The filtered martingale problem

Thomas G. Kurtz [*]
Departments of Mathematics and Statistics
University of Wisconsin - Madison
480 Lincoln Drive
Madison, WI 53706-1388
USA
kurtz@math.wisc.edu
http://www.math.wisc.edu/∼kurtz/

Giovanna Nappo [†]
Dipartimento di Matematica "Guido Castelnuovo"
Università degli Studi LA SAPIENZA
piazzale A. Moro 2
00185 Roma
ITALY
nappo@mat.uniroma1.it
http://www.mat.uniroma1.it/people/nappo/nappo.html

November 18, 2009

Abstract

Let X be a Markov process characterized as the solution of a martingale problem with generator A, and let Y be a related *observation process*. The conditional distribution π_t of $X(t)$ given observations of Y up to time t satisfies certain martingale properties, and it is shown that any probability-measure-valued process with the appropriate martingale properties can be interpreted as the conditional distribution of X for some observation process. In particular, if $Y(t) = \gamma(X(t))$ for some measurable mapping γ, the conditional distribution of $X(t)$ given observations of Y up to time t is characterized as the solution of a *filtered martingale problem*. Uniqueness for the original martingale problem implies uniqueness for the filtered martingale problem which in turn implies the Markov property for the conditional distribution considered as a probability-measure-valued process. Other applications include a Markov mapping theorem and uniqueness for filtering equations.

MSC 2000 subject classifications: 60J25, 93E11, 60G35, 60J35, 60G44
Keywords: partial observation, conditional distribution, martingale problems, filtering, filtering equation, Markov process, Markov mapping

[*]Research supported in part by NSF grants DMS 05-03983 and 08-05793

[†]Research supported in part by Progetto Università - 2008 - prot.C26A08H8YM, Interazione e dipendenza nei modelli stocastici

1 Introduction

The notion of a *filtered martingale problem* was introduced in Kurtz and Ocone (1988) and extended to a more general setting in Kurtz (1998). The basic idea is that the conditional distribution of the state of a Markov process given the information from related observations satisfies a kind of *martingale problem*. The fundamental results give conditions under which every solution of the filtered martingale problem arises from a solution of the original martingale problem, and hence, uniqueness for the original martingale problem implies uniqueness for the filtered martingale problem. These results have a variety of consequences, most notably, the uniqueness for filtering equations and general results on Markov mappings, that is, conditions under which a transformation of a Markov process is still Markov.

The current paper is concerned with extension of these ideas to further settings. The filtering literature contains a number of results (for example, Bhatt, Budhiraja, and Karandikar (2000); Budhiraja (2003); Kunita (1971)) showing that the conditional distribution for the classical filtering problem is itself a Markov process. In order to address this question for general solutions of filtered martingale problems, we need to generalize the earlier definition to include information available at time zero. We are then able to show the Markov property for the conditional distributions for a large class of partially observed Markov processes.

We also extend the earlier results to local martingale problems which in turn allows us to generalize previous uniqueness results for filtering equations. The basic results can also be extended to constrained martingale problems, that is, martingale problems for processes in which the behavior of the process on the boundary of the state space is determined by a second operator (see Anderson (1976); Kurtz (1990, 1991); Kurtz and Stockbridge (2001); Stroock and Varadhan (1971)). Reflecting diffusion processes provide one example.

We also relax some technical conditions present in the earlier work.

Throughout this paper, all filtrations are assumed complete, all processes are assumed to be progressively measurable, $\{\mathcal{F}_t^Y\}$ denotes the completion of the filtration generated by the observed process Y, and assuming Y takes values in $\mathbb{S}_0$, $\widehat{\mathcal{F}}_t^Y$ denotes the completion of $\sigma(\int_0^r h(Y(s))ds : r \le t, h \in B(\mathbb{S}_0)) \vee \sigma(Y(0))$.

2 Martingale properties of conditional distributions

Let $\mathbb{S}$ be a complete, separable metric space. $C(\mathbb{S})$ will denote the space of $\mathbb{R}$-valued, continuous functions on $\mathbb{S}$, $M(\mathbb{S})$ the Borel measurable functions, $C_b(\mathbb{S})$ the bounded continuous functions, $B(\mathbb{S})$ the bounded measurable functions, and $\mathcal{P}(\mathbb{S})$ the space of probability measures on $\mathbb{S}$. $M_{\mathbb{S}}[0,\infty)$ will denote the space of measurable functions $x : [0,\infty) \to \mathbb{S}$ topologized by convergence in Lebesgue measure, $D_{\mathbb{S}}[0,\infty) \subset M_{\mathbb{S}}[0,\infty)$ the space of $\mathbb{S}$-valued, cadlag functions with the Skorohod topology, and $C_{\mathbb{S}}[0,\infty) \subset D_{\mathbb{S}}[0,\infty)$ the subspace of continuous functions. We consider martingale problems for operators satisfying the following condition:

Condition 2.1 *i) $A : \mathcal{D}(A) \subset C_b(\mathbb{S}) \to M(\mathbb{S})$ with $1 \in \mathcal{D}(A)$ and $A1 = 0$.*

ii) Either $\mathcal{R}(A) \subset C(\mathbb{S})$ or there exists a complete separable metric space $\mathbb{U}$, a transition *function* η *from* $\mathbb{S}$ *to* $\mathbb{U}$*, and an operator* $A_1 : \mathcal{D}(A) \subset C_b(\mathbb{S}) \to C(\mathbb{S} \times \mathbb{U})$ *such that*

$$Af(x) = \int_{\mathbb{U}} A_1 f(x,z)\eta(x,dz), \quad f \in \mathcal{D}(A). \tag{2.1}$$

iii) There exist $\psi \in C(\mathbb{S})$*,* $\psi \geq 1$*, and constants* a_f *such that* $f \in \mathcal{D}(A)$ *implies*

$$|Af(x)| \leq a_f \psi(x),$$

or if A *is of the form (2.1), there exist* $\psi_1 \in C(\mathbb{S} \times \mathbb{U})$*,* $\psi_1 \geq 1$*, and constants* a_f *such that, for all* $(x,z) \in \mathbb{S} \times \mathbb{U}$

$$|A_1 f(x,z)| \leq a_f \psi_1(x,z).$$

(If A *is of the form (2.1), then define* $\psi(x) \equiv \int_{\mathbb{U}} \psi_1(x,z)\eta(x,dz)$*.)*

iv) Defining $A_0 = \{(f, \psi^{-1}Af) : f \in \mathcal{D}(A)\}$ *(or* $\{(f, \psi_1^{-1}A_1 f), f \in \mathcal{D}(A)\}$*),* A_0 *is* separable *in the sense that there exists a countable collection* $\{g_k\} \subset \mathcal{D}(A)$ *such that* A_0 *is contained in the bounded, pointwise closure of the linear span of* $\{(g_k, A_0 g_k) = (g_k, \psi^{-1}Ag_k)\}$ *in* $B(\mathbb{S}) \times B(\mathbb{S})$ *(or in* $B(\mathbb{S}) \times B(\mathbb{S} \times \mathbb{U})$*).*

v) A_0 *is a pre-generator (for each fixed* z*, if* A *is of the form (2.1)), that is,* A_0 *is dissipative and there are sequences of functions* $\mu_n : \mathbb{S} \to \mathcal{P}(\mathbb{S})$ *and* $\lambda_n : \mathbb{S} \to [0,\infty)$ *such that for each* $(f,g) \in A$

$$g(x) = \lim_{n\to\infty} \lambda_n(x) \int_{\mathbb{S}} (f(y) - f(x))\mu_n(x,dy) \tag{2.2}$$

for each $x \in \mathbb{S}$*.*

vi) $\mathcal{D}(A)$ *is closed under multiplication and separates points.*

Remark 2.2 *Suppose that we are interested in a diffusion* X *in a closed set* $\mathbb{S} \subset \mathbb{R}^d$ *with absorbing boundary conditions, that is,*

$$X(t) = X(0) + \int_0^{t\wedge\tau} \sigma(X(s))dW(s) + \int_0^{t\wedge\tau} b(X(s))ds, \tag{2.3}$$

where $\tau = \inf\{t : X(t) \in \partial\mathbb{S}\} = \inf\{t : X(t) \notin \mathbb{S}^o\}$*, where* $\partial\mathbb{S}$ *is the topological boundary of* $\mathbb{S}$ *and* $\mathbb{S}^o$ *is the interior of* $\mathbb{S}$*. Setting* $a(x) = \sigma(x)\sigma(x)^\top$ *and* $Lf(x) = \sum \frac{1}{2}a_{ij}(x)\partial_i\partial_j f(x) + \sum b_i(x)\partial_i f(x)$ *for* $f \in C^2(\mathbb{R}^d)$*, assuming sufficient smoothness, the natural generator would be* Lf *with domain being the* C^2*-functions satisfying* $Lf(x) = 0$*,* $x \in \partial\mathbb{S}$*. This domain does not satisfy Condition 2.1(vi). However, if we take* $\mathcal{D}(A) = C_b^2(\mathbb{S})$*,* $\mathbb{U} = \{0,1\}$*,* $A_1 f(x,u) = uLf(x)$ *and* $\eta(x,du) = \mathbf{1}_{\mathbb{S}^o}(x)\delta_1(du) + \mathbf{1}_{\partial\mathbb{S}}(x)\delta_0(du)$*, where* δ_0 *and* δ_1 *are the Dirac measures at* 0 *and* 1 *respectively, we have* $Af(x) = \mathbf{1}_{\mathbb{S}^o}(x)Lf(x)$ *with domain satisfying Condition 2.1(vi). Any solution of (2.3) will be a solution of the martingale problem for* A*, and any solution of the martingale problem for* A *will be a solution of the martingale problem for the natural generator.*

Definition 2.3 *Let A satisfy Condition 2.1. A measurable, $\mathbb{S}$-valued process X is a solution of the* martingale problem *for A, if there exists a filtration $\{\mathcal{F}_t\}$ such that X is $\{\mathcal{F}_t\}$-adapted,*

$$\mathbb{E}[\int_0^t \psi(X(s))ds] < \infty, \quad t \geq 0, \tag{2.4}$$

and for each $f \in \mathcal{D}(A)$,

$$f(X(t)) - f(X(0)) - \int_0^t Af(X(s))ds \tag{2.5}$$

is an $\{\mathcal{F}_t\}$-martingale. For $\nu_0 \in \mathcal{P}(\mathbb{S})$, X is a solution of the martingale problem *for (A, ν_0), if X is a solution of the martingale problems for A and $X(0)$ has distribution ν_0.*

A measurable, $\mathbb{S}$-valued process X and a nonnegative random variable τ are a solution of the stopped martingale problem *for A, if there exists a filtration $\{\mathcal{F}_t\}$ such that X is $\{\mathcal{F}_t\}$-adapted, τ is a $\{\mathcal{F}_t\}$-stopping time,*

$$\mathbb{E}[\int_0^{t\wedge\tau} \psi(X(s))ds] < \infty, \quad t \geq 0, \tag{2.6}$$

and for each $f \in \mathcal{D}(A)$,

$$f(X(t\wedge\tau)) - f(X(0)) - \int_0^{t\wedge\tau} Af(X(s))ds \tag{2.7}$$

is an $\{\mathcal{F}_t\}$-martingale.

A measurable, $\mathbb{S}$-valued process X is a solution of the local-martingale problem *for A, if there exists a filtration $\{\mathcal{F}_t\}$ such that X is $\{\mathcal{F}_t\}$-adapted and a sequence $\{\tau_n\}$ of $\{\mathcal{F}_t\}$-stopping times such that $\tau_n \to \infty$ a.s. and for each n, (X, τ_n) is a solution of the stopped martingale problem for A using the filtration $\{\mathcal{F}_t\}$.*

Remark 2.4 *Note that (2.4) ensures the integrability of (2.5) and similarly for the forward equation (2.8). Furthermore, if $M^f(t)$ denotes the process in (2.5), then (2.4) together with Condition 2.1(iii) imply that M^f is a martingale if and only if it is a local martingale, since*

$$\sup_{s\in[0,t]} |M^f(s)| \leq 2\|f\| + \int_0^t \psi(X(s))ds$$

has finite expectation for all $t \geq 0$.

If X is a solution of the local martingale problem for A, then the localizing sequence $\{\tau_n\}$ can be taken to be predictable. In particular, we can take

$$\tau_n = \inf\{t : \int_0^t \psi(X(s))ds \geq n\}.$$

Definition 2.5 Uniqueness *holds for the (local) martingale problem for (A, ν_0) if and only if all solutions have the same finite-dimensional distributions.* Stopped uniqueness *holds if for any two solutions, (X_1, τ_1), (X_2, τ_2), of the stopped martingale problem for (A, ν_0), there exists a stochastic process $\widetilde{X}$ and nonnegative random variables $\widetilde{\tau}_1$, $\widetilde{\tau}_2$ such that $(\widetilde{X}, \widetilde{\tau}_1 \vee \widetilde{\tau}_2)$ is a solution of the stopped martingale problem for (A, ν_0), $(\widetilde{X}(\cdot\wedge\widetilde{\tau}_1), \widetilde{\tau}_1)$ has the same distribution as $(X_1(\cdot\wedge\tau_1), \tau_1)$, and $(\widetilde{X}(\cdot\wedge\widetilde{\tau}_2), \widetilde{\tau}_2)$ has the same distribution as $(X_2(\cdot\wedge\tau_2), \tau_2)$.*

Remark 2.6 *Note that stopped uniqueness implies uniqueness. Stopped uniqueness holds if uniqueness holds and every solution of the stopped martingale problem can be extended (beyond the stopping time) to a solution of the (local) martingale problem. (See Lemma 4.5.16 of Ethier and Kurtz (1986) for conditions under which this extension can be done.)*

Definition 2.7 *A $\mathcal{P}(\mathbb{S})$-valued function $\{\nu_t, t \geq 0\}$ is a solution of the* forward equation *for A if for each $t > 0$, $\int_0^t \nu_s \psi ds < \infty$ and for each $f \in \mathcal{D}(A)$,*

$$\nu_t f = \nu_0 f + \int_0^t \nu_s A f ds. \tag{2.8}$$

A pair of measure-valued functions $\{(\nu_t^0, \nu_t^1), t \geq 0\}$ is a solution of the stopped forward equation *for A if for each $t \geq 0$, $\nu_t \equiv \nu_t^0 + \nu_t^1 \in \mathcal{P}(\mathbb{S})$ and $\int_0^t \nu_s^1 \psi ds < \infty$, $t \to \nu_t^0(C)$ is nondecreasing for all $C \in \mathcal{B}(\mathbb{S})$, and for each $f \in \mathcal{D}(A)$,*

$$\nu_t f = \nu_0 f + \int_0^t \nu_s^1 A f ds. \tag{2.9}$$

A $\mathcal{P}(\mathbb{S})$-valued function $\{\nu_t, t \geq 0\}$ is a solution of the local forward equation *for A if there exists a sequence $\{(\nu^{0,n}, \nu^{1,n})\}$ of solutions of the stopped forward equation for A such that for each $C \in \mathcal{B}(\mathbb{S})$ and $t \geq 0$, $\{\nu_t^{1,n}(C)\}$ is nondecreasing and $\lim_{n\to\infty} \nu_t^{1,n}(C) = \nu_t(C)$.*

Clearly, any solution X of the martingale problem for A gives a solution of the forward equation for A, that is $\nu_t f = \mathbb{E}[f(X(t)]$, and any solution of the stopped martingale problem for A gives a solution of the stopped forward equation for A, that is, $\nu_t^0 f = \mathbb{E}[\mathbf{1}_{[\tau,\infty)}(t) f(X(\tau))]$ and $\nu_t^1 f = \mathbb{E}[\mathbf{1}_{[0,\tau)}(t) f(X(t))]$. The primary consequence of Condition 2.1 is the converse.

Lemma 2.8 *If A satisfies Condition 2.1 and $\{\nu_t, t \geq 0\}$ is a solution of the forward equation for A, then there exists a solution X of the martingale problem for A satisfying $\nu_t f = \mathbb{E}[f(X(t))]$.*

If A satisfies Condition 2.1 and $\{(\nu_t^0, \nu_t^1), t \geq 0\}$ is a solution of the stopped forward equation for A, then there exists a solution (X, τ) of the stopped martingale problem for A such that $\nu_t^0 f = \mathbb{E}[\mathbf{1}_{[\tau,\infty)}(t) f(X(\tau))]$ and $\nu_t^1 f = \mathbb{E}[\mathbf{1}_{[0,\tau)}(t) f(X(t))]$.

If A satisfies Condition 2.1 and $\{\nu_t, t \geq 0\}$ is a solution of the local forward equation for A, then there exists a solution X of the local martingale problem for A satisfying $\nu_t f = \mathbb{E}[f(X(t)]$.

Proof. Various forms of the first part of this result exist in the literature beginning with the result of Echeverría (1982) for the stationary case, that is, $\nu_t \equiv \nu_0$ and $\nu_0 A f = 0$. Extension of Echeverria's result to the forward equation is given in Theorem 4.9.19 of Ethier and Kurtz (1986) for locally compact spaces and in Theorem 3.1 of Bhatt and Karandikar (1993) for general complete separable metric spaces. The version given here is a special case of Corollary 1.12 of Kurtz and Stockbridge (2001).

The the result for stopped forward equations also follows by the same corollary. First enlarge the state space $\widetilde{\mathbb{S}} = \mathbb{S} \times \{0, 1\}$ and define $\widetilde{\nu}_t h = \nu_t^0 h(\cdot, 0) + \nu_t^1 h(\cdot, 1)$. Setting $\mathcal{D}(\widetilde{A}) =$

$\{f(x)g(y) : f \in \mathcal{D}(A), g \in B(\{0,1\})\}$, for $h = fg \in \mathcal{D}(\widetilde{A})$, define $\widetilde{A}h(x,y) = yAh(x,y) = yg(y)Af(x)$ and $Bh(x,y) = y(h(x,0) - h(x,y))$. Then

$$\begin{aligned} 0 &= \nu_t^0 h(\cdot,1) + \nu_t^1 h(\cdot,1) - \nu_0^0 h(\cdot,1) - \nu_0^1 h(\cdot,1) - \int_0^t \widetilde{\nu}_s \widetilde{A} h ds \\ &= \widetilde{\nu}_t h - \widetilde{\nu}_0 h - \int_0^t \widetilde{\nu}_s \widetilde{A} h ds + \nu_t^0 h(\cdot,1) - \nu_t^0 h(\cdot,0) + \nu_0^0 h(\cdot,1) - \nu_0^0 h(\cdot,0) \\ &= \widetilde{\nu}_t h - \widetilde{\nu}_0 h - \int_0^t \widetilde{\nu}_s \widetilde{A} h ds - \int_{\mathbb{S}\times\{0,1\}\times[0,t]} Bh(x,y)\mu(dx \times dy \times ds), \end{aligned}$$

where, noting that $\nu_t^0(C)$ is an increasing function of t, μ is the measure determined by

$$\mu(C \times \{1\} \times [0,t_2]) = \nu_{t_2}^0(C) - \nu_0^0(C), \quad \mu(C \times \{0\} \times [t_1,t_2]) = 0.$$

Corollary 1.12 of Kurtz and Stockbridge (2001) then implies the existence of a process $(\widetilde{X}, Y)$ in $(\mathbb{S} \times \{0,1\})$ such that $\nu_t^0 f = \mathbb{E}[(1 - Y(t))f(\widetilde{X}(t))]$, $\nu_t^1 f = \mathbb{E}[Y(t)f(\widetilde{X}(t))]$, and

$$f(\widetilde{X}(t)) - f(\widetilde{X}(0)) - \int_0^t Y(s)Af(\widetilde{X}(s))ds$$

is a martingale for each $f \in \mathcal{D}(A)$. Following the arguments in Section 2 of Kurtz and Stockbridge (2001), the process can be constructed in such a way that $Y(s) = 0$ implies $Y(t) = 0$ for $t > s$, and hence $\tau = \inf\{t : Y(t) = 0\}$. Note that $\widetilde{X}(t) = \widetilde{X}(\tau)$ for $t \geq \tau$.

Similarly, suppose $\{(\nu^{0,n}, \nu^{1,n})\}$ is the sequence of solutions of the stopped forward equation associated with a solution of the local forward equation and take $(\nu_t^{0,0}, \nu_t^{1,0}) \equiv (\nu_0, 0)$. For $f \in B(\mathbb{S} \times \mathbb{Z}_+)$, define

$$\widehat{\nu}_t f = \sum_{n=1}^{\infty} (\nu_t^{1,n} f(\cdot,n) - \nu_t^{1,n-1} f(\cdot,n))$$

and

$$\int_{\mathbb{S}\times\mathbb{Z}_+\times[0,t]} f(x,n)\widehat{\mu}(dx \times dn \times ds) = \sum_{n=1}^{\infty} \nu_t^{0,n} f(\cdot,n).$$

Note that $\widehat{\nu}_t$ is a probability measure with $\mathbb{S}$-marginal $\nu_t = \lim_{n\to\infty} \nu_t^{1,n}$.

Setting

$$\mathcal{D}(\widetilde{A}) = \mathcal{D}(B) = \{gf : g \in C_c(\mathbb{Z}_+), f \in \mathcal{D}(A)\}$$

(where, of course, $C_c(\mathbb{Z}_+)$ is the collection of functions with finite support) and defining $\widetilde{A}gf(x,n) = g(n)Af(x)$ and $Bgf(x,n) = f(x)(g(n+1) - g(n))$,

$$\widehat{\nu}_t gf = \widehat{\nu}_0 gf + \int_0^t \widehat{\nu}_s \widetilde{A} gf ds + \int_{\mathbb{S}\times\mathbb{Z}_+\times[0,t]} Bgf d\widehat{\mu}.$$

Let $0 < \psi_0(n) < 1$ satisfy

$$\sum_n \psi_0(n) \int_0^n \nu_s^{1,n} \psi ds < \infty.$$

Then $\widetilde{A}$ satisfies Condition (2.1) with ψ replaced by $\widetilde{\psi}(x,n)=\psi(x)\psi_0(n)$, and

$$\int_0^t \widehat{\nu}_s\widetilde{\psi}ds<\infty, \quad t>0.$$

Corollary 1.12 of Kurtz and Stockbridge (2001) then implies the existence of a process (X,N) such that $(X(t),N(t))$ has distribution $\widehat{\nu}_t$ and a random measure Γ on $\mathbb{S}\times\mathbb{Z}_+\times[0,\infty)$ satisfying

$$E[\int_{\mathbb{S}\times\mathbb{Z}_+\times[0,t]} f(x,n)\Gamma(dx\times dn\times ds)]=\int_{\mathbb{S}\times\mathbb{Z}_+\times[0,t]} f(x,n)\widehat{\mu}(dx\times dn\times ds)$$

such that for each $gf\in\mathcal{D}(\widetilde{A})$,

$$g(N(t))f(X(t))-\int_0^t g(N(s))Af(X(s))ds-\int_{\mathbb{S}\times\mathbb{Z}_+\times[0,t]} f(x)Bg(n)\Gamma(dx\times dn\times ds) \tag{2.10}$$

is a $\{\mathcal{F}_t^{X,N}\}$-martingale.

Let $\tau_k=\inf\{t:\int_0^t\psi(X(s))ds\geq k\}$. Let $g_m(n)=\mathbf{1}_{[0,m]}(n)$, and consider the limit of the sequence of martingales

$$\begin{aligned} g_m(N(t\wedge\tau_k))f(X(t\wedge\tau_k))-\int_0^{t\wedge\tau_k} g_m(N(s))Af(X(s))ds \\ -\int_{\mathbb{S}\times\mathbb{Z}_+\times[0,t\wedge\tau_k]} f(x)Bg_m(n)\Gamma(dx\times dn\times ds) \end{aligned} \tag{2.11}$$

as $m\to\infty$. The first two terms converge in L^1 by the dominated convergence theorem, and the third term satisfies

$$\begin{aligned} \mathbb{E}[|\int_{\mathbb{S}\times\mathbb{Z}_+\times[0,t\wedge\tau_k]} f(x)Bg_m(n)\Gamma(dx\times dn\times ds)|] &\leq \|f\|\widehat{\mu}(\mathbb{S}\times\{m\}\times[0,t]) \\ &= \|f\|\nu_t^{0,m}(\mathbb{S}) \end{aligned}$$

and hence converges to zero in L^1. It follows that

$$f(X(t\wedge\tau_k))-\int_0^{t\wedge\tau_k} Af(X(s))ds$$

is a martingale, and consequently, X is a solution of the local martingale problem for A such that $X(t)$ has distribution ν_t. □

Let X be a solution of the martingale problem for A with respect to a filtration $\{\mathcal{F}_t\}$, and let $\{\mathcal{G}_t\}$ be a filtration with $\mathcal{G}_t\subset\mathcal{F}_t$. Then letting π_t denote the conditional distribution of $X(t)$ given $\mathcal{G}_t$, Lemma A.1 implies that for each $f\in\mathcal{D}(A)$,

$$\pi_tf-\pi_0f-\int_0^t\pi_sAfds \tag{2.12}$$

is a $\{\mathcal{G}_t\}$-martingale.

Let $\mathbb{S}_0$ and S_0 be complete, separable metric spaces, and let $\gamma : \mathbb{S} \to \mathbb{S}_0$ be Borel measurable. Let X be a solution of the martingale problem for A, and let Z be a S_0-valued random variable. Assume that $\mathcal{F}_t \supset \sigma(Z)$ for all $t \geq 0$. Define $Y(t) = \gamma(X(t))$,

$$\widehat{\mathcal{F}}_t^Y = \text{completion of } \sigma(\int_0^r g(Y(s))ds : r \leq t, g \in B(\mathbb{S}_0)) \vee \sigma(Y(0), \tag{2.13}$$

$\widehat{\mathcal{F}}_t^{Y,Z} = \widehat{\mathcal{F}}_t^Y \vee \sigma(Z)$, $\pi_t(C) = \mathbb{P}\{X(t) \in C | \widehat{\mathcal{F}}_t^{Y,Z}\}$, where by Theorem A.3 of Kurtz (1998), we can assume that π is a progressively measurable, $\mathcal{P}(\mathbb{S})$-valued process and $\pi_{t\wedge\tau}(C) = \mathbb{P}\{X(t\wedge\tau) \in C | \widehat{\mathcal{F}}_{t\wedge\tau}^{Y,Z}\}$ for every $\{\widehat{\mathcal{F}}_t^{Y,Z}\}$-stopping time τ.

Remark 2.9 *If Y is cadlag with no fixed points of discontinuity, then by Lemma A.5, $\mathcal{F}_t^Y = \widehat{\mathcal{F}}_t^Y$.*

Note that

$$\int_0^t \pi_s(g\circ\gamma)ds = \int_0^t g(Y(s))ds, \quad \text{for each } g \in B(\mathbb{S}_0), \tag{2.14}$$

and if (2.4) holds,

$$\pi_t f - \int_0^t \pi_s Af ds$$

is a $\{\widehat{\mathcal{F}}_t^{Y,Z}\}$-martingale for each $f \in \mathcal{D}(A)$. With these properties in mind, we work with a definition of the filtered martingale problem slightly more general than that of Kurtz (1998).

Definition 2.10 *Let $\widehat{\mu}_0 \in \mathcal{P}(\mathbb{S}\times S_0)$. $(\widetilde{Y}, \widetilde{\pi}, \widetilde{Z}, \widetilde{\tau}) \in M_{\mathbb{S}_0}[0,\infty) \times M_{\mathcal{P}(\mathbb{S})}[0,\infty) \times S_0 \times [0,\infty]$ is a solution of the* stopped, filtered martingale problem *for $(A, \gamma, \widehat{\mu}_0)$, if*

$$\mathbb{E}[\widetilde{\pi}_0(C)\mathbf{1}_D(\widetilde{Z})] = \widehat{\mu}_0(C\times D), \tag{2.15}$$

$\widetilde{\pi}$ is $\{\widehat{\mathcal{F}}_t^{\widetilde{Y}\widetilde{Z}}\}$-adapted, $\widetilde{\tau}$ is a $\{\widehat{\mathcal{F}}_t^{\widetilde{Y},\widetilde{Z}}\}$-stopping time, for each $g \in B(\mathbb{S}_0)$ and $t \geq 0$,

$$\int_0^t \widetilde{\pi}_s(g\circ\gamma)ds = \int_0^t g(\widetilde{Y}(s))ds \;, \tag{2.16}$$

$$\mathbb{E}[\int_0^{t\wedge\widetilde{\tau}} \widetilde{\pi}_s\psi ds] < \infty, \; t > 0, \tag{2.17}$$

and for each $f \in \mathcal{D}(A)$,

$$\widetilde{M}_f(t\wedge\widetilde{\tau}) \equiv \widetilde{\pi}_{t\wedge\widetilde{\tau}} f - \int_0^{t\wedge\widetilde{\tau}} \widetilde{\pi}_s Af ds \tag{2.18}$$

is a $\{\widehat{\mathcal{F}}_t^{\widetilde{Y}\widetilde{Z}}\}$-martingale.

If $(\widetilde{Y}, \widetilde{\pi}, \widetilde{Z}, \widetilde{\tau})$ satisfies all the conditions except (2.15), we will refer to it as a solution of the stopped, filtered martingale problem for (A, γ).

If $\widetilde{\tau} = \infty$ a.s., then $(\widetilde{Y}, \widetilde{\pi}, \widetilde{Z})$ is a solution of the filtered martingale problem *for (A, γ).*

$(\widetilde{Y}, \widetilde{\pi}, \widetilde{Z}) \in M_{\mathbb{S}_0}[0,\infty) \times M_{\mathcal{P}(\mathbb{S})}[0,\infty) \times S_0$ is a solution of the filtered local-martingale problem *for (A, γ) if there exists a sequence $\{\widetilde{\tau}_n\}$ of $\{\widehat{\mathcal{F}}_t^{\widetilde{Y},\widetilde{Z}}\}$-stopping times such that $\widetilde{\tau}_n \to \infty$ a.s. and for each n, $(\widetilde{Y}, \widetilde{\pi}, \widetilde{Z}, \widetilde{\tau}_n)$ is a solution of the stopped, filtered martingale problem for (A, γ).*

Remark 2.11 *By the optional projection theorem (see Theorem A.3 of Kurtz (1998)), there exists a modification of $\widetilde{\pi}$ such that for all $t \geq 0$ and all $\{\widehat{\mathcal{F}}_t^{\widetilde{Y},\widetilde{Z}}\}$-stopping times τ, $\widetilde{\pi}_{t\wedge\tau}$ is $\widehat{\mathcal{F}}_{t\wedge\tau}^{\widetilde{Y},\widetilde{Z}}$-measurable. Consequently, we will assume that $\widetilde{\pi}$ has this property.*

Remark 2.12 *Kurtz and Ocone (1988) consider the filtered martingale problem with $\mathbb{S} = \mathbb{S}_1 \times \mathbb{S}_0$ and γ the projection onto $\mathbb{S}_0$.*

The following lemma is an immediate consequence of (2.16).

Lemma 2.13 *If $(\widetilde{Y}, \widetilde{\pi}, \widetilde{Z}, \widetilde{\tau})$ is a solution of the stopped, filtered martingale problem for $(A, \gamma, \widehat{\mu}_0)$, then $\widehat{\mathcal{F}}_t^{\widetilde{Y}}$ is contained in the completion of $\sigma(\widetilde{\pi}_s, s \leq t) \vee \sigma(\widetilde{Y}(0))$.*

If X is a solution of the martingale problem for A, then $X^{(r)}$ given by $X^{(r)}(t) = X(r+t)$ is also a solution of the martingale problem for A. The following lemma gives the analogous result for filtered martingale problems. Let $\widetilde{Y}_r$ denote the restriction of $\widetilde{Y}$ to $[0, r]$.

Lemma 2.14 *Suppose $(\widetilde{Y}, \widetilde{\pi}, \widetilde{Z}, \widetilde{\tau}) \in M_{\mathbb{S}_0}[0,\infty) \times M_{\mathcal{P}(\mathbb{S})}[0,\infty) \times S_0 \times [0,\infty]$ is a solution of the stopped, filtered martingale problem for (A, γ). For $r \geq 0$ such that $\widetilde{Y}(r)$ is $\widehat{\mathcal{F}}_r^{\widetilde{Y},\widetilde{Z}}$-measurable, let $\widehat{Y}(t) = \widetilde{Y}(r+t)$, $\widehat{Z} = (\widetilde{Z}, \widetilde{Y}_r) \in S_0 \times M[0,r]$, and $\widehat{\pi}_t = \widetilde{\pi}_{r+t}$. Then*

$$(\widehat{Y}, \widehat{\pi}, \widehat{Z}, (\widetilde{\tau} - r) \vee 0) \in M_{\mathbb{S}_0}[0,\infty) \times M_{\mathcal{P}(\mathbb{S})}[0,\infty) \times S_0 \times M[0,r] \times [0,\infty]$$

is a solution of the stopped, filtered martingale problem for (A, γ).

Suppose $\widetilde{\tau} = \infty$, a.s. (that is, $(\widetilde{Y}, \widetilde{\pi}, \widetilde{Z})$ is a solution of the filtered martingale problem for (A, γ)). For $r \geq 0$ such that $\widetilde{Y}(r)$ is $\widehat{\mathcal{F}}_r^{\widetilde{Y},\widetilde{Z}}$-measurable, let $\widehat{Y}(t) = Y(r+t)$, $\widehat{Z} = \widetilde{\pi}_r = \widehat{\pi}_0$, and

$$\widehat{\pi}_t = \mathbb{E}[\widetilde{\pi}_{r+t} | \widehat{\mathcal{F}}_t^{\widehat{Y}} \vee \sigma(Y(r)) \vee \sigma(\widetilde{\pi}_r)].$$

Then $(\widehat{Y}, \widehat{\pi}, \widehat{Z}) \in M_{\mathbb{S}_0}[0,\infty) \times M_{\mathcal{P}(\mathbb{S})}[0,\infty) \times \mathcal{P}(\mathbb{S})$ is a solution of the filtered martingale problem for (A, γ).

Proof. In the second part of the lemma, the existence of $\widehat{\pi}$ as an adapted, $\mathcal{P}(\mathbb{S})$-valued process follows by Theorem A.3 of Kurtz (1998) and

$$\mathbb{E}[\int_0^t \widehat{\pi}_s \psi ds] = \mathbb{E}[\int_r^{r+t} \widetilde{\pi}_s \psi ds] < \infty.$$

In both parts, the required martingale properties follow by Lemma A.1. □

3 Conditional distributions and solutions of martingale problems

Of course, the forward equation is a special case of (2.12) in which the "martingale" is identically zero. Consequently, the following proposition can be viewed as an extension of Lemma 2.8.

Proposition 3.1 *Let A satisfy Condition 2.1. Suppose that $\widetilde{Y}$ is a cadlag, $\mathbb{S}_0$-valued process with no fixed points of discontinuity, $\{\widetilde{\pi}_t, t \geq 0\}$ is a $\mathcal{P}(\mathbb{S})$-valued process, adapted to $\{\mathcal{F}_t^{\widetilde{Y}}\}$, $\int_0^t \widetilde{\pi}_s \psi ds < \infty$ a.s., $t \geq 0$, and*

$$\widetilde{\pi}_t f - \widetilde{\pi}_0 f - \int_0^t \widetilde{\pi}_s A f ds \tag{3.1}$$

is a $\{\mathcal{F}_t^{\widetilde{Y}}\}$-local-martingale for each $f \in \mathcal{D}(A)$. Then there exist a solution X of the local martingale problem for A, a cadlag, $\mathbb{S}_0$-valued process Y, and a $\mathcal{P}(\mathbb{S})$-valued process $\{\pi_t, t \geq 0\}$ such that (Y, π) has the same finite-dimensional distributions as $(\widetilde{Y}, \widetilde{\pi})$ and π_t is the conditional distribution of $X(t)$ given $\mathcal{F}_t^Y$.

For each $t \geq 0$, there exists a Borel measurable mapping $H_t : D_{\mathbb{S}_0}[0,\infty) \to \mathcal{P}(\mathbb{S})$ such that $\pi_t = H_t(Y)$ and $\widetilde{\pi}_t = H_t(\widetilde{Y})$ almost surely.

Proof. As in Kurtz (1998), we begin by enlarging the state space so that the current state of the process contains all information about the past of the observation $\widetilde{Y}$. Let $\{b_k\}, \{c_k\} \subset C_b(\mathbb{S}_0)$ satisfy $0 \leq b_k, c_k \leq 1$, and suppose that the spans of $\{b_k\}$ and $\{c_k\}$ are bounded, pointwise dense in $B(\mathbb{S}_0)$. (Existence of $\{b_k\}$ and $\{c_k\}$ follows from the separability of $\mathbb{S}_0$.) Let $a_1, a_2, \ldots$ be an ordering of the rationals with $a_i \geq 1$. For $k, i \geq 1$, let

$$\begin{aligned} \widetilde{U}_{ki}(t) &= c_k(\widetilde{Y}(0)) - a_i \int_0^t \widetilde{U}_{ki}(s) ds + \int_0^t b_k(\widetilde{Y}(s)) ds \\ &= c_k(\widetilde{Y}(0)) e^{-a_i t} + \int_0^t e^{-a_i(t-s)} b_k(\widetilde{Y}(s)) ds. \end{aligned} \tag{3.2}$$

If we assume $c_1 = 1$ and $b_1 = 0$, $\widetilde{U}_{1i}(t) = e^{-a_i t}$ and $\widetilde{U}_{1i}(t)$ determines the value of t. Let $\widetilde{U}(t) = (\widetilde{U}_{ki}(t) : k, i \geq 1) \in [0,1]^\infty$.

Define $F : (r, y) \in [0,\infty) \times D_{\mathbb{S}_0}[0,\infty) \to u \in [0,1]^\infty$ by

$$u_{ki}(r, y) = c_k(y(0)) e^{-a_i r} + \int_0^r e^{-a_i(r-s)} b_k(y(s)) ds,$$

so $\widetilde{U}(t) = F(t, \widetilde{Y})$. Properties of Laplace transforms and the assumption that $\widetilde{Y}$ has no fixed points of discontinuity imply that there are measurable mappings $\Lambda : [0,1]^\infty \to D_{\mathbb{S}_0}[0,\infty)$ and $\Lambda_0 : [0,1]^\infty \to \mathbb{S}_0$ such that $\Lambda(\widetilde{U}(t)) = \widetilde{Y}(\cdot \wedge t)$ and $\Lambda_0(\widetilde{U}(t)) = \widetilde{Y}(t)$ almost surely. We can define Λ so that if $u_{1,i} = e^{-a_i t}$, then $y = \Lambda(u)$ satisfies $y(s) = y(t-)$ for $s \geq t$. Note that these observations imply that

$$\text{the completion of } \sigma(\widetilde{U}(t)) = \widehat{\mathcal{F}}_t^{\widetilde{Y}} = \mathcal{F}_t^{\widetilde{Y}},$$

where the second equality follows by Lemma A.5.

Let $\widehat{\mathbb{S}} = \mathbb{S} \times [0,1]^\infty$, and let $\mathcal{D}(\widehat{A})$ be the collection of functions on $\widehat{\mathbb{S}}$ given by

$$\{f(x) \prod_{k,i=1}^m g_{ki}(u_{ki}) : f \in \mathcal{D}(A), g_{ki} \in C^1[0,1], m = 1, 2, \ldots\}.$$

Writing $g(u)$ instead of $\prod_{ki} g_{ki}(u_{ki})$ and denoting the partial derivative with respect to u_{ki} by $\partial_{ki} g$, for $fg \in \mathcal{D}(\widehat{A})$,

$$\begin{aligned}
&\widetilde{\pi}_t f g(\widetilde{U}(t)) - \widetilde{\pi}_0 f g(\widetilde{U}(0)) \\
&\quad - \int_0^t \Big(g(\widetilde{U}(s))\widetilde{\pi}_s A f + \widetilde{\pi}_s f \sum (-a_i \widetilde{U}_{ki}(s) + b_k(\widetilde{Y}(s)))\partial_{ki} g(\widetilde{U}(s)) \Big) ds
\end{aligned}$$

is a $\{\mathcal{F}_t^{\widetilde{Y}}\}$-local-martingale. Note that without loss of generality, we can take the localizing sequence to be $\widetilde{\tau}_n = \inf\{t : \int_0^t \widetilde{\pi}_s \psi ds \geq n\}$.

Define

$$\widehat{A}_1(fg)(x,u,z) = g(u)Af(x) + f(x) \sum (-a_i u + b_k(z))\partial_{ki} g(u),$$

and

$$\widehat{A}(fg)(x,u) = \int \widehat{A}_1(fg)(x,u,z)\eta(x,u,dz), \tag{3.3}$$

where, with reference to (2.1) and the definition of Λ_0, we define $\eta(x,u,dz) = \delta_{\Lambda_0(u)}(dz)$.

Define $\widetilde{\nu}_t^{0,n}, \widetilde{\nu}_t^{1,n} \in \mathcal{M}(\mathbb{S} \times [0,1]^\infty)$ by

$$\begin{aligned}
\widetilde{\nu}_t^{0,n} h &= \mathbb{E}[\mathbf{1}_{[\widetilde{\tau}_n,\infty)}(t) \int_{\mathbb{S}} h(z, \widetilde{U}(\widetilde{\tau}_n))\widetilde{\pi}_{\widetilde{\tau}_n}(dz)] \\
\widetilde{\nu}_t^{1,n} h &= \mathbb{E}[\mathbf{1}_{[0,\widetilde{\tau}_n)}(t) \int_{\mathbb{S}} h(z, \widetilde{U}(t))\widetilde{\pi}_t(dz)] \, .
\end{aligned}$$

Setting $\widetilde{\nu}_t^n = \widetilde{\nu}_t^{0,n} + \widetilde{\nu}_t^{1,n}$, for $fg \in \mathcal{D}(\widehat{A})$,

$$\begin{aligned}
\widetilde{\nu}_t^n(fg) &= \mathbb{E}[\widetilde{\pi}_{t\wedge\widetilde{\tau}_n} f g(\widetilde{U}(t \wedge \widetilde{\tau}_n))] \\
&= \mathbb{E}[\widetilde{\pi}_0 f g(\widetilde{U}(0))] \\
&\quad + \mathbb{E}[\textstyle\int_0^{t\wedge\widetilde{\tau}_n} \Big(g(\widetilde{U}(s))\widetilde{\pi}_s A f + \widetilde{\pi}_s f \sum (-a_i \widetilde{U}_{ki}(s) + b_k(\widetilde{Y}(s)))\partial_{ki} g(\widetilde{U}(s)) \Big) ds] \\
&= \widetilde{\nu}_0(fg) + \textstyle\int_0^t \widetilde{\nu}_s^{1,n} \widehat{A}(fg) ds.
\end{aligned}$$

Consequently, $(\widetilde{\nu}^{0,n}, \widetilde{\nu}^{1,n})$ is a solution of the stopped forward equation for $\widehat{A}$, and $\widetilde{\nu} = \lim_{n\to\infty} \widetilde{\nu}^{1,n}$ is a solution of the local forward equation. By Lemma 2.8, there exists a solution (X, U) of the local martingale problem for $\widehat{A}$, such that

$$\begin{aligned}
\mathbb{E}[f(X(t)) \prod_{k,i=1}^m g_{ki}(U_{ki}(t))] &= \widetilde{\nu}_t(f \prod_{k,i=1}^m g_{ki}) \\
&= \mathbb{E}[\widetilde{\pi}_t f \prod_{k,i=1}^m g_{ki}(\widetilde{U}_{ki}(t))] \, .
\end{aligned} \tag{3.4}$$

It follows that for each t, $U(t)$ and $\widetilde{U}(t)$ have the same distribution. If we define $Y(\cdot \wedge t) = \Lambda(U(t))$, $Y(\cdot \wedge t)$ and $\widetilde{Y}(\cdot \wedge t)$ have the same distribution on $D_{\mathbb{S}_0}[0,\infty)$.

Define π_t as the conditional distribution of $X(t)$ given $\mathcal{F}_t^Y$. Then, for any bounded measurable function g on $[0,1]^\infty$

$$\begin{aligned}
\mathbb{E}[f(X(t))g(U(t)] &= \mathbb{E}[\pi_t f g(U(t))] \\
&= \mathbb{E}[\widetilde{\pi}_t f g(\widetilde{U}(t))] \, .
\end{aligned} \tag{3.5}$$

Since $\mathcal{F}_t^Y$ is the completion of $\sigma(U(t))$ and $\mathcal{F}_t^{\widetilde{Y}}$ is the completion of $\sigma(\widetilde{U}(t))$, for every t, there exist mappings $G_t, \widetilde{G}_t : [0,1]^\infty \to \mathcal{P}(\mathbb{S})$ such that $\pi_t = G_t(U(t))$ a.s. and $\widetilde{\pi}_t = \widetilde{G}_t(\widetilde{U}(t))$ a.s. By (3.4),

$$\begin{aligned} \mathbb{E}[G_t(U(t))fh(U(t))] &= \mathbb{E}[\widetilde{G}_t(\widetilde{U}(t))fh(\widetilde{U}(t))] \\ &= \mathbb{E}[\widetilde{G}_t(U(t))fh(U(t))] \end{aligned} \tag{3.6}$$

for all $h \in B(S_0 \times [0,1]^\infty)$, where the last equality follows from the fact that $U(t)$ and $\widetilde{U}(t)$ have the same distribution. Applying (3.6) with $h = G_t(\cdot)f$ and with $h = \widetilde{G}_t(\cdot)f$, we have

$$\mathbb{E}[G_t(U(t))f\widetilde{G}_t(U(t))f] = \mathbb{E}[\Big(\widetilde{G}_t(U(t))f\Big)^2] = \mathbb{E}[(G_t(U(t))f))^2],$$

and it follows that

$$\mathbb{E}[\Big(G_t(U(t))f - \widetilde{G}_t(U(t))f\Big)^2] = 0 .$$

Consequently, $\widetilde{\pi}_t f = G_t(\widetilde{U}(t))f$ a.s., and hence $(\pi_t, U(t))$ has the same distribution as $(\widetilde{\pi}_t, \widetilde{U}(t))$.

Since $U(t)$ $(\widetilde{U}(t))$ determines $U(s)$ $(\widetilde{U}(s))$ for $s < t$, U and $\widetilde{U}$ have the same distribution on $C_{[0,1]^\infty}[0,\infty)$. Consequently, (π, Y) and $(\widetilde{\pi}, \widetilde{Y})$ have the same finite-dimensional distributions.

The mapping H_t is given by $H_t(y) \equiv G_t(F(t,y))$. □

Corollary 3.2 *Let A satisfy Condition 2.1. Suppose that $\{\widetilde{\pi}_t, t \geq 0\}$ is a cadlag, $\mathcal{P}(\mathbb{S})$-valued process with no fixed points of discontinuity adapted to a complete filtration $\{\widetilde{\mathcal{G}}_t\}$ such that $\int_0^t \widetilde{\pi}_s \psi ds < \infty$ a.s., $t \geq 0$, and*

$$\widetilde{\pi}_t f - \widetilde{\pi}_0 f - \int_0^t \widetilde{\pi}_s A f ds$$

is a $\{\widetilde{\mathcal{G}}_t\}$-local martingale for each $f \in \mathcal{D}(A)$. Then there exists a solution X of the local martingale problem for A, a $\mathcal{P}(\mathbb{S})$-valued process $\{\pi_t, t \geq 0\}$ such that $\{\pi_t, t \geq 0\}$ has the same distribution as $\{\widetilde{\pi}_t, t \geq 0\}$, and a filtration $\{\mathcal{G}_t\}$ such that π_t is the conditional distribution of $X(t)$ given $\mathcal{G}_t$.

Proof. The corollary follows by taking $\widetilde{Y}(t) = \widetilde{\pi}_t$ and applying Proposition 3.1. □

The next corollary extends Corollary 3.5 of Kurtz (1998).

Corollary 3.3 *Let A satisfy Condition 2.1. Let $\gamma : \mathbb{S} \to \mathbb{S}_0$ be Borel measurable, and let α be a transition function from $\mathbb{S}_0$ into $\mathbb{S}$ ($y \in \mathbb{S}_0 \to \alpha(y,\cdot) \in \mathcal{P}(\mathbb{S})$ is Borel measurable) satisfying $\alpha(y, \gamma^{-1}(y)) = 1$. Assume that $\widetilde{\psi}(y) \equiv \int_{\mathbb{S}} \psi(z)\alpha(y,dz) < \infty$ for each $y \in \mathbb{S}_0$ and define*

$$C = \{(\int_{\mathbb{S}} f(z)\alpha(\cdot,dz), \int_{\mathbb{S}} Af(z)\alpha(\cdot,dz)) : f \in \mathcal{D}(A)\} .$$

Let $\mu_0 \in \mathcal{P}(\mathbb{S}_0)$, and define $\nu_0 = \int \alpha(y,\cdot)\mu_0(dy)$.

a) If $\widetilde{Y}$ is a solution of the local-martingale problem for (C,μ_0) satisfying $\int_0^t \widetilde{\psi}(\widetilde{Y}(s))ds < \infty$ a.s., then there exists a solution X of the local-martingale problem for (A,ν_0) such that $\widetilde{Y}$ has the same distribution on $M_{\mathbb{S}_0}[0,\infty)$ as $Y=\gamma\circ X$. If Y and $\widetilde{Y}$ are cadlag, then Y and $\widetilde{Y}$ have the same distribution on $D_{\mathbb{S}_0}[0,\infty)$.

b) If $Y(t)$ is $\widehat{\mathcal{F}}_t^Y$-measurable (which by Lemma A.4 holds for almost every t), then $\alpha(Y(t),\cdot)$ is the conditional distribution of $X(t)$ given $\widehat{\mathcal{F}}_t^Y$.

c) If, in addition, uniqueness holds for the martingale problem for (A,ν_0), then uniqueness holds for the $M_{\mathbb{S}_0}[0,\infty)$-martingale problem for (C,μ_0). If $\widetilde{Y}$ has sample paths in $D_{\mathbb{S}_0}[0,\infty)$, then uniqueness holds for the $D_{\mathbb{S}_0}[0,\infty)$-martingale problem for (C,μ_0).

d) If uniqueness holds for the martingale problem for (A,ν_0), then Y restricted to $\mathbf{T}^Y = \{t : Y(t) \text{ is } \widehat{\mathcal{F}}_t^Y\text{-measurable}\}$ is a Markov process.

Proof. We are not assuming that $\widetilde{Y}$ is cadlag, so to apply Proposition 3.1, replace $\widetilde{Y}$ by the continuous process $\widetilde{U}$ given by (3.2). Observing that $\mathcal{F}_t^{\widetilde{U}} = \widehat{\mathcal{F}}_t^{\widetilde{Y}}$, define

$$\widetilde{\pi}_t = \mathbb{E}[\alpha(\widetilde{Y}(t),\cdot)|\mathcal{F}_t^{\widetilde{U}}] = \mathbb{E}[\alpha(\widetilde{Y}(t),\cdot)|\widehat{\mathcal{F}}_t^{\widetilde{Y}}],$$

and note that $\widetilde{\pi}_t = \alpha(Y(t),\cdot)$ for $t\in\mathbf{T}^{\widetilde{Y}}$. Then

$$\widetilde{\pi}_t f - \widetilde{\pi}_0 f - \int_0^t \widetilde{\pi}_s Af ds = \widetilde{\pi}_t f - \alpha f(\widetilde{Y}(0)) - \int_0^t \alpha Af(\widetilde{Y}(s))ds$$

is a $\{\mathcal{F}_t^{\widetilde{U}}\}$-local martingale for each $f\in\mathcal{D}(A)$ and Proposition 3.1 gives the existence of the processes X and U such that X is a solution of the local-martingale problem for A and π_t, the conditional distribution of $X(t)$ given $\mathcal{F}_t^U$ has the distribution as $\widetilde{\pi}_t$. Consequently, for almost every t, $\pi_t = \alpha(\gamma(X(t)),\cdot)$ and it follows that $Y=\gamma\circ X$ has the same distribution on $M_{\mathbb{S}_0}[0,\infty)$ as $\widetilde{Y}$.

Since the finite-dimensional distributions of X are uniquely determined, the distribution of $\gamma\circ X$ (and hence of $\widetilde{Y}$) on $M_{\mathbb{S}_0}[0,\infty)$ is uniquely determined. If $\widetilde{Y}$ has sample paths in $D_{\mathbb{S}_0}[0,\infty)$, then its distribution on $D_{\mathbb{S}_0}[0,\infty)$ is determined by its distribution on $M_{\mathbb{S}_0}[0,\infty)$.

Since X is the unique solution of a martingale problem, by Lemma A.13, it is Markov. The Markov property for Y for $t\in\mathbf{T}^Y$ follows from the Markov property for X by

$$\begin{aligned}
\mathbb{E}[f(Y(t+s))|\widehat{\mathcal{F}}_t^Y] &= \mathbb{E}[\mathbb{E}[f(\gamma(X(t+s)))|\mathcal{F}_t^X]|\widehat{\mathcal{F}}_t^Y] \\
&= \mathbb{E}[h_{f,t,s}(X(t))|\widehat{\mathcal{F}}_t^Y] \\
&= \int_{\mathbb{S}} h_{f,t,s}(x)\alpha(Y(t),dx).
\end{aligned}$$

□

We will also need a stopped version of Proposition 3.1.

Proposition 3.4 *Let A satisfy Condition 2.1. Suppose that $\widetilde{Y}$ is a cadlag, $\mathbb{S}_0$-valued process with no fixed points of discontinuity, $\widetilde{\tau}$ is a $\{\mathcal{F}_t^{\widetilde{Y}}\}$-stopping time, $\{\widetilde{\pi}_t, t \geq 0\}$ is a $\mathcal{P}(\mathbb{S})$-valued process, adapted to $\{\mathcal{F}_t^{\widetilde{Y}}\}$, $\int_0^{t\wedge\tau} \widetilde{\pi}_s \psi ds < \infty$ a.s., $t \geq 0$, and*

$$\widetilde{\pi}_{t\wedge\widetilde{\tau}} f - \widetilde{\pi}_0 f - \int_0^{t\wedge\widetilde{\tau}} \widetilde{\pi}_s A f ds \tag{3.7}$$

is a $\{\mathcal{F}_t^{\widetilde{Y}}\}$-martingale for each $f \in \mathcal{D}(A)$. Then there exist a solution (X,τ) of the stopped martingale problem for A, a cadlag, $\mathbb{S}_0$-valued process Y, and a $\mathcal{P}(\mathbb{S})$-valued process $\{\pi_t, t \geq 0\}$ such that $\{(Y(t\wedge\tau), \pi_t \mathbf{1}_{\{\tau\geq t\}}), t \geq 0\}$ has the same distribution as $\{(\widetilde{Y}(\cdot\wedge\widetilde{\tau}), \widetilde{\pi}_t \mathbf{1}_{\{\widetilde{\tau}\geq t\}})$ and $\pi_{t\wedge\tau}$ is the conditional distribution of $X(t)$ given $\mathcal{F}_{t\wedge\tau}^Y$.

Proof. With $\widehat{A}$ and $\widetilde{U}$ defined as in the proof or Proposition 3.1,

$$\begin{aligned} \widetilde{\nu}_t^0 h &= \mathbb{E}[\mathbf{1}_{[\widetilde{\tau},\infty)}(t) \int_{\mathbb{S}} h(z, \widetilde{U}(\widetilde{\tau}))\widetilde{\pi}_{\widetilde{\tau}}(dz)] \\ \widetilde{\nu}_t^1 h &= \mathbb{E}[\mathbf{1}_{[0,\widetilde{\tau})}(t) \int_{\mathbb{S}} h(z, \widetilde{U}(t))\widetilde{\pi}_t(dz)] \end{aligned}$$

defines a solution of the stopped martingale problem for $\widehat{A}$. Lemma 2.8 ensures the existence of a solution (X, U, τ) of the stopped martingale problem for $\widehat{A}$ such that

$$\mathbb{E}[f(X(t\wedge\tau))g(U(t\wedge\tau))] = \mathbb{E}[\widetilde{\pi}_{t\wedge\widetilde{\tau}} f g(\widetilde{U}(t\wedge\widetilde{\tau}))].$$

Then for $t \geq 0$, $U(t\wedge\tau)$ has the same distribution as $\widetilde{U}(t\wedge\widetilde{\tau})$ and hence

$$Y(\cdot\wedge\tau) = \lim_{t\to\infty} Y(\cdot\wedge t\wedge\tau) \equiv \lim_{t\to\infty} \Lambda(U(t\wedge\tau))$$

has the same distribution as $\widetilde{Y}(\cdot\wedge\widetilde{\tau})$. With reference to Section A.3,

$$\mathbb{E}[f(X(t\wedge\tau))|\sigma(U(t\wedge\tau))] = \mathbb{E}[f(X(t\wedge\tau))|\mathcal{G}_{t\wedge\tau}^Y]$$

has the same distribution as

$$\mathbb{E}[\widetilde{\pi}_{t\wedge\widetilde{\tau}} f|\sigma(\widetilde{U}(t\wedge\widetilde{\tau}))] = \mathbb{E}[\widetilde{\pi}_{t\wedge\widetilde{\tau}} f|\mathcal{G}_{t\wedge\widetilde{\tau}}^{\widetilde{Y}}],$$

and by Lemma A.10, $\widetilde{\pi}_t \mathbf{1}_{\{\widetilde{\tau}\geq t\}}$ has the same distribution as $\pi_t \mathbf{1}_{\{\tau\geq t\}}$, where π_t is the conditional distribution of $X(t\wedge\tau)$ given $\mathcal{F}_{t\wedge\tau}^Y$. □

The only place that Condition 2.1 is used in the proof of Proposition 3.1 is to conclude that every solution of the local forward equation for $\widehat{A}$ defined in (3.3) corresponds to a solution of the local martingale problem. For the filtered martingale problem, $\widehat{A}$ can be given explicitly by

$$\widehat{A}(fg)(x,u) = g(u)Af(x) + f(x)\sum(-a_i u + b_k \circ \gamma(x))\partial_{ki} g(u), \tag{3.8}$$

Consequently, we state the next result under the following hypothesis.

Condition 3.5 *For $\widehat{A}$ defined by (3.8), each solution of the local forward equation for $\widehat{A}$ corresponds to a solution of the local martingale problem for $\widehat{A}$.*

We have the following generalization of Theorem 3.2 of Kurtz (1998).

Theorem 3.6 *Let $A \subset B(\mathbb{S}) \times M(\mathbb{S})$, $\widehat{\mu}_0 \in \mathcal{P}(\mathbb{S} \times S_0)$, and $\gamma : \mathbb{S} \to \mathbb{S}_0$ be Borel measurable, and assume Condition 3.5. Let $(\widetilde{Y}, \widetilde{\pi}, \widetilde{Z})$ be a solution of the local filtered martingale problem for $(A, \gamma, \widehat{\mu}_0)$. Then the following hold:*

a) *There exists a solution X of the local-martingale problem for A and an S_0-valued random variable Z such that $(X(0), Z)$ has distribution $\widehat{\mu}_0$ and $Y = \gamma \circ X$ has the same distribution on $M_{\mathbb{S}_0}[0,\infty)$ as $\widetilde{Y}$.*

b) *Let π_t be the conditional distribution of $X(t)$ given $\widehat{\mathcal{F}}_t^{Y,Z}$. For each $t \geq 0$, there exists a Borel measurable mapping $H_t : M_{\mathbb{S}_0}[0,\infty) \times S_0 \to \mathcal{P}(\mathbb{S})$ such that $\pi_t = H_t(Y, Z)$ and $\widetilde{\pi}_t = H_t(\widetilde{Y}, \widetilde{Z})$.*

c) *If Y and $\widetilde{Y}$ have sample paths in $D_{\mathbb{S}_0}[0,\infty)$, then Y and $\widetilde{Y}$ have the same distribution on $D_{\mathbb{S}_0}[0,\infty)$ and H_t is a Borel measurable mapping from $D_{\mathbb{S}_0}[0,\infty) \times S_0$ to $\mathcal{P}(\mathbb{S})$.*

d) *If uniqueness holds for the local martingale problem for (A, ν_0), then uniqueness holds for the filtered local-martingale problem for $(A, \gamma, \widehat{\mu}_0)$ in the sense that if (Y, π, Z) and $(\widetilde{Y}, \widetilde{\pi}, \widetilde{Z})$ are solutions, then for each $0 \leq t_1 < \cdots < t_m$, $(\pi_{t_1}, \ldots, \pi_{t_m}, Y, Z)$ and $(\widetilde{\pi}_{t_1}, \ldots, \widetilde{\pi}_{t_m}, \widetilde{Y}, \widetilde{Z})$ have the same distribution on $\mathcal{P}(\mathbb{S})^m \times M_{\mathbb{S}_0}[0,\infty) \times S_0$.*

Remark 3.7 *Note that the theorem does not assume that γ is continuous.*

Proof. In the definition of $\widetilde{U}$ in (3.2), replace $c_k(\widetilde{Y}(0))$ by $c_k(\widetilde{Y}(0), \widetilde{Z})$. Note that for a.e. t,

$$f_1(\widetilde{Y}(t \wedge \widetilde{\tau}))\widetilde{\pi}_{t\wedge\widetilde{\tau}} f_2 = \widetilde{\pi}_{t\wedge\widetilde{\tau}}(f_2 f_1 \circ \gamma) \quad a.s. \tag{3.9}$$

(First consider $f_1 = \mathbf{1}_C$, $C \in \mathcal{B}(\mathbb{S}_0)$.)

With $\widetilde{\nu}_t^{n,0}$ and $\widetilde{\nu}_t^{n,1}$ defined as before,

$$\begin{aligned}
\widetilde{\nu}_t^n(fg) &= \mathbb{E}[\widetilde{\pi}_{t\wedge\widetilde{\tau}_n} f g(\widetilde{U}(t \wedge \widetilde{\tau}_n))] \\
&= \mathbb{E}[\widetilde{\pi}_0 f g(\widetilde{U}(0))] \\
&\quad + \mathbb{E}[\textstyle\int_0^{t\wedge\widetilde{\tau}_n} \Big(g(\widetilde{U}(s))\widetilde{\pi}_s Af + \widetilde{\pi}_s f \sum(-a_i \widetilde{U}_{ki}(s) + b_k(\widetilde{Y}(s)))\partial_{ki} g(\widetilde{U}(s)) \Big) ds] \\
&= \mathbb{E}[\widetilde{\pi}_0 f g(\widetilde{U}(0))] \\
&\quad + \mathbb{E}[\textstyle\int_0^{t\wedge\widetilde{\tau}_n} \Big(g(\widetilde{U}(s))\widetilde{\pi}_s Af + \sum(-a_i \widetilde{U}_{ki}(s) + \widetilde{\pi}_s(f b_k \circ \gamma))\partial_{ki} g(\widetilde{U}(s)) \Big) ds] \\
&= \widetilde{\nu}_0(fg) + \textstyle\int_0^t \widetilde{\nu}_s^{1,n} \widehat{A}(fg) ds,
\end{aligned}$$

where the third equality follows from (3.9) and $\widehat{A}$ is defined in (3.8).

We are not assuming that $\widetilde{Y}$ is cadlag, but we still conclude that the completion of $\sigma(\widetilde{U}(t))$ is $\widehat{\mathcal{F}}_t^{\widetilde{Y},\widetilde{Z}}$ and there exist $\Lambda : [0,1]^\infty \to M_{\mathbb{S}_0}[0,\infty) \times S_0$ and $\Lambda_1 : [0,1]^\infty \to S_0$ such that

$\Lambda(\widetilde{U}(t)) = (\widetilde{Y}(\cdot \wedge t), \widetilde{Z})$ and $\Lambda_1(\widetilde{U}(0)) = \widetilde{Z}$. Condition 3.5 ensures the existence of a solution (X, U) of the local martingale problem for $\widehat{A}$ such that U and $\widetilde{U}$ have the same distribution,

$$\begin{aligned} U_{ki}(t) &= U_{ki}(0) - a_i \int_0^t U_{ki}(s)ds + \int_0^t b_k(Y(s))ds \\ &= U_{ki}(0)e^{-a_i t} + \int_0^t e^{-a_i(t-s)} b_k(Y(s))ds, \end{aligned} \tag{3.10}$$

and defining $Z = \Lambda_1(U(0))$, Parts (a) and (b) hold.

Part (c) follows from the fact that the distribution of a cadlag process is determined by its distribution on $M_{\mathbb{S}_0}[0, \infty)$.

Finally, for Part (d), uniqueness for the local martingale problem for (A, ν_0) implies uniqueness for the local martingale problem for $(\widehat{A}, \widehat{\nu}_0)$, where $\widehat{\nu}_0 h = E[\widetilde{\pi}_0 h(\cdot, \widetilde{U}(0))]$. Uniqueness of the distribution of (X, U) in Part (b) implies uniqueness of the distribution of $(\widetilde{Y}, \widetilde{\pi}, \widetilde{Z})$. □

3.1 The Markov property

Uniqueness for martingale problems usually implies the Markov property for solutions, and a similar result holds for filtered martingale problems.

Theorem 3.8 *Let $A \subset B(\mathbb{S}) \times M(\mathbb{S})$, $\widehat{\mu}_0 \in \mathcal{P}(\mathbb{S} \times S_0)$, and $\gamma : \mathbb{S} \to \mathbb{S}_0$ be Borel measurable, and assume Condition 3.5. Let $(\widetilde{Y}, \widetilde{\pi}, \widetilde{Z})$ be a solution of the filtered martingale problem for $(A, \gamma, \widehat{\mu}_0)$. ($\widetilde{\tau} = \infty$.) If uniqueness holds for the martingale problem for (A, ν_0), then $\widetilde{\pi}$ is a $\mathcal{P}(\mathbb{S})$-valued Markov process.*

Proof. Fix $r \geq 0$, and let $(\widehat{Y}, \widehat{\pi})$ be as in the second part of Lemma 2.14. Since $\widehat{\pi}_0 = \widetilde{\pi}_r$, they have the same distribution. By Lemma A.12, a process (Y^*, π^*, Z^*) can be constructed so that $(Y^*(r+\cdot), \pi^*_{r+\cdot}, \pi^*_r)$ has the same distribution on $M_{\mathbb{S}_0 \times \mathcal{P}(\mathbb{S})}[0, \infty) \times \mathcal{P}(\mathbb{S})$ as $(\widehat{Y}, \widehat{\pi}, \widehat{\pi}_0)$, $(Y^*(\cdot \wedge r), \pi^*_{\cdot \wedge r}, Z^*, \pi^*_r)$ has the same distribution on $M_{\mathbb{S}_0 \times \mathcal{P}(\mathbb{S})}[0, r] \times S_0 \times \mathcal{P}(\mathbb{S})$ as $(\widetilde{Y}(\cdot \wedge r), \widetilde{\pi}_{\cdot \wedge r}, \widetilde{Z}, \widetilde{\pi}_r)$, and

$$\mathbb{E}[g(Y^*(r+\cdot), \pi^*_{r+\cdot}) | Y^*(\cdot \wedge r), \pi^*_{\cdot \wedge r}, Z^*, \pi^*_r] = \mathbb{E}[g(Y^*(r+\cdot), \pi^*_{r+\cdot}) | \pi^*_r]. \tag{3.11}$$

We claim that (Y^*, π^*, Z^*) is a solution of the filtered martingale problem for $(A, \gamma, \widehat{\mu}_0)$.

(π^*_0, Z^*) has the same distribution as $(\widetilde{\pi}_0, \widetilde{Z})$, so (2.15) holds. Since $(Y^*(\cdot \wedge r), \pi^*_{\cdot \wedge r})$ has the same distribution as $(\widetilde{Y}(\cdot \wedge r), \widetilde{\pi}_{\cdot \wedge r})$, for $g \in B(\mathbb{S}_0)$ and $t \leq r$,

$$\int_0^t \pi^*_s(g \circ \gamma)ds = \int_0^t g(Y^*(s))ds \quad a.s.$$

For $t > r$, $(\int_r^t \pi^*_s(g \circ \gamma)ds, \int_r^t g(Y^*(s))ds)$ has the same distribution as $(\int_0^{t-r} \widehat{\pi}_s(g \circ \gamma)ds, \int_0^{t-r} g(\widehat{Y}(s))ds)$, so $\int_r^t \pi^*_s(g \circ \gamma)ds = \int_r^t g(Y^*(s))ds$ a.s. Consequently, (2.16) follows.

For $f \in \mathcal{D}(A)$, let $M_f^*(t) = \pi_t^* f - \int_0^t \pi_s^* Af ds$. For $r \leq t < t+h$, let H_1 be a bounded random variable measurable with respect to the completion of

$$\sigma(\int_r^u h(Y^*(s))ds : r \leq u \leq t, h \in B(\mathbb{S}_0)) \vee \sigma(\pi_r^*)$$

and H_2 a bounded random variable measurable with respect to $\widehat{\mathcal{F}}_r^{Y^*,Z^*} = \widehat{\mathcal{F}}_r^{Y^*} \vee \sigma(Z^*)$. Then by (3.11),

$$\mathbb{E}[(M_f^*(t+h) - M_f^*(t))H_1 H_2] = \mathbb{E}[(M_f^*(t+h) - M_f^*(t))H_1 \mathbb{E}[H_2|\pi_r^*]],$$

and the right side is zero by the fact that $(Y^*(r+\cdot), \pi_{r+\cdot}^*)$ has the same distribution as $(\widehat{Y}, \widehat{\pi})$. It follows that

$$\mathbb{E}[M_f^*(t+h) - M_f^*(t)|\widehat{\mathcal{F}}_t^{Y^*,Z^*}] = 0. \tag{3.12}$$

If $t < t+h \leq r$, then (3.12) follows from the fact that $(Y^*(\cdot \wedge r), \pi_{\cdot\wedge r}^*, Z^*, \pi_r^*)$ has the same distribution as $(\widetilde{Y}(\cdot \wedge r), \widetilde{\pi}_{\cdot\wedge r}, \widetilde{Z}, \widetilde{\pi}_r)$, and for $t < r < t+h$,

$$\mathbb{E}[M_f^*(t+h) - M_f^*(t)|\widehat{\mathcal{F}}_t^{Y^*,Z^*}] = \mathbb{E}[M_f^*(t+h) - M_f^*(r) + M_f^*(r) - M_f^*(t)|\widehat{\mathcal{F}}_t^{Y^*,Z^*}] = 0$$

verifying (2.18).

By uniqueness, (π^*, Y^*, Z^*) and $(\widetilde{\pi}, \widetilde{Y}, \widetilde{Z})$ have the same distribution. Consequently, (3.11) implies

$$\mathbb{E}[g(\widetilde{Y}(r+\cdot), \widetilde{\pi}_{r+\cdot})|\widetilde{Y}(\cdot \wedge r), \widetilde{\pi}_{\cdot\wedge r}, \widetilde{Z}, \widetilde{\pi}_r] = \mathbb{E}[g(\widetilde{Y}(r+\cdot), \widetilde{\pi}_{r+\cdot})|\widetilde{\pi}_r],$$

giving the Markov property. □

In the classical setting, $X = (X_1, Y) \in \mathbb{S}_1 \times \mathbb{S}_0$ and $\gamma(X) = Y$, $\pi_t = \pi_t^1 \times \delta_{Y(t)}$, where π_t^1 is the conditional distribution of $X_1(t)$ given $\mathcal{F}_t^Y$. In the observations in additive white noise setting, a number of authors (Kunita (1971); Bhatt, Budhiraja, and Karandikar (2000); Stettner (1989)) have given conditions under which π^1 is Markov. The following example shows that the conclusion does not hold in general, and hence the Markov property for π^1 does not immediately follow from Theorem 3.8.

Example 3.9 *Let (X_1, Y) be the Markov process with values in $\{-1, +1\} \times \mathbb{N}$ and generator*

$$Af(x,y) = \lambda y[f(-x, y+1) - f(x,y)] + \mu[f(-x,y) - f(x,y)].$$

Given a pure jump Markov counting process (a Yule process) Y with intensity λY and an independent Poisson process Z with intensity μ, the process (X_1, Y) can be represented by

$$X_1(t) = (-1)^{Y(t)-Y(0)+Z(t)} X_1(0).$$

Then the conditional distribution of $X_1(t)$ given $\mathcal{F}_t^Y$ is

$$\begin{aligned}\pi_t^1(dx) &= \mathbf{1}_E(Y(t) - Y(0)) \Big(\alpha_t \delta_{+1}(dx) + (1-\alpha_t)\delta_{-1}(dx)\Big) \\ &\quad + \mathbf{1}_O(Y(t) - Y(0)) \Big((1-\alpha_t)\delta_{+1}(dx) + \alpha_t \delta_{-1}(dx)\Big),\end{aligned}$$

where E is the set of even integers, O the set of odd integers, and

$$\alpha_t = \frac{1 + (2\alpha_0 - 1)e^{-2\mu t}}{2},$$

where $\alpha_0 = \mathbb{P}\{X_1(0) = 1|Y(0)\}$.

If $\alpha_0 = \frac{1}{2}$, then $\alpha_t = \frac{1}{2}$ and $\pi^1_t(dx) = \frac{1}{2}\delta_{+1}(dx) + \frac{1}{2}\delta_{-1}(dx)$, for all $t \geq 0$, and π^1 is trivially Markov; however, if $\mathbb{P}\{\alpha_0 \neq \frac{1}{2}\} > 0$, in general π^1 is not Markov. Assuming, for example, that $Y(0) = 1$ and $\alpha_0 = \mathbb{P}\{X_1(0) = 1|Y(0)\} = \mathbb{P}\{X_1(0) = 1\} \neq \frac{1}{2}$, then $\mathcal{F}^{\pi^1}_t \equiv \sigma(\pi^1_s : s \leq t) = \mathcal{F}^Y_t$ ($= \widehat{\mathcal{F}}^Y_t$ by Lemma A.5). Consequently, α_t is deterministic and $\pi^1_t f = g(t, Y(t))$, with

$$g(t,y) = \mathbf{1}_E(y-1)\Big(\alpha_t f(1) + (1-\alpha_t)f(-1)\Big) + \mathbf{1}_O(y-1)\Big((1-\alpha_t)f(1) + \alpha_t f(-1)\Big).$$

Taking into account that $\alpha'_t = \mu(1 - 2\alpha_t)$ and that $\mathbf{1}_E(y-1) = \mathbf{1}_O(y)$, we have

$$\frac{\partial}{\partial t}g(t,y) + \lambda\, y\,[g(t,y+1) - g(t,y)] = (\lambda\, y + \mu)\,(1 - 2\alpha_t)\,[f(-1) - f(1)]\,\big[\mathbf{1}_E(y) - \mathbf{1}_O(y)\big],$$

and therefore

$$\lim_{h\to 0} h^{-1}\mathbb{E}[\pi^1_{t+h}f - \pi^1_t f|\mathcal{F}^{\pi^1}_t] = (\lambda Y(t) + \mu)\Big(\mathbf{1}_E(Y(t)) - \mathbf{1}_O(Y(t))\Big)(1 - 2\alpha_t)\big[f(-1) - f(1)\big].$$

The right side is not just a function of π^1_t, and it follows that π^1 is not a Markov process.

Of course, there is additional structure in the classical example with

$$Y(t) = \sigma W(t) + \int_0^t h(X_1(s))ds, \tag{3.13}$$

W a standard Brownian motion. With this example in mind, we have the following definition.

Definition 3.10 *For $\mathbb{S} = \mathbb{S}_1 \times \mathbb{S}_0$ with $\mathbb{S}_0 = \mathbb{R}^d$ and γ the projection of $\mathbb{S}$ onto $\mathbb{S}_0$, the filtered martingale problem for (A, γ) has* additive observations*, if for each solution $(\widetilde{Y}, \widetilde{\pi}, \widetilde{Z})$, each $r \geq 0$, and each $y \in \mathbb{R}^d$,*

$$\begin{aligned} \widehat{Y}(t) &= \widetilde{Y}(r+t) - \widetilde{Y}(r) + y \\ \widehat{\pi}^1_t &= \mathbb{E}[\widetilde{\pi}^1_{r+t}|\widehat{\mathcal{F}}^{\widehat{Y}}_t \vee \sigma(\widetilde{\pi}^1_r)] \end{aligned} \tag{3.14}$$

determines a solution $(\widehat{Y}, \widehat{\pi}^1 \times \delta_{\widehat{Y}}, \widehat{\pi}^1_0 \times \delta_y)$ of the filtered martingale problem for (A, γ).

Lemma 3.11 *Let $\mathbb{S} = \mathbb{S}_1 \times \mathbb{S}_0$ with $\mathbb{S}_0 = \mathbb{R}^d$ and γ be the projection of $\mathbb{S}$ onto $\mathbb{S}_0$, and suppose that A satisfies Condition 2.1. Assume that every solution $X = (X_1, Y)$ of the martingale problem for A has a version such that Y is cadlag with no fixed points of discontinuity and for each $r \geq 0$ and $y \in \mathbb{R}^d$, $(X_1(\cdot + r), Y(\cdot + r) - Y(r) + y)$ is a solution of the martingale problem for A. Then the filtered martingale problem for (A, γ) has additive observations.*

Remark 3.12 *If X_1 is the solution of the martingale problem for a generator L satisfying Condition 2.1 with $\mathbb{S}$ replaced by $\mathbb{S}_1$, $((a_{ij})) = \sigma\sigma^\top$ and*

$$A[f_1 f_2] = f_2 L f_1 + f_1(\frac{1}{2}\sum_{i,j} a_{ij}\partial_i\partial_j f_2 + h\cdot\nabla f_2)$$

for $f_1 \in \mathcal{D}(L)$ and $f_2 \in C_c^2(\mathbb{R}^d)$, that is, the $\mathbb{S}_0 = \mathbb{R}^d$ component satisfies (3.13) with W independent of X_1, then the hypotheses of the lemma are satisfied.

Proof. Suppose that $(\widetilde{Y}, \widetilde{\pi}, \widetilde{Z})$ is a solution of the filtered martingale problem for $(A, \gamma, \widehat{\mu}_0)$. Then there exists a solution $X = (X_1, Y)$ of the martingale problem for A and a random variable Z such that $(X_1(0), Y(0), Z)$ has distribution $\widehat{\mu}_0$, Y has the same distribution as $\widetilde{Y}$, and for $t \geq 0$, there exist $H_t : M_{\mathbb{S}_0}[0,\infty) \times S_0 \to \mathcal{P}(\mathbb{S})$ such that $\widetilde{\pi}_t = H_t(\widetilde{Y}, \widetilde{Z})$ and $\pi_t = \pi_t^1 \times \delta_{Y(t)} = H_t(Y, Z)$ is the conditional distribution of $X(t)$ given $\widehat{\mathcal{F}}_t^{Y,Z}$.

By assumption, $(\widehat{X}_1, \widehat{Y}) = (X_1(\cdot + r), Y(\cdot + r) - Y(r) + y)$ is a solution of the martingale problems for A. Consequently, defining $\widehat{\pi}^1$ by

$$\widehat{\pi}_t^1 g = \mathbb{E}[g(X_1(r+t))|\mathcal{F}_t^{\widehat{Y}} \vee \sigma(\pi_r)] = \mathbb{E}[\pi_{r+t}^1 g|\mathcal{F}_t^{\widehat{Y}} \vee \sigma(\pi_r)],$$

$(\widehat{Y}, \widehat{\pi}^1 \times \delta_{\widehat{Y}(\cdot)}, \pi_r)$ is a solution of the filtered martingale problem for (A, γ). Since $(\widehat{Y}, \pi)$ has the same distribution as $(\widetilde{Y}(\cdot + r) - \widetilde{Y}(r) + y, \widetilde{\pi})$, the filtered martingale problem for (A, γ) has additive observations. □

Theorem 3.13 *Let $A \subset B(\mathbb{S}) \times M(\mathbb{S})$ and $\gamma : (x, y) \in \mathbb{S}_1 \times \mathbb{R}^d \to y \in \mathbb{R}^d$, and assume Condition 3.5. Suppose that the filtered martingale problem for (A, γ) has additive observations. Let $(\widetilde{Y}, \widetilde{\pi}^1 \times \delta_{\widetilde{Y}}, \widetilde{Z})$ be a solution of the filtered martingale problem for $(A, \gamma, \widehat{\mu}_0)$. If uniqueness holds for the martingale problem for (A, ν_0), then $\widetilde{\pi}^1$ is a $\mathcal{P}(\mathbb{S}_1)$-valued Markov process.*

Proof. As in the proof of Theorem 3.8, fix $r \geq 0$, and let $(\widehat{Y}, \widehat{\pi})$ be as in (3.14) with $y = 0$. Since $\widehat{\pi}_0^1 = \widetilde{\pi}_r{}^1$, they have the same distribution. By Lemma A.12, a process (Y^*, π^{1*}, Z^*) can be constructed so that $(Y^*(r+\cdot) - Y^*(r), \pi_{r+\cdot}^{1*}, \pi_r^{1*})$ has the same distribution on $M_{\mathbb{R}^d \times \mathcal{P}(\mathbb{S}_1)}[0,\infty) \times \mathcal{P}(\mathbb{S}_1)$ as $(\widehat{Y}, \widehat{\pi}^1, \widehat{\pi}_0^1)$, $(Y^*(\cdot \wedge r), \pi_{\cdot\wedge r}^{1*}, Z^*, \pi_r^{1*})$ has the same distribution on $M_{\mathbb{R}^d \times \mathcal{P}(\mathbb{S}_1)}[0, r] \times S_0 \times \mathcal{P}(\mathbb{S}_1)$ as $(\widetilde{Y}(\cdot \wedge r), \widetilde{\pi}_{\cdot\wedge r}^1, \widetilde{Z}, \widetilde{\pi}_r^1)$, and

$$\begin{aligned}\mathbb{E}[g(Y^*(r+\cdot) - Y^*(r), \pi_{r+\cdot}^{1*})|Y^*(\cdot \wedge r), \pi_{\cdot\wedge r}^{1*}, Z^*, \pi_r^{1*}]\\ = \mathbb{E}[g(Y^*(r+\cdot) - Y^*(r), \pi_{r+\cdot}^{1*})|\pi_r^{1*}].\end{aligned}$$

Employing the assumption of additive observations, the proof that (Y^*, π^*, Z^*) is a solution of the filtered martingale problem for $(A, \gamma, \widehat{\mu}_0)$ and the proof of the Markov property are essentially the same as before. □

Remark 3.14 *Theorem 3.13 can be extended to processes in which $\mathbb{S}_0$ is a group and the definition of $\widehat{Y}$ in (3.14) is replaced by $\widehat{Y}(t) = y\widetilde{Y}(r)^{-1}\widetilde{Y}(r+t)$.*

4 Filtering equations

In the nonlinear filtering literature, a filtering equation is a collection of identities satisfied by $\{\pi_t f, f \in \mathcal{D}\}$ and the observation process Y for a set of test functions $\mathcal{D}$. The set $\mathcal{D}$ should be small enough to handle easily, but large enough to insure that the identities uniquely determine π as a function of Y. Uniqueness means that if $\widetilde{\pi}$ is another $\mathcal{P}(\mathbb{S})$-valued process adapted to $\{\mathcal{F}_t^Y\}$ and $\{\widetilde{\pi} f, f \in \mathcal{D}\}$ and Y satisfy the identities, then $\widetilde{\pi} = \pi$.

The results of Section 3 can be exploited to prove uniqueness for a filtering equation provided each solution of the filtering equation has the appropriate martingale properties. Then, Proposition 3.1 ensures that

$$\widetilde{\pi}_t = H_t(Y) = \pi_t \quad a.s. \tag{4.1}$$

In practice, it frequently turns out that verifying that $\widetilde{\pi}$ "has the appropriate martingale properties" requires a change of measure, but since the new measure is equivalent to the original measure, (4.1) still holds. In the next section, we illustrate this argument in the classical setting of a signal in additive white noise.

4.1 Filtering equations for a signal in additive white noise

We consider the classical *Markov model with additive white noise*, in which

$$Y(t) = Y(0) + W(t) + \int_0^t h(X(s))ds, \tag{4.2}$$

where W is a standard d-dimensional Brownian motion and $h = (h_1, \cdots, h_d)^\top$ is a Borel function. Note that we are assuming $\mathbb{S}_0 = \mathbb{R}^d$.

Condition 4.1 *The process X is a cadlag process with values in $\mathbb{S}$ and is a solution of the martingale problem for (A, ν_0), where the operator A satisfies Condition 2.1 and uniqueness holds for the stopped martingale problem for (A, ν_0).*

Note that we are not assuming the independence of X and W.

Define

$$M_f(t) := f(X(t)) - \int_0^t Af(X(s))\,ds, \qquad f \in \mathcal{D}(A),$$

and assume that, for $i = 1, \cdots, d$,

$$\langle M_f, W_i\rangle_t = \int_0^t C_i f(X(s), Y(s))\,ds,$$

where C_i is an operator mapping $\mathcal{D}(A)$ into $M(\mathbb{S}_0 \times \mathbb{R}^d)$. We assume the following condition on the C_i.

Condition 4.2 *There exist a function $\psi_0(x, y) \geq 1$ and constants c_f such that for each $i = 1, \cdots, d$ and $f \in \mathcal{D}(A)$,*

$$|C_i f(x, y)| \leq c_f \psi_0(x, y). \tag{4.3}$$

Remark 4.3 *Note that if* $f = 1$*, then* $M_f(t) = 1$ *and therefore*

$$\langle M_f, W_i \rangle_t = 0, \qquad t \geq 0,$$

that is, $C_i 1 = 0$.

In the uncorrelated case, that is, when X *and* W *are independent,* $C_i f = 0$*, for all* $f \in \mathcal{D}(A)$.

Example 4.4 *Let* X *be a diffusion process with values in* $\mathbb{R}^m$ *solving*

$$dX(t) = b(X(t))\,dt + \sigma(X(t))\,dW(t) + \overline{\sigma}(X(t))\,d\overline{W}(t),$$

with $\overline{W}$ *a Wiener process, independent of* W*. Then*

$$C_i f(x, y) = \sum_{j=1}^{m} \partial_j f(x) \sigma_{j,i}(x), \qquad f \in C_c^2(\mathbb{R}^m),$$

where $\partial_j f$ *denotes the partial derivative of* f *with respect to the* j*-th component of* x.

We assume the following integrability conditions.

Condition 4.5 *For each* $t > 0$*,*

$$\mathbb{E}\left[\int_0^t (\psi(X(s)) + \psi_0(X(s), Y(s)) + |h(X(s))|)\,ds\right] < \infty, \quad t \geq 0, \tag{4.4}$$

and

$$\int_0^t (\pi_s \psi + (\pi_s \psi_0(\cdot, Y(s)))^2 + (\pi_s |h|)^2) ds < \infty \quad a.s. \tag{4.5}$$

Under Condition 4.5, the innovation process $I^\pi(t) = Y(t) - Y(0) - \int_0^t \pi_s h\,ds$ is a Brownian motion. (Note that we are not assuming that $\mathbb{E}\left[\int_0^t |h(X(s))|^2\,ds\right] < \infty$. $\mathbb{E}\left[\int |h(X(s)|ds\right] < \infty$ is sufficient to ensure that I^π is a martingale.) For all $f \in \mathcal{D}(A)$, π satisfies

$$\begin{aligned} \pi_t f &= \pi_0 f + \int_0^t \pi_s A f\,ds \\ &\quad + \int_0^t \left[\pi_s(hf + Cf(\cdot, Y(s)) - \pi_s h\,\pi_s f\right] \left[dY(s) - \pi_s h\,ds\right], \end{aligned}$$

where $Cf(x_1, y) = (C_1 f(x_1, y), \cdots, C_p f(x_1, y))$, or equivalently,

$$\pi_t f = \pi_0 f + \int_0^t \pi_s A f\,ds + \int_0^t \left[\pi_s(hf + Cf(\cdot, Y(s)) - \pi_s h\,\pi_s f\right] dI^\pi(s). \tag{4.6}$$

Define the unnormalized filter for X as the $\mathcal{M}(\mathbb{S})$-valued process

$$\rho_t(dx) := \mathcal{Z}_t^\pi\,\pi_t(dx), \tag{4.7}$$

where

$$\mathcal{Z}_t^\pi := \exp\left\{\int_0^t \pi_s(h)\,dY(s) - \frac{1}{2}\int_0^t |\pi_s(h)|^2\,ds\right\}. \tag{4.8}$$

By Itô's formula, one can show that ρ satisfies the Duncan-Mortensen-Zakai unnormalized filtering equation,

$$\rho_t f = \pi_0 f + \int_0^t \rho_s Af\,ds + \int_0^t \rho_s\big(hf + Cf(\cdot, Y(s))\big)\,dY(s), \quad \forall f \in \mathcal{D}(A).$$

The following result, essentially Theorem 9.1 of Bhatt, Kallianpur, and Karandikar (1995), extends Theorems 4.1 and 4.5 of Kurtz and Ocone (1988) and gives uniqueness of the Kushner-Stratonovich and Fujisaki-Kallianpur-Kunita equations.

Theorem 4.6 *Assume that Conditions 4.1, 4.2, and 4.5, are satisfied. Let $\{\mu_t\}$ be a $\{\mathcal{F}_t^Y\}$-adapted, cadlag $\mathcal{P}(\mathbb{S})$-valued process satisfying*

$$\int_0^t (\mu_s\psi + (\mu_s\psi_0(\cdot, Y(s)))^2 + (\mu_s|h|)^2)ds < \infty, \quad a.s., \tag{4.9}$$

and for $f \in \mathcal{D}(A)$,

$$\begin{aligned}\mu_t f &= \pi_0 f + \int_0^t \mu_s Af\,ds \\ &\quad + \int_0^t \left[\mu_s\big(hf + Cf(\cdot, Y(s))\big) - \mu_s h \mu_s f\right]\,dI^\mu(s),\end{aligned} \tag{4.10}$$

where $I^\mu(t) = Y(t) - \int_0^t \mu_s h\,ds$. Then $\mu_t = \pi_t$, $t \geq 0$, a.s.

Remark 4.7 *There is a large literature on uniqueness for the filtering equations with varying assumptions depending on the techniques used by the authors. Kurtz and Ocone (1988), Bhatt, Kallianpur, and Karandikar (1995), Lucic and Heunis (2001), and Rozovskiĭ (1991) provide a reasonable sampling of results and methods. Our introduction and exploitation of the filtered local-martingale problem allows us to avoid a number of assumptions that appear in many of the earlier results. In particular, we do not assume that h is continuous. We only require the first moment assumption $\mathbb{E}[\int_0^t |h(X_1(s))|ds] < \infty$ rather than a second moment assumption. (Note that there is no expectation in (4.5) and (4.9).) There are no a priori moment assumptions on the solution μ (only on the true conditional distribution).*

Proof. If I^μ is a $\{\mathcal{F}_t^Y\}$-local martingale, then

$$\mu_t f - \pi_0 f - \int_0^t \mu_s Af\,ds$$

is a $\{\mathcal{F}_t^Y\}$-local martingale, and Proposition 3.1 implies that there exist a solution X^* of the local martingale problem for (A, ν_0), a $\mathbb{S}_0$-valued process Y^*, and a $\mathcal{P}(\mathbb{S})$-valued process π^* such that (π^*, Y^*) and (μ, Y) have the same distribution and π_t^* is the conditional distribution of $X^*(t)$ given $\mathcal{F}_t^{Y^*}$. In addition, there exists H_t such that $\pi_t^* = H_t(Y^*)$ and $\mu_t = H_t(Y)$, and the assumption of uniqueness for the martingale problem for (A, ν_0) ensures that $\mu_t = H_t(Y) = \pi_t$.

Unfortunately, it is not immediately clear that I^μ is a local martingale. However, since $I^\pi(t) = Y(t) - Y(0) - \int_0^t \pi_s h ds$ is a Brownian motion, if we define

$$\xi(t) = \mu_t h - \pi_t h,$$

$$\tau_n = \inf\{t : \int_0^t (|\xi(s)|^2 + |\mu_s h|^2) ds \geq n\},$$

and let Q_n be the probability measure on $\mathcal{Y}_{n\wedge\tau_n}$ given by

$$\frac{dQ_n}{dP} = \exp\{\int_0^{n\wedge\tau_n} \xi(s)^T dI^\pi(s) - \frac{1}{2}\int_0^{n\wedge\tau_n} |\xi(s)|^2 ds\},$$

then under Q_n,

$$I^\mu(t\wedge\tau_n) = I^\pi(t\wedge\tau_n) - \int_0^{t\wedge\tau_n} \xi(s) ds,$$

is a martingale for $0 \leq t \leq n$. Consequently, under Q_n,

$$\mu_{t\wedge\tau_n} f - \pi_0 f - \int_0^{t\wedge\tau_n} \mu_s Af\, ds \tag{4.11}$$

is a $\{\mathcal{F}_t^Y\}$-martingale, and by Proposition 3.4 and uniqueness of the stopped martingale problem for A,

$$\mu_t \mathbf{1}_{\{n\wedge\tau_n \geq t\}} = H_t(Y)\mathbf{1}_{\{n\wedge\tau_n \geq t\}} = \pi_t \mathbf{1}_{\{n\wedge\tau_n \geq t\}}.$$

It follows that $Q_n = P$ on $\mathcal{F}^Y_{n\wedge\tau_n}$, $n = 1, 2, \ldots$ and hence by (4.9), $\tau_n \to \infty$ a.s. and $\mu_t = \pi_t$ a.s. □

Corollary 4.8 *Assume that Conditions 4.1, 4.2, and 4.5 are satisfied. Let $\{\theta_t\}$ be a $\{\mathcal{F}_t^Y\}$-adapted, cadlag $\mathcal{M}(\mathbb{S})$-valued process satisfying*

$$\int_0^t (\theta_s \psi + (\theta_s \psi_0(\cdot, Y(s)))^2 + (\theta_s |h|)^2) ds < \infty, \quad a.s.,$$

such that for every $f \in \mathcal{D}(A)$,

$$\theta_t f = \pi_0 f + \int_0^t \theta_s Af\, ds + \int_0^t \theta_s\big(hf + Cf(\cdot, Y(s))\big)\, dY(s). \tag{4.12}$$

Then $\theta_t = \rho_t$ a.s. for all $t \geq 0$.

Proof. For $\epsilon > 0$, define $\beta_\epsilon = \inf\{t > 0 : \theta_t 1 \leq \epsilon\}$ and set

$$\mu_t = \frac{\theta_t}{\theta_t 1} \text{ and } 0 \leq t < \beta_0 \equiv \lim_{\epsilon\to 0} \beta_\epsilon.$$

Then, by Itô's formula, μ satisfies (4.10) on $[0, \beta_0)$. Defining τ_n and the appropriate change of measure as in the proof of the previous theorem, it follows that under the new measure, (4.11), with τ_n replaced by $\tau_n \wedge \beta_\epsilon$ is a martingale, and as before, $\mu_t \mathbf{1}_{\{\tau_n \wedge \beta_\epsilon \geq t\}} = \pi_t \mathbf{1}_{\{\tau_n \wedge \beta_\epsilon \geq t\}}$. Letting $n \to \infty$, $\mu_{t\wedge\beta_\epsilon} = \pi_{t\wedge\beta_\epsilon}$.

Observe that, for $t \leq \beta_\epsilon$,

$$\theta_t 1 = 1 + \int_0^t \theta_s 1\, \pi_s h dY(s),$$

so

$$\theta_t 1 = \mathcal{Z}_t^\pi = \exp\{\int_0^t \pi_s h dY(s) - \frac{1}{2}\int_0^t |\pi_s h|^2 ds\}, \quad t < \beta_0.$$

Since for $T > 0$, $\inf_{t\leq T} \mathcal{Z}_t^\pi > 0$, it follows that $\beta_0 = \infty$ and hence

$$\mu_t = \pi_t \text{ and } \theta_t = \theta_t 1\, \pi_t = \rho_t.$$

□

4.2 Related results on filtering equations

The filtered martingale problem was first introduced in Kurtz and Ocone (1988) in the special case considered here, following a question raised by Giorgio Koch: For all f in the domain of A, the process $\pi_t f - \int_0^t \pi_s Af\, ds$ is a $\{\mathcal{F}_t^Y\}$-martingale; can this observation be used to study nonlinear filtering problems? Under the condition that the state space $\mathbb{S}$ is locally compact, the results on the filtered martingale problem in Kurtz and Ocone (1988) give uniqueness of Zakai and Kushner-Stratonovich equations in the natural class of $\{\mathcal{F}_t^Y\}$-adapted measure-valued processes. Stochastic equations relate known random inputs (in our case Y) to unknown random outputs (in our case, the conditional distribution π). Weak uniqueness (or more precisely, joint uniqueness in law) says that the joint distribution of the input and the output is uniquely determined by the distribution of the input. Strong or pathwise uniqueness says that there is a unique, appropriately measurable transformation that maps the input into the output. In our case, since the equations are derived to be satisfied by the conditional distribution, existence of a transformation (H_t of Proposition 3.1 and Theorem 3.6) is immediate. Consequently, it follows by a generalization of a theorem of Engelbert (1991) (see Kurtz (2007), Theorem 3.14) that weak and strong uniqueness are equivalent.

The uniqueness results derived using the filtered martingale problem in Kurtz and Ocone (1988) were extended in Bhatt, Kallianpur, and Karandikar (1995), in particular, eliminating the local compactness assumption on the state space of the signal. Still in the framework of signals observed in Gaussian white noise, Bhatt and Karandikar (1999) goes beyond the classical Markov model with additive white noise and considers diffusive, non-Markovian signal/observation systems. The systems solve stochastic differential equations with coefficients that may depend on the signal and on the whole past trajectory of the observation process. Therefore, in particular, the signal need not be a Markov process. By enlarging the observation space to a suitable space of continuous functions, such a system can be seen as the solution of a martingale problem, and the filtered martingale problem approach can be used.

In Kallianpur and Mandal (2002) the signal is the solution of a stochastic delay-differential equation, and the Wiener processes driving the signal and the observations are independent. In this case the signal state space is enlarged to a suitable space of continuous functions, and the system is seen as the solution of a martingale problem, and again the filtered martingale problem approach can be used.

Filtering models with point process observations can also be analyzed using the filtered martingale problem approach. In Kliemann, Koch, and Marchetti (1990), the signal/observation system is a Markov process, and the signal is a jump-diffusion process (not necessarily Markovian by itself), while the observation process is a counting process, with unbounded intensity. Strong uniqueness for the filtering equation is obtained in a class of probability-valued processes characterized by a suitable second moment growth condition. In Ceci and Gerardi (2001a), the observation is still a counting process, while the signal/observation system is a Markov jump process with values in $\mathbb{S} = \mathbb{R}^d \times \mathbb{N}$ (see also Ceci and Gerardi (2000), where the signal process itself is a jump Markov process); the weak uniqueness of the Kushner-Stratonovich equation is obtained by the filtered martingale problem approach, while the pathwise uniqueness is obtained by a direct method. (Note that the notion of weak solution considered in this paper places additional restrictions on the solution beyond adaptedness and satisfying the identity. Consequently, the equivalence of weak and strong solutions mentioned above does not hold in this context.) Uniqueness of the filtered martingale problem has also been used for a partially observable control problem of a jump Markov system with counting observations in Ceci, Gerardi, and Tardelli (2002) (see also Ceci and Gerardi (1998) and Ceci and Gerardi (2001b)). The case of a marked point observation process has been considered in various papers. In Fan (1996), the signal/observation system is a continuous-time Markov chain, with $\mathbb{S}$ a finite set. In Ceci and Gerardi (2005) and Ceci and Gerardi (2006a), partially observed branching processes are discussed, while Ceci and Gerardi (2006b) focuses on the financial applications of filtering. In all these examples, the observation state space $\mathbb{S}_0$ is discrete but not necessarily finite.

5 Constrained Markov processes

In this section, $\mathcal{M}(\mathbb{S} \times [0,\infty))$ will denote the space of Borel measures μ on $\mathbb{S} \times [0,\infty)$ such that $\mu(\mathbb{S} \times [0,t]) < \infty$ for each $t > 0$.

Let A and B be operators satisfying Condition 2.1 with ψ replaced by ψ_A and ψ_B, respectively, and $\mathcal{D}(A) = \mathcal{D}(B) = \mathcal{D}$, and let $\mathbb{D}$ and $\partial\mathbb{D}$ be closed subsets of $\mathbb{S}$. In many situations, $\partial\mathbb{D}$ will be the topological boundary of $\mathbb{D}$, but that is not necessary.

Definition 5.1 *A measurable, $\mathbb{D}$-valued process X and a random measure Γ in $\mathcal{M}(\partial\mathbb{D} \times [0,\infty))$ give a solution of the* constrained martingale problem *for $(A, B, \mathbb{D}, \partial\mathbb{D})$ if there exists a filtration $\{\mathcal{F}_t\}$ such that X and Γ are $\{\mathcal{F}_t\}$-adapted,*

$$\mathbb{E}[\int_0^t \psi_A(X(s))ds + \int_{\partial\mathbb{D}\times[0,t]} \psi_B(x)\Gamma(dx \times ds)] < \infty, \quad t \geq 0,$$

and for each $f \in \mathcal{D}$,

$$f(X(t)) - \int_0^t Af(X(s))ds - \int_{\partial\mathbb{D}\times[0,t]} Bf(x)\Gamma(dx \times ds)$$

is an $\{\mathcal{F}_t\}$*-martingale. For* $\nu_0 \in \mathcal{P}(\mathbb{D})$*,* (X, Γ) *is a solution of the* constrained martingale problem *for* $(A, B, \mathbb{D}, \partial\mathbb{D}, \nu_0)$ *if* (X, Γ) *is a solution of the* constrained martingale problem *for* $(A, B, \mathbb{D}, \partial\mathbb{D})$ *and* $X(0)$ *has distribution* ν_0*.*

A $\mathcal{P}(\mathbb{D})$*-valued function* $\{\nu_t\}$ *is a solution of the* forward equation *for* $(A, B, \mathbb{D}, \partial\mathbb{D})$ *if for each* $t > 0$*,* $\int_0^t \nu_s \psi_A ds < \infty$ *and there exists* $\mu \in \mathcal{M}(\partial\mathbb{D} \times [0, \infty))$ *such that* $\int_{\partial\mathbb{D}\times[0,t]} \psi_B d\mu < \infty$*,* $t > 0$*, and for each* $f \in \mathcal{D}$*,*

$$\nu_t f = \nu_0 f + \int_0^t \nu_s A f ds + \int_{\partial\mathbb{D}\times[0,t]} Bf(x)\mu(dx \times ds), \quad t \geq 0.$$

Remark 5.2 *By uniqueness for a constrained martingale problem, we mean uniqueness of the finite-dimensional distributions of* X*. We do not expect* Γ *to be unique. For example, there may be a measure* $\widehat{\mu}$ *such that* $\int_{\partial\mathbb{D}} Bf d\widehat{\mu} \equiv 0$*. Then, if* (X, Γ) *is a solution of the constrained martingale problem and*

$$\widehat{\Gamma}(dx \times ds) = \Gamma(dx \times ds) + \widehat{\mu}(dx)ds,$$

then $(X, \widehat{\Gamma})$ *is also a solution.*

The most familiar examples of constrained Markov processes are reflecting diffusion processes satisfying equations of the form

$$X(t) = X(0) + \int_0^t \sigma(X(s))dW(s) + \int_0^t b(X(s))ds + \int_0^t m(X(s))d\xi(s),$$

where X is required to remain in the closure of a domain $\mathbb{D} \subset \mathbb{R}^d$ and ξ is a nondecreasing process that increases only when X is on the topological boundary $\partial\mathbb{D}$ of $\mathbb{D}$. Then

$$Af(x) = \frac{1}{2}\sum_{i,j} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} f(x) + b(x) \cdot \nabla f(x),$$

where $a(x) = ((a_{ij}(x))) = \sigma(x)\sigma(x)^T$,

$$Bf(x) = m(x) \cdot \nabla f(x),$$

and

$$\Gamma(C \times [0, t]) = \int_0^t \mathbf{1}_C(X(s))d\xi(s).$$

As before, let $\gamma : \mathbb{S} \to \mathbb{S}_0$, $Y(t) = \gamma(X(t))$, and

$$\pi_t(\Gamma) = \mathbb{P}\{X(t) \in \Gamma | \widehat{\mathcal{F}}_t^{Y,Z}\}.$$

Then

$$\pi_t f - \int_0^t \pi_s A f ds - \int_{\partial\mathbb{D}\times[0,t]} Bf(x)\widehat{\Gamma}(dx \times ds)$$

is a $\{\widehat{\mathcal{F}}_t^{Y,Z}\}$-martingale, where $\widehat{\Gamma}$ is the dual predictable projection of Γ with respect to $\{\widehat{\mathcal{F}}_t^{Y,Z}\}$. (See Kurtz and Stockbridge (2001), Lemma 6.1.)

Definition 5.3 *Let $\widehat{\mu}_0 \in \mathcal{P}(\mathbb{D} \times S_0)$.*

$$(\widetilde{Y}, \widetilde{\Gamma}, \widetilde{\pi}, \widetilde{Z}) \in M_{\mathbb{S}_0}[0,\infty) \times \mathcal{M}(\partial\mathbb{D} \times [0,\infty)) \times M_{\mathcal{P}(\mathbb{D})}[0,\infty) \times S_0$$

is a solution of the filtered martingale problem *for $(A, B, \mathbb{D}, \partial\mathbb{D}, \gamma, \widehat{\mu}_0)$, if*

$$\mathbb{E}[\widetilde{\pi}_0(C)\mathbf{1}_D(\widetilde{Z})] = \widehat{\mu}_0(C \times D), \quad C \in \mathcal{B}(\mathbb{D}), D \in \mathcal{B}(S_0), \tag{5.1}$$

$\widetilde{\pi}_0$ is $\sigma(\widetilde{Z})$-measurable, $\widetilde{\pi}$ and $\widetilde{\Gamma}$ are $\{\widehat{\mathcal{F}}_t^{\widetilde{Y},\widetilde{Z}}\}$-adapted, for each $g \in B(\mathbb{D})$ and $t \geq 0$,

$$\int_0^t \widetilde{\pi}_s(g \circ \gamma)ds = \int_0^t g(\widetilde{Y}(s))ds ,$$

$$\mathbb{E}[\int_0^t \widetilde{\pi}_s \psi_A ds + \int_{\partial\mathbb{D}\times[0,t]} \psi_B(x)\widetilde{\Gamma}(dx \times ds)] < \infty, \quad t \geq 0,$$

and for each $f \in \mathcal{D}$,

$$\widetilde{\pi}_t f - \int_0^t \widetilde{\pi}_s Af ds - \int_{\partial\mathbb{D}\times[0,t]} Bf(x)\widetilde{\Gamma}(dx \times ds)$$

is an $\{\widehat{\mathcal{F}}_t^{\widetilde{Y},\widetilde{Z}}\}$-martingale.

If $(\widetilde{Y}, \widetilde{\Gamma}, \widetilde{\pi}, \widetilde{Z})$ satisfies all the conditions except (5.1), we will refer to it as a solution of the filtered martingale problems for $(A, B, \mathbb{D}, \partial\mathbb{D}, \gamma)$.

The following extension of Lemma 2.8 follows from Corollary 1.12 of Kurtz and Stockbridge (2001).

Lemma 5.4 *Let A and B satisfy Condition 2.1 with ψ replaced by ψ_A and ψ_B respectively. If $\{\nu_t\}$ is a solution of the forward equation for $(A, B, \mathbb{D}, \partial\mathbb{D})$, then there exists a solution (X, Γ) of the constrained martingale problem for $(A, B, \mathbb{D}, \partial\mathbb{D})$ such that $\nu_t f = \mathbb{E}[f(X(t))]$.*

Define $\widehat{A}$ as in Section 3, and define $\widehat{B}$ by

$$\widehat{B}[ff_0 \prod_{k,i=1}^m g_{ki}](x, z, u) = \left(f_0(z) \prod_{k,i=1}^m g_{ki}(u_{ki}) \right) Bf(x) .$$

As before, if A and B satisfy Condition 2.1, then so do $\widehat{A}$ and $\widehat{B}$. Any solution $(\widetilde{Y}, \widetilde{\Gamma}, \widetilde{\pi}, \widetilde{Z})$ of the filtered martingale problem for $(A, B, \mathbb{D}, \partial\mathbb{D}, \gamma, \widehat{\mu}_0)$ determines a solution of the forward equation for $(\widehat{A}, \widehat{B}, \mathbb{D} \times S_0 \times [0,1]^\infty, \partial\mathbb{D} \times [0,1]^\infty, \widehat{\nu}_0)$, which, by Lemma 5.4 corresponds to a solution of the constrained martingale problem satisfying (3.4). This observation then gives the following analog of Theorem 3.6.

Theorem 5.5 *Let A and B satisfy Condition 2.1, $\widehat{\mu}_0 \in \mathcal{P}(\mathbb{S} \times S_0)$, and $\gamma : \mathbb{S} \to \mathbb{S}_0$ be Borel measurable. Let $(\widetilde{Y}, \widetilde{\Gamma}, \widetilde{\pi}, \widetilde{Z})$ be a solution of the filtered martingale problem for $(A, B, \mathbb{D}, \partial\mathbb{D}, \gamma, \widehat{\mu}_0)$. Then the following hold:*

a) *There exists a solution* (X, Γ) *of the constrained martingale problem for* $(A, B, \mathbb{D}, \partial\mathbb{D})$ *and an* S_0*-valued random variable* Z *such that* $(X(0), Z)$ *has distribution* $\widehat{\mu}_0$ *and* $\widetilde{Y}$ *has the same distribution on* $M_{\mathbb{S}_0}[0,\infty)$ *as* $Y = \gamma \circ X$*.*

b) *For each* $t \geq 0$*, there exists a Borel measurable mapping* $H_t : M_{\mathbb{S}_0}[0,\infty) \times S_0 \to \mathcal{P}(\mathbb{S})$ *such that* $\pi_t = H_t(Y, Z)$ *is the conditional distribution of* $X(t)$ *given* $\widehat{\mathcal{F}}_t^{Y,Z}$*, and* $\widetilde{\pi}_t = H_t(\widetilde{Y}, \widetilde{Z})$ *a.s. In particular,* $\widetilde{\pi}$ *has the same finite-dimensional distributions as* π*.*

c) *If* Y *and* $\widetilde{Y}$ *have sample paths in* $D_{\mathbb{S}_0}[0,\infty)$*, then* Y *and* $\widetilde{Y}$ *have the same distribution on* $D_{\mathbb{S}_0}[0,\infty)$ *and* H_t *is Borel measurable mapping from* $D_{\mathbb{S}_0}[0,\infty) \times S_0$ *to* $\mathcal{P}(\mathbb{S})$*.*

d) *If uniqueness holds for the constrained martingale problem for* $(A, B, \mathbb{D}, \partial\mathbb{D}, \nu_0)$*, then uniqueness holds for the filtered martingale problem for* $(A, B, \mathbb{D}, \partial\mathbb{D}, \gamma, \widehat{\mu}_0)$ *in the sense that if* (Y, Γ, π, Z) *and* $(\widetilde{Y}, \widetilde{\Gamma}, \widetilde{\pi}, \widetilde{Z})$ *are solutions, then for each* $0 \leq t_1 < \cdots < t_m$*,* $(\pi_{t_1}, \ldots, \pi_{t_m}, Y, Z)$ *and* $(\widetilde{\pi}_{t_1}, \ldots, \widetilde{\pi}_{t_m}, \widetilde{Y}, \widetilde{Z})$ *have the same distribution on* $\mathcal{P}(\mathbb{S})^m \times M_{\mathbb{S}_0}[0,\infty) \times S_0$*.*

The analogs of Theorems 3.8 and 3.13 hold by essentially the same arguments as before.

A Appendix

A.1 A martingale lemma

Let $\{\mathcal{F}_t\}$ and $\{\mathcal{G}_t\}$ be filtrations with $\mathcal{G}_t \subset \mathcal{F}_t$.

Lemma A.1 *Suppose* U *and* V *are measurable and* $\{\mathcal{F}_t\}$*-adapted,* $\mathbb{E}[|U(t)| + \int_0^t |V(s)|ds] < \infty$*,* $t \geq 0$*, and*

$$U(t) - \int_0^t V(s)ds$$

is an $\{\mathcal{F}_t\}$*-martingale. Then*

$$\mathbb{E}[U(t)|\mathcal{G}_t] - \int_0^t \mathbb{E}[V(s)|\mathcal{G}_s]ds$$

is a $\{\mathcal{G}_t\}$*-martingale, where we take* $\mathbb{E}[V(s)|\mathcal{G}_s]$ *to be the optional projection of* V*.*

Proof. The lemma follows by the definition and properties of conditional expectations. □

Example A.2 *If* X *is a solution (wrt* $\{\mathcal{F}_t\}$*) of the martingale problem for* A *and*

$$\pi_t(\Gamma) = \mathbb{P}\{X(t) \in \Gamma|\mathcal{G}_t\},$$

then

$$\pi_t f - \int_0^t \pi_s A f ds$$

is a $\{\mathcal{G}_t\}$*-martingale.*

A.2 Filtrations generated by processes

Let Y be a measurable stochastic process. $\mathcal{F}_t^Y$ will denote the completion of $\sigma(Y(s), s \leq t)$ and $\widehat{\mathcal{F}}_t^Y$ will denote the completion of

$$\sigma(\int_0^s h(Y(r))dr, s \leq t, h \in B(\mathbb{S})) \vee \sigma(Y(0)).$$

Lemma A.3 *If Y is $\{\mathcal{F}_t^Y\}$−progressively measurable, then $\widehat{\mathcal{F}}_t^Y \subset \mathcal{F}_t^Y$.*

Lemma A.4 *For almost every t, $Y(t)$ is $\widehat{\mathcal{F}}_t^Y$.*

Proof. For $g \in B(\mathbb{S}_0)$,

$$M_g(t) = \int_0^t g(Y(s))ds - \int_0^t \mathbb{E}[g(Y(s))|\widehat{\mathcal{F}}_s^Y]ds$$

is a continuous $\{\widehat{\mathcal{F}}_t^Y\}$-martingale. (Take $\mathbb{E}[g(Y(s))|\widehat{\mathcal{F}}_s^Y]$ to be the optional projection of $g \circ Y$.) Since M_g is a finite variation process, it must be zero with probability one, and hence, with probability one

$$g(Y(t)) = \mathbb{E}[g(Y(t))|\widehat{\mathcal{F}}_t^Y] \text{ for almost every } t,$$

which in turn implies that for almost every t, $g(Y(t)) = \mathbb{E}[g(Y(t))|\widehat{\mathcal{F}}_t^Y]$ a.s. and $g(Y(t))$ is $\widehat{\mathcal{F}}_t^Y$-measurable. Since $\mathbb{S}_0$ is separable, there exists a countable separating set $\{g_k\}$ such that for almost every t, $g_1(Y(t)), g_2(Y(t)), \ldots$ are $\widehat{\mathcal{F}}_t^Y$-measurable and hence $Y(t)$ is $\widehat{\mathcal{F}}_t^Y$-measurable. □

Lemma A.5 *If Y is cadlag with no fixed points of discontinuity (that is, $\mathbb{P}\{Y(t) = Y(t-)\} = 1$ for all t), then $\mathcal{F}_t^Y = \widehat{\mathcal{F}}_t^Y$, $t \geq 0$.*

Proof. If Y is cadlag, it is $\{\mathcal{F}_t^Y\}$-progressively measurable so $\widehat{\mathcal{F}}_t^Y \subset \mathcal{F}_t^Y$. $Y(0)$ is $\widehat{\mathcal{F}}_0^Y$ measurable by definition. For $t > 0$, since $\mathbb{P}\{Y(t) = Y(t-)\} = 1$, for $g \in C_b(\mathbb{S})$,

$$g(Y(t)) = \lim_{\epsilon \to} \frac{1}{\epsilon} \int_{t-\epsilon}^t g(Y(r))dr \quad a.s.,$$

and hence $Y(t)$ is $\widehat{\mathcal{F}}_t^Y$-measurable. Consequently, $\mathcal{F}_t^Y \subset \widehat{\mathcal{F}}_t^Y$. □

The following lemma implies that most cadlag processes of interest will have no fixed points of discontinuity.

Lemma A.6 *Let U and V be $\mathbb{S}$-valued random variables and let $\mathcal{G}$ be a σ-algebra of events. Suppose that $M \subset C_b(\mathbb{S})$ is separating and*

$$\mathbb{E}[f(U)|\mathcal{G}] = f(V) \tag{A.1}$$

for all $f \in M$*. Then* $U = V$ *a.s.*

In particular, if Y *is cadlag and adapted to* $\{\mathcal{G}_t\}$ *and*

$$\lim_{s\to t-} \mathbb{E}[f(Y(t))|\mathcal{G}_s] = \mathbb{E}[f(Y(t))| \vee_{s<t} \mathcal{G}_s] = f(Y(t-)), \tag{A.2}$$

for $f \in M$*, then* $Y(t) = Y(t-)$ *a.s.*

Remark A.7 *Note that if*

$$f(Y(t)) - \int_0^t Z_f(s)ds$$

is a martingale for Z_f *satisfying* $\mathbb{E}[\int_0^t |Z_f(s)|ds] < \infty$*, then (A.2) holds.*

Proof. Let Z be a nonnegative $\mathcal{G}$-measurable random variable. Then $\mathbb{E}[f(U)Z] = \mathbb{E}[f(V)Z]$ for $f \in M$ and since M is separating, this identity must hold for all $f \in B(\mathbb{S})$. Consequently, (A.1) holds for all $f \in B(\mathbb{S})$, and replacing f by f^2,

$$\mathbb{E}[f^2(U)] = \mathbb{E}[f^2(V)] = \mathbb{E}[\mathbb{E}[f(U)|\mathcal{G}]f(V)] = \mathbb{E}[f(U)f(V)]$$

and hence that $\mathbb{E}[(f(U) - f(V))^2] = 0$ for all $f \in B(\mathbb{S})$. □

A.3 Stopped filtrations and filtrations generated by stopped processes

Let V be a cadlag process, $\mathcal{F}_t^V$ be the completion of $\sigma(V(s), s \le t)$, and $\mathcal{S}$ be the collection of finite, $\{\mathcal{F}_t^V\}$-stopping times. For $\tau \in \mathcal{S}$, define $\mathcal{G}_\tau^V$ to be the completion of $\sigma(V(t \wedge \tau) : t \ge 0) \vee \sigma(\tau)$. Of course, $\mathcal{G}_t^V = \mathcal{F}_t^V$, and more generally $\mathcal{G}_\tau^V = \mathcal{F}_\tau^V$ for all discrete stopping times in $\mathcal{S}$ and $\mathcal{G}_\tau^V \subset \mathcal{F}_\tau^V$ for all $\tau \in \mathcal{S}$, but we do not know whether or not $\mathcal{G}_\tau^V = \mathcal{F}_\tau^V$ for all $\tau \in \mathcal{S}$.

Lemma A.8 *If* $\tau_1, \tau_2 \in \mathcal{S}$*, then*

$$\mathcal{F}_{\tau_1\wedge\tau_2}^V = \{(A_1 \cap \{\tau_1 < \tau_2\}) \cup (A_2 \cap \{\tau_1 \ge \tau_2\}) : A_1 \in \mathcal{F}_{\tau_1}^V, A_2 \in \mathcal{F}_{\tau_2}^V\} \tag{A.3}$$

and for $\tau \in \mathcal{S}$ *and* $t \ge 0$*,*

$$\mathcal{G}_{\tau\wedge t}^V = \{(A_1 \cap \{\tau < t\}) \cup (A_2 \cap \{\tau \ge t\}) : A_1 \in \mathcal{G}_\tau^V, A_2 \in \mathcal{G}_t^V = \mathcal{F}_t^V\}. \tag{A.4}$$

Remark A.9 *The first assertion holds for arbitrary filtrations.*

Proof. Since $\mathcal{F}_{\tau_1\wedge\tau_2}^V \subset \mathcal{F}_{\tau_1}^V \cap \mathcal{F}_{\tau_2}^V$, is follows that $\mathcal{F}_{\tau_1\wedge\tau_2}^V$ is contained in the right side of (A.3). (Take $A_1 = A_2 \in \mathcal{F}_{\tau_1\wedge\tau_2}^V$.)

Observe that $\{\tau_1 < \tau_2\} \in \mathcal{F}_{\tau_1\wedge\tau_2}^V$ since

$$\{\tau_1 < \tau_2\} \cap \{\tau_1 \wedge \tau_2 \le t\} = \{\tau_1 \le t < \tau_2\} \cup \cup_{r\in\mathbb{Q},r\le t}\{\tau_1 \le r < \tau_2\} \in \mathcal{F}_t^V.$$

Now let $A_1 \in \mathcal{F}_{\tau_1}^V$. Then

$$A_1 \cap \{\tau_1 < \tau_2\} \cap \{\tau_1 \wedge \tau_2 \le t\} = A_1 \cap \{\tau_1 < \tau_2\} \cap \{\tau_1 \le t\} \in \mathcal{F}_t^V,$$

so $A_1 \cap \{\tau_1 < \tau_2\} \in \mathcal{F}^V_{\tau_1\wedge\tau_2}$, and for $A_2 \in \mathcal{F}^V_{\tau_2}$,

$$A_2 \cap \{\tau_1 \geq \tau_2\} \cap \{\tau_1 \wedge \tau_2 \leq t\} = A_2 \cap \{\tau_1 \geq \tau_2\} \cap \{\tau_2 \leq t\} \in \mathcal{F}^V_t,$$

so $A_2 \cap \{\tau_1 \geq \tau_2\} \in \mathcal{F}^V_{\tau_1\wedge\tau_2}$.

If $A_1 \in \mathcal{G}^V_\tau$, then A_1 differs from a set of the form $\{(V(\cdot\wedge\tau),\tau) \in C\}$, $C \in \mathcal{B}(D_E[0,\infty)) \times \mathcal{B}([0,\infty))$, by an event of probability zero, and, noting that $\{\tau \wedge t < t\} = \{\tau < t\}$,

$$\{(V(\cdot\wedge\tau),\tau) \in C\} \cap \{\tau < t\} = \{(V(\cdot\wedge\tau\wedge t),\tau\wedge t) \in C\} \cap \{\tau < t\} \in \mathcal{G}^V_{\tau\wedge t}.$$

Similarly,

$$\{V(\cdot\wedge t) \in C\} \cap \{\tau \geq t\} = \{V(\cdot\wedge\tau\wedge t) \in C\} \cap \{\tau \geq t\} \in \mathcal{G}^V_{\tau\wedge t},$$

and the right side of (A.4) is contained in the left.

Finally,

$$\{(V(\cdot\wedge\tau\wedge t),\tau\wedge t) \in C\} \cap \{\tau < t\} = \{(V(\cdot\wedge\tau),\tau) \in C\} \cap \{\tau < t\} \in \mathcal{G}^V_\tau$$

and

$$\{(V(\cdot\wedge\tau\wedge t),\tau\wedge t) \in C\} \cap \{\tau \geq t\} = \{(V(\cdot\wedge t),t) \in C\} \cap \{\tau \geq t\} \in \mathcal{G}^V_t,$$

so the left side of (A.4) is contained in the right. □

We have the following consequence of the previous lemma.

Lemma A.10 *Let $\mathbb{E}[|Z|] < \infty$ and $\tau \in \mathcal{S}$. Then*

$$\mathbb{E}[Z|\mathcal{F}^V_{\tau\wedge t}] = \mathbb{E}[Z|\mathcal{F}^V_\tau]\mathbf{1}_{\{\tau<t\}} + \mathbb{E}[Z|\mathcal{F}^V_t]\mathbf{1}_{\{\tau\geq t\}}$$

and

$$\mathbb{E}[Z|\mathcal{G}^V_{\tau\wedge t}] = \mathbb{E}[Z|\mathcal{G}^V_\tau]\mathbf{1}_{\{\tau<t\}} + \mathbb{E}[Z|\mathcal{G}^V_t]\mathbf{1}_{\{\tau\geq t\}},$$

and since $\mathcal{G}^V_t = \mathcal{F}^V_t$,

$$\mathbb{E}[Z|\mathcal{F}^V_{\tau\wedge t}]\mathbf{1}_{\{\tau\geq t\}} = \mathbb{E}[Z|\mathcal{G}^V_{\tau\wedge t}]\mathbf{1}_{\{\tau\geq t\}}.$$

A.4 Random probability measures as conditional distributions

Lemma A.11 *Let $\widetilde{\pi}$ be a $\mathcal{P}(\mathbb{S})$-valued random variable and $\widetilde{Z}$ a $\mathbb{S}_0$-valued random variable on a probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$. Then there exists a probability space with random variables (X, Z, π) in $\mathbb{S} \times \mathbb{S}_0 \times \mathcal{P}(\mathbb{S})$, and a sub-$\sigma$-algebra $\mathcal{D}$ such that (Z, π) has the same distribution as $(\widetilde{Z}, \widetilde{\pi})$, Z is $\mathcal{D}$-measurable, and π is the conditional distribution of X given $\mathcal{D}$.*

Proof. For $C \in \mathcal{B}(\mathbb{S})$ and $D \in \mathcal{B}(\mathbb{S}_0 \times \mathcal{P}(\mathbb{S}))$,

$$\nu(C, D) = \mathbb{E}[\widetilde{\pi}(C)\mathbf{1}_D(\widetilde{Z}, \widetilde{\pi})]$$

defines a bimeasure on $\mathbb{S} \times (\mathbb{S}_0 \times \mathcal{P}(\mathbb{S}))$. By Morando's theorem (see, for example, Ethier and Kurtz (1986), Appendix 8), ν extends to a probability measure on $\mathbb{S} \times \mathbb{S}_0 \times \mathcal{P}(\mathbb{S})$. Let (X, Z, π) be the coordinate random variables on $(\mathbb{S} \times \mathbb{S}_0 \times \mathcal{P}(\mathbb{S}), \mathcal{B}(\mathbb{S} \times \mathbb{S}_0 \times \mathcal{P}(\mathbb{S})), \nu)$. Then by definition, (Z, π) has the same distribution as $(\widetilde{Z}, \widetilde{\pi})$, and defining $\mathcal{D} = \sigma(Z, \pi)$, the fact that

$$\mathbb{E}[\mathbf{1}_C(X)\mathbf{1}_D(Z, \pi)] = \mathbb{E}[\widetilde{\pi}(C)\mathbf{1}_D(\widetilde{Z}, \widetilde{\pi})] = \mathbb{E}[\pi(C)\mathbf{1}_D(Z, \pi)]$$

implies that π is the conditional distribution of X given $\mathcal{D}$. □

A.5 A coupling lemma

Lemma A.12 *Let S_0, S_1, and S_2 be complete, separable metric spaces, and let $\mu \in \mathcal{P}(S_0, S_1)$ and $\nu \in \mathcal{P}(S_0, S_2)$ satisfy $\mu(\cdot \times S_1) = \nu(\cdot \times S_2)$. Then there exists a probability space and random variables X_0, X_1, X_2 such that (X_0, X_1) has distribution μ, (X_0, X_2) has distribution ν, and*

$$\mathbb{E}[g(X_2)|X_0, X_1] = \mathbb{E}[g(X_2)|X_0], \quad g \in B(S_2),$$

which is equivalent to

$$\mathbb{E}[g_1(X_1)g_2(X_2)|X_0] = \mathbb{E}[g_1(X_1)|X_0]\mathbb{E}[g_2(X_2)|X_0], \quad g_1 \in B(S_1), g_2 \in B(S_2).$$

Proof. The lemma is a consequence of Ethier and Kurtz (1986), Lemma 4.5.15 □

A.6 The Markov property

Lemma A.13 *Let $A \subset B(\mathbb{S}) \times M(\mathbb{S})$ and $\nu_0 \in \mathcal{P}(\mathbb{S})$. Suppose that there exists $\psi \geq 1$ such that for each $f \in \mathcal{D}(A)$, there exists a constant a_f such that $|Af(x)| \leq a_f\psi(x)$. Assume that uniqueness holds for the local-martingale problem for (A, ν_0). If X is a solution of the local-martingale problem for (A, ν_0) with respect to the filtration*

$$\mathcal{F}_t = \sigma(X(s) : s \leq t) \vee \sigma(\int_0^s h(X(r))dr : s \leq t, h \in B(\mathbb{S})),$$

then X is an $\{\mathcal{F}_t\}$-Markov process.

Proof. Let $Y(t) = \int_0^t \psi(X(s))ds$, and for $f \in \mathcal{D}(A)$ and $g \in C_c^1[0, \infty)$, define

$$\widehat{A}(fg)(x, y) = g(y)Af(x) + \psi(x)f(x)g'(y).$$

Note that since g has compact support, there exists a constant $a_{f,g}$ such that $|\widehat{A}(fg)(x, y)| \leq a_{f,g}(1 + y)^{-2}\psi(x)$, and

$$\int_0^t (1 + Y(s))^{-2}\psi(X(s))ds \leq 1.$$

Consequently, if X is the unique solution of the local-martingale problem for (A, ν_0), then (X, Y) is the unique solution of the martingale problem for $(\widehat{A}, \nu_0 \times \delta_0)$. The proof of Theorem A.5 in Kurtz (1998) remains valid after replacing the assumption $A \subset B(\mathbb{S}) \times B(\mathbb{S})$ with $A \subset B(\mathbb{S}) \times M(\mathbb{S})$, □

References

Robert F. Anderson. Diffusions with second order boundary conditions. I,II. *Indiana Univ. Math. J.*, 25(4):367–395,403–441, 1976. ISSN 0022-2518. 1

Abhay G. Bhatt and Rajeeva L. Karandikar. Invariant measures and evolution equations for Markov processes characterized via martingale problems. *Ann. Probab.*, 21(4):2246–2268, 1993. ISSN 0091-1798. 2

Abhay G. Bhatt and Rajeeva L. Karandikar. Characterization of the optimal filter: the non-Markov case. *Stochastics Stochastics Rep.*, 66(3-4):177–204, 1999. ISSN 1045-1129. 4.2

Abhay G. Bhatt, G. Kallianpur, and Rajeeva L. Karandikar. Uniqueness and robustness of solution of measure-valued equations of nonlinear filtering. *Ann. Probab.*, 23(4):1895–1938, 1995. ISSN 0091-1798. 4.1, 4.7, 4.2

Abhay G. Bhatt, Amarjit Budhiraja, and Rajeeva L. Karandikar. Markov property and ergodicity of the nonlinear filter. *SIAM J. Control Optim.*, 39(3):928–949 (electronic), 2000. ISSN 0363-0129. 1, 3.1

Amarjit Budhiraja. Asymptotic stability, ergodicity and other asymptotic properties of the nonlinear filter. *Ann. Inst. H. Poincaré Probab. Statist.*, 39(6):919–941, 2003. ISSN 0246-0203. 1

Claudia Ceci and Anna Gerardi. Multitype branching processes observing particles of a given type. *J. Appl. Probab.*, 42(2):446–462, 2005. ISSN 0021-9002. 4.2

Claudia Ceci and Anna Gerardi. Modelling a multitype branching Brownian motion: filtering of a measure-valued process. *Acta Appl. Math.*, 91(1):39–66, 2006a. ISSN 0167-8019. 4.2

Claudia Ceci and Anna Gerardi. A mode for high frequency data under partial information: a filtering approach. *Int. J. Theor. Appl. Finance*, 9(4):555–576, 2006b. ISSN 0219-0249. 4.2

Claudia Ceci and Anna Gerardi. Filtering of a Markov jump process with counting observations. *Appl. Math. Optim.*, 42(1):1–18, 2000. ISSN 0095-4616. 4.2

Claudia Ceci and Anna Gerardi. Nonlinear filtering equation of a jump process with counting observations. *Acta Appl. Math.*, 66(2):139–154, 2001a. ISSN 0167-8019. 4.2

Claudia Ceci and Anna Gerardi. Controlled partially observed jump processes: dynamics dependent on the observed history. In *Proceedings of the Third World Congress of Nonlinear Analysts, Part 4 (Catania, 2000)*, volume 47, pages 2449–2460, 2001b. 4.2

Claudia Ceci and Anna Gerardi. Partially observed control of a Markov jump process with counting observations: equivalence with the separated problem. *Stochastic Process. Appl.*, 78(2):245–260, 1998. ISSN 0304-4149. 4.2

Claudia Ceci, Anna Gerardi, and Paola Tardelli. Existence of optimal controls for partially observed jump processes. *Acta Appl. Math.*, 74(2):155–175, 2002. ISSN 0167-8019. 4.2

Pedro Echeverría. A criterion for invariant measures of Markov processes. *Z. Wahrsch. Verw. Gebiete*, 61(1):1–16, 1982. ISSN 0044-3719. 2

H. J. Engelbert. On the theorem of T. Yamada and S. Watanabe. *Stochastics Stochastics Rep.*, 36(3-4):205–216, 1991. ISSN 1045-1129. 4.2

Stewart N. Ethier and Thomas G. Kurtz. *Markov processes: Characterization and convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. ISBN 0-471-08186-8. 2.6, 2, A.4, A.5

Kaisheng Fan. On a new approach to the solution of the nonlinear filtering equation of jump processes. *Probab. Engrg. Inform. Sci.*, 10(1):153–163, 1996. ISSN 0269-9648. 4.2

G. Kallianpur and P. K. Mandal. Nonlinear filtering with stochastic delay equations. In *Advances on theoretical and methodological aspects of probability and statistics (Hamilton, ON, 1998)*, pages 3–36. Taylor & Francis, London, 2002. 4.2

W. Kliemann, G. Koch, and F. Marchetti. On the unnormalized solution of the filtering problem with counting observations. *IEEE Trans. Inform. Theory*, 316(6):1415–1425, 1990. 4.2

Hiroshi Kunita. Asymptotic behavior of the nonlinear filtering errors of Markov processes. *J. Multivariate Anal.*, 1:365–393, 1971. ISSN 0047-259x. 1, 3.1

Thomas G. Kurtz. The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities. *Electron. J. Probab.*, 12:951–965 (electronic), 2007. ISSN 1083-6489. 4.2

Thomas G. Kurtz. Martingale problems for constrained Markov problems. In *Recent advances in stochastic calculus (College Park, MD, 1987)*, Progr. Automat. Info. Systems, pages 151–168. Springer, New York, 1990. 1

Thomas G. Kurtz. A control formulation for constrained Markov processes. In *Mathematics of random media (Blacksburg, VA, 1989)*, volume 27 of *Lectures in Appl. Math.*, pages 139–150. Amer. Math. Soc., Providence, RI, 1991. 1

Thomas G. Kurtz. Martingale problems for conditional distributions of Markov processes. *Electron. J. Probab.*, 3:no. 9, 29 pp. (electronic), 1998. ISSN 1083-6489. 1, 2, 2, 2.11, 2, 3, 3, 3, A.6

Thomas G. Kurtz and Daniel L. Ocone. Unique characterization of conditional distributions in nonlinear filtering. *Ann. Probab.*, 16(1):80–107, 1988. ISSN 0091-1798. 1, 2.12, 4.1, 4.7, 4.2

Thomas G. Kurtz and Richard H. Stockbridge. Stationary solutions and forward equations for controlled and singular martingale problems. *Electron. J. Probab.*, 6:no. 17, 52 pp. (electronic), 2001. ISSN 1083-6489. 1, 2, 5, 5

Vladimir M. Lucic and Andrew J. Heunis. On uniqueness of solutions for the stochastic differential equations of nonlinear filtering. *Ann. Appl. Probab.*, 11(1):182–209, 2001. ISSN 1050-5164. 4.7

B. L. Rozovskiĭ. A simple proof of uniqueness for Kushner and Zakai equations. In *Stochastic analysis*, pages 449–458. Academic Press, Boston, MA, 1991. 4.7

Łukasz Stettner. On invariant measures of filtering processes. In *Stochastic differential systems (Bad Honnef, 1988)*, volume 126 of *Lecture Notes in Control and Inform. Sci.*, pages 279–292. Springer, Berlin, 1989. 3.1

Daniel W. Stroock and S. R. S. Varadhan. Diffusion processes with boundary conditions. *Comm. Pure Appl. Math.*, 24:147–225, 1971. ISSN 0010-3640. 1