ver-filt_2017d-2_0.tex

On SPDE and backward filtering equations for SDE systems (direct approach) – 2

A. Yu. Veretennikov*†

August 25, 2017

Abstract

A direct approach to linear backward filtering equations for SDE systems is proposed. This preprint contains an extended and corrected version of the paper 1995 in the LMS Lecture Notes [10] combined with another paper by the author on the direct approach to linear SPDEs for SDEs [9]. The first part of this extension was presented in [11]. In this part 2, a more general diffusion filtering model with some mild inter-relation between the signal and the observation components – via the same second Wiener process – leading to a more general SPDE is presented.

1 Introduction

Filtering theory is one of the main sources of stochastic partial differential equations (SPDE's). In this paper the filtering problem for the two-dimensional stochastic

*University of Leeds, UK; National Research University Higher School of Economics, and Institute for Information Transmission Problems, Moscow, Russia, e-mail: a.veretennikov @ leeds.ac.uk

†The author is grateful to the financial support by the DFG through the CRC 1283 "Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications" at Bielefeld University and for highly stimulating atmosphere during his stay there in August 2017. Also, this study has been funded by the Russian Academic Excellence Project '5-100' and by the Russain Science Foundation project no. 17-11-01098.

differential equation system is considered,

$$
\begin{aligned}
dX_t &= f(X_t)dt + \sigma(X_t)dw_t^1 + \sigma_2(X_t)dw_t^2, \quad X_0 = x, \\
dY_t &= h(X_t)dt + dw_t^2, \quad Y_0 = y, \quad t \ge 0,
\end{aligned}
\tag{1}
$$

where functions f and h are smooth and bounded (see the details below in the assumptions), w^1 and w^2 are independent standard Wiener processes on some probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t,\, t \ge 0), \mathbb{P})$. Note that non-degeneracy of (σ, σ_2) is not assumed; however, it *may* be required for the uniqueness of solution of the the problem (3)–(4) below which uniqueness is not the goal of this paper (see, e.g., [4] concerning this issue). Initial data $X_0 = x$ and $Y_0 = y$ are assumed non-random. (In fact, both X_0 and Y_0 *may* be distributed being mutually independent with (w^1, w^2) and this might be helpful in filtering, but we do not pursue this goal here.) The problem is to describe the estimate of the unobservable signal process X_t via the observable component Y_t, $0 \le s \le t$, which would be optimal in the mean–square sense, i.e.,

$$
m_t \equiv m_{0,t} = \mathbb{E}[g(X_t)|\mathcal{F}_t^Y], \quad t \ge 0.
$$

It will be also useful to let the variable t change from s to T (or to ∞). In this case, initial data will be

$$
X_s = x, \quad Y_s = y.
$$

In fact, the answer is known even for more general situations, see [3, 4, 8], et al. In particular, m_t may be represented via backward stochastic differential equations, which makes sense if we are interested in the optimal estimation for some fixed time t; in this case we should find the solution of our backward SPDE and substitute there the trajectory of our observation process.

In this paper we present a *direct* approach to such a representation, using a similar idea for an equation without filtering, that is, for a completely observed SDE trajectory. This preprint is an improved version of the paper [10] presented along with the main lines of the calculus from [9]. The matter is that the standard way – as in [4, 7, 8] – is to write down the SPDE, then establish existence and uniqueness of solution in appropriate (Sobolev) classes, then apply Ito's (or Ito–Wentzell's) formula and, hence, justify that this solution, indeed, coincides with the desired conditional expectation. Apparently, this way assumes that somehow the equation should be known in advance. What the direct approach provides is exactly how to derive the equation "by hand" without reference to any big theory. Note that there is a paper [5] with a very similar title; yet, this is a quite different *direct approach,* which also stems from Krylov's idea of representing solutions of SDEs as solutions of linear SPDEs, see [3], [8], [9]. The proof of the Theorem 1 actually contains some additional information about convergence of discretised approximate solutions toward the limit.

The paper consists of three sections. Number one is the Introduction; the second one contains the main result about filtering SPDEs as well as two auxiliary Lemmata; for the proof of the Lemma 1 see [11] while the second Lemma is a well known result with a reference provided; and the third section contains the proof of the main result – the Theorem 1. The first part of this publication was a preprint [11] – already mentioned a few lines earlier – devoted to a more elementary SDE system with the signal X which also evolved functionally independently of the observations Y, but there was no dw^2 term in the equation of the signal. In this part we allow the term dw^2 in the signal equation, yet the resulting SPDE remains only "in x variable"; however, it also contains an additional gradient stochastic term with $\star dY_t$ which makes this presentation more technically involved in comparison to the case in [11] . We also keep both Lemmata (without proofs) for the readers' convenience. The author believes that *replacing* the preprint [11] would not be a reasonable step because the latter more elementary version is less technical and, hence, offers a more explicit presentation of the ideas. On the other hand, the present version covers a much wider class of signals and, thus, also has a rights to be presented.

It should be also added a little about why this paper is written now. On the one hand, a new and hopefully simple approach which does not require any big theory is a good thing. Yet, the results established are known. However, on the other hand, for new problems which emerge with the development of new engineering techniques and devices, this "new" approach may be rather helpful. Hence, the paper is written with a hope to be useful for further new problems such as filtering given information at discrete moments of time. These new problems are postponed till further research.

2 Main result and auxiliary lemmata

Due to Girsanov's theorem, process Y_t, $0 \le t \le T$ is a Wiener process on some the probability space with some new measure $(\Omega, \mathcal{F}, (\mathcal{F}_t, 0 \le t \le T), \tilde{\mathbb{P}})$ (see below).

Theorem 1 (backward SPDE) *Let $\sigma, \sigma_2, f, h \in C_b^3$. Then the process m_t may be represented as follows:*

$$m_T = \frac{v^g(0,x)}{v^1(0,x)}, \qquad (2)$$

where the processes v^g and v^1 satisfy the following linear backward stochastic differ-

ential equation (the same for both functions):

$$
\begin{aligned}
-dv^g(t,x) = &\left[\frac{\sigma^2(x)+\sigma_2^2(x)}{2} v^g_{xx}(t,x) + f(x)v^g_x(t,x)\right] dt \\
&+ [h(x)v^g(t,x) + \sigma_2(x)v^g_x(t,x)] \star dY_t, \quad 0 \le t \le T,
\end{aligned}
\tag{3}
$$

with terminal data

$$
v^g(T,x) = g(x), \quad x \in \mathbb{R}^1. \tag{4}
$$

Note that the denominator in (2) is strictly positive a.s. as a conditional expectation of a strictly positive random variable with respect to some new probability measure. This will be commented in the proof.

Here in (3) $\int \cdot \star dY_t$ means "backward" stochastic Ito integral, i.e., a normal "regular" stochastic Ito integral with inverse time, see [3, 8]. It may be formally defined, for example, by the formula

$$
\begin{aligned}
&\int_0^T h(x)v^g(t,x) \star dY_t := \int_0^T h(x)\tilde{v}^g(t,x)d\tilde{Y}_t, \\
&\tilde{Y}_t = Y_T - Y_{T-t}, \quad \tilde{v}^g(t,x) = v^g(T-t,x),
\end{aligned}
\tag{5}
$$

where $\int_0^T h(x)\tilde{v}^g(t,x)d\tilde{Y}_t$ is a standard Itô's integral. (The only small nuance is that this integral *might* be naturally defined up to the $\pm$ sign – which relates simply to how a Wiener process in the inverse time is defined – and, clearly, this sign would also affect the sign in the last term of the equation (3); this will be commented later.) The function v^1 has its terminal condition $v^1(T,x) \equiv 1$ and satisfies the same SPDE (3). Notice that the random function $v^g(t,x)$ is, in fact, $F^{w^1,w^2}_{t,T}$-adapted (*not* $F^{w^1,w^2}_{0,t}$-adapted); therefore, the integral above makes sense exactly as a classical standard Itô's one (cf. [8, Theorem 6.3.1]). The system (3)–(4) is a parabolic SPDE with random coefficients; see [8] about solutions of such equations.

Before the proof we recall one more Krylov and Rozovsky's result – the Lemma 1 below – concerning multidimensional SDEs (see [3], [8], [9]).

Let $(Z^{s,z}_t,\ t \ge s,\ s \ge 0,\ z \in R^d)$ be the family of d-dimensional processes depending on the parameters (s,z) and satisfying the following multidimensional SDEs:

$$
dZ^{s,z}_t = b(Z^{s,z}_t)dt + \sigma(Z^{s,z}_t)dw_t, \quad t \ge s, \quad Z^{s,z}_s = z, \tag{6}
$$

where b is a bounded smooth d-dimensional vector, σ is a matrix $d \times d_1$, w_t is a d_1-dimensional Wiener process, $d, d_1 \geq 1$; there are neither any other restrictions on the values d and d_1, nor any non-degenerabilty condition is assumed. We will use the following different notations for the same value:

$$Z_t^{s,z} \equiv Z(s,t,z),$$

and for $t = T$ also

$$Z_T^{s,z} = u(s,z).$$

Recall that here T is fixed throughout the text, and that the multidimensional setting is essential: we will need it in the proof of the Theorem 1 with $d = 2$, $d_1 = 1$.

Lemma 1 *Let $b, \sigma \in C_b^3$. Then the random field $Z_T^{s,z}$ is continuous in all variables (s, T, z). Moreover, continuous partial derivatives exist, the gradient vector $\partial_z Z_t^{s,z} =: Z_z(s,t,z)$ and the Hessian matrix $\partial_{zz}^2 Z_t^{s,z} =: Z_{zz}(s,t,z)$, and the process $u(s,z)$ satisfies an SPDE*

$$\begin{aligned} -du(t,z) = &\left[\frac{1}{2}(\sigma\sigma^*)_{ij}(z)u_{z_iz_j}(t,z) + b^i(z)u_{z_i}(t,z)\right] dt \\ &+\sigma_{ij}(z)u_{z_i}(t,z) \star dw_t^j, \end{aligned} \tag{7}$$

with a terminal condition

$$u(T,z) \equiv z. \tag{8}$$

Here σ^* means the matrix σ transposed, and the equation (7) holds true for each component of the vector $u(t,z) = (u^1(t,z), \ldots, u^d(t,z))^*$.

The *direct* approach to this result may be found in [9] and [11].

Further, we will use the Bayes representation for conditional expectations, also known as Kallianpur–Striebel's formula, see [8].

Lemma 2 *Let the Borel functions h, g be bounded. Then the following representation is valid a.s.:*

$$m_T = \frac{\tilde{\mathbb{E}}[g(X_T)\rho^{-1}|\mathcal{F}_T^Y]}{\tilde{\mathbb{E}}[\rho^{-1}|\mathcal{F}_T^Y]},$$

where $\tilde{\mathbb{E}}$ is the expectation with respect to the measure $\tilde{\mathbb{P}}$: $d\tilde{\mathbb{P}} = \rho P$, with

$$\rho \equiv \rho_{0T} = \exp\left(-\int_0^T h(X_t)dw_t^2 - \frac{1}{2}\int_0^T |h(X_t)|^2 dt\right).$$

The rest in this section is a technical preparation to the proof of the main result. Due to Girsanov's theorem the process $(\tilde{w}_t = w_t^2 + \int_0^t h(X_s)\,ds,\, 0 \le t \le T)$ is a Wiener process on probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t,\, 0 \le t \le T), \tilde{\mathbb{P}})$, independent of w^1. On this space, our system (1) has the form

$$\begin{aligned} dX_t &= f(X_t, Y_t)dt + \sigma(X_t)dw_t^1 + \sigma_2(X_t)dw_t^2, \quad X_0 = x, \\ dY_t &= d\tilde{w}_t, \quad Y_0 = y, \end{aligned} \tag{9}$$

with two independent Wiener processes $(w^1, \tilde{w})$. Replacing w^2 by its expression via $\tilde{w}$, we get,

$$\begin{aligned} dX_t &= f(X_t, Y_t)dt + \sigma(X_t)dw_t^1 + \sigma_2(X_t)(d\tilde{w}_t - h(X_t)\,dt), \quad X_0 = x, \\ dY_t &= d\tilde{w}_t, \quad Y_0 = y, \end{aligned} \tag{10}$$

or, equivalently,

$$\begin{aligned} dX_t &= (f(X_t) - \sigma_2(X_t)h(X_t))dt + \sigma(X_t)dw_t^1 + \sigma_2(X_t)d\tilde{w}_t, \quad X_0 = x, \\ dY_t &= d\tilde{w}_t, \quad Y_0 = y, \end{aligned} \tag{11}$$

or, if we denote $\bar{h}(x) = -\sigma_2(x)h(x)$,

$$\begin{aligned} dX_t &= (f(X_t) - \bar{h}(X_t))dt + \sigma(X_t)dw_t^1 + \sigma_2(X_t)d\tilde{w}_t, \quad X_0 = x, \\ dY_t &= d\tilde{w}_t, \quad Y_0 = y, \end{aligned} \tag{12}$$

or, with a new notation $\bar{f}(x) = f(x) - \sigma_2(x)h(x)$,

$$\begin{aligned} dX_t &= \bar{f}(X_t)dt + \sigma(X_t)dw_t^1 + \sigma_2(X_t)d\tilde{w}_t, \quad X_0 = x, \\ dY_t &= d\tilde{w}_t, \quad Y_0 = y. \end{aligned} \tag{13}$$

3 Direct proof of Theorem 1

1. Recall, under the new measure $\tilde{\mathbb{P}}$,

$$dX_t = \bar{f}(X_t)dt + \sigma(X_t)dw_t^1 + \sigma_2(X_t)d\tilde{w}_t, \quad X_0 = x,$$

$$dY_t = d\tilde{w}_t, \quad Y_0 = y.$$

We will also need a system with a variable initial time s:

$$dX_t = \bar{f}(X_t)dt + \sigma(X_t)dw_t^1 + \sigma_2(X_t)d\tilde{w}_t, \quad X_s = x,$$

$$dY_t = d\tilde{w}_t, \quad Y_s = y, \quad t \ge s.$$

The solution of this system is denoted by $(X_t^{s,x}, Y_t^{s,y})$. What we are to prove is,

$$-dv^g(t,x) = \left[\frac{\sigma^2(x) + \sigma_2^2(x)}{2} v_{xx}^g(t,x) + f(x)v_x^g(t,x)\right] dt$$

$$+ [h(x)v^g(t,x) + \sigma_2(x)v_x^g(t,x)] \star dY_t,$$

as $0 \le t \le T$, with terminal data

$$v^g(T,x) = g(x), \quad x \in \mathbb{R}^1.$$

Note that this equation does not depend on y.

2. Denote

$$v^g(s,x) = \tilde{\mathbb{E}}[g(X_T^{s,x})\rho_{s,T}^{-1}|\mathcal{F}_{s,T}^Y].$$

The expression in the right hand side clearly does not depend on y, since the sigma-algebra here is generated only by the *increments* of Y, which trajectory is independent of the component X. Under the conditional expectation we have ρ^{-1}, as well, which also involves only the increments of $\tilde{w} \equiv Y$ and does not depend on the initial state y either. Hence, the notation $v^g(s,x)$ is well justified.

The terminal condition is easily checked,

$$v^g(T,x) = \tilde{\mathbb{E}}[g(X_T^{T,x})\rho_{T,T}^{-1}|\mathcal{F}_{T,T}^Y] = g(x).$$

Now, what we want to establish is precisely the following integral equality (for each $T > 0$ and any $0 \le t_0 \le T$):

$$v^g(t_0,x) - v^g(T,x)$$

$$= \int_{t_0}^{T} \left[\frac{\sigma^2(x)+\sigma_2^2(x)}{2} v^g_{xx}(t,x) + f(x) v^g_x(t,x) \right] dt \tag{14}$$

$$+ \int_{t_0}^{T} [h(x) v^g(t,x) + \sigma_2(x) v^g_x(t,x)] \star d\tilde{w}_t.$$

Note that during the calculus for a while the "shifted" drift $\bar{f}(x)$ will show up; however, at some point it will be transformed back to $f(x)$.

Let us use the identity

$$v^g(t_0,x) - v^g(T,x) = \sum_{i=1}^{N} (v^g(t_{i-1},x) - v^g(t_i,x)),$$

for any partition $t_0 < t_1 < \ldots < t_N = T$. Consider one term from this sum: we have,

$$v^g(t_{i-1},x) - v^g(t_i,x)$$

$$= \tilde{\mathbb{E}}[\rho^{-1}_{t_{i-1},T} g(X(t_{i-1},T,x)) | \mathcal{F}^Y_{t_{i-1},T}] - \tilde{\mathbb{E}}[\rho^{-1}_{t_i,T} g(X(t_i,T,x)) | \mathcal{F}^Y_{t_i,T}]$$

$$= \tilde{\mathbb{E}}[\rho^{-1}_{t_{i-1},T} g(X(t_{i-1},T,x)) | \mathcal{F}^{\tilde{w}}_{t_{i-1},T}] - \tilde{\mathbb{E}}[\rho^{-1}_{t_i,T} g(X(t_i,T,x)) | \mathcal{F}^{\tilde{w}}_{t_i,T}].$$

3. Using continuity of the family $X(s,T,x)$ with respect to all variables and existence of two continuous partial derivatives with respect to x (see [1]) we get a.s. by virtue of Taylor's expansion,

$$X(t_{i-1},T,x) = X(t_i,T,X^{t_{i-1},x}_{t_i})$$

$$= X^{t_i,x}_T + X_x(t_i,T,x)(X^{t_{i-1},x}_{t_i} - x)$$

$$+\frac{1}{2} X_{xx}(t_i,T,x)(X^{t_{i-1}}_{t_i} - x)^2 + \alpha^1_i$$

$$= X_T^{t_i,x} + X_x(t_i,T,x)(\bar{f}(x)\Delta t_i + \sigma(x)\Delta w_{t_i}^1 + \sigma_2(x)\Delta\tilde{w}_{t_i})$$

$$+\frac{\sigma^2(x)+\sigma_2^2(x)}{2}X_{xx}(t_i,T,x)\Delta t_i + \alpha_i^2,$$

where $\Delta t_i = t_i - t_{i-1}$, $\Delta w_{t_i}^j = w_{t_i}^j - w_{t_{i-1}}^j$, and $|\alpha_i^1| + |\alpha_i^2| = o(\Delta t_i)$ in the mean-square sense. We used, in particular, the almost sure convergence of the approximations to the quadratic variation for stochastic integrals, which implies not only $(\Delta w_{t_i}^1)^2 \approx \Delta t_i$, but also $(\Delta\tilde{w}_{t_i})^2 \approx \Delta t_i$ a.s. under $\tilde{\mathbb{P}}$. As was mentioned earier, the drift coefficient $\bar{f}$ appears here. Hence, we have,

$$g(X(t_{i-1},T,x)) = g(X(t_i,T,X_{t_i}^{t_{i-1},x}))$$

$$= g\Big(X_T^{t_{i-1},x} + X_x(t_i,T,x)(X_{t_i}^{t_{i-1},x} - x)$$

$$+\frac{1}{2}X_{xx}(t_i,T,x)(X_{t_i}^{t_{i-1},x} - x)^2 + \alpha_i^1\Big)$$

$$= g\left(X_T^{t_i,x} + X_x(t_i,T,x)(\bar{f}(x)\Delta t_i + \sigma(x)\Delta w_{t_i}^1 + \sigma_2(x)\Delta\tilde{w}_{t_i})\right.$$

$$\left.+\frac{\sigma^2(x)+\sigma_2^2(x)}{2}X_{xx}(t_i,T,x)\Delta t_i + \alpha_i^2\right)$$

$$= g(X_T^{t_i,x})$$

$$+g_x(X_T^{t_i,x})\left(X_x(t_i,T,x)(\bar{f}(x)\Delta t_i + \sigma(x)\Delta w_{t_i}^1 + \sigma_2(x)\Delta\tilde{w}_{t_i}) + \frac{\sigma^2(x)+\sigma_2^2(x)}{2}X_{xx}(t_i,T,x)\Delta t_i\right)$$

$$+\frac{\sigma^2(x)+\sigma_2^2(x)}{2}g_{xx}(X_T^{t_i,x})\Delta t_i + o(\Delta t_i).$$

Denote $V(s,t,x) = g(X(s,t,x))$. Then, assuming that $g \in C^2$, we have,

$$V_x = g_x X_x; \quad V_{xx} = g_x X_{xx} + g_{xx} X_x^2,$$

where we dropped the arguments in g_x, g_{xx}, X_x, and X_{xx} for brevity. So,

$$g(X(t_{i-1},T,x)) = V(t_{i-1},T,x)$$

$$= g(X_T^{t_i,x})$$

$$+g_x(X_T^{t_i,x})(X_x(t_i,T,x)(\bar f(x)\Delta t_i + \Delta w^1_{t_i} + \sigma_2(x)\Delta\tilde w_{t_i}) + \frac{\sigma^2(x)+\sigma_2^2(x)}{2}X_{xx}(t_i,T,x)\Delta t_i)$$

$$+\frac{\sigma^2(x)+\sigma_2^2(x)}{2}(X_x(t_i,T,x))^2 g_{xx}(X_T^{t_i,x})\Delta t_i + o(\Delta t_i)$$

$$= V(t_i,T,x) + V_x(t_i,T,x)(\bar f(x)\Delta t_i + \sigma(x)\Delta w^1_{t_i} + \sigma_2(x)\Delta\tilde w_{t_i})$$

$$+\frac{\sigma^2(x)+\sigma_2^2(x)}{2}V_{xx}(t_i,T,x)\Delta t_i + o(\Delta t_i).$$

This is an indication that conditional expectation of $V = g(X)$ should satisfy the same equation as for X itself – that is, in the case $g(x) \equiv x$ – just with another terminal condition.

4. Thus,

$$\tilde{\mathbb{E}}[\rho^{-1}_{t_{i-1},T} g(X(t_{i-1},T,x))|\mathcal{F}^{\tilde w}_{t_{i-1},T}]$$

$$= \tilde{\mathbb{E}}[\rho^{-1}_{t_{i-1},T} g(X(t_i,T,X_{t_i}^{t_{i-1},x}))|\mathcal{F}^{\tilde w}_{t_{i-1},T}] = \tilde{\mathbb{E}}[\rho^{-1}_{t_{i-1},T} V(t_{i-1},T,x)|\mathcal{F}^{\tilde w}_{t_{i-1},T}]$$

$$= \tilde{\mathbb{E}}[\rho^{-1}_{t_{i-1},T}\{V(t_i,T,x) + V_x(t_i,T,x)\bar f(x)\Delta t_i$$

$$+V_x(t_i,T,x)(\sigma(x)\Delta w^1_{t_i} + \sigma_2(x)\Delta\tilde w_{t_i}) + \frac{\sigma^2(x)+\sigma_2^2(x)}{2}V_{xx}(t_i,T,x)\Delta t_i\}|\mathcal{F}^{\tilde w}_{t_{i-1},T}] + \alpha_i^3.$$

Here again, $\alpha_i^3 = o(\Delta t_i)$ in the mean square sense, i.e., $(\mathbb{E}|\alpha_i^3|^2)^{1/2} = o(\Delta t_i)$.

5. Now, we would like to replace $\rho^{-1}_{t_{i-1},T}$ by $\rho^{-1}_{t_i,T}$. For this aim we apply the Lemma 1 to the process $(X_t^{s,x},\ \rho^{-1}_{s,t},\ t \geq s)$. More precisely, let us note that this two-dimensional process satisfies the following SDE system:

$$dX_t^{s,x} = \bar f(X_t^{s,x})dt + \sigma(X_t^{s,x})dw^1_t + \sigma_2(X_t^{s,x})d\tilde w_t, \quad X_s^{s,x} = x,$$

$$d\rho_{s,t}^{-1} = h(X_t^{s,x})\rho_{s,t}^{-1} d\tilde{w}_t, \quad \rho_{s,s}^{-1} = 1, \tag{15}$$

with $s \le t \le T$. Indeed, $\rho_{s,t}^{-1}$ has the following representation:

$$\rho_{s,t}^{-1} = \exp\left(\int_s^t h(X_r^{s,x}) d\tilde{w}_r - \frac{1}{2}\int_s^t |h(X_r^{s,x})|^2 dr\right).$$

Let us consider a bit more general set of processes $\{(X_t^{s,x}, \rho_{s,t}^{-1,\xi})\}$ which satisfy SDE's

$$\begin{gathered} dX_t^{s,x} = \bar{f}(X_t^{s,x})dt + \sigma(X_t^{s,x})dw_t^1 + \sigma_2(X_t^{s,x})d\tilde{w}_t, \quad X_s^{s,x} = x, \\ d\rho_{s,t}^{-1,\xi} = h(X_t^{s,x})\rho_{s,t}^{-1,\xi} d\tilde{w}_t, \quad \rho_{s,s}^{-1,\xi} = \xi, \end{gathered} \tag{16}$$

for $s \le t \le T$, with $\xi > 0$. In fact, $X_t^{s,x}$ here is the same as earlier, and $\rho_t^{-1,\xi}$ has the following representation:

$$\rho_{s,t}^{-1,\xi} = \xi \exp\left(\int_s^t h(X_r^{s,x}) d\tilde{w}_r - \frac{1}{2}\int_s^t |h(X_r^{s,x})|^2 \, dr\right) = \xi \rho_{s,t}^{-1}. \tag{17}$$

Recall that due to the Lemma 1,

$$\begin{gathered} -du(t,z) = \left[\frac{1}{2}(\sigma\sigma^*)_{ij}(z) u_{z_i z_j}(t,z) + b^i(z) u_{z_i}(t,z)\right] dt \\ + \sigma_{ij}(z) u_{z_i}(t,z) \star dw_t^j, \end{gathered}$$

with (the argument x in σ, σ_2, h are dropped for brevity)

$$(\sigma\sigma^*)(x,\xi) = \begin{pmatrix} \sigma & \sigma_2 \\ 0 & h\xi \end{pmatrix} \times \begin{pmatrix} \sigma & 0 \\ \sigma_2 & h\xi \end{pmatrix} = \begin{pmatrix} \sigma^2 + \sigma_2^2 & \sigma_2 h\xi \\ \sigma_2 h\xi & h^2\xi^2 \end{pmatrix}.$$

Thus,

$$-d_s \rho_{s,t}^{-1,\xi}$$

$$= \left[\frac{1}{2}h^2(x)\xi^2(\rho_{s,t}^{-1,\xi})_{\xi\xi} + \sigma_2(x)h(x)\xi(\rho_{s,t}^{-1,\xi})_{\xi}x + \frac{\sigma^2(x)+\sigma_2^2(x)}{2}(\rho_{s,t}^{-1,\xi})_{xx} + \bar{f}(x)(\rho_{s,t}^{-1,\xi})_x\right] dt$$

$$+(\rho_{s,t}^{-1,\xi})_x(\sigma(x)\star dw_t^1+\sigma_2(x)\star d\tilde{w}_t)+h(x)\xi(\rho_{s,t}^{-1,\xi})_\xi\star d\tilde{w}_t.$$

$$=\left[\frac{1}{2}h^2(x)\xi^2(\rho_{s,t}^{-1,\xi})_{\xi\xi}+\sigma_2(x)h(x)\xi(\rho_{s,t}^{-1})_x+\frac{\sigma^2(x)+\sigma_2^2(x)}{2}(\rho_{s,t}^{-1,\xi})_{xx}+\bar{f}(x)\xi(\rho_{s,t}^{-1})_x\right]dt$$

$$+\xi(\rho_{s,t}^{-1})_x(\sigma(x)\star dw_t^1+\sigma_2(x)\star d\tilde{w}_t)+h(x)\xi\rho_{s,t}^{-1}\star d\tilde{w}_t.$$

We used that $(\rho_{s,t}^{-1,\xi})_\xi=\rho_{s,t}^{-1}$ and, hence, $(\rho_{s,t}^{-1,\xi})_{\xi x}=(\rho_{s,t}^{-1})_x$; also, $(\rho_{s,t}^{-1,\xi})_x=\xi(\rho_{s,t}^{-1})_x$. Recall that

$$\bar{f}(x)+\sigma_2(x)h(x)=f(x).$$

This allows us to transform $\bar{f}$ in the expressions related to $(\rho_{s,t}^{-1,\xi})_\xi$ into f. However, the term $\bar{f}$ will still remain in the expressions related to V; this will be tackled a bit later. So, from the equality above and because $(\rho_{s,t}^{-1,\xi})_{\xi\xi}=0$, we get,

$$-d_s\rho_{s,t}^{-1,\xi}=\left[\frac{\sigma^2(x)+\sigma_2^2(x)}{2}(\rho_{s,t}^{-1,\xi})_{xx}+f(x)\xi(\rho_{s,t}^{-1})_x\right]dt$$

$$+(\rho_{s,t}^{-1,\xi})_x(\sigma(x)\star dw_t^1+\sigma_2(x)\star d\tilde{w}_t)+h(x)\xi(\rho_{s,t}^{-1})\star d\tilde{w}_t.$$

Note that due to the linear in ξ representation (17), all derivatives with respect to ξ disappear from the expression in the right hand side. So, we may write

$$\rho_{t_{i-1},T}^{-1,\xi}-\rho_{t_i,T}^{-1,\xi}\equiv-\Delta\rho_{t_i}^{-1,\xi}$$

$$=\left[\frac{\sigma^2(x)+\sigma_2^2(x)}{2}(\rho_{t_i,T}^{-1,\xi})_{xx}+f(x)\xi(\rho_{s,t}^{-1})_x\right]\Delta t_i$$

$$+(\rho_{t_i,T}^{-1,\xi})_x\sigma(x)\Delta w_{t_i}^1+(\rho_{t_i,T}^{-1,\xi})_x\sigma_2(x)\Delta\tilde{w}_{t_i}+h(x)\xi\rho_{t_i,T}^{-1}\Delta\tilde{w}_{t_i}+\alpha_i^4,$$

with a similar $o(\Delta t_i)$ property for α_i^4 as for previous $\alpha_i^1,\alpha_i^2,\alpha_i^3$. Below we will use this assertion with $\xi=1$, that is,

$$\rho_{t_{i-1},T}^{-1}-\rho_{t_i,T}^{-1}\equiv-\Delta\rho_{t_i}^{-1}$$

$$=\left[\frac{\sigma^2(x)+\sigma_2^2(x)}{2}(\rho_{t_i,T}^{-1})_{xx}+f(x)(\rho_{t_i,T}^{-1})_x\right]\Delta t_i$$

$$+(\rho_{t_i,T}^{-1})_x\sigma(x)\Delta w_{t_i}^1+(\rho_{t_i,T}^{-1})_x\sigma_2(x)\Delta\tilde{w}_{t_i}+h(x)\rho_{t_i,T}^{-1}\Delta\tilde{w}_{t_i}+\tilde{\alpha}_i^4,$$

with $\tilde{\alpha}_i^4 = o(\Delta t_i)$, as usual, in the mean square sense.

6. Now, we obtain

$$\tilde{\mathbb{E}}[\rho_{t_i,T}^{-1} V(t_{i-1},T,x)|\mathcal{F}_{t_{i-1},T}^{\tilde{w}}]$$

$$= \tilde{\mathbb{E}}\left[\left\{V(t_i,T,x) + \left(\bar{f}(x)V_x(t_i,T,x) + \frac{\sigma^2(x)+\sigma_2^2(x)}{2}V_{xx}(t_i,T,x)\right)\Delta t_i\right.\right.$$

$$\left. + V_x(t_i,T,x)(\sigma(x)\Delta w_{t_i}^1 + \sigma_2(x)\Delta\tilde{w}_{t_i})\right\} \times$$

$$\times\left\{\rho_{t_i,T}^{-1} + \left(\frac{\sigma^2(x)+\sigma_2^2(x)}{2}(\rho_{t_i,T}^{-1})_{xx} + f(x)(\rho_{t_i,T}^{-1})_x\right)\Delta t_i\right.$$

$$\left.\left. + (\rho_{t_i,T}^{-1})_x\sigma(x)\Delta w_{t_i}^1 + (\rho_{t_i,T}^{-1})_x\sigma_2(x)\Delta\tilde{w}_{t_i} + h(x)\rho_{t_i,T}^{-1}\Delta\tilde{w}_{t_i} + \alpha_i^5\right\}|\mathcal{F}_{t_{i-1},T}^{\tilde{w}}\right],$$

where again, $\alpha_i^5 = o(\Delta t_i)$ in the same sense. We multiply using the Itô calculus rules $((\Delta w_t)^2 \approx \Delta t)$,

$$\tilde{\mathbb{E}}[\rho_{t_i,T}^{-1} V(t_{i-1},T,x)|\mathcal{F}_{t_{i-1},T}^{\tilde{w}}]$$

$$= \tilde{\mathbb{E}}\left[V(t_{i-1},T,x)\rho_{t_i,T}^{-1}|\mathcal{F}_{t_{i-1},T}^{\tilde{w}}\right]$$

$$+\tilde{\mathbb{E}}\left[V_x(t_i,T,x)\rho_{t_i,T}^{-1}\left(\bar{f}(x)+\sigma_2(x)h(x)\right)|\mathcal{F}_{t_{i-1},T}^{\tilde{w}}\right]$$

$$+\tilde{\mathbb{E}}\left[V(t_i,T,x)(\rho_{t_i,T}^{-1})_x f(x)|\mathcal{F}_{t_{i-1},T}^{\tilde{w}}\right]$$

$$+\tilde{\mathbb{E}}\left[\frac{\sigma^2(x)+\sigma_2^2(x)}{2}\left(V_{xx}(t_i,T,x)\rho_{t_i,T}^{-1} + 2V_x(t_i,T,x)(\rho_{t_i,T}^{-1})_x + V(t_i,T,x)(\rho_{t_i,T}^{-1})_{xx}\right)\Delta t_i|\mathcal{F}_{t_{i-1},T}^{\tilde{w}}\right]$$

$$+\tilde{\mathbb{E}}\left[\left\{V(t_i,T,x) + \left(\bar{f}(x)V_x(t_i,T,x) + \frac{\sigma^2(x)+\sigma_2^2(x)}{2}V_{xx}(t_i,T,x)\right)\Delta t_i\right.\right.$$

$$\left. + V_x(t_i,T,x)(\sigma(x)\Delta w_{t_i}^1 + \sigma_2(x)\Delta\tilde{w}_{t_i})\right\} \times$$

$$\times \left\{ \rho^{-1}_{t_i,T} + \left(\frac{\sigma^2(x)+\sigma_2^2(x)}{2} (\rho^{-1}_{t_i,T})_{xx} + f(x)(\rho^{-1}_{t_i,T})_x \right) \Delta t_i \right.$$

$$\left. + (\rho^{-1}_{t_i,T})_x \sigma(x) \Delta w^1_{t_i} + (\rho^{-1}_{t_i,T})_x \sigma_2(x) \Delta \tilde{w}_{t_i} + h(x) \rho^{-1}_{t_i,T} \Delta \tilde{w}_{t_i} + \alpha_i^5 \right\} |\mathcal{F}^{\tilde{w}}_{t_{i-1},T} \Bigg],$$

7. Now, note that $\mathcal{F}^{\tilde{w}}_{t_{i-1},T} = \mathcal{F}^{\tilde{w}}_{t_{i-1},t_i} \bigvee \mathcal{F}^{\tilde{w}}_{t_i,T}$ and, moreover this σ-field is independent from w^1. Using the regular calculus for conditional expectations (cf. [8]), we get

$$\tilde{\mathbb{E}}[V(t_i,T,x)\rho^{-1}_{t_i,T} | \mathcal{F}^{\tilde{w}}_{t_{i-1},T}] = \tilde{\mathbb{E}}[V(t_i,T,x)\rho^{-1}_{t_i,T} | \mathcal{F}^{\tilde{w}}_{t_i,T}],$$

and in the same manner we can replace σ-fields $\mathcal{F}^{\tilde{w}}_{t_{i-1},T}$ by $\mathcal{F}^{\tilde{w}}_{t_i,T}$ in *all* terms in the previous step. Also, $\tilde{\mathbb{E}}\left[\Delta w^1_{t_i} | \mathcal{F}^{\tilde{w}}_{t_{i-1},T}\right] = 0$ due to the independence of w^1 and $\tilde{w}$ with respect to the measure $\tilde{\mathbb{P}}$, and $(\Delta w^1_{t_i})^2 \approx \Delta t_i$ as well as and $(\Delta \tilde{w}_{t_i})^2 \approx \Delta t_i$. Hence, we obtain (during this calculus the last remaining term $\bar{f}$ will be replaced by f according to the rule $\bar{f}(x) = f(x) - \sigma_2(x)h(x)$),

$$\tilde{\mathbb{E}}[V(t_i,T,x)\rho^{-1}_{t_{i-1},T} | \mathcal{F}^{\tilde{w}}_{t_{i-1},T}] = \tilde{\mathbb{E}}[V(t_i,T,x)\rho^{-1}_{t_i,T} | \mathcal{F}^{\tilde{w}}_{t_{i-1},T}]$$

$$+\tilde{\mathbb{E}}\left[\frac{\sigma^2(x)+\sigma_2^2(x)}{2} V(t_i,T,x)(\rho^{-1}_{t_i,T})_{xx} + (\sigma^2(x)+\sigma_2^2(x)) V_x(t_i,T,x)(\rho^{-1}_{t_i,T})_x \right.$$

$$\left. + \frac{\sigma^2(x)+\sigma_2^2(x)}{2} V_{xx}(t_i,T,x)\rho^{-1}_{t_i,T} | \mathcal{F}^{\tilde{w}}_{t_{i-1},T} \right] \Delta t_i$$

$$+\tilde{\mathbb{E}}[f(x)(V_x(t_i,T,x)\rho^{-1}_{t_i,T} + V(t_i,T,x)(\rho^{-1}_{t_i,T})_x) | \mathcal{F}^{\tilde{w}}_{t_{i-1},T}] \Delta t_i$$

$$+\tilde{\mathbb{E}}\left[V(t_i,T,x) \left\{ (\rho^{-1}_{t_i,T})_x (\sigma(x)\Delta w^1_{t_i} + \sigma_2(x)\Delta \tilde{w}_{t_i}) + h(x)\rho^{-1}_{t_i,T} \Delta \tilde{w}_{t_i} \right\} + \alpha_i^5 | \mathcal{F}^{\tilde{w}}_{t_{i-1},T} \right]$$

$$+\tilde{\mathbb{E}}\left[\rho^{-1}_{t_i,T} \left\{ V_x(t_i,T,x)(\sigma(x)\Delta w^1_{t_i} + \sigma_2(x)\Delta \tilde{w}_{t_i}) \right\} | \mathcal{F}^{\tilde{w}}_{t_{i-1},T} \right] + \alpha_i^6$$

$$= \tilde{\mathbb{E}}[V(t_i,T,x)\rho^{-1}_{t_i,T} | \mathcal{F}^{\tilde{w}}_{t_{i-1},T}] \Delta t_i + \tilde{\mathbb{E}}\left[\frac{\sigma^2(x)+\sigma_2^2(x)}{2} (V(t_{i-1},T,x)\rho^{-1}_{t_i,T})_{xx} | \mathcal{F}^{\tilde{w}}_{t_{i-1},T} \right] \Delta t_i$$

$$+\tilde{\mathbb{E}}[f(x)(V(t_i,T,x)\rho^{-1}_{t_i,T})_x | \mathcal{F}^{\tilde{w}}_{t_{i-1},T}] \Delta t_i$$

$$+\tilde{\mathbb{E}}\left[\left\{V(t_i,T,x)(\rho^{-1}_{t_i,T})_x + V_x(t_i,T,x)\rho^{-1}_{t_i,T}\right\}\sigma(x)\Delta w^1_{t_i}|\mathcal{F}^{\tilde{w}}_{t_{i-1},T}\right]$$

$$+\tilde{\mathbb{E}}\left[\left\{V(t_i,T,x)(\rho^{-1}_{t_i,T})_x + V_x(t_i,T,x)\rho^{-1}_{t_i,T}\right\}\sigma_2(x)\Delta \tilde{w}_{t_i}|\mathcal{F}^{\tilde{w}}_{t_{i-1},T}\right] + \alpha^6_i$$

$$= \tilde{\mathbb{E}}[V(t_i,T,x)\rho^{-1}_{t_i,T}|\mathcal{F}^{\tilde{w}}_{t_i,T}]\Delta t_i + \tilde{\mathbb{E}}\left[\frac{\sigma^2(x)+\sigma^2_2(x)}{2}(V(t_i,T,x)\rho^{-1}_{t_i,T})_{xx}|\mathcal{F}^{\tilde{w}}_{t_i,T}\right]\Delta t_i$$

$$+\tilde{\mathbb{E}}[f(x)(V(t_i,T,x)\rho^{-1}_{t_i,T})_x|\mathcal{F}^{\tilde{w}}_{t_i,T}]\Delta t_i$$

$$+\tilde{\mathbb{E}}\left[\left\{V(t_i,T,x)(\rho^{-1}_{t_i,T})_x + V_x(t_i,T,x)\rho^{-1}_{t_i,T}\right\}\sigma_2(x)|\mathcal{F}^{\tilde{w}}_{t_i,T}\right]\Delta \tilde{w}_{t_i} + \alpha^6_i$$

$$= v^g(t_i,x) + \frac{\sigma^2(x)+\sigma^2_2(x)}{2}v^g_{xx}(t_i,x)\Delta t_i + f(x)v^g_x(t_i,x)\Delta t_i$$

$$+\left(h(x)v^g(t_i,x) + \sigma_2(x)v^g_x(t_i,x)\right)\Delta \tilde{w}_{t_i} + \alpha^6_i,$$

with a similar property for α^6_i: $\alpha^6_i = o(\Delta t_i)$ in the mean square sense. In the last equality in this calculus we have changed the order of integration and derivation with respect to the x variable.

8. Therefore, we obtain the equality

$$v^g(t_0,x) - v^g(T,x)$$

$$= \sum\left\{\frac{\sigma^2(x)+\sigma^2_2(x)}{2}v^g_{xx}(t_i,x) + f(x)v^g_x(t_i,x)\right\}\Delta t_i$$

$$+\sum\left(\sigma_2(x)v^g_x(t_i,x) + h(x)v^g(t_i,x)\right)\Delta \tilde{w}_{t_i} + \alpha^7,$$

with $\alpha^7 = o(1)$ in the mean square sense as $\sup_i \Delta t_i \to 0$. Letting $\sup_i \Delta t_i \to 0$, we get from here the desired integral equality (14). The Theorem 1 is proved.

References

[1] Yu.N. Blagoveshchenskii, M.I. Freidlin, Certain properties of diffusion processes depending on a parameter. Math. USSR Doklady, 2(3) (1961), 633-636.

[2] N.V. Krylov, On the selection of a Markov process from a system of processes and the construction of quasi-diffusion process, Math. USSR Izv., 7(3) (1973), 691-709.

[3] N.V. Krylov, On explicit formulas for solutions of evolutionary SPDE's (a kind of introduction to the theory). Lecture Notes in Control and Information Sciences, 1992, vol. 176, p. 153-164.

[4] N.V. Krylov, B.L. Rozovsky, Stochastic differential equations and diffusion processes, Russian Math. Surveys, 37(6) (1982), 81-105.

[5] N.V. Krylov, A. Zatezalo, A Direct Approach to Deriving Filtering Equations for Diffusion Processes, Appl. Math. Optimiz., 42(3) (2000), 315-332.

[6] H. Kunita, Stochastic differential equations and stochastic flows of diffeomorphisms, École d'Été de Probabilités de Saint-Flour XII (1982), Lecture Notes in Mathematics, vol. 1097 (1984), 143-303.

[7] E. Pardoux, Stochastic partial differential equations and filtering of diffusion processes, Stochastics, 3(1-4) (1980), 127-167.

[8] B.L. Rozovski, Stochastic Evolution Systems. Kluwer Acad. Publ.: Dordreht et al., 1990.

[9] A.Yu. Veretennikov, Inverse Diffusion and Direct Derivation of Stochastic Liouville Equations. Soviet Math. Notes, 33(5-6) (1983), 397-400.

[10] A.Yu. Veretennikov, On backward filtering equations for SDE systems (direct approach). In: Stochastic Partial Differential Equations, ed. by A. Etheridge. LMS Lecture Notes Series, Cambridge Univ. Press, vol. 216 (1995), 304-311.

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