Electronic Journal of Probability

Vol. 3 (1998) Paper no. 9, pages 1–29.

Journal URL
http://www.math.washington.edu/~ejpecp/
Paper URL
http://www.math.washington.edu/~ejpecp/EjpVol3/paper9.abs.html

Martingale problems for conditional distributions of Markov processes

Thomas G. Kurtz [1]
Departments of Mathematics and Statistics
University of Wisconsin - Madison
480 Lincoln Drive
Madison, WI 53706-1388
kurtz@math.wisc.edu

Abstract

Let X be a Markov process with generator A and let $Y(t) = \gamma(X(t))$. The conditional distribution π_t of $X(t)$ given $\sigma(Y(s) : s \leq t)$ is characterized as a solution of a *filtered martingale problem.* As a consequence, we obtain a generator/martingale problem version of a result of Rogers and Pitman on Markov functions. Applications include uniqueness of filtering equations, exchangeability of the state distribution of vector-valued processes, verification of quasireversibility, and uniqueness for martingale problems for measure-valued processes. New results on the uniqueness of forward equations, needed in the proof of uniqueness for the filtered martingale problem are also presented.

AMS 1991 subject classifications: 60G35, 60J25, 60J35, 93E11, 60G09, 60G44
Keywords: partial observation, conditional distribution, filtering, forward equation, martingale problem, Markov process, Markov function, quasireversibility, measure-valued process

Submitted to EJP on November 21, 1996.

Final version accepted on July 6, 1998.

[1] Research supported in part by NSF grant DMS 96-26116 at the University of Wisconsin and ONR grant N00014-96-1-0262 at Purdue University.

1 Introduction

Suppose Z is a Markov process with state space E and that γ is a function mapping E into another space E_0. We are interested in characterizing the conditional distribution of Z given the information obtained by observing the process Y defined by $Y(t) = \gamma(Z(t))$. This problem is, of course, the fundamental problem of filtering theory, but it is also closely related to such issues as identification of conditions under which Y is a Markov process and verification of quasireversibility in queueing networks. We approach these questions at a general level, characterizing the process Z as the solution of a martingale problem. The fundamental result is the characterization of the desired conditional distribution as the solution of another martingale problem.

Throughout, E will be a complete, separable metric space, $B(E)$ will denote the bounded, Borel measurable functions on E, and $\mathcal{P}(E)$ the Borel probability measures on E. Let A be a mapping $A : \mathcal{D}(A) \subset B(E) \to B(E)$, and $\nu_0 \in \mathcal{P}(E)$. (Note that we do not assume that A is linear; however, usually it is or can easily be extended to be linear.) In general, we will denote the domain of an operator A by $\mathcal{D}(A)$ and its range by $\mathcal{R}(A)$.

A progressively measurable, E-valued process Z is a solution of the *martingale problem* for (A, ν_0) if $\nu_0 = PZ(0)^{-1}$ and there exists a filtration $\{\mathcal{F}_t\}$ such that

$$f(Z(t)) - \int_0^t Af(Z(s))ds$$

is an $\{\mathcal{F}_t\}$-martingale for each $f \in \mathcal{D}(A)$. A measurable $\mathcal{P}(E)$-valued function ν on $[0, \infty)$ is a solution of the *forward equation* for (A, ν_0) if

$$\nu_t f = \nu_0 f + \int_0^t \nu_s Af ds$$

for each $f \in \mathcal{D}(A)$. (For $\mu \in \mathcal{P}(E)$, we set $\mu f = \int_E f d\mu$.) Note that if Z is a solution of the martingale problem for (A, ν_0), then the one-dimensional distributions

$$\nu_t = PZ(t)^{-1} \tag{1.1}$$

give a solution of the forward equation for (A, ν_0). A critical aspect of our discussion is conditions under which the converse of this statement holds, that is, given a solution of the forward equation ν, when does there exist a solution of the martingale problem Z satisfying (1.1).

Let E_0 be a second complete, separable metric space, and let $\gamma : E \to E_0$ be Borel measurable. Let Z be a solution of the martingale problem for (A, ν_0) defined on a complete probability space $(\Omega, \mathcal{F}, P)$, and define $Y(t) = \gamma(Z(t))$. We are interested in the filtering problem of estimating Z from observations of Y. Because we don't want to place regularity assumptions on Z and γ, we formulate this problem as follows: Let $\hat{\mathcal{F}}_t^Y$ be the completion of $\sigma(\int_0^s h(Y(u))du : s \leq t, h \in B(E_0))$, and define π by $\pi_t(\Gamma) = P\{Z(t) \in \Gamma | \hat{\mathcal{F}}_t^Y\}$. π can be taken to be an $\{\hat{\mathcal{F}}_t^Y\}$-progressive (in fact, $\{\hat{\mathcal{F}}_t^Y\}$-optional), $\mathcal{P}(E)$-valued process. (See the appendix.) Note that

$$\pi_t f - \int_0^t \pi_s Af ds \tag{1.2}$$

is an $\{\hat{\mathcal{F}}_t^Y\}$-martingale for each $f \in \mathcal{D}(A)$ and that for each $t \geq 0$ and $h \in B(E_0)$,

$$\int_0^t \pi_s h \circ \gamma ds = \int_0^t h(Y(s))ds \qquad a.s. \tag{1.3}$$

Furthermore, if Y is cadlag and has no fixed points of discontinuity, that is, $P\{Y(t) = Y(t-)\} = 1$ for all $t \geq 0$, then $\hat{\mathcal{F}}_t^Y$ equals the completion of $\sigma(Y(s) : s \leq t)$ for all $t \geq 0$ and (1.3) can be replaced by

$$\pi_t h \circ \gamma = h(Y(t)) \; .$$

Our primary goal is to give conditions underwhich (1.2) and (1.3) along with the requirement that $E[\pi_0] = \nu_0$ determine the joint distribution of (Y, π). To be precise, we will say that $(\tilde{Y}, \tilde{\pi})$ is a solution of the *filtered martingale problem* for (A, ν_0, γ), if $E[\tilde{\pi}_0] = \nu_0$, for each $h \in B(E)$ and $t \geq 0$,

$$\int_0^t \tilde{\pi}_s h \circ \gamma = \int_0^t h(\tilde{Y}(s))ds \; ,$$

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

$$\tilde{\pi}_t f - \int_0^t \tilde{\pi}_s A f ds \tag{1.4}$$

is an $\{\hat{\mathcal{F}}_t^{\tilde{Y}}\}$-martingale.

The filtered martingale problem was defined and studied in Kurtz and Ocone (1988) in the special case of Z of the form $Z = (X, Y)$ and $\gamma(x, y) = y$, following a question raised to the author by Georgio Koch regarding the use of (1.2) as an approach to studying nonlinear filtering problems. The primary application of the results of that paper were to the Zakai and Kushner-Stratonovich equations for filtering of signals observed in Gaussian white noise. Bhatt, Kallianpar, and Karandikar (1995) extend the results of Kurtz and Ocone, in particular, eliminating an assumption of local compactness on the state space; Klieman, Koch and Marchetti (1990) apply the results to filtering problems with counting process observations; and Fan (1996) considers more general jump process observations. Other applications include an interesting proof of Burke's output theorem in the Kliemann, Koch, and Marchetti paper and a similar application in Donnelly and Kurtz (1996) verifying the conditional independence of a permutation-valued process from another Markov chain. The more general results given in the present paper are motivated by applications to proofs of uniqueness of martingale problems for measure-valued processes in Donnelly and Kurtz (1997). See Corollaries 3.5 and 3.7.

Throughout, $C_E[0, \infty)$ will denote the space of continuous, E-valued functions with the compact uniform topology, $D_E[0, \infty)$, the space of cadlag, E-valued functions with the Skorohod topology, and $M_E[0, \infty)$, the space of measurable, E-valued functions topologized by convergence in measure. Note that the Borel σ-algebra for each of these spaces is generated by functions of the form $x \to \int_0^\infty f(x(s), s)ds$ for $f \in C_c(E \times [0, \infty))$. Consequently, any probability distribution on $C_E[0, \infty)$ uniquely determines a probability distribution on $D_E[0, \infty)$ and any probability distribution on $D_E[0, \infty)$ uniquely determines a probability distribution on $M_E[0, \infty)$.

2 Uniqueness for the forward equation

The primary hypothesis of our main results will be uniqueness for the forward equation for (A, ν_0). There are a variety of conditions that imply this uniqueness, and if $\mathcal{D}(A)$ is separating, uniqueness for the forward equation implies uniqueness for the martingale problem. (See Ethier and Kurtz (1986), Theorem 4.4.2.) We identify A with its graph $A = \{(f, Af) : f \in \mathcal{D}(A)\}$. See Ethier and Kurtz (1986) for definitions and results on generators and semigroups.

Let $A \subset B(E) \times B(E)$ and let A_S be the linear span of A. We will say that A is *dissipative* if and only if A_S is dissipative, that is, for $(f, g) \in A_S$ and $\lambda > 0$,

$$\|\lambda f - g\| \geq \lambda \|f\|.$$

Note that a solution of the martingale problem (forward equation) for A will be a solution of the martingale problem (forward equation) for A_S.

We say that $A \subset B(E) \times B(E)$ is a *pre-generator* if A is dissipative and there are sequences of functions $\mu_n : E \to \mathcal{P}(E)$ and $\lambda_n : E \to [0, \infty)$ such that for each $(f, g) \in A$

$$g(x) = \lim_{n\to\infty} \lambda_n(x) \int_E (f(y) - f(x))\mu_n(x, dy) \tag{2.1}$$

for each $x \in E$. Note that we have not assumed that μ_n and λ_n are measurable.

If $A \subset \bar{C}(E) \times \bar{C}(E)$ ($\bar{C}(E)$ denotes the bounded continuous functions on E) and for each $x \in E$, there exists a solution ν^x of the forward equation for (A, δ_x) that is right-continuous (in the weak topology) at zero, then A is a pre-generator. In particular, if $(f, g) \in A$, then

$$\begin{aligned}
\int_0^\infty e^{-\lambda t} \nu_t^x (\lambda f - g) dt &= \int_0^\infty \lambda e^{-\lambda t} \nu_t^x f dt - \int_0^\infty \lambda e^{-\lambda t} \int_0^t \nu_s^x g ds dt \\
&= f(x)
\end{aligned}$$

which implies $\|\lambda f - g\| \geq \lambda f(x)$ and hence dissipativity, and if we take $\lambda_n = n$ and $\mu_n(x, \cdot) = \nu_{1/n}^x$,

$$n \int_E (f(y) - f(x) \nu_{1/n}^x = n(\nu_{1/n}^x f - f(x)) = n \int_0^{\frac{1}{n}} \nu_s^x g ds \to g(x).$$

(We do not need to verify that ν_t^x is a measurable function of x for either of these calculations.)

If E is locally compact and $\mathcal{D}(A) \subset \hat{C}(E)$ ($\hat{C}(E)$, the continuous functions vanishing at infinity), then the existence of λ_n and μ_n satisfying (2.1) implies A is dissipative. In particular, A_S will satisfy the positive maximum principle, that is, if $(f, g) \in A_S$ and $f(x_0) = \|f\|$, then $g(x_0) \leq 0$ which implies

$$\|\lambda f - g\| \geq \lambda f(x_0) - g(x_0) \geq \lambda f(x_0) = \lambda \|f\|.$$

If E is compact, $A \subset C(E) \times C(E)$, and A satisfies the positive maximum principle, then A is a pre-generator. If E is locally compact, $A \subset \hat{C}(E) \times \hat{C}(E)$, and A satisfies the positive maximum principle, then A can be extended to a pre-generator on E^Δ, the one-point compactification of E. See Ethier and Kurtz (1986), Theorem 4.5.4.

Suppose $A \subset \bar{C}(E) \times \bar{C}(E)$. If $\mathcal{D}(A)$ is convergence determining, then every solution of the forward equation is continuous. Of course, if for each $x \in E$ there exists a cadlag solution of the martingale problem for A, then there exists a right continuous solution of the forward equation, and hence, A is a pre-generator.

Theorem 2.1 *(Semigroup conditions.) Let $A \subset B(E) \times B(E)$ be linear and dissipative. Suppose that there exists $A' \subset A$ such that $\overline{\mathcal{R}(\lambda - A')} \supset \mathcal{D}(A')$ for some $\lambda > 0$ (where the closure of A is in the sup norm) and that $\mathcal{D}(A')$ is separating. Let $\overline{A'}$ denote the closure in the sup norm of A'. Then $T(t)f = \lim_{n\to\infty}(I - \frac{1}{n}\overline{A'})^{-[nt]}f$ defines a semigroup of linear operators on $L = \overline{\mathcal{D}(A')}$. If ν is a solution of the forward equation for A (and hence for A'), then $\nu_t f = \nu_0 T(t)f$, $f \in \mathcal{D}(A')$ and uniqueness holds for the forward equation and for the martingale problem for (A, ν_0).*

Remark 2.2 *Proposition 4.9.18 of Ethier and Kurtz (1986) gives closely related conditions based on the assumption that $\mathcal{R}(\lambda - A)$ is separating for each $\lambda > 0$, but without assuming that $\mathcal{D}(A)$ is separating.*

Proof. Existence of the semigroup follows by the Hille-Yosida theorem. Uniqueness for the martingale problem follows by Theorem 4.4.1 of Ethier and Kurtz (1986), and uniqueness for the forward equation follows by a similar argument. In particular, integration by parts gives

$$\nu_0 f = \int_0^\infty e^{-\lambda s}(\lambda \nu_s f - \nu_s A' f)ds. \tag{2.2}$$

Since if the range condition holds for some $\lambda > 0$ it holds for all $\lambda > 0$ (see Ethier and Kurtz (1986), Lemma 1.2.3), (2.2) implies for $h \in \mathcal{D}(A')$

$$\nu_0(I - n^{-1}A')^{-1}h = n\int_0^\infty e^{-ns}\nu_s h ds = \int_0^\infty e^{-s}\nu_{n^{-1}s} h ds\ , \quad h \in L,$$

and hence, iterating this identity,

$$\begin{aligned}\nu_0(I - n^{-1}A')^{-[nt]}h &= \int_0^\infty \cdots \int_0^\infty e^{-s_1 - \cdots - s_{[nt]}}\nu_{n^{-1}(s_1 + \cdots + s_{[nt]})} h ds \\ &= E[\nu_{n^{-1}S_{[nt]}} h]\ , \quad h \in L,\end{aligned}$$

where $S_{[nt]}$ is the sum of $[nt]$ independent unit exponential random variables. Letting $n \to \infty$, the law of large numbers and the continuity of $\nu_t h$ for $h \in \mathcal{D}(A')$ implies

$$\nu_0 T(t)h = \nu_t h\ .$$

□

The *bp-closure* of the set $H \subset B(E)$ is the smallest set $\bar{H}$ containing H that is closed under bounded, pointwise convergence of sequences. A set H is *bp-dense* in a set $G \subset B(E)$ if G is the *bp*-closure of H.

Theorem 2.3 *Let $A \subset B(E) \times B(E)$ be linear and dissipative. Suppose that for each $\nu_0 \in \mathcal{P}(E)$, there exists a cadlag solution of the martingale problem for (A, ν_0) and that there exists a $\lambda > 0$ for which $\mathcal{R}(\lambda - A)$ is bp-dense in $B(E)$. Then for each $\nu_0 \in \mathcal{P}(E)$, uniqueness holds for the martingale problem for (A, ν_0) and for the forward equation for (A, ν_0).*

Proof. Let X_x denote a cadlag solution of the martingale problem for (A, δ_x). Then for $h \in \mathcal{R}(\lambda - A)$, that is for $h = \lambda f - g$ for some $(f, g) \in A$,

$$f(x) = E[\int_0^\infty e^{-\lambda t} h(X_x(t))dt] \tag{2.3}$$

(cf. (4.3.20) in Ethier and Kurtz (1986)). Let $\bar{A}$ denote the *bp*-closure of A. Note that if $h_n = \lambda f_n - g_n$ for $(f_n, g_n) \in A$ and $bp - \lim_{n\to\infty} h_n = h$, then $bp - \lim_{n\to\infty} f_n = f$, where f is given by (2.3). It follows that $\mathcal{R}(\lambda - \bar{A}) = B(E)$. Since $\bar{A}$ is still dissipative, by (2.3), we must have $\mathcal{R}(\lambda - \bar{A}) = B(E)$ for all $\lambda > 0$. Since a process is a solution of the martingale problem for A if and only if it is a solution of the martingale problem for $\bar{A}$, we may as well assume $A = \bar{A}$.

If $h \in \bar{C}(E)$,

$$\lambda(\lambda - \bar{A})^{-1} h(x) = E[\int_0^\infty \lambda e^{-\lambda t} h(X_x(t))dt] \to h(x)$$

as $\lambda \to \infty$. Consequently, we see that $\mathcal{D}(\bar{A})$ is *bp*-dense in $B(E)$ and hence is separating. Consequently, the theorem follows by Theorem 2.1. □

As noted above, uniqueness for the forward equation typically implies uniqueness for the martingale problem. The converse of this assertion does not hold in general. For example, let $E = [0, 1]$,

$$\mathcal{D}(A) = \{f \in C^2[0,1] : f'(0) = f'(1) = 0, f'(\frac{1}{3}) = f'(\frac{2}{3})\},$$

and $Af(x) = \frac{1}{2} f''(x)$. Then for any $\nu_0 \in \mathcal{P}[0,1]$, the unique solution of the martingale problem for (A, ν_0) is reflecting Brownian motion on $[0, 1]$. Note, however, that $\nu_t(dx) = 3I_{[\frac{1}{3},\frac{2}{3}]}(x)dx$ is a stationary solution of the forward equation that does not correspond to a solution of the martingle problem. In particular, the only stationary distribution for the martingale problem is the uniform distribution on $[0, 1]$. (See Ethier and Kurtz (1986), Problem 4.11.4.)

We next consider conditions under which the converse does hold. We will need the following separability hypothesis.

Hypothesis 2.4 *There exists a countable subset $\{g_k\} \subset \mathcal{D}(A) \cap \bar{C}(E)$ such that the graph of A is contained in the bounded, pointwise closure of the linear span of $\{(g_k, Ag_k)\}$.*

Remark 2.5 *If $L \subset \bar{C}(E)$ is separable and $A \subset L \times L$, Hypothesis 2.4 is satisfied with the bounded, pointwise closure replaced by the (sup-norm) closure. In particular, if E is locally compact and $A \subset \hat{C}(E) \times \hat{C}(E)$, then the hypothesis is satisfied.*

If $E = \mathbb{R}^d$, $\mathcal{D}(A) = C_c^\infty(\mathbb{R}^d)$ (the infinitely differentiable functions with compact support),

$$Af(x) = \frac{1}{2} \sum_{i,j} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial_j} f(x) + \sum_i b_i(x) \frac{\partial}{\partial x_i} f(x),$$

and a_{ij} and b_i are bounded on compact sets, then the hypothesis is satisfied. (Note that $\{f, \partial_i, \partial_i\partial_j f : 1 \le i, j \le d\} \subset \hat{C}(\mathbb{R}^d)^{1+d+d^2}$.)

If Hypothesis 2.4 holds and we define $A_0 = \{(g_k, Ag_k) : k = 1, 2, \ldots\}$, then any solution of the martingale problem (forward equation) for A_0 will be a solution of the martingale problem (forward equation) for A. (See Ethier and Kurtz (1986) Proposition 4.3.1.) For some of the consequences of Hypothesis 2.4, see Ethier and Kurtz (1986), Theorem 4.6.

Theorem 2.6 *(Conditions based on the martingale problem.) Let $A \subset \bar{C}(E) \times \bar{C}(E)$ be a pre-generator and satisfy the separability Hypothesis 2.4. Suppose that $\mathcal{D}(A)$ is closed under multiplication (that is, $f_1, f_2 \in \mathcal{D}(A)$ implies $f_1 f_2 \in \mathcal{D}(A)$) and separates points. Then*

a) Each solution of the forward equation for A corresponds to a solution of the martingale problem for A.

b) If $\pi \in \mathcal{P}(E)$ satisfies $\int_E Af d\pi = 0$ for each $f \in \mathcal{D}(A)$, (that is, $\nu_t \equiv \pi$ is a stationary solution of the forward equation), then there is a stationary solution of the martingale problem for (A, π).

c) If uniqueness holds for the martingale problem for (A, ν_0), then uniqueness holds for the forward equation for (A, ν_0).

Proof. This result is essentially Theorem 4.1 of Bhatt and Karandikar (1993), or if E is locally compact and $A \subset \hat{C}(E) \times \hat{C}(E)$, Proposition 4.9.19 of Ethier and Kurtz (1986). The technical modifications of the earlier results are discussed in the appendix. □

Theorem 2.6c can be extended to any situation in which one can show that each solution of the forward equation corresponds to a solution of the martingale problem in the sense that the solution of the forward equation gives the one-dimensional distributions of the solution of the martingale problem. Uniqueness for the martingale problem then implies uniqueness for the forward equation. As noted previously, uniqueness for the forward equation for every initial distribution implies uniqueness for the martingale problem (Ethier and Kurtz (1986), Theorem 4.4.2).

A result of Kurtz and Stockbridge (1997) weakens the continuity hypothesis on the range of A in Theorem 2.6.

Theorem 2.7 *Let E and F be complete and separable, and let $A^0 \subset \bar{C}(E) \times \bar{C}(E \times F)$. Let η be a transition function η from E to F, and define*

$$A_\eta f(x) = \int_F A^0 f(x, y) \eta(x, dy)\,, \quad f \in \mathcal{D}(A^0).$$

Suppose that $\mathcal{D}(A^0)$ is closed under multiplication and separates points, that for each $y \in F$, $A_y f \equiv A^0 f(\cdot, y)$ is a pre-generator, and that A_η satisfies Hypothesis 2.4.
Then

a) Each solution of the forward equation for A_η corresponds to a solution of the martingale problem for A_η.

b) If $\pi \in \mathcal{P}(E)$ satisfies $\int_E A_\eta f d\pi = 0$ for each $f \in \mathcal{D}(A^0)$, (that is, $\nu_t \equiv \pi$ is a stationary solution of the forward equation), then there is a stationary solution of the martingale problem for (A_η, π).

c) If uniqueness holds for the martingale problem for (A_η, ν_0), then uniqueness holds for the forward equation for (A_η, ν_0).

Remark 2.8 *a) Note that we are not assuming that $A_\eta f$ is continuous. In particular, any diffusion operator with bounded coefficients can be represented in this form. (With a little more care, the boundedness assumption on the coefficients can be relaxed.)*

b) Under the above conditions existence for the forward equation implies existence for the martingale problem and uniqueness for the forward equation for every initial distribution implies uniqueness for the martingale problem.

Proof. The result extends Theorem 3.1 of Kurtz and Stockbridge (1997) to the setting of Bhatt and Karandikar (1993) eliminating the assumption of local compactness. See also Bhatt and Borkar (1996). Technical details are discussed in the appendix. □

In the development that follows, we will need to supplement the process Z with additional components obtained by solving differential equations of the form

$$\dot{U}(t) = -aU(t) + b \circ \gamma(Z(t)) \,,$$

where $b \in \bar{C}(E_0)$. If we assume that $a \geq 1$, $0 \leq b \leq 1$, and $0 \leq U(0) \leq 1$, then $0 \leq U(t) \leq 1$, and we can take the state space of (Z, U) to be $E \times [0,1]$. If Z is a solution of the martingale problem for A, then (Z, U) is a solution of the martingale problem for $\hat{A}$ defined as follows: let

$$\mathcal{D}(\hat{A}) = \{fg : f \in \mathcal{D}(A), g \in C^1[0,1]\}$$

and set

$$\hat{A}(fg)(z,u) = g(u)Af(z) + f(z)(-au + b \circ \gamma(z))g'(u) \,.$$

The following theorem will be needed.

Theorem 2.9 *Let $\mu_0 \in \mathcal{P}(E \times [0,1])$ satisfy $\mu_0(\cdot \times [0,1]) = \nu_0$.*

a) If uniqueness holds for the martingale problem for (A, ν_0), then uniqueness holds for the martingale problem for $(\hat{A}, \mu_0)$.

b) If A satisfies the conditions of Theorem 2.3, then the linear span of $\hat{A}$ satisfies the conditions of Theorem 2.3 and existence and uniqueness hold for the forward equation for $\hat{A}$ and for the martingale problem for $\hat{A}$.

c) If A satisfies the conditions of Theorem 2.6 and γ is continuous, then $\hat{A}$ satisfies the conditions of Theorem 2.6.

d) If A satisfies the conditions of Theorem 2.6 or if $A = A_\eta$ where

$$A_\eta f(x) = \int_F A^0 f(x,y)\eta(x,dy)$$

and A^0 and η satisfy the conditions of Theorem 2.7, then (without assuming continuity of γ) there exists $\tilde{A}^0$ and $\tilde{\eta}$ satisfying the conditions of Theorem 2.7 such that $\hat{A}f(x,u) = \int_{\tilde{F}} \tilde{A}^0 f(x,u,y)\eta(x,u,dy)$. In particular, if uniqueness holds for the martingale problem for (A, ν_0) and $\mu_0 \in \mathcal{P}(E \times [0,1])$ has E-marginal ν_0, then uniqueness holds for the forward equation for $(\hat{A}, \mu_0)$.

Proof. Let (Z,U) be a solution of the martingale problem for $(\hat{A},\mu_0)$. By the same argument used in the proof of Ethier and Kurtz (1986), Theorem 4.3.6, U has a cadlag modification, which we will continue to denote by U. (Note that the assertions of the theorem only depend on the finite dimensional distributions of (Z,U).)

For $g\in C^1[0,1]$

$$\begin{aligned}&E[(g(U(t+r))-g(U(t)))^2]\\&\quad=E[\int_t^{t+r}(-aU(s)+b\circ\gamma(Z(s)))(2g(U(s))g'(U(s))-2g(U(t))g'(U(s)))ds].\end{aligned}$$

It follows that for each $t\geq 0$ and partitions $0=t_0<t_1<\cdots$ with $t_i\to\infty$

$$\lim_{\max|t_{i+1}-t_i|\to 0}E\left[\sum(g(X(t_{i+1}\wedge t))-g(X(t_i\wedge t)))^2\right]=0,$$

and hence U is continuous. A similar calculation shows that

$$g(U(t))=g(U(0))+\int_0^t(-aU(s)+b\circ\gamma(Z(s)))g'(U(s))ds,$$

which in turn implies

$$U(t)=U(0)e^{-at}+\int_0^t e^{-a(t-s)}b\circ\gamma(Z(s))ds\;. \tag{2.4}$$

This identity and Fubini's theorem imply that

$$E[h(Z(t_0))\prod_{i=1}^k U(t_i)^{m_i}]$$

is determined by the finite dimensional distributions of Z, and hence, the finite dimensional distributions of (Z,U) are determined by the finite dimensional distributions of Z and part a) follows.

Now consider part b). As in the proof of Theorem 2.3, we may as well assume that $\mathcal{R}(\lambda-A)=B(E)$. This condition implies that A is the full generator for the semigroup on $B(E)$ determined by

$$T(t)f(x)=E[f(X_x(t))]$$

where X_x is a solution of the martingale problem for (A,δ_x). (See Ethier and Kurtz (1986), Section 1.5.) Then $(f,g)\in A$ if and only if

$$T(t)f=f+\int_0^t T(s)gds\;. \tag{2.5}$$

Existence of cadlag solutions of the martingale problem for $\hat{A}$ follows from existence for A and (2.4). Let $\hat{A}_S$ denote the linear span of $\hat{A}$. To complete the proof, we need to verify that $\mathcal{R}(\lambda-\hat{A}_S)$ is *bp*-dense in $B(E\times[0,1])$.

Let f' denote the partial derivative of f with respect to u. Let $\hat{A}_r$ be the collection of $(f,g)\in B(E\times[0,1])\times B(E\times[0,1])$ such that $f(x,\cdot)\in C^1[0,1]$ and $g(x,\cdot)\in C[0,1]$ for each x, $f'\in B(E\times[0,1])$, and

$$(f(\cdot,u),g(\cdot,u)-(-au+b\circ\gamma(\cdot))f'(\cdot,u))\in A$$

for each $u \in [0,1]$. We claim that $\hat{A}_r$ is in the bp-closure of $\hat{A}_S$. To see that this claim holds, let

$$\tilde{g}(x,u) = g(x,u) - (-au + b\circ\gamma(x))f'(x,u) \ ,$$

approximate f by the Bernstein polynomial

$$f_n(x,u) = \sum_{k=0}^{n} f(x,\frac{k}{n})\binom{n}{k}u^k(1-u)^{n-k},$$

and set

$$g_n(x,u) = \sum_{k=0}^{n} \tilde{g}(x,\frac{k}{n})\binom{n}{k}u^k(1-u)^{n-k} + (-au + b\circ\gamma(x))f_n'(x,u).$$

Then $(f_n, g_n) \in \hat{A}_S$ and $bp\text{-}\lim_{n\to\infty}(f_n, g_n) = (f,g)$.

Now let $h \in B(E\times[0,1])$ satisfy $h(x,\cdot) \in C^1[0,1]$ and $h' \in B(E\times[0,1])$, and set

$$f(x,u) = E[\int_0^\infty e^{-\lambda t}h(X_x(t), ue^{-at} + \int_0^t e^{-a(t-s)}b\circ\gamma(X_x(s))ds)dt].$$

We now claim that $(f, \lambda f - h) \in \hat{A}_r$. It is easy to check that f is differentiable with respect to u. Let

$$\tilde{g}(x,u) = \lambda f(x,u) - h(x,u) - (-au + b\circ\gamma(x))f'(x,u) \ .$$

We need to show that for each $u \in [0,1]$,

$$(f(\cdot,u), \tilde{g}(\cdot,u)) \in A \ .$$

By (2.5) we must verify that

$$E[f(X_x(r),u)] = f(x,u) + E[\int_0^r \tilde{g}(X_x(s),u)ds] \ . \tag{2.6}$$

By the Markov property

$$\begin{aligned} E[f(X_x(r),u)] &= E[\int_0^\infty e^{-\lambda t}h(X_x(r+t), ue^{-at} + \int_0^t e^{-a(t-s)}b\circ\gamma(X_x(r+s))ds)dt] \\ &= E[e^{\lambda r}\int_r^\infty e^{-\lambda t}h(X_x(t), ue^{-a(t-r)} + \int_r^t e^{-a(t-s)}b\circ\gamma(X_x(s))ds)dt] \ , \end{aligned}$$

and differentiating by r verifies (2.6).

Finally, the collection of h satisfying the above properties is bp-dense in $B(E\times[0,1])$, so $\mathcal{R}(\lambda - \hat{A}_r)$ (and hence $\mathcal{R}(\lambda - \hat{A}_S)$) is bp-dense in $B(E\times[0,1])$.

To verify part c), let $\mathcal{D}(A^0) = \{fg : f \in \mathcal{D}(A), g \in C^1[0,1]\}$ and note that $\mathcal{D}(A^0)$ is closed under multiplication and separates points in $E\times[0,1]$. Define

$$A^0 fg(x,u,y) = g(u)Af(x) + f(x)(-au + b(y))g'(u)$$

and

$$\eta(x,u,dy) = \delta_{\gamma(x)}(dy) \ .$$

Then $\hat{A}f(x,u) = A_\eta f(x,u)$ and the conditions of Theorem 2.7 are satisfied.

The proof of part d) is similar. Assume $A = A_\eta$ where $A_\eta f(x) = \int_F A^0 f(x,y)\eta(x,dy)$ where A^0 and η satisfy the conditions of Theorem 2.7. Define $\mathcal{D}(\tilde{A}^0) = \{fg : f \in \mathcal{D}(A^0), g \in C^1[0,1]\}$,

$$\tilde{A}^0 fg(x,u,y_1,y_2) = g(u)A^0 f(x,y_1) + f(x)(-au + b(y_2))g'(u),$$

and

$$\tilde{\eta}(x,u,dy) = \eta(x,dy_1)\delta_{\gamma(x)}(dy_2)\ .$$

Then $\hat{A}f(x,u) = \tilde{A}_{\tilde{\eta}}f(x,u)$ and the conditions of Theorem 2.7 are satisfied. □

3 Uniqueness for the filtered martingale problem

Let $\{b_k\} \subset \bar{C}(E_0)$ satisfy $0 \le b_k \le 1$, and suppose that the span of $\{b_k\}$ is bounded, pointwise dense in $B(E_0)$. (Existence of $\{b_k\}$ follows from the separability of E_0.) Let $a_1, a_2, \ldots$ be an ordering of the rationals with $a_i \ge 1$. Let U_{ki} satisfy

$$U_{ki}(t) = -a_i \int_0^t U_{ki}(s)ds + \int_0^t b_k \circ \gamma(Z(s))ds\ . \tag{3.1}$$

Then

$$U_{ki}(t) = \int_0^t e^{-a_i(t-s)} b_k \circ \gamma(Z(s))ds = \int_0^t e^{-a_i(t-s)} b_k(Y(s))ds. \tag{3.2}$$

Let $U(t) = ((U_{ki}(t)\ ))$. By the properties of Laplace transforms, it follows that the completion of $\sigma(U(t))$ equals $\hat{\mathcal{F}}_t^Y$. Let $\hat{E} = E \times [0,1]^\infty$ denote the state space of $\hat{Z} = (Z, U_{ki} : k, i \ge 1)$. Let $\mathcal{D}(\hat{A})$ be the collection of functions on $\hat{E}$ given by

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

and define $\hat{A}$ by

$$\begin{aligned} \hat{A}f \prod_{k,i=1}^m g_{ki}(z,u) &= \left(\prod_{k,i=1}^m g_{ki}(u_{ki})\right) Af(z) \\ &\quad + \sum_{l,j=1}^m f(z) \left(\prod_{k,i=1:(k,i)\neq(l,j)}^m g_{k,i}(u_{ki})\right)(-a_j u_{lj} + b_l \circ \gamma(z))g'_{lj}(u_{lj})\ . \end{aligned}$$

The proof of the following lemma is essentially the same as the proof of Theorem 2.9.

Lemma 3.1 *Let A and $\hat{A}$ be as above.*

a) If uniqueness holds for the martingale problem for A, then uniqueness holds for the martingale problem for $\hat{A}$.

b) If A satisfies the conditions of Theorem 2.3, 2.6 or 2.7, then $\hat{A}$ will satisfy the conditions of one of these theorems. In particular, if A satisfies the conditions of Theorem 2.3, 2.6, or 2.7, then each solution of the forward equation for $\hat{A}$ corresponds to a solution of the martingale problem for $\hat{A}$, and hence, uniqueness for the martingale problem implies uniqueness for the forward equation.

Theorem 3.2 *Let $A \subset B(E) \times B(E)$, $\nu_0 \in \mathcal{P}(E)$, and $\gamma : E \to E_0$ be Borel measurable. Suppose that each solution of the forward equation for $(\hat{A}, \nu_0 \times \delta_0)$ corresponds to a solution of the martingale problem. (This condition will hold if A satisfies the conditions of Theorem 2.3, 2.6 or 2.7.) Let $(\tilde{\pi}, \tilde{Y})$ be a solution of the filtered martingale problem for (A, ν_0, γ). The following hold:*

a) There exists a solution Z of the martingale problem for (A, ν_0), such that $\tilde{Y}$ has the same distribution on $M_{E_0}[0, \infty)$ as $Y = \gamma \circ Z$.

b) For each $t \geq 0$, there exists a Borel measurable mapping $H_t : M_{E_0}[0, \infty) \to \mathcal{P}(E)$ such that $\pi_t = H_t(Y)$ is the conditional distribution of $Z(t)$ given $\hat{\mathcal{F}}_t^Y$, and $\tilde{\pi}_t = H_t(\tilde{Y})$ a.s. In particular, $\tilde{\pi}$ has the same finite dimensional distributions as π.

c) If Y and $\tilde{Y}$ have sample paths in $D_{E_0}[0, \infty)$, then Y and $\tilde{Y}$ have the same distribution on $D_{E_0}[0, \infty)$ and the H_t are Borel measurable mappings from $D_{E_0}[0, \infty)$ to $\mathcal{P}(E)$.

d) If uniqueness holds for the martingale problem for (A, ν_0), then uniqueness holds for the filtered martingale problem for (A, ν_0, γ) in the sense that if (π, Y) and $(\tilde{\pi}, \tilde{Y})$ are solutions, then for each $0 \leq t_1 < \cdots < t_m$, $(\pi_{t_1}, \ldots, \pi_{t_m}, Y)$ and $(\tilde{\pi}_{t_1}, \ldots, \tilde{\pi}_{t_m}, \tilde{Y})$ have the same distribution on $\mathcal{P}(E)^m \times M_{E_0}[0, \infty)$.

Proof. As in (3.2), let

$$\tilde{U}_{ki}(t) = -a_i \int_0^t \tilde{U}_{ki}(s)ds + \int_0^t b_k(\tilde{Y}(s))ds = \int_0^t e^{-a_i(t-s)} b_k(\tilde{Y}(s))ds.$$

Define $\tilde{\nu}_t \in \mathcal{P}(E \times [0,1]^\infty)$ by

$$\tilde{\nu}_t h = E[\int_E h(z, \tilde{U}(t))\tilde{\pi}_t(dz)] \ .$$

Note that for a.e. t,

$$f_1(\tilde{Y}(t))\tilde{\pi}_t f_2 = \tilde{\pi}_t(f_2 f_1 \circ \gamma) \ . \tag{3.3}$$

For $fg \in \mathcal{D}(\hat{A})$,

$$\begin{aligned}
\tilde{\nu}_t fg &= E[\tilde{\pi}_t fg(\tilde{U}(t))] \\
&= E[\tilde{\pi}_0 fg(\tilde{U}(0))] \\
&\qquad + E[\textstyle\int_0^t \left(g(\tilde{U}(s))\tilde{\pi}_s Af + \tilde{\pi}_s f \sum(-a_i \tilde{U}_{ki}(s) + b_k(\tilde{Y}(s)))\partial_{ki} g(\tilde{U}(s))\right) ds] \\
&= \tilde{\nu}_0 fg + \textstyle\int_0^t \tilde{\nu}_s \hat{A} fg ds
\end{aligned}$$

where $\tilde{\nu}_0 = \nu_0 \times \delta_0$ and the last equality follows from the definition of $\tilde{\nu}$ and (3.3). Consequently, $\tilde{\nu}$ is a solution of the forward equation for $(\hat{A}, \nu_0 \times \delta_0)$, and by assumption, there exists a solution (Z, U) of the martingale problem for $(\hat{A}, \nu_0 \times \delta_0)$, such that

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

It follows that for each t, $U(t)$ and $\tilde{U}(t)$ have the same distribution. Since $\hat{\mathcal{F}}_t^Y$ equals the completion of $\sigma(U(t))$ and $\hat{\mathcal{F}}_t^{\tilde{Y}}$ equals the completion of $\sigma(\tilde{U}(t))$, there exist mappings $G_t, \tilde{G}_t : [0,1]^\infty \to \mathcal{P}(E)$ such that $\pi_t = G_t(U(t))$ a.s. and $\tilde{\pi}_t = \tilde{G}_t(\tilde{U}(t))$ a.s. By (3.4)

$$E[G_t(U(t))fg(U(t))] = E[\tilde{G}_t(\tilde{U}(t))fg(\tilde{U}(t))] = E[\tilde{G}_t(U(t))fg(U(t))] \tag{3.5}$$

for all $g \in B([0,1]^\infty)$, where the last equality follows from the fact that $U(t)$ and $\tilde{U}(t)$ have the same distribution. Applying (3.5) with $g = G_t(\cdot)f$ and with $g = \tilde{G}_t(\cdot)f$, we have

$$E[G_t(U(t))f\tilde{G}_t(U(t))f] = E[\left(\tilde{G}_t(U(t))f\right)^2] = E[(G_t(U(t))f))^2]$$

and it follows that

$$E[\left(G_t(U(t))f - \tilde{G}_t(U(t))f\right)^2] = 0\ .$$

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

Since $U(t)$ $(\tilde{U}(t))$ determines $U(s)$ $(\tilde{U}(s))$ for $s < t$, U and $\tilde{U}$ have the same distribution on $C_{[0,1]^\infty}[0,\infty)$. Consequently, (π, U) and $(\tilde{\pi}, \tilde{U})$ have the same finite dimensional distributions, and (U, Y) and $(\tilde{U}, \tilde{Y})$ have the same distribution on $C_{[0,1]^\infty}[0,\infty) \times M_{E_0}[0,\infty)$. The mapping $F : y \in M_{E_0}[0,\infty) \to u \in C_{[0,1]^\infty}[0,\infty)$ determined by

$$u_{ki}(t) = \int_0^t e^{-a_i(t-s)} b_k(y(s))ds$$

is Borel measurable, and $U(t) = F(Y,t)$ and $\tilde{U}(t) = F(\tilde{Y},t)$. Consequently, $H_t(y) \equiv G_t(F(y,t))$ is Borel measurable, $\pi_t = H_t(Y)$ a.s. and $\tilde{\pi}_t = H_t(\tilde{Y})$ a.s.

Finally, if uniqueness holds for the martingale problem for (A, ν_0), then the distribution of (Z, Y, π) is uniquely determined and uniqueness holds for the filtered martingale problem for (A, ν_0, γ). □

We say that the filtered martingale problem for (A, γ) is *well-posed* if for each $\nu_0 \in \mathcal{P}(E)$ there exists a unique (in the sense of Theorem 3.2(d)) solution of the filtered martingale problem for (A, ν_0, γ).

Corollary 3.3 *Let A satisfy the conditions of Theorem 2.3, and let $\gamma : E \to E_0$ be Borel measurable. Then the filtered martingale problem for A is well-posed.*

Proof. For each $\nu_0 \in \mathcal{P}(E)$, existence holds for the martingale problem for (A, ν_0) by assumption, so existence holds for the filtered martingale problem for (A, ν_0, γ). Theorem 2.3 and Theorem 2.9 imply A satisfies the conditions of Theorem 3.2 which establishes the corollary. □

Corollary 3.4 *Let $A \subset B(E) \times B(E)$, let $\gamma : E \to E_0$ be Borel measurable. Suppose A satisfies the conditions of Theorem 2.6 or $A = A_\eta$ where*

$$A_\eta f(x) = \int_F A^0 f(x,y)\eta(x,dy)$$

and A^0 and η satisfy the conditions of Theorem 2.7. For each solution $(\tilde{\pi}, \tilde{Y})$ of the filtered martingale problem for (A, γ), there exists a solution Z of the martingale problem for A such that if $Y = \gamma \circ Z$ and π_t is the conditional distribution of $Z(t)$ given $\hat{\mathcal{F}}_t^Y$, then for each $0 \leq t_1 < \cdots < t_m$, $(\tilde{\pi}_{t_1}, \ldots, \tilde{\pi}_{t_m}, \tilde{Y})$ and $(\pi_{t_1}, \ldots, \pi_{t_m}, Y)$ have the same distribution on $\mathcal{P}(E)^m \times M_{E_0}[0, \infty)$. If uniqueness holds for the martingale problem for (A, ν_0), then uniqueness holds for the filtered martingale problem for (A, ν_0, γ).

Proof. Lemma 3.1 implies the conditions of Theorem 3.2 which establishes the corollary.□

The following corollaries address the question of when a function of a Markov process is Markov. For earlier results on this question, see Cameron (1973) and Rosenblatt (1966), Rogers and Pitman (1981), Kelly (1982), and Glover (1991). The results given here have application in proving uniqueness for martingale problems, in particular, for measure-valued processes.

Corollary 3.5 *Let $A \subset B(E) \times B(E)$, and let $\gamma : E \to E_0$ be Borel measurable. Let α be a transition function from E_0 into E ($y \in E_0 \to \alpha(y, \cdot) \in \mathcal{P}(E)$ is Borel measurable) satisfying $\int h \circ \gamma(z) \alpha(y, dz) = h(y)$, $h \in B(E_0)$, that is, $\alpha(y, \gamma^{-1}(y)) = 1$. Define*

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

Let $\mu_0 \in \mathcal{P}(E_0)$, and define $\nu_0 = \int \alpha(y, \cdot)\mu_0(dy)$. Suppose that each solution of the forward equation for $(\hat{A}, \nu_0 \times \delta_0)$ corresponds to a solution of the martingale problem. If $\tilde{Y}$ is a solution of the martingale problem for (C, μ_0), then there exists a solution Z of the martingale problem for (A, ν_0) such that $\tilde{Y}$ has the same distribution on $M_{E_0}[0, \infty)$ as $Y = \gamma \circ Z$ and for $\Gamma \in \mathcal{B}(E)$,

$$P\{Z(t) \in \Gamma | \hat{\mathcal{F}}_t^Y\} = \alpha(Y(t), \Gamma). \tag{3.6}$$

a) If, in addition, uniqueness holds for the martingale problem for (A, ν_0), then uniqueness holds for the $M_{E_0}[0, \infty)$-martingale problem for (C, μ_0). If $\tilde{Y}$ has sample paths in $D_{E_0}[0, \infty)$, then uniqueness holds for the $D_{E_0}[0, \infty)$-martingale problem for (C, μ_0).

b) If uniqueness holds for the martingale problem for (A, ν_0) then Y is a Markov process.

Remark 3.6 *Part b) is essentially a generator/martingale problem version of Theorem 2 of Rogers and Pitman (1981). Let $P(t, z, \cdot)$ be a transitition function corresponding to A. They define*

$$Q(t, y, \Gamma) = \int_E P(t, z, \gamma^{-1}(\Gamma))\alpha(y, dz), \quad \Gamma \in \mathcal{B}(E_0)$$

and assume

$$\int_E P(t, z, \Gamma)\alpha(y, dz) = \int_{E_0} \alpha(w, \Gamma)Q(t, y, dw), \quad \Gamma \in \mathcal{B}(E).$$

In particular, (3.6) is just (1) of Rogers and Pitman (1981).

Proof. Define $\tilde{\pi}_t = \alpha(\tilde{Y}(t), \cdot)$. Then $(\tilde{\pi}, \tilde{Y})$ is a solution of the filtered martingale problem for (A, ν_0, γ). Except for part (b), the theorem then follows from Theorem 3.2.

The assumptions of part (b) imply Z is a Markov process (see Appendix A.4). By (3.6), $P\{Z(t) \in \Gamma | \mathcal{F}_t^Y\} = P\{Z(t) \in \Gamma | Y(t)\}$, and since Z is Markov,

$$\begin{aligned} E[f(Y(t+s))|\mathcal{F}_t^Y] &= E[E[f(\gamma(Z(t+s))|\mathcal{F}_t^Z]|\mathcal{F}_t^Y] \\ &= E[E[f(\gamma(Z(t+s))|Z(t)]|\mathcal{F}_t^Y] \\ &= E[E[f(\gamma(Z(t+s))|Z(t)]|Y(t)] \\ &= E[f(Y(t+s))|Y(t)] \end{aligned}$$

which is the Markov property for Y $\square$

In the next corollary we consider martingale problems with "side conditions", that is, in addition to requiring

$$f(Z(t)) - \int_0^t Af(Z(s))ds$$

to be a martingale for all $f \in \mathcal{D}(A)$ we require

$$E[h(Z(t)] = 0 \,, \quad t \geq 0, \quad h \in \mathcal{H},$$

for a specified collection $\mathcal{H} \subset B(E)$. We will refer to this problem as the *restricted martingale problem* for $(A, \mathcal{H}, \nu_0)$ (cf. Dawson (1993), Section 5). Of course, ν_0 must satisfy $\int h d\nu_0 = 0$, $h \in \mathcal{H}$. We will denote the collection of probability measures satisfying this condition $\mathcal{P}_{\mathcal{H}}(E)$. The *restricted forward equation* has the obvious definition. The motivation for introducing this notion is a family of problems in infinite product spaces in which we want the coordinate random variables to be exchangeable. (See Dawson (1993), Donnelly and Kurtz (1996, 1997).)

Note that if A satisfies the conditions of Theorem 2.6 or 2.7, then $\hat{A}$ satisfies the conditions of Theorem 2.7 and, consequently, if there exists a solution $\{\mu_t\}$ of the restricted forward equation for $(\hat{A}, \mathcal{H})$, then there exists a solution (Z, U) of the martingale problem for $\hat{A}$ that satisfies $E[h(Z(t), U(t))] = \mu_t h$ for all $h \in B(E \times [0,1]^\infty)$. It follows that (Z, U) is a solution of the restricted martingale problem for $(\hat{A}, \mathcal{H})$, and if uniqueness holds for the restricted martingale problem for $(\hat{A}, \mathcal{H}, \mu_0)$, then uniqueness holds for the restricted forward equation.

Corollary 3.7 *Let $\mathcal{H} \subset B(E)$, and let $\gamma : E \to E_0$ be Borel measurable. Let α be a transition function from E_0 into E satisfying $\alpha(y, \cdot) \in \mathcal{P}_{\mathcal{H}}(E)$, $y \in E_0$, and $\int h \circ \gamma(z)\alpha(y, dz) = h(y)$, $h \in B(E_0)$, that is $\alpha(y, \gamma^{-1}(y)) = 1$. Define*

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

Let $\mu_0 \in \mathcal{P}(E_0)$, and define $\nu_0 = \int \alpha(y, \cdot)\mu_0(dy)$. (Note that $\nu_0 \in \mathcal{P}_{\mathcal{H}}(E)$.) Suppose that each solution of the restricted forward equation for $(\hat{A}, \nu_0 \times \delta_0)$ corresponds to a solution of the restricted martingale problem. If $\tilde{Y}$ is a solution of the martingale problem for (C, μ_0), then there exists a solution Z of the restricted martingale problem for $(A, \mathcal{H}, \nu_0)$ such that $\tilde{Y}$ has the same distribution on $M_{E_0}[0, \infty)$ as $Y = \gamma \circ Z$ and for $\Gamma \in \mathcal{B}(E)$,

$$P\{Z(t) \in \Gamma | \mathcal{F}_t^Y\} = \alpha(Y(t), \Gamma) \tag{3.7}$$

a) If, in addition, uniqueness holds for the restricted martingale problem for $(A, \mathcal{H}, \nu_0)$, then uniqueness holds for the $M_{E_0}[0,\infty)$-martingale problem for (C, μ_0). If $\tilde{Y}$ has sample paths in $D_{E_0}[0,\infty)$, then uniqueness holds for the $D_{E_0}[0,\infty)$-martingale problem for (C, μ_0).

b) If uniqueness holds for the restricted martingale problem for $(A, \mathcal{H}, \nu_0)$ and there exists a solution with sample paths in $D_E[0,\infty)$ or if uniqueness holds for the restricted martingale problem for $(A, \mathcal{H}, \mu)$ for each $\mu \in \mathcal{P}_{\mathcal{H}}(E)$, then Y is a Markov process.

Proof. Let $\tilde{\pi}_t = \alpha(\tilde{Y}(t), \cdot)$, and let $\tilde{U}$ be defined as in the proof of Theorem 3.2. Then $(\tilde{\pi}, \tilde{Y}, \tilde{U})$ is a solution of the filtered martingale problem for $(\hat{A}, \nu_0 \times \delta_0, \gamma)$ and $\hat{\nu}_t = E[\tilde{\pi}_t \times \delta_{\tilde{U}(t)}]$ defines a solution of the restricted forward equation for $(\hat{A}, \mathcal{H}, \nu_0 \times \delta_0)$. By the hypotheses on A and $\hat{A}$, there exists a solution (Z, U) of the martingale problem for $(\hat{A}, \nu_0 \times \delta_0)$ satisfying $E[h(Z(t), U(t))] = \int h d\hat{\nu}_t$ (so it is also a solution of the restricted martingale problem), and as in the proof of Corollary 3.5, $\tilde{Y}$ has the same distribution on $M_{E_0}[0,\infty)$ as $\gamma \circ Z$. The proof of Part b) is the same as in Corollary 3.5. □

4 Applications.

The original motivation for studying the filtered martingale problem comes from classical filtering theory, and the uniqueness theorem has been used to prove uniqueness for a variety of filtering equations. (See Kurtz and Ocone (1988), Kliemann, Koch and Marchetti (1990), Bhatt, Kallianpur, and Karandikar (1995), and Fan (1996).) Theorem 3.2 and Corollary 3.4 can be used to improve slightly on these results. For example, the continuity assumption on h in Theorems 4.1, 4.2, 4.5, and 4.6 of Kurtz and Ocone (1988) can be dropped.

Corollaries 3.5 and 3.7, however, provide the motivation for the present work, and the following examples illustrate their application.

4.1 Exchangeability.

Let A^n be the generator of a Markov process with state space E^n. Let $\mathcal{P}_n(E) = \{n^{-1}\sum_{k=1}^n \delta_{x_k} : x \in E^n\}$ and $\gamma_n(x) = n^{-1}\sum_{k=1}^n \delta_{x_k}$. Define

$$\alpha_n(n^{-1}\sum_{k=1}^n \delta_{x_k}, dz) = \frac{1}{n!}\sum_{\sigma \in \Sigma_n} \delta_{(x_{\sigma_1},\ldots,x_{\sigma_n})}(dz),$$

where Σ_n is the collection of permutations of $(1,\ldots,n)$, and

$$C^n = \{(\alpha_n f, \alpha_n g) : (f,g) \in A^n\} \subset B(\mathcal{P}_n(E)) \times B(\mathcal{P}_n(E))\}. \tag{4.1}$$

Theorem 4.1 *Suppose A^n satisfies the conditions of Theorem 2.6 or Theorem 2.7. For $\mu_0 \in \mathcal{P}(\mathcal{P}_n(E))$, define*

$$\nu_0(dx) = \int_{E_0} \alpha_n(z, dx)\mu_0(dz) \in \mathcal{P}(E^n).$$

If there exists a solution of the martingale problem for (C^n, μ_0), then there exists a solution X of the martingale problem for (A^n, ν_0). If the solution of the martingale problem for (A^n, ν_0) is unique, then the solution for (C^n, μ_0) is unique and X satisfies

$$E[h(X_1(t), \ldots, X_n(t))|\mathcal{F}_t^{\gamma(X)}] = \frac{1}{n!} \sum_{\sigma \in \Sigma_n} h(X_{\sigma_1}(t), \ldots, X_{\sigma_n}(t)), . \tag{4.2}$$

In particular, (4.2) implies that for each $t \geq 0$, $(X_1(t), \ldots, X_n(t))$ is exchangeable.

Proof. The result is an immediate application of Corollary 3.5. □

We illustrate Theorem 4.1 with the following simple application. Let B be the generator of a Markov process with state space E, let $\mathcal{D}(A_0^n) = \{\prod_{i=1}^n f_i(x_i) : f_i \in \mathcal{D}(B)\}$, and for $f \in \mathcal{D}(A_0^n)$, define

$$B_i f(x) = B f_i(x_i) \prod_{j \neq i} f_j(x_j)$$

and

$$A_0^n f(x) = \sum_{i=1}^n B_i f(x) + \frac{1}{2}\lambda \sum_{1 \leq i \neq j \leq n} (f(\eta_{ij}(x)) - f(x))$$

where $\eta_{ij}(x)$ is the element of E^n obtained by replacing x_j by x_i in $x = (x_1, \ldots, x_n)$. A_0^n models a system of n particles that move independently in E according to the Markov process with generator B, live independent, exponentially distributed lifetimes, and at death are instantaneously replaced by a copy of one of the remaining $n-1$ particles, selected at random. By symmetry, if $(X_1, \ldots, X_n)$ is a solution of the martingale problem for A_0^n, then $(X_{\sigma_1}, \ldots, X_{\sigma_n})$ is a solution, and if we define C^n as in (4.1), it is easy to check that $\gamma_n(X)$ is a solution of the martingale problem for C^n.

The generator

$$A^n f(x) = \sum_{i=1}^n B_i f(x) + \lambda \sum_{1 \leq i < j \leq n} (f(\eta_{ij}(x)) - f(x)) \tag{4.3}$$

is not symmetric; however,

$$C^n = \{(\alpha_n f, \alpha_n g) : (f, g) \in A_0^n\} = \{(\alpha_n f, \alpha_n g) : (f, g) \in A^n\}. \tag{4.4}$$

If the martingale problem for B is well-posed, then the martingale problems for A_0^n and A^n will be well-posed. Theorem 4.1 implies that if $\tilde{X}$ is a solution of the martingale problem for A^n and $\tilde{X}(0) = (\tilde{X}_1(0), \ldots, \tilde{X}_n(0))$ is exchangeable, then for each $t \geq 0$ $\tilde{X}(t)$ is exchangeable and

$$E[h(\tilde{X}_1(t), \ldots, \tilde{X}_n(t))|\mathcal{F}_t^{\gamma(\tilde{X})}] = \frac{1}{n!} \sum_{\sigma} h(\tilde{X}_{\sigma_1}(t), \ldots, \tilde{X}_{\sigma_n}(t)).$$

In addition, if X is a solution of the martingale problem for A_0^n, and $X(0)$ has the same exchangeable distribution as $\tilde{X}(0)$, then $\gamma_n(X)$ and $\gamma_n(\tilde{X})$ have the same distribution. See Donnelly and Kurtz (1996, 1997) for further discussion and motivation for models of this type.

4.2 Uniqueness for measure-valued processes.

As above, let B be the generator of a Markov process with state space E. Let $\mathcal{D}(A) \subset B(E^\infty)$ be given by $\mathcal{D}(A) = \{\prod_{i=1}^n f_i(x_i) : f_i \in \mathcal{D}(B), n = 1, 2, \ldots\}$, and define

$$Af(x) = \sum_{i=1}^\infty B_i f(x) + \lambda \sum_{1 \le i < j < \infty} (f(\eta_{ij}(x)) - f(x)). \tag{4.5}$$

Note that since $f \in \mathcal{D}(A)$ depends on only finitely many components of $x \in E^\infty$, the sums in (4.5) are, in fact, finite. In addition, if $X = (X_1, X_2, \ldots)$ is a solution of the martingale problem for A, then $X^n = (X_1, \ldots, X_n)$ is a solution of the martingale problem for A^n given by (4.3). Consequently, by the above results on exchangeability, if $X(0) = (X_1(0), X_2(0), \ldots)$ is exchangeable, then for each $t \ge 0$, $X(t) = (X_1(t), X_2(t), \ldots)$ is exchangeable. (The (finite) exchangeability of $(X_1(t), \ldots, X_n(t))$ for each n implies the (infinite) exchangeability of $X(t)$.) Let

$$Z^n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}.$$

If $X(0)$ is exchangeable (so $X(t)$ is exchangeable of all $t \ge 0$), then by deFinetti's theorem

$$Z(t) = \lim_{n\to\infty} Z^n(t) \qquad a.s.$$

where convergence is in the weak topology on $\mathcal{P}(E)$. Recall that Z^n is a solution of the martingale problem for C^n given by (4.4), and it follows that Z is a solution of the martingale problem for $C \subset B(\mathcal{P}(E)) \times B(\mathcal{P}(E))$ given by

$$C = \{(\langle f, \mu^\infty\rangle, \langle Af, \mu^\infty\rangle) : f \in \mathcal{D}(A)\},$$

where $\langle f, \mu^\infty\rangle$ denotes integration of f by the product measure μ^∞. Since the martingale problem for A is well-posed (assuming the martingale problem for B is well-posed) and exchangeability of $X(0)$ implies exchangeability of $X(t)$, it follows that the restricted martingale problem for $(A, \mathcal{H})$ is well-posed, where $\mathcal{H} \subset B(E^\infty)$ is the collection of all functions h of the form

$$h(x_1, x_2, \ldots) = f(x_1, \ldots, x_m) - f(x_{\sigma_1}, \ldots, x_{\sigma_m})$$

for some $f \in \cup_{m=1}^\infty B(E^m)$ and some permutation $(\sigma_1, \ldots, \sigma_m)$. Note that the restriction is just the requirement of the exchangeability of $X(t)$. Fix $\eta_0 \in \mathcal{P}(E)$, and define $\gamma : E^\infty \to \mathcal{P}(E)$ by

$$\gamma(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \delta_{x_i}$$

if the limit exists and $\gamma(x) = \eta_0$ otherwise. Define $\pi_t = Z^\infty(t)$. It follows that (Z, π) is a solution of the filtered martingale problem for (A, γ), and by Corollary 3.7, the solution of the martingale problem for C is well-posed. In particular, we have

$$E[f(X_1(t), \ldots, X_m(t)) | \mathcal{F}_t^Z] = \langle f, Z^m(t)\rangle.$$

Note that C is the generator for the Fleming-Viot process with mutation operator B. (See, for example, Ethier and Kurtz (1993).) For further applications of Corollary 3.7 to proofs of uniqueness of martingale problems for measure-valued processes, see Donnelly and Kurtz (1997).

4.3 Burke's theorem and quasireversibility.

One of the motivating examples for Corollary 3.5 is the proof of Burke's output theorem given by Kliemann, Koch, and Marchetti (1990). For the simplest example, let Q be an $M/M/1$ queue length process and D the corresponding departure process. Then the generator for $X = (Q, D)$ is

$$\begin{aligned} Af(k,l) = \lambda(f(k+1,l) - f(k,l)) \\ +\mu I_{[1,\infty)}(k)(f(k-1,l+1) - f(k,l)). \end{aligned}$$

Assuming $\lambda < \mu$, the stationary distribution for Q is given by

$$\pi_0 = \sum_{k=0}^{\infty} \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^k \delta_k.$$

Define $\gamma(k,l) = l$ and $\alpha(l,\cdot) = \pi_0 \times \delta_l$. Then $\alpha Af(\cdot,l) = \lambda(\pi_0 f(\cdot,l+1) - \pi_0 f(\cdot,l))$, and we see that C in Corollary 3.5 is just the generator of the Poisson process with parameter λ. Consequently, if D is a Poisson process with parameter λ, then (π_0, D) is a solution of the filtered martingale problem for (A, π_0, γ), and we have Burke's theorem.

Theorem 4.2 *Let (Q, D) be the solution of the martingale problem for $(A, \pi_0 \times \delta_0)$. Then D is a Poisson process with intensity λ, Q is stationary, and $Q(t)$ is independent of $D(\cdot \wedge t)$.*

This theorem has been generalized in a variety of ways. For example, Serfozo (1989) gives conditions under which a variety of counting process functionals of a Markov chain are Poisson. Serfozo's arguments are based on time-reversal and Watanabe's characterization of a Poisson process. To see how these results can be obtained using Corollary 3.5, let

$$A_1 f(x) = \lambda(x) \int_E (f(y) - f(x))\mu(x,dy)$$

be the generator of a pure-jump Markov process on E with stationary distribution π_0. Let $\Phi \subset E \times E - \{(x,x) : x \in E\}$ and $\varphi : \Phi \to E'$. Let $J(E')$ be the collection of integer-valued measures on E'. For a solution X of the martingale problem for A_1, define a $J(E')$-valued process by

$$N(\Gamma,t) = \sum_{s\leq t} I_\Phi(X(s-),X(s))\delta_{\varphi(X(s-),X(s))}(\Gamma),$$

that is, $N(\Gamma,t)$ counts the number of times $(X(s-),X(s)) \in \Phi$ and $\varphi(X(s-),X(s)) \in \Gamma$. Then (X,N) is a solution of the martingale problem for

$$Af(x,z) = \lambda(x) \int_E (f(y, z + I_\Phi(x,y)\delta_{\varphi(x,y)}) - f(x,z))\mu(x,dy).$$

As above, let $\gamma(x,z) = z$ and $\alpha(z,\cdot) = \pi_0 \times \delta_z$. Then

$$\begin{aligned} \alpha Af(z) &= \int_E \lambda(x) \int_E (f(y, z + I_\Phi(x,y)\delta_{\varphi(x,y)}) - f(x,z))\mu(x,dy)\pi_0(dx) \\ &= \int_E \lambda(x) \int_E (f(y, z + I_\Phi(x,y)\delta_{\varphi(x,y)}) - f(y,z))\mu(x,dy)\pi_0(dx) \\ &\quad + \int_E \lambda(x) \int_E (f(y,z) - f(x,z))\mu(x,dy)\pi_0(dx). \end{aligned}$$

Note that the second term on the right is zero since π_0 is a stationary distribution for A_1. Suppose that there exists a measure β on E' such that for each $h \in B(E \times E')$,

$$\int_{E\times E} \lambda(x) I_\Phi(x,y) h(y,\varphi(x,y)) \mu(x,dy)\pi_0(dx) = \int_{E\times E'} h(y,u)\pi_0(dy)\beta(du). \qquad (4.6)$$

Then

$$\alpha A f(z) = \int_{E'} (\pi_0 f(z+\delta_u) - \pi_0 f(z))\beta(du).$$

Consequently, if X is a solution of the martingale problem for (A_1, π_0), then (N, π_0) is a solution of the filtered martingale problem for $(A, \pi_0 \times \delta_0, \gamma)$ which implies N is a solution of the martingale problem for A_2 given by

$$A_2 f(z) = \int_{E'} (f(z+\delta_u) - f(z))\beta(du).$$

It follows that $\xi(\Gamma \times [0,t]) = N(\Gamma, t)$ defines a Poisson random measure on $E' \times [0,\infty)$ with mean measure $\beta(du) \times dt$ and $X(t)$ is independent of $\sigma(N(\Gamma, s) : s \le t, \Gamma \in \mathcal{B}(E'))$.

Note that (4.6) is essentially the condition $\alpha^*(x, B) = aF(B)$ in Theorem 3.2 of Serfozo (1989). Serfozo also shows that this condition is necessary for N to be Poisson and the independence property to hold.

4.4 Reflecting Brownian motion.

Harrison and Williams (1990, 1992) have considered analogues of the Poisson output theorems for Brownian network models. Here we give a version of their result for a single service station.

Let $W = (W_1, \ldots, W_m)^T$ be standard Brownian motion, and let X be the one-dimensional reflecting Brownian motion satisfying the Skorohod equation

$$X(t) = X(0) + \sum_{i=1}^m a_i W_i(t) - bt + \Lambda(t).$$

Let C be a $d \times m$ matrix, $c \in \mathbb{R}^d$, and $D = CC^T$. Define

$$Y(t) = Y(0) + CW(t) + c\Lambda(t)$$

By Itô's formula

$$\begin{aligned} f(X(t), Y(t)) &= f(X(0), Y(0)) + MG \\ &\quad + \int_0^t Af(X(s), Y(s))ds \\ &\quad + \int_0^t Bf(0, Y(s))d\Lambda(s) \end{aligned}$$

where

$$Af = \frac{1}{2}\sum a_i^2 f_{xx} + \sum_j (\sum_i c_{ji} a_i) f_{xy_j} + \frac{1}{2}\sum_{jk} d_{jk} f_{y_j y_k} - b f_x$$

and

$$Bf = f_x + \sum c_j f_{y_j}.$$

Consequently, if we take $\mathcal{D}(A) = \{f \in C_c^2(\mathbb{R}^{d+1}) : Bf = 0\}$, then (X, Y) is a solution for the martingale problem for A.

For $\beta = 2b/\sum a_i^2$, let $\pi_0(dx) = \beta e^{-\beta x}dx$. Then π_0 is the stationary distribution for X. As above, define $\alpha(y, \cdot) = \pi_0 \times \delta_y$. Then

$$\begin{aligned}\alpha Af(y) = {} & \frac{1}{2}\sum_{j,k} d_{jk}\pi_0 f_{y_j y_k}(y) + \sum_j(\beta\sum_i c_{ji}a_i)\pi_0 f_{y_j}(y) \\ & -\beta\left(\frac{1}{2}\sum_i a_i^2 f_x(0,y) + \sum_j(\sum_i c_{ji}a_i)f_{y_j}(0,y)\right)\end{aligned} \tag{4.7}$$

If c, C, and the a_i satisfy

$$c_j = \frac{2\sum_i c_{ji}a_i}{\sum_i a_i^2},$$

then the second term on the right of (4.7) is zero by the boundary condition $Bf = 0$. Let $\tilde{Y}$ be a solution of the martingale problem for

$$A_0 f(y) = \frac{1}{2}\sum_{j,k} d_{jk} f_{y_j y_k}(y) + \sum_j(\alpha\sum_i c_{ji}a_i)f_{y_j}(y)$$

with $\tilde{Y}(0) = 0$. Then $(\tilde{Y}, \pi_0)$ is a solution of the filtered martingale problem for $(A, \pi_0 \times \delta_0, \gamma)$, $\gamma(x, y) = y$. It follows that if (X, Y) is a solution of the martingale problem for $(A, \pi_0 \times \delta_0)$, then Y is Brownian motion with generator A_0 and $X(t)$ is independent of $\sigma(Y(s) : s \leq t)$.

A Appendix

A.1 Proof of Theorem 2.6

Let A satisfy the conditions of Theorem 2.6. Since A is a pre-generator, we can assume $(1, 0) \in A$. The domain of the linear span of A is then an algebra that separates points and vanishes nowhere. In addition, we can assume that $\{g_k\}$ is closed under multiplication. Let $\mathcal{I}$ be the collection of finite subsets of positive integers, and for $I \in \mathcal{I}$, let $k(I)$ satisfy $g_{k(I)} = \prod_{i\in I} g_i$. For each k, there exists $a_k \geq |g_k|$. Let

$$\hat{E} = \{z \in \prod_{i=1}^{\infty}[-a_i, a_i] : z_{k(I)} = \prod_{i\in I} z_i, I \in \mathcal{I}\}.$$

Note that $\hat{E}$ is compact. Following Bhatt and Karandikar (1993), define $G : E \to \hat{E}$ by

$$G(x) = (g_1(x), g_2(x), \ldots).$$

Then G has a measurable inverse defined on the (measurable) set $G(E)$. We will need the following lemma.

Lemma A.1 *Let $\nu \in \mathcal{P}(E)$. Then there exists a unique measure $\mu \in \mathcal{P}(\hat{E})$ satisfying $\int_E g_k d\nu = \int_{\hat{E}} z_k \mu(dz)$. In particular, if Z has distribution μ, then $G^{-1}(Z)$ has distribution ν.*

Proof. Existence is immediate. Take $\mu = \nu G^{-1}$. Since $\hat{E}$ is compact, $\{\prod_{i\in I} z_i : I \in \mathcal{I}\}$ is separating. Consequently, uniqueness follows from the fact that

$$\int_{\hat{E}} \prod_{i\in I} z_i \mu(dz) = \int_{\hat{E}} z_{k(I)}\mu(dz) = \int_E g_{k(I)} d\nu.$$

□

Suppose that $\mu \in \mathcal{P}(E)$ satisfies

$$\int_E Af d\mu = 0, \quad f \in \mathcal{D}(A).$$

Then there exists a stationary process that is a solution of the martingale problem for (A, μ). This assertion is essentially Theorem 3.1 of Bhatt and Karandikar (1993) with the assumption of existence of solutions of the $D_E[0,\infty)$ martingale problem replaced by the assumption that A is a pre-generator (which is, as we noted previously, implied by the existence of cadlag solutions). The proof of the earlier result only needs to be modified using the fact that a pre-generator is dissipative (needed in the definition of A_n) and the following generalization of Lemma 4.9.16 of Ethier and Kurtz (1986) (needed to show that Λ is a positive functional).

Lemma A.2 *Let $A \subset B(E) \times B(E)$ be a pre-generator. Suppose that φ is continuously differentiable and convex on $D \subset \mathbb{R}^m$, that $f_1, \ldots, f_m \in \mathcal{D}(A)$ and $(f_1, \ldots, f_m) : E \to D$, and that $(\varphi(f_1, \ldots, f_m), h) \in A$. Then*

$$h \geq \nabla\varphi(f_1, \ldots, f_m) \cdot (Af_1, \ldots, Af_m).$$

Proof. Since A is a pre-generator, there exists λ_n and μ_n such that

$$\begin{aligned} h(x) &= \lim_{n\to\infty} \lambda_n(x) \int_E (\varphi(f_1(y), \ldots, f_m(y)) - \varphi(f_1(x), \ldots, f_m(x))\mu_n(x, dy) \\ &\geq \lim_{n\to\infty} \nabla\varphi(f_1(x), \ldots, f_m(x)) \cdot \lambda_n(x) \int_E (f_1(y) - f_1(x), \ldots, f_m(y) - f_m(x))\mu_n(x, dy) \\ &= \nabla\varphi(f_1(x), \ldots, f_m(x)) \cdot (Af_1(x), \ldots, Af_m(x)) \end{aligned}$$

□

To complete the proof of Theorem 2.6, let ν be a solution of the forward equation, that is

$$\nu_t f = \nu_0 f + \int_0^t \nu_s Af ds, \quad f \in \mathcal{D}(A). \tag{A.1}$$

Fix $\alpha > 0$, and define the generator $\tilde{A}$ on $E_0 = [0, \infty) \times E$ by

$$\tilde{A}(\varphi f)(s, x) = \varphi(s)Af(x) + \varphi'(s)f(x) + \alpha \left[\varphi(0)\int_E f(y)\nu_0(dy) - \varphi(s)f(x)\right] \tag{A.2}$$

for $f \in \mathcal{D}(A)$ and $\varphi \in \hat{C}^1[0, \infty)$, and observe that $\tilde{A}$ also satisfies the conditions of Theorem 2.6.

Define the measure $\pi \in \mathcal{P}([0,\infty)\times E)$ by

$$\int_{[0,\infty)\times E} h(s,x)\pi(ds\times dx) = \alpha\int_0^\infty e^{-\alpha s}\int_E h(s,x)\nu_s(dx)ds \tag{A.3}$$

for $h \in B([0,\infty)\times E)$. Then, using (A.1), one can show $\int_{[0,\infty)\times E}\tilde{A}fd\pi = 0$ for $f \in \mathcal{D}(\tilde{A})$, and hence, by the construction of Bhatt and Karandikar, there exists a cadlag, stationary process (S,Z) with values in $[0,\infty)\times\hat{E}$, such that $(S,Y) = (S,G^{-1}(Z))$ is a solution of the martingle problem for $(\tilde{A},\pi)$. In particular, for $k = 1,2,\ldots$

$$\varphi(S(t))Z_k(t) - \int_0^t [\varphi(S(u))Ag_k(Y(u)) + \varphi'(S(u))Z_k(u) \\ +\alpha\left(\varphi(0)\int_E g_k d\nu_0 - \varphi(S(u))Z_k(u)\right)]du$$

is a martingale.

We assume, without loss of generality, that (S,Z) is defined for all $t \in (-\infty,\infty)$ on a probability space $(\Omega,\mathcal{F},P)$, and define $\mathcal{G}_t = \sigma(S(s),Z(s) : s \le t)$. Let $\tau_{-1} = \sup\{t < 0 : S(t) = 0\}$, $\tau_1 = \inf\{t > 0 : S(t) = 0\}$, and $\tau_2 = \{t > \tau_1 : S(t) = 0\}$. Taking $\varphi_n(s) = e^{-ns}$ and applying the optional sampling theorem, it is easy to check that

$$E[Z_k(\tau_1)] = \lim_{n\to\infty} E[\varphi_n(S(\tau_1))Z_k(\tau_1)] = E[\tau_1]\alpha\int_E g_k d\nu_0 = \int_E g_k d\nu_0.$$

By Lemma A.1, it follows that $P\{Z(\tau_1) \in G(E)\} = 1$ and $Y(\tau_1)$ has distribution ν_0.

For $t \ge 0$, set $\tilde{Z}(t) = Z(\tau_1 + t)$, $X(t) = Y(\tau_1 + t)$, $L(t) = [\alpha(\tau_1 - \tau_{-1})]^{-1}e^{\alpha t}I_{[0,\tau_2-\tau_1]}(t)$, and $\mathcal{F}_t = \mathcal{G}_{\tau_1+t}$. Then L is an $\{\mathcal{F}_t\}$-martingale with $E[L(t)] = 1$, and for $k = 1,2,\ldots$,

$$\tilde{Z}_k(t) - \int_0^t \left[Ag_k(X(s)) + \alpha\left(\int_E f d\nu_0 - \tilde{Z}_k(s)\right)\right] ds$$

is an $\{\mathcal{F}_t\}$-martingale. Let $\hat{P}$ be given by $\hat{P}(C) = E[I_C L(t)]$ for $C \in \mathcal{F}_t$. Then, under $\hat{P}$, we claim that X is a solution of the martingale problem for A and $\nu_t = \hat{P}X^{-1}(t)$.

Following the argument in the proof of Theorem 4.1 in Kurtz and Stockbridge (1997),

$$L(t)\tilde{Z}_k(t) - \int_0^t L(s)Ag_k(X(s))ds$$

is an $\{\mathcal{F}_t\}$-martingale under P and

$$\tilde{Z}_k(t) - \int_0^t Ag_k(X(s))ds$$

is an $\{\mathcal{F}_t\}$-martingale under $\hat{P}$. Finally, as in (4.8) in Kurtz and Stockbridge (1997),

$$E^{\hat{P}}\left[\alpha\int_0^\infty \varphi(t)\tilde{Z}_k(t)dt\right] = \alpha\int_0^\infty e^{-\alpha t}\varphi(t)\int_E g_k d\nu_t dt.$$

By the right continuity of $\tilde{Z}_k$ and the continuity of $\int_E g_k d\nu_t$, for each $t > 0$,

$$E[\tilde{Z}_k(t)] = \int_E g_k d\nu_t.$$

As before, it follows that $P\{\tilde{Z}(t) \in G(E)\} = 1$ and that $X(t)$ has distribution ν_t which completes the proof of Theorem 2.6. □

A.2 Proof of Theorem 2.7

The argument is similar to Section A.1, again following the proof of Theorem 4.1 of Stockbridge and Kurtz (1997). We now take

$$\tilde{A}_0(\varphi f)(s,x,y) = \varphi(s)A_0 f(x,y) + \varphi'(s)f(x) + \alpha\left[\varphi(0)\int_E f d\nu_0 - \varphi(s)f(x)\right] \tag{A.4}$$

for $f \in \mathcal{D}(A_0)$ and $\varphi \in \hat{C}^1[0,\infty)$, and observe that $\tilde{A}_0$ also satisfies the conditions of Theorem 2.7. Define the measure $\pi \in \mathcal{P}([0,\infty)\times E\times F)$ by

$$\int_{[0,\infty)\times E} h(s,x,y)\pi(ds\times dx\times dy) = \alpha\int_0^\infty e^{-\alpha s}\int_E\int_F h(s,x,y)\eta(x,dy)\nu_s(dx)ds. \tag{A.5}$$

As before, $\int_{[0,\infty)\times E\times F}\tilde{A}_0 f d\pi = 0$, and we can use the controlled martingale problem results of Bhatt and Borkar (1996) and Kurtz and Stockbridge (1997) with modifications similar to those outlined in Section A.1. □

A.3 Optional projections.

The following is essentially a result of Yor (1977). We do not, however, assume that the filtrations are right continuous. See Lenglart (1983) for more general results of this type.

Theorem A.3 *Let $(\Omega,\mathcal{F},P)$ be a complete probability space, let E be a complete, separable metric space, and let η be a measurable, $\mathcal{P}(E)$-valued process. Let $\{\mathcal{F}_t\}$ be a complete filtration. Then there exists an $\{\mathcal{F}_t\}$-optional, $\mathcal{P}(E)$-valued process π such that for each finite (that is, $P\{\tau<\infty\}=1$) $\{\mathcal{F}_t\}$-stopping time τ and each $A\in\mathcal{B}(E)$, $E[\eta_\tau(A)|\mathcal{F}_\tau]=\pi_\tau(A)$ a.s. If, in addition, $\{\mathcal{F}_t\}$ is right continuous and η is right continuous (cadlag), then π is right continuous (cadlag).*

Proof. We follow the proof of Morando's theorem. (See for example, Ethier and Kurtz (1986), Appendix 8.) Let $\{x_i, i\geq 1\}$ be a countable dense subset of E. Let $C_1, C_2,\ldots$ be some ordering of the collection of balls $\{B_{k^{-1}}(x_i): i,k=1,2,\ldots\}$. Let $\mathcal{B}_n$ be the (finite) σ-algebra generated by $C_1,\ldots,C_n$, and let $\mathcal{P}_n$ be the collection of probability measures defined on $\mathcal{B}_n$. The optional projection theorem (see Ethier and Kurtz (1986) Theorem 2.4.2 for a version not requiring right continuity of $\{\mathcal{F}_t\}$) implies the existence of an optional, $\mathcal{P}_n$-valued process $\hat{\pi}^n$ satisfying $E[\eta_\tau(A)|\mathcal{F}_\tau]=\hat{\pi}^n_\tau(A)$ a.s. for each $A\in\mathcal{B}_n$, and each finite $\{\mathcal{F}_t\}$-stopping time τ. The sequence can be constructed to be consistent in the sense that $\hat{\pi}^n(A)=\hat{\pi}^{n+1}(A)$ for $A\in\mathcal{B}_n$. Let $A_1^n,\ldots,A_{m_n}^n$ be the partition generating $\mathcal{B}_n$ and let $x_i^n\in A_i^n$. Then $\hat{\pi}^n$ can be extended to a measure on $\mathcal{B}(E)$ by defining

$$\hat{\pi}^n(A) = \sum_{i=1}^{m_n}\delta_{x_i^n}(A)\hat{\pi}^n(A_i^n)\,,\quad A\in\mathcal{B}(E).$$

Taking the usual weak topology on $\mathcal{P}(E)$, the set $O=\{(t,\omega):\lim_{n\to\infty}\hat{\pi}^n_t(\cdot,\omega)\text{ exists in }\mathcal{P}(E)\}$ is optional. Define $\pi_t(\cdot,\omega)=\lim_{n\to\infty}\hat{\pi}^n_t(\cdot,\omega)$ for $(t,\omega)\in O$ and $\pi_t(\cdot,\omega)=\pi_*$ for $(t,\omega)\in O^c$ for some fixed $\pi_*\in\mathcal{P}(E)$. Then π is an optional, $\mathcal{P}(E)$-valued process. We will show that π satisfies the conclusion of the theorem.

For each compact $K \subset E$ and each $\epsilon > 0$, there exists $n_{K,\epsilon}$ and $A \in \mathcal{B}_{n_{K,\epsilon}}$ such that $K \subset A \subset K^\epsilon$. Then for any finite stopping time τ and $n \geq n_{K,\epsilon}$ we have $\hat{\pi}_\tau^n(K^\epsilon) \geq E[\eta_\tau(K)|\mathcal{F}_\tau]$ a.s. Since $E[\eta_\tau]$ is a probability measure, there exists an increasing sequence of compact sets $K_1 \subset K_2 \subset \cdots$ such that $\lim_{k\to\infty} E[\eta_\tau(K_k)] = 1$ and hence $\lim_{k\to\infty} E[\eta_\tau(K_k)|\mathcal{F}_\tau] = 1$ a.s. It follows that the sequence $\{\hat{\pi}_\tau^n\}$ is almost surely relatively compact. The consistency on $\cup\mathcal{B}_n$ ensures that there can be at most one limiting measure, so we must have $\lim_{n\to\infty} \hat{\pi}_\tau^n = \pi_\tau$ a.s. For any open set $G \in \cup\mathcal{B}_n$, we have $E[\eta_\tau(G)|\mathcal{F}_\tau] = \lim_{n\to\infty} \hat{\pi}_\tau^n(G) \geq \pi_\tau(G)$. But any open set can be written as an increasing union of open sets in $\cup\mathcal{B}_n$, so the inequality $E[\eta_\tau(G)|\mathcal{F}_\tau] \geq \pi_\tau(G)$ a.s. holds for all open sets. Taking decreasing intersections, the inequality holds for all closed sets, that is for G^c for each open G, so in fact equality holds. The monotone class theorem gives equality for all Borel sets.

If $\{\mathcal{F}_t\}$ is right continuous, then the optional projection of a right continuous (cadlag) process is right continuous (cadlag) (Dellacherie (1972), Theorem V-T20). Consequently, if η is right continuous (cadlag) and $f \in \bar{C}(E)$, then

$$\pi_t f = E[\eta_t f|\mathcal{F}_t]$$

is right continuous (cadlag). Since there is a countable convergence determining subset of $\bar{C}(E)$, it follows that π must be right continuous (cadlag). □

Corollary A.4 *Let X be an E-valued, measurable, $\{\mathcal{F}_t\}$-adapted process. Then there exists an $\{\mathcal{F}_t\}$-optional process $\hat{X}$ such that $X(t) = \hat{X}(t)$ a.s. for each $t \geq 0$.*

Proof. By the previous theorem, there exists a $\mathcal{P}(E)$-valued optional process π such that $E[I_A(X(t))|\mathcal{F}_t] = \pi_t(A)$. But since X is $\{\mathcal{F}_t\}$-adapted, $\pi_t = \delta_{X(t)}$ a.s. The set $D = \{(t,\omega) : \pi_t(\cdot,\omega)$ is degenerate$\}$ is optional. Fix x_0, and define $\hat{X}(t,\omega)$ so that $\pi_t(\cdot,\omega) = \delta_{\hat{X}(t,\omega)}$ on D and $\hat{X}(t,\omega) = x_0$ on D^c. It follows that $\hat{X}$ is optional and that $\hat{X}(t) = X(t)$ a.s. for each t. □

A.4 Uniqueness for martingale problem implies Markov property.

Let $A \subset B(E) \times B(E)$ and let $\nu_0 \in \mathcal{P}(E)$. A measurable process X is a solution of the martingale problem for (A, ν_0) if $X(0)$ has distribution ν_0 and there exists a filtration $\{\mathcal{F}_t\}$ such that X is $\{\mathcal{F}_t\}$-adapted and

$$f(X(t)) - \int_0^t g(X(s))ds \tag{A.6}$$

is an $\{\mathcal{F}_t\}$-martingale for each $(f, g) \in A$. We say that the solution of the martingale problem for (A, ν_0) is unique if any two solutions have the same finite dimensional distributions. Ethier and Kurtz (1986), Theorem 4.4.2, states that if uniqueness holds for every $\nu_0 \in \mathcal{P}(E)$, then every solution is a Markov process. In fact, uniqueness for a fixed initial distribution implies the solution with that initial distribution is Markov. Recall that we are assuming the (E, r) is complete and separable. (The author thanks Ely Merzbach for helpful communications regarding the results on bimeasures in Karni and Merzbach (1990).)

Theorem A.5 *Let $A \subset B(E) \times B(E)$ and $\nu_0 \in \mathcal{P}(E)$, and assume that uniqueness holds for the martingale problem for (A, ν_0). Suppose that X is a solution of the martingale problem for (A, ν_0) with respect to the filtration*

$$\mathcal{F}_t = \sigma(X(s) : s \leq t) \vee \sigma(\int_0^s h(X(r))dr : s \leq t, h \in B(E)). \tag{A.7}$$

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

Remark A.6 *Let $\mathcal{F}_t^X = \sigma(X(s) : s \leq t)$. If X is cadlag, or more generally, if X is progressive with respect to $\{\mathcal{F}_t^X\}$, then $\mathcal{F}_t = \mathcal{F}_t^X$, $t \geq 0$.*

Proof. Suppose that the solution X is defined on the probability space $(\Omega, \mathcal{F}, P)$. We must show that for each $r, s \geq 0$, $f \in B(E)$,

$$E[f(X(r+s))|\mathcal{F}_r] = E[f(X(r+s))|X(r)],$$

or, equivalently, for each bounded $\mathcal{F}_r$-measurable random variable H

$$E[f(X(r+s))H] = E[E[f(X(r+s))|X(r)]H]. \tag{A.8}$$

Fix $r \geq 0$, and define

$$Q_r(F \cap G) = E^P[I_F E^P[I_G|X(r)]], \quad F, G \in \mathcal{F}.$$

Since the argument will involve more than one probability measure, we use E^P to denote the expectation (or conditional expectation) under P. Then Q_r is a bimeasure on $(\Omega, \mathcal{F}) \times (\Omega, \mathcal{F})$, which by Theorem 2.8b of Karni and Merzbach (1990) extends to a countably additive probability measure on the product σ-algebra $\mathcal{F} \times \mathcal{F}$. Define $\tilde{X}_r$ on $(\Omega \times \Omega, \mathcal{F} \times \mathcal{F}, Q_r)$ by

$$\tilde{X}_r(s, \omega_1, \omega_2) = \begin{cases} X(s, \omega_1), & s \leq r \\ X(s, \omega_2), & s > r. \end{cases}$$

Define $X_i(s, \omega_1, \omega_2) = X(s, \omega_i)$, and note that for $\Gamma \in \mathcal{B}(E)$,

$$Q_r(X_1(r) \in \Gamma, X_2(r) \in \Gamma) = E^P[I_\Gamma(X(r)) E^P[I_\Gamma(X(r))|X(r)]] = P\{X(r) \in \Gamma\}$$

and it follows that

$$Q_r\{X_1(r) = X_2(r)\} = 1.$$

We claim that $\tilde{X}_r$ is a solution of the martingale problem for (A, ν_0). Since $Q_r\{\tilde{X}_r(0) \in \Gamma)\} = E^P[I_\Gamma(X(0))] = \nu_0(\Gamma)$, $\tilde{X}_r$ has the correct initial distribution. We need to show that for $f \in \mathcal{D}(A)$, $t_1 < \ldots < t_{m+1}$ and $g_i \in B(E)$,

$$E^{Q_r}[(f(\tilde{X}_r(t_{m+1}) - f(\tilde{X}_r(t_m)) - \int_{t_m}^{t_{m+1}} Af(\tilde{X}_r(s))ds) \prod_{i=1}^m g_i(\tilde{X}_r(t_i))] = 0. \tag{A.9}$$

If $t_{m+1} \le r$, the equality is immediate. Suppose $t_m \ge r$. Then

$$\begin{aligned}
& E^{Q_r}[(f(\tilde{X}_r(t_{m+1}) - f(\tilde{X}_r(t_m)) - \int_{t_m}^{t_{m+1}} Af(\tilde{X}_r(s))ds) \prod_{i=1}^{m} g_i(\tilde{X}_r(t_i))] \\
&= E^P[E^P[(f(X(t_{m+1}) - f(X(t_m)) - \int_{t_m}^{t_{m+1}} Af(X(s))ds) \prod_{t_i \ge r} g_i(X(t_i))|X(r)] \prod_{t_i < r} g_i(X(t_i))] \\
&= 0.
\end{aligned}$$

If $t_m < r < t_{m+1}$, write

$$\begin{aligned}
f(\tilde{X}_r(t_{m+1}) - f(\tilde{X}_r(t_m)) - \int_{t_m}^{t_{m+1}} Af(\tilde{X}_r(s))ds \\
= f(\tilde{X}_r(t_{m+1}) - f(\tilde{X}_r(r)) - \int_{r}^{t_{m+1}} Af(\tilde{X}_r(s))ds \\
+ f(\tilde{X}_r(r) - f(\tilde{X}_r(t_m)) - \int_{t_m}^{r} Af(\tilde{X}_r(s))ds,
\end{aligned}$$

and check the identity for each term separately.

The identity (A.9) actually only verifies

$$E[f(\tilde{X}_r(t_{m+1}) - f(\tilde{X}_r(t_m)) - \int_{t_m}^{t_{m+1}} Af(\tilde{X}_r(s))ds|\mathcal{F}_{t_m}^{\tilde{X}_r}] = 0$$

where $\mathcal{F}_t^{\tilde{X}_r} = \sigma(\tilde{X}_r(s) : s \le t)$. The conditioning for the (possibly) larger σ-algebra

$$\tilde{\mathcal{F}}_t = \sigma(\tilde{X}_r(s) : s \le t) \vee \sigma(\int_0^s h(\tilde{X}_r(u))du : s \le t, h \in B(E))$$

can be obtained by including factors of the form $g(\int_0^t h(\tilde{X}_r(s))ds)$ for $t \le r$ and $g(\int_r^t h(\tilde{X}_r(s))ds)$ for $t \ge r$ in the product.

Since $\tilde{X}_r$ is a solution of the martingale problem for (A, ν_0), $\tilde{X}_r$ and X have the same finite dimensional distributions. Note that the finite dimensional distributions of X determine the joint distribution of any collection of random variables of the form

$$(X(t_1), \ldots, X(t_m), \int_0^{t_1} h_1(X(s))ds, \ldots, \int_0^{t_m} h_m(X(s))ds),$$

$h_1, \ldots, h_m \in B(E)$. (Compute joint moments of $g_i(X(t_i))$ and the integral terms.) Consequently, if

$$H = H(X(t_1), \ldots, X(t_m), \int_0^{t_1} h_1(X(s))ds, \ldots, \int_0^{t_m} h_m(X(s))ds),$$

where $t_1, \ldots, t_m \le r$, and H_r is defined similarly with X replaced by X_r, then

$$E^P[f(X(r+s))H] = E^{Q_r}[f(X_r(r+s))H_r] = E^P[E^P[f(X(r+s))|X(r)]H],$$

which gives (A.8). $\square$

References

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

Bhatt, Abhay G. and Borkar, V. S. (1996). Occupation measures for controlled Markov processes: Characterization and optimality. *Ann. Probab.* 24, 1531-1562.

Bhatt, Abhay G., Kallianpur, G., Karandikar, Rajeeva L. (1995). Uniqueness and robustness of solution of measure-valued equations of nonlinear filtering. *Ann. Probab.* 23, 1895-1938.

Cameron, Murray A. (1973). A note on functions of Markov processes with an application to a sequence of χ^2 statistics. *J. Appl. Probab.* 10, 895-900.

Dawson, Donald A. (1993). Measure-valued Markov processes. *École d'Été de Probabilités de Saint-Flour XXI, 1991*, 1-260, *Lecture Notes in Math.*, 1541, Springer, Berlin.

Dellacherie, Claude (1972). *Capacités et processus stochastiques.* Springer, Berlin.

Donnelly, Peter E. and Kurtz, Thomas G. (1996). A countable representation of the Fleming-Viot measure-valued diffusion. *Ann. Probab.* 24, 698-742.

Donnelly, Peter E. and Kurtz, Thomas G. (1997). Particle representations for measure-valued population models. *Ann. Probab.* (to appear)

Ethier, Stewart N. and Kurtz, Thomas G. (1986). *Markov Processes: Characterization and Convergence.* Wiley, New York.

Ethier, Stewart N. and Kurtz, Thomas G. (1993). Fleming-Viot processes in population genetics. *SIAM J. Control Optim.* 31 (1993), 345-386.

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

Glover, Joseph (1991). Markov functions. *Ann. Inst. H. Poincaré Probab. Statist. 27, 221-238.*

Harrison, J. M. and Williams, R. J. (1990). On the quasireversibility of a multiclass Brownian service station. *Ann. Probab.* 18, 1249-1268.

Harrison, J. M. and Williams, R. J. (1992). Brownian models of feedforward queueing networks: Quasireversibility and product form solutions. *Ann. Appl. Probab.* 2, 263-293.

Karni, Shaul and Merzbach, Ely (1990). On the extension of bimeasures. *J. Analyse Math.* 55, 1-16.

Kelly, Frank P. (1982). Markovian functions of a Markov chain. *Sankya Ser A* 44, 372-379.

Kliemann, Wolfgang H., Koch, Georgio, and Marchetti, Federico (1990). On the unnormalized solution of the filtering problem with counting process observations. *IEEE Trans. Inform. Theory* 36, 1415-1425.

Kurtz, Thomas G. and Ocone, Daniel L. (1988). Unique characterization of conditional distributions in nonlinear filtering. *Ann. Probab.* 16, 80-107.

Kurtz, Thomas G. and Stockbridge, Richard H. (1997). Existence of Markov controls and characterization of optimal Markov controls. *SIAM J. Cont. Optim.* (to appear)

Lenglart, E. Désintégration régulière de mesure sans conditions habituelles. *Seminar on probability, XVII,* 321-345, *Lecture Notes in Math.* 986, Springer-Verlag, Berlin-New York.

Rogers, L. C. G. and Pitman, J. W. (1981). Markov functions. *Ann. Probab.* 9, 573-582.

Rosenblatt, Murray (1966). Functions of Markov processes. *Z. Wahrscheinlichkeitstheorie* 5, 232-243.

Serfozo, Richard F. (1989). Poisson functionals of Markov processes and queueing networks. *Adv. in Appl. Probab.* 21, 595-611.

Yor, Marc (1977). Sur les théories du filtrage et de la prédiction. *Séminaire de Probabilités XI. Lecture Notes in Math.* 581, 257-297. Springer, Berlin.