Appl Math Optim 48:93–128 (2003)
DOI: 10.1007/s00245-003-0772-8

Applied Mathematics & Optimization

Convergence of Nonlinear Filters for Randomly Perturbed Dynamical Systems*

Vladimir M. Lucic[1] and Andrew J. Heunis[2]

[1]TD Securities, Toronto Dominion Bank,
14–18 Finsbury Square, London, EC2A 1DB, England
vladimir.lucic@tdsecurities.com

[2]Department of Electrical and Computer Engineering, University of Waterloo,
Waterloo, Ontario, Canada N2L 3G1
heunis@kingcong.uwaterloo.ca

Abstract. We establish convergence of the nonlinear filter of the state of a randomly perturbed dynamical system in which the perturbation is a rapidly fluctuating ergodic Markov process, and the observation process conditions the state of the system. The limiting nonlinear filter is completely characterized.

Key Words. Nonlinear filter equations, Martingale problems, Measure-valued processes, Weak convergence.

AMS Classification. Primary 60G35, Secondary 60G44, 60G57.

1. Introduction

Suppose that $Z := \{Z_t\}$ is a Markov process on a probability space $(\Omega, \mathcal{F}, P)$, taking values in a metric space S, with a unique invariant probability measure $\bar{m}$. Let Z^ε be a "rescaled process" defined by

$$Z_t^\varepsilon := Z_{t/\varepsilon^2}, \qquad \forall t \in [0, \infty), \tag{1.1}$$

for some *small parameter* $\varepsilon \in (0, 1]$, and consider the random ordinary differential equation

$$\frac{dX_t^\varepsilon}{dt} = \frac{1}{\varepsilon} F(X_t^\varepsilon, Z_t^\varepsilon) + H(X_t^\varepsilon, Z_t^\varepsilon), \qquad X_0^\varepsilon := \text{nonrandom } \ x_0 \in \mathbb{R}^d, \tag{1.2}$$

* This research was supported by MITACS-NSERC of Canada.

where the mappings $F, H\colon \mathbb{R}^d \times S \to \mathbb{R}^d$ are sufficiently well behaved to ensure that (1.2) has a unique solution X^ε for each sample-path of the process Z^ε. The ODE (1.2) with *small* ε is frequently used as a model for randomly perturbed dynamical systems, the interpretation of the terms in (1.2) being, briefly, as follows: the process Z^ε models "internal randomness" in the system, the vector field H models the aggregate drift of the dynamical system, and the vector field F models rapid local fluctuations around the paths determined by this drift. If the Markov process Z in (1.1) is "weak mixing," then $F(x, Z_s^\varepsilon)$ and $F(x, Z_t^\varepsilon)$ are "almost independent" for fixed $x \in \mathbb{R}^d$ and $s \neq t$ (since $\varepsilon > 0$ is small), so that the local fluctuations are indeed fast. We typically want these local fluctuations to be essentially "aimless," without any overall drift, at least asymptotically for small ε (when the distribution of Z_t^ε should be given approximately by the invariant probability measure $\bar{m}$), and hence it is also assumed that F satisfies the condition

$$\int_S F(x, z)\bar{m}(dz) = 0, \qquad \forall x \in \mathbb{R}^d. \tag{1.3}$$

When $\varepsilon > 0$ is small enough then the process $\{(1/\varepsilon)F(x, Z_t^\varepsilon)\}$ should be "almost a white noise" for each $x \in \mathbb{R}^d$, and hence one would expect that, as $\varepsilon \to 0$, the solution process X^ε of (1.2) converges weakly to the solution of a random differential equation driven by "perfect" white noise, or, more precisely, to the solution $\bar{X}$ of an Itô stochastic differential equation (SDE) of the form

$$d\bar{X}_t = b(\bar{X}_t)\,dt + c(\bar{X}_t)\,d\bar{V}_t, \qquad X_0^\varepsilon := \text{nonrandom } \ x_0 \in \mathbb{R}^d, \tag{1.4}$$

for some standard $\mathbb{R}^d$-valued Wiener process $\bar{V}$ on a probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P})$. Indeed, this convergence was established by Blankenship and Papanicolaou [6, Section 4] when the Markov process Z takes values in a *compact* metric space S, and the coefficients $b(\cdot)$ and $c(\cdot)$ in (1.4) are calculated in terms of the vector fields F and H in (1.2) (this result also appears as Theorem 12.2.4 on p. 475 of [8]).

Now suppose that, related to the process X^ε, we have an $\mathbb{R}^r$-valued *observation process* Y^ε defined by

$$Y_t^\varepsilon := \int_0^t h(X_s^\varepsilon)\,ds + W_t, \tag{1.5}$$

where W is an $\mathbb{R}^r$-valued Wiener process on $(\Omega, \mathcal{F}, P)$ independent of the process Z in (1.1), and $h\colon \mathbb{R}^d \to \mathbb{R}^r$ is a *sensor function* which is characteristic of the technical apparatus used to "measure" the signal X^ε. Without being too exact at this point about existence and technical measurability issues, we shall regard the *nonlinear filter of the signal* X^ε *corresponding to the observation process* Y^ε, as a process π^ε taking values in $\mathcal{P}(\mathbb{R}^d)$ (this denotes the set of all probability measures on $\mathbb{R}^d$), which is adapted to the self-filtration of the observation process Y^ε, and such that

$$\pi_t^\varepsilon f = \mathrm{E}[f(X_t^\varepsilon) \mid Y_s^\varepsilon,\ 0 \le s \le t] \quad \text{a.s.} \tag{1.6}$$

for each bounded Borel-measurable $f\colon \mathbb{R}^d \to \mathbb{R}$ and $t \ge 0$. If we regard $\bar{X}$ given by (1.4) as a "limiting signal," then we can introduce a corresponding observation process $\bar{Y}$ defined by

$$\bar{Y}_t := \int_0^t h(\bar{X}_s)\,ds + \bar{W}_t, \tag{1.7}$$

where $\bar{W}$ is an $\mathbb{R}^r$-valued Wiener process on $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P})$ which is independent of the Wiener process $\bar{V}$ in (1.4). Again, we shall regard the nonlinear filter of the signal $\bar{X}$ corresponding to the observation process $\bar{Y}$ as a $\mathcal{P}(\mathbb{R}^d)$-valued process $\bar{\pi}$ which is adapted to the self-filtration of $\bar{Y}$, and such that

$$\bar{\pi}_t f = \mathrm{E}[f(\bar{X}_t) \mid \bar{Y}_s,\ 0 \le s \le t] \quad \text{a.s.} \tag{1.8}$$

for each bounded Borel-measurable $f\colon \mathbb{R}^d \to \mathbb{R}$ and $t \ge 0$. The weak convergence of X^ε to $\bar{X}$ as $\varepsilon \to 0$ suggests that the nonlinear filter π^ε will likewise converge weakly to the nonlinear filter $\bar{\pi}$, and our goal is to establish this result. In fact, we look at a problem which is somewhat more general that the one just outlined and which is motivated as follows: In many applications (especially in aerospace problems) it is usual for the observation Y^ε to be *fed back* to the dynamics of the signal X^ε, so that the process X^ε is given, not by (1.2), but rather by

$$\begin{aligned} dX_t^\varepsilon &= \frac{1}{\varepsilon} F(X_t^\varepsilon, Z_t^\varepsilon)\,dt + H(X_t^\varepsilon, Z_t^\varepsilon)\,dt + B(X_t^\varepsilon)\,dY_t^\varepsilon, \\ X_0^\varepsilon &= \text{nonrandom } x_0 \in \mathbb{R}^d, \end{aligned} \tag{1.9}$$

or, equivalently, by

$$\begin{aligned} dX_t^\varepsilon &= \frac{1}{\varepsilon} F(X_t^\varepsilon, Z_t^\varepsilon)\,dt + G(X_t^\varepsilon, Z_t^\varepsilon)\,dt + B(X_t^\varepsilon)\,dW_t, \\ X_0^\varepsilon &= \text{nonrandom } x_0 \in \mathbb{R}^d, \end{aligned} \tag{1.10}$$

for $G(x,z) := H(x,z) + B(x)h(x), \quad (x,z) \in \mathbb{R}^d \times S$. Here $B\colon \mathbb{R}^d \to \mathbb{R}^{d\times r}$ is a mapping which defines the feedback of the observation Y^ε to the dynamics of X^ε. It seems plausible that the weak convergence of X^ε given by (1.10) to a limit $\bar{X}$ continues to hold provided that we modify the model (1.4) of the "limiting" signal $\bar{X}$ by adding an extra term to account for feedback of the observation process $\bar{Y}$ given by (1.7), namely

$$\begin{aligned} d\bar{X}_t &= b(\bar{X}_t)\,dt + c(\bar{X}_t)\,d\bar{V}_t + B(\bar{X}_t)\,d\bar{W}_t, \\ \bar{X}_0 &= \text{nonrandom } x_0 \in \mathbb{R}^d. \end{aligned} \tag{1.11}$$

The general question of convergence (as $\varepsilon \to 0$) of the nonlinear filter π^ε given by (1.5), (1.6), to a limiting nonlinear filter $\bar{\pi}$ given by (1.7), (1.8), continues to make sense when X^ε and $\bar{X}$ are given by (1.10) and (1.11), respectively, and our goal is to show this convergence.

In Section 2 we formulate the basic conditions which will always be postulated, and state the main results of this work (see Theorems 2.12 and 2.18). In Section 3 we present a convergence theorem of Bhatt and Karandikar [3] which is the main technical result that we use for establishing weak convergence. Sections 4 and 5 give proofs of the main results set forth in Section 2, and in Section 6 we compare our result on convergence of nonlinear filters with other works in the established literature which have also addressed this question. Finally, the remaining Sections 7–9 are appendices for the proofs of Theorems 2.12 and 2.18.

2. Conditions and Main Results

We always use the following notation and terminology:

(i) For a separable 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 define the supremum norm on $B(E)$ by $\|\varphi\| := \sup\{|\varphi(x)| : x \in E\}$, $\forall\varphi \in B(E)$. Also write $C(E)$ for the set of all real-valued continuous functions on E, put $\bar{C}(E) := B(E) \cap C(E)$, and write $C_c(E)$ for the set of all members of $\bar{C}(E)$ which have compact support. When E is locally compact let $\hat{C}(E)$ be the collection of all members of $\bar{C}(E)$ which vanish at infinity.

(ii) For a vector x in a finite-dimensional Euclidean space $\mathbb{R}^q$, write x^k for the kth scalar entry of x, and write $|x|$ for the Euclidean norm on $\mathbb{R}^q$. For a positive integer r, let $C^r(\mathbb{R}^q)$ denote the collection of all members of $C(\mathbb{R}^q)$ with continuous derivatives of each order, up to and including r. Let $C^\infty(\mathbb{R}^q)$ denote the collection of all members of $C(\mathbb{R}^q)$ with continuous derivatives of all orders. Put $C_c^\infty(\mathbb{R}^q) := C_c(\mathbb{R}^q) \cap C^\infty(\mathbb{R}^q)$ and $C_c^r(\mathbb{R}^q) := C_c(\mathbb{R}^q) \cap C^r(\mathbb{R}^q)$ for positive integers r. For E a metric space and r some positive integer, write $C^{r,0}(\mathbb{R}^q \times E)$ for the collection of all mappings $f \in C(\mathbb{R}^q \times E)$ whose partial derivatives of every order up to and including r, with respect to its first q real-valued arguments, exist and are members of $C(\mathbb{R}^q \times E)$, and put $C_c^{r,0}(\mathbb{R}^q \times E) := C^{r,0}(\mathbb{R}^q \times E) \cap C_c(\mathbb{R}^q \times E)$.

(iii) When E is a complete separable metric space let $\mathcal{P}(E)$ denote the collection of all probability measures on the measurable space $(E, \mathcal{B}(E))$ with the usual topology of weak (or narrow) convergence; and if $X\colon (\Omega, \mathcal{F}, P) \to E$ is $\mathcal{F}/\mathcal{B}(E)$-measurable, then let $\mathcal{L}(X)$ or, more precisely, $\mathcal{L}_P(X)$ denote the member of $\mathcal{P}(E)$ which is the distribution of X on $(E, \mathcal{B}(E))$. Also, for a $\mathcal{B}(E)$-measurable mapping $f\colon E \to \mathbb{R}$ which is integrable with respect to $\mu \in \mathcal{P}(E)$, we put $\mu f := \int_E f\,d\mu$. We now formulate conditions on the SDEs (1.10), (1.11), and the observations (1.5), (1.7):

Condition 2.1. For each $\varepsilon \in (0, 1]$ the process $\{Z_t^\varepsilon,\ t \in [0, \infty)\}$ is defined by (1.1), where $\{Z_t,\ t \in [0, \infty)\}$ is a Markov process on a complete probability space $(\Omega, \mathcal{F}, P)$, taking values in a compact metric space S, with a transition probability function $P_t(z, \Gamma)$ and initial distribution $\mu_0 \in \mathcal{P}(S)$. The transition probability function $P_t(z, \Gamma)$ is conservative, i.e. $P_t(z, S) = 1$, $\forall(t, z) \in [0, \infty) \times S$, and the semigroup $\{T_t\}$ defined by $T_t\Psi(z) := \int_S \Psi(z')P_t(z, dz')$, $z \in S$, maps $C(S)$ into $C(S)$ for each $t \in [0, \infty)$, and is a strongly continuous semigroup on $C(S)$ with infinitesimal generator $\mathcal{Q}$ (that is, $\{T_t\}$ is a *Feller semigroup* in the terminology of Ethier and Kurtz [8, p. 166]).

Remark 2.2. Write $\mathcal{D}(\mathcal{Q})$ and $\mathcal{R}(\mathcal{Q})$ for the domain and range of $\mathcal{Q}$. Without loss of generality we suppose that the sample paths of Z are corlol (cadlag) (see Theorem 4.2.7 of [8]).

Condition 2.3 (Ergodicity of $\{Z_t\}$). The transition probability $P_t(z, \Gamma)$ has unique invariant probability $\bar{m} \in \mathcal{P}(S)$, and $\int_0^\infty \|T_t\Psi - \bar{m}\Psi\|\,dt < \infty$, $\forall\Psi \in C(S)$.

Condition 2.4. The mappings $F\colon \mathbb{R}^d \times S \to \mathbb{R}^d$, $G\colon \mathbb{R}^d \times S \to \mathbb{R}^d$, and $B\colon \mathbb{R}^d \to \mathbb{R}^{d\times r}$ in (1.10) satisfy $F^i \in C^{3,0}(\mathbb{R}^d \times S)$, $G^i \in C^{2,0}(\mathbb{R}^d \times S)$, $B^{ij} \in C^2(\mathbb{R}^d)$,

$i = 1, 2, \ldots, d,\ j = 1, 2, \ldots, r$, and have uniformly bounded first x-derivatives, namely

$$\sup_{\mathbb{R}^d \times S} (|\partial_k F^i(x, z)| + |\partial_k G^i(x, z)| + |\partial_k B^{ij}(x)|) < \infty,$$
$$\forall i, k = 1, 2, \ldots, d, \quad j = 1, 2, \ldots, r. \tag{2.12}$$

Moreover, F is such that (1.3) holds.

Condition 2.5. The process $\{W_t,\ t \in [0, \infty\}$ in (1.5) and (1.10) is a standard $\mathbb{R}^r$-valued Wiener process on $(\Omega, \mathcal{F}, P)$, and is independent of the Markov process Z in Condition 2.1.

The following result, which is established in Section 9, will be needed to formulate the remaining conditions:

Lemma 2.6. *Suppose that Conditions* 2.1 *and* 2.3 *hold. Then, for each* $z \in S$, *there is a unique finite signed regular Borel measure* $\chi(z, \cdot)$ *on* S *such that* $\sup_z \|\chi(z, \cdot)\|_{\mathrm{TV}} < \infty$ (*where* $\|\cdot\|_{\mathrm{TV}}$ *is the total variation norm*) *and*

$$\int_S \Psi(z')\chi(\cdot, dz') = \int_0^\infty [T_t\Psi(\cdot) - \bar{m}\Psi]\, dt \in C(S), \qquad \forall \Psi \in C(S). \tag{2.13}$$

Moreover, the following hold:

(i) *If* $\Psi \in C(S)$ *is such that* $\bar{m}\Psi = 0$, *then* $\Psi \in \mathcal{R}(\mathfrak{Q})$, *and the function* $\Phi(z) := \int_S \Psi(z')\chi(z, dz'), z \in S$, *is a member of* $\mathcal{D}(\mathfrak{Q})$ *and solves the "Poisson equation"* $\mathfrak{Q}\Phi = -\Psi$.

(ii) *For every* $g \in C^{1,0}(\mathbb{R}^d \times S)$, *the function* $f(x, z) := \int_S g(x, z')\ \chi(z, dz')$, $\forall(x, z) \in (\mathbb{R}^d \times S)$, *is a member of* $C^{1,0}(\mathbb{R}^d \times S)$, *and* $\partial_j f(x, z) = \int_S \partial_j g(x, z')\ \chi(z, dz'), \forall(x, z) \in \mathbb{R}^d \times S,\ j = 1, 2, \ldots, d$.

Remark 2.7. The mapping $t \in [0, \infty) \longmapsto [T_t\Psi - \bar{m}\Psi] \in C(S)$ is continuous (since $\{T_t\}$ is strongly continuous). Thus Condition 2.3 and Lemma 1.1.4 of [8] ensures that the Riemann integral on the right-hand side of (2.13) exists in $C(S)$.

Remark 2.8. Condition 2.4, together with standard existence and uniqueness results for SDEs (see Theorem 5.1.1 of [13]), ensures that (1.10) has an $\mathbb{R}^d$-valued pathwise unique strong solution $\{X_t^\varepsilon\}$ adapted to the filtration $\{\mathcal{F}_t^{W,Z^\varepsilon}\}$ defined by

$$\mathcal{F}_t^{W,Z^\varepsilon} := \sigma\{W_s, Z_s^\varepsilon,\ s \in [0, t]\} \vee \mathcal{N}(P), \qquad t \in [0, \infty),$$

where $\mathcal{N}(P)$ denotes the set of P-null events in $(\Omega, \mathcal{F}, P)$.

We next formulate coefficients $b(x)$ and $c(x)$ of (1.11) so that $\bar{X}$ is the weak limit of X^ε given by (1.10). To this end, use the signed Borel measures $\chi(z, \cdot)$ from Lemma 2.6 to define

$$b^i(x) := \int_S \left\{ G^i(x, z) + \int_S \sum_{j=1}^d \partial_j F^i(x, z')\chi(z, dz')F^j(x, z) \right\} \bar{m}(dz), \tag{2.14}$$

$$a^{ij}(x) := \int_S \left\{ \int_S F^i(x, z')\chi(z, dz')F^j(x, z) + \int_S F^j(x, z')\chi(z, dz')F^i(x, z) \right\}$$
$$\times \bar{m}(dz) + [BB^T(x)]^{ij}, \tag{2.15}$$

for each $i, j = 1, \ldots, d$, and $x \in \mathbb{R}^d$. Also, define diffusion operator $\mathcal{L}$ with domain $\mathcal{D}(\mathcal{L})$ by

$$\mathcal{D}(\mathcal{L}) := C_c^\infty(\mathbb{R}^d), \tag{2.16a}$$

$$\mathcal{L}\varphi(x) := \sum_{i=1}^d b^i(x)\,\partial_i\varphi(x) + \frac{1}{2}\sum_{i,j=1}^d a^{ij}(x)\,\partial_i\partial_j\varphi(x),$$
$$x \in \mathbb{R}^d, \quad \varphi \in \mathcal{D}(\mathcal{L}). \tag{2.16b}$$

For each $x \in \mathbb{R}^d$ we have $\bar{m}F^i(x, \cdot) = 0$ (see Condition 2.4). Thus, upon taking $\Psi(\cdot) := F^i(x, \cdot)$ in (2.13), we get $\int_0^\infty (T_t F^i(x, \cdot))(z)\,dt = \int_S F^i(x, z')\chi(z, dz')$, $\forall z \in S$; now multiply each side by $F^j(x, z)$, integrate with respect to $\bar{m}(dz)$, and use Fubini's theorem to get

$$\int_S \left[\int_S F^i(x, z')\chi(z, dz') \right] F^j(x, z)\bar{m}(dz)$$
$$= \int_0^\infty \left[\int_S (T_t F^i(x, \cdot))(z) F^j(x, z)\bar{m}(dz) \right] dt$$
$$= \int_0^\infty \mathrm{E}[F^i(x, \hat{Z}_t)F^j(x, \hat{Z}_0)]\,dt, \tag{2.17}$$

where $\hat{Z}$ is a stationary S-valued process corresponding to the transition probability $P_t(z, \Gamma)$ and initial measure $\bar{m}$. It follows from (2.17) that the first term on the right-hand side of (2.15) defines a $d \times d$ nonnegative definite symmetric matrix which we denote by $\hat{a}(x)$, so that $a(x)$ is nonnegative definite and symmetric for each $x \in \mathbb{R}^d$. Let c: $\mathbb{R}^d \to \mathbb{R}^{d\times d}$ in (1.11) be a fixed Borel-measurable square root of $\hat{a}(\cdot)$; then from (2.15) we have

$$a(x) = [c(x) \quad B(x)] \begin{bmatrix} c^T(x) \\ B^T(x) \end{bmatrix}, \qquad x \in \mathbb{R}^d. \tag{2.18}$$

Remark 2.9. From (2.12), (2.14), (2.15), and Lemma 2.6 it follows that $b^i(\cdot)$ and $a^{ij}(\cdot)$ are continuous functions on $\mathbb{R}^d$, and there exists a constant $C \in [0, \infty)$ such that $|b^i(x)| \le C[1 + |x|]$, $|a^{ij}(x)| \le C[1 + |x|^2]$, $\forall x \in \mathbb{R}^d$, $\forall i, j = 1, 2, \ldots, d$.

Condition 2.10. For each $\mu \in \mathcal{P}(\mathbb{R}^d)$ there exists a *unique* $P_\mu \in \mathcal{P}(C_{\mathbb{R}^d}[0, \infty))$ which solves the Stroock–Varadhan martingale problem associated with $\mathcal{L}$ (see Definition 5.4.10 of [15]) and which has initial distribution μ, namely $P_\mu[\omega \in C_{\mathbb{R}^d}[0, \infty) : \omega_0 \in \Gamma] = \mu(\Gamma)$, $\forall \Gamma \in \mathcal{B}(\mathbb{R}^d)$.

It follows from Condition 2.10 and Corollary 5.3.4 of [8] that, for each $\mu \in \mathcal{P}(\mathbb{R}^d)$, there is a complete filtered probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{P})$, on which is defined an

$\mathbb{R}^{d+r}$-valued $\{\bar{\mathcal{F}}_t\}$-standard Wiener process $\{(\bar{V}_t, \bar{W}_t),\ t \in [0,\infty)\}$ and an $\mathbb{R}^d$-valued $\{\bar{\mathcal{F}}_t\}$-adapted continuous process $\{\bar{X}_t,\ t \in [0,\infty)\}$, such that (i) $\{(\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{P}), (\bar{X}_t, (\bar{V}_t, \bar{W}_t))\}$ is a (weak) solution of the stochastic integral equation corresponding to $((c\ B), b, \mu)$ (i.e. solves the SDE in (1.11) with $\mathcal{L}(\bar{X}_0) = \mu$—see the terminology on p. 291 of [8]) and (ii) $\mathcal{L}(\bar{X}) = P_\mu$. Moreover, for every solution $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{X}_t, (\tilde{V}_t, \tilde{W}_t))\}$ of the stochastic integral equation corresponding to $((c\ B), b, \mu)$ we have $\mathcal{L}(\tilde{X}) = P_\mu$ (by a "complete filtered probability space" $(\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{P})$ we always mean that $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P})$ is a complete probability space carrying a filtration $\{\bar{\mathcal{F}}_t\}$ such that each $\bar{\mathcal{F}}_t$ includes all $\bar{P}$-null events in $\bar{\mathcal{F}}$).

Remark 2.11. Conditions 2.1, 2.3–2.5, and 2.10 will typically be invoked together, and we therefore refer to these conditions collectively as Condition **AI**. The next result will be needed to formulate the main result on convergence of nonlinear filters.

Theorem 2.12. *Suppose Condition* AI *and fix an arbitrary Borel-measurable locally bounded mapping* c: $\mathbb{R}^d \to \mathbb{R}^{d\times d}$ *such that* (2.18) *holds. If* $\{(\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{P}), (\bar{X}_t, (\bar{V}_t, \bar{W}_t))\}$ *is a solution of* (1.11), *then* $\lim_{\varepsilon\to 0} \mathcal{L}(X^\varepsilon) = \mathcal{L}(\bar{X})$ *in* $\mathcal{P}(C_{\mathbb{R}^d}[0,\infty))$.

Remark 2.13. The conditions postulated for Theorem 2.12 are similar to the conditions postulated by Blankenship and Papanicolaou [6] who establish the preceding result in the case where $B \equiv 0$ in (1.10) and (1.11): see the (unnumbered) theorem on p. 449 of [6].

To study the convergence of nonlinear filters we postulate the additional two conditions:

Condition 2.14. The mapping c: $\mathbb{R}^d \to \mathbb{R}^{d\times d}$ in the factorization (2.18) (which is the coefficient $c(\cdot)$ in (1.11)) is continuous, $c(x)c^T(x)$ is strictly positive definite for each $x \in \mathbb{R}^d$, and $|c^{ik}(x)| \le C[1+|x|]$, $\forall x \in \mathbb{R}^d$, for a constant $C \in [0,\infty)$.

Condition 2.15. The sensor function h: $\mathbb{R}^d \to \mathbb{R}^r$ in (1.5) is continuous and uniformly bounded.

Remark 2.16. Fix some $T \in (0,\infty)$, and define the filtration of the observation process $\{Y_t^\varepsilon\}$ in (1.5) by

$$\mathcal{F}_t^{Y^\varepsilon} := \sigma\{Y_s^\varepsilon,\ 0 \le s \le t\} \vee \mathcal{N}(P), \qquad \forall t \in [0,T]. \tag{2.19}$$

By a standard Girsanov measure-change one finds a probability measure on $(\Omega, \mathcal{F})$, having the same null events as P, and with respect to which $\{Y_t^\varepsilon\}$ is a Brownian motion; now it follows (see, e.g. no. 4, pp. 21–22 of [13]) that the filtration $\{\mathcal{F}_t^{Y^\varepsilon}\}$ is right-continuous. Then Lemma 1.1 of [17] yields a $\mathcal{P}(\mathbb{R}^d)$-valued corlol process $\{\pi_t^\varepsilon,\ t \in [0,T]\}$ which is $\{\mathcal{F}_t^{Y^\varepsilon}\}$-adapted and satisfies

$$\pi_t^\varepsilon f = \mathrm{E}[f(X_t^\varepsilon) \mid \mathcal{F}_t^{Y^\varepsilon}] \quad \text{a.s.} \tag{2.20}$$

for each $t \in [0,T]$ and $f \in B(\mathbb{R}^d)$. This process is called the *nonlinear filter* for the signal process $\{X_t^\varepsilon,\ t \in [0,T]\}$ given the observation $\{Y_t^\varepsilon,\ t \in [0,T]\}$. Our goal is to study and characterize the asymptotic limit, in the sense of weak convergence, of the nonlinear filter $\{\pi_t^\varepsilon,\ t \in [0,T]\}$ as $\varepsilon \to 0$. To this end, fix some some solution $\{(\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{P}), (\bar{X}_t, (\bar{W}_t, \bar{V}_t))\}$ of (1.11) (as in Theorem 2.12), and define the $\mathbb{R}^r$-valued observation process $\{\bar{Y}_t,\ t \in [0,T]\}$ as in (1.7), together with its filtration

$$\mathcal{F}_t^{\bar{Y}} := \sigma\{\bar{Y}_s,\ 0 \le s \le t\} \vee \mathcal{N}(\bar{P}), \qquad \forall t \in [0,T]. \tag{2.21}$$

Then the filtration $\{\mathcal{F}_t^{\bar{Y}}\}$ is right-continuous (in the same way that $\{\mathcal{F}_t^{Y^\varepsilon}\}$ is seen to be right-continuous) and, again by Lemma 1.1 of [17], there is a $\mathcal{P}(\mathbb{R}^d)$-valued, corlol, and $\{\mathcal{F}_t^{\bar{Y}}\}$-adapted process $\{\bar{\pi}_t,\ t \in [0,T]\}$ such that

$$\bar{\pi}_t f = \mathrm{E}[f(\bar{X}_t) \mid \mathcal{F}_t^{\bar{Y}}] \quad \text{a.s.} \tag{2.22}$$

for each $t \in [0,T]$ and $f \in B(\mathbb{R}^d)$. The process $\{\bar{\pi}_t,\ t \in [0,T]\}$ is the *nonlinear filter* for the signal $\{\bar{X}_t,\ t \in [0,T]\}$ corresponding to the observation $\{\bar{Y}_t,\ t \in [0,T]\}$.

Remark 2.17. We refer to the set of conditions given by Condition AI (see Remark 2.11) together with Conditions 2.14 and 2.15 as Condition **AII**. With the preceding established, we can state the main result as follows:

Theorem 2.18. *Suppose that Condition* AII *holds, and fix* $T \in (0,\infty)$. *Then* $\{(\pi_t^\varepsilon, Y_t^\varepsilon),\ t \in [0,T]\}$ *and* $\{(\bar{\pi}_t, \bar{Y}_t),\ t \in [0,T]\}$, *defined by* (2.20) *and* (2.22), *respectively, are* $(\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r)$*-valued continuous processes, and* $\lim_{\varepsilon\to 0+} \mathcal{L}((\pi^\varepsilon, Y^\varepsilon)) = \mathcal{L}((\bar{\pi}, \bar{Y}))$ *in* $\mathcal{P}(C_{\mathcal{P}(\mathbb{R}^d)\times\mathbb{R}^r}[0,T])$.

3. A Convergence Theorem of Bhatt and Karandikar

In this section we state a result of Bhatt and Karandikar [3] which is a very effective tool for establishing weak convergence of stochastic processes, and which will be used to prove Theorems 2.12 and 2.18.

Remark 3.19. The following notation and terminology will be needed for this and later sections. Suppose that E is a separable metric space:

(i) The set $M \subset B(E)$ *separates points in* E when the equality $f(x) = f(y)$, $\forall f \in M$, for some $x, y \in E$, implies $x = y$; and the set $M \subset B(E)$ *strongly separates points in* E when the convergence $\lim_n f(x_n) = f(x)$, $\forall f \in M$, for some $x, x_n \in E$, implies $\lim_n x_n = x$.

(ii) For $\varphi, \varphi_n \in B(E)$ write b.p.-$\lim_n \varphi_n = \varphi$ to indicate $\sup_n \|\varphi_n\| < \infty$ and $\lim_n \varphi_n(x) = \varphi(x)$, $\forall x \in E$. A set $M \subset B(E)$ is called *b.p.-closed* when b.p.-$\lim_n \varphi_n = \varphi$ for some sequence $\{\varphi_n\} \subset M$ implies $\varphi \in M$. If $M_\lambda \subset B(E)$ is a given collection of b.p.-closed sets, then $\cap_\lambda M_\lambda$ is of course b.p.-closed; the *b.p.-closure* of $M \subset B(E)$ is defined to be the intersection of all b.p.-closed $M_\lambda \subset B(E)$ such that $M \subset M_\lambda$. Likewise, a set $M \subset B(E) \times B(E)$ is called *b.p.-closed* when b.p.-$\lim_n \varphi_n = \varphi$ and b.p.-$\lim_n \psi_n = \psi$ for a sequence $\{(\varphi_n, \psi_n)\} \subset M$ implies $(\varphi, \psi) \in M$; and the *b.p.-closure* of a set $M \subset B(E) \times B(E)$ is defined to be the intersection of all b.p.-closed sets $M_\lambda \subset B(E) \times B(E)$ such that $M \subset M_\lambda$.

(iii) A set $M \subset \bar{C}(E)$ is called *separating* when the equality $\nu_1\varphi = \nu_2\varphi, \forall\varphi \in M$, for some $\nu_1, \nu_2 \in \mathcal{P}(E)$, implies that $\nu_1 = \nu_2$; and is called *convergence determining* when the convergence $\lim_{n\to\infty} \nu_n\varphi = \nu_0\varphi, \forall\varphi \in M$, for some sequence $\{\nu_n,\ n = 0, 1, 2, \ldots\} \subset \mathcal{P}(E)$, implies weak convergence of ν_n to ν_0 as $n \to \infty$.

(iv) Let $\mathcal{A} \subset B(E) \times B(E)$ be a relation with domain $\mathcal{D}(\mathcal{A})$, and let $\mu \in \mathcal{P}(E)$. Then a *progressively measurable solution* of the martingale problem for $\mathcal{A}$ (for $(\mathcal{A}, \mu)$) is some pair $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{X}_t)\}$, in which $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P})$ is a complete filtered probability space and $\{\tilde{X}_t\}$ is an E-valued $\{\tilde{\mathcal{F}}_t\}$-progressively measurable process such that $f(\tilde{X}_t) - \int_0^t \mathcal{A}f(\tilde{X}_s)\,ds$ is an $\{\tilde{\mathcal{F}}_t\}$-martingale for each $f \in \mathcal{D}(\mathcal{A})$ (and $\mathcal{L}(\tilde{X}_0) = \mu$). If $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{X}_t)\}$ is a progressively measurable solution of the martingale problem for $\mathcal{A}$ (for $(\mathcal{A}, \mu)$) and the E-valued process $\{\tilde{X}_t\}$ has corlol paths, then $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{X}_t)\}$ is a solution of the *corlol martingale problem* for $\mathcal{A}$ (for $(\mathcal{A}, \mu)$). The martingale problem for $(\mathcal{A}, \mu)$ has the property of *existence* when there exists some progressively measurable solution of the martingale problem for $(\mathcal{A}, \mu)$, and has the property of *uniqueness* when, given any two progressively measurable solutions $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{X}_t)\}$ and $\{(\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{P}), (\bar{X}_t)\}$ of the martingale problem for $(\mathcal{A}, \mu)$, the E-valued processes $\tilde{X}$ and $\bar{X}$ necessarily have identical finite-dimensional distributions. The martingale problem for $(\mathcal{A}, \mu)$ is called *well-posed* when it has the properties of both existence and uniqueness. Finally, the martingale problem for $\mathcal{A}$ is *well-posed* when the martingale problem for $(\mathcal{A}, \mu)$ is well-posed for each $\mu \in \mathcal{P}(E)$. The notions of existence, uniqueness, and well-posedness of the corlol martingale problem for $(\mathcal{A}, \mu)$ and for $\mathcal{A}$ are similarly formulated.

Theorem 3.20 (Theorem 2.1 and Remark 2.2 of [3]). *Suppose that E is a complete separable metric space, $\mu \in \mathcal{P}(E)$, and $\mathcal{A}$: $\mathcal{D}(\mathcal{A}) \subset \bar{C}(E) \to \bar{C}(E)$ is a linear operator such that the following conditions hold*:

(I) *There is a countable set $\{g_k\} \subset \mathcal{D}(\mathcal{A})$ such that* (i) $\{(f, \mathcal{A}f) :\ f \in \mathcal{D}(\mathcal{A})\}$ *is a subset of the b.p.-closure of* $\{(g_k, \mathcal{A}g_k)\}$, *and* (ii) *the set $\{g_k\}$ strongly separates points in E.*

(II) $\mathcal{D}(\mathcal{A})$ *is an algebra that vanishes nowhere.*

(III) *The martingale problem for $\mathcal{A}$ is well-posed* (*see Remark* 3.19(iv)).

(IV) *The martingale problem for $(\mathcal{A}, \mu)$ has a solution $\{X_t,\ t \in [0, \infty)\}$ with corlol paths.*

(V) *There is a sequence $\{X_n(t),\ t \in [0, \infty)\}$, $n = 1, 2, \ldots,$ of E-valued processes with corlol paths such that $\{\mathcal{L}(X_n(t)),\ n = 1, 2, \ldots\}$ is a tight sequence in $\mathcal{P}(E)$ for each $t \in [0, \infty)$, and $\lim_n \mathcal{L}(X_n(0)) = \mu$ in $\mathcal{P}(E)$.*

(VI) *For each $f \in \mathcal{D}(\mathcal{A})$, there exist $\mathbb{R}$-valued progressively measurable processes $\{(U_n(t), \mathcal{F}_t^n), t \in [0, \infty)\}$ and $\{(V_n(t), \mathcal{F}_t^n),\ t \in [0, \infty)\}$ on $(\Omega_n, \mathcal{F}^n, P_n)$, $n = 1, 2, \ldots,$ such that*

$$U_n(t) - \int_0^t V_n(s)\,ds, \qquad t \in [0, \infty) \quad \text{is an } \{\mathcal{F}_t^n\}\text{-martingale}, \tag{3.23}$$

$$\lim_n \mathrm{E}_n \left[\sup_{t\in[0,t_1]} |U_n(t) - f(X_n(t))| \right] = 0, \qquad \text{for each} \quad t_1 \in [0, \infty), \tag{3.24}$$

$$\sup_n \mathrm{E}_n\left[\left\{\int_0^{t_1} |V_n(s)|^p\, ds\right\}^{1/p}\right] < \infty, \qquad \textit{for some} \quad p \in (1,\infty)$$
$$\textit{and each} \quad t_1 \in [0,\infty), \tag{3.25}$$

$$\lim_n \mathrm{E}_n[|\mathcal{A} f(X_n(t)) - V_n(t)|] = 0, \qquad \textit{for each} \quad t \in [0,\infty). \tag{3.26}$$

Then $\lim_{n\to\infty} \mathcal{L}(X_n) = \mathcal{L}(X)$ *in* $\mathcal{P}(D_E[0,\infty))$.

4. Proof of Theorem 2.12

In this section we establish Theorem 2.12, which extends the theorem of Blankenship and Papanicolaou [6, Section 4, p. 449] to include the third term on the right of (1.10) (recall Remark 2.13). Our tool of proof is Theorem 3.20 (which will also be used for the main result on convergence of nonlinear filters). In fact, the full strength of Theorem 3.20 is not actually needed in this section, since it is really meant for proving weak convergence of processes taking values in "large" (i.e. not locally compact) spaces, and in Theorem 2.12 we are only looking at convergence of processes taking values in $\mathbb{R}^d$. Nevertheless, even in this case, the use of Theorem 3.20 results in a considerable gain in simplicity and transparency with respect to the arguments in [6], which rely on the more classical methods of weak convergence found in [5]. In the course of this section we make explicit the construction of certain functions which will be essential not only for proving Theorem 2.12, but also for establishing the main result on convergence of nonlinear filters (see the functions $f^{\varepsilon,\varphi}$ in Proposition 4.22).

Recalling the infinitesimal generator $\mathcal{Q}$ of the Markov process $\{Z_t\}$ in Condition 2.1, put

$$\begin{aligned}\tilde{\mathcal{D}} := \{ f \in C_c^{2,0}(\mathbb{R}^d \times S) \;:\; & f(x,\cdot) \in \mathcal{D}(\mathcal{Q}), \quad \forall x \in \mathbb{R}^d, \text{ and the mapping} \\ & (x,z) \in \mathbb{R}^d \times S \mapsto \mathcal{Q}[f(x,\cdot)](z) \in \mathbb{R} \text{ is a member of} \\ & C_c^{2,0}(\mathbb{R}^d \times S)\}. \end{aligned} \tag{4.27}$$

For each $\varepsilon \in (0,1]$ define the operator $\mathcal{C}^\varepsilon$ with domain $\tilde{\mathcal{D}}$ as follows:

$$\begin{aligned}\mathcal{C}^\varepsilon f(x,z) := & \frac{1}{\varepsilon^2}\mathcal{Q}[f(x,\cdot)](z) + \sum_{i=1}^d \left[\frac{1}{\varepsilon}F^i(x,z) + G^i(x,z)\right]\partial_i f(x,z) \\ & + \frac{1}{2}\sum_{i,j=1}^d [BB^T(x)]^{ij}\,\partial_i\partial_j f(x,z), \end{aligned} \tag{4.28}$$

for all $(x,z) \in \mathbb{R}^d \times S$ and $f \in \tilde{\mathcal{D}}$. Note that $\tilde{\mathcal{D}}$ is a linear subspace of $C_c^{2,0}(\mathbb{R}^d \times S)$, and that $\mathcal{C}^\varepsilon f$ is an element of $C_c(\mathbb{R}^d \times S)$ for each $f \in \tilde{\mathcal{D}}$, so that $(\mathcal{C}^\varepsilon, \tilde{\mathcal{D}})$ is a linear operator on $C_c(\mathbb{R}^d \times S)$ for every $\varepsilon \in (0,1]$. The next proposition, whose proof is given in Section 7, establishes that $(X^\varepsilon, Z^\varepsilon)$ solves the martingale problem for the operator $(\mathcal{C}^\varepsilon, \tilde{\mathcal{D}})$:

Proposition 4.21. *Suppose Condition* AI (*see Remark* 2.11). *For every* $\varepsilon \in (0, 1]$ *and* $f \in \tilde{\mathcal{D}}$ *the process* $M_t^{\varepsilon,f} := f(X_t^\varepsilon, Z_t^\varepsilon) - \int_0^t \mathcal{C}^\varepsilon f(X_s^\varepsilon, Z_s^\varepsilon)\,ds$, $\forall t \in [0, \infty)$, *is an* $\{\mathcal{F}_t^{W,Z^\varepsilon}\}$*-martingale, where* $\mathcal{F}_t^{W,Z^\varepsilon}$ *is defined in Remark* 2.8.

The next result, which is also proved in Section 7, gives functions which will be needed for checking (VI) of Theorem 3.20 (and which will also be needed for establishing Theorem 2.18):

Proposition 4.22. *Suppose Condition* AI (*see Remark* 2.11). *Then, for each* $\varphi \in C_c^\infty(\mathbb{R}^d)$, *there exist functions* $f_1^\varphi \in C_c^{3,0}(\mathbb{R}^d \times S)$ *and* $f_2^\varphi \in C_c^{2,0}(\mathbb{R}^d \times S)$ *with the following property: If* $f^{\varepsilon,\varphi}(\cdot)$ *is defined for each* $\varepsilon \in (0, 1]$ *by*

$$f^{\varepsilon,\varphi}(x, z) := \varphi(x) + \varepsilon f_1^\varphi(x, z) + \varepsilon^2 f_2^\varphi(x, z), \qquad \forall (x, z) \in \mathbb{R}^d \times S, \tag{4.29}$$

then $f^{\varepsilon,\varphi} \in \tilde{\mathcal{D}}$, *and* $\mathcal{C}^\varepsilon f^{\varepsilon,\varphi}$ *has the form* $\mathcal{C}^\varepsilon f^{\varepsilon,\varphi}(x, z) = \mathcal{L}\varphi(x) + \varepsilon\gamma_1^\varphi(x, z) + \varepsilon 2\gamma_2^\varphi(x, z)$, $\forall (x, z, \varepsilon) \in \mathbb{R}^d \times S \times (0, 1]$, *for some* $\gamma_1^\varphi, \gamma_2^\varphi \in \bar{C}(\mathbb{R}^d \times S)$.

Now we are ready to establish Theorem 2.12. We do so using Theorem 3.20, and identify E, μ, $\mathcal{A}$, and $\mathcal{D}(\mathcal{A})$ in Theorem 3.20 with $\mathbb{R}^d$, δ_{x_0} (for x_0 in (1.10) and (1.11)), $\mathcal{L}$, and $\mathcal{D}(\mathcal{L}) := C_c^\infty(\mathbb{R}^d)$, respectively (see (2.16)).

Verification of (I) *and* (II) *in Theorem* 3.20. (I) is a consequence of the fact that $\{(\varphi, \mathcal{L}\varphi),\ \varphi \in C_c^\infty(\mathbb{R}^d)\}$ is separable in the supremum norm of $\bar{C}(\mathbb{R}^d) \times \bar{C}(\mathbb{R}^d)$ (which follows from Lemma 9.48) and the fact that $C_c^\infty(\mathbb{R}^d)$ strongly separates points in $\mathbb{R}^d$. (II) is clear.

Verification of (III) *and* (IV) *in Theorem* 3.20. From Condition 2.10 and Proposition 5.3.5 of [8] (together with the bounds on b^i and a^{ij} in Remark 2.9) it follows that the martingale problem for $\mathcal{L}$ is well-posed (in the sense of Remark 3.19(iv)), as required to check (III). As for (IV), we identify X in Theorem 3.20(IV) with the continuous (hence corlol) process $\bar{X}$ in the postulated solution $\{(\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{P}), (\bar{X}_t, (\bar{V}_t, \bar{W}_t))\}$ of (1.11). Now fix a sequence $\{\varepsilon_n\} \subset (0, 1]$ with $\lim_n \varepsilon_n = 0$, and identify the continuous (hence corlol) process X^{ε_n} given by (1.10) with X_n in Theorem 3.20(V) and (VI).

Verification of (V) *in Theorem* 3.20. From (1.10) we have $\mathcal{L}(X_0^{\varepsilon_n}) = \delta_{x_0}$. The remainder of (V) is verified by the next result which is established in Section 7:

Proposition 4.23. *Suppose that Condition* AI *holds* (*see Remark* 2.11). *Then the sequence* $\{X_t^{\varepsilon_n},\ n = 1, 2, \ldots\}$ *is tight in* $\mathcal{P}(\mathbb{R}^d)$ *for each* $t \in [0, \infty)$.

Verification of (VI) *in Theorem* 3.20. Fix some $\varphi \in C_c^\infty(\mathbb{R}^d)$. For each $n = 1, 2, \ldots,$ put

$$U_n(t) := f^{\varepsilon_n,\varphi}(X_t^{\varepsilon_n}, Z_t^{\varepsilon_n}), \qquad V_n(t) := \mathcal{C}^{\varepsilon_n} f^{\varepsilon_n,\varphi}(X_t^{\varepsilon_n}, Z_t^{\varepsilon_n}),$$

where $f^{\varepsilon_n,\varphi}$ is given by Proposition 4.22 and $\mathcal{C}^{\varepsilon_n}$ is the operator defined by (4.28). Then U_n and V_n are $\{\mathcal{F}_t^{W,Z^{\varepsilon_n}}\}$-progressively measurable (see Remark 2.8), and (3.23) follows from Propositions 4.21 and 4.22. By Proposition 4.22 we have $U_n(t)-\varphi(X_t^{\varepsilon_n}) = \varepsilon_n f_1^{\varphi}(X_t^{\varepsilon_n}, Z_t^{\varepsilon_n}) + \varepsilon_n 2 f_2^{\varphi}(X_t^{\varepsilon_n}, Z_t^{\varepsilon_n})$, and (3.24) follows from the uniform boundedness of f_1^{φ} and f_2^{φ} over $\mathbb{R}^d \times S$, and $\varepsilon_n \to 0$. Again by Proposition 4.22, we have $V_n(t) = \mathcal{L}\varphi(X_t^{\varepsilon_n}) + \varepsilon_n \gamma_1^{\varphi}(X_t^{\varepsilon_n}, Z_t^{\varepsilon_n}) + \varepsilon_n 2\gamma_2^{\varphi}(X_t^{\varepsilon_n}, Z_t^{\varepsilon_n})$. Now (3.25) and (3.26) follow from the uniform boundedness of γ_1^{φ} and γ_2^{φ} over $\mathbb{R}^d \times S$.

All conditions of Theorem 3.20 have now been checked, and thus $\lim_n \mathcal{L}(X^{\varepsilon_n}) = \mathcal{L}(\bar{X})$ in $\mathcal{P}(D_{\mathbb{R}^d}[0,\infty))$. Since X^{ε_n} and $\bar{X}$ are continuous processes we then get $\lim_n \mathcal{L}(X^{\varepsilon_n}) = \mathcal{L}(\bar{X})$ in $\mathcal{P}(C_{\mathbb{R}^d}[0,\infty))$ (see Corollary 3.3.2 of [8]) as required.

5. Proof of Theorem 2.18

We begin by formulating the general notion of a (weak) solution of the normalized (or Fujisaki–Kallianpur–Kunita–Kushner–Stratonovich) equation of nonlinear filtering:

Condition 5.24. Let E be a metric space, and let $c^k \in B(E),\ k = 1,2,\ldots,r$. Also let $\mathcal{G}, \mathcal{H}_k \subset \bar{C}(E) \times \bar{C}(E),\ k = 1,2,\ldots,r$, be linear operators with a common domain $\mathcal{D}(\mathcal{G},\mathcal{H}) \subset \bar{C}(E)$.

Definition 5.25. Suppose Condition 5.24. The pair $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\nu}_t, \tilde{W}_t)\}$ is a solution of the normalized filter equation corresponding to $(E; \mathcal{G}, \mathcal{H}, c)$, when the following hold:

1. $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P})$ is a complete filtered probability space;
2. $\{\tilde{W}_t,\ t \in [0,T]\}$ is a standard $\mathbb{R}^r$-valued $\{\tilde{\mathcal{F}}_t\}$-Wiener process on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$;
3. $\{\tilde{\nu}_t,\ t \in [0,T]\}$ is a $\mathcal{P}(E)$-valued, corlol, $\{\tilde{\mathcal{F}}_t\}$-adapted process, such that for every $\varphi \in \mathcal{D}(\mathcal{G},\mathcal{H})$, we have

$$\tilde{\nu}_t\varphi = \tilde{\nu}_0\varphi + \int_0^t \tilde{\nu}_s(\mathcal{G}\varphi)\,ds + \sum_{k=1}^r \int_0^t R_{\mathcal{H}_k}(\varphi, c^k, \tilde{\nu}_s)\,d\tilde{W}_s^k, \quad t \in [0,T], \tag{5.30}$$

where

$$R_{\mathcal{H}_k}(\varphi, c^k, \nu) := \nu(c^k\varphi + \mathcal{H}_k\varphi) - (\nu c^k)(\nu\varphi), \qquad \forall \nu \in \mathcal{P}(E), \quad \forall \varphi \in \mathcal{D}(\mathcal{G},\mathcal{H}). \tag{5.31}$$

Remark 5.26. We have given a fairly general formulation of the normalized filter equation because later we will need to interpret E as both $\mathbb{R}^d \times S$ and $\mathbb{R}^d$, and will have to make suitable choices for the operators $\mathcal{G}$ and $\mathcal{H}$ in Condition 5.24. When $E := \mathbb{R}^d$ and $\mathcal{D}(\mathcal{G},\mathcal{H}) := C_c^\infty(\mathbb{R}^d)$ in Condition 5.24, then the $\mathcal{P}(\mathbb{R}^d)$-valued process $\{\tilde{\nu}_t\}$ in

Definition 5.25 is continuous. This follows from (5.30), since $\{\tilde{\nu}_t\varphi\}$ is continuous for each φ in the set $C_c^\infty(\mathbb{R}^d)$, which is convergence determining (by Problem 3.11.11 of [8] and the fact that $C_c^\infty(\mathbb{R}^d)$ is dense in $\hat{C}(\mathbb{R}^d)$ with respect to the supremum norm).

Remark 5.27. Define $\mathbb{R}^r$-valued innovations processes I^ε, $\varepsilon \in (0, 1]$, and $\bar{I}$ by

$$I_t^\varepsilon := Y_t^\varepsilon - \int_0^t \pi_u^\varepsilon h\,du, \qquad \bar{I}_t := \bar{Y}_t - \int_0^t \bar{\pi}_u h\,du, \qquad t \in [0, T]. \tag{5.32}$$

Theorem VI.8.4(i) of [19] shows that $\{(I_t^\varepsilon, \mathcal{F}_t^{Y^\varepsilon})\}$ and $\{(\bar{I}_t, \mathcal{F}_t^{\bar{Y}})\}$ are $\mathbb{R}^r$-valued Wiener processes on $(\Omega, \mathcal{F}, P)$ and $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P})$, respectively.

We next introduce a $\mathcal{P}(\mathbb{R}^d \times S)$-valued process μ_t^ε, which is auxiliary to the nonlinear filter π^ε defined by (2.20). The $(\mathbb{R}^d \times S)$-valued process $(X^\varepsilon, Z^\varepsilon)$ is corlol (recall Remark 2.2), hence Lemma 1.1 of [17] gives a $\mathcal{P}(\mathbb{R}^d \times S)$-valued, corlol, and $\{\mathcal{F}_t^{Y^\varepsilon}\}$-adapted process $\{\mu_t^\varepsilon,\ t \in [0, T]\}$ on $(\Omega, \mathcal{F}, P)$ such that

$$\mu_t^\varepsilon g = \mathrm{E}[g(X_t^\varepsilon, Z_t^\varepsilon) \mid \mathcal{F}_t^{Y^\varepsilon}] \quad \text{a.s.,} \qquad \forall t \in [0, T], \quad g \in B(\mathbb{R}^d \times S). \tag{5.33}$$

Thus, the nonlinear filter $\{\pi_t^\varepsilon\}$ and the process $\{\mu_t^\varepsilon\}$ are related by

$$\pi_t^\varepsilon f = \mu_t^\varepsilon (f \otimes 1) \quad \text{a.s.,} \qquad \forall t \in [0, T], \quad f \in B(\mathbb{R}^d). \tag{5.34}$$

Define the operators $\mathcal{E}_k$, $k = 1, 2, \ldots, r$, on $\tilde{\mathcal{D}}$ (recall (4.27)) by

$$\mathcal{E}_k f(x, z) := \sum_{j=1}^d B^{jk}(x)\,\partial_j f(x, z), \qquad (x, z) \in \mathbb{R}^d \times S, \quad f \in \tilde{\mathcal{D}}, \\ k = 1, \ldots, r. \tag{5.35}$$

We are going to show that $\{(\mu_t^\varepsilon, I_t^\varepsilon)\}$ is a solution of a normalized filter equation of the form in Definition 5.25. The key is the following lemma whose proof is given in Section 8:

Lemma 5.28. *Suppose Condition* AII (*Remark* 2.17). *Then, for each* $\varepsilon \in (0, 1]$ *and* $f \in \tilde{\mathcal{D}}$, *we have that* $\langle M^{\varepsilon,f}, W^k\rangle_t = \int_0^t \mathcal{E}_k f(X_s^\varepsilon, Z_s^\varepsilon)\,ds,\ 0 \le t \le T,\ k = 1, \ldots, r,$ ($M_t^{\varepsilon,f}$ *is defined in Proposition* 4.21).

Remark 5.29. (i) From (5.32) and (5.34) we have

$$I_t^\varepsilon = Y_t^\varepsilon - \int_0^t \mu_s^\varepsilon (h \otimes 1)\,ds, \qquad t \in [0, T]. \tag{5.36}$$

Put $E := \mathbb{R}^d \times S$, $c^k := h^k \otimes 1$, $\mathcal{H}_k := \mathcal{E}_k$, $k = 1, 2, \ldots, r$, $\mathcal{G} := \mathcal{C}^\varepsilon$, and

$$\mathcal{D}(\mathcal{G}, \mathcal{H}) \equiv \mathcal{D}(\mathcal{C}^\varepsilon, \mathcal{E}) := \tilde{\mathcal{D}} \tag{5.37}$$

in Condition 5.24. Then, from (5.36), Lemma 5.28, Proposition 4.21, Remark 5.27, and Theorem 4.1 of [10], we see that μ^ε satisfies the following relation: for each $f \in \tilde{\mathcal{D}}$

we have

$$\mu_t^\varepsilon f = \mu_0^\varepsilon f + \int_0^t \mu_s^\varepsilon(\mathcal{C}^\varepsilon f)\,ds + \int_0^t \sum_{k=1}^r [\mu_s^\varepsilon((h^k \otimes 1)f + \mathcal{E}_k f) - (\mu_s^\varepsilon(h^k \otimes 1))(\mu_s^\varepsilon f)]\,d(I^\varepsilon)_s^k \tag{5.38}$$

for every $t \in [0, T]$. That is, $\{(\Omega, \mathcal{F}, \{\mathcal{F}_t^{Y^\varepsilon}\}, P), (\mu_t^\varepsilon, I_t^\varepsilon)\}$ is a solution of the normalized filter equation corresponding to $(\mathbb{R}^d \times S; \mathcal{C}^\varepsilon, \mathcal{E}, h \otimes 1)$ (in the sense of Definition 5.25).

(ii) It follows from (i) that the process $\{\mu_t^\varepsilon f\}$ is $\mathbb{R}$-valued and continuous for each $f \in \tilde{\mathcal{D}}$. Now fix some $\varphi \in C_c^\infty(\mathbb{R}^d)$. Since $1 \in \mathcal{D}(\mathcal{Q})$ with $\mathcal{Q}1 \equiv 0$ (see Condition 2.1), by (4.27) we have $f := \varphi \otimes 1 \in \tilde{\mathcal{D}}$, thus, from (5.34), we find $\pi_t^\varepsilon \varphi = \mu_t^\varepsilon f$, $\forall t \in [0, T]$, and so the $\mathbb{R}$-valued process $\{\pi_t^\varepsilon \varphi\}$ is continuous for each $\varphi \in C_c^\infty(\mathbb{R}^d)$. Then, as in Remark 5.26, we see that the $\mathcal{P}(\mathbb{R}^d)$-valued process $\{\pi_t^\varepsilon\}$ is continuous.

Remark 5.30. Since $\{(\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{P}), (\bar{X}_t, (\bar{W}_t, \bar{V}_t))\}$ is a solution of (1.11) (recall Remark 2.16) it follows from (2.18) and Itô's formula that the process $\bar{M}_t^\varphi := \varphi(\bar{X}_t) - \int_0^t \mathcal{L}\varphi(\bar{X}_s)\,ds$, $t \in [0, T]$, is an $\{\bar{\mathcal{F}}_t\}$-martingale for each $\varphi \in C_c^\infty(\mathbb{R}^d)$, where $\mathcal{L}$ is defined by (2.16), (2.15), and (2.14). Moreover, for

$$\mathcal{B}_k\varphi(x) := \sum_{j=1}^d B^{jk}(x)\,\partial_j\varphi(x), \qquad \forall x \in \mathbb{R}^d,$$
$$\forall \varphi \in \mathcal{D}(\mathcal{B}) := C_c^\infty(\mathbb{R}^d), \quad \forall k = 1, 2, \ldots, r, \tag{5.39}$$

one sees, again from Itô's formula, that $\langle \bar{M}^\varphi, \bar{W}^k \rangle_t = \int_0^t \mathcal{B}_k\varphi(X_s)\,ds$, for each $\varphi \in C_c^\infty(\mathbb{R}^d)$, $k = 1, 2, \ldots, r$, $t \in [0, T]$. Now identify $E := \mathbb{R}^d$, $c^k := h^k$, $\mathcal{H}_k := \mathcal{B}_k$, $k = 1, 2, \ldots, r$, $\mathcal{G} := \mathcal{L}$, and

$$\mathcal{D}(\mathcal{G}, \mathcal{H}) \equiv \mathcal{D}(\mathcal{L}, \mathcal{B}) := C_c^\infty(\mathbb{R}^d) \tag{5.40}$$

in Condition 5.24. It follows from Remark 5.27 and Theorem 4.1 of [10] that $\{(\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t^{\bar{Y}}\}, \bar{P}), (\bar{\pi}_t, \bar{I}_t)\}$ is a solution of the normalized filter equation corresponding to $(\mathbb{R}^d; \mathcal{L}, \mathcal{B}, h)$ (in the sense of Definition 5.25), and (see Remark 5.26) the $\mathcal{P}(\mathbb{R}^d)$-valued process $\{\bar{\pi}_t\}$ is continuous.

Remark 5.31. From Remarks 5.29(ii) and 5.30, we know that $\{(\pi_t^\varepsilon, Y_t^\varepsilon),\ t \in [0, T]\}$ and $\{(\bar{\pi}_t, \bar{Y}_t),\ t \in [0, T]\}$ are $(\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r)$-valued continuous processes. Since the mapping

$$(\nu_\cdot, V_\cdot) \in C_{\mathcal{P}(\mathbb{R}^d)\times\mathbb{R}^r}[0, T] \longmapsto \left(\nu_\cdot, V_\cdot + \int_0^\cdot \nu_s h\,ds\right) \in C_{\mathcal{P}(\mathbb{R}^d)\times\mathbb{R}^r}[0, T]$$

is easily seen to be continuous, we see from (5.32) and the Continuous Mapping Theorem (e.g. Corollary 3.1.9 of [8]) that Theorem 2.18 follows when we have shown:

Theorem 5.32. *Suppose Condition* AII *holds* (*see Remark* 2.17). *Then* $\lim_{\varepsilon\to 0+} \mathcal{L}((\pi^\varepsilon, I^\varepsilon)) = \mathcal{L}((\bar{\pi}, \bar{I}))$ *in* $\mathcal{P}(C_{\mathcal{P}(\mathbb{R}^d)\times\mathbb{R}^r}[0, T])$.

We are going to use Theorem 3.20 to establish Theorem 5.32, in a way that is very similar to the earlier application of Theorem 3.20 in the proof of Theorem 2.12. To this end, we use a martingale problem introduced by Hijab [12, page 132] which has a nice relationship with solutions of the normalized filter equation (this martingale problem is also useful for studying the Fleming–Viot process of population genetics—see [7] and the references therein). Suppose Condition 5.24, and define a linear operator $\mathbb{H}(\mathcal{G}, \mathcal{H}, c)$ on $B(\mathcal{P}(E))$ as follows: Put

$$\begin{aligned}\mathcal{D}(\mathbb{H}(\mathcal{G}, \mathcal{H}, c)) := \{\Phi \in \bar{C}(\mathcal{P}(E)) : \Phi(\nu) = H(\nu\varphi_1, \nu\varphi_2, \dots, \nu\varphi_n),\ \forall \nu \in \mathcal{P}(E),\\ \text{for some positive integer } n,\ (\varphi_i)_{i=1}^n \subset \mathcal{D}(\mathcal{G}, \mathcal{H}),\\ H \in C_c^\infty(\mathbb{R}^n)\},\end{aligned} \tag{5.41}$$

and for each $\Phi \in \mathcal{D}(\mathbb{H}(\mathcal{G}, \mathcal{H}, c))$ define (recall (5.31))

$$\begin{aligned}\mathbb{H}(\mathcal{G}, \mathcal{H}, c)(\Phi)(\nu) := \sum_{i=1}^n \partial_i H(\nu\varphi_1, \dots, \nu\varphi_n)\nu(\mathcal{G}\varphi_i)\\ + \frac{1}{2}\sum_{k=1}^r \sum_{i,j=1}^n \partial_i\partial_j H(\nu\varphi_1, \dots, \nu\varphi_n)\\ \times R_{\mathcal{H}_k}(\varphi_i, c^k, \nu) R_{\mathcal{H}_k}(\varphi_j, c^k, \nu), \qquad \forall \nu \in \mathcal{P}(E).\end{aligned} \tag{5.42}$$

Properties of $\mathcal{D}(\mathbb{H}(\mathcal{G}, \mathcal{H}, c))$ are summarized in the following lemma, which is proved in Section 8.

Lemma 5.33. *Suppose that Condition* 5.24 *holds*. *Then*: (i) $\mathcal{D}(\mathbb{H}(\mathcal{G}, \mathcal{H}, c))$ *is a subalgebra of* $\bar{C}(\mathcal{P}(E))$ *that includes constant functions*; (ii) *if* $\mathcal{D}(\mathcal{G}, \mathcal{H})$ *is separating*, *then* $\mathcal{D}(\mathbb{H}(\mathcal{G}, \mathcal{H}, c))$ *separates points in* $\mathcal{P}(E)$; (iii) *if* $\mathcal{D}(\mathcal{G}, \mathcal{H})$ *is convergence determining*, *then* $\mathcal{D}(\mathbb{H}(\mathcal{G}, \mathcal{H}, c))$ *strongly separates points in* $\mathcal{P}(E)$ (*recall Remark* 3.19(i), (iii)).

Again following Hijab [12], put

$$\begin{aligned}&\mathcal{D}(\hat{\mathbb{H}}(\mathcal{G}, \mathcal{H}, c))\\ &\quad := \operatorname{span}\{\Phi \otimes g : \Phi \in \mathcal{D}(\mathbb{H}(\mathcal{G}, \mathcal{H}, c)),\ g \in \operatorname{span}\{1, C_c^\infty(\mathbb{R}^r)\}\},\end{aligned} \tag{5.43a}$$

$$\begin{aligned}\hat{\mathbb{H}}(\mathcal{G}, \mathcal{H}, c)(\Psi)(\nu, y) := g(y)\mathbb{H}(\mathcal{G}, \mathcal{H}, c)(\Phi)(\nu)\\ + \sum_{i=1}^n \sum_{k=1}^r \partial_i H(\nu\varphi_1, \dots, \nu\varphi_n) R_{\mathcal{H}_k}(\varphi_i, c^k, \nu)\, \partial_k g(y)\\ + \tfrac{1}{2}\Phi(\nu)\Delta g(y),\\ \forall \Psi := \Phi \otimes g \in \mathcal{D}(\hat{\mathbb{H}}(\mathcal{G}, \mathcal{H}, c)), \quad (\nu, y) \in \mathcal{P}(E) \times \mathbb{R}^r.\end{aligned} \tag{5.43b}$$

Remark 5.34. Observe that $\hat{\mathbb{H}}(\mathcal{G}, \mathcal{H}, c) \subset \bar{C}(\mathcal{P}(E) \times \mathbb{R}^r) \times \bar{C}(\mathcal{P}(E) \times \mathbb{R}^r)$ in the case where $c^k \in \bar{C}(E)$, $k = 1, 2, \dots, r$, in Condition 5.24.

The following result, which is a direct consequence of Itô's formula, shows that the $(\mathcal{P}(E) \times \mathbb{R}^r)$-valued process (ν, V) in Definition 5.25 solves the martingale problem for the operator $\hat{\mathbb{H}}(\mathcal{G}, \mathcal{H}, c)$:

Lemma 5.35. *Suppose that Condition* 5.24 *holds*, *and let* $\{(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P), (\nu_t, V_t)\}$ *be a solution of the normalized filter equation corresponding to* $(E; \mathcal{G}, \mathcal{H}, c)$ (*see Definition* 5.25). *Then the process* $\hat{M}_t^{\Psi} := \Psi(\nu_t, V_t) - \int_0^t \hat{\mathbb{H}}(\mathcal{G}, \mathcal{H}, c)(\Psi)(\nu_u, V_u)\, du,\ t \in [0, T]$, *is an* $\{\mathcal{F}_t\}$-*martingale for each* $\Psi \in \mathcal{D}(\hat{\mathbb{H}}(\mathcal{G}, \mathcal{H}, c))$.

Remark 5.36. With these preliminaries in place we next establish Theorem 5.32 by verifying the conditions of Theorem 3.20. Put $E := \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r$ (which is a complete separable metric space with the Prohorov metric for $\mathcal{P}(\mathbb{R}^d)$), let $\mu \in \mathcal{P}(E)$ be given by the Dirac measure in $\mathcal{P}(\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r)$ at the point $(\delta_{x_0}, 0)$, and let $(\mathcal{A}, \mathcal{D}(\mathcal{A})) := (\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h), \mathcal{D}(\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)))$ (see (5.43)) in Theorem 3.20. Since $h^k \in \bar{C}(\mathbb{R}^d)$, $k = 1, \ldots, r$ (Condition 2.15), it follows from Remark 5.34 that $\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)$ is a linear operator on $\bar{C}(\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r)$.

Verification of (I) *in Theorem* 3.20. We need the next result, which follows from Lemma 9.48 (the elementary proof is omitted):

Lemma 5.37. *Suppose Condition* AII. *Then* $\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h) \subset \bar{C}(\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r) \times \bar{C}(\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r)$, *and* $\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)$ *is separable* (*in the supremum norm of* $\bar{C}(\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r) \times \bar{C}(\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r)$).

Thus there is a sequence $\{\Psi_k\} \subset \mathcal{D}(\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h))$ such that the sequence $\{(\Psi_k, \hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)(\Psi_k))\}$ is a dense subset of $\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)$, and part (i) of (I) is verified when we identify $\{g_k\}$ with $\{\Psi_k\}$. As for part (ii) of (I), since $\mathcal{D}(\mathcal{L}, \mathcal{B}) = C_c^{\infty}(\mathbb{R}^d)$ (see (5.40)) is convergence determining (see Problem 3.11.11 of [8]), we see from Lemma 5.33(iii) that $\mathcal{D}(\mathbb{H}(\mathcal{L}, \mathcal{B}, h))$ strongly separates points in $\mathcal{P}(\mathbb{R}^d)$, and it is clear that $C_c^{\infty}(\mathbb{R}^r)$ strongly separates points in $\mathbb{R}^r$. Now it follows from (5.43a) that $\mathcal{D}(\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h))$ strongly separates points in $\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r$. However, $\{\Psi_k\}$ has been seen to be a dense subset of $\mathcal{D}(\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h))$, so that $\{\Psi_k\}$ strongly separates points in $\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r$.

Verification of (II) *in Theorem* 3.20. By Lemma 5.33 we know that $\mathcal{D}(\mathbb{H}(\mathcal{L}, \mathcal{B}, h))$ is an algebra that includes constant functions, hence vanishes nowhere.

Verification of (III) *in Theorem* 3.20. This is the most important of the conditions associated with Theorem 3.20, and is verified by the next result, established in Section 8:

Theorem 5.38. *Suppose that Condition* AII *holds*. *Then the martingale problem for* $\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)$ *defined by* (5.43) *is well-posed*.

Items (IV)–(VI) of Theorem 3.20 are stated for processes defined over the semi-infinite interval $[0, \infty)$ and the theorem delivers weak convergence of stochastic processes defined over $[0, \infty)$. It is evident that if (IV)–(VI) are verified for processes over the finite interval $[0, T]$, rather than over $[0, \infty)$, then Theorem 3.20 gives weak convergence of stochastic processes also defined over $[0, T]$. From now on we use this restriction of Theorem 3.20 to the interval $[0, T]$ without further mention.

Verification of (IV) *in Theorem* 3.20. As noted in Remark 5.36 we take $\mu \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r)$ to be the Dirac measure of the point $(\delta_{x_0}, 0)$. Then $\mathcal{L}(\bar{\pi}_0, \bar{I}_0) = \mu$. In Remark 5.30 we saw that $\{(\bar{\Omega}, \bar{\mathcal{F}}, \{\mathcal{F}_t^{\bar{Y}}\}, \bar{P}), (\bar{\pi}_t, \bar{I}_t)\}$ is a solution of the normalized filter equation corresponding to $(\mathbb{R}^d; \mathcal{L}, \mathcal{B}, h)$, thus, from Lemma 5.35, it follows that the martingale problem for $(\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h), \mu)$ has a solution $\{(\bar{\pi}_t, \bar{I}_t),\ t \in [0, T]\}$, whose paths are continuous, hence corlol. This verifies (IV) of Theorem 3.20 when we identify $\{X_t\}$ with $\{(\bar{\pi}_t, \bar{I}_t)\}$.

Remark 5.39. Fix an arbitrary sequence $\{\varepsilon_n\} \subset (0, 1]$ such that $\lim_n \varepsilon_n = 0$, and, to simplify the notation, let π_t^n and I_t^n stand for $\pi_t^{\varepsilon_n}$ and $I_t^{\varepsilon_n}$, respectively. Likewise, put $\mathcal{C}^n$ for $\mathcal{C}^{\varepsilon_n}$ (see (4.28)), μ_t^n for $\mu_t^{\varepsilon_n}$ (see (5.33)), and $\mathcal{F}_t^n$ for $\mathcal{F}_t^{Y^{\varepsilon_n}}$ (see (2.19)).

Verification of (V) *in Theorem* 3.20. Identify X_n in Theorem 3.20(V) with (π^n, I^n). We see from (1.10) and (5.32) that $\mathcal{L}(\pi_0^n, I_0^n) = \mu$, $\forall n = 1, 2, \ldots$. Since the I_t^n, $n = 1, 2, \ldots$, have a common Gaussian law for each t, the verification of (V) follows from the next result which is proved in Section 8:

Lemma 5.40. *Suppose that Condition* AII *holds. Then the sequence* $\{\mathcal{L}(\pi_t^n),\ n = 1, 2, \ldots\}$ *is tight in* $\mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ *for each* $t \in [0, T]$.

Verification of (VI) *in Theorem* 3.20. Here we use the functions $f^{\varepsilon,\varphi}$ furnished by Proposition 4.22, and the $\mathcal{P}(\mathbb{R}^d \times S)$-valued process $\{\mu_t^\varepsilon, t \in [0, T]\}$ given by (5.33). In keeping with Remark 5.39, we simplify the notation and put $f^{n,\varphi}$ for $f^{\varepsilon_n,\varphi}$ (see (4.29)). Fix arbitrary $\Psi \in \mathcal{D}(\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h))$ (see (5.43a)) of the form $\Psi(\nu, y) = (\Phi \otimes g)(\nu, y)$, $\forall(\nu, y) \in \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r$, for some $\Phi \in \mathcal{D}(\mathbb{H}(\mathcal{L}, \mathcal{B}, h))$ and $g \in \mathrm{span}\{1, C_c^\infty(\mathbb{R}^r)\}$, with Φ given by

$$\Phi(\nu) := H(\nu\varphi_1, \ldots, \nu\varphi_m), \qquad \nu \in \mathcal{P}(\mathbb{R}^d), \tag{5.44}$$

for some positive integer m, some $H \in C_c^\infty(\mathbb{R}^m)$, $\varphi_i \in C_c^\infty(\mathbb{R}^d)$, $i = 1, 2, \ldots, m$ (see (5.41)). For each positive integer n define

$$\Phi_n(\mu) := H(\mu f^{n,\varphi_1}, \ldots, \mu f^{n,\varphi_m}), \qquad \mu \in \mathcal{P}(\mathbb{R}^d \times S), \tag{5.45}$$

where the $f^{n,\varphi_i} \in \tilde{\mathcal{D}}$ are given by Proposition 4.22 with $\varepsilon := \varepsilon_n$ and $\varphi := \varphi_i$. Then $f^{n,\varphi_i} \in \mathcal{D}(\mathcal{C}^n, \mathcal{E})$ (by (5.37)), from which we find $\Phi_n \in \mathcal{D}(\mathbb{H}(\mathcal{C}^n, \mathcal{E}, h \otimes 1))$, $\forall n = 1, 2, \ldots$ (recall (5.41)). In view of (5.43a), we then have

$$\Phi_n \otimes g \in \mathcal{D}(\hat{\mathbb{H}}(\mathcal{C}^n, \mathcal{E}, h \otimes 1)), \qquad \forall n = 1, 2, \ldots. \tag{5.46}$$

For each $t \in [0, T]$, $n = 1, 2, \ldots$, put

$$\begin{aligned} U_n(t) &:= (\Phi_n \otimes g)(\mu_t^n, I_t^n), \\ V_n(t) &:= \hat{\mathbb{H}}(\mathcal{C}^n, \mathcal{E}, h \otimes 1)(\Phi_n \otimes g)(\mu_t^n, I_t^n). \end{aligned} \tag{5.47}$$

Now Condition 5.24 holds for $E := \mathbb{R}^d \times S$, $\mathcal{G} := \mathcal{C}^n$, $\mathcal{H}_k := \mathcal{E}_k$, and $c := h \otimes 1$, and, from Remark 5.29(i), we know that $\{(\Omega, \mathcal{F}, \{\mathcal{F}_t^n\}, P), (\mu_t^n, I_t^n)\}$ is a solution of the normalized filter equation corresponding to $(\mathbb{R}^d \times S; \mathcal{C}^n, \mathcal{E}, h \otimes 1)$. Thus, Lemma 5.35 shows that the $(\mathcal{P}(\mathbb{R}^d \times S) \times \mathbb{R}^r)$-valued process $\{(\mu_t^n, I_t^n),\ t \in [0, T]\}$ is a solution of the martingale problem for $\hat{\mathbb{H}}(\mathcal{C}^n, \mathcal{E}, h \otimes 1)$, and hence (5.46) and (5.47) show that $U_n(t) - \int_0^t V_n(s)\, ds,\ t \in [0, T]$, is an $\{\mathcal{F}_t^n\}$-martingale, as required for (3.23). To verify the remaining conditions (3.24)–(3.26) we need the following fact, the elementary proof of which is omitted:

Fact 5.41. *With reference to* (5.44), (5.45), *and* (5.47), *we have*

(a) $\sup_{t\in[0,T],\omega\in\Omega} |\Phi_n(\mu_t^n) - \Phi(\pi_t^n)| \le C\varepsilon_n, \forall n = 1, 2, \ldots,$ *for a constant* $C \in [0, \infty)$;
(b) $\sup_{n,\omega,t} |V_n(t)| < \infty$;
(c) $\sup_{n,\omega} |V_n(t) - \hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)(\Psi)(\pi_t^n, I_t^n)| < C\,\varepsilon_n, \forall t \in [0, T]$, *for a constant* $C \in [0, \infty)$.

From (5.47) we find $U_n(t) - \Psi(\pi_t^n, I_t^n) = [\Phi_n(\mu_t^n) - \Phi(\pi_t^n)]g(I_t^n), \forall t \in [0, T]$, $n = 1, 2, \ldots,$ and (3.24) follows from this, together with Fact 5.41(a) and uniform boundedness of $g(\cdot)$. Now (3.25) is an immediate consequence of Fact 5.41(b) and (3.26) follows from Fact 5.41(c) and the Dominated Convergence Theorem. This verifies (VI) of Theorem 3.20, from which it follows that $\lim_{n\to\infty} \mathcal{L}((\pi^n, I^n)) = \mathcal{L}((\bar{\pi}, \bar{I}))$ in $\mathcal{P}(D_{\mathcal{P}(\mathbb{R}^d)\times\mathbb{R}^r}[0, T])$. However, $\{(\pi_t^n, I_t^n),\ t \in [0, T]\}$ and $\{(\bar{\pi}_t, \bar{I}_t),\ t \in [0, T]\}$ are *continuous* $(\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r)$-valued processes (see Remarks 5.29(ii) and 5.30), thus this convergence takes place in $\mathcal{P}(C_{\mathcal{P}(\mathbb{R}^d)\times\mathbb{R}^r}[0, T])$ (see Corollary 3.3.2 of [8]). Now Theorem 5.32 follows by the arbitrary choice of $\{\varepsilon_n\}$ in Remark 5.39, and, as noted in Remark 5.31, this gives Theorem 2.18. □

6. Discussion

In this section we compare Theorem 2.18 with some results on convergence of nonlinear filters that have recently been established.

Bhatt et al. [1] study the nonlinear filter for the observation equation

$$Y_t = W_t + \int_0^t h(X_s)\, ds,$$

in the case where the $\mathbb{R}^r$-valued Wiener process W is independent of the signal X, which takes values in a complete separable metric space E, and the sensor mapping $h\colon E \to \mathbb{R}^r$ is subject to only weak integrability conditions. The main result of [1] establishes that the nonlinear filter of the signal X based on the observation Y depends continuously on the law of X, and is achieved by an elegant argument that involves applying only elementary ideas from integration theory and stochastic analysis to the Kallianpur–Striebel formula for the nonlinear filter; in particular, there is no need to postulate a model for X (e.g. that X be a diffusion or a Markov process), and no use is made of the normalized or unnormalized filter equations.

In a subsequent work, Bhatt and Karandikar [4] extend the general approach of [1] to the case in which the signal X and observation process Y are correlated. In contrast to the independence case of [1], here it is necessary to postulate a model for the signal X and its relation to the observation process Y. We suppose that these are given by

$$dX_t = a(X_t)\,dW_t^1 + b(X_t)\,dY_t + g(X_t)\,dt,$$

$$X_0 = \text{nonrandom } \; x_0 \in \mathbb{R}^d, \tag{6.48}$$

$$dY_t = h(X_t)\,dt + dW_t^2, \qquad Y_0 = 0 \in \mathbb{R}^r, \tag{6.49}$$

where (W_t^1, W_t^2) is an $\mathbb{R}^{d+r}$-valued standard Wiener process on $(\Omega, \mathcal{F}, P)$, and the coefficients $a(\cdot)$, $b(\cdot)$, $g(\cdot)$, and $h(\cdot)$ are globally Lipschitz continuous. Suppose, also, that there is a signal/observation pair (X_t^n, Y_t^n) given by

$$dX_t^n = a^n(X_t^n)\,dW_t^{n,1} + b^n(X_t^n)\,dY_t^n + g^n(X_t^n)\,dt,$$

$$X_0^n = \text{nonrandom } \; x_0 \in \mathbb{R}^d, \tag{6.50}$$

$$dY_t^n = h(X_t^n)\,dt + dW_t^{n,2}, \qquad Y_0^n = 0 \in \mathbb{R}^r, \tag{6.51}$$

where $(W_t^{n,1}, W_t^{n,2})$ is an $\mathbb{R}^{d+r}$-valued standard Wiener process on $(\Omega^n, \mathcal{F}^n, P^n)$, and the coefficients $a^n(\cdot)$, $b^n(\cdot)$, and $g^n(\cdot)$ are again globally Lipschitz continuous. Let $\{\pi_t\}$ and $\{\pi_t^n\}$ be the $\mathcal{P}(\mathbb{R}^d)$-valued processes which are, respectively, the nonlinear filter of the signal X given the observation process Y, and the nonlinear filter of the signal X^n given the observation process Y^n. If the coefficients $a^n(\cdot)$ converge to $a(\cdot)$ uniformly on compacta, and likewise for the remaining coefficients $b^n(\cdot)$, $g^n(\cdot)$, and the Lipschitz constants associated with $a^n(\cdot)$, $b^n(\cdot)$, $g^n(\cdot)$ are uniform with respect to n, then one sees that the signal/observation pair (X^n, Y^n) converges weakly to the signal/observation pair (X, Y) (see Theorem 5.1 of [4]), and it then becomes natural to try to establish weak convergence of π^n to π. The key technical result used in [4] to show this convergence are Kallianpur–Striebel representations for π^n and π, which have the following form: in view of the Lipschitz continuity of the coefficients in (6.48) and (6.50), there are nonanticipative mappings $\mathfrak{e}, \mathfrak{e}^n\colon\ C_{\mathbb{R}^d}[0,T] \times C_{\mathbb{R}^r}[0,T] \to C_{\mathbb{R}^d}[0,T]$ which give pathwise representations for X and X^n in terms of the "driving" processes (W^1, Y) and $(W^{n,1}, Y^n)$, namely

$$X_t = \mathfrak{e}(W^1, Y)(t), \quad P\text{-a.e.} \qquad \text{and} \qquad X_t^n = \mathfrak{e}^n(W^{n,1}, Y^n)(t), \quad P^n\text{-a.e.} \tag{6.52}$$

(see Theorem 4.3 of [14]). Then, subject to only mild boundedness restrictions on $h(\cdot)$, it follows from the Girsanov theorem (see pp. 44 and 45 of [4]) that, for each $f \in \bar{C}(\mathbb{R}^d)$, one has the Kallianpur–Striebel representations

$$\pi_t^n f = \frac{\tilde{\sigma}_t^n(f, Y^n)}{\tilde{\sigma}_t^n(1, Y^n)} \quad \text{and} \quad \pi_t f = \frac{\tilde{\sigma}_t(f, Y)}{\tilde{\sigma}_t(1, Y)}, \tag{6.53}$$

where, for arbitrary $w \in C_{\mathbb{R}^d}[0,T]$, $y \in C_{\mathbb{R}^r}[0,T]$, we have defined

$$\tilde{\sigma}_t^n(f, y) := \int_{C_{\mathbb{R}^d}[0,T]} \chi_t^n(f, w, y)\, dQ_1(w), \tag{6.54}$$

$$\chi_t^n(f, w, y) := f(\mathfrak{e}^n(w, y)(t))\tilde{q}_t^n(w, y), \tag{6.55}$$

$$\tilde{q}_t^n(w, y) := \exp\left\{\int_0^t h'(\mathfrak{e}^n(w, y)(s))\; dy(s) - \frac{1}{2}\int_0^t |h(\mathfrak{e}^n(w, y)(s))|^2\, ds\right\}, \tag{6.56}$$

with Q_1 indicating Wiener measure on $C_{\mathbb{R}^d}[0,T]$, and $\tilde{\sigma}_t(f,y)$ being defined in the same way as $\tilde{\sigma}^n_t(f,y)$, but with the superscript "n" suppressed everywhere in (6.54)–(6.56). From these representations one sees that if convergence of the integrand $\chi^n(f,\cdot,\cdot)$ in (6.54) to the limiting integrand $\chi(f,\cdot,\cdot)$ can be established in a sufficiently strong sense, then we should have convergence of the functional $\tilde{\sigma}^n(f,\cdot)$ to the functional $\tilde{\sigma}(f,\cdot)$, which, in view of the ratios in (6.53), should lead to weak convergence of π^n to π. Indeed, for an arbitrary sequence $t_n \to t \in [0,T]$, it follows easily from Scheffé's lemma that

$$\lim_{n\to\infty} \int_{C_{\mathbb{R}^r}[0,T]} \left\{ \int_{C_{\mathbb{R}^d}[0,T]} |\chi^n_{t_n}(f,w,y) - \chi_t(f,w,y)| \, dQ_1(w) \right\} dQ_2(y) = 0, \tag{6.57}$$

where Q_2 denotes the Wiener measure on $C_{\mathbb{R}^r}[0,T]$. This is the essential step in establishing weak convergence of π^n, since, from (6.57), (6.54), and Fubini's theorem, we see that the functional $\tilde{\sigma}^n_{t_n}(f,\cdot)$ converges to the functional $\tilde{\sigma}_t(f,\cdot)$ in Q_2-probability, and weak convergence of π^n to π follows from this latter convergence and easy integration theory (see pp. 50 and 51 in [4]).

Remark 6.42. Note that Bhatt and Karandikar [4] use models for (X,Y) and (X^n,Y^n) which are more general than the pairs (6.48), (6.49) and (6.50), (6.51) that we postulate here—see the pairs (2.1a), (2.1b) and (5.1a), (5.1b) in [4]—and also allow random initial conditions X_0 and X^n_0, with X^n_0 converging weakly to X_0. However, to focus on the essential ideas and facilitate comparison with Theorem 2.18, we limit attention to the model (6.48)–(6.51).

The method just summarized completely eschews the nonlinear filter equations and martingale problems, together with their rather complicated technical machinery, and raises the question of whether one can establish Theorem 2.18 by the same simple and elegant approach. To explore this possibility we next establish a Kallianpur–Striebel formula for the nonlinear filter π^ε_t (recall (2.20)) in much the same spirit as the representation given by (6.53)–(6.56).

Remark 6.43. For present and later use put

$$\mathcal{M}_t := \sigma\{Z_u, W_s,\ u \in [0,\infty), s \in [0,t]\} \vee \mathcal{N}(P),$$

$\forall t \in [0,\infty)$, and note from the independence of Z and W (Condition 2.5) that $\{W_t\}$ is an $\{\mathcal{M}_t\}$-Wiener process on $(\Omega,\mathcal{F},P)$.

For each $\varepsilon \in (0,1]$ and $t \in [0,\infty)$ put

$$q^\varepsilon_t := \exp\left\{ \int_0^t h'(X^\varepsilon_s)\, dY^\varepsilon_s - \frac{1}{2}\int_0^t |h(X^\varepsilon_s)|^2\, ds \right\}, \tag{6.58}$$

where X^ε and Y^ε are given by (1.9) and (1.5), respectively. Then $E[(q_T^\varepsilon)^{-1}] = 1$ (by Condition 2.15), so that

$$\frac{dP^\varepsilon}{dP} := (q_T^\varepsilon)^{-1} \tag{6.59}$$

defines a probability measure P^ε on $(\Omega, \mathcal{F})$, with P and P^ε having identical null events. Then, from the Girsanov theorem, we see that $\{(Y^\varepsilon, \mathcal{M}_t),\ t \in [0, T]\}$ is an $\mathbb{R}^r$-valued standard Wiener process on $(\Omega, \mathcal{F}, P^\varepsilon)$, and therefore

(i) $\{(q_t^\varepsilon, \mathcal{M}_t),\ t \in [0, T]\}$ is a martingale on $(\Omega, \mathcal{F}, P^\varepsilon)$,
(ii) Z and Y^ε are P^ε-independent,
(iii) $P^\varepsilon \circ Z^{-1} = P \circ Z^{-1}$,

for each $\varepsilon \in (0, 1]$ ((ii) and (iii) follow since $\sigma\{Z_t,\ t \in [0, \infty)\} = \mathcal{M}_0$). Moreover, from (i), (6.59), (2.20), and the Bayes formula (see, e.g., Lemma 7.1, p. 243 of [21]), for each $f \in \bar{C}(\mathbb{R}^d)$ we have

$$\pi_t^\varepsilon f = \frac{\sigma_t^\varepsilon f}{\sigma_t^\varepsilon 1}, \quad P^\varepsilon\text{-a.e.}, \tag{6.60}$$

where

$$\sigma_t^\varepsilon f := \mathrm{E}^{P^\varepsilon}[f(X_t^\varepsilon) q_t^\varepsilon | \mathcal{F}_t^{Y^\varepsilon}]. \tag{6.61}$$

In view of (1.9), (1.1), and a straightforward extension of Theorem 4.3 of [14], for each $\varepsilon \in (0, 1]$ we get a mapping $\mathfrak{e}^\varepsilon\colon D_S[0, \infty) \times C_{\mathbb{R}^r}[0, T] \to C_{\mathbb{R}^d}[0, T]$ such that $(z, y) \to \mathfrak{e}^\varepsilon(z, y)(t)$ is $\mathcal{B}_{t/\varepsilon^2}\{D_S[0, \infty)\} \otimes \mathcal{B}_t\{C_{\mathbb{R}^r}[0, T]\}$-measurable, and we have the "pathwise" representation

$$X_t^\varepsilon = \mathfrak{e}^\varepsilon(Z, Y^\varepsilon)(t), \quad P^\varepsilon\text{-a.e.} \tag{6.62}$$

for each $t \in [0, T]$. Let Q_Z denote the distribution in $D_S[0, \infty)$ of the Markov process Z, with initial distribution μ_0 specified in Condition 2.1, and, for each $\varepsilon \in (0, 1]$, $t \in [0, T]$, $f \in \bar{C}(\mathbb{R}^d)$, $y \in C_{\mathbb{R}^r}[0, T]$, and $z \in D_S[0, \infty)$, define

$$\tilde{\sigma}_t^\varepsilon(f, y) := \int_{D_S[0,\infty)} \xi_t^\varepsilon(f, z, y)\, dQ_Z(z), \tag{6.63}$$

$$\xi_t^\varepsilon(f, z, y) := f(\mathfrak{e}^\varepsilon(z, y)(t))\, \tilde{q}_t^\varepsilon(z, y), \tag{6.64}$$

$$\tilde{q}_t^\varepsilon(z, y) := \exp\left\{\int_0^t h'(\mathfrak{e}^\varepsilon(z, y)(s))\, dy(s) - \frac{1}{2}\int_0^t |h(\mathfrak{e}^\varepsilon(z, y)(s))|^2\, ds\right\}. \tag{6.65}$$

In view of (6.58), (6.62), and (6.65), we find that $q_t^\varepsilon = \tilde{q}_t^\varepsilon(Z, Y^\varepsilon)$, P^ε-a.e., hence, from (ii), (iii), and (6.61), we have $\sigma_t^\varepsilon f = \tilde{\sigma}_t^\varepsilon(f, Y^\varepsilon)$, P^ε-a.e.; and thus, from (6.60), we have the Kallianpur–Striebel representation

$$\pi_t^\varepsilon f = \frac{\tilde{\sigma}_t^\varepsilon(f, Y^\varepsilon)}{\tilde{\sigma}_t^\varepsilon(1, Y^\varepsilon)}, \quad P^\varepsilon\text{-a.e.}, \tag{6.66}$$

where the functional $\tilde{\sigma}_t^\varepsilon(f, \cdot)$ is defined on $C_{\mathbb{R}^r}[0, T]$ by (6.63)–(6.65). We next write a Kallianpur–Striebel representation for the nonlinear filter $\bar{\pi}_t$ (see (2.22)). From (1.11) and (1.7) we have

$$d\bar{X}_t = [b(\bar{X}_t) - B(\bar{X}_t)h(\bar{X}_t)]\,dt + c(\bar{X}_t)\,d\bar{V}_t + B(\bar{X}_t)\,d\bar{Y}_t,$$
$$\bar{X}_0 = \text{nonrandom } x_0 \ \in \mathbb{R}^d. \tag{6.67}$$

We need a "pathwise" representation for $\bar{X}$ in terms of the pair $(\bar{V}, \bar{Y})$, analogous to the representations in (6.52). However, the conditions postulated for Theorem 2.18 are not quite enough for this, since these entail only uniqueness-in-distribution for the pair $(\bar{X}, \bar{Y})$ given by (1.11), (1.7), rather than pathwise uniqueness, as required for such a representation to hold. We therefore add the further condition that the mappings $b(\cdot)$, $c(\cdot)$, $B(\cdot)$, and $h(\cdot)$ be globally Lipschitz continuous, and then, from (6.67) and Theorem 4.3 of [14], we have a nonanticipative mapping $\bar{\mathfrak{e}}$: $C_{\mathbb{R}^d}[0,T]\times C_{\mathbb{R}^r}[0,T] \to C_{\mathbb{R}^d}[0,T]$ such that

$$\bar{X}_t = \bar{\mathfrak{e}}(\bar{V}, \bar{Y})(t), \quad \bar{P}\text{-a.e.} \tag{6.68}$$

Then, similarly to (6.53)–(6.56), we have the following Kallianpur–Striebel representation for $\bar{\pi}_t$: for each $t \in [0, T]$ and $f \in \bar{C}(\mathbb{R}^d)$,

$$\bar{\pi}_t f = \frac{\tilde{\sigma}_t(f, \bar{Y})}{\tilde{\sigma}_t(1, \bar{Y})}, \tag{6.69}$$

where, for arbitrary $v \in C_{\mathbb{R}^d}[0,T]$, $y \in C_{\mathbb{R}^r}[0,T]$, we have defined

$$\tilde{\sigma}_t(f, y) := \int_{C_{\mathbb{R}^d}[0,T]} \xi_t(f, v, y)\, dQ_1(v), \tag{6.70}$$

$$\xi_t(f, v, y) := f(\bar{\mathfrak{e}}(v, y)(t))\tilde{q}_t(v, y), \tag{6.71}$$

$$\tilde{q}_t(v, y) := \exp\left\{\int_0^t h'(\bar{\mathfrak{e}}(v, y)(s))\, dy(s) - \frac{1}{2}\int_0^t |h(\bar{\mathfrak{e}}(v, y)(s))|^2\, ds\right\}, \tag{6.72}$$

with Q_1 indicating Wiener measure on $C_{\mathbb{R}^d}[0,T]$. The goal now is to try to use the Kallianpur–Striebel representations (6.63)–(6.66) and (6.69)–(6.72) to show weak convergence of π^ε to $\bar{\pi}$ as $\varepsilon \to 0$. To this end, fix sequences $\{\varepsilon_n\} \subset (0, 1]$ and $\{t_n\} \subset [0, T]$ such that $\varepsilon_n \to 0$ and $t_n \to t$. Then, as in [4], it is enough to show that the functional $\tilde{\sigma}_{t_n}^{\varepsilon_n}(f, \cdot)$ (see (6.63)) converges to the functional $\tilde{\sigma}_t(f, \cdot)$ (see (6.70)) in Q_2-probability (where Q_2 denotes the Wiener measure in $C_{\mathbb{R}^r}[0,T]$), in order to conclude weak convergence of π^{ε_n} to $\bar{\pi}$. However, we now encounter a fundamental difficulty in following the method of [4], which relied on (6.57) to secure convergence in the Q_2-measure of the functional $\tilde{\sigma}_{t_n}^n(f, \cdot)$ (see (6.54)) to $\tilde{\sigma}_t(f, \cdot)$, since, in the present case, the analogue of (6.57) does not even make sense. This is because the integrands $\xi_{t_n}^{\varepsilon_n}(f, z, y)$ and $\xi_t(f, v, y)$ in (6.63) and (6.70) are defined on *different* domains, namely $(z, y) \in D_S[0, \infty) \times C_{\mathbb{R}^r}[0,T]$ and $(v, y) \in C_{\mathbb{R}^d}[0,T] \times C_{\mathbb{R}^r}[0,T]$, respectively, so that one does not have available a double-integral of the kind appearing in (6.57), where the integrands $\chi_{t_n}^n(f, w, y)$ and $\chi_t(f, w, y)$ are defined on, and can be integrated over, the *common* domain $(w, y) \in C_{\mathbb{R}^d}[0,T] \times C_{\mathbb{R}^r}[0,T]$. The reason for this is to be found in the dynamics of (6.48) (for X) and (6.50) (for X^n), in which the "driving" random pairs (W_t^1, Y_t) and $(W_t^{n,1}, Y_t^n)$ have sample-paths in the *same* space $C_{\mathbb{R}^d}[0,T] \times C_{\mathbb{R}^r}[0,T]$, while on the other hand, in (6.67) (for $\bar{X}$) and (1.9) (for X^ε), the driving random pairs

$(\bar{V}_t, \bar{Y}_t)$ and (Z_t, Y_t^ε) have sample-paths in the *different* spaces $C_{\mathbb{R}^d}[0,T] \times C_{\mathbb{R}^r}[0,T]$ and $D_S[0,\infty) \times C_{\mathbb{R}^r}[0,T]$, respectively. This seems to make it rather difficult to establish Theorem 2.18 by the elegant method of Bhatt and Karandikar [4], and, at least at our current level of understanding, appears to necessitate the martingale-problem approach taken here.

7. Appendix for Section 4

For each $f \in \tilde{\mathcal{D}}$, put

$$(\mathcal{V}f)(x, z_1, z_2) := \sum_{i=1}^{d} \left[\frac{1}{\varepsilon} F^i(x, z_1) + G^i(x, z_1) \right] \partial_i f(x, z_2) + \frac{1}{2} \sum_{i,j=1}^{d} [BB^T(x)]^{ij} \partial_i \partial_j f(x, z_2), \tag{7.73}$$

for all $(x, z_1, z_2) \in \mathbb{R}^d \times S \times S$, and observe from (4.28) and (7.73) that

$$(\mathcal{C}^\varepsilon f)(x, z) = (\mathcal{V}f)(x, z, z) + \varepsilon^{-2} \mathcal{Q}[f(x, \cdot)](z), \qquad \forall (x, z) \in \mathbb{R}^d \times S. \tag{7.74}$$

Proof of Proposition 4.21. Fix some $\varepsilon \in (0, 1]$ and $f \in \tilde{\mathcal{D}}$. Since $\{(W_t, \mathcal{M}_t)\}$ is a Wiener process on $(\Omega, \mathcal{F}, P)$ (Remark 6.43), it easily follows from (1.10), (7.73), and Itô's formula that, for each $z \in S$,

$$f(X_t^\varepsilon, z) - \int_0^t (\mathcal{V}f)(X_s^\varepsilon, Z_s^\varepsilon, z)\, ds \qquad \text{is an} \quad \{\mathcal{M}_t\}\text{-martingale},$$

thus, for $0 \le t_1 < t_2 < \infty$ and $z \in S$, we have

$$\mathrm{E}\left[f(X_{t_2}^\varepsilon, z) - f(X_{t_1}^\varepsilon, z) - \int_{t_1}^{t_2} (\mathcal{V}f)(X_s^\varepsilon, Z_s^\varepsilon, z)\, ds \,\middle|\, \mathcal{M}_{t_1} \right] = 0 \quad \text{a.s.} \tag{7.75}$$

By Problem 1.5.7 of [20] and the $\mathcal{M}_{t_1}$-measurability of $Z_{t_2}^\varepsilon$, we see that (7.75) still holds when z is replaced with $Z_{t_2}^\varepsilon$, and then, since $\mathcal{F}_t^{W,Z^\varepsilon} \subset \mathcal{M}_t$ (for $\{\mathcal{F}_t^{W,Z^\varepsilon}\}$ defined in Remark 2.8), we get

$$\mathrm{E}\left[f(X_{t_2}^\varepsilon, Z_{t_2}^\varepsilon) - f(X_{t_1}^\varepsilon, Z_{t_2}^\varepsilon) - \int_{t_1}^{t_2} (\mathcal{V}f)(X_s^\varepsilon, Z_s^\varepsilon, Z_{t_2}^\varepsilon)\, ds \,\middle|\, \mathcal{F}_{t_1}^{W,Z^\varepsilon} \right] = 0 \quad \text{a.s.} \tag{7.76}$$

If $\{T_t^\varepsilon\}$ is the Feller semigroup on $C(S)$ corresponding to Markov process $\{Z_t^\varepsilon\}$ defined by (1.1), then $\varepsilon^{-2}\mathcal{Q}$ is the generator of $\{T_t^\varepsilon\}$ with domain $\mathcal{D}(\mathcal{Q})$ (see Condition 2.1). Since $f \in \tilde{\mathcal{D}}$, one has $f(x, \cdot) \in \mathcal{D}(\mathcal{Q})$, $\forall x \in \mathbb{R}^d$ (see (4.27)), and thus Proposition 4.1.7 of [8] shows that, for each $x \in \mathbb{R}^d$,

$$U_t^\varepsilon(x) := f(x, Z_t^\varepsilon) - \varepsilon^{-2} \int_0^t \mathcal{Q}[f(x, \cdot)](Z_u^\varepsilon)\, du, \qquad t \in [0, \infty) \quad \text{is an } \mathcal{F}_t^{Z^\varepsilon}\text{-martingale.} \tag{7.77}$$

However, Z and W are independent, so that for $0 \le t_1 < t_2 < \infty$, we have $\mathrm{E}[U^\varepsilon_{t_2}(x) - U^\varepsilon_{t_1}(x) \mid \mathcal{F}^{W,Z^\varepsilon}_{t_1}] = 0$ a.s. for each $x \in \mathbb{R}^d$. Since $X^\varepsilon_{t_1}$ is $\mathcal{F}^{W,Z^\varepsilon}_{t_1}$-measurable (Remark 2.8) it follows from Problem 1.5.7 of [20] that $\mathrm{E}[U^\varepsilon_{t_2}(X^\varepsilon_{t_1}) - U^\varepsilon_{t_1}(X^\varepsilon_{t_1}) \mid \mathcal{F}^{W,Z^\varepsilon}_{t_1}] = 0$ a.s. or (see (7.77))

$$\mathrm{E}\left[f(X^\varepsilon_{t_1}, Z^\varepsilon_{t_2}) - f(X^\varepsilon_{t_1}, Z^\varepsilon_{t_1}) - \varepsilon^{-2}\int_{t_1}^{t_2} \mathcal{Q}[f(X^\varepsilon_{t_1}, \cdot)](Z^\varepsilon_s)\,ds \,\middle|\, \mathcal{F}^{W,Z^\varepsilon}_{t_1}\right] = 0 \quad \text{a.s.} \tag{7.78}$$

Since (7.78) and (7.76) hold for all $0 \le t_1 < t_2 < \infty$, by Lemma 4.3.4(a) of [8] (with $E := S$, $X := Z^\varepsilon$, $\mathcal{G}_t := \mathcal{F}^{W,Z^\varepsilon}_t$, and mappings u, v: $[0,\infty) \times S \times \Omega \to \mathbb{R}$, w: $[0,\infty) \times [0,\infty) \times S \times \Omega \to \mathbb{R}$ defined by $u(t,z,\omega) := f(X^\varepsilon_t(\omega), z)$, $v(t,z,\omega) := (\mathcal{V}f)(X^\varepsilon_t(\omega), Z^\varepsilon_t(\omega), z)$, $w(t,s,z,\omega) := \varepsilon^{-2}\mathcal{Q}[f(X^\varepsilon_t(\omega),\cdot)](z)$), it follows that

$$u(t, Z^\varepsilon_t) - \int_0^t [v(s, Z^\varepsilon_s) + w(s,s,Z^\varepsilon_s)]\,ds \tag{7.79}$$

is an $\{\mathcal{F}^{W,Z^\varepsilon}_t\}$-martingale. However, from (7.74), the quantity in (7.79) is $M^{\varepsilon,f}_t$ in the statement of Proposition 4.21, as required. □

Proof of Proposition 4.22. Fix $\varepsilon \in (0,1]$ and $\varphi \in \mathcal{D}(\mathcal{L}) := C^\infty_{\mathrm{c}}(\mathbb{R}^d)$ (recall (2.16)), and put

$$g^\varphi_1(x,z) := \sum_{i=1}^d F^i(x,z)\,\partial_i\varphi(x),$$
$$f^\varphi_1(x,z) := \int_S g^\varphi_1(x,z')\chi(z,dz'), \qquad \forall (x,z) \in \mathbb{R}^d \times S. \tag{7.80}$$

Now Condition 2.4 (see (1.3)) shows that $\bar{m}g^\varphi_1(x) = 0$, $\forall x \in \mathbb{R}^d$, so that Lemma 2.6(i) gives $f^\varphi_1(x,\cdot) \in \mathcal{D}(\mathcal{Q})$, $\forall x \in \mathbb{R}^d$, with

$$\mathcal{Q}[f^\varphi_1(x,\cdot)](z) = -g^\varphi_1(x,z), \qquad \forall (x,z) \in \mathbb{R}^d \times S. \tag{7.81}$$

From (7.80) with Condition 2.4 and Lemma 2.6(ii), we see that $g^\varphi_1, f^\varphi_1 \in C^{3,0}_{\mathrm{c}}(\mathbb{R}^d \times S)$. Put

$$g^\varphi_2(x,z) := \sum_{i=1}^d F^i(x,z)\,\partial_i f^\varphi_1(x,z) + \sum_{i=1}^d G^i(x,z)\,\partial_i\varphi(x) + \frac{1}{2}\sum_{i,j=1}^d [BB^T(x)]^{ij}\,\partial_i\partial_j\varphi(x), \tag{7.82}$$

$$f^\varphi_2(x,z) := \int_S [g^\varphi_2(x,z') - \bar{m}g^\varphi_2(x)]\chi(z,dz'), \qquad \forall (x,z) \in \mathbb{R}^d \times S. \tag{7.83}$$

Then Lemma 2.6(i) gives $f^\varphi_2(x,\cdot) \in \mathcal{D}(\mathcal{Q})$, $\forall x \in \mathbb{R}^d$, with

$$\mathcal{Q}[f^\varphi_2(x,\cdot)](z) = \bar{m}g^\varphi_2(x) - g^\varphi_2(x,z), \qquad \forall (x,z) \in \mathbb{R}^d \times S. \tag{7.84}$$

Now define $f^{\varepsilon,\varphi}$ as in (4.29). The semigroup $\{T_t\}$ is conservative, so that $(1, 0) \in \mathcal{Q}$ (see Condition 2.1), and we have seen that $f_1^{\varphi}(x,\cdot)$, $f_2^{\varphi}(x,\cdot) \in \mathcal{D}(\mathcal{Q})$, $\forall x \in \mathbb{R}^d$. Thus $f^{\varepsilon,\varphi}(x,\cdot) \in \mathcal{D}(\mathcal{Q})$, $\forall x \in \mathbb{R}^d$, and it follows from (7.81) and (7.84) that

$$\mathcal{Q}[f^{\varepsilon,\varphi}(x,\cdot)](z) = -\varepsilon g_1^{\varphi}(x,z) + \varepsilon^2[\bar{m} g_2^{\varphi}(x) - g_2^{\varphi}(x,z)]. \tag{7.85}$$

We have already seen that $g_1^{\varphi}, f_1^{\varphi} \in C_c^{3,0}(\mathbb{R}^d \times S)$. Then $\partial_i f_1^{\varphi} \in C_c^{2,0}(\mathbb{R}^d \times S)$, hence Condition 2.4 and (7.82) show that $g_2^{\varphi} \in C_c^{2,0}(\mathbb{R}^d \times S)$. Thus, by standard results on exchanging the order of integrals and derivatives (see, e.g. Theorem 2.27 in [9]), we get $\bar{m} g_2^{\varphi} \in C_c^2(\mathbb{R}^d)$. It therefore follows that $g_2^{\varphi} - \bar{m} g_2^{\varphi} \in C_c^{2,0}(\mathbb{R}^d \times S)$, and so, by (7.83) and Lemma 2.6(ii), we have $f_2^{\varphi} \in C_c^{2,0}(\mathbb{R}^d \times S)$. To summarize, we have seen that

$$g_1^{\varphi},\, f_1^{\varphi},\, g_2^{\varphi} - \bar{m} g_2^{\varphi},\;\; f_2^{\varphi} \in C_c^{2,0}(\mathbb{R}^d \times S). \tag{7.86}$$

Thus (see (4.29)) we have $f^{\varepsilon,\varphi} \in C_c^{2,0}(\mathbb{R}^d \times S)$, and (see (7.85)) the mapping $(x,z) \to \mathcal{Q}[f^{\varepsilon,\varphi}(x,\cdot)](z)$ defines a member of $C_c^{2,0}(\mathbb{R}^d \times S)$. This shows that $f^{\varepsilon,\varphi} \in \tilde{\mathcal{D}}$ (see (4.27)). We next evaluate $\mathcal{C}^{\varepsilon} f^{\varepsilon,\varphi}$. From $\mathcal{C}^{\varepsilon}$ in (4.28), (7.85), and some simplification, we find

$$\begin{aligned}\mathcal{C}^{\varepsilon} f^{\varepsilon,\varphi}(x,z) &= \frac{1}{\varepsilon}\left[\sum_i F^i(x,z)\,\partial_i\varphi(x) - g_1^{\varphi}(x,z)\right]\\ &\quad + \Bigg[\sum_i F^i(x,z)\,\partial_i f_1^{\varphi}(x,z) + \sum_i G^i(x,z)\,\partial_i\varphi(x)\\ &\qquad + \frac{1}{2}\sum_{i,j}[BB^T(x)]^{ij}\,\partial_i\partial_j\varphi(x) - g_2^{\varphi}(x,z)\Bigg]\\ &\quad + \bar{m} g_2^{\varphi}(x) + \varepsilon\gamma_1^{\varphi}(x,z) + \varepsilon^2\gamma_2^{\varphi}(x,z),\end{aligned} \tag{7.87}$$

for some $\gamma_1^{\varphi}, \gamma_2^{\varphi} \in C_c(\mathbb{R}^d \times S)$. Now the expressions in square brackets on the right-hand side cancel to zero (by (7.80) and (7.82)). An easy calculation involving Lemma 2.6(ii), (7.80), and (7.82), shows that

$$\bar{m} g_2^{\varphi}(x) = \mathcal{L}\varphi(x), \tag{7.88}$$

as required. □

Proof of Proposition 4.23. Fix a mapping $g\colon [0,\infty) \to [0,1]$ such that $g(r) := 1$, $\forall r \in [0,1]$; $g(r) := 0$, $\forall r \in [2,\infty)$; and $g(\cdot)$ has continuous derivatives of all orders. For each $k = 1, 2, \ldots$ define $g_k\colon [0,\infty) \to [0,1]$ by

$$g_k(r) := g\left(\frac{1}{k}\log[1+r]\right), \qquad \forall r \in [0,\infty), \tag{7.89}$$

and put $\varphi_k(x) := g_k(|x|)$, $\forall x \in \mathbb{R}^d$. Then $\varphi_k \in C_c^{\infty}(\mathbb{R}^d)$. For notational convenience, let f_1^k, f_2^k, γ_1^k, and γ_2^k, denote respectively the functions $f_1^{\varphi_k}$, $f_2^{\varphi_k}$, $\gamma_1^{\varphi_k}$, and $\gamma_2^{\varphi_k}$ in Proposition 4.22; let $f^{k,n}$ denote the member of $\tilde{\mathcal{D}}$ given by (4.29) with $\varepsilon := \varepsilon_n$, $\varphi = \varphi_k$;

and let $M_t^{k,n}$ denote the martingale in Proposition 4.21 when $\varepsilon := \varepsilon_n$ and $f := f^{k,n}$, so that

$$M_t^{k,n} := \varphi_k(X_t^{\varepsilon_n}) + \varepsilon_n f_1^k(X_t^{\varepsilon_n}, Z_t^{\varepsilon_n}) + \varepsilon_n^2 f_2^k(X_t^{\varepsilon_n}, Z_t^{\varepsilon_n}) - \int_0^t [\mathcal{L}\varphi_k(X_s^{\varepsilon_n}) + \varepsilon_n \gamma_1^k(X_s^{\varepsilon_n}, Z_s^{\varepsilon_n}) + \varepsilon_n^2 \gamma_2^k(X_s^{\varepsilon_n}, Z_s^{\varepsilon_n})]\, ds. \tag{7.90}$$

Now fix some $t \in [0, \infty)$, some (small) $\eta \in (0, 1)$, and some positive integer k_0 such that $\varphi_k(x_0) = 1$, $\forall k \geq k_0$ (for x_0 in (1.10)). In view of the bounds on $b^i(\cdot)$ and $a^{ij}(\cdot)$ in Remark 2.9, an easy calculation shows that $\lim_{k\to\infty} \|\mathcal{L}\varphi_k\| = 0$, thus fix integer $k_1 := k_1(\eta) \geq k_0$ such that

$$\left| \int_0^t \mathcal{L}\varphi_{k_1}(X_s^{\varepsilon_n})\, ds \right| < \eta, \qquad \forall n = 1, 2, \ldots. \tag{7.91}$$

Since $f_1^{k_1}, f_2^{k_1}, \gamma_1^{k_1}, \gamma_2^{k_1} \in \bar{C}(\mathbb{R}^d \times S)$ (see Proposition 4.22), it follows from (7.90) that

$$M_t^{k_1,n} = \varphi_{k_1}(X_t^{\varepsilon_n}) + O(\varepsilon_n) - \int_0^t \mathcal{L}\varphi_{k_1}(X_s^{\varepsilon_n})\, ds, \qquad n = 1, 2, \ldots. \tag{7.92}$$

Then, from (7.92), (7.91), and the facts that $\varphi_{k_1}(x_0) = 1$ and $\mathrm{E}[M_t^{k_1,n}] = \mathrm{E}[M_0^{k_1,n}]$ (by Proposition 4.21), we find that $\mathrm{E}[\varphi_{k_1}(X_t^{\varepsilon_n})] \geq 1 + O(\varepsilon_n) - \eta$. Since the support of φ_{k_1} is the ball of radius $R_\eta := e^{2k_1} - 1$, we have $P[|X_t^{\varepsilon_n}| \leq R_\eta] \geq 1 + O(\varepsilon_n) - \eta$, $\forall n = 1, 2, \ldots$, and tightness of the sequence $\{X_t^{\varepsilon_n},\ n = 1, 2, \ldots\}$ follows. □

8. Appendix for Section 5

Proof of Proposition 5.28. Fix some $\varepsilon \in (0, 1]$, $f \in \tilde{\mathcal{D}}$, and $k = 1, 2, \ldots, d$. Since $\{W_t\}$ is continuous, the result follows when we establish that

$$\left\{ \left[f(X_t^\varepsilon, Z_t^\varepsilon) - \int_0^t (\mathcal{C}^\varepsilon f)(X_s^\varepsilon, Z_s^\varepsilon)\, ds \right] W_t^k \right\} - \int_0^t (\mathcal{E}_k f)(X_s^\varepsilon, Z_s^\varepsilon)\, ds \qquad \text{is an } \{\mathcal{F}_t^{W,Z^\varepsilon}\}\text{-martingale} \tag{8.93}$$

(for $\{\mathcal{F}_t^{W,Z^\varepsilon}\}$ defined in Remark 2.8), since the quantity in square braces is just $M_t^{\varepsilon,f}$. We show this by an argument similar to that used in the proof of Proposition 4.21. Since $\{(W_t, \mathcal{M}_t)\}$ is a Wiener process on $(\Omega, \mathcal{F}, P)$ (Remark 6.43), it easily follows from (1.10), (5.35), (7.73), and Itô's formula that

$$f(X_t^\varepsilon, z) W_t^k - \int_0^t [W_s^k (\mathcal{V} f)(X_s^\varepsilon, Z_s^\varepsilon, z) + (\mathcal{E}_k f)(X_s^\varepsilon, z)]\, ds$$

is an $\{\mathcal{M}_t\}$-martingale

for each $z \in S$, thus, for $0 \leq t_1 < t_2 < \infty$,

$$\mathrm{E}\Bigg[f(X^\varepsilon_{t_2}, z)W^k_{t_2} - f(X^\varepsilon_{t_1}, z)W^k_{t_1} \\ - \int_{t_1}^{t_2} [W^k_s\,(\mathcal{V} f)(X^\varepsilon_s, Z^\varepsilon_s, z) + (\mathcal{E}_k f)(X^\varepsilon_s, z)]\,ds \;\Bigg|\; \mathcal{M}_{t_1} \Bigg] = 0 \quad \text{a.s.} \tag{8.94}$$

By Problem 1.5.7 of [20] and the $\mathcal{M}_{t_1}$-measurability of $Z^\varepsilon_{t_2}$ (see Remark 6.43), it is easily seen that (8.94) still holds when z is replaced with $Z^\varepsilon_{t_2}$, and then, since $\mathcal{F}^{W,Z^\varepsilon}_{t_1} \subset \mathcal{M}_{t_1}$, we get

$$\mathrm{E}\Bigg[f(X^\varepsilon_{t_2}, Z^\varepsilon_{t_2})W^k_{t_2} - f(X^\varepsilon_{t_1}, Z^\varepsilon_{t_2})W^k_{t_1} \\ - \int_{t_1}^{t_2} [W^k_s\,(\mathcal{V} f)(X^\varepsilon_s, Z^\varepsilon_s, Z^\varepsilon_{t_2}) + (\mathcal{E}_k f)(X^\varepsilon_s, Z^\varepsilon_{t_2})]\,ds \;\Bigg|\; \mathcal{F}^{W,Z^\varepsilon}_{t_1} \Bigg] = 0 \quad \text{a.s.} \tag{8.95}$$

Now, exactly as in the proof of Proposition 4.21, for $0 \leq t_1 < t_2 < \infty$ we have (see (7.78))

$$\mathrm{E}\Bigg[f(X^\varepsilon_{t_1}, Z^\varepsilon_{t_2}) - f(X^\varepsilon_{t_1}, Z^\varepsilon_{t_1}) - \varepsilon^{-2} \int_{t_1}^{t_2} \mathcal{Q}[f(X^\varepsilon_{t_1}, \cdot)](Z^\varepsilon_s)\,ds \;\Bigg|\; \mathcal{F}^{W,Z^\varepsilon}_{t_1} \Bigg] = 0 \quad \text{a.s.}$$

and therefore

$$\mathrm{E}\Bigg[f(X^\varepsilon_{t_1}, Z^\varepsilon_{t_2})W^k_{t_1} - f(X^\varepsilon_{t_1}, Z^\varepsilon_{t_1})W^k_{t_1} - \varepsilon^{-2}\, W^k_{t_1} \int_{t_1}^{t_2} \mathcal{Q}[f(X^\varepsilon_{t_1}, \cdot)](Z^\varepsilon_s)\,ds \;\Bigg|\; \mathcal{F}^{W,Z^\varepsilon}_{t_1} \Bigg] \\ = 0 \quad \text{a.s.} \tag{8.96}$$

Since (8.95) and (8.96) hold for all $0 \leq t_1 < t_2 < \infty$, by Lemma 4.3.4(a) of [8] (with $E := S, X := Z^\varepsilon, \mathcal{G}_t := \mathcal{F}^{W,Z^\varepsilon}_t$, and mappings $u, v\colon [0,\infty) \times E \times \Omega \to \mathbb{R}$, $w\colon [0,\infty) \times [0,\infty) \times E \times \Omega \to \mathbb{R}$ defined by $u(t,z,\omega) := f(X^\varepsilon_t(\omega), z)W^k_t(\omega)$, $v(t,z,\omega) := W^k_t(\omega)(\mathcal{V} f)(X^\varepsilon_t(\omega), Z^\varepsilon_t(\omega), z) + (\mathcal{E}_k f)(X^\varepsilon_t(\omega), z)$, $w(t,s,z,\omega) := \varepsilon^{-2}\, W^k_t(\omega)\, \mathcal{Q}[f(X^\varepsilon_t(\omega), \cdot)](z)$) one sees that

$$f(X^\varepsilon_t, Z^\varepsilon_t)\, W^k_t - \int_0^t W^k_s\,(\mathcal{C}^\varepsilon f)(X^\varepsilon_s, Z^\varepsilon_s)\,ds - \int_0^t (\mathcal{E}_k f)(X^\varepsilon_s, Z^\varepsilon_s)\,ds \\ \text{is an } \{\mathcal{F}^{W,Z^\varepsilon}_t\}\text{-martingale,} \tag{8.97}$$

where we have used (7.74). Now $\{W^k_t\}$ is an $\{\mathcal{F}^{W,Z^\varepsilon}_t\}$-Wiener process (by Condition 2.5), thus

$$\int_0^t W^k_s(\mathcal{C}^\varepsilon f)(X^\varepsilon_s, Z^\varepsilon_s)\,ds - W^k_t \int_0^t (\mathcal{C}^\varepsilon f)(X^\varepsilon_s, Z^\varepsilon_s)\,ds \\ \text{is an } \{\mathcal{F}^{W,Z^\varepsilon}_t\}\text{-martingale}$$

(see Problem 2.9.22 of [8]), so that (8.93) follows from (8.97). □

Proof of Lemma 5.33. (i) Fix $\Phi_1, \Phi_2 \in \mathcal{D}(\mathbb{H}(\mathcal{G}, \mathcal{H}, c))$. Then it is clear that $\Phi_1\Phi_2 \in \mathcal{D}(\mathbb{H}(\mathcal{G}, \mathcal{H}, c))$, and it is easily checked that $\alpha_1\Phi_1 + \alpha_2\Phi_2$ has the form of a member of $\mathcal{D}(\mathbb{H}(\mathcal{G}, \mathcal{H}, c))$, $\forall\alpha_1, \alpha_2 \in \mathbb{R}$. Thus $\mathcal{D}(\mathbb{H}(\mathcal{G}, \mathcal{H}, c))$ is a subalgebra of $\bar{C}(\mathcal{P}(E))$. To see that $\mathcal{D}(\mathbb{H}(\mathcal{G}, \mathcal{H}, c))$ includes constant functions fix $\varphi \in \bar{C}(E)$ and $H \in C_c^\infty(\mathbb{R})$ such that $H(x) = 1$, $\forall|x| \le \|\varphi\|$. Then $\Phi(\nu) := H(\nu\varphi)$, $\forall\nu \in \mathcal{P}(E)$, defines a member of $\mathcal{D}(\mathbb{H}(\mathcal{G}, \mathcal{H}, c))$ and has the constant value of unity.

(ii) Suppose that $\Phi(\nu_1) = \Phi(\nu_2)$, $\forall\Phi \in \mathcal{D}(\mathbb{H}(\mathcal{G}, \mathcal{H}, c))$, for some $\nu_1, \nu_2 \in \mathcal{P}(E)$. Then $H(\nu_1\varphi) = H(\nu_2\varphi)$, $\forall H \in C_c^\infty(\mathbb{R})$, $\forall\varphi \in \mathcal{D}(\mathcal{G}, \mathcal{H})$. Since $C_c^\infty(\mathbb{R})$ separates points in $\mathbb{R}$, we get $\nu_1\varphi = \nu_2\varphi$, $\forall\varphi \in \mathcal{D}(\mathcal{G}, \mathcal{H})$, so that $\nu_1 = \nu_2$ (since $\mathcal{D}(\mathcal{G}, \mathcal{H}) \subset \bar{C}(E)$ is separating).

(iii) The argument is similar to that for (ii) but just uses the elementary fact that $C_c^\infty(\mathbb{R})$ strongly separates points in $\mathbb{R}$. □

Proof of Theorem 5.38. Fix some $\xi \in \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r$, and observe from Lemma 5.35 that there exists a corlol (indeed continuous) solution of the martingale problem for $(\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h), \delta_\xi)$. To see uniqueness of the solution we need the next two results, the first of which is established later in the present section:

Theorem 8.44. *Suppose that* $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\nu}_t, \tilde{V}_t)\}$ *is a progressively measurable solution of the martingale problem for* $\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)$. *Then there exists some* $(\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r)$*-valued* $\{\tilde{\mathcal{F}}_t\}$*-adapted process* $\{(\tilde{\nu}_t', \tilde{V}_t')\}$ *on the filtered probability space* $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P})$ *such that* $\{(\tilde{\nu}_t', \tilde{V}_t')\}$ *is a modification of* $\{(\tilde{\nu}_t, \tilde{V}_t)\}$ *and* $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\nu}_t', \tilde{V}_t' - \tilde{V}_0')\}$ *is a solution of the normalized filter equation corresponding to* $(\mathbb{R}^d; \mathcal{L}, \mathcal{B}, h)$ (*in the sense of Definition* 5.25); *namely for each* $\varphi \in C_c^\infty(\mathbb{R}^d)$ *we have a.s.*

$$\tilde{\nu}_t'\varphi = \tilde{\nu}_0'\varphi + \int_0^t \tilde{\nu}_s'(\mathcal{L}\varphi)\,ds + \sum_{k=1}^r \int_0^t R_{\mathcal{B}_k}(\varphi, h^k, \tilde{\nu}_s')\,d(\tilde{V}_s' - \tilde{V}_0')^k,$$
$$\forall t \in [0, T]. \tag{8.98}$$

Remark 8.45. Theorem 8.44 shows that any progressively measurable solution of the martingale problem for $\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)$ has a corlol modification (in fact, by Remark 5.26, the modification is actually continuous, and not just corlol).

The next result is an immediate consequence of Theorem 2.21 of [18], which establishes the property of uniqueness in law for the normalized filter equation when the signal and observation are given by the pair (1.11) and (1.7), with $b(\cdot)$, $c(\cdot)$, and $B(\cdot)$ being continuous and linearly bounded, and $h(\cdot)$ being Borel-measurable and uniformly bounded. In the problem under consideration here these hypotheses are implied by Conditions AII.

Theorem 8.46. *Suppose Conditions* AII (*see Remark* 2.17). *If* $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\eta}_t, \tilde{W}_t)\}$ *and* $\{(\hat{\Omega}, \hat{\mathcal{F}}, \{\hat{\mathcal{F}}_t\}, \hat{P}), (\hat{\eta}_t, \hat{W}_t)\}$ *are solutions of the normalized filter equation corresponding to* $(\mathbb{R}^d; \mathcal{L}, \mathcal{B}, h)$ (*in the sense of Definition* 5.25), *with* $\mathcal{L}_{\tilde{P}}(\tilde{\eta}_0) = \mathcal{L}_{\hat{P}}(\hat{\eta}_0)$, *then the* $(\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r)$*-valued processes* $\{(\tilde{\eta}_t, \tilde{W}_t)\}$ *and* $\{(\hat{\eta}_t, \hat{W}_t)\}$ *are identically distributed.*

From Theorems 8.44 and 8.46 we see that any two progressively measurable solutions of the martingale problem for $(\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h), \delta_\xi)$ have identical finite-dimensional distributions. However, we have already seen that there exists a corlol solution of the martingale problem for $(\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h), \delta_\xi)$, so it follows that the *corlol* martingale problem for $(\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h), \delta_\xi)$ is well-posed, $\forall \xi \in \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r$. Since $\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)$ has been shown to satisfy the separability condition (I) of Theorem 3.20, we then see from Theorem 2.1 of [2] that the *corlol* martingale problem for $\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)$ is well-posed. Consequently, one sees from Remark 8.45 that, for each $\mu \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r)$, the finite-dimensional distributions of a progressively measurable solution of the martingale problem for $(\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h), \mu)$ are uniquely determined, as required to establish Theorem 5.38. □

Proof of Theorem 8.44. The proof is given in a series of steps:

Step 1. Since $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\nu}_t, \tilde{V}_t)\}$ is a progressively measurable solution of the martingale problem for $\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)$, and since $\mathbb{H}(\mathcal{L}, \mathcal{B}, h)$ includes the constant functions (by Lemma 5.33(i)), it follows that $g(\tilde{V}_t) - \frac{1}{2}\int_0^t \Delta g(\tilde{V}_s)\, ds$ is an $\{\tilde{\mathcal{F}}_t\}$-martingale for each $g \in C_c^\infty(\mathbb{R}^r)$. Thus, Proposition 5.3.5 of [8] ensures that $\{\tilde{V}_t\}$ has a *continuous* modification $\{\tilde{V}'_t\}$, which is $\{\tilde{\mathcal{F}}_t\}$-adapted, thus $\{(\tilde{V}'_t - \tilde{V}'_0, \tilde{\mathcal{F}}_t)\}$ is a standard $\mathbb{R}^r$-valued Wiener process.

Step 2. We next show that the $\mathcal{P}(\mathbb{R}^d)$-valued process $\{\tilde{\nu}_t\}$ has a $\mathcal{P}(\mathbb{R}^d)$-valued corlol modification. Let $\mathbb{R}^{d*}$ be the one-point compactification of $\mathbb{R}^d$, with the point at infinity denoted by Δ, and let $\mathcal{L}^\Delta$ and $\mathcal{B}^\Delta$ be linear operators on $\bar{C}(\mathbb{R}^{d*})$, with common domain $\mathcal{D}(\mathcal{L}^\Delta, \mathcal{B}^\Delta)$, defined by

$$\mathcal{D}(\mathcal{L}^\Delta, \mathcal{B}^\Delta) := \{\varphi \in \bar{C}(\mathbb{R}^{d*}) : \varphi|_{\mathbb{R}^d} \in C_c^\infty(\mathbb{R}^d),\ \varphi(\Delta) = 0\}, \tag{8.99a}$$

$$\mathcal{L}^\Delta := \{(\varphi, \psi) \in \mathcal{D}(\mathcal{L}^\Delta, \mathcal{B}^\Delta) \times \bar{C}(\mathbb{R}^{d*}) : (\varphi|_{\mathbb{R}^d}, \psi|_{\mathbb{R}^d}) \in \mathcal{L},\ \varphi(\Delta) = \psi(\Delta) = 0\}, \tag{8.99b}$$

$$\mathcal{B}^\Delta := \{(\varphi, \psi) \in \mathcal{D}(\mathcal{L}^\Delta, \mathcal{B}^\Delta) \times \bar{C}(\mathbb{R}^{d*}) : (\varphi|_{\mathbb{R}^d}, \psi|_{\mathbb{R}^d}) \in \mathcal{B},\ \varphi(\Delta) = \psi(\Delta) = 0\} \tag{8.99c}$$

(recall (2.16) and (5.39)). Also, define (Borel-measurable) $h^\Delta\colon \mathbb{R}^{d*} \to \mathbb{R}^r$ by

$$h^\Delta(x) := h(x), \quad \forall x \in \mathbb{R}^d, \qquad h^\Delta(\Delta) := 0. \tag{8.100}$$

Then we have (A) that $\mathbb{H}(\mathcal{L}^\Delta, \mathcal{B}^\Delta, h^\Delta) \subset \bar{C}(\mathcal{P}(\mathbb{R}^{d*})) \times B(\mathcal{P}(\mathbb{R}^{d*}))$. From Problem 5.4.25 of [15] it easily follows that $\mathcal{D}(\mathcal{L}^\Delta, \mathcal{B}^\Delta) \subset \bar{C}(\mathbb{R}^{d*})$ is separating, and therefore Lemma 5.33 shows that $\mathcal{D}(\mathbb{H}(\mathcal{L}^\Delta, \mathcal{B}^\Delta, h^\Delta))$ is an algebra in $\bar{C}(\mathcal{P}(\mathbb{R}^{d*}))$ which includes constant functions and separates points in $\mathcal{P}(\mathbb{R}^{d*})$. Since $\mathcal{P}(\mathbb{R}^{d*})$ is compact, the Stone–Weierstrass theorem establishes that $\mathcal{D}(\mathbb{H}(\mathcal{L}^\Delta, \mathcal{B}^\Delta, h^\Delta))$ is a dense subset of $\bar{C}(\mathcal{P}(\mathbb{R}^{d*}))$, and therefore we have (B) the set $\mathcal{D}(\mathbb{H}(\mathcal{L}^\Delta, \mathcal{B}^\Delta, h^\Delta))$ is separating. Moreover, since $\mathcal{P}(\mathbb{R}^{d*})$ is a compact metric space, we know that $\bar{C}(\mathcal{P}(\mathbb{R}^{d*}))$ is separable, and thus, since $\mathcal{D}(\mathbb{H}(\mathcal{L}^\Delta, \mathcal{B}^\Delta, h^\Delta))$ has been seen to be a dense subset of $\bar{C}(\mathcal{P}(\mathbb{R}^{d*}))$, it follows (C) that $\mathcal{D}(\mathbb{H}(\mathcal{L}^\Delta, \mathcal{B}^\Delta, h^\Delta))$ contains a countable subset which separates points in $\mathcal{P}(\mathbb{R}^{d*})$. Now $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\nu}_t, \tilde{V}_t)\}$ is a progressively measurable solution of the martingale problem for $\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)$, from which we see that $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\nu}_t)\}$ is a progressively measurable solution of the martingale problem for $\mathbb{H}(\mathcal{L}, \mathcal{B}, h)$ (take $g \equiv 1$ in (5.43)). We regard $\{\tilde{\nu}_t\}$ as a $\mathcal{P}(\mathbb{R}^{d*})$-valued process with $\tilde{\nu}_t(\Delta) = 0$, $\forall t$, and then it follows from (8.99) that (D) $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\nu}_t)\}$ solves the martingale problem for $\mathbb{H}(\mathcal{L}^\Delta, \mathcal{B}^\Delta, h^\Delta)$. In view of (A)–(D), compactness of $\mathcal{P}(\mathbb{R}^{d*})$, and Theorem 4.3.6 of [8] (with $E := \mathcal{P}(\mathbb{R}^{d*})$, $A := \mathbb{H}(\mathcal{L}^\Delta, \mathcal{B}^\Delta, h^\Delta)$, and $X := \tilde{\nu}$), there exists a $\mathcal{P}(\mathbb{R}^{d*})$-valued corlol modification of $\{\tilde{\nu}_t\}$, which we denote by $\{\tilde{\nu}'_t\}$. In particular, since $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P})$ is a complete filtered probability space, the corlol process $\{\tilde{\nu}'_t\}$ is $\{\tilde{\mathcal{F}}_t\}$-progressively measurable, thus $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\nu}'_t)\}$ is a solution of the corlol martingale problem for $\mathbb{H}(\mathcal{L}^\Delta, \mathcal{B}^\Delta, h^\Delta)$, and $\tilde{\nu}'_0(\mathbb{R}^d) = 1$ a.s. Next we need the following result, which is a variant of Theorem 4.3.8 of [8], and which is established later in this section:

Lemma 8.47. *Suppose that Condition* AII *holds, and let* $\{(\hat{\Omega}, \hat{\mathcal{F}}, \{\hat{\mathcal{F}}_t\}, \hat{P}), (\hat{\nu}_t)\}$ *be a solution of the corlol martingale problem for* $\mathbb{H}(\mathcal{L}^\Delta, \mathcal{B}^\Delta, h^\Delta)$. *If* $\hat{P}(\hat{\nu}_0(\mathbb{R}^d) = 1) = 1$, *then* $\hat{P}(\hat{\nu}_t(\mathbb{R}^d) = 1,\ \forall t \geq 0) = 1$.

Lemma 8.47 shows that $\tilde{P}(\tilde{\nu}'_t(\mathbb{R}^d) = 1,\ \forall t \in [0, T]) = 1$, thus $\{\tilde{\nu}'_t\}$ is a $\mathcal{P}(\mathbb{R}^d)$-valued, $\{\tilde{\mathcal{F}}_t\}$-adapted, corlol modification of $\{\tilde{\nu}_t\}$. We conclude that $\{(\tilde{\nu}'_t, \tilde{V}'_t)\}$ is a $(\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^r)$-valued, $\{\tilde{\mathcal{F}}_t\}$-adapted, corlol modification of $\{(\tilde{\nu}_t, \tilde{V}_t)\}$.

Step 3. Fix $\varphi \in C_c^\infty(\mathbb{R}^d)$ and put

$$\tilde{\eta}_t := \tilde{\nu}'_t\varphi - \int_0^t \tilde{\nu}'_s(\mathcal{L}\varphi)\, ds. \tag{8.101}$$

Since $\{\tilde{\nu}'_t\}$ is $\mathcal{P}(\mathbb{R}^d)$-valued, corlol, and $\{\tilde{\mathcal{F}}_t\}$-adapted, it follows that $\{\tilde{\eta}_t\}$ is $\mathbb{R}$-valued, corlol, and $\{\tilde{\mathcal{F}}_t\}$-adapted. Fix $H_2 \in C_c^\infty(\mathbb{R})$ such that $H_2(x) = x$, $|x| \leq \|\varphi\|$, and define $\Phi_2 \in \mathcal{D}(\mathbb{H}(\mathcal{L}, \mathcal{B}, h))$ by

$$\Phi_2(\nu) := H_2(\nu\varphi), \qquad \forall \nu \in \mathcal{P}(\mathbb{R}^d). \tag{8.102}$$

Then (see (5.42))

$$\Phi_2(\nu) = \nu\varphi \quad \text{and} \quad \mathbb{H}(\mathcal{L}, \mathcal{B}, h)(\Phi_2)(\nu) = \nu(\mathcal{L}\varphi), \qquad \forall \nu \in \mathcal{P}(\mathbb{R}^d). \tag{8.103}$$

Now $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\nu}'_t, \tilde{V}'_t)\}$ solves the corlol martingale problem for $\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)$, and therefore $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\nu}'_t)\}$ solves the corlol martingale problem for $\mathbb{H}(\mathcal{L}, \mathcal{B}, h)$, and hence from (8.101) and (8.103), it follows that $\{\tilde{\eta}_t\}$ is a corlol $\{\tilde{\mathcal{F}}_t\}$-martingale. However, $\{(\tilde{V}'_t)^k\}$ is a *continuous* $\{\tilde{\mathcal{F}}_t\}$-martingale (in fact, Wiener process) for each $k = 1, 2, \ldots, r$, so that Theorem VI.37.8 of [19] gives a *continuous* co-variation process $\{[\tilde{\eta}, (\tilde{V}')^k]_t\}$ which is unique to within indistinguishability. We will show that

$$[\tilde{\eta}, (\tilde{V}')^k]_t = \int_0^t R_k(s)\, ds, \qquad \text{where} \quad R_k(s) := R_{\mathcal{B}_k}(\varphi, h^k, \tilde{\nu}'_s) \tag{8.104}$$

(recall (5.31), (5.39), and (5.40)). For each $n = 1, 2, \ldots$ fix some $g_n \in C_c^\infty(\mathbb{R}^r)$ with $g_n(y) = y^k$, $\forall y \in \mathbb{R}^r$, $|y| \leq n$, and put $\Psi_n := \Phi_2 \otimes g_n$, $n = 3, 4, \ldots$, where Φ_2 is given by (8.102). Then $\Psi_n \in \mathcal{D}(\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h))$, and it is easily seen that

$$\Psi_n(\nu, y) = (y^k)(\nu\varphi) \quad \text{and} \quad \hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)(\Psi_n)(\nu, y) = (y^k)\nu(\mathcal{L}\varphi) + R_{\mathcal{B}_k}(\varphi, h^k, \nu),$$
$$\forall |y| \le n, \quad \nu \in \mathcal{P}(\mathbb{R}^d), \quad y \in \mathbb{R}^r. \tag{8.105}$$

Put $T_n := \inf\{t : |(\tilde{V}_t')^k| \ge n\}$, $n = 1, 2, \ldots$. Then T_n is an $\{\tilde{\mathcal{F}}_t\}$-stopping time, and since $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\nu}_t', \tilde{V}_t')\}$ solves the corlol martingale problem for $\hat{\mathbb{H}}(\mathcal{L}, \mathcal{B}, h)$, we see from the optional sampling theorem and (8.105) that

$$M_n(t) := (\tilde{V}'_{t\wedge T_n})^k (\tilde{\nu}'_{t\wedge T_n}\varphi) - \int_0^{t\wedge T_n} [(\tilde{V}_s')^k\, \tilde{\nu}_s'(\mathcal{L}\varphi) + R_k(s)]\, ds \tag{8.106}$$

is a corlol $\{\tilde{\mathcal{F}}_t\}$-martingale. From (8.106) and (8.101) we have

$$\tilde{\eta}_{t\wedge T_n}(\tilde{V}'_{t\wedge T_n})^k - \int_0^{t\wedge T_n} R_k(s)\, ds$$
$$= M_n(t) + \left[\int_0^{t\wedge T_n} (\tilde{V}_s')^k\, \tilde{\nu}_s'(\mathcal{L}\varphi)\, ds - \left(\int_0^{t\wedge T_n} \tilde{\nu}_s'(\mathcal{L}\varphi)\, ds\right)(\tilde{V}'_{t\wedge T_n})^k\right], \tag{8.107}$$

and since $\{(\tilde{V}_t')^k\}$ is a continuous $\{\tilde{\mathcal{F}}_t\}$-martingale it follows from Problem 2.9.22 in [8] that

$$\int_0^t (\tilde{V}_s')^k\, \tilde{\nu}_s'(\mathcal{L}\varphi)\, ds - \left(\int_0^t \tilde{\nu}_s'(\mathcal{L}\varphi)\, ds\right)(\tilde{V}_t')^k$$

is a continuous $\{\tilde{\mathcal{F}}_t\}$-martingale. Thus, the process on the left-hand side of (8.107) is a corlol $\{\tilde{\mathcal{F}}_t\}$-martingale for all positive integers n, whence $\tilde{\eta}_t(\tilde{V}_t')^k - \int_0^t R_k(s)\, ds$ is a corlol $\{\tilde{\mathcal{F}}_t\}$-local martingale. Now (8.104) follows by the uniqueness of $[\eta, \tilde{V}']_t$.

Step 4. Put

$$\tilde{\rho}_t := \tilde{\eta}_t - \sum_{k=1}^r \int_0^t R_{\mathcal{B}_k}(\varphi, h^k, \tilde{\nu}_s')\, d(\tilde{V}_s' - \tilde{V}_0')^k. \tag{8.108}$$

We shall establish that $\tilde{P}(\tilde{\rho}_t = \tilde{\rho}_0,\ \forall t \ge 0) = 1$, which shows that $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\nu}_t', \tilde{V}_t' - \tilde{V}_0')\}$ is a solution of the normalized filter equation corresponding to $(\mathbb{R}^d; \mathcal{L}, \mathcal{B}, h)$. Since $\mathcal{D}(\mathbb{H}(\mathcal{L}, \mathcal{B}, h))$ is a subalgebra of $\bar{C}(\mathcal{P}(\mathbb{R}^d))$ (by Lemma 5.33) we have $\Phi_2^2 \in \mathcal{D}(\mathbb{H}(\mathcal{L}, \mathcal{B}, h))$ (for Φ_2 given by (8.102)), and it is easily seen that

$$\Phi_2^2(\nu) = (\nu\varphi)^2 \quad \text{and} \quad \mathbb{H}(\mathcal{L}, \mathcal{B}, h)(\Phi_2^2)(\nu) = 2(\nu\varphi)\nu(\mathcal{L}\varphi) + \sum_{k=1}^r R^2_{\mathcal{B}_k}(\varphi, h^k, \nu),$$
$$\forall \nu \in \mathcal{P}(\mathbb{R}^d). \tag{8.109}$$

Now in Step 3 we have seen that $\{(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P}), (\tilde{\nu}_t')\}$ solves the corlol martingale problem for $\mathbb{H}(\mathcal{L}, \mathcal{B}, h)$, and hence from (8.109) and (8.104) we know that

$$(\tilde{\nu}_t'\varphi)^2 - \int_0^t \left[2(\tilde{\nu}_s'\varphi)\tilde{\nu}_s'(\mathcal{L}\varphi) + \sum_{k=1}^r R_k^2(s)\right] ds \tag{8.110}$$

is an $\{\tilde{\mathcal{F}}_t\}$-martingale. From (8.101), (8.110), and Problem 2.9.29 of [8], we then get

$$\tilde{\eta}_t^2 - \int_0^t \left[2(\tilde{\nu}_s'\varphi)\tilde{\nu}_s'(\mathcal{L}\varphi) + \sum_{k=1}^r R_k^2(s) - 2(\tilde{\nu}_s'\varphi)\tilde{\nu}_s'(\mathcal{L}\varphi) \right] ds$$
$$\equiv \tilde{\eta}_t^2 - \int_0^t \sum_{k=1}^r R_k^2(s)\, ds \tag{8.111}$$

is an $\{\tilde{\mathcal{F}}_t\}$-martingale. Squaring (8.108) and taking expectations yields

$$\mathrm{E}[\tilde{\rho}_t^2] = \mathrm{E}[\tilde{\eta}_t^2] - 2\sum_{k=1}^r \mathrm{E}\left[\tilde{\eta}_t \int_0^t R_k(u)\, d(\tilde{V}_u' - \tilde{V}_0')^k \right]$$
$$+ \sum_{k=1}^r \mathrm{E}\left[\int_0^t R_k^2(u)\, du \right]. \tag{8.112}$$

Since the quantity in (8.111) is an $\{\tilde{\mathcal{F}}_t\}$-martingale, we have $\mathrm{E}[\tilde{\eta}_t^2 - \int_0^t \sum_{k=1}^r R_k^2(u)\, du] = \mathrm{E}[\tilde{\eta}_0^2] \equiv \mathrm{E}[\tilde{\rho}_0^2]$, and combining this with (8.112) gives

$$\mathrm{E}[\tilde{\rho}_t^2] = \mathrm{E}[\tilde{\rho}_0^2] - 2\sum_{k=1}^r \mathrm{E}\left[\tilde{\eta}_t \int_0^t R_k(u)\, d(\tilde{V}_u' - \tilde{V}_0')^k \right]$$
$$+ 2\sum_{k=1}^r \mathrm{E}\left[\int_0^t R_k^2(u)\, du \right]. \tag{8.113}$$

Since $\{\tilde{\nu}_t'\}$ is corlol and $h^k \in \bar{C}(\mathbb{R}^d)$, it follows that $\{R_k(t)\}$ defined in (8.104) is corlol, so that $\{R_k(t-)\}$ is left-continuous and $\{\tilde{\mathcal{F}}_t\}$-adapted, thus it is $\{\tilde{\mathcal{F}}_t\}$-predictable. Then, since $\{\tilde{\eta}_t\}$ has been seen to be a corlol $\{\tilde{\mathcal{F}}_t\}$-martingale, it follows from (8.104) and Theorem VI.37.9(vi) of [19] that

$$\mathrm{E}\left[\tilde{\eta}_t \int_0^t R_k(s-)\, d(\tilde{V}_s' - \tilde{V}_0')^k \right] = \mathrm{E}\left[\int_0^t R_k(s-)\, R_k(s)\, ds \right]. \tag{8.114}$$

Again since $\{R_k(t)\}$ is corlol, the set $\{t \in [0, T] : R_k(t) \neq R_k(t-)\}$ is countable, hence has Lebesgue measure zero, for each $\tilde{\omega}$, so that (8.114) gives

$$\mathrm{E}\left[\tilde{\eta}_t \int_0^t R_k(s)\, d(\tilde{V}_s' - \tilde{V}_0')^k \right] = \mathrm{E}\left[\int_0^t R_k^2(s)\, ds \right]. \tag{8.115}$$

Combining (8.115) with (8.113) gives $\mathrm{E}[\tilde{\rho}_t^2 - \tilde{\rho}_0^2] = 0$, and, since $\{\tilde{\rho}_t\}$ is an $\{\tilde{\mathcal{F}}_t\}$-martingale, this gives $\mathrm{E}[(\tilde{\rho}_t - \tilde{\rho}_0)^2] = 0$, and thus $\tilde{\rho}_t = \tilde{\rho}_0$ a.s. for each $t \geq 0$. However, the $\mathbb{R}$-valued process $\{\tilde{\rho}_t\}$ is corlol, hence we find $\tilde{P}(\tilde{\rho}_t = \tilde{\rho}_0,\ \forall t \geq 0) = 1$, as required to show that (8.98) holds. □

Proof of Lemma 8.47. Fix some $\varphi \in C_c^\infty(\mathbb{R}^d)$ such that $I[0, 1](|x|) \leq \varphi(x) \leq I[0, 2](|x|)$, $x \in \mathbb{R}^d$, and define $\varphi_n \in \bar{C}(\mathbb{R}^{d*})$ by $\varphi_n(x) := \varphi(x/n)$, $\forall x \in \mathbb{R}^d$, $\varphi_n(\Delta) := 0$, $\forall n = 1, 2, \ldots$. Fix $H \in C_c^\infty(\mathbb{R})$ such that $H(x) = x$, for all $|x| \leq 1$, and put

$\Phi_n(\nu) := H(\nu\varphi_n)$, $\forall \nu \in \mathcal{P}(\mathbb{R}^{d*})$. Since $\varphi_n \in \mathcal{D}(\mathcal{L}^\Delta, \mathcal{B}^\Delta)$ (recall (8.99a)), we see that $\Phi_n \in \mathcal{D}(\mathbb{H}(\mathcal{L}^\Delta, \mathcal{B}^\Delta, h^\Delta))$, with $\Phi_n(\nu) = \nu\varphi_n$ and $\mathbb{H}(\mathcal{L}^\Delta, \mathcal{B}^\Delta, h^\Delta)(\Phi_n)(\nu) = \nu(\mathcal{L}^\Delta\varphi_n)$, $\forall \nu \in \mathcal{P}(\mathbb{R}^{d*})$. However, $\{(\hat{\Omega}, \hat{\mathcal{F}}, \{\hat{\mathcal{F}}_t\}, \hat{P}), (\hat{\nu}_t)\}$ solves the corlol martingale problem for $\mathbb{H}(\mathcal{L}^\Delta, \mathcal{B}^\Delta, h^\Delta)$, thus

$$N_t^n := \hat{\nu}_t\varphi_n - \int_0^t \hat{\nu}_s(\mathcal{L}^\Delta\varphi_n)\, ds, \qquad t \geq 0,$$

is a corlol $\{\hat{\mathcal{F}}_t\}$-martingale, and therefore $\{N_t^n\}$ is an $\{\hat{\mathcal{F}}_{t+}\}$-martingale. Now put

$$M_k := \left\{\mu \in \mathcal{P}(\mathbb{R}^{d*}) :\ \mu(\{\Delta\}) < \frac{1}{k}\right\}, \quad k = 1, 2, \ldots, \qquad \text{and}$$

$$M := \bigcap_{k=1}^{\infty} M_k, \tag{8.116}$$

$$T_m^k := \inf\left\{t \geq 0 :\ \inf_{y \in \mathcal{P}(\mathbb{R}^{d*})\setminus M_k} d(y, \hat{\nu}_t) < \frac{1}{m}\right\}, \qquad \forall\, k, m = 1, 2, \ldots \tag{8.117}$$

(here $d(\cdot, \cdot)$ denotes the Prohorov metric on $\mathcal{P}(\mathbb{R}^{d*})$). Since the $\mathcal{P}(\mathbb{R}^{d*})$-valued process $\{\hat{\nu}_t\}$ is corlol and $\{\hat{\mathcal{F}}_t\}$-adapted, the T_m^k are $\{\hat{\mathcal{F}}_{t+}\}$-stopping times, with $T_1^k \leq T_2^k \leq \cdots$, for each $k = 1, 2, \ldots$. Fix some $t > 0$, put $T^k := \lim_{m\to\infty} T_m^k$, and observe that the limit

$$\hat{\mu}_t^k \equiv \lim_{m\to\infty} \hat{\nu}_{t\wedge T_m^k} \tag{8.118}$$

exists in $\mathcal{P}(\mathbb{R}^{d*})$ for each $k = 1, 2, \ldots$, since $\{\hat{\nu}_t\}$ has left limits. Now $M_k \subset \mathcal{P}(\mathbb{R}^{d*})$ is open, from which it easily follows that

$$\{\hat{\mu}_t^k(\omega) \in M_k\} \subset \{T^k(\omega) > t\}, \tag{8.119}$$

for all $k = 1, 2, \ldots, \omega \in \hat{\Omega}$. We have seen that $\{N_t^n\}$ is a corlol $\{\hat{\mathcal{F}}_{t+}\}$-martingale, thus the Optional Sampling Theorem gives

$$\mathrm{E}^{\hat{P}}[\hat{\nu}_{t\wedge T_m^k}\varphi_n] = \mathrm{E}^{\hat{P}}[\hat{\nu}_0\varphi_n] + \mathrm{E}^{\hat{P}}\left[\int_0^{t\wedge T_m^k} \hat{\nu}_s(\mathcal{L}^\Delta\varphi_n)\, ds\right],$$
$$\forall k,\ m,\ n = 1, 2, \ldots, \tag{8.120}$$

and, from dominated convergence (with $m \to \infty$, n fixed) and (8.118), we then get

$$\mathrm{E}^{\hat{P}}[\hat{\mu}_t^k\varphi_n] = \mathrm{E}^{\hat{P}}[\hat{\nu}_0\varphi_n] + \mathrm{E}^{\hat{P}}\left[\int_0^{t\wedge T^k} \hat{\nu}_s(\mathcal{L}^\Delta\varphi_n)\, ds\right],$$
$$\forall k,\ n = 1, 2, \ldots. \tag{8.121}$$

Now b.p.-$\lim_{n\to\infty}(\varphi_n, \mathcal{L}^\Delta\varphi_n) = (I_{\mathbb{R}^d}, 0)$ (from the upper-bounds in Remark 2.9 and Problem 4.11.12 of [8]), so that $n \to \infty$ in (8.121) gives $\mathrm{E}^{\hat{P}}[\hat{\mu}_t^k(\mathbb{R}^d)] = \mathrm{E}^{\hat{P}}[\hat{\nu}_0(\mathbb{R}^d)] = 1$, that is $\hat{P}[\hat{\mu}_t^k \in M_k] = 1$, and therefore (see (8.119)) we have $\hat{P}[T^k > t] = 1$, which, by the definition of T^k, implies $\hat{P}[\hat{\nu}_s \in M_k,\ \forall s \in [0,t]] = 1$, $\forall k = 1, 2, \ldots$. Thus $\hat{P}[\hat{\nu}_s \in M,\ \forall s \in [0,t]] = 1$ (see (8.116)), and since $M = \{\mu \in \mathcal{P}(\mathbb{R}^{d*}) : \mu(\{\Delta\}) = 0\}$ the result follows from the arbitrary choice of t. □

Proof of Lemma 5.40 (motivated by Goggin [11, p. 1100]). Fix $t \geq 0$. It is enough to show that, for each $\eta \in (0,1)$, there is a sequence of compacta $K_k^\eta \subset \mathbb{R}^d$, $k = 1, 2, \ldots$, such that

$$\inf_n P\left[\pi_t^n(K_k^\eta) > 1 - \frac{1}{k}\right] \geq 1 - \frac{\eta}{2^k}, \qquad \forall k = 1, 2, \ldots, \tag{8.122}$$

for then it follows that

$$\inf_n\ P\left(\bigcap_{k=1}^{\infty}\left[\pi_t^n(K_k^\eta) > 1 - \frac{1}{k}\right]\right) \geq 1 - \eta. \tag{8.123}$$

Now put $\mathcal{K}^\eta := \bigcap_{k=1}^\infty \{\pi \in \mathcal{P}(\mathbb{R}^d) : \pi(K_k^\eta) > 1-(1/k)\}$. Then it is clear that $\mathcal{K}^\eta$ is a tight (hence relatively compact) subset of $\mathcal{P}(\mathbb{R}^d)$, and $P[\pi_t^n \in \mathcal{K}^\eta] = P[\bigcap_{k=1}^\infty \{\pi_t^n(K_k^\eta) > 1 - (1/k)\}]$, so that (8.123) gives

$$\inf_n\ P\left[\pi_t^n \in \bar{\mathcal{K}^\eta}\right] \geq 1 - \eta. \tag{8.124}$$

Since $\bar{\mathcal{K}^\eta} \subset \mathcal{P}(\mathbb{R}^d)$ is compact for each $\eta \in (0,1)$, it follows from (8.124) and the Prohorov theorem (Theorem 3.3.2 of [8]) that the sequence $\{\mathcal{L}(\pi_t^n),\ n = 1, 2, \ldots\}$ is relatively compact in $\mathcal{P}(\mathcal{P}(\mathbb{R}^d))$, as required. Thus, fix some $\eta \in (0,1)$. It remains to construct a sequence of compacta $K_k^\eta \subset \mathbb{R}^d$, $k = 1, 2, \ldots$, such that (8.122) holds. By Proposition 4.23 and the Prohorov theorem, for each $k = 1, 2, \ldots$ there is a compact $K_k^\eta \subset \mathbb{R}^d$ such that

$$1 - \frac{\eta}{k2^k} < \inf_n P[X_t^n \in K_k^\eta] = \inf_n \mathrm{E}[\pi_t^n(K_k^\eta)], \qquad k = 1, 2, \ldots, \tag{8.125}$$

where the last equality follows since $\pi_t^n(K_k^\eta) = P\{X_t^n \in K_k^\eta | \mathcal{F}_t^n\}$. Thus, for all $n, k = 1, 2, \ldots$, we get

$$\begin{aligned}
1 - \frac{\eta}{k2^k} &< \mathrm{E}\left[\pi_t^n(K_k^\eta) :\ \pi_t^n(K_k^\eta) > 1 - \frac{1}{k}\right] + \mathrm{E}\left[\pi_t^n(K_k^\eta) :\ \pi_t^n(K_k^\eta) \leq 1 - \frac{1}{k}\right] \\
&\leq P\left[\pi_t^n(K_k^\eta) > 1 - \frac{1}{k}\right] + \left(1 - \frac{1}{k}\right) P\left[\pi_t^n(K_k^\eta) \leq 1 - \frac{1}{k}\right] \\
&\leq 1 - \frac{1}{k} + \frac{1}{k} P\left[\pi_t^n(K_k^\eta) > 1 - \frac{1}{k}\right],
\end{aligned}$$

from which (8.122) follows. □

9. Appendix of Miscellaneous Proofs and Results

Proof of Lemma 2.6. The existence of the Borel measures $\chi(z, \cdot)$ such that (2.13) and (i) hold is suggested by Remark 12.2.3 of [8]. For completeness we briefly summarize the proof:

Put $U_n\Psi := \int_0^n [T_t\Psi - \bar{m}\Psi]\,dt$ and $U\Psi := \int_0^\infty [T_t\Psi - \bar{m}\Psi]\,dt$, for all $\Psi \in C(S)$, $n = 1, 2, \ldots$. Then U_n is a bounded linear operator on $C(S)$ and Remark 2.7 shows that U is a linear operator on $C(S)$ with $\lim_n \|U_n\Psi - U\Psi\| = 0$ for each $\Psi \in C(S)$. Then U is a bounded operator (by the Uniform Boundedness Principle), thus $\Psi \to (U\Psi)(z)$ is a bounded linear functional on $C(S)$ for each $z \in S$. Now the Riesz Representation Theorem gives a unique signed regular Borel measure $\chi(z,\cdot)$ on $\mathcal{B}(S)$ such that (2.13) holds and $\sup_z \|\chi(z,\cdot)\|_{\mathrm{TV}} \le |||U|||$ (operator norm of U).

(i) Fix $\Psi \in C(S)$ with $\bar{m}\Psi = 0$. Then, from (2.13), we have $\Phi = \int_0^\infty T_t\Psi\,dt$. If $\{R_\lambda,\ \lambda \in (0,\infty)\}$ is the resolvent of the semigroup $\{T_t\}$, then, from the Dominated Convergence Theorem and $\int_0^\infty \|T_t\Psi\|\,dt < \infty$ (recall Condition 2.3), we have $\lim_{\lambda\to 0+}\|R_\lambda\Psi - \Phi\| = 0$, so that we get $\lim_{\lambda\to 0+}(R_\lambda\Psi, \lambda R_\lambda\Psi - \Psi) = (\Phi, -\Psi)$ in $C(S)\times C(S)$. However, $R_\lambda\Psi \in \mathcal{D}(\mathcal{Q})$ with $\mathcal{Q}(R_\lambda\Psi) = \lambda R_\lambda\Psi - \Psi$ (see p. 11 of [8]) and $\mathcal{Q}$ is a closed operator (Corollary 1.1.6 of [8]), thus $(\Phi, -\Psi) \in \mathcal{Q}$, as required.

(ii) By making a Hahn decomposition of the signed measure $\chi(z,\cdot)$ and applying standard results on interchanging derivatives and integrals (see, e.g. Theorem 2.27 of [9]) to the integrals with respect to the positive and negative parts of $\chi(z,\cdot)$, we find that $f(\cdot,z) \in C^1(\mathbb{R}^d), \forall z \in S$, with $(\partial_j f)(x,z) = \int_S (\partial_j g)(x,z')\chi(x,dz'), \forall (x,z) \in \mathbb{R}^d \times S$. Thus, to establish the result, it must be shown that

$$v(x,z) := \int_S h(x,z')\chi(z,dz'), \qquad \forall (x,z) \in \mathbb{R}^d \times S, \tag{9.126}$$

is a member of $C(\mathbb{R}^d \times S)$ when $h \in C(\mathbb{R}^d \times S)$. Fix such an h, put $C := \sup_z \|\chi(z,\cdot)\|_{\mathrm{TV}} < \infty$, and fix some $(x_0, z_0) \in \mathbb{R}^d \times S$ and (small) $\eta > 0$. Let K denote the closed ball in $\mathbb{R}^d$ of unit radius centered at x_0. Since $K \times S$ is compact, the Stone–Weierstrass theorem gives some positive integer n and $\varphi_i \in C(K)$ and $\Psi_i \in C(S)$, $i = 1, 2, \ldots, n$, such that

$$\left| h(x,z') - \sum_{i=1}^n \varphi_i(x)\Psi_i(z') \right| < \frac{\eta}{3C}, \qquad \forall (x,z') \in K \times S \tag{9.127}$$

(e.g. see Exercise 4.68 of [9]). Now put $\Phi_i(z) := \int_S \Psi_i(z')\chi(z,dz')$, $\forall z \in S$, and observe, from (2.13), that $\Phi_i \in C(S)$. Thus there is some $\delta > 0$ such that

$$\left| \sum_{i=1}^n [\varphi_i(x)\Phi_i(z) - \varphi_i(x_0)\Phi_i(z_0)] \right| < \frac{\eta}{3}, \tag{9.128}$$

for all $(x,z) \in \mathbb{R}^d \times S$ with $|x_0 - x| + r(z, z_0) < \delta$ (where r denotes the metric on S). However, from (9.126),

$$\begin{aligned} v(x,z) &- v(x_0,z_0) \\ &= \int_S \left[h(x,z') - \sum_i \varphi_i(x)\Psi_i(z') \right] \chi(z,dz') \\ &\quad + \sum_i [\varphi_i(x)\Phi_i(z) - \varphi_i(x_0)\Phi_i(z_0)] \\ &\quad + \int_S \left[\sum_i \varphi_i(x_0)\Psi_i(z') - h(x_0,z') \right] \chi(z_0,dz'), \end{aligned} \tag{9.129}$$

for all $(x, z) \in K \times S$. From (9.127), (9.128), and (9.129) we see that if $|x - x_0| + r(z, z_0) < \delta$, then $(x, z) \in K \times S$ and $|v(x, z) - v(x_0, z_0)| < \eta$. $\square$

The following result is an immediate consequence of the separability of $\hat{C}(\mathbb{R}^D)$ in the supremum norm (see also Remark 2.5 of [16]):

Lemma 9.48. *For every* $n = 1, 2, \ldots$, *there exists a countable set* $\mathcal{H}_n \subset C_c^\infty(\mathbb{R}^n)$ *with the following property*: *for each* $\varphi \in C_c^\infty(\mathbb{R}^n)$ *there is some sequence* $\{\varphi_k\} \subset \mathcal{H}_n$ *such that*, *when* $g\colon \mathbb{R}^n \to \mathbb{R}$ *is bounded on bounded sets*, *we have* $\lim_{k\to\infty} \|g\varphi_k - g\varphi\| = 0$, $\lim_{k\to\infty} \|g\, \partial_i \varphi_k - g\, \partial_i \varphi\| = 0$, *and* $\lim_{k\to\infty} \|g\, \partial_i \partial_j \varphi_k - g\, \partial_i \partial_j \varphi\| = 0, \forall i, j = 1, 2, \ldots, n$.

References

1. Bhatt A.G., Kallianpur G., Karandikar R.L. (1999) Robustness of the nonlinear filter. Stochastic Process Appl., 81: 247–254.
2. Bhatt A.G., Karandikar R.L. (1993) Invariant measures and evolution equations for Markov processes characterized via martingale problems. Ann. Probab., 21: 2246–2268.
3. Bhatt A.G., Karandikar R.L. (1993) Weak convergence to a Markov process: the martingale approach. Probab. Theory Rel. Fields, 96: 335–351.
4. Bhatt A.G., Karandikar R.L. (2002) Robustness of the nonlinear filter: the correlated case. Stochastic Process Appl., 97: 41–58.
5. Billingsley P. (1968) Convergence of Probability Measures. Wiley, New York.
6. Blankenship G., Papanicolaou G.C. (1968) Stability and control of stochastic systems with wide-band noise disturbances I. SIAM J. Appl. Math., 34: 437–476.
7. Dawson D.A. (1993) Measure-Valued Markov Processes. Lecture Notes in Mathematics, no. 1541, pages 1–260. Springer-Verlag, Berlin.
8. Ethier S.N., Kurtz T.G. (1986) Markov Processes: Characterization and Convergence. Wiley, New York.
9. Folland G. (1984) Real Analysis: Modern Techniques and Their Applications. Wiley, New York.
10. Fujisaki M., Kallianpur G., Kunita H. (1972) Stochastic differential equations for the non-linear filtering problem. Osaka J. Math., 9: 19–40.
11. Goggin E.M. (1994) Convergence in distribution of conditional expectations. Ann. Probab., 22: 1097–1114.
12. Hijab O. (1989) Partially observed control of Markov processes I. Stochastics, 28: 123–144.
13. Kallianpur G. (1980) Stochastic Filtering Theory. Springer-Verlag, Berlin.
14. Karandikar R.L. (1989) On the Metivier–Pellaumail inequality, Emery topology and pathwise formulae in stochstic calculus. Sankhyā Ser. A, 51: 121–143.
15. Karatzas I., Shreve S.E. (1991) Brownian Motion and Stochastic Calculus, 2nd edn. Springer-Verlag, Berlin.
16. Kurtz T.G. (1998) Martingale problems for conditional distributions of Markov processes. Electron J. Probab., 3: 1–29.
17. Kurtz T.G., Ocone D.L. (1988) Unique characterization of conditional distributions in nonlinear filtering. Ann. Probab., 16: 80–107.
18. Lucic V.M., Heunis A.J. (2001) On uniqueness of solutions for the stochastic differential equations of nonlinear filtering. Ann. Appl. Probab., 11: 182–209.
19. Rogers L.C.G., Williams D. (1987) Diffusions, Markov Processes and Martingales, Vol. 2: Itô Calculus. Wiley, New York.
20. Stroock D.W., Varadhan S.R.S. (1979) Multidimensional Diffusion Processes. Springer-Verlag, Berlin.
21. Wong E., Hajek B.E. (1985) Stochastic Processes in Engineering Systems. Springer-Verlag, Berlin.

Accepted 27 *March* 2003. *Online publication* 7 *August* 2003.