On Uniqueness of Solutions for the Stochastic Differential Equations of Nonlinear Filtering [1] [2]

Vladimir M. Lucic and Andrew J. Heunis
Department of Electrical and Computer Engineering
University of Waterloo

[1] Research supported by MITACS/NSERC
[2] Appeared in Annals of Applied Probability, v.11(1), pp. 182-209, 2001

Abbreviated Title: Uniqueness of Solutions

AMS Subject Classifications: Primary 60G35; Secondary 60G44, 60G57.

Key Words: nonlinear filter equations, weak solutions, uniqueness

Abstract

We study a nonlinear filtering problem in which the signal to be estimated is conditioned by the observations. The main results establish pathwise uniqueness for the unnormalized filter equation and uniqueness in law for the normalized and unnormalized filter equations.

1 Introduction

An early work on uniqueness for the stochastic differential equations of nonlinear filtering is that of Szpirglas [15]. The basic viewpoint adopted in [15] is to regard the measure-valued stochastic differential equations of nonlinear filtering as entities quite separate from the original nonlinear filtering problem, for which one can formulate the notions of solution (or weak solution), pathwise uniqueness and uniqueness in law, by essentially adapting these concepts from the theory of Itô stochastic differential equations (for which see Section IV.1 of Ikeda and Watanabe [4] or Section IX.1 of Revuz and Yor [12]). With these notions at hand, it is then established in [15] that pathwise uniqueness and uniqueness in law hold for both the normalized (Fujisaki-Kallianpur-Kunita) and unnormalized (Duncan-Mortensen-Zakai) filter equations, in the case of a nonlinear filtering problem where the signal is a Markov process which is independent of the Wiener process in the observation equation, and the sensor function in the observation equation is uniformly bounded.

Our goal is to look at uniqueness for the stochastic differential equations of nonlinear filtering from a point of view very similar to that of Szpirglas [15], but for a nonlinear filtering problem in which there is dependence of the signal on the observations. In fact, we shall look at the specific nonlinear filtering problem where the signal $\{X_t\}$ is an $\mathbb{R}^d$-valued process solving an equation of the form

$$dX_t = b(X_t)\,dt + B(X_t)\,dW_t + c(X_t)\,dV_t, \tag{1.1}$$

the $\mathbb{R}^{d_1}$-valued observation process $\{Y_t\}$ is defined by

$$Y_t = W_t + \int_0^t h(X_s)\,ds, \tag{1.2}$$

and $\{(W_t, V_t)\}$ is a standard $\mathbb{R}^{d_1+d_2}$-valued Wiener process (precise conditions on the mappings $b(\cdot)$, $B(\cdot)$, $c(\cdot)$ and $h(\cdot)$ will be stated in Section 2). The pair (1.1) and (1.2) represents a simple model of a signal and observation in which the signal $\{X_t\}$ depends on the Wiener process $\{W_t\}$ of the observation equation.

Motivated by Szpirglas [15], we shall regard the normalized and unnormalized filter equations for this nonlinear filtering problem as measure-valued stochastic differential equations, defined quite independently of the filtering problem, and will formulate the notions of weak solution, pathwise uniqueness, and uniqueness in law for the filter equations. Our main result (see Theorem 2.21 to follow) establishes pathwise uniqueness for the unnormalized filter equation, together with uniqueness in law for the normalized and unnormalized filter equations, subject to reasonably general conditions on the mappings $b(\cdot)$, $B(\cdot)$, and $c(\cdot)$ in the signal equation (1.1), and a uniform boundedness condition on the sensor function $h(\cdot)$ in the observation equation (1.2). As will be seen from the discussion of Section 2 (see Remark 2.22) the elegant semigroup ideas used in Szpirglas [15] to establish pathwise uniqueness do not seem to extend to the filtering problem represented by (1.1) and (1.2), where the signal $\{X_t\}$ depends on the observation Wiener process $\{W_t\}$, and our approach necessarily involves a different method of proof.

In Section 2 we review the normalized and unnormalized filter equations for the nonlinear filtering problem given by (1.1) and (1.2), define weak solutions, pathwise uniqueness, and uniqueness in law for the filter equations, and state the main result, namely Theorem 2.21. We also discuss the relationship of this result with other works on uniqueness for the nonlinear filter equations in Remark 2.22, Remark 2.23, and Remark 2.24. Section 3 is devoted to the proof of the main result, while the proofs of various technical facts and lemmas needed in Section 3 are relegated to Section 4, Section 5, and Section 6.

2 Stochastic Differential Equations of Nonlinear Filtering

Remark 2.1. For easy access we first summarize most of the basic notation which will be used in the sequel:

(i) For a metric space E, let $\mathcal{B}(E)$ denote the Borel σ-algebra on E, let $B(E)$ denote the set of all real-valued uniformly-bounded Borel measurable mappings on E, and, for $\phi \in B(E)$, define the supremum norm by $\|\phi\| := \sup_{x\in E} |\phi(x)|$. Likewise, write $C(E)$ for the set of all real-valued continuous mappings on E, and write $\bar{C}(E)$ for the collection of all members of $C(E)$ which are uniformly bounded.

(ii) For a complete separable metric space E, let $\mathcal{M}^+(E)$ denote the space of all positive bounded measures on the measurable space $(E, \mathcal{B}(E))$, with the usual topology of weak (or narrow) convergence. Then $\mathcal{M}^+(E)$ is separable and metrically topologically complete, and Exercise 9.5.6 of Ethier and Kurtz [3] shows that a simple variant of the Prohorov metric turns the topological space $\mathcal{M}^+(E)$ into a complete separable metric space. Also, let $\mathcal{P}(E)$ denote the collection of all members of $\mathcal{M}^+(E)$ which are probability measures. For $\mu \in \mathcal{M}^+(E)$ and a $\mathcal{B}(E)$-measurable and μ-integrable mapping ϕ from E into $\mathbb{R}$, write $\mu(\phi)$ or $\mu\phi$ for the integral $\int_E \phi d\mu$.

(iii) For a vector x in a finite-dimensional Euclidean space $\mathbb{R}^q$, write x^k for the k-th scalar entry of x, $\forall\, k = 1, \ldots, q$, and write $|x|$ for the Euclidean norm of x, namely $|x|^2 := \sum_{k=1}^q (x^k)^2$. Also, let $C^\infty(\mathbb{R}^q)$ denote the set of all infinitely smooth real-valued mappings on $\mathbb{R}^q$, and let $C_c^\infty(\mathbb{R}^q)$ be the collection of all members of $C^\infty(\mathbb{R}^q)$ with compact support. Finally, let $\hat{C}(\mathbb{R}^q)$ denote the collection of all members of $\bar{C}(\mathbb{R}^q)$ which vanish at infinity.

(iv) For the positive integers q, r, let $\mathbb{R}^{q\times r}$ denote the set of q by r matrices with real entries. Likewise, let $\mathbb{S}_+^{q\times q}$ denote the collection of all members of $\mathbb{R}^{q\times q}$ which are symmetric non-negative definite, and let $\mathbb{S}_{++}^{q\times q}$ denote the collection of all members of $\mathbb{S}_+^{q\times q}$ which are strictly positive definite.

(v) For the positive integer q, let $\mathbb{R}^{q*}$ denote the compact metric space which is the one-point compactification of the Euclidean space $\mathbb{R}^q$.

Now consider a nonlinear filtering problem made up of the following basic elements:

E.1 A fixed interval of interest $[0, T]$, with $T \in (0, \infty)$.

E.2 A complete probability space $(\Omega, \mathcal{F}, P)$ carrying a filtration $\{\mathcal{F}_t,\ t \in [0,T]\}$ such that $\mathcal{F}_0$ includes all null events of $(\Omega, \mathcal{F}, P)$. Defined on $(\Omega, \mathcal{F}, P)$ is an $\mathbb{R}^d$-valued continuous $\{\mathcal{F}_t\}$-adapted process $\{X_t,\ t \in [0,T]\}$ and an $\mathbb{R}^{d_1+d_2}$-valued $\{\mathcal{F}_t\}$-Wiener process $\{(W_t, V_t),\ t \in [0,T]\}$ such that (1.1) holds, where $b : \mathbb{R}^d \to \mathbb{R}^d$, $B : \mathbb{R}^d \to \mathbb{R}^{d\times d_1}$, and $c : \mathbb{R}^d \to \mathbb{R}^{d\times d_2}$ are Borel-measurable and locally bounded functions (that is, uniformly bounded over bounded subsets of $\mathbb{R}^d$).

E.3 an $\mathbb{R}^{d_1}$-valued *observation process* $\{Y_t,\ t \in [0,T]\}$ defined by (1.2), where $h : \mathbb{R}^d \to \mathbb{R}^{d_1}$ is Borel-measurable, with

$$\mathrm{E}\left[\sum_{k=1}^{d_1}\int_0^T |h^k(X_u)|^2\,du\right] < \infty. \tag{2.3}$$

Define the *observation filtration* $\{\mathcal{F}_t^Y,\ t \in [0,T]\}$ by

$$\mathcal{F}_t^Y := \sigma\{Y_u,\ u \in [0,t]\} \vee \mathcal{N}(P), \qquad \text{where} \qquad \mathcal{N}(P) := \{N \in \mathcal{F} :\ P(N) = 0\}. \tag{2.4}$$

From Lemma 1.1 of Kurtz and Ocone [10] there exists some $\mathcal{P}(\mathbb{R}^d)$-valued $\{\mathcal{F}_{t+}^Y\}$-optional process $\{\pi_t,\ t \in [0,T]\}$, called the *optimal filter*, which is defined on $(\Omega, \mathcal{F}, P)$ and satisfies

$$\pi_t\phi = \mathrm{E}[\phi(X_t)|\mathcal{F}_{t+}^Y] \quad a.s., \quad \forall t \in [0,T], \quad \forall \phi \in B(\mathbb{R}^d). \tag{2.5}$$

From (2.3) and Jensen's inequality we see that

$$\mathrm{E}\left[\sum_{k=1}^{d_1}\int_0^T [\pi_u(|h^k|)]^2\,du\right] < \infty,$$

and we can therefore define the $\mathbb{R}^{d_1}$-valued *innovations process* $\{I_t,\ t \in [0,T]\}$ by

$$I_t^k := Y_t^k - \int_0^t \pi_s h^k\,ds, \quad \forall\, t \in [0,T],\ k = 1,\dots,d_1. \tag{2.6}$$

An important property of the innovations process is that $\{I_t,\ t \in [0,T]\}$ is an $\mathbb{R}^{d_1}$-valued $\{\mathcal{F}_{t+}^Y\}$-Wiener process (see Theorem VI.8.4 of Rogers and Williams [13], observing that the filtration $\{\mathcal{Y}_t\}$ on p. 322 of [13] corresponds to our $\{\mathcal{F}_{t+}^Y\}$). Since $\{I_t\}$ is continuous, it is necessarily $\{\mathcal{F}_t^Y\}$-adapted, thus $\{I_t,\ t \in [0,T]\}$ is a $\{\mathcal{F}_t^Y\}$-Wiener process. Now define $m : \mathbb{R}^d \to \mathbb{R}^{d\times(d_1+d_2)}$ by

$$m(x) := \begin{bmatrix} B(x) & c(x) \end{bmatrix}, \quad \forall\, x \in \mathbb{R}^d,$$

and put

$$\mathcal{A}\phi(x) := \sum_{i=1}^d b^i(x)\partial_i\phi(x) + \frac{1}{2}\sum_{i,j=1}^d [m(x)m^T(x)]^{ij}\partial_i\partial_j\phi(x), \quad \forall x \in \mathbb{R}^d,\ \phi \in C^\infty(\mathbb{R}^d); \tag{2.7a}$$

$$\mathcal{B}_k\phi(x) := \sum_{i=1}^d B^{ik}(x)\partial_i\phi(x), \quad \forall x \in \mathbb{R}^d,\ \phi \in C^\infty(\mathbb{R}^d),\ k = 1,\dots,d_1. \tag{2.7b}$$

For each $\phi \in C_c^\infty(\mathbb{R}^d)$ one sees from (1.1), (1.2), and Itô's formula that the process

$$M_t^\phi := \phi(X_t) - \int_0^t \mathcal{A}\phi(X_s)\,ds, \quad t \in [0,T], \tag{2.8}$$

is a square-integrable $\{\mathcal{F}_t\}$-martingale with

$$\langle M^\phi, W^k\rangle_t = \int_0^t \mathcal{B}_k\phi(X_u)\,du, \quad t \in [0,T],\ k = 1,\dots,d_1. \tag{2.9}$$

This observation, together with Theorem VI.8.11 of Rogers and Williams [13], establishes

Theorem 2.2. *For the nonlinear filtering problem given by* **E.1**, **E.2**, *and* **E.3**, *the optimal filter* $\{\pi_t,\ t \in [0,T]\}$ *satisfies*

$$\pi_t\phi = \pi_0\phi + \int_0^t \pi_s(\mathcal{A}\phi)\,ds + \int_0^t \sum_{k=1}^{d_1}[\pi_s(h^k\phi + \mathcal{B}_k\phi) - (\pi_s h^k)(\pi_s\phi)]\,dI_s^k, \quad \forall\, \phi \in C_c^\infty(\mathbb{R}^d). \tag{2.10}$$

The relation (2.10) is known variously as the Fujisaki-Kallianpur-Kunita equation, the Kushner-Stratonovich equation, or the normalized filter equation.

Remark 2.3. Since $C_c^\infty(\mathbb{R}^d)$ is dense in $\hat{C}(\mathbb{R}^d)$, with respect to the supremum norm, it must be convergence determining (see Problem 3.11.11 of Ethier and Kurtz [3]). Now it follows from (2.10) that $\{\pi_t,\ t \in [0,T]\}$ is a continuous $\mathcal{P}(\mathbb{R}^d)$-valued process, and therefore $\{\mathcal{F}_t^Y\}$-adapted. Thus, we can replace $\mathcal{F}_{t+}^Y$ in (2.5) by $\mathcal{F}_t^Y$.

The characterization of $\{\pi_t\}$ given by Theorem 2.2 becomes useful when some form of uniqueness is established for (2.10). The approach adopted here is suggested by the work of Szpirglas [15], which in turn is motivated by the results of Yamada and Watanabe [16] on weak solutions, pathwise uniqueness, and uniqueness in law for Itô stochastic differential equations (see Section IV.1 of Ikeda and Watanabe [4] or Section IX.1 of Revuz and Yor [12] for a comprehensive account of these ideas). Taking advantage of the fact that the innovations process $\{I_t\}$ which "drives" (2.10) is a standard $\{\mathcal{F}_t^Y\}$-Wiener process, we can follow [15] and regard the normalized filter equation as an entity quite separate from the nonlinear filtering problem, namely as a probability-measure valued stochastic differential equation driven by a standard Wiener process, for which one can formulate the notions of weak solution, pathwise uniqueness, and uniqueness in law as follows: (compare Szpirglas [15], Définition III.1, V.1, V.2, and Bhatt, Kallianpur and Karandikar [1], Definition 9.1):

Definition 2.4. The pair $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\pi}_t, \tilde{I}_t)\}$ is a *weak solution* of the normalized filter equation when:

1. $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P})$ is a complete filtered probability space;
2. $\{\tilde{I}_t,\ t \in [0,T]\}$ is an $\mathbb{R}^{d_1}$-valued $\{\tilde{\mathcal{F}}_t\}$-Wiener process on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$;

3. $\{\tilde{\pi}_t,\ t \in [0,T]\}$ is a $\mathcal{P}(\mathbb{R}^d)$-valued continuous $\{\tilde{\mathcal{F}}_t\}$-adapted process such that

$$\tilde{P}\left(\int_0^T \sum_{k=1}^{d_1} [\tilde{\pi}_s|h^k|]^2\, ds < \infty\right) = 1, \tag{2.11}$$

and, for each $\phi \in C_c^\infty(\mathbb{R}^d)$, the following holds to within indistinguishability

$$\tilde{\pi}_t\phi = \tilde{\pi}_0\phi + \int_0^t \tilde{\pi}_s(\mathcal{A}\phi)\, ds + \sum_{k=1}^{d_1} \int_0^t [\tilde{\pi}_s(h^k\phi + \mathcal{B}_k\phi) - (\tilde{\pi}_s h^k)(\tilde{\pi}_s\phi)]\, d\tilde{I}_s^k, \quad t \in [0,T]. \tag{2.12}$$

Remark 2.5. The terminology that $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P})$ is a "complete filtered probability space" will always be understood to mean that $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ is a complete probability space carrying the filtration $\{\tilde{\mathcal{F}}_t,\ t \in [0,T]\}$, and $\tilde{\mathcal{F}}_0$ includes all P-null events in $\tilde{\mathcal{F}}$.

Remark 2.6. In view of Definition 2.4, it follows that $\{(\Omega, \mathcal{F}, \{\mathcal{F}_t^Y\}, P), (\pi_t, I_t)\}$ for $\{\mathcal{F}_t^Y,\ t \in [0,T]\}$, $\{\pi_t,\ t \in [0,T]\}$, and $\{I_t,\ t \in [0,T]\}$ defined by (2.4), (2.5), and (2.6) is a weak solution of the normalized filter equation.

Definition 2.7. The normalized filter equation has the property of *pathwise uniqueness* when the following holds: If $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\pi}_t^1, \tilde{I}_t)\}$ and $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\pi}_t^2, \tilde{I}_t)\}$ are weak solutions of the normalized filter equation such that $\tilde{P}(\tilde{\pi}_0^1 = \tilde{\pi}_0^2) = 1$, then

$$\tilde{P}\left(\tilde{\pi}_t^1 = \tilde{\pi}_t^2 \quad \forall t \in [0,T]\right) = 1.$$

Remark 2.8. For the next definition we shall need the following notation: if ξ is a measurable mapping from some probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ into a separable metric space E, then $\mathcal{L}_{\tilde{P}}(\xi)$ is the probability measure on the Borel σ-algebra $\mathcal{B}(E)$ defined by $\mathcal{L}_{\tilde{P}}(\xi)(\Gamma) := \tilde{P}\{\xi \in \Gamma\}$ for each $\Gamma \in \mathcal{B}(E)$.

Definition 2.9. The normalized filter equation has the property of *uniqueness in joint law* when the following holds: If $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\pi}_t, \tilde{I}_t)\}$ and $\{(\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{P}), (\bar{\pi}_t, \bar{I}_t)\}$ are weak solutions of the normalized filter equation such that $\mathcal{L}_{\tilde{P}}(\tilde{\pi}_0) = \mathcal{L}_{\bar{P}}(\bar{\pi}_0)$, then the processes $\{(\tilde{\pi}_t, \tilde{I}_t),\ t \in [0,T]\}$ and $\{(\bar{\pi}_t, \bar{I}_t),\ t \in [0,T]\}$ have the same finite-dimensional distributions.

Remark 2.10. Under certain conditions one can associate a simpler *unnormalized filter equation* with the normalized filter equation. For this purpose the following additional notation is useful: If $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P})$ is a complete filtered probability space, $\{\tilde{M}_t\}$ is a continuous $\{\tilde{\mathcal{F}}_t\}$-semimartingale, and $\{\tilde{\gamma}_t\}$ is a locally bounded $\{\tilde{\mathcal{F}}_t\}$-progressively measurable process, then $\tilde{\gamma} \bullet \tilde{M}$ denotes the stochastic integral of $\tilde{\gamma}$ with respect to $\tilde{M}$. Also, put

$$\mathcal{E}(\tilde{M})_t := \exp\left(\tilde{M}_t - \frac{1}{2}\langle\tilde{M}\rangle_t\right).$$

Now let $\{(\tilde\Omega, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde P), (\tilde\pi_t, \tilde I_t)\}$ be a weak solution of the normalized filter equation, and define

$$\tilde Y_t^k := \tilde I_t^k + \int_0^t \tilde\pi_s h^k \, ds, \quad \forall\, t \in [0,T],\ k = 1, \ldots, d_1; \tag{2.13}$$

$$\tilde\chi_t := \mathcal{E}\left(-\sum_{k=1}^{d_1} (\tilde\pi h^k) \bullet \tilde I^k\right)_t, \quad \forall\, t \in [0,T]. \tag{2.14}$$

Since $\{\tilde I_t,\ t \in [0,T]\}$ is a $\{\tilde{\mathcal{F}}_t\}$-Wiener process, it follows that $\{\tilde\chi_t,\ t \in [0,T]\}$ is a continuous strictly-positive $\{\tilde{\mathcal{F}}_t\}$-local martingale on $(\tilde\Omega, \tilde{\mathcal{F}}, \tilde P)$, and

$$\frac{1}{\tilde\chi_t} = \mathcal{E}\left(\sum_{k=1}^{d_1} (\tilde\pi h^k) \bullet \tilde Y^k\right)_t, \quad \forall\, t \in [0,T]. \tag{2.15}$$

Define the $\mathcal{M}^+(\mathbb{R}^d)$-valued process $\{\tilde\sigma_t,\ t \in [0,T]\}$ on $(\tilde\Omega, \tilde{\mathcal{F}}, \tilde P)$ by

$$\tilde\sigma_t\phi := \frac{\tilde\pi_t\phi}{\tilde\chi_t}, \quad \forall\, t \in [0,T],\ \phi \in B(\mathbb{R}^d). \tag{2.16}$$

Hence

$$\tilde\sigma_t\phi := (\tilde\pi_t\phi)\, \mathcal{E}\left(\sum_{k=1}^{d_1} (\tilde\pi h^k) \bullet \tilde Y^k\right)_t, \quad \forall\, t \in [0,T],\ \phi \in B(\mathbb{R}^d), \tag{2.17}$$

and, in light of (2.11), we see that

$$\tilde P\left(\int_0^T \sum_{k=1}^{d_1} [\tilde\sigma_s |h^k\phi + \mathcal{B}_k\phi|]^2 \, ds < \infty\right) = 1, \quad \forall\, \phi \in C_c^\infty(\mathbb{R}^d) \cup \{1\}.$$

Using Itô's formula and the relation (2.12), we easily arrive at the *Duncan-Mortensen-Zakai equation* or *unnormalized filter equation*: for each $\phi \in C_c^\infty(\mathbb{R}^d) \cup \{1\}$ we have

$$\tilde\sigma_t\phi = \tilde\sigma_0\phi + \int_0^t \tilde\sigma_s(\mathcal{A}\phi)\, ds + \sum_{k=1}^{d_1} \int_0^t \tilde\sigma_s(h^k\phi + \mathcal{B}_k\phi)\, d\tilde Y_s^k, \quad \forall\, t \in [0,T]. \tag{2.18}$$

Remark 2.11. From Remark 2.3 and (2.17) we see that $t \to \tilde\sigma_t\phi : [0,T] \to \mathbb{R}$ is continuous for each bounded continuous $\phi : \mathbb{R}^d \to \mathbb{R}$, thus $\{\tilde\sigma_t\}$ is a *continuous* $\mathcal{M}^+(\mathbb{R}^d)$-valued process which is $\{\mathcal{F}_t\}$-adapted. Moreover, from (2.17), we see that the random element $\tilde\sigma_0$ takes values in $\mathcal{P}(\mathbb{R}^d)$, the set of *probability* measures on $\mathbb{R}^d$.

Remark 2.12. If, in (2.17), we use the optimal filter $\{\pi_t\}$ in place of $\{\tilde\pi_t\}$ and the observation process $\{Y_t\}$ in place of $\{\tilde Y_t\}$ to get an $\mathcal{M}^+(\mathbb{R}^d)$-valued and $\{\mathcal{F}_t^Y\}$-adapted process $\{\sigma_t\}$, namely

$$\sigma_t\phi := (\pi_t\phi)\, \mathcal{E}\left(\sum_{k=1}^{d_1} (\pi h^k) \bullet Y^k\right)_t, \quad \forall\, t \in [0,T],\ \phi \in B(\mathbb{R}^d), \tag{2.19}$$

then $\{\sigma_t\}$ is called the *unnormalized optimal filter* for the filtering problem given by (1.1) and (1.2).

Remark 2.13. In (2.18) the "driving process" $\{\tilde{Y}_t\}$ is the continuous $\{\tilde{\mathcal{F}}_t\}$-semimartingale defined on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ by (2.13). The equation (2.18) becomes more tractable if we can replace $\tilde{P}$ with some equivalent probability measure $\tilde{Q}$ such that $\{\tilde{Y}_t,\ t \in [0,T]\}$ is an $\{\tilde{\mathcal{F}}_t\}$-Wiener process with respect to $\tilde{Q}$. To this end, observe from (2.14) that $\{(\tilde{\chi}_t, \tilde{\mathcal{F}}_t),\ t \in [0,T]\}$ is a continuous local martingale on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, and that, *if it is a martingale*, then

$$\tilde{Q}(A) := \mathrm{E}^{\tilde{P}}[\tilde{\chi}_T; A], \quad \forall A \in \tilde{\mathcal{F}}, \tag{2.20}$$

defines a probability measure $\tilde{Q}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ which is equivalent to the probability measure $\tilde{P}$, namely

$$\tilde{P} \equiv \tilde{Q} \quad [\tilde{\mathcal{F}}]. \tag{2.21}$$

From (2.13), (2.14), and the Girsanov theorem, it then follows that $\{(\tilde{Y}_t, \tilde{\mathcal{F}}_t),\ t \in [0,T]\}$ is a Wiener process on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{Q})$.

Remark 2.14. A sufficient condition on the weak solution $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\pi}_t, \tilde{I}_t)\}$ and sensor function $h(\cdot)$ which ensures $\{(\tilde{\chi}_t, \tilde{\mathcal{F}}_t),\ t \in [0,T]\}$ is a martingale on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ is that

$$\mathrm{E}^{\tilde{P}}\left[\exp\left(\frac{1}{2}\sum_{k=1}^{d_1}\int_0^T [\tilde{\pi}_s h^k]^2\, ds\right)\right] < \infty$$

(see Corollary 3.5.13 of Karatzas and Shreve [5]). In particular, this condition always holds when $h^k \in B(\mathbb{R}^d)$, $k = 1, \ldots, d_1$.

With the preceding discussion in mind, we next formulate the notion of weak solution of the unnormalized filter equation, pathwise uniqueness and uniqueness in law (compare with Définition IV.1 of Szpirglas [15]):

Definition 2.15. A pair $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q}), (\tilde{\sigma}_t, \tilde{Y}_t)\}$ is a *weak solution* of the unnormalized filter equation when

1. $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q})$ is a complete filtered probability space;
2. $\{\tilde{Y}_t,\ t \in [0,T]\}$ is an $\mathbb{R}^{d_1}$-valued $\{\tilde{\mathcal{F}}_t\}$-Wiener process;
3. $\{\tilde{\sigma}_t,\ t \in [0,T]\}$ is a $\mathcal{M}^+(\mathbb{R}^d)$-valued continuous $\{\tilde{\mathcal{F}}_t\}$-adapted process such that the random element $\tilde{\sigma}_0$ takes values in $\mathcal{P}(\mathbb{R}^d)$, and, for each $\phi \in C_c^\infty(\mathbb{R}^d) \cup \{1\}$, we have the following:

   (a)
   $$\tilde{Q}\left(\int_0^T \sum_{k=1}^{d_1} [\tilde{\sigma}_s|h^k\phi + \mathcal{B}_k\phi|]^2\, ds < \infty\right) = 1; \tag{2.22}$$

   (b) the LHS and RHS of (2.18) are indistinguishable.

Definition 2.16. The unnormalized filter equation has the property of *pathwise uniqueness* when the following holds: If $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q}), (\tilde{\sigma}_t^1, \tilde{Y}_t)\}$ and $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q}), (\tilde{\sigma}_t^2, \tilde{Y}_t)\}$ are weak solutions of the unnormalized filter equation such that $\tilde{Q}(\tilde{\sigma}_0^1 = \tilde{\sigma}_0^2) = 1$, then

$$\tilde{Q}\left(\tilde{\sigma}_t^1 = \tilde{\sigma}_t^2 \quad \forall t \in [0,T]\right) = 1.$$

Definition 2.17. The unnormalized filter equation has the property of *uniqueness in joint law* when the following holds: If $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q}), (\tilde{\sigma}_t, \tilde{Y}_t)\}$ and $\{(\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{Q}), (\bar{\sigma}_t, \bar{Y}_t)\}$ are weak solutions of the unnormalized filter equation such that $\mathcal{L}_{\tilde{Q}}(\tilde{\sigma}_0) = \mathcal{L}_{\bar{Q}}(\bar{\sigma}_0)$, then $\{(\tilde{\sigma}_t, \tilde{Y}_t),\ t \in [0,T]\}$ and $\{(\bar{\sigma}_t, \bar{Y}_t),\ t \in [0,T]\}$ have the same finite-dimensional distributions.

In this paper our goal is to establish pathwise uniqueness for the unnormalized filter equation and uniqueness in joint law for both the normalized and unnormalized filter equations. To this end we postulate the following conditions on the mappings $b(\cdot)$, $B(\cdot)$, $c(\cdot)$ in (1.1), and the mapping $h(\cdot)$ in (1.2):

Condition 2.18. *The mapping $b : \mathbb{R}^d \to \mathbb{R}^d$ is Borel-measurable, and the mappings $B : \mathbb{R}^d \to \mathbb{R}^{d\times d_1}$ and $c : \mathbb{R}^d \to \mathbb{R}^{d\times d_2}$ are continuous. There exists a constant $C \in [0,\infty)$ such that*

$$\max_{i,j,k}\{|b^i(x)|,\ |B^{ij}(x)|, |c^{ik}(x)|\} \le C[1+|x|], \quad \forall x \in \mathbb{R}^d.$$

Condition 2.19. *The mapping $c : \mathbb{R}^d \to \mathbb{R}^{d\times d_2}$ is such that the matrix $c(x)c^T(x)$ is strictly positive definite for every $x \in \mathbb{R}^d$.*

Condition 2.20. *The mapping $h : \mathbb{R}^d \to \mathbb{R}^{d_1}$ is Borel-measurable and uniformly bounded.*

We can now state our main result:

Theorem 2.21. *Suppose that Conditions 2.18, 2.19, and 2.20 hold for the nonlinear filtering problem given by* **E.1***,* **E.2** *and* **E.3**. *Then:*
(i) The unnormalized filter equation has the property of pathwise uniqueness;
(ii) The normalized filter equation has the property of uniqueness in joint law;
(iii) The unnormalized filter equation has the property of uniqueness in joint law.

Remark 2.22. Szpirglas [15] establishes pathwise uniqueness and uniqueness in law for the normalized and unnormalized filter equations corresponding to the following nonlinear filtering problem: The signal $\{X_t\}$ is a homogeneous Markov process with values in a complete separable metric space E and weak infinitesimal generator $\mathcal{A}$, the observation process is

$$Y_t := W_t + \int_0^t h(X_u)\, du, \qquad t \in [0,T],$$

where $\{W_t\}$ is an $\mathbb{R}^{d_1}$-valued Wiener process *independent* of the Markov process $\{X_t\}$, and the sensor function $h : E \to \mathbb{R}^{d_1}$ is uniformly bounded and $\mathcal{B}(E)$-measurable. In this context, by a weak solution of the unnormalized filter equation is meant a pair $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q}), (\tilde{\sigma}_t, \tilde{Y}_t)\}$ such that
(a) $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q})$ is a complete filtered probability space;
(b) $\{\tilde{Y}_t,\ t \in [0,T]\}$ is an $\mathbb{R}^{d_1}$-valued $\{\tilde{\mathcal{F}}_t\}$-Wiener process;
(c) $\{\tilde{\sigma}_t,\ t \in [0,T]\}$ is a $\mathcal{M}^+(E)$-valued, cadlag (or r.c.l.l.) $\{\tilde{\mathcal{F}}_t\}$-adapted process, the random element $\tilde{\sigma}_0$ takes values in $\mathcal{P}(E)$, and $\sup_{t\in[0,T]} E[|\tilde{\sigma}_t 1|^2] < \infty$;

(d) For each $\phi \in \mathcal{D}(\mathcal{A})$ [the domain of the generator $\mathcal{A}$] one has to within indistinguishability that

$$\tilde{\sigma}_t\phi = \tilde{\sigma}_0\phi + \int_0^t \tilde{\sigma}_s(\mathcal{A}\phi)\, ds + \sum_{k=1}^{d_1} \int_0^t \tilde{\sigma}_s(h^k\phi)\, d\tilde{Y}_s^k, \quad \forall\, t \in [0,T]. \tag{2.23}$$

(See Définition IV.1 of Szpirglas [15]). The nice thing about (2.23) is that it includes reference to just one unbounded linear operator, namely the infinitesimal generator $\mathcal{A}$ of the signal process, and the resolvent identity can be used to eliminate $\mathcal{A}$ and re-write (2.23) in the form

$$\tilde{\sigma}_t\phi = \tilde{\sigma}_0(P_t\phi) + \sum_{k=1}^{d_1} \int_0^t \tilde{\sigma}_s(h^k P_{t-s}\phi) d\tilde{Y}_s^k, \quad \forall\, t \in [0,T], \tag{2.24}$$

where $\{P_t\}$ is the Borel semigroup with infinitesimal generator $\mathcal{A}$. There is complete equivalence between (2.23) and (2.24) in the sense that if the pair $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q}), (\tilde{\sigma}_t, \tilde{Y}_t)\}$ is subject to (a), (b), (c), then (2.23) holds for each $\phi \in \mathcal{D}(\mathcal{A})$ if and only if (2.24) holds for each $\phi \in B(E)$ (see Théorème IV.1 of [15]). Consequently, it is enough to establish pathwise uniqueness for (2.24) in order to conclude pathwise uniqueness for the unnormalized filter equation. The advantage of (2.24) is that it involves only the *bounded* linear operators $\{P_t\}$, and this structure makes it possible to establish pathwise uniqueness for solutions of (2.24) by iterating a simple integral inequality (see Section V.2 of Szpirglas [15]). Comparing (2.23) with the unnormalized filter equation (2.18) for the nonlinear filtering problem defined by (1.1) and (1.2), we see that (2.18) includes two unbounded linear operators, namely the first-order differential operator $\mathcal{B}_k$ which results from dependence of the signal $\{X_t\}$ on the Wiener process $\{W_t\}$ of the observation equation, as well as the second-order differential operator $\mathcal{A}$ corresponding to the signal process $\{X_t\}$. In this case there seems to be no clear way of adapting the elegant semigroup ideas of [15] to remove both of these unbounded operators and get an equivalent equation involving just bounded linear operators. Accordingly, the approach that we shall use to establish Theorem 2.21(i) is different from that of Szpirglas [15], and relies on a uniqueness theorem for measure-valued evolution equations (see Theorem 3.32 to follow).

Remark 2.23. Uniqueness for the normalized and unnormalized filter equations has also been studied by Bhatt, Kallianpur and Karandikar [1], Kurtz and Ocone [10], and Rozovskii [14] from a somewhat different point of view than that taken by Szpirglas [15] and the present work. To see this in the context of the filtering problem given by (1.1) and (1.2), observe from Remark 2.12 that the unnormalized optimal filter $\{\sigma_t\}$ solves the Duncan-Mortensen-Zakai equation, namely for each $\phi \in C_c^\infty(\mathbb{R}^d) \cup \{1\}$ we have

$$\sigma_t\phi = \pi_0\phi + \int_0^t \sigma_s(\mathcal{A}\phi)\, ds + \sum_{k=1}^{d_1} \int_0^t \sigma_s(h^k\phi + \mathcal{B}_k\phi)\, dY_s^k, \quad \forall\, t \in [0,T]. \tag{2.25}$$

With this in mind, the following question is natural: Suppose that $\{\rho_t\}$ is some $\mathcal{M}^+(\mathbb{R}^d)$-valued, cadlag, and $\{\mathcal{F}_t^Y\}$-adapted process on $(\Omega, \mathcal{F}, P)$, such that for each $\phi \in C_c^\infty(\mathbb{R}^d) \cup \{1\}$ we have

$$\rho_t\phi = \pi_0\phi + \int_0^t \rho_s(\mathcal{A}\phi)\, ds + \sum_{k=1}^{d_1} \int_0^t \rho_s(h^k\phi + \mathcal{B}_k\phi)\, dY_s^k, \quad \forall\, t \in [0,T]. \tag{2.26}$$

Does it follow that $\{\sigma_t\}$ and $\{\rho_t\}$ are indistinguishable? The works of Bhatt, Kallianpur and Karandikar ([1], Theorem 3.1), Kurtz and Ocone ([10], Theorems 4.2 and 4.7), and Rozovskii ([14], Theorem 3.1) provide conditions on the nonlinear filtering problem for which the answer is in the affirmative. Uniqueness in this sense is useful for the following reason: the observation process $\{Y_t\}$ is the random data that "drives" the unnormalized filter equation (2.25), and if we can "non-anticipatively" use the individual paths of $\{Y_t\}$ as data to compute a measure-valued process $\{\rho_t\}$ which satisfies (2.26) - e.g. by a numerical method - then uniqueness ensures that $\{\rho_t\}$ is in fact the desired unnormalized optimal filter $\{\sigma_t\}$. It should be noted that uniqueness in this sense can be established for much more general nonlinear filtering problems than that represented by the simple model (1.1) and (1.2). In fact, Theorem 3.1 of [1] deals with a filtering problem in which the signal process takes values in a complete separable metric space (not necessarily locally compact), the sensor function $h(\cdot)$ need not be uniformly bounded but only satisfy a finite-energy condition similar to (2.3), the dependence of the signal $\{X_t\}$ on the Wiener process $\{W_t\}$ is more general than that given by the explicit model (1.1), (1.2) (see (1.3) of [1]), and the joint signal/observation process $\{(X_t, Y_t)\}$ is the cadlag solution of a well-posed martingale problem.

The sense of pathwise uniqueness in the preceding paragraph is different from that established by Theorem 2.21(i), since the candidate solution $\{\rho_t\}$ of the filter equation (2.26) is postulated to be adapted specifically to the observation filtration $\{\mathcal{F}_t^Y\}$ (in fact, the arguments used in [1], [10], and [14] rely crucially on this restriction). In contrast, Theorem 2.21(i) establishes pathwise uniqueness in the more general sense of Definition 2.16, where the candidate solutions $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q}), (\tilde{\sigma}_t^1, \tilde{Y}_t)\}$ and $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q}), (\tilde{\sigma}_t^2, \tilde{Y}_t)\}$ are defined on an arbitrary filtered probability space, and there is no insistence that the measure-valued components $\{\tilde{\sigma}_t^1\}$ and $\{\tilde{\sigma}_t^2\}$ of the two solutions be adapted to the self-filtration of $\{\tilde{Y}_t\}$. The usefulness of this latter notion of pathwise uniqueness is that, by an adaptation to the filter equations of the construction of Yamada and Watanabe [16], it leads to uniqueness in law for the normalized and unnormalized filter equations (see Theorem 2.21(ii) and (iii), and recall Definitions 2.9 and 2.17). Uniqueness in law turns out to be essential for studying weak limits and approximations of the nonlinear filter equations by the method of martingale problems and weak convergence.

Remark 2.24. Using the method of stochastic flows and backward equations Kunita ([8], Theorem 6.2.8) and [7] establishes a form of pathwise uniqueness for the unnormalized filter equation. Rather restrictive boundedness and smoothness conditions on the coefficients of (1.1) and (1.2) seem necessary for this method to work. One should also note that the approach of Szpirglas [15], which is based on the equivalence of (2.23) and (2.24), finds its inspiration in an earlier study of pathwise uniqueness for the filter equations due to Kunita [6].

Remark 2.25. A basic property of Itô stochastic differential equations due to Yamada and Watanabe [16] is that pathwise uniqueness implies uniqueness in joint law, so that pathwise uniqueness is the stronger of the two uniqueness properties. It is shown in Szpirglas [15] that the basic Yamada-Watanabe argument extends to the measure-valued filter equations, so that pathwise uniqueness is again the stronger property (this is how we will conclude (ii) and (iii) from (i) in Theorem 2.21). Linearity of the unnormalized filter equation in fact implies the converse, so that for this equation the two uniqueness properties are actually equivalent:

Theorem 2.26. *Suppose that Conditions 2.18, 2.19, and 2.20 hold for the nonlinear filtering problem given by* **E.1**, **E.2** *and* **E.3**. *Then uniqueness in joint law implies pathwise uniqueness for the unnormalized filter equation.*

3 Proofs of Theorems 2.21 and 2.26:

The terminology in the next remark will be useful for dealing with measure-valued evolution equations:

Remark 3.27. Suppose that E is a complete separable metric space, and $\mathcal{Q} : \mathcal{D}(\mathcal{Q}) \to B(E)$ is a mapping with domain $\mathcal{D}(\mathcal{Q}) \subset B(E)$. Then $\{\mu_t,\ t \in [0, \infty)\}$ is an $\mathcal{M}^+(E)$-valued solution of the evolution equation for $(\mathcal{Q}, \mathcal{D}(\mathcal{Q}))$, when (i) $\mu_t \in \mathcal{M}^+(E)$, $\forall\, t \in [0, \infty)$, and $\mu_0 \in \mathcal{P}(E)$; (ii) for each $\Gamma \in \mathcal{B}(E)$, the mapping $t \to \mu_t(\Gamma) : [0, \infty) \to [0, \infty)$ is Borel-measurable; (iii) for each $f \in \mathcal{D}(\mathcal{Q})$ we have $\int_0^t |\mu_s(\mathcal{Q}f)|ds < \infty$, $\forall t \in [0, \infty)$, and

$$\mu_t f = \mu_0 f + \int_0^t \mu_s(\mathcal{Q}f)\, ds, \quad \forall t \in [0, \infty). \tag{3.27}$$

Moreover, $\{\mu_t,\ t \in [0, \infty)\}$ is a $\mathcal{P}(E)$-valued solution of the evolution equation for $(\mathcal{Q}, \mathcal{D}(\mathcal{Q}))$ when it is an $\mathcal{M}^+(E)$-valued solution with $\mu_t(E) = 1$, $\forall\, t \in [0, \infty)$. The evolution equation for $(\mathcal{Q}, \mathcal{D}(\mathcal{Q}))$ is said to have uniqueness in the class of $\mathcal{M}^+(E)$-valued solutions over the interval $[0, \infty)$ when, for any two such solutions $\{\mu_t^i,\ t \in [0, \infty)\}$, $i = 1, 2$, with $\mu_0^1 = \mu_0^2$, it follows that $\mu_t^1 = \mu_t^2, \quad \forall t \in [0, \infty)$. The notion of uniqueness within the class of $\mathcal{P}(E)$-valued solutions over the interval $[0, \infty)$ has an analogous formulation. Finally, the preceding terminology adapts in an obvious way to solutions $\{\mu_t,\ t \in [0, T]\}$ defined over the finite interval $[0, T]$.

Proof of Theorem 2.21(i) We shall need the following result, the proof of which is given in Section 5:

Fact 3.28. *Suppose that Conditions 2.18-2.20 hold and let* $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q}), (\tilde{\sigma}, \tilde{Y})\}$ *be a weak solution of the unnormalized filter equation. Then, for every* $\alpha \in (1, \infty)$ *there exists a constant* $\gamma(\alpha) \in [0, \infty)$ *such that*

$$\mathrm{E}^{\tilde{Q}} \left[\sup_{0 \le s \le T} |\tilde{\sigma}_s 1|^\alpha \right] \le \gamma(\alpha). \tag{3.28}$$

Now fix two weak solutions $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q}), (\tilde{\sigma}_t^i, \tilde{Y}_t)\}$, $i = 1, 2$, of the unnormalized filter equation, such that

$$\tilde{Q} \left[\tilde{\sigma}_0^1 = \tilde{\sigma}_0^2 \right] = 1, \tag{3.29}$$

and define product measures on $(\mathbb{R}^{2d}, \mathcal{B}(\mathbb{R}^{2d}))$ by

$$\mu_t^{12}(\cdot, \tilde{\omega}) := (\tilde{\sigma}_t^1 \times \tilde{\sigma}_t^2)(\cdot, \tilde{\omega}), \quad \forall (t, \tilde{\omega}) \in [0,T] \times \tilde{\Omega}.$$

A simple application of the Dynkin class theorem establishes

Fact 3.29. *For every $\Gamma \in \mathcal{B}(\mathbb{R}^{2d})$, the mapping $(t, \tilde{\omega}) \to \mu_t^{12}(\Gamma, \tilde{\omega}) : \tilde{\Omega} \times [0,T] \to [0,\infty)$ is measurable with respect to the $\{\tilde{\mathcal{F}}_t\}$-progressive σ-algebra.*

Also put

$$\nu_t^{12}(\Gamma) := \mathrm{E}^{\tilde{Q}}[\mu_t^{12}(\Gamma)], \quad \forall \Gamma \in \mathcal{B}(\mathbb{R}^{2d}),\ t \in [0,T]. \tag{3.30}$$

It readily follows that ν_t^{12} defines a positive measure on $(\mathbb{R}^{2d}, \mathcal{B}(\mathbb{R}^{2d}))$ for every $t \in [0,T]$. By Fact 3.28,

$$\begin{aligned} \nu_t^{12}(\mathbb{R}^{2d}) &= \mathrm{E}^{\tilde{Q}}[(\tilde{\sigma}_t^1 1)(\tilde{\sigma}_t^2 1)] \\ &\leq \left(\mathrm{E}^{\tilde{Q}}[\sup_{0\leq s\leq T} |\tilde{\sigma}_s^1 1|^2]\right)^{\frac{1}{2}} \left(\mathrm{E}^{\tilde{Q}}[\sup_{0\leq s\leq T} |\tilde{\sigma}_s^2 1|^2]\right)^{\frac{1}{2}} \leq \gamma(2), \quad \forall t \in [0,T]. \end{aligned} \tag{3.31}$$

This shows that ν_t^{12} is a positive measure on $(\mathbb{R}^{2d}, \mathcal{B}(\mathbb{R}^{2d})$, uniformly bounded with respect to $t \in [0,T]$, while Fact 3.29 with Fubini's theorem shows that the mapping $t \to \nu_t^{12}(\Gamma) : [0,T] \to \mathbb{R}$ is Borel-measurable for each $\Gamma \in \mathcal{B}(\mathbb{R}^{2d})$. Next, define $\nu_t^{11}, \nu_t^{22} \in \mathcal{M}^+(\mathbb{R}^{2d})$, $t \in [0,T]$, analogously to ν_t^{12}, by

$$\nu_t^{ii}(\Gamma) := \mathrm{E}^{\tilde{Q}}[(\tilde{\sigma}_t^i \times \tilde{\sigma}_t^i)(\Gamma)], \quad \forall \Gamma \in \mathcal{B}(\mathbb{R}^{2d}),\ t \in [0,T],\ i = 1, 2. \tag{3.32}$$

In the same way as for $\nu_\cdot^{12}$, we see that $\nu_\cdot^{ii}$ are positive measures on $(\mathbb{R}^{2d}, \mathcal{B}(\mathbb{R}^{2d}))$, uniformly bounded with respect to $t \in [0,T]$, and the mappings $t \to \nu_t^{ii}(\Gamma) : [0,T] \to \mathbb{R}$ are Borel-measurable for each $\Gamma \in \mathcal{B}(\mathbb{R}^{2d})$, $i = 1, 2$.

Remark 3.30. For mappings $f_1, f_2 \in B(\mathbb{R}^d)$ define the tensor product of f_1 with f_2 to be the mapping $f_1 \otimes f_2 : \mathbb{R}^{2d} \to \mathbb{R}$ given by

$$f_1 \otimes f_2(x_1, x_2) := f_1(x_1) f_2(x_2), \quad \forall\, x_1, x_2 \in \mathbb{R}^d.$$

In view of (3.30) and (3.32), for each $f_1, f_2 \in B(\mathbb{R}^d)$ we have

$$\nu_t^{12}(f_1 \otimes f_2) = \mathrm{E}^{\tilde{Q}}[(\tilde{\sigma}_t^1 f_1)(\tilde{\sigma}_t^2 f_2)], \tag{3.33}$$

$$\nu_t^{ii}(f_1 \otimes f_2) = \mathrm{E}^{\tilde{Q}}[(\tilde{\sigma}_t^i f_1)(\tilde{\sigma}_t^i f_2)], \quad i = 1, 2. \tag{3.34}$$

From (3.29), (3.30), and (3.32) we see that

$$\nu_0^{11}, \nu_0^{22} \text{ and } \nu_0^{12} \text{ are probability measures on } \mathcal{B}(\mathbb{R}^{2d}) \text{ and } \nu_0^{11} = \nu_0^{22} = \nu_0^{12}. \tag{3.35}$$

Using this fact, we shall establish

$$\nu_t^{11} = \nu_t^{22} = \nu_t^{12}, \quad \forall t \in [0,T], \tag{3.36}$$

from which pathwise uniqueness follows. Indeed, if (3.36) holds, then for each $f \in B(\mathbb{R}^d)$ we have

$$\nu_t^{11}(f \otimes f) = \nu_t^{22}(f \otimes f) = \nu_t^{12}(f \otimes f), \quad \forall t \in [0,T],$$

and therefore from (3.33) and (3.34),

$$\begin{aligned} \mathrm{E}^{\tilde{Q}}[(\tilde{\sigma}_t^1 f - \tilde{\sigma}_t^2 f)^2] &= \mathrm{E}^{\tilde{Q}}[(\tilde{\sigma}_t^1 f)(\tilde{\sigma}_t^1 f)] - 2\,\mathrm{E}^{\tilde{Q}}[(\tilde{\sigma}_t^1 f)(\tilde{\sigma}_t^2 f)] + \mathrm{E}^{\tilde{Q}}[(\tilde{\sigma}_t^2 f)(\tilde{\sigma}_t^2 f)] \\ &= \nu_t^{11}(f \otimes f) - 2\nu_t^{12}(f \otimes f) + \nu_t^{22}(f \otimes f) = 0. \end{aligned}$$

Thus, for each $t \in [0,T]$ and $f \in B(\mathbb{R}^d)$, we have

$$\tilde{Q}\left[\tilde{\sigma}_t^1 f = \tilde{\sigma}_t^2 f\right] = 1. \tag{3.37}$$

Now $\hat{C}(\mathbb{R}^d)$ equipped with the supremum norm $\|\cdot\|$ is separable. Thus, from (3.37), for each $t \in [0,T]$ there is a $\tilde{Q}$-null event $N_t \in \tilde{\mathcal{F}}$ such that, for each $\tilde{\omega} \notin N_t$, we have

$$\tilde{\sigma}_t^1(\tilde{\omega}) f = \tilde{\sigma}_t^2(\tilde{\omega}) f, \quad \forall\, f \in \hat{C}(\mathbb{R}^d). \tag{3.38}$$

But $\hat{C}(\mathbb{R}^d)$ separates bounded positive measures on $\mathcal{B}(\mathbb{R}^d)$ (see Problem 5.4.25 of Karatzas and Shreve [5]), thus (3.38) establishes $\tilde{Q}[\tilde{\sigma}_t^1 = \tilde{\sigma}_t^2] = 1$ for each $t \in [0,T]$. Now Theorem 2.21(i) follows from the fact that $\{\tilde{\sigma}_t^i,\ t \in [0,T]\}$ are continuous (recall Definition 2.15).

It therefore remains to establish (3.36) in order to prove Theorem 2.21(i). To this end, for each $x_1, x_2 \in \mathbb{R}^d$ define the $2d \times 2d$ matrix $\bar{a}(x_1,x_2)$, the $2d$ vector $\bar{b}(x_1,x_2)$, and the real number $\bar{h}(x_1,x_2)$ by

$$\bar{a}(x_1,x_2) := \begin{bmatrix} cc^T(x_1) & 0 \\ 0 & cc^T(x_2) \end{bmatrix} + \begin{bmatrix} B(x_1) \\ B(x_2) \end{bmatrix} \begin{bmatrix} B^T(x_1) & B^T(x_2) \end{bmatrix} \tag{3.39a}$$

$$\bar{b}(x_1,x_2) := \begin{bmatrix} b(x_1) + B(x_1)h(x_2) \\ b(x_2) + B(x_2)h(x_1) \end{bmatrix} \tag{3.39b}$$

$$\bar{h}(x_1,x_2) := \sum_{k=1}^{d_1} h^k(x_1)h^k(x_2). \tag{3.39c}$$

Observe that the matrix $\bar{a}(x_1,x_2)$ is symmetric and strictly positive-definite (see Condition 2.19), and let $\bar{\mathcal{A}}$ be the second order linear differential operator corresponding to the matrices $\bar{a}$ and $\bar{b}$, namely

$$\bar{\mathcal{A}}\phi(x) := \sum_{i=1}^{2d} \bar{b}^i(x)\partial_i\phi(x) + \frac{1}{2}\sum_{i,j=1}^{2d} \bar{a}^{ij}(x)\partial_i\partial_j\phi(x), \quad \forall\, \phi \in C^\infty(\mathbb{R}^{2d}),\ x \in \mathbb{R}^{2d}. \tag{3.40}$$

From (3.39a), (3.39b), Condition 2.18, and Condition 2.20, there is a constant $K \in [0,\infty)$ such that

$$\max_i |\bar{b}^i(x)| \le K[1+|x|], \quad \max_{i,j} |\bar{a}^{ij}(x)| \le K[1+|x|^2], \quad \forall x \in \mathbb{R}^{2d}, \tag{3.41}$$

and the operator $\bar{\mathcal{A}}$ has the following property, which is established in Section 4:

Lemma 3.31. *Suppose that Conditions 2.18–2.20 hold. Then* $\{\nu_t^{11},\ t\in[0,T]\}$, $\{\nu_t^{12},\ t\in[0,T]\}$, *and* $\{\nu_t^{22},\ t\in[0,T]\}$, *given by* (3.30) *and* (3.32), *are* $\mathcal{M}^+(\mathbb{R}^{2d})$*-valued solutions of the evolution equation for* $(\bar{\mathcal{A}}+\bar{h}, \operatorname{span}\{1, C_c^\infty(\mathbb{R}^{2d})\})$.

It remains to show that the evolution equation for $(\bar{\mathcal{A}}+\bar{h}, \operatorname{span}\{1, C_c^\infty(\mathbb{R}^{2d})\})$ has uniqueness in the class of $\mathcal{M}^+(\mathbb{R}^{2d})$-valued solutions over the interval $[0,T]$, since this fact, along with (3.35) and Lemma 3.31, gives (3.36), as required to establish Theorem 2.21(i). To this end we need the following result on uniqueness of measure-valued solutions of the evolution equation corresponding to a multiplicatively perturbed linear second-order differential operator on Euclidean space:

Theorem 3.32. *Let* $\mathcal{C}$ *be the linear second-order differential operator on the finite-dimensional Euclidean space* $\mathbb{R}^q$ *defined by*

$$\mathcal{D}(\mathcal{C}) := \operatorname{span}\{1, C_c^\infty(\mathbb{R}^q)\}; \tag{3.42a}$$

$$\mathcal{C}f(x) := \sum_i \beta^i(x)\partial_i f(x) + \frac{1}{2}\sum_{i,j}\alpha^{ij}(x)\partial_i\partial_j f(x), \quad \forall\, x\in\mathbb{R}^q,\ \forall\, f\in\mathcal{D}(\mathcal{C}), \tag{3.42b}$$

where $\beta:\mathbb{R}^q\to\mathbb{R}^q$ *is Borel measurable,* $\alpha:\mathbb{R}^q\to\mathbb{S}_{++}^{q\times q}$ *is continuous, and there exists a constant* $K\in[0,\infty)$ *such that*

$$|\beta^i(x)|\le K(1+|x|),\quad |\alpha^{ij}(x)|\le K(1+|x|^2),\quad \forall x\in\mathbb{R}^q. \tag{3.43}$$

If $\lambda\in B(\mathbb{R}^q)$ *then the evolution equation for* $(\mathcal{C}-\lambda, \mathcal{D}(\mathcal{C}))$ *has uniqueness in the class of* $\mathcal{M}^+(\mathbb{R}^q)$*-valued solutions over the interval* $[0,T]$.

To complete the proof of Theorem 2.21(i) we note from (3.39) that $\bar{a}(\cdot)$ is continuous on $\mathbb{R}^{2d}$, $\bar{b}(\cdot)$ is Borel-measurable on $\mathbb{R}^{2d}$, and $\bar{h}\in B(\mathbb{R}^{2d})$. That the evolution equation for $(\bar{\mathcal{A}}+\bar{h}, \operatorname{span}\{1, C_c^\infty(\mathbb{R}^{2d})\})$ has uniqueness in the class of $\mathcal{M}^+(\mathbb{R}^{2d})$-valued solutions over the interval $[0,T]$ now follows from (3.41) and Theorem 3.32 with $q:=2d$, $\beta(\cdot):=\bar{b}(\cdot)$, $\alpha(\cdot):=\bar{a}(\cdot)$, and $\lambda(\cdot):=-\bar{h}(\cdot)$. □

Remark 3.33. When $\beta(\cdot)$ and $\lambda(\cdot)$ in Theorem 3.32 are *continuous* then $\mathcal{C}$ is a linear operator on $\bar{C}(\mathbb{R}^q)$, and Theorem 3.32 is just a very special consequence of a general theorem of Bhatt and Karandikar (see Theorem 3.4 and Remark 1 of [2]) on uniqueness of measure-valued solutions of perturbed evolution equations. However, when $\beta(\cdot)$ and $\lambda(\cdot)$ are only Borel-measurable, then $\mathcal{C}f(\cdot)$ is not continuous for $f\in\mathcal{D}(\mathcal{C})$, and we cannot directly use the result of [2]. We prove Theorem 3.32 in Section 6.

Remark 3.34. The proof just given for Theorem 2.21(i) relies on the special structure of the unnormalized filter equation (2.18) and does not appear to extend to the normalized filter equation (2.10). We have therefore not been able to establish pathwise uniqueness in the sense of Definition 2.7 under conditions comparable to those of Theorem 2.21.

Proof of Theorem 2.21(ii): Let $\{(\tilde{\Omega},\tilde{\mathcal{F}},\{\tilde{\mathcal{F}}_t\},\tilde{P}),(\tilde{\pi}_t,\tilde{I}_t)\}$ and $\{(\bar{\Omega},\bar{\mathcal{F}},\{\bar{\mathcal{F}}_t\},\bar{P}),(\bar{\pi}_t,\bar{I}_t)\}$ be two weak solutions of the normalized filter equation. By an argument similar to that used for

Proposition IX.1.4 of Revuz and Yor [12], to establish uniqueness in joint law it is enough to show that the processes $\{(\tilde{\pi}_t, \tilde{I}_t)\}$ and $\{(\bar{\pi}_t, \bar{I}_t)\}$ are identically distributed when

$$\tilde{\pi}_0 = \bar{\pi}_0 = \mu, \qquad \text{for each } \mu \in \mathcal{P}(\mathbb{R}^d). \tag{3.44}$$

Thus suppose (3.44) holds for some $\mu \in \mathcal{P}(\mathbb{R}^d)$. Put

$$\tilde{\chi}_t := \mathcal{E}(-\sum_{k=1}^{d_1}(\tilde{\pi}h^k) \bullet \tilde{I}^k)_t \qquad \text{and} \qquad \bar{\chi}_t := \mathcal{E}(-\sum_{k=1}^{d_1}(\bar{\pi}h^k) \bullet \bar{I}^k)_t, \quad \forall t \in [0,T],$$

and define the measures $\tilde{Q}$ and $\bar{Q}$ on the measurable spaces $(\tilde{\Omega}, \tilde{\mathcal{F}})$ and $(\bar{\Omega}, \bar{\mathcal{F}})$ respectively by

$$\tilde{Q}(A) := \mathrm{E}^{\tilde{P}}[\tilde{\chi}_T; A], \quad \forall A \in \tilde{\mathcal{F}}, \tag{3.45}$$

$$\bar{Q}(A) := \mathrm{E}^{\bar{P}}[\bar{\chi}_T; A], \quad \forall A \in \bar{\mathcal{F}}. \tag{3.46}$$

Then, with

$$\tilde{Y}_t := \tilde{I}_t + \sum_{k=1}^{d_1}\int_0^t \tilde{\pi}_u h^k \, du, \qquad \bar{Y}_t := \bar{I}_t + \sum_{k=1}^{d_1}\int_0^t \bar{\pi}_u h^k \, du, \quad t \in [0,T], \tag{3.47}$$

and

$$\tilde{\sigma}_t := \tilde{\pi}_t/\tilde{\chi}_t, \qquad \bar{\sigma}_t := \bar{\pi}_t/\bar{\chi}_t, \quad \forall t \in [0,T], \tag{3.48}$$

we see, as in Remark 2.10 and Remark 2.13, that the pairs $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q}), (\tilde{\sigma}_t, \tilde{Y}_t)\}$ and $\{(\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{Q}), (\bar{\sigma}_t, \bar{Y}_t)\}$ are weak solutions of the unnormalized filter equation, with

$$\tilde{\sigma}_0 = \bar{\sigma}_0 = \mu.$$

For a complete separable metric space E, let $C_E[0,T]$ denote the complete separable metric space of all continuous mappings from $[0,T]$ into E with the usual metric giving uniform convergence over $[0,T]$. Define

$$\hat{\Omega} := C_{\mathcal{M}^+(\mathbb{R}^d)}[0,T] \times C_{\mathcal{M}^+(\mathbb{R}^d)}[0,T] \times C_{\mathbb{R}^{d_1}}[0,T],$$

which is a complete separable metric space with the usual product metric, and let $\hat{\omega} = (\omega^1, \omega^2, \omega^3)$ be a generic member of $\hat{\Omega}$. By the Yamada-Watanabe construction (see Theorem IV.1.1 of Ikeda and Watanabe [4]), there exists $\hat{P} \in \mathcal{P}(\hat{\Omega})$ such that

YW.1: $\mathcal{L}_{\hat{P}}(\omega^1, \omega^3) = \mathcal{L}_{\tilde{Q}}(\tilde{\sigma}, \tilde{Y})$;

YW.2: $\mathcal{L}_{\hat{P}}(\omega^2, \omega^3) = \mathcal{L}_{\bar{Q}}(\bar{\sigma}, \bar{Y})$;

YW.3: If $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$ is the completion of $(\hat{\Omega}, \mathcal{B}(\hat{\Omega}), \hat{P})$, and $\hat{\mathcal{F}}_t$ is the augmentation of the σ-algebra $\sigma\{\hat{\omega}(s),\ s \in [0,t]\}$ with the null events of $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$, $\forall\, t \in [0,T]$, then $\{\omega^3_t,\ t \in [0,T]\}$ is a $\{\hat{\mathcal{F}}_t\}$-Wiener process on $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$.

From (YW.1), (YW.2), and (YW.3), along with Exercise IV.5.16 of Revuz and Yor [12], it follows that $\{(\hat{\Omega}, \hat{\mathcal{F}}, \{\hat{\mathcal{F}}_t\}, \hat{P}), (\omega^1, \omega^3)\}$ and $\{(\hat{\Omega}, \hat{\mathcal{F}}, \{\hat{\mathcal{F}}_t\}, \hat{P}), (\omega^2, \omega^3)\}$ are weak solutions for the unnormalized filter equation with

$$\omega_0^1 = \omega_0^2 = \mu,$$

and hence, from Theorem 2.21(i),

$$\hat{P}(\omega_t^1 = \omega_t^2 \quad \forall t \in [0, T]) = 1. \tag{3.49}$$

From (3.48) we see that

$$\tilde{\pi}_t \phi = (\tilde{\sigma}_t \phi)(\tilde{\sigma}_t 1), \quad \forall\, t \in [0, T], \quad \phi \in B(\mathbb{R}^d), \tag{3.50}$$

and so, from (3.47),

$$\tilde{I}_t^k = \tilde{Y}_t^k - \sum_{k=1}^{d_1} \int_0^t (\tilde{\sigma}_u 1)(\tilde{\sigma}_u h^k)\, du, \quad \forall\, k = 1, 2, \ldots, d_1, \quad t \in [0, T]. \tag{3.51}$$

From (3.50) and (3.51) there exists a measurable mapping $\Phi : C_{\mathcal{M}^+(\mathbb{R}^d)}[0, T] \times C_{\mathbb{R}^{d_1}}[0, T] \to C_{\mathcal{P}(\mathbb{R}^d)}[0, T] \times C_{\mathbb{R}^{d_1}}[0, T]$ such that

$$(\tilde{\pi}, \tilde{I}) = \Phi(\tilde{\sigma}, \tilde{Y}). \tag{3.52}$$

Now (3.50) and (3.51) continue to hold with "overbar" in place of "tilde", and hence

$$(\bar{\pi}, \bar{I}) = \Phi(\bar{\sigma}, \bar{Y}). \tag{3.53}$$

Thus, for each $\Gamma \in \mathcal{B}(C_{\mathcal{P}(\mathbb{R}^d)}[0, T] \times C_{\mathbb{R}^{d_1}}[0, T])$, we see from (3.45), (3.52), and (YW.1), that

$$\tilde{P}((\tilde{\pi}, \tilde{I}) \in \Gamma) = \mathrm{E}^{\tilde{Q}}[(\tilde{\sigma}_T 1)^{-1} I_\Gamma(\Phi(\tilde{\sigma}, \tilde{Y}))] = \mathrm{E}^{\hat{P}}[(\omega_T^1 1)^{-1} I_\Gamma(\Phi(\omega^1, \omega^3))], \tag{3.54}$$

and, from (3.46), (3.53), and (YW.2), we similarly have

$$\bar{P}((\bar{\pi}, \bar{I}) \in \Gamma) = \mathrm{E}^{\bar{Q}}[(\bar{\sigma}_T 1)^{-1} I_\Gamma(\Phi(\bar{\sigma}, \bar{Y}))] = \mathrm{E}^{\hat{P}}[(\omega_T^2 1)^{-1} I_\Gamma(\Phi(\omega^2, \omega^3))]. \tag{3.55}$$

Now (3.49), (3.54), and (3.55) show that $\tilde{P}((\tilde{\pi}, \tilde{I}) \in \Gamma) = \bar{P}((\bar{\pi}, \bar{I}) \in \Gamma)$, as required.

Proof of Theorem 2.21(iii): The proof is an obvious simplification of the proof of Theorem 2.21(ii) and is omitted. □

Proof of Theorem 2.26: Let $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q}), (\tilde{\sigma}_t^i, \tilde{Y}_t)\}$, $i = 1, 2$ be two weak solutions of the unnormalized filter equation. Define

$$\tilde{\sigma}_t^3(\cdot) := \frac{\tilde{\sigma}_t^1(\cdot) + \tilde{\sigma}_t^2(\cdot)}{2}, \quad t \in [0, T].$$

It then follows that $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q}), (\tilde{\sigma}^3_t, \tilde{Y}_t)\}$ is a weak solution of the unnormalized filter equation. Therefore, the postulated uniqueness in joint law together with Fact 3.28 implies that for an arbitrary $\phi \in C^\infty_c(\mathbb{R}^d) \cup \{1\}$ we have

$$2\,\mathrm{E}^{\tilde{Q}}\left[\left(\frac{\tilde{\sigma}^1_t\phi + \tilde{\sigma}^2_t\phi}{2}\right)^2\right] - \mathrm{E}^{\tilde{Q}}[(\tilde{\sigma}^1_t\phi)^2] - \mathrm{E}^{\tilde{Q}}[(\tilde{\sigma}^2_t\phi)^2] = 0, \quad \forall t \in [0,T].$$

Rearanging this expression gives

$$\mathrm{E}^{\tilde{Q}}[(\tilde{\sigma}^1_t\phi - \tilde{\sigma}^2_t\phi)^2] = 0, \quad \forall t \in [0,T]. \tag{3.56}$$

Since $\hat{C}(\mathbb{R}^d)$ is separable (in the supremum norm), it follows that $C^\infty_c(\mathbb{R}^d)$ is likewise separable, and therefore, by Problem 5.4.25 of Karatzas and Shreve [5], there is a countable determining set for $\mathcal{M}^+(\mathbb{R}^d)$ in $C^\infty_c(\mathbb{R}^d) \cup \{1\}$. Then (3.56) shows that $\{\tilde{\sigma}^1_t\}$ and $\{\tilde{\sigma}^2_t\}$ are modifications of each other, hence indistinguishable (since $\{\tilde{\sigma}^1_t\}$ and $\{\tilde{\sigma}^2_t\}$ are continuous). □

4 Proof of Lemma 3.31

For arbitrary $f_1, f_2 \in C^\infty_c(\mathbb{R}^d)$ put

$$\begin{aligned}\tilde{\mathcal{A}}(f_1 \otimes f_2) := f_1 \otimes (\mathcal{A}f_2) + (\mathcal{A}f_1) \otimes f_2 + \sum_{k=1}^{d_1}\big[&(h^k f_1) \otimes (h^k f_2) + (h^k f_1) \otimes (\mathcal{B}_k f_2) \\ &+ (\mathcal{B}_k f_1) \otimes (h^k f_2) + (\mathcal{B}_k f_1) \otimes (\mathcal{B}_k f_2)\big],\end{aligned} \tag{4.57}$$

where $\mathcal{A}$ and $\mathcal{B}_k$ are given by (2.7). We need the following lemma, the proof of which is given in Section 5:

Lemma 4.35. *Suppose Conditions 2.18–2.20 hold, let $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q}), (\tilde{\sigma}^i_t, \tilde{Y}_t)\}$, $i = 1, 2$, be weak solutions of the unnormalized filter equation, and define the $\mathcal{M}^+(\mathbb{R}^{2d})$-valued functions $\{\nu^{12}_t,\ t \in [0,T]\}$, $\{\nu^{11}_t,\ t \in [0,T]\}$, and $\{\nu^{22}_t,\ t \in [0,T]\}$, as in* (3.30) *and* (3.32)*. Then, for each $f_1, f_2 \in C^\infty_c(\mathbb{R}^d)$, we have*

$$\nu^{12}_t(f_1 \otimes f_2) = \nu^{12}_0(f_1 \otimes f_2) + \int_0^t \nu^{12}_u(\tilde{\mathcal{A}}(f_1 \otimes f_2))\,du, \quad \forall\, t \in [0,T], \tag{4.58}$$

with identical relations for $\nu^{11}_\cdot$ and $\nu^{22}_\cdot$ in place of $\nu^{12}_\cdot$.

By direct evaluation of the right-hand side of (4.57) it is easy to establish

Lemma 4.36. *For $\tilde{\mathcal{A}}$ and $\bar{\mathcal{A}}$ defined in* (4.57) *and* (3.40) *respectively, we have*

$$\tilde{\mathcal{A}}(f_1 \otimes f_2)(x) = \bar{\mathcal{A}}(f_1 \otimes f_2)(x) + \bar{h}(x)(f_1 \otimes f_2)(x), \quad \forall\, x \in \mathbb{R}^{2d}, \tag{4.59}$$

for each $f_1, f_2 \in C^\infty_c(\mathbb{R}^d)$.

Define

$$\tilde{\mathcal{D}} := \text{span}\{f_1 \otimes f_2 \,:\, f_1, f_2 \in C_c^\infty(\mathbb{R}^d)\}. \tag{4.60}$$

Putting Lemma 4.36 and Lemma 4.35 together, we see that the mappings $\{\nu_t^{12},\ t \in [0,T]\}$, $\{\nu_t^{11},\ t \in [0,T]\}$, and $\{\nu_t^{22},\ t \in [0,T]\}$, defined at (3.30) and (3.32), are $\mathcal{M}^+(\mathbb{R}^{2d})$-valued solutions of the evolution equation for $(\bar{\mathcal{A}} + \bar{h}, \tilde{\mathcal{D}})$, that is,

$$\nu_t^{12} f = \nu_0^{12} f + \int_0^t \nu_u^{12}(\bar{\mathcal{A}} f + \bar{h} f)\, du, \quad \forall\, t \in [0,T],\ \forall f \in \tilde{\mathcal{D}}, \tag{4.61}$$

with identical relations for $\nu_\cdot^{11}$ and $\nu_\cdot^{22}$ in place of $\nu_\cdot^{12}$. In order to prove Lemma 3.31, it remains to show that (4.61) holds not only for $f \in \tilde{\mathcal{D}}$, but for each f in the larger domain $\text{span}\{1, C_c^\infty(\mathbb{R}^{2d})\}$. That is, it must be shown that the mappings $\{\nu_t^{12},\ t \in [0,T]\}$, $\{\nu_t^{11},\ t \in [0,T]\}$, and $\{\nu_t^{22},\ t \in [0,T]\}$ are $\mathcal{M}^+(\mathbb{R}^{2d})$-valued solutions of the evolution equation for $(\bar{\mathcal{A}} + \bar{h}, \text{span}\{1, C_c^\infty(\mathbb{R}^{2d})\})$. To this end we need the following result, whose proof is deferred to Section 5:

Lemma 4.37. *Suppose Conditions 2.18–2.20 hold. Then the closure of the relation $\{(f, \bar{\mathcal{A}} f) : f \in \tilde{\mathcal{D}}\}$ in the supremum norm of $B(\mathbb{R}^{2d}) \times B(\mathbb{R}^{2d})$ contains the relation $\{(f, \bar{\mathcal{A}} f) : f \in C_c^\infty(\mathbb{R}^{2d})\}$.*

From Lemma 4.37 and the notions of bp-closedness and bp-closure of a relation (see foot of page 111 of Ethier and Kurtz [3]), we see that

$$\{(f, \bar{\mathcal{A}} f) :\ f \in C_c^\infty(\mathbb{R}^{2d})\} \subset \text{bp-closure}\{(f, \bar{\mathcal{A}} f) :\ f \in \tilde{\mathcal{D}}\}. \tag{4.62}$$

Now put

$$S^{12} := \left\{(f,g) \in B(\mathbb{R}^{2d}) \times B(\mathbb{R}^{2d}) \,:\, \nu_t^{12} f = \nu_0^{12} f + \int_0^t \nu_s^{12}(g + \bar{h} f)\, ds,\ \forall t \in [0,T]\right\}, \tag{4.63}$$

and observe that S^{12} is a linear relation. By (3.31) we have

$$\sup_{0 \le t \le T} \nu_t^{12}(\mathbb{R}^{2d}) < \infty,$$

and therefore, since $\bar{h} \in B(\mathbb{R}^{2d})$, it follows from the dominated convergence theorem that the linear relation S^{12} is bp-closed in $B(\mathbb{R}^{2d}) \times B(\mathbb{R}^{2d})$. Since the $\mathcal{M}^+(\mathbb{R}^{2d})$-valued mapping $\{\nu_t^{12},\ t \in [0,T]\}$ solves the evolution equation for $(\bar{\mathcal{A}} + \bar{h}, \tilde{\mathcal{D}})$, we have $\{(f, \bar{\mathcal{A}} f) :\ f \in \tilde{\mathcal{D}}\} \subset S^{12}$, and therefore, from the bp-closedness of S^{12} and (4.62), we have

$$\{(f, \bar{\mathcal{A}} f) :\ f \in C_c^\infty(\mathbb{R}^{2d})\} \subset S^{12}. \tag{4.64}$$

Next, observe from (3.41) and Problem 4.11.12 of Ethier and Kurtz [3] that the operator $(\bar{\mathcal{A}}, C_c^\infty(\mathbb{R}^{2d}))$ is conservative, and hence (see page 166 of [3]) we have

$$(1,0) \in \text{bp-closure}\{(f, \bar{\mathcal{A}} f) :\ f \in C_c^\infty(\mathbb{R}^{2d})\}. \tag{4.65}$$

In the light of (4.65), (4.64), and the bp-closedness of S^{12}, we then get

$$(1,0) \in S^{12}. \tag{4.66}$$

Now (4.64), (4.66), and linearity of the relation S^{12} shows that

$$\{(f, \bar{\mathcal{A}}f) : f \in \mathrm{span}\{1, C_c^\infty(\mathbb{R}^{2d})\}\} \subset S^{12},$$

which, in view of (4.63), shows that $\{\nu_t^{12},\ t \in [0,T]\}$ is an $\mathcal{M}^+(\mathbb{R}^{2d})$-valued solution of the evolution equation for $(\bar{\mathcal{A}} + \bar{h}, \mathrm{span}\{1, C_c^\infty(\mathbb{R}^{2d})\})$. Defining S^{11} as in (4.63), but with ν^{11} in place of ν^{12}, we can similarly show that $\{\nu_t^{11},\ t \in [0,T]\}$ is an $\mathcal{M}^+(\mathbb{R}^{2d})$-valued solution of the evolution equation for $(\bar{\mathcal{A}} + \bar{h}, \mathrm{span}\{1, C_c^\infty(\mathbb{R}^{2d})\})$, and likewise for $\{\nu_t^{22},\ t \in [0,T]\}$. □

5 Proofs of Technical Results

Proof of Fact 3.28: Fix some $\alpha \in (1,\infty)$. Since $\tilde{\sigma}_0$ takes values in $\mathcal{P}(\mathbb{R}^d)$, we see from (2.18) with $\phi \equiv 1$ that

$$\tilde{\sigma}_t 1 = 1 + \int_0^t \sum_{k=1}^{d_1} \left(\frac{\tilde{\sigma}_s h^k}{\tilde{\sigma}_s 1}\right) (\tilde{\sigma}_s 1)\, d\tilde{Y}_s^k.$$

This gives (see Exercise IV.3.10(1) of Revuz and Yor [12])

$$\tilde{\sigma}_t 1 = \mathcal{E}\left(\sum_{k=1}^{d_1} \left(\frac{\tilde{\sigma} h^k}{\tilde{\sigma} 1} \bullet \tilde{Y}^k\right)\right)_t, \tag{5.67}$$

and hence

$$\begin{aligned} |\tilde{\sigma}_t 1|^\alpha &= \tilde{M}_t \exp\left(\frac{\alpha(\alpha-1)}{2} \sum_{k=1}^{d_1} \int_0^t \left|\frac{\tilde{\sigma}_s h^k}{\tilde{\sigma}_s 1}\right|^2 ds\right) \\ &\le \tilde{M}_t \exp\left(\frac{\alpha(\alpha-1)}{2} \|h\| T\right), \quad \forall\, t \in [0,T], \end{aligned} \tag{5.68}$$

for

$$\tilde{M}_t := \mathcal{E}\left(\alpha \sum_{k=1}^{d_1} \left(\frac{\tilde{\sigma} h^k}{\tilde{\sigma} 1} \bullet \tilde{Y}^k\right)\right)_t \qquad \text{and} \qquad \|h\| := \sup_{x \in \mathbb{R}^d} |h(x)|.$$

Now Condition 2.20 ensures that the processes $\{(\tilde{\sigma}_t h^k)/(\tilde{\sigma}_t 1),\ t \in [0,T]\}$ are uniformly bounded (by $\|h^k\|$), and therefore $\{(\tilde{M}_t, \tilde{\mathcal{F}}_t),\ t \in [0,T]\}$ is a continuous martingale on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{Q})$, with $\tilde{M}_0 = 1$. Taking $\tilde{Q}$-expectations in (5.68) then gives

$$\mathrm{E}^{\tilde{Q}}[|\tilde{\sigma}_t 1|^\alpha] \le \exp\left(\frac{\alpha(\alpha-1)}{2} \|h\| T\right), \quad \forall\, t \in [0,T]. \tag{5.69}$$

Again, by (5.67) and uniform-boundedness of the processes $\{(\tilde{\sigma}_t h^k)/(\tilde{\sigma}_t 1),\ t \in [0,T]\}$ we see that $\{(\tilde{\sigma}_t 1, \tilde{\mathcal{F}}_t),\ t \in [0,T]\}$ is a continuous martingale on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{Q})$, which, in light of (5.69), is L_α-bounded. Thus, by Doob's inequality, there is some $\gamma(\alpha) \in (0, \infty)$ such that (3.28) holds. □

Proof of Lemma 4.35: Fix $f_1, f_2 \in C_c^\infty(\mathbb{R}^d)$. Since $\{\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{Q}), (\tilde{\sigma}_t^i, \tilde{Y}_t)\}$, $i = 1, 2$ are weak solutions of the unnormalized filter equations, we have

$$\tilde{\sigma}_t^i f_i = \tilde{\sigma}_0^i f_i + \int_0^t \tilde{\sigma}_u^i(\mathcal{A} f_i)\, du + \sum_k \int_0^t \tilde{\sigma}_u^i(h^k f_i + \mathcal{B}_k f_i)\, d\tilde{Y}_u^k, \quad i = 1, 2. \tag{5.70}$$

Expand the product of semimartingales $\{(\tilde{\sigma}_t^1 f_1)(\tilde{\sigma}_t^2 f_2)\}$ by Itô's formula, and note that each stochastic integral in this expansion is with respect to a $\tilde{Q}$-standard Wiener process $\{(\tilde{Y}_t^k, \tilde{\mathcal{F}}_t)\}$, $k = 1, 2, \ldots, d_1$. Fact 3.28 ensures that these stochastic integrals are genuine $\tilde{Q}$-martingales, and therefore have $\tilde{Q}$-expectation identically zero. Upon taking $\tilde{Q}$-expectations on each side of the resulting expansion of $\{(\tilde{\sigma}_t^1 f_1)(\tilde{\sigma}_t^2 f_2)\}$ and using (4.57) and (3.33), we get (4.58). The corresponding identities for $\nu_\cdot^{11}$ and $\nu_\cdot^{22}$ are similarly obtained. □

Proof of Lemma 4.37: Fix arbitrary $\epsilon \in (0, \infty)$ and $g \in C_c^\infty(\mathbb{R}^{2d})$. Put

$$B_R := \{x \in \mathbb{R}^{2d} \;:\; |x| \leq R\}, \quad R \in [0, \infty),$$

and fix R such that $\operatorname{supp}(g) \subset B_R$. Also fix some $q \in C_c^\infty(\mathbb{R}^d)$ such that

$$\|q\| \leq 1; \tag{5.71a}$$
$$q(z) = 1, \quad \forall\, z \in \mathbb{R}^d, \text{ with } |z| \leq R; \tag{5.71b}$$
$$q(z) = 0, \quad \forall\, z \in \mathbb{R}^d, \text{ with } |z| \geq R\sqrt{2}. \tag{5.71c}$$

By Proposition 7.1 in Appendix 7 of Ethier and Kurtz [3], there exists a polynomial $p : \mathbb{R}^{2d} \to \mathbb{R}$ such that

$$\max_{x \in B_{2R}} |g(x) - p(x)| < \epsilon; \tag{5.72a}$$
$$\max_{x \in B_{2R}} |\partial_i g(x) - \partial_i p(x)| < \epsilon, \quad \forall\, i = 1, \ldots, 2d; \tag{5.72b}$$
$$\max_{x \in B_{2R}} |\partial_i \partial_j g(x) - \partial_i \partial_j p(x)| < \epsilon, \quad \forall\, i, j = 1, \ldots, 2d. \tag{5.72c}$$

Since $g(x) = 0$ when $x \notin B_R$, we note from (5.72a) that

$$\sup_{x \in B_{2R} \setminus B_R} |p(x)| < \epsilon. \tag{5.73}$$

For all $x \in \mathbb{R}^{2d}$, put $x := (x_1, x_2)$, $x_1, x_2 \in \mathbb{R}^d$, and define

$$\bar{q}(x) := q(x_1) q(x_2),$$
$$f(x) := \bar{q}(x)\, p(x).$$

Since $q \in C_c^\infty(\mathbb{R}^d)$ and $p(x)$ is a polynomial in $x = (x_1, x_2)$, it follows that $f \in \tilde{\mathcal{D}}$ (recall (4.60)). From (5.71), we have $\bar{q}(x) = 0$ when $x \notin B_{2R}$, and $\bar{q}(x) = 1$ when $x \in B_R$. Thus

$$\begin{aligned}\|f - g\| &= \sup_{x\in\mathbb{R}^{2d}} |\bar{q}(x)p(x) - g(x)| = \max_{x\in B_{2R}} |\bar{q}(x)p(x) - g(x)| \\ &\le \max_{x\in B_R} |p(x) - g(x)| + \sup_{x\in B_{2R}\setminus B_R} |\bar{q}(x)p(x) - g(x)|,\end{aligned}$$

hence (5.72a) and (5.73) give

$$\|f - g\| \le \epsilon + \sup_{x\in B_{2R}\setminus B_R} |p(x)| \le 2\epsilon. \tag{5.74}$$

Next, consider $\|\bar{\mathcal{A}}f - \bar{\mathcal{A}}g\|$. From (3.40) we have

$$\bar{\mathcal{A}}f(x) = \bar{q}(x)\bar{\mathcal{A}}p(x) + p(x)\bar{\mathcal{A}}\bar{q}(x) + (\nabla p(x))^T\bar{a}(x)\nabla\bar{q}(x). \tag{5.75}$$

By the choice of R we have $g(x) = 0$ and therefore $\bar{\mathcal{A}}g(x) = 0$, $\forall x \notin B_R$. Moreover, from (5.71c), we have $\bar{q}(x) = 1$, and therefore $\nabla\bar{q}(x) = 0$ and $\bar{\mathcal{A}}\bar{q}(x) = 0$, $\forall\, x \in B_R$. Similarly, $\bar{q}(x) = 0$, and therefore $\bar{\mathcal{A}}\bar{q}(x) = 0$, $\forall\, x \notin B_{2R}$. Then, it follows from (5.75) that

$$\begin{aligned}\|\bar{\mathcal{A}}f - \bar{\mathcal{A}}g\| &= \sup_{x\in B_{2R}} |\bar{q}(x)\bar{\mathcal{A}}p(x) + p(x)\bar{\mathcal{A}}\bar{q}(x) + (\nabla p(x))^T\bar{a}(x)\nabla\bar{q}(x) - \bar{\mathcal{A}}g(x)| \\ &\le \sup_{x\in B_R} |\bar{q}(x)\bar{\mathcal{A}}p(x) + p(x)\bar{\mathcal{A}}\bar{q}(x) + (\nabla p(x))^T\bar{a}(x)\nabla\bar{q}(x) - \bar{\mathcal{A}}g(x)| \\ &+ \sup_{x\in B_{2R}\setminus B_R} |\bar{q}(x)\bar{\mathcal{A}}p(x) + p(x)\bar{\mathcal{A}}\bar{q}(x) + (\nabla p(x))^T\bar{a}(x)\nabla\bar{q}(x)| \\ &= \sup_{x\in B_R} |\bar{\mathcal{A}}p(x) - \bar{\mathcal{A}}g(x)| \\ &+ \sup_{x\in B_{2R}\setminus B_R} (|\bar{q}(x)\bar{\mathcal{A}}p(x)| + |p(x)\bar{\mathcal{A}}\bar{q}(x)| + |(\nabla p(x))^T\bar{a}(x)\nabla\bar{q}(x)|).\end{aligned} \tag{5.76}$$

Since $\bar{a}$ and $\bar{b}$ are locally bounded, we have

$$C_1 := \sup_{x\in B_{2R}} \left(\sum_{i=1}^{2d} |\bar{b}_i(x)| + \frac{1}{2}\sum_{i,j=1}^{2d} |\bar{a}_{ij}(x)| \right) < \infty.$$

Also let

$$C_2 := \|\bar{q}\| + \sum_i \|\partial_i\bar{q}\| + \sum_{i,j} \|\partial_i\partial_j\bar{q}\| < \infty.$$

Then by (5.72)

$$\sup_{x\in B_R} |\bar{\mathcal{A}}p(x) - \bar{\mathcal{A}}g(x)| < C_1\epsilon. \tag{5.77}$$

Similarly, by (5.72) and the fact that $g(x) = 0$, $\forall x \in B_{2R} \setminus B_R$, we obtain

$$\left.\begin{aligned}|\partial_i p(x)| &< \epsilon, \\ |\bar{\mathcal{A}}p(x)| &< C_1\epsilon,\end{aligned}\right\} \quad \forall\, x \in B_{2R} \setminus B_R,$$

and hence, from these bounds and (5.73),

$$\sup_{x \in B_{2R} \setminus B_R} (|\bar{q}(x)\bar{\mathcal{A}}p(x)| + |p(x)\bar{\mathcal{A}}\bar{q}(x)| + |(\nabla p(x))^T \bar{a}(x) \nabla \bar{q}(x)|) \leq \epsilon C_1 + \epsilon C_1 C_2 + \epsilon C_1 C_2. \quad (5.78)$$

Now, upon combining (5.76), (5.77), and (5.78) we have

$$\|\bar{\mathcal{A}}g - \bar{\mathcal{A}}f\| \leq 2\epsilon\,(C_1 + C_1 C_2), \quad (5.79)$$

and the result follows. □

6 Proof of Theorem 3.32

Remark 6.38. In view of Remark 1 on page 345 of Bhatt and Karandikar [2], with no loss of generality we suppose that $\lambda \in B(\mathbb{R}^q)$ is a nonnegative mapping, and that the constant K in (3.43) is such that $0 \leq \lambda(x) \leq K$, $\forall\, x \in \mathbb{R}^q$.

We first show that the evolution equation for $(\mathcal{C} - \lambda, \mathcal{D}(\mathcal{C}))\})$ has uniqueness in the class of $\mathcal{M}^+(\mathbb{R}^q)$-valued solutions over the interval $[0, \infty)$. Thus, let $\{\mu_t,\ t \in [0, \infty)\}$ be such a solution, hence

$$\mu_t f = \mu_0 f + \int_0^t \mu_s(\mathcal{C}f - \lambda f)ds, \quad \forall\, t \in [0, \infty), \quad \forall\, f \in \mathcal{D}(\mathcal{C}). \quad (6.80)$$

Since $(1, 0) \in \mathcal{C}$, it follows from (6.80) that

$$\mu_t(\mathbb{R}^q) = \mu_0(\mathbb{R}^q) - \int_0^t \mu_s \lambda\, ds, \quad \forall\, t \in [0, \infty),$$

and therefore $\mu_t(\mathbb{R}^q) \in [0, 1]$, $\forall\, t \in [0, \infty)$ (since μ_0 is a probability measure and, by Remark 6.38, $\lambda(\cdot)$ is nonnegative); that is, each μ_t is a sub-probability measure on $\mathcal{B}(\mathbb{R}^q)$.

Remark 6.39. We are going to use Theorem 6.41 (which follows) to establish uniqueness in the class of $\mathcal{M}^+(\mathbb{R}^q)$-valued solutions of the evolution equation for $(\mathcal{C} - \lambda, \mathcal{D}(\mathcal{C}))$. However, Theorem 6.41 provides uniqueness in the class of *probability measure-valued* solutions of an evolution equation (recall Remark 3.27), and we have seen that, for an $\mathcal{M}^+(E)$-valued solution $\{\mu_t,\ t \in [0, \infty)\}$ of the evolution equation for $(\mathcal{C} - \lambda, \mathcal{D}(\mathcal{C}))$, the μ_t are only *sub-probability measures* on $\mathbb{R}^q$. We shall therefore use an idea of Bhatt and Karandikar (see page 344 of [2]), and add a "point at infinity" to get a one-point compactification $\mathbb{R}^{q*}$ of $\mathbb{R}^q$. Then the sub-probability measures μ_t on $\mathbb{R}^q$ are extended to *probability measures* μ_t^* on $\mathbb{R}^{q*}$ (see (6.81)), and it will be seen that the resulting function $\{\mu_t^*,\ t \in [0, \infty)\}$ is a $\mathcal{P}(\mathbb{R}^{q*})$-valued solution of the evolution equation for an operator $\mathcal{G}^* \subset B(\mathbb{R}^{q*}) \times B(\mathbb{R}^{q*})$ which is an "extension" of the operator $(\mathcal{C} - \lambda) \subset B(\mathbb{R}^q) \times B(\mathbb{R}^q)$ (see (6.83)). Theorem 6.41 will then be used to establish uniqueness in the class of $\mathcal{P}(\mathbb{R}^{q*})$-valued solutions of the evolution equation for $\mathcal{G}^*$, and this in turn will yield uniqueness in the class of the $\mathcal{M}^+(\mathbb{R}^q)$-valued solutions of the evolution equation (6.80), as required.

We use Δ to denote the point at infinity in the one-point compactification $\mathbb{R}^{q*}$ of $\mathbb{R}^q$. Also, members of $B(\mathbb{R}^{q*})$ and operators on $B(\mathbb{R}^{q*})$ will be superscripted with "$*$", and, for $f^* \in B(\mathbb{R}^{q*})$, we write $f^*|_{\mathbb{R}^q}$ to denote the restriction of f^* to the domain $\mathbb{R}^q$.

Using the postulated $\mathcal{M}^+(\mathbb{R}^q)$-valued solution $\{\mu_t,\ t \in [0,\infty)\}$ of the evolution equation for $(\mathcal{C} - \lambda, \mathcal{D}(\mathcal{C}))$, define $\mu_t^* \in \mathcal{P}(\mathbb{R}^{q*})$, $\forall\, t \in [0,\infty)$, as follows:

$$\mu_t^*(\Gamma) := \mu_t(\Gamma \cap \mathbb{R}^q) + (1 - \mu_t(\mathbb{R}^q)) I_\Gamma(\Delta), \quad \forall\, \Gamma \in \mathcal{B}(\mathbb{R}^{q*}). \tag{6.81}$$

Next, define the linear operator $\mathcal{C}^* : \mathcal{D}(\mathcal{C}^*) \subset C(\mathbb{R}^{q*}) \to B(\mathbb{R}^{q*})$ by

$$\mathcal{D}(\mathcal{C}^*) := \{f^* \in C(\mathbb{R}^{q*}) : f^*|_{\mathbb{R}^q} - f^*(\Delta) \in C_c^\infty(\mathbb{R}^q)\}, \tag{6.82a}$$

$$\mathcal{C}^* f^*(x) := \mathcal{C}(f^*|_{\mathbb{R}^q} - f^*(\Delta))(x), \quad \forall x \in \mathbb{R}^q, \quad \forall f^* \in \mathcal{D}(\mathcal{C}^*), \tag{6.82b}$$

$$\mathcal{C}^* f^*(\Delta) := 0, \quad \forall f^* \in \mathcal{D}(\mathcal{C}^*). \tag{6.82c}$$

Finally, define linear operator $\mathcal{G}^* : \mathcal{D}(\mathcal{G}^*) \subset C(\mathbb{R}^{q*}) \to B(\mathbb{R}^{q*})$ by

$$\mathcal{G}^* f^*(x) := \mathcal{C}^* f^*(x) - \lambda(x)(f^*(x) - f^*(\Delta)), \quad \forall\, x \in \mathbb{R}^{q*}, \quad \forall\, f^* \in \mathcal{D}(\mathcal{G}^*) := \mathcal{D}(\mathcal{C}^*). \tag{6.83}$$

From (6.81) and (6.80) one easily checks that $\{\mu_t^*,\ t \in [0,\infty)\}$ solves the evolution equation for $(\mathcal{G}^*, \mathcal{D}(\mathcal{G}^*))$, namely

$$\mu_t^* f^* = \mu_0^* f^* + \int_0^t \mu_s^*(\mathcal{G}^* f^*)\, ds, \quad \forall t \in [0,\infty), \quad \forall f^* \in \mathcal{D}(\mathcal{G}^*), \tag{6.84}$$

Remark 6.40. We clearly have the following: if the evolution equation for $(\mathcal{G}^*, \mathcal{D}(\mathcal{G}^*))$ has uniqueness in the class of $\mathcal{P}(\mathbb{R}^{q*})$-valued solutions over $[0,\infty)$, then the evolution equation for $(\mathcal{C} - \lambda, \mathcal{D}(\mathcal{C}))$ must have uniqueness in the class of $\mathcal{M}^+(\mathbb{R}^q)$-valued solutions over $[0,\infty)$. It therefore remains to establish the former type of uniqueness, and for this purpose we shall use the following special case of Theorem 2.7(c) from Kurtz [9]:

Theorem 6.41. *Suppose that F is a compact metric space, $\mathbb{A}^0 : \mathcal{D}(\mathbb{A}^0) \subset \bar{C}(\mathbb{R}^{q*}) \to \bar{C}(\mathbb{R}^{q*} \times F)$ is a linear operator, and η is a transition function from $\mathbb{R}^{q*}$ to F. Define*

$$\mathbb{A}_\eta^0 f^*(x) := \int_F \mathbb{A}^0 f^*(x,y)\eta(x,dy), \quad \forall\, x \in \mathbb{R}^{q*}, \quad \forall\, f^* \in \mathcal{D}(\mathbb{A}_\eta^0) := \mathcal{D}(\mathbb{A}^0). \tag{6.85}$$

Suppose also that (i) $\mathcal{D}(\mathbb{A}^0)$ is closed under multiplication and separates points, (ii) $\mathbb{A}_y^0 f^* \equiv \mathbb{A}^0 f^*(\cdot, y)$ is a pre-generator for each $y \in F$, and (iii) $\mathbb{A}_\eta^0$ satisfies the following separability hypothesis: there exists some countable $\{g_k^*\} \subset \mathcal{D}(\mathbb{A}_\eta^0)$ such that the graph of $\mathbb{A}_\eta^0$ is included within the bp-closure of the linear span of $\{(g_k^*, \mathbb{A}_\eta^0 g_k^*)\}$. With these conditions we have the following: if uniqueness holds for the martingale problem for $\mathbb{A}_\eta^0$ then the evolution equation for $(\mathbb{A}_\eta^0, \mathcal{D}(\mathbb{A}_\eta^0))$ has uniqueness in the class of $\mathcal{P}(\mathbb{R}^{q*})$-valued solutions over the interval $[0,\infty)$.

Define the compact metric space

$$F := \big\{ y = (y_1, y_2, y_3) \in \mathbb{S}_+^{q\times q} \times \mathbb{R}^q \times \mathbb{R} \,:\, |y_1^{ij}| \le K,\ |y_2^i| \le K,\ 0 \le y_3 \le K \big\},$$

and, motivated by Example 3.4 of Kurtz and Stockbridge [11], for each $y \in F$ define the linear operator $\mathcal{L}_y$ on $\bar{C}(\mathbb{R}^q)$ by

$$\begin{aligned} \mathcal{L}_y f(x) \; &:= \textstyle\sum_i (1+|x|)\, y_2^i\, \partial_i f(x) + \frac{1}{2} \sum_{i,j} (1+|x|^2)\, y_1^{ij}\, \partial_i \partial_j f(x), \\ &\qquad \forall x \in \mathbb{R}^q, \quad f \in \mathcal{D}(\mathcal{L}_y) := C_c^\infty(\mathbb{R}^q). \end{aligned} \tag{6.86}$$

Also, put $\mathcal{D}(\mathbb{A}^0) := \mathcal{D}(\mathcal{C}^*)$ (see (6.82a)) and

$$\begin{aligned} \mathbb{A}^0 f^*(x,y) &:= \mathcal{L}_y(f^*|_{\mathbb{R}^q} - f^*(\Delta))(x) - y_3(f^*(x) - f^*(\Delta)), \\ &\qquad \forall\, f^* \in \mathcal{D}(\mathbb{A}^0), \quad \forall\, (x,y) \in \mathbb{R}^q \times F, \end{aligned} \tag{6.87a}$$

$$\mathbb{A}^0 f^*(\Delta, y) := 0, \quad \forall\, f^* \in \mathcal{D}(\mathbb{A}^0), \quad \forall\, y \in F. \tag{6.87b}$$

This defines a linear mapping $\mathbb{A}^0 : \mathcal{D}(\mathbb{A}^0) \subset \bar{C}(\mathbb{R}^{q*}) \to \bar{C}(\mathbb{R}^{q*} \times F)$. Next, fix an arbitrary $\bar{y} \in F$, and define a transition function η from $\mathbb{R}^q$ to F by

$$\eta(x, \Gamma) := \delta_{\left(\frac{a(x)}{1+|x|^2}, \frac{b(x)}{1+|x|}, \lambda(x)\right)}(\Gamma), \quad x \in \mathbb{R}^q,\ \Gamma \in \mathcal{B}(F), \tag{6.88a}$$

$$\eta(\Delta, \Gamma) := \delta_{\bar{y}}(\Gamma), \quad \Gamma \in \mathcal{B}(F). \tag{6.88b}$$

Putting together (6.83), (6.86), (6.87) and (6.88), we get

$$\mathcal{G}^* \equiv \mathbb{A}_\eta^0, \tag{6.89}$$

(where $\mathbb{A}_\eta^0$ is given by (6.85) in terms of the operator $\mathbb{A}^0$ in (6.87) and the transition function η in (6.88)). We next check the conditions of Theorem 6.41 for $\mathbb{A}^0$ and η given by (6.87) and (6.88): (i) From (6.82a) it follows that $\mathcal{D}(\mathbb{A}^0) := \mathcal{D}(\mathcal{C}^*)$ is closed under multiplication and separates points. (ii) Fix $y \in F$ and $\nu^* \in \mathcal{P}(\mathbb{R}^{q*})$, and define the linear operator $\mathcal{L}_y^*$ on $C(\mathbb{R}^{q*})$ by $\mathcal{D}(\mathcal{L}_y^*) := \mathcal{D}(\mathcal{C}^*)$ and

$$\mathcal{L}_y^* f^*(x) := \mathcal{L}_y(f^*|_{\mathbb{R}^q} - f^*(\Delta))(x), \quad \forall\, x \in \mathbb{R}^q, \qquad \mathcal{L}_y^* f^*(\Delta) := 0, \quad \forall\, f^* \in \mathcal{D}(\mathcal{L}_y^*).$$

Clearly $\mathcal{L}_y$ given by (6.86) satisfies the positive maximum principle, thus Theorem 4.5.4 of Ethier and Kurtz [3] gives existence of a solution of the $D_{\mathbb{R}^{q*}}[0,\infty)$-martingale problem for $(\mathcal{L}_y^*, \nu^*)$. Then, for the linear operator $\mathbb{A}_y^0$ on $C(\mathbb{R}^{q*})$ given by $\mathcal{D}(\mathbb{A}_y^0) := \mathcal{D}(\mathcal{C}^*)$ and

$$\mathbb{A}_y^0 f^*(x) := \mathcal{L}_y^* f^*(x) - y_3(f^*(x) - f^*(\Delta)), \quad \forall\, x \in \mathbb{R}^{q*}, \quad \forall\, f^* \in \mathcal{D}(\mathbb{A}_y^0),$$

it follows from Theorem 4.10.2 of Ethier and Kurtz [3] that there exists a solution of the $D_{\mathbb{R}^{q*}}[0,\infty)$-martingale problem for $(\mathbb{A}_y^0, \delta_x)$, $\forall\ x \in \mathbb{R}^{q*}$, and therefore $\mathbb{A}_y^0$ is a pre-generator (see remark at foot of page 4 in [9]). (iii) From Remark 2.5 of Kurtz [9] there is a sequence $\{g_k\} \subset C_c^\infty(\mathbb{R}^q)$ such that the graph of $\mathcal{C}$ is included within the bp-closure of the linear span of $\{(g_k, \mathcal{C}g_k)\}$. Now define $g_{k,q}^*(x) := g_k(x) + q$, $\forall\ x \in \mathbb{R}^q$, $g_{k,q}^*(\Delta) := q$, $k = 1, 2, \ldots$, q rational.

From (6.83) it follows that the countable set $\{g^*_{k,q}\}$ is a subset of $\mathcal{D}(\mathcal{G}^*)$ and the graph of $\mathcal{G}^*$ is contained in the bp-closure of the linear span of $\{(g^*_{k,q}, \mathcal{G}^* g^*_{k,q})\}$. In view of (6.89) we have thus verified condition (iii) of Theorem 6.41. Finally, note that uniqueness holds for the martingale problem for $\mathcal{G}^*$. Indeed, observe from Theorem 8.1.7 of Ethier and Kurtz [3] that the martingale problem for $\mathcal{C}$ is well-posed, from which it easily follows that the $D_{\mathbb{R}^{q*}}[0,\infty)$-martingale problem for $\mathcal{C}^*$ is well-posed, and hence Theorem 4.10.2 of [3] shows that the $D_{\mathbb{R}^{q*}}[0,\infty)$-martingale problem for $\mathcal{G}^*$ is well-posed. Now it easily follows from Theorem 4.3.6 of [3] that uniqueness holds for the martingale problem for $\mathcal{G}^*$. We thus conclude from Theorem 6.41 and (6.89) that the evolution equation for $(\mathcal{G}^*, \mathcal{D}(\mathcal{G}^*))$ has uniqueness in the class of $\mathcal{P}(\mathbb{R}^{q*})$-valued solutions over $[0,\infty)$, and hence (see Remark 6.40) the evolution equation for $(\mathcal{C} - \lambda, \mathcal{D}(\mathcal{C}))$ has uniqueness in the class of $\mathcal{M}^+(\mathbb{R}^q)$-valued solutions over $[0,\infty)$.

It finally remains to establish that the evolution equation for $(\mathcal{C} - \lambda, \mathcal{D}(\mathcal{C}))$ has uniqueness in the class of $\mathcal{M}^+(\mathbb{R}^q)$-valued solutions over the finite interval $[0,T]$. Let $\{\mu_t,\ t \in [0,T]\}$ be such a solution. Extend $\mu_\cdot$ from $[0,T]$ to $[0,\infty)$ by defining

$$\mu_t(\Gamma) := \int_{\mathbb{R}^q} E_x\left[I_\Gamma(\omega_t) \exp\left(\int_0^{t-T} -\lambda(\omega_s)ds \right) \right] \mu_T(dx), \quad \forall\, t \in [T,\infty), \quad \forall\, \Gamma \in \mathcal{B}(\mathbb{R}^q),$$

where $P_x \in \mathcal{P}(C_{\mathbb{R}^q}[0,\infty))$, $x \in \mathbb{R}^q$, is the probability measure on the space $C_{\mathbb{R}^q}[0,\infty)$ (of continuous functions from $[0,\infty)$ into $\mathbb{R}^q$) which solves the martingale problem for $(\mathcal{C}, \delta_x)$, and $\omega_\cdot$ denotes a generic element of $C_{\mathbb{R}^q}[0,\infty)$ (existence and uniqueness of P_x follow from Theorem 8.1.7 of Ethier and Kurtz [3]). It is easily checked that the mapping $\{\mu_t,\ t \in [0,\infty)\}$ is an $\mathcal{M}^+(\mathbb{R}^q)$-valued solution of the evolution equation for $(\mathcal{C} - \lambda, \mathcal{D}(\mathcal{C}))$, and therefore the evolution equation for $(\mathcal{C} - \lambda, \mathcal{D}(\mathcal{C}))$ must have uniqueness in the class of $\mathcal{M}^+(\mathbb{R}^q)$-valued solutions over the finite interval $[0,T]$. □

Acknowledgements: The authors thank Professor T.G. Kurtz for valuable comments on this work.

References

[1] A. G. Bhatt, G. Kallianpur and R. L. Karandikar, Uniqueness and robustness of solution of measure-valued equations of nonlinear filtering, *Ann. Prob.*, 23, 1895–1938, 1995.

[2] A. G. Bhatt and R. L. Karandikar, Evolution equations for Markov processes: application to the white-noise theory of filtering, *Appl. Math. Optim.*, 31, 327–348, 1995.

[3] Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, J. Wiley & Sons, 1986.

[4] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland, 1989.

[5] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer–Verlag, 1991.

[6] H. Kunita, Asymptotic behavior of the nonlinear filtering errors of Markov processes, *J. Multivariate Anal.*, 1, 365–393, 1971.

[7] H. Kunita, Stochastic partial differential equations connected with nonlinear filtering, *Nonlinear Filtering and Stochastic Control, Lecture Notes in Mathematics, Vol. 972*, 100–169, Springer-Verlag, 1981.

[8] H. Kunita, *Stochastic Flows and Stochastic Differential Equations*, Cambridge University Press, 1990.

[9] T. G. Kurtz, Martingale problems for conditional distributions of Markov processes, *Elec. J. Prob.*, 3, 1–29, 1998.

[10] T. G. Kurtz and D. L. Ocone, Unique characterization of conditional distributions in nonlinear filtering, *Ann. Prob.*, 16, 80–107, 1988.

[11] T.G. Kurtz and R.H. Stockbridge, Existence of Markov controls and characterization of optimal Markov controls, *SIAM J. Control Optim.*, 36, 609–653, 1998.

[12] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Springer-Verlag, 1994.

[13] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales Vol. 2: Itô Calculus*, J. Wiley & Sons, 1987.

[14] B. L. Rozovskii, A simple proof of uniqueness for Kushner and Zakai equations, *Stochastic Analysis*, E. Mayer-Wolf, E. Merzbach and A. Shwartz, eds. 449–458, Academic Press, 1991.

[15] J. Szpirglas, Sur l'équivalence d'équations différentielles stochastiques á valeurs mesures intervenant dans le filtrage markovien non linéaire, *Ann. Inst. Henri Poincaré*, XIV, 33–59, 1978.

[16] T. Yamada and S. Watanabe, On the uniqueness of solutions of stochastic differential equations, *J. Math. Kyoto Univ.*, 11, 155–167, 1971.

Postal and E-mail addresses:

Vladimir M. Lucic
Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada
e-mail: lucic@kingcong.uwaterloo.ca

Andrew J. Heunis
Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada
e-mail: heunis@kingcong.uwaterloo.ca