The Annals of Applied Probability
2013, Vol. 23, No. 5, 2139–2160
DOI: 10.1214/12-AAP896

ROBUST FILTERING: CORRELATED NOISE AND MULTIDIMENSIONAL OBSERVATION

BY D. CRISAN[1], J. DIEHL[2], P. K. FRIZ[3] AND H. OBERHAUSER[3]

Imperial College London, Technical University Berlin, Technical University Berlin and Weierstrass Institute for Applied Analysis and Stochastics Berlin, and Technical University Berlin

In the late seventies, Clark [In *Communication Systems and Random Process Theory* (*Proc. 2nd NATO Advanced Study Inst.*, *Darlington*, 1977) (1978) 721–734, Sijthoff & Noordhoff] pointed out that it would be natural for π_t, the solution of the stochastic filtering problem, to depend continuously on the observed data $Y = \{Y_s, s \in [0,t]\}$. Indeed, if the signal and the observation noise are independent one can show that, for any suitably chosen test function f, there exists a continuous map θ_t^f, defined on the space of continuous paths $C([0,t],\mathbb{R}^d)$ endowed with the uniform convergence topology such that $\pi_t(f) = \theta_t^f(Y)$, almost surely; see, for example, Clark [In *Communication Systems and Random Process Theory* (*Proc. 2nd NATO Advanced Study Inst.*, *Darlington*, 1977) (1978) 721–734, Sijthoff & Noordhoff], Clark and Crisan [*Probab. Theory Related Fields* **133** (2005) 43–56], Davis [*Z. Wahrsch. Verw. Gebiete* **54** (1980) 125–139], Davis [*Teor. Veroyatn. Primen.* **27** (1982) 160–167], Kushner [*Stochastics* **3** (1979) 75–83]. As shown by Davis and Spathopoulos [*SIAM J. Control Optim.* **25** (1987) 260–278], Davis [In *Stochastic Systems*: *The Mathematics of Filtering and Identification and Applications*, *Proc. NATO Adv. Study Inst. Les Arcs, Savoie, France 1980* 505–528], [In *The Oxford Handbook of Nonlinear Filtering* (2011) 403–424 Oxford Univ. Press], this type of *robust* representation is also possible when the signal and the observation noise are correlated, provided the observation process is scalar. For a general correlated noise and multidimensional observations such a representation does not exist. By using the theory of rough paths we provide a solution to this deficiency: the observation process Y is "lifted" to the process $\mathbf{Y}$ that consists of Y and its corresponding Lévy area process, and we show that there exists a continuous map θ_t^f, defined on a suitably chosen space of Hölder continuous paths such that $\pi_t(f) = \theta_t^f(\mathbf{Y})$, almost surely.

Received January 2012; revised August 2012.

[1]Supported in part by the EPSRC Grant EP/H0005500/1.

[2]Supported by DFG Grant SPP-1324.

[3]Supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement 258237.

MSC2010 subject classifications. 60G35, 93E11.

Key words and phrases. Filtering, robustness, rough path theory.

1. Introduction. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space on which we have defined a two-component diffusion process (X, Y) solving a stochastic differential equation driven by a multidimensional Brownian motion. One assumes that the first component X is unobservable, and the second component Y is observed. The filtering problem consists of computing the conditional distribution of the unobserved component, called the *signal* process, given the *observation* process Y. Equivalently, one is interested in computing

$$\pi_t(f) = \mathbb{E}[f(X_t, Y_t)|\mathcal{Y}_t],$$

where $\mathcal{Y} = \{\mathcal{Y}_t, t \geq 0\}$ is the observation filtration, and f is a suitably chosen test function. An elementary measure theoretic result tells us[4] that there exists a Borel-measurable map $\theta_t^f : C([0, t], \mathbb{R}^{d_Y}) \to \mathbb{R}$, such that

$$\pi_t(f) = \theta_t^f(Y_\cdot), \qquad \mathbb{P}\text{-a.s.,} \tag{1}$$

where d_Y is the dimension of the observation state space, and $Y_\cdot$ is the path-valued random variable

$$Y_\cdot : \Omega \to C\big([0, t], \mathbb{R}^{d_Y}\big), \qquad Y_\cdot(\omega) = \big(Y_s(\omega), 0 \leq s \leq t\big).$$

Of course, θ_t^f is not unique. Any other function $\bar{\theta}_t^f$ such that

$$\mathbb{P} \circ Y_\cdot^{-1}(\bar{\theta}_t^f \neq \theta_t^f) = 0,$$

where $\mathbb{P} \circ Y_\cdot^{-1}$ is the distribution of $Y_\cdot$ on the path space $C([0, t], \mathbb{R}^{d_Y})$ can replace θ_t^f in (1). It would be desirable to solve this ambiguity by choosing a suitable representative from the class of functions that satisfy (1). A *continuous* version, if it exists, would enjoy the following uniqueness property: if the law of the observation $\mathbb{P} \circ Y_\cdot^{-1}$ positively charges all nonempty open sets in $C([0, t], \mathbb{R}^{d_Y})$, then there exists a unique continuous function θ_t^f that satisfies (1). In this case, we call $\theta_t^f(Y_\cdot)$ the *robust version* of $\pi_t(f)$ and equation (1) is the robust representation formula for the solution of the stochastic filtering problem.

The need for this type of representation arises when the filtering framework is used to model and solve "real-life" problems. As explained in a substantial number of papers (e.g., [7, 8, 10–14, 26]) the model chosen for the "real life" observation process $\bar{Y}$ may not be a perfect one. However, if θ^f is continuous (or even locally Lipschitz, as in the setting of [8]), and as long as the distribution of $\bar{Y}_\cdot$ is close in a weak sense to that of $Y_\cdot$ (and some integrability assumptions hold), the estimate $\theta_t^f(\bar{Y}_\cdot)$ computed on the actual observation will still be reasonable, as $\mathbb{E}[(f(X_t, Y_t) - \theta_t^f(\bar{Y}_\cdot))^2]$ is close to the idealized error $\mathbb{E}[(f(X_t, Y_t) - \theta_t^f(Y_\cdot))^2]$.

[4] See, for example, Proposition 4.9, page 69, in [5].

Moreover, even when Y and $\bar{Y}$ actually coincide, one is never able to obtain and exploit a continuous stream of data as modeled by the continuous path $Y_{\cdot}(\omega)$. Instead the observation arrives and is processed at discrete moments in time

$$0 = t_0 < t_1 < t_2 < \cdots < t_n = t.$$

However, the continuous path $\hat{Y}_{\cdot}(\omega)$ obtained from the discrete observations $(Y_{t_i}(\omega))_{i=1}^n$ by linear interpolation is close to $Y_{\cdot}(\omega)$ (with respect to the supremum norm on $C([0,t],\mathbb{R}^{d_Y})$); hence, by the same argument, $\theta_t^f(\hat{Y}_{\cdot})$ will be a sensible approximation to $\pi_t(f)$. To conclude the discussion on the un-correlated framework, let us also mention that Kushner introduces in [27] a robust computable approximation for the filtering solution.

In the following, we will assume that the pair of processes (X,Y) satisfy the equation

$$dX_t = l_0(X_t, Y_t)\,dt + \sum_k Z_k(X_t, Y_t)\,dW_t^k + \sum_j L_j(X_t, Y_t)\,dB_t^j, \tag{2}$$

$$dY_t = h(X_t, Y_t)\,dt + dW_t \tag{3}$$

with X_0 being a bounded random variable and $Y_0 = 0$. In (2) and (3), the process X is the d_X-dimensional signal, Y is the d_Y-dimensional observation, B and W are independent d_B-dimensional, respectively, d_Y-dimensional Brownian motions independent of X_0. Suitable assumptions on the coefficients $l_0, L_1, \ldots, L_{d_B} : \mathbb{R}^{d_X+d_Y} \to \mathbb{R}^{d_X}$, $Z_1, \ldots, Z_{d_Y} : \mathbb{R}^{d_X+d_Y} \to \mathbb{R}^{d_X}$ and $h = (h^1, \ldots, h^{d_Y}) : \mathbb{R}^{d_X+d_Y} \to \mathbb{R}^{d_Y}$ will be introduced later on. This framework covers a wide variety of applications of stochastic filtering (see, e.g., [9] and the references therein) and has the added advantage that, within it, $\pi_t(f)$ admits an alternative representation that is crucial for the construction of its robust version. Let us detail this representation first.

Let $u = \{u_t, t > 0\}$ be the process defined by

$$u_t = \exp\left[-\sum_{i=1}^{d_Y}\left(\int_0^t h^i(X_s, Y_s)\,dW_s^i - \frac{1}{2}\int_0^t \left(h^i(X_s, Y_s)\right)^2 ds\right)\right]. \tag{4}$$

Then, under suitable assumptions,[5] u is a martingale which is used to construct the probability measure $\mathbb{P}_0$ equivalent to $\mathbb{P}$ on $\bigcup_{0\le t<\infty}\mathcal{F}_t$ whose Radon–Nikodym derivative with respect to $\mathbb{P}$ is given by u, namely,

$$\left.\frac{d\mathbb{P}_0}{d\mathbb{P}}\right|_{\mathcal{F}_t} = u_t.$$

[5] For example, if Novikov's condition is satisfied, that is, if $\mathbb{E}[\exp(\frac{1}{2}\int_0^t \|h^i(X_s, Y_s)\|^2\,ds)] < \infty$ for all $t > 0$, then u is a martingale. In particular it will be satisfied in our setting, in which h is bounded.

Under $\mathbb{P}_0$, Y is a Brownian motion independent of B. Moreover the equation for the signal process X becomes

$$dX_t = \bar{l}_0(X_t, Y_t)\,dt + \sum_k Z_k(X_t, Y_t)\,dY_t^k + \sum_j L_j(X_t, Y_t)\,dB_t^j. \tag{5}$$

Observe that equation (5) is now written in terms of the pair of Brownian motions (Y, B) and the coefficient $\bar{l}_0$ is given by $\bar{l}_0 = l_0 + \sum_k Z_k h_k$. Moreover, for any measurable, bounded function $f : \mathbb{R}^{d_X + d_Y} \to \mathbb{R}$, we have the following formula, called the Kallianpur–Striebel formula:

$$\pi_t(f) = \frac{p_t(f)}{p_t(1)}, \qquad p_t(f) := \mathbb{E}_0\big[f(X_t, Y_t)v_t \,|\, \mathcal{Y}_t\big], \tag{6}$$

where $v = \{v_t, t > 0\}$ is the process defined as $v_t := \exp(I_t)$, $t \geq 0$ and

$$I_t := \sum_{i=1}^{d_Y} \left(\int_0^t h^i(X_r, Y_r)\,dY_r^i - \frac{1}{2}\int_0^t \big(h^i(X_r, Y_r)\big)^2\,dr \right), \qquad t \geq 0. \tag{7}$$

The representation (6) suggests the following three-step methodology to construct a robust representation formula for π_t^f:

Step 1. We construct the triplet of processes (X^y, Y^y, I^y)[6] corresponding to the pair (y, B) where y is now a *fixed* observation path $y_\cdot = \{y_s, s \in [0, t]\}$ belonging to a suitable class of continuous functions and prove that the random variable $f(X^y, Y^y)\exp(I^y)$ is $\mathbb{P}_0$-integrable.

Step 2. We prove that the function $y_\cdot \to g_t^f(y_\cdot)$ defined as

$$g_t^f(y_\cdot) = \mathbb{E}_0\big[f\big(X_t^y, Y_t^y\big)\exp\big(I_t^y\big)\big] \tag{8}$$

is continuous.

Step 3. We prove that $g_t^f(Y_\cdot)$ is a version of $p_t(f)$. Then, following (6), the function, $y_\cdot \to \theta_t^f(y_\cdot)$ defined as

$$\theta_t^f = \frac{g_t^f}{g_t^1} \tag{9}$$

provides the robust version of of $\pi_t(f)$.

We emphasize that step 3 cannot be omitted from the methodology. Indeed one has to prove that $g_t^f(Y_\cdot)$ is a version of $p_t(f)$ as this fact is not immediate from the definition of g_t^f.

---

[6] As we shall see momentarily, in the uncorrelated case the choice of Y^y will trivially be y. In the correlated case we make it part of the *SDE with rough drift*, for (notational) convenience.

Step 1 is immediate in the particular case when only the Brownian motion B drives X (i.e., the coefficient $Z = 0$) and X is itself a diffusion, that is, it satisfies an equation of the form

$$dX_t = l_0(X_t)\,dt + \sum_j L_j(X_t)\,dB_t^j, \tag{10}$$

and h does only depend on X. In this case the process (X^y, Y^y) can be taken to be the pair (X, y). Moreover, we can define I^y by the formula

$$I_t^y := \sum_{i=1}^{d_Y}\biggl(h^i(X_t)y_t^i - \int_0^t y_r^i\,dh^i(X_r) - \frac{1}{2}\int_0^t \bigl(h^i(X_r, Y_r)\bigr)^2\,dr\biggr), \qquad t \geq 0, \tag{11}$$

provided the processes $h^i(X)$ are semi-martingales. In (11), the integral $\int_0^t y_r^i\,dh^i(X_r)$ is the Itô integral of the nonrandom process y^i with respect to $h^i(X)$. Note that the formula for I_t^y is obtained by applying integration by parts to the stochastic integral in (7)

$$\int_0^t h^i(X_r)\,dY_r^i = h^i(X_t)Y_r^i - \int_0^t Y_r^i\,dh^i(X_r), \tag{12}$$

and replacing the process Y by the fixed path y in (12). This approach has been successfully used to study the robustness property for the filtering problem for the above case in a number of papers [7, 8, 26].

The construction of the process (X^y, Y^y, I^y) is no longer immediate in the case when $Z \neq 0$, that is, when the signal is driven by both B and W (the correlated noise case). In the case when the observation is one-dimensional, one can solve this problem by using a method akin with the Doss–Sussmann "pathwise solution" of a stochastic differential equation; see [20, 32]. This approach has been employed by Davis to extend the robustness result to the correlated noise case with scalar observation; see [10, 12–14]. In this case one constructs first a diffeomorphism which is a pathwise solution of the equation[7]

$$\phi(t, x) = x + \int_0^t Z\bigl(\phi(s, x)\bigr) \circ dY_t. \tag{13}$$

The diffeomorphism is used to express the solution X of equation (5) as a composition between the diffeomorphism ϕ and the solution of a stochastic differential equation driven by B only and whose coefficients depend continously on Y. As a result, we can make sense of X^y. I^y is then defined by a suitable (formal) integration by parts that produces a pathwise interpretation of the stochastic integral appearing in (7), and Y^y is chosen to be y, as before. The robust representation formula is then introduced as per (9). Additional results for the correlated noise case with scalar observation can be found in [22]. The extension of the robustness

[7]Here $d^Y = 1$ and Y is scalar.

result to special cases of the correlated noise and multidimensional observation has been tackled in several works. Robustness results in the correlated setting have been obtained by Davis in [10, 13] and Elliott and Kohlmann in [21], under a commutativity condition on the signal vector fields. Florchinger and Nappo [23] do not have correlated noise, but allow the coefficients to depend on the signal and the observation.[8] To sum up, all previous works on the robust representation problem either treat the uncorrelated case, the case with one-dimensional observation or the case where the Lie brackets of the relevant vector fields vanish. In parallel, Bagchi and Karandikar treat in [1] a different model with “finitely additive” state white noise and “finitely additive” observation noise. Robustness there is virtually built into the problem.

An alternative framework is that where the signal and the observation run in discrete time. In this case the filtering problem is well understood and has been studied in many works, including the monograph [15] and the articles [16–18]. These works include an analysis of discrete time filtering problems and their approximation models, including particle approximation, approximate Bayesian computation, filtering models, etc. We note that in this context the continuity of the filter with respect to the observation data holds true[9] provided very natural conditions are imposed on the model: for example, the likelihood functions are assumed to be continuous and bounded (which includes the Gaussian case).

To our knowledge, the general correlated noise and multidimensional observation case has not been studied, and it is the subject of the current work. In this case it turns out that we cannot hope to have robustness in the sense advocated by Clark. More precisely, there may not exist a map continuous map $\theta_t^f : C([0,t],\mathbb{R}^{d_Y}) \to \mathbb{R}$, such that the representation (1) holds almost surely. The following is a simple example that illustrates this.

EXAMPLE 1. Consider the filtering problem where the signal and the observation process solve the following pair of equations:

$$X_t = X_0 + \int_0^t X_r \, d\left[Y_r^1 + Y_r^2\right] + \int_0^t X_r \, dr,$$
$$Y_t = \int_0^t h(X_r)\, dr + W_t,$$

where Y is two-dimensional and $\mathbb{P}(X_0 = 0) = \mathbb{P}(X_0 = 1) = \frac{1}{2}$. Then with f, h such that $f(0) = h_1(0) = h_2(0) = 0$ one can explicitly compute

$$\mathbb{E}\left[f(X_t)|\mathcal{Y}_t\right] = \frac{f(\exp(Y_t^1 + Y_t^2))}{1 + \exp(-\sum_{k=1,2}\int_0^t h^k(\exp(Y_r))\, dY_r^k + \int_0^t \|h(\exp(Y_r))\|^2 \, dr/2)}. \tag{14}$$

[8] We thank the anonymous referee for these references.

[9] which can be easily seen using the representation of Lemma 2.1 in [18].

Following the findings of rough path theory (see, e.g., [25, 28–30]) the expression on the right-hand side of (14) is not continuous in supremum norm (nor in any other metric on path space) because of the stochastic integral. Explicitly, this follows, for example, from Theorem 1.1.1 in [29] by rewriting the exponential term as the solution to a stochastic differential equation driven by Y.

Nevertheless, we can show that a variation of the robustness representation formula still exists in this case. For this we need to "enhance" the original process Y by adding a second component to it which consists of its iterated integrals (that, knowing the path, is in a one-to-one correspondance with the Lévy area process). Explicitly we consider the process $\mathbf{Y} = \{\mathbf{Y}_t, t \geq 0\}$ defined as

$$\mathbf{Y}_t = \left(Y_t, \begin{pmatrix} \int_0^t Y_r^1 \circ dY_r^1 & \cdots & \int_0^t Y_r^1 \circ dY_r^{d^Y} \\ \vdots & \vdots & \vdots \\ \int_0^t Y_r^{d^Y} \circ dY_r^1 & \cdots & \int_0^t Y_r^{d^Y} \circ dY_r^{d^Y} \end{pmatrix} \right), \qquad t \geq 0. \tag{15}$$

The stochastic integrals in (15) are Stratonovich integrals. The state space of $\mathbf{Y}$ is $G^2(\mathbb{R}^{d_Y}) \cong \mathbb{R}^{d_Y} \oplus \mathrm{so}(d_Y)$, where $\mathrm{so}(d_Y)$ is the set of anti-symmetric matrices of dimension d_Y.[10] Over this state space we consider not the space of continuous function, but a subspace $\mathcal{C}^{0,\alpha}$ that contains paths $\eta : [0,t] \to G^2(\mathbb{R}^{d_Y})$ that are α-Hölder in the $\mathbb{R}^{d_Y}$-component and somewhat "2α-Hölder" in the $\mathrm{so}(d_Y)$-component, where α is a suitably chosen constant $\alpha < 1/2$. Note that there exists a modification of $\mathbf{Y}$ such that $\mathbf{Y}(\omega) \in \mathcal{C}^{0,\alpha}$ for all ω (Corollary 13.14 in [25]).

The space $\mathcal{C}^{0,\alpha}$ is endowed with the α-Hölder rough path metric under which $\mathcal{C}^{0,\alpha}$ becomes a complete metric space. The *main result of the paper* (captured in Theorems 6 and 7) is that there exists a continuous map $\theta_t^f : \mathcal{C}^{0,\alpha} \to \mathbb{R}$, such that

$$\pi_t(f) = \theta_t^f(\mathbf{Y}_\cdot), \qquad \mathbb{P}\text{-a.s.} \tag{16}$$

Even though the map is defined on a slightly more abstract space, it nonetheless enjoys the desirable properties described above for the case of a continuous version on $C([0,t],\mathbb{R}^d)$. Since $\mathbb{P} \circ \mathbf{Y}^{-1}$ positively charges all nonempty open sets of $\mathcal{C}^{0,\alpha}$,[11] the continuous version we construct will be unique. Also, it provides a certain model robustness, in the sense that $\mathbb{E}[(f(X_t) - \theta_t^f(\bar{\mathbf{Y}}_\cdot))^2]$ is well approximated by the idealized error $\mathbb{E}[(f(X_t) - \theta_t^f(\mathbf{Y}_\cdot))^2]$, if $\bar{\mathbf{Y}}_\cdot$ is close in distribution to $\mathbf{Y}_\cdot$. The problem of discrete observation is a little more delicate. One one

[10] More generally, $G^{[1/\alpha]}(\mathbb{R}^d)$ is the "correct" state space for a geometric α-Hölder rough path; the space of such paths subject to α-Hölder regularity (in rough path sense) yields a complete metric space under α-Hölder rough path metric. Technical details of geometric rough path spaces (as found, e.g., in Section 9 of [25]) are not required for understanding the results of the present paper.

[11] This fact is a consequence of the support theorem of Brownian motion in Hölder rough path topology [24]; see also Chapter 13 in [25].

hand, it is true that the rough path lift $\hat{\mathbf{Y}}$ calculated from the linearly interpolated Brownian motion $\hat{Y}$ will converge to the true rough path $\mathbf{Y}$ in probability as the mesh goes to zero (Corollary 13.21 in [25]), which implies that $\theta_t^f(\hat{\mathbf{Y}})$ is close in probability to $\theta_t^f(\mathbf{Y})$ (we provide local Lipschitz estimates for θ^f). Actually, most sensible approximations will do, as is, for example, shown in Chapter 13 in [25] (although, contrary to the uncorrelated case, not all interpolations that converge in uniform topology will work; see, e.g., Theorem 13.24 ibid). But these are probabilistic statements, that somehow miss the pathwise stability that one wants to provide with θ_t^f. If, on the other hand, one is able to observe at discrete time points not only the process itself, but also its second level, that is, the area, one can construct an interpolating rough path using geodesics (see, e.g., Chapter 13.3.1 in [25]) which is close to the true (lifted) observation path $\mathbf{Y}$ in the relevant metric *for all realizations* $\mathbf{Y} \in \mathcal{C}^{0,\alpha}$.

The following is the outline of the paper: In the next section, we enumerate the common notation used throughout the paper. In Section 3 we introduce the notion of a stochastic differential equation with rough drift, which is necessary for our main result and correspond to step 1 above. We present it separately of the filtering problem, since we believe this notion to be of independent interest. The proof of the existence of a solution of a stochastic differential equation with rough drifts and its properties is postponed to Section 5. Section 4 contains the main results of the paper and the assumptions under which they hold true. Steps 2 and 3 of above mentioned methodology are carried out in Theorems 6 and 7.

2. Nomenclature. Lip^γ is the set of γ-Lipschitz[12] functions $a : \mathbb{R}^m \to \mathbb{R}^n$ where m and n are chosen according to the context.

$G^2(\mathbb{R}^{d_Y}) \cong \mathbb{R}^d \oplus \mathrm{so}(d_Y)$ is the state space for a d_Y-dimensional Brownian motion (or, in general for an arbitrary semi-martingale) and its corresponding Lévy area.

$\mathcal{C}^{0,\alpha} := C_0^{0,\alpha\text{-Höl}}([0,t], G^2(\mathbb{R}^{d_Y}))$ is the set of geometric α-Hölder rough paths $\eta : [0,t] \to G^2(\mathbb{R}^{d_Y})$ starting at 0. We shall use the nonhomogenous metric $\rho_{\alpha\text{-Höl}}$ on this space.

In the following we will make use of an auxiliary filtered probability space $(\bar{\Omega}, \bar{\mathcal{F}}, (\bar{F}_t)_{t\geq 0}, \bar{\mathbb{P}})$ carrying a d_B-dimensional Brownian motion $\bar{B}$.[13]

Let $\mathcal{S}^0 = \mathcal{S}^0(\bar{\Omega})$ denote the space of adapted, continuous processes in $\mathbb{R}^{d_S}$, with the topology of uniform convergence in probability.

[12] In the sense of E. Stein, that is, bounded kth dervative for $k = 0, \dots, \lfloor\gamma\rfloor$ and $\gamma - \lfloor\gamma\rfloor$-Hölder continuous $\lfloor\gamma\rfloor$th derivative.

[13] We introduce this auxiliary probability space, since in the proof of Theorem 7 it will be easier to work on a product space separating the randomness coming from Y and B. A similar approach was followed in the proof of Theorem 1 in [2].

For $q \geq 1$ we denote by $\mathcal{S}^q = \mathcal{S}^q(\bar{\Omega})$ the space of processes $X \in \mathcal{S}^0$ such that

$$\|X\|_{\mathcal{S}^q} := \Big(\bar{\mathbb{E}}\Big[\sup_{s \leq t} |X_t|^q\Big]\Big)^{1/q} < \infty.$$

3. SDE with rough drift. For the statement and proof of the main results we shall use the notion (and the properties) of an *SDE with rough drift* captured in the following theorems. The proofs are postponed to Section 5.

As defined above, let $(\bar{\Omega}, \bar{\mathcal{F}}, (\bar{F}_t)_{t \geq 0}, \bar{\mathbb{P}})$ be a filtered probability space carrying a d_B-dimensional Brownian motion $\bar{B}$ and a bounded d_S-dimensional random vector S_0 independent of $\bar{B}$. In the following, we fix $\epsilon \in (0, 1)$ and $\alpha \in (\frac{1}{2+\epsilon}, \frac{1}{2})$. Let $\eta^n : [0, t] \to \mathbb{R}^{d_Y}$ be smooth paths, such that $\eta^n \to \boldsymbol{\eta}$ in α-Hölder, for some $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$, and let S^n be a d_S-dimensional process which is the unique solution to the classical SDE

$$S^n_t = S_0 + \int_0^t a(S^n_r)\, dr + \int_0^t b(S^n_r)\, d\bar{B}_r + \int_0^t c(S^n_r)\, d\eta^n_r,$$

where we assume that[14]

(a1) $a \in \mathrm{Lip}^1(\mathbb{R}^{d_S})$, $b_1, \ldots, b_{d_B} \in \mathrm{Lip}^1(\mathbb{R}^{d_S})$ and $c_1, \ldots, c_{d_Y} \in \mathrm{Lip}^{4+\epsilon}(\mathbb{R}^{d_S})$;
(a1′) $a \in \mathrm{Lip}^1(\mathbb{R}^{d_S})$, $b_1, \ldots, b_{d_B} \in \mathrm{Lip}^1(\mathbb{R}^{d_S})$ and $c_1, \ldots, c_{d_Y} \in \mathrm{Lip}^{5+\epsilon}(\mathbb{R}^{d_S})$.

THEOREM 2. *Under assumption* (a1), *there exists a d_S-dimensional process $S^\infty \in \mathcal{S}^0$ such that*

$$S^n \to S^\infty \qquad \text{in } \mathcal{S}^0.$$

In addition, the limit $\Xi(\boldsymbol{\eta}) := S^\infty$ only depends on $\boldsymbol{\eta}$ and not on the approximating sequence.

Moreover, for all $q \geq 1$, $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$ it holds that $\Xi(\boldsymbol{\eta}) \in \mathcal{S}^q$ and the corresponding mapping $\Xi : \mathcal{C}^{0,\alpha} \to \mathcal{S}^q$ is locally uniformly continuous [*and locally Lipschitz under assumption* (a1′)].

Following Theorem 2, we say that $\Xi(\boldsymbol{\eta})$ is a solution of the *SDE with rough drift*

$$\Xi(\boldsymbol{\eta})_t = S_0 + \int_0^t a\big(\Xi(\boldsymbol{\eta})_r\big)\, dr + \int_0^t b\big(\Xi(\boldsymbol{\eta})_r\big)\, d\bar{B}_r + \int_0^t c\big(\Xi(\boldsymbol{\eta})_r\big)\, d\boldsymbol{\eta}_r. \tag{17}$$

The following result establishes some of the salient properties of solutions of SDEs with rough drift. Recall that $(\Omega, \mathcal{F}, \mathbb{P}_0)$ carries, as above, the d_Y-dimensional

[14] In the forthcoming publication [19] we show existence of solutions to SDEs with rough drift under additional sets of assumptions. The corresponding proofs do not rely on the technique of flow decomposition used in the present work, but require more elements of rough path theory and would lead us too far astray from the topic of filtering.

Brownian motion Y, and let $\hat{\Omega} = \Omega \times \bar{\Omega}$ be the product space, with product measure $\hat{\mathbb{P}} := \mathbb{P}_0 \otimes \bar{\mathbb{P}}$. Let S be the unique solution on this probability space to the SDE

$$S_t = S_0 + \int_0^t a(S_r)\,dr + \int_0^t b(S_r)\,d\bar{B}_r + \int_0^t c(S_r) \circ dY_r. \tag{18}$$

Denote by $\mathbf{Y}$ the rough path lift of Y (i.e., the enhanced Brownian Motion over Y).

THEOREM 3. *Under assumption* (a1) *we have that*:

- *For every* $R > 0, q \geq 1$

$$\sup_{\|\boldsymbol{\eta}\|_{\alpha\text{-Höl}} < R} \mathbb{E}\left[\exp\left(q \left|\Xi(\boldsymbol{\eta})\right|_{\infty;[0,t]}\right)\right] < \infty. \tag{19}$$

- *For* $\mathbb{P}_0$-*a.e.* ω

$$\bar{\mathbb{P}}\left[S_s(\omega, \cdot) = \Xi\left(\mathbf{Y}(\omega)\right)_s(\cdot), s \leq t\right] = 1. \tag{20}$$

4. Assumptions and main results. In the following we will make use of the Stratonovich version of equation (5); that is, we will consider that the signal satisfies the equation

$$\begin{aligned} X_t &= X_0 + \int_0^t L_0(X_r, Y_r)\,dr + \sum_k \int_0^t Z_k(X_r, Y_r) \circ dY_r^k \\ &\quad + \sum_j \int_0^t L_j(X_r, Y_r)\,dB_r^j, \\ Y_t &= \int_0^t h(X_r, Y_r)\,dr + W_t, \end{aligned} \tag{21}$$

where $L_0^j(x, y) = \bar{l}_0^j(x, y) - \frac{1}{2}\sum_k \sum_i \partial_{x_i} Z_k^j(x, y) Z_k^i(x, y) - \frac{1}{2}\sum_k \partial_{y_k} Z_k^j(x, y)$. We remind the reader that under $\mathbb{P}_0$ the observation Y is a Brownian motion independent of B.

We will assume that f is a bounded Lipschitz function, and we fix $\epsilon \in (0, 1)$ $\alpha \in (\frac{1}{2+\epsilon}, \frac{1}{2})$, $t > 0$, and X_0 is a bounded random vector independent of B and Y. We will use one of the following assumptions:

(A1) $Z_1, \ldots, Z_{d_Y} \in \mathrm{Lip}^{4+\epsilon}$, $h^1, \ldots, h^{d_Y} \in \mathrm{Lip}^{4+\epsilon}$ and $L_0, L_1, \ldots, L_{d_B} \in \mathrm{Lip}^1$;
(A1′) $Z_1, \ldots, Z_{d_Y} \in \mathrm{Lip}^{5+\epsilon}$, $h^1, \ldots, h^{d_Y} \in \mathrm{Lip}^{5+\epsilon}$ and $L_0, L_1, \ldots, L_{d_B} \in \mathrm{Lip}^1$.

REMARK 4. Assumption (A1) and (A1′) lead to the existence of a solution of an *SDEs with rough driver* (Theorem 2). Under (A1) the solution mapping is locally uniformly continuous, and under (A1′) it is locally Lipschitz (Theorem 3).

Assume either (A1) or (A1′). For $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$ there exists by Theorem 2 a solution $(X^{\boldsymbol{\eta}}, I^{\boldsymbol{\eta}})$ to the following *SDE with rough drift*:

$$\begin{aligned} X_t^{\boldsymbol{\eta}} &= X_0 + \int_0^t L_0(X_r^{\boldsymbol{\eta}}, Y_r^{\boldsymbol{\eta}})\,dr + \int_0^t Z(X_r^{\boldsymbol{\eta}}, Y_r^{\boldsymbol{\eta}})\,d\boldsymbol{\eta}_r \\ &\quad + \sum_j \int_0^t L_j(X_r^{\boldsymbol{\eta}}, Y_r^{\boldsymbol{\eta}})\,d\bar{B}_r^j, \\ Y_t^{\boldsymbol{\eta}} &= \int_0^t d\boldsymbol{\eta}_r, \\ I_t^{\boldsymbol{\eta}} &= \int_0^t h(X_r^{\boldsymbol{\eta}}, Y_r^{\boldsymbol{\eta}})\,d\boldsymbol{\eta}_r - \frac{1}{2}\sum_k \int_0^t D_k h^k(X_r^{\boldsymbol{\eta}}, Y_r^{\boldsymbol{\eta}})\,dr. \end{aligned} \tag{22}$$

REMARK 5. Note that formally (!) when replacing the rough path $\boldsymbol{\eta}$ with the process Y, $X^{\boldsymbol{\eta}}, Y^{\boldsymbol{\eta}}$ yields the solution to the SDE (21) and $\exp(I_t^{\boldsymbol{\eta}})$ yields the (Girsanov) multiplicator in (6). This observation is made precise in the statement of Theorem 2.

We introduce the functions $g^f, g^1, \theta : \mathcal{C}^{0,\alpha} \to \mathbb{R}$ defined as

$$g^f(\boldsymbol{\eta}) := \bar{\mathbb{E}}[f(X_t^{\boldsymbol{\eta}}, Y_t^{\boldsymbol{\eta}})\exp(I_t^{\boldsymbol{\eta}})], \qquad g^1(\boldsymbol{\eta}) := \bar{\mathbb{E}}[\exp(I_t^{\boldsymbol{\eta}})],$$

$$\theta(\boldsymbol{\eta}) := \frac{g^f(\boldsymbol{\eta})}{g^1(\boldsymbol{\eta})}, \qquad \boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}.$$

THEOREM 6. *Assume that* (A1) *holds*; *then θ is locally uniformly continuous*. *Moreover if* (A1′) *holds*, *then θ is locally Lipschitz*.

PROOF. From Theorem 2 we know that for $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$ the SDE with rough drift (22) has a unique solution $(X^{\boldsymbol{\eta}}, Y^{\boldsymbol{\eta}}, I^{\boldsymbol{\eta}})$ belonging to $\mathcal{S}^2$.

Let now $\boldsymbol{\eta}, \boldsymbol{\eta}' \in \mathcal{C}^{0,\alpha}$. Denote $X = X^{\boldsymbol{\eta}}$, $Y = Y^{\boldsymbol{\eta}}$, $I = I^{\boldsymbol{\eta}}$ and analogously for $\boldsymbol{\eta}'$. Then

$$\begin{aligned} &|g^f(\boldsymbol{\eta}) - g^f(\boldsymbol{\eta})| \\ &\quad \le \mathbb{E}[|f(X_t, Y_t)\exp(I_t) - f(X_t', Y"_t)\exp(I_t')|] \\ &\quad \le \mathbb{E}[|f(X_t, Y_t)||\exp(I_t) - \exp(I_t')|] \\ &\qquad + \mathbb{E}[|f(X_t, Y_t) - f(X_t', Y_t')|\exp(I_t')] \\ &\quad \le |f|_\infty \mathbb{E}[|\exp(I_t) - \exp(I_t')|] \\ &\qquad + \mathbb{E}[|f(X_t, Y_t) - f(X_t', Y_t')|^2]^{1/2}\mathbb{E}[|\exp(I_t')|^2]^{1/2} \\ &\quad \le |f|_\infty \mathbb{E}[|\exp(I_t) + \exp(I_t')|^2]^{1/2}\mathbb{E}[|I_t - I_t'|]^{1/2} \\ &\qquad + \mathbb{E}[|f(X_t, Y_t) - f(X_t', Y_t')|^2]^{1/2}\mathbb{E}[|\exp(I_t')|^2]^{1/2}. \end{aligned}$$

Hence, using from Theorems 2 and 3 the continuity statements as well as the boundedness of exponential moments, we see that g^f is locally uniformly continuous under (A1), and it is locally Lipschitz under (A1′).

The same then holds true for g^1 and moreover $g^1(\boldsymbol{\eta}) > 0$. Hence θ is locally uniformly continuous under (A1) and locally Lipschitz under (A1′). $\square$

Denote by $\mathbf{Y}_\cdot$, as before, the canonical rough path lift of Y to $\mathcal{C}^{0,\alpha}$. We then have

THEOREM 7. *Assume either* (A1) *or* (A1′). *Then* $\theta(\mathbf{Y}_\cdot) = \pi_t(f)$, $\mathbb{P}$-*a.s.*

PROOF. To prove the statement it is enough to show that

$$g^f(\mathbf{Y}_\cdot) = p_t(f), \qquad \mathbb{P}\text{-a.s.},$$

which is equivalent to

$$g^f(\mathbf{Y}_\cdot) = p_t(f), \qquad \mathbb{P}_0\text{-a.s.}$$

For that, it suffices to show that

$$\mathbb{E}_0\big[p_t(f)\Upsilon(Y_\cdot)\big] = \mathbb{E}_0\big[g^f(\mathbf{Y}_\cdot)\Upsilon(Y_\cdot)\big] \tag{23}$$

for an arbitrary continuous bounded function $\Upsilon : C([0,t], \mathbb{R}^{d_Y}) \to \mathbb{R}$.

Let $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ be the auxiliary probability space from before, carrying an d_B-dimensional Brownian motion $\bar{B}$. Let $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}) := (\Omega \times \bar{\Omega}, \mathcal{F} \otimes \bar{\mathcal{F}}, \mathbb{P}_0 \otimes \bar{\mathbb{P}})$. By Y and X_0 we denote also the "lift" of Y to $\hat{\Omega}$, that is, $Y(\omega, \bar{\omega}) = Y(\omega)$, $X_0(\omega, \bar{\omega}) = X_0(\omega)$. Then (Y, B) (on Ω under $\mathbb{P}_0$) has the same distribution as $(Y, \bar{B})$ (on $\hat{\Omega}$ under $\hat{\mathbb{P}}$).

Denote by $(\hat{X}, \hat{I})$ the solution on $(\hat{\Omega}, \hat{F}, \hat{\mathbb{P}})$ to the SDE

$$\begin{aligned}\hat{X}_t &= X_0 + \int_0^t L_0(\hat{X}_r, Y_r)\,dr + \sum_k \int_0^t Z_k(\hat{X}_r, Y_r) \circ dY_r^k \\ &\quad + \sum_j \int_0^t L_j(\hat{X}_r, Y_r)\,d\bar{B}_r^j, \\ \hat{I}_t &= \sum_k \int_0^t h^k(\hat{X}_r, Y_r) \circ dY_r^k - \frac{1}{2}\sum_k \int_0^t D_k h^k(\hat{X}_r, Y_r)\,dr.\end{aligned}$$

Then

$$(Y, \hat{X}, \hat{I})_{\hat{\mathbb{P}}} \sim \left(Y, X, \sum_k \int_0^\cdot h^k(X_r, Y_r) \circ dY_r^k - \frac{1}{2}\sum_k \int_0^\cdot D_k h^k(X_r, Y_r)\,dr\right)_{\mathbb{P}_0}.$$

Hence, for the left-hand side of (23),

$$\begin{aligned}
&\mathbb{E}_0\big[p_t(f)\Upsilon(Y_\cdot)\big] \\
&\quad = \mathbb{E}_0\bigg[f(X_t, Y_t)\exp\bigg(\sum_k \int_0^t h^k(X_r, Y_r)\circ dY_r^k \\
&\qquad\qquad - \frac{1}{2}\sum_k \int_0^t D_k h^k(X_r, Y_r)\,dr\bigg)\Upsilon(Y_\cdot)\bigg] \\
&\quad = \hat{\mathbb{E}}\big[f(\hat{X}_t, Y_t)\exp(\hat{I}_t)\Upsilon(Y_\cdot)\big].
\end{aligned}$$

On the other hand, from Theorem 3 we know that for $\mathbb{P}_0$-a.e. ω

$$\begin{aligned}
&X^{\mathbf{Y}_\cdot(\omega)}(\bar{\omega})_t = \hat{X}_t(\omega, \bar{\omega}), \qquad Y^{\mathbf{Y}_\cdot(\omega)}(\bar{\omega})_t = \hat{Y}_t(\omega, \bar{\omega}), \\
&I^{\mathbf{Y}_\cdot(\omega)}(\bar{\omega})_t = \hat{I}_t(\omega, \bar{\omega}), \qquad \bar{\mathbb{P}}\text{-a.e. } \bar{\omega}.
\end{aligned}$$

Hence, for the right-hand side of (23) we get (using Fubini for the last equality)

$$\begin{aligned}
\mathbb{E}_0\big[g^f(\mathbf{Y}_\cdot)\Upsilon(Y_\cdot)\big] &= \mathbb{E}_0\big[\bar{\mathbb{E}}\big[f\big(X_t^{\mathbf{Y}_\cdot}, Y_t^{\mathbf{Y}_\cdot}\big)\exp\big(I_t^{\mathbf{Y}_\cdot}\big)\big]\Upsilon(Y_\cdot)\big] \\
&= \mathbb{E}_0\big[\bar{\mathbb{E}}\big[f(\hat{X}_t, Y_t)\exp(\hat{I}_t)\big]\Upsilon(Y_\cdot)\big] \\
&= \hat{\mathbb{E}}\big[f(\hat{X}_t, Y_t)\exp(\hat{I}_t)\Upsilon(Y_\cdot)\big],
\end{aligned}$$

which yields (23). $\square$

5. Proofs of Theorems 2 and 3.

PROOF OF THEOREM 2. Let $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$ be the lift of a smooth path η. Let $S^{\boldsymbol{\eta}}$ be the unique solution of the SDE

$$S_t^{\boldsymbol{\eta}} = S_0 + \int_0^t a\big(S_r^{\boldsymbol{\eta}}\big)\,dr + \int_0^t b\big(S_r^{\boldsymbol{\eta}}\big)\,d\bar{B}_r + \int_0^t c\big(S_r^{\boldsymbol{\eta}}\big)\,d\eta_r.$$

Define $\tilde{S}^{\boldsymbol{\eta}} := (\phi^{\boldsymbol{\eta}})^{-1}(t, S_t^{\boldsymbol{\eta}})$, where $\phi^{\boldsymbol{\eta}}$ is the ODE flow

$$\phi^{\boldsymbol{\eta}}(t,x) = x + \int_0^t c\big(\phi^{\boldsymbol{\eta}}(r,x)\big)\,d\boldsymbol{\eta}_r. \tag{24}$$

By Lemma 9, we have that $\tilde{S}^{\boldsymbol{\eta}}$ satisfies the SDE

$$\tilde{S}_t^{\boldsymbol{\eta}} = S_0 + \int_0^t \tilde{a}^{\boldsymbol{\eta}}\big(r, \tilde{S}_r^{\boldsymbol{\eta}}\big)\,dr + \int_0^t \tilde{b}^{\boldsymbol{\eta}}\big(r, \tilde{S}_r^{\boldsymbol{\eta}}\big)\,d\bar{B}_r \tag{25}$$

with $\tilde{a}^{\boldsymbol{\eta}}, \tilde{b}^{\boldsymbol{\eta}}$ defined as in Lemma 9.

This equation makes sense, even if $\boldsymbol{\eta}$ is a generic rough path in $\mathcal{C}^{0,\alpha}$ [in which case (24) is now really an RDE]. Indeed, since the first two derivatives of $\phi^{\boldsymbol{\eta}}$ and its inverse are bounded (Proposition 11.11 in [25]) we have that $\tilde{a}^{\boldsymbol{\eta}}(t,\cdot), \tilde{b}^{\boldsymbol{\eta}}(t,\cdot)$ are also in Lip^1. Hence by Theorem V.7 in [31], there exists a unique strong solution to (25).

We define the mapping introduced in Theorem 2 as

$$\Xi(\boldsymbol{\eta})_t := \phi^{\boldsymbol{\eta}}(t, \tilde{S}_t^{\boldsymbol{\eta}}).$$

To show continuity of the mapping we restrict ourselves to the case $q = 2$. Moreover we shall assume $c_1, \dots, c_{d_Y} \in \mathrm{Lip}^{5+\epsilon}(\mathbb{R}^{d_S})$, and we will hence prove the local Lipschitz property of the respective maps.

Let $\boldsymbol{\eta}^1, \boldsymbol{\eta}^2 \in \mathcal{C}^{0,\alpha}$ with $|\boldsymbol{\eta}^1|_{\alpha\text{-Höl}}, |\boldsymbol{\eta}^2|_{\alpha\text{-Höl}} < R$. By Lemma 12 we have

$$\bar{\mathbb{E}}\Big[\sup_{s \le t} |\tilde{S}_s^1 - \tilde{S}_s^2|^2\Big]^{1/2} \le C_{\mathrm{Lem}\ 12}(R)\rho_{\alpha\text{-Höl}}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2).$$

Hence

$$\begin{aligned}
&\bar{\mathbb{E}}\Big[\sup_{s \le t} |\Xi(\boldsymbol{\eta}^1)_s - \Xi(\boldsymbol{\eta}^2)_s|^2\Big]^{1/2} \\
&\quad = \bar{\mathbb{E}}\Big[\sup_{s \le t} |\phi^1(s, \tilde{S}_s^1) - \phi^2(s, \tilde{S}_s^2)|^2\Big]^{1/2} \\
&\quad \le \bar{\mathbb{E}}\Big[\sup_{s \le t} |\phi^1(s, \tilde{S}_s^1) - \phi^1(s, \tilde{S}_s^2)|^2\Big]^{1/2} + \bar{\mathbb{E}}\Big[\sup_{s \le t} |\phi^1(s, \tilde{S}_s^2) - \phi^2(s, \tilde{S}_s^2)|^2\Big]^{1/2} \\
&\quad \le \bar{\mathbb{E}}\Big[\sup_{s \le t} |\phi^1(s, \tilde{S}_s^1) - \phi^1(s, \tilde{S}_s^2)|^2\Big]^{1/2} + \sup_{s \le t, x \in \mathbb{R}^{d_S}} |\phi^1(s, x) - \phi^1(s, x)| \\
&\quad \le K(R)\bar{\mathbb{E}}\Big[\sup_{s \le t} |\tilde{S}_s^1 - \tilde{S}_s^2|^2\Big]^{1/2} + C_{\mathrm{Lem}\ 13}(R)\rho_{\alpha\text{-Höl}}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2) \\
&\quad \le C_1 \rho_{\alpha\text{-Höl}}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2)
\end{aligned}$$

as desired, where

$$C_1 = K_{\mathrm{Lem}\ 13}(R) C_{\mathrm{Lem}\ 12}(R) + C_{\mathrm{Lem}\ 13}(R),$$

where $K_{\mathrm{Lem}\ 13}(R)$ and $C_{\mathrm{Lem}\ 13}(R)$ are the constants from Lemma 13. □

PROOF OF THEOREM 3. In order to show (19), pick $k \in 1, \dots, d_S$. We first note that by simply scaling (the coefficients of) $S^{\boldsymbol{\eta}}$ it is sufficient to argue for $q = 1$. And consider the kth component of Ξ.

Then

$$\begin{aligned}
&\mathbb{E}[\exp(|\Xi^{(k)}(\boldsymbol{\eta})|_{\infty;[0,t]})] \\
&\quad \le \mathbb{E}[\exp(|D\psi^{\boldsymbol{\eta}}|_\infty(|\phi^{\boldsymbol{\eta}}(0, S_0)| + |\tilde{S}^{(k);\boldsymbol{\eta}}|_{\infty;[0,t]}))] \\
&\quad \le \exp\Big(|D\psi^{\boldsymbol{\eta}}|_\infty \sup_{|x| \le |S_0|_{L^\infty}} |\phi^{\boldsymbol{\eta}}(0, x)|\Big)\mathbb{E}\Big[\exp\Big(|D\psi^{\boldsymbol{\eta}}|_\infty \sup_{s \le t} |\tilde{S}_t^{(k);\boldsymbol{\eta}}|\Big)\Big]
\end{aligned}$$

$$
\begin{aligned}
&\le \exp\Bigl(|D\psi^{\eta}|_{\infty} \sup_{|x|\le|S_0|_{L^\infty}} |\phi^{\eta}(0,x)|\Bigr) \\
&\quad \times \Bigl(\mathbb{E}\Bigl[\exp\Bigl(|D\psi^{\eta}|_{\infty} \sup_{s\le t} \tilde{S}_s^{(k);\eta}\Bigr)\Bigr] + \mathbb{E}\Bigl[-\exp\Bigl(|D\psi^{\eta}|_{\infty} \sup_{s\le t} \tilde{S}_s^{(k);\eta}\Bigr)\Bigr]\Bigr) \\
&= \exp\Bigl(|D\psi^{\eta}|_{\infty} \sup_{|x|\le|S_0|_{L^\infty}} |\phi^{\eta}(0,x)|\Bigr) \\
&\quad \times \Bigl(\mathbb{E}\Bigl[\sup_{s\le t} \exp(|D\psi^{\eta}|_{\infty} \tilde{S}_s^{(k);\eta})\Bigr] + \mathbb{E}\Bigl[-\sup_{s\le t} \exp(|D\psi^{\eta}|_{\infty} \tilde{S}_s^{(k);\eta})\Bigr]\Bigr).
\end{aligned}
$$

Now, only the boundedness of the last two terms remains to be shown, for η bounded.

By applying Itô's formula we get that

$$
\begin{aligned}
\exp(\tilde{S}_t^{(k);\eta}) &= 1 + \int_0^t \exp(\tilde{S}_r^{(k);\eta})\, d\tilde{S}_r^{(k);\eta} + \int_0^t \exp(\tilde{S}_r^{(k);\eta})\, d\langle \tilde{S}^{(k);\eta}\rangle_r \\
&= 1 + \int_0^t \exp(\tilde{S}_r^{(k);\eta}) \tilde{a}_k^{\eta}(\tilde{S}_r^{(k);\eta})\, dr + \sum_{i=1}^{d_B} \int_0^t \exp(\tilde{S}_r^{(k);\eta}) \tilde{b}_{ki}^{\eta}(\tilde{S}_r^{(k);\eta})\, d\bar{B}_r^i \\
&\quad + \sum_{i=1}^{d_B} \int_0^t \exp(\tilde{S}_r^{(k);\eta}) |\tilde{b}_{ki}^{\eta}(\tilde{S}_r^{(k);\eta})|^2\, dr.
\end{aligned}
$$

Hence the process $\exp(\tilde{S}^{(k);\eta})$ satisfies an SDE with Lipschitz coefficients and by an application of Gronwalls lemma and the Burkholder–Davis–Gundy inequality (see also Lemma V.2 in [31]) one arrives at

$$
\sup_{|\eta|_{\alpha\text{-Höl}}<R} \sup_{s\le t} \bar{\mathbb{E}}[\exp(|D\psi^{\eta}|_{\infty} \tilde{S}_t^{(k);\eta})] \le C \exp(C_2), \tag{26}
$$

where C is universal and

$$
C_2 := \sup_{\eta:\|\eta\|_{\alpha\text{-Höl}}<R} |\tilde{a}^{\eta}|_{\infty} + |\tilde{b}^{\eta}|_{\infty}^2,
$$

which is finite because of Lemma 10. One argues analogously for

$$
\sup_{s\le t} \bar{\mathbb{E}}[-\exp(|D\psi^{\eta}|_{\infty} \tilde{S}_t^{(k);\eta})],
$$

which then gives (19).

Now, for the correspondence to an SDE solution let Ω be the additional probability space as given in the statement. Let S be the solution to the SDE (18).

In Section 3 in [6] it was shown (see also Theorem 2 in [3]), that if we let Θ be the stochastic (Stratonovich) flow

$$
\Theta(\omega; t, x) = x + \int_0^t c(\Theta(\omega; r, x)) \circ dY_r(\omega),
$$

then with $\hat{S}_t := \Theta^{-1}(t, S_t)$ we have $\hat{\mathbb{P}}$-a.s.

$$\begin{aligned}\hat{S}_s(\omega, \omega^{\bar{B}}) &= S_0 + \int_0^s \hat{a}(r, \hat{S}_r)\, dr \\ &\quad + \int_0^s \hat{b}(r, \hat{S}_r)\, d\bar{B}_r, \qquad s \in [0,t], \qquad \hat{\mathbb{P}}\text{-a.e. } (\omega, \omega^{\bar{B}}).\end{aligned} \tag{27}$$

Here, componentwise,

$$\begin{aligned}\hat{a}(t,x)_i &:= \sum_k \partial_{x_k} \Theta_i^{-1}\big(t, \Theta(t,x)\big) a_k\big(\Theta(t,x)\big) \\ &\quad + \frac{1}{2} \sum_{j,k} \partial_{x_j x_k} \Theta_i^{-1}\big(t, \Theta(t,x)\big) \sum_l b_{jl}\big(\Theta(t,x)\big) b_{kl}\big(\Theta(t,x)\big), \\ \hat{b}(t,x)_{ij} &:= \sum_k \partial_{x_k} \Theta_i^{-1}(t,x) b_{kj}\big(\Theta(t,x)\big).\end{aligned}$$

Especially, by a Fubini-type theorem (e.g., Theorem 3.4.1 in [4]), there exists Ω_0 with $\mathbb{P}_0(\Omega_0) = 1$ such that for $\omega \in \Omega_0$ equation (27) holds true $\bar{\mathbb{P}}$-a.s.

Let $\mathbf{Y} \in \mathcal{C}^{0,\alpha}$ be the enhanced Brownian motion over Y. We can then construct ω-wise the rough flow $\phi^{\mathbf{Y}(\omega)}$ as given in (24). By the very definition of Ξ we know that $\tilde{S}_t^{\mathbf{Y}(\omega)}(\omega) := (\phi^{\mathbf{Y}(\omega)})^{-1}(\omega; t, \Xi(\boldsymbol{\omega})_t)$ satifies the SDE

$$\begin{aligned}\tilde{S}_t^{\mathbf{Y}(\omega)} &= S_0 + \int_0^t \hat{b}^{\mathbf{Y}(\omega)}\big(r, \tilde{S}_r^{\mathbf{Y}(\omega)}\big)\, dr \\ &\quad + \int_0^t \hat{b}^{\mathbf{Y}(\omega)}\big(r, \tilde{S}_r^{\mathbf{Y}(\omega)}\big)\, d\bar{B}_r, \qquad \bar{\mathbb{P}}\text{-a.e. } \omega^{\bar{B}},\end{aligned} \tag{28}$$

where

$$\begin{aligned}\tilde{a}^{\mathbf{Y}(\omega)}(t,x)_i &:= \sum_k \partial_{x_k} \big(\phi^{\mathbf{Y}(\omega)}\big)_i^{-1}\big(t, \phi^{\mathbf{Y}(\omega)}(t,x)\big) a_k\big(t, \phi^{\mathbf{Y}(\omega)}(t,x)\big) \\ &\quad + \frac{1}{2} \sum_{j,k} \partial_{x_j x_k} \big(\phi^{\mathbf{Y}(\omega)}\big)_i^{-1}\big(t, \phi^{\mathbf{Y}(\omega)}(t,x)\big) \\ &\quad \times \sum_l b_{jl}\big(t, \phi^{\mathbf{Y}(\omega)}(t,x)\big) b_{kl}\big(t, \phi^{\mathbf{Y}(\omega)}(t,x)\big), \\ \tilde{b}^{\mathbf{Y}(\omega)}(t,x)_{ij} &:= \sum_k \partial_{x_k} \big(\phi^{\mathbf{Y}(\omega)}\big)_i^{-1}(t,x) b_{kj}\big(t, \phi^{\mathbf{Y}(\omega)}(t,x)\big).\end{aligned}$$

It is a classical rough path result (see, e.g., Section 17.5 in [25]), that there exists Ω_1 with $\mathbb{P}^Y(\Omega_1) = 1$ such that for $\omega \in \Omega_1$, we have

$$\phi^{\mathbf{Y}(\omega)}(\cdot,\cdot) = \Theta(\omega; \cdot,\cdot).$$

Hence for $\omega \in \Omega_1$ we have that $\hat{a} = \tilde{a}^{\mathbf{Y}(\omega)}$, $\hat{b} = \tilde{b}^{\mathbf{Y}(\omega)}$. Hence for $\omega \in \Omega_0 \cap \Omega_1$ the processes $\hat{S}_t(\omega, \cdot)$, $\tilde{S}_t^{\mathbf{Y}(\omega)}(\cdot)$ satisfy the same Lipschitz SDE (with respect to $\bar{\mathbb{P}}$).[15] By strong uniqueness we hence have for $\omega \in \Omega_0 \cap \Omega_1$ that $\bar{\mathbb{P}}$-a.s.

$$\hat{S}_s(\omega, \cdot) = \tilde{S}_s^{\mathbf{Y}(\omega)}(\cdot), \qquad s \le t.$$

Hence for $\omega \in \Omega_0 \cap \Omega_1$

$$S_s(\omega, \cdot) = \Xi(\mathbf{Y}(\omega))(\cdot)_s, \qquad s \le t, \qquad \bar{\mathbb{P}}\text{-a.s.} \qquad \square$$

REMARK 8. We remark that the above idea of a flow decomposition is also used in the work by Davis [10, 12–14]. Without rough path theory this approach is restricted to one-dimensional observation, since, for multidimensional flows, one cannot hope for continuous dependence on the driving signal in supremum norm.

LEMMA 9. *Let η be a smooth d_Y-dim path η and S be the solution of the the following classical SDE*

$$S_t = S_0 + \int_0^t a(S_r)\,dr + \int_0^t b(S_r)\,d\bar{B}_r + \int_0^t c(S_r)\,d\eta_r,$$

where $\bar{B}$ is a d_B-dimensional Brownian motion,

$$\int_0^t c(S_r)\,d\eta_r := \sum_{i=1}^{d_S} \int_0^t c_i(S_r)\dot{\eta}_r^i\,dr,$$

$a \in \mathrm{Lip}^1(\mathbb{R}^{d_S})$, $b_1, \dots, b_{d_B} \in \mathrm{Lip}^1(\mathbb{R}^{d_S})$, $c_1, \dots, c_{d_Y} \in \mathrm{Lip}^{4+\epsilon}(\mathbb{R}^{d_S})$, *and* $S_0 \in L^\infty(\bar{\Omega}; \mathbb{R}^{d_S})$ *independent of $\bar{B}$. Consider the flow*

$$\phi(t, x) = x + \int_0^t c(\phi(r, x))\,d\eta_r. \tag{29}$$

Then $\tilde{S}_t := \phi^{-1}(t, S_t)$ satisfies the following SDE:

$$\tilde{S}_t = S_0 + \int_0^t \tilde{a}(r, \tilde{S}_r)\,dr + \int_0^t \tilde{b}(r, \tilde{S}_r)\,d\bar{B}_r,$$

where we define componentwise

$$\begin{aligned}
\tilde{a}(t, x)_i &:= \sum_k \partial_{x_k}\phi_i^{-1}(t, \phi(t, x)) a_k(\phi(t, x)) \\
&\quad + \frac{1}{2}\sum_{j,k} \partial_{x_j x_k}\phi_i^{-1}(t, \phi(t, x)) \sum_l b_{jl}(\phi(t, x)) b_{kl}(\phi(t, x)), \\
\tilde{b}(t, x)_{ij} &:= \sum_k \partial_{x_k}\phi_i^{-1}(t, \phi(t, x)) b_{kj}(\phi(t, x)).
\end{aligned}$$

[15] Here one has to argue that fixing ω in equation (28) gives ($\mathbb{P}_0$-a.s.) the solution to the respective SDE on $\Omega^{\bar{B}}$.

PROOF. Denote $\psi(t,x) := \phi^{-1}(t,x)$. Then

$$\psi(r,x) = x - \int_0^t \partial_x \psi(r,x) c(x)\, d\eta_r.$$

By Itô's formula,

$$\begin{aligned}
&\psi_i(t,S_t) - \psi_i(0,S_0) \\
&\quad = \int_0^t \partial_t \psi_i(r,S_r)\, dr + \sum_j \int_0^t \partial_{x_j} \psi_i(r,S_r)\, dS_j(r) \\
&\qquad + \sum_{j,k} \frac{1}{2} \int_0^t \partial_{x_j x_k} \psi_i(r,S_r)\, d\langle S_k, S_j\rangle_r \\
&\quad = \sum_j \int_0^t \partial_{x_j} \psi_i(r,S_r) a_j(S_r)\, dr + \sum_j \int_0^t \partial_{x_j} \psi_i(r,S_r) \sum_k b_{jk}(S_r)\, d\bar{B}_k(r) \\
&\qquad + \sum_{j,k} \frac{1}{2} \int_0^t \partial_{x_j x_k} \psi_i(S_r) \sum_l b_{kl}(S_r) b_{jl}(S_r)\, dr.
\end{aligned}$$

□

LEMMA 10. *Consider for a rough path $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$ the coefficients transformed analogously to Lemma 9, $\tilde{a}^{\boldsymbol{\eta}}$, $\tilde{b}^{\boldsymbol{\eta}}$; that is, consider the rough flow*

$$\phi(t,x) = \phi^{\boldsymbol{\eta}}(t,x) = x + \int_0^t c\big(\phi(r,x)\big)\, d\boldsymbol{\eta}_r \tag{30}$$

and define

$$\begin{aligned}
\tilde{a}^{\boldsymbol{\eta}}(t,x)_i &:= \sum_k \partial_{x_k} \phi_i^{-1}\big(t,\phi(t,x)\big) a_k\big(\phi(t,x)\big) \\
&\qquad + \frac{1}{2} \sum_{j,k} \partial_{x_j x_k} \phi_i^{-1}\big(t,\phi(t,x)\big) \sum_l b_{jl}\big(\phi(t,x)\big) b_{kl}\big(\phi(t,x)\big), \\
\tilde{b}^{\boldsymbol{\eta}}(t,x)_{ij} &:= \sum_k \partial_{x_k} \phi_i^{-1}\big(t,\phi(t,x)\big) b_{kj}\big(\phi(t,x)\big).
\end{aligned}$$

Then for every $R > 0$ there exists $K_{\text{Lem 10}} = K_{\text{Lem 10}}(R) < \infty$ such that

$$\begin{aligned}
\sup_{\boldsymbol{\eta}:|\boldsymbol{\eta}|_{\alpha\text{-Höl}} < R} \big|\tilde{a}^{\boldsymbol{\eta}}\big|_\infty &\le K_{\text{Lem 10}}, \\
\sup_{\boldsymbol{\eta}:|\boldsymbol{\eta}|_{\alpha\text{-Höl}} < R} \big|\tilde{b}^{\boldsymbol{\eta}}\big|_\infty &\le K_{\text{Lem 10}}, \\
\sup_{\boldsymbol{\eta}:|\boldsymbol{\eta}|_{\alpha\text{-Höl}} < R} \sup_{s \le t} \big|D\tilde{a}^{\boldsymbol{\eta}}(s,\cdot)\big|_\infty &\le K_{\text{Lem 10}}, \\
\sup_{\boldsymbol{\eta}:|\boldsymbol{\eta}|_{\alpha\text{-Höl}} < R} \sup_{s \le t} \big|D\tilde{b}^{\boldsymbol{\eta}}(s,\cdot)\big|_\infty &\le K_{\text{Lem 10}},
\end{aligned}$$

and such that if $\boldsymbol{\eta}, \tilde{\boldsymbol{\eta}}$ *are two rough paths with* $|\boldsymbol{\eta}|_{\alpha\text{-Höl}}, |\tilde{\boldsymbol{\eta}}|_{\alpha\text{-Höl}} < R$, *we have*

$$\sup_{t,x}\left|\tilde{a}^1(t,x) - \tilde{a}^2(t,x)\right| \le K_{\text{Lem 10}}(R)\rho_{\alpha\text{-Höl}}\left(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2\right),$$

$$\sup_{t,x}\left|\tilde{b}^1(t,x) - \tilde{b}^2(t,x)\right| \le K_{\text{Lem 10}}(R)\rho_{\alpha\text{-Höl}}\left(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2\right).$$

PROOF. This is a straightforward calculation using Lemma 13 and the properties of a, b. $\square$

The following is a standard result for continuous dependence of SDEs on parameters.

LEMMA 11. *Let* $\tilde{a}^i(t,x)$, $\tilde{b}^i(t,x)$, $i = 1, 2$, *be bounded and uniformly Lipschitz in* x.

Let $\tilde{S}^i$ *be the corresponding unique solutions to the SDEs*

$$\tilde{S}^i_t = S_0 + \int_0^t \tilde{a}^i(r, \tilde{S}^i_r)\,dr + \int_0^t \tilde{b}^i(r, \tilde{S}^i_r)\,d\bar{B}_r, \qquad i = 1, 2.$$

Assume

$$\sup_{s\le t}\left|D\tilde{a}^1(s,\cdot)\right|_\infty, \qquad \sup_{s\le t}\left|D\tilde{b}^1(s,\cdot)\right|_\infty < K < \infty,$$

$$\sup_{r,x}\left(\tilde{a}^1(r,x) - \tilde{a}^2(r,x)\right), \qquad \sup_{r,x}\left(\tilde{b}^1(r,x) - \tilde{b}^2(r,x)\right) < \varepsilon < \infty.$$

Then there exists $C_{\text{Lem 11}} = C_{\text{Lem 11}}(K)$ *such that*

$$\mathbb{E}\left[\sup_{s\le t}\left|\tilde{S}^1_s - \tilde{S}^2_s\right|^2\right]^{1/2} \le C_{\text{Lem 11}}\varepsilon.$$

PROOF. This is a straightforward application of Itô's formula and the Burkholder–Davis–Gundy inequality. $\square$

We now apply the previous lemma to our concrete setting.

LEMMA 12. *Let* $\boldsymbol{\eta}^1, \boldsymbol{\eta}^2 \in \mathcal{C}^{0,\alpha}$ *and let* $\tilde{S}^1, \tilde{S}^2$ *be the corresponding unique solutions to the SDEs*

$$\tilde{S}^i_t = S_0 + \int_0^t \tilde{a}^i(r, \tilde{S}^i_r)\,dr + \int_0^t \tilde{b}^i(r, \tilde{S}^i_r)\,d\bar{B}_r, \qquad i = 1, 2,$$

where $\tilde{a}^i, \tilde{b}^i$ *are given as in Lemma* 10.

Assume $R > \max\{|\boldsymbol{\eta}^1|_{\alpha\text{-Höl}}, |\boldsymbol{\eta}^2|_{\alpha\text{-Höl}}\}$. *Then there exists* $C_{\text{Lem 12}} = C_{\text{Lem 12}}(R)$:

$$\mathbb{E}\left[\sup_{s\le t}\left|\tilde{S}^1_s - \tilde{S}^2_s\right|^2\right]^{1/2} \le C_{\text{Lem 12}}\rho_{\alpha\text{-Höl}}\left(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2\right).$$

PROOF. Fix $R > 0$. Let $\|\boldsymbol{\eta}^1\|_{\alpha\text{-Höl}}, \|\boldsymbol{\eta}^2\|_{\alpha\text{-Höl}} < R$.
From Lemma 10 we know that

$$\sup_{t,x} \bigl|\tilde{b}^1(t,x) - \tilde{b}^2(t,x)\bigr| \le K_{\text{Lem } 10}(R)\rho_{\alpha\text{-Höl}}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2).$$

Analogously, we get

$$\sup_{t,x} \bigl|\tilde{a}^1(t,x) - \tilde{a}^2(t,x)\bigr| \le L_2 \rho_{\alpha\text{-Höl}}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2)$$

for a $L_2 = L_2(R)$. □

LEMMA 13. *Let $\alpha \in (0,1)$. Let $\gamma > \frac{1}{\alpha} \ge 1$, $k \in \{1, 2, \ldots\}$ and assume that $V = (V_1, \ldots, V_d)$ is a collection of $\mathrm{Lip}^{\gamma+k}$-vector fields on $\mathbb{R}^e$. Write $n = (n_1, \ldots, n_e) \in \mathbb{N}^e$ and assume $|n| := n_1 + \cdots + n_e \le k$.*

Then, for all $R > 0$ there exist $C = C(R, |V|_{\mathrm{Lip}^{\gamma+k}})$, $K = K(R, |V|_{\mathrm{Lip}^{\gamma+k}})$ such that if $\mathbf{x}^1, \mathbf{x}^2 \in C^{\alpha\text{-Höl}}([0,t], G^{[p]}(\mathbb{R}^d))$ with $\max_i \|\mathbf{x}^i\|_{\alpha\text{-Höl};[0,t]} \le R$, then

$$\sup_{y_0 \in \mathbb{R}^e} \bigl|\partial_n \pi_{(V)}(0, y_0; \mathbf{x}^1) - \partial_n \pi_{(V)}(0, y_0; \mathbf{x}^2)\bigr|_{\alpha\text{-Höl};[0,t]} \le C\rho_{\alpha\text{-Höl}}(\mathbf{x}^1, \mathbf{x}^2),$$

$$\sup_{y_0 \in \mathbb{R}^e} \bigl|\partial_n \pi_{(V)}(0, y_0; \mathbf{x}^1)^{-1} - \partial_n \pi_{(V)}(0, y_0; \mathbf{x}^2)^{-1}\bigr|_{\alpha\text{-Höl};[0,t]} \le C\rho_{\alpha\text{-Höl}}(\mathbf{x}^1, \mathbf{x}^2),$$

$$\sup_{y_0 \in \mathbb{R}^e} \bigl|\partial_n \pi_{(V)}(0, y_0; \mathbf{x}^1)\bigr|_{\alpha\text{-Höl};[0,t]} \le K,$$

$$\sup_{y_0 \in \mathbb{R}^e} \bigl|\partial_n \pi_{(V)}(0, y_0; \mathbf{x}^1)^{-1}\bigr|_{\alpha\text{-Höl};[0,t]} \le K.$$

PROOF. The fact that $V \in \mathrm{Lip}^{\gamma+k}$ (instead of just $\mathrm{Lip}^{\gamma+k-1}$) entails that the derivatives up to order k are unique, nonexplosive solutions to RDEs with $\mathrm{Lip}^{\gamma}_{\mathrm{loc}}$ vector fields; see Section 11 in [25]. Localization (uniform for driving paths bounded in α-Hölder norm) then yields the desired results. □

REFERENCES

[1] BAGCHI, A. and KARANDIKAR, R. (1994). White noise theory of robust nonlinear filtering with correlated state and observation noises. *Systems Control Lett.* **23** 137–148. MR1287606

[2] BAIN, A. and CRISAN, D. (2009). *Fundamentals of Stochastic Filtering*. *Stochastic Modelling and Applied Probability* **60**. Springer, New York. MR2454694

[3] BEN AROUS, G. and CASTELL, F. (1996). Flow decomposition and large deviations. *J. Funct. Anal.* **140** 23–67. MR1404573

[4] BOGACHEV, V. I. (2007). *Measure Theory. Vol. I, II*. Springer, Berlin. MR2267655

[5] BREIMAN, L. (1992). *Probability*. *Classics in Applied Mathematics* **7**. SIAM, Philadelphia, PA. MR1163370

[6] BUCKDAHN, R. and MA, J. (2007). Pathwise stochastic control problems and stochastic HJB equations. *SIAM J. Control Optim.* **45** 2224–2256 (electronic). MR2285722

[7] CLARK, J. M. C. (1978). The design of robust approximations to the stochastic differential equations of nonlinear filtering. In *Communication Systems and Random Process Theory* (*Proc*. 2*nd NATO Advanced Study Inst*., *Darlington*, 1977). *NATO Advanced Study Inst*. *Ser*., *Ser*. *E*: *Appl*. *Sci*. **25** 721–734. Sijthoff & Noordhoff, Alphen aan den Rijn. MR0529130

[8] CLARK, J. M. C. and CRISAN, D. (2005). On a robust version of the integral representation formula of nonlinear filtering. *Probab*. *Theory Related Fields* **133** 43–56. MR2197136

[9] CRISAN, D. and ROZOVSKIĬ, B., eds. (2011). *The Oxford Handbook of Nonlinear Filtering*. Oxford Univ. Press, Oxford. MR2882749

[10] DAVIS, M. H. A. (1980). On a multiplicative functional transformation arising in nonlinear filtering theory. *Z*. *Wahrsch*. *Verw*. *Gebiete* **54** 125–139. MR0597335

[11] DAVIS, M. H. A. (1981). Pathwise nonlinear filtering. In *Stochastic Systems*: *The Mathematics of Filtering and Identification and Applications*, *Proc. NATO Adv. Study Inst. Les Arcs, Savoie, France 1980* 505–528. NATO Advanced Study Institute Series.

[12] DAVIS, M. H. A. (1982). A pathwise solution of the equations of nonlinear filtering. *Teor*. *Veroyatn*. *Primen*. **27** 160–167. MR0645139

[13] DAVIS, M. H. A. (2011). Pathwise nonlinear filtering with correlated noise. In *The Oxford Handbook of Nonlinear Filtering* 403–424. Oxford Univ. Press, Oxford. MR2884603

[14] DAVIS, M. H. A. and SPATHOPOULOS, M. P. (1987). Pathwise nonlinear filtering for nondegenerate diffusions with noise correlation. *SIAM J*. *Control Optim*. **25** 260–278. MR0877062

[15] DEL MORAL, P. (2004). *Feynman–Kac Formulae*: *Genealogical and Interacting Particle Systems with Applications*. Springer, New York. MR2044973

[16] DEL MORAL, P., HU, P. and WU, L. (2010). On the concentration properties of interacting particle processes. *Machine Learning* **3** 225–389.

[17] DEL MORAL, P., JACOD, J. and PROTTER, P. (2001). The Monte-Carlo method for filtering with discrete-time observations. *Probab*. *Theory Related Fields* **120** 346–368. MR1843179

[18] DEL MORAL, P. and MICLO, L. (2000). Branching and interacting particle systems approximations of Feynman–Kac formulae with applications to non-linear filtering. In *Séminaire de Probabilités*, *XXXIV*. *Lecture Notes in Math*. **1729** 1–145. Springer, Berlin. MR1768060

[19] DIEHL, J. and OBERHAUSER, H. (2013). A Levy-area between Brownian motion and rough paths with applications to robust non-linear filtering and RPDEs. Preprint available at http://arxiv.org/abs/1301.3799.

[20] DOSS, H. (1977). Liens entre équations différentielles stochastiques et ordinaires. *Ann*. *Inst*. *H*. *Poincaré Sect*. *B* (*N.S.*) **13** 99–125. MR0451404

[21] ELLIOTT, R. J. and KOHLMANN, M. (1981). Robust filtering for correlated multidimensional observations. *Math*. *Z*. **178** 559–578. MR0638818

[22] FLORCHINGER, P. (1993). Zakai equation of nonlinear filtering with unbounded coefficients. The case of dependent noises. *Systems Control Lett*. **21** 413–422. MR1244766

[23] FLORCHINGER, P. and NAPPO, G. (2011). Continuity of the filter with unbounded observation coefficients. *Stoch*. *Anal*. *Appl*. **29** 612–630. MR2812519

[24] FRIZ, P. K. (2005). Continuity of the Itô-map for Hölder rough paths with applications to the support theorem in Hölder norm. In *Probability and Partial Differential Equations in Modern Applied Mathematics*. *The IMA Volumes in Mathematics and Its Applications* **140** 117–135. Springer, New York. MR2202036

[25] FRIZ, P. K. and VICTOIR, N. B. (2010). *Multidimensional Stochastic Processes as Rough Paths*: *Theory and Applications*. *Cambridge Studies in Advanced Mathematics* **120**. Cambridge Univ. Press, Cambridge. MR2604669

[26] KUSHNER, H. J. (1979). A robust discrete state approximation to the optimal nonlinear filter for a diffusion. *Stochastics* **3** 75–83. MR0553906
[27] KUSHNER, H. J. (1979). A robust discrete state approximation to the optimal nonlinear filter for a diffusion. *Stochastics* **3** 75–83. MR0553906
[28] LYONS, T. (1994). Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young. *Math*. *Res*. *Lett*. **1** 451–464. MR1302388
[29] LYONS, T. and QIAN, Z. (2002). *System Control and Rough Paths*. Oxford Univ. Press, Oxford. MR2036784
[30] LYONS, T. J., CARUANA, M. and LÉVY, T. (2007). *Differential Equations Driven by Rough Paths*. *Lecture Notes in Math*. **1908**. Springer, Berlin. MR2314753
[31] PROTTER, P. E. (2004). *Stochastic Integration and Differential Equations*: *Stochastic Modelling and Applied Probability*, 2nd ed. *Applications of Mathematics* (*New York*) **21**. Springer, Berlin. MR2020294
[32] SUSSMANN, H. J. (1978). On the gap between deterministic and stochastic ordinary differential equations. *Ann*. *Probab*. **6** 19–41. MR0461664

D. CRISAN
DEPARTMENT OF MATHEMATICS
IMPERIAL COLLEGE LONDON
180 QUEEN'S GATE
LONDON SW7 2BZ
UNITED KINGDOM
E-MAIL: d.crisan@ic.ac.uk

J. DIEHL
H. OBERHAUSER
INSTITUT FÜR MATHEMATIK, MA 7-2
TECHNICAL UNIVERSITY BERLIN
STRASSE DES 17. JUNI 136
10623 BERLIN
GERMANY
E-MAIL: diehl@math.tu-berlin.de
h.oberhauser@gmail.com

P. FRIZ
INSTITUT FÜR MATHEMATIK, MA 7-2
STRASSE DES 17. JUNI 136
10623 BERLIN
GERMANY
AND
WEIERSTRASS-INSTITUT FÜR ANGEWANDTE ANALYSIS UND STOCHASTIK
MOHRENSTRASSE 39
10117 BERLIN
GERMANY
E-MAIL: friz@math.tu-berlin.de
friz@wias-berlin.de