimate (1.3.6) is obtained in a similar way from the identity $\partial_h Z_s^{ij} = \partial_{hj}^2 \gamma_s^i$, with $\partial_{hj}^2 \gamma_s^i$ satisfying SDE (1.1.5), and employing estimate (1.3.9) with $|\beta| = 2$. Hence, by Lemma

1.3.3, Z has a $\mathbf{bC}_{0,T}^{1-\frac{d}{p}}$-modification and therefore $\nabla\gamma$ is bounded as a function of $(s,\xi) \in [0,T]\times\mathbb{R}^d$, P-a.e. by (1.3.7).

Next we prove that $\det \nabla\gamma_s(\xi)$ is bounded from above and below by a positive random variable for all (s,ξ), P-a.s. By Itô formula (see [37], Lemma 3.1 for more details), with probability one we have

$$\det \nabla\gamma_s(\xi) = \exp\left(-\int_0^s \operatorname{tr}(\nabla\sigma_\tau^k)(\gamma_\tau(\xi))dW_\tau^k + \frac{1}{2}\int_0^s \operatorname{tr}\left((\nabla\sigma_\tau^k)^2\right)(\gamma_\tau(\xi))d\tau\right). \tag{1.3.10}$$

Since both parts of the equality are continuous w.r.t (s,ξ), the equality holds for all (s,ξ) at once with probability one. Thus the assertion follows from the boundedness of the integrals appearing in (1.3.10), which again can be proved as an application of Lemma 1.3.3, estimate (1.3.8) and Assumption 1.1.4.

Then the matrix $Y_t(x)$ is well defined and $\det Y_s(\xi)$ is bounded from below by a positive random variable for all (s,ξ), P-a.s. This fact, together with the uniform boundedness of the entries of $\nabla\gamma_s(\xi)$, implies (1.3.4).

It remains to prove that $\nabla\gamma_s(\xi)$ and Y_s have uniformly bounded spatial derivatives P-a.s. Again, this is a consequence of formula (1.1.5), Lemma 1.3.3 and the simple equality

$$\partial_j Y_s(\xi) = -Y_s(\xi)\partial_j(\nabla\gamma_s(\xi))Y_s(\xi).$$

□

Theorem 1.3.4. *The function u is a classical solution of SPDE* (1.1.1) *if and only if $\hat{u}$ in* (1.3.1) *solves*

$$d\hat{u}_s(\xi) = \left(L_s\hat{u}_s(\xi) + \hat{f}_s(\xi)\right)ds \tag{1.3.11}$$

where

$$L_s = \frac{1}{2}\bar{a}_s^{ij}\partial_{ij} + \bar{b}_s^i\partial_i + \bar{c}_s \tag{1.3.12}$$

is the parabolic operator with coefficients $\bar{a}^{ij}, \bar{b}^j, \bar{c} \in \mathbf{bC}_{0,T}^\alpha$ given explicitly by

$$\begin{aligned} \bar{a}_s^{ij} &= (Y_s\hat{\alpha}_sY_s^*)_{ij}\,, \\ \bar{b}_s^i &= Y_s^{ir}\left(\hat{b}_s^r - \widehat{\partial_j\sigma_s^{rk}\sigma_s^{jk}} - \hat{a}_s^{jh}\left(Y_s^*(\nabla^2\gamma_s^r)Y_s\right)_{jh}\right), \\ \bar{c}_s &= \hat{c}_s. \end{aligned} \tag{1.3.13}$$

Moreover, for some positive random variable μ, the following coercivity condition is satisfied

$$\langle\bar{a}_s(\xi)x,x\rangle \geqslant \mu|x|^2, \qquad s\in[0,T],\ x,\xi\in\mathbb{R}^d,\ P\text{-a.s.} \tag{1.3.14}$$

Proof. By assumption, u_s satisfies (1.3.2) with $h_s = \mathcal{L}_s u_s + f_s \in \mathbf{C}^{\alpha}_{0,T}$ and $g^k_s = \mathcal{G}^k_s u_s \in \mathbf{C}^{1+\alpha}_{0,T}$. Thus, by the Itô-Wentzell formula (1.3.3) we get

$$d\hat{u}_s = \left(\frac{1}{2}\hat{\alpha}^{ij}_s \widehat{\partial_{ij} u_s} + \left(\hat{b}^j_s - \widehat{\partial_i \sigma^{jk}_s \sigma^{ik}_s}\right)\widehat{\partial_j u}_s + \hat{c}_t \hat{u}_s + \hat{f}\right) ds. \tag{1.3.15}$$

Now, we have

$$\begin{aligned}
\partial_j \hat{u}_a(\xi) &= \widehat{\partial_i u_s}(\xi)\partial_j \gamma^i_s(\xi) = \left(\widehat{\nabla u_s}(\xi)\nabla\gamma_s(\xi)\right)_j, \\
\partial_{ij}\hat{u}_s(\xi) &= \left(\nabla\gamma^*_s(\xi)\widehat{\nabla^2 u_s}(\xi)\nabla\gamma_s(\xi)\right)_{ij} + \left(\widehat{\partial_h u_s}(\xi)\nabla^2\gamma^h_s(\xi)\right)_{ij},
\end{aligned}$$

or equivalently

$$\begin{aligned}
\widehat{\nabla u_s}(\xi) &= \nabla\hat{u}_s(\xi)Y_s(\xi), \\
\widehat{\nabla^2 u_s}(\xi) &= Y^*_s(\xi)\nabla^2\hat{u}_s(\xi)Y_s(\xi) - \left(Y^*_s(\xi)\nabla^2\gamma^h_s(\xi)Y_s(\xi)\right)\widehat{\partial_h u_s}(\xi).
\end{aligned}$$

Plugging these formulas into (1.3.15) and rearranging the indexes, we get (1.3.11)-(1.3.12)-(1.3.13). Moreover, from expressions (1.3.13) combined with Assumption 1.1.2 and Proposition 1.3.2 it is straightforward to see that $a^{ij}, b^j, c \in \mathbf{bC}^{\alpha}_{0,T}$. Eventually, by Assumption 1.1.3 and estimate (1.3.4) of Proposition 1.3.2 we have

$$\langle\hat{\alpha}_s(\xi)Y^*_s(\xi)x, Y^*_s(\xi)x\rangle \geqslant \mathbf{m}|Y^*_s(\xi)x|^2 \geqslant \mathbf{m}\tilde{\mathbf{m}}|x|^2$$

for any $s \in [0,T]$, $x, \xi \in \mathbb{R}^d$, P-a.s. and this proves (1.3.14). □

1.4 Time-dependent parametrix

In this section we consider equation (1.3.11) for fixed $\omega \in \Omega$; more generally we consider the (deterministic) parabolic PDE

$$\mathcal{K}_s u(s,\xi) := \mathcal{L}_s u(s,\xi) - \partial_s u(s,\xi) = 0 \tag{1.4.1}$$

where

$$\mathcal{L}_s u(s,\xi) = \frac{1}{2}a^{ij}(s,\xi)\partial_{ij}u(s,\xi) + b^i(s,\xi)\partial_i u(s,\xi) + c(s,\xi)u(s,\xi). \tag{1.4.2}$$

Since the coefficients will be assumed only measurable in the time variable, equation (1.4.1) has to be understood in the integral sense: a solution to the Cauchy problem

$$\begin{cases}
\mathcal{K}_s u(s,\xi) + f(s,\xi) = 0, & \xi \in \mathbb{R}^d, \text{ a.e. } s \in (t,T], \\
u(t,\xi) = \varphi(\xi), & \xi \in \mathbb{R}^d,
\end{cases} \tag{1.4.3}$$

is a function $u \in C^2_{t,T}(\mathbb{R}^d)$ that satisfies

$$u(s,\xi) = \varphi(x) + \int_t^s (\mathcal{L}_\tau u(\tau,\xi) + f(\tau,\xi))d\tau, \qquad (s,\xi) \in [t,T] \times \mathbb{R}^d.$$

The main idea of the parametrix method is to construct the fundamental solution $\Gamma = \Gamma(t,x;s,\xi)$ of $\mathcal{K}_s$ using as a first approximation the so-called *parametrix*, that is the Gaussian kernel of the heat operator obtained by freezing the coefficients of $\mathcal{K}_s$ at the pole (t,x). If $Z = Z(t,x;s,\xi)$ denotes the parametrix, one looks for the fundamental solution of $\mathcal{K}_s$ in the form

$$\Gamma(t,x;s,\xi) = Z(t,x;s,\xi) + (\Phi \otimes Z)(t,x;s,\xi), \tag{1.4.4}$$

where the symbol $\otimes$ denotes the convolution operator

$$(\Phi \otimes Z)(t,x;s,\xi) := \int_t^s \int_{\mathbb{R}^d} \Phi(t,x;\tau,y) Z(\tau,y;s,\xi) dy d\tau. \tag{1.4.5}$$

The unknown function Φ is determined by imposing $\mathcal{K}_s\Gamma(t,x;s,\xi) = 0$: this implies that Φ should satisfy the integral equation

$$\Phi(t,x;s,\xi) = \mathcal{K}_s Z(t,x;s,\xi) + (\Phi \otimes \mathcal{K}_s Z)(t,x;s,\xi) \tag{1.4.6}$$

for any $x,\xi \in \mathbb{R}^d$ and a.e. $s \in (t,T]$. By recursive approximation we have

$$\Phi(t,x;s,\xi) = \sum_{k=1}^{+\infty} H^{\otimes k}(t,x;s,\xi) \tag{1.4.7}$$

where

$$\begin{aligned} H^{\otimes 1}(t,x;s,\xi) &= H(t,x;s,\xi) := \mathcal{K}_s Z(t,x;s,\xi), \\ H^{\otimes (k+1)}(t,x;s,\xi) &:= (H^{\otimes k} \otimes H)(t,x;s,\xi), \quad k = 2,3,\cdots. \end{aligned}$$

To prove convergence of the series (1.4.7) and show that the candidate Γ in (1.4.4)-(1.4.6) is indeed a fundamental solution for $\mathcal{K}_s$, we need to impose some conditions.

Assumption 1.4.1. *There exists a positive constants λ, $\alpha \in (0,1)$ such that $a^{ij} \in C^\alpha_{0,T}$ with Hölder constant λ for every $i,j = 1,\cdots,d$ and*

$$\lambda^{-1}|x|^2 \leqslant \langle a(s,\xi)x,x\rangle \leqslant \lambda|x|^2, \quad |b^j(s,\xi)| + |c(s,\xi)| \leqslant \lambda, \qquad s \in [0,T],\ x,\xi \in \mathbb{R}^d.$$

Notation 1.4.2. *We introduce for notational convenience the following parameter set which gathers important quantities appearing in the assumptions:*

$$\Theta := (T,\alpha,\lambda,d),$$

where again $T > 0$ stands for the fixed considered final time.

As opposed to the classical parametrix method, in Assumption 1.4.1 we do not require any regularity of the coefficients in the time variable. Instead, here we only require Hölder continuity in the spatial variables. The reason lies in the fact that we are going to adopt a time-dependent definition of parametrix: namely, we do not freeze the time variable in the definition of Z (see (1.4.14) below) and take as parametrix the fundamental solution of a parabolic equation with coefficients depending on s.

Remark 1.4.3. *Using the enhanced version of the parametrix method proposed in [13], we can weaken the conditions on the first- and zero-order coefficients that can be supposed to be unbounded with sub-linear growth at infinity.*

Definition 1.4.4. *A fundamental solution* $\Gamma = \Gamma(t, x; s, \xi)$ *for equation* (1.4.1) *is a function defined for* $0 \leqslant t < s \leqslant T$ *and* $x, \xi \in \mathbb{R}^d$*, such that for any* $(t, x) \in [0, T) \times \mathbb{R}^d$ *we have:*

i) $\Gamma(t, x; \cdot, \cdot) \in C^2_{t_0,T}(\mathbb{R}^d)$ *for any* $t_0 \in (t, T)$ *and satisfies* $\mathcal{K}_s\Gamma(t, x; s, \xi) = 0$ *for any* $\xi \in \mathbb{R}^d$ *and a.e.* $s \in (t, T]$*;*

ii) for any continuous and non-rapidly increasing function φ *on* $\mathbb{R}^d$ *and* $x_0 \in \mathbb{R}^d$

$$\lim_{\substack{(s,\xi)\to(t,x_0)\\ s>t}} \int_{\mathbb{R}^d} \varphi(x)\Gamma(t, x; s, \xi)dx = \varphi(x_0).$$

Next we state the main result of this section.

Theorem 1.4.5 (Existence of the fundamental solution)**.** *Under Assumption 1.4.1, there exists a fundamental solution* Γ *for equation* (1.4.1)*. Moreover, assume that* $\varphi = \varphi(x)$ *is continuous and non-rapidly increasing on* $\mathbb{R}^d$*, and* $f = f(s, x)$ *is non-rapidly increasing uniformly on* $[t, T] \times \mathbb{R}^d$ *and such that* $f \in C^{\alpha'}_{t,T}$ *for some* $\alpha' \in (0, 1)$*. Then*

$$u(s, \xi) = \int_{\mathbb{R}^d} \varphi(x)\Gamma(t, x; s, \xi)dx + \int_t^s \int_{\mathbb{R}^d} f(\tau, x)\Gamma(\tau, x; s, \xi)dxd\tau \tag{1.4.8}$$

is a solution to the Cauchy problem (1.4.3)*. Such a solution is unique in the class of functions with quadratic exponential growth (cf. Corollary 1.1.6).*

Theorem 1.4.6 (Properties of the fundamental solution)**.** *Under the same assumptions of Theorem 1.4.5, the fundamental solution* Γ *enjoys the following properties:*

i) Γ *verifies the Chapman-Kolmogorov identity*

$$\Gamma(t_0, x_0; s, \xi) = \int_{\mathbb{R}^d} \Gamma(t_0, x_0; t, x)\Gamma(t, x; s, \xi)dx, \qquad t_0 < t < s, \ \xi, x_0 \in \mathbb{R}^d;$$

and, if $c = c_s$ is independent of ξ, we have

$$\int_{\mathbb{R}^d} \Gamma(t,x;s,\xi)dx = e^{\int_t^s c_\tau d\tau}, \qquad t \leqslant s \leqslant T, \ \xi \in \mathbb{R}^d. \tag{1.4.9}$$

In particular, if $c \equiv 0$ then $\Gamma(t,\cdot;s,\xi)$ is a density;

ii) there exist two positive constants $\mu = \mu(\Theta) \geqslant 1$, $C = C(\Theta) \geqslant 1$, such that

$$C^{-1}\Gamma^{\text{heat}}\left(\mu^{-1}\mathcal{I}_{s-t}, \xi - x\right) \leqslant \Gamma(t,x;s,\xi) \leqslant C\Gamma^{\text{heat}}\left(\mu\mathcal{I}_{s-t}, \xi - x\right), \tag{1.4.10}$$

$$|\nabla_\xi \Gamma(t,x;s,\xi)| \leqslant \frac{C}{\sqrt{s-t}}\Gamma^{\text{heat}}\left(\mu\mathcal{I}_{s-t}, \xi - x\right), \tag{1.4.11}$$

$$\left|\nabla_\xi^2 \Gamma(t,x;s,\xi)\right| \leqslant \frac{C}{s-t}\Gamma^{\text{heat}}\left(\mu\mathcal{I}_{s-t}, \xi - x\right), \tag{1.4.12}$$

for every $0 \leqslant t < s \leqslant T$ and $x, \xi \in \mathbb{R}^d$.

1.4.1 Preliminary Gaussian and potential estimates

We freeze the coefficients of $\mathcal{L}_s$ in (1.4.2) at a fixed point $x_0 \in \mathbb{R}^d$ and consider the operator with time-dependent coefficients

$$\widetilde{\mathcal{L}}_s^{x_0} = \frac{1}{2}a^{ij}(s,x_0)\partial_{\xi_i\xi_j}$$

acting in the ξ-variable. We denote by

$$\widetilde{\Gamma}^{x_0}(t,x;s,\xi) = \Gamma^{\text{heat}}(\mathcal{C}_{t,s}(x_0), \xi - x), \qquad \mathcal{C}_{t,s}(x_0) := \int_t^s a(\tau, x_0)d\tau, \tag{1.4.13}$$

the fundamental solution of $\widetilde{\mathcal{L}}_s^{x_0} - \partial_s$. Notice that $\widetilde{\Gamma}^{x_0}$ is well defined for $0 \leqslant t < s \leqslant T$ in virtue of Assumption 1.4.1 and solves

$$\partial_s \widetilde{\Gamma}^{x_0}(t,x;s,\xi) = \widetilde{\mathcal{L}}_s^{x_0}\widetilde{\Gamma}^{x_0}(t,x;s,\xi)$$

for any $x, \xi \in \mathbb{R}^d$ and almost every $s \in (t,T]$. Finally, we define the parametrix for $\mathcal{K}_s$ as

$$Z(t,x;s,\xi) = \widetilde{\Gamma}^{x}(t,x;s,\xi), \qquad 0 \leqslant t < s \leqslant T, \ x, \xi \in \mathbb{R}^d. \tag{1.4.14}$$

The following Gaussian estimates are standard consequences of Assumption 1.4.1.

Lemma 1.4.7. *We have*

$$\lambda^{-d}\Gamma^{\text{heat}}\left(\lambda^{-1}\mathcal{I}_{s-t}, \xi - x\right) \leqslant \widetilde{\Gamma}^{x_0}(t,x;s,\xi) \leqslant \lambda^d\Gamma^{\text{heat}}(\lambda\mathcal{I}_{s-t}, \xi - x), \tag{1.4.15}$$

for any $0 \leqslant t < s \leqslant T$ and $x, \xi, x_0 \in \mathbb{R}^d$. Moreover, $\widetilde{\Gamma}^{x_0}(t,x;s,\xi)$ verifies the Gaussian estimates (1.4.11)-(1.4.12) *for some positive constants μ, C dependent on Θ.*

Proposition 1.4.8. *There exists $k_0 \in \mathbb{N}$ such that, for every $t \in [0,T[$ and $x \in \mathbb{R}^d$, the series*

$$\sum_{k=k_0}^{\infty} H^{\otimes k}(t,x;\cdot,\cdot)$$

converges in $L^\infty((t,T] \times \mathbb{R}^d)$. The function Φ defined by (1.4.7) solves the integral equation (1.4.6) and there exist two positive constants μ, C dependent on Θ such that

$$|\Phi(t,x;s,\xi)| \leqslant C(s-t)^{-1+\frac{\alpha}{2}}\Gamma^{\mathrm{heat}}(\mu\mathcal{I}_{s-t}, \xi - x), \tag{1.4.16}$$

$$\begin{aligned} |\Phi(t,x;s,\xi) - \Phi(t,x;s,y)| &\leqslant C\frac{|\xi - y|^{\frac{\alpha}{2}}}{(s-t)^{1-\frac{\alpha}{4}}} \times \\ &\quad \times \left(\Gamma^{\mathrm{heat}}(\mu\mathcal{I}_{s-t}, \xi - x) + \Gamma^{\mathrm{heat}}(\mu\mathcal{I}_{s-t}, y - x)\right), \end{aligned} \tag{1.4.17}$$

for every $x, y, \xi \in \mathbb{R}^d$ and almost every $s \in (t,T]$.

Proof. We first establish the following elementary inequality: for any $\varepsilon > 0$ and $n \in \mathbb{N}$ there exists a positive constant $c_{\varepsilon,n}$ such that

$$|\lambda|^n e^{-\frac{\lambda^2}{\mu}} \leqslant c_{n,\varepsilon} e^{-\frac{\lambda^2}{\mu+\varepsilon}}, \qquad \lambda \in \mathbb{R}. \tag{1.4.18}$$

Next prove the preliminary estimate

$$\left|H^{\otimes k}(t,x;s,\xi)\right| \leqslant M_k (s-t)^{-1+\alpha k/2}\Gamma^{\mathrm{heat}}(\mu\mathcal{I}_{s-t}, \xi - x) \quad x, \xi \in \mathbb{R}^d, \text{ a.e. } s \in (t,T],\ k \in \mathbb{N}, \tag{1.4.19}$$

where $C = C(\Theta)$ is a positive constant, $M_k = C^k \frac{\Gamma_E^k\left(\frac{\alpha}{2}\right)}{\Gamma_E\left(\frac{\alpha k}{2}\right)}$ and Γ_E is the Euler Gamma function.

For $k = 1$, we have

$$\begin{aligned} |H(t,x;s,\xi)| &= \left|(\mathcal{L}_s - \widetilde{\mathcal{L}}_s^x)Z(t,x;s,\xi)\right| \\ &\leqslant \frac{1}{2}\left|a^{ij}(s,\xi) - a^{ij}(s,x)\right| |\partial_{ij}Z(t,x;s,\xi)| + \\ &\quad + \left|b^i(s,\xi)\right| |\partial_i Z(t,x;s,\xi)| + |c(s,\xi)Z(t,x;s,\xi)| \\ &=: I_1 + I_2 + I_3. \end{aligned}$$

By Assumption 1.4.1, Lemma 1.4.7 and (1.4.18), we have

$$I_1 \leqslant C(s-t)^{-1+\frac{\alpha}{2}}\left(\frac{|\xi - x|}{\sqrt{s-t}}\right)^{\alpha}\Gamma^{\mathrm{heat}}(\mu\mathcal{I}_{s-t}, \xi - x) \leqslant C(s-t)^{-1+\frac{\alpha}{2}}\Gamma^{\mathrm{heat}}((\mu+1)\mathcal{I}_{s-t}, \xi - x).$$

Since the coefficients are bounded, by Lemma 1.4.7 we also have

$$I_2 \leqslant C(s-t)^{-\frac{1}{2}}\Gamma^{\mathrm{heat}}(\mu\mathcal{I}_{s-t}, \xi - x), \qquad I_3 \leqslant C\Gamma^{\mathrm{heat}}(\mu\mathcal{I}_{s-t}, \xi - x),$$

and this proves (1.4.19) for $k = 1$. Now we assume that (1.4.19) holds for k and prove it for $k+1$: by inductive hypothesis and the Chapman-Kolmogorov property of the heat kernel we have

$$\begin{aligned}
\left|H^{\otimes(k+1)}(t,x;s,\xi)\right| &= \left|\int_t^s \int_{\mathbb{R}^d} H^{\otimes k}(t,x;\tau,y)H(\tau,y;s,\xi)dyd\tau\right| \leqslant \\
&\leqslant M_k M_1 \int_t^s (\tau-t)^{-1+\frac{\alpha k}{2}}(s-\tau)^{-1+\frac{\alpha}{2}} \times \\
&\qquad \times \int_{\mathbb{R}^d} \Gamma^{\text{heat}}(\mu\mathcal{I}_{\tau-t}, y-x)\Gamma^{\text{heat}}(\mu\mathcal{I}_{s-\tau}, \xi-y)dyd\tau \leqslant \\
&\leqslant M_k M_1 (t-s)^{-1+\frac{\alpha(k+1)}{2}} \Gamma^{\text{heat}}(\mu\mathcal{I}_{s-t}, \xi-x) \int_0^1 \tau^{-1+\frac{\alpha k}{2}}(1-\tau)^{-1+\frac{\alpha}{2}} d\tau
\end{aligned}$$

that yields (1.4.19) thanks to the properties of the Gamma function. From (1.4.19) we directly deduce the uniform convergence of the series and estimate (1.4.16). The proof of (1.4.17) follows the same lines as in the classical case (see [24], Ch.1, Theor.7) and is omitted. □

We close this section by stating a generalization of a classical result about the so-called *volume potential* defined as

$$V_f(s,\xi) = \int_{t_0}^s \int_{\mathbb{R}^d} f(t,x)Z(t,x;s,\xi)dxdt, \qquad (s,\xi) \in [t_0,T] \times \mathbb{R}^d, \tag{1.4.20}$$

where Z denotes the parametrix. The proof is based on classical arguments (see [24], Ch.1, Sec.3 and [28]) that can be applied to the time-dependent parametrix Z in (1.4.14) without any significant change.

Lemma 1.4.9. *Let V_f be the volume potential in* (1.4.20) *with $f \in C^{\alpha}_{t_0,T}(\mathbb{R}^d)$, non-rapidly increasing uniformly w.r.t. t. Then $V_f \in C^2_{t_0,T}\left(\mathbb{R}^d\right)$ satisfies*

$$\begin{aligned}
\nabla^j_\xi V_f(s,\xi) &= \int_{t_0}^s \int_{\mathbb{R}^d} f(t,x)\nabla^j_\xi Z(t,x;s,\xi)dxdt, \quad j = 1,2, \\
\partial_s V_f(t,x) &= f(s,\xi) + \int_{t_0}^s \int_{\mathbb{R}^d} f(t,x)\partial_s Z(t,x;s,\xi)dxdt,
\end{aligned}$$

for any $\xi \in \mathbb{R}^d$ and a.e. $s \in (t_0,T]$.

1.4.2 Proof of Theorem 1.4.5

Let $\Gamma = \Gamma(t,x;s,\xi)$ be the function defined by (1.4.4)-(1.4.7) for $0 \leqslant t < s \leqslant T$ and $x,\xi \in \mathbb{R}^d$. By Proposition 1.4.8, it is clear that $\Gamma(t,x;\cdot,\cdot) \in C^0_{t,T}(\mathbb{R}^d)$ for any $(t,x) \in [0,T) \times \mathbb{R}^d$. Next, we fix $t_0 \in (t,s)$ and notice that by (1.4.16)-(1.4.17) the function $f := \Phi(t,x;\cdot,\cdot)$, defined on $[t_0,T] \times \mathbb{R}^d$, satisfies the conditions of Lemma 1.4.9: hence the volume potential

$$V_\Phi(s,\xi) := \int_{t_0}^s \int_{\mathbb{R}^d} \Phi(t,x;\tau,y)Z(\tau,y;s,\xi)dyd\tau$$

is twice continuously differentiable in ξ and satisfies

$$\mathcal{K}_s V_\Phi(s,\xi) = \int_{t_0}^{s}\int_{\mathbb{R}^d} \Phi(t,x;\tau,y)H(\tau,y;s,\xi)dyd\tau - \Phi(t,x;s,\xi), \qquad \text{a.e. } s \in (t_0,T].$$

On the other hand, we also have

$$\mathcal{K}_s \int_t^{t_0}\int_{\mathbb{R}^d} \Phi(t,x;\tau,y)Z(\tau,y;s,\xi)dyd\tau = \int_t^{t_0}\int_{\mathbb{R}^d} \Phi(t,x;\tau,y)H(\tau,y;s,\xi)dyd\tau$$

by the dominated convergence theorem. Consequently, we have

$$\mathcal{K}_s\Gamma(t,x;s,\xi) = H(t,x;s,\xi) + (\Phi \otimes H)(t,x;s,\xi) - \Phi(t,x;s,\xi) = 0$$

for a.e. $s \in (t,T]$, because Φ solves equation (1.4.6). This proves property i) of Definition 1.4.4 of fundamental solution. To prove property ii), it suffices to notice that

$$\int_{\mathbb{R}^d} \varphi(x)\Gamma(t,x;s,\xi)dx = I_1(t,s,\xi) + I_2(t,s,\xi)$$

where

$$\begin{aligned}
\lim_{\substack{(s,\xi)\to(t,x_0)\\ s>t}} I_1(t,s,\xi) &= \lim_{\substack{(s,\xi)\to(t,x_0)\\ s>t}} \int_{\mathbb{R}^d} \varphi(x)Z(t,x;s,\xi)dx = \varphi(x_0),\\
\lim_{\substack{(s,\xi)\to(t,x_0)\\ s>t}} |I_2(t,s,\xi)| &\leqslant \lim_{\substack{(s,\xi)\to(t,x_0)\\ s>t}} \int_{\mathbb{R}^d}\int_t^s\int_{\mathbb{R}^d} |\varphi(x)\Phi(t,x;\tau,y)|\, Z(\tau,y;s,\xi)dyd\tau dx \leqslant
\end{aligned}$$

(by (1.4.15)-(1.4.16) and since φ is non-rapidly increasing, taking $\delta > 0$ suitably small, with $C = C(\lambda,\delta)$)

$$\begin{aligned}
&\leqslant \lim_{\substack{(s,\xi)\to(t,x_0)\\ s>t}} \int_{\mathbb{R}^d}\int_t^s Ce^{\delta|x|^2}(\tau-t)^{-1+\frac{\alpha}{2}} \times\\
&\qquad \times \int_{\mathbb{R}^d} \Gamma^{\text{heat}}(\mu\mathcal{I}_{\tau-t}, y-x)\Gamma^{\text{heat}}(\mu\mathcal{I}_{s-\tau}, \xi-y)dyd\tau dx\\
&\leqslant \lim_{\substack{(s,\xi)\to(t,x_0)\\ s>t}} \int_{\mathbb{R}^d}\int_t^s C(\tau-t)^{1-\frac{\alpha}{2}}\Gamma^{\text{heat}}(\mu\mathcal{I}_{s-t}, \xi-x)e^{\delta|x|^2}d\tau dx = 0.
\end{aligned} \tag{1.4.21}$$

Finally, the standard proof of existence for the Cauchy problem (see for instance [24], Ch.1, Theor.12, or [28]) applies without modification. Uniqueness follows from the maximum principle.

1.4.3 Proof of Theorem 1.4.6

The Chapman-Kolmogorov identity follows from uniqueness of the Cauchy problem (1.4.3) and representation (1.4.8) with $f \equiv 0$ and $\varphi = \Gamma(t_0, x_0; t, \cdot)$, for fixed $(t_0, x_0) \in [0, t) \times \mathbb{R}^d$. Analogously, formula (1.4.9) follows from uniqueness of the Cauchy problem (1.4.3) with $f \equiv 0$ and $\varphi \equiv 1$.

Next we prove the Gaussian estimates for Γ. By the Chapman-Kolmogorov property for Γ^{heat} we have

$$
\begin{aligned}
|(\Phi \otimes Z)(t, x; \tau, \xi)| &\leqslant \int_t^s \int_{\mathbb{R}^d} |\Phi(t, x; \tau, y)|\, Z(\tau, y; s, \xi) dy d\tau \\
&\leqslant C\Gamma^{\text{heat}}(\mu \mathcal{I}_{s-t}, \xi - x) \int_t^s (\tau - t)^{-1+\frac{\alpha}{2}} d\tau \leqslant C(s-t)^{\frac{\alpha}{2}} \Gamma^{\text{heat}}(\mu \mathcal{I}_{s-t}, \xi - x)
\end{aligned}
$$

for some positive C, μ. Since $\Gamma = Z + \Phi \otimes Z$, the previous estimate combined with (1.4.15) proves

$$
|\Gamma(t, x; s, \xi)| \leqslant C_1 \Gamma^{\text{heat}}(\mu \mathcal{I}_{s-t}, \xi - x)
$$

and in particular, the upper bound for Γ in (1.4.10). The proof of (1.4.11)-(1.4.12) is similar. Notice that by the maximum principle (in the form of Lemma 5 p.43 in [24]) applied to $u(s, \xi) = \int_{\mathbb{R}^d} \varphi(x) \Gamma(t, x; s, \xi) dx$, where φ is any bounded, non-negative and continuous function, one easily infers that Γ is non-negative.

To prove the Gaussian lower bound we adapt a procedure due to Aronson that is essentially based on a crucial Nash's lower bound (see [20], Sect. 2). The main difference is that in our setting we replace Nash's estimate with a bound that we directly derive from the parametrix method. Let us first notice that, for $\mu > 1$, we have $\Gamma^{\text{heat}}(\mu \mathcal{I}_t, x) \leqslant \Gamma^{\text{heat}}(\mu^{-1} \mathcal{I}_t, x)$ if $|x| \leqslant \varrho_\mu \sqrt{t}$ where $\varrho_\mu = \sqrt{\frac{\mu d}{\mu^2 - 1} \log \mu}$. Thus, by (1.4.15) and (1.4.16) we have

$$
\Gamma(t, x; s, \xi) \geqslant Z(t, x; s, \xi) - |(\Phi \otimes Z)(t, x; s, \xi)| \geqslant
$$

(if $|\xi - x| \leqslant \varrho_\mu \sqrt{s - t}$)

$$
\begin{aligned}
&\geqslant \left(\mu^{-d} - C(s-t)^{\frac{\alpha}{2}}\right) \Gamma^{\text{heat}}(\mu^{-1} \mathcal{I}_{s-t}; \xi - x) \\
&\geqslant \frac{1}{2} \mu^{-d} \Gamma^{\text{heat}}(\mu^{-1} \mathcal{I}_{s-t}; \xi - x) \qquad (1.4.22)
\end{aligned}
$$

if $0 < s - t \leqslant T_\mu := (2C\lambda^d)^{-\frac{2}{\alpha}} \wedge T$.

For any $(t, x) \in [0, T) \times \mathbb{R}^d$, $(s, \xi) \in (t, T] \times \mathbb{R}^d$ we set m to be the smallest natural number greater than

$$
\max\left\{4\varrho_\mu^{-2} \frac{|\xi - x|^2}{(s-t)}, \frac{T}{T_\mu}\right\}.
$$

Then we set

$$t_i = t + i\frac{s-t}{m+1}, \quad x_i = x + i\frac{\xi - x}{m+1}, \qquad i = 0, \dots, m+1.$$

Denoting by $D(x;r) = \{y \in \mathbb{R}^d, |x-y| < r\}$ the Euclidean ball centered at x with radius $r > 0$, by the Chapman-Kolmogorov equation we have

$$\Gamma(t,x;\tau,\xi) = \int_{\mathbb{R}^{md}} \Gamma(t,x;t_1,y_1) \prod_{i=1}^{m-1} \Gamma(t_i,y_i;t_{i+1},y_{i+1})\Gamma(t_m,y_m;s,\xi)dy_1 \cdots dy_m$$

(since Γ is non-negative)

$$\begin{aligned}\geqslant \int_{\mathbb{R}^{md}} \Gamma(t,x;t_1,y_1) \prod_{i=1}^{m-1} \Gamma(t_i,y_i;t_{i+1},y_{i+1}) \mathbb{1}_{D(x_i;r)}(y_i) \times \\ \times \Gamma(t_m,y_m;s,\xi)\mathbb{1}_{D(x_m;r)}(y_m)dy_1 \cdots dy_m.\end{aligned} \tag{1.4.23}$$

Now we have

$$t_{i+1} - t_i = \frac{s-t}{m+1} \leqslant \frac{T}{m+1} \leqslant T_\mu, \qquad i = 0, \dots, m$$

by definition of m. Moreover, if $y_i \in D(x_i;r)$ for $i = 1, \dots, m$, by the triangular inequality we have

$$|y_{i+1} - y_i| \leqslant 2r + |x_{i+1} - x_i| = 2r + \frac{|\xi - x|}{m+1} \leqslant$$

(again, by definition of m)

$$\leqslant 2r + \frac{\varrho_\mu}{2}\sqrt{\frac{s-t}{m+1}} \leqslant \varrho_\mu \sqrt{\frac{s-t}{m+1}}, \tag{1.4.24}$$

if we set

$$r = \frac{\varrho_\mu}{4}\sqrt{\frac{s-t}{m+1}} > 0.$$

For such a choice of r, we can use (1.4.22) repeatedly in (1.4.23) and get

$$\begin{aligned}\Gamma(t,x;s,\xi) \geqslant (2\lambda^d)^{-(m+1)} \int_{\mathbb{R}^{md}} & \Gamma^{\text{heat}}\left(\mu^{-1}\mathcal{I}_{\frac{s-t}{m+1}}, y_1 - x\right) \times \\ & \times \prod_{i=1}^{m-1} \Gamma^{\text{heat}}\left(\mu^{-1}\mathcal{I}_{\frac{s-t}{m+1}}, y_{i+1} - y_i\right) \mathbb{1}_{D(x_i;r)}(y_i) \times \\ & \times \Gamma^{\text{heat}}\left(\mu^{-1}\mathcal{I}_{\frac{s-t}{m+1}}, \xi - y_m\right) \mathbb{1}_{D(x_m;r)}(y_m)dy_1 \cdots dy_m\end{aligned}$$

(by (1.4.24) and denoting by ω_d the volume of the unit ball in $\mathbb{R}^d$)

$$\geqslant (2\lambda^d)^{-(m+1)}(\omega_d r^d)^m \left(\frac{\mu(m+1)}{2\pi(s-t)}\right)^{\frac{d}{2}(m+1)} \exp\left(-\frac{\mu\varrho_\mu^2}{2}(m+1)\right).$$

It follows that there exists a positive constant $C = C(\Theta)$ such that

$$\Gamma(t,x;s,\xi) \geqslant C_0^{-1}(s-t)^{-\frac{d}{2}} \exp\left(-Cm\right),$$

and this implies the required estimate: indeed if $|\xi - x|^2 \geqslant \frac{T\varrho_\mu^2}{4T_\mu}(s-t)$, then, by the definition of m it follows that $m < 8|\xi - x|^2(\varrho_\mu^2(s-t))^{-1}$ and then

$$\Gamma(t,x;s,\xi) \geqslant C_1(2\pi(s-t))^{-\frac{d}{2}} \exp\left(-\frac{\bar\mu}{2}\frac{|\xi-x|^2}{s-t}\right) = C_2\Gamma^{\text{heat}}(\bar\mu^{-1}\mathcal{I}_{s-t}, \xi - x).$$

Otherwise if we have $m < 2T/T_\mu$, and then

$$\Gamma(t,x;s,\xi) \geqslant C_3(2\pi(s-t))^{-\frac{d}{2}} \geqslant C_3(2\pi(s-t))^{-\frac{d}{2}} \exp\left(-\frac{\bar\mu}{2}\frac{|\xi-x|^2}{s-t}\right) = C_4\Gamma^{\text{heat}}(\bar\mu^{-1}\mathcal{I}_{s-t}, \xi - x).$$

1.4.4 Proof of Theorem 1.1.5

For any fixed $t \in [0,T)$, we consider the stochastic flow $\gamma_{t,s}^{\text{IW}}$ defined as in (1.1.3) for $s \in [t,T]$. Let $L_{t,s}$ be the operator defined as in (1.3.12)-(1.3.13) through the random change of variable $\gamma_{t,s}^{\text{IW}}$. By Theorem 1.4.5, $K_{t,s} = L_{t,s} - \partial_s$ is a parabolic operator on the strip $[t,T] \times \mathbb{R}^d$ with random coefficients, that satisfies Assumption 1.4.1 on $[t,T]\times\mathbb{R}^d$ for almost every $\omega \in \Omega$. Then, by Theorem 1.4.5, $K_{t,s}$ admits a fundamental solution $\Gamma^{(t)}(t,x;s,\xi)$ defined for $s \in (t,T]$ and $x,\xi \in \mathbb{R}^d$. We put

$$\mathbf{\Gamma}(t,x;s,\xi) := \Gamma^{(t)}\left(t,x;s,\gamma_{t,s}^{\text{IW},-1}(\xi)\right), \qquad s \in (t,T],\ x,\xi \in \mathbb{R}^d. \tag{1.4.25}$$

Combining Theorems 1.3.4 and 1.4.5, we infer that $\mathbf{\Gamma}(t,x;\cdot,\cdot) \in \mathbf{C}^2_{\tau,T}(\mathbb{R}^d)$ and satisfies (1.1.2) with probability one. Moreover, let us consider a continuous and non-rapidly increasing function φ on $\mathbb{R}^d$; proceeding as in the proof of Theorem 1.4.5 we have

$$\int_{\mathbb{R}^d} \varphi(x)\mathbf{\Gamma}(t,x;s,\xi)dx - \varphi(\xi) = I_1(t,s,\xi) + I_2(t,s,\xi)$$

where $I_2(t,s,\xi)$ is defined and can be estimated as in (1.4.21); whereas, recalling the definition of parametrix in (1.4.13), (1.4.14), we have

$$\begin{aligned}\lim_{\substack{(s,\xi)\to(t,x_0)\\ s>t}} I_1(t,s,\xi) &= \lim_{\substack{(s,\xi)\to(t,x_0)\\ s>t}} \int_{\mathbb{R}^d} (\varphi(x) - \varphi(x_0))\,\Gamma^{\text{heat}}\left(\mathcal{C}_{t,s}(x), \gamma_{t,s}^{\text{IW},-1}(\xi) - x\right) dx \\ &= \lim_{\substack{(s,\xi)\to(t,x_0)\\ s>t}} \int_{\mathbb{R}^d} \left(\varphi(\gamma_{t,s}^{\text{IW},-1}(\xi) - x) - \varphi(x_0)\right) \Gamma^{\text{heat}}\left(\mathcal{C}_{t,s}(\gamma_{t,s}^{\text{IW},-1}(\xi) - x), x\right) dx = 0\end{aligned}$$

by the dominated convergence theorem. This proves that $\mathbf{\Gamma}$ is a fundamental solution for the SPDE (1.1.1).

The Gaussian bounds (1.1.8) follow directly from the definition (1.4.25) and the analogous estimates (1.4.10) for $\Gamma^{(t)}$ in Theorem 1.4.6. Moreover, since

$$\nabla_\xi \mathbf{\Gamma}(t,x;s,\xi) = (\nabla \Gamma^{(t)})\big(t,x;s,\gamma_{t,s}^{\text{IW},-1}(\xi)\big)\nabla \gamma_{t,s}^{\text{IW},-1}(\xi),$$

the gradient estimate (1.1.9) follows from the analogous estimate (1.4.11) for $\Gamma^{(t)}$ and from Proposition 1.3.2. The proof of (1.4.10) is analogous.

Chapter 2

On a class of Langevin and Fokker-Plank SPDEs

2.1 Introduction

In this chapter we expand on the first study and consider a stochastic version of the Fokker-Plank equation

$$\partial_s u + \sum_{j=1}^{n} \nu_j \partial_{\xi_j} u = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij} \partial_{\nu_i \nu_j} u. \tag{2.1.1}$$

Here the variables $s \geqslant 0$, $\xi \in \mathbb{R}^n$ and $\nu \in \mathbb{R}^n$ respectively stand for time, position and velocity, and the unknown $u = u_s(\xi, \nu) \geqslant 0$ stands for the density of particles in phase space.

We study a kinetic model where the position and the velocity of a particle are stochastic processes (X_s, V_s) only partially observable through some observation process Y_s. We consider the two-dimensional case, $n = 1$, and propose an approach that hopefully can be extended to the multi-dimensional case. If $\mathcal{F}_s^Y = \sigma(Y_\tau,\, \tau \leqslant s)$ denotes the filtration of the observations then, under natural assumptions, the *conditional* density given $\mathcal{F}_s^Y$ solves a linear SPDE of the form

$$d_{\mathbf{B}} u_s(\zeta) = \mathcal{A}_{s,\zeta} u_s(\zeta) ds + \mathcal{G}_{s,\zeta} u_s(\zeta) dW_s, \quad \zeta = (\xi, \nu_1, \cdots \nu_d) \in \mathbb{R}^{d+1}, \tag{2.1.2}$$

where $\mathbf{B} = \partial_s + \nu_1 \partial_\xi$ and

$$\mathcal{A}_{s,\zeta} u_s(\zeta) := \frac{1}{2} a_s^{ij}(\zeta) \partial_{\nu_i \nu_j} u_s(\zeta) + b_s^i(\zeta) \partial_{\nu_i} u_s(\zeta) + c_s(\zeta) u_s(\zeta);$$
$$\mathcal{G}_{s,\zeta} u_s(\zeta) := \sigma_s^i(\zeta) \partial_{\nu_i} u_s(\zeta) + h_s(\zeta) u_s(\zeta).$$

In (2.1.2) W is a Wiener process defined on a complete probability space $(\Omega, \mathcal{F}, P)$ endowed with a filtration $(\mathcal{F}_s)_{s \geqslant 0}$ satisfying the usual conditions. The symbol $d_{\mathbf{B}}$ indicates that the equation is solved in the Itô (or strong) sense:

Definition 2.1.1. *A solution to* (2.1.2) *on* $[t,T]$ *is a process* $u_s = u_s(\xi,\nu) \in \mathbf{C}^0_{t,T}$ *that is twice differentiable in the variables* ν *and solves the equation*

$$u_s\left(\gamma^{\mathbf{B}}_{s-t}(\zeta)\right) = u_t(\zeta) + \int_t^s \mathcal{A}_{\tau,\gamma^{\mathbf{B}}_{\tau-t}(\zeta)} u_\tau(\gamma^{\mathbf{B}}_{\tau-t}(\zeta))d\tau + \int_t^s \mathcal{G}_{\tau,\gamma^{\mathbf{B}}_{\tau-t}(\zeta)} u_\tau(\gamma^{\mathbf{B}}_{\tau-t}(\zeta))dW_\tau, \qquad s \in [t,T],$$

where $s \mapsto \gamma^{\mathbf{B}}_s(\xi,\nu)$ *denotes the integral curve, starting from* (ξ,ν), *of the advection vector field* $\nu_1\partial_\xi$, *that is*

$$\gamma^{\mathbf{B}}_s(\xi,\nu) = e^{sB}(\xi,\nu) = (\xi + s\nu_1,\nu), \qquad B = \begin{pmatrix} 0 & 1 & \mathbf{0}_{1\times(d-1)} \\ \mathbf{0}_{d\times 1} & \mathbf{0}_{d\times 1} & \mathbf{0}_{d\times(d-1)} \end{pmatrix},$$

where $\mathbf{0}_{n\times m}$ *denotes the* $n \times m$ *null matrix.*

Definition 2.1.2. *A stochastic fundamental solution* $\mathbf{\Gamma} = \mathbf{\Gamma}(t,z;s,\zeta)$ *for the SPDE* (2.1.2) *is a function defined for* $0 \leqslant t < s \leqslant T$ *and* $z,\zeta \in \mathbb{R}^{d+1}$, *such that for any* $(t,z) \in [0,T) \times \mathbb{R}^{d+1}$ *and* $t_0 \in (t,T)$ *we have:*

i) $\mathbf{\Gamma}(t,z,\cdot,\cdot)$ *is a solution to* (2.1.2) *on* $[t_0,T]$*;*

ii) for any $\varphi \in bC(\mathbb{R}^{d+1})$ *and* $z_0 \in \mathbb{R}^{d+1}$, *we have*

$$\lim_{\substack{(s,\zeta)\to(t,z_0)\\ s>t}} \int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;s,\zeta)\varphi(z)dz = \varphi(z_0), \qquad P\text{-}a.s.$$

The actual dimension of the equation can be possibly greater than two: this is still coherent with the two-dimensional model and it is due to the fact that the set of variables carried by Y may also enter the equation, as will become clear in Chapter 3. On the other hand, this only affects the elliptic part of (2.1.2), and in this sense, the cases $d = 1$ and $d > 1$ are completely analogous. In case the observation process Y is independent of X and V, the SPDE (2.1.2) boils down to the deterministic PDE (2.1.1) with $n = 1$.

Compared to the uniformly parabolic case, two main new difficulties arise:

i) the Itô-Wentzell transform drastically affects the drift $\mathbf{B}$: in particular, after the random change of coordinates, the new drift has no longer polynomial coefficients. Consequently, a careful analysis is needed to check the validity of the Hörmander condition in the new coordinates. This question is discussed in more detail in Section 2.2;

ii) in the deterministic case, the parametrix method has been applied to degenerate Fokker-Planck equations, including (2.1.2) with $\sigma \equiv 0$, by several authors, [58], [16], using *intrinsic* Hölder spaces. Loosely speaking, the intrinsic Hölder regularity reflects the geometry of the PDE and is defined in terms of the translations and homogeneous norm associated to the Hörmander vector fields: this kind of regularity is natural for the study of the singular kernels that come into play in the parametrix iterative procedure. Now, under the weak Hörmander condition, the intrinsic regularity properties in space and time are closely intertwined and cannot be studied separately. However, assuming that the coefficients are merely predictable, we have no good control on the regularity in the time variable; for instance, even in the deterministic case, the coefficients are only measurable in s and consequently they cannot be Hölder continuous in (x, v) in the intrinsic sense. On the other hand, assuming that the coefficients are Hölder continuous in (x, v) in the classical Euclidean sense, the parametrix method still works as long as we use a suitable *time-dependent* parametrix and exploit the fact that the intrinsic translations coincide with the Euclidean ones for points (s, x, v) and (s, ξ, η) at the same time level. We comment on this question more thoroughly in Section 2.3.

The chapter is organized as follows. In the remaining part of this section we set the assumptions on (2.1.2) and state the main result, Theorem 2.1.6. In Sections 2.2 and 2.3 we go deeper into the issues mentioned above. In Section 2.4 we prove some crucial estimates for stochastic flows of diffeomorphisms: these estimates, which can be of independent interest, are based on the ideas introduced in Section 1.3 and extend some result of [44]. In Section 2.5 we exploit the results of Section 2.4 to perform the reduction of the SPDE to a PDE with random coefficients. In Section 2.6 we build on the work by [14] to develop a parametrix method for Kolmogorov PDEs with general drift (Theorem 2.6.6) and in Section 2.6.2 we complete the proof of Theorem 2.1.6 for $d = 1$. In Section 2.7 we explain how our methods can be tweaked to apply to the backward version of equation (2.1.2).

2.1.1 Assumptions and main results

We start by setting the standing assumptions on the coefficients of the SPDE (2.1.2).

Assumption 2.1.3 (Regularity)**.** *For some $\alpha \in (0, 1)$ we have: $a \in \mathbf{bC}^{\alpha}_{0,T}$, $\sigma \in \mathbf{bC}^{3+\alpha}_{0,T}$, $b, c \in \mathbf{bC}_{0,T}$ and $h \in \mathbf{bC}^{2}_{0,T}$.*

Assumption 2.1.4 (Coercivity)**.** *There exists a random, finite and positive constant $\mathbf{m}$ such that*

$$\langle a_t(z) - \sigma_t(z)\sigma_t^*(z)\zeta, \zeta\rangle \geqslant \mathbf{m}|\zeta|^2, \qquad t \in [0, T],\ z, \zeta \in \mathbb{R}^{d+1},\ P\text{-a.s.}$$

We make again use of an Itô-Wentzell transform, but this time we only need to operate on the directions ν_j: for fixed $t \in [0, T]$ we consider the SDE in $\mathbb{R}^d$

$$\gamma^{\mathrm{IW}}_{t,s}(x, v) = v - \int_t^s \sigma_\tau(x, \gamma^{\mathrm{IW}}_{t,\tau}(x, v))dW_\tau. \tag{2.1.3}$$

Assumption 2.1.3 ensures that (2.1.3) is solvable in the strong sense and the map $(x,v) \mapsto \big(x, \gamma^{\mathrm{IW}}_{t,s}(x,v)\big)$ is a stochastic flow of diffeomorphisms of $\mathbb{R}^{d+1}$ (see Theorem 2.4.1 below). To have a control on its gradient, recalling the notation (1.1.6), we impose the following additional

Assumption 2.1.5. *There exist $\varepsilon > 0$ and two random variables $M_1 \in L^p(\Omega)$, with $p > \max\left\{2, \frac{1}{\varepsilon}\right\}$, and $M_2 \in L^\infty(\Omega)$, such that with probability one*

$$\sup_{t\in[0,T]} \big(\{\sigma_t\}_{\varepsilon,\beta} + \{\sigma_t\}_{1/2+\varepsilon,\beta'}\big) \leqslant M_1, \qquad |\beta| = 1,\ |\beta'| = 2,3,$$

$$\sup_{t\in[0,T]} \{h_t\}_{1/2,\beta} \leqslant M_2, \qquad |\beta| = 1.$$

In order to state the main result of this section, Theorem 2.1.6 below, we need to introduce some additional notation: we denote by $g^{\mathrm{IW},-1}$ the inverse of the Itô-Wentzell stochastic flow $(x,v) \mapsto g^{\mathrm{IW}}_{t,s}(x,v) := \big(x, \gamma^{\mathrm{IW}}_{t,s}(x,v)\big)$ defined by (2.1.3). Moreover, we consider the vector field

$$\mathbf{Y}_{t,s}(z) := \Big((\gamma^{\mathrm{IW}}_{t,s})_1(z), -(\gamma^{\mathrm{IW}}_{t,s}(z))_1(\nabla_v \gamma^{\mathrm{IW}}_{t,s})^{-1}(z)\partial_x \gamma^{\mathrm{IW}}_{t,s}(z)\Big), \tag{2.1.4}$$

with $\nabla_v \gamma^{\mathrm{IW}} = (\partial_{v_j}\gamma^{\mathrm{IW}}_i)_{i,j=1,\cdots d}$ and $\partial_x \gamma^{\mathrm{IW}} = (\partial_x \gamma^{\mathrm{IW}}_i)_{i=1,\cdots d}$. Eventually, equation

$$\gamma_{t,s}(z) = z + \int_t^s \mathbf{Y}_{t,\tau}(\gamma_{t,\tau}(z))d\tau, \qquad s \in [t,T],$$

defines the integral curve of $\mathbf{Y}_{t,s}$ starting from (t,z).

The central result of this chapter is the following theorem whose proof is postponed to Section 2.6.

Theorem 2.1.6. *Under Assumptions 2.1.3, 2.1.4 and 2.1.5, the SPDE* (2.1.2) *has a fundamental solution $\mathbf{\Gamma}$ and there exist two positive random variables μ_1 and μ_2 such that, with probability one we have*

$$\mathbf{\Gamma}(t,z;s,\zeta) \geqslant \mu_2^{-1}\Gamma^{\mathrm{heat}}\left(\mu_1^{-1}\mathcal{D}_{s-t}, g^{\mathrm{IW},-1}_{t,s}(\zeta) - \gamma_{t,s}(z)\right) \tag{2.1.5}$$

$$\mathbf{\Gamma}(t,z;s,\zeta) \leqslant \mu_2\Gamma^{\mathrm{heat}}\left(\mu_1\mathcal{D}_{s-t}, g^{\mathrm{IW},-1}_{t,s}(\zeta) - \gamma_{t,s}(z)\right), \tag{2.1.6}$$

$$|\nabla_\nu \mathbf{\Gamma}(t,z;s,\xi,\nu)| \leqslant \frac{\mu_2}{\sqrt{s-t}}\Gamma^{\mathrm{heat}}\left(\mu_1\mathcal{D}_{s-t}, g^{\mathrm{IW},-1}_{t,s}(\xi,\nu) - \gamma_{t,s}(z)\right), \tag{2.1.7}$$

$$\left|\nabla^2_\nu \mathbf{\Gamma}(t,z;s,\xi,\nu)\right| \leqslant \frac{\mu_2}{s-t}\Gamma^{\mathrm{heat}}\left(\mu_1\mathcal{D}_{s-t}, g^{\mathrm{IW},-1}_{t,s}(\xi,\nu) - \gamma_{t,s}(z)\right), \tag{2.1.8}$$

for every $0 \leqslant t < s \leqslant T$, $\zeta = (\xi,\nu), z \in \mathbb{R}^{d+1}$, where $\mathcal{D}_\lambda$ is the $(d+1)\times(d+1)$ matrix $\mathrm{diag}(\lambda^3, \lambda, \cdots, \lambda)$.

Remark 2.1.7. *We would like to emphasize that Theorem 2.1.6 is new even in the deterministic case, i.e. when $\sigma \equiv 0$, $h \equiv 0$ and the coefficients are deterministic functions. In fact, a study of*

Kolmogorov PDEs with coefficients measurable in time was only recently proposed in [7]: however in [7] the coefficients are assumed to be independent of the spatial variables that is a very particular case where the fundamental solution is known explicitly.

2.2 Stochastic Langevin equation and the Hörmander condition

For illustrative purposes, we examine the case of constant coefficients and introduce the stochastic counterpart of the classical Langevin PDE.

Let B, W be independent real Brownian motions, $a > 0$ and $\sigma \in [0, \sqrt{a}]$. The Langevin model is defined in terms of the system of SDEs

$$\begin{cases} dX_t = V_t dt, \\ dV_t = \sqrt{a-\sigma^2}dB_t - \sigma dW_t. \end{cases} \tag{2.2.1}$$

We interpret W as the observation process: if $\sigma = 0$ the velocity V is unobservable, while for $\sigma = \sqrt{a}$ the velocity V is completely observable, being equal to W. To shorten notations, we denote by $\zeta = (\xi, \nu)$ and by $z = (x, v)$ the points in $\mathbb{R}^2$. Setting $Z_t = (X_t, Y_t)$, equation (2.2.1) can be rewritten as

$$dZ_t = BZ_t dt + \mathbf{e}_2 d(\sqrt{a-\sigma^2}B_t - \sigma W_t), \tag{2.2.2}$$

with

$$B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.2.3}$$

In this section we show in two different ways that the SPDE

$$d_{\mathbf{B}} u_s = \frac{a}{2}\partial_{\nu\nu} u_s ds + \sigma \partial_\nu u_s dW_s, \qquad \mathbf{B} := \partial_s + \nu\partial_\xi, \tag{2.2.4}$$

is the *forward Kolmogorov* (or *Fokker-Planck*) equation of the SDE (2.2.1) conditioned to the Brownian observation given by $\mathcal{F}_s^W = \sigma(W_t,\, t \leqslant s)$. In the uniformly parabolic case, this is a well-known fact, proved under diverse assumptions by several authors (see, for instance, [72], [39] and [54]).

In the first approach, we solve explicitly the linear SDE (2.2.2) and find the expression of the conditional transition density $\mathbf{\Gamma}$ of the solution Z: by Itô formula, we directly infer that $\mathbf{\Gamma}$ is the fundamental solution of the SPDE (2.2.4). The second approach, inspired by [42], is much more general because it does not require the explicit knowledge of $\mathbf{\Gamma}$: we first prove the existence of the fundamental solution of the SPDE (2.2.4) and then show that it is the conditional transition density of the solution of (2.2.1).

The following result is a consequence of the Itô formula and isometry.

Proposition 2.2.1. *The solution $Z = Z^z$ of* (2.2.2)*, with initial condition $z = (x, v) \in \mathbb{R}^2$, is given by*

$$Z_s^z = e^{sB}\left(z + \int_0^s e^{-\tau B}\mathbf{e}_2\, d(\sqrt{a-\sigma^2}B_\tau - \sigma W_\tau)\right)$$

with $\mathbf{e}_2$ as in (2.2.2)*. Conditioned to $\mathcal{F}_s^W$, Z_s^z has normal distribution with mean and covariance matrix given by*

$$m_s(z) := E\left[Z_s^z \mid \mathcal{F}_s^W\right] = e^{sB}\left(z - \sigma\int_0^s e^{-\tau B}\mathbf{e}_2 dW_\tau\right) = \begin{pmatrix} x + sv - \sigma\int_0^s (s-\tau)dW_\tau \\ v - \sigma W_s \end{pmatrix}, \quad (2.2.5)$$

$$\mathcal{C}_s := cov\left(Z_s^z \mid \mathcal{F}_s^W\right) = (a-\sigma^2)Q_s, \qquad Q_s := \int_0^s \left(e^{\tau B}\mathbf{e}_2\right)\left(e^{\tau B}\mathbf{e}_2\right)^* d\tau = \begin{pmatrix} \frac{s^3}{3} & \frac{s^2}{2} \\ \frac{s^2}{2} & s \end{pmatrix}.$$

In particular, if $\sigma = \sqrt{a}$ then the distribution of Z_s^z conditioned to $\mathcal{F}_s^W$ is a Dirac delta centered at $m_s(z)$; if $\sigma \in [0, \sqrt{a})$ and $s > 0$ then Z_s^z has density, conditioned to $\mathcal{F}_s^W$, given by

$$\mathbf{\Gamma}(0, z; s, \zeta) = \frac{1}{2\pi\sqrt{\det \mathcal{C}_s}}\exp\left(-\frac{1}{2}\langle \mathcal{C}_s^{-1}(\zeta - m_s(z)), (\zeta - m_s(z))\rangle\right), \qquad \zeta \in \mathbb{R}^2. \quad (2.2.6)$$

More explicitly, we have $\mathbf{\Gamma}(0, z; s, \zeta) = \mathbf{\Gamma}_0(s, \zeta - m_s(z))$ where

$$\mathbf{\Gamma}_0(s, \xi, \nu) = \frac{\sqrt{3}}{\pi s^2(a-\sigma^2)}\exp\left(-\frac{2}{a-\sigma^2}\left(\frac{\nu^2}{s} - \frac{3\nu\xi}{s^2} + \frac{3\xi^2}{s^3}\right)\right), \qquad s > 0, \ (\xi, \nu) \in \mathbb{R}^2. \quad (2.2.7)$$

By the Itô formula, $\mathbf{\Gamma}(0, z; s, \zeta)$ is the stochastic fundamental solution of SPDE (2.2.4)*, with pole at* $(0, z)$*.*

As an alternative approach, we construct the fundamental solution of the SPDE (2.2.4). First we transform (2.2.4) into a PDE with random coefficients, satisfying the weak Hörmander condition; by a second change of variables, we remove the drift of the equation and transform it into a deterministic heat equation; going back to the original variables, we find the stochastic fundamental solution of (2.2.4), which obviously coincides with $\mathbf{\Gamma}$ in (2.2.6). Eventually, we prove that $\mathbf{\Gamma}(0, z; s, \cdot)$ is a density of Z_s^z conditioned to $\mathcal{F}_s^W$. We split the proof in three steps.

[Step 1] We set

$$\hat{u}_s(\xi, \nu) = u_s(\xi, \nu - \sigma W_s). \quad (2.2.8)$$

By Itô formula, u solves (2.2.4) if and only if $\hat{u}$ solves the Langevin PDE

$$\partial_s \hat{u} + (\nu - \sigma W_s)\partial_\xi \hat{u} = \frac{a-\sigma^2}{2}\partial_{\nu\nu}\hat{u}. \quad (2.2.9)$$

By this change of coordinates we get rid of the stochastic part of the SPDE; however, this is done at the cost of introducing a random drift term. For the moment, this is not a big issue because σ

is constant and, in particular, independent of ν: for this reason, the weak Hörmander condition is preserved since the vector fields ∂_ν, $\partial_s + (\nu - \sigma W_s)\partial_\xi$ and their Lie bracket

$$[\partial_\nu, \partial_s + (\nu - \sigma W_s)\partial_\xi] = \partial_\xi$$

span $\mathbb{R}^3$ at any point.

[Step 2] In order to remove the random drift, we perform a second change of variables:

$$\bar{u}_s(\xi,\nu) = \hat{u}_s(\gamma_s(\xi,\nu)), \qquad \gamma_s(\xi,\nu) := \left(\xi + s\nu - \sigma \int_0^s W_\tau d\tau, \nu\right). \tag{2.2.10}$$

The spatial Jacobian of γ_s equals

$$\nabla\gamma_s(\xi,\nu) = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$$

so that γ_s is one-to-one and onto $\mathbb{R}^2$ for any s. Then, (2.2.9) is transformed into the deterministic heat equation with time-dependent coefficients

$$\partial_s \bar{u}_s(\xi,\nu) = \frac{a-\sigma^2}{2}\left(s^2\partial_{\xi\xi} - 2s\partial_{\xi\nu} + \partial_{\nu\nu}\right)\bar{u}_s(\xi,\nu). \tag{2.2.11}$$

Equation (2.2.11) is not uniformly parabolic because the matrix of coefficients of the second order part

$$a_s := (a-\sigma^2)\begin{pmatrix} s^2 & -s \\ -s & 1 \end{pmatrix}$$

is singular. However, in case of partial observation, that is $\sigma \in [0, \sqrt{a})$, the diffusion matrix

$$\mathcal{C}_s = \int_0^s a_\tau d\tau = (a-\sigma^2)\begin{pmatrix} \frac{s^3}{3} & -\frac{s^2}{2} \\ -\frac{s^2}{2} & s \end{pmatrix}$$

is positive definite for any $s > 0$ and therefore (2.2.11) admits a Gaussian fundamental solution. For $\sigma = 0$, this result was originally proved by [31] (see also the introduction in [27]). Going back to the original variables we recover the explicit expression of $\mathbf{\Gamma}$ in (2.2.6).

Incidentally, we notice that (2.2.11) also reads

$$\partial_s \bar{u}_s(\xi,\nu) = \frac{a-\sigma^2}{2}\mathbf{V}_s^2 \bar{u}_s(\xi,\nu), \qquad \mathbf{V}_s := \partial_\nu - s\partial_\xi,$$

where the vector fields ∂_s and $\mathbf{V}_s$ satisfy the weak Hörmander condition in $\mathbb{R}^3$ because $[\mathbf{V}_s, \partial_s] = \partial_\xi$.

[Step 3] We show that $\mathbf{\Gamma}$ is the conditional transition density of Z: the proof is based on a combination of the arguments of [42] with the gradient estimates for Kolmogorov equations proved in [17] and anticipates some of the arguments we will develop in greater generality in Chapter 3.

Theorem 2.2.2. *Let Z^z denote the solution of the linear SDE* (2.2.2) *starting from $z \in \mathbb{R}^2$ and let $\mathbf{\Gamma} = \mathbf{\Gamma}(0,z;s,\cdot)$ in* (2.2.6) *be the fundamental solution of the Langevin SPDE* (2.2.4) *with $\sigma \in [0,\sqrt{a})$. For any bounded and measurable function φ on $\mathbb{R}^2$, we have*

$$E\left[\varphi(Z_s^z) \mid \mathcal{F}_s^W\right] = \int_{\mathbb{R}^2} \varphi(\zeta)\mathbf{\Gamma}(0,z;s,\zeta)d\zeta.$$

Proof. Let

$$I_s(z) := \int_{\mathbb{R}^2} \varphi(\zeta)\mathbf{\Gamma}(0,z;s,\zeta)d\zeta, \qquad s > 0,\ \zeta \in \mathbb{R}^2.$$

By (2.2.5)-(2.2.6), $I_s(z)$ is $\mathcal{F}_s^W$-measurable: then, to prove the thesis we show that for any bounded and $\mathcal{F}_s^W$-measurable random variable G we have

$$E\left[G\varphi(Z_s^z)\right] = E\left[GI_s(z)\right].$$

By an approximation argument it suffices to take φ in the class of test functions and G of the form $G = e^{-\int_0^s c_\tau(W_\tau)d\tau}$, where $c = c_s(w)$ is a smooth, bounded and non negative function on $[0,T] \times \mathbb{R}$. Let

$$\mathcal{L}^{(\sigma)} = \frac{1}{2}\left(a\partial_{\nu\nu} - 2\sigma\partial_{\nu\eta} + \partial_{\eta\eta}\right) + \nu\partial_\xi$$

be the infinitesimal generator of the three-dimensional process (X,V,W). For $\sigma \in [0,\sqrt{a})$, $\partial_s + \mathcal{L}^{(\sigma)}$ satisfies the weak Hörmander condition in $\mathbb{R}^4$ and has a Gaussian fundamental solution (see, for instance, formula (2.9) in [17]). We denote by $f = f_t(\xi,\nu,\eta)$ the classical solution of the *backward* Cauchy problem

$$\begin{cases} \left(\partial_t + \mathcal{L}^{(\sigma)}\right) f_t(\xi,\nu,\eta) - c_t(\eta) f_t(\xi,\nu,\eta) = 0, & (t,\xi,\nu,\eta) \in [0,s) \times \mathbb{R}^3, \\ f_s(\xi,\nu,\eta) = \varphi(x,v), & (\xi,\nu,\eta) \in \mathbb{R}^3, \end{cases}$$

and set

$$M_t := e^{-\int_0^t c_\tau(W_\tau)d\tau} \int_{\mathbb{R}^2} f_t(\zeta, W_t)\mathbf{\Gamma}(0,z;t,\zeta)dz, \qquad t \in [0,s].$$

By definition, we have

$$E\left[M_s\right] = E\left[e^{-\int_0^s c_\tau(W_\tau)d\tau} I_s(z)\right]$$

and, by the Feynman-Kac representation of f,

$$E\left[M_0\right] = f_0(z,0) = E\left[e^{-\int_0^s c_\tau(W_\tau)d\tau} \varphi(Z_s^z)\right].$$

Hence the thesis follows by proving that M is a martingale. By the Itô formula, we have

$$\begin{aligned} df_t(\xi,\nu,W_t) &= \left(\partial_t f_t + \frac{1}{2}\partial_{\eta\eta} f_t\right)(\xi,\nu,W_t)dt + (\partial_\eta f_t)(\xi,\nu,W_t)dW_t \\ &= \left(-\mathcal{L}^{(\sigma)} f_t + c_t f_t + \frac{1}{2}\partial_{\eta\eta} f_t\right)(\xi,\nu,W_t)dt + (\partial_\eta f_t)(\xi,\nu,W_t)dW_t. \end{aligned}$$

Moreover, since $\mathbf{\Gamma}$ solves the SPDE (2.2.4), setting $e_t := e^{-\int_0^t c_\tau(W_\tau)d\tau}$ for brevity, we get

$$\begin{aligned} dM_t = -c_t(W_t)M_t dt &+ e_t \int_{\mathbb{R}^2} \left(-\mathcal{L}^{(\sigma)} f_t + c_t f_t + \frac{1}{2}\partial_{\eta\eta} f_t\right)(\xi,\nu,W_t)\mathbf{\Gamma}(0,z;t,\xi,\nu)d\xi d\nu\, dt \\ &+ e_t \int_{\mathbb{R}^2} (\partial_\eta f_t)(\xi,\nu,W_t)\mathbf{\Gamma}(0,z;t,\xi,\nu)d\xi d\nu\, dW_t \\ &+ e_t \int_{\mathbb{R}^2} f_t(\xi,\nu,W_t)\left(\frac{a}{2}\partial_{\nu\nu} - \nu\partial_\xi\right)\mathbf{\Gamma}(0,z;t,\xi,\nu)d\xi d\nu\, dt \\ &+ e_t\sigma \int_{\mathbb{R}^2} f_t(\xi,\nu,W_t)\partial_\nu\mathbf{\Gamma}(0,z;t,\xi,\nu)d\xi d\ v\, dW_t \\ &+ e_t\sigma \int_{\mathbb{R}^2} \partial_\eta f_t(\xi,\nu,W_t)\partial_\nu\mathbf{\Gamma}(0,z;t,\xi,\nu)d\xi d\nu\, dt. \end{aligned}$$

Integrating by parts, we find

$$dM_t = e_t \int_{\mathbb{R}^2} (\partial_\eta f_s - \sigma\partial_\nu f_t)(\xi,\nu,W_t)\mathbf{\Gamma}(0,z;t,\xi,\nu)d\xi d\nu\, dW_t,$$

which shows that M is at least a local martingale.

To conclude, we recall the gradient estimates proved in [17], Proposition 3.3: for any test function φ there exist two positive constants ε, C such that

$$|\partial_\nu f_t(\xi,\nu,\eta)| + |\partial_\eta f_t(\xi,\nu,\eta)| \leqslant C(s-t)^{\varepsilon-\frac{1}{2}}, \qquad (t,\xi,\nu,\eta) \in [0,s) \times \mathbb{R}^3.$$

Thus, we have

$$\begin{aligned} &E\left[\int_0^s \left(\int_{\mathbb{R}^2} (\partial_\eta f_t - \sigma\partial_\nu f_t)(\xi,\nu,W_t)\mathbf{\Gamma}(0,z;t,\xi,\nu)d\xi d\nu\right)^2 dt\right] \\ &\leqslant \int_0^s C(s-t)^{2\varepsilon-1} E\left[\left(\int_{\mathbb{R}^2} \mathbf{\Gamma}(0,z;t,\xi,\nu)d\xi d\nu\right)^2\right] dt \\ &= \int_0^s C(s-t)^{2\varepsilon-1} dt < \infty \end{aligned}$$

and this proves that M is a true martingale. $\square$

2.3 Intrinsic vs Euclidean Hölder spaces for the deterministic Langevin equation

The parametrix method requires some assumption on the regularity of the coefficients of the PDE: in the uniformly parabolic case, it suffices to assume that the coefficients are bounded, Hölder continuous in the spatial variables and measurable in time (cf. [24]).

In this study, we apply the parametrix method assuming that the coefficients of the Langevin SPDE (2.1.2) are predictable processes that are Hölder continuous in the Euclidean sense. From the analytical perspective this is not the natural choice: indeed, it is well known that the natural framework for the study of Hörmander operators is the analysis on Lie groups (see, for instance, [23]). In this section, we motivate our choice to use Euclidean Hölder spaces rather than the intrinsic ones. We recall that [45] first studied the intrinsic geometry of the Langevin operator in (2.2.4) with $\sigma = 0$:

$$\mathcal{L}_a := \frac{a}{2}\partial_{\nu\nu} - \nu\partial_\xi - \partial_s.$$

They noticed that $\mathcal{L}_a$ is invariant with respect to the homogeneous Lie group $(\mathbb{R}^3, *, \delta)$ where the group law is given by

$$(t, x, v) * (s, \xi, \nu) = (t + s, x + \xi + sv, v + \nu), \qquad (t, x, v), (s, \xi, \nu) \in \mathbb{R}^3, \tag{2.3.1}$$

and $\delta = (\delta_\lambda)_{\lambda>0}$ is the ultra-parabolic dilation operator defined as

$$\delta_\lambda(t, x, v) = (\lambda^2 t, \lambda^3 x, \lambda v), \qquad (t, x, v) \in \mathbb{R}^3,\ \lambda > 0.$$

More precisely, $\mathcal{L}_a$ is invariant with respect to the left-$*$-translations $\ell_{(t,x,v)}(s, \xi, \nu) = (t, x, v) * (s, \xi, \nu)$, in the sense that

$$\mathcal{L}_a(f \circ \ell_{(t,x,v)}) = (\mathcal{L}_a f) \circ \ell_{(t,x,v)}, \qquad (t, x, v) \in \mathbb{R}^3,$$

and is δ-homogeneous of degree two, in that

$$\mathcal{L}_a(f \circ \delta_\lambda) = \lambda^2 (\mathcal{L}_a f) \circ \delta_\lambda, \qquad \lambda > 0.$$

It is natural to endow $(\mathbb{R}^3, *, \delta)$ with the δ-homogeneous norm

$$|(t, x, v)|_{\mathcal{L}} = |t|^{\frac{1}{2}} + |x|^{\frac{1}{3}} + |v| \tag{2.3.2}$$

and the distance

$$d_{\mathcal{L}}\left((s, \xi, \nu), (t, x, v)\right) = |(t, x, v)^{-1} * (s, \xi, \nu)|_{\mathcal{L}}. \tag{2.3.3}$$

The intrinsic Hölder spaces associated to $d_{\mathcal{L}}$ are particularly beneficial for the study of existence and regularity properties of solutions to the Langevin equation because they comply with the asymptotic properties of its fundamental solution $\mathbf{\Gamma}$ near the pole: let us recall that

$$\mathbf{\Gamma}(t,x,v;s,\xi,\nu) = \mathbf{\Gamma}_0\left((t,x,v)^{-1} * (s,\xi,\nu)\right), \qquad t<s,$$

where $\mathbf{\Gamma}_0$ is the fundamental solution of $\mathcal{L}$ in (2.2.7) with $\sigma = 0$ and $(t,x,v)^{-1} = (-t,-x+tv,-v)$ is the $*$-inverse of (t,x,v). Notice also that $\mathbf{\Gamma}$ is δ-homogeneous of degree four, where four is the so-called δ-homogeneous dimension of $\mathbb{R}^2$.

Based on the use of intrinsic Hölder spaces defined in terms of $d_{\mathcal{L}}$, a stream of literature has built a complete theory of existence and regularity, analogous to that for uniformly parabolic PDEs: we mention some of the main contributions [58], [59], [47], [46], [16], [18], [53] and, in particular, [58], [16], [32] where the parametrix method for Kolmogorov PDEs was developed.

On the other hand, intrinsic Hölder regularity can be a rather restrictive property as shown by the following example.

Example 2.3.1. *For $x,\xi \in \mathbb{R}$ and $s \neq t$, let*

$$z = \left(x, -\frac{x-\xi}{s-t}\right), \qquad \zeta = \left(\xi, -\frac{x-\xi}{s-t}\right) \tag{2.3.4}$$

Then we have

$$(t,z)^{-1} * (s,\zeta) = (s-t,0,0),$$

and therefore

$$d_{\mathcal{L}}((t,z),(s,\zeta)) = |s-t|^{\frac{1}{2}}.$$

Since x and ξ are arbitrary real numbers, we see that points in $\mathbb{R}^3$ that are far from each other in the Euclidean sense, can be very close in the intrinsic sense. It follows that, if a function $f(t,x,v) = f(x)$ depends only on x and is Hölder continuous in the intrinsic sense (i.e. with respect to $d_{\mathcal{L}}$), then it must be constant: in fact, for z,ζ as in (2.3.4), we have

$$|f(\xi)-f(x)| = |f(s,\zeta)-f(t,z)| \leqslant C|s-t|^{\alpha}$$

for some positive constants C,α and for any $x,\xi \in \mathbb{R}$ and $s \neq t$.

When it comes to studying the stochastic Langevin equation, the use of Euclidean Hölder spaces seems unavoidable. The problem is that we have to deal with functions $f = f_t(x,v)$ that are:

- *Hölder continuous* with respect to the space variables (x,v) in order to apply the parametrix method;

- *measurable* with respect to the time variable t because f plays the role of a coefficient of the SPDE that is a predictable process (i.e. merely measurable in t).

As opposed to the standard parabolic case, in terms of the metric $d_{\mathcal{L}}$ there doesn't seem to be a clear way to separate regularity in (x, v) from regularity in t: indeed this is due to the definition of $*$-translation that mixes up spatial and time variables (see (2.3.1)). On the other hand, we may observe that the Euclidean- and $*$- differences of points at the same time level coincide:

$$(t, x, v)^{-1} * (t, \xi, \nu) = (0, \xi - x, \nu - v), \qquad x, v, \xi, \nu \in \mathbb{R}.$$

Thus, to avoid using $*$-translations, the idea is to combine this property with a suitable definition of *time-dependent parametrix* that makes the parametrix procedure work: this will be done in Section 2.6.

Concerning the use of the Euclidean or homogeneous norm in $\mathbb{R}^2$, let us denote by $bC^\alpha(\mathbb{R}^2)$ and $bC^\alpha_{\mathcal{L}}(\mathbb{R}^2)$ the space of bounded and Hölder continuous functions with respect to the Euclidean norm and the homogeneous norm $|x|^{\frac{1}{3}} + |v|$, respectively. Since $|(x, v)| \leqslant |x|^{\frac{1}{3}} + |v|$ for $|(x, v)| \leqslant 1$, we have the inclusion

$$bC^\alpha(\mathbb{R}^2) \subseteq bC^\alpha_{\mathcal{L}}(\mathbb{R}^2). \tag{2.3.5}$$

Preferring simplicity to generality, we shall use Hölder spaces defined in terms of the Euclidean norm (cf. Assumption 2.1.3): by (2.3.5), this results in a slightly more restrictive condition compared to the analogous one given in terms of the homogeneous norm. On the other hand, all the results of this Chapter can be proved using the homogeneous norm $|x|^{\frac{1}{3}} + |v|$ as in [58], [16] and [32]: this would be more natural but would increase the technicalities.

We close this section by giving some standard Gaussian estimates that play a central role in the parametrix construction. After the change of variables (2.2.10) with $\sigma = 0$, the Langevin operator $\mathcal{L}_a$ is transformed into

$$L_a = \frac{a}{2}\mathbf{V}_s^2 - \partial_s, \qquad \mathbf{V}_s := \partial_\nu - s\partial_\xi.$$

Since L_a is a heat operator with time dependent coefficients, its fundamental solution is the Gaussian function $\mathbf{\Gamma}_a(t, z; s, \zeta) = \mathbf{\Gamma}_a(t, 0; s, \zeta - z)$ where

$$\mathbf{\Gamma}_a(t, 0, 0; s, \xi, \eta) = \frac{\sqrt{3}}{a\pi(s-t)^2} \exp\left(-\frac{2}{a(s-t)^3}\left(3\xi^2 + 3\xi\eta(t+s) + \eta^2(t^2 + ts + s^2)\right)\right) \tag{2.3.6}$$

for $t < s$ and $\xi, \eta \in \mathbb{R}$.

Lemma 2.3.2. *For every $\varepsilon > 0$ there exists a positive constant c such that*

$$\begin{aligned} |\mathbf{V}_s \mathbf{\Gamma}_a(t,0,0;s,\xi,\eta)| &\leqslant \frac{c}{\sqrt{s-t}} \mathbf{\Gamma}_{a+\varepsilon}(t,0,0;s,\xi,\eta), \\ \left|\mathbf{V}_s^2 \mathbf{\Gamma}_a(t,0,0;s,\xi,\eta)\right| &\leqslant \frac{c}{s-t} \mathbf{\Gamma}_{a+\varepsilon}(t,0,0;s,\xi,\eta), \end{aligned} \tag{2.3.7}$$

for every $0 \leqslant t < s \leqslant T$ and $\xi, \eta \in \mathbb{R}$.

Proof. We remark that $\mathbf{\Gamma}_a(t,0,0;s,\xi,\eta)$ has different asymptotic regimes as $s \to t^+$ depending on whether or not t is zero: in fact, if $t = 0$ then the quadratic form in the exponent of $\mathbf{\Gamma}_a$ is similar to that of the Langevin operator, that is

$$\frac{1}{a} \begin{pmatrix} \frac{6}{s^3} & \frac{3}{s^2} \\ \frac{3}{s^2} & \frac{2}{s} \end{pmatrix}.$$

On the other hand, if $t \neq 0$ we see in (2.3.6) that all the components of the quadratic form are $O((s-t)^{-3})$ as $s \to t^+$. We have

$$|\mathbf{V}_s \mathbf{\Gamma}_a(t,0,0;s,\xi,\eta)| = \frac{1}{\sqrt{s-t}} \left| \frac{6\xi + 2\eta(s+2t)}{a(s-t)^{\frac{3}{2}}} \right| \mathbf{\Gamma}_a(t,0,0;s,\xi,\eta) \leqslant$$

(by (1.4.18) with $n = 1$)

$$\leqslant \frac{C}{\sqrt{s-t}} \mathbf{\Gamma}_{a+\varepsilon}(t,0,0;s,\xi,\eta).$$

The proof of (2.3.7) is similar, using that

$$\mathbf{V}_s^2 \mathbf{\Gamma}_a(t,0,0;s,\xi,\eta) = \frac{4}{a(s-t)} \left(\frac{(3\xi + \eta(s+2t))^2}{a(s-t)^3} - 1 \right) \mathbf{\Gamma}_a(t,0,0;s,\xi,\eta).$$

$\square$

2.4 Pointwise estimates for Itô processes

In this section we build on the ideas of Lemma 1.3.3 and Proposition 1.3.2 and prove some estimates for stochastic flows of diffeomorphisms that will play a central role in the following sections. Information about stochastic flows in a more general framework can be found in [44]. Since the following results are of a general nature and may be of independent interest, in this section we reset the notations and give the proofs in the more general multi-dimensional setting.

Specifically, until the end of the section, the point of $\mathbb{R}^d$ is denoted by $z = (z_1, \dots, z_d)$ and we set $\nabla_z = (\partial_{z_1}, \dots, \partial_{z_d})$ and $\partial^\beta = \partial_z^\beta = \partial_{z_1}^{\beta_1} \cdots \partial_{z_d}^{\beta_d}$ for any multi-index β. We will also employ the notation

$$\langle z \rangle := \sqrt{1 + |z|^2}, \qquad z \in \mathbb{R}^d.$$

First, we recall some basic facts about stochastic flows of diffeomorphisms. Let $k \in \mathbb{N}$. A $\mathbb{R}^d$-valued random field $\varphi_{t,s}(z)$, with $0 \leqslant t \leqslant s \leqslant T$ and $z \in \mathbb{R}^d$, defined on $(\Omega, \mathcal{F}, P)$, is called a (forward) stochastic flow of C^k-diffeomorphisms if there exists a set of probability one where:

i) $\varphi_{t,t}(z) = z$ for any $t \in [0,T]$ and $z \in \mathbb{R}^d$;

ii) $\varphi_{t,s} = \varphi_{\tau,s} \circ \varphi_{t,\tau}$ for $0 \leqslant t \leqslant \tau \leqslant s \leqslant T$;

iii) $\varphi_{t,s} : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is a C^k-diffeomorphism for all $0 \leqslant t \leqslant s \leqslant T$.

Stochastic flows can be constructed as solutions of stochastic differential equations. Let B a n-dimensional Brownian motion and consider the stochastic differential equation

$$\varphi_s(z) = z + \int_t^s b_\tau(\varphi_\tau(z))d\tau + \int_t^s \sigma_\tau(\varphi_\tau(z))dB_\tau \tag{2.4.1}$$

where $b = (b_t^i(z))$, $\sigma = (\sigma_s^{ij}(z))$ are a d-valued and $(d \times n)$-valued processes respectively, on $[0,T] \times \mathbb{R}^d \times \Omega$. The following theorem summarizes the results of Lemmas 4.5.3-7 and Theorems 4.6.4-5 in [44].

Theorem 2.4.1. *Let b, $\sigma \in \mathbf{bC}_{0,T}^{k,\alpha}$ for some $k \in \mathbb{N}$ and $\alpha \in (0,1)$. Then the solution of the stochastic differential equation* (2.4.1) *has a modification $\varphi_{t,s}$ that is a forward stochastic flow of C^k-diffeomorphisms. Moreover, $\varphi_{t,\cdot} \in \mathbf{C}_{t,T}^{k,\alpha'}$ for any $\alpha' \in [0,\alpha)$ and $t \in [0,T)$, and we have the following estimates: for each $p \in \mathbb{R}$ there exists a positive constant $\mathbf{c}_{1,p}$ such that*

$$E\left[\langle \varphi_{t,s}(z)\rangle^p\right] \leqslant \mathbf{c}_{1,p}\langle z\rangle^p, \qquad z \in \mathbb{R}^d, \tag{2.4.2}$$

and for each $p \geqslant 1$ there exists a positive constant $\mathbf{c}_{2,p}$ such that

$$E\left[\left|\partial^\beta \varphi_{t,s}(z)\right|^p\right] \leqslant \mathbf{c}_{2,p}, \qquad z \in \mathbb{R}^d,\ p \geqslant 1,\ 1 \leqslant |\beta| \leqslant k. \tag{2.4.3}$$

Now, consider $\varphi_{t,s}$ as in Theorem 2.4.1, $F_i = F_{i,t}(z;\zeta) \in \mathbf{C}_{0,T}^k(\mathbb{R}^{2d})$, $i = 1,2$, and a real Brownian motion W. The goal of this section is to prove some pointwise estimate for the Itô process

$$I_{t,s}(z) := \int_t^s F_{1,\tau}(z;\varphi_{t,\tau}(z))dW_\tau + \int_t^s F_{2,\tau}(z;\varphi_{t,\tau}(z))d\tau, \qquad 0 \leqslant t \leqslant s \leqslant T,\ z \in \mathbb{R}^d, \tag{2.4.4}$$

in terms of the usual Hölder norm in $\mathbb{R}^d$

$$|f|_\alpha = \sup_{z\in\mathbb{R}^d} |f(z)| + \sup_{\substack{z,\zeta\in\mathbb{R}^d \\ z\neq\zeta}} \frac{|f(z) - f(\zeta)|}{|z-\zeta|^\alpha}, \qquad \alpha \in (0,1),$$

under the following

Assumption 2.4.2. *There exist $\varepsilon_1, \varepsilon_2 \in \mathbb{R}$ with $\varepsilon := \varepsilon_1 + \varepsilon_2 > 0$ and a random variable $M \in L^{\bar{p}}(\Omega)$, with $\bar{p} > \left(2 \vee d \vee \frac{d}{\varepsilon}\right)$, such that*

$$\sum_{|\beta| \leqslant k} \sup_{\substack{t \in [0,T] \\ z,\zeta \in \mathbb{R}^d}} \langle z \rangle^{\varepsilon_1} \langle \zeta \rangle^{\varepsilon_2} |\partial^{\beta}_{z,\zeta} F_{i,t}(z;\zeta)| \leqslant M \qquad i = 1,2, \ P\text{-a.s.}$$

The main result of this section is the following theorem which provides global-in-space pointwise estimates for the process in (2.4.4).

Theorem 2.4.3. *Let $\varphi_{t,s}$ be as in Theorem 2.4.1 and $F_i \in \mathbf{C}^k_{0,T}(\mathbb{R}^{2d})$, $i = 1,2$, for some $k \in \mathbb{N}$. Let $I = I_{t,s}(z)$ be as in (2.4.4) and set*

$$I^{(\delta)}_{t,s}(z) := \langle z \rangle^{\delta} I_{t,s}(z).$$

Under Assumption 2.4.2, for any p, α and δ such that

$$\left(2 \vee d \vee \frac{d}{\varepsilon}\right) < p < \bar{p}, \qquad 0 \leqslant \alpha < \frac{1}{2} - \frac{1}{p}, \qquad 0 \leqslant \delta < \varepsilon - \frac{d}{p},$$

there exists a (random, finite) constant $\mathbf{m}$ such that

$$\sum_{|\beta| \leqslant k-1} |\partial^{\beta} I^{(\delta)}_{t,s}|_{1-\frac{d}{p}} \leqslant \mathbf{m}(s-t)^{\alpha} \qquad P\text{-a.s.} \tag{2.4.5}$$

Proof. The proof is based on a combination of sharp L^p-estimates, Kolmogorov continuity theorem in Banach spaces and Sobolev embedding theorem.

Let us first consider the case $k = 1$. We prove some preliminary L^p-estimates for $I_{t,s}$ and $\partial^{\beta} I_{t,s}$ with $|\beta| = 1$. Below we denote by $\bar{c}$ various positive constants that depend only on p, d, T and the flow φ. By Burkölder's inequality we have

$$E\left[|I^{(\delta)}_{t,s}(z)|^p\right] \leqslant \bar{c}\langle z \rangle^{\delta p} E\left[\left(\int_t^s F^2_{1,\tau}(z;\varphi_{t,\tau}(z))d\tau\right)^{\frac{p}{2}}\right] + \bar{c}\langle z \rangle^{\delta p} E\left[\left(\int_t^s F_{2,\tau}(z;\varphi_{t,\tau}(z))d\tau\right)^{p}\right] \leqslant$$

(by Hölder's inequality)

$$\begin{aligned} &\leqslant \bar{c}\langle z \rangle^{\delta p}(s-t)^{\frac{p-2}{2}} \int_t^s E\left[|F_{1,\tau}(z;\varphi_{t,\tau}(z))|^p\right] d\tau \\ &\quad + \bar{c}\langle z \rangle^{\delta p}(s-t)^{p-1} \int_t^s E\left[|F_{2,\tau}(z;\varphi_{t,\tau}(z))|^p\right] d\tau \leqslant \end{aligned}$$

(by Assumption 2.4.2)

$$\leqslant \bar{c}\langle z \rangle^{(\delta-\varepsilon_1)p}(s-t)^{\frac{p-2}{2}} \int_t^s E\left[M^p \langle \varphi_{t,\tau}(z) \rangle^{-\varepsilon_2 p}\right] d\tau \leqslant$$

(by Hölder's inequality with conjugate exponents $q := \frac{\bar{p}}{p}$ and r)

$$\leqslant \bar{c}\langle z\rangle^{(\delta-\varepsilon_1)p}(s-t)^{\frac{p-2}{2}}\|M\|^p_{L^{\bar{p}}(\Omega)}\int_t^s E\left[\langle\varphi_{t,\tau}(z)\rangle^{-\varepsilon_2 pr}\right]^{\frac{1}{r}} d\tau \leqslant$$

(by (2.4.2))

$$= \bar{c}\langle z\rangle^{(\delta-\varepsilon)p}(s-t)^{\frac{p}{2}}. \tag{2.4.6}$$

The same estimate holds for the gradient of $I_{t,s}$, that is

$$E\left[|\nabla I^{(\delta)}_{t,s}(z)|^p\right] \leqslant \bar{c}\langle z\rangle^{(\delta-\varepsilon)p}(s-t)^{\frac{p}{2}}. \tag{2.4.7}$$

Indeed, let us consider for simplicity only the case $\delta = 0$ since the general case is a straightforward consequence of the product rule: for $j = 1, \dots, d$, we have

$$\begin{aligned} E\left[|\partial_{z_j} I_{t,s}(z)|^p\right] &\leqslant \bar{c}E\left[\left|\int_t^s \Big((\partial_{z_j}F_{1,\tau})(z;\varphi_{t,\tau}(z)) + \langle\nabla_\zeta F_{1,\tau}(z;\varphi_{t,\tau}(z)), \partial_{z_j}\varphi_{t,\tau}(z)\rangle\Big) dW_\tau\right|^p\right] \\ &\quad + \bar{c}E\left[\left|\int_t^s \Big((\partial_{z_j}F_{2,\tau})(z;\varphi_{t,\tau}(z)) + \langle\nabla_\zeta F_{2,\tau}(z;\varphi_{t,\tau}(z)), \partial_{z_j}\varphi_{t,\tau}(z)\rangle\Big) d\tau\right|^p\right] \\ &\leqslant \bar{c}(s-t)^{\frac{p-2}{2}}\int_t^s E\left[|(\partial_{z_j}F_{i,\tau})(z;\varphi_{t,\tau}(z))|^p + \left|\langle\nabla_\zeta F_{i,\tau}(z;\varphi_{t,\tau}(z)), \partial_{z_j}\varphi_{t,\tau}(z)\rangle\right|^p\right] d\tau. \end{aligned}$$

The terms containing $\partial_{z_j}F_{i,\tau}$ can be estimated as before, by means of Assumption 2.4.2. On the other hand, by Hölder's inequality with conjugate exponents q and r with $1 < q < \frac{\bar{p}}{p}$, for every $i, j = 1, \cdots, d$ we have

$$E\left[\left|\langle\nabla_\zeta F_{i,\tau}(z;\varphi_{t,\tau}(z)), \partial_{z_j}\varphi_{t,\tau}(z)\rangle\right|^p\right] \leqslant E\left[|\nabla_\zeta F_{i,\tau}(z;\varphi_{t,\tau}(z))|^{pq}\right]^{\frac{1}{q}} E\left[\left|\partial_{z_j}\varphi_{t,\tau}(z)\right|^{pr}\right]^{\frac{1}{r}} \leqslant$$

(by Assumption 2.4.2 and (2.4.3))

$$\leqslant \bar{c}_{2,pr}^{\frac{1}{r}} E\left[M^{pq}\langle\varphi_{t,\tau}(z)\rangle^{-\varepsilon_2 pq}\right]^{\frac{1}{q}}\langle z\rangle^{-\varepsilon_1 p} \leqslant$$

(by Hölder inequality with conjugate exponents $\bar{q} := \frac{\bar{p}}{pq} > 1$ and $\bar{r}$)

$$\leqslant \bar{c}_{2,pr}^{\frac{1}{r}}\|M\|^p_{L^{\bar{p}}(\Omega)} E\left[\langle\varphi_{t,\tau}(z)\rangle^{-\varepsilon_2 pq\bar{r}}\right]^{\frac{1}{q\bar{r}}}\langle z\rangle^{-\varepsilon_1 p} \leqslant$$

(by (2.4.2))

$$\leqslant \bar{c}\|M\|^p_{L^{\bar{p}}(\Omega)}\langle z\rangle^{-\varepsilon p}.$$

This proves (2.4.7) with $\delta = 0$.

Now, we have

$$E\left[\|I^{(\delta)}_{t,s}\|^p_{W^{1,p}(\mathbb{R}^d)}\right] = E\left[\int_{\mathbb{R}^d}\Big(|I^{(\delta)}_{t,s}(z)|^p + |\nabla I^{(\delta)}_{t,s}(z)|^p\Big) dz\right] \leqslant$$

(by (2.4.6) and (2.4.7))

$$\leqslant \bar{c}(s-t)^{\frac{p}{2}} \int_{\mathbb{R}^d} \langle z\rangle^{(\delta-\varepsilon)p} dz =$$

(since $(\varepsilon-\delta)p > d$ by assumption)

$$= \bar{c}(s-t)^{\frac{p}{2}}. \tag{2.4.8}$$

Estimate (2.4.8) and Kolmogorov's continuity theorem for processes with values in the Banach space $W^{1,p}(\mathbb{R}^d)$ (see, for instance, [44], Theor.1.4.1) yield

$$\|I_{t,s}^{(\delta)}\|_{W^{1,p}(\mathbb{R}^d)} \leqslant \mathbf{m}(s-t)^{\alpha}, \qquad 0 \leqslant t \leqslant s \leqslant T, \ P\text{-a.s.}$$

for some positive and finite random variable $\mathbf{m}$ and for $\alpha \in [0, \frac{p-2}{2p})$. This is sufficient to prove (2.4.5) with $k=1$: in fact, by the Sobolev embedding theorem, we have the following estimate of the Hölder norm

$$|I_{t,s}^{(\delta)}|_{1-\frac{d}{p}} \leqslant N\|I_{t,s}^{(\delta)}\|_{W^{1,p}(\mathbb{R}^d)} \tag{2.4.9}$$

where N is a positive constant that depends only on p and d. Thus, combining (2.4.5) and (2.4.9), we get the thesis with $k=1$.

Noting that

$$\begin{aligned}\partial_{z_j} I_{t,s}(z) = &\int_t^s \Big((\partial_{z_j} F_{1,\tau})(z;\varphi_{t,\tau}(z)) + \langle \nabla_\zeta F_{1,\tau}(z;\varphi_{t,\tau}(z)), \partial_{z_j}\varphi_{t,\tau}(z)\rangle \Big) dW_\tau \\ &+ \int_t^s \Big((\partial_{z_j} F_{2,\tau})(z;\varphi_{t,\tau}(z)) + \langle \nabla_\zeta F_{2,\tau}(z;\varphi_{t,\tau}(z)), \partial_{z_j}\varphi_{t,\tau}(z)\rangle \Big) d\tau,\end{aligned}$$

for $j=1,\dots,d$, the thesis with $k=2$ can be proved repeating the previous arguments and using (2.4.5) for $k=1$ and Assumption 2.4.2 with $k=2$.

We omit the complete proof for brevity and since in the rest of the Chapter we will use (2.4.5) only for $k=1,2$. The general result can be proved by induction, using the multi-variate Faà di Bruno's formula. □

Remark 2.4.4. *Let $I_{t,s}$ as in (2.4.4) with coefficients $\widetilde{F}_1, \widetilde{F}_2 \in b\mathbf{C}^1_{0,T}(\mathbb{R}^{2d})$ and let $\delta > 0$ and $\alpha \in [0, \frac{1}{2})$. Applying Theorem 2.4.3 with $F_{i,s}(z;\zeta) := \langle z\rangle^{-\delta}\widetilde{F}_{i,s}(z;\zeta)$, $i=1,2$, we get the existence of a (random, finite) constant $\mathbf{m}$ such that, with probability one,*

$$|I_{t,s}(z)| \leqslant \mathbf{m}\langle z\rangle^{\delta}(s-t)^{\alpha}, \qquad 0 \leqslant t \leqslant s \leqslant T, \ z \in \mathbb{R}^d.$$

2.5 Itô-Wentzell change of coordinates

We go back to the main SPDE (2.1.2) and suppose that Assumptions 1.1.3, 1.1.2, 1.1.4 are satisfied and $d = 1$. In this section we study the properties of a random change of variables which plays the same role as transformation (2.2.8) in Step 1 of Section 2.2 for the Langevin SPDE. The main result of this section is Theorem 4.1.5 which shows that this change of variables transforms SPDE (2.1.2) into a PDE with random coefficients.

We denote by $(\xi, \gamma^{\mathrm{IW}}_{t,s}(\xi, \nu))$ the stochastic flow of diffeomorphisms of $\mathbb{R}^2$ defined by equation (2.1.3), that is

$$\gamma^{\mathrm{IW}}_{t,s}(\xi, \nu) = \nu - \int_t^s \sigma_\tau(\xi, \gamma^{\mathrm{IW}}_{t,\tau}(\xi, \nu))dW_\tau, \qquad s \in [t, T],\ (\xi, \nu) \in \mathbb{R}^2. \tag{2.5.1}$$

By Theorem 2.4.1, $\gamma^{\mathrm{IW}}_{t,\cdot} \in \mathbf{C}^{3,\alpha'}_{t,T}$ for any $\alpha' \in [0, \alpha)$. Global estimates for γ^{IW} and its derivatives are provided in the next:

Lemma 2.5.1. *There exists $\varepsilon \in \left(0, \frac{1}{2}\right)$ and a random, finite constant $\mathbf{c}$ such that, with probability one,*

$$|\gamma^{\mathrm{IW}}_{t,s}(\xi, \nu)| \leqslant \mathbf{c}\sqrt{1 + \xi^2 + \nu^2}, \tag{2.5.2}$$

$$e^{-\mathbf{c}(s-t)^\varepsilon} \leqslant \partial_\nu \gamma^{\mathrm{IW}}_{t,s}(\xi, \nu) \leqslant e^{\mathbf{c}(s-t)^\varepsilon}, \tag{2.5.3}$$

$$|\partial_\xi \gamma^{\mathrm{IW}}_{t,s}(\xi, \nu)| \leqslant \mathbf{c}(s-t)^\varepsilon, \tag{2.5.4}$$

$$|\partial^\beta \gamma^{\mathrm{IW}}_{t,s}(\xi, \nu)| \leqslant \frac{\mathbf{c}(s-t)^\varepsilon}{\sqrt{1 + \xi^2 + \nu^2}}, \tag{2.5.5}$$

for any $(\xi, \nu) \in \mathbb{R}^2$, $0 \leqslant t \leqslant s \leqslant T$ and $|\beta| = 2$.

Proof. Estimate (2.5.2) follows directly from Remark 2.4.4 (with $\delta = 1$). Differentiating (2.5.1), we find that $\partial_\nu \gamma^{\mathrm{IW}}_{t,s}$ solves the linear SDE

$$\partial_\nu \gamma^{\mathrm{IW}}_{t,s}(\xi, \nu) = 1 - \int_t^s (\partial_2 \sigma_\tau)(\xi, \gamma^{\mathrm{IW}}_{t,\tau}(\xi, \nu))\partial_\nu \gamma^{\mathrm{IW}}_{t,\tau}(\xi, \nu)dW_\tau,$$

where $\partial_2 \sigma_s$ denotes the partial derivative of $\sigma_s(\cdot, \cdot)$ with respect to its second argument. Hence we have

$$\partial_\nu \gamma^{\mathrm{IW}}_{t,s}(\xi, \nu) = \exp\left(-\int_t^s (\partial_2 \sigma_\tau)(\xi, \gamma^{\mathrm{IW}}_{t,\tau}(\xi, \nu))dW_\tau - \frac{1}{2}\int_t^s (\partial_2 \sigma_\tau)^2(\xi, \gamma^{\mathrm{IW}}_{t,\tau}(\xi, \nu))d\tau\right).$$

Now we apply Theorem 2.4.3 with $\varphi_{t,s}(\xi, \nu) = (\xi, \gamma^{\mathrm{IW}}_{t,s}(\xi, \nu))$ and $F_{i,s}(\zeta; \xi, V) = (\partial_2 \sigma_s(\xi, V))^i$, $i = 1, 2$: thanks to Assumption 1.1.4, we get estimate (2.5.3). Incidentally, from Theorem 2.4.3 we also deduce that the first order derivatives of $\partial_\nu \gamma^{\mathrm{IW}}_{t,s}$ are bounded:

$$|\partial^\beta \partial_\nu \gamma^{\mathrm{IW}}_{t,s}(\xi, \nu)| \leqslant \mathbf{c}(s-t)^\varepsilon, \qquad |\beta| = 1. \tag{2.5.6}$$

This last estimate is used in the next step, for the proof of (2.5.4).

Similarly, we have

$$\partial_\xi \gamma^{\mathrm{IW}}_{t,s}(\xi,\nu) = -\int_t^s \big((\partial_1\sigma_\tau)(\xi,\gamma^{\mathrm{IW}}_{t,\tau}(\xi,\nu)) + (\partial_2\sigma_\tau)(\xi,\gamma^{\mathrm{IW}}_{t,\tau}(\xi,\nu))\partial_\xi\gamma^{\mathrm{IW}}_{t,\tau}(\xi,\nu)\big)dW_\tau.$$

Thus, we have a linear SDE whose solution is given by

$$\begin{aligned}\partial_\xi \gamma^{\mathrm{IW}}_{t,s}(\xi,\nu) = {}& -\partial_\nu\gamma^{\mathrm{IW}}_{t,s}(\xi,\nu)\int_t^s \frac{(\partial_1\sigma_\tau)(\xi,\gamma^{\mathrm{IW}}_{t,\tau}(\xi,\nu))}{\partial_\nu\gamma^{\mathrm{IW}}_{t,\tau}(\xi,\nu)}dW_\tau \\ & - \partial_\nu\gamma^{\mathrm{IW}}_{t,s}(\xi,\nu)\int_t^s \frac{(\partial_1\sigma_\tau)(\xi,\gamma^{\mathrm{IW}}_{t,\tau}(\xi,\nu))(\partial_2\sigma_\tau)(\xi,\gamma^{\mathrm{IW}}_{t,\tau}(\xi,\nu))}{\partial_\nu\gamma^{\mathrm{IW}}_{t,\tau}(\xi,\nu)}d\tau,\end{aligned}$$

Again, (2.5.4) follows from Theorem 2.4.3 thanks to Assumption 1.1.4 and estimates (2.5.3) and (2.5.6).

Eventually, the same argument can be used to prove (2.5.5): indeed, differentiating (2.5.1) we have that $\partial^\beta\gamma^{\mathrm{IW}}_{t,s}$ satisfies a linear SDE whose solution is explicit. Thus, for $|\beta| = 2$, $\partial^\beta\gamma^{\mathrm{IW}}_{t,s}$ can be expressed in the form (2.4.4) with the coefficients satisfying Assumption 2.4.2 for some $\varepsilon > 1$. Applying Theorem 2.4.3 with $\delta = 1$ we get estimate (2.5.5). □

We introduce the "hat" operator which transforms any function $f_s(\xi,\nu)$, $s \in [\tau,T]$, into

$$\hat f_{\tau,s}(\xi,\nu) := f_s(\xi,\gamma^{\mathrm{IW}}_{\tau,s}(\xi,\nu)).$$

Let $u_s(\xi,\nu)$ a solution to (2.1.2) on $[\tau,T]$. Then we define

$$v_{\tau,s}(\zeta) := \varrho_{\tau,s}(\zeta)\hat u_{\tau,s}(\zeta), \quad \varrho_{\tau,s}(\zeta) := \exp\left(-\int_\tau^s \hat h_t(\zeta)dW_t - \frac{1}{2}\int_\tau^s \hat h_t^2(\zeta)dt\right).$$

We have the following

Theorem 2.5.2. *u_s is a solution to the SPDE (2.1.2) on $[\tau,T]$ if and only if $v_{\tau,s}$ is a solution on $[\tau,T]$ to the PDE with random coefficients*

$$d_{\hat{\mathbf{B}}}v_{\tau,s}(\zeta) = \big(a^*_{\tau,s}(\zeta)\partial_{\nu\nu}v_{\tau,s} + b^*_{\tau,s}(\zeta)\partial_\nu v_{\tau,s}(\zeta) + c^*_{\tau,s}(\zeta)v_{\tau,s}(\zeta)\big)\,ds, \qquad \hat{\mathbf{B}} = \partial_s + \mathbf{Y}_{\tau,s}, \tag{2.5.7}$$

where

$$\mathbf{Y}_{\tau,s} = \mathbf{Y}_{\tau,s}(\xi,\nu) := (\gamma^{\mathrm{IW}}_{\tau,s})_1(\xi,\nu)\partial_\xi - (\gamma^{\mathrm{IW}}_{\tau,s}(\xi,\nu))_1(\partial_\nu\gamma^{\mathrm{IW}}_{\tau,s})^{-1}(\xi,\nu)\partial_\xi\gamma^{\mathrm{IW}}_{\tau,s}(\xi,\nu)\partial_\nu, \tag{2.5.8}$$

is the first order operator identified with the vector field in (2.1.4) (with $d = 1$) and the coefficients $a^*_{\tau,\cdot}$, $b^*_{\tau,\cdot}$, $c^*_{\tau,\cdot}$ are defined in (2.5.11) below. Moreover, $a^*_{\tau,\cdot} \in \mathbf{bC}^\alpha_{\tau,T}$, $b^*_{\tau,\cdot}, c^*_{\tau,\cdot} \in \mathbf{bC}^0_{\tau,T}$, $\mathbf{Y}_{\tau,\cdot} \in \mathbf{C}^{0,1}_{\tau,T}$, $\partial_\nu(\mathbf{Y}_{\tau,\cdot})_1 \in \mathbf{bC}^{\bar\alpha}_{\tau,T}$ for any $\bar\alpha \in [0,\alpha)$, and there exist two random, finite and positive constants $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_2$ such that, for $s \in [\tau,T]$, $\zeta \in \mathbb{R}^2$, we have

$$\boldsymbol{\lambda}_1^{-1} \leqslant a^*_{\tau,s}(\zeta) \leqslant \boldsymbol{\lambda}_1, \qquad \boldsymbol{\lambda}_2^{-1} \leqslant \partial_\nu(\mathbf{Y}_{\tau,s}(\zeta))_1 \leqslant \boldsymbol{\lambda}_2, \tag{2.5.9}$$

with probability one.

Proof. By a standard regularization argument, we may assume $u \in \mathbf{C}^2_{\tau,T}$ so that equation (2.1.2) can be written in the usual Itô sense, namely

$$du_s(\zeta) = (\mathcal{A}_{s,\zeta} - \nu_1\partial_\xi)u_s(\zeta)ds + \mathcal{G}_{s,\zeta}u_s(s)dW_s.$$

By the Itô-Wentzell formula 1.3.1 and the chain rule we have

$$\begin{aligned} d\hat{u}_{\tau,s} &= \left(\widehat{\mathcal{A}_{s,\zeta}u}_{\tau,s} - \gamma^{\mathrm{IW}}_{\tau,s}\widehat{\partial_1 u}_{\tau,s} + \frac{1}{2}\hat{\sigma}^2_{\tau,s}\widehat{\partial_2^2 u}_{\tau,s} - \widehat{\partial_2\mathcal{G}}_{s,\zeta}\hat{\sigma}_{\tau,s}\right)ds + \hat{h}_{\tau,s}\hat{u}_{\tau,s}dW_s \\ &= (\mathfrak{L}_{\tau,s} - \mathbf{Y}_{\tau,s})\,\hat{u}_{\tau,s}dt + \hat{h}_{\tau,s}\hat{u}_{\tau,s}dW_s, \end{aligned}$$

where $\mathfrak{L}_{\tau,s} := \bar{a}_{\tau,s}\partial_{vv} + \bar{b}_{\tau,s}\partial_v + \bar{c}_{\tau,s}$ with

$$\begin{aligned} \bar{a}_{\tau,s} &= \frac{1}{2}(\partial_\nu\gamma^{\mathrm{IW}}_{\tau,s})^{-2}(\hat{a}_{\tau,s} - \hat{\sigma}^2_{\tau,s}), \\ \bar{b}_{\tau,s} &= (\partial_\nu\gamma^{\mathrm{IW}}_{\tau,s})^{-1}\left(\hat{b}_{\tau,s} - \hat{\sigma}_{\tau,s}\hat{h}_{\tau,s} - (\partial_\nu\gamma^{\mathrm{IW}}_{\tau,s})^{-1}\hat{\sigma}_{\tau,s}\partial_\nu\hat{\sigma}_{\tau,s} - \bar{a}_{\tau,s}\partial_{\nu\nu}\gamma^{\mathrm{IW}}_{\tau,s}\right), \\ \bar{c}_{\tau,s} &= \hat{c}_{\tau,s} - (\partial_\nu\gamma^{\mathrm{IW}}_{\tau,s})^{-1}\hat{\sigma}_{\tau,s}\partial_\nu\hat{h}_{\tau,s}. \end{aligned} \tag{2.5.10}$$

Notice that the change of variable is well defined by the estimates of Lemma 2.5.1. Next we compute the product $v_{\tau,s} = \hat{\varrho}_{\tau,s}\hat{u}_{\tau,s}$: by the Itô formula $d\varrho_{\tau,s} = -\varrho_{\tau,s}\hat{h}_{\tau,s}dW_s$ and therefore

$$\begin{aligned} dv_{\tau,s}(\zeta) &= \varrho_{\tau,s}(\zeta)d\hat{u}_{\tau,s}(\zeta) + \hat{u}_{\tau,s}(\zeta)d\varrho_{\tau,s} + d\langle\hat{u}_{\tau,\cdot}(\zeta)\varrho_{\tau,\cdot}(\zeta)\rangle_s \\ &= \left(\varrho_{\tau,s}(\zeta)\mathfrak{L}_{\tau,s}(\varrho^{-1}_{\tau,s}v_{\tau,s})(\zeta) - \varrho_{\tau,s}(\zeta)(\mathbf{Y}_{\tau,s}(\varrho^{-1}_{\tau,s}v_{\tau,s}))(\zeta) - \bar{h}^2_{\tau,s}(\zeta)v_{\tau,s}(\zeta)\right)ds. \end{aligned}$$

Now we notice that

$$\varrho_{\tau,s}(\zeta)(\mathbf{Y}_{\tau,s}(\varrho^{-1}_{\tau,s}v_{\tau,s}))(\zeta) = (\mathbf{Y}_{\tau,s}v_{\tau,s})(\zeta) + (\mathbf{Y}_{\tau,s}\ln\varrho^{-1}_{\tau,s})(\zeta)v_{\tau,s}(\zeta),$$

and eventually, by a standard application of the Leibniz rule, we get

$$dv_{\tau,s}(\zeta) = \left(a^*_{\tau,s}(\zeta)\partial_{vv}v_{\tau,s}(\zeta) - (\mathbf{Y}_{\tau,s}v_{\tau,s})(\zeta) + b^*_{\tau,s}(\zeta)\partial_v v_{\tau,s}(\zeta) + c^*_{\tau,s}(\zeta)v_{\tau,s}(\zeta)\right)ds,$$

where

$$\begin{aligned} a^*_{\tau,s} &= \bar{a}_{\tau,s} = \frac{1}{2}(\partial_\nu\gamma^{\mathrm{IW}}_{\tau,s})^{-2}(\hat{a}_{\tau,s} - \hat{\sigma}^2_{\tau,s}), \\ b^*_{\tau,s} &= \bar{b}_{\tau,s} + 2\bar{a}_{\tau,s}\partial_\nu\ln\varrho^{-1}_{\tau,s}, \\ c^*_{\tau,s} &= \bar{c}_{\tau,s} + \bar{b}_{\tau,s}\partial_\nu\ln\varrho^{-1}_{\tau,s} + \bar{a}_{\tau,s}\left(\partial_\nu\ln\varrho^{-1}_{\tau,s} + \partial^2_\nu\ln\varrho^{-1}_{\tau,s}\right) + \mathbf{Y}_{\tau,s}\ln\varrho^{-1}_{\tau,s} - \hat{h}^2_{\tau,s}. \end{aligned} \tag{2.5.11}$$

The regularity of the coefficients and (2.5.9) follows directly from (2.5.11), Assumption 2.1.4 and Lemma 2.5.1. $\square$

2.6 Time-dependent and drift adapted parametrix method

In this section we study equation (2.5.7) for fixed $\omega \in \Omega$ and $0 \leqslant \tau < T < \infty$. More generally, we consider a deterministic equation of the form

$$\mathcal{K}_s u(s,\zeta) = \mathcal{L}_s u(s,\zeta) - \partial_s u(s,\zeta) = 0 \tag{2.6.1}$$

where

$$\mathcal{L}_s u(s,\zeta) := \frac{1}{2} a(s,\zeta)\partial_{\nu\nu} u(s,\zeta) + b(s,\zeta)\partial_\nu u(s,\zeta) - \langle Y(s,\zeta), \nabla_\zeta u(s,\zeta)\rangle + c(s,\zeta)u(s,\zeta),$$

for $s \in [\tau, T]$, $\zeta = (\xi, \nu) \in \mathbb{R}^2$, and $Y = (Y_1, Y_2)$ is a generic vector field. We assume the following conditions on the coefficients.

Assumption 2.6.1. *There exist positive constants α, λ_1 such that $a \in C^\alpha_{\tau,T}$ with Hölder constant λ_1 and*

$$\lambda_1^{-1} \leqslant a(s,\zeta) \leqslant \lambda_1, \quad |b(s,\zeta)| + |c(s,\zeta)| \leqslant \lambda_1 \qquad (s,\zeta) \in [\tau, T] \times \mathbb{R}^2. \tag{2.6.2}$$

Assumption 2.6.2. *$Y \in C_{\tau,T}$ and is uniformly Lipschitz continuous in the sense that*

$$\sup_{\substack{s \in [\tau,T] \\ z \neq \zeta}} \frac{|Y(s,z) - Y(s,\zeta)|}{|z - \zeta|} \leqslant \lambda_2$$

for some positive constant λ_2. Moreover $\partial_\nu Y_1 \in C^\alpha_{\tau,T}$ and

$$\lambda_2^{-1} \leqslant \partial_\nu Y_1(s,\zeta) \leqslant \lambda_2, \qquad (s,\zeta) \in [\tau, T] \times \mathbb{R}^2. \tag{2.6.3}$$

Notation 2.6.3. *Similarly to Section 1.4 we introduce the parameter*

$$\Theta := (\alpha, \lambda_1, \lambda_2, T)$$

which gathers the important quantities appearing in the assumptions.

Remark 2.6.4. *When the coefficients are smooth, conditions* (2.6.2) *and* (2.6.3) *ensure the validity of the weak Hörmander condition: indeed the vector fields $\sqrt{a}\partial_\nu$ and Y, together with their commutator, span $\mathbb{R}^3$ at any point. In this case a smooth fundamental solution exists by Hörmander's theorem.*

Since the coefficients are assumed to be only measurable in time, a solution to (2.6.1) has to be understood in the integral sense according to the following definition.

Definition 2.6.5. *A fundamental solution $\Gamma = \Gamma(t,z;s,\zeta)$ for equation* (2.1.2) *is a function defined for $\tau \leqslant t < s \leqslant T$ and $z, \zeta \in \mathbb{R}^2$, such that for any $(t,z) \in [\tau, T) \times \mathbb{R}^2$ we have:*

i) for $t < t_0 \leqslant s \leqslant T$ and $z \in \mathbb{R}^2$, $\Gamma(t, z; \cdot, \cdot)$ belongs to $C_{t_0,T}$, is twice continuously differentiable in ν and satisfies

$$\begin{aligned}\Gamma(t, z; s, \gamma_{t_0,s}(\zeta)) =&\Gamma(t, z; t_0, \zeta) + \int_{t_0}^{s} \Big(a(\varrho, \gamma_{t_0,\varrho}(\zeta))\partial_{\nu\nu}\Gamma(t, z; \varrho, \gamma_{t_0,\varrho}(\zeta))+\\ &+ b(\varrho, \gamma_{t_0,\varrho}(\zeta))\partial_{\nu}\Gamma(t, z; \varrho, \gamma_{t_0,\varrho}(\zeta)) + c(\varrho, \gamma_{t_0,\varrho}(\zeta))\Gamma(t, z; \varrho, \gamma_{t_0,\varrho}(\zeta))\Big) d\varrho\end{aligned}$$

where $\gamma_{t_0,s}(z)$ stands for the integral curve of the field Y with initial datum $\gamma_{t_0,t_0}(\zeta) = \zeta$;

ii) for any bounded and continuous function φ and $z_0 \in \mathbb{R}^2$, we have

$$\lim_{\substack{(s,\zeta)\to(t,z_0)\\ s>t}} \int_{\mathbb{R}^2} \Gamma(t, z; s, \zeta)\varphi(z)dz = \varphi(z_0).$$

The main result of this section is the following

Theorem 2.6.6. *Under Assumptions 2.6.1 and 2.6.2 there exists a fundamental solution Γ for the PDE* (2.6.1)*. Moreover, there exist two constants $\mu = \mu(\Theta) \geqslant 1$, $C = C(\Theta) \geqslant 1$ such that, for any $\zeta = (\xi, \nu), z \in \mathbb{R}^2$ and $\tau \leqslant t < s \leqslant T$,*

$$C^{-1}\Gamma^{\text{heat}}\left(\mu^{-1}\mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)\right) \leqslant \Gamma(t, z; s, \zeta) \leqslant C\Gamma^{\text{heat}}\left(\mu\mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)\right), \tag{2.6.4}$$

$$|\partial_\nu\Gamma(t, z; s, \xi, \nu)| \leqslant \frac{C}{\sqrt{s-t}}\Gamma^{\text{heat}}\left(\mu\mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)\right), \tag{2.6.5}$$

$$|\partial_{\nu\nu}\Gamma(t, z; s, \xi, \nu)| \leqslant \frac{C}{s-t}\Gamma^{\text{heat}}\left(\mu\mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)\right). \tag{2.6.6}$$

where $\mathcal{D}_\lambda$ is as in Theorem 2.1.6 and $\gamma_{t,s}(\zeta)$ is as in Definition 2.6.5.

2.6.1 Proof of Theorem 2.6.6

Parametrix expansion

For fixed $(t_0, z_0) \in [\tau, T) \times \mathbb{R}^2$, let

$$\gamma_{t,s}(z_0) = z_0 + \int_t^s Y\left(\varrho, \gamma_{t,\varrho}(z_0)\right) d\varrho, \qquad s \in [\tau, T], \tag{2.6.7}$$

be the integral curve of Y starting from (t_0, z_0). Following [14] we linearize $Y = Y(s, \zeta)$ at (t_0, z_0) setting

$$\widetilde{Y}^{t_0,z_0}(s, \zeta) = Y(s, \gamma_{t_0,s}(z_0)) + DY(s, \gamma_{t_0,s}(z_0))\left(\zeta - \gamma_{t_0,s}(z_0)\right), \qquad s \in [t_0, T],\ \zeta \in \mathbb{R}^2.$$

where DY stands for a reduced Jacobian defined as

$$DY := \begin{pmatrix} 0 & \partial_v Y_1 \\ 0 & 0 \end{pmatrix}.$$

Then we consider the linear approximation of $\mathcal{L}_s$ defined as

$$\widetilde{\mathcal{L}}_s^{t_0,z_0} := \frac{1}{2}a(s,\gamma_{t_0,s}(z_0))\partial_{\nu\nu} - \langle \widetilde{Y}^{t_0,z_0}(s,\zeta), \nabla\rangle.$$

The diffusion coefficient of $\widetilde{\mathcal{L}}_s^{t_0,z_0}$ depends on s only (apart from t_0, z_0 that are fixed parameters), while the drift coefficients depend on s and linearly on ξ, ν. Notice that $\widetilde{\mathcal{L}}_s^{t_0,z_0} - \partial_s$ is the forward Kolmogorov operator of the system of linear SDEs

$$dZ_s = \widetilde{Y}^{t_0,z_0}(s, Z_s)\, ds + \sqrt{a(s,\gamma_{t_0,s}(z_0))}\mathbf{e_2} dB_s. \tag{2.6.8}$$

Let $Z_s^{t,z}$ denote the solution of (2.6.8) starting from z at time $t \in [t_0, T)$. Then $Z_s^{t,z}$ is a Gaussian process: the mean $\widetilde{\gamma}_{t,s}^{t_0,z_0}(z) := E\left[Z_s^{t,z}\right]$ solves the ODE

$$\widetilde{\gamma}_{t,s}^{t_0,z_0}(z) = z + \int_t^s \widetilde{Y}^{t_0,z_0}(\varrho, \widetilde{\gamma}_{t,s}^{t_0,z_0}(z))d\varrho, \qquad s \in [t,T], \tag{2.6.9}$$

and the covariance matrix is given by

$$\widetilde{\mathcal{C}}_{t,s}^{t_0,z_0} = \int_t^s a(\varrho, \gamma_{t_0,\varrho}(z_0)) \left(E_{\varrho,s}^{t_0,z_0}\mathbf{e}_2\right)\left(E_{\varrho,s}^{t_0,z_0}\mathbf{e}_2\right)^* d\varrho, \tag{2.6.10}$$

where $E_{\varrho,s}^{t_0,z_0}$ is the fundamental matrix associated with $(DY)(s,\gamma_{t_0,s}(z_0))$, that is the solution of

$$E_{\varrho,s}^{t_0,z_0} = \mathrm{Id} + \int_\varrho^s (DY)(u,\gamma_{t_0,u}(z_0))E_{\varrho,u}^{t_0,z_0} du, \qquad s \in [\varrho, T],$$

with Id equal to the (2×2)-identity matrix.

Lemma 2.6.7. *For any $z_0 \in \mathbb{R}^2$ and $\tau \leqslant t_0 \leqslant t < s \leqslant T$, we have* $\det \widetilde{\mathcal{C}}_{t,s}^{t_0,z_0} > 0$.

Proof. By Assumption 2.6.1 it is enough to prove the assertion for $a \equiv 1$. Suppose that there exist $z \in \mathbb{R}^2 \setminus \{0\}$, $z_0 \in \mathbb{R}^2$ and $\tau \leqslant t_0 \leqslant t < s \leqslant T$ such that $\langle \widetilde{\mathcal{C}}_{t,s}^{t_0,z_0} z, z\rangle = 0$. Since $\widetilde{\mathcal{C}}_{t,s}^{t_0,z_0}$ is positive semi-definite, this is equivalent to the condition

$$|(E_{\varrho,s}^{t_0,z_0}\mathbf{e}_2)^* z|^2 = 0, \quad \text{a.e. } \varrho \in (t,s),$$

that is $((E_{\varrho,s}^{t_0,z_0})^* z)_2 = 0$, for a.e. $\varrho \in (t,s)$. We have

$$\partial_\varrho (E_{\varrho,s}^{t_0,z_0})^* z = -DY^*(\varrho, \gamma_{t_0,\varrho}(z_0))(E_{\varrho,s}^{t_0,z_0})^* z,$$

and therefore

$$0 = \partial_\varrho((E_{\varrho,s}^{t_0,z_0})^* z)_2 = \partial_\nu Y_1(\varrho, \gamma_{t_0,\varrho}(z_0))((E_{\varrho,s}^{t_0,z_0})^* z)_1.$$

Since $((E_{\varrho,s}^{t_0,z_0})^* z)_2 = 0$ and $\partial_\nu Y_1 \in [\lambda_2^{-1}, \lambda_2]$ by Assumption 2.6.2 we have $(E_{\varrho,s}^{t_0,z_0})^* z \equiv 0$, for a.e. $\varrho \in (t,s)$, which is absurd. $\square$

Lemma 2.6.7 ensures that the Gaussian process in (2.6.8) admits a transition density that is the fundamental solution of $\widetilde{\mathcal{L}}_s^{t_0,z_0} - \partial_s$. To be more precise we have the following:

Proposition 2.6.8. *For any $0 \leqslant \tau \leqslant t_0 \leqslant t < s \leqslant T$ and $z, \zeta, z_0 \in \mathbb{R}^2$, the function*

$$\widetilde{\Gamma}^{t_0,z_0}(t,z;s,\zeta) := \Gamma^{\text{heat}}\left(\widetilde{\mathcal{C}}_{t,s}^{t_0,z_0}, \zeta - \widetilde{\gamma}_{t,s}^{t_0,z_0}(z)\right)$$

is the fundamental solution of $\widetilde{\mathcal{L}}_s^{t_0,z_0} - \partial_s$, evaluated at (s,ζ) and with pole at (t,z).

We are now in position to define the parametrix Z for $\mathcal{K}_s$ in (2.6.1). We set

$$Z(t,z;s,\zeta) := \widetilde{\Gamma}^{t,z}(t,z;s,\zeta), \qquad \tau \leqslant t < s \leqslant T,\ z,\zeta \in \mathbb{R}^2.$$

Since

$$\gamma_{t,s}(z) = z + \int_t^s Y(\varrho, \gamma_{t,\varrho}(z))d\varrho = z + \int_t^s \widetilde{Y}^{t,z}(\varrho, \gamma_{t,\varrho}(z))d\varrho$$

we have $\gamma_{t,s}(z) = \widetilde{\gamma}_{t,s}^{t,z}(z)$ and therefore the parametrix reads

$$Z(t,z;s,\zeta) = \Gamma^{\text{heat}}\left(\widetilde{\mathcal{C}}_{t,s}^{t,z}, \zeta - \gamma_{t,s}(z)\right) \tag{2.6.11}$$

for $\tau \leqslant t < s \leqslant T$ and $z, \zeta \in \mathbb{R}^2$.

Finally, in analogy to Section 1.4 we set

$$\Gamma(t,z;s,\zeta) = Z(t,z;s,\zeta) + (\Phi \otimes Z)(t,z;s,\zeta), \tag{2.6.12}$$

with

$$\Phi(t,z;s,\zeta) = \sum_{k \geqslant 1} H^{\otimes k}(t,z;s,\zeta), \tag{2.6.13}$$

for $H(t,z;s,\zeta) = (\mathcal{K}_s Z)(t,z;s,\zeta)$, and we are going to prove that Γ is indeed the fundamental solution for $\mathcal{K}_s$.

Gaussian bounds for the parametrix

Proposition 2.6.9. *There exists a positive constant $\mu = \mu(\Theta) \geqslant 1$ such that*

$$\mu^{-1}|\mathcal{D}_{\sqrt{s-t}}\zeta|^2 \leqslant \langle \widetilde{\mathcal{C}}_{t,s}^{t_0,z_0}\zeta, \zeta\rangle \leqslant \mu|\mathcal{D}_{\sqrt{s-t}}\zeta|^2, \qquad \tau \leqslant t < s \leqslant T,\ z_0, \zeta \in \mathbb{R}^2. \tag{2.6.14}$$

Proof. By Assumptions 2.6.1 it is enough to prove the assertion for $a \equiv 1$. For $\lambda > 0$, let $\mathcal{U}_\lambda$ be the set of 2×2, time-dependent matrices $\mathcal{Y}(s)$, with entries uniformly bounded by λ, and such that $(\mathcal{Y}(s))_{1,2} \in [\lambda^{-1}, \lambda]$. Let $\mathcal{Y}(s) \in \mathcal{U}_\lambda$ and

$$\mathcal{C}_{t,s} := \int_t^s (\mathcal{E}_{\varrho,s}\mathbf{e_2})(\mathcal{E}_{\varrho,s}\mathbf{e_2})^* \, d\varrho, \qquad \tau \leqslant t < s \leqslant T,$$

where $\mathcal{E}_{\varrho,s}$ denotes the resolvent associated with $\mathcal{Y}(s)$. We split the proof in two steps.

Step 1. First we prove that

$$c^{-1}|\zeta|^2 \leqslant \langle \mathcal{C}_{0,1}\zeta,\zeta\rangle \leqslant c|\zeta|^2, \tag{2.6.15}$$

where $c = c(\lambda) > 0$. As in [14] (see Proposition 3.4), we consider the map

$$\Psi : L^2([0,1],\mathcal{M}_2(\mathbb{R})) \longrightarrow \mathbb{R}, \qquad \Psi(\mathcal{Y}) := \det \mathcal{C}_{0,1},$$

where $\mathcal{M}_2(\mathbb{R})$ is the space of 2×2 matrices with real entries. Notice that $\mathcal{U}_\lambda$ is compact in the weak topology of $L^2([0,1],\mathcal{M}_2(\mathbb{R}))$ because it is bounded, convex and closed in the strong topology (cf., for instance, [8], Corollary III.19). On the other hand, Ψ is continuous from $L^2([0,1],\mathcal{M}_2(\mathbb{R}))$, equipped with the weak topology, to $\mathbb{R}$. Therefore the image $\Psi(\mathcal{U}_\lambda)$ is a compact subset of $\mathbb{R}_{>0}$ by Lemma 2.6.7. Thus there exists $\bar\lambda > 0$ such that $\inf\{\det \mathcal{C}_{0,1} \mid \mathcal{Y}\in\mathcal{U}_\lambda\} \geqslant \bar\lambda^{-1}$ and $\sup\{\|\mathcal{C}_{0,1}\| \mid \mathcal{Y}\in\mathcal{U}_\lambda\} \leqslant \bar\lambda$. This suffices to prove (2.6.15).

Step 2. We use a scaling argument. For every $\tau \leqslant t < s \leqslant T$ we show that $\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\mathcal{C}_{t,s}\mathcal{D}_{\frac{1}{\sqrt{s-t}}}$ coincides with some matrix $\hat{\mathcal{C}}_{0,1}$ to which we can apply the result of Step 1. We have

$$\begin{aligned}
\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\mathcal{C}_{t,s}\mathcal{D}_{\frac{1}{\sqrt{t-s}}} &= \int_t^s \left(\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\mathcal{E}_{\varrho,s}\mathcal{D}_{\sqrt{s-t}}\mathbf{e_2}\right)\left(\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\mathcal{E}_{\varrho,s}\mathcal{D}_{\sqrt{s-t}}\mathbf{e_2}\right)^* \frac{d\varrho}{s-t}\\
&= \int_0^1 \left(\hat{\mathcal{E}}^{t,s}_{\varrho,1}\mathbf{e_2}\right)\left(\hat{\mathcal{E}}^{t,s}_{\varrho,1}\mathbf{e_2}\right)^* =: \hat{\mathcal{C}}^{t,s}_{0,1}
\end{aligned}$$

where

$$\hat{\mathcal{E}}^{t,s}_{\varrho_1,\varrho_2} = \mathcal{D}_{\frac{1}{\sqrt{s-t}}}\mathcal{E}_{t+\varrho_1(s-t),t+\varrho_2(s-t)}\mathcal{D}_{\sqrt{s-t}},$$

solves the differential system

$$\partial_{\varrho_2}\hat{\mathcal{E}}^{t,s}_{\varrho_1,\varrho_2} = (s-t)\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\mathcal{Y}(t+\varrho_2(s-t))\mathcal{D}_{\sqrt{s-t}}\hat{\mathcal{E}}^{t,s}_{\varrho_1,\varrho_2} =: \hat{\mathcal{Y}}^{t,s}(\varrho_2)\hat{\mathcal{E}}^{t,s}_{\varrho_1,\varrho_2}$$

with $\hat{\mathcal{E}}^{t,s}_{\varrho,\varrho} = I_2$. A direct computation shows that

$$(\hat{\mathcal{Y}}^{t,s}(\varrho))_{1,2} = (\mathcal{Y}(t+\varrho(s-t)))_{1,2} \in [\lambda^{-1},\lambda], \qquad \|\hat{\mathcal{Y}}^{t,s}(\varrho)\|_\infty \leqslant (1+T^2)\|\mathcal{Y}(\varrho)\|_\infty.$$

Therefore (2.6.15) holds for $\hat{\mathcal{C}}^{t,s}_{1,0}$, uniformly in t,s, with c dependent only on λ and T. □

Remark 2.6.10. *Since, for $\tau \leqslant t < s \leqslant T$, $\widetilde{\mathcal{C}}^{t,z}_{t,s}$ is a symmetric and positive definite matrix,* (2.6.14) *also yields an analogous estimate for the inverse: we have*

$$\mu^{-1}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\zeta\right|^2 \leqslant \langle(\widetilde{\mathcal{C}}^{t_0,z_0}_{t,s})^{-1}\zeta,\zeta\rangle \leqslant \mu\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\zeta\right|^2, \qquad \tau\leqslant t<s\leqslant T,\ z_0,\zeta\in\mathbb{R}^2. \tag{2.6.16}$$

The following result is a standard consequence of (2.6.14) and (2.6.16) (cf., for instance, Proposition 3.1 in [16]).

Proposition 2.6.11. *There exists a positive constant $\mu = \mu(\Theta) \geqslant 1$, such that*

$$\mu^{-2}\Gamma^{\text{heat}}\left(\mu^{-1}\mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)\right) \leqslant Z(t,z;s,\zeta) \leqslant \mu^2\Gamma^{\text{heat}}\left(\mu\mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)\right), \tag{2.6.17}$$

for every $\tau \leqslant t < s \leqslant T$ and $z, \zeta \in \mathbb{R}^2$.

Next we prove some estimates for the derivatives of $Z(t,z;s,\zeta)$. We start with the following

Lemma 2.6.12. *We have*

$$(s-t)^{2-i}\left|\left((\widetilde{\mathcal{C}}_{t,s}^{t_0,z_0})^{-1}w\right)_i\right| \leqslant \frac{C}{\sqrt{s-t}}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}w\right|, \tag{2.6.18}$$

$$(s-t)^{4-i-j}\left|\left((\widetilde{\mathcal{C}}_{t,s}^{t_0,z_0})^{-1}\right)_{ij}\right| \leqslant \frac{C}{s-t} \tag{2.6.19}$$

for every $i,j \in \{1,2\}$, $\tau \leqslant s < t \leqslant T$ and $w, \zeta, z_0 \in \mathbb{R}^2$.

Proof. We have

$$\begin{aligned}(s-t)^{2-i}\left|\left((\widetilde{\mathcal{C}}_{t,s}^{t_0,z_0})^{-1}w\right)_i\right| &= \frac{1}{\sqrt{s-t}}\left|\left(\mathcal{D}_{\sqrt{s-t}}(\widetilde{\mathcal{C}}_{t,s}^{t_0,z_0})^{-1}\mathcal{D}_{\sqrt{s-t}}\mathcal{D}_{\frac{1}{\sqrt{s-t}}}w\right)_i\right| \\ &\leqslant \frac{1}{\sqrt{s-t}}\left\|\mathcal{D}_{\sqrt{s-t}}(\widetilde{\mathcal{C}}_{t,s}^{t_0,z_0})^{-1}\mathcal{D}_{\sqrt{s-t}}\right\|\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}w\right|.\end{aligned}$$

In order to get (2.6.18) it suffice to notice that, by (2.6.16), we have

$$\left\|\mathcal{D}_{\sqrt{s-t}}(\widetilde{\mathcal{C}}_{t,s}^{t_0,z_0})^{-1}\mathcal{D}_{\sqrt{t-s}}\right\| \leqslant C.$$

Taking $w = \mathbf{e}_j$ we also get (2.6.19). □

We are ready to state the last result for this section, which is a standard consequence of estimates (2.6.18), (2.6.19) and Proposition 2.6.11 (cf., for instance, Proposition 3.6 in [16]).

Proposition 2.6.13. *For any $j = 0, 1, \cdots$ there exist two positive constants C_j, $\mu_j \geqslant 1$ depending on Θ such that, for any $\tau \leqslant t < s \leqslant T$ and $\zeta = (\xi, \nu), z \in \mathbb{R}^2$ we have*

$$|\partial_\xi^i\partial_\nu^j Z(t,z;s,\xi,\nu)| \leqslant C_{ij}(s-t)^{-\frac{3i+j}{2}}\Gamma^{\text{heat}}\left(\mu_{ij}\mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)\right), \tag{2.6.20}$$

and, for every $\zeta, \zeta' \in \mathbb{R}^2$ such that $|\zeta - \zeta'|_B \leqslant \sqrt{s-t}$,

$$|\partial_\xi^i\partial_\nu^j Z(t,z;s,\zeta) - \partial_\xi^i\partial_\nu^j Z(t,z;s,\zeta')| \leqslant C_j\frac{|z-\zeta'|_B}{(s-t)^{\frac{1+3i+j}{2}}}\Gamma^{\text{heat}}\left(\mu_j\mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)\right), \tag{2.6.21}$$

where $|(x,v)|_B := |x|^{\frac{1}{3}} + |v|$ is the spatial part of the δ-homogeneous norm in (2.3.2). Notice as well that $|\zeta - \zeta'|_B = d_{\mathcal{L}}((s,\zeta),(s,\zeta'))$ where $d_{\mathcal{L}}$ is the intrinsic distance in (2.3.3).

Proof. For simplicity consider the case $i = 1$, $j = 0$, since the other cases are analogous. By (2.6.11) we have

$$|\partial_\xi Z(t,z;s,\xi,\nu)| \leqslant \left|\left((\widetilde{\mathcal{C}}^{t,z}_{t,s})^{-1}w\right)_1\right| Z(t,z;s,\zeta)$$

(by (2.6.18) and Proposition 2.6.11)

$$\leqslant C(s-t)^{-\frac{3}{2}}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}w\right|\Gamma^{\text{heat}}\left(\mu\text{Id},\mathcal{D}_{\frac{1}{\sqrt{s-t}}}w\right)$$

(by (1.4.18))

$$\begin{aligned}&\leqslant C'(s-t)^{-\frac{3}{2}}\Gamma^{\text{heat}}\left((\mu+\varepsilon)\text{Id},\mathcal{D}_{\frac{1}{\sqrt{s-t}}}w\right)\\&= C'(s-t)^{-\frac{3}{2}}\Gamma^{\text{heat}}\left((\mu+\varepsilon)\mathcal{D}_{s-t},\zeta-\gamma_{t,s}(z)\right).\end{aligned}$$

Let us now turn to estimate (2.6.21):

$$\begin{aligned}&|\partial_\xi Z(t,z;s,\zeta)-\partial_\xi Z(t,z;s,\zeta')|\\&\leqslant \sup_{\lambda\in[0,1]}\left(|\partial^2_\xi Z(t,z;s,\zeta+\lambda(\zeta'-\zeta))||(\zeta-\zeta')_1|+|\partial_\xi\partial_\nu Z(t,z;s,\zeta+\lambda(\zeta'-\zeta))||(\zeta-\zeta')_2|\right)\\&\leqslant \left((s-t)^{-3}|\zeta-\zeta'|^3_B+(s-t)^{-2}|\zeta-\zeta'|_B\right)\sup_{\lambda\in[0,1]}\Gamma^{\text{heat}}\left(\mu\mathcal{D}_{s-t},\zeta+\lambda(\zeta'-\zeta)-\gamma_{t,s}(z)\right)\\&\leqslant (s-t)^{-2}|\zeta-\zeta'|_B\Gamma^{\text{heat}}\left(\mu'_j\mathcal{D}_{s-t},\zeta-\gamma_{t,s}(z)\right),\end{aligned}$$

where in the last inequality we noticed that, for $|\zeta-\zeta'|_B \leqslant \sqrt{s-t}$,

$$\begin{aligned}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(\zeta+\lambda(\zeta'-\zeta)-\gamma_{t,s}(z))\right| &\geqslant \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(\zeta-\gamma_{t,s}(z))\right|-\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(\lambda(\zeta'-\zeta))\right|\\&\geqslant \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(\zeta-\gamma_{t,s}(z))\right|-2.\end{aligned}$$

The proof is complete. □

Gaussian bounds for the parametrix series

Next we need some estimates for the iterated kernels which appear in the parametrix expansion. Recall from Chapter 1, Proposition 1.4.8 the crucial role of the reproduction property of the Gaussian kernel Γ^{heat} to obtain estimates that are uniform with respect to the iteration parameter. This passage here is not trivial, since it becomes necessary to handle both the presence of the dilation matrix $\mathcal{D}_{s-t}$ and, most importantly, the transport term $\gamma_{t,s}$. We start with some preliminary lemmas. For simplicity we assume $\tau = 0$ throughout the section.

Lemma 2.6.14 (Reproduction formula)**.** *For any $c', c'' > 0$ we have*

$$\Lambda(c',c'')^{-1}\Gamma^{\text{heat}}\left(\frac{c'\wedge c''}{2}\mathcal{D}_{s-t}, \zeta''-\zeta'\right) \leqslant \int_{\mathbb{R}^2}\Gamma^{\text{heat}}\left(c'\mathcal{D}_{s-\varrho}, \eta-\zeta''\right)\Gamma^{\text{heat}}\left(c''\mathcal{D}_{\varrho-t}, \zeta'-\eta\right)d\eta$$
$$\leqslant \Lambda(c',c'')\Gamma^{\text{heat}}\left((c'\vee c'')\mathcal{D}_{s-t}, \zeta''-\zeta'\right),$$

for every $0 \leqslant t < \varrho < s \leqslant T$, $\zeta', \zeta'' \in \mathbb{R}^2$, where $\Lambda(c',c'') = \sqrt{\frac{2(c'\vee c'')}{c'\wedge c''}}$.

Proof. It is a direct consequence (see also [14], Lemma B.1) of the following trivial estimate

$$\frac{c'\wedge c''}{2}\mathcal{D}_{s-t} \leqslant c'\mathcal{D}_{s-\varrho} + c''\mathcal{D}_{\varrho-t} \leqslant (c'\vee c'')\mathcal{D}_{s-t}.$$

□

Remark 2.6.15. *Let $\tau = 0$, $T = 1$. If $\hat{Y}$ is a vector field satisfying Assumption 2.6.2 and $\hat{\gamma}_s$ is the integral curve*

$$\hat{\gamma}_s(z) = z + \int_0^s \hat{Y}_\varrho(\hat{\gamma}_\varrho(z))d\varrho, \qquad s \in [0,1],$$

then $\hat{\gamma}_1(\cdot)$ is a diffeomorphism of $\mathbb{R}^2$. Moreover, since $\hat{Y}$ is Lipschitz continuous, we have

$$m^{-1}|z-\hat{\gamma}_1(\zeta)| \leqslant |\hat{\gamma}_1^{-1}(z)-\zeta| \leqslant m|z-\hat{\gamma}_1(\zeta)|, \qquad z,\zeta \in \mathbb{R}^2, \tag{2.6.22}$$

for a constant m which depends only on λ_2.

Lemma 2.6.16. *Let $\gamma_{t,s}(z)$ be as in (2.6.7). There exists a positive constant m, only dependent on λ_2 and T, such that*

$$m^{-1}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(z-\gamma_{t,s}(\zeta))\right| \leqslant \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(\gamma_{s,t}(z)-\zeta)\right| \leqslant m\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(z-\gamma_{t,s}(\zeta))\right|, \tag{2.6.23}$$

for every $0 \leqslant t < s \leqslant T$ and $z, \zeta \in \mathbb{R}^2$.

Proof. We use again a scaling argument: we set $z' = \mathcal{D}_{\sqrt{s-t}}z$ and

$$\hat{\gamma}_\varrho(z) = \mathcal{D}_{\frac{1}{\sqrt{s-t}}}\gamma_{t,t+\varrho(s-t)}(z'), \quad \hat{Y}(\varrho,z) = (s-t)\mathcal{D}_{\frac{1}{\sqrt{s-t}}}Y(t+\varrho(s-t),z'), \qquad \varrho \in [0,1].$$

Then we have

$$\hat{\gamma}_\varrho(z) = z + \int_0^\varrho \hat{Y}_u(\hat{\gamma}_u(z))du, \qquad \varrho \in [0,1].$$

As in the proof of Proposition 2.6.9, we have that $\hat{Y}$ satisfies Assumption 2.6.2. By Remark 2.6.15, estimate (2.6.22) holds for $\hat{\gamma}_\varrho(z)$. To conclude, it suffices to substitute z and ζ with $\bar{z} = \mathcal{D}_{\frac{1}{\sqrt{s-t}}}z$ and $\bar{\zeta} = \mathcal{D}_{\frac{1}{\sqrt{s-t}}}\zeta$ in (2.6.22). □

Assume now we need to compute

$$I(t,z;s,\zeta)=\int_{\mathbb{R}^2}\Gamma^{\text{heat}}\left(\mu\mathcal{D}_{\tau-t},y-\gamma_{t,\tau}(z)\right)\Gamma^{\text{heat}}\left(\mu\mathcal{D}_{s-\tau},\zeta-\gamma_{\tau,s}(y)\right)dy.$$

In order to apply Lemma 2.6.14, we need to use (2.6.23) and get

$$\begin{aligned}I(t,z;s,\zeta)&\leqslant m\int_{\mathbb{R}^2}\Gamma^{\text{heat}}\left(\mu\mathcal{D}_{\tau-t},y-\gamma_{t,\tau}(z)\right)\Gamma^{\text{heat}}\left(m\mu\mathcal{D}_{s-\tau},\gamma_{s,\tau}(\zeta)-y\right)dy\\&\leqslant C(m)\Gamma^{\text{heat}}\left(m\mu\mathcal{D}_{s-t},\gamma_{s,\tau}(\zeta)-\gamma_{t,\tau}(z)\right)\\&\leqslant C'(m)\Gamma^{\text{heat}}\left(m^2\mu\mathcal{D}_{s-t},\zeta-\gamma_{t,s}(z)\right).\end{aligned}$$

This precisely show how the diffusion constant grows at each iteration, and therefore a direct estimate of the series (2.6.13) seems not possible. This problem has already been addressed in the work of Delarue and Menozzi [14]: here the authors truncate the series at a suitable iteration and proceed by estimating the remainder through some stochastic control techniques. We will resort to similar computations to get some crucial bounds in chapter four.

Here we notice that, by the results in [14] it is possible to verify that the full parametrix series converges: the idea is to exploit the lower bound in [14] to rewrite the controls (2.6.17), (2.6.20)-(2.6.21) in terms of the transition density of some auxiliary diffusion, for which an exact reproduction formula holds. More precisely, for some $\delta>0$ we introduce the SDE

$$dZ_s=Y(s,Z_s)ds+\delta\mathbf{e}_2dB_s,\qquad s\geqslant 0.$$

By [14], Theorem 1.1 and [48], Theorem 1.1 there exists the corresponding density $p_\delta=p_\delta(t,z;s,\zeta)$, satisfying the two sided bounds

$$C^{-1}\Gamma^{\text{heat}}\left(c^{-1}\mathcal{D}_{s-t},\zeta-\gamma_{t,s}(z)\right)\leqslant p_\delta(t,z;s,\zeta)\leqslant C\Gamma^{\text{heat}}\left(c\mathcal{D}_{s-t},\zeta-\gamma_{t,s}(z)\right),$$

for all $0\leqslant t<s\leqslant T$ and $z,\zeta\in\mathbb{R}^2$, for some constants $c,C\geqslant 1$ depending on λ_2, T and δ. Therefore, with the notations of Propositions 2.6.11 and 2.6.13, and noting that

$$|z|_B=\sqrt{s}\left|\mathcal{D}_{\frac{1}{\sqrt{s}}}(z)\right|_B,\qquad z\in\mathbb{R}^2,\ s\in[t,T],$$

we can chose μ and then $\delta=\delta(\mu)$ such that, for all $i=0,1$, $j=0,1,2$, $\beta\in[0,3]$, $0\leqslant t<s\leqslant T$ and $z,\zeta\in\mathbb{R}^2$ we have

$$|\zeta-\gamma_{t,s}(z)|_B^\beta|\partial_\xi^i\partial_\nu^j Z(t,z;s,\xi,\nu)|\leqslant C(s-t)^{\frac{\beta-3i-j}{2}}p_\delta(t,z;s,\zeta),\tag{2.6.24}$$

and, for any $\zeta,\zeta'\in\mathbb{R}^2$ such that $|\zeta-\zeta'|_B\leqslant\sqrt{s-t}$,

$$|\zeta-\gamma_{t,s}(z)|_B^\beta|\partial_\xi^i\partial_\nu^j Z(t,z;s,\zeta)-\partial_\xi^j\partial_\nu^j Z(t,z;s,\zeta')|\leqslant C\frac{|\zeta-\zeta'|_B}{(s-t)^{\frac{1+3i+j-\beta}{2}}}p_\delta(t,z;s,\zeta),$$

where C only depends on Θ, δ and β. Similarly, we also have

$$|\zeta - \gamma_{t,s}(z)|_B^\beta |\partial_\xi^i \partial_\nu^j \widetilde{\Gamma}^{t_0,z_0}(t,z;s,\xi,\nu)|_{(t_0,z_0)=(s,\zeta)} \leqslant C(s-t)^{\frac{\beta-3i-j}{2}} p_\delta(t,z;s,\zeta), \tag{2.6.25}$$

Indeed, by Lemma 2.6.16

$$\begin{aligned} \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\left(\zeta - \widetilde{\gamma}^{s,\zeta}_{t,s}(z)\right)\right| &\geqslant m^{-1}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\left(\widetilde{\gamma}^{s,\zeta}_{s,t}(\zeta) - z\right)\right| = m^{-1}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\left(\gamma_{s,t}(\zeta) - z\right)\right| \\ &\geqslant m^{-2}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\left(\zeta - \gamma_{t,s}(z)\right)\right|, \end{aligned}$$

and analogously

$$\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\left(\zeta - \widetilde{\gamma}^{s,\zeta}_{t,s}(z)\right)\right| \leqslant m^2\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\left(\zeta - \gamma_{t,s}(z)\right)\right|,$$

and therefore one can argue as in Propositions 2.6.11 and 2.6.13.

We fix from now δ such that (2.6.24)-(2.6.25) hold, and write $p = p_\delta$. In particular we have

$$\int_{\mathbb{R}^2} p(t,z;\varrho,y)p(\varrho,y;s,\zeta)dy = p(t,z;s,\zeta), \qquad 0 \leqslant t < \varrho < s \leqslant T,\ z,\zeta \in \mathbb{R}^2. \tag{2.6.26}$$

We are now ready to prove the main result of the section.

Proposition 2.6.17. *For every $t \in [0,T]$, $z \in \mathbb{R}^2$ the series* (2.6.13) *is uniformly convergent in $]0,T] \times \mathbb{R}^2$. Moreover there exists a constant $C \geqslant 1$ such that*

$$|\Phi(t,z;s,\zeta)| \leqslant \frac{C}{(s-t)^{1-\frac{\alpha}{2}}} p(t,z;s,\zeta), \tag{2.6.27}$$

$$|\Phi(t,z;s,\zeta) - \Phi(t,z;s,\zeta')| \leqslant C\frac{|\zeta-\zeta'|_B^{\frac{\alpha}{2}}}{(s-t)^{1-\frac{\alpha}{4}}}\left(p(t,z;s,\zeta) + p(t,z;s,\zeta')\right) \tag{2.6.28}$$

for every $0 \leqslant t < s \leqslant T$ and $z, \zeta, \zeta' \in \mathbb{R}^2$.

Proof. The proof follows the same lines as for the parabolic case (cf. Proposition 1.4.8), as well as for the Kolmogorov operators with linear drifts (see [58], Lemma 2.3 and Corollary 2.3, and [16], Lemma 4.3). In the former case, the proxy operator $\widetilde{\mathcal{L}}^{t,z}$ does not need to include any first order derivative, since they don't add any singular contribution to the estimate of $H = \mathcal{K}_s Z$. On the other hand, for the Kolmogorov case, the first order derivatives in the degenerate directions carry the most singular contributions: in [58] and [16] the authors include the full drift Y in the definition of the proxy $\widetilde{\mathcal{L}}^{t,z}$ so that they don't add in the estimate; in the present case however, $\widetilde{\mathcal{L}}^{t,z}$ include as approximated version of Y which poses the main concerns, because it leaves some critical singular terms to handle. It turns out that Assumption 2.6.2 is enough to make the procedure work.

We prove the preliminary estimate

$$\left|H^{\otimes k}(t,z;s,\zeta)\right| \leqslant \frac{M_k}{(s-t)^{1-\frac{k\alpha}{2}}} p(t,z;s,\zeta), \qquad 0 \leqslant t < s \leqslant T,\ z,\zeta \in \mathbb{R}^2, \tag{2.6.29}$$

where $M_k = C^k \frac{\Gamma_E^k\left(\frac{\alpha}{2}\right)}{\Gamma_E\left(\frac{k\alpha}{2}\right)}$ and Γ_E is the Euler Gamma function.

For $k = 1$, we have

$$\begin{aligned}
H(t,z;s,\zeta) &= (\mathcal{L}_s - \widetilde{\mathcal{L}}_s^{t,z})Z(t,z;s,\zeta)\\
&= \frac{1}{2}\left(a(s,\zeta) - a(s,\gamma_{t,s}(z))\right)\partial_{\nu\nu}Z(t,z;s,\zeta) + b(s,\zeta)\partial_\nu Z(t,z;s,\zeta) +\\
&\quad + \langle Y(s,\zeta) - \widetilde{Y}^{t,z}(s,\zeta), \nabla Z(t,z;s,\zeta)\rangle + c(s,\zeta)Z(t,z;s,\zeta)\\
&=: E_1 + E_2 + E_3 + E_4.
\end{aligned}$$

By Assumption 2.6.1 and (2.6.24) we have

$$\begin{aligned}
|E_1| &\leqslant \lambda_1|\zeta - \gamma_{t,s}(z)|^\alpha \partial_{\nu\nu}Z(t,z;s,\zeta)\\
&\leqslant C(|\zeta - \gamma_{t,s}(z)|_B^\alpha + |\zeta - \gamma_{t,s}(z)|_B^{3\alpha})\partial_{\nu\nu}Z(t,z;s,\zeta) \leqslant C(s-t)^{\frac{\alpha}{2}-1}p(t,z;s,\zeta).
\end{aligned}$$

By Assumption 2.6.1 and (2.6.24) we also have

$$|E_2 + E_4| \leqslant C(s-t)^{-\frac{1}{2}}p(t,z;s,\zeta).$$

As for E_3, we have

$$\begin{aligned}
|(Y(s,\zeta) - \widetilde{Y}^{t,z}(s,\zeta))_1| &= |Y_1(s,\zeta) - Y_1(s,\gamma_{t,s}(z)) - \partial_\nu Y_1(s,\gamma_{t,s}(z))(\zeta - \gamma_{t,s}(z))_2|\\
&\leqslant |Y_1(s,\zeta) - Y_1(s,(\gamma_{t,s}(z))_1,\zeta_2)| +\\
&\quad + |Y_1(s,(\gamma_{t,s}(z))_1,\zeta_2) - Y_1(s,\gamma_{t,s}(z)) - \partial_\nu Y_1(s,\gamma_{t,s}(z))(\zeta - \gamma_{t,s}(z))_2|\\
&\leqslant C\left(|(\zeta - \gamma_{t,s}(z))_1| + |(\zeta - \gamma_{t,s}(z))_2|^{1+\alpha}\right),
\end{aligned}$$

because $\partial_\nu Y_1$ is Hölder continuous by Assumption 2.6.2: here we use the elementary inequality

$$\left|\int_0^1 (f'(y + t(x-y)) - f'(y))(x-y)dt\right| \leqslant c_\alpha|x-y|^{1+\alpha}.$$

which is valid for $f \in C^{1+\alpha}$. On the other hand, we have

$$|(Y(s,\zeta) - \widetilde{Y}^{t,z}(s,\zeta))_2| \leqslant \lambda_2|\zeta - \gamma_{t,s}(z)|.$$

Therefore, by (2.6.24), we have

$$\begin{aligned}
|E_3| &\leqslant C\left(|\zeta - \gamma_{t,s}(z)|_B^{1+\alpha}\partial_\xi Z(t,z;s,\zeta) + |\zeta - \gamma_{t,s}(z)|_B \partial_\nu Z(t,z;s,\zeta)\right) +\\
&\quad + C|\zeta - \gamma_{t,s}(z)|_B^3\left(\partial_\xi Z(t,z;s,\zeta) + \partial_\nu Z(t,z;s,\zeta)\right)\\
&\leqslant C(s-t)^{\frac{\alpha}{2}-1}p(t,z;s,\zeta).
\end{aligned}$$

The general case $k > 1$ follows by induction as in the proof of Proposition 1.4.8, exploiting the reproduction property (2.6.26). Then the convergence of the series and (2.6.27) follow immediately.

Estimate (2.6.28) can be directly derived from (2.6.27) if $|\zeta - \zeta'|_B \geqslant \sqrt{s-t}$. When $|\zeta - \zeta'|_B \leqslant \sqrt{s-t}$ we may again proceed by induction. We consider the case $k = 1$ only. Write

$$
\begin{aligned}
&\left|H(t,z;s,\zeta) - H(t,z;s,\zeta')\right| \\
&= \left|\left(\mathcal{L}_s - \widetilde{\mathcal{L}}_s^{t,z}\right) Z(t,z;s,\zeta) - \left(\mathcal{L}_s - \widetilde{\mathcal{L}}_s^{t,z}\right) Z(t,z;s,\zeta')\right| \\
&\leqslant \left|a(s,\zeta) - a(s,\zeta')\right| \left|\partial_{\nu\nu} Z(t,z;s,\zeta')\right| \\
&\qquad + \left|a(s,\zeta) - a(s,\gamma_{t,s}(z))\right| \left|\partial_{\nu\nu} Z(t,z;s,\zeta) - \partial_{\nu\nu} Z(t,z;s,\zeta')\right| \\
&\qquad + \left|Y_1(s,\zeta) - Y_1(s,\zeta') + \partial_\nu Y_1(s,\gamma_{t,s}(z))(\zeta - \zeta')_2\right| \left|\partial_\xi Z(t,z;s,\zeta')\right| \\
&\qquad + \left|Y_1(s,\zeta) - Y_1(s,\gamma_{t,s}(z)) + \partial_\nu Y_1(s,\gamma_{t,s}(z))(\zeta - \gamma_{t,s}(z))_2\right| \left|\partial_\xi Z(t,z;s,\zeta) - \partial_\xi Z(t,z;s,\zeta')\right| \\
&\qquad + \{\text{non singular terms}\}
\end{aligned}
$$

(by (2.6.24)-(2.6.25) and neglecting the terms in the parentheses for brevity)

$$
\begin{aligned}
&\leqslant C\Big(\frac{|\zeta - \zeta'|^{\alpha}}{s-t} p(t,z;s,\zeta') + \frac{|\zeta - \zeta'|_B}{(s-t)^{\frac{3}{2} - \frac{\alpha}{2}}} p(t,z;s,\zeta) + \\
&\qquad + \frac{|\zeta - \zeta'|^{1+\alpha}}{(s-t)^{\frac{3}{2}}} p(t,z;s,\zeta') + \frac{|\zeta - \zeta'|_B}{(s-t)^{2-\frac{1}{2}(1+\alpha)}} p(t,z;s,\zeta) + \cdots \Big)
\end{aligned}
$$

(using that $|\zeta - \zeta'|_B \leqslant (s-t)^{\frac{1}{2}}$)

$$
\leqslant C \frac{|\zeta - \zeta'|_B^{\frac{\alpha}{2}}}{(s-t)^{1-\frac{\alpha}{4}}} \left(p(t,z;s,\zeta) + p(t,z;s,\zeta')\right).
$$

□

Potential estimates and C^2_ν regularity

Proposition 2.6.18. *There exist positive constants C, $\mu \geqslant 1$ depending on Θ such that, for any $j = 1,2$, $0 \leqslant t < s \leqslant T$ and $\zeta = (\xi,\nu), z \in \mathbb{R}^2$ we have*

$$
|\partial_\nu^j \Gamma(t,z;s,\xi,\nu)| \leqslant C(s-t)^{-\frac{j}{2}} \Gamma^{\text{heat}}\left(\mu \mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)\right).
$$

Proof. The derivative of the main term of the parametrix expansion $\partial_\nu^j Z$ is readily controlled using the bounds 2.6.20. The main difficulty is therefore to prove that $\partial_\nu^j(\Phi \otimes Z)$ exists and satisfies the indicated bounds. We proceed in two steps:

1. We prove that $\Phi \otimes \partial_\nu^j Z$ is indeed well defined and satisfies the stated estimates.

2. We justify that actually $\Phi\otimes\partial_\nu^j Z = \partial_\nu^j(\Phi\otimes Z)$. This last identity amounts to say that the spatial derivatives can be exchanged with the integrals of the time-space convolution $\otimes$. This point is established by taking the limit in a suitable cut-off in time procedure which allows to get away from the pole.

Such an approach was already used for instance in [24].

The first derivative estimate directly follows from the parametrix representation (2.6.12), estimate (2.6.24) and the controls of Proposition 2.6.17. Namely

$$\begin{aligned}|(\Phi\otimes\partial_\nu Z)(t,z;s,\zeta)| &\leqslant C\int_t^s(\varrho-t)^{\frac{\alpha}{2}-1}(s-\varrho)^{-\frac12}\int_{\mathbb{R}^2}p(t,z;\varrho,y)p(\varrho,y;s,\zeta)dyd\varrho\\ &\leqslant C(s-t)^{-\frac12+\frac{\alpha}{2}}p(t,z;s,\zeta).\end{aligned}$$

When trying to get controls for higher order derivatives, some time singularities appear in the integrals. A way to overcome such a problem is to exploit cancellation properties of the derivatives of the Gaussian kernels, namely for fixed $(t_0,z_0)\in[0,T]\times\mathbb{R}^d$:

$$\int_{\mathbb{R}^2}\partial_\nu^j\widetilde{\Gamma}^{t_0,z_0}(t,z;s,\zeta)dz = 0, \qquad j=\{1,2\}. \tag{2.6.30}$$

Then we have

$$\begin{aligned}(\Phi\otimes\partial_{\nu\nu}Z)(t,z;s,\zeta) = &\int_t^s\int_{\mathbb{R}^2}\Phi(t,z;\varrho,y)\left(\partial_{\nu\nu}Z-\partial_{\nu\nu}\widetilde{\Gamma}^{t_0,z_0}\right)(\varrho,y;s,\zeta)dyd\varrho+\\ &+\int_t^s\int_{\mathbb{R}^2}\left(\Phi(t,z;\varrho,y)-\Phi(t,z;\varrho,y')\right)\partial_{\nu\nu}\widetilde{\Gamma}^{t_0,z_0}(\varrho,y;s,\zeta)dyd\varrho.\end{aligned}$$

It then remains to appropriately choose the freezing parameters (t_0,z_0) and y' to exploit the regularity of the terms in the above r.h.s. in order to balance the singularities deriving from the spatial differentiation $\partial_{\nu\nu}$.

To do so, and conclude the proof of Proposition 2.6.18, we need the following lemma, whose proof is postponed to the Appendix A.

Lemma 2.6.19. *There exist positive constants $C,\mu\geqslant 1$ depending on Θ such that, for any $j=0,1,2$, $\zeta=(\xi,\nu), z\in\mathbb{R}^2$ and $0\leqslant t<s\leqslant T$, we have*

$$\left|\partial_\nu^j\widetilde{\Gamma}^{t_0,z_0}(t,z;s,\zeta)-\partial_\nu^j Z(t,z;s,\zeta)\right|_{(t_0,z_0)=(s,\zeta)}\leqslant C(s-t)^{\frac{\alpha-j}{2}}\Gamma^{\text{heat}}\left(\mu\mathcal{D}_{s-t},\zeta-\gamma_{t,s}(z)\right). \tag{2.6.31}$$

The above Lemma suggests that a natural choice consists in choosing $(t_0,z_0)=(s,\zeta)$, $y'=\gamma_{s,\tau}(\zeta)$. Indeed one gets from (2.6.30) and Proposition (2.6.17)

$$
\begin{aligned}
&|(\Phi \otimes \partial_{\nu\nu} Z)(t,z;s,\zeta)| \\
&\leqslant C \int_t^s \int_{\mathbb{R}^2} (\varrho - t)^{\frac{\alpha}{2}-1} \Gamma^{\text{heat}}\left(\mu \mathcal{D}_{\varrho-t}, y - \gamma_{t,\varrho}(z)\right) (s-\varrho)^{\frac{\alpha}{2}-1} \Gamma^{\text{heat}}\left(\mu \mathcal{D}_{s-\varrho}, \zeta - \gamma_{\varrho,s}(y)\right) dy d\varrho \\
&\qquad + C \int_t^s \int_{\mathbb{R}^2} \frac{|y - \gamma_{s,\varrho}(\zeta)|^{\frac{\alpha}{2}}}{(\varrho - t)^{1-\frac{\alpha}{4}}} \Big(\Gamma^{\text{heat}}\left(\mu \mathcal{D}_{\varrho-t}, y - \gamma_{t,\varrho}(z)\right) + \Gamma^{\text{heat}}\left(\mu \mathcal{D}_{\varrho-t}, \gamma_{s,\varrho}(\zeta) - \gamma_{t,\varrho}(z)\right) \Big) \times \\
&\qquad\qquad \times (s-\varrho)^{-1} \Gamma^{\text{heat}}\left(\mu \mathcal{D}_{s-\varrho}, \zeta - \gamma_{\varrho,s}(y)\right) dy d\varrho
\end{aligned}
$$

(reasoning as in the proof of Proposition 2.6.17)

$$
\leqslant C(s-t)^{\frac{3\alpha}{4}-1} \Gamma^{\text{heat}}\left(\mu \mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)\right) \tag{2.6.32}
$$

Observe now that the previous computations could be reproduced to estimate for all $\varepsilon > 0$

$$
\partial_\nu^j G_\varepsilon(t,z;s,\zeta) := \partial_\nu^j \int_{t+\varepsilon}^s \int_{\mathbb{R}^d} \Phi(t,z;\varrho,y) Z(\varrho,y;s,\zeta) dy d\varrho = \int_{t+\varepsilon}^s \int_{\mathbb{R}^d} \Phi(t,z;\varrho,y) \partial_\nu^j Z(\varrho,y;s,\zeta) dy d\varrho,
$$

where the last equality follows from the bounded convergence Theorem. We would obtain, uniformly in $\epsilon \in (0, \epsilon_0)$, $\epsilon_0 > 0$ meant to be small, a control similar to (2.6.32) for $\partial_\nu^j G_\varepsilon(t,z;s,\zeta)$.

Letting now ϵ go to zero, we derive that $\partial_\nu^j (\Phi \otimes Z)(t,z;s,\zeta) = (\Phi \otimes \partial_\nu^j Z)(t,z;s,\zeta)$, which together with (2.6.32) completes the proof of the statement for the derivatives.

Proof of Theorem 2.6.6

Let us now derive, under Assumptions 2.6.1 and 2.6.2 the main result of Theorem 2.6.6. We already proved that the function $\Gamma = \Gamma(t,z;s,\zeta)$ defined in (2.6.12) belongs to $C_{t_0,T}$ for any $\tau < t < t_0 < T$, $z \in \mathbb{R}^2$, is twice continuously differentiable in the variable ν and satisfies the Gaussian upper bounds in (2.6.4)-(2.6.6). It remains to prove that Γ is indeed the fundamental solution for $\mathcal{K}_s$ and that it satisfies the lower bound in (2.6.4).

We introduce a regularized version of the PDE (2.6.1): let $a^\varepsilon := a \star \varphi_\varepsilon$, $b^\varepsilon := b \star \varphi_\varepsilon$, $c^\varepsilon := c \star \varphi_\varepsilon$, $Y^\varepsilon := Y \star \varphi_\varepsilon$ where $\varphi_\varepsilon(\cdot) = \varepsilon^{-3} \varphi(\cdot/\varepsilon)$, $\varphi \in C_0^\infty(\mathbb{R}^3; \mathbb{R}_+)$ is a standard time-space mollifier in $\mathbb{R}^3$ and $\star$ denotes the time-space convolution. It follows from the Hörmander theorem that the PDE

$$
a^\varepsilon(s,\zeta) \partial_{\nu\nu} u(s,\zeta) + b^\varepsilon(s,\zeta) \partial_\nu u(s,\zeta) + c^\varepsilon(s,\zeta) u(s,\zeta) = \partial_s u(s,\zeta) + \langle Y^\varepsilon(s,\zeta), \nabla_\zeta u(s,\zeta) \rangle,
$$

admits a smooth fundamental solution $\Gamma^\varepsilon = \Gamma^\varepsilon(t,z;s,\zeta)$. In particular Γ^ε satisfies

$$
\begin{aligned}
\Gamma^\varepsilon(t,z;s,\gamma^\varepsilon_{t_0,s}(\zeta)) = &\Gamma^\varepsilon(t,z;t_0,\zeta) + \int_{t_0}^s \Big(a^\varepsilon(\varrho, \gamma^\varepsilon_{t_0,\varrho}(\zeta)) \partial_{\nu\nu} \Gamma^\varepsilon(t,z;\varrho,\gamma^\varepsilon_{t_0,\varrho}(\zeta)) + \\
&+ b^\varepsilon(\varrho, \gamma^\varepsilon_{t_0,\varrho}(\zeta)) \partial_\nu \Gamma^\varepsilon(t,z;\varrho,\gamma^\varepsilon_{t_0,\varrho}(\zeta)) + c^\varepsilon(\varrho, \gamma^\varepsilon_{t_0,\varrho}(\zeta)) \Gamma^\varepsilon(t,z;\varrho,\gamma^\varepsilon_{t_0,\varrho}(\zeta)) \Big) d\varrho
\end{aligned} \tag{2.6.33}
$$

for every $t < t_0 \leqslant s \leqslant T$ and $z \in \mathbb{R}^2$, where $\gamma^{\varepsilon}_{t_0,s}(\zeta)$ stands for the integral curve of Y^{ε} with initial datum $\gamma^{\varepsilon}_{t_0,t_0} = \zeta$, and

$$\lim_{\substack{(s,\zeta)\to(t,z_0)\\ s>t}} \int_{\mathbb{R}^2} \Gamma^{\varepsilon}(t,z;s,\zeta) f(z) dz = f(z_0), \tag{2.6.34}$$

for any bounded and continuous function f and $z_0 \in \mathbb{R}^2$. Importantly, it also satisfies the upper bounds in (2.6.4)-(2.6.6), uniformly in $\varepsilon \in (0, \varepsilon_0]$, $\varepsilon_0 > 0$ meant to be small.

On the other hand, it is clear that we can write

$$\Gamma^{\varepsilon}(t,z;s,\zeta) = Z^{\varepsilon}(t,z;s,\zeta) + \left(\Phi^{\varepsilon} \otimes Z^{\varepsilon}\right)(t,z;s,\zeta), \tag{2.6.35}$$

with the obvious definitions of Z^{ε} and Φ^{ε}, which satisfy as well, uniformly in $\varepsilon \in (0, \varepsilon_0]$ the estimates in (2.6.17), (2.6.20), (2.6.27), (2.6.28).

Observe now that the RHS of (2.6.35) converges pointwise to Γ by construction. Similarly $\partial^j_{\nu}\Gamma^{\varepsilon}$ converges pointwise to $\partial^j_{\nu}\Gamma$ for $j = 1, 2$. Moreover, a direct computation shows that, for any $z \in \mathbb{R}^3$, $t_0 \in (t, T)$, there exists a constant c such that

$$\left|\mathcal{D}_{\frac{1}{\sqrt{\varrho - t}}}\left(\zeta - \gamma_{t,\varrho}(z)\right)\right| - c \leqslant \left|\mathcal{D}_{\frac{1}{\sqrt{\varrho - t}}}\left(\zeta - \gamma^{\varepsilon}_{t,\varrho}(z)\right)\right| \leqslant \left|\mathcal{D}_{\frac{1}{\sqrt{\varrho - t}}}\left(\zeta - \gamma_{t,\varrho}(z)\right)\right| + c,$$

for any $\varepsilon \in (0, \varepsilon_0]$, $\varrho \in [t_0, s]$ and $\zeta \in \mathbb{R}^2$. Therefore we deduce that for any $z \in \mathbb{R}^2$, $\tau \leqslant t < t_0 < T$, the functions $\partial^j_{\nu}\Gamma^{\varepsilon}(t,z;\cdot,\cdot)$ converge pointwise and *boundedly* to $\partial^j_{\nu}\Gamma(t,z;\cdot,\cdot)$ in $[t_0,T] \times \mathbb{R}^3$. Thus we can take the limit for $\varepsilon \to 0$ in equations (2.6.33) and (2.6.34) under the integral signs by the bounded convergence theorem and get the first part of the thesis.

Lower bound for the fundamental solution. Similarly to the parabolic case, we first derive a local bound, starting from the parametrix expansion (2.6.12) and exploiting the results of the previous Sections . We have

$$\Gamma(t,z;s,\zeta) \geqslant Z(t,z;s,\zeta) - \int_t^s \int_{\mathbb{R}^2} |\Phi(t,z;\varrho,y) Z(\varrho,y;s,\zeta)| d\varrho dy \geqslant$$

(by Proposition 2.6.11 and Lemma 2.6.17)

$$\begin{aligned}
&\geqslant C^{-1}\Gamma^{\text{heat}}\left(\mu^{-1}\mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)\right) \\
&\quad - \int_t^s C(\varrho - t)^{\frac{\alpha}{2}-1} \int_{\mathbb{R}^2} \Gamma^{\text{heat}}(\mu\mathcal{D}_{\varrho-t}, y - \gamma_{t,\varrho}(z)) \Gamma^{\text{heat}}(\mu\mathcal{D}_{s-\varrho}, \zeta - \gamma_{\varrho,s}(y)) d\varrho dy \geqslant \\
&\geqslant C^{-1}\Gamma^{\text{heat}}\left(\mu^{-1}\mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)\right) - \frac{C}{2}(s-t)^{\frac{\alpha}{2}} \Gamma^{\text{heat}}\left(\mu\mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)\right).
\end{aligned}$$

Let $d_{t_1,t_2}(z_1,z_2) := \left|\mathcal{D}_{\frac{1}{\sqrt{t_2 - t_1}}}\left(z_2 - \gamma_{t_1,t_2}(z_1)\right)\right|$ denote the "control metric" of the system. A direct computation shows that $\Gamma^{\text{heat}}(c\mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)) \leqslant \Gamma^{\text{heat}}(c^{-1}\mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z))$ if $d_{t,s}(z,\zeta) \leqslant \varrho_c$ where

$\varrho_c = \sqrt{\frac{4c\ln c}{c^2-1}}$. Then we have

$$\Gamma(t,z;s,\zeta) \geqslant \left(\frac{1}{C^2} - \frac{(s-t)^{\frac{\alpha}{2}}}{2}\right)\Gamma^{\text{heat}}(\mu^{-1}\mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)) \geqslant \frac{1}{2C}\Gamma^{\text{heat}}(\mu^{-1}\mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)) \tag{2.6.36}$$

if $d_{t,s}(z,\zeta) \leqslant \varrho_\mu$ and $0 < s-t \leqslant T_C := C^{-\frac{4}{\alpha}}$.

In order to pass from the local to the global bound, we use a chaining procedure: we first define a sequence of points (t_k, z_k) such that $t_0 = t, z_0 = z, t_{M+1} = s, z_{M+1} = \zeta$ for some integer M (to be defined later), along which we can control the increments with respect to the control metric $d_{t_{k-1},t_k}(z_k, z_{k+1})$. Let us consider the controlled version of the system (2.6.7):

$$\psi_{t,\varrho}(z) = z + \int_t^\varrho (Y(u, \psi_{t,u}(z)) + v(u)\mathbf{e}_2)\,du, \qquad \varrho \in [t,s].$$

We have the following (see [14], Propositions 4.1 and 4.2):

Lemma 2.6.20. *There exists a control $(v(\varrho))_{t\leqslant\varrho\leqslant s}$ with values in $\mathbb{R}^2$ such that*

i) the solution $\psi_{t,\varrho}(z)$ associated with $v(\varrho)$ reaches ζ at time s, that is $\psi_{t,s}(z) = \zeta$;

ii) there exist two constants $m_1, m_2 > 0$, only dependent on the constants of Assumptions 2.6.1-2.6.2, such that

$$\int_t^s |v(\varrho)|^2 d\varrho \geqslant m_1 \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(\zeta - \gamma_{t,s}(z))\right|^2, \qquad \sup_{t\leqslant\varrho\leqslant s} |v(\varrho)|^2 \leqslant \frac{m_2}{s-t}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(\zeta - \gamma_{t,s}(z))\right|^2.$$

We set

$$t_i = t + i\frac{s-t}{M+1} = t + i\varepsilon, \quad z_k = \psi_{t,t_k}(z), \qquad i = 1, \cdots, M,$$

where $\psi_{t,\varrho}(z)$ is the optimal path of Lemma 2.6.20 and M is the smallest integer greater than

$$\max\left\{\frac{K^2 d_{t,s}^2(z,\zeta)}{\varrho_\mu^2}, \frac{T}{T_C}\right\}.$$

with $K = \frac{12m^2m_2}{m_1}$, where m, m_1 and m_2 are the constants in Lemmas 2.6.16 and 2.6.20. Finally we define the sets

$$B_i(r) := \left\{y \in \mathbb{R}^2 \mid \left|\mathcal{D}_{\frac{1}{\sqrt{\varepsilon}}}(y - \gamma_{t_{i-1},t_i}(z_{i-1}))\right| + \left|\mathcal{D}_{\frac{1}{\sqrt{\varepsilon}}}(z_{i+1} - \gamma_{t_i,t_{i+1}}(y))\right| \leqslant r\right\},$$

and write

$$\Gamma(t,z;s,\zeta) \geqslant \int_{B_1(\varrho_c/3)} \cdots \int_{B_M(\varrho_c/3)} \Gamma(t,z;t_1,\zeta_1) \prod_{j=1}^{M-1} \Gamma(t_j,\zeta_j;t_{j+1},\zeta_{j+1})\Gamma(t_M,\zeta_M;s,\zeta)d\zeta_1 \ldots d\zeta_M. \tag{2.6.37}$$

By definition of M we have

$$t_{j+1}-t_j=\frac{s-t}{M+1}\leqslant\frac{T}{M+1}\leqslant T_C.$$

On the other hand, if $\zeta_i\in B_i\left(\frac{\varrho_c}{3}\right)$ for $i=1,\dots,M-1$ we have

$$\begin{aligned}
&d_{t_i,t_{i+1}}(\zeta_i,\zeta_{i+1})\\
&\quad=\left|\mathcal{D}_{\frac{1}{\sqrt{\varepsilon}}}\left(\zeta_{i+1}-\gamma_{t_i,t_{i+1}}(\zeta_i)\right)\right|\\
&\quad\leqslant\left|\mathcal{D}_{\frac{1}{\sqrt{\varepsilon}}}\left(\zeta_{i+1}-\gamma_{t_i,t_{i+1}}(z_i)\right)\right|+\left|\mathcal{D}_{\frac{1}{\sqrt{\varepsilon}}}\left(z_{i+1}-\gamma_{t_i,t_{i+1}}(z_i)\right)\right|+\left|\mathcal{D}_{\frac{1}{\sqrt{\varepsilon}}}\left(z_{i+1}-\gamma_{t_i,t_{i+1}}(\zeta_i)\right)\right|\\
&\quad=:E_1+E_2+E_3,
\end{aligned}$$

where $E_1+E_3\leqslant\frac{2}{3}\varrho_\mu$. By Lemma 2.6.20, we have

$$\begin{aligned}
E_2&\leqslant m_1^{-1}\left(\int_{t_i}^{t_{i+1}}|v(\varrho)|^2d\varrho\right)^{\frac{1}{2}}\\
&\leqslant\frac{m_2}{m_1}\sqrt{\frac{\varepsilon}{s-t}}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\left(\zeta-\gamma_{t,s}(z)\right)\right|=\frac{m_2}{m_1}\frac{d_{t,s}(z,\zeta)}{\sqrt{M+1}}\leqslant\frac{\varrho_\mu}{12m^2}.
\end{aligned}\tag{2.6.38}$$

Therefore $d_{t_i,t_{i+1}}(\zeta_i,\zeta_{i+1})\leqslant\varrho_\mu$ and we can use (2.6.36) repeatedly in (2.6.37) to get

$$\Gamma(t,z;s,\zeta)\geqslant(2C)^{-(M+1)}\left|\prod_{i=1}^{M}B_i\left(\frac{\varrho_\mu}{3}\right)\right|\left(\frac{C(M+1)^2}{(s-t)^2}\right)^{M+1}\exp\left(-\frac{C}{2}\varrho_\mu^2(M+1)\right).$$

Assume for a moment the validity of the inequality

$$\left|B_i\left(\frac{\varrho_\mu}{3}\right)\right|\geqslant C_0\pi\left(\frac{s-t}{M+1}\right)^2\varrho_\mu^2\tag{2.6.39}$$

for some positive constant C_0 (only dependent on the constants of Assumptions 2.6.1-2.6.2). Then we have

$$\Gamma(t,z;s,\zeta)\geqslant C_1C_2^M\frac{1}{2\pi\sqrt{\det\mathcal{D}_{s-t}}}\exp\left(-\frac{C}{2}\varrho_\mu^2M\right)\geqslant\frac{C_3}{2\pi\sqrt{\det\mathcal{D}_{s-t}}}\exp\left(-\frac{C_4}{2}M\right),$$

for some positive constants $C_1,\dots,C_4$. Now, if $TT_C^{-1}\leqslant\frac{K^2d_{t,s}^2(z,\zeta)}{\varrho_\mu^2}$ and $M<2\frac{K^2d_{t,s}^2(z,\zeta)}{\varrho_\mu^2}$, we have

$$\Gamma(t,z;s,\zeta)\geqslant\frac{C_3}{2\pi\sqrt{\det\mathcal{D}_{s-t}}}\exp\left(-\frac{C_5}{2}d_{t,s}^2(z,\zeta)\right)=C_6\Gamma^{\text{heat}}(C_5^{-1}\mathcal{D}_{s-t},\zeta-\gamma_{t,s}(z)).$$

On the other hand, if $M<2TT_C^{-1}$ then

$$\Gamma(t,z;s,\zeta)\geqslant\frac{C_7}{2\pi\sqrt{\det\mathcal{D}_{s-t}}}\geqslant\frac{C_7}{2\pi\sqrt{\det\mathcal{D}_{s-t}}}\exp\left(-\frac{C_5}{2}d_{t,s}^2(z,\zeta)\right)=C_8\Gamma^{\text{heat}}(C_5^{-1}\mathcal{D}_{s-t},\zeta-\gamma_{t,s}(z)),$$

and this proves the lower bound.

We are left with the proof of (2.6.39). Let $\widetilde{B}_i(r) = \{y \in \mathbb{R}^2,\ |\mathcal{D}_{\frac{1}{\sqrt{\varepsilon}}}(y - z_i)| \leqslant r\}$: a direct computation shows $|\widetilde{B}_i(r)| = \pi\varepsilon^2 r^2$. Then it is enough to show that $B_i\left(\frac{\varrho_\mu}{3}\right) \supseteq \widetilde{B}_i\left(\frac{\varrho_\mu}{C}\right)$ for a positive constant C (only dependent on Θ). For any $y \in \widetilde{B}_i(r)$ we have

$$\begin{aligned}&\left|\mathcal{D}_{\frac{1}{\sqrt{\varepsilon}}}\left(y - \gamma_{t_{i-1},t_i}(z_{i-1})\right)\right| + \left|\mathcal{D}_{\frac{1}{\sqrt{\varepsilon}}}\left(z_{i+1} - \gamma_{t_i,t_{t+1}}(y)\right)\right| \leqslant \\ &\quad \leqslant \left|\mathcal{D}_{\frac{1}{\sqrt{\varepsilon}}}(z - z_i)\right| + \left|\mathcal{D}_{\frac{1}{\sqrt{\varepsilon}}}\left(z_i - \gamma_{t_{i-1},t_i}(z_{i-1})\right)\right| + m\left|\mathcal{D}_{\frac{1}{\sqrt{\varepsilon}}}(z - z_i)\right| + m^2\left|\mathcal{D}_{\frac{1}{\sqrt{\varepsilon}}}(z_{i+1} - \gamma_{t_i,t_{i+1}}(z_i))\right|\end{aligned}$$

(by (2.6.38))

$$\leqslant (1+m)r + \frac{\varrho_\mu}{6}.$$

Then it is sufficient to take $r \leqslant \frac{\varrho_\mu}{6(1+m)}$ and this concludes the proof. □

2.6.2 Proof of Theorem 2.1.6

For any fixed $t \in [0,T)$ and $\omega \in \Omega$, let $\mathcal{K}^{(t)}$ the operator of the form (2.6.1), as defined by (2.5.7) and (2.5.11) through the random change of variable $\gamma_{t,s}^{\mathrm{IW}}$. By Assumptions 2.1.3-2.1.4 and Lemma 2.5.1, $\mathcal{K}^{(t)}$ satisfies Assumptions 2.6.1-2.6.2 for a.e. $\omega \in \Omega$. Then, by Theorem 2.6.6, $\mathcal{K}^{(t)}$ admits a fundamental solution $\Gamma^{(t)}$: we set

$$\mathbf{\Gamma}(t,z;s,\xi,\nu) = \Gamma^{(t)}(t,z;s,\xi,\gamma_{t,s}^{\mathrm{IW},-1}(\xi,\nu)), \qquad t < s \leqslant T,\ \xi,\nu \in \mathbb{R},\ z \in \mathbb{R}^2. \tag{2.6.40}$$

Combining Theorems 2.5.2, 2.6.6 and Lemma 2.5.1 we infer that $\mathbf{\Gamma}(t,z;\cdot,\cdot) \in \mathbf{C}^0_{t_0,T}$ for any $t_0 \in (t,T]$, is twice continuously differentiable in the variable ν and is a solution to (2.1.2) in $[t_0,T]$ (in the sense of definition 2.1.1). Now, for any bounded and continuous function φ and $z_0 \in \mathbb{R}^2$, we have

$$\begin{aligned}\int_{\mathbb{R}^2}\varphi(z)\mathbf{\Gamma}(t,z;s,\zeta)dz - \varphi(z_0) &= \int_{\mathbb{R}^2}\varphi(z)\Gamma^{(t)}(t,z;s,\zeta)dz - \varphi(z_0) + \\ &\quad + \int_{\mathbb{R}^2}\varphi(z)\left(\Gamma^{(t)}(t,z;s,\xi,\gamma_{t,s}^{\mathrm{IW},-1}(\xi,\nu)) - \Gamma^{(t)}(t,z;s,\zeta)\right)dz = \\ &= I_1(t,s,\zeta) + I_2(t,s,\zeta).\end{aligned}$$

Now, by Theorem 2.6.6 and the dominated convergence theorem, we have

$$\lim_{\substack{(s,\zeta)\to(t,z_0)\\ s>t}} I_i(t,s,\zeta) = 0, \qquad i = 1,2.$$

This proves the first part of the thesis.

The Gaussian bounds (2.1.5), (2.1.6) follow directly from the definition (2.6.40) and the analogous estimates (2.6.4) for $\Gamma^{(t)}$ in Theorem 2.6.6. Moreover, since

$$\partial_\nu \mathbf{\Gamma}(t,z;s,\zeta) = (\partial_\nu \Gamma^{(t)})\left(t,z;s,\xi,\gamma^{\mathrm{IW},-1}_{t,s}(\xi,\nu)\right)\partial_\nu \gamma^{\mathrm{IW},-1}_{t,s}(\xi,\nu),$$

the gradient estimate (2.1.7) follows from the analogous estimate (2.6.5) for $\Gamma^{(t)}$ and from Lemma 2.5.1. The proof of (2.1.8) is analogous. □

2.7 The backward Langevin SPDE

In this section we show how the general results from Section 2.1.1 can be derived without significant modifications to our methods, for the backward version of equation (2.1.2), that is

$$-d_{\mathbf{B}} u_t(z) = \mathcal{A}_{t,z} u_t(z) dt + \mathcal{G}_{t,z} u_t(z) \star dW_t, \qquad \mathbf{B} = \partial_t + v_1 \partial_x. \tag{2.7.1}$$

Here the symbol $\star dW_t$ means that (2.7.1) is written in terms of the *backward Itô integral*: for reader convenience we recall its definition and some basic results about the backward Itô calculus in Appendix B.

We denote by $\overleftarrow{\mathbf{C}}^{k+\alpha}_{t,T}$ (and $\mathbf{b}\overleftarrow{\mathbf{C}}^{k+\alpha}_{t,T}$) the stochastic Hölder spaces formally defined as in Section 1.1.1 with $\mathcal{P}_{t,T}$ in condition ii) replaced by the backward predictable σ-algebra $\overleftarrow{\mathcal{P}}_{t,T}$ defined in terms of the backward Brownian filtration (cf. Section B). Again, (2.7.1) is understood in the strong sense:

Definition 2.7.1. *A solution to* (2.7.1) *on* $[0,s]$ *is a process* $u = u_t(x,v) \in \overleftarrow{\mathbf{C}}^{0}_{0,s}$ *that is twice continuously differentiable in the variables* v *and such that*

$$u_t\left(\gamma^{\mathbf{B}}_{s-t}(z)\right) = u_s(z) + \int_t^s \mathcal{A}_{\tau,\gamma^{\mathbf{B}}_{s-\tau}(z)} u_\tau(\gamma^{\mathbf{B}}_{s-\tau}(z)) d\tau + \int_t^s \mathcal{G}_{\tau,\gamma^{\mathbf{B}}_{s-\tau}(z)} u_\tau(\gamma^{\mathbf{B}}_{s-\tau}(z)) \star dW_\tau, \qquad t \in [0,s].$$

Definition 2.7.2. *A fundamental solution for the backward SPDE* (2.7.1) *is a stochastic process* $\overleftarrow{\mathbf{\Gamma}} = \overleftarrow{\mathbf{\Gamma}}(t,z;s,\zeta)$ *defined for* $0 \leqslant t < s \leqslant T$ *and* $z,\zeta \in \mathbb{R}^{d+1}$, *such that for any* $(s,\zeta) \in (0,T] \times \mathbb{R}^{d+1}$ *and* $t_0 \in (0,s)$ *we have:*

i) $\overleftarrow{\mathbf{\Gamma}}(\cdot,\cdot;s,\zeta)$ *is a solution to* (2.7.1) *on* $[0,t_0]$*;*

ii) for any $\varphi \in bC(\mathbb{R}^{d+1})$ *and* $z_0 \in \mathbb{R}^{d+1}$, *we have*

$$\lim_{\substack{(t,z)\to(s,z_0)\\ t<s}} \int_{\mathbb{R}^2} \overleftarrow{\mathbf{\Gamma}}(t,z;s,\zeta)\varphi(\zeta)d\zeta = \varphi(z_0), \qquad P\text{-a.s.}$$

For a fixed $s \in [0,T)$ and $(x,v) \in \mathbb{R}^{d+1}$ the backward SDE

$$\overleftarrow{\gamma}^{\mathrm{IW}}_{t,s}(x,v) = v + \int_t^s \sigma_\tau(x, \overleftarrow{\gamma}^{\mathrm{IW}}_{\tau,s}(x,v)) \star dW_\tau, \qquad t \in [0,s], \tag{2.7.2}$$

defines a backward flow of diffeomorphism $(x,v) \mapsto \overleftarrow{g}^{\mathrm{IW}}_{t,s}(x,v) := (x, \overleftarrow{\gamma}^{\mathrm{IW}}_{t,s}(x,v))$, which replaces $g^{\mathrm{IW}}_{t,s}$ in the analysis; moreover

$$\overleftarrow{\gamma}_{t,s}(\zeta) = \zeta + \int_t^s \overleftarrow{\mathbf{Y}}_{\tau,s}(\overleftarrow{\gamma}_{\tau,s}(\zeta)) d\tau, \qquad t \in [0,s],$$

defines the integral curve, *ending* at (s,ζ), of the vector field $\overleftarrow{\mathbf{Y}}_{t,s}$, which is defined analogously to $\mathbf{Y}_{t,s}$ in (2.1.4), formally replacing γ^{IW} with $\overleftarrow{\gamma}^{\mathrm{IW}}$.

Finally we replace Assumption 2.1.3 in Section 2.1.1 with the following:

Assumption 2.7.3. *For some $\alpha \in (0,1)$, we have: $a \in \mathbf{b}\,\overleftarrow{\mathbf{C}}^{\alpha}_{0,T}$, $\sigma \in \mathbf{b}\,\overleftarrow{\mathbf{C}}^{3+\alpha}_{0,T}$, $b, c \in \mathbf{b}\,\overleftarrow{\mathbf{C}}^{0}_{0,T}$, $h \in \mathbf{b}\,\overleftarrow{\mathbf{C}}^{2}_{0,T}$.*

Theorem 2.7.4. *Under Assumptions 2.7.3, 2.1.4 and 2.1.5, the backward SPDE* (2.7.1) *has a fundamental solution $\overleftarrow{\mathbf{\Gamma}}$ satisfying estimates*

$$\overleftarrow{\mathbf{\Gamma}}(t,z;s,\zeta) \geqslant \mu_2^{-1}\Gamma^{\mathrm{heat}}\lambda^{-1}\left(\mu_1^{-1}\mathcal{D}_{s-t}, \overleftarrow{g}^{\mathrm{IW},-1}_{t,s}(z) - \overleftarrow{\gamma}_{t,s}(\zeta)\right),$$

$$\overleftarrow{\mathbf{\Gamma}}(t,z;s,\zeta) \leqslant \mu_2\Gamma^{\mathrm{heat}}\left(\mu_1\mathcal{D}_{s-t}, \overleftarrow{g}^{\mathrm{IW},-1}_{t,s}(z) - \overleftarrow{\gamma}_{t,s}(\zeta)\right), \tag{2.7.3}$$

$$\left|\partial_{v_i}\overleftarrow{\mathbf{\Gamma}}(t,x,v;s,\zeta)\right| \leqslant \frac{\mu_2}{\sqrt{s-t}}\Gamma^{\mathrm{heat}}\left(\mu_1\mathcal{D}_{s-t}, \overleftarrow{g}^{\mathrm{IW},-1}_{t,s}(x,v) - \overleftarrow{\gamma}_{t,s}(\zeta)\right), \tag{2.7.4}$$

$$\left|\partial_{v_iv_j}\overleftarrow{\mathbf{\Gamma}}(t,x,v;s,\zeta)\right| \leqslant \frac{\mu_2}{s-t}\Gamma^{\mathrm{heat}}\left(\mu_1\mathcal{D}_{s-t}, \overleftarrow{g}^{\mathrm{IW},-1}_{t,s}(x,v) - \overleftarrow{\gamma}_{t,s}(\zeta)\right), \tag{2.7.5}$$

for every $0 \leqslant t < s \leqslant T$, $z = (x,v), \zeta \in \mathbb{R}^{d+1}$ and $i,j = 1, \dots d$, with probability one.

In the next chapter, we will use a deterministic backward Kolmogorov PDE to which Theorem 2.7.4 applies. Precisely, we will use the following

Corollary 2.7.5. *Let Assumption 2.1.4 with $\sigma \equiv 0$ be satisfied and let $a \in bC^{\alpha}_{0,T}$, $b, c \in bC^0_{0,T}$, for some $\alpha \in (0,1)$, and $\varphi \in bC(\mathbb{R}^{d+1})$. Then there exists a bounded solution of the backward Cauchy problem*

$$\begin{cases} -d_{\mathbf{B}} f(t,z) = \mathcal{A}_{t,z} f(t,z) dt, \\ f(T,\cdot) = \varphi, \end{cases} \tag{2.7.6}$$

in the sense of Definition 2.7.1, that is

$$f\left(t, \gamma^{\mathbf{B}}_{T-t}(z)\right) = \varphi(z) + \int_t^T \mathcal{A}_{s,\gamma^{\mathbf{B}}_{T-s}(z)} f\left(s, \gamma^{\mathbf{B}}_{T-s}(z)\right) ds, \qquad (t,z) \in [0,T] \times \mathbb{R}^{d+1}, \tag{2.7.7}$$

where $\gamma_s^{\mathbf{B}}(x,v) = (x + sv_1, v)$*. Moreover, if* $\varphi \in bC^{\alpha}(\mathbb{R}^{d+1})$ *for some* $\alpha \in (0,1)$ *then there exists a positive constant* C *such that,*

$$\sup_{(x,v)\in\mathbb{R}\times\mathbb{R}^d} |\partial_v^{\beta} f(t,x,v)| \leqslant C(T-t)^{-\frac{|\beta|-\alpha}{2}}, \qquad 1 \leqslant |\beta| \leqslant 2.$$

Proof of Theorem 2.7.4. In the backward case the computations are completely analogous to the forward one, since it only suffices to reverse the time in equations (2.1.2) and (2.5.7). Precisely, we introduce the "check" transform

$$\check{f}_{t,s}(x,v) := f_t(\xi, \overleftarrow{\gamma}^{\mathrm{IW}}_{t,s}(x,v)), \qquad t \in [0,s],$$

with $\overleftarrow{\gamma}^{\mathrm{IW}}_{t,s}$ as in (2.7.2). For a solution $u_t = u_t(z)$ to (2.7.1) on $[0,s]$, we define

$$v_{t,s}(z) := \varrho_{t,s}(z)\check{u}_{t,s}(z), \quad \varrho_{t,s}(z) := \exp\left(-\int_t^s \check{h}_\tau(z) \star dW_\tau - \frac{1}{2}\int_t^s \check{h}^2_\tau(z) d\tau\right),$$

which solves, on $[0,s]$, the deterministic equation with random coefficients

$$-d_{\overleftarrow{\mathbf{B}}} v_{t,s}(z) = \left(\overleftarrow{a}^*_{t,s}(z)\partial_{vv}v_{t,s} + \overleftarrow{b}^*_{t,s}(z)\partial_v v_{t,s}(z) + \overleftarrow{c}^*_{t,s}(z)v_{t,s}(z)\right)dt, \qquad \overleftarrow{\mathbf{B}} = \partial_t + \overleftarrow{\mathbf{Y}}_{t,s}, \tag{2.7.8}$$

where $\overleftarrow{\mathbf{Y}}_{t,s}$ and the coefficients are defined similarly to (2.5.8) and (2.5.10), exchanging the *hat*- and *check*-transforms in the definitions. As for the forward case, by Assumption 2.7.3 and Lemma 2.5.1, $a^*_{\cdot,s} \in \mathbf{b}\overleftarrow{\mathbf{C}}^{\alpha}_{0,s}$, $b^*_{\cdot,s}, c^*_{\cdot,s} \in \mathbf{b}\overleftarrow{\mathbf{C}}^{0}_{0,s}$, $\overleftarrow{\mathbf{Y}}_{\cdot,s} \in \overleftarrow{\mathbf{C}}^{0,1}_{0,s}$, $\partial_v(\overleftarrow{\mathbf{Y}}_{\cdot,s})_1 \in \mathbf{b}\overleftarrow{\mathbf{C}}^{\bar{\alpha}}_{0,s}$, for any $\bar{\alpha} \in [0,\alpha)$, and there exist two random, finite and positive constant $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_2$ such that, for $t \in [0,s]$ and $z \in \mathbb{R}^2$, we have

$$\boldsymbol{\lambda}_1^{-1} \leqslant \overleftarrow{a}^*_{t,s}(z) \leqslant \boldsymbol{\lambda}_1, \qquad \boldsymbol{\lambda}_2^{-1} \leqslant \partial_v(\overleftarrow{\mathbf{Y}}_{t,s}(z))_1 \leqslant \boldsymbol{\lambda}_2,$$

with probability one, which ensures the weak Hörmander condition to hold.

Next, we reset the notations and rewrite equation (2.7.8) as

$$\overleftarrow{\mathcal{A}}_t u_t(z) + \overleftarrow{\mathbf{Y}}_t u_t(z) + \partial_t u_t(z) = 0, \qquad t \in [0,s),\ z = (x,v) \in \mathbb{R}^2, \tag{2.7.9}$$

where $\overleftarrow{\mathcal{A}}_t$ is a second order operator of the form

$$\overleftarrow{\mathcal{A}}_t = \overleftarrow{a}_t\partial_{vv} + \overleftarrow{b}_t\partial_v + \overleftarrow{c}_t, \qquad z = (x,v) \in \mathbb{R}^2,$$

and $\overleftarrow{\mathbf{Y}}_t = (\overleftarrow{\mathbf{Y}}_t)_1\partial_x + (\overleftarrow{\mathbf{Y}}_t)_2\partial_v$. For a fixed $(s_0,\zeta_0) \in (0,s] \times \mathbb{R}^2$, we define the linearized version of (2.7.9), that is

$$\overleftarrow{\mathcal{A}}^{s_0,\zeta_0}_t u_t(z) + \overleftarrow{\mathbf{Y}}^{s_0,\zeta_0}_t u_t(z) + \partial_t u_t(z) = 0, \qquad t \in [0,s),\ z \in \mathbb{R}^2, \tag{2.7.10}$$

where the definition of $\overleftarrow{\mathbf{Y}}_t^{s_0,\zeta_0}$ is analogous to that of $\mathbf{Y}_s^{t_0,z_0}$ in (2.1.4) and

$$\overleftarrow{\mathcal{A}}_t^{s_0,\zeta_0} := \overleftarrow{a}_t(\overleftarrow{\gamma}_{t,s_0}(\zeta_0))\partial_{vv}, \qquad \overleftarrow{\gamma}_{t,s_0}(\zeta_0) = \zeta_0 + \int_t^{s_0} \overleftarrow{\mathbf{Y}}_\tau(\overleftarrow{\gamma}_{\tau,s_0}(\zeta_0))d\tau, \qquad t \in [0, s_0].$$

Equation (2.7.10) has an explicit fundamental solution $\overleftarrow{\Gamma}^{s_0,\zeta_0} = \overleftarrow{\Gamma}^{s_0,\zeta_0}(t, z; s, \zeta)$ of Gaussian type, that satisfies estimates analogous to (2.6.17) and (2.6.20). The *backward parametrix* for (2.7.9) is defined as

$$\overleftarrow{Z}(t, z; s, \zeta) = \overleftarrow{\mathbf{\Gamma}}^{s,\zeta}(t, z; s, \zeta), \qquad 0 \leqslant t < s \leqslant T, \ z, \zeta \in \mathbb{R}^2.$$

As for the forward case, we set

$$\overleftarrow{\Gamma}(t, z; s, \zeta) = \overleftarrow{Z}(t, z; s, \zeta) + (\overleftarrow{Z} \otimes \overleftarrow{\Phi})(t, z; s, \zeta), \tag{2.7.11}$$

with

$$\Phi(t, z; s, \zeta) := \sum_{k=1}^{\infty} \overleftarrow{H}^{\otimes k}(t, z; s, \zeta), \tag{2.7.12}$$

where $\overleftarrow{H}(t, z; s, \zeta) = \left(\overleftarrow{\mathcal{A}}_t + \overleftarrow{\mathbf{Y}}_t - \overleftarrow{\mathcal{A}}_t^{s_0,\zeta_0} - \overleftarrow{\mathbf{Y}}_t^{s_0,\zeta_0}\right) \overleftarrow{Z}(t, z; s, \zeta)$ and the rest of the proof proceeds as in the forward case. In particular, existence and estimates for the fundamental solution of (2.7.8) (in the sense of Definitions 2.7.2) follow from the parametrix expansions (2.7.11) and (2.7.12). Eventually, it suffices to go back to the original variables to conclude the proof. □

Proof of Corollary 2.7.5. By Theorem 2.7.4 there exists a fundamental solution $\overleftarrow{\mathbf{\Gamma}}$ of equation (2.7.7), in the sense of Definition 2.7.2. Moreover, since $\sigma \equiv 0$, $\overleftarrow{\mathbf{\Gamma}}$ satisfies estimates (2.7.3), (2.7.4) and (2.7.5) with $\overleftarrow{g}_{t,s}^{\mathrm{IW},-1} \equiv \mathrm{Id}$ and $\overleftarrow{\gamma}_{t,s}(\zeta) = \gamma_{t-s}^{\mathbf{B}}(\zeta)$ as in Definition 2.7.1. Then, the function

$$f_t(z) := \int_{\mathbb{R}^{d+1}} \overleftarrow{\mathbf{\Gamma}}(t, z, T, \zeta)\varphi(\zeta)d\zeta, \qquad (t, z) \in [0, T] \times \mathbb{R}^{d+1},$$

solves problem (2.7.6). Since $\varphi \in bC(\mathbb{R}^{d+1})$, we have

$$\sup_{z\in\mathbb{R}^{d+1}} |f_t(z)| \leqslant \|\varphi\|_\infty \sup_{z\in\mathbb{R}^{d+1}} \int_{\mathbb{R}^{d+1}} \overleftarrow{\mathbf{\Gamma}}(t, z, T, \zeta)d\zeta \leqslant C$$

for a positive constant C. Moreover, since

$$\int_{\mathbb{R}^{d+1}} \nabla_v^j \mathbf{\Gamma}(t, x, v, T, \zeta)d\zeta = \nabla_v^j \int_{\mathbb{R}^{d+1}} \mathbf{\Gamma}(t, x, v, T, \zeta)d\zeta = 0, \qquad 1 \leqslant |j| \leqslant 2,$$

for any $w \in \mathbb{R}^3$ we have

$$\left|\nabla_v^j f_t(x, v)\right| \leqslant \int_{\mathbb{R}^{d+1}} \left|\nabla_v^j \mathbf{\Gamma}(t, x, v, T, \zeta)\right| |\varphi(\zeta) - \varphi(w)| \, dz$$

(choosing $w = \gamma^{\mathbf{B}}_{T-t}(z)$)

$$\begin{aligned}
&\leqslant C(T-t)^{-\frac{j}{2}} \int_{\mathbb{R}^{d+1}} \Gamma^{\text{heat}}\left(\mu \mathcal{D}_{T-t}, \gamma^{\mathbf{B}}_{T-t}(z) - \zeta\right) \left|\gamma^{\mathbf{B}}_{T-t}(z) - \zeta\right|^{\alpha} d\zeta \\
&\leqslant C'(T-t)^{\frac{\alpha-j}{2}} \int_{\mathbb{R}^{d+1}} \Gamma^{\text{heat}}\left(\mu' \mathcal{D}_{T-t}, \gamma^{\mathbf{B}}_{T-t}(z) - \zeta\right) d\zeta \leqslant C''(T-t)^{\frac{\alpha-j}{2}}.
\end{aligned}$$

The proof is complete. $\square$

Chapter 3

Filtering under the weak Hörmander condition

3.1 Introduction

In this chapter we study the filtering problem for the partially observable kinetic model we introduced in the previous chapter. Having an existence and regularity theory for degenerate SPDEs at hand, we can pursue the "direct" approaches proposed by Krylov and Zatezalo [42] and Veretennikov [70] to derive both the *forward and backward filtering equations*, avoiding the use of general results from filtering theory. In particular, as in [70] we derive the backward filtering equation "by hand", without resorting to prior knowledge of the SPDE, in a more direct way compared to the classical approach in [54], [29], [44] or [62].

To be more specific, we consider the following general setup: we assume that the position X_t and the velocity V_t of a particle are scalar stochastic processes only partially observable through some observation process Y_t. The joint dynamics of X, V and Y is given by the system of SDEs

$$\begin{cases} dX_t = V_t dt, \\ dV_t = b(t, X_t, V_t, Y_t)dt + \sigma^i(t, X_t, V_t, Y_t)dW_t^i, \\ dY_t = h(t, X_t, V_t, Y_t)dt + {}_0\sigma^i(t, Y_t)dW_t^i, \end{cases} \tag{3.1.1}$$

where, as usual, $W_t = (W_t^1, \cdots, W_t^{d_1})$ denotes a d_1-dimensional Brownian motion, with $d_1 \geqslant 2$, defined on a complete probability space $(\Omega, \mathcal{F}, P)$ with a filtration $(\mathcal{F}_t)_{t\in[0,T]}$ satisfying the usual assumptions. Hereafter, for simplicity we set $Z_t = (X_t, V_t)$ and denote by $z = (x, v)$ and $\zeta = (\xi, \nu)$ the points in $\mathbb{R}^2$.

Let $\mathcal{F}_{t,T}^Y = \sigma(Y_s, t \leqslant s \leqslant T)$ define the filtration of observations and let φ be a bounded and continuous function, $\varphi \in bC(\mathbb{R}^2)$. The filtering problem consists in finding the best $\mathcal{F}^Y_{t,T}$-measurable least-square estimate of $\varphi(Z_T)$, that is the conditional expectation $E\left[\varphi(Z_T) \mid \mathcal{F}^Y_{t,T}\right]$.

Consider the case when $h \equiv {}_0\sigma \equiv 0$, that is no observation is available on the solution $Z^{t,z}$ starting from z at time t. Then, it is well known that, under suitable regularity and non-degeneracy assumptions on σ, we have

$$E\left[\varphi(Z^{t,z}_T)\right] = \int_{\mathbb{R}^2} \Gamma(t,z;T,\zeta)\varphi(\zeta)d\zeta, \tag{3.1.2}$$

where the density $\Gamma = \Gamma(t,z;T,\zeta)$ is the fundamental solution of the backward Kolmogorov operator

$$\mathcal{K} = \frac{|\sigma|^2}{2}\partial_{vv} + b\partial_v + v\partial_x + \partial_t \tag{3.1.3}$$

with respect to the variables (t,x,v) and of its adjoint, the Fokker-Plank operator $\mathcal{K}^*$, w.r.t the forward variables (T,ξ,ν). When Y is not trivial, we prove a representation formula for $E\left[\varphi(Z_T) \mid \mathcal{F}^Y_{t,T}\right]$ that is analogous to (3.1.2) in the sense that it is written in terms of the fundamental solution of a backward and a forward SPDE, whose existence is guaranteed by Theorems 2.1.6 and 2.7.4.

In Section 2.2 we already derived the forward equation in the particular case of a kinetic system with constant coefficients, by adapting the direct approach by Krylov and Zatezalo [42]. Such approach mimicks the derivation of the standard Kolmogorov operator (3.1.3): roughly speaking, *assuming that the filtering SPDE is known in advance*, one takes a solution u_t (whose existence is guaranteed by Theorem 2.1.6), applies the Itô formula to $u_t(Z_t)$ and finally takes expectations. This is the approach we follow again in Section 3.2 to prove the existence of the *forward filtering density* and the representation of the conditional expectation $E\left[\varphi(Z_T) \mid \mathcal{F}^Y_{t,T}\right]$ in terms of it.

On the other hand, the direct approach by Veretennikov [69], [70], allows to derive the backward filtering SPDE "by hand", without knowing the equation in advance: the main tools are the backward Itô calculus and the remarkable *backward diffusion SPDE* of Theorem B.1. We follow this approach in Section 3.3 to derive the backward filtering SPDE and the corresponding *filtering density*. Note however that in Section 3.3 we only provide an informal, yet quite detailed, derivation: a full proof would require a generalization of the results of Section B to degenerate diffusions. This is certainly possible but would require some additional effort and is postponed to future research.

Throughout this section we assume the following non-degeneracy condition: there exists a positive constant m such that

$$|{}_0\sigma(t,y)|^2 \geqslant m, \qquad \left\langle \left(I - \frac{{}_0\sigma(t,y){}_0\sigma^*(t,y)}{|{}_0\sigma(t,y)|^2} \right) \sigma(t,z,y), \sigma(t,z,y) \right\rangle \geqslant m, \qquad t \in [0,T],\ z \in \mathbb{R}^2,\ y \in \mathbb{R}. \tag{3.1.4}$$

Equivalently, $|Q(t,y)\sigma(t,z,y)|^2 \geqslant m$, where Q is the orthogonal projector on $\mathrm{Ker}_0\sigma$.

Under condition (3.1.4), up to a straightforward transformation (see [62], Section 6.1), system (3.1.1) can be written in the canonical form

$$\begin{cases} dX_t = V_t dt, \\ dV_t = b(t,X_t,V_t,Y_t)dt + \sigma^i(t,X_t,V_t,Y_t)dW_t^i, \\ dY_t = h(t,Z_t,Y_t)dt + {}_0\sigma(t,Y_t)dW_t^1, \end{cases} \tag{3.1.5}$$

where $W = (W^1, \cdots, W^{d_1})$ is a d_1-dimensional Brownian motion. Setting ${}_1\sigma := \left(\sigma^2, \ldots, \sigma^{d_1}\right)$ so that $\sigma = (\sigma^1, {}_1\sigma)$, assumption (3.1.4) becomes

Assumption 3.1.1 (Coercivity)**.** *There exists a positive constant m such that*

$$_0\sigma(t,y)^2 \geqslant m, \qquad |_1\sigma(t,z,y)|^2 \geqslant m, \qquad t \in [0,T],\ z \in \mathbb{R}^2,\ y \in \mathbb{R}.$$

Moreover, system (3.1.5) can be written more conveniently as

$$\begin{cases} dZ_t = BZ_t dt + \mathbf{e}_2 \left(b(t,Z_t,Y_t)dt + \sigma^i(t,Z_t,Y_t)dW_t^i \right), \\ dY_t = h(t,Z_t,Y_t)dt + {}_0\sigma(t,Y_t)dW_t^1, \end{cases} \tag{3.1.6}$$

with B and $\mathbf{e}_2$ as in (2.2.3).

3.2 Forward filtering SPDE

We consider the solution $(Z_s^{t,z}, Y_s)_{s\in[t,T]}$ of system (3.1.6) with initial condition $Z_t^{t,z} = z \in \mathbb{R}^2$; we do not impose any initial condition on the Y-component. We introduce the stochastic processes

$$\sigma_s(\zeta) := \sigma(s,\zeta,Y_s), \qquad {}_0\sigma_s := {}_0\sigma(s,Y_s), \qquad b_s(\zeta) := b_s(\zeta,Y_s), \qquad \widetilde{h}_s(\zeta) := \frac{h(s,\zeta,Y_s)}{{}_0\sigma(s,Y_s)},$$

The *forward filtering SPDE* for system (3.1.6) reads as follows

$$d_{\mathbf{B}} v_s(\xi,\nu) = \mathcal{A}_s^* v_s(\xi,\nu) ds + \mathcal{G}_s^* v_s(\xi,\nu) \frac{dY_s}{{}_0\sigma_s}, \qquad \mathbf{B} = \partial_s + \nu\partial_\xi, \tag{3.2.1}$$

where $\mathcal{A}^*$ and $\mathcal{G}^*$ are the adjoints of the differential operators (with random coefficients)

$$\mathcal{A}_s := \frac{|\sigma_s(\xi,\nu)|^2}{2}\partial_{\nu\nu} + b_s(\xi,\nu)\partial_\nu, \qquad \mathcal{G}_s := \sigma_s(\xi,\nu)\partial_\nu + \widetilde{h}_s(\xi,\nu),$$

respectively.

In order to apply to (3.2.1) the general results of Sections 2.1.1 and 2.7, in particular Theorem 2.1.6 and Corollary 2.7.5, we assume the following conditions. We recall notation (1.1.6) and that $\sigma = (\sigma^1, \hat{\sigma}) \equiv (\sigma^1, \sigma^2, \ldots, \sigma^{d_1})$.

Assumption 3.2.1 (Regularity). *The coefficients of* (3.1.6) *are such that* $\sigma^1 \in bC^{3+\alpha}_{0,T}(\mathbb{R}^3)$, ${}_1\sigma \in bC^{2+\alpha}_{0,T}(\mathbb{R}^3)$, ${}_0\sigma \in bC^{\alpha}_{0,T}(\mathbb{R})$, $b \in bC^{1}_{0,T}(\mathbb{R}^3)$, $h \in bC^{2}_{0,T}(\mathbb{R}^3)$.

Assumption 3.2.2 (Flattening at infinity). *There exist two positive constants* ε, M *such that*

$$\sup_{\substack{t\in[0,T]\\ y\in\mathbb{R}}} \left(\{\sigma^1(t,\cdot,y)\}_{\varepsilon,\beta} + \{\sigma^1(t,\cdot,y)\}_{1/2+\varepsilon,\beta'} + \{h(t,\cdot,y)\}_{1/2,\beta}\right) \leqslant M$$

for $|\beta| = 1$ *and* $|\beta'| = 2, 3$.

Remark 3.2.3. *With regard to the existence of solutions to* (3.2.1), *let us introduce the process*

$$\widetilde{W}_s := \int_t^s {}_0\sigma_\tau^{-1} dY_\tau = W^1_s - W^1_t + \int_t^s \widetilde{h}_\tau(Z^{t,z}_\tau)d\tau, \qquad s \in [t,T].$$

By Girsanov's theorem, $(\widetilde{W}_s)_{s\in[t,T]}$ *is a Brownian motion w.r.t the measure* Q *defined by* $dQ = (\varrho^{t,z}_T)^{-1}dP$ *where*

$$d\varrho^{t,z}_s = \widetilde{h}_s(Z^{t,z}_s)^2 \varrho^{t,z}_s dt + \widetilde{h}_s(Z^{t,z}_s)\varrho^{t,z}_s dW^1_s, \qquad \varrho^{t,z}_t = 1. \tag{3.2.2}$$

Moreover, $(\widetilde{W}_s)_{s\in[t,T]}$ *is adapted to* $(\mathcal{F}^Y_{t,s})_{s\in[t,T]}$. *Then, equation* (3.2.1) *can be written in the equivalent form*

$$d_{\mathbf{B}} v_s(\zeta) = \mathcal{A}^*_s v_s(\zeta)ds + \mathcal{G}^*_s v_s(\zeta)d\widetilde{W}_s \tag{3.2.3}$$

under Q. *Under Assumptions 3.1.1, 3.2.1 and 3.2.2, by Theorem 2.1.6 a fundamental solution* $\mathbf{\Gamma} = \mathbf{\Gamma}(t,z;s,\zeta)$ *for* (3.2.3) *exists, satisfies estimates* (2.1.5) - (2.1.8) *and* $s \mapsto \mathbf{\Gamma}(t,z;s,\zeta)$ *is adapted to* $(\mathcal{F}^Y_{t,s})_{s\in[t,T]}$. *We say that the stochastic process*

$$\hat{\mathbf{\Gamma}}(t,z;s,\zeta) = \frac{\mathbf{\Gamma}(t,z;s,\zeta)}{\int_{\mathbb{R}^2}\mathbf{\Gamma}(t,z;s,\zeta_1)d\zeta_1}, \qquad 0 \leqslant t < s \leqslant T,\ z,\zeta \in \mathbb{R}^2,$$

is the forward filtering density for system (3.1.6). *This definition is motivated by the following*

Theorem 3.2.4. *Let* $(Z^{t,z}_s, Y_s)_{s\in[t,T]}$ *denote the solution of system* (3.1.6) *with initial condition* $Z^{t,z}_t = z$. *Under Assumptions 3.1.1, 3.2.1 and 3.2.2, for any* $\varphi \in bC(\mathbb{R}^2)$ *we have*

$$E\left[\varphi(Z^{t,z}_T) \mid \mathcal{F}^Y_{t,T}\right] = \int_{\mathbb{R}^2} \hat{\mathbf{\Gamma}}(t,z;T,\zeta)\varphi(\zeta)d\zeta, \qquad (t,z) \in [0,T]\times\mathbb{R}^2. \tag{3.2.4}$$

Proof. By Remark 3.2.3, $\int_{\mathbb{R}^2}\hat{\mathbf{\Gamma}}(t,z;T,\zeta)\varphi(\zeta)d\zeta \in m\mathcal{F}^Y_{t,T}$. We prove that, for any bounded and $\mathcal{F}^Y_{t,T}$-measurable random variable G, we have

$$E\left[G\varphi(Z^{t,z}_T)\right] = E\left[G(\varrho^{t,z}_T)^{-1}\int_{\mathbb{R}^2}\mathbf{\Gamma}(t,z;T,\zeta)\varphi(\zeta)d\zeta\right], \tag{3.2.5}$$

with $\varrho^{t,z}$ as in (3.2.2). From (3.2.5) with $\varphi \equiv 1$ it will follow that

$$E\left[(\varrho_T^{t,z})^{-1} \mid \mathcal{F}_{t,T}^Y\right] = \left(\int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;T,\zeta)d\zeta\right)^{-1}$$

and therefore also (3.2.4) will follow from (3.2.5).

By a standard approximation argument, it is enough to take φ in the class of test functions and G of the form $G = e^{-\int_t^T c_s ds}$ where $c_s = c(s, Y_s)$ with $c = c(s,y)$ being a smooth, bounded and non-negative function on $[t,T] \times \mathbb{R}$. Thus, we are left with the proof of the following identity:

$$E\left[e^{-\int_t^T c_s ds}\varphi(Z_T^{t,z})\right] = E\left[e^{-\int_t^T c_s ds}(\varrho_T^{t,z})^{-1}\int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;T,\zeta)\varphi(\zeta)d\zeta\right]. \tag{3.2.6}$$

To this end, we consider the deterministic backward Cauchy problem

$$f\left(s, e^{(T-s)B}\zeta, y\right) = \varphi(\zeta) + \int_s^T (\widetilde{\mathcal{A}}_\tau - c(\tau,y))f\left(\tau, e^{(T-\tau)B}\zeta, y\right)d\tau, \qquad (s,\zeta,y) \in [t,T] \times \mathbb{R}^2 \times \mathbb{R}, \tag{3.2.7}$$

where

$$\widetilde{\mathcal{A}}_\tau := \frac{1}{2}\left(|\sigma(\tau,\zeta,y)|^2\partial_{\nu\nu} + 2{}_0\sigma(\tau,y)\sigma(\tau,\zeta,y)\partial_{\nu y} + {}_0\sigma^2(\tau,y)\partial_{yy}\right) + b(\tau,\zeta,y)\partial_\nu + h(\tau,\zeta,y)\partial_y.$$

In differential form, (3.2.7) reads as

$$\begin{cases} -d_{\mathbf{B}}f(s,\zeta,y) = \left(\widetilde{\mathcal{A}}_s f(s,\zeta,y) - c(s,y)f(s,\zeta,y)\right)ds, \\ f(T,\zeta,y) = \varphi(\zeta). \end{cases}$$

Corollary 2.7.5 ensures existence and estimates of a strong solution f to (3.2.7).

Next, we consider the process

$$M_s^{t,z} := e^{-\int_t^s c_\tau d\tau}(\varrho_s^{t,z})^{-1}\int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;s,\zeta)f(s,\zeta,Y_s)d\zeta, \qquad s \in [t,T].$$

By definition, we have

$$M_T^{t,z} = e^{-\int_t^T c_s ds}(\varrho_T^{t,z})^{-1}\int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;T,\zeta)\varphi(\zeta)d\zeta.$$

On the other hand, by the Feynman-Kac theorem we have

$$M_t^{t,z} = f(t,z,Y_t) = E\left[e^{-\int_t^T c_s ds}\varphi(Z_T^{t,z}) \mid Y_t\right].$$

Hence to prove (3.2.6) it suffices to check that $M = (M_s^{t,z})_{s\in[t,T]}$ is a martingale: to this end, we prove the representation

$$\begin{aligned} M_T^{t,z} &= M_t^{t,z} + \int_t^T G_s^{t,z} dW_s^1, \\ G_s^{t,z} &= e^{-\int_t^s c_\tau d\tau}(\varrho_s^{t,z})^{-1}\int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;s,\zeta)\left(\mathcal{G}_s + {}_0\sigma_s\partial_y\right)f(s,\zeta,Y_s)d\zeta, \qquad s \in [t,T], \end{aligned} \tag{3.2.8}$$

and conclude by showing that

$$E\left[\int_t^T |G_s^{t,z}|^2 ds\right] < \infty. \tag{3.2.9}$$

We first compute the stochastic differential $d_{\mathbf{B}} f(s,\zeta,Y_s)$: by Corollary 2.7.5 we have

$$\begin{aligned} d_{\mathbf{B}} f(s,\zeta,Y_s) &= \left(-\widetilde{\mathcal{A}}_s + \frac{1}{2}{}_0\sigma_s^2\partial_{yy} + c_s\right) f(s,\zeta,Y_s)ds + \partial_y f(s,\zeta,Y_s)dY_s \\ &= \left(-\widetilde{\mathcal{A}}_s + \frac{1}{2}{}_0\sigma_s^2\partial_{yy} + h_s(Z_s)\partial_y + c_s\right) f(s,\zeta,Y_s)ds + {}_0\sigma_s\partial_y f(s,\zeta,Y_s)dW_s^1. \end{aligned}$$

On the other hand, we have

$$\begin{aligned} d_{\mathbf{B}}\mathbf{\Gamma}(t,z;s,\zeta) &= \mathcal{A}_s^*\mathbf{\Gamma}(t,z;s,\zeta)ds + \mathcal{G}_s^*\mathbf{\Gamma}(t,z;s,\zeta)\frac{dY_s}{{}_0\sigma_s} \\ &= \left(\mathcal{A}_s^* + \widetilde{h}_s(Z_s)\mathcal{G}_s^*\right)\mathbf{\Gamma}(t,z;s,\zeta)ds + \mathcal{G}_s^*\mathbf{\Gamma}(t,z;s,\zeta)dW_s^1. \end{aligned}$$

Then, by Itô formula we have

$$d_{\mathbf{B}}\left(f(s,\zeta,Y_s)\mathbf{\Gamma}(t,z;s,\zeta)\right) = I_1(t,z;s,\zeta)ds + I_2(t,z;s,\zeta)dW_s^1$$

where

$$\begin{aligned} I_1(t,z;s,\zeta) &= f(s,\zeta,Y_s)\left(\mathcal{A}_s^* + \widetilde{h}_s(Z_s)\mathcal{G}_s^*\right)\mathbf{\Gamma}(t,z;s,\zeta) \\ &\quad + \mathbf{\Gamma}(t,z;s,\zeta)\left(-\widetilde{\mathcal{A}}_s + \frac{1}{2}{}_0\sigma_s^2\partial_{yy} + h_s(Z_s)\partial_y + c_s\right) f(s,\zeta,Y_s) \\ &\quad + {}_0\sigma_s\mathcal{G}_s^*\mathbf{\Gamma}(t,z;s,\zeta)\partial_y f(s,\zeta,Y_s), \\ I_2(t,z;s,\zeta) &= f(s,\zeta,Y_s)\mathcal{G}_s^*\mathbf{\Gamma}(t,z;s,\zeta) + {}_0\sigma_s\mathbf{\Gamma}(t,z;s,\zeta)\partial_y f(s,\zeta,Y_s). \end{aligned}$$

This means that for any $s \in (t,T]$ we have

$$\begin{aligned} f(T,e^{(T-s)B}\zeta,Y_T)\mathbf{\Gamma}(t,z;T,e^{(T-s)B}\zeta) &= f(s,\zeta,Y_s)\mathbf{\Gamma}(t,z;s,\zeta) \\ &\quad + \int_s^T I_1(t,z;\tau,e^{(\tau-s)B}\zeta)d\tau + \int_s^T I_2(t,z;\tau,e^{(\tau-s)B}\zeta)dW_\tau^1. \end{aligned}$$

Next, we integrate over $\mathbb{R}^2$ the previous identity and apply the standard and stochastic Fubini's theorems (see, for instance, [62], Chapter 1) to get

$$\begin{aligned} \int_{\mathbb{R}^2} f(T,e^{(T-s)B}\zeta,Y_T)\mathbf{\Gamma}(t,z;T,e^{(T-s)B}\zeta)d\zeta &= \int_{\mathbb{R}^2} f(s,\zeta,Y_s)\mathbf{\Gamma}(t,z;s,\zeta)d\zeta \\ &\quad + \int_s^T\int_{\mathbb{R}^2} I_1(t,z;\tau,e^{(\tau-s)B}\zeta)d\zeta d\tau \\ &\quad + \int_s^T\int_{\mathbb{R}^2} I_2(t,z;\tau,e^{(\tau-s)B}\zeta)d\zeta dW_\tau^1. \end{aligned}$$

By the estimates of the fundamental solution in Theorem 2.1.6, the estimates of the solution f and its derivatives in Corollary 2.7.5, the boundedness of the coefficients and the non-degeneracy condition of Assumption 3.1.1, we have

$$\begin{aligned}\int_s^T \int_{\mathbb{R}^2} |I_1(t,z;\tau,\zeta)| d\zeta d\tau &\leqslant \int_s^T \frac{C}{(T-\tau)^{\frac{1}{2}}(s-t)} \int_{\mathbb{R}^2} \Gamma^{\text{heat}}\left(\mu \mathcal{D}_{\tau-t}, g_{t,\tau}^{\text{IW},-1}(\zeta) - \gamma_{t,\tau}(z)\right) d\zeta d\tau \\ &\leqslant C'(T-s)^{\frac{1}{2}}(s-t)^{-1},\end{aligned}$$

and, analogously

$$\begin{aligned}\int_s^T \left(\int_{\mathbb{R}^2} |I_2(t,z;\tau,\zeta)| d\zeta\right)^2 d\tau &\leqslant \int_s^T \left(\frac{C}{(s-t)^{\frac{1}{2}}} \int_{\mathbb{R}^2} \Gamma^{\text{heat}}\left(\mu \mathcal{D}_{\tau-t}, g_{t,\tau}^{\text{IW},-1}(\zeta) - \gamma_{t,\tau}(z)\right) d\zeta\right)^2 d\tau \\ &\leqslant C'(s-t)^{-1},\end{aligned}$$

for some positive constants C, C'. This justifies the use of Fubini's theorems.

Now, since the Jacobian of the transformation $\zeta' = e^{sB}\zeta$ equals one for any s, the previous equality yields

$$\begin{aligned}\int_{\mathbb{R}^2} f(T,\zeta,Y_T)\mathbf{\Gamma}(t,z;T,\zeta)d\zeta = &\int_{\mathbb{R}^2} f(s,\zeta,Y_s)\mathbf{\Gamma}(t,z;s,\zeta)d\zeta \\ &+ \int_s^T \int_{\mathbb{R}^2} I_1(t,z;\tau,\zeta)\zeta d\tau + \int_s^T \int_{\mathbb{R}^2} I_2(t,z;\tau,\zeta) d\zeta dW_\tau^1.\end{aligned}$$

Integrating by parts and using the identity

$$\begin{aligned}&\int_{\mathbb{R}^2} \left(f(s,\zeta,Y_s)\mathcal{A}_s^*\mathbf{\Gamma}(t,z;s,\zeta) + \mathbf{\Gamma}(t,z;s,\zeta)\frac{1}{2}{}_0\sigma_s^2\partial_{yy}f(s,\zeta,Y_s) + {}_0\sigma_s\mathcal{G}_s^*\mathbf{\Gamma}(t,z;s,\zeta)\partial_y f(s,\zeta,Y_s)\right) d\zeta \\ &= \int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;s,\zeta)\left(\mathcal{A}_s + \frac{1}{2}{}_0\sigma_s^2\partial_{yy} + {}_0\sigma_s\sigma_s\partial_{y\nu} + h_s(\zeta,Y_s)\partial_y\right) f(s,\zeta,Y_s) d\zeta \\ &= \int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;s,\zeta)\widetilde{\mathcal{A}}_s f(s,\zeta,Y_s) d\zeta,\end{aligned}$$

we get

$$\begin{aligned}\int_{\mathbb{R}^2} f(T,\zeta,Y_T)\mathbf{\Gamma}(t,z;T,\zeta)d\zeta = &\int_{\mathbb{R}^2} f(s,\zeta,Y_s)\mathbf{\Gamma}(t,z;s,\zeta)d\zeta \\ &+ \int_s^T \int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;\tau,\zeta)\left(\widetilde{h}_\tau(Z_\tau)\mathcal{G}_\tau + h_\tau(Z_\tau)\partial_y + c_\tau\right) f(\tau,\zeta,Y_\tau) d\zeta d\tau \\ &+ \int_s^T \int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;\tau,\zeta)\left(\mathcal{G}_\tau + {}_0\sigma_\tau\partial_y\right) f(\tau,\zeta,Y_\tau) d\zeta dW_\tau^1.\end{aligned}$$

Eventually, we multiply the expression above by $e^{-\int_t^s c_\tau d\tau}(\varrho_s^{t,z})^{-1}$: since

$$d\left(e^{-\int_t^s c_\tau d\tau}(\varrho_s^{t,z})^{-1}\right) = e^{-\int_t^s c_\tau d\tau}(\varrho_s^{t,z})^{-1}\left(-c_s ds - \widetilde{h}_s(Z_s)dW_s^1\right),$$

$$d\langle e^{-\int_t^\cdot c_\tau d\tau}(\varrho_\cdot^{t,z})^{-1}, \int_{\mathbb{R}^2} f(\cdot,\zeta,Y_\cdot)\mathbf{\Gamma}(t,z;\cdot,\zeta)d\zeta\rangle_s = -\int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;s,\zeta)\left(\widetilde{h}_s(Z_s)\mathcal{G}_s + h_s(Z_s)\partial_y\right) f(s,\zeta,Y_s) d\zeta ds,$$

by Itô formula, for $s \in (t, T]$ we have

$$\begin{aligned}
M_T^{t,z} &= e^{-\int_t^T c_\tau d\tau} (\varrho_T^{t,z})^{-1} \int_{\mathbb{R}^2} f(T,\zeta,Y_T)\mathbf{\Gamma}(t,z;T,\zeta)d\zeta \\
&= M_s^{t,z} + \int_s^T e^{-\int_t^\tau c_\varrho d\varrho} (\varrho_\tau^{t,z})^{-1} \int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;\tau,\zeta) \left(\mathcal{G}_\tau + {}_0\sigma_\tau \partial_y\right) f(\tau,\zeta,Y_\tau) d\zeta dW_\tau^1 \\
&= M_s^{t,z} + \int_s^T G_\tau^{t,z} dW_\tau^1.
\end{aligned}$$

with $G_\tau^{t,z}$ as in (3.2.8). Now, again by the estimates of the fundamental solution (cf. Theorem 2.1.6), the estimates of the solution f and its derivatives (cf. Corollary 2.7.5), the boundedness of the coefficients and the non-degeneracy condition of Assumption 3.1.1, we deduce the estimate

$$|G_\tau^{t,z}| \leqslant C(\varrho_\tau^{t,z})^{-1} \int_{\mathbb{R}^2} \Gamma^{\text{heat}}(\mu \mathcal{D}_{\tau-t}, g_{\tau,t}^{\text{IW},-1}(\zeta) - \gamma_{t,\tau}(z)) d\zeta \leqslant C'$$

for some positive constants C, C'. This implies (3.2.9) and concludes the proof. □

3.3 Backward filtering SPDE

As in the previous section, in order to apply the general results of Sections 2.1.1 and 2.7 to the filtering SPDE for system (3.1.6), we impose the following conditions:

Assumption 3.3.1 (Regularity)**.** *The coefficients of* (3.1.6) *are such that* $\sigma^1 \in bC_{0,T}^{3+\alpha}(\mathbb{R}^3)$, ${}_1\sigma \in bC_{0,T}^{\alpha}(\mathbb{R}^3)$, ${}_0\sigma \in bC_{0,T}^{3+\alpha}(\mathbb{R})$, $b \in bC_{0,T}^{0}(\mathbb{R}^3)$, $h \in bC_{0,T}^{2}(\mathbb{R}^3)$.

Assumption 3.3.2 (Flattening at infinity)**.** *There exist two positive constants* ε, M *such that*

$$\sup_{t\in[0,T]} \left(\{\sigma^1(t,\cdot,\cdot)\}_{\varepsilon,\beta} + \{\sigma^1(t,\cdot,\cdot)\}_{1/2+\varepsilon,\beta'} + \{{}_0\sigma(t,\cdot)\}_{\varepsilon,\beta} + \{{}_0\sigma(t,\cdot)\}_{1/2+\varepsilon,\beta'} + \{h(t,\cdot,\cdot)\}_{1/2,\beta}\right) \leqslant M$$

for $|\beta| = 1$ *and* $|\beta'| = 2, 3$.

The backward filtering SPDE for system (3.1.6) reads

$$-d_{\mathbf{B}} u_t(z,y) = \widetilde{\mathcal{A}}_t u_t(z,y)dt + \widetilde{\mathcal{G}}_t u_t(z,y) \star \frac{dY_t}{{}_0\sigma(t,y)}, \qquad \mathbf{B} := \partial_t + v\partial_x, \tag{3.3.1}$$

where $z = (x, v)$ and

$$\begin{aligned}
\widetilde{\mathcal{A}}_t &:= \frac{1}{2}\left(|\sigma(t,z,y)|^2 \partial_{vv} + 2{}_0\sigma(t,y)\sigma(t,z,y)\partial_{vy} + {}_0\sigma^2(t,y)\partial_{yy}\right) \\
&\quad + b(t,z,y)\partial_v + h(t,z,y)\partial_y,
\end{aligned} \tag{3.3.2}$$

$$\widetilde{\mathcal{G}}_t := \sigma(t,z,y)\partial_v + {}_0\sigma(t,y)\partial_y + \widetilde{h}(t,z,y), \qquad \widetilde{h}(t,z,y) := \frac{h(t,z,y)}{{}_0\sigma(t,y)}.$$

Before presenting the main result of this section, we comment on the existence of solutions to (3.3.1). Let $(Z_s^{t,z,y}, Y_s^{t,z,y}, \varrho_s^{t,z,y,\eta})_{s\in[t,T]}$ be the solution, starting at time t from (z, y, η), of the system of SDEs

$$\begin{cases} dZ_t = BZ_t dt + \mathbf{e}_2(b(t, Z_t, Y_t)dt + \sigma^i(t, Z_t, Y_t)dW_t^i, \\ dY_t = h(t, Z_t, Y_t)dt + {}_0\sigma(t, Y_t)dW_t^1, \\ d\varrho_t = \widetilde{h}(t, Z_t, Y_t)^2 \varrho_t dt + \widetilde{h}(t, Z_t, Y_t)\varrho_t dW_t^1. \end{cases} \tag{3.3.3}$$

By Girsanov's theorem, the process

$$\begin{aligned} \widetilde{W}_s^{t,z,y} &:= \int_t^s {}_0\sigma^{-1}(\tau, Y_\tau^{t,z,y})dY_\tau^{t,z,y} \\ &= W_s^1 - W_t^1 + \int_t^s \widetilde{h}(\tau, Z_\tau^{t,z,y}, Y_\tau^{t,z,y})d\tau, \qquad s \in [t, T], \end{aligned}$$

is a Brownian motion w.r.t the measure $Q^{t,z,y}$ defined by $dQ^{t,z,y} = (\varrho_T^{t,z,y,1})^{-1}dP$. Notice also that $(\widetilde{W}_s^{t,z,y})_{s\in[t,T]}$ is adapted to $(\mathcal{F}_{t,s}^Y)_{s\in[t,T]}$ where $\mathcal{F}_{t,s}^Y = \sigma(Y_\tau^{t,z,y},\ t \leqslant \tau \leqslant s)$. Then equation (3.3.1) can be written in the equivalent form

$$-d_{\mathbf{B}}u_s(z, y) = \widetilde{\mathcal{A}}_s u_s(z, y)ds + \widetilde{\mathcal{G}}_s u_s(z, y) \star d\widetilde{W}_s^t \tag{3.3.4}$$

or, more explicitly,

$$u_t\left(\gamma_{T-t}^{\mathbf{B}}(z, y)\right) = u_T(z, y) + \int_t^T \widetilde{\mathcal{A}}_s u_s(\gamma_{T-s}^{\mathbf{B}}(z, y))ds + \int_t^T \widetilde{\mathcal{G}}_s u_s(\gamma_{T-s}^{\mathbf{B}}(z, y)) \star d\widetilde{W}_s^t, \qquad t \in [0, T], \tag{3.3.5}$$

where $\gamma_s^{\mathbf{B}}(z, y) = \gamma_s^{\mathbf{B}}(x, v, y) = (x + sv, v, y)$. In (3.3.4) and (3.3.5), we simply write $\widetilde{W}_s^t$ instead of $\widetilde{W}_s^{t,z,y}$ because the starting point of the Brownian motion is irrelevant in the stochastic integration. Theorem 2.7.4 guarantees that a fundamental solution $\overleftarrow{\mathbf{\Gamma}} = \overleftarrow{\mathbf{\Gamma}}(t, z, y; s, \zeta, \eta)$ for (3.3.4) exists and satisfies estimates (2.7.3), (2.7.4) and (2.7.5). Moreover, $t \mapsto \overleftarrow{\mathbf{\Gamma}}(t, z, y; T, \zeta, \eta)$ is adapted to $(\mathcal{F}_{t,T}^Y)_{t\in[0,T]}$. The main result of this section is the following

Theorem 3.3.3. *Let $(Z_T^{t,z,y}, Y_T^{t,z,y})$ denote the solution of system (3.1.6) starting from (z, y) at time $t \in [0, T)$ and $\varphi \in bC(\mathbb{R}^3)$. Under Assumptions 3.1.1, 3.3.1 and 3.3.2, we have*

$$E\left[\varphi(Z_T^{t,z,y}, Y_T^{t,z,y}) \mid \mathcal{F}_{t,T}^Y\right] = \frac{u_t^{(\varphi)}(z, y)}{u_t^{(1)}(z, y)}, \qquad (t, z, y) \in [0, T] \times \mathbb{R}^2 \times \mathbb{R}, \tag{3.3.6}$$

where $u_t^{(\varphi)}$ denotes the solution to (3.3.1) with final datum $u_T^{(\varphi)} = \varphi$.

Definition 3.3.4 (Backward filtering density). *The normalized process*

$$\bar{\mathbf{\Gamma}}(t, z, y; T, \zeta, \eta) = \frac{\overleftarrow{\mathbf{\Gamma}}(t, z, y; T, \zeta, \eta)}{\int_{\mathbb{R}^3} \overleftarrow{\mathbf{\Gamma}}(t, z, y; T, \zeta_1, \eta_1)d\zeta_1 d\eta_1},$$

for $0 \leqslant t < T$ *and* $(z,y),(\zeta,\eta) \in \mathbb{R}^2 \times \mathbb{R}$*, is called the backward filtering density of system* (3.1.6)*. By Theorem 3.3.3, we have*

$$E\left[\varphi(Z_T^{t,z,y}, Y_T^{t,z,y}) \mid \mathcal{F}_{t,T}^Y\right] = \int_{\mathbb{R}^3} \bar{\mathbf{\Gamma}}(t,z,y;T,\zeta,\eta)\varphi(\zeta,\eta)d\zeta d\eta, \qquad (t,z,y) \in [0,T]\times\mathbb{R}^2\times\mathbb{R}, \quad (3.3.7)$$

for any $\varphi \in bC(\mathbb{R}^3)$.

Remark 3.3.5. *Notice that formulas* (3.3.6) *and* (3.3.7) *represent the conditional expectation in terms of solutions to the Cauchy problem for the backward filtering SPDE. This is not the case for formula* (3.2.4) *in the forward case.*

In the rest of the section we sketch the proof of Theorem 3.3.3. First, notice that under $Q^{t,z,y}$ we have

$$\varrho_s^{t,z,y,\eta} = \eta \exp\left(\int_t^s \widetilde{h}(\tau, Z_\tau^{t,z,y}, Y_\tau^{t,z,y})d\widetilde{W}_\tau^t - \frac{1}{2}\int_t^s \widetilde{h}(\tau, Z_\tau^{t,z,y}, Y_\tau^{t,z,y})^2 d\tau\right), \qquad s \in [t,T],$$

and system (3.3.3) reads

$$\begin{cases} dZ_s^{t,z,y} = \widetilde{B}(s, Z_s^{t,z,y}, Y_s^{t,z,y})ds + \mathbf{e}_2\left({}_1\sigma^i(s, Z_s^{t,z,y}, Y_s^{t,z,y})dW_t^i + \sigma(s, Z_s^{t,z,y}, Y_s^{t,z,y})d\widetilde{W}_s^t\right), \\ dY_s^{t,z,y} = {}_0\sigma(s, Y_s^{t,z,y})d\widetilde{W}_s^t, \\ d\varrho_s^{t,z,y,\eta} = \widetilde{h}(s, Z_s^{t,z,y}, Y_s^{t,z,y})\varrho_s^{t,z,y,\eta}d\widetilde{W}_s^t, \end{cases} \quad (3.3.8)$$

where $\widetilde{B}(s,z,y) = Bz + \mathbf{e}_2(b(s,z,y) - \widetilde{h}(s,z,y)\sigma(s,z,y))$. Recalling the notation $z = (x,v) \in \mathbb{R}^2$ and omitting the arguments of the coefficients for brevity, the correspondent characteristic operator is

$$\mathcal{L} = \frac{1}{2}\left(|\sigma|^2\partial_{vv} + {}_0\sigma^2\partial_{yy} + \eta^2\widetilde{h}^2\partial_{\eta\eta} + 2\sigma_0\sigma\partial_{vy} + 2\eta\sigma\widetilde{h}\partial_{v\eta} + 2\eta_0\sigma\widetilde{h}\partial_{y\eta}\right) + \langle \widetilde{B}, \nabla_z\rangle.$$

We write the backward diffusion SPDE for system (3.3.8). Assuming that φ is smooth and letting $V_s(z,y) := \varphi(Z_T^{s,z,y}, Y_T^{s,z,y})$, by Corollary B.4 we have

$$\begin{aligned} -d(V_s(z,y)\varrho_T^{s,z,y,\eta}) = {} & \mathcal{L}(V_s(z,y)\varrho_T^{s,z,y,\eta})ds \\ & + \partial_v(V_s(z,y)\varrho_T^{s,z,y,\eta})\left({}_1\sigma^i(s,z,y) \star dW_s^i + \sigma(s,z,y) \star d\widetilde{W}_s^t\right) \\ & + \partial_y(V_s(z,y)\varrho_T^{s,z,y,\eta})_0\sigma(s,z,y) \star d\widetilde{W}_s^t + \partial_\eta(V_s(z,y)\varrho_T^{s,z,y,\eta})\eta\widetilde{h}(s,z,y) \star d\widetilde{W}_s^t \end{aligned}$$

(noting that $\partial_\eta Z_T^{t,z,y} = \partial_\eta Y_T^{t,z,y} = \partial_{\eta\eta}\varrho_T^{t,z,y,\eta} = 0$ and $\eta\partial_\eta \varrho_T^{t,z,y,\eta} = \varrho_T^{t,z,y,\eta}$)

$$
\begin{aligned}
&= \frac{1}{2}\left(|\sigma(t,z,y)|^2\partial_{vv} + {}_0\sigma^2(t,z,y)\partial_{yy} + 2\sigma_0\sigma(t,z,y)\partial_{vy}\right)(V_s(z,y)\varrho_T^{s,z,y,\eta})ds \\
&\quad + \left(\widetilde{h}(s,z,y)(\sigma\partial_v + {}_0\sigma\partial_y) + \langle \widetilde{B}(t,z,y), \nabla_z\rangle\right)(V_s(z,y)\varrho_T^{s,z,y,\eta})ds \\
&\quad + {}_1\sigma^i(s,z,y)\partial_v(V_s(z,y)\varrho_T^{s,z,y,\eta}) \star dW_s^i \\
&\quad + \left(\sigma(s,z,y)\partial_v + {}_0\sigma(s,z,y)\partial_y + \widetilde{h}(s,z,y)\right)(V_s(z,y)\varrho_T^{s,z,y,\eta}) \star d\widetilde{W}_s^t
\end{aligned}
$$

(noting that $\widetilde{h}(s,z,y)(\sigma\partial_v + {}_0\sigma\partial_y) + \langle \widetilde{B}(t,z,y), \nabla_z\rangle = v\partial_x + b(t,z,y)\partial_v + h(t,z,y)\partial_y$)

$$
\begin{aligned}
&= \widetilde{\mathcal{L}}(V_s(z,y)\varrho_T^{s,z,y,\eta})ds + {}_1\sigma^i(s,z,y)\partial_v(V_s(z,y)\varrho_T^{s,z,y,\eta}) \star dW_s^i \\
&\quad + \left(\sigma(s,z,y)\partial_v + {}_0\sigma(s,z,y)\partial_y + \widetilde{h}(s,z,y)\right)(V_s(z,y)\varrho_T^{s,z,y,\eta}) \star d\widetilde{W}_s^t.
\end{aligned}
$$

where $\widetilde{\mathcal{L}} = \widetilde{\mathcal{A}}_t + v\partial_x$, with $\widetilde{\mathcal{A}}_t$ as in (3.3.2), is the infinitesimal generator of (Z_t, Y_t). Therefore we have

$$
\begin{aligned}
\varphi(Z_T^{t,z,y}, Y_T^{t,z,y})\varrho_T^{t,z,y,1} &= \varphi(z,y) + V_t(z,y)\varrho_T^{t,z,y,1} - V_T(z,y)\varrho_T^{T,z,y,1} \\
&= \int_t^T \widetilde{\mathcal{L}}(V_s(z,y)\varrho_T^{s,z,y,\eta})ds + \int_t^T {}_1\sigma^i(s,z,y)\partial_v(V_s(z,y)\varrho_T^{s,z,y,\eta}) \star dW_s^i \\
&\quad + \int_t^T \widetilde{\mathcal{G}}_s(V_s(z,y)\varrho_T^{s,z,y,\eta}) \star d\widetilde{W}_s^t. \qquad (3.3.9)
\end{aligned}
$$

Now we take the conditional expectation in (3.3.9) and exploit the fact that $(W^2, \dots, W^{d_1})$ is independent of $\mathcal{F}_{t,T}^Y$ under $Q^{t,z,y}$ (this follows from the crucial assumption that ${}_0\sigma$ is a function of t, y only): setting

$$
u_t^{(\varphi)}(z,y) = E^{Q^{t,z,y}}\left[V_t(z,y)\varrho_T^{t,z,y,1} \mid \mathcal{F}_{t,T}^Y\right],
$$

and applying the standard and stochastic Fubini's theorems, we directly get the filtering equation

$$
u_t^{(\varphi)}(z,y) = \varphi(z,y) + \int_t^T \widetilde{\mathcal{L}}_s u_s^{(\varphi)}(z,y)ds + \int_t^T \widetilde{\mathcal{G}}_s u_s^{(\varphi)}(z,y) \star \frac{dY_s^{t,z,y}}{{}_0\sigma(s,y)}
$$

which is equivalent to (3.3.1). Analogously,

$$
u_t^{(1)}(z,y) := E^{Q^{t,z,y}}\left[\varrho_T^{t,z,y,1} \mid \mathcal{F}_{t,T}^Y\right]
$$

solves the same SPDE with terminal condition $u_T^{(1)}(z,y) \equiv 1$. To conclude, it suffices recall the Bayes representation for conditional expectations or the Kallianpur-Striebel's formula (cf. [62], Lemma 6.1) according to which we have

$$
E\left[\varphi(Z_T^{t,z,y}, Y_T^{t,z,y}) \mid \mathcal{F}_{t,T}^Y\right] = \frac{E^{Q^{t,z,y}}\left[\varphi(Z_T^{t,z,y}, Y_T^{t,z,y})\varrho_T^{t,z,y,1} \mid \mathcal{F}_{t,T}^Y\right]}{E^{Q^{t,z,y}}\left[\varrho_T^{t,z,y,1} \mid \mathcal{F}_{t,T}^Y\right]}.
$$

Chapter 4

Density and gradient estimates for non degenerate Brownian SDEs with unbounded measurable drift

4.1 Introduction

In this chapter we provide Aronson-like bounds and corresponding pointwise estimates for the derivatives up to order two for the transition probability density of the following d-dimensional, non-degenerate diffusion

$$dX_s = b(s, X_s)ds + \sigma(s, X_s)dW_s, \quad s \geqslant 0, \ X_0 = x, \tag{4.1.1}$$

where $(W_s)_{s\geqslant 0}$ is a standard d-dimensional Brownian motion on the probability space $(\Omega, \mathcal{F}, P)$ with filtration $(\mathcal{F}_s)_{s\geqslant 0}$, satisfying the usual assumptions. The diffusion coefficient σ is assumed to be rough in time, and Hölder continuous in space. The drift b is assumed to be measurable and to have linear growth in space. Importantly, we will always assume throughout the chapter that the diffusion coefficient σ is bounded and separated from 0 (usual uniform ellipticity condition).

The chapter is organized as follows. Our main results are stated in details in Section 2.1.1; Section 4.2 is dedicated to the proof of our main results when the coefficients satisfy our previous assumptions and are also *smooth*. Importantly, we prove that the two-sided heat-kernel bounds do not depend on the smoothness of the coefficients but only on constants appearing in Assumptions 4.1.1 and 4.1.2 below, the fixed final time horizon $T > 0$ and the dimension d. We also establish there bounds for the derivatives through Malliavin calculus techniques which is precisely possible because the coefficients are smooth. Those bounds serve as *a priori* controls to derive in Section

4.3, through a circular type argument based on the *Duhamel-parametrix type* representation of the density, that those bounds actually do not depend on the smoothness of the coefficients. We then deduce the main results passing to the limit in a mollification procedure through convergence in law and compactness arguments. We eventually discuss in Section 4.4 some possible extensions for the estimation of higher order derivatives of the heat kernel when the coefficients have some additional smoothness properties.

4.1.1 Assumptions and main results

We make the following assumptions on the coefficients of (4.1.1).

Assumption 4.1.1 (Non degeneracy). *There exists a positive constant $\lambda_1 \geqslant 1$, such that*

$$\lambda_1^{-1}|\xi|^2 \leqslant \langle \sigma\sigma^*(t,x)\xi,\xi\rangle \leqslant \lambda_1|\xi|^2, \quad x,\xi \in \mathbb{R}^d,\ t \geqslant 0.$$

Assumption 4.1.2 (Regularity). *For some $\alpha \in (0,1)$ we have $\sigma \in bC^\alpha_{0,T}$. Moreover there exists a positive constant $\lambda_2 > 0$ and $\beta \in [0,1]$ such that for all $x,y \in \mathbb{R}^d$ and $t \geqslant 0$,*

$$|b(t,0)| \leqslant \lambda_2, \quad |b(t,x)-b(t,y)| \leqslant \lambda_2(|x-y|^\beta \vee |x-y|). \tag{4.1.2}$$

It should be noticed that when $\beta = 0$, b can possibly be an *unbounded* measurable function with linear growth. For instance, $b(t,x) = x + b_0(t,x)$ with b_0 being bounded measurable satisfies (4.1.2). The drift $b(t,x) = c_1(t) + c_2(t)|x|^\beta,\ \beta \in [0,1]$ where c_1, c_2 are bounded measurable functions of time, also joins this class.

Under Assumptions 4.1.1 and 4.1.2, for any $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, it is well known that there exists a unique weak solution to (4.1.1) starting from x at time t (see e.g. [67], [3], [14], [48]).

To state our main result, we prepare some deterministic regularized flow associated with the drift b. Let ρ be a non-negative smooth function with support in the unit ball of $\mathbb{R}^d$ and such that $\int_{\mathbb{R}^d} \rho(x)dx = 1$. For $\varepsilon \in (0,1]$, define

$$\rho_\varepsilon(x) := \varepsilon^{-d}\rho(\varepsilon^{-1}x), \quad b_\varepsilon(t,x) := b(t,\cdot) * \rho_\varepsilon(x) = \int_{\mathbb{R}^d} b(t,y)\rho_\varepsilon(x-y)dy, \tag{4.1.3}$$

i.e. $*$ stands for the usual spatial convolution. Then for each $j = 1,2,\cdots$, it is easy to see that

$$\begin{aligned} |\nabla_x^j b_\varepsilon(t,x)| &= \left|\int_{\mathbb{R}^d} (b(t,y)-b(t,x))\nabla_x^j \rho_\varepsilon(x-y)dy\right| \\ &\leqslant \int_{\mathbb{R}^d} |b(t,y)-b(t,x)||\nabla_x^j \rho_\varepsilon|(x-y)dy \\ &\leqslant \lambda_2\varepsilon^\beta \int_{\mathbb{R}^d} |\nabla_x^j\rho_\varepsilon|(x-y)dy \leqslant c\varepsilon^{\beta-j}. \end{aligned} \tag{4.1.4}$$

On the other hand, from (4.1.2) we also have

$$|b_\varepsilon(t,x) - b(t,x)| \leqslant \int_{\mathbb{R}^d} |b(t,y) - b(t,x)|\rho_\varepsilon(x-y)dy \leqslant \lambda_2 \varepsilon^\beta. \tag{4.1.5}$$

For fixed $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, we denote by $\theta^{(\varepsilon)}_{t,s}(x)$ the deterministic flow solving

$$\dot\gamma^{(\varepsilon)}_{t,s}(x) = b_\varepsilon(s, \gamma^{(\varepsilon)}_{t,s}(x)),\ s \geqslant 0, \quad \gamma^{(\varepsilon)}_{t,t}(x) = x. \tag{4.1.6}$$

Note that $(\gamma^{(\varepsilon)}_{t,s}(x))_{s\geqslant t}$ stands for a forward flow and $(\gamma^{(\varepsilon)}_{t,s}(x))_{s\leqslant t}$ stands for a backward flow. Moreover, after the regularization, equation (4.1.6) is well posed.

The following lemma, which provides a kind of equivalence between mollified flows, is our starting point for treating the unbounded rough drifts.

Lemma 4.1.3 (Equivalence of flows)**.** *Let Assumptions 4.1.1 and 4.1.2 be in force. For any $\varepsilon \in (0,1]$, the mapping $x \mapsto \gamma^{(\varepsilon)}_{t,s}(x)$ is a C^∞-diffeomorphism and its inverse is given by $x \mapsto \gamma^{(\varepsilon)}_{s,t}(x)$. Moreover, for any $T > 0$, there exists a constant $C = C(T, \lambda_2, d) \geqslant 1$ such that for any $\varepsilon \in (0,1]$, all $|s-t| \leqslant T$ and $x, y \in \mathbb{R}^d$,*

$$|\gamma^{(1)}_{t,s}(x) - y| + |s-t| \asymp_C |\gamma^{(\varepsilon)}_{t,s}(x) - y| + |s-t| \asymp_C |x - \gamma^{(\varepsilon)}_{s,t}(y)| + |s-t|,$$

where $Q_1 \asymp_C Q_2$ means that $C^{-1}Q_2 \leqslant Q_1 \leqslant CQ_2$.

Proof. By (4.1.4), it is a classical fact that $x \mapsto \gamma^{(\varepsilon)}_{t,s}(x)$ is a C^∞-diffemorphism and its inverse is given by $x \mapsto \gamma^{(\varepsilon)}_{s,t}(x)$. Below, without loss of generality, we assume $t < s$. By (4.1.6) and (4.1.3), (4.1.4), (4.1.5), we have

$$\begin{aligned}|\gamma^{(\varepsilon)}_{t,s}(x) - \gamma^{(1)}_{t,s}(x)| &\leqslant \int_t^s \Big|b_\varepsilon(\tau, \gamma^{(\varepsilon)}_{t,\tau}(x)) - b_1(\tau, \gamma^{(\varepsilon)}_{t,\tau}(x))\Big| d\tau + \int_t^s \Big|b_1(\tau, \gamma^{(\varepsilon)}_{t,\tau}(x)) - b_1(\tau, \gamma^{(1)}_{t,\tau}(x))\Big| d\tau \\ &\leqslant 2\lambda_2(s-t) + \|\nabla b_1\|_\infty \int_t^s |\gamma^{(\varepsilon)}_{t,\tau}(x) - \gamma^{(1)}_{t,\tau}(x)| d\tau.\end{aligned}$$

By the Gronwall inequality we get

$$|\gamma^{(\varepsilon)}_{t,s}(x) - \gamma^{(1)}_{t,s}(x)| \leqslant 2\lambda_2(s-t)\varepsilon^{\|\nabla b_1\|_\infty (s-t)},$$

and therefore we have

$$|\gamma^{(1)}_{t,s}(x) - y| \leqslant |\gamma^{(\varepsilon)}_{t,s}(x) - y| + 2\lambda_2 \varepsilon^{\|\nabla b_1\|_\infty (s-t)}|s-t|.$$

By symmetry, we obtain the first $\asymp_C$. As for the second one, note that by the Gronwall inequality we derive

$$|\gamma^{(1)}_{t,s}(x) - \gamma^{(1)}_{t,s}(y)| \asymp_{\varepsilon^{\|\nabla b_1\|_\infty (s-t)}} |x-y| \Rightarrow |\gamma^{(1)}_{t,s}(x) - y| \asymp_{\varepsilon^{\|\nabla b_1\|_\infty (s-t)}} |x - \gamma^{(1)}_{s,t}(y)|. \tag{4.1.7}$$

By (4.1.7) and by the first $\asymp_C$, we obtain the second $\asymp_C$. $\square$

Notation 4.1.4. *For notational convenience we introduce the parameter*

$$\Theta := (T, \alpha, \beta, \lambda_1, \lambda_2, d), \tag{4.1.8}$$

where again $T > 0$ stands for the fixed considered final time. It is not restrictive to assume that λ_1 is also the Hölder modulus of σ.

Our main result is the following theorem.

Theorem 4.1.5. *Under Assumptions 4.1.1 and 4.1.2 with $\beta = 0$, for any $T > 0$, $x, y \in \mathbb{R}^d$ and $0 \leqslant t < s \leqslant T$, the unique weak solution $X_{t,s}(x)$ of (4.1.1) starting from x at time t admits a density $p(t, x; s, y)$ which is continuous in $x, y \in \mathbb{R}^d$. Moreover, $p(t, x; s, y)$ enjoys the following estimates:*

(i) (Two-sided density bounds) There exist two constants μ_0, $C_0 \geqslant 1$ depending on Θ such that for any $x, y \in \mathbb{R}^d$ and $0 \leqslant t < s \leqslant T$ we have

$$C_0^{-1}\Gamma^{\text{heat}}(\mu_0^{-1}\mathcal{I}_{s-t}, \gamma_{t,s}^{(1)}(x) - y) \leqslant p(t, x; s, y) \leqslant C_0\Gamma^{\text{heat}}(\mu_0\mathcal{I}_{s-t}, \gamma_{t,s}^{(1)}(x) - y). \tag{4.1.9}$$

(ii) (Gradient estimate in x) There exist two constants μ_1, $C_1 \geqslant 1$ depending on Θ such that for any $x, y \in \mathbb{R}^d$ and $0 \leqslant t < s \leqslant T$ we have

$$|\nabla_x p(t, x; s, y)| \leqslant \frac{C_1}{\sqrt{s-t}}\Gamma^{\text{heat}}(\mu_1\mathcal{I}_{s-t}, \gamma_{t,s}^{(1)}(x) - y). \tag{4.1.10}$$

(iii) (Second order derivative estimate in x) If Assumption 4.1.2 holds for some $\beta \in (0, 1]$, then there exist two constants μ_2, $C_2 \geqslant 1$ depending on Θ such that for any $x, y \in \mathbb{R}^d$ and $0 \leqslant t < s \leqslant T$ we have

$$\left|\nabla_x^2 p(t, x; s, y)\right| \leqslant \frac{C_2}{s-t}\Gamma^{\text{heat}}(\mu_2\mathcal{I}_{s-t}, \gamma_{t,s}^{(1)}(x) - y). \tag{4.1.11}$$

(iv) (Gradient estimate in y) If Assumption 4.1.2 holds for some $\beta \in (0, 1]$ and $\sigma \in bC_{0,T}^{1,\alpha}$ for some $\alpha \in (0, 1)$, then there exist two constants μ_3, $C_3 \geqslant 1$ depending on Θ and the Hölder modulus of $\nabla\sigma$ such that, for any $x, y \in \mathbb{R}^d$ and $0 \leqslant t < s \leqslant T$ we have

$$|\nabla_y p(t, x; s, y)| \leqslant \frac{C_3}{\sqrt{s-t}}\Gamma^{\text{heat}}(\mu_3\mathcal{I}_{s-t}, \gamma_{t,s}^{(1)}(x) - y). \tag{4.1.12}$$

Remark 4.1.6. *By Lemma 4.1.3, the above $\gamma_{t,s}^{(1)}(x)$ can be replaced by any regularized flow $\gamma_{t,s}^{(\varepsilon)}(x)$. Importantly, if b satisfies Assumption 4.1.2 for some $\beta \in (0, 1]$, then $\gamma_{t,s}^{(1)}(x)$ can be replaced as well by **any** Peano flow solving $\dot{\gamma}_{t,s}(x) = b(s, \gamma_{t,s}(x))$, $\gamma_{t,t}(x) = x$. Indeed, it is plain to check that, in this case, the result of Lemma 4.1.3 still holds with $\gamma_{t,s}(x)$ instead of $\gamma_{t,s}^{(\varepsilon)}(x)$.*

Remark 4.1.7. *Under the assumptions of the theorem, in fact, we can show the Hölder continuiy of $\nabla_x p$, $\nabla_x^2 p$ and $\nabla_y p$ in the variables x and y (see Appendix C).*

4.2 A priori heat kernel estimates for SDEs with smooth coefficients

In this section we suppose Assumptions 4.1.1 and 4.1.2 to be in force, and consider the mollified b_ε and σ_ε. In particular, we have

$$\lambda_j^{(\varepsilon)} := \sum_{k=1,\cdots j} \Big(\|\nabla_x^k b_\varepsilon\|_\infty + \|\nabla_x^k \sigma_\varepsilon\|_\infty \Big) < \infty, \qquad j \in \mathbb{N}. \tag{4.2.1}$$

In the following, for ease of notations, we shall drop the subscripts ε. In other words, we assume that the coefficients b and σ satisfy Assumptions 4.1.1, 4.1.2 and (4.2.1), and call (**S**) this set of Assumptions. Under (**S**), it is well known that for each $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, the following SDE has a unique strong solution:

$$dX_{t,s} = b(s, X_{t,s})ds + \sigma(s, X_{t,s})\, dW_s, \quad s \geqslant t, \qquad X_{t,t} = x. \tag{4.2.2}$$

The following theorem is well known in the theory of the Malliavin calculus. We refer to [51], [71, Remarks 2.1 and 2.2] or [75, Theorem 5.4] for more details.

Theorem 4.2.1. *Under (**S**), for any $j, j' \in \mathbb{N}$, $p > 1$ and $T > 0$, there is a constant $C = C(\Theta, j, j', \kappa_{j+j'})$ such that for all $x \in \mathbb{R}^d$, $0 \leqslant t < s \leqslant T$ and $f \in bC^\infty(\mathbb{R}^d)$,*

$$\left|\nabla^j E\left[\nabla^{j'} f(X_{t,s}(x))\right]\right| \leqslant \frac{C_p}{(s-t)^{\frac{j+j'}{2}}} E\left[|f(X_{t,s}(x))|^p\right]^{1/p}. \tag{4.2.3}$$

In particular, $X_s^{t,x}$ has a density $p(t,x;s,y)$, which is smooth in x, y.

Remark 4.2.2. *By Itô's formula, one sees that $p(t,x,s,y)$ satisfies the backward Kolmogorov equation*

$$\partial_t p(t,x;s,y) + \mathcal{L}_{t,x} p(t,x;s,y) = 0, \quad \lim_{t\to s^-} \int_{\mathbb{R}^d} p(t,x;s,y) f(x)dx = f(y), \tag{4.2.4}$$

and the forward Kolmogorov equation (Fokker-Planck equation):

$$\partial_s p(s,x;t,y) - \mathcal{L}^*_{s,y} p(t,x;s,y) = 0, \quad \lim_{s\to t^+} \int_{\mathbb{R}^d} p(t,x;s,y) f(y)dy = f(x), \tag{4.2.5}$$

where, setting $a = \sigma\sigma^*/2$,

$$\mathcal{L}_{t,x} f(x) = \mathrm{tr}\big(a(t,x)\nabla_x^2 f(x)\big) + \langle b(t,x), \nabla_x f(x)\rangle$$

and

$$\mathcal{L}^*_{s,y} f(y) = \partial_{y_i}\partial_{y_j}(a_{ij}(s,y)f(y)) - \mathrm{div}(b(s,\cdot)f)(y).$$

4.2.1 The Duhamel representation

Fix now $(t_0, x_0) \in \mathbb{R}_+ \times \mathbb{R}^d$ as *freezing parameters* to be chosen later on. Let $\widetilde{X}_{t,s}^{(t_0,x_0)}(x)$ denote the process starting at x at time t, with dynamics

$$d\widetilde{X}_{t,s}^{(t_0,x_0)} = b(s, \gamma_{t_0,s}(x_0))ds + \sigma(s, \gamma_{t_0,s}(\xi))dW_s, \ s \geqslant t, \quad \widetilde{X}_{t,t}^{(t_0,x_0)} = x, \tag{4.2.6}$$

i.e. $\widetilde{X}_{t,s}^{(t_0,x_0)}$ denotes the process derived from (4.2.2), when freezing the spatial coefficients along the flow $\gamma_{t_0,\cdot}(x_0)$, where $\gamma_{t_0,\cdot}(x_0)$ is the unique solution of ODE (4.1.6) corresponding to b. For any choice of freezing couple (t_0, x_0), $\widetilde{X}_{t,s}^{(t_0,x_0)}$ has a Gaussian density

$$\widetilde{p}^{t_0,x_0}(t, x; s, y) = \frac{\exp\{-\langle(\widetilde{\mathcal{C}}_{t,s}^{t_0,x_0})^{-1}(\vartheta_{t,s}^{t_0,x_0} + x - y), \vartheta_{t,s}^{t_0,x_0} + x - y\rangle/2\}}{\sqrt{(2\pi)^d \det(\widetilde{\mathcal{C}}_{t,s}^{t_0,x_0})}}, \tag{4.2.7}$$

where

$$\vartheta_{t,s}^{t_0,x_0} := \int_t^s b(r, \gamma_{t_0,r}(x_0))dr, \qquad \widetilde{\mathcal{C}}_{t,s}^{t_0,x_0} := \int_t^s \sigma\sigma^*(r, \gamma_{t_0,r}(x_0))dr.$$

In particular, $\widetilde{p}^{t_0,x_0}(t, x, s, y)$ satisfies for fixed $(s, y) \in \mathbb{R}_+ \times \mathbb{R}^d$:

$$\partial_t \widetilde{p}^{t_0,x_0}(t, x; s, y) + \widetilde{\mathcal{L}}_{t,x}^{t_0,x_0} \widetilde{p}^{t_0,x_0}(t, x; s, y) = 0, \quad (t, x) \in [0, s) \times \mathbb{R}^d, \tag{4.2.8}$$

subjected to the final condition

$$\lim_{t \to s^-} \int_{\mathbb{R}^d} \widetilde{p}^{t_0,x_0}(t, x; s, y) f(x) dx = f(y), \tag{4.2.9}$$

where

$$\widetilde{\mathcal{L}}_{t,x}^{t_0,x_0} = \mathrm{tr}\big(a(t, \gamma_{t_0,t}(x_0)) \cdot \nabla_x^2\big) + \langle b(t, \gamma_{t_0,t}(x_0)), \nabla_x\rangle$$

denotes the generator of the diffusion with frozen coefficients in (4.2.6).

The following lemma is direct by the explicit representation (4.2.7), the uniform ellipticity condition (4.1.1) and the chain rule.

Lemma 4.2.3 (A priori controls for the frozen Gaussian density)**.** *For any $j = 0, 1, 2, \cdots$, there exist constants $\mu_j, C_j > 0$ depending only on j, λ_1, d such that for all $(t_0, x_0) \in \mathbb{R}_+ \times \mathbb{R}^d$, $x, y \in \mathbb{R}^d$ and $0 \leqslant t < s < \infty$,*

$$\widetilde{p}^{t_0,x_0}(t, x; s, y) \geqslant C_0 \Gamma^{\mathrm{heat}}\big(\mu_0^{-1} \mathcal{I}_{s-t}, \vartheta_{t,s}^{t_0,x_0} + x - y\big),$$

and

$$\big|\nabla_x^j \widetilde{p}^{t_0,x_0}(t, x; s, y)\big| = \big|\nabla_y^j \widetilde{p}^{t_0,x_0}(t, x; s, y)\big| \leqslant C_j (s-t)^{-\frac{j}{2}} \Gamma^{\mathrm{heat}}\big(\mu_j \mathcal{I}_{s-t}, \vartheta_{t,s}^{t_0,x_0} + x - y\big)$$

Moreover, for each $j, j' \in \mathbb{N}$, there are constants C', μ' depending on Θ and $\lambda_{j'}$ such that

$$\left|\nabla_x^j \nabla_\xi^{j'} \widetilde{p}^{t_0,x_0}(t,x;s,y)\right| \leqslant C'(s-t)^{-\frac{j}{2}} \Gamma^{\mathrm{heat}}\big(\mu' \mathcal{I}_{s-t}, \vartheta_{t,s}^{t_0,x_0} + x - y\big). \tag{4.2.10}$$

Proof. We focus on (4.2.10) for which it suffices to note that for any $k \in \mathbb{N}$, $T > 0$,

$$|\nabla_\xi^k \vartheta_{t,s}^{t_0,x_0}| + |\nabla_\xi^k \widetilde{\mathcal{C}}_{t,s}^{t_0,x_0}| \leqslant C_k |s-t|, \quad 0 \leqslant t < s \leqslant T,\ (\tau,\xi) \in [s,t] \times \mathbb{R}^d,$$

where the constant C_k depends on the bound of $\nabla^j b$ and $\nabla^j \sigma$, $j = 1, \cdots, k$. □

The starting point of our analysis is the following Duhamel type representation formula which readily follows in the current *smooth coefficients* setting from (4.2.4)-(4.2.5) and (4.2.8)-(4.2.9):

$$\begin{aligned} p(t,x;s,y) &= \widetilde{p}^{t_0,x_0}(t,x;s,y) + \int_t^s \int_{\mathbb{R}^d} \widetilde{p}^{t_0,x_0}(t,x;r,z)(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{t_0,x_0}) p(r,z;s,y) dz dr && (4.2.11) \\ &= \widetilde{p}^{t_0,x_0}(t,x;s,y) + \int_t^s \int_{\mathbb{R}^d} p(t,x;r,z)(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{t_0,x_0}) \widetilde{p}^{t_0,x_0}(r,z;s,y) dz dr, && (4.2.12) \end{aligned}$$

where

$$\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{t_0,x_0} = \mathrm{Tr}(A_{r,z}^{t_0,x_0} \cdot \nabla_z^2) + B_{r,z}^{t_0,x_0} \cdot \nabla_z \tag{4.2.13}$$

and

$$A_{r,z}^{t_0,x_0} := a(r,z) - a(r, \gamma_{t_0,r}(x_0)), \quad B_{r,z}^{t_0,x_0} := b(r,z) - b(r, \gamma_{t_0,r}(x_0)). \tag{4.2.14}$$

If we take $(t_0, x_0) = (t, x)$ in (4.2.11) and set $Z_0(t,x,s,y) := \widetilde{p}^{t,x}(t,x,s,y)$, then we obtain the *forward* representation

$$p(t,x;s,y) = Z_0(t,x;s,y) + \int_t^s \int_{\mathbb{R}^d} Z_0(t,x;r,z)(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{t,x}) p(r,z;s,y) dz dr,$$

and in this case

$$\vartheta_{t,s}^{t,x} + x - y = \int_t^s b(r, \gamma_{t,r}(x)) dr + x - y = \gamma_{t,s}(x) - y; \tag{4.2.15}$$

it involves the *forward* deterministic flow $\gamma_{t,s}(x)$ in the frozen Gaussian density. If we now take $(t_0, x_0) = (s, y)$ in (4.2.12) and set $Z_1(t,x;s,y) := \widetilde{p}^{s,y}(t,x;s,y)$, we then obtain the *backward* (cf. Section 2.7) representation

$$p(t,x;s,y) = Z_1(t,x;s,y) + \int_t^s \int_{\mathbb{R}^d} p(t,x;r,z)(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{s,y}) Z_1(r,z;s,y) dz dr, \tag{4.2.16}$$

and in this case

$$\vartheta_{t,s}^{s,y} + x - y = \int_t^s b(r, \gamma_{s,r}(y)) dr + x - y = x - \gamma_{s,t}(y).$$

It involves the *backward* deterministic flow $\gamma_{s,t}(y)$ in the frozen Gaussian density.

4.2.2 Two-sided Estimates

We first deal here with the two-sided estimates for the density in the current *smooth coefficients* setting. Importantly, we emphasize as much as possible that all the controls obtained are actually *independent* of the derivatives of the coefficients, or even of the continuity of the drift b, but only depend on the parameters gathered in Θ introduced in (4.1.8). We first iterate in Section 4.2.2 the Duhamel representation (4.2.16) to obtain the *parametrix series* expansion of the density. We then give some controls related to the smoothing effects in time of the parametrix kernel.

As seen is Section 2.6.1, a specific feature of the heat kernels associated with unbounded drifts is that the corresponding parametrix series needs to be handled with care. Indeed, it is not direct to prove that it converges unless a suitable lower bound for the density is already available, and therefore some truncation step is needed. Here we use a similar kind of argument than in [14], based on slightly different techniques deriving from the stochastic control representation of some Brownian functionals, see [6], [74] and Section 4.2.2 below.

We can assume here without loss of generality that $T \leqslant 1$. Indeed, once the two-sided estimates are established in this case, they can be easily extended to any compact time interval $[0, T]$ through Gaussian convolutions using the scaling properties (see Lemma 4.2.9).

Two-sided heat kernel estimates parametrix series

For notational convenience, we write from now on for $x, y \in \mathbb{R}^d$ and $0 \leqslant t < s \leqslant T$

$$Z_1(t,x;s,y) = \widetilde{p}^{(s,y)}(t,x;s,y), \quad H(s,x;t,y) := (\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}^{s,y}_{r,z}) Z_1(t,x;s,y). \tag{4.2.17}$$

Thus, from the Duhamel representation (4.2.16), recalling notation (1.4.5), we have

$$p(t,x;s,y) = Z_1(t,x;s,y) + (p \otimes H)(t,x;s,y). \tag{4.2.18}$$

For $N \geqslant 2$, by iterating $N-1$-times the identity (4.2.18), we obtain

$$p(s,x;t,y) = Z_1(t,x;s,y) + \sum_{j=1}^{N-1} (Z_1 \otimes H^{\otimes j})(t,x;s,y) + (p \otimes H^{\otimes N})(t,x;s,y). \tag{4.2.19}$$

We shall now use the following notational convention without mentioning the flow $\gamma^{(1)}_{t,s}(x)$. For $x, y \in \mathbb{R}^d$ and $0 \leqslant t < s \leqslant T$, we define for $\mu > 0$:

$$\Gamma_\mu(t,x;s,y) := \Gamma^{\text{heat}}(\mu \mathcal{I}_{s-t}, \gamma^{(1)}_{t,s}(x) - y) = \frac{1}{(2\pi)^{\frac{d}{2}} (s-t)^{\frac{d}{2}}} \exp\left(-\frac{|\gamma^{(1)}_{t,s}(x) - y|^2}{\mu(s-t)} \right), \tag{4.2.20}$$

recalling (1.1.7) for the last equality. From Lemmas 4.2.3 and 4.1.3 we derive

Lemma 4.2.4. *For any $T > 0$ and $j = 0, 1, 2, \cdots$, there exist constants $\widetilde{\mu}_j, \widetilde{C}_j > 0$ depending only on Θ such that for all $x, y \in \mathbb{R}^d$ and $0 \leqslant t < s \leqslant T$,*

$$Z_1(t,x;s,y) \geqslant \widetilde{C}_0 \Gamma_{\widetilde{\mu}_0^{-1}}(t,x;s,y),$$

and for all $\alpha \in [0,1]$,

$$|x - \theta_{s,t}(y)|^\alpha \left|\nabla_x^j Z_1(t,x;s,y)\right| \leqslant \widetilde{C}_j (s-t)^{\frac{\alpha}{2}-\frac{j}{2}} \Gamma_{\widetilde{\mu}_j}(t,x;s,y). \tag{4.2.21}$$

The following convolution type inequality is also an easy consequence of Lemma 4.1.3.

Lemma 4.2.5. *For any $T > 0$, there is an $\varepsilon = \varepsilon(\Theta) > 1$ such that for any $\mu > 0$, there is a $C_\varepsilon = C_\varepsilon(\Theta, \mu) > 0$ such that for all $x, y \in \mathbb{R}^d$ and $0 \leqslant t < s \leqslant T$ and $r \in [s,t]$,*

$$\int_{\mathbb{R}^d} \Gamma_\mu(t,x;r,z)\Gamma_\mu(r,z;s,y)dz \leqslant C_\varepsilon \Gamma_{\varepsilon\mu}(t,x;s,y).$$

Proof. By definition and Lemma 4.1.3, we have for some $\varepsilon > \varepsilon' > 1$,

$$\begin{aligned}
\int_{\mathbb{R}^d} \Gamma_\mu(t,x;r,z)\Gamma_\mu(r,z;s,y)dz &= \int_{\mathbb{R}^d} \Gamma^{\text{heat}}\big(\mu\mathcal{I}_{r-t}, \gamma_{t,r}^{(1)}(x) - z\big)\Gamma^{\text{heat}}\big(\mu\mathcal{I}_{s-r}, \gamma_{r,s}^{(1)}(z) - y\big)dz \\
&\leqslant C_{\varepsilon'} \int_{\mathbb{R}^d} \Gamma^{\text{heat}}\big(\varepsilon'\mu\mathcal{I}_{r-t}, \gamma_{t,r}^{(1)}(x) - z\big)\Gamma^{\text{heat}}\big(\varepsilon'\mu\mathcal{I}_{s-r}, z - \gamma_{s,r}^{(1)}(y)\big)dz \\
&= C_{\varepsilon'}\Gamma^{\text{heat}}\big(\varepsilon'\mu\mathcal{I}_{s-t}, \gamma_{t,r}^{(1)}(x) - \gamma_{s,r}^{(1)}(y)\big) \leqslant C_\varepsilon \Gamma_{\varepsilon\mu}(t,x;s,y),
\end{aligned}$$

where the second equality is due to the Chapman-Kolmogorov property for the Gaussian semigroup, and the last inequality again follows from Lemma 4.1.3 and the following control

$$|\gamma_{t,s}^{(1)}(x) - y| = |\gamma_{r,s}^{(1)} \circ \gamma_{t,r}^{(1)}(x) - \gamma_{r,s}^{(1)} \circ \gamma_{s,r}^{(1)}(y)| \leqslant C|\gamma_{t,r}^{(1)}(x) - \gamma_{s,r}^{(1)}(y)|. \tag{4.2.22}$$

The proof is complete. □

Next we give the control for the iterated convolutions of the parametrix kernel H which appears in the expansion (4.2.19), that is similar to (1.4.19) and (2.6.29), but with a crucial difference. Indeed, because of the flow we are not able to recover an estimate that is uniform in the iteration parameter.

Lemma 4.2.6. *Under Assumptions 4.1.1 and 4.1.2, for any $T > 0$ and $N \in \mathbb{N}$, there are constants $C_N, \mu_N > 0$ depending only on Θ such that, for all $x, y \in \mathbb{R}^d$ and $0 \leqslant t < s \leqslant T$,*

$$|H^{\otimes N}(t,x;s,y)| \leqslant C_N (s-t)^{-1+\frac{N\alpha}{2}} \Gamma_{\mu_N}(t,x;s,y),$$

where $\mu_N \to \infty$ as $N \to \infty$.

The proof proceeds in the same manner as for Propositions 1.4.8 and 2.6.17, starting from the definition of H in (4.2.17) and exploiting Lemmas 4.2.4 and 4.2.5, and therefore is omitted. From the above lemma, (4.2.19) and (4.2.21), we thus derive that for all $N \in \mathbb{N}$, $x, y \in \mathbb{R}^d$ and $0 \leqslant t < s \leqslant T$:

$$p(t,x;s,y) \leqslant \bar{C}\Gamma_{\mu_{N-1}}(t,x;s,y) + |p \otimes H^{\otimes N}(t,x;s,y)|, \tag{4.2.23}$$

which is *almost* the expected upper-bound except that we explicitly have to control the remainder to stop the iteration at some fixed N to avoid the collapse to ∞ of μ_N as N goes to infinity. This is precisely the purpose of the next subsection.

Stochastic control arguments and truncation of the parametrix series

In this section, we aim at controlling the remainder term $(p \otimes H^{\otimes N})(t,x;s,y)$ in the almost Gaussian upper-bound (4.2.23).

To this end, we use the variational representation formula to show the a priori derivative estimates of the density when the coefficients are smooth, following a similar idea as in [14]. The following variational representation formula was first proved by Boué and Dupuis [6].

Theorem 4.2.7. *Fix $T > 0$ and let F be a bounded Wiener functional on the classical Wiener space $(\Omega, \mathcal{F}, P)$*[1] *which is $\mathcal{F}_T$ measurable. Then it holds that*

$$-\ln E\left[e^F\right] = \inf_{h \in \mathcal{S}} E\left[\frac{1}{2}\int_0^T |\dot{h}(\tau)|^2 \mathrm{d}s - F(\omega + h)\right],$$

where $\mathcal{S}$ denotes the set of all $\mathbb{R}^d$-valued $\mathcal{F}_t$-adapted and absolutely continuous processes with

$$E\left[\int_0^T |\dot{h}(\tau)|^2 d\tau\right] < \infty.$$

Using the above variational representation formula, we obtain the following important lemma.

Lemma 4.2.8. *Let $\ell : \mathbb{R}^d \to (0,\infty)$ be a bounded measurable function s.t. for all $x \in \mathbb{R}^d, \zeta^{-1} \leqslant \ell(x) \leqslant \zeta$ for some $\zeta \geqslant 1$. Under Assumptions 4.1.1 and 4.1.2, for any $T > 0$, there is a constant $C = C(\Theta) > 0$ such that for all $x \in \mathbb{R}^d$ and $0 \leqslant t < s \leqslant T$,*

$$E\left[\ell(X_{t,s}(x))\right] \leqslant C \sup_{z \in \mathbb{R}^d} \exp\left\{\ln \ell(z) - C^{-1}|z - \gamma_{t,s}^{(1)}(x)|^2\right\}.$$

[1] We recall that for the Wiener space, the fundamental set $\Omega = C(\mathbb{R}_+, \mathbb{R}^d)$ and in this framework $\omega \in \Omega$ simply stands here for an $\mathbb{R}^d$-valued continuous function. Also, the canonical process $\big(\omega(t)\big)_{t \geqslant 0}$ is a standard d-dimensional Brownian motion under the measure $\mathbb{P}$.

Proof. Without loss of generality, we assume $t = 0$ and write $X_s := X_{0,s}(x)$. Since in the current smooth coefficients setting X_s can be viewed as a functional of the Brownian path, taking $F = \ln(\ell(X_s))$, we derive from Theorem 4.2.7 that

$$-\ln E\left[\ell(X_s)\right] = \inf_{h\in\mathcal{S}} E\left[\frac{1}{2}\int_0^s |\dot{h}(\tau)|^2 d\tau - \ln \ell(X_s^h)\right],$$

where X^h solves the following SDE:

$$dX_s^h = \Big(b(s, X_s^h) + \sigma(s, X_s^h)\dot{h}(s)\Big)ds + \sigma(s, X_s^h)dW_s, \quad X_0^h = x,$$

i.e. the control process h enters the dynamics in the drift part. Note that $\gamma_s := \gamma_{0,s}(x)$ solves the following ODE:

$$\dot{\gamma}_s = b(s, \gamma_s), \quad \gamma_0 = x.$$

By Itô's formula, we have

$$E\left[|X_s^h - \gamma_s|^2\right] = E\left[\int_0^s \Big(2\langle X_\tau^h - \gamma_\tau, b(\tau, X_\tau^h) - b(\tau, \gamma_\tau) + \sigma(\tau, X_\tau^h)\dot{h}(\tau)\rangle + (\sigma\sigma^*)(\tau, X_\tau^h)\Big)d\tau\right].$$

Recalling

$$|b(s,x) - b(s,y)| \leqslant \lambda_2(1 + |x-y|),$$

the Young inequality yields

$$E\left[|X_s^h - \gamma_s|^2\right] \leqslant CE\left[\int_0^s \Big(|X_\tau^h - \gamma_\tau|^2 + |\dot{h}(\tau)|^2\Big)\, d\tau\right] + Cs.$$

From the Gronwall inequality, we thus obtain

$$E\left[|X_s^h - \gamma_s|^2\right] \leqslant CE\left[\int_0^s |\dot{h}(\tau)|^2 d\tau\right] + Cs.$$

Hence we have,

$$\frac{1}{2}E\left[\int_0^s |\dot{h}(\tau)|^2 d\tau\right] \geqslant C^{-1}E\left[|X_s^h - \gamma_s|^2\right] - Cs.$$

Therefore, we eventually derive

$$-\ln E\left[\ell(X_s)\right] \geqslant \inf_{h\in\mathcal{S}} E\left[C^{-1}|X_s^h - \gamma_s|^2 - \ln \ell(X_s^h)\right] - C \geqslant \inf_{z\in\mathbb{R}^d}\big(C^{-1}|z - \gamma_s|^2 - \ln \ell(z)\big) - C.$$

The desired estimate eventually follows from Lemma 4.1.3. □

Next we state a direct yet important scaling lemma. We refer to Section 2.3 of [14] for additional details.

Lemma 4.2.9 (Scaling property of the density)**.** *Fix $0 \leqslant t < s \leqslant T$ and let $\lambda := s - t$. Introduce for $u \in [0,1]$, $\widehat{X}_u^\lambda := \lambda^{-\frac{1}{2}} X_{t,t+u\lambda}$. Then, $(\widehat{X}_u^\lambda)_{u\in[0,1]}$ satisfies the SDE*

$$d\widehat{X}_u^\lambda = \lambda^{\frac{1}{2}} b\big(t+u\lambda, \widehat{X}_u^\lambda \lambda^{\frac{1}{2}}\big)du + \sigma\big(t+u\lambda, \widehat{X}_u^\lambda\lambda^{\frac{1}{2}}\big)d\widehat{W}_u^\lambda = \widehat{b}^\lambda(u, \widehat{X}_u^\lambda)du + \widehat{\sigma}^\lambda(u, \widehat{X}_u^\lambda)d\widehat{W}_u^\lambda,$$

where $\widehat{W}_u^\lambda = \lambda^{-\frac{1}{2}} W_{u\lambda}$ is a Brownian motion. It also holds that:

$$p(t,x;s,y) = \lambda^{-\frac{d}{2}} \widehat{p}^\lambda\Big(0, \lambda^{-\frac{1}{2}}x; 1, \lambda^{-\frac{1}{2}}y\Big),$$

and introducing for $z \in \mathbb{R}^d$, $u \in [0,1]$, $\partial_u \widehat{\gamma}_{0,u}^\lambda(z) = \widehat{b}^\lambda(u, \widehat{\gamma}_{0,u}(z))$, $\widehat{\gamma}_{0,0}^\lambda(z) = z$,

$$\big|\widehat{\gamma}_{0,1}^\lambda(\lambda^{-\frac{1}{2}}x) - \lambda^{-\frac{1}{2}}y\big|^2 = \lambda^{-1}|\gamma_{t,s}(x) - y|^2.$$

Proof. We only focus on the last statement. The other ones readily follow from the change of variable. Write:

$$\lambda^{-\frac{1}{2}}\gamma_{t,s}(x) = \lambda^{-\frac{1}{2}}x + \lambda^{-\frac{1}{2}}\int_t^s b(r, \gamma_{t,r}(x))dr = \lambda^{-\frac{1}{2}}x + \lambda^{\frac{1}{2}}\int_0^1 b(t+u\lambda, \gamma_{t,t+u\lambda}(x))du.$$

Setting now for $u \in [0,1]$, $\bar{\gamma}_{0,u}(x) = \gamma_{t,t+u\lambda}(x)$, the above equation rewrites:

$$\lambda^{-\frac{1}{2}}\bar{\gamma}_{0,1}(x) = \lambda^{-\frac{1}{2}}x + \lambda^{\frac{1}{2}}\int_0^1 b(t+u\lambda, \bar{\gamma}_{0,u}(x))du = \lambda^{-\frac{1}{2}}x + \int_0^1 \widehat{b}^\lambda\big(u, \lambda^{-\frac{1}{2}}\bar{\gamma}_{0,u}(x)\big)du$$

from which we readily derive by uniqueness of the solution to the ODE that for $u \in [0,1]$,

$$\lambda^{-\frac{1}{2}}\bar{\gamma}_{u,0}(x) = \widehat{\gamma}_{u,0}^\lambda(\lambda^{-\frac{1}{2}}x) = \lambda^{-\frac{1}{2}}\gamma_{t,t+u\lambda}(x),$$

which gives the statement. □

We will now use the previous Lemmas 4.2.8 and 4.2.9 to establish the following result from which the Gaussian upper-bound will readily follow.

Lemma 4.2.10 (Control of the remainder)**.** *Choose N large enough in order to have:*

$$-1 + \frac{N\alpha}{2} > \frac{d}{2}. \tag{4.2.24}$$

There exist constants $C_0, \mu_0 > 0$ depending only on Θ such that for all $x, y \in \mathbb{R}^d$ and $0 \leqslant t < s \leqslant T$,

$$|(p \otimes H^{\otimes N})(t,x;s,y)| \leqslant C_0\Gamma_{\mu_0}(t,x;s,y).$$

Proof. From the scaling property of Lemma 4.2.9 above, we can assume without loss of generality that $t = 0$ and $s = 1$. By Lemmas 4.2.6, 4.2.8, we have

$$\begin{aligned}
|(p \otimes H^{\otimes N})(0,x;1,y)| &\leqslant \int_0^1 \left| \int_{\mathbb{R}^d} p(0,x,r,z) H^{\otimes N}(r,z;1,y) dz \right| dr \\
&= \int_0^1 \left| E\left[H^{\otimes N}(r, X_{0,r}(x); 1, y) \right] \right| dr \\
&\leqslant C_N \int_0^1 (1-r)^{-1+\frac{N\alpha}{2}} E\left[\Gamma_{\mu_N}(r, X_{0,r}(x); 1, y) \right] dr \\
&\leqslant C_N \int_0^1 (1-r)^{-1+\frac{N\alpha}{2}} \sup_{z \in \mathbb{R}^d} \exp\left\{ \ln \Gamma_{\mu_N}(r,z;1,y) - C^{-1}|z - \gamma_{0,r}^{(1)}(x)|^2 \right\} dr.
\end{aligned}$$

Since by (0.0.7),

$$\ln \Gamma_{\mu_N}(r,z;1,y) = \ln \Gamma^{\text{heat}}(\mu_N \mathcal{I}_{1-r}, \gamma_{r,1}^{(1)}(z) - y) = -\frac{d}{2}\ln(1-r) - \mu_N |\gamma_{r,1}^{(1)}(z) - y|^2/(1-r),$$

we have

$$\begin{aligned}
&\sup_z (\ln \Gamma^{\text{heat}}(\mu_N \mathcal{I}_{1-r}, \gamma_{r,1}^{(1)}(z) - y) - C^{-1}|z - \gamma_{0,r}^{(1)}(x)|^2) \\
&\leqslant -\frac{d}{2}\ln(1-r) - \inf_z (\mu_N |\gamma_{r,1}^{(1)}(z) - y|^2/(1-r) + C^{-1}|z - \gamma_{0,r}^{(1)}(x)|^2) \\
&\leqslant -\frac{d}{2}\ln(1-r) - \mu_N' \inf_z (|z - \gamma_{1,r}^{(1)}(y)|^2/(1-r) - C(1-r) + |z - \gamma_{0,r}^{(1)}(x)|^2) \\
&\leqslant -\frac{d}{2}\ln(1-r) - \mu_N' |\gamma_{1,r}^{(1)}(y) - \gamma_{0,r}^{(1)}(x)|^2/2 + C \\
&\leqslant -\frac{d}{2}\ln(1-r) - \mu_N'' |\gamma_{0,1}^{(1)}(x) - y|^2 + C,
\end{aligned}$$

where the last step is due to (4.2.22). Therefore, from the condition (4.2.24) and the above computations, there exist constants $C_0, \lambda_0 > 0$ depending only on Θ such that

$$|(p \otimes H^{\otimes N})(0,x;1,y)| \leqslant C_0 \Gamma^{\text{heat}}(\mu_0 \mathcal{I}_1, \gamma_{0,1}^{(1)}(x) - y) = C_0 \Gamma_{\mu_0}(0,x;1,y).$$

The general statement for arbitrary $s,\ t$ again follows from the scaling arguments of Lemma 4.2.9.

□

Final derivation of the two-sided heat kernel estimates

We are now in position to prove the following two-sided estimates.

Theorem 4.2.11. *Under Assumptions 4.1.1 and 4.1.2, for any $T > 0$, there exist constants $C_0, \mu_0 \geqslant 1$ depending only on Θ such that for all $x, y \in \mathbb{R}^d$ and $0 \leqslant t < s \leqslant T$,*

$$C_0^{-1} \Gamma_{\mu_0^{-1}}(t,x;s,y) \leqslant p(t,x;s,y) \leqslant C_0 \Gamma_{\mu_0}(t,x;s,y).$$

Proof. (i) *(Upper bound)* The upper bound is a direct consequence of the expansion (4.2.23) and the previous Lemma 4.2.10 up to a possible modification of the constants C_0, μ_0 that anyhow still only depend on Θ.

(ii) *(Lower bound)* By the upper bound and Lemmas 4.2.6 and 4.2.5, we get for some $\mu_1 > \mu_0$ and $\varepsilon > 1$:

$$\begin{aligned}|p \otimes H(t,x;s,y)| &\leqslant C \int_t^s (s-r)^{-1+\frac{\alpha}{2}} \int_{\mathbb{R}^d} \Gamma_{\mu_1}(t,x;r,z)\Gamma_{\mu_1}(r,z;s,y)dzdr \\ &\leqslant C_2(s-t)^{\frac{\alpha}{2}}\Gamma_{\varepsilon\mu_1}(t,x;s,y).\end{aligned}$$

Hence, for $|\gamma_{t,s}^{(1)}(x) - y| \leqslant \sqrt{s-t}$, recalling (4.2.18) and (4.2.20), we have

$$p(t,x;s,y) \geqslant \left(C_1 - C_2(s-t)^{\frac{\alpha}{2}}\right)\Gamma_{\varepsilon\mu_1}(t,x;s,y) \geqslant \left(C_1 - C_2(s-t)^{\frac{\alpha}{2}}\right)(s-t)^{-d/2}e^{-\varepsilon\mu_1}.$$

In particular, letting $s-t \leqslant \delta$ with δ small enough, we obtain that

$$p(t,x;s,y) \geqslant C_3(s-t)^{-d/2} \quad \text{on} \quad |\gamma_{t,s}^{(1)}(x) - y| \leqslant \sqrt{s-t}. \tag{4.2.25}$$

Next we propose yet another chaining argument to obtain the lower bound when $|\gamma_{t,s}^{(1)}(x) - y| \geqslant \sqrt{s-t}$. The idea is again to consider a suitable sequence of balls between the points x and y, for which the diagonal lower estimate (4.2.25) holds, and which also have a large enough volume to consent to derive the global *off-diagonal* lower bound. The usual strategy to build such balls consists in considering the "geodesic" line between x and y. In the non-degenerate case, when the coefficients are bounded, this is nothing but the straight-line joining x and y: this is precisely the strategy used in Section 1.4.3. When dealing with unbounded coefficients, a possibility is to consider the optimal path associated with the deterministic controllability problem $\dot{\varphi}_u = b(u, \varphi_u) + \varphi_u,\ u \in [t,s],\ \varphi_r = x, \varphi_s = y$ with $\varphi \in L^2([t,s], \mathbb{R}^d)$. This is the strategy adopded in Section 2.6.1 for a Lipschitz continuous b; the constants in the lower bound estimates obtained therein actually depend on the Lipschitz modulus b.

We adopt here a slightly different strategy which only involves the *mollified* flow $\gamma^{(1)}$ but which will have the main advantage to provide constants that will again only depend on Θ and not on the smoothness of b, using thoroughly the controls established in Lemma 4.1.3. We now detail such a construction which is in some sense *original* though pretty natural. From the scaling arguments of Lemma 4.2.9, we can assume without loss of generality that $\delta = 1$, $t = 0$ and $s = 1$. Suppose $|\gamma_{0,1}^{(1)}(x) - y| > 1$ and let M be the smallest integer greater than $4\varepsilon^{2\|\nabla b_1\|_\infty}|\gamma_{0,1}^{(1)}(x) - y|^2$, i.e.,

$$M - 1 \leqslant 4\varepsilon^{2\|\nabla b_1\|_\infty}|\gamma_{0,1}^{(1)}(x) - y|^2 < M. \tag{4.2.26}$$

Importantly, we recall from (4.1.4) that under Assumptions 4.1.1 and 4.1.2, $\|\nabla b_1\|_\infty \leqslant C(\lambda_1)$. Let

$$t_j := j/M, \quad j = 0, 1, \cdots, M.$$

The important point for the proof is the following claim.

Claim: Set $\xi_0 := x$ and $\xi_M := y$. There exist $(M+1)$-points $\xi_0, \xi_1, \cdots, \xi_M$ such that

$$|\xi_{j+1} - \gamma^{(1)}_{t_j,t_{j+1}}(\xi_j)| \leqslant \tfrac{1}{2\sqrt{M}}, \ j = 0, 1, \cdots, M-1.$$

Indeed, let $Q_1 := B_{1/(2\sqrt{M})}(\gamma^{(1)}_{0,t_1}(x))$ and recursively define for $j = 2, \cdots, M$,

$$Q_j := \bigcup_{z \in Q_{j-1}} B_{1/(2\sqrt{M})}(\gamma^{(1)}_{t_{j-1},t_j}(z)) = \Big\{ z : \mathrm{dist}\Big(z, \gamma^{(1)}_{t_{j-1},t_j}(Q_{j-1})\Big) \leqslant 1/(2\sqrt{M}) \Big\}.$$

Letting $\lambda := \|\nabla b_1\|_\infty$ and noting that (see (4.1.7))

$$e^{-\lambda/M}|z - z'| \leqslant |\gamma^{(1)}_{t_j,t_{j+1}}(z) - \gamma^{(1)}_{t_j,t_{j+1}}(z')| \leqslant e^{\lambda/M}|z - z'|,$$

by the previous induction method and noting that $\gamma^{(1)}_{t_j,t_{j+1}} \circ \gamma^{(1)}_{0,t_j}(x) = \gamma^{(1)}_{0,t_{j+1}}(x)$, we have

$$B_{j\varepsilon^{-(j-1)\lambda/M}/(2\sqrt{M})}(\gamma^{(1)}_{0,t_j}(x)) \subset Q_j, \quad j = 1, 2, \cdots, M.$$

Intuitively, the image of a ball with radius r under the flow $\gamma^{(1)}_{t_{j-1},t_j}$ contains a ball with radius $\varepsilon^{-\lambda/M} r$. In particular, by (4.2.26),

$$\xi_M = y \in B_{\sqrt{M}\varepsilon^{-\lambda/2}}(\gamma^{(1)}_{0,1}(x)) \subset B_{M\varepsilon^{-(M-1)\lambda/M}/(2\sqrt{M})}(\gamma^{(1)}_{0,t_M}(x)) \subset Q_M.$$

The claim then follows. The idea of the construction is illustrated in Figure 4.1.

Now let $\kappa := 1/(2(\varepsilon^{\|\nabla b_1\|_\infty} + 1))$ and $z_0 := x$, $z_{M+1} := y$ and $\Sigma_j := B_{\kappa/\sqrt{M}}(\xi_j)$. From the previous claim, we have that for $z_j \in \Sigma_j$ and $z_{j+1} \in \Sigma_{j+1}$,

$$\begin{aligned}
|\gamma^{(1)}_{t_j,t_{j+1}}(z_j) - z_{j+1}| &\leqslant |\gamma^{(1)}_{t_j,t_{j+1}}(z_j) - \gamma^{(1)}_{t_j,t_{j+1}}(\xi_j)| + |\gamma^{(1)}_{t_j,t_{j+1}}(\xi_j) - \xi_{j+1}| + |\xi_{j+1} - z_{j+1}| \\
&\leqslant \varepsilon^{\|\nabla b_1\|_\infty}|z_j - \xi_j| + |\gamma^{(1)}_{t_j,t_{j+1}}(\xi_j) - \xi_{j+1}| + |\xi_{j+1} - z_{j+1}| \\
&\leqslant \frac{\kappa(\varepsilon^{\|\nabla b_1\|_\infty} + 1)}{\sqrt{M}} + \frac{1}{2\sqrt{M}} = \frac{1}{\sqrt{M}} = \sqrt{t_{j+1} - t_j}.
\end{aligned}$$

This precisely means that the previous diagonal lower bound holds for $p(t_j, z_j, t_{j+1}, z_{j+1})$. Thus, by the Chapman-Kolmogorov equation and (4.2.25), we have

$$\begin{aligned}
p(0, x; 1, y) &= \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} p(t_0, z_0; t_1, z_1) \cdots p(t_{M-1}, z_{M-1}; t_M, z_M) dz_1 \cdots dz_{M-1} \\
&\geqslant \int_{\Sigma_1} \cdots \int_{\Sigma_{M-1}} p(t_0, z_0; t_1, z_1) \cdots p(t_{M-1}, z_{M-1}; t_M, z_M) dz_1 \cdots dz_{M-1} \\
&\geqslant (C_3 M^{d/2})^M \int_{\Sigma_1} \cdots \int_{\Sigma_{M-1}} dz_1 \cdots dz_{M-1} = (C_3 M^{d/2})^M (M^{-d/2}\kappa^d |B_1|)^{M-1} \\
&= C_3^M M^{d/2} (\kappa^d |B_1|)^{M-1} = M^{d/2} \exp\{M \ln(C_3 \kappa^d |B_1|)\}/(\kappa^d |B_1|) \\
&\geqslant C_4 \exp\{M \log(C_3 \kappa^d |B_1|)\} \geqslant C_5 \exp\{-C_6 |\gamma^{(1)}_{0,1}(x) - y|^2\},
\end{aligned}$$

recalling the definition of M in (4.2.26) and that $C_3 \kappa^d |B_1| \leqslant 1$ for the last inequality. $\square$

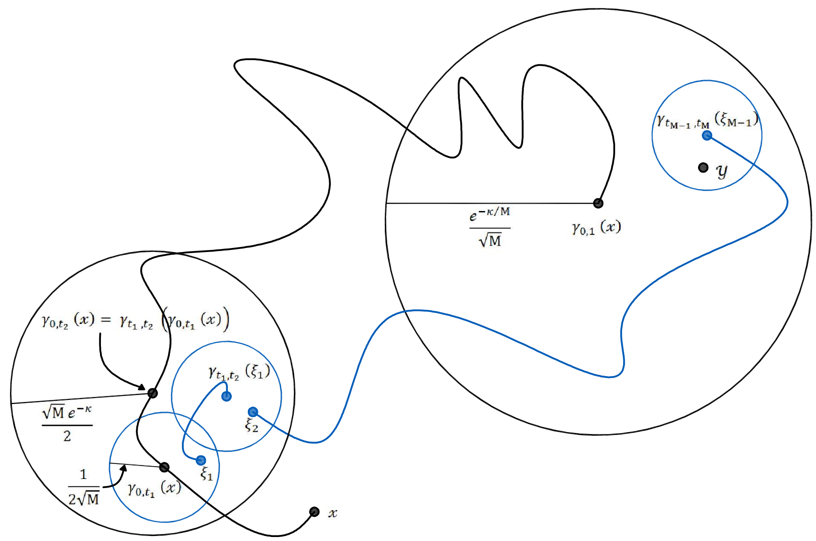

Figure 4.1: Construction of the chaining balls for the lower bound.

4.2.3 Estimates for the derivatives of the heat kernel with smooth coefficients

We insisted in the previous section on the fact that, even though we considered smooth coefficients, all our estimates for the two-sided Gaussian bounds were actually uniform w.r.t. Θ which only depends on parameters appearing in Assumptions 4.1.1 and 4.1.2.

Our point of view is here different since we mainly want to derive some *a priori bounds* on the derivatives of the heat-kernel when the coefficients are smooth which will then serve in a second time, namely in the circular argument developed in Section 4.3, to prove that those estimates are actually again independent of the smoothness of the coefficients. Anyhow, in the current section, we fully exploit such a smoothness and obtain controls on the derivatives which *do* depend on the derivatives of b, σ. To this end, we restart from the representation (4.2.18) of the density and exploit the gradient estimate (4.2.3).

Proof of the main estimates

Theorem 4.2.12 (Controls on the derivatives of the heat kernel with smooth coefficients)**.** *Under* **(S)**, *for* $j \in \{1, 2\}$, *there exist constants* $C_j := C_j\big(\Theta, \lambda_j\big)$, $\mu_j := \mu_j(\Theta) > 0$, *such that, for every* $x, y \in \mathbb{R}^d$ *and* $0 \leqslant t < s \leqslant T$,

$$|\nabla_x^j p(t,x;s,y)| \leqslant C_j (s-t)^{-\frac{j}{2}} \Gamma_{\mu_j}(t,x;s,y), \quad |\nabla_y p(t,x;s,y)| \leqslant C_1 (s-t)^{-\frac{1}{2}} \Gamma_{\mu_1}(t,x;s,y).$$

Proof. (i) Let us first establish the estimates on the derivatives w.r.t. the backward variable x. Write from (4.2.18):

$$\nabla_x^j p(t,x;s,y) = \nabla_x^j \widetilde{p}_1(t,x;s,y) + \nabla_x^j \big(p \otimes H\big)(t,x;s,y).$$

From Lemma 4.2.4 it readily follows that

$$|\nabla_x^j Z_1(t,x;s,y)| \leqslant C_j (s-t)^{-\frac{j}{2}} \Gamma_{\mu_j}(t,x;s,y).$$

For the other contribution, setting $u := \frac{s+t}{2}$,

$$p \otimes H(t,x;s,y) = \int_u^s E\left[H(r, X_{t,r}(x); s,y)\right] dr + \int_t^u E\left[H(r, X_{t,r}(x); s,y)\right] dr =: I_1(x) + I_2(x).$$

Consider I_1 first: choosing $p>1$ such that $\frac{d+\alpha-2}{2p} > \frac{d}{2} - 1$, by (4.2.3) and Lemmas 4.2.6 and 4.2.5, we get

$$\begin{aligned}
|\nabla_x^j I_1(x)| &\leqslant C_j \int_u^s (r-t)^{-j/2} E\left[|H(r, X_{t,r}(x); s,y)|^p\right]^{1/p} dr \qquad (4.2.27)\\
&= C_j \int_u^s (r-t)^{-j/2} \left(\int_{\mathbb{R}^d} p(t,x,r,z) |H(r,z;s,y)|^p dz\right)^{1/p} dr\\
&\leqslant C_j' (s-t)^{-j/2} \int_u^s (s-r)^{-\frac{1}{p} + \frac{\alpha}{2p} + \frac{d}{2p} - \frac{d}{2}} \left(\int_{\mathbb{R}^d} \Gamma_{\mu_0}(t,x;r,z) \Gamma_{\frac{\mu_1}{p}}(r,z;s,y) dz\right)^{1/p} dr\\
&\leqslant C_j'' (s-t)^{-j/2} \left(\int_u^s (s-r)^{-\frac{1}{p} + \frac{\alpha}{2p} + \frac{d}{2p} - \frac{d}{2}} dr\right) \Gamma_{\mu_2}^{1/p}(t,x;s,y)\\
&\leqslant C_j''' (s-t)^{-j/2} (s-t)^{1-\frac{1}{p} + \frac{\alpha}{2p} + \frac{d}{2p} - \frac{d}{2}} \Gamma_{\mu_2}^{1/p}(t,x;s,y)\\
&\leqslant \widetilde{C}_j (s-t)^{-j/2} \Gamma_{\mu_2 p}(t,x;s,y).
\end{aligned}$$

To treat $I_2(x)$, we only consider $j=1$ since the case $j=2$ is similar. By the chain rule, we have

$$\nabla_x E\left[H(r, X_{r,s}(x); t,y)\right] = E\left[\left(\nabla_x H\right)(r, X_{t,r}(x); s,y) \cdot \nabla_x X_{t,r}(x)\right],$$

and for all $k \in \{1, \cdots, d\}$,

$$\begin{aligned}
\partial_{x_k} H(t,x;s,y) := \operatorname{tr}\left(\partial_{x_k} a(t,x) \cdot \nabla_x^2 Z_1(t,x;s,y)\right) + \partial_{x_k} b(t,x) \cdot \nabla_x Z_1(t,x;s,y)\\
+ \operatorname{tr}(a(t,x) - a(t, \gamma_{s,t}(y))) \cdot \partial_{x_k} \nabla_x^2 Z_1(t,x;s,y)\\
+ (b(t,x) - b(t, \gamma_{s,t}(y))) \cdot \partial_{x_k} \nabla_x Z_1(t,x;s,y).
\end{aligned}$$

Thus by Lemma 4.2.4, (4.2.1) and (4.2.21), it is easy to see that for some $\mu_3 > 0$,

$$|\nabla_x H(t,x;s,y)| \leqslant C(s-t)^{-1} \Gamma_{\mu_3}(t,x;s,y).$$

We carefully emphasize that the constants denoted by C above, **do** depend on the smoothness of the coefficients. Using the same argument as above, from the Hölder inequality, one sees that for $p = \frac{d}{d-1}$,

$$\begin{aligned} |\nabla_x I_2(x)| &\leqslant C_1 \int_t^u E\left[|(\nabla_x H)(r, X_{t,r}(x); s, y)|^p\right]^{1/p} dr \\ &\leqslant C_1' \int_t^u (s-r)^{-1} E\left[\Gamma_{\mu_3}^p(r, X_{t,r}(x); s, y)\right]^{1/p} dr \leqslant (s-t)^{-\frac{1}{2}} \Gamma_{\mu_4}(t, x; s, y). \end{aligned}$$

We thus obtain the gradient estimate in the variable x.

(ii) Let us now turn to the gradient estimate w.r.t. y. We restart from (4.2.16) differentiating first w.r.t. y. This can be done for arbitrary freezing parameters (t_0, x_0). Write:

$$\begin{aligned} \nabla_y p(t, x; s, y) &= \nabla_y \widetilde{p}^{t_0,x_0}(t, x; s, y) + \int_t^s \int_{\mathbb{R}^d} p(t, x; r, z)(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{t_0,x_0}) \nabla_y \widetilde{p}^{t_0,x_0}(r, z; s, y) dz dr \\ &= -\nabla_x \widetilde{p}^{t_0,x_0}(t, x; s, y) - \int_t^s \int_{\mathbb{R}^d} p(t, x; r, z)(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{t_0,x_0}) \nabla_z \widetilde{p}^{t_0,x_0}(r, z; s, y) dz dr, \end{aligned} \tag{4.2.28}$$

where we have used the explicit expression (4.2.7) for the second equality. Letting again $u = \frac{s+t}{2}$ and taking $(t_0, x_0) = (s, y)$, we can split

$$\nabla_y p(t, x; s, y) = -\nabla_x Z_1(t, x; s, y) - J_1(y) - J_2(y),$$

where

$$\begin{aligned} J_1(y) &:= \int_t^u \int_{\mathbb{R}^d} p(t, x; r, z)(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{s,y}) \nabla_z Z_1(r, z; s, y) dz dr, \\ J_2(y) &:= \int_u^s \int_{\mathbb{R}^d} p(t, x; r, z)(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{s,y}) \nabla_z Z_1(r, z; s, y) dz dr. \end{aligned}$$

For $J_1(y)$, from the Gaussian upper-bound of Theorem 4.2.11, (4.2.21) and Lemma 4.2.5 (see also Lemma 4.2.6), we have

$$|J_1(y)| \leqslant C_1 \int_t^u (s-r)^{-3/2} \int_{\mathbb{R}^d} \Gamma_{\mu_0}(t, x; r, z) \Gamma_{\mu_5}(r, z; s, y) dz dr \leqslant C_1'(s-t)^{-1/2} \Gamma_{\mu_6}(t, x; s, y).$$

Consider now J_2: integrating by parts and recalling (4.2.13) and (4.2.14), we have

$$\begin{aligned} |J_2(y)| \leqslant& \int_u^s \left| \int_{\mathbb{R}^d} \nabla_z p(t, x; r, z)(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{s,y}) Z_1(r, z; s, y) dz \right| dr \\ &+ \int_u^s \left| \int_{\mathbb{R}^d} p(t, x; r, z) \nabla_z b \cdot \nabla_z Z_1(r, z; s, y) dz \right| dr \\ &+ \int_u^s \left| \int_{\mathbb{R}^d} p(t, x; r, z) \nabla_z a \cdot \nabla_z^2 Z_1(r, z; s, y) dz \right| dr \\ =:& J_{21}(y) + J_{22}(y) + J_{23}(y). \end{aligned}$$

For $J_{21}(y)$, recalling from (4.2.17) that $(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}^{s,y}_{r,z})Z_1(r,z;s,y) = H(r,z;s,y)$, we derive from (4.2.3) and as in (4.2.27) that

$$J_{21}(y) = \int_u^s \left| \int_{\mathbb{R}^d} p(t,x;r,z)\nabla_z H(r,z;s,y)dz \right| dr = \int_u^s \left| E\left[(\nabla_z H)(r, X_{r,s}(x); s,y) \right] \right| dr$$
$$\leqslant C_1 \int_u^s (r-t)^{-1/2} E\left[|H(r, X_{r,s}(x); s,y)|^p \right]^{1/p} dr \leqslant (s-t)^{-1/2}\Gamma_{\mu_7}(t,x;s,y).$$

For $J_{22}(y)$, from the upper bound in Theorem 4.2.11 and (4.2.21), we have

$$J_{22}(y) \leqslant C_1 \int_u^t (s-r)^{-1/2} \int_{\mathbb{R}^d} \Gamma_{\mu_0}(t,x;r,z)\Gamma_{\mu_1}(r,z;s,y)dzdr \leqslant C_1'(s-t)^{-1/2}\Gamma_{\mu_8}(t,x;s,y).$$

For J_{23}, since $|\nabla_z^2 \widetilde{p}_1(r,z;s,y)|$ has the singularity $(s-r)^{-1}$, noting that

$$\nabla_z a \cdot \nabla_z^2 Z_1 = \nabla_z^2(\nabla_z a \cdot Z_1) - \nabla_z^3 a \cdot Z_1 - \nabla_z^2 a \cdot \nabla_z Z_1,$$

as above, by (4.2.3) we still have

$$J_{23}(y) \leqslant C(s-t)^{-1/2}\Gamma_{\mu_9}(t,x;s,y).$$

Combining the above estimates, we obtain the derivative estimate in y. The proof is complete. □

Remark 4.2.13. *We point out that Theorem 4.2.12 anyhow has some interest by itself. A careful reading of the proof shows that actually the statements about the derivatives w.r.t. x hold true if additionally to Assumption 4.1.1 and 4.1.2, the coefficients b, σ are twice continuously differentiable with bounded derivatives and that the second order derivatives are themselves Hölder continuous. In this framework, the Duhamel representation* (4.2.18) *coupled to the heat-kernel estimates of Theorem 4.2.11 provides an alternative approach to the full Malliavin calculus viewpoint developed in [26].*

4.3 Proof of Main Theorem

In the following proof, the final time horizon $T > 0$ is fixed. We first work under the assumptions **(S)** aiming at obtaining constants in the estimates of Section 4.2.3 that only depend on $\Theta := (T, \alpha, \beta, \lambda_1, \lambda_2, d)$ introduced in (4.1.8).

With the same reasoning as for Section 2.6.1, we introduce for $\delta > 0$ the SDE (4.2.2) with diffusion coefficient $\sigma(t,x) = \delta\mathbb{I}_{d\times d}$ and denote by $\bar{p}_\delta$ the corresponding density. By the lower bound estimate proven in Theorem 4.2.11 and scaling techniques similar to those presented in Lemma 4.2.9, it holds that for any $\mu > 0$, there exists $\delta := \delta(\mu)$ large enough and $\bar{C}_\delta > 0, \mu'$ depending on $\bar{\Theta} = (T, \beta, \delta, \lambda_2, d)$ such that for all $0 \leqslant t < s \leqslant T$ and $x, y \in \mathbb{R}^d$,

$$\bar{C}_\delta^{-1}\Gamma_\mu(t,x;s,y) \leqslant \bar{p}_\delta(t,x;s,y) \leqslant \bar{C}_\delta\Gamma_{\mu'}(t,x;s,y). \tag{4.3.1}$$

Importantly, with the notations of Section 4.2.1, we choose μ, and then $\delta := \delta(\mu)$ s.t. for all $\theta \in [0,1]$, $0 \leqslant t < s \leqslant T$, $x, y \in \mathbb{R}^d$ and $j \in \{0,1,2\}$,

$$|\gamma_{t,s}(x) - y|^\theta |\nabla_y^j Z_0(t,x;s,y)| + |x - \gamma_{s,t}(y)|^\theta |\nabla_x^j Z_1(t,x;s,y)| \leqslant C_\delta (s-t)^{\frac{\theta}{2} - \frac{j}{2}} \bar{p}_\delta(t,x;s,y), \quad (4.3.2)$$

where C_δ here only depends on Θ and δ, θ.

Without further declaration, we shall fix from now on a δ such that (4.3.2) holds. From the definition of H in (4.2.17) and the proof of Lemma 4.2.6, we also derive from this choice of δ that, under the sole Assumptions 4.1.1 and 4.1.2, there exists $C := C(\Theta)$ such that, for all $0 \leqslant t < s \leqslant T$, $x, y \in \mathbb{R}^d$:

$$|H(t,x;s,y)| \leqslant C(s-t)^{-1+\frac{\alpha}{2}} \bar{p}_\delta(t,x;s,y). \quad (4.3.3)$$

For simplicity we will write from now on $\bar{p} = \bar{p}_\delta$. In particular, for all $0 \leqslant t < s \leqslant T$, $x, y \in \mathbb{R}^d$, $r \in [t,s]$:

$$\int_{\mathbb{R}^d} \bar{p}(t,x;r,z)\bar{p}(r,z;s,y)dz = \bar{p}(t,x;s,y). \quad (4.3.4)$$

For the rest of the section, we use the convention that all the constants appearing below only depend on Θ. Again, we have shown in the previous section that for smooth coefficients the expected bounds for the derivatives hold. The constants in Theorem 4.2.12 however do depend on the derivatives of the coefficients, since we use the gradient estimate (4.2.3). We aim here at proving that we can obtain the same type of estimates as in Theorem 4.2.12 under Assumptions 4.1.1, 4.1.2 and (4.2.1) but for constants that only depend on Θ. This is the purpose of Sections 4.3.1 to 4.3.3. We will then eventually derive in Section 4.3.4 the main results of Theorem 4.1.5 thanks to some compactness arguments (Ascoli-Arzelà theorem) to the uniformity of the controls obtained for mollified parameters.

4.3.1 First order derivative estimates in the backward variable x

Without loss of generality we shall assume $t = 0$ and for $s \in (0,T]$, we define

$$f_1(s) := \sup_{x,y} |\nabla_x p(0,x;s,y)|/\bar{p}(0,x;s,y).$$

From Theorem 4.2.12 and (4.3.1), we know that

$$\int_0^T f_1(s)ds < \infty.$$

By the forward representation formula (4.2.18), we have

$$|\nabla_x p(0,x;s,y)| \leqslant |\nabla_x Z_1(0,x;s,y)| + |\nabla_x p| \otimes |H|(0,x;s,y).$$

Observe first that, from Lemma 4.2.4 and (4.3.2)

$$|\nabla_x Z_1(0,x;s,y)| \leqslant C_1 t^{-1/2}\Gamma_\lambda(0,x;s,y) \leqslant C_1' t^{-1/2}\bar{p}(0,x;s,y).$$

Secondly, (4.3.3) yields

$$\begin{aligned}
|\nabla_x p| \otimes |H|(0,x;s,y) &\leqslant \int_0^s \int_{\mathbb{R}^d} f_1(r)\bar{p}(0,x;r,z)|H(r,z;s,y)|dzdr \\
&\leqslant C_1 \int_0^s f_1(r)(s-r)^{-1+\frac{\alpha}{2}} \int_{\mathbb{R}^d} \bar{p}(0,x;r,z)\bar{p}(r,z;s,y)dzdr \\
&= C_1 \left(\int_0^s f_1(r)(s-r)^{-1+\frac{\alpha}{2}} dr \right) \bar{p}(0,x;s,y),
\end{aligned}$$

using also (4.3.4) for the last identity. Thus,

$$f_1(t) \leqslant C_1 \left(s^{-\frac{1}{2}} + \int_0^s (s-r)^{-1+\frac{\alpha}{2}} f_1(r)dr \right).$$

By the Volterra type Gronwall inequality, we obtain

$$f_1(t) \leqslant C_1' s^{-\frac{1}{2}} \Rightarrow |\nabla_x p(0,x;s,y)| \leqslant C_1' s^{-\frac{1}{2}} \bar{p}(0,x;s,y). \tag{4.3.5}$$

4.3.2 Second order derivative estimates in the backward variable x

We assume for this section that Assumption 4.1.2 holds for some $\beta \in (0,1]$. It is crucial to take here $\beta > 0$. Below we fix $s \in (0,T]$ and define for $t < s$

$$f_2(t) := (s-t) \cdot \sup_{x,y} |\nabla_x^2 p(t,x;s,y)|/\bar{p}(t,x;s,y). \tag{4.3.6}$$

By Theorem 4.2.12 and (4.3.1), we have

$$\sup_{t \leqslant s} f_2(t) < \infty.$$

To derive the estimate of the second order derivative of the density, we use the backward Duhamel representation (4.2.11). And for fixed freezing parameters (t_0, x_0) we differentiate twice w.r.t. x to derive:

$$\begin{aligned}
\nabla_x^2 p(t,x;s,y) &= \nabla_x^2 \widetilde{p}^{t_0,x_0}(t,x;s,y) + \int_t^s \int_{\mathbb{R}^d} \nabla_x^2 \widetilde{p}^{t_0,x_0}(t,x;r,z) \left(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{t_0,x_0} \right) p(r,z;s,y)dzdr \\
&= \nabla_y^2 \widetilde{p}^{t_0,x_0}(t,x;s,y) + \int_t^s \int_{\mathbb{R}^d} \nabla_z^2 \widetilde{p}^{t_0,x_0}(t,x;r,z) \left(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{t_0,x_0} \right) p(r,z;s,y)dzdr,
\end{aligned} \tag{4.3.7}$$

using again the explicit expression (4.2.7) for the second equality. Let us now denote for a parameter $\varepsilon > 0$ that might depend on r to be specified later on,

$$A^{\varepsilon,t_0,x_0}_{r,z} := a_\varepsilon(r,z) - a_\varepsilon(r,\gamma_{t_0,r}(x_0)), \quad \bar{A}^{t_0,x_0}_{r,z} := A^{t_0,x_0}_{r,z} - A^{\varepsilon,t_0,x_0}_{r,z}, \tag{4.3.8}$$

where similarly to (4.1.3), $a_\varepsilon(r,z) = a(r,\cdot) * \rho_\varepsilon(z)$. Choosing the freezing point $(t_0,x_0) = (t,x)$ and setting as well

$$Z_0(t,x;s,y) = \widetilde{p}^{t,x}(t,x;s,y), \quad u := (t+s)/2,$$

we decompose the expression in (4.3.7) as follows:

$$\nabla^2_x p(t,x,s,y) =: \sum_{i=1}^{5} I_i(t,x,s,y), \tag{4.3.9}$$

where $I_1(t,x;s,y) := \nabla^2_y Z_0(t,x;s,y)$ and

$$\begin{aligned}
I_2(t,x;s,y) &:= \int_s^u \int_{\mathbb{R}^d} \nabla^2_z Z_0(t,x;r,z)\mathrm{tr}(A^{t,x}_{r,z} \cdot \nabla^2_z p(r,z;s,y))dzdr\\
I_3(t,x;s,y) &:= \int_u^s \int_{\mathbb{R}^d} \nabla^2_z Z_0(t,x;r,z)\mathrm{tr}(A^{\varepsilon,t,x}_{r,z} \cdot \nabla^2_z p(r,z;s,y))dzdr\\
I_4(t,x;s,y) &:= \int_u^s \int_{\mathbb{R}^d} \nabla^2_z Z_0(t,x;r,z)\mathrm{tr}(\bar{A}^{\varepsilon,t,x}_{r,z} \cdot \nabla^2_z p(r,z;s,y))dzdr\\
I_5(t,x;s,y) &:= \int_t^s \int_{\mathbb{R}^d} \nabla^2_z Z_0(t,x;r,z)B^{t,x}_{r,z} \cdot \nabla_z p(r,z;s,y)dzdr.
\end{aligned}$$

By Lemma 4.2.3, (4.2.15) and (4.3.2), it is easy to see that

$$|I_1(t,x;s,y)| \leqslant C(s-t)^{-1}\Gamma^{\mathrm{heat}}(\mu\mathcal{I}_{s-t}, \gamma_{t,s}(x)-y) \leqslant C(s-t)^{-1}\bar{p}(t,x;s,y).$$

For I_2, by Assumption 4.1.1 and again (4.3.2), we have

$$\begin{aligned}
|I_2(t,x;s,y)| &\leqslant C\int_t^u \int_{\mathbb{R}^d} \frac{\Gamma^{\mathrm{heat}}(\mu\mathcal{I}_{r-t}, \gamma_{t,r}(x)-z)}{r-t}|z-\theta_{t,r}(x)|^\alpha|\nabla^2_x p(r,z;s,y)|dzdr\\
&\leqslant C\int_t^u \frac{(r-t)^{\frac{\alpha}{2}} f_2(r)}{(r-t)(s-r)}\int_{\mathbb{R}^d} \bar{p}(t,x;r,z)\bar{p}(r,z;s,y)dzdr\\
&\leqslant (s-t)^{-1}\bar{p}(t,x;s,y)\int_t^s (r-t)^{-1+\frac{\alpha}{2}} f_2(r)dr.
\end{aligned}$$

For I_3, integrating by parts, we have

$$\begin{aligned}
|I_3(t,x;s,y)| \leqslant C&\int_u^s \int_{\mathbb{R}^d} \left|\nabla^3_z Z_0(t,x;r,z)\right| \cdot |A^{\varepsilon,t,x}_{r,z}| \cdot |\nabla_z p(r,z;s,y)|dzdr\\
&+\int_u^s \int_{\mathbb{R}^d} \left|\nabla^2_z Z_0(t,x;r,z)\right| \cdot |\nabla_z A^{\varepsilon,t,x}_{r,z}| \cdot |\nabla_z p(r,z;s,y)|dzdr.
\end{aligned}$$

Note that by the property of convolutions,

$$|\nabla_z A^{\varepsilon,t,x}_{r,z}| \leqslant C\varepsilon^{-1+\alpha}, \quad |A^{\varepsilon,t,x}_{r,z}| \leqslant C|z-\gamma_{t,r}(x)|^\alpha, \quad |\bar{A}^{\varepsilon,t,x}_{r,z}| \leqslant C\varepsilon^\alpha.$$

In particular, taking $\varepsilon = (s-r)^{\frac{1}{2}}$, by Lemma 4.2.3, (4.2.15), (4.3.2) and using as well the bound (4.3.5) on the gradient established in the previous section, we obtain

$$\begin{aligned}
|I_3(t,x;s,y)| &\leqslant C\int_u^s\int_{\mathbb{R}^d} \frac{\bar{p}(t,x;r,z)}{(r-t)^{\frac{3}{2}}}\cdot(r-t)^{\frac{\alpha}{2}}\cdot\frac{\bar{p}(r,z;s,y)}{(s-r)^{\frac{1}{2}}}dzdr \\
&\quad + \int_u^s\int_{\mathbb{R}^d} \frac{\bar{p}(t,x;r,z)}{r-t}\cdot(s-r)^{\frac{\alpha}{2}}\cdot\frac{\bar{p}(r,z;s,y)}{s-r}dzdr \\
&\leqslant C\bar{p}(t,x;s,y)\left(\int_u^s (r-t)^{-\frac{3-\alpha}{2}}(s-r)^{-\frac{1}{2}}dr + \int_u^s (r-t)^{-1}(s-r)^{-1+\frac{\alpha}{2}}\right)dr \\
&\leqslant C\bar{p}(t,x;s,y)(s-t)^{-1+\frac{\alpha}{2}},
\end{aligned}$$

and

$$\begin{aligned}
|I_4(t,x;s,y)| &\leqslant C\int_u^s\int_{\mathbb{R}^d} \frac{\bar{p}(t,x;r,z)}{r-t}\cdot(s-r)^{\frac{\alpha}{2}}\cdot\frac{f_2(r)\bar{p}(r,z;s,y)}{s-r}dzdr \\
&\leqslant C\bar{p}(t,x;s,y)(s-t)^{-1}\int_u^s (s-r)^{-1+\frac{\alpha}{2}}f_2(r)dr.
\end{aligned}$$

For I_5, from (4.3.2), we derive similarly to I_2 that

$$\begin{aligned}
|I_5(t,x;s,y)| &\leqslant C\int_t^s\int_{\mathbb{R}^d} \frac{\Gamma^{\text{heat}}(\mu\mathcal{I}_{r-t},\gamma_{t,r}(x)-z)}{r-t}(|z-\gamma_{t,r}(x)|^\beta + |z-\theta_{t,r}(x)|)\frac{\bar{p}(r,z;s,y)}{(s-r)^{\frac{1}{2}}}dzdr \\
&\leqslant C\bar{p}(t,x;s,y)\int_t^s \frac{(r-t)^{\frac{\beta}{2}}+(r-t)^{\frac{1}{2}}}{(r-t)(s-r)^{\frac{1}{2}}}dr \leqslant C\bar{p}(t,x;s,y)(s-t)^{-1}.
\end{aligned}$$

Combining the above estimates for the $(I_j)_{j\in\{1,\cdots,5\}}$, we obtain from (4.3.9) and (4.3.6) that:

$$f_2(t) \leqslant C\left(1+\int_t^s (r-t)^{-1+\frac{\alpha}{2}}f_2(r)dr + \int_t^s (s-r)^{-1+\frac{\alpha}{2}}f_2(r)dr\right).$$

Finally, from the Volterra type Gronwall inequality, we obtain

$$\sup_{t\in[0,s]} f_2(t) \leqslant C \Rightarrow |\nabla_x^2 p(t,x;s,y)| \leqslant C(s-t)^{-1}\bar{p}(t,x;s,y). \tag{4.3.10}$$

4.3.3 First order derivative estimate in y

We assume for this section that Assumption 4.1.2 holds for some $\beta > 0$ and that the diffusion coefficient $\sigma \in bC^{1,\alpha}_{0,T}$. Fix $t > 0$. For $s \in (t,T]$, we define

$$f_3(s) := \sup_{x,y} |\nabla_y p(t,x;s,y)|/\bar{p}(t,x;s,y). \tag{4.3.11}$$

By Theorem 4.2.12 and (4.3.1) we know that

$$\int_t^T f_3(s)ds < \infty.$$

In (4.2.28), taking $(t_0, x_0) = (s, y)$ and recalling the notations of (4.2.14) and $Z_1(t, x; s, y) = \widetilde{p}^{s,y}(t, x; s, y)$, by the integration by parts, we have

$$\begin{aligned}\nabla_y p(t, x; s, y) = &-\nabla_x Z_1(t, x; s, y) + \int_t^s \int_{\mathbb{R}^d} \nabla_z p(t, x; r, z) \mathrm{tr}(A^{s,y}_{r,z} \cdot \nabla^2_z Z_1)(r, z; s, y) dz dr \\ &+ \int_s^t \int_{\mathbb{R}^d} p(t, x; r, z) \mathrm{tr}((\nabla_z a)(r, z) \cdot \nabla^2_z Z_1)(r, z; s, y) dz dr \\ &- \int_s^t \int_{\mathbb{R}^d} p(t, x; r, z) B^{s,y}_{r,z} \cdot \nabla^2_z Z_1(r, z; s, y) dz dr =: \sum_{i=1}^4 J_i(t, x; s, y). \end{aligned} \tag{4.3.12}$$

For J_1, we readily get from (4.3.2)

$$|J_1(s, x; t, y)| \leqslant C(s-t)^{-\frac{1}{2}} \bar{p}(t, x; s, y).$$

For J_2, using again (4.3.2) and (4.3.11) gives:

$$\begin{aligned}|J_2(t, x; s, y)| &\leqslant C \int_t^s \int_{\mathbb{R}^d} |\nabla_z p(t, x; r, z)| \cdot (s-r)^{-1+\frac{\alpha}{2}} \bar{p}(r, z; s, y) dz dr \\ &\leqslant C \int_t^s f_3(r) \int_{\mathbb{R}^d} \bar{p}(t, x; r, z) \cdot (s-r)^{-1+\frac{\alpha}{2}} \bar{p}(r, z; s, y) dz dr \\ &\leqslant C \bar{p}(t, x; s, y) \int_s^t f_3(r)(s-r)^{-1+\frac{\alpha}{2}} dr.\end{aligned}$$

For J_3, we further write

$$\begin{aligned}J_3(t, x; s, y) = &\int_t^s \int_{\mathbb{R}^d} p(t, x; r, z) \mathrm{tr}\Big(((\nabla_z a)(r, z) - (\nabla_z a)(r, \gamma_{s,r}(y))) \cdot \nabla^2_z Z_1\Big)(r, z; s, y) dz dr \\ &+ \int_t^s \int_{\mathbb{R}^d} p(t, x; r, z) \mathrm{tr}\Big((\nabla_z a)(r, \gamma_{s,r}(y)) \cdot \nabla^2_z Z_1\Big)(r, z; s, y) dz dr \\ =&: J_{31}(t, x, s, y) + J_{32}(t, x; s, y).\end{aligned}$$

For J_{31}, as above, by (4.3.2) we have

$$|J_{31}(t, x; s, y)| \leqslant C \int_t^s \int_{\mathbb{R}^d} \bar{p}(t, x; r, z) \cdot (s-r)^{\frac{\alpha}{2}-1} \bar{p}(r, z; s, y) dz dr \leqslant C \bar{p}(t, x; s, y).$$

For J_{32}, again by the integration by parts, we derive

$$\begin{aligned}|J_{32}(t, x; s, y)| &\leqslant C \int_t^s \int_{\mathbb{R}^d} |\nabla_z p(t, x; r, z)| \cdot |\nabla_z Z_1(r, z; s, y)| dz dr \\ &\leqslant C \int_s^t f_3(r) \int_{\mathbb{R}^d} \bar{p}(t, x; r, z) \cdot (s-r)^{-\frac{1}{2}} \bar{p}(r, z; s, y) dz dr \\ &\leqslant C \bar{p}(t, x; s, y) \int_s^t f_3(r)(s-r)^{-\frac{1}{2}} dr.\end{aligned}$$

Finally, we derive similarly to the term J_{31} that

$$|J_4(t,x;s,y)| \leqslant C\bar{p}(t,x;s,y).$$

Combining all the estimates above for $(J_i)_{i\in\{1,\cdots,4\}}$, from (4.3.12) and (4.3.11) we get

$$f_3(t) \leqslant C\left((s-t)^{-\frac{1}{2}} + \int_t^s f_3(r)(s-r)^{-1+\frac{\alpha}{2}}dr\right),$$

which in turn yields

$$f_3(t) \leqslant C(s-t)^{-\frac{1}{2}} \Rightarrow |\nabla_y p(t,x;s,y)| \leqslant C(s-t)^{-\frac{1}{2}}\bar{p}(t,x;s,y). \tag{4.3.13}$$

4.3.4 Proof of Theorem 4.1.5

Now we go back the notations of Section 2 and keep the index ε, associated with the spatial mollification of the coefficients. Thus, let p_ε be the corresponding heat kernel and $X^\varepsilon_{t,s}(x)$ the solution of SDE (4.2.2) corresponding to b_ε and σ_ε. It is well known, see e.g. Theorem 11.1.4 in [67], that under Assumptions 4.1.1 and 4.1.2, for any $f \in bC^\infty(\mathbb{R}^d)$

$$\lim_{\varepsilon\to 0} E\left[f(X^\varepsilon_{t,s}(x))\right] = E\left[f(X_{t,s}(x))\right].$$

Moreover, from Theorem 4.2.11 we have the following uniform estimate: there exist constants $\mu_0, C_0 > 0$ depending only on Θ such that for all $\varepsilon \in (0,1)$,

$$C_0^{-1}\Gamma_{\mu_0^{-1}}(t,x;s,y) \leqslant p_\varepsilon(t,x;s,y) \leqslant C_0\Gamma_{\mu_0}(t,x;s,y).$$

Similarly, we derive from (4.3.5), (4.3.10) and (4.3.13) that under Assumptions 4.1.1 and 4.1.2,

$$\sup_\varepsilon |\nabla_x p_\varepsilon(t,x;s,y)| \leqslant C_1(s-t)^{-1/2}\Gamma_{\mu_1}(t,x;s,y), \tag{4.3.14}$$

and under Assumptions 4.1.1 and 4.1.2 with $\beta \in (0,1)$, $j \in \{1,2\}$,

$$\sup_\varepsilon |\nabla^j_x p_\varepsilon(t,x;s,y)| \leqslant C_2(s-t)^{-j/2}\Gamma_{\mu_2}(t,x;s,y), \tag{4.3.15}$$

and under Assumptions 4.1.1 and 4.1.2 with $\beta \in (0,1)$ and $\sigma \in bC^{1,\alpha}_{0,T}$,

$$\sup_\varepsilon |\nabla_y p_\varepsilon(t,x;s,y)| \leqslant C_1'(s-t)^{-1/2}\Gamma_{\mu_1'}(t,x;s,y), \tag{4.3.16}$$

where in the above equations (4.3.14)-(4.3.16) the constants C_1, C_2, C_1' only depend on Θ and *not* on the mollification parameter ε.

In particular, for every non-negative measurable function f, we eventually derive

$$C_0^{-1}\int_{\mathbb{R}^d}\Gamma_{\mu_0^{-1}}(t,x;s,y)f(y)dy \leqslant E\left[f(X_{t,s}(x))\right] \leqslant C_0\int_{\mathbb{R}^d}\Gamma_{\mu_0}(t,x;s,y)f(y)dy,$$

which implies that $X_{t,s}(x)$ has a density $p(t,x;s,y)$ having lower and upper bound as in (4.1.9). This proves point *(i)* of the theorem.

Moreover, for each $t<s$, we now aim at proving that

$$(x,y)\mapsto \nabla_x p_\varepsilon(t,x;s,y) \text{ is equi-continuous on any compact subset of } \mathbb{R}^d\times\mathbb{R}^d, \qquad (\mathbf{C}_1)$$

and

$$(x,y)\mapsto \nabla_x^2 p_\varepsilon(t,x;s,y) \text{ is equi-continuous on any compact subset of } \mathbb{R}^d\times\mathbb{R}^d, \qquad (\mathbf{C}_2)$$

$$(x,y)\mapsto \nabla_y p_\varepsilon(t,x;s,y) \text{ is equi-continuous on any compact subset of } \mathbb{R}^d\times\mathbb{R}^d. \qquad (\mathbf{C}_3)$$

Assume for a while that such a continuity condition holds. Then, from the Ascoli-Arzelà theorem, one can find a subsequence ε_k such that for each $x,y\in\mathbb{R}^d$,

$$\nabla_x^j p_{\varepsilon_k}(t,x;s,y)\to\nabla_x^j p(t,x;s,y), \quad j=0,1,2,\ \nabla_y p_{\varepsilon_k}(t,x;s,y)\to\nabla_y p(t,x;s,y).$$

The gradient and second order derivative estimates follow, under the previously recalled additional assumptions when needed, from (4.3.14), (4.3.15) and (4.3.16). This completes the proof of points *(ii)* to *(iv)* of the theorem up to the proof of $(\mathbf{C}_1)$, $(\mathbf{C}_2)$ and $(\mathbf{C}_3)$. This equicontinuity property is proved in Appendix C.

4.4 Extension to higher order derivatives

We explain here how the estimates (4.1.10), (4.1.11), (4.1.12) can be extended for higher order derivatives in our analysis. We claim that under **(S)** the *a-priori* bounds of Theorem 4.2.12 can be obtained for any $j\in\mathbb{N}$, using the same techniques based on the Duhamel representation of the density and (4.2.3). On the other hand the circular arguments used in Section 4.3 can be repeated as well, provided that the coefficients are smooth enough.

For instance, let us assume **(S)** to be in force; assume as well that $\|\nabla\sigma\|_\infty+\|\nabla b\|_\infty<\infty$ and for some $\alpha,\beta\in(0,1]$, $\lambda_3\geqslant 1$,

$$|\nabla\sigma(t,x)-\nabla\sigma(t,y)|\leqslant\lambda_3|x-y|^\alpha, \quad |\nabla b(t,x)-\nabla b(t,y)|\leqslant\lambda_3|x-y|^\beta, \quad x,y\in\mathbb{R}^d. \qquad (4.4.1)$$

We aim here at proving that we can obtain bounds on the third order derivatives which only depend on Assumptions 4.1.1 and 4.1.2 (with Hölder indexes equal to one) and the constants in (4.4.1).

Namely, we want to illustrate a kind of *parabolic bootstrap property*, i.e. in (4.4.1) we give some Hölder conditions on the first derivatives of the coefficients which together with the assumptions 4.1.1, 4.1.2 lead to a uniform control of the third order derivatives.

As in (4.3.7), for the choice of the freezing parameters $(t_0, x_0) = (t, x)$ and recalling $Z_0(t, x; s, y) = \widetilde{p}^{t,x}(t, x; s, y)$, we have the following representation for the derivatives of order three:

$$\nabla_x^3 p(t, x; s, y) = -\nabla_y^3 Z_0(t, x; s, y) - \int_t^s \int_{\mathbb{R}^d} \nabla_z^3 Z_0(t, x; r, z) \left(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{t,x}\right) p(r, z; s, y) dz dr. \quad (4.4.2)$$

Let us now concentrate on the most singular term in (4.4.2). Setting $u = (t + s)/2$ and $A_{r,z}^{t_0,x_0}$, $A_{r,z}^{\varepsilon,t_0,x_0}$, $\bar{A}_{r,z}^{\varepsilon,t_0,x_0}$ as in (4.2.14), (4.3.8) we write

$$\begin{aligned}
&\int_t^s \int_{\mathbb{R}^d} \nabla_z^3 Z_0(t, x; r, z) \mathrm{tr}\left(A_{r,z}^{t,x} \cdot \nabla_z^2 p(r, z; s, y)\right) dz dr \\
&= \int_u^s \int_{\mathbb{R}^d} \nabla_z^3 Z_0(t, x; r, z) \mathrm{tr}\left((A_{r,z}^{\varepsilon,t,x} + \bar{A}_{r,z}^{\varepsilon,t,x}) \cdot \nabla_z^2 p(r, z; s, y)\right) dz dr \\
&\quad + \int_t^u \int_{\mathbb{R}^d} \nabla_z^3 Z_0(t, x; r, z) \mathrm{tr}\left(A_{r,z}^{t,x} \cdot \nabla_z^2 p(r, z; s, y)\right) dz dr =: G_1(t, x; s, y) + G_2(t, x; s, y).
\end{aligned}$$

When $r \in [u, s]$, $(r - t)^{-\frac{3}{2}} \asymp (s - t)^{-\frac{3}{2}}$ is not singular. Therefore we may control G_1 similarly to the terms I_3 and I_4 appearing in Section 4.3.2, owing to the fact that the upper bound on $\nabla_z^2 p$ is already available at this point.

When $r \in [t, u]$, then $(r - t)^{-\frac{3}{2}}$ is indeed singular. Thus, to control G_2 the point is precisely to exploit the regularity of the coefficients and perform an integration by parts to balance the singularity. We write

$$\begin{aligned}
G_2(t, x; s, y) = &- \int_t^u \int_{\mathbb{R}^d} \nabla_z^2 Z_0(t, x; r, z) \mathrm{tr}\left(\nabla_z A_{r,z}^{t,x} \cdot \nabla_z^2 p(r, z; s, y)\right) dz dr \\
&- \int_t^u \int_{\mathbb{R}^d} \nabla_z^2 Z_0(t, x; r, z) \mathrm{tr}\left(A_{r,z}^{t,x} \cdot \nabla_z^3 p(r, z; s, y)\right) dz dr,
\end{aligned}$$

and define

$$f_3(t) := (s - t)^{\frac{3}{2}} \sup_{x,y} |\nabla_x^3 p(t, x; s, y)| / \bar{p}(t, x; s, y);$$

Then, exploiting the uniform bounds for the derivatives of order lower or equal than 2 obtained in Section 4.3, we eventually derive

$$f_3(s) \leqslant C\left(1 + \int_t^u (r - t)^{-1+\frac{\alpha}{2}} f_3(t) dt\right) \Rightarrow \sup_{t \in [0,s]} f_3(s) \leqslant C,$$

which yields the desired estimate for $\nabla_x^3 p$. In the same manner, starting from the Duhamel expansion (4.2.16), and assuming in addition that $\|\nabla^2 \sigma\|_\infty < \infty$ and $|\nabla^2 \sigma(t, x) - \nabla^2 \sigma(t, y)| \leqslant \lambda_4 |x - y|^\alpha$ for some $\alpha \in (0, 1)$ we could derive

$$|\nabla_y^2 p(t, x; s, y)| \leqslant C(s - t)^{-1} \bar{p}(t, x; s, y).$$

A careful reading of the proof suggests that the above arguments may be repeated for any derivative of order $j > 3$ in the backward variable x as soon as we have appropriate regularity assumptions on $\nabla^{j-2}\sigma$ and $\nabla^{j-2}b$. More precisely, assuming that

$$\|\nabla^{j'}\sigma\|_\infty + \|\nabla^{j'}b\|_\infty < \infty,\ j' = 1, \cdots, j-2,$$

and for some $\alpha, \beta \in (0,1]$, $\mu_{j-2} \geqslant 1$,

$$|\nabla^{j-2}\sigma(t,x) - \nabla^{j-2}\sigma(t,y)| \leqslant \kappa_{j-2}|x-y|^\alpha, \quad |\nabla^{j-2}b(t,x) - \nabla^{j-2}b(t,y)| \leqslant \mu_{j-2}|x-y|^\beta, \quad x, y \in \mathbb{R}^d,$$

then we may derive

$$|\nabla_x^j p(t,x;s,y)| \leqslant C(s-t)^{-\frac{j}{2}}\bar{p}(t,x;s,y).$$

On the other hand, the derivative with respect to the forward variable ∇_y^{j-1} requires an additional assumption on $\nabla^{j-1}\sigma$. Again, assuming that for some $\alpha \in (0,1)$, $|\nabla^{j-1}\sigma(t,x) - \nabla^{j-1}\sigma(t,y)| \leqslant \mu_{j-1}|x-y|^\alpha$ for any $x, y \in \mathbb{R}^d$, then we may derive

$$|\nabla_y^{j-1} p(t,x;s,y)| \leqslant C(s-t)^{-\frac{j-1}{2}}\bar{p}(t,x;s,y).$$

Chapter 5

Appendix

A Proof of Lemma 2.6.19 of Chapter 2

We start this Section by proving two important technical results about the sensitivity of the flow and the covariance matrix with respect to the choice of the freezing parameters.

Lemma A.1. *There exists a constant $C = C(\Theta) \geqslant 1$ such that, for every $0 \leqslant t < s \leqslant T$ and $z, \zeta \in \mathbb{R}^2$ with $(s-t)^{\frac{1}{2}} \left| \mathcal{D}_{\frac{1}{\sqrt{s-t}}} (\zeta - \gamma_{t,s}(z)) \right| \leqslant c_0$, we have:*

$$\left| \mathcal{D}_{\frac{1}{\sqrt{s-t}}} \left(\gamma_{t,s}(z) - \widetilde{\gamma}^{s,\zeta}_{t,s}(z) \right) \right| \leqslant C (s-t)^{\frac{\alpha}{2}} \left(1 + \left| \mathcal{D}_{\frac{1}{\sqrt{s-t}}} (\zeta - \gamma_{t,s}(z)) \right|^{1+\alpha} \right). \tag{A.1}$$

Proof. By (2.6.7) and (2.6.9) we write

$$\begin{aligned}
&\mathcal{D}_{\frac{1}{\sqrt{s-t}}} \left(\gamma_{t,s}(z) - \widetilde{\gamma}^{s,\zeta}_{t,s}(z) \right) \\
&\quad = \mathcal{D}_{\frac{1}{\sqrt{s-t}}} \Big\{ \int_t^s \left(Y(\varrho, \gamma_{t,\varrho}(z)) - Y(\varrho, \gamma_{s,\varrho}(\zeta)) \right) d\varrho + \int_t^s (DY)(\varrho, \gamma_{s,\varrho}(\zeta)) \left(\gamma_{\varrho,s}(\zeta) - \widetilde{\gamma}^{s,\zeta}_{t,\varrho}(z) \right) d\varrho \Big\} \\
&\quad = \mathcal{D}_{\frac{1}{\sqrt{s-t}}} \Big\{ \int_t^s \begin{pmatrix} Y_1(\varrho, \gamma_{t,\varrho}(z)) - Y_1(\varrho, (\gamma_{s,\varrho}(\zeta))_1, (\gamma_{t,\varrho}(z))_2) \\ Y_2(\varrho, \gamma_{t,\varrho}(z)) - Y_2(\varrho, \gamma_{s,\varrho}(\zeta)) \end{pmatrix} d\varrho + \\
&\qquad + \int_t^s \Big[\left(Y_1(\varrho, (\gamma_{s,\varrho}(\zeta))_1, (\gamma_{t,\varrho}(z))_2) - Y_1(\varrho, \gamma_{s,\varrho}(\zeta)) \right) \mathbf{e}_1 \\
&\qquad - (DY)(\varrho, \gamma_{s,\varrho}(\zeta)) \left(\gamma_{t,\varrho}(z) - \gamma_{\varrho,s}(\zeta) \right) \Big] d\varrho + \int_t^s (DY)(\varrho, \gamma_{s,\varrho}(\zeta)) \left(\gamma_{t,\varrho}(z) - \widetilde{\gamma}^{s,\zeta}_{t,\varrho}(z) \right) d\varrho \Big\} \\
&\quad =: I_1 + I_2 + I_3
\end{aligned}$$

Then, by Assumption 2.6.2 we have

$$\begin{aligned}
|I_1| &\leqslant C\int_t^s \left(\frac{|(\gamma_{t,\varrho}(z)-\gamma_{s,\varrho}(\zeta))_1|}{(s-t)^{\frac{3}{2}}}+\frac{|\gamma_{t,\varrho}(z)-\gamma_{s,\varrho}(\zeta)|}{(s-t)^{\frac{1}{2}}}\right)d\varrho\\
&\leqslant C\int_t^s \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\left(\gamma_{t,\varrho}(z)-\gamma_{s,\varrho}(\zeta)\right)\right| \leqslant (s-t)^{\frac{1}{2}};\\
|I_2| &\leqslant C\int_t^s \frac{1}{(s-t)^{\frac{3}{2}}}\left|(\gamma_{t,\varrho}(z)-\gamma_{s,\varrho}(\zeta))_2\right|^{1+\alpha}d\varrho\\
&\leqslant C\int_t^s \frac{1}{(s-t)^{\frac{3}{2}-\frac{1}{2}(1+\alpha)}}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\left(\gamma_{t,\varrho}(z)-\gamma_{s,\varrho}(\zeta)\right)\right|^{1+\alpha}d\varrho\\
&\leqslant C(s-t)^{\frac{\alpha}{2}}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\left(\zeta-\gamma_{t,s}(z)\right)\right|^{1+\alpha}.
\end{aligned}$$

Lastly we notice that by the upper diagonal structure of DY and Assumption (2.6.2), for any $\varrho\in[t,s]$ we have

$$\left\|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(DY)(\varrho,\gamma_{s,\varrho}(\zeta))\mathcal{D}_{\sqrt{s-t}}\right\| \leqslant C(s-t)^{-1}$$

Therefore, we have

$$\begin{aligned}
|I_3| &\leqslant \int_t^s \left\|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(DY)(\varrho,\gamma_{s,\varrho}(\zeta))\mathcal{D}_{\sqrt{s-t}}\right\|\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\left(\gamma_{t,\varrho}(z)-\widetilde{\gamma}^{s,\zeta}_{t,\varrho}(z)\right)\right|d\varrho\\
&\leqslant C(s-t)^{-1}\int_t^s \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\left(\gamma_{t,\varrho}(z)-\widetilde{\gamma}^{s,\zeta}_{t,\varrho}(z)\right)\right|d\varrho.
\end{aligned}$$

Gathering all the terms together we get (A.1) by the Gronwall inequality. □

Lemma A.2. *There exists a constant $C=C(\Theta)\geqslant 1$ such that, for every $0\leqslant t<s\leqslant T$ and $z,\zeta\in\mathbb{R}^2$:*

$$\left\|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\left(\widetilde{\mathcal{C}}^{t,z}_{t,s}-\widetilde{\mathcal{C}}^{s,\zeta}_{t,s}\right)\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\right\| \leqslant C(s-t)^{\frac{\alpha}{2}}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\left(\zeta-\gamma_{t,s}(z)\right)\right|^{\alpha}.$$

Proof. By (2.6.10) we write

$$\begin{aligned}
&\left\|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\left(\widetilde{\mathcal{C}}^{t,z}_{t,s}-\widetilde{\mathcal{C}}^{s,\zeta}_{t,s}\right)\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\right\|\\
&= \sup_{|y|=1}\Big\{\int_t^s \left(a(\varrho,\gamma_{t,\varrho}(z))-a(\varrho,\gamma_{s,\varrho}(\zeta))\right)\langle (E^{t,z}_{\varrho,s}\mathbf{e}_2)(E^{t,z}_{\varrho,s}\mathbf{e}_2)^*\mathcal{D}_{\frac{1}{\sqrt{s-t}}}y,\mathcal{D}_{\frac{1}{\sqrt{s-t}}}y\rangle d\varrho+\\
&\qquad +\int_t^s\langle a(\varrho,\gamma_{s,\varrho}(\zeta)\left[(E^{t,z}_{\varrho,s}\mathbf{e}_2)(E^{t,z}_{\varrho,s}\mathbf{e}_2)^*-(E^{s,\zeta}_{\varrho,s}\mathbf{e}_2)(E^{s,\zeta}_{\varrho,s}\mathbf{e}_2)^*\right]\mathcal{D}_{\frac{1}{\sqrt{s-t}}}y,\mathcal{D}_{\frac{1}{\sqrt{s-t}}}y\rangle d\varrho\Big\}\\
&=: I_1+I_2
\end{aligned}$$

By Assumption 2.6.1 and Proposition 2.6.9 it is easy to see that

$$|I_1| \leqslant C(s-t)^{\frac{\alpha}{2}} \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}} \left(\zeta - \gamma_{t,s}(z)\right)\right|^{\alpha}.$$

For I_2 we have

$$\begin{aligned} |I_2| &\leqslant (s-t)^{-1} \int_t^s \left\| \mathcal{D}_{\frac{1}{\sqrt{s-t}}} \left(E^{t,z}_{\varrho,s} - E^{s,\zeta}_{\varrho,s}\right) \mathcal{D}_{\sqrt{s-t}} \right\| \left\| \mathcal{D}_{\frac{1}{\sqrt{s-t}}} \left(E^{t,z}_{\varrho,s} + E^{s,\zeta}_{\varrho,s}\right) \mathcal{D}_{\sqrt{s-t}} \right\| d\varrho \\ &\leqslant C(s-t)^{-1} \int_t^s \left\| \mathcal{D}_{\frac{1}{\sqrt{s-t}}} \left(E^{t,z}_{\varrho,s} - E^{s,\zeta}_{\varrho,s}\right) \mathcal{D}_{\sqrt{s-t}} \right\| d\varrho \end{aligned}$$

where we used that $\mathcal{D}_{\frac{1}{\sqrt{s-t}}} E^{t_0,z_0}_{\varrho,s} \mathcal{D}_{\sqrt{s-t}}$ is a positive, bounded matrix, uniformly in $t_0 \in [0,T]$, $z_0 \in \mathbb{R}^2$ and $\varrho \in [t,s]$, by the structure of the resolvent. Moreover, by the upper diagonal structure of (DY) and Assumption 2.6.2 we have

$$\begin{aligned} &\left\| \mathcal{D}_{\frac{1}{\sqrt{s-t}}} \left(E^{t,z}_{\varrho,s} - E^{s,\zeta}_{\varrho,s}\right) \mathcal{D}_{\sqrt{s-t}} \right\| \\ &\quad\leqslant \int_\varrho^s \left\| \mathcal{D}_{\frac{1}{\sqrt{s-t}}} (DY)(u,\gamma_{t,u}(z)) \mathcal{D}_{\sqrt{s-t}} \right\| \left\| \mathcal{D}_{\frac{1}{\sqrt{s-t}}} \left(E^{t,z}_{u,s} - E^{s,\zeta}_{u,s}\right) \mathcal{D}_{\sqrt{s-t}} \right\| du \\ &\qquad + \int_\varrho^s \left\| \mathcal{D}_{\frac{1}{\sqrt{s-t}}} \left((DY)(u,\gamma_{t,u}(z)) - (DY)(u,\gamma_{s,u}(\zeta))\right) \mathcal{D}_{\sqrt{s-t}} \right\| \left\| \mathcal{D}_{\frac{1}{\sqrt{s-t}}} E^{s,\zeta}_{u,s} \mathcal{D}_{\sqrt{s-t}} \right\| du \\ &\quad\leqslant C(s-t)^{-1} \int_\varrho^s \left\| \mathcal{D}_{\frac{1}{\sqrt{s-t}}} \left(E^{t,z}_{u,s} - E^{s,\zeta}_{u,s}\right) \mathcal{D}_{\sqrt{s-t}} \right\| du \\ &\qquad + C(s-t)^{-1} \int_\varrho^s \left|\partial_v Y_1(u,\gamma_{t,u}(z)) - \partial_v Y_1(u,\gamma_{s,u}(\zeta))\right| du \\ &\quad\leqslant C(s-t)^{-1} \int_\varrho^s \left\| \mathcal{D}_{\frac{1}{\sqrt{s-t}}} \left(E^{t,z}_{u,s} - E^{s,\zeta}_{u,s}\right) \mathcal{D}_{\sqrt{s-t}} \right\| du + (s-t)^{-1} \int_\varrho^s |\gamma_{t,u}(z) - \gamma_{s,u}(\zeta)|^{\alpha} du. \\ &\quad\leqslant C(s-t)^{\frac{\alpha}{2}} \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}} \left(\zeta - \gamma_{t,s}(z)\right)\right|^{\alpha}. \end{aligned}$$

where we used the Gronwall inequality in the last step. Coming back to I_2 we directly derive

$$|I_2| \leqslant C(s-t)^{\frac{\alpha}{2}} \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}} \left(\zeta - \gamma_{t,s}(z)\right)\right|^{\alpha},$$

and this proves the assertion. □

Proof of Lemma 2.6.19. Assume that $(s-t)^{\frac{1}{2}} \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}} \left(\zeta - \gamma_{t,s}(z)\right)\right| \geqslant c_0$. Then (2.6.31) can be directly derived by (2.6.24) and (2.6.25). Indeed we can write

$$\begin{aligned} \left| \partial^j_\nu \widetilde{\Gamma}^{t_0,\eta}(t,z;s,\zeta) - \partial^j_\nu Z(t,z;s,\zeta) \right|_{(t_0,\eta)=(s,\zeta)} &\leqslant C \frac{\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}} \left(\zeta - \gamma_{t,s}(z)\right)\right|^{\alpha}}{(s-t)^{\frac{j-\alpha}{2}}} \Gamma^{\text{heat}} \left(\mu \mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)\right) \\ &\leqslant C(s-t)^{\frac{\alpha-j}{2}} \Gamma^{\text{heat}} \left(\mu' \mathcal{D}_{s-t}, \zeta - \gamma_{t,s}(z)\right). \end{aligned}$$

Assume now that $(s-t)^{\frac{1}{2}}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(\zeta-\gamma_{t,s}(z))\right| \leqslant c_0$ to be determined. We will prove the statement when $j=0$. The statement for $j=1,2$ would be derived similarly from usual computations on gaussian kernels. We denote for simplicity $C_1=\widetilde{\mathcal{C}}^{t,z}_{t,s}$, $C_2=\widetilde{\mathcal{C}}^{s,\zeta}_{t,s}$, $w_1=\zeta-\gamma_{t,s}(z)$, $w_2=\zeta-\widetilde{\gamma}^{s,\zeta}_{t,s}(z)$. Then the thesis for $j=0$ follows from the following estimates:

$$\left|(\det C_2)^{-\frac{1}{2}}-(\det C_2)^{-\frac{1}{2}}\right| \leqslant C(s-t)^{-2+\frac{\alpha}{2}}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}w_1\right|^{\alpha};$$

$$\left|\exp\left(-\frac{1}{2}\langle C_1^{-1}w_1,w_1\rangle\right)-\exp\left(-\frac{1}{2}\langle C_2^{-1}w_2,w_2\rangle\right)\right| \leqslant (s-t)^{\frac{\alpha}{2}}\exp\left(-\frac{1}{2\mu}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}w_1\right|^2\right). \quad (A.2)$$

By proposition 2.6.9 we have

$$\begin{aligned}\left|(\det C_1)^{-\frac{1}{2}}-(\det C_2)^{-\frac{1}{2}}\right| &\leqslant (s-t)^{-2}\frac{|\det C_1-\det C_2|}{(s-t)^4}\\ &\leqslant (s-t)^{-2}\left|\det\mathcal{D}_{\frac{1}{\sqrt{s-t}}}C_1\mathcal{D}_{\frac{1}{\sqrt{s-t}}}-\det\mathcal{D}_{\frac{1}{\sqrt{s-t}}}C_2\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\right|\\ &\leqslant (s-t)^{-2}\|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(C_1-C_2)\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\|\end{aligned}$$

(by Lemma A.2)

$$\leqslant (s-t)^{-2+\frac{\alpha}{2}}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}w_1\right|^{\alpha}.$$

Let us now turn to the proof of (A.2). Write:

$$\begin{aligned}&\left|\exp\left(-\frac{1}{2}\langle C_1^{-1}w_1,w_1\rangle\right)-\exp\left(-\frac{1}{2}\langle C_2^{-1}w_2,w_2\rangle\right)\right|\\ &\quad\leqslant\left|\exp\left(-\frac{1}{2}\langle C_1^{-1}w_1,w_1\rangle\right)-\exp\left(-\frac{1}{2}\langle C_2^{-1}w_1,w_1\rangle\right)\right|\\ &\qquad+\left|\exp\left(-\frac{1}{2}\langle C_2^{-1}w_1,w_1\rangle\right)-\exp\left(-\frac{1}{2}\langle C_2^{-1}w_2,w_2\rangle\right)\right|\\ &\quad=: I_1+I_2.\end{aligned}$$

For the first term we have

$$I_1\leqslant\left|\langle(C_1^{-1}-C_2^{-1})w_1,w_1\rangle\right|\int_0^1\exp\left\{-\frac{1}{2}\left[\langle C_1^{-1}w_1,w_1\rangle+\lambda\left(\langle C_2^{-1}w_1,w_1\rangle-\langle C_1^{-1}w_1,w_1\rangle\right)\right]\right\}d\lambda.$$

Exploiting the equality $C_2^{-1} - C_1^{-1} = C_2^{-1}(C_1 - C_2)C_1^{-1}$, Remark 2.6.10 and Lemma A.2 we get

$$\begin{aligned}
&\left|\langle (C_1^{-1} - C_2^{-1})w_1, w_1\rangle\right| \\
&\leqslant \left\|\mathcal{D}_{\sqrt{s-t}}(C_1^{-1} - C_2^{-1})\mathcal{D}_{\sqrt{s-t}}\right\| \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}} w_1\right|^2 \\
&\leqslant \left\|\mathcal{D}_{\sqrt{s-t}}C_2^{-1}\mathcal{D}_{\sqrt{s-t}}\right\| \left\|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(C_2 - C_1)\mathcal{D}_{\frac{1}{\sqrt{s-t}}}\right\| \left\|\mathcal{D}_{\sqrt{s-t}}C_1^{-1}\mathcal{D}_{\sqrt{s-t}}\right\| \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}} w_1\right|^2 \\
&\leqslant (s-t)^{\frac{\alpha}{2}} \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}} w_1\right|^{2+\alpha}.
\end{aligned}$$

and, for every $\lambda \in [0,1]$

$$\exp\left\{\rho\left(\langle (C_2^{-1} - C_1^{-1})w_1, w_1\rangle\right)\right\} \leqslant \exp\left\{Cc_0 \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}} w_1\right|^2\right\},$$

which eventually yields

$$I_1 \leqslant C(s-t)^{\frac{\alpha}{2}} \exp\left\{-\frac{1}{2\mu}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}} w_1\right|^2\right\}$$

provided $c_0 = c_0(\Theta)$ is small enough. On the other hand we have

$$\begin{aligned}
I_2 &\leqslant \left|\langle C_2^{-1}(w_1 + w_2), (w_1 - w_2)\rangle\right| \times \\
&\times \int_0^1 \exp\left\{-\frac{1}{2}\left[\langle C_2^{-1}w_1, w_1\rangle + \lambda\left(\langle C_2^{-1}w_2, w_2\rangle - \langle C_2^{-1}w_1, w_1\rangle\right)\right]\right\} d\lambda.
\end{aligned}$$

By Lemma A.1 we get

$$\begin{aligned}
&\left|C_2^{-\frac{1}{2}}(w_1 - w_2)\right| \leqslant C\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(w_1 - w_2)\right| \leqslant (s-t)^{\frac{1}{2}}\left(1 + \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}} w_1\right|^{1+\alpha}\right); \\
&\left|C_2^{-\frac{1}{2}}(w_1 + w_2)\right| \leqslant C\left(\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}} w_1\right| + \left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}}(w_1 - w_2)\right|\right) \leqslant C.
\end{aligned}$$

Then it suffices to notice that for all $\lambda \in [0,1]$ and $\eta \in (0,1]$:

$$\exp\left\{-\frac{1}{2}\varrho\left(\langle C_2^{-1}\omega_2, \omega_2\rangle - \langle C_2^{-1}\omega_1, \omega_1\rangle\right)\right\} \leqslant \exp\left\{\frac{C}{2}\left[\eta\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}} w_1\right|^2 + (\eta^{-1} + 1)c_0^2 T\right]\right\},$$

which eventually gives, for η small enough:

$$I_2 \leqslant C(s-t)^{\frac{\alpha}{2}} \exp\left(-\frac{1}{2\mu}\left|\mathcal{D}_{\frac{1}{\sqrt{s-t}}} w_1\right|^2\right).$$

The proof is complete. $\square$

B Backward Itô calculus

In this section we collect some basic result about *backward Itô integrals* and the *backward diffusion SPDE* (or *Krylov equation* according to [62]). This is standard material which resume the original results in [33], [40], [41], [43], [68] (see also the monographs [62] and [44]).

Let $W = (W_t)_{t\in[0,T]}$ be a d-dimensional Brownian motion on $(\Omega, \mathcal{F}, P, \mathcal{F}^W)$ where $\mathcal{F}^W$ denotes the standard Brownian filtration satisfying the usual assumptions. We consider

$$\mathcal{F}_T^{W,t} = \sigma(\mathcal{G}_t \cup \mathcal{N}), \qquad \mathcal{G}_t = \sigma(W_s - W_t,\, t \leqslant s \leqslant T), \qquad t \in [0,T],$$

the augmented σ-algebra of Brownian increments between t and T. Notice that $(\mathcal{F}_T^{W,t})_{0\leqslant t\leqslant T}$ is a decreasing family of σ-algebras. Then the process

$$\overleftarrow{W}_t := W_T - W_{T-t}, \qquad t \in [0,T],$$

is a Brownian motion on $(\Omega, \mathcal{F}, P, \overleftarrow{\mathcal{F}})$ where

$$\overleftarrow{\mathcal{F}}_t := \mathcal{F}_T^{W,T-t}, \qquad t \in [0,T],$$

is the "backward" Brownian filtration. The backward stochastic Itô integral is defined as

$$\int_t^s u_r \star dW_r := \int_{T-s}^{T-t} u_{T-r} d\overleftarrow{W}_r, \qquad 0 \leqslant t \leqslant s \leqslant T, \tag{B.1}$$

under the assumptions on u for which the RHS of (B.1) is defined in the usual Itô sense, that is

i) $t \mapsto u_{T-t}$ is $\overleftarrow{\mathcal{F}}$-progressively measurable (thus $u_t \in m\mathcal{F}_T^{W,t}$ for any $t \in [0,T]$);

ii) $u \in L^2([0,T])$ a.s.

For practical purposes, if u is continuous, the backward integral is the limit

$$\int_t^s u_r \star dW_r := \lim_{|\pi|\to 0^+} \sum_{k=1}^n u_{t_k}\left(W_{t_k} - W_{t_{k-1}}\right) \tag{B.2}$$

in probability, where $\pi = \{t = t_0 < t_1 < \cdots < t_n = s\}$ denotes a partition of $[t,s]$.

A backward Itô process is a process of the form

$$X_t = X_T + \int_t^T b_s ds + \int_t^T \sigma_s \star dW_s, \qquad t \in [0,T],$$

also written in differential form as

$$-dX_t = b_t dt + \sigma_t \star dW_t. \tag{B.3}$$

Theorem B.1 (Backward Itô formula). *Let $v = v(t,x) \in C^{1,2}(\mathbb{R}_{\geqslant 0} \times \mathbb{R}^d)$ and let X be the process in* (B.3)*. Then*

$$-dv(t,X_t) = \left((\partial_t v)(t,X_t) + \frac{1}{2}(\sigma_t\sigma_t^*)_{ij}(\partial_{x_ix_j}v)(t,X_t) + (b_t)_i(\partial_{x_i}v)(t,X_t)\right)dt + (\sigma_t)_{ij}\left(\partial_{x_i}v\right)(t,X_t)\star dW_t^j. \tag{B.4}$$

A crucial tool in our analysis is the following

Theorem B.2 (Backward diffusion SPDE)**.** *Assume $b,\sigma \in bC^3(\mathbb{R}_{\geqslant 0} \times \mathbb{R}^d)$ and denote by $s \mapsto X_s^{t,x}$ the solution of the SDE*

$$dX_s^{t,x} = b(s, X_s^{t,x})ds + \sigma(s, X_s^{t,x})dW_s \tag{B.5}$$

with initial condition $X_t^{t,x} = x$. Then the process $(t,x) \mapsto X_T^{t,x}$ solves the backward SPDE

$$\begin{cases} -dX_T^{t,x} = \mathcal{L}X_T^{t,x}dt + \sigma_{ij}(t,x)\partial_{x_i}X_T^{t,x} \star dW_t^j, \\ X_T^{T,x} = x, \end{cases} \tag{B.6}$$

where

$$\mathcal{L} = \frac{1}{2}(\sigma(t,x)\sigma^*(t,x))_{ij}\partial_{x_jx_i} + b_i(t,x)\partial_{x_i}$$

is the characteristic operator of X. More explicitly, in (B.6) *we have*

$$\mathcal{L}X_T^{t,x} \equiv \frac{1}{2}(\sigma(t,x)\sigma^*(t,x))_{ij}\partial_{x_jx_i}X_T^{t,x} + b_i(t,x)\partial_{x_i}X_T^{t,x}.$$

Remark B.3. *The regularity assumption of Theorem B.2 on the coefficients is by no means optimal: [62], Theorem 5.1, proves that $(t,x) \mapsto X_T^{t,x}$ is a generalized (or classical, under non-degeneracy conditions) solution of* (B.6) *if $b,\sigma \in bC^1(\mathbb{R}_{\geqslant 0} \times \mathbb{R}^d)$.*

Proof. For illustrative purposes we only consider the one-dimensional, autonomous case. A general proof can be found in [62], Proposition 5.3. Here we follow the "direct" approach proposed in [68]. By standard results for stochastic flows (cf. [44]), $x \mapsto X_T^{t,x}$ is sufficiently regular to support the derivatives in the classical sense. We use the Taylor expansion for C^2-functions:

$$f(\delta) - f(0) = \delta f'(0) + \frac{\delta^2}{2}f''(\lambda\delta), \qquad \lambda \in [0,1]. \tag{B.7}$$

We have

$$\begin{aligned} X_T^{t,x} - x &= X_T^{t,x} - X_T^{T,x} \\ &= \sum_{k=1}^n \left(X_T^{t_{k-1},x} - X_T^{t_k,x}\right) = \end{aligned}$$

(by the flow property)

$$= \sum_{k=1}^{n} \left(X_T^{t_k, X_{t_k}^{t_{k-1},x}} - X_T^{t_k,x} \right) =$$

(by (B.7) with $f(\delta) = X_T^{t_k,x+\delta}$ and $\delta = \Delta_k X := X_{t_k}^{t_{k-1},x} - x$)

$$= \sum_{k=1}^{n} \left(\Delta_k X \partial_x X_T^{t_k,x} + \frac{(\Delta_k X)^2}{2} \partial_{xx} X_T^{t_k,x+\lambda_k \Delta_k X} \right) \tag{B.8}$$

for some $\lambda_k = \lambda_k(\omega) \in [0,1]$. Now, we have

$$\Delta_k X = X_{t_k}^{t_{k-1},x} - x = \int_{t_{k-1}}^{t_k} b(X_s^{t_{k-1},x}) ds + \int_{t_{k-1}}^{t_k} \sigma(X_s^{t_{k-1},x}) dW_s.$$

Thus, setting

$$\Delta_k t = t_k - t_{k-1}, \qquad \Delta_k W = W_{t_k} - W_{t_{k-1}}, \qquad \widetilde{\Delta}_k X = b(x)\Delta_k t + \sigma(x)\Delta_k W,$$

by standard estimates for solutions of SDEs, we have

$$\Delta_k X - \widetilde{\Delta}_k X = \int_{t_{k-1}}^{t_k} \left(b(X_s^{t_{k-1},x}) - b(x) \right) ds + \int_{t_{k-1}}^{t_k} \left(\sigma(X_s^{t_{k-1},x}) - \sigma(x) \right) dW_s = \mathrm{O}(\Delta_k t),$$
$$\partial_{xx} X_T^{t_k,x+\lambda_k \Delta_k X} - \partial_{xx} X_T^{t_k,x} = \mathrm{O}(\Delta_k t),$$

in the square mean sense or, more precisely,

$$E\left[|\Delta_k X - \widetilde{\Delta}_k X|^2 + \left| \partial_{xx} X_T^{t_k,x+\lambda_k \Delta_k X} - \partial_{xx} X_T^{t_k,x} \right|^2 \right] \leqslant c(1+|x|^2)(\Delta_k t)^2$$

with c depending only on T and the Lipschitz constants of b, σ. From (B.8) we get

$$X_T^{t,x} - x = \sum_{k=1}^{n} \left(\widetilde{\Delta}_k X \partial_x X_T^{t_k,x} + \frac{(\widetilde{\Delta}_k X)^2}{2} \partial_{xx} X_T^{t_k,x} \right) + \mathrm{O}(\Delta_k t).$$

Next we recall (B.2) and notice that $\partial_x X_T^{t_k,x}, \partial_{xx} X_T^{t_k,x} \in m\mathcal{F}_T^{W,t_k}$. Thus, passing to the limit, we have

$$\sum_{k=1}^{n} \widetilde{\Delta}_k X \partial_x X_T^{t_k,x} \longrightarrow \int_t^T b(t,x) \partial_x X_T^{s,x} ds + \int_t^T \sigma(x) \partial_x X_T^{s,x} \star dW_s,$$
$$\sum_{k=1}^{n} (\widetilde{\Delta}_k X)^2 \partial_{xx} X_T^{t_k,x} \longrightarrow \int_t^T \sigma^2(x) \partial_{xx} X_T^{s,x} ds,$$

in the square mean sense and this concludes the proof. □

We have a useful corollary of Theorem B.2.

Corollary B.4 (Invariance of the backward diffusion SPDE). *For $v \in bC^2(\mathbb{R}^d)$ and X as in* (B.5)*, let $V_T^{t,x} = v(X_T^{t,x})$. Then $V_T^{t,x}$ satisfies the same SPDE* (B.6)*, that is*

$$-dV_T^{t,x} = \mathcal{L}V_T^{t,x}dt + \sigma_{ij}(t,x)\partial_{x_i}V_T^{t,x} \star dW_t^j$$

with terminal condition $V_T^{T,x} = g(x)$.

Proof. To fix ideas, we first consider the one-dimensional case: by the backward SPDE (B.6) and the backward Itô formula (B.4), we have

$$\begin{aligned}-dv(X_T^{t,x}) = &\left(\frac{\sigma^2(t,x)}{2}v''(X_T^{t,x})(\partial_x X_T^{t,x})^2 + \frac{\sigma^2(t,x)}{2}v'(X_T^{t,x})\partial_{xx}X_T^{t,x} + b(t,x)v'(X_T^{t,x})\partial_x X_T^{t,x}\right)dt\\ &+ \sigma(t,x)v'(X_T^{t,x})\partial_x X_T^{t,x} \star dW_t =\end{aligned}$$

(using the identities $\partial_x V_T^{t,x} = v'(X_T^{t,x})\partial_x X_T^{t,x}$ and $\partial_{xx}V_T^{t,x} = v''(X_T^{t,x})(\partial_x X_T^{t,x})^2 + v'(X_T^{t,x})\partial_{xx}X_T^{t,x}$)

$$= \left(\frac{\sigma^2(t,x)}{2}\partial_{xx}V_T^{t,x} + b(t,x)\partial_x V_T^{t,x}\right)dt + \sigma(t,x)\partial_x V_T^{t,x} \star dW_t$$

and this proves the thesis. In general, we have

$$\begin{aligned}\partial_{x_h}V_T^{t,x} &= (\nabla v)(X_T^{t,x})\partial_{x_h}X_T^{t,x},\\ \partial_{x_h x_k}V_T^{t,x} &= (\partial_{ij}v)(X_T^{t,x})(\partial_{x_h}X_T^{t,x})_i(\partial_{x_k}X_T^{t,x})_j + (\nabla v)(X_T^{t,x})(\partial_{x_h x_k}X_T^{t,x}),\end{aligned} \tag{B.9}$$

and by (B.6) and (B.4)

$$\begin{aligned}-dv(X_T^{t,x}) = &\left(\frac{1}{2}\left((\nabla X_T^{t,x})\sigma(t,x)((\nabla X_T^{t,x})\sigma(t,x))^*\right)_{ij}(\partial_{ij}v)(X_T^{t,x})\right)dt\\ &+ \left(\frac{1}{2}(\sigma(t,x)\sigma^*(t,x))_{ij}\partial_{x_j x_i}X_T^{t,x} + b(t,x)\nabla X_T^{t,x}\right)(\nabla v)(X_T^{t,x})dt\\ &+ (\nabla v)(X_T^{t,x})(\nabla X_T^{t,x})\sigma(t,x) \star dW_t =\end{aligned}$$

(by (B.9))

$$= \left(\frac{1}{2}(\sigma(t,x)\sigma^*(x))_{ij}\partial_{x_j x_i}V_T^{t,x} + b(t,x)\nabla V_T^{t,x}\right)dt + \nabla V_T^{t,x}\sigma(t,x) \star dW_t.$$

□

C Proof of the equicontinuity ($\mathbf{C}_1$), ($\mathbf{C}_2$) and ($\mathbf{C}_3$) of Chapter 4

In this section, we drop the subscripts and superscripts in ε for notational convenience. However, it must be recalled that we aim at proving some equicontinuity properties for the densities associated with the SDE (4.2.2) with mollified coefficients and their derivatives.

In this section we devote to proving the following Hölder continuity of the derivatives.

Lemma C.1. *Suppose that Assumptions 4.1.1 and 4.1.2 hold. Let $T > 0$, $\alpha_1 \in (0,1)$, $\alpha_2 \in (0,\alpha)$ and $\alpha_3 \in (0, \alpha \wedge \beta)$.*

(C$_1$) *There exist constants $C, \mu > 0$ depending only on Θ, α_1, α_2 such that for all $0 \leqslant t < s \leqslant T$ and $x, x', y, y' \in \mathbb{R}^d$,*

$$
\begin{aligned}
|\nabla_x p(t,x;s,y) - \nabla_x p(t,x';s,y)| &\leqslant C \frac{|x-x'|^{\alpha_1}}{(s-t)^{(1+\alpha_1)/2}} \Big(\Gamma_\mu(t,x;s,y) + \Gamma_\mu(t,x';s,y) \Big) \\
|\nabla_x p(t,x;s,y) - \nabla_x p(t,x;s,y')| &\leqslant C \frac{|y-y'|^{\alpha_2}}{(s-t)^{(1+\alpha_2)/2}} \Big(\Gamma_\mu(t,x;s,y) + \Gamma_\mu(t,x;s,y') \Big).
\end{aligned}
$$

(C$_2$) *If $\beta \in (0,1]$, there exist constants $C, \mu > 0$ depending only on Θ such that for all $0 \leqslant t < s \leqslant T$ and $x, x', y, y' \in \mathbb{R}^d$,*

$$
\begin{aligned}
|\nabla_x^2 p(t,x;s,y) - \nabla_x^2 p(t,x';s,y)| &\leqslant C \left(\frac{|x-x'|}{(s-t)^{\frac{3}{2}}} + \frac{|x-x'|^\alpha + |x-x'|^\beta}{s-t} \right) \times \qquad \text{(C.1)} \\
&\quad \times \Big(\Gamma_\mu(t,x;s,y) + \Gamma_\mu(t,x';s,y) \Big), \\
|\nabla_x^2 p(t,x;s,y) - \nabla_x^2 p(t,x;s,y')| &\leqslant C \left(\frac{|y-y'|^{\alpha_2}}{(s-t)^{1+\frac{\alpha_2}{2}}} + \frac{|y-y'|^\alpha + |y-y'|^\beta}{s-t} \right) \times \\
&\quad \times \Big(\Gamma_\mu(t,x;s,y) + \Gamma_\mu(t,x';s,y) \Big).
\end{aligned}
$$

(C$_3$) *If $\sigma \in bC^{1,\alpha}_{0,T}$ for some $\alpha \in (0,1)$ and $\beta \in (0,1)$, then there exist constants $C, \mu > 0$ depending only on Θ, α_1, α_3 such that for all $0 \leqslant t < s \leqslant T$ and $x, x', y, y' \in \mathbb{R}^d$,*

$$
\begin{aligned}
|\nabla_y p(t,x;s,y) - \nabla_y p(t,x;s,y')| &\leqslant C \frac{|y-y'|^{\alpha_3}}{(s-t)^{(1+\alpha_3)/2}} \Big(\Gamma_\mu(t,x;s,y) + \Gamma_\mu(t,x;s,y') \Big), \\
|\nabla_y p(t,x;s,y) - \nabla_y p(t,x';s,y)| &\leqslant C \frac{|x-x'|^{\alpha_1}}{(s-t)^{(1+\alpha_1)/2}} \Big(\Gamma_\mu(t,x;s,y) + \Gamma_\mu(t,x;s,y') \Big).
\end{aligned}
$$

Proof. We only prove **(C$_2$)** and focus on the sensitivity w.r.t the variable x. The sensitivity w.r.t. the variable y could be established similarly. Also, the inequalities in conditions **(C$_1$)** and **(C$_3$)** could be shown more directly.

First of all, if $|x-x'|^2 > (t-s)/4$, then by (4.3.15), we clearly have

$$
|\nabla_x^2 p(t,x;s,y) - \nabla_x^2 p(t,x';s,y)| \leqslant C(s-t)^{-1} \Big(\Gamma_\mu(t,x;s,y) + \Gamma_\mu(t,x';s,y) \Big) \lesssim \text{r.h.s. of (C.1)}.
$$

Next we restrict to the so-called *diagonal case*

$$
|x-x'|^2 \leqslant (t-s)/4.
$$

For any fixed freezing point (t_0, x_0) and $r \in (t, s)$, by (4.2.11), one sees that

$$p(t, x; s, y) = \widetilde{P}^{t_0,x_0}_{t,r} p(r, \cdot; s, y)(x) + \int_t^r \int_{\mathbb{R}^d} \widetilde{p}^{t_0,x_0}(t, x; u, z)(\mathcal{L}_{u,z} - \widetilde{\mathcal{L}}^{t_0,x_0}_{u,z})p(u, z; s, y)dzdu,$$

where, with the notations of (4.2.6),

$$\widetilde{P}^{t_0,x_0}_{t,r} f(x) = \int_{\mathbb{R}^d} \widetilde{p}^{t_0,x_0}(t, x; r, z)f(z)dz.$$

Let us now differentiate w.r.t. r. We obtain for all $(t_0, x_0) \in [0, T] \times \mathbb{R}^d$:

$$0 = \partial_r[\widetilde{P}^{t_0,x_0}_{t,r} p(r, \cdot; s, y)(x)] + \int_{\mathbb{R}^d} \widetilde{p}^{t_0,x_0}(t, x; r, z)(\mathcal{L}_{u,z} - \widetilde{\mathcal{L}}^{t_0,x_0}_{u,z})p(r, z; s, y)dz. \tag{C.2}$$

Fix $\bar{t} \in (t, s)$. Now, integrating (C.2) between $\bar{t}$ and s and taking $x_0 = x_0'$, we get

$$0 = \widetilde{p}^{t_0,x_0'}(t, x; s, y) - \widetilde{P}^{t_0,x_0'}_{t,\bar{t}} p(\bar{t}, \cdot; s, y)(x) + \int_{\bar{t}}^s dr \int_{\mathbb{R}^d} \widetilde{p}^{t_0,x_0'}(t, x; r, z)(\mathcal{L}_{u,z} - \widetilde{\mathcal{L}}^{t_0,x_0'}_{u,z})p(r, z; s, y)dz.$$

Moreover, integrating (C.2) between t and $\bar{t}$, we obtain

$$0 = \widetilde{P}^{t_0,x_0}_{t,\bar{t}} p(\bar{t}, \cdot; s, y)(x) - p(t, x; s, y) + \int_t^{\bar{t}} dr \int_{\mathbb{R}^d} \widetilde{p}^{t_0,x_0}(t, x; r, z)(\mathcal{L}_{u,z} - \widetilde{\mathcal{L}}^{t_0,x_0}_{u,z})p(r, z; s, y)dz;$$

Summing up the two equalities we get the following new representation for $p(t, x, s, y)$:

$$\begin{aligned} p(t, x; s, y) =& \widetilde{p}^{t_0,x_0'}(t, x; s, y) + \Big(\widetilde{P}^{t_0,x_0}_{t,\bar{t}} - \widetilde{P}^{t_0,x_0'}_{t,\bar{t}}\Big)p(\bar{t}, \cdot; s, y)(x) \\ &+ \int_{\bar{t}}^s dr \int_{\mathbb{R}^d} \widetilde{p}^{t_0,x_0'}(t, x; r, z)(\mathcal{L}_{u,z} - \widetilde{\mathcal{L}}^{t_0,x_0'}_{u,z})p(r, z; s, y)dz \\ &+ \int_t^{\bar{t}} dr \int_{\mathbb{R}^d} \widetilde{p}^{t_0,x_0}(t, x; r, z)(\mathcal{L}_{u,z} - \widetilde{\mathcal{L}}^{t_0,x_0}_{u,z})p(r, z; s, y)dz, \end{aligned}$$

which, together with (4.2.11) yields

$$\begin{aligned} p(t, x; s, y) - p(t, x'; s, y) =& \widetilde{p}^{t_0,x_0'}(t, x; s, y) - \widetilde{p}^{t_0,x_0'}(t, x'; s, y) + \Big(\widetilde{P}^{t_0,x_0}_{t,\bar{t}} - \widetilde{P}^{t_0,x_0'}_{t,\bar{t}}\Big)p(\bar{t}, \cdot; s, y)(x) \\ &+ \Delta^{t_0,x_0',x_0'}_{\mathrm{diag}}(t, s, x, x', y) + \Delta^{t_0,x_0,x_0'}_{\mathrm{off-diag}}(t, s, x, x', y), \end{aligned} \tag{C.3}$$

where

$$\Delta^{t_0,x_0',x_0'}_{\mathrm{diag}}(t, s, x, x', y) = \int_{\bar{t}}^s dr \int_{\mathbb{R}^d} \Big[\widetilde{p}^{t_0,x_0'}(t, x; r, z) - \widetilde{p}^{t_0,x_0'}(t, x'; r, z)\Big](\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}^{t_0,x_0'}_{r,z})p(r, z; s, y)dz$$

and

$$\begin{aligned} \Delta^{t_0,x_0,x_0'}_{\mathrm{off-diag}}(t, s, x, x', y) = \int_t^{\bar{t}} dt \int_{\mathbb{R}^d} \Big[&\widetilde{p}^{t_0,x_0}(t, x; r, z)(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}^{t_0,x_0}_{r,z})p(r, z; t, y) + \\ &- \widetilde{p}^{t_0,x_0'}(t, x'; r, z)(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}^{t_0,x_0'}_{r,z})p(r, z; s, y)\Big]dz. \end{aligned}$$

Observe that for any freezing couple (t_0, x_0) and $h \in \mathbb{R}^d$,

$$\nabla_x^2 \widetilde{p}^{t_0,x_0}(t, x+h; s, y) = \nabla_y^2 \widetilde{p}^{t_0,x_0}(t, x; s, y-h).$$

After differentiating twice in x for both sides of (C.3) and taking $t_0 = t$ and $x_0 = x,\ x_0' = x'$, we obtain

$$\nabla_x^2 p(t, x; s, y) - \nabla_x^2 p(t, x'; s, y) = \sum_{i=1}^{4} I_i(t, s, x, x', y),$$

where, with the notation $Z_0(s, x; t, y) = \widetilde{p}^{t,x}(t, x; s, y)$,

$$\begin{aligned}
I_1(t, s, x, x', y) &:= \nabla_y^2 Z_0(t, x'; s, y + x' - x) - \nabla_y^2 Z_0(t, x'; s, y),\\
I_2(t, s, x, x', y) &:= \int_{\mathbb{R}^d} \nabla_z^2 (\widetilde{p}^{t,x}(t, x; \bar{t}, z) - \widetilde{p}^{t,x'}(t, x; \bar{t}, z)) p(\bar{t}, z; s, y) dz,\\
I_3(t, s, x, x', y) &:= \int_{\bar{t}}^{s} dr \int_{\mathbb{R}^d} \Big[\nabla_z^2 Z_0(t, x', r, z + x' - x) - \nabla_z^2 Z_0(t, x'; r, z)\Big](\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{t,x'}) p(r, z; s, y) dz,\\
I_4(t, s, x, x', y) &:= \int_{t}^{\bar{t}} dr \int_{\mathbb{R}^d} \Big[\nabla_z^2 Z_0(t, x; r, z)(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{t,x}) p(r, z; s, y) +\\
&\qquad - \nabla_z^2 Z_0(t, x'; r, z)(\mathcal{L}_{r,z} - \widetilde{\mathcal{L}}_{r,z}^{t,x'}) p(r, z; s, y)\Big] dz.
\end{aligned}$$

Note that by Lemma 4.2.3, for $j \in \mathbb{N}$ and $h \in \mathbb{R}^d$ with $|h|^2 \leqslant (s-t)/4$,

$$\begin{aligned}
\Big|\nabla_y^j Z_0(t, x; s, y+h) - \nabla_y^j Z_0(t, x; s, y)\Big| &\leqslant |h| \sup_{\varrho \in [0,1]} \big|\nabla_y^{j+1} Z_0(t, x; s, y + \varrho h)\big|\\
&\leqslant C|h|(s-t)^{-(j+1)/2} \sup_{\varrho \in [0,1]} \Gamma^{\text{heat}}(\mu \mathcal{I}_{s-t}, \gamma_{t,s}(x) - (y + \varrho h))\\
&\leqslant C'|h|(s-t)^{-(j+1)/2} \Gamma^{\text{heat}}(\mu \mathcal{I}_{s-t}, \gamma_{t,s}^{(1)}(x) - y), \qquad \text{(C.4)}
\end{aligned}$$

using Lemma 4.1.3 for the last step. On the other hand, we also have

$$\Big|\nabla_y^j Z_0(t, x; s, y+h)\Big| \leqslant C(s-t)^{-j/2} \Gamma^{\text{heat}}(\mu \mathcal{I}_{s-t}, \gamma_{t,s}^{(1)}(x) - y).$$

Thus, by interpolation, we get for any $\theta \in (0, 1)$,

$$\Big|\nabla_y^j Z_0(t, x; s, y+h) - \nabla_y^j Z_0(t, x; s, y)\Big| \leqslant C|h|^\theta (t-s)^{-(j+\theta)/2} \Gamma^{\text{heat}}(\mu \mathcal{I}_{s-t}, \gamma_{t,s}^{(1)}(x) - y).$$

Hence,

$$|I_1(t, s, x, x', y)| \leqslant C|x - x'|^\theta (s-t)^{-1-\frac{\theta}{2}} \Gamma_\mu(t, x; s, y).$$

To treat the remaining terms, we take $\bar{t} = t + |x-x'|^2$. We have the following claim:

$$\left|\widetilde{p}^{t,x}(t,x;\bar{t},y) - \widetilde{p}^{t,x'}(t,x;\bar{t},y)\right| \leqslant C\big(|x-x'|^\alpha + |x-x'|^\beta\big)\Gamma_\mu(t,x;\bar{t},y). \tag{C.5}$$

Indeed, by Lemma 4.1.3, there is a constant $C = C(\Theta)$ such that

$$|\gamma_{t,r}(x) - \gamma_{t,r}(x')| \leqslant C\left(|x-x'| + |t-r|\right), \quad x, x' \in \mathbb{R}^d,\ r \in [t,\bar{t}].$$

Recalling $\vartheta^{t_0,x_0}_{t,\bar{t}} = \int_t^{\bar{t}} b(r,\gamma_{t_0,r}(x_0))dr$ from the notations of Section 4.2.1, we have:

$$\begin{aligned}|\vartheta^{t,x}_{t,\bar{t}} - \vartheta^{t,x'}_{t,\bar{t}}| &\leqslant \int_t^{\bar{t}} |b(r,\gamma_{t,r}(x)) - b(r,\gamma_{t,r}(x'))|dr \\ &\leqslant \lambda_2 \int_t^{\bar{t}} |\gamma_{t,r}(x) - \gamma_{t,r}(x')|^\beta dr \leqslant C|x-x'|^{2+\beta},\end{aligned}$$

where the last step is due to $|r-t| \leqslant |x-x'|^2 \leqslant |s-t|/4$. Then desired claim (C.5) follows by (4.2.7), reasoning as in the proof of Lemma 2.6.19.

Now, integrating by parts, we get from (4.3.10), (C.5) and Lemma 4.2.5

$$\begin{aligned}|I_2(t,s,x,x',y)| &\leqslant \int_{\mathbb{R}^d} |\widetilde{p}^{t,x}(t,x;\bar{t},z) - \widetilde{p}^{t,x'}(t,x;\bar{t},z)| \cdot |\nabla_z^2 p(\bar{t},z;s,y)|dz, \\ &\leqslant C\big(|x-x'|^\alpha + |x-x'|^\beta\big)(\bar{t}-t)^{-1}\int_{\mathbb{R}^d} \Gamma_\mu(t,x;\bar{t},z)\Gamma_{\mu'}(\bar{t},z;s,y)dz \\ &\leqslant C'\big(|x-x'|^\alpha + |x-x'|^\beta\big)(s-t)^{-1}\Gamma_{\mu''}(t,x;s,y).\end{aligned}$$

For I_3, by (C.4) and using arguments completely similar to those of Section 4.3.2, we have

$$|I_3(t,s,x,x',y)| \leqslant C|x-x'|^\alpha (s-t)^{-1}\Gamma_{\mu''}(t,x;s,y).$$

Finally, for I_4, from (4.3.2), we have

$$\begin{aligned}|I_4(t,s,x,x',y)| &\leqslant C\int_t^{\bar{t}}\int_{\mathbb{R}^d} \frac{(\Gamma_{\mu_1}(t,x;r,z) + \Gamma_{\mu_1}(t,x';r,z))\Gamma_{\mu_2}(r,z;s,y)}{(r-t)^{1-\frac{\alpha}{2}}(s-r)}dzdr \\ &\leqslant C'\big(\Gamma_{\mu_3}(t,x;s,y) + \Gamma_{\mu_3}(t,x';s,y)\big)\int_t^{\bar{t}} \frac{dr}{(r-t)^{1-\frac{\alpha}{2}}(s-r)} \\ &\leqslant C''\frac{|x-x'|^\alpha}{s-t}\big(\Gamma_{\mu_3}(t,x;s,y) + \Gamma_{\mu_3}(t,x';s,y)\big),\end{aligned}$$

where we have used that $|x-x'|^2 \leqslant (s-t)/4$ and $\bar{t} = t + |x-x'|$. Combining the above calculations, we obtain (C.1).

$\square$

Bibliography

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

[2] Aronson, D. G. Bounds for the fundamental solution of a parabolic equation. Bull. Amer. Math. Soc. 73 (1967), 890–896.

[3] Bass, R. F., and Perkins, E. A new technique for proving uniqueness for martingale problems. Astérisque, 327 (2009), 47–53 (2010).

[4] Bensoussan, A., and Temam, R. Équations stochastiques du type Navier-Stokes. J. Functional Analysis 13 (1973), 195–222.

[5] Bonfiglioli, A., Lanconelli, E., and Uguzzoni, F. Fundamental solutions for non-divergence form operators on stratified groups. Transactions of the American Mathematical Society 356, 7 (2004), 2709–2737.

[6] Boué, M., and Dupuis, P. A variational representation for certain functionals of Brownian motion. Ann. Probab. 26, 4 (1998), 1641–1659.

[7] Bramanti, M., and Polidoro, S. Fundamental solutions for Kolmogorov-Fokker-Planck operators with time-depending measurable coefficients. arXiv:2002.12042 (2020).

[8] Brezis, H. Analyse fonctionnelle. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1983. Théorie et applications. [Theory and applications].

[9] Chen, Z., Hu, E., Xie, L., and Zhang, X. Heat kernels for non-symmetric diffusion operators with jumps. J. Differential Equations 263 (2017), 6576–6634.

[10] Chow, P. L. Stochastic partial differential equations in turbulence related problems. In Probabilistic analysis and related topics, Vol. 1. 1978, pp. 1–43.

[11] Chow, P. L., and Jiang, J.-L. Stochastic partial differential equations in Hölder spaces. Probab. Theory Related Fields 99, 1 (1994), 1–27.

[12] Dawson, D. A. Stochastic evolution equations and related measure processes. J. Multivariate Anal. 5 (1975), 1–52.

[13] Deck, T., and Kruse, S. Parabolic differential equations with unbounded coefficients - a generalization of the parametrix method. Acta Appl. Math. 74, 1 (2002), 71–91.

[14] Delarue, F., and Menozzi, S. Density estimates for a random noise propagating through a chain of differential equations. J. Funct. Anal. 259, 6 (2010), 1577–1630.

[15] Denis, L., Matoussi, A., and Stoica, L. L^p estimates for the uniform norm of solutions of quasilinear SPDE's. Probab. Theory Related Fields 133, 4 (2005), 437–463.

[16] Di Francesco, M., and Pascucci, A. On a class of degenerate parabolic equations of Kolmogorov type. AMRX Appl. Math. Res. Express 3 (2005), 77–116.

[17] Di Francesco, M., and Pascucci, A. A continuous dependence result for ultraparabolic equations in option pricing. J. Math. Anal. Appl. 336, 2 (2007), 1026–1041.

[18] Di Francesco, M., and Polidoro, S. Schauder estimates, Harnack inequality and Gaussian lower bound for Kolmogorov-type operators in non-divergence form. Adv. Differential Equations 11, 11 (2006), 1261–1320.

[19] Du, K., and Liu, J. A Schauder estimate for stochastic PDEs. C. R. Math. Acad. Sci. Paris 354, 4 (2016), 371–375.

[20] Fabes, E. B., and Stroock, D. W. A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash. Arch. Rational Mech. Anal. 96, 4 (1986), 327–338.

[21] Fleming, W. H., and Rishel, R. W. Deterministic and stochastic optimal control. Springer-Verlag, Berlin-New York, 1975. Applications of Mathematics, No. 1.

[22] Fleming, W. H., and Sheu, S. J. Stochastic variational formula for fundamental solutions of parabolic PDE. Appl. Math. Optim. 13, 3 (1985), 193–204.

[23] Folland, G. B., and Stein, E. M. Hardy spaces on homogeneous groups, vol. 28 of Mathematical Notes. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1982.

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

[25] Friedman, A. Stochastic differential equations. Chapmann-Hall, 1975.

[26] Gobet, E. LAN property for ergodic diffusions with discrete observations. Ann. Inst. H. Poincaré Probab. Statist. 38, 5 (2002), 711–737.

[27] Hörmander, L. Hypoelliptic second order differential equations. Acta Math. 119 (1967), 147–171.

[28] Il′ in, A. M., Kalahsnikov, A. S., and Oleinik, O. A. Second-order linear equations of parabolic type. Uspehi Mat. Nauk 17, 3 (105) (1962), 3–146.

[29] Kallianpur, G. Stochastic filtering theory, vol. 13 of Applications of Mathematics. Springer-Verlag, New York-Berlin, 1980.

[30] Keller, J. B. Stochastic equations and wave propagation in random media. In Proc. Sympos. Appl. Math., Vol. XVI (1964), Amer. Math. Soc., Providence, R.I., pp. 145–170.

[31] Kolmogorov, A. Zufallige Bewegungen. (Zur Theorie der Brownschen Bewegung.). Ann. of Math., II. Ser. 35 (1934), 116–117.

[32] Konakov, V., Menozzi, S., and Molchanov, S. Explicit parametrix and local limit theorems for some degenerate diffusion processes. Ann. Inst. Henri Poincaré Probab. Stat. 46, 4 (2010), 908—923.

[33] Krylov, N. V. The selection of a Markov process from a Markov system of processes, and the construction of quasidiffusion processes. Izv. Akad. Nauk SSSR Ser. Mat. 37 (1973), 691–708.

[34] Krylov, N. V. On L_p-theory of stochastic partial differential equations in the whole space. SIAM J. Math. Anal. 27, 2 (1996), 313–340.

[35] Krylov, N. V. An analytic approach to SPDEs. In Stochastic partial differential equations: six perspectives, vol. 64 of Math. Surveys Monogr. Amer. Math. Soc., Providence, RI, 1999, pp. 185–242.

[36] Krylov, N. V. Hörmander's theorem for parabolic equations with coefficients measurable in the time variable. SIAM J. Math. Anal. 46, 1 (2014), 854–870.

[37] Krylov, N. V. Hörmander's theorem for stochastic partial differential equations. Algebra i Analiz 27, 3 (2015), 157–182.

[38] Krylov, N. V. Hypoellipticity for filtering problems of partially observable diffusion processes. Probab. Theory Related Fields 161, 3-4 (2015), 687–718.

[39] Krylov, N. V., and Rozovskii, B. L. The Cauchy problem for linear stochastic partial differential equations. Izv. Akad. Nauk SSSR Ser. Mat. 41, 6 (1977), 1329–1347, 1448.

[40] Krylov, N. V., and Rozovsky, B. L. On the first integrals and Liouville equations for diffusion processes. In Stochastic differential systems (Visegrád, 1980), vol. 36 of Lecture Notes in Control and Information Sci. Springer, Berlin-New York, 1981, pp. 117–125.

[41] Krylov, N. V., and Rozovsky, B. L. Characteristics of second-order degenerate parabolic Itô equations. Trudy Sem. Petrovsk., 8 (1982), 153–168.

[42] Krylov, N. V., and Zatezalo, A. A direct approach to deriving filtering equations for diffusion processes. Appl. Math. Optim. 42, 3 (2000), 315–332.

[43] Kunita, H. On backward stochastic differential equations. Stochastics 6, 3-4 (1981/82), 293–313.

[44] Kunita, H. Stochastic flows and stochastic differential equations, vol. 24 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 1990.

[45] Lanconelli, E., and Polidoro, S. On a class of hypoelliptic evolution operators. Rend. Sem. Mat. Univ. Politec. Torino 52, 1 (1994), 29–63.

[46] Lunardi, A. Schauder estimates for a class of degenerate elliptic and parabolic operators with unbounded coefficients in $\mathbb{R}^N$. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 24, 1 (1997), 133–164.

[47] Manfredini, M. The Dirichlet problem for a class of ultraparabolic equations. Adv. Differential Equations 2, 5 (1997), 831–866.

[48] Menozzi, S. Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. Electron. Commun. Probab. 16 (2011), 234–250.

[49] Menozzi, S., Pesce, A., and Zhang, X. Density and gradient estimates for non degenerate brownian sdes with unbounded measurable drift. Journal of Differential Equations 272 (2021), 330 – 369.

[50] Mikulevicius, R. On the Cauchy problem for parabolic SPDEs in Hölder classes. Ann. Probab. 28, 1 (2000), 74–103.

[51] Nualart, D. The Malliavin Calculus and Related Topics. Springer, 2006.

[52] Pagès, G., and Panloup, F. Total Variation and Wasserstein bounds for the ergodic Euler-Naruyama scheme for diffusions. Preprint (2020).

[53] Pagliarani, S., Pascucci, A., and Pignotti, M. Intrinsic Taylor formula for Kolmogorov-type homogeneous groups. J. Math. Anal. Appl. 435, 2 (2016), 1054–1087.

[54] Pardoux, E. Stochastic partial differential equations and filtering of diffusion processes. Stochastics 3, 2 (1979), 127–167.

[55] Pascucci, A., and Pesce, A. On stochastic Langevin and Fokker-Planck equations: the two-dimensional case. arXiv:1910.05301 (2019).

[56] Pascucci, A., and Pesce, A. Backward and forward filtering under the weak Hörmander condition. arXiv:2006.13325 (2020).

[57] Pascucci, A., and Pesce, A. The parametrix method for parabolic spdes. Stochastic Processes and their Applications 130, 10 (2020), 6226 – 6245.

[58] Polidoro, S. On a class of ultraparabolic operators of Kolmogorov-Fokker-Planck type. Matematiche (Catania) 49, 1 (1994), 53–105.

[59] Polidoro, S. A global lower bound for the fundamental solution of Kolmogorov-Fokker-Planck equations. Arch. Rational Mech. Anal. 137, 4 (1997), 321–340.

[60] Qiu, J. Hörmander-type theorem for Itô processes and related backward SPDEs. Bernoulli 24, 2 (2018), 956–970.

[61] Rozovskii, B. L. Stochastic partial differential equations. Mat. Sb. (N.S.) 96(138) (1975), 314–341, 344.

[62] Rozovsky, B. L., and Lototsky, S. V. Stochastic evolution systems, vol. 89 of Probability Theory and Stochastic Modelling. Springer, Cham, 2018. Linear theory and applications to non-linear filtering, Second edition of [MR1135324].

[63] Sheu, S. J. Some estimates of the transition density of a nondegenerate diffusion Markov process. Ann. Probab. 19, 2 (1991), 538–561.

[64] Shimizu, A. Fundamental solutions of stochastic partial differential equations arising in nonlinear filtering theory. In Probability theory and mathematical statistics (Tbilisi, 1982), vol. 1021 of Lecture Notes in Math. Springer, Berlin, 1983, pp. 594–602.

[65] Sowers, R. B. Recent results on the short-time geometry of random heat kernels. Math. Res. Lett. 1, 6 (1994), 663–675.

[66] Sowers, R. B. Short-time geometry of random heat kernels. Mem. Amer. Math. Soc. 132, 629 (1998), viii+130.

[67] Stroock, D. W., and Varadhan, S. R. S. Multidimensional diffusion processes, vol. 233 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1979.

[68] Veretennikov, A. Y. "Inverse diffusion" and direct derivation of stochastic Liouville equations. Mat. Zametki 33, 5 (1983), 773–779.

[69] Veretennikov, A. Y. On backward filtering equations for SDE systems (direct approach). In Stochastic partial differential equations (Edinburgh, 1994), vol. 216 of London Math. Soc. Lecture Note Ser. Cambridge Univ. Press, Cambridge, 1995, pp. 304–311.

[70] Veretennikov, A. Y. On SPDE and backward filtering equations for SDE systems (direct approach). https://arxiv.org/abs/1607.00333 (July 2016).

[71] Wang, F. Y., and Zhang, X. Derivative formula and applications for degenerate diffusion semigroups. J. Math. Pures Appl. 99 (2013), 726–740.

[72] Zakai, M. On the optimal filtering of diffusion processes. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 11 (1969), 230–243.

[73] Zatezalo, A. Filtering of partially observable stochastic processes. ProQuest LLC, Ann Arbor, MI, 1998. Thesis (Ph.D.)–University of Minnesota.

[74] Zhang, X. A variational representation for random functionals on abstract Wiener spaces. J. Math. Kyoto Univ. 49, 3 (2009), 475–490.

[75] Zhang, X. Fundamental solutions of nonlocal Hörmander's operators. Commun. Math. Stat. 4 (2016), 359–402.