Backward and forward filtering under the weak Hörmander condition

(Article begins on next page)

11 September 2024

Backward and forward filtering under the weak Hörmander condition

Andrea Pascucci · **Antonello Pesce**

Abstract We derive the forward and backward filtering equations for a class of degenerate partially observable diffusions, satisfying the weak Hörmander condition. Our approach is based on the Hölder theory for degenerate SPDEs that allows to pursue the direct approaches proposed by N. V. Krylov and A. Zatezalo, and A. Yu. Veretennikov, avoiding the use of general results from filtering theory. As a by-product we also provide existence, regularity and estimates for the filtering density.

Keywords filtering · Hölder theory of SPDEs · Langevin equation · weak Hörmander condition

Mathematics Subject Classification (2010) 60G35 · 60H15 · 60J60 · 35H20

1 Introduction

The classical kinetic model

$$
\begin{cases}
dX_t = V_t dt, \\
dV_t = \sigma dW_t, \qquad \sigma > 0,
\end{cases}
\tag{1.1}
$$

is a remarkable example of a system of SDEs whose Kolmogorov equation

$$
\frac{\sigma^2}{2}\partial_{vv}f + v\partial_x f + \partial_t f = 0, \qquad (t,x,v) \in \mathbb{R}^3,
\tag{1.2}
$$

is hypoelliptic but not uniformly parabolic. Precisely, (1.2) satisfies the *weak* Hörmander condition in that the drift plays a key role in the noise propagation (see [14] and the introduction in [12]). In (1.1) W is a Brownian motion and X, V represent position and velocity of a particle. This type of SDEs arises in several

A. Pascucci

Department of Mathematics, University of Bologna, Piazza di Porta san Donato, 5 Bologna

Tel.: +39-0512094428

E-mail: andrea.pascucci@unibo.it

A. Pesce

Department of Mathematics, University of Bologna, Piazza di Porta san Donato, 5 Bologna

E-mail: antonello.pesce2@unibo.it

linear and non-linear models in physics (see, for instance, [4], [23], [8], [11]) and in mathematical finance (see, for instance, [2], [26]).

In this paper we study the filtering problem for (1.1). To the best of our knowledge, this kind of problem was never considered in the literature, possibly because the known results for hypoelliptic SPDEs (e.g. [5], [21], [16], [30] and [29]) do not apply in this case. Here we propose a unified approach for the derivation of the *backward and forward filtering equations* based on the Hölder theory for degenerate SPDEs recently developed in [28] and [27] (see also [6] and [24] for similar results for uniformly parabolic SPDEs). Having an existence and regularity theory at hand, we can pursue the "direct" approaches proposed by Krylov and Zatezalo [19] and Veretennikov [33], thus avoiding the use of general results from filtering theory. In particular, as in [33] 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 [25], [13], [22] or [30].

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

$$\begin{cases} dX_t = V_t dt, \\ dV_t = b(t, X_t, V_t, Y_t)dt + \bar{\sigma}(t, X_t, V_t, Y_t)dW_t^1 + \hat{\sigma}(t, X_t, V_t, Y_t)dW_t^2, \\ dY_t = h(t, X_t, V_t, Y_t)dt + \theta(t, Y_t)dW_t^1, \end{cases} \tag{1.3}$$

where $W_t = (W_t^1, W_t^2)$ denotes a bi-dimensional Brownian motion 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 $\mathcal{F}_{t,T}^Y$-measurable least-square estimate of $\varphi(Z_T)$, that is the conditional expectation $E\left[\varphi(Z_T) \mid \mathcal{F}_{t,T}^Y\right]$. Our first result, Theorem 8 shows that

$$E\left[\varphi(Z_T^{t,z}) \mid \mathcal{F}_{t,T}^Y\right] = \int_{\mathbb{R}^2} \hat{\mathbf{\Gamma}}(t, z; T, \zeta)\varphi(\zeta)d\zeta,$$

where $\hat{\mathbf{\Gamma}}$ is the (normalized) fundamental solution of the *forward filtering equation*; the latter is a SPDE of the form

$$d_{\mathbf{B}} u_s(\zeta) = \mathcal{A}_{s,\zeta} u_s(\zeta)ds + \mathcal{G}_{s,\zeta} u_s(\zeta)dW_s^1 \tag{1.4}$$

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

$$\mathcal{A}_{s,\zeta} u_s(\zeta) = \frac{1}{2}\left(\bar{\sigma}^2 + \hat{\sigma}^2\right)(s, \zeta, Y_s)\partial_{\nu\nu} u_s(\zeta) + \text{``lower order terms''},$$

$$\mathcal{G}_{s,\zeta} u_s(\zeta) = \bar{\sigma}(s, \zeta, Y_s)\partial_\nu u_s(\zeta) + \text{``lower order terms''}.$$

The forward filtering SPDE is precisely formulated in (3.4). The symbol $d_{\mathbf{B}}$ in (1.4) indicates that the SPDE is understood in the Itô (or strong) sense, that is

$$u_s\left(\gamma_{s-t}^{\mathbf{B}}(\zeta)\right) = u_t(\zeta) + \int_t^s \mathcal{A}_{\tau, \gamma_{\tau-t}^{\mathbf{B}}(\zeta)} u_\tau(\gamma_{\tau-t}^{\mathbf{B}}(\zeta))d\tau + \int_t^s \mathcal{G}_{\tau, \gamma_{\tau-t}^{\mathbf{B}}(\zeta)} u_\tau(\gamma_{\tau-t}^{\mathbf{B}}(\zeta))dW_\tau^1, \qquad s \in [t, T],$$

where $s \mapsto \gamma_s^{\mathbf{B}}(\xi, \nu)$ denotes the integral curve, starting from (ξ, ν), of the advection vector field $\nu\partial_\xi$ or, more explicitly, $\gamma_s^{\mathbf{B}}(\xi, \nu) = (\xi + s\nu, \nu)$.

Example 1 The prototype of (1.4) is the Langevin SPDE

$$d_{\mathbf{B}}u_s(\xi,\nu) = \frac{\sigma^2}{2}\partial_{\nu\nu}u_s(\xi,\nu)ds + \beta\partial_\nu u_s(\xi,\nu)dW_s^1, \qquad (1.5)$$

with σ, β constant parameters. Clearly, if $u_s = u_s(\xi,\nu)$ is a smooth function then (1.5) can be written in the usual Itô form

$$du_s(\xi,\nu) = \left(\frac{\sigma^2}{2}\partial_{\nu\nu}u_s(\xi,\nu) - \nu\partial_\xi u_s(\xi,\nu)\right)ds + \beta\partial_\nu u_s(\xi,\nu)dW_s^1.$$

Notice that ∂_ξ, being equal to the Lie bracket $[\partial_\nu, \mathbf{B}]$, has to be regarded as a *third order derivative* in the intrinsic sense of subelliptic operators (cf. [10]): this motivates the use of the "Lie stochastic differential" $d_{\mathbf{B}}$ instead of the standard Itô differential in (2.1). Notice also that (1.5) reduces to the forward Kolmogorov (or Fokker-Planck) equation for (1.1) when $\beta = 0$.

Analogously, in Section 3.2 we prove that

$$E\left[\varphi(Z_T^{t,z,y}, Y_T^{t,z,y}) \mid \mathcal{F}_{t,T}^Y\right] = \int_{\mathbb{R}^3} \bar{\mathbf{\Gamma}}(t,z,y;T,\zeta,\eta)\varphi(\zeta,\eta)d\zeta d\eta, \qquad (t,z,y) \in [0,T] \times \mathbb{R}^2 \times \mathbb{R},$$

where $\bar{\mathbf{\Gamma}}$ denotes the (normalized) fundamental solution of the *backward filtering equation* that is a SPDE of the form

$$-d_{\mathbf{B}}u_t(z,y) = \widetilde{\mathcal{A}}_t u_t(z,y)dt + \widetilde{\mathcal{G}}_t u_t(z,y) \star dW_t^1. \qquad (1.6)$$

We refer to (3.14) for the precise formulation of the backward filtering SPDE. The symbol $\star$ means that (1.6) is written in terms of the *backward Itô integral* whose definition is recalled in Section 5 for reader's convenience. We shall see that the coefficients of the forward filtering SPDE are random, while the coefficients of the backward filtering SPDE are deterministic. Moreover, (1.4) is posed in $\mathbb{R}^3$ (including the time variable) while (1.6) is posed in $\mathbb{R}^4$.

The rest of the paper is organized as follows. In Section 2 we resume and extend the Hölder theory for degenerate SPDEs satisfying the weak Hörmander condition, developed in [28] and [27]. In Section 3, which is the core of the paper, we state the filtering problem and derive the forward and backward filtering SPDEs. Section 4 contains the proof of the results about the existence and Gaussian estimates for the fundamental solutions of the filtering SPDEs. In Section 5 we review the definition and some basic result about backward stochastic integration. For reader's convenience, in Section 6 we collect the main notations systematically used throughout the paper.

2 Fundamental solution of Langevin-type SPDEs

We present the Hölder theory for degenerate SPDEs that will be used in the derivation of the filtering equations. Compared to [27], here we state our results in a slightly more general setting where the dimension of the non degenerate variable v can be possibly greater than one. This is done with the purpose of being able to handle differential operators constructed from the generator of the full process (X_t, V_t, Y_t) which will appear in Section 3, both in the derivation of the forward and the backward filtering SPDE, and it is not related to the number n of Brownian motions considered in the model (1.3). On the other hand

the dimension lift does not bring any additional difficulty in the analysis since it is performed in the non degenerate directions.

We first introduce some general notation and the functional spaces used throughout the paper.

We denote by $z = (x, v_1, \ldots, v_d)$ and $\zeta = (\xi, \nu_1, \ldots, \nu_d)$ the points in $\mathbb{R} \times \mathbb{R}^d$. Moreover, for any $k \in \mathbb{N}$, $0 < \alpha < 1$ and $0 \leqslant t < T$,

i) $m\mathcal{B}_{t,T}$ (resp. $b\mathcal{B}_{t,T}$) is the space of all real-valued (resp. bounded) Borel measurable functions $f = f_s(z)$ on $[t, T] \times \mathbb{R}^{d+1}$;

ii) $C^0_{t,T}$ (resp. $bC^0_{t,T}$) is the space of functions $f \in m\mathcal{B}_{t,T}$ (resp. $f \in b\mathcal{B}_{t,T}$) that are continuous in z and $C^\alpha_{t,T}$ (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 z uniformly with respect to s, that is

$$\sup_{\substack{s \in [t,T] \\ z \neq \zeta}} \frac{|f_s(z) - f_s(\zeta)|}{|z - \zeta|^\alpha} < \infty.$$

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 z uniformly with respect to $s \in [t, T]$;

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

We use boldface to denote the stochastic version of the previous functional spaces. Let $(W_t)_{t \in [0,T]}$ be a one-dimensional Brownian motion on a complete probability space $(\Omega, \mathcal{F}, P)$, endowed with a filtration $(\mathcal{F}_t)_{t \in [0,T]}$ satisfying the usual conditions, and let $\mathcal{P}_{t,T}$ be the predictable σ-algebra on $[t, T] \times \Omega$.

Definition 1 We denote by $\mathbf{C}^{k+\alpha}_{t,T}$ the family of functions $f = f_s(z, \omega)$ on $[t, T] \times \mathbb{R}^{d+1} \times \Omega$ such that:

i) $(s, z) \mapsto f_s(z, \omega) \in C^{k+\alpha}_{t,T}$ for any $\omega \in \Omega$;

ii) $(s, \omega) \mapsto f_s(z, \omega)$ is $\mathcal{P}_{t,T}$-measurable for any $z \in \mathbb{R}^{d+1}$.

Similarly, we define $\mathbf{b}\mathbf{C}^{k+\alpha}_{t,T}$.

We consider a class of degenerate SPDEs of the form

$$d_{\mathbf{B}} u_s(\zeta) = \mathcal{A}_{s,\zeta} u_s(\zeta) ds + \mathcal{G}_{s,\zeta} u_s(\zeta) dW_s \tag{2.1}$$

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

$$\mathcal{A}_{s,\zeta} u_s(\zeta) := \frac{1}{2} a^{ij}_s(\zeta) \partial_{\nu_i \nu_j} u_s(\zeta) + b^i_s(\zeta) \partial_{\nu_i} u_s(\zeta) + c_s(\zeta) u_s(\zeta),$$

$$\mathcal{G}_{s,\zeta} u_s(\zeta) := \sigma^i_s(\zeta) \partial_{\nu_i} u_s(\zeta) + h_s(\zeta) u_s(\zeta).$$

Definition 2 A solution to (2.1) on $[t, T]$ is a process $u = u_s(\xi, \nu) \in \mathbf{C}^0_{t,T}$ that is twice continuously 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 $\gamma^{\mathbf{B}}_s(\xi, \nu) = (\xi + s\nu_1, \nu)$.

Definition 3 A fundamental solution of the forward SPDE (2.1) is a stochastic process $\mathbf{\Gamma} = \mathbf{\Gamma}(t, z; s, \zeta)$, defined for $0 \leqslant t < s \leqslant T$ and $z, \zeta \in \mathbb{R}^{d+1}$, such that for any $(t, z) \in [0, T) \times \mathbb{R}^{d+1}$ and $t_0 \in (t, T)$ we have:

i) $\mathbf{\Gamma}(t, z; \cdot, \cdot)$ is a solution to (2.1) on $[t_0, T]$;
ii) for any $\varphi \in bC(\mathbb{R}^{d+1})$ and $z_0 \in \mathbb{R}^{d+1}$, we have

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

In [27], under suitable assumptions on the coefficients, we proved existence and Gaussian-type estimates of a fundamental solution for (2.1) when $b_s \equiv c_s \equiv h_s \equiv 0$ and $d = 1$. Here we slightly extend those results to an SPDE of the general form (2.1) and to the backward version of it, that is

$$-d_{\mathbf{B}}u_t(z) = \mathcal{A}_{t,z}u_t(z)dt + \mathcal{G}_{t,z}u_t(z) \star dW_t, \qquad \mathbf{B} = \partial_t + v_1\partial_x. \tag{2.2}$$

We denote by $\overset{\leftarrow}{\mathbf{C}}_{t,T}^{k+\alpha}$ (and $\mathbf{b}\overset{\leftarrow}{\mathbf{C}}_{t,T}^{k+\alpha}$) the stochastic Hölder spaces formally defined as in Definition 1 with $\mathcal{P}_{t,T}$ in condition ii) replaced by the backward predictable σ-algebra $\overset{\leftarrow}{\mathcal{P}}_{t,T}$ defined in terms of the backward Brownian filtration (cf. Section 5). Again, (2.2) is understood in the strong sense:

Definition 4 A solution to (2.2) on $[0, s]$ is a process $u = u_t(x, v) \in \overset{\leftarrow}{\mathbf{C}}_{0,s}^0$ that is twice continuously differentiable in the variables v and such that

$$u_t\left(\gamma_{s-t}^{\mathbf{B}}(z)\right) = u_s(z) + \int_t^s \mathcal{A}_{\tau,\gamma_{s-\tau}^{\mathbf{B}}(z)}u_\tau(\gamma_{s-\tau}^{\mathbf{B}}(z))d\tau + \int_t^s \mathcal{G}_{\tau,\gamma_{s-\tau}^{\mathbf{B}}(z)}u_\tau(\gamma_{s-\tau}^{\mathbf{B}}(z)) \star dW_\tau, \qquad t \in [0, s].$$

Definition 5 A fundamental solution for the backward SPDE (2.2) is a stochastic process $\overset{\leftarrow}{\mathbf{\Gamma}} = \overset{\leftarrow}{\mathbf{\Gamma}}(t, z; s, \zeta)$ 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) $\overset{\leftarrow}{\mathbf{\Gamma}}(\cdot, \cdot; s, \zeta)$ is a solution to (2.2) on $[0, t_0]$;
ii) for any $\varphi \in bC(\mathbb{R}^{d+1})$ and $z_0 \in \mathbb{R}^{d+1}$, we have

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

Next we pose the standing assumptions on the coefficients of (2.1) and (2.2).

Assumption 1 (Regularity) *For some $\alpha \in (0, 1)$, we have:*

i) $a \in \mathbf{bC}_{0,T}^\alpha$, $\sigma \in \mathbf{bC}_{0,T}^{3+\alpha}$, $b, c \in \mathbf{bC}_{0,T}^0$, $h \in \mathbf{bC}_{0,T}^2$ *in the forward SPDE (2.1);*
ii) $a \in \mathbf{b}\overset{\leftarrow}{\mathbf{C}}_{0,T}^\alpha$, $\sigma \in \mathbf{b}\overset{\leftarrow}{\mathbf{C}}_{0,T}^{3+\alpha}$, $b, c \in \mathbf{b}\overset{\leftarrow}{\mathbf{C}}_{0,T}^0$, $h \in \mathbf{b}\overset{\leftarrow}{\mathbf{C}}_{0,T}^2$ *in the backward SPDE (2.2).*

Assumption 2 (Coercivity) *There exists a random, finite and positive constant $\mathbf{m}$ such that*

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

In our analysis we make use of the Itô-Wentzell transform. Let $(x, v) \in \mathbb{R}^{d+1}$. For a fixed $t \in [0, T)$ we consider the SDE in $\mathbb{R}^d$

$$\gamma_{t,s}^{\mathrm{IW}}(x, v) = v - \int_t^s \sigma_\tau(x, \gamma_{t,\tau}^{\mathrm{IW}}(x, v))dW_\tau, \qquad s \in [t, T], \tag{2.3}$$

and, for a fixed $s \in (0, T]$, the SDE

$$\breve{\gamma}_{t,s}^{\mathrm{IW}}(x,v) = v + \int_t^s \sigma_\tau(x, \breve{\gamma}_{\tau,s}^{\mathrm{IW}}(x,v)) \star dW_\tau, \qquad t \in [0, s]. \tag{2.4}$$

Assumption 1 ensures that (2.3) and (2.4) have strong solutions and the maps $(x,v) \mapsto (x, \gamma_{t,s}^{\mathrm{IW}}(x,v))$ and $(x,v) \mapsto (x, \breve{\gamma}_{t,s}^{\mathrm{IW}}(x,v))$ define forward and backward flows of diffeomorphisms of $\mathbb{R}^{d+1}$ respectively. These changes of coordinates allow to transform the SPDEs (2.1) and (2.4) into PDEs with random coefficients whose properties depend on the gradient of the stochastic flow: to have a control on it, we impose the some additional condition. For any suitably regular function $f = f(w) : \mathbb{R}^N \longrightarrow \mathbb{R}$, $\varepsilon > 0$ and multi-index $\beta \in \mathbb{N}_0^N$, we set

$$\langle f \rangle_{\varepsilon, \beta} := \sup_{w \in \mathbb{R}^N} (1 + |w|^2)^\varepsilon |\partial_w^\beta f(w)|. \tag{2.5}$$

Assumption 3 *There exist $\varepsilon > 0$ and two random variables $M_1 \in L^p(\Omega)$, with $p > \max\{2, \frac{1}{\varepsilon}\}$, and $M_2 \in L^\infty(\Omega)$ such that with probability one*

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

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

Assumption 3 requires that $\sigma_t(z)$ and $h_t(z)$ flatten as $z \to \infty$. In particular, this condition is clearly satisfied if σ and h depend only on t or, more generally, if the spatial gradients of σ and h have compact support.

In order to state the main result of this section, Theorem 4 below, we need to introduce some additional notation: we consider the Gaussian kernel

$$\Gamma_\lambda(t, x, v) = \frac{1}{t^{\frac{d+3}{2}}} \exp\left(-\frac{1}{2\lambda} \left(\frac{x^2}{t^3} + \frac{|v|^2}{t} \right) \right), \qquad t > 0, \ (x, v) \in \mathbb{R} \times \mathbb{R}^d, \ \lambda > 0. \tag{2.6}$$

To fix ideas, for $d = 1$ and up to some renormalization, Γ_λ is the fundamental solution of the degenerate Langevin equation (1.2). For a recent survey on the theory of this kind of ultra-parabolic operators and the related sub-elliptic structure, we refer to [1].

In the following statement, we denote by $g^{\mathrm{IW}, -1}$ (and $\bar{g}^{\mathrm{IW}, -1}$) the inverse of the Itô-Wentzell stochastic flow $(x, v) \mapsto g^{\mathrm{IW}}(x, v) := (x, \gamma_{t,s}^{\mathrm{IW}}(x, v))$ defined by (2.3) (and $(x, v) \mapsto \bar{g}^{\mathrm{IW}}(x, v) := (x, \breve{\gamma}_{t,s}^{\mathrm{IW}}(x, v))$ defined by (2.4), respectively). Moreover, we consider the vector field

$$\mathbf{Y}_{t,s}(z) := \left((\gamma_{t,s}^{\mathrm{IW}})_1(z), -(\gamma_{t,s}^{\mathrm{IW}}(z))_1 (\nabla_v \gamma_{t,s}^{\mathrm{IW}})^{-1}(z) \partial_x \gamma_{t,s}^{\mathrm{IW}}(z) \right), \tag{2.7}$$

with $\nabla_v \gamma^{\mathrm{IW}} = (\partial_{v_j} \gamma_i^{\mathrm{IW}})_{i,j=1,\cdots d}$ and $\partial_x \gamma^{\mathrm{IW}} = (\partial_x \gamma_i^{\mathrm{IW}})_{i=1,\cdots d}$, and define $\overleftarrow{\mathbf{Y}}_{t,s}$ analogously, namely

$$\overleftarrow{\mathbf{Y}}_{t,s}(z) := \left((\breve{\gamma}_{t,s}^{\mathrm{IW}})_1(z), -(\breve{\gamma}_{t,s}^{\mathrm{IW}}(z))_1 (\nabla_v \breve{\gamma}_{t,s}^{\mathrm{IW}})^{-1}(z) \partial_x \breve{\gamma}_{t,s}^{\mathrm{IW}}(z) \right).$$

Eventually, equation

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

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

$$\overleftarrow{\gamma}_t^{s,\zeta} = \zeta + \int_t^s \overleftarrow{\mathbf{Y}}_{\tau,s}(\overleftarrow{\gamma}_\tau^{s,\zeta})d\tau, \qquad t \in [0, s],$$

defines the integral curve of $\overleftarrow{\mathbf{Y}}_{t,s}$ ending at (s, ζ). The main result of this section is the following theorem whose proof is postponed to Section 4; for reader's convenience, in Section 6 we collect the main notations used hereafter.

Theorem 4 *Under Assumptions 1-i), 2 and 3, the forward SPDE* (2.1) *has a fundamental solution* $\mathbf{\Gamma}$ *and there exist two positive random variables* λ, μ *such that*

$$\mu^{-1}\Gamma_{\lambda^{-1}}\left(s - t, g_{t,s}^{\mathrm{IW},-1}(\zeta) - \gamma_s^{t,z}\right) \leqslant \mathbf{\Gamma}(t, z; s, \zeta) \leqslant \mu\Gamma_\lambda\left(s - t, g_{t,s}^{\mathrm{IW},-1}(\zeta) - \gamma_s^{t,z}\right), \tag{2.8}$$

$$|\partial_{\nu_i}\mathbf{\Gamma}(t, z; s, \xi, \nu)| \leqslant \frac{\mu}{\sqrt{s-t}}\Gamma_\lambda\left(s - t, g_{t,s}^{\mathrm{IW},-1}(\xi, \nu) - \gamma_s^{t,z}\right), \tag{2.9}$$

$$\left|\partial_{\nu_i\nu_j}\mathbf{\Gamma}(t, z; s, \xi, \nu)\right| \leqslant \frac{\mu}{s-t}\Gamma_\lambda\left(s - t, g_{t,s}^{\mathrm{IW},-1}(\xi, \nu) - \gamma_s^{t,z}\right), \tag{2.10}$$

for every $0 \leqslant t < s \leqslant T$, $z, \zeta = (\xi, \nu) \in \mathbb{R}^{d+1}$ *and* $i, j = 1, \cdots d$, *with probability one.*

Similarly, under Assumptions 1-ii), 2 and 3, the backward SPDE (2.2) *has a fundamental solution* $\overleftarrow{\mathbf{\Gamma}}$ *satisfying estimates*

$$\mu^{-1}\Gamma_{\lambda^{-1}}\left(s - t, \overleftarrow{g}_{t,s}^{\mathrm{IW},-1}(z) - \overleftarrow{\gamma}_t^{s,\zeta}\right) \leqslant \overleftarrow{\mathbf{\Gamma}}(t, z; s, \zeta) \leqslant \mu\Gamma_\lambda\left(s - t, \overleftarrow{g}_{t,s}^{\mathrm{IW},-1}(z) - \overleftarrow{\gamma}_t^{s,\zeta}\right), \tag{2.11}$$

$$\left|\partial_{v_i}\overleftarrow{\mathbf{\Gamma}}(t, x, v; s, \zeta)\right| \leqslant \frac{\mu}{\sqrt{s-t}}\Gamma_\lambda\left(s - t, \overleftarrow{g}_{t,s}^{\mathrm{IW},-1}(x, v) - \overleftarrow{\gamma}_t^{s,\zeta}\right), \tag{2.12}$$

$$\left|\partial_{v_iv_j}\overleftarrow{\mathbf{\Gamma}}(t, x, v; s, \zeta)\right| \leqslant \frac{\mu}{s-t}\Gamma_\lambda\left(s - t, \overleftarrow{g}_{t,s}^{\mathrm{IW},-1}(x, v) - \overleftarrow{\gamma}_t^{s,\zeta}\right), \tag{2.13}$$

for every $0 \leqslant t < s \leqslant T$, $z = (x, v), \zeta \in \mathbb{R}^{d+1}$ *and* $i, j = 1, \ldots d$, *with probability one.*

Remark 1 We would like to emphasize that Theorem 4 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 [3]: however in [3] the coefficients are assumed to be independent of the spatial variables that is a very particular case where the fundamental solution is known explicitly.

In the derivation of the forward filtering SPDE, we will use a deterministic backward Kolmogorov PDE to which Theorem 4 applies. Precisely, we will use the following

Corollary 1 *Let Assumption 2 with* $\sigma \equiv 0$ *be satisfied and let* $a \in bC_{0,T}^\alpha$, $b, c \in bC_{0,T}^0$, *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*

$$\begin{cases} -d_\mathbf{B}f(t, z) = \mathcal{A}_{t,z}f(t, z)dt, \\ f(T, \cdot) = \varphi, \end{cases} \tag{2.14}$$

in the sense of Definition 4, that is

$$f\left(t, \gamma_{T-t}^\mathbf{B}(z)\right) = \varphi(z) + \int_t^T \mathcal{A}_{s,\gamma_{T-s}^\mathbf{B}(z)}f\left(s, \gamma_{T-s}^\mathbf{B}(z)\right)ds, \qquad (t, z) \in [0, T] \times \mathbb{R}^{d+1}, \tag{2.15}$$

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,

$$\sup_{(x,v)\in\mathbb{R}\times\mathbb{R}^d} |\partial_v^\beta f(t,x,v)| \leqslant C(T-t)^{-\frac{|\beta|-\alpha}{2}}, \qquad 1 \leqslant |\beta| \leqslant 2. \tag{2.16}$$

3 The filtering problem

Consider system (1.3) and suppose that $h \equiv \theta \equiv \bar{\sigma} \equiv 0$, 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 $\hat{\sigma}$, we have

$$E\left[\varphi(Z_T^{t,z})\right] = \int_{\mathbb{R}^2} \Gamma(t,z;T,\zeta)\varphi(\zeta)d\zeta, \tag{3.1}$$

where $\Gamma = \Gamma(t,z;T,\zeta)$ is the fundamental solution of the backward Kolmogorov operator

$$\mathcal{K} = \frac{\hat{\sigma}^2}{2}\partial_{vv} + b\partial_v + v\partial_x + \partial_t, \tag{3.2}$$

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).

In this section we study the filtering problem for system (1.3) and, assuming that Y is not trivial, we prove a representation formula for $E\left[\varphi(Z_T) \mid \mathcal{F}_{t,T}^Y\right]$ that is analogous to (3.1) 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 Theorem 4. Actually, in filtering theory, the derivation itself of the filtering SPDE is a non-trivial task.

As already mentioned, in our analysis we will adopt a *direct approach*. However, we should acknowledge that there are at least two quite different *direct approaches* proposed in the literature: both of them are meant to avoid the purely probabilistic techniques of the general filtering theory.

The direct approach by Krylov and Zatezalo [19] mimics the derivation of the standard Kolmogorov operator (3.2): roughly speaking, *assuming that the filtering SPDE is known in advance*, one takes a solution u_t (whose existence is guaranteed by Theorem 4), applies the Itô formula to $u_t(Z_t)$ and finally takes expectations. This is the approach we follow in Section 3.1 to prove the existence of the *forward filtering density* and the representation of the conditional expectation $E\left[\varphi(Z_T) \mid \mathcal{F}_{t,T}^Y\right]$ in terms of it.

On the other hand, the direct approach by Veretennikov [32], [33], 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 12. We follow this approach in Section 3.2 to derive the backward filtering SPDE and the corresponding *filtering density*. Note however that in Section 3.2 we only provide an informal, yet quite detailed, derivation: a full proof is outside the scope of the present paper and would require a generalization of the results of Section 5 to degenerate diffusions. This is certainly possible but would require some additional effort and is postponed to future research.

We notice that system (1.3) can be written more conveniently as

$$\begin{cases} dZ_t = BZ_t dt + \mathbf{e}_2\left(b(t,Z_t,Y_t)dt + \bar{\sigma}(t,Z_t,Y_t)dW_t^1 + \hat{\sigma}(t,Z_t,Y_t)dW_t^2\right), \\ dY_t = h(t,Z_t,Y_t)dt + \theta(t,Y_t)dW_t^1, \end{cases} \tag{3.3}$$

with

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

Hereafter we assume the following non-degeneracy condition:

Assumption 5 (Coercivity) *There exists a positive constant m such that*

$$\theta(t,y) \geqslant m, \qquad \hat{\sigma}(t,z,y) \geqslant m, \qquad t \in [0,T], \ z \in \mathbb{R}^2, \ y \in \mathbb{R}.$$

3.1 Forward filtering SPDE

We consider the solution $(Z_s^{t,z}, Y_s)_{s \in [t,T]}$ of system (3.3) with initial condition $Z_t^{t,z} = z \in \mathbb{R}^2$; we do not impose any initial condition on the Y-component. We set $\sigma = (\bar{\sigma}, \hat{\sigma})$ and introduce the stochastic processes

$$\sigma_s(\zeta) := \sigma(s, \zeta, Y_s), \qquad \theta_s := \theta(s, Y_s), \qquad b_s(\zeta) := b_s(\zeta, Y_s), \qquad \widetilde{h}_s(\zeta) := \frac{h(s, \zeta, Y_s)}{\theta(s, Y_s)},$$

The *forward filtering SPDE* for system (3.3) reads as follows

$$d_{\mathbf{B}} v_s(\xi, \nu) = \mathcal{A}_s^* v_s(\xi, \nu) ds + \mathcal{G}_s^* v_s(\xi, \nu) \frac{dY_s}{\theta_s}, \qquad \mathbf{B} = \partial_s + \nu \partial_\xi, \qquad (3.4)$$

where $\mathcal{A}^*$ and $\mathcal{G}^*$ are the adjoints of the differential operators (with random coefficients)

$$\mathcal{A}_s := \frac{|\sigma_s(\xi, \nu)|^2}{2} \partial_{\nu\nu} + b_s(\xi, \nu) \partial_\nu, \qquad \mathcal{G}_s := \bar{\sigma}_s(\xi, \nu) \partial_\nu + \widetilde{h}_s(\xi, \nu),$$

respectively.

In order to apply to (3.4) the general results of Section 1, in particular Theorem 4 and Corollary 1, we assume the following conditions. We recall notation (2.5).

Assumption 6 (Regularity) *The coefficients of (3.3) are such that* $\bar{\sigma} \in bC_{0,T}^{3+\alpha}(\mathbb{R}^3)$, $\hat{\sigma} \in bC_{0,T}^{2+\alpha}(\mathbb{R}^3)$, $\theta \in bC_{0,T}^{\alpha}(\mathbb{R})$, $b \in bC_{0,T}^1(\mathbb{R}^3)$, $h \in bC_{0,T}^2(\mathbb{R}^3)$.

Assumption 7 (Flattening at infinity) *There exist two positive constants* ε, M *such that*

$$\sup_{\substack{t \in [0,T] \\ y \in \mathbb{R}}} \left(\langle \bar{\sigma}(t, \cdot, y) \rangle_{\varepsilon, \beta} + \langle \bar{\sigma}(t, \cdot, y) \rangle_{1/2+\varepsilon, \beta'} + \langle h(t, \cdot, y) \rangle_{1/2, \beta} \right) \leqslant M$$

for $|\beta| = 1$ *and* $|\beta'| = 2, 3$.

Remark 2 With regard to the existence of solutions to (3.4), let us introduce the process

$$\widetilde{W}_s := \int_t^s \theta_\tau^{-1} dY_\tau = W_s^1 - W_t^1 + \int_t^s \widetilde{h}_\tau(Z_\tau^{t,z}) d\tau, \qquad s \in [t,T].$$

By Girsanov's theorem, $(\widetilde{W}_s)_{s \in [t,T]}$ is a Brownian motion w.r.t the measure Q defined by $dQ = (\varrho_T^{t,z})^{-1} dP$ where

$$d\varrho_s^{t,z} = \widetilde{h}_s(Z_s^{t,z})^2 \varrho_s^{t,z} dt + \widetilde{h}_s(Z_s^{t,z}) \varrho_s^{t,z} dW_s^1, \qquad \varrho_t^{t,z} = 1. \qquad (3.5)$$

Moreover, $(\widetilde{W}_s)_{s\in[t,T]}$ is adapted to $(\mathcal{F}_{t,s}^Y)_{s\in[t,T]}$. Then, equation (3.4) can be written in the equivalent form

$$d_{\mathbf{B}}v_s(\zeta) = \mathcal{A}_s^* v_s(\zeta)ds + \mathcal{G}_s^* v_s(\zeta)d\widetilde{W}_s \qquad (3.6)$$

under Q. Under Assumptions 5, 6 and 7, by Theorem 4 a fundamental solution $\mathbf{\Gamma} = \mathbf{\Gamma}(t,z;s,\zeta)$ for (3.6) exists, satisfies estimates (2.8), (2.9), (2.10) and $s \mapsto \mathbf{\Gamma}(t,z;s,\zeta)$ is adapted to $(\mathcal{F}_{t,s}^Y)_{s\in[t,T]}$. 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 (3.3). This definition is motivated by the following

Theorem 8 *Let $(Z_s^{t,z}, Y_s)_{s\in[t,T]}$ denote the solution of system (3.3) with initial condition $Z_t^{t,z} = z$. Under Assumptions 5, 6 and 7, for any $\varphi \in bC(\mathbb{R}^2)$ we have*

$$E\left[\varphi(Z_T^{t,z}) \mid \mathcal{F}_{t,T}^Y\right] = \int_{\mathbb{R}^2} \hat{\mathbf{\Gamma}}(t,z;T,\zeta)\varphi(\zeta)d\zeta, \qquad (t,z) \in [0,T] \times \mathbb{R}^2. \qquad (3.7)$$

Proof By Remark 2, $\int_{\mathbb{R}^2} \hat{\mathbf{\Gamma}}(t,z;T,\zeta)\varphi(\zeta)d\zeta \in m\mathcal{F}_{t,T}^Y$. We prove that, for any bounded and $\mathcal{F}_{t,T}^Y$-measurable random variable G, we have

$$E\left[G\varphi(Z_T^{t,z})\right] = E\left[G(\varrho_T^{t,z})^{-1} \int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;T,\zeta)\varphi(\zeta)d\zeta\right], \qquad (3.8)$$

with $\varrho^{t,z}$ as in (3.5). From (3.8) 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 (3.7) will follow from (3.8).

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:

$$E\left[e^{-\int_t^T c_s ds}\varphi(Z_T^{t,z})\right] = E\left[e^{-\int_t^T c_s ds}(\varrho_T^{t,z})^{-1} \int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;T,\zeta)\varphi(\zeta)d\zeta\right]. \qquad (3.9)$$

To this end, we consider the deterministic backward Cauchy problem

$$f\left(s, e^{(T-s)B}\zeta, y\right) = \varphi(\zeta) + \int_s^T (\tilde{\mathcal{A}}_\tau - c(\tau,y))f\left(\tau, e^{(T-\tau)B}\zeta, y\right)d\tau, \qquad (s,\zeta,y) \in [t,T] \times \mathbb{R}^2 \times \mathbb{R}, \quad (3.10)$$

where

$$\tilde{\mathcal{A}}_\tau := \frac{1}{2}\left(|\sigma(\tau,\zeta,y)|^2\partial_{\nu\nu} + 2\theta(\tau,y)\bar{\sigma}(\tau,\zeta,y)\partial_{\nu y} + \theta^2(\tau,y)\partial_{yy}\right) + b(\tau,\zeta,y)\partial_\nu + h(\tau,\zeta,y)\partial_y.$$

In differential form, (3.10) reads as

$$\begin{cases} -d_{\mathbf{B}}f(s,\zeta,y) = \left(\tilde{\mathcal{A}}_s f(s,\zeta,y) - c(s,y)f(s,\zeta,y)\right)ds, \\ f(T,\zeta,y) = \varphi(\zeta). \end{cases}$$

Corollary 1 ensures existence and estimates of a strong solution f to (3.10).

Next, we consider the process

$$M_s^{t,z} := e^{-\int_t^s c_\tau d\tau}(\varrho_s^{t,z})^{-1}\int_{\mathbb{R}^2}\mathbf{\Gamma}(t,z;s,\zeta)f(s,\zeta,Y_s)d\zeta, \qquad s \in [t,T],$$

where $M_t^{t,z}$ is defined by continuity. 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 (3.9) it suffices to check that $M = (M_s^{t,z})_{s\in[t,T]}$ is a martingale: to this end, we prove the representation

$$M_T^{t,z} = M_t^{t,z} + \int_t^T G_s^{t,z}dW_s^1,$$

$$G_s^{t,z} = e^{-\int_t^s c_\tau d\tau}(\varrho_s^{t,z})^{-1}\int_{\mathbb{R}^2}\mathbf{\Gamma}(t,z;s,\zeta)\left(\mathcal{G}_s + \theta_s\partial_y\right)f(s,\zeta,Y_s)d\zeta, \qquad s \in [t,T], \tag{3.11}$$

and conclude by showing that

$$E\left[\int_t^T |G_s^{t,z}|^2 ds\right] < \infty. \tag{3.12}$$

We first compute the stochastic differential $d_{\mathbf{B}}f(s,\zeta,Y_s)$: by Corollary 1 we have

$$d_{\mathbf{B}}f(s,\zeta,Y_s) = \left(-\widetilde{\mathcal{A}}_s + \frac{1}{2}\theta_s^2\partial_{yy} + c_s\right)f(s,\zeta,Y_s)ds + \partial_y f(s,\zeta,Y_s)dY_s$$

$$= \left(-\widetilde{\mathcal{A}}_s + \frac{1}{2}\theta_s^2\partial_{yy} + h_s(Z_s)\partial_y + c_s\right)f(s,\zeta,Y_s)ds + \theta_s\partial_y f(s,\zeta,Y_s)dW_s^1.$$

On the other hand, we have

$$d_{\mathbf{B}}\mathbf{\Gamma}(t,z;s,\zeta) = \mathcal{A}_s^*\mathbf{\Gamma}(t,z;s,\zeta)ds + \mathcal{G}_s^*\mathbf{\Gamma}(t,z;s,\zeta)\frac{dY_s}{\theta_s}$$

$$= \left(\mathcal{A}_s^* + \widetilde{h}_s(Z_s)\mathcal{G}_s^*\right)\mathbf{\Gamma}(t,z;s,\zeta)ds + \mathcal{G}_s^*\mathbf{\Gamma}(t,z;s,\zeta)dW_s^1.$$

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

$$I_1(t,z;s,\zeta) = f(s,\zeta,Y_s)\left(\mathcal{A}_s^* + \widetilde{h}_s(Z_s)\mathcal{G}_s^*\right)\mathbf{\Gamma}(t,z;s,\zeta)$$

$$+ \mathbf{\Gamma}(t,z;s,\zeta)\left(-\widetilde{\mathcal{A}}_s + \frac{1}{2}\theta_s^2\partial_{yy} + h_s(Z_s)\partial_y + c_s\right)f(s,\zeta,Y_s) + \theta_s\mathcal{G}_s^*\mathbf{\Gamma}(t,z;s,\zeta)\partial_y f(s,\zeta,Y_s),$$

$$I_2(t,z;s,\zeta) = f(s,\zeta,Y_s)\mathcal{G}_s^*\mathbf{\Gamma}(t,z;s,\zeta) + \theta_s\mathbf{\Gamma}(t,z;s,\zeta)\partial_y f(s,\zeta,Y_s).$$

This means that for any $s \in (t,T]$ we have

$$f(T,\gamma_{T-s}^{\mathbf{B}}(\zeta),Y_T)\mathbf{\Gamma}(t,z;T,\gamma_{T-s}^{\mathbf{B}}(\zeta)) = f(s,\zeta,Y_s)\mathbf{\Gamma}(t,z;s,\zeta)$$

$$+ \int_s^T I_1(t,z;\tau,\gamma^{\mathbf{B}}_{\tau-s}(\zeta))d\tau + \int_s^T I_2(t,z;\tau,\gamma^{\mathbf{B}}_{\tau-s}(\zeta))dW^1_\tau.$$

Next, we integrate over $\mathbb{R}^2$ the previous identity and apply the standard and stochastic Fubini's theorems (see, for instance, [30], Chapter 1) to get

$$\int_{\mathbb{R}^2} f(T,\gamma^{\mathbf{B}}_{T-s}(\zeta),Y_T)\mathbf{\Gamma}(t,z;T,\gamma^{\mathbf{B}}_{T-s}(\zeta))d\zeta = \int_{\mathbb{R}^2} f(s,\zeta,Y_s)\mathbf{\Gamma}(t,z;s,\zeta)d\zeta$$
$$+ \int_s^T \int_{\mathbb{R}^2} I_1(t,z;\tau,\gamma^{\mathbf{B}}_{\tau-s}(\zeta))d\zeta d\tau$$
$$+ \int_s^T \int_{\mathbb{R}^2} I_2(t,z;\tau,\gamma^{\mathbf{B}}_{\tau-s}(\zeta))d\zeta dW^1_\tau. \qquad (3.13)$$

By the upper bounds (2.8)-(2.10) of the fundamental solution, the estimates of the solution f and its derivatives in Corollary 1, the boundedness of the coefficients and the non-degeneracy Assumption 5, we have

$$\int_s^T \int_{\mathbb{R}^2} |I_1(t,z;\tau,\zeta)|d\zeta d\tau \leqslant \int_s^T \frac{C}{(T-\tau)^{\frac{1}{2}}(s-t)} \int_{\mathbb{R}^2} \Gamma_\lambda(\tau-t,z;\tau,g^{\mathrm{IW},-1}_{t,\tau}(\zeta)-\gamma^{t,z}_\tau)d\zeta d\tau \leqslant C' \frac{(T-s)^{\frac{1}{2}}}{s-t},$$

and, analogously

$$\int_s^T \left(\int_{\mathbb{R}^2} |I_2(t,z;\tau,\zeta)|d\zeta \right)^2 d\tau \leqslant \int_s^T \left(\frac{C}{(s-t)^{\frac{1}{2}}} \int_{\mathbb{R}^2} \Gamma_\lambda(\tau-t,z;\tau,g^{\mathrm{IW},-1}_{t,\tau}(\zeta)-\gamma^{t,z}_\tau)d\zeta \right)^2 d\tau \leqslant \frac{C'}{s-t},$$

for some positive random variables C, C'. This justifies the use of Fubini's theorems.

Now, from equality (3.13) we perform the change of variable $\zeta' = \gamma^{\mathbf{B}}_{\tau-s}(\zeta)$, which has Jacobian matrix $\mathrm{Id}_{2\times 2} + (\tau-s)B$; since its determinant is equal to one for any τ, we get

$$\int_{\mathbb{R}^2} f(T,\zeta,Y_T)\mathbf{\Gamma}(t,z;T,\zeta)d\zeta = \int_{\mathbb{R}^2} f(s,\zeta,Y_s)\mathbf{\Gamma}(t,z;s,\zeta)d\zeta$$
$$+ \int_s^T \int_{\mathbb{R}^2} I_1(t,z;\tau,\zeta)\zeta d\tau + \int_s^T \int_{\mathbb{R}^2} I_2(t,z;\tau,\zeta)d\zeta dW^1_\tau.$$

Integrating by parts and using the identity

$$\int_{\mathbb{R}^2} \left(f(s,\zeta,Y_s)\mathcal{A}^*_s\mathbf{\Gamma}(t,z;s,\zeta) + \mathbf{\Gamma}(t,z;s,\zeta)\frac{1}{2}\theta^2_s\partial_{yy}f(s,\zeta,Y_s) + \theta_s\mathcal{G}^*_s\mathbf{\Gamma}(t,z;s,\zeta)\partial_y f(s,\zeta,Y_s) \right) d\zeta$$
$$= \int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;s,\zeta) \left(\mathcal{A}_s + \frac{1}{2}\theta^2_s\partial_{yy} + \theta_s\sigma^1_s\partial_{y\nu} + h_s(\zeta,Y_s)\partial_y \right) f(s,\zeta,Y_s)d\zeta$$
$$= \int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;s,\zeta)\tilde{\mathcal{A}}_s f(s,\zeta,Y_s)d\zeta,$$

we get

$$\int_{\mathbb{R}^2} f(T,\zeta,Y_T)\mathbf{\Gamma}(t,z;T,\zeta)d\zeta = \int_{\mathbb{R}^2} f(s,\zeta,Y_s)\mathbf{\Gamma}(t,z;s,\zeta)d\zeta$$
$$+ \int_s^T \int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;\tau,\zeta) \left(\tilde{h}_\tau(Z_\tau)\mathcal{G}_\tau + h_\tau(Z_\tau)\partial_y + c_\tau \right) f(\tau,\zeta,Y_\tau)d\zeta d\tau$$
$$+ \int_s^T \int_{\mathbb{R}^2} \mathbf{\Gamma}(t,z;\tau,\zeta) \left(\mathcal{G}_\tau + \theta_\tau\partial_y \right) f(\tau,\zeta,Y_\tau)d\zeta dW^1_\tau.$$

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

$$M_T^{t,z} = e^{-\int_t^T c_\tau d\tau}(\varrho_T^{t,z})^{-1}\int_{\mathbb{R}^2} f(T,\zeta,Y_T)\mathbf{\Gamma}(t,z;T,\zeta)d\zeta$$

$$= M_s^{t,z} + \int_s^T e^{-\int_t^\tau c_\varrho d\varrho}(\varrho_\tau^{t,z})^{-1}\int_{\mathbb{R}^2}\mathbf{\Gamma}(t,z;\tau,\zeta)\left(\mathcal{G}_\tau + \theta_\tau\partial_y\right)f(\tau,\zeta,Y_\tau)d\zeta dW_\tau^1$$

$$= M_s^{t,z} + \int_s^T G_\tau^{t,z}dW_\tau^1.$$

with $G_\tau^{t,z}$ as in (3.11). Now, again by the estimates of the fundamental solution (cf. Theorem 4), the estimates of the solution f and its derivatives (cf. Corollary 1), the boundedness of the coefficients and the non-degeneracy condition (5), we deduce the estimate

$$|G_\tau^{t,z}| \leqslant C(\varrho_\tau^{t,z})^{-1}\int_{\mathbb{R}^2}\Gamma_\lambda(\tau - t, g_{\tau,t}^{\mathrm{IW},-1}(\zeta) - \gamma_\tau^{t,z})d\zeta \leqslant C'$$

for some positive constants C, C'. This implies (3.12) and concludes the proof.

3.2 Backward filtering SPDE

As in the previous section, in order to apply the general results of Section 1 to the filtering SPDE for system (3.3), we impose the following conditions:

Assumption 9 (Regularity) *The coefficients of (3.3) are such that $\bar{\sigma} \in bC_{0,T}^{3+\alpha}(\mathbb{R}^3)$, $\hat{\sigma} \in bC_{0,T}^\alpha(\mathbb{R}^3)$, $\theta \in bC_{0,T}^{3+\alpha}(\mathbb{R})$, $b \in bC_{0,T}^0(\mathbb{R}^3)$, $h \in bC_{0,T}^2(\mathbb{R}^3)$.*

Assumption 10 (Flattening at infinity) *There exist two positive constants ε, M such that*

$$\sup_{t\in[0,T]}\left(\langle\bar{\sigma}(t,\cdot,\cdot)\rangle_{\varepsilon,\beta} + \langle\bar{\sigma}(t,\cdot,\cdot)\rangle_{1/2+\varepsilon,\beta'} + \langle\theta(t,\cdot)\rangle_{\varepsilon,\beta} + \langle\theta(t,\cdot)\rangle_{1/2+\varepsilon,\beta'} + \langle h(t,\cdot,\cdot)\rangle_{1/2,\beta}\right) \leqslant M$$

for $|\beta| = 1$ and $|\beta'| = 2,3$.

The backward filtering SPDE for system (3.3) reads

$$-d_{\mathbf{B}}u_t(z,y) = \widetilde{\mathcal{A}}_t u_t(z,y)dt + \widetilde{\mathcal{G}}_t u_t(z,y) \star \frac{dY_t}{\theta(t,y)}, \qquad \mathbf{B} := \partial_t + v\partial_x, \tag{3.14}$$

where $z = (x,v)$ and

$$\widetilde{\mathcal{A}}_t := \frac{1}{2}\left(|\sigma(t,z,y)|^2\partial_{vv} + 2\theta(t,y)\bar{\sigma}(t,z,y)\partial_{vy} + \theta^2(t,y)\partial_{yy}\right) + b(t,z,y)\partial_v + h(t,z,y)\partial_y, \tag{3.15}$$

$$\widetilde{\mathcal{G}}_t := \bar{\sigma}(t,z,y)\partial_v + \theta(t,y)\partial_y + \widetilde{h}(t,z,y), \qquad \widetilde{h}(t,z,y) := \frac{h(t,z,y)}{\theta(t,y)}.$$

Before presenting the main result of this section, we comment on the existence of solutions to (3.14). 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

$$\begin{cases} dZ_t = BZ_tdt + \mathbf{e}_2(b(t, Z_t, Y_t)dt + \bar{\sigma}(t, Z_t, Y_t)dW_t^1 + \hat{\sigma}(t, Z_t, Y_t)dW_t^2), \\ dY_t = h(t, Z_t, Y_t)dt + \theta(t, Y_t)dW_t^1, \\ d\varrho_t = \widetilde{h}(t, Z_t, Y_t)^2\varrho_t dt + \widetilde{h}(t, Z_t, Y_t)\varrho_t dW_t^1. \end{cases} \tag{3.16}$$

By Girsanov's theorem, the process

$$\widetilde{W}_s^{t,z,y} := \int_t^s \theta^{-1}(\tau, Y_\tau^{t,z,y})dY_\tau^{t,z,y}$$

$$= W_s^1 - W_t^1 + \int_t^s \widetilde{h}(\tau, Z_\tau^{t,z,y}, Y_\tau^{t,z,y})d\tau, \qquad s \in [t, T],$$

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 (3.14) can be written in the equivalent form

$$-d_{\mathbf{B}}u_s(z, y) = \widetilde{\mathcal{A}}_s u_s(z, y)ds + \widetilde{\mathcal{G}}_s u_s(z, y) \star d\widetilde{W}_s^t \tag{3.17}$$

or, more explicitly,

$$u_t\left(\gamma_{T-t}^{\mathbf{B}}(z, y)\right) = u_T(z, y) + \int_t^T \widetilde{\mathcal{A}}_s u_s(\gamma_{T-s}^{\mathbf{B}}(z, y))ds + \int_t^T \widetilde{\mathcal{G}}_s u_s(\gamma_{T-s}^{\mathbf{B}}(z, y)) \star d\widetilde{W}_s^t, \qquad t \in [0, T], \tag{3.18}$$

where $\gamma_s^{\mathbf{B}}(z, y) = \gamma_s^{\mathbf{B}}(x, v, y) = (x + sv, v, y)$. In (3.17) and (3.18), 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 4 guarantees that a fundamental solution $\overleftarrow{\boldsymbol{\Gamma}} = \overleftarrow{\boldsymbol{\Gamma}}(t, z, y; s, \zeta, \eta)$ for (3.17) exists and satisfies estimates (2.11), (2.12) and (2.13). Moreover, $t \mapsto \overleftarrow{\boldsymbol{\Gamma}}(t, z, y; T, \zeta, \eta)$ is adapted to $(\mathcal{F}_{t,T}^Y)_{t\in[0,T]}$. The main result of this section is the following

Theorem 11 *Let $(Z_T^{t,z,y}, Y_T^{t,z,y})$ denote the solution of system (3.3) starting from (z, y) at time $t \in [0, T)$ and $\varphi \in bC(\mathbb{R}^3)$. Under Assumptions 5, 9 and 10, we have*

$$E\left[\varphi(Z_T^{t,z,y}, Y_T^{t,z,y}) \mid \mathcal{F}_{t,T}^Y\right] = \frac{u_t^{(\varphi)}(z, y)}{u_t^{(1)}(z, y)}, \qquad (t, z, y) \in [0, T] \times \mathbb{R}^2 \times \mathbb{R}, \tag{3.19}$$

where $u_t^{(\varphi)}$ denotes the solution to (3.14) with final datum $u_T^{(\varphi)} = \varphi$.

Definition 6 (Backward filtering density) The normalized process

$$\bar{\boldsymbol{\Gamma}}(t, z, y; T, \zeta, \eta) = \frac{\overleftarrow{\boldsymbol{\Gamma}}(t, z, y; T, \zeta, \eta)}{\int_{\mathbb{R}^3} \overleftarrow{\boldsymbol{\Gamma}}(t, z, y; T, \zeta_1, \eta_1)d\zeta_1 d\eta_1},$$

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 (3.3). By Theorem 11, we have

$$E\left[\varphi(Z_T^{t,z,y}, Y_T^{t,z,y}) \mid \mathcal{F}_{t,T}^Y\right] = \int_{\mathbb{R}^3} \bar{\boldsymbol{\Gamma}}(t, z, y; T, \zeta, \eta)\varphi(\zeta, \eta)d\zeta d\eta, \qquad (t, z, y) \in [0, T] \times \mathbb{R}^2 \times \mathbb{R}, \tag{3.20}$$

for any $\varphi \in bC(\mathbb{R}^3)$.

Remark 3 Notice that formulas (3.19) and (3.20) represent the conditional expectation in terms of solutions to the Cauchy problem for the backward filtering SPDE. This is not the case for formula (3.7) in the forward case.

In the rest of the section we sketch the proof of Theorem 11. 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 (3.16) reads

$$\begin{cases} dZ_s^{t,z,y} = \widetilde{B}(s, Z_s^{t,z,y}, Y_s^{t,z,y})ds + \mathbf{e}_2\left(\hat{\sigma}(s, Z_s^{t,z,y}, Y_s^{t,z,y})dW_t^2 + \bar{\sigma}(s, Z_s^{t,z,y}, Y_s^{t,z,y})d\widetilde{W}_s^t\right), \\ dY_s^{t,z,y} = \theta(s, Y_s^{t,z,y})d\widetilde{W}_s^t, \\ d\varrho_s^{t,z,y,\eta} = \widetilde{h}(s, Z_s^{t,z,y}, Y_s^{t,z,y})\varrho_s^{t,z,y,\eta}d\widetilde{W}_s^t, \end{cases} \tag{3.21}$$

where $\widetilde{B}(s,z,y) = Bz + \mathbf{e}_2(b(s,z,y) - \widetilde{h}(s,z,y)\bar{\sigma}(s,z,y))$. Recalling the notation $z = (x,v) \in \mathbb{R}^2$ and omitting the arguments of the coefficients for brevity, the correspondent characteristic operator is

$$\mathcal{L} = \frac{1}{2}\left(|\sigma|^2 \partial_{vv} + \theta^2 \partial_{yy} + \eta^2\widetilde{h}^2 \partial_{\eta\eta} + 2\bar{\sigma}\theta\partial_{vy} + 2\eta\bar{\sigma}\widetilde{h}\partial_{v\eta} + 2\eta\theta\widetilde{h}\partial_{y\eta}\right) + \langle\widetilde{B}, \nabla_z\rangle.$$

We write the backward diffusion SPDE for system (3.21). Assuming that φ is smooth and letting $V_s(z,y) := \varphi(Z_T^{s,z,y}, Y_T^{s,z,y})$, by Corollary 2 we have

$$-d(V_s(z,y)\varrho_T^{s,z,y,\eta}) = \mathcal{L}(V_s(z,y)\varrho_T^{s,z,y,\eta})ds + \partial_v(V_s(z,y)\varrho_T^{s,z,y,\eta})\left(\hat{\sigma}(s,z,y)\star dW_s^2 + \bar{\sigma}(s,z,y)\star d\widetilde{W}_s^t\right)$$
$$+ \partial_y(V_s(z,y)\varrho_T^{s,z,y,\eta})\theta(s,z,y)\star d\widetilde{W}_s^t + \partial_\eta(V_s(z,y)\varrho_T^{s,z,y,\eta})\eta\widetilde{h}(s,z,y)\star d\widetilde{W}_s^t$$

(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 $\eta\partial_\eta\varrho_T^{t,z,y,\eta} = \varrho_T^{t,z,y,\eta}$)

$$= \frac{1}{2}\left(|\sigma(t,z,y)|^2 \partial_{vv} + \theta^2(t,z,y)\partial_{yy} + 2\bar{\sigma}(t,z,y)\theta(t,y)\partial_{vy}\right)(V_s(z,y)\varrho_T^{s,z,y,\eta})ds$$
$$+ \left(\widetilde{h}(s,z,y)(\bar{\sigma}(t,z,y)\partial_v + \theta(t,y)\partial_y) + \langle\widetilde{B}(t,z,y), \nabla_z\rangle\right)(V_s(z,y)\varrho_T^{s,z,y,\eta})ds$$
$$+ \hat{\sigma}(s,z,y)\partial_v(V_s(z,y)\varrho_T^{s,z,y,\eta})\star dW_s^2$$
$$+ \left(\bar{\sigma}(s,z,y)\partial_v + \theta(s,z,y)\partial_y + \widetilde{h}(s,z,y)\right)(V_s(z,y)\varrho_T^{s,z,y,\eta})\star d\widetilde{W}_s^t$$

(noting that $\widetilde{h}(s,z,y)(\bar{\sigma}(t,z,y)\partial_v + \theta(t,y)\partial_y) + \langle\widetilde{B}(t,z,y), \nabla_z\rangle = v\partial_x + b(t,z,y)\partial_v + h(t,z,y)\partial_y$)

$$= \widetilde{\mathcal{L}}(V_s(z,y)\varrho_T^{s,z,y,\eta})ds + \hat{\sigma}(s,z,y)\partial_v(V_s(z,y)\varrho_T^{s,z,y,\eta})\star dW_s^2$$
$$+ \left(\bar{\sigma}(s,z,y)\partial_v + \theta(s,y)\partial_y + \widetilde{h}(s,z,y)\right)(V_s(z,y)\varrho_T^{s,z,y,\eta})\star d\widetilde{W}_s^t.$$

where $\widetilde{\mathcal{L}} = \widetilde{\mathcal{A}}_t + v\partial_x$, with $\widetilde{\mathcal{A}}_t$ as in (3.15), is the infinitesimal generator of (Z_t, Y_t). Therefore we have

$$\varphi(Z_T^{t,z,y}, Y_T^{t,z,y})\varrho_T^{t,z,y,1} - \varphi(z,y) = V_t(z,y)\varrho_T^{t,z,y,1} - V_T(z,y)\varrho_T^{T,z,y,1}$$
$$= \int_t^T \widetilde{\mathcal{L}}(V_s(z,y)\varrho_T^{s,z,y,\eta})ds + \int_t^T \hat{\sigma}(s,z,y)\partial_v(V_s(z,y)\varrho_T^{s,z,y,\eta})\star dW_s^2$$
$$+ \int_t^T \widetilde{\mathcal{G}}_s(V_s(z,y)\varrho_T^{s,z,y,\eta})\star d\widetilde{W}_s^t. \tag{3.22}$$

Now we take the conditional expectation in (3.22) and exploit the fact that W^2 is independent of $\mathcal{F}_{t,T}^Y$ under $Q^{t,z,y}$ (this follows from the crucial assumption that θ 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

$$u_t^{(\varphi)}(z,y) = \varphi(z,y) + \int_t^T \widetilde{\mathcal{L}}_s u_s^{(\varphi)}(z,y)ds + \int_t^T \widetilde{\mathcal{G}}_s u_s^{(\varphi)}(z,y) \star \frac{dY_s^{t,z,y}}{\theta(s,y)}$$

which is equivalent to (3.14). 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. [30], Lemma 6.1) according to which we have

$$E\left[\varphi(Z_T^{t,z,y}, Y_T^{t,z,y}) \mid \mathcal{F}_{t,T}^Y\right] = \frac{E^{Q^{t,z,y}}\left[\varphi(Z_T^{t,z,y}, Y_T^{t,z,y})\varrho_T^{t,z,y,1} \mid \mathcal{F}_{t,T}^Y\right]}{E^{Q^{t,z,y}}\left[\varrho_T^{t,z,y,1} \mid \mathcal{F}_{t,T}^Y\right]}.$$

4 Proof of Theorem 4

As in [27] the main ingredient in the proof of Theorem 4 is the Itô-Wentzell formula that transforms the original SPDE into a PDE with random coefficients. In this section we explain how to tweak the change of variables introduced in [27] to deal with the additional term h and we also consider the backward equation. We set $d = 1$ for simplicity.

4.1 Itô-Wentzell change of variables

We first recall some global estimates, proved in [27], Section 4, for $\gamma^{\mathrm{IW}}, \bar{\gamma}^{\mathrm{IW}}$ in (2.3)-(2.4) and their derivatives under Assumptions 1, 2 and 3.

Lemma 1 *For any $\bar{\alpha} \in [0, \alpha)$, we have $\gamma_{t,\cdot}^{\mathrm{IW}} \in \mathbf{C}_{t,T}^{3+\bar{\alpha}}$. Moreover there exists $\varepsilon \in \left(0, \frac{1}{2}\right)$ and a random, finite constant $\mathbf{c}$ such that, with probability one,*

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

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

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

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

for any $(\xi, \nu) \in \mathbb{R}^2$, $0 \leqslant t \leqslant s \leqslant T$ and $|\beta| = 2$. Analogous estimates hold for $\bar{\gamma}_{\cdot,s}^{\mathrm{IW}} \in \mathbf{C}_{0,s}^{3,\bar{\alpha}}$.

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

$$\hat{f}_{t,s}(\xi, \nu) := f_s(\xi, \gamma_{t,s}^{\mathrm{IW}}(\xi, \nu)).$$

Let $u_s(\xi, \nu)$ a solution to (2.1) on $[t, T]$. Then we define

$$v_{t,s}(\zeta) := \hat{\varrho}_{t,s}(\zeta)\hat{u}_{t,s}(\zeta), \qquad \hat{\varrho}_{t,s}(\zeta) := \exp\left(-\int_t^s \hat{h}_\tau(\zeta)dW_\tau - \frac{1}{2}\int_t^s \hat{h}_\tau^2(\zeta)d\tau\right).$$

We have the following

Proposition 1 u_s *is a solution to the SPDE (2.1) on $[t, T]$ if and only if $v_{t,s}$ is a solution on $[t, T]$ to the PDE with random coefficients*

$$d_{\hat{\mathbf{B}}} v_{t,s}(\zeta) = \left(a_{t,s}^*(\zeta)\partial_{\nu\nu}v_{t,s} + b_{t,s}^*(\zeta)\partial_\nu v_{t,s}(\zeta) + c_{t,s}^*(\zeta)v_{t,s}(\zeta)\right) ds, \qquad \hat{\mathbf{B}} = \partial_s + \mathbf{Y}_{t,s}, \tag{4.1}$$

where

$$\mathbf{Y}_{t,s} = \mathbf{Y}_{t,s}(\xi, \nu) := (\gamma_{t,s}^{\mathrm{IW}})_1(\xi, \nu)\partial_\xi - (\gamma_{t,s}^{\mathrm{IW}}(\xi, \nu))_1(\partial_\nu\gamma_{t,s}^{\mathrm{IW}})^{-1}(\xi, \nu)\partial_\xi\gamma_{t,s}^{\mathrm{IW}}(\xi, \nu)\partial_\nu, \tag{4.2}$$

is the first order operator identified with the vector field in (2.7) (with $d = 1$) and the coefficients $a_{t,\cdot}^*$, $b_{t,\cdot}^*$, $c_{t,\cdot}^*$ are defined in (4.7) below. Moreover, $a_{t,\cdot}^* \in \mathbf{bC}_{t,T}^\alpha$, $b_{t,\cdot}^*, c_{t,\cdot}^* \in \mathbf{bC}_{t,T}^0$, $\mathbf{Y}_{t,\cdot} \in \mathbf{C}_{t,T}^{0,1}$, $\partial_\nu(\mathbf{Y}_{t,\cdot})_1 \in \mathbf{bC}_{t,T}^{\bar{\alpha}}$ for any $\bar{\alpha} \in [0, \alpha)$, and there exist two random, finite and positive constants $\mathbf{m}_1$, $\mathbf{m}_2$ such that, for $s \in [t, T]$, $\zeta \in \mathbb{R}^2$, we have

$$\mathbf{m}_1^{-1} \leqslant a_{t,s}^*(\zeta) \leqslant \mathbf{m}_1, \qquad \mathbf{m}_2^{-1} \leqslant \partial_\nu(\mathbf{Y}_{t,s}(\zeta))_1 \leqslant \mathbf{m}_2, \tag{4.3}$$

with probability one.

Proof By a standard regularization argument, we may assume $u \in \mathbf{C}_{t,T}^2$ so that equation (2.1) 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 standard Itô-Wentzell formula (see for instance [30], Theorem 1.17), and the chain rule we have

$$d\hat{u}_{t,s} = \left(\widehat{\mathcal{A}_{s,\zeta}u}_{t,s} - \gamma_{t,s}^{\mathrm{IW}}\widehat{\partial_1 u}_{t,s} + \frac{1}{2}\hat{\sigma}_{t,s}^2\widehat{\partial_2^2 u}_{t,s} - \widehat{\partial_2\mathcal{G}_{s,\zeta}}\hat{\sigma}_{t,s}\right) ds + \hat{h}_{t,s}\hat{u}_{t,s}dW_s \tag{4.4}$$

$$= \left(\mathfrak{L}_{t,s} - \mathbf{Y}_{t,s}\right)\hat{u}_{t,s}dt + \hat{h}_{t,s}\hat{u}_{t,s}dW_s,$$

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

$$\begin{aligned}
\bar{a}_{t,s} &= \frac{1}{2}(\partial_\nu\gamma_{t,s}^{\mathrm{IW}})^{-2}(\hat{a}_{t,s} - \hat{\sigma}_{t,s}^2), \\
\bar{b}_{t,s} &= (\partial_\nu\gamma_{t,s}^{\mathrm{IW}})^{-1}\left(\hat{b}_{t,s} - \hat{\sigma}_{t,s}\hat{h}_{t,s} - (\partial_\nu\gamma_{t,s}^{\mathrm{IW}})^{-1}\hat{\sigma}_{t,s}\partial_\nu\hat{\sigma}_{t,s} - \bar{a}_{t,s}\partial_{\nu\nu}\gamma_{t,s}^{\mathrm{IW}}\right), \\
\bar{c}_{t,s} &= \hat{c}_{t,s} - (\partial_\nu\gamma_{t,s}^{\mathrm{IW}})^{-1}\hat{\sigma}_{t,s}\partial_\nu\hat{h}_{t,s}.
\end{aligned} \tag{4.5}$$

Notice that the change of variable is well defined by the estimates of Lemma 1. Next we compute the product $v_{t,s} = \hat{\varrho}_{t,s}\hat{u}_{t,s}$: by the Itô formula $d\hat{\varrho}_{t,s} = -\hat{\varrho}_{t,s}\hat{h}_{t,s}dW_s$ and therefore

$$dv_{t,s}(\zeta) = \hat{\varrho}_{t,s}(\zeta)d\hat{u}_{t,s}(\zeta) + \hat{u}_{t,s}(\zeta)d\hat{\varrho}_{t,s} + d\langle \hat{u}_{t,\cdot}(\zeta)\hat{\varrho}_{t,\cdot}(\zeta)\rangle_s \tag{4.6}$$

$$= \big(\hat{\varrho}_{t,s}(\zeta) \mathfrak{L}_{t,s}(\hat{\varrho}_{t,s}^{-1} v_{t,s})(\zeta) - \hat{\varrho}_{t,s}(\zeta)(\mathbf{Y}_{t,s}(\hat{\varrho}_{t,s}^{-1} v_{t,s}))(\zeta) - \bar{h}_{t,s}^2(\zeta) v_{t,s}(\zeta) \big) \, ds.$$

Now we notice that

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

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

$$dv_{t,s}(\zeta) = \big(a_{t,s}^*(\zeta) \partial_{vv} v_{t,s}(\zeta) - (\mathbf{Y}_{t,s} v_{t,s})(\zeta) + b_{t,s}^*(\zeta) \partial_v v_{t,s}(\zeta) + c_{t,s}^*(\zeta) v_{t,s}(\zeta) \big) \, ds,$$

where

$$
\begin{aligned}
a_{t,s}^* &= \bar{a}_{t,s} = \frac{1}{2}(\partial_\nu \gamma_{t,s}^{\mathrm{IW}})^{-2}(\hat{a}_{t,s} - \hat{\sigma}_{t,s}^2), \\
b_{t,s}^* &= \bar{b}_{t,s} + 2\bar{a}_{t,s} \partial_\nu \ln \hat{\varrho}_{t,s}^{-1}, \\
c_{t,s}^* &= \bar{c}_{t,s} + \bar{b}_{t,s} \partial_\nu \ln \hat{\varrho}_{t,s}^{-1} + \bar{a}_{t,s} \left(\partial_\nu \ln \hat{\varrho}_{t,s}^{-1} + \partial_\nu^2 \ln \hat{\varrho}_{t,s}^{-1} \right) + \mathbf{Y}_{t,s} \ln \hat{\varrho}_{t,s}^{-1} - \hat{h}_{t,s}^2.
\end{aligned}
\tag{4.7}
$$

The regularity of the coefficients and (4.3) follow directly from (4.7), Assumption 2 and Lemma 1.

Remark 4 When the coefficients are smooth, condition (4.3) ensures the validity of the weak Hörmander condition: indeed the vector fields $\sqrt{a_{t,\cdot}^*} \partial_\nu$ and $\mathbf{Y}_{t,\cdot}$, together with their commutator, span $\mathbb{R}^2$ at any point. In this case a smooth fundamental solution to (4.1) exists by Hörmander's theorem.

In the backward case the computations are completely analogous since it only suffices to reverse the time in equations (4.4) and (4.6). Precisely, we introduce the "check" transform

$$\check{f}_{t,s}(x,v) := f_t(\xi, \overleftarrow{\gamma}_{t,s}^{\mathrm{IW}}(x,v)), \qquad t \in [0,s],$$

with $\overleftarrow{\gamma}_{t,s}^{\mathrm{IW}}$ as in (2.4). For a solution $u_t = u_t(z)$ to (2.2) on $[0,s]$, we define

$$v_{t,s}(z) := \check{\varrho}_{t,s}(z) \check{u}_{t,s}(z), \qquad \check{\varrho}_{t,s}(z) := \exp\left(-\int_t^s \check{h}_\tau(z) \star dW_\tau - \frac{1}{2}\int_t^s \check{h}_\tau^2(z) d\tau \right),$$

which solves, on $[0,s]$, the deterministic equation with random coefficients

$$-d_{\overleftarrow{\mathbf{B}}} v_{t,s}(z) = \left(\bar{a}_{t,s}^*(z) \partial_{vv} v_{t,s} + \overleftarrow{b}_{t,s}^*(z) \partial_v v_{t,s}(z) + \overleftarrow{c}_{t,s}^*(z) v_{t,s}(z) \right) dt, \qquad \overleftarrow{\mathbf{B}} = \partial_t + \overleftarrow{\mathbf{Y}}_{t,s}, \tag{4.8}$$

where $\overleftarrow{\mathbf{Y}}_{t,s}$ and the coefficients are defined similarly to (4.2) and (4.5), exchanging the *hat-* and *check-* transforms in the definitions. As for the forward case, by Assumption 2 and Lemma 1, $a_{t,\cdot}^* \in \mathbf{b}\overleftarrow{\mathbf{C}}_{t,T}^\alpha$, $b_{t,\cdot}^*, c_{t,\cdot}^* \in \mathbf{b}\overleftarrow{\mathbf{C}}_{t,T}^0$, $\overleftarrow{\mathbf{Y}}_{t,\cdot} \in \overleftarrow{\mathbf{C}}_{t,T}^{0,1}$, $\partial_v(\overleftarrow{\mathbf{Y}}_{t,\cdot})_1 \in \mathbf{b}\overleftarrow{\mathbf{C}}_{t,T}^{\bar{\alpha}}$, for any $\bar{\alpha} \in [0,\alpha)$, and there exist two random, finite and positive constant $\mathbf{m}_1$, $\mathbf{m}_2$ such that, for $t \in [0,s]$ and $z \in \mathbb{R}^2$, we have

$$\mathbf{m}_1^{-1} \leqslant \bar{a}_{t,s}^*(z) \leqslant \mathbf{m}_1, \qquad \mathbf{m}_2^{-1} \leqslant \partial_v(\overleftarrow{\mathbf{Y}}_{t,s}(z))_1 \leqslant \mathbf{m}_2,$$

with probability one, which ensures the weak Hörmander condition to hold.

4.2 The parametrix expansion

Equations of the form (4.1) have been studied in [27] by means of a time-dependent parametrix expansion which takes into account the unbounded drift $\mathbf{Y}$. The only minor difference here is the presence of a term of order zero c^* which, as we shall see, does not modify the analysis substantially.

In this section we briefly resume the parametrix construction and show how it works in the backward framework. For the sake of readability, here we reset the notations and rewrite equation (4.1) as

$$\mathcal{A}_s v_s(\zeta) - \mathbf{Y}_s v_s(\zeta) - \partial_s v_s(\zeta) = 0, \qquad s \in (t, T], \ \zeta \in \mathbb{R}^2, \tag{4.9}$$

where $\mathcal{A}_s$ is a second order operator of the form

$$\mathcal{A}_s = a_s \partial_{\nu\nu} + b_s \partial_\nu + c_s$$

and $\mathbf{Y}_s = (\mathbf{Y}_s)_1 \partial_\xi + (\mathbf{Y}_s)_2 \partial_\nu$ is the vector field in (4.2). For fixed $(t_0, z_0) \in [t, T] \times \mathbb{R}^2$, we linearize $\mathbf{Y}_s$ by setting

$$\mathbf{Y}_s^{t_0, z_0}(\zeta) = \mathbf{Y}_s(\gamma_s^{t_0, z_0}) + (D\mathbf{Y}_s)(\gamma_s^{t_0, z_0})\left(\zeta - \gamma_s^{t_0, z_0}\right) \tag{4.10}$$

where

$$\gamma_s^{t_0, z_0} = z_0 + \int_{t_0}^s \mathbf{Y}_\tau(\gamma_\tau^{t_0, z_0}) d\tau, \qquad s \in [t_0, T],$$

and $D\mathbf{Y}_s$ is the reduced Jacobian defined as

$$D\mathbf{Y}_s := \begin{pmatrix} 0 & \partial_v(\mathbf{Y}_s)_1 \\ 0 & 0 \end{pmatrix}.$$

Then we consider the linearized version of (4.9), that is

$$\mathcal{A}_s^{t_0, z_0} v_s(\zeta) - \mathbf{Y}_s^{t_0, z_0} v_s(\zeta) - \partial_s v_s(\zeta) = 0, \qquad s \in (t, T], \ \zeta \in \mathbb{R}^2,$$

where

$$\mathcal{A}_s^{t_0, z_0} := a_s(\gamma_s^{t_0, z_0}) \partial_{\nu\nu}. \tag{4.11}$$

It turns out (cf. [27], Section 5) that, for any choice of the parameters (t_0, z_0), equation (4.11) has a fundamental solution $\mathbf{\Gamma}^{t_0, z_0} = \mathbf{\Gamma}^{t_0, z_0}(t, z; s, \zeta)$. Moreover, $\mathbf{\Gamma}^{t_0, z_0}$ has an explicit Gaussian expression and satisfies the following estimates for some $\lambda, \mu > 0$ that depend only on the general constants of Assumptions 1, 2 and 3:

$$\Gamma_{\lambda^{-1}}\left(s - t, \zeta - \gamma_{t,s}^{t_0, z_0}(z)\right) \leqslant \mu \mathbf{\Gamma}^{t_0, z_0}(t, z; s, \zeta) \leqslant \Gamma_\lambda\left(s - t, \zeta - \gamma_{t,s}^{t_0, z_0}(z)\right),$$

$$|\partial_\nu \mathbf{\Gamma}^{t_0, z_0}(t, z; s, \xi, \nu)| \leqslant \frac{\mu}{\sqrt{s - t}} \Gamma_\lambda\left(s - t, (\xi, \nu) - \gamma_{t,s}^{t_0, z_0}(z)\right), \tag{4.12}$$

$$|\partial_{\nu\nu} \mathbf{\Gamma}^{t_0, z_0}(t, z; s, \xi, \nu)| \leqslant \frac{\mu}{s - t} \Gamma_\lambda\left(s - t, (\xi, \nu) - \gamma_{t,s}^{t_0, z_0}(z)\right),$$

for $0 \leqslant t < s \leqslant T$ and $z, \zeta \in \mathbb{R}^2$, with Γ_λ as in (2.6) and $s \mapsto \gamma_{t,s}^{t_0, z_0}(z)$ defined by

$$\gamma_{t,s}^{t_0, z_0}(z) = z + \int_t^s \mathbf{Y}_\tau^{t_0, z_0}(\gamma_{\tau,t}^{t_0, z_0}(z)) d\tau, \qquad s \in [t, T].$$

We introduce the so-called *forward parametrix*

$$Z(t,z;s,\zeta) := \mathbf{\Gamma}^{t,z}(t,z;s,\zeta), \qquad 0 \leqslant t < s \leqslant T, \ z,\zeta \in \mathbb{R}^2,$$

that will be used as a first approximation of a fundamental solution $\mathbf{\Gamma}$ of (4.9). Owing to the fact that $\gamma_{t,s}^{t,z}(z) = \gamma_s^{t,z}$, Z satisfies estimates (4.12).

Now we set

$$H(t,z;s,\zeta) := \left(\mathcal{A}_s - \mathbf{Y}_s - (\mathcal{A}_s^{t,z} - \mathbf{Y}_s^{t,z})\right) Z(t,z;s,\zeta)$$

and notice that

$$|H(t,z;s,\zeta)| \leqslant |a_s(\zeta) - a_s(\gamma_s^{t,z})||\partial_{\nu\nu} Z(t,z;s,\zeta)|$$
$$+ |(\mathbf{Y}_s - \mathbf{Y}_s^{t,z})Z(t,z;s,\zeta)| + |b_s(\zeta)||\partial_\nu Z(t,z;s,\zeta)| + |c_s(\zeta)||Z(t,z;s,\zeta)|$$

(since $\partial_\nu(\mathbf{Y}_s)_1$ is $\bar{\alpha}$-Hölder continuous by Proposition 1)

$$\leqslant \mu \left(\frac{|\zeta - \gamma_s^{t,z}|^\alpha}{s-t} + \frac{|\zeta - \gamma_s^{t,z}|^{1+\bar{\alpha}}}{(s-t)^{3/2}} + \frac{|\zeta - \gamma_s^{t,z}|^\alpha}{(s-t)^{1/2}} + 1 \right) \Gamma_\lambda(s-t, \zeta - \gamma_s^{t,z})$$

(for some $\bar{\lambda} > \lambda$ and $\bar{\mu} > \mu$)

$$\leqslant \frac{\bar{\mu}}{(s-t)^{1-\bar{\alpha}/2}} \Gamma_{\bar{\lambda}}(s-t, \zeta - \gamma_s^{t,z}).$$

Next, we set

$$\mathbf{\Gamma} \otimes H(t,z;s,\zeta) := \int_t^s \int_{\mathbb{R}^2} H(t,z;\tau,w)\mathbf{\Gamma}(\tau,w;s,\zeta)dwd\tau.$$

A recursive application of the Duhamel principle shows that

$$\mathbf{\Gamma}(t,z;s,\zeta) = Z(t,z;s,\zeta) + \mathbf{\Gamma} \otimes H(t,z;s,\zeta) \tag{4.13}$$
$$= Z(t,z;s,\zeta) + \sum_{k=1}^{N-1} Z \otimes H^{\otimes k}(t,z;s,\zeta) + \mathbf{\Gamma} \otimes H^{\otimes N}(t,z;s,\zeta), \qquad N \geqslant 1.$$

As N tends to infinity we formally obtain a representation of $\mathbf{\Gamma}$ as a series of convolution kernels. Unfortunately, as already noticed in [7] and [27], the presence of the transport term makes it hard to control the iterated kernels uniformly in N, as opposed to the classical parametrix method for uniformly parabolic PDEs. Thus the remainder $\mathbf{\Gamma} \otimes H^{\otimes N}$ must be handled with a different technique, borrowed from stochastic control theory: the rest of the proof proceeds exactly in the same way as in [27] to which we refer for a detailed explanation.

Next, we consider the backward equation

$$\breve{\mathcal{A}}_t u_t(z) + \breve{\mathbf{Y}}_t u_t(z) + \partial_t u_t(z) = 0, \qquad t \in [0,s), \ z = (x,v) \in \mathbb{R}^2, \tag{4.14}$$

where $\breve{\mathcal{A}}_t$ is a second order operator of the form

$$\breve{\mathcal{A}}_t = \breve{a}_t \partial_{vv} + \breve{b}_t \partial_v + \breve{c}_t, \qquad z = (x,v) \in \mathbb{R}^2,$$

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 (4.14), that is

$$\overleftarrow{\tilde{\mathcal{A}}}_t^{s_0,\zeta_0} u_t(z) + \overleftarrow{\tilde{\mathbf{Y}}}_t^{s_0,\zeta_0} u_t(z) + \partial_t u_t(z) = 0, \qquad t \in [0, s), \ z \in \mathbb{R}^2, \tag{4.15}$$

where the definition of $\overleftarrow{\tilde{\mathbf{Y}}}_t^{s_0,\zeta_0}$ is analogous to that of $\mathbf{Y}_s^{t_0,z_0}$ in (4.10) and

$$\overleftarrow{\tilde{\mathcal{A}}}_t^{s_0,\zeta_0} := \bar{a}_t(\overleftarrow{\tilde{\gamma}}_t^{s_0,\zeta_0}) \partial_{vv}, \qquad \overleftarrow{\tilde{\gamma}}_t^{s_0,\zeta_0} = \zeta_0 + \int_t^{s_0} \overleftarrow{\tilde{\mathbf{Y}}}_\tau(\overleftarrow{\tilde{\gamma}}_\tau^{s_0,\zeta_0}) d\tau, \qquad t \in [0, s_0].$$

Equation (4.15) has an explicit fundamental solution $\overleftarrow{\tilde{\mathbf{\Gamma}}}^{s_0,\zeta_0} = \overleftarrow{\tilde{\mathbf{\Gamma}}}^{s_0,\zeta_0}(t, z; s, \zeta)$ of Gaussian type, that satisfies estimates analogous to (4.12). The *backward parametrix* for (4.14) is defined as

$$\overleftarrow{\tilde{Z}}(t, z; s, \zeta) = \overleftarrow{\tilde{\mathbf{\Gamma}}}^{s,\zeta}(t, z; s, \zeta), \qquad 0 \leqslant t < s \leqslant T, \ z, \zeta \in \mathbb{R}^2.$$

As in the forward case, Duhamel principle yields the expansion

$$\overleftarrow{\tilde{\mathbf{\Gamma}}}(t, z; s, \zeta) = \overleftarrow{\tilde{Z}}(t, z; s, \zeta) + \sum_{k=1}^{N-1} \overleftarrow{\tilde{Z}} \otimes \overleftarrow{\tilde{H}}^{\otimes k}(t, z; s, \zeta) + \overleftarrow{\tilde{\mathbf{\Gamma}}} \otimes \overleftarrow{\tilde{H}}^{\otimes N}(t, z; s, \zeta), \qquad N \geqslant 1, \tag{4.16}$$

where $\overleftarrow{\tilde{H}}(t, z; s, \zeta) = \left(\overleftarrow{\tilde{\mathcal{A}}}_t + \overleftarrow{\tilde{\mathbf{Y}}}_t - \overleftarrow{\tilde{\mathcal{A}}}_t^{s_0,\zeta_0} - \overleftarrow{\tilde{\mathbf{Y}}}_t^{s_0,\zeta_0}\right) \overleftarrow{\tilde{Z}}(t, z; s, \zeta)$ and the rest of the proof proceeds as in the forward case. In particular, existence and estimates for the fundamental solutions of (4.1) and (4.8) (in the sense of Definitions 3 and 5) follow from the parametrix expansions (4.13) and (4.16). Eventually, it suffices to go back to the original variables to conclude the proof: we refer to [27], Section 6, for full details.

4.3 Proof of Corollary 1

By Theorem 4 there exists a fundamental solution $\overleftarrow{\tilde{\mathbf{\Gamma}}}$ of equation (2.15), in the sense of Definition 5. Moreover, since $\sigma \equiv 0$, $\overleftarrow{\tilde{\mathbf{\Gamma}}}$ satisfies estimates (2.11), (2.12) and (2.13) with $\bar{g}_{t,s}^{\mathrm{IW},-1} \equiv \mathrm{Id}$ and $\overleftarrow{\tilde{\gamma}}_t^{s,\zeta} = \gamma_{t-s}^{\mathbf{B}}(\zeta)$ as in Definition 2. Then, the function

$$f_t(z) := \int_{\mathbb{R}^{d+1}} \overleftarrow{\tilde{\mathbf{\Gamma}}}(t, z, T, \zeta) \varphi(\zeta) d\zeta, \qquad (t, z) \in [0, T] \times \mathbb{R}^{d+1},$$

solves problem (2.14). Since $\varphi \in bC(\mathbb{R}^{d+1})$, we have

$$\sup_{z \in \mathbb{R}^{d+1}} |f_z(z)| \leqslant \|\varphi\|_\infty \sup_{z \in \mathbb{R}^{d+1}} \int_{\mathbb{R}^{d+1}} \overleftarrow{\tilde{\mathbf{\Gamma}}}(t, z, T, \zeta) d\zeta \leqslant C$$

for a positive constant C. Estimate (2.16) is proven in [9], Proposition 3.3.

5 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 [30]). This is standard material which resumes the original results in [15], [17], [18], [20], [31] (see also the monographs [30] and [22]).

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

$$\int_t^s u_r \star dW_r := \int_{T-s}^{T-t} u_{T-r} d\overleftarrow{W}_r, \qquad 0 \leqslant t \leqslant s \leqslant T, \tag{5.1}$$

under the assumptions on u for which the RHS of (5.1) 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

$$\int_t^s u_r \star dW_r := \lim_{|\pi| \to 0^+} \sum_{k=1}^n u_{t_k} \left(W_{t_k} - W_{t_{k-1}} \right) \tag{5.2}$$

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

$$-dX_t = b_t dt + \sigma_t \star dW_t. \tag{5.3}$$

Theorem 12 (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 (5.3). Then*

$$-dv(t, X_t) = \left((\partial_t v)(t, X_t) + \frac{1}{2}(\sigma_t \sigma_t^*)_{ij}(\partial_{x_i x_j} v)(t, X_t) + (b_t)_i (\partial_{x_i} v)(t, X_t) \right) dt + (\sigma_t)_{ij} \left(\partial_{x_i} v \right)(t, X_t) \star dW_t^j. \tag{5.4}$$

A crucial tool in our analysis is the following

Theorem 13 (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*

$$dX_s^{t,x} = b(s, X_s^{t,x})ds + \sigma(s, X_s^{t,x})dW_s \qquad (5.5)$$

with initial condition $X_t^{t,x} = x$. Then the process $(t,x) \mapsto X_T^{t,x}$ solves the backward SPDE

$$\begin{cases} -dX_T^{t,x} = \mathcal{L}X_T^{t,x}dt + \sigma_{ij}(t,x)\partial_{x_i}X_T^{t,x} \star dW_t^j, \\ X_T^{T,x} = x, \end{cases} \qquad (5.6)$$

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 (5.6) 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 5 The regularity assumption of Theorem 13 on the coefficients is by no means optimal: [30], Theorem 5.1, proves that $(t,x) \mapsto X_T^{t,x}$ is a generalized (or classical, under non-degeneracy conditions) solution of (5.6) 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 [30], Proposition 5.3. Here we follow the "direct" approach proposed in [31]. By standard results for stochastic flows (cf. [22]), $x \mapsto X_T^{t,x}$ is sufficiently regular to support the derivatives in the classical sense. We use the Taylor expansion for C^2-functions:

$$f(\delta) - f(0) = \delta f'(0) + \frac{\delta^2}{2}f''(\lambda\delta), \qquad \lambda \in [0,1]. \qquad (5.7)$$

We have

$$X_T^{t,x} - x = X_T^{t,x} - X_T^{T,x}$$
$$= \sum_{k=1}^n \left(X_T^{t_{k-1},x} - X_T^{t_k,x}\right)$$

(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 (5.7) with $f(\delta) = X_T^{t_k, x+\delta}$ and $\delta = \Delta_k X := X_{t_k}^{t_{k-1},x} - x$)

$$= \sum_{k=1}^n \left(\Delta_k X \partial_x X_T^{t_k,x} + \frac{(\Delta_k X)^2}{2}\partial_{xx}X_T^{t_k, x+\lambda_k\Delta_k X}\right) \qquad (5.8)$$

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 (5.8) 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 (5.2) 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 13.

Corollary 2 (Invariance of the backward diffusion SPDE) *For $v \in bC^2(\mathbb{R}^d)$ and X as in (5.5), let* $V_T^{t,x} = v(X_T^{t,x})$. *Then $V_T^{t,x}$ satisfies the same SPDE (5.6), 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 (5.6) and the backward Itô formula (5.4), we have

$$-dv(X_T^{t,x}) = \left(\frac{\sigma^2(t,x)}{2} v''(X_T^{t,x})(\partial_x X_T^{t,x})^2 + \frac{\sigma^2(t,x)}{2} v'(X_T^{t,x}) \partial_{xx} X_T^{t,x} + b(t,x) v'(X_T^{t,x}) \partial_x X_T^{t,x} \right) dt$$

$$+ \sigma(t,x) v'(X_T^{t,x}) \partial_x X_T^{t,x} \star dW_t =$$

(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

$$
\begin{aligned}
\partial_{x_h} V_T^{t,x} &= (\nabla v)(X_T^{t,x})\partial_{x_h} X_T^{t,x}, \\
\partial_{x_h x_k} V_T^{t,x} &= (\partial_{ij} v)(X_T^{t,x})(\partial_{x_h} X_T^{t,x})_i(\partial_{x_k} X_T^{t,x})_j + (\nabla v)(X_T^{t,x})(\partial_{x_h x_k} X_T^{t,x}),
\end{aligned}
\tag{5.9}
$$

and by (5.6) and (5.4)

$$
\begin{aligned}
-dv(X_T^{t,x}) = & \left(\frac{1}{2} \left((\nabla X_T^{t,x})\sigma(t,x)((\nabla X_T^{t,x})\sigma(t,x))^* \right)_{ij} (\partial_{ij} v)(X_T^{t,x}) \right) dt \\
& + \left(\frac{1}{2}(\sigma(t,x)\sigma^*(t,x))_{ij}\partial_{x_j x_i} X_T^{t,x} + b(t,x)\nabla X_T^{t,x} \right) (\nabla v)(X_T^{t,x})dt \\
& + (\nabla v)(X_T^{t,x})(\nabla X_T^{t,x})\sigma(t,x) \star dW_t =
\end{aligned}
$$

(by (5.9))

$$
= \left(\frac{1}{2}(\sigma(t,x)\sigma^*(x))_{ij}\partial_{x_j x_i} V_T^{t,x} + b(t,x)\nabla V_T^{t,x} \right) dt + \nabla V_T^{t,x}\sigma(t,x) \star dW_t.
$$

6 Summary of notations

The points in $\mathbb{R}^2$ are denoted by $z = (x,v)$ and $\zeta = (\xi,\nu)$, with z generally standing for the *initial point* and ζ for the *final point*. Analogously the points in $\mathbb{R}^{d+1}$ are denoted by $z = (x,v_1,\dots,v_d)$ and $\zeta = (\xi,\nu_1,\dots,\nu_d)$. 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 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. In particular, as in Section 2 we denote by:

- $t \mapsto \gamma_t^{\mathbf{B}}(z) = (x + tv, v)$ the integral curve, starting from z, of the advection vector field $v\partial_x$;
- $g^{\mathrm{IW}}(x,v) := (x,\gamma_{t,s}^{\mathrm{IW}}(x,v))$ is the forward stochastic flow of diffeomorphism defined by the SDE

$$
\gamma_{t,s}^{\mathrm{IW}}(x,v) = v - \int_t^s \sigma_\tau(x,\gamma_{t,\tau}^{\mathrm{IW}}(x,v))dW_\tau, \qquad s \in [t,T],
$$

and $g^{\mathrm{IW},-1}$ is its inverse;
- $\bar{g}^{\mathrm{IW}}(x,v) := \left(x, \bar{\gamma}_{t,s}^{\mathrm{IW}}(x,v) \right)$ is the backward stochastic flow of diffeomorphism defined by

$$
\bar{\gamma}_{t,s}^{\mathrm{IW}}(x,v) = v + \int_t^s \sigma_\tau(x,\bar{\gamma}_{\tau,s}^{\mathrm{IW}}(x,v)) \star dW_\tau, \qquad t \in [0,s],
$$

and $\bar{g}^{\mathrm{IW}}(x,v)$ is its inverse;
- $\gamma_s^{t,z}$ is the integral curve, starting at z at time t, defined by the ODE

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

where $\mathbf{Y}_{t,s}(z) := \left((\gamma_{t,s}^{\mathrm{IW}})_1(z), -(\gamma_{t,s}^{\mathrm{IW}}(z))_1(\nabla_v \gamma_{t,s}^{\mathrm{IW}})^{-1}(z)\partial_x \gamma_{t,s}^{\mathrm{IW}}(z) \right)$;

- $\overleftarrow{\gamma}_t^{s,\zeta}$ is the integral curve, ending at ζ at time s defined by

$$\overleftarrow{\gamma}_t^{s,\zeta} = \zeta + \int_t^s \overleftarrow{\mathbf{Y}}_{\tau,s}(\overleftarrow{\gamma}_\tau^{s,\zeta})d\tau, \qquad t \in [0,s],$$

where $\overleftarrow{\mathbf{Y}}_{t,s}(z) := \left((\overleftarrow{\gamma}_{t,s}^{\mathrm{IW}})_1(z), -(\overleftarrow{\gamma}_{t,s}^{\mathrm{IW}}(z))_1 (\nabla_v \overleftarrow{\gamma}_{t,s}^{\mathrm{IW}})^{-1}(z)\partial_x \overleftarrow{\gamma}_{t,s}^{\mathrm{IW}}(z) \right).$

Lastly, $\Gamma_\lambda(t,x,v)$ denotes the Gaussian kernel

$$\Gamma_\lambda(t,x,v) = \frac{1}{t^{\frac{d+3}{2}}} \exp\left(-\frac{1}{2\lambda}\left(\frac{x^2}{t^3} + \frac{|v|^2}{t} \right) \right), \qquad t > 0, \ (x,v) \in \mathbb{R} \times \mathbb{R}^d, \ \lambda > 0.$$

References

1. Anceschi, F., Polidoro, S.: A survey on the classical theory for Kolmogorov equation. Matematiche (Catania) **75**(1), 221–258 (2020). DOI 10.4418/2020.75.1.11. URL https://doi-org.ezproxy.unibo.it/10.4418/2020.75.1.11

2. Barucci, E., Polidoro, S., Vespri, V.: Some results on partial differential equations and Asian options. Math. Models Methods Appl. Sci. **11**(3), 475–497 (2001)

3. Bramanti, M., Polidoro, S.: Fundamental solutions for Kolmogorov-Fokker-Planck operators with time-depending measurable coefficients. Math. Eng. **2**(4), 734–771 (2020). DOI 10.3934/mine.2020035. URL https://doi-org.ezproxy.unibo.it/10.3934/mine.2020035

4. Cercignani, C.: The Boltzmann equation and its applications. Springer-Verlag, New York (1988)

5. Chaleyat-Maurel, M., Michel, D.: Hypoellipticity theorems and conditional laws. Z. Wahrsch. Verw. Gebiete **65**(4), 573–597 (1984). DOI 10.1007/BF00531840. URL https://doi.org/10.1007/BF00531840

6. Chow, P.L., Jiang, J.L.: Stochastic partial differential equations in Hölder spaces. Probab. Theory Related Fields **99**(1), 1–27 (1994). DOI 10.1007/BF01199588. URL http://dx.doi.org/10.1007/BF01199588

7. Delarue, F., Menozzi, S.: Density estimates for a random noise propagating through a chain of differential equations. J. Funct. Anal. **259**(6), 1577–1630 (2010). DOI 10.1016/j.jfa.2010.05.002. URL https://doi.org/10.1016/j.jfa.2010.05.002

8. Desvillettes, L., Villani, C.: On the trend to global equilibrium in spatially inhomogeneous entropy-dissipating systems: the linear Fokker-Planck equation. Comm. Pure Appl. Math. **54**(1), 1–42 (2001)

9. Di Francesco, M., Pascucci, A.: A continuous dependence result for ultraparabolic equations in option pricing. J. Math. Anal. Appl. **336**(2), 1026–1041 (2007). DOI 10.1016/j.jmaa.2007.03.031. URL https://doi-org.ezproxy.unibo.it/10.1016/j.jmaa.2007.03.031

10. Folland, G.B., Stein, E.M.: Hardy spaces on homogeneous groups, Mathematical Notes, vol. 28. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo (1982)

11. Helffer, B., Nier, F.: Hypoelliptic estimates and spectral theory for Fokker-Planck operators and Witten Laplacians, Lecture Notes in Mathematics, vol. 1862. Springer-Verlag, Berlin (2005). DOI 10.1007/b104762. URL https://doi-org.ezproxy.unibo.it/10.1007/b104762

12. Hörmander, L.: Hypoelliptic second order differential equations. Acta Math. **119**, 147–171 (1967)

13. Kallianpur, G.: Stochastic filtering theory, Applications of Mathematics, vol. 13. Springer-Verlag, New York-Berlin (1980)

14. Kolmogorov, A.: Zufallige Bewegungen. (Zur Theorie der Brownschen Bewegung.). Ann. of Math., II. Ser. **35**, 116–117 (1934)

15. 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**, 691–708 (1973)

16. Krylov, N.V.: Hörmander's theorem for stochastic partial differential equations. Algebra i Analiz **27**(3), 157–182 (2015)

17. Krylov, N.V., Rozovsky, B.L.: On the first integrals and Liouville equations for diffusion processes. In: Stochastic differential systems (Visegrád, 1980), Lecture Notes in Control and Information Sci., vol. 36, pp. 117–125. Springer, Berlin-New York (1981)

18. Krylov, N.V., Rozovsky, B.L.: Characteristics of second-order degenerate parabolic Itô equations. Trudy Sem. Petrovsk. (8), 153–168 (1982)

19. Krylov, N.V., Zatezalo, A.: A direct approach to deriving filtering equations for diffusion processes. Appl. Math. Optim. **42**(3), 315–332 (2000). DOI 10.1007/s002450010015. URL http://dx.doi.org/10.1007/s002450010015

20. Kunita, H.: On backward stochastic differential equations. Stochastics **6**(3-4), 293–313 (1981/82). DOI 10.1080/17442508208833209. URL https://doi-org.ezproxy.unibo.it/10.1080/17442508208833209

21. Kunita, H.: Stochastic partial differential equations connected with nonlinear filtering. In: Nonlinear filtering and stochastic control (Cortona, 1981), Lecture Notes in Math., vol. 972, pp. 100–169. Springer, Berlin (1982). DOI 10.1007/BFb0064861. URL https://doi.org/10.1007/BFb0064861

22. Kunita, H.: Stochastic flows and stochastic differential equations, Cambridge Studies in Advanced Mathematics, vol. 24. Cambridge University Press, Cambridge (1990)

23. Lions, P.L.: On Boltzmann and Landau equations. Philos. Trans. Roy. Soc. London Ser. A **346**(1679), 191–204 (1994)

24. Mikulevicius, R.: On the Cauchy problem for parabolic SPDEs in Hölder classes. Ann. Probab. **28**(1), 74–103 (2000). DOI 10.1214/aop/1019160112. URL https://doi-org.ezproxy.unibo.it/10.1214/aop/1019160112

25. Pardoux, E.: Stochastic partial differential equations and filtering of diffusion processes. Stochastics **3**(2), 127–167 (1979). DOI 10.1080/17442507908833142. URL https://doi.org/10.1080/17442507908833142

26. Pascucci, A.: PDE and martingale methods in option pricing. Bocconi&Springer Series. Springer-Verlag, New York (2011)

27. Pascucci, A., Pesce, A.: On stochastic Langevin and Fokker-Planck equations: the two-dimensional case. arXiv:1910.05301 (2019)

28. Pascucci, A., Pesce, A.: The parametrix method for parabolic SPDEs. Stochastic Process. Appl. **130**(10), 6226–6245 (2020). DOI 10.1016/j.spa.2020.05.008. URL https://doi-org.ezproxy.unibo.it/10.1016/j.spa.2020.05.008

29. Qiu, J.: Hörmander-type theorem for Itô processes and related backward SPDEs. Bernoulli **24**(2), 956–970 (2018). DOI 10.3150/16-BEJ816. URL https://doi.org/10.3150/16-BEJ816

30. Rozovsky, B.L., Lototsky, S.V.: Stochastic evolution systems, Probability Theory and Stochastic Modelling, vol. 89. Springer, Cham (2018). DOI 10.1007/978-3-319-94893-5. URL https://doi-org.ezproxy.unibo.it/10.1007/978-3-319-94893-5. Linear theory and applications to non-linear filtering, Second edition of [MR1135324]

31. Veretennikov, A.Y.: "Inverse diffusion" and direct derivation of stochastic Liouville equations. Mat. Zametki **33**(5), 773–779 (1983)

32. Veretennikov, A.Y.: On backward filtering equations for SDE systems (direct approach). In: Stochastic partial differential equations (Edinburgh, 1994), London Math. Soc. Lecture Note Ser., vol. 216, pp. 304–311. Cambridge Univ. Press, Cambridge (1995). DOI 10.1017/CBO9780511526213.019. URL https://doi-org.ezproxy.unibo.it/10.1017/CBO9780511526213.019

33. Veretennikov, A.Y.: On SPDE and backward filtering equations for SDE systems (direct approach). https://arxiv.org/abs/1607.00333 (July 2016)