Electronic Journal of Probability

Vol. 6 (2001) Paper no. 17, pages 1–52.

Journal URL
http://www.math.washington.edu/~ejpecp/
Paper URL
http://www.math.washington.edu/~ejpecp/EjpVol6/paper17.abs.html

STATIONARY SOLUTIONS AND FORWARD EQUATIONS FOR CONTROLLED AND SINGULAR MARTINGALE PROBLEMS [1]

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

Richard H. Stockbridge
Department of Statistics, University of Kentucky
Lexington, KY 40506–0027
stockb@ms.uky.edu

Abstract Stationary distributions of Markov processes can typically be characterized as probability measures that annihilate the generator in the sense that $\int_E Af d\mu = 0$ for $f \in \mathcal{D}(A)$; that is, for each such μ, there exists a stationary solution of the martingale problem for A with marginal distribution μ. This result is extended to models corresponding to martingale problems that include absolutely continuous and singular (with respect to time) components and controls. Analogous results for the forward equation follow as a corollary.

Keywords singular controls, stationary processes, Markov processes, martingale problems, forward equations, constrained Markov processes.

MSC subject classifications Primary: 60J35, 93E20 Secondary: 60G35, 60J25.

Submitted to EJP on August 16, 2000. Final version accepted on January 17, 2001.

[1] THIS RESEARCH IS PARTIALLY SUPPORTED BY NSF UNDER GRANTS DMS 9626116 AND DMS 9803490.

1 Introduction

Stationary distributions for Markov processes can typically be characterized as probability measures that annihilate the corresponding generator. Suppose A is the generator for a Markov process X with state space E, where X is related to A by the requirement that

$$f(X(t)) - f(X(0)) - \int_0^t Af(X(s))ds \tag{1.1}$$

be a martingale for each $f \in \mathcal{D}(A)$. (We say that X is a solution of the *martingale problem* for A.) If μ is a stationary distribution for A, that is, there exists a stationary solution of the martingale problem with marginal distribution μ, then since (1.1) has expectation zero, we have

$$\int_E Af\,d\mu = 0, \quad f \in \mathcal{D}(A). \tag{1.2}$$

More generally, if $\{\nu_t : t \geq 0\}$ are the one-dimensional distributions of a solution, then they satisfy the *forward equation*

$$\int_E f\,d\nu_t = \int_E f\,d\nu_0 + \int_0^t \int_E Af\,d\nu_s ds, \quad f \in \mathcal{D}(A). \tag{1.3}$$

Conversely, if μ satisfies (1.2), then under mild additional assumptions, there exists a stationary solution of the martingale problem for A with marginal distribution μ, and if $\{\nu_t : t \geq 0\}$ satisfies (1.3), then there should exist a corresponding solution of the martingale problem. (See [11], Section 4.9.)

Many processes of interest in applications (see, for example, the survey paper by Shreve [24]) can be modelled as solutions to a stochastic differential equation of the form

$$dX(t) = b(X(s), u(s))ds + \sigma(X(s), u(s))dW(s) + m(X(s-), u(s-))d\xi_s \tag{1.4}$$

where X is the state process with $E = \mathbb{R}^d$, u is a control process with values in U_0, ξ is a nondecreasing process arising either from the boundary behavior of X (e.g., the local time on the boundary for a reflecting diffusion) or from a singular control, and W is a Brownian motion. (Throughout, we will assume that the state space and control space are complete, separable metric spaces.) A corresponding martingale problem can be derived by applying Itô's formula to $f(X(t))$. In particular, setting $a(x,u) = ((a_{ij}(x,u))) = \sigma(x,u)\sigma(x,u)^T$, we have

$$\begin{aligned} f(X(t)) - f(X(0)) - \int_0^t Af(X(s), u(s))ds - \int_0^t Bf(X(s-), u(s-), \delta\xi(s))d\xi(s) \\ = \int_0^t \nabla f(X(s))^T \sigma(X(s), u(s))dW(s), \end{aligned} \tag{1.5}$$

where $\delta\xi(s) = \xi(s) - \xi(s-)$,

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

and

$$Bf(x,u,\delta) = \frac{f(x+\delta m(x,u)) - f(x)}{\delta}, \quad \delta > 0, \tag{1.6}$$

with the obvious extension to $Bf(x,u,0) = m(x,u)\cdot\nabla f(x)$. We will refer to A as the generator of the *absolutely continuous part* of the process and B as the generator of the *singular part*, since frequently in applications ξ increases on a set of times that are singular with respect to Lebesgue measure. In general, however, ξ may be absolutely continuous or have an absolutely continuous part.

Suppose the state process X and control process u are stationary and that the nondecreasing process ξ has stationary increments and finite first moment. Then there exist measures μ_0 and μ_1 satisfying

$$\mu_0(H) = E[I_H(X(s),u(s))], \quad H \in \mathcal{B}(\mathbb{R}^d \times U_0),$$

for each s and

$$\mu_1(H_1 \times H_2) = \frac{1}{t}E[\int_0^t I_{H_1\times H_2}(X(s-),u(s-),\delta\xi(s))d\xi_s], \quad H_1 \in \mathcal{B}(\mathbb{R}^d \times U_0),\ H_2 \in \mathcal{B}[0,\infty),$$

for each t. Let $\mathcal{D}$ be the collection of $f \in \overline{C}^2(\mathbb{R}^d)$ for which (1.5) is a martingale. Then the martingale property implies

$$\frac{E[f(X(t))]}{t} - \frac{1}{t}E\left[\int_0^t Af(X(s),u(s))ds\right] - \frac{1}{t}E\left[\int_{[0,t]} Bf(X(s-),u(s-),\delta\xi(s))d\xi_s\right] = \frac{E[f(X(0))]}{t}$$

and, under appropriate integrability assumptions,

$$\int_{\mathbb{R}^d\times U_0} Af(x,u)\mu_0(dx\times du) + \int_{\mathbb{R}^d\times U_0\times[0,\infty)} Bf(x,u,v)\mu_1(dx\times du\times dv) = 0, \tag{1.7}$$

for each $f \in \mathcal{D}$.

As with (1.2), we would like to show that measures μ_0 and μ_1 satisfying (1.7) correspond to a stationary solution of a martingale problem defined in terms of A and B. The validity of this assertion is, of course, dependent on having the correct formulation of the martingale problem.

1.1 Formulation of martingale problem

For a complete, separable, metric space S, we define $M(S)$ to be the space of Borel measurable functions on S, $B(S)$ to be the space of bounded, measurable functions on S, $C(S)$ to be the space of continuous functions on S, $\overline{C}(S)$ to be the space of bounded, continuous functions on S, $\mathcal{M}(S)$ to be the space of finite Borel measures on S, and $\mathcal{P}(S)$ to be the space of probability measures on S. $\mathcal{M}(S)$ and $\mathcal{P}(S)$ are topologized by weak convergence.

Let $\mathcal{L}_t(S) = \mathcal{M}(S\times[0,t])$. We define $\mathcal{L}(S)$ to be the space of measures ξ on $S\times[0,\infty)$ such that $\xi(S\times[0,t]) < \infty$, for each t, and topologized so that $\xi_n \to \xi$ if and only if $\int f d\xi_n \to \int f d\xi$, for every $f \in \overline{C}(S\times[0,\infty))$ with $\text{supp}(f) \subset S\times[0,t_f]$ for some $t_f < \infty$. Let $\xi_t \in \mathcal{L}_t(S)$ denote the restriction of ξ to $S\times[0,t]$. Note that a sequence $\{\xi^n\} \subset \mathcal{L}(S)$ converges to a $\xi \in \mathcal{L}(S)$ if and only if there exists a sequence $\{t_k\}$, with $t_k \to \infty$, such that, for each t_k, $\xi^n_{t_k}$ converges weakly to ξ_{t_k}, which in turn implies ξ^n_t converges weakly to ξ_t for each t satisfying $\xi(S \times \{t\}) = 0$. Finally, we define $\mathcal{L}^{(m)}(S) \subset \mathcal{L}(S)$ to be the set of ξ such that $\xi(S \times [0,t]) = t$ for each $t > 0$.

Throughout, we will assume that the state space E and control space U are complete, separable, metric spaces.

It is sometimes convenient to formulate martingale problems and forward equations in terms of multi-valued operators. For example, even if one begins with a single-valued operator, certain closure operations lead naturally to multi-valued operators. Let $A \subset B(E) \times B(E)$. A measurable process X is a solution of the martingale problem for A if there exists a filtration $\{\mathcal{F}_t\}$ such that, for each $(f,g) \in A$,

$$f(X(t)) - f(X(0)) - \int_0^t g(X(s))ds \tag{1.8}$$

is an $\{\mathcal{F}_t\}$-martingale. Similarly, $\{\nu_t : t \geq 0\}$ is a solution of the forward equation for A if, for each $(f,g) \in A$,

$$\int_E f d\nu_t = \int_E f d\nu_0 + \int_0^t \int_E g d\nu_s ds, \quad t \geq 0. \tag{1.9}$$

Note that if we have a single valued operator $A : \mathcal{D}(A) \subset B(E) \to B(E)$, the "$A$" of (1.8) and (1.9) is simply the graph $\{(f, Af) \in B(E) \times B(E) : f \in \mathcal{D}(A)\}$.

Let A_S be the linear span of A. Note that a solution of the martingale problem or forward equation for A is a solution for A_S. We will say that A is *dissipative* if and only if A_S is dissipative, that is, for $(f,g) \in A_S$ and $\lambda > 0$,

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

An operator $A \subset B(E) \times B(E)$ 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 $(f,g) \in A$,

$$g(x) = \lim_{n\to\infty} \lambda_n(x) \int_E (f(y) - f(x))\mu_n(x, dy), \tag{1.10}$$

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

Remark 1.1 *If $A \subset \overline{C}(E) \times \overline{C}(E)$ ($\overline{C}(E)$ 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 with $\nu^x_0 = \delta_x$ that is right-continuous (in the weak topology) at zero, then A is a pre-generator. In particular, if $(f,g) \in A$, then*

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

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

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

(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 $\mathcal{D}(A) \subset \widehat{C}(E)$ ($\widehat{C}(E)$, the continuous functions vanishing at infinity), then the existence of λ_n and μ_n satisfying (1.10) 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, $A \subset \widehat{C}(E) \times \widehat{C}(E)$, 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 $A \subset \overline{C}(E) \times \overline{C}(E)$. 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.

To obtain results of the generality we would like, we must allow relaxed controls (controls represented by probability distributions on U) and a relaxed formulation of the singular part. We now give a precise formulation of the martingale problem we will consider. To simplify notation, we will assume that A and B are single-valued.

Let $A, B : \mathcal{D} \subset \overline{C}(E) \to C(E \times U)$ and $\nu_0 \in \mathcal{P}(E)$. (Note that the example above with B given by (1.6) will be of this form for $\mathcal{D} = C_c^2$ and $U = U_0 \times [0, \infty)$.) Let (X, Λ) be an $E \times \mathcal{P}(U)$-valued process and Γ be an $\mathcal{L}(E \times U)$-valued random variable. Let Γ_t denote the restriction of Γ to $E \times U \times [0,t]$. Then (X, Λ, Γ) is a relaxed solution of the *singular, controlled martingale problem* for (A, B, ν_0) if there exists a filtration $\{\mathcal{F}_t\}$ such that (X, Λ, Γ_t) is $\{\mathcal{F}_t\}$-progressive, $X(0)$ has distribution ν_0, and for every $f \in \mathcal{D}$,

$$f(X(t)) - \int_0^t \int_U Af(X(s), u)\Lambda_s(du)ds - \int_{E \times U \times [0,t]} Bf(x,u)\Gamma(dx \times du \times ds) \tag{1.11}$$

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

For the model (1.4) above, the $\mathcal{L}(E \times U)$-valued random variable Γ of (1.11) is given by $\Gamma(H \times [0,t]) = \int_0^t I_H(X(s-), u(s-), \delta\xi(s))d\xi_s$.

Rather than require all control values $u \in U$ to be available for every state $x \in E$, we allow the availability of controls to depend on the state. Let $\mathcal{U} \subset E \times U$ be a closed set, and define

$$U_x = \{u : (x,u) \in \mathcal{U}\}.$$

Let (X, Λ, Γ) be a solution of the singular, controlled martingale problem for (A, B, μ_0). The control Λ and the singular control process Γ are *admissible* if for every t,

$$\int_0^t I_{\mathcal{U}}(X(s), u)\Lambda_s(du)ds = t, \text{ and} \tag{1.12}$$

$$\Gamma(\mathcal{U} \times [0,t]) = \Gamma(E \times U \times [0,t]). \tag{1.13}$$

Note that condition (1.12) essentially requires Λ_s to have support in U_x when $X(s) = x$.

1.2 Conditions on A and B

We assume that the absolutely continuous generator A and the singular generator B have the following properties.

Condition 1.2 *i)* $A, B : \mathcal{D} \subset \overline{C}(E) \to C(E \times U)$, $1 \in \mathcal{D}$, *and* $A1 = 0, B1 = 0$.

ii) There exist $\psi_A, \psi_B \in C(E \times U)$, $\psi_A, \psi_B \geq 1$, *and constants* a_f, b_f, $f \in \mathcal{D}$, *such that*

$$|Af(x,u)| \leq a_f \psi_A(x,u), \qquad |Bf(x,u)| \leq b_f \psi_B(x,u), \qquad \forall (x,u) \in \mathcal{U}.$$

iii) Defining $(A_0, B_0) = \{(f, \psi_A^{-1} Af, \psi_B^{-1} Bf) : f \in \mathcal{D}\}$, (A_0, B_0) *is* separable *in the sense that there exists a countable collection* $\{g_k\} \subset \mathcal{D}$ *such that* (A_0, B_0) *is contained in the bounded, pointwise closure of the linear span of* $\{(g_k, A_0 g_k, B_0 g_k) = (g_k, \psi_A^{-1} A g_k, \psi_B^{-1} B g_k)\}$.

iv) For each $u \in U$, *the operators* A_u *and* B_u *defined by* $A_u f(x) = Af(x,u)$ *and* $B_u f(x) = Bf(x,u)$ *are pre-generators.*

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

Remark 1.3 *Condition (ii), which will establish uniform integrability, has been used in [27] with* ψ *only depending on the control variable and in [4] with dependence on both the state and control variables. The separability of condition (iii), which allows the embedding of the processes in a compact space, was first used in [2] for uncontrolled processes. The relaxation to the requirement that* A *and* B *be pre-generators was used in [19].*

The generalization of (1.7) is

$$\int_{E \times U} Af(x,u) \mu_0(dx \times du) + \int_{E \times U} Bf(x,u) \mu_1(dx \times du) = 0, \tag{1.14}$$

for each $f \in \mathcal{D}$. Note that if ψ_A is μ_0-integrable and ψ_B is μ_1-integrable, then the integrals in (1.14) exist.

Example 1.4 *Reflecting diffusion processes.*

The most familiar class of processes of the kind we consider 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 D (assumed smooth for the moment) and ξ increases only when X is on the boundary of D. Then there is no control, so

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

where $a(x) = ((a_{ij}(x))) = \sigma(x)\sigma(x)^T$. In addition, under reasonable conditions ξ will be continuous, so

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

If μ_0 is a stationary distribution for X, then (1.14) must hold with the additional restrictions that μ_0 is a probability measure on $\overline{D}$ and μ_1 is a measure on ∂D.

If m is not continuous (which is typically the case for the reflecting Brownian motions that arise in heavy traffic limits for queues), then a natural approach is to introduce a "control" in the singular/boundary part so that $Bf(x,u) = u \cdot \nabla f(x)$ and the set $\mathcal{U} \subset \overline{D} \times U$ that determines the admissible controls is the closure of $\{(x,u) : x \in \partial D, u = m(x)\}$. Then

$$X(t) = X(0) + \int_0^t \sigma(X(s))dW(s) + \int_0^t b(X(s))ds + \int_0^t \int_U u\Lambda_s(du)d\xi(s),$$

where again, under reasonable conditions, ξ is continuous and by admissiblity, Λ_s is a probability measure on $U_{X(s)}$. In particular, if m is continuous at $X(s)$, then $\int_U u\Lambda_s(du) = m(X(s))$, and if m is not continuous at $X(s)$, then the direction of reflection $\int_U u\Lambda_s(du)$ is a convex combination of the limit points of m at $X(s)$.

Example 1.5 *Diffusion with jumps away from the boundary.*

Assume that D is an open domain and that for $x \in \partial D$, $m(x)$ satisfies $x + m(x) \in D$. Assume that

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

where ξ is required to be the counting process that "counts" the number of times that X has hit the boundary of D, that is, assuming $X(0) \in D$, X diffuses until the first time τ_1 that X hits the boundary ($\tau_1 = \inf\{s > 0 : X(s-) \in \partial D\}$) and then jumps to $X(\tau_1) = X(\tau_1-) + m(X(\tau_1-))$. The diffusion then continues until the next time τ_2 that the process hits the boundary, and so on. (In general, this model may not be well-defined since the $\{\tau_k\}$ may have a finite limit point, but we will not consider that issue.) Then A is the ordinary diffusion operator, $Bf(x) = f(x + m(x)) - f(x)$, and $\Gamma(H \times [0,t]) = \int_0^t I_H(X(s-))d\xi(s)$.

Models of this type arise naturally in the study of optimal investment in the presence of transaction costs. (See, for example, [8, 25].) In the original control context, the model is of the form

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

where ξ counts the number of transactions. Note that ξ is forced to be a counting process, since otherwise the investor would incur infinite transaction costs in a finite amount of time. We then have A as before and $Bf(x,u) = f(x+u) - f(x)$. D and m are then determined by the solution of the optimal control problem.

Example 1.6 *Tracking problems.*

A number of authors (see, for example, [14, 26]) have considered a class of tracking problems that can be formulated as follows: Let the location of the object to be tracked be given by a Brownian motion W and let the location of the tracker be given by

$$Y(t)=Y(0)+\int_0^t u(s-)d\xi(s),$$

where $|u(s)|\equiv 1$. The object is to keep $X\equiv W-Y$ small while not consuming too much fuel, measured by ξ. Setting $X(0)=-Y(0)$, we have

$$X(t)=X(0)+W(t)-\int_0^t u(s-)d\xi(s),$$

so $Af(x)=\frac{1}{2}\Delta f(x)$ and

$$Bf(x,u,\delta)=\frac{f(x-u\delta)-f(x)}{\delta},$$

extending to $Bf(x,u,\delta)=-u\cdot\nabla f(x)$ for $\delta=0$. As before, δ represents discontinuities in ξ, that is the martingale problem is

$$f(X(t))-f(X(0))-\int_0^t\frac{1}{2}\Delta f(X(s))ds-\int_0^t Bf(X(s-),u(s-),\delta\xi(s))d\xi(s).$$

For appropriate cost functions, the optimal solution is a reflecting Brownian motion in a domain D.

1.3 Statement of main results.

In the context of Markov processes (no control), results of the type we will give appeared first in work of Weiss [29] for reflecting diffusion processes. He worked with a submartingale problem rather than a martingale problem, but ordinarily, it is not difficult to see that the two approaches are equivalent. For reflecting Brownian motion, (1.7) is just the *basic adjoint relationship* consider by Harrison et. al. (See, for example, [7].) Kurtz [16] extended Weiss's result to very general Markov processes and boundary behavior.

We say that an $\mathcal{L}(E)$-valued random variable has stationary increments if for $a_i<b_i$, $i=1,\ldots,m$, the distribution of $(\Gamma(H_1\times(t+a_1,t+b_1]),\ldots,\Gamma(H_m\times(t+a_m,t+b_m]))$ does not depend on t. Let X be a measurable stochastic process defined on a complete probability space $(\Omega,\mathcal{F},P)$, and let $\mathcal{N}\subset\mathcal{F}$ be the collection of null sets. Then $\mathcal{F}_t^X=\sigma(X(s):s\le t)$, $\overline{\mathcal{F}}_t^X=\mathcal{N}\vee\mathcal{F}_t^X$ will denote the completion of $\mathcal{F}_t^X$, and $\overline{\mathcal{F}}_{t+}^X=\cap_{s>t}\overline{\mathcal{F}}_s^X$. Let E_1 and E_2 be complete, separable metric spaces. $q:E_1\times\mathcal{B}(E_2)\to[0,1]$ is a *transition function* from E_1 to E_2 if for each $x\in E_1$, $q(x,\cdot)$ is a Borel probability measure on E_2, and for each $D\in\mathcal{B}(E_2)$, $q(\cdot,D)\in B(E_1)$. If $E=E_1=E_2$, then we say that q is a transition function on E.

Theorem 1.7 *Let A and B satisfy Condition 1.2. Suppose that $\mu_0\in\mathcal{P}(E\times U)$ and $\mu_1\in\mathcal{M}(E\times U)$ satisfy*

$$\mu_0(\mathcal{U})=\mu_0(E\times U)=1,\qquad \mu_1(\mathcal{U})=\mu_1(E\times U),\tag{1.15}$$

$$\int\psi_A(x,u)\mu_0(dx\times du)+\int\psi_B(x,u)\mu_1(dx\times du)<\infty,\tag{1.16}$$

and

$$\int_{E\times U} Af(x,u)\,\mu_0(dx\times du) + \int_{E\times U} Bf(x,u)\,\mu_1(dx\times du) = 0, \qquad \forall f\in\mathcal{D}. \tag{1.17}$$

For $i=0,1$, let μ_i^E be the state marginal μ_i and let η_i be the transition function from E to U such that $\mu_i(dx\times du) = \eta_i(x,du)\mu_i^E(dx)$.

Then there exist a process X and a random measure Γ on $E\times[0,\infty)$, adapted to $\{\overline{\mathcal{F}}_{t+}^X\}$, such that

- *X is stationary and $X(t)$ has distribution μ_0^E.*
- *Γ has stationary increments, $\Gamma(E\times[0,t])$ is finite for each t, and $E[\Gamma(\cdot\times[0,t])] = t\mu_1(\cdot)$.*
- *For each $f\in\mathcal{D}$,*

$$\begin{aligned} f(X(t)) &- \int_0^t\int_U Af(X(s),u)\,\eta_0(X(s),du)ds \\ &- \int_{E\times[0,t]}\int_U Bf(y,u)\,\eta_1(y,du)\,\Gamma(dy\times ds) \end{aligned} \tag{1.18}$$

  is an $\{\overline{\mathcal{F}}_{t+}^X\}$-martingale.

Remark 1.8 *The definition of the solution of a singular, controlled martingale problem did not require that Γ be adapted to $\{\overline{\mathcal{F}}_{t+}^X\}$, and it is sometimes convenient to work with solutions that do not have this property. Lemma 6.1 ensures, however, that for any solution with a nonadapted Γ, an adapted Γ can be constructed.*

Theorem 1.7 can in turn be used to extend the results in the Markov (uncontrolled) setting to operators with range in $M(E)$, the (not necessarily bounded) measurable functions on E, that is, we relax both the boundedness and the continuity assumptions of earlier results.

Corollary 1.9 *Let E be a complete, separable metric space. Let $\widehat{A},\widehat{B}:\mathcal{D}\subset\overline{C}(E)\to M(E)$, and suppose $\widehat{\mu}_0\in\mathcal{P}(E)$ and $\widehat{\mu}_1\in\mathcal{M}(E)$ satisfy*

$$\int_E \widehat{A}f(x)\,\widehat{\mu}_0(dx) + \int_E \widehat{B}f(x)\,\widehat{\mu}_1(dx) = 0, \qquad \forall f\in\mathcal{D}. \tag{1.19}$$

Assume that there exist a complete, separable, metric space U, operators $A,B:\mathcal{D}\to C(E\times U)$, satisfying Condition 1.2, and transition functions η_0 and η_1 from E to U such that

$$\widehat{A}f(x) = \int_U Af(x,u)\,\eta_0(x,du), \qquad \widehat{B}f(x) = \int_U Bf(x,u)\,\eta_1(x,du), \qquad \forall f\in\mathcal{D},$$

and

$$\int_{E\times U}\psi_A(x,u)\,\eta_0(x,du)\widehat{\mu}_0(dx) + \int_{E\times U}\psi_B(x,u)\,\eta_1(x,du)\widehat{\mu}_1(dx) < \infty.$$

Then there exists a solution (X,Γ) of the (uncontrolled) singular martingale problem for $(\widehat{A},\widehat{B},\widehat{\mu}_0)$ such that X is stationary and Γ has stationary increments.

Remark 1.10 *For $E = \mathbb{R}^d$, by appropriate selection of the control space and the transition functions η_i, $\widehat{A}$ and $\widehat{B}$ can be general operators of the form*

$$\frac{1}{2}\sum_{ij} a_{ij}(x)\frac{\partial^2}{\partial_i\partial_j}f(x) + b(x)\cdot\nabla f(x) + \int_{\mathbb{R}^d}(f(x+y) - f(x) - \frac{1}{1+|x|^2}y\cdot\nabla f(x))\nu(x,dy),$$

where $a = ((a_{ij}))$ is a measurable function with values in the space of nonnegative-definite $d\times d$-matrices, b is a measurable $\mathbb{R}^d$-valued function, and ν is an appropriately measurable mapping from $\mathbb{R}^d$ into the space of measures satisfying $\int_{\mathbb{R}^d}|y|^2\wedge 1\gamma(dy) < \infty$.

Proof. Define $\mu_0(dx\times du) = \eta_0(x,du)\widehat{\mu}_0(dx)$ and $\mu_1(dx\times du) = \eta_1(x,du)\widehat{\mu}_1(dx)$. The corollary follows immediately from Theorem 1.7. $\square$

Applying Corollary 1.9, we give a corresponding generalization of Proposition 4.9.19 of Ethier and Kurtz [11] and Theorem 3.1 of Bhatt and Karandikar [2] regarding solutions of the forward equation (1.3). With the singular operator B, the forward equation takes the form

$$\int_E f d\nu_t = \int_E f d\nu_0 + \int_0^t\int_E \widehat{A}f d\nu_s ds + \int_{E\times[0,t]}\widehat{B}f d\mu\ , \quad f\in\mathcal{D}, \tag{1.20}$$

where $\{\nu_t : t\geq 0\}$ is a measurable $\mathcal{P}(E)$-valued function and μ is a measure on $E\times[0,\infty)$ such that $\mu(E\times[0,t]) < \infty$ for every t.

Theorem 1.11 *Let $\widehat{A},\widehat{B}\subset\overline{C}(E)\times M(E)$, η_0, η_1, A, B, ψ_A and ψ_B be as in Corollary 1.9. Let $\{\nu_t : t\geq 0\}$ and μ satisfy (1.20) and*

$$\int_0^\infty e^{-\alpha s}\int_{E\times U}\psi_A(x,u)\,\eta_0(x,du)\nu_s(dx)ds + \int_{E\times U\times[0,\infty)} e^{-\alpha s}\psi_B(x,u)\,\eta_1(x,du)\mu(dx\times ds) < \infty, \tag{1.21}$$

for all sufficiently large $\alpha > 0$. Then there exists a solution (X,Γ) of the singular martingale problem for $(\widehat{A},\widehat{B},\mu_0)$ such that for each $t\geq 0$, $X(t)$ has distribution ν_t and $E[\Gamma] = \mu$.

If uniqueness holds for the martingale problem for $(\widehat{A},\widehat{B},\nu_0)$ in the sense that the distribution of X is uniquely determined, then (1.20) uniquely determines $\{\nu_t\}$ among solutions satisfying the integrability condition (1.21).

The standard approach of adding a "time" component to the state of the process allows us to extend Theorem 1.11 to time inhomogeneous processes and also relax the integrability condition (1.21).

Corollary 1.12 *Let E be a complete, separable metric space. For $t\geq 0$, let $\widehat{A}_t,\widehat{B}_t : \mathcal{D}\subset\overline{C}(E)\to M(E)$. Assume that there exist a complete, separable, metric space U, operators $A,B:\mathcal{D}\to C(E\times[0,\infty)\times U)$, satisfying Condition 1.2 with x replaced by (x,t), and transition functions η_0 and η_1 from $E\times[0,\infty)$ to U such that for each $t\geq 0$,*

$$\widehat{A}_t f(x) = \int_U Af(x,t,u)\,\eta_0(x,t,du), \qquad \widehat{B}_t f(x) = \int_U Bf(x,t,u)\,\eta_1(x,t,du), \qquad \forall f\in\mathcal{D}.$$

Suppose $\{\nu_t : t \geq 0\}$ is a measurable $\mathcal{P}(E)$-valued function and μ is a measure on $E \times [0,\infty)$ such that for each $t > 0$, $\mu(E \times [0,t]) < \infty$,

$$\int_{E\times[0,t]}\int_U \psi_A(x,s,u)\,\eta_0(x,s,du)\nu_s(dx)ds+\int_{E\times[0,t]}\int_U \psi_B(x,u)\,\eta_1(x,du)\mu(dx\times ds) < \infty, \quad (1.22)$$

and

$$\int_E f d\nu_t = \int_E f d\nu_0 + \int_0^t \int_E \widehat{A}_s f d\nu_s ds + \int_{E\times[0,t]} \widehat{B}_s f(x)\mu(dx\times dx)\ , \quad f \in \mathcal{D}. \qquad (1.23)$$

Then there exists a solution (X,Γ) of the singular martingale problem for $(\widehat{A},\widehat{B},\nu_0)$, that is, there exists a filtration $\{\mathcal{F}_t\}$ such that

$$f(X(t)) - f(X(0)) - \int_0^t \widehat{A}_s f(X(s))ds - \int_{E\times[0,t]} \widehat{B}_s f(x)\Gamma(dx\times ds)$$

is a $\{\mathcal{F}_t\}$-martingale for each $f \in \mathcal{D}$, such that for each $t \geq 0$, $X(t)$ has distribution ν_t and $E[\Gamma] = \mu$.

If uniqueness holds for the martingale problem for $(\widehat{A},\widehat{B},\nu_0)$ in the sense that the distribution of X is uniquely determined, then (1.23) uniquely determines $\{\nu_t\}$ among solutions satisfying the integrability condition (1.22).

Proof. Let $\beta(t) > 0$ and define $\tau : [0,\infty) \to [0,\infty)$ so that

$$\int_0^{\tau(t)} \frac{1}{\beta(s)} ds = t,$$

that is, $\dot{\tau}(t) = \beta(\tau(t))$. Defining $\widehat{\nu}_t = \nu_{\tau(t)}$ and $\widehat{\mu}$ so that

$$\int_{E\times[0,t]} \beta(\tau(s))h(x,\tau(s))\widehat{\mu}(dx\times ds) = \int_{E\times[0,\tau(t)]} h(x,s)\mu(dx\times ds),$$

we have

$$\int_E f d\widehat{\nu}_t = \int_E f d\widehat{\nu}_0 + \int_0^t \int_E \beta(\tau(s))\widehat{A}_{\tau(s)} f d\widehat{\nu}_s ds + \int_{E\times[0,t]} \beta(\tau(s))\widehat{B}_{\tau(s)} f(x)\widehat{\mu}(dx\times dx)\ , \quad f \in \mathcal{D}.$$

Note also that β can be selected so that $\tau(t) \to \infty$ slowly enough to give

$$\begin{aligned}
&\int_0^\infty e^{-t}\Big[\int_{E\times[0,\tau(t)]}\int_U \psi_A(x,s,u)\,\eta_0(x,s,du)\nu_s(dx)ds \\
&\qquad + \int_{E\times[0,\tau(t)]}\int_U \psi_B(x,s,u)\,\eta_1(x,s,du)\mu(dx\times ds)dt \\
&= \int_{E\times[0,\infty)} e^{-s}\beta(\tau(s))\int_U \psi_A(x,\tau(s),u)\,\eta_0(x,\tau(s),du)\widehat{\nu}_s(dx)ds \\
&\qquad + \int_{E\times[0,\infty)} e^{-s}\beta(\tau(s))\int_U \psi_B(x,\tau(s),u)\,\eta_1(x,\tau(s),du)\widehat{\mu}(dx\times ds) \\
&< \infty.
\end{aligned}$$

It follows that $\{\widehat{\nu}_t\}$ and $\widehat{\mu}$ satisfy (1.21) for $\widehat{\psi}_A(x,s,u) = \beta(\tau(s))\psi_A(x,\tau(s),u)$ and $\widehat{\psi}_B(x,s,u) = \beta(\tau(s))\psi_B(x,\tau(s),u)$. Note also that if

$$f(\widehat{X}(t)) - f(\widehat{X}(0)) - \int_0^t \beta(\tau(s)\widehat{A}_{\tau(s)}f(\widehat{X}(s))ds - \int_{E\times[0,t]} \beta(\tau(s))\widehat{B}_{\tau(s)}f(x)\widehat{\Gamma}(dx\times ds)$$

is a $\{\mathcal{F}_t\}$-martingale for each $f \in \mathcal{D}$, $\widehat{X}(t)$ has distribution $\widehat{\nu}_t$, and $E[\widehat{\Gamma}] = \widehat{\mu}$, then (X,Γ) given by $X(t) \equiv \widehat{X}(\tau^{-1}(t))$ and $\Gamma(G\times[0,t]) = \int_0^{\tau^{-1}(t)} \beta(\tau(s))\widehat{\Gamma}(G\times ds)$ is a solution of the martingale problem for $(\widehat{A},\widehat{B},\nu_0)$, $X(t)$ has distribution ν_t and $E[\Gamma] = \mu$.

For simplicity, we assume that we can take $\beta \equiv 1$ in the above discussion. Let $\mathcal{D}_0$ be the collection of continuously differentiable functions with compact support in $[0,\infty)$. For $\gamma \in \mathcal{D}_0$ and $f \in \mathcal{D}$, define $\tilde{A}$ and $\tilde{B}$ by

$$\tilde{A}(\gamma f)(x,s) = \gamma(s)\widehat{A}_s f(x) + f(x)\gamma'(s), \quad \tilde{B}(\gamma f)(x,s) = \gamma(s)B_s f(x),$$

and define $\tilde{\nu}_t(dx\times ds) = \delta_t(ds)\nu_t(dx)$ and $\tilde{\mu}(dx\times ds\times dt) = \delta_t(ds)\mu(dx\times dt)$. It follows that

$$\int_{E\times[0,\infty)} \gamma f d\tilde{\nu}_t = \int_{E\times[0,\infty)} \gamma f \tilde{\nu}_0 + \int_0^t \int_{E\times[0,\infty)} \tilde{A}(\gamma f)d\tilde{\nu}_s ds + \int_{E\times[0,\infty)\times[0,t]} \tilde{B}(\gamma f)d\tilde{\mu}$$
$$\gamma \in \mathcal{D}_0, f \in \mathcal{D}.$$

Applying Theorem 1.11 with $\widehat{A}$ and $\widehat{B}$ replaced by $\tilde{A}$ and $\tilde{B}$ gives the desired result. $\square$

The results in the literature for models without the singular term B have had a variety of applications including an infinite dimensional linear programming formulation of stochastic control problems [1, 21, 28], uniqueness for filtering equations [3, 5, 20], uniqueness for martingale problems for measure-valued processes [9], and characterization of Markov functions (that is, mappings of a Markov process under which the image is still Markov) [19]. We anticipate a similar range of applications for the present results. In particular, in a separate paper, we will extend the results on the linear programming formulation of stochastic control problems to models with singular controls. A preliminary version of these results applied to queueing models is given in [22].

The paper is organized as follows. Properties of the measure Γ (or more precisely, the nonadapted precurser of Γ) are discussed in Section 2. A generalization of the existence theorem *without the singular operator* B is given in Section 3. Theorem 1.7 is proved in Section 4, using the results of Section 3. Theorem 1.11 is proved in Section 5.

2 Properties of Γ

Theorems 1.7 and 1.11 say very little about the random measure Γ that appears in the solution of the martingale problem other than to relate its expectation to the measures μ_1 and μ. The solution, however, is constructed as a limit of approximate solutions, and under various conditions, a more careful analysis of this limit reveals a great deal about Γ.

Essentially, the approximate solutions X_n are obtained as solutions of regular martingale problems corresponding to operators of the form

$$C_n f(x) = \int_U \beta_0^n(x) Af(x,u)\eta_0(x,du) + \int_U n\beta_1^n(x) Bf(x,u)\eta_1(x,du),$$

where η_0 and η_1 are defined in Theorem 1.7 and β_0^n and β_1^n are defined as follows: For $n > 1$, let $\mu_n^E = K_n^{-1}(\mu_0^E + \frac{1}{n}\mu_1^E) \in \mathcal{P}(E)$, where $K_n = \mu_0^E(E) + \frac{1}{n}\mu_1^E(E)$. Noting that μ_0^E and μ_1^E are absolutely continuous with respect to μ_n^E, we define

$$\beta_0^n = \frac{d\mu_0^E}{d\mu_n^E} \text{ and } \beta_1^n = \frac{1}{n}\frac{d\mu_1^E}{d\mu_n^E}\ ,$$

which makes $\beta_0^n + \beta_1^n = K_n$.

Remark 2.1 *In many examples (e.g., the stationary distribution for a reflecting diffusion), μ_0 and μ_1 will be mutually singular. In that case, $\beta_0^n = K_n$ on the support of μ_0 and $\beta_1^n = K_n$ on the support of μ_1. We do not, however, require μ_0 and μ_1 to be mutually singular.*

It follows that

$$\int_E C_n f d\mu_n^E = 0, \quad f \in \mathcal{D},$$

and the results of Section 3 give a stationary solution X_n of the martingale problem for C_n, where X_n has marginal distribution μ_n^E.

The proofs of the theorems in the generality they are stated involves the construction of an abstract compactification of E. In this section, we avoid that technicality by assuming that E is already compact or that we can verify a compact containment condition for $\{X_n\}$. Specifically, we assume that for each $\epsilon > 0$ and $T > 0$, there exists a compact set $K_{\epsilon,T} \subset E$ such that

$$\inf_n P\{X_n(t) \in K_{\epsilon,T}, t \leq T\} \geq 1 - \epsilon. \tag{2.1}$$

Set

$$\Gamma_n(H \times [0,t]) = \int_0^t n\beta_1^n(X_n(s)) I_H(X_n(s))ds,$$

and observe that

$$E[\Gamma_n(H \times [0,t])] = \mu_1^E(H)t.$$

Then $\{(X_n, \Gamma_n)\}$ is relatively compact, in an appropriate sense (see the proof of Theorem 1.7), and any limit point (X, Γ^*) is a solution of the singular, controlled martingale problem. Since Γ^* need not be $\{\overline{\mathcal{F}}_t^X\}$-adapted, the Γ of Theorem 1.7 is obtained as the dual predictable projection of Γ^*. (See Lemma 6.1.)

To better understand the properties of Γ^*, we consider a change of time given by

$$\int_0^{\tau_n(t)} (\beta_0^n(X_n(s)) + n\beta_1^n(X_n(s)))ds = t.$$

Note that since $\beta_0^n + \beta_1^n = K_n$, $\tau_n(t) \leq t/K_n$. Define

$$\gamma_0^n(t) = \int_0^{\tau_n(t)} \beta_0^n(X_n(s))ds \text{ and } \gamma_1^n(t) = \int_0^{\tau_n(t)} n\beta_1^n(X_n(s))ds.$$

Define

$$\widehat{A}f(x) = \int_U Af(x,u)\eta_0(x,du), \quad \widehat{B}f(x) = \int_U Bf(x,u)\eta_1(x,du),$$

and set $Y_n = X_n \circ \tau_n$. Then

$$f(Y_n(t)) - f(Y_n(0)) - \int_0^t \widehat{A}f(Y_n(s))d\gamma_0^n(s) - \int_0^t \widehat{B}f(Y_n(s))d\gamma_1^n(s) \tag{2.2}$$

is a martingale for each $f \in \mathcal{D}$. Since $\gamma_0^n(t) + \gamma_1^n(t) = t$, the derivatives $\dot{\gamma}_0^n$ and $\dot{\gamma}_1^n$ are both bounded by 1. It follows that $\{Y_n, \gamma_0^n, \gamma_1^n)\}$ is relatively compact in the Skorohod topology. (Since $\{Y_n\}$ satisfies the compact containment condition and γ_0^n and γ_1^n are uniformly Lipschitz, relative compactness follows by Theorems 3.9.1 and 3.9.4 of [11].)

We can select a subsequence along which (X_n, Γ_n) converges to (X, Γ^*) and $(Y_n, \gamma_0^n, \gamma_1^n)$ converges to a process (Y, γ_0, γ_1). Note that, in general, X_n does not converge to X in the Skorohod topology. (The details are given in Section 4.) In fact, one way to describe the convergence is that $(X_n \circ \tau_n, \tau_n) \Rightarrow (Y, \gamma_0)$ in the Skorohod topology and $X = Y \circ \gamma_0^{-1}$. The nature of the convergence is discussed in [17], and the corresponding topology is given in [13]. In particular, the finite dimensional distributions of X_n converge to those of X except for a countable set of time points.

Theorem 2.2 *Let (X, Γ^*) and (Y, γ_0, γ_1) be as above. Then*

a) (X, Γ^*) is a solution of the singular, controlled martingale problem for (A, B).

b) X is stationary with marginal distribution μ_0^E, and Γ^* has stationary increments with $E[\Gamma^*(\cdot \times [0,t]) = t\mu_1^E(\cdot)$.

c) $\lim_{t\to\infty} \gamma_0(t) = \infty$ a.s.

d) Setting $\gamma_0^{-1}(t) = \inf\{u : \gamma_0(u) > t\}$,

$$X = Y \circ \gamma_0^{-1} \tag{2.3}$$

and

$$\Gamma^*(H \times [0,t]) = \int_0^{\gamma_0^{-1}(t)} I_H(Y(s))d\gamma_1(s). \tag{2.4}$$

e) $E[\int_0^t I_H(Y(s))d\gamma_1(s)] \le t\mu_1^E(H)$, and if K_1 is the closed support of μ_1^E, then γ_1 increases only when Y is in K_1, that is,

$$\int_0^t I_{K_1}(Y(s))d\gamma_1(s) = \gamma_1(t) \quad a.s. \tag{2.5}$$

f) If γ_0^{-1} is continuous (that is, γ_0 is strictly increasing), then

$$\Gamma^*(H \times [0,t]) = \int_0^t I_H(X(s))d\lambda(s), \tag{2.6}$$

where $\lambda = \gamma_1 \circ \gamma_0^{-1}$. Since Γ^* has stationary increments, λ will also.

Proof. By invoking the Skorohod representation theorem, we can assume that the convergence of $(X_n, \Gamma_n, X_n \circ \tau_n, \gamma_0^n, \gamma_1^n)$ is almost sure, in the sense that $X_n(t) \to X(t)$ a.s. for all but countably many t, $\Gamma_n \to \Gamma^*$ almost surely in $\mathcal{L}(E)$, and $(X_n \circ \tau_n, \gamma_0^n, \gamma_1^n) \to (Y, \gamma_0, \gamma_1)$ a.s. in $D_{E\times\mathbb{R}^2}[0,\infty)$. Parts (a) and (b) follow as in the Proof of Theorem 1.7 applying (2.1) to avoid having to compactify E.

Note that $K_n\tau_n(t) \geq \gamma_0^n(t)$ and

$$\begin{aligned} E[K_n\tau_n(t) - \gamma_0^n(t)] &= E[\int_0^{\tau_n(t)} (K_n - \beta_0^n(X_n(s)))ds] \\ &\leq E[\int_0^{t/K_n} \beta_1^n(X_n(s)))ds] \\ &= \frac{\mu_1^E(E)t}{K_n n} \to 0. \end{aligned}$$

Since $\gamma_0^n(t) + \gamma_1^n(t) = t$, for $t > T$,

$$\begin{aligned} (t-T)P\{K_n\tau_n(t) \leq T\} &\leq E[(t - \gamma_0^n(t))I_{\{K_n\tau_n(t)\leq T\}}] \\ &\leq E[\int_0^T n\beta_1^n(X_n(s))ds] \\ &= T\mu_1^E(E), \end{aligned}$$

and since $\gamma_0^n(t)$ and $K_n\tau_n(t)$ are asymptotically the same, we must have that

$$P\{\gamma_0(t) \leq T\} \leq \frac{T\mu_1^E(E)}{t-T}.$$

Consequently, $\lim_{t\to\infty} \gamma_0(t) = \infty$ a.s.

The fact that $X = Y \circ \gamma_0^{-1}$ follows from Theorem 1.1 of [17]. Let

$$\Gamma_n(g,t) = \int_E g(x)\Gamma_n(dx \times [0,t]) = \int_0^t n\beta_1^n(X_n(s))g(X_n(s))ds.$$

Then for bounded continuous g and all but countably many t,

$$\Gamma_n(g,t) \to \Gamma^*(g,t) = \int_E g(x)\Gamma^*(dx \times [0,t]) \quad a.s.$$

Since

$$\Gamma_n(g,\tau_n(t)) = \int_0^t g(X_n \circ \tau_n(s))d\gamma_1^n(s) \to \int_0^t g(Y(s))d\gamma_1(s) \quad a.s.,$$

Theorem 1.1 of [17] again gives

$$\Gamma^*(g,t) = \int_0^{\gamma_0^{-1}(t)} g(Y(s))d\gamma_1(s),$$

which implies (2.4).

Since $\gamma_0(t) \leq t$, $\gamma_0^{-1}(t) \geq t$, so

$$E[\int_0^t I_H(Y(s))d\gamma_1(s)] \leq E[\Gamma^*(H \times [0,t])] = t\mu_1^E(H),$$

and Part (e) follows.

The representation (2.6) follows immediately from (2.4). □

Lemma 2.3 *Let (Y, γ_0, γ_1) be as above. Then for each $f \in \mathcal{D}$,*

$$f(Y(t)) - f(Y(0)) - \int_0^t \widehat{A}f(Y(s))d\gamma_0(s) - \int_0^t \widehat{B}f(Y(s))d\gamma_1(s) \tag{2.7}$$

is a $\{\mathcal{F}_t^{Y,\gamma_0}\}$-martingale.

Proof. We show that (2.2) converges in distribution to (2.7). Then (2.7) can be shown to be a martingale by essentially the same argument used in the proof of Theorem 1.7. If $\widehat{A}f$ and $\widehat{B}f$ were continuous, then the convergence in distribution would be immediate. Let g be continuous. Then, recalling that $\tau_n(T) \leq T$,

$$\begin{aligned} & E\left[\sup_{t\leq T}\left|\int_0^t \widehat{A}f(Y_n(s))d\gamma_0^n(s) - \int_0^t g(Y_n(s))d\gamma_0^n(s)\right|\right] \\ & \leq E\left[\int_0^T \left|\widehat{A}f(Y_n(s)) - g(Y_n(s))\right| d\gamma_0^n(s)\right] \\ & \leq E\left[\int_0^{\tau_n(T)} \left|\widehat{A}f(X_n(s)) - g(X_n(s))\right| \beta_0^n(X_n(s))ds\right] \\ & \leq T\int_E |\widehat{A}f(x) - g(x)|\mu_0^E(dx). \end{aligned} \tag{2.8}$$

The right side can be made arbitrarily small by selecting the appropriate $g \in \overline{C}(E)$. Note that for any nonnegative, bounded continuous function h,

$$E\left[\int_0^T h(Y(s))d\gamma_0(s)\right] = \lim_{n\to\infty} E\left[\int_0^T h(Y_n(s))d\gamma_0^n(s)\right] \leq T\int_E h(x)\mu_0^E(dx),$$

and the inequality between the left and right sides extends to all nonnegative measurable h. It follows that

$$E\left[\sup_{t\leq T}\left|\int_0^t \widehat{A}f(Y(s))d\gamma_0(s) - \int_0^t g(Y(s))d\gamma_0(s)\right|\right] \leq T\int_E |\widehat{A}f(x) - g(x)|\mu_0^E(dx),$$

and the convergence of (2.2) to (2.7) follows. □

In general, γ_0^{-1} need not be continuous. Continuity of γ_0^{-1} is equivalent to γ_0 being strictly increasing. The following lemma, which is a simple extension of Lemma 6.1.6 of [15], gives conditions for γ_0 to be strictly increasing. We say that (Z, ζ) is a solution of the *stopped* martingale problem for an operator C if there exists a filtration $\{\mathcal{F}_t\}$ such that Z is $\{\mathcal{F}_t\}$-adapted, ζ is an $\{\mathcal{F}_t\}$-stopping time, and for each $f \in \mathcal{D}(C)$,

$$f(Z(t \wedge \zeta)) - f(Z(0)) - \int_0^{t\wedge\zeta} Cf(Z(s))ds$$

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

Lemma 2.4 *Let K_1 be the closed support of μ_1^E. Suppose that every solution (Z, ζ) of the stopped martingale problem for $\widehat{B}$ satisfies*

$$\zeta \wedge \inf\{t : Z(t) \notin K_1\} = 0 \quad a.s. \tag{2.9}$$

Then γ_0 is strictly increasing.

Remark 2.5 *In the case of reflecting diffusions in a domain D (Example 1.4), $K_1 = \partial D$, $\widehat{B}f(x) = m(x) \cdot \nabla f(x)$, and any solution of the stopped martingale problem for $\widehat{B}$ satisfies*

$$Z(t \wedge \zeta) = Z(0) + \int_0^{t\wedge\zeta} m(Z(s))ds.$$

Results on solutions of ordinary differential equations can then be used to verify (2.9).

Proof. For $t_0 \geq 0$, let $\zeta_0 = \inf\{t > t_0 : \gamma_0(t) > \gamma_0(t_0)$. Either γ_0 is a.s. strictly increasing or there exists t_0 such that $P\{\zeta_0 > t_0\} > 0$. For such a t_0, define $Z(t) = Y(t_0 + t)$ and $\zeta = \zeta_0 - t_0$. Then, since γ_0 is constant on $[t_0, \zeta_0]$, and hence $d\gamma_1(s) = ds$ on $[t_0, \zeta_0]$,

$$f(Z(t \wedge \zeta)) - f(Z(0)) - \int_0^{t\wedge\zeta} \widehat{B}f(Z(s))ds$$

is an $\{\mathcal{F}_{t_0+t}^{Y,\gamma_0}\}$-martingale. In particular, (Z, ζ) is a solution of the stopped martingale problem for $\widehat{B}$. Since, with probability one, γ_1 increases only when $Y \in K_1$, and by assumption $\zeta \wedge \inf\{t : Z(t) \notin K_1\} = 0$ a.s., it follows that $\zeta = 0$ a.s. contradicting the assumption that $P\{\zeta_0 > t_0\} > 0$. $\square$

Theorem 2.2 and Lemma 2.4 give a good description of Γ^* for many interesting examples in which Γ^* is continuous. The next result addresses examples in which the natural version of Γ^* is discontinuous.

Theorem 2.6 *Suppose that $Bf(x,u) = \alpha(x,u)\int_E (f(z) - f(x))q(x,u,dz)$, where $0 \leq \alpha(x,u) \leq \sup_{z,v} \alpha(z,v) < \infty$, and q is a transition function from $E \times U$ to E. Define $\widehat{\alpha} : E \to [0,\infty)$ and $\widehat{q}$, a transition function on E, so that*

$$\widehat{B}f(x) = \int_U Bf(x,u)\eta_1(x,du) = \widehat{\alpha}(x)\int_E (f(z) - f(x))\widehat{q}(x,dz).$$

(In particular, $\widehat{\alpha}(x) = \int_U \alpha(x,u)\eta_1(x,du)$.) Then there exist $(X, \Gamma^*, Y, \gamma_0, \gamma_1)$ satisfying (2.3) and (2.4) and a counting process N such that

a) (X,Γ^*) *is a solution of the singular, controlled martingale problem for* (A,B).

b) X *is stationary with marginal distribution* μ_0^E.

c) Γ^* *has stationary increments, and for each* $t\geq 0$, $E[\Gamma^*(\cdot\times[0,t])]=t\mu_1(\cdot)$.

d) There exists a filtration $\{\mathcal{G}_t\}$ *such that* (Y,γ_0,γ_1,N) *is adapted to* $\{\mathcal{G}_t\}$,

$$\tilde{N}(t)\equiv N(t)-\int_0^t\widehat{\alpha}(Y(s))d\gamma_1(s) \tag{2.10}$$

is a $\{\mathcal{G}_t\}$*-martingale, and for each* $f\in\mathcal{D}$,

$$f(Y(t))-f(Y(0))-\int_0^t\widehat{A}f(Y(s))d\gamma_0(s)-\int_0^t\widehat{B}f(Y(s))d\gamma_1(s) \tag{2.11}$$

and

$$f(Y(t))-f(Y(0))-\int_0^t\widehat{A}f(Y(s))d\gamma_0(s)-\int_0^t\int_E(f(z)-f(Y(s-))\widehat{q}(Y(s-),dz)dN(s) \tag{2.12}$$

are $\{\mathcal{G}_t\}$*-martingales.*

e) Letting K_1 *be the closed support of* μ_1^E,

$$\int_0^t I_{K_1}(Y(s-))dN(s)=N(t)\quad a.s.,\quad t\geq 0.$$

Proof. Let $\tilde{E}=E\times\{-1,1\}$ and $\tilde{\mathcal{D}}=\{f:f(x,\theta)=f_1(x)f_2(\theta),f_1\in\mathcal{D},f_2\in B(\{-1,1\})\}$. For $f\in\tilde{\mathcal{D}}$, define

$$\tilde{A}f(x,\theta,u)=f_2(\theta)Af_1(x,u)$$

and

$$\tilde{B}f(x,\theta,u)=\alpha(x,u)\int_E(f_1(z)f_2(-\theta)-f_1(x)f_2(\theta))q(x,u,dz).$$

Let

$$\tilde{\mu}_0(dx\times d\theta\times du)=\mu_0(dx\times du)\times(\frac{1}{2}\delta_{-1}(d\theta)+\frac{1}{2}\delta_1(d\theta))$$

and

$$\tilde{\mu}_1(dx\times d\theta\times du)=\mu_1(dx\times du)\times(\frac{1}{2}\delta_{-1}(d\theta)+\frac{1}{2}\delta_1(d\theta)).$$

Then

$$\int_{\tilde{E}}\tilde{A}fd\tilde{\mu}_0+\int_{\tilde{E}}\tilde{B}fd\tilde{\mu}_1=0,\quad f\in\tilde{\mathcal{D}},$$

and $\tilde{A}$ and $\tilde{B}$ satisfy Condition 1.2 with

$$\tilde{\psi}_A(x,\theta,u)=\psi_A(x,u),\quad \tilde{\psi}_B(x,\theta,u)=\psi_B(x,u).$$

By Theorems 1.7 and 2.2, there exist $(X, \Theta, \tilde{\Gamma}^*)$ and $(Y, \Phi, \gamma_0, \gamma_1)$ satisfying $X = Y \circ \gamma_0^{-1}$ and $\Theta = \Phi \circ \gamma_0^{-1}$ and a filtration $\{\mathcal{G}_t\}$ such that for each $f \in \tilde{\mathcal{D}}$,

$$f(X(t), \Theta(t)) - \int_0^t \int_U \tilde{A} f(X(s), \Theta(s), u) \eta_0(X(s), du) ds \\ - \int_{\tilde{E} \times [0,t]} \int_U \tilde{B} f(x, \theta, u) \eta_1(x, du) \tilde{\Gamma}^*(dx \times d\theta \times ds)$$

is a $\{\mathcal{G}_{\gamma_0^{-1}(t)}\}$-martingale and

$$f(Y(t), \Phi(t)) - \int_0^t \int_U \tilde{A} f(Y(s), \Phi(s), u) \eta_0(Y(s), du) d\gamma_0(s) \\ - \int_0^t \int_U \tilde{B} f(Y(s), \Phi(s), u) \eta_1(Y(s), du) d\gamma_1(s) \tag{2.13}$$

is a $\{\mathcal{G}_t\}$-martingale, (X, Θ) is stationary with marginal distribution $\tilde{\mu}_0^{\tilde{E}}$ (and hence, X has marginal distribution μ_0^E), and $\tilde{\Gamma}^*$ has stationary increments and satisfies $E[\tilde{\Gamma}^*(H_1 \times H_2 \times [0,t])] = t\mu_1(H_1)(\frac{1}{2}\delta_{-1}(H_2) + \frac{1}{2}\delta_1(H_2))$. Parts (a), (b), and (c) follow by taking $f(x, \theta)$ to depend only on x.

For f depending only on θ, we have that

$$f(\Phi(t)) - \int_0^t \widehat{\alpha}(Y(s))(f(-\Phi(s)) - f(\Phi(s))) d\gamma_1(s)$$

is a $\{\mathcal{G}_t\}$-martingale. Let $\beta(t) = \inf\{r : \int_0^r \widehat{\alpha}(Y(s)) d\gamma_1(s) > t\}$, for $0 \leq t \leq \overline{\beta} = \int_0^\infty \widehat{\alpha}(Y(s)) d\gamma_1(s)$. It follows that $(\Phi \circ \beta, \overline{\beta})$ is a solution of the stopped martingale problem for $Cf(\theta) = (f(-\theta) - f(\theta))$. Since the martingale problem for C is well-posed, it follows that $\Phi \circ \beta$ can be extended to a solution Ψ of the martingale problem for C (see Lemma 4.5.16 of [11]), and we can write

$$\Phi(t) = \Psi\left(\int_0^t \widehat{\alpha}(Y(s)) d\gamma_1(s)\right).$$

But $\Psi(t) = \Phi(0)(-1)^{N_0(t)}$, where N_0 is a unit Poisson process, so $\Phi(t) = \Phi(0)(-1)^{N(t)}$, where

$$N(t) = N_0\left(\int_0^t \widehat{\alpha}(Y(s)) d\gamma_1(s)\right).$$

Note that N is $\{\mathcal{G}_t\}$-adapted and that (2.10) is a $\{\mathcal{G}_t\}$-martingale. Since (2.11) is a martingale by Lemma 2.3, and the difference of (2.11) and (2.12) is

$$\int_0^t \int_E (f(z) - f(Y(s-)) \widehat{q}(Y(s-), dz) d(N(s) - \int_0^s \widehat{\alpha}(Y(r)) d\gamma_1(r)),$$

it follows that (2.12) is a martingale.

By Theorem 2.2, Part (e),

$$\int_0^t \widehat{\alpha}(Y(s)) d\gamma_1(s) = \int_0^t \widehat{\alpha}(Y(s)) I_{K_1}(Y(s)) d\gamma_1(s) \quad a.s.$$

Define

$$\tilde{N}(t) = N(t) - \int_0^t \widehat{\alpha}(Y(r))d\gamma_1(r).$$

Then, with probability one,

$$N(t) - \int_0^t I_{K_1}(Y(s-))dN(s) = N(t) - \int_0^t \widehat{\alpha}(Y(s))d\gamma_1(s) - \int_0^t I_{K_1}(Y(s-))d\tilde{N}(s),$$

and since the right side is a local martingale, the left side must be zero. □

In the context of Theorem 2.6, the right analog of Lemma 2.4 would be a condition that implies $N \circ \gamma_0^{-1}$ is still a counting process.

Lemma 2.7 *Let K_1 be the closed support of μ_1^E, and suppose that for each $x \in K_1$, $\widehat{q}(x, K_1) = 0$. Let $\sigma_1, \sigma_2 \ldots$ be the jump times of N. Then $P\{\sigma_{k+1} < \infty, \gamma_0(\sigma_k) = \gamma_0(\sigma_{k+1})\} = 0$, and hence, $N \circ \gamma_0^{-1}$ is a counting process.*

Proof. Since $\int_0^t I_{K_1}(Y(s-))dN(s) = N(t)$ and $\gamma_0(t+r) - \gamma_0(t) \geq \int_t^{t+r} I_{K_1^c}(Y(s))ds$, it is enough to show that $\int_0^t I_{K_1^c}(Y(s))dN(s) = N(t)$, that is, that every boundary jump lands inside the open set K_1^c, and hence that Y is in K_1^c for a positive time interval after each boundary jump. Let M_Φ denote the martingale

$$\Phi(t) + \int_0^t 2\widehat{\alpha}(Y(s))\Phi(s)ds$$

and M_f denote the martingale (2.11). Then

$$[M_\Phi, M_f]_t = -\int_0^t 2\Phi(s-)(f(Y(s)) - f(Y(s-))dN(s)$$

and, using the fact that (2.13) is a martingale,

$$\langle M_\Phi, M_f\rangle_t = \int_0^t 2\widehat{\alpha}(Y(s))\Phi(s)(f(Y(s)) - \int_E f(z)\widehat{q}(Y(s), dz))d\gamma_1(s).$$

Since $[M_\Phi, M_f]_t - \langle M_\Phi, M_f\rangle_t$ is a martingale, it follows that

$$\begin{aligned}&\int_0^t 2\Phi(s-)(\int_E f(Y(s)) - f(z)\widehat{q}(Y(s-), dz)dN(s)\\ &\quad = \langle M_\Phi, M_f\rangle_t - [M_\Phi, M_f]_t + \int_0^t 2\Phi(s-)(\int_E f(z)\widehat{q}(Y(s-), dz) - f(Y(s-))d\tilde{N}(s)\end{aligned}$$

is a martingale, and integrating against Φ,

$$\int_0^t 2(f(Y(s)) - \int_E f(z)\widehat{q}(Y(s-), dz)dN(s) \tag{2.14}$$

is a martingale for every $f \in \mathcal{D}$. But the collection of f for which (2.14) is a martingale is closed under bounded pointwise convergence, so is all of $B(E)$. Taking $f = I_{K_1^c}$, we have that

$$2\int_0^t (I_{K_1^c}(Y(s)) - 1)dN(s)$$

is a martingale, but, since the integrand is non positive, that can hold only if the integral is identically zero and hence

$$\int_0^t I_{K_1^c}(Y(s))dN(s) = N(t).$$

□

3 Stationary solutions of controlled martingale problems

The objective of this section is to establish the existence of a particular form of stationary solution for the controlled martingale problem for a generator A. The formulation is obtained by taking $Bf \equiv 0$ for each $f \in \mathcal{D}$ above, so we drop any reference to B. We also denote μ_0 of (1.7) by μ and ψ_A by ψ, since there will not be a μ_1 or a ψ_B.

The first result of this type was by Echeverria [10] in the context of an uncontrolled Markov process (see also Ethier and Kurtz [11, Theorem 4.9.15]). Stockbridge [27] extended the result to controlled processes. In [27], the state and control spaces were locally compact, complete, separable, metric spaces and the control process was only shown to be adapted to the past of the state process. Bhatt and Karandikar [2] removed the local compactness assumption (on the state space) for uncontrolled processes. The stationary control process was shown to be a feedback control of the current state of the process (where the particular control is determined from the stationary measure) by Kurtz and Stockbridge [21] and Bhatt and Borkar [1]. Kurtz and Stockbridge also established this result for generators whose range consisted of bounded, measurable (not necessarily continuous) functions. The results were proved by Kurtz and Stockbridge under the assumption that the state and control spaces are locally compact and by Bhatt and Borkar under the assumption that the state space E is a complete, separable metric space and that the control space U is compact.

Here we make certain that the results are valid if both the state and control spaces are complete, separable metric spaces. Many of the recent proofs simply refer back to previous results when needed. In this section, we compile the previous results and provide complete details.

Suppose μ is a probability measure on $E \times U$ with

$$\mu(\mathcal{U}) = 1 \tag{3.1}$$

and which satisfies

$$\int_{E\times U} Af(x,u)\, \mu(dx \times du) = 0, \quad \forall f \in \mathcal{D}. \tag{3.2}$$

Denote the state marginal by $\mu_E = \mu(\cdot \times U)$, and let η be the regular conditional distribution of u given x, that is, η satisfies

$$\mu(H_1 \times H_2) = \int_{H_1} \eta(x, H_2)\, \mu_E(dx), \quad H_1 \in \mathcal{B}(E), H_2 \in \mathcal{B}(U). \tag{3.3}$$

Implicit in (3.3) is the requirement that $\eta(x, U_x) = 1$ *a.e.* $\mu_E(dx)$.

Our goal is to show that there exists a stationary process X such that the $E \times \mathcal{P}(U)$-valued process $(X, \eta(X,\cdot))$ is a stationary, relaxed solution of the controlled martingale problem for

(A, μ_E). Note that if X is a stationary process with $X(0)$ having distribution μ_E, the pair $(X, \eta(X, \cdot))$ is stationary and the one-dimensional distributions satisfy

$$E[I_{H_1}(X(t))\,\eta(X(t), H_2)] = \mu(H_1 \times H_2), \quad t \geq 0.$$

Following Bhatt and Karandikar [2], we construct an embedding of the state space E in a compact space $\widehat{E}$. Without loss of generality, we can assume that $\{g_k\}$ in the separability condition 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

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

Note that $\widehat{E}$ is compact. Define $G : E \to \widehat{E}$ by

$$G(x) = (g_1(x), g_2(x), \ldots). \tag{3.4}$$

Then G has a measurable inverse defined on the (measurable) set $G(E)$. In this section and the next, we will typically denote measures on E by μ, $\tilde{\mu}$, μ_0, μ_1, etc. and the corresponding measures on $\widehat{E}$ by ν, $\tilde{\nu}$, ν_0, ν_1, etc.

We will need the following lemmas.

Lemma 3.1 *Let $\mu_0 \in \mathcal{P}(E)$. Then there exists a unique measure $\nu_0 \in \mathcal{P}(\widehat{E})$ satisfying $\int_E g_k d\mu_0 = \int_{\widehat{E}} z_k \nu_0(dz)$. In particular, if Z has distribution ν_0, then $G^{-1}(Z)$ has distribution μ_0.*

Proof. Existence is immediate: take $\nu_0 = \mu_0 G^{-1}$. Since $\widehat{E}$ is compact, $\{\prod_{i \in I} z_i : I \in \mathcal{I}\}$ is separating. Consequently, uniqueness follows from the fact that

$$\int_{\widehat{E}} \prod_{i \in I} z_i \nu_0(dz) = \int_{\widehat{E}} z_{k(I)} \nu_0(dz) = \int_E g_{k(I)} d\mu_0.$$

□

Lemma 3.2 *Let $C \subset B(E) \times M(E)$ be a pre-generator. Suppose that φ is continuously differentiable and convex on $D \subset R^m$, that $f_1, \ldots, f_m \in \mathcal{D}(C)$ and $(f_1, \ldots, f_m) : E \to D$, and that $(\varphi(f_1, \ldots, f_m), h) \in C$. Then*

$$h \geq \nabla\varphi(f_1, \ldots, f_m) \cdot (Cf_1, \ldots, Cf_m).$$

Proof. Since C is a pre-generator, there exist λ_n and μ_n such that

$$\begin{aligned} h(x) &= \lim_{n\to\infty} \lambda_n(x) \int_E (\varphi(f_1(y), \ldots, f_m(y)) - \varphi(f_1(x), \ldots, f_m(x))\mu_n(x, dy) \\ &\geq \lim_{n\to\infty} \nabla\varphi(f_1(x), \ldots, f_m(x)) \cdot \lambda_n(x) \int_E (f_1(y) - f_1(x), \ldots, f_m(y) - f_m(x))\mu_n(x, dy) \\ &= \nabla\varphi(f_1(x), \ldots, f_m(x)) \cdot (Cf_1(x), \ldots, Cf_m(x)). \end{aligned}$$

□

Lemma 3.3 *Let X_n, X be processes in $D_E[0,\infty)$ with $X_n \Rightarrow X$, and let $D_X = \{t : P\{X(t) \neq X(t-)\} > 0\}$. Suppose for each $t \geq 0$, $X_n(t)$ and $X(t)$ have a common distribution $\mu_t \in \mathcal{P}(E)$. Let g be Borel measurable on $[0,\infty) \times E$ and satisfy*

$$\int_0^t \int_E |g(s,x)|\mu_s(dx)ds < \infty$$

for each $t > 0$. Then

$$\int_0^{\cdot} g(s, X_n(s))\, ds \Rightarrow \int_0^{\cdot} g(s, X(s))\, ds \tag{3.5}$$

and, in particular, for each $m \geq 1, 0 \leq t_1 \leq \cdots \leq t_m < t_{m+1}$, $t_i \notin D_X$, and $h_1, \ldots, h_m \in \overline{C}(E)$,

$$\begin{aligned}\lim_{n\to\infty} E\left[\int_{t_m}^{t_{m+1}} g(s, X_n(s))\, ds \prod_{i=1}^m h_i(X_n(t_i))\right] \\ = E\left[\int_{t_m}^{t_{m+1}} g(s, X(s))\, ds \prod_{i=1}^m h_i(X(t_i))\right].\end{aligned} \tag{3.6}$$

Proof. See Kurtz and Stockbridge (1997), Lemma 2.1. Note that the proof there does not use the assumption that E is locally compact. □

Theorem 3.4 *Let A satisfy Condition 1.2. Suppose $\mu \in \mathcal{P}(E \times U)$ satisfies (3.1), (3.2) and*

$$\int \psi(x,u)\, \mu(dx \times du) < \infty, \tag{3.7}$$

and define μ_E and η by (3.3). Then there exists a stationary process X such that $(X, \eta(X,\cdot))$ is a stationary relaxed solution of the controlled martingale problem for (A, μ_E), $\eta(X,\cdot)$ is an admissible (absolutely continuous) control, and for each $t \geq 0$,

$$E[I_{H_1}(X(t))\eta(X(t), H_2)] = \mu(H_1 \times H_2) \tag{3.8}$$

for every $H_1 \in \mathcal{B}(E)$ and $H_2 \in \mathcal{B}(U)$.

Remark 3.5 *We will obtain X in the form $X = G^{-1}(Z)$. It will be clear from the proof that there always exists a modification of Z with sample paths in $D_{\widehat{E}}[0,\infty)$, but our assumptions do not imply that X will have sample paths in $D_E[0,\infty)$. For example, let $Af = (1+x^4)(f''(x) + f'(x))$. It is easy to check that $\mu(dx) = c(1+x^4)^{-1}dx$ satisfies $\int_R Af(x)\mu(dx) = 0$, but the corresponding process will repeatedly "go out" at $+\infty$ and "come back in" at $-\infty$.*

We first consider the case $\psi \equiv 1$.

Theorem 3.6 *Let A satisfy Condition 1.2 with $\psi \equiv 1$. Suppose $\mu \in \mathcal{P}(E \times U)$ satisfies (3.1) and (3.2), and define μ_E and η by (3.3). Then there exists a stationary process X such that $(X, \eta(X,\cdot))$ is a stationary relaxed solution of the controlled martingale problem for (A, μ_E) satisfying (3.8) and $\eta(X,\cdot)$ is an admissible absolutely continuous control.*

Proof. For $n = 1, 2, 3, \dots$, define the Yosida approximations A_n by

$$A_n g = n[(I - n^{-1}A)^{-1} - I]g$$

for $g \in \mathcal{R}(I - n^{-1}A)$, and note that for $f \in \mathcal{D}(A)$ and $g = (I - n^{-1}A)f$, $A_n g = Af$.
Let M be the linear subspace of functions of the form

$$\begin{aligned} F(x_1, x_2, u_1, u_2) &= \sum_{i=1}^{m} \left\{ h_i(x_1) \left[(I - n^{-1}A) f_i(x_2, u_1) + g_i(x_2, u_2) - g_i(x_2, u_1) \right] \right\} \\ &\qquad + h_0(x_2, u_1, u_2), \end{aligned} \tag{3.9}$$

where $h_1, \dots, h_m \in \overline{C}(E)$, $h_0 \in \overline{C}(E \times U \times U)$, $f_1, \dots, f_m \in \mathcal{D}(A)$, and $g_1, \dots, g_m \in \overline{C}(E \times U)$.
Define the linear functional Ψ on M by

$$\begin{aligned} \Psi F &= \int_{E\times U} \int_U \sum_{i=1}^{m} \{ h_i(x_2)[f_i(x_2) + g_i(x_2, u_2) - g_i(x_2, u_1)] \} \, \eta(x_2, du_2) \, \mu(dx_2 \times du_1) \\ &\quad + \int_{E\times U} \int_U h_0(x_2, u_1, u_2) \, \eta(x_2, du_2) \, \mu(dx_2 \times du_1) \\ &= \int_{E\times U} \int_U \left[\sum_{i=1}^{m} h_i(x_2) f_i(x_2) + h_0(x_2, u_1, u_2) \right] \eta(x_2, du_2) \, \mu(dx_2 \times du_1) \end{aligned} \tag{3.10}$$

in which the second representation follows from the fact that

$$\int_{E\times U} \int_U h(x_2)[g(x_2, u_2) - g(x_2, u_1)] \eta(x_2, du_2) \, \mu(dx_2 \times du_1) = 0 \tag{3.11}$$

(write $\mu(dx_2 \times du_1) = \eta(x_2, du_1)\mu_E(dx_2)$). Also define the linear operator

$$\Pi : B(E \times E \times U \times U) \to B(E \times E \times U)$$

by

$$\Pi F(x_1, x_2, u_1) = \int_U F(x_1, x_2, u_1, u_2) \, \eta(x_2, du_2) \tag{3.12}$$

and the functional p on $B(E \times E \times U \times U)$ by

$$p(F) = \int_{E\times U} \sup_{x_1} |\Pi F(x_1, x_2, u_1)| \mu(dx_2 \times du_1). \tag{3.13}$$

Observe that $\Pi(\Pi F) = \Pi F$ so

$$p(F - \Pi F) = 0. \tag{3.14}$$

In order to simplify notation, define the operator C on $\overline{C}(E \times U)$ by

$$Cg(x_2, u_1) = \int_U [g(x_2, u_2) - g(x_2, u_1)] \, \eta(x_2, du_2). \tag{3.15}$$

We claim $|\Psi F| \leq p(F)$. To verify this claim, fix $F \in M$. For $\alpha_i \geq ||(I-n^{-1}A)f_i+Cg_i||\vee||f_i||, i = 1,\ldots,m$, let ϕ be a polynomial on R^m that is convex on $\prod_{i=1}^m[-\alpha_i,\alpha_i]$. By the convexity of ϕ and Lemma 3.2

$$
\begin{aligned}
&\phi((I-n^{-1}A)f_1+Cg_1,\ldots,(I-n^{-1}A)f_m+Cg_m) \\
&\quad \geq \phi(f_1,\ldots,f_m) - n^{-1}\nabla\phi(f_1,\ldots,f_m)\cdot(Af_1,\ldots,Af_m) \\
&\qquad + \nabla\phi(f_1,\ldots,f_m)\cdot(Cg_1,\ldots,Cg_m) \\
&\quad \geq \phi(f_1,\ldots,f_m) - n^{-1}A\phi(f_1,\ldots,f_m) + \nabla\phi(f_1,\ldots,f_m)\cdot(Cg_1,\ldots,Cg_m).
\end{aligned}
$$

In light of (3.2) and (3.11), integration with respect to μ yields

$$
\int \phi((I-n^{-1}A)f_1+Cg_1,\ldots,(I-n^{-1}A)f_m+Cg_m)\,d\mu \geq \int \phi(f_1,\ldots,f_m)\,d\mu, \tag{3.16}
$$

and this inequality can be extended to arbitrary convex functions. Consider, in particular, the convex function $\phi(r_1,\ldots,r_m) = \sup_{x_1}\sum_{i=1}^m h_i(x_1)r_i$. It follows that

$$
\begin{aligned}
\Psi F &\leq \int_{E\times U}\left\{\sup_{x_1}\sum_{i=1}^m h_i(x_1)f_i(x_2) + \int_U h_0(x_2,u_1,u_2)\,\eta(x_2,du_2)\right\}\mu(dx_2\times du_1) \\
&= \int_{E\times U}\left\{\phi(f_1,\ldots,f_m)(x_2) + \int_U h_0(x_2,u_1,u_2)\,\eta(x_2,du_2)\right\}\mu(dx_2\times du_1) \\
&\leq \int_{E\times U}\left\{\phi((I-n^{-1}A)f_1+Cg_1,\ldots,(I-n^{-1}A)f_m+Cg_m)(x_2,u_1))\right. \\
&\qquad\qquad \left. + \int_U h_0(x_2,u_1,u_2)\,\eta(x_2,du_2)\right\}\mu(dx_2\times du_1) \\
&= \int_{E\times U}\sup_{x_1}\Pi F(x_1,x_2,u_1)\,\mu(dx_2\times du_1) \\
&\leq p(F),
\end{aligned}
$$

and $-\Psi F = \Psi(-F) \leq p(-F) = p(F)$, so $|\Psi F| \leq p(F)$.

As a result, we can apply the Hahn-Banach theorem to extend Ψ to the entire space $\overline{C}(E\times E\times U\times U)$, still satisfying $|\Psi F| \leq p(F)$ (see [23, p. 187]). Since $\Psi 1 = 1$, for $F \geq 0$,

$$
||F|| - \Psi F = \Psi(||F|| - F) \leq p(\|F\| - F) \leq ||F||, \tag{3.17}
$$

so $\Psi F \geq 0$. By the extension of the Riesz representation theorem in Theorem 2.3 of Bhatt and Karandikar [2], there exists a measure $\tilde{\mu} \in \mathcal{P}(E\times E\times U\times U)$ such that

$$
\Psi F = \int_{E\times E\times U\times U} F(x_1,x_2,u_1,u_2)\,\tilde{\mu}(dx_1\times dx_2\times du_1\times du_2). \tag{3.18}
$$

Considering F of the form $F(x_1,x_2,u_1,u_2) = h(x_1)(I-n^{-1}A)1(x_2,u_1)$, (1 being the constant function), we see that $\tilde{\mu}(\cdot\times E\times U\times U) = \mu_E(\cdot)$. Taking $F(x_1,x_2,u_1,u_2) = h(x_1)(I-$

$n^{-1}A)f(x_2,u_1)$ and writing $\tilde{\mu}(dx_1\times dx_2\times du_1\times du_2)=\tilde{\eta}(x_1,dx_2\times du_1\times du_2)\mu_E(dx_1)$, we have

$$\begin{aligned}
&\int_E h(x_1)f(x_1)\,\mu_E(dx_1)\\
&\quad=\Psi F\\
&\quad=\int_{E\times E\times U\times U} h(x_1)(I-n^{-1}A)f(x_2,u_1)\,\tilde{\mu}(dx_1\times dx_2\times du_1\times du_2)\\
&\quad=\int_E h(x_1)\left[\int_{E\times U}(I-n^{-1}A)f(x_2,u_1)\,\tilde{\eta}(x_1,dx_2\times du_1\times U)\right]\,\mu_E(dx_1).
\end{aligned}$$

Letting $\widehat{\eta}(x_1,dx_2\times du_1)=\tilde{\eta}(x_1,dx_2\times du_1\times U)$, it follows that

$$\int_{E\times U}(I-n^{-1}A)f(x_2,u_1)\widehat{\eta}(x_1,dx_2\times du_1)=f(x_1)\quad\text{a.e. }\mu_E(dx_1). \tag{3.19}$$

Observe that $\Psi F=\Psi(\Pi F)$ by (3.14) and the fact that $|\Psi F|\leq p(F)$. Again by considering $F(x_1,x_2,u_1,u_2)=f(x_1,x_2,u_1)g(u_2)$ and writing $\tilde{\mu}(dx_1\times dx_2\times du_1\times du_2)=\overline{\eta}(x_1,x_2,u_1,du_2)\overline{\mu}(dx_1\times dx_2\times du_1)$, we have

$$\begin{aligned}
&\int_{E\times E\times U} f(x_1,x_2,u_1)\left[\int_U g(u_2)\,\overline{\eta}(x_1,x_2,u_1,du_2)\right]\,\overline{\mu}(dx_1\times dx_2\times du_1)\\
&\quad=\Psi F\\
&\quad=\Psi(\Pi F)\\
&\quad=\int_{E\times E\times U} f(x_1,x_2,u_1)\left[\int_U g(u_2)\eta(x_2,du_2)\right]\,\overline{\mu}(dx_1\times dx_2\times du_1).
\end{aligned}$$

Therefore

$$\begin{aligned}
\tilde{\mu}(dx_1\times dx_2\times du_1\times du_2) &= \eta(x_2,du_2)\overline{\mu}(dx_1\times dx_2\times du_1)\\
&= \eta(x_2,du_2)\widehat{\eta}(x_1,dx_2\times du_1)\mu_E(dx_1).
\end{aligned}$$

Furthermore, using $F(x_1,x_2,u_1,u_2)=h(x_1)[g(x_2,u_2)-g(x_2,u_1)]$, it follows that

$$\begin{aligned}
0 &= \Psi F\\
&= \int_E h(x_1)\left[\int_{E\times U}\int_U\{g(x_2,u_2)-g(x_2,u_1)\}\eta(x_2,du_2)\widehat{\eta}(x_1,dx_2\times du_1)\right]\mu_E(dx_1),
\end{aligned}$$

so

$$\int_{E\times U}\int_U\{g(x_2,u_2)-g(x_2,u_1)\}\eta(x_2,du_2)\widehat{\eta}(x_1,dx_2\times du_1)=0\quad\text{a.e. }\mu_E(dx_1). \tag{3.20}$$

Let $\{(X_k,u_k)\colon k=1,2,\ldots\}$ be a Markov chain on $E\times U$ having initial distribution μ and transition function $\widehat{\eta}$. A straightforward computation shows that the Markov chain is stationary, and by (3.19) and (3.20), for each $f\in\mathcal{D}(A)$ and $g\in\overline{C}(E\times U)$,

$$[(I-n^{-1}A)f](X_k,u_k)-\sum_{i=0}^{k-1}n^{-1}A_n[(I-n^{-1}A)f](X_i,u_i)$$

and

$$\sum_{i=0}^{k} Cg(X_i, u_i)$$

are martingales with respect to the filtration $\mathcal{F}_k = \sigma((X_i, u_i)\colon\ 0 \le i \le k)$.

Define $X_n(\cdot) = X_{[n\cdot]}$, $u_n(\cdot) = u_{[n\cdot]}$ and $\mathcal{F}_t^n = \sigma((X_n(s), u_n(s))\colon\ 0 \le s \le t)$. It immediately follows (recall $A_n(I - n^{-1}A)f = Af$) that

$$[(I - n^{-1}A)f](X_n(t), u_n(t)) - \int_0^{[nt]/n} Af(X_n(s), u_n(s))\,ds \tag{3.21}$$

and

$$\int_0^{[nt]/n} Cg(X_n(s), u_n(s))\,ds \tag{3.22}$$

are $\mathcal{F}_t^n$-martingales.

Define the measure-valued random variable Λ_n by

$$\Lambda_n([0,t] \times H) = \int_0^t I_H(u_n(s))\,ds, \qquad \forall H \in \mathcal{B}(U).$$

Note that

$$E[\Lambda_n([0,t] \times H)] = t\mu(E \times H), \qquad \forall H \in \mathcal{B}(U),$$

so, by the tightness of a measure on a complete, separable metric space, for each $\epsilon > 0$, there exists a compact K_ϵ such that

$$E[\Lambda_n([0,t] \times K_\epsilon^c)] = t\mu(E \times K_\epsilon^c) \le t\epsilon. \tag{3.23}$$

Recall the definition of $\mathcal{L}^{(m)}(U)$ from Section 1.1. Relative compactness of $\{\Lambda_n\}$ on $\mathcal{L}^{(m)}(U)$ follows from (3.23) by Lemma 1.3 of Kurtz [18].

Let $Z^n = G(X_n)$, with G defined by (3.4). Then by (3.21) and the definition of Λ_n, for $k = 1, 2, \ldots$,

$$\begin{aligned} &g_k(X_n(t)) - \frac{[nt] + 1 - nt}{n} Ag_k(X_n(t), u_n(t)) - \int_{[0,t]} Ag_k(X_n(s), u)\Lambda_n(ds \times du) \\ &= Z_k^n(t) - \frac{[nt] + 1 - nt}{n} Ag_k(G^{-1}(Z^n(t)), u_n(t)) - \int_{[0,t]} Ag_k(G^{-1}(Z^n(s)), u)\Lambda_n(ds \times du) \end{aligned} \tag{3.24}$$

is a martingale. Recalling that $\prod_{k\in I} Z_k^n = Z_{k(I)}^n$, Theorems 3.9.4 and 3.9.1 of Ethier and Kurtz [11] imply the relative compactness of $\{Z^n\}$ in $D_{\widehat{E}}[0,\infty)$, and hence, (Z^n, Λ_n) is relatively compact in $D_{\widehat{E}}[0,\infty) \times \mathcal{L}^{(m)}(U)$. Define $\nu \in \mathcal{P}(\widehat{E} \times U)$ by $\int f(z,u)\nu(dz \times du) = \int f(G(x),u)\mu(dx \times du)$. Then $Ag_k(G^{-1}(\cdot),\cdot)$ can be approximated in $L_1(\nu)$ by bounded, continuous functions in $\overline{C}(\widehat{E} \times U)$, and as in Lemma 3.3, we see that for any limit point (Z, Λ), (3.24) converges in distribution, at least along a subsequence, to

$$Z_k(t) - \int_{[0,t]\times U} Ag_k(G^{-1}(Z(s)), u)\Lambda(ds \times du), \tag{3.25}$$

which will be a martingale with respect to the filtration $\{\mathcal{F}_t^{Z,\Lambda}\}$. Note that Z is a stationary process (even though as continuous time processes the X_n are not). Since for each $t \geq 0$, $Z(t)$ has distribution $\nu_{\widehat{E}} = \nu(\cdot \times U)$ which satisfies $\int f d\nu_{\widehat{E}} = \int f \circ G d\mu_E$, by Lemma 3.1, $Z(t) \in G(E)$ a.s. and hence we can define $X(t) = G^{-1}(Z(t))$, and we have that

$$M_{g_k}(t) = g_k(X(t)) - \int_{[0,t]\times U} Ag_k(X(s),u)\Lambda(ds \times du)$$

is a $\{\mathcal{F}_t^{Z,\Lambda}\}$-martingale.

By the same argument, (3.22) converges in distribution to

$$M_g^C(t) = \int_{[0,t]\times U} Cg(X(s),u)\Lambda(ds \times du), \tag{3.26}$$

for every $g \in B(E \times U)$. Since (3.22) is a martingale, it follows that (3.26) is an $\{\mathcal{F}_t^{X,\Lambda}\}$-martingale. But (3.26) is a continuous, finite variation process, and every martingale with these properties is a constant implying $M_g^C \equiv 0$. Consequently, recalling that $\Lambda \in \mathcal{L}^{(m)}(U)$ implies $\Lambda(ds \times U) = ds$, we have the identity

$$\int_0^t \int_U g(X(s),u)\eta(X(s),du)ds = \int_{[0,t]\times U} g(X(s),u)\Lambda(ds \times du), \quad g \in B(E \times U),$$

so

$$M_{g_k}(t) = g_k(X(t)) - \int_0^t \int_U Ag_k(X(s),u)\,\eta(X(s),du)\,ds.$$

Since A is contained in the bounded pointwise closure of the linear span of $\{(g_k, Ag_k)\}$, we see that $(X, \eta(X,\cdot))$ is a solution of the controlled martingale problem for (A, μ_E). Finally, $\eta(X,\cdot)$ is admissible since $\mu(\mathcal{U}) = 1$ implies $\eta(x, U_x) = 1$ *a.e.* $\mu(dx)$. □

Proof of Theorem 3.4. For each $n \geq 1$, let

$$\begin{aligned}
\psi_n &= 2^{-n}(2^n \vee \psi)\ , \\
k_n(x) &= \int \psi_n(x,u)\,\eta(x,du)\ , \\
c_n &= \int k_n(x)\,\mu_0(dx) = \int \psi_n(x,u)\,\mu(dx \times du).
\end{aligned}$$

Observe that $\psi_n \geq 1$ for all n, and that as $n \to \infty$ $\psi_n(x,u) \searrow 1$, $c_n \searrow 1$ and $k_n \searrow 1$. Define the operators A_n on $\mathcal{D}(A)$ by

$$A_n f(x,u) = Af(x,u)/\psi_n(x,u),$$

and note that $A_n \subset \overline{C}(E) \times \overline{C}(E \times U)$. Defining $\mu_n \in \mathcal{P}(E \times U)$ by

$$\mu_n(H) = c_n^{-1} \int_H \psi_n(x,u)\mu(dx \times du) \qquad \forall H \in \mathcal{B}(E \times U),$$

we have

$$\int_{E\times U} A_n f d\mu_n = c_n^{-1} \int_{E\times U} Af d\mu = 0.$$

For each n, A_n and μ_n satisfy the conditions of Theorem 3.6 and η_n of (3.3) is given by

$$\eta_n(x,du) = \frac{\psi_n(x,u)}{k_n(x)}\eta(x,du).$$

Note, in particular, that $A_n = \frac{\psi}{\psi_n}A_0$ so Condition 1.1(*iii*) is satisfied since $\frac{\psi}{\psi_n}$ is bounded and A_0 satisfies the condition. Thus there exist stationary processes $\{Z^n\}$ with sample paths in $D_{\widehat{E}}[0,\infty)$ such that, setting $X_n = G^{-1}(Z^n)$, $(X_n, \eta_n(X_n,\cdot))$ is a solution of the controlled martingale problem for (A_n, μ_n).

Let φ be nonnegative and convex on $[0,\infty)$ with $\varphi(0) = 0$, $\lim_{r\to\infty}\varphi(r)/r = \infty$, and

$$\int_{E\times U} \varphi(\psi(x,u))\mu(dx \times du) < \infty.$$

(Existence of φ follows from (3.7).) Since φ and φ' are nondecreasing on $[0,\infty)$, it follows that if $z \geq 0$ and $y \geq 1$, then $\varphi(\frac{z}{y})y \leq \varphi(z)$. Consequently, for $f \in \mathcal{D}(A)$,

$$\varphi(|A_n f(x,u)|/a_f) \leq \varphi(\psi(x,u)/\psi_n(x,u)) \leq \frac{\varphi(\psi(x,u))}{\psi_n(x,u)},$$

and

$$\int \varphi(|A_n f(x,u)|/a_f)\mu_n(dx\times du) \leq \frac{1}{c_n}\int \varphi(\psi(x,u))\mu(dx\times du) \leq \int \varphi(\psi(x,u))\mu(dx\times du).$$

In particular, this inequality ensures the uniform integrability of

$$\left\{\left|\int_U A_n f(X_n(t),u)\eta_n(X_n(t),du)\right|\right\}.$$

The relative compactness of $\{Z^n\}$ is established by applying Theorem 3.9.1 of Ethier and Kurtz [11] and Theorem 4.5 Stockbridge [27] exactly as in the proof of Theorem 4.7 of Stockbridge [27]. Let Z be a weak limit point of $\{Z^n\}$, and to simplify notation, assume that the original sequence converges. As before, set $X = G^{-1}(Z)$.

For each k,

$$g_k(X(t)) - \int_0^t \int_U Ag_k(X(s),u)\,\eta(X(s),du)\,ds$$

is an $\{\mathcal{F}_t^X\}$-martingale if and only if

$$E\Big[\Big(Z_k(t_{m+1}) - Z_k(t_m) - \int_{t_m}^{t_{m+1}} \int_U Ag_k(G^{-1}(Z(s)),u)\,\eta(G^{-1}(Z(s)),du)\,ds\Big) \times \prod_{i=1}^m h_i(Z(t_i))\Big] = 0 \qquad (3.27)$$

for each $m \geq 1$, $0 \leq t_1 \leq \ldots \leq t_m < t_{m+1}$, and $h_1,\ldots,h_m \in \overline{C}(\widehat{E})$. Note that condition (3.27) is satisfied with A_n, η_n and Z^n replacing A, η, and Z .

Let $t_1, \ldots, t_{m+1} \in \{t \geq 0 : P(Z(t) = Z(t-)) = 1\}$ and $h_1, \ldots, h_m \in \overline{C}(\widehat{E})$. Since $Z^n \Rightarrow Z$, as $n \to \infty$,

$$
\begin{aligned}
E\left[(Z_k^n(t_{m+1}) - Z_k^n(t_m)) \prod_{i=1}^m h_i(Z^n(t_i))\right] \\
\to \quad & E\left[(Z_k(t_{m+1}) - Z_k(t_m)) \prod_{i=1}^m h_i(Z(t_i))\right] \\
= \quad & E\left[(g_k(X(t_{m+1})) - g_k(X(t_m))) \prod_{i=1}^m h_i(Z(t_i))\right]
\end{aligned}
$$

Lemma 3.3 does not apply directly to the integral term, but a similar argument works using the fact that $\mu_n(dx \times du) = c_n^{-1}\psi_n(x,u)\mu(dx \times du) \leq \psi(x,u)\mu(dx \times du)$. □

Theorem 3.4 establishes the existence of stationary processes on the complete, separable, metric space E. The proof involves embedding E in the compact space $\widehat{E}$, demonstrating existence of appropriate stationary processes Z on $\widehat{E}$, and then obtaining the solution by applying the inverse G^{-1}. In the next section, it will be necessary to work directly with the processes Z. We therefore state the corresponding existence result in terms of these processes; the proof, of course, has already been given.

Theorem 3.7 *Let A satisfy Condition 1.2. Suppose $\mu \in \mathcal{P}(E \times U)$ satisfies (3.1) and (3.2), ψ satisfies (3.7), and define μ_E and η by (3.3). Define $\nu \in \mathcal{P}(\widehat{E} \times U)$ by $\nu(H_1 \times H_2) = \mu(G^{-1}(H_1), H_2)$ for every $H_1 \in \mathcal{B}(\widehat{E})$ and $H_2 \in \mathcal{B}(U)$. Then there exists a cadlag, stationary, $\widehat{E}$-valued process Z such that*

$$
f(G^{-1}(Z(t))) - \int_0^t \int_U Af(G^{-1}(Z(s)), u)\eta(G^{-1}(Z(s), du)ds
$$

is an $\mathcal{F}_t^Z$-martingale for every $f \in \mathcal{D}$ and

$$
E[I_{H_1}(Z(t))\eta(G^{-1}(Z(t)), H_2)] = \nu(H_1 \times H_2)
$$

for every $H_1 \in \mathcal{B}(\widehat{E})$ and $H_2 \in \mathcal{B}(U)$.

4 Singular martingale problems

In this section, we characterize the marginal distributions of stationary solutions of the singular controlled martingale problem. Previous work of this nature includes the papers by Weiss [29] and Kurtz [16] which considered constrained processes. Weiss [29] characterized the marginal distribution of a stationary solution to a submartingale problem for diffusions in a bounded domain. Inspired by Weiss, Kurtz [16] used the results of Stockbridge [27] to characterize the stationary marginals for general constrained processes. The results of this section are more general than the previous results in that they apply to processes with singular control, constrained processes being a subclass of such processes, and the controls are identified in feedback form.

Let $\widehat{E}$ be the compact space constructed in Section 3, and let G be the mapping from E into $\widehat{E}$ given by (3.4).

Lemma 4.1 *Let A and B satisfy Condition 1.2. Suppose that $\mu_0 \in \mathcal{M}(E \times U)$ and $\mu_1 \in \mathcal{M}(E \times U)$ satisfy conditions (1.15), (1.17), and (1.16). For $i = 0, 1$, let μ_i have a state marginal μ_i^E and kernel $\eta_i(x, \cdot)$ on the control space so that $\mu_i(dx \times du) = \eta_i(x, du)\mu_i^E(dx)$. Define the measure $\mu \in \mathcal{P}(E \times U)$ by*

$$\mu(H) = K^{-1}\left(\mu_0(H) + \mu_1(H)\right), \qquad \forall H \in \mathcal{B}(E \times U), \tag{4.1}$$

where $K = \mu_0(E \times U) + \mu_1(E \times U)$ is the normalizing constant. Let

$$\nu = \mu \circ G^{-1}, \qquad \nu_0 = \mu_0 \circ G^{-1}, \qquad \nu_1 = \mu_1 \circ G^{-1}, \tag{4.2}$$

and let $\nu^{\widehat{E}}$, $\nu_0^{\widehat{E}}$ and $\nu_1^{\widehat{E}}$ denote the corresponding marginals on $\widehat{E}$. Then there exist a stationary process Z on $\widehat{E}$ and non-negative, continuous, non-decreasing processes λ_0 and λ_1 such that

- *$Z(0)$ has distribution $\nu^{\widehat{E}}$,*
- *λ_0 and λ_1 have stationary increments,*
- *$\lambda_0(t) + \lambda_1(t) = t$,*
- *$(Z, \lambda_0, \lambda_1)$ is $\{\mathcal{F}_t^Z\}$-adapted and*
- *for each $f \in \mathcal{D}$,*

$$\begin{aligned} f(G^{-1}(Z(t))) &- \int_0^t \int_U KAf(G^{-1}(Z(s)), u)\, \eta_0(G^{-1}(Z(s)), du) d\lambda_0(s) \\ &- \int_0^t \int_U KBf(G^{-1}(Z(s)), u)\, \eta_1(G^{-1}(Z(s)), du) d\lambda_1(s) \end{aligned}$$

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

Remark 4.2 *By defining $X = G^{-1}(Z)$, the conclusions of Lemma 4.1 can be stated in terms of a stationary E-valued process X. Since we will need to use the process Z in the sequel, we have chosen to express Lemma 4.1 in terms of this process.*

Proof. Let $\{\kappa_0, \kappa_1\}$ be distinct points not contained in U and define $\tilde{U} = U \times \{\kappa_0, \kappa_1\}$. For $f \in \mathcal{D}$, define

$$Cf(x, u, \kappa) = KAf(x, u)I_{\{\kappa_0\}}(\kappa) + KBf(x, u)I_{\{\kappa_1\}}(\kappa)$$

and

$$\psi(x, u, \kappa) = \psi_A(x, u)I_{\{\kappa_0\}}(\kappa) + \psi_B(x, u)I_{\{\kappa_1 0\}}(\kappa).$$

We redefine μ so that it is a probability measure on $E \times \tilde{U}$ by setting

$$\begin{aligned} &\int h(x, u, \kappa)\mu(dx \times du \times d\kappa) \\ &\qquad = K^{-1}\left(\int h(x, u, \kappa_0)\, \mu_0(dx \times du) + \int h(x, u, \kappa_1)\, \mu_1(dx \times du)\right), \end{aligned}$$

where $K = \mu_0(E \times U) + \mu_1(E \times U)$ is the normalizing constant above.

Observe that μ has marginal μ^E and that both μ_0^E and μ_1^E are absolutely continuous with respect to μ^E. Hence we can write

$$\begin{aligned}
&\int_{E\times\tilde{U}} h(x,u,\kappa)\,\mu(dx\times du\times d\kappa)\\
&\quad= \int_E\int_{\tilde{U}} h(x,u,\kappa)\Big(\eta_0(x,du)\delta_{\{\kappa_0\}}(d\kappa)K^{-1}\frac{d\mu_0^E}{d\mu^E}(x)\\
&\qquad\qquad +\eta_1(x,du)\delta_{\{\kappa_1\}}(d\kappa)K^{-1}\frac{d\mu_1^E}{d\mu^E}(x)\Big)\,\mu^E(dx).
\end{aligned}$$

Thus, when μ is decomposed as $\mu(dx\times du\times d\kappa)=\eta(x,du\times d\kappa)\mu^E(dx)$, the conditional distribution η satisfies

$$\eta(x,du\times d\kappa)=\eta_0(x,du)\delta_{\{\kappa_0\}}(d\kappa)K^{-1}\frac{d\mu_0^E}{d\mu^E}(x)+\eta_1(x,du)\delta_{\{\kappa_1\}}(d\kappa)K^{-1}\frac{d\mu_1^E}{d\mu^E}(x)\qquad a.e.\ \mu^E.$$

It follows that for each $f\in\mathcal{D}$,

$$\begin{aligned}
\int Cf(x,u,\kappa)\,\mu(dx\times du\times d\kappa) &= \left[\int Af(x,u)\mu_0(dx\times du)+\int Bf(x,u)\mu_1(dx\times du)\right]\\
&= 0.
\end{aligned}$$

This identity, together with the conditions on A and B, imply that the conditions of Theorem 3.7 are satisfied. Therefore there exists a stationary process Z such that

$$\begin{aligned}
&f(G^{-1}(Z(t)))-\int_0^t\int_{\tilde{U}} Cf(G^{-1}(Z(s)),u,\kappa)\,\eta(G^{-1}(Z(s)),du\times d\kappa)\,ds\\
&\quad= f(G^{-1}(Z(t)))-\int_0^t\int_U Af(G^{-1}(Z(s)),u)\,\eta_0(G^{-1}(Z(s)),du)\frac{d\mu_0^E}{d\mu^E}(G^{-1}(Z(s)))\,ds\\
&\qquad -\int_0^t\int_U Bf(G^{-1}(Z(s)),u)\,\eta_1(G^{-1}(Z(s)),du)\frac{d\mu_1^E}{d\mu^E}(G^{-1}(Z(s)))\,ds
\end{aligned}$$

is an $\{\mathcal{F}_t^Z\}$-martingale for each $f\in\mathcal{D}$ and $E[I_{H_1}(Z(t))\eta(G^{-1}(Z(t)),H_2)]=\nu(H_1\times H_2)$. Observe that for each $H_1\in\mathcal{B}(\widehat{E})$ and $H_2\in\mathcal{B}(U)$,

$$E[I_{H_1}(Z(t))\frac{d\mu_0^E}{d\mu^E}(G^{-1}(Z(s)))\eta_0(G^{-1}(Z(t)),H_2)]=\nu_0(H_1\times H_2)$$

and

$$E[I_{H_1}(Z(t))\frac{d\mu_1^E}{d\mu^E}(G^{-1}(Z(s)))\eta_1(G^{-1}(Z(t)),H_2)]=\nu_1(H_1\times H_2)\ .$$

For $i=0,1$, define

$$\lambda_i(t)=K^{-1}\int_0^t\frac{d\mu_i^E}{d\mu^E}(G^{-1}(Z(s)))\,ds=K^{-1}\int_0^t\frac{d\nu_i^E}{d\nu^E}(Z(s))\,ds.$$

Then λ_0 and λ_1 have stationary increments and $\lambda_0(t)+\lambda_1(t)=t$. $\square$

4.1 Proof of Theorem 1.7

Proof. For $n = 1, 2, 3, \ldots$, consider the operators A and $B_n = nB$. By (1.17), the measures μ_0 and $\mu_{n,1} = (1/n)\mu_1$ satisfy

$$\int Af\, d\mu_0 + \int B_n f\, d\mu_{n,1} = 0, \quad \forall f \in \mathcal{D},$$

and A and B_n satisfy the conditions of Lemma 4.1. Define the probability measure $\mu_n = K_n^{-1}(\mu_0 + (1/n)\mu_1)$, where K_n is a normalizing constant, and the measures ν_n, ν_0 and $\nu_{n,1}$ as in (4.2). Let $\nu_n^{\widehat{E}}$, $\nu_0^{\widehat{E}}$, and $\nu_{n,1}^{\widehat{E}}$ denote the corresponding marginals on $\widehat{E}$. Then for each n, Lemma 4.1 implies that there exist a stationary process Z^n and non-negative, continuous, non-decreasing processes λ_0^n and λ_1^n having stationary increments such that $\lambda_0^n(t) + \lambda_1^n(t) = t$ and for each $f \in \mathcal{D}$,

$$\begin{aligned} f(G^{-1}(Z^n(t))) &- \int_0^t \int_U K_n A f(G^{-1}(Z^n(s)), u)\eta_0(G^{-1}(Z^n(s)), du)d\lambda_0^n(s) \\ &- \int_0^t \int_U K_n B_n f(G^{-1}(Z^n(s)), u)\eta_1(G^{-1}(Z^n(s)), du)d\lambda_1^n(s) \end{aligned} \tag{4.3}$$

is an $\{\mathcal{F}_t^{Z^n}\}$-martingale and $Z^n(t)$ has distribution $\nu_n^{\widehat{E}}(\cdot) = \nu_n(\cdot \times U)$. In particular, by considering $f = g_k$, we have

$$\begin{aligned} Z_k^n(t) &- \int_0^t \int_U K_n A g_k(G^{-1}(Z^n(s)), u)\eta_0(G^{-1}(Z^n(s)), du)d\lambda_0^n(s) \\ &- \int_0^t \int_U K_n B g_k(G^{-1}(Z^n(s)), u)\eta_1(G^{-1}(Z^n(s)), du)d\lambda_1^n(s) \end{aligned}$$

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

Observe that $K_n = \mu_0(E \times U) + (1/n)\mu_1(E \times U) > 1$ and $K_n \searrow 1$ as $n \to \infty$. Thus $\nu_n^{\widehat{E}} \Rightarrow \nu_0^{\widehat{E}}$ as $n \to \infty$.

Also note that

$$\lambda_0^n(t) = K_n^{-1} \int_0^t \frac{d\mu_0^E}{d\mu_n^E}(G^{-1}(Z^n(s)))\, ds \quad \text{and} \quad \lambda_1^n(t) = K_n^{-1} \int_0^t \frac{1}{n}\frac{d\mu_1^E}{d\mu_n^E}(G^{-1}(Z^n(s)))\, ds. \tag{4.4}$$

Now observe that

$$\begin{aligned} nE[\lambda_1^n(t)] &= nE\left[\int_0^t \frac{1}{n} K_n^{-1} \frac{d\mu_1^E}{d\mu_n^E}(G^{-1}(Z^n(s)))\, ds\right] \\ &= t\int_E K_n^{-1} \frac{d\mu_1^E}{d\mu_n^E}(x)\, \mu_n^E(dx) \\ &= K_n^{-1} t \mu_1^E(E) \\ &< t\mu_1^E(E) =: C_t. \end{aligned}$$

Therefore $E[\lambda_1^n(t)] < C_t/n$ and converges to zero as $n \to \infty$, which implies

$$E[\lambda_0^n(t)] \to t, \qquad \text{as } n \to \infty, \tag{4.5}$$

since $\lambda_0^n(t) + \lambda_1^n(t) = t$. Note that $\dfrac{d\mu_0^E}{d\mu_n^E} \le K_n$, and hence (4.4) and (4.5) imply

$$\lambda_0^n(t) \to t \text{ in probability} \qquad \text{as } n \to \infty.$$

We now show existence of a limiting process Z. We verify that the conditions of Corollary 1.4 of [17] are satisfied.

Consider the collection of coordinate functions $\{z_k\}$ Note that the compact containment condition is trivially satisfied and $\{z_k\} \subset \overline{C}(\widehat{E})$ separates points in $\widehat{E}$.

For $t > 0$, consider any partition $\{t_i\}$ of $[0,t]$. Then

$$\begin{aligned}
&E\left[\sum_i |E\left[Z_k^n(t_{i+1}) - Z_k^n(t_i)|\mathcal{F}_{t_i}\right]|\right] \\
&= E\left[\sum_i \left|E\left[\int_{t_i}^{t_{i+1}} \int_U K_n A g_k(G^{-1}(Z^n(s)),u)\,\eta_0(G^{-1}(Z^n(s)),du)\,d\lambda_0^n(s)\right.\right.\right. \\
&\qquad \left.\left.\left. + \int_{t_i}^{t_{i+1}} \int_U K_n B_n g_k(G^{-1}(Z^n(s)),u)\,\eta_1(G^{-1}(Z^n(s)),du)\,d\lambda_1^n(s)\,|\mathcal{F}_{t_i}\right]\right|\,\right] \\
&\le \sum_i (t_{i+1}-t_i)\left(\int_E \int_U |Ag_k(G^{-1}(z),u)|\eta_0(G^{-1}(z),du)\frac{d\mu_0^E}{d\mu_n^E}(G^{-1}(z))\,\nu_n^E(dz)\right) \\
&\qquad + \sum_i (t_{i+1}-t_i)\left(\int_E \int_U n\,|Bg_k(x,u)|\eta_1(x,du)\frac{1}{n}\frac{d\mu_1^E}{d\nu_n^E}(x)\,\nu_n^E(dx)\right) \\
&= t\cdot\left(\int_{E\times U} |Ag_k(x,u)|\,\mu_0(dx\times du) + \int_{E\times U} |Bg_k(x,u)|\,\mu_1(dx\times du)\right) \\
&= t\cdot\left(||Ag_k||_{\mathcal{L}^1(\mu_0)} + ||Bg_k||_{\mathcal{L}^1(\mu_1)}\right) < \infty,
\end{aligned}$$

where the last inequality follows from (1.16) and Condition 1.2. Thus condition (1.7) of [17, Corollary 1.4] is satisfied. By selecting a weakly convergent subsequence and applying the Skorohod representation theorem, if necessary, we may assume that there exists a process Z such that $Z^n(t) \to Z(t)$ *a.s.*, for all but countably many t.

Now for each n, define the random measure $\widehat{\Gamma}^n$ on $\widehat{E} \times [0,\infty)$ satisfying

$$\begin{aligned}
\widehat{\Gamma}^n(H_1 \times H_2) &= \int_0^\infty nK_n\, I_{H_1\times H_2}(Z^n(s),s)d\lambda_1^n(s) \\
&= \int_0^\infty I_{H_1\times H_2}(Z^n(s),s)\frac{d\mu_1^E}{d\mu_n^E}(G^{-1}(Z^n(s)))\,ds,
\end{aligned} \tag{4.6}$$

for all $H_1 \in \mathcal{B}(\widehat{E}), H_2 \in \mathcal{B}[0,\infty)$. Then $\{\widehat{\Gamma}^n\}$ is a sequence of $\mathcal{L}(\widehat{E})$-valued random variables. We show that this sequence of measure-valued random variables is relatively compact.

Note that for a complete, separable metric space S, a collection of measures $\mathcal{K} \subset \mathcal{L}(S)$ is relatively compact if $\sup_{\mu\in\mathcal{K}} \mu(S\times[0,T]) < \infty$ for each T, and for each T and $\epsilon > 0$, there exists a compact set $K_{T,\epsilon} \subset S$ such that $\sup_{\mu\in\mathcal{K}} \mu(K^c_{T,\epsilon} \times [0,T]) < \epsilon$.

Recall, $\widehat{E}$ is compact, so the second condition is trivially satisfied by each $\widehat{\Gamma}^n$. Now observe that

$$\begin{aligned} E[\widehat{\Gamma}^n(H\times[0,T])] &= E\left[\int_0^T I_H(Z^n(s))\frac{d\mu_1^E}{d\mu_n^E}(G^{-1}(Z^n(s)))\,ds\right] \\ &= T\int_{G^{-1}(H)} \frac{d\mu_1^E}{d\mu_n^E}(x)\mu_n^E(dx) \\ &= T\,\mu_1^E(G^{-1}(H)). \end{aligned} \tag{4.7}$$

Taking $H = \widehat{E}$ and applying Markov's inequality,

$$P(\widehat{\Gamma}^n(\widehat{E}\times[0,T]) \geq M_T) \leq T\mu_1^E(E)M_T^{-1}.$$

Given $\epsilon > 0$, taking a sequence $\{T_j\}$ with $T_j \to \infty$ and setting $M_{T_j} = T_j\mu_1^E(E)\epsilon/2^j$ shows that the sequence $\{\widehat{\Gamma}^n\}$ of random measures is tight and hence relatively compact. By passing to an appropriate subsequence $\{n_k\}$, if necessary, and applying the Skorohod representation theorem, we can assume that there exists a random measure $\widehat{\Gamma}$ on $\widehat{E}\times[0,\infty)$ such that $\widehat{\Gamma}^{n_k} \to \widehat{\Gamma}$ *a.s.* in $\mathcal{L}(\widehat{E})$ and, for all but countably many t, $Z^{n_k}(t) \to Z(t)$ *a.s.* in $\widehat{E}$ and $\widehat{\Gamma}^{n_k}_t \to \widehat{\Gamma}_t$ *a.s.* in $\mathcal{M}(\widehat{E}\times[0,t])$, where $\widehat{\Gamma}^{n_k}_t$ and $\widehat{\Gamma}_t$ are the restriction of the measures to $\widehat{E}\times[0,t]$. The stationarity of the time increments of $\widehat{\Gamma}$ follows from the definition of Γ^n and the fact that the Z^n are stationary. The finiteness of $E[\widehat{\Gamma}(\widehat{E}\times[0,t])]$ for each t follows from Fatou's Lemma, which implies the finiteness of $\widehat{\Gamma}(\widehat{E}\times[0,t])$ *a.s.*.

We now show that for each k, each $m \geq 1$, and each choice of $0 \leq t_1 \leq t_2 \leq \cdots \leq t_m < t_{m+1}$ and $h_i \in \overline{C}(\widehat{E}\times\mathcal{L}_{t_i}(\widehat{E}))$, $i = 1,\ldots,m$

$$\begin{aligned} E\Bigg[\Bigg\{ & Z_k(t_{m+1}) - Z_k(t_m) - \int_{t_m}^{t_{m+1}}\int_U Ag_k(G^{-1}(Z(s)),u)\,\eta_0(G^{-1}(Z(s)),du)ds \\ & - \int_{\widehat{E}\times(t_m,t_{m+1}]}\int_U Bg_k(G^{-1}(z),u)\,\eta_1(G^{-1}(z),du)\widehat{\Gamma}(dz\times ds)\Bigg\}\cdot\prod_{i=1}^m h_i(Z(t_i),\widehat{\Gamma}_{t_i})\Bigg] = 0, \end{aligned} \tag{4.8}$$

which is true if and only if for each k,

$$\begin{aligned} Z_k(t) &- \int_0^t\int_U Ag_k(G^{-1}(Z(s)),u)\eta_0(G^{-1}(Z(s)),du)ds \\ &- \int_{\widehat{E}\times[0,t]}\int_U Bg_k(G^{-1}(z),u)\eta_1(G^{-1}(z),du)\widehat{\Gamma}(dz\times ds) \end{aligned} \tag{4.9}$$

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

The analog of (4.8) for $(Z^n,\widehat{\Gamma}^n)$ is

$$\begin{aligned} E\Bigg[\Bigg\{ & Z_k^n(t_{m+1}) - Z_k^n(t_m) \\ & - \int_{t_m}^{t_{m+1}}\int_U \frac{d\mu_0^E}{d\mu_n^E}(G^{-1}(Z^n(s))Ag_k(G^{-1}(Z^n(s)),u)\,\eta_0(G^{-1}(Z^n(s)),du)ds \end{aligned} \tag{4.10}$$

$$-\int_{\widehat{E}\times(t_m,t_{m+1}]}\int_U Bg_k(G^{-1}(z),u)\,\eta_1(G^{-1}(z),du)\widehat{\Gamma}^n(dz\times ds)\Big\} \tag{4.11}$$
$$\cdot\prod_{i=1}^m h_i(Z^n(t_i),\widehat{\Gamma}^n_{t_i})\Big] = 0.$$

The idea is to let $n\to\infty$ to establish (4.8). However, care needs to be taken since the $\{\widehat{\Gamma}^n\}$ are not necessarily bounded measures. To overcome this difficulty, for each $n\geq 0$ and $M\geq 0$, we define the stopping time $\tau^{M,n}$ by

$$\tau^{M,n} = \inf\left\{t>0 : \widehat{\Gamma}^n(\widehat{E}\times[0,t])\geq M\right\}. \tag{4.12}$$

Note that for $M_1\leq M_2$, $\tau^{M_1,n}\leq\tau^{M_2,n}$ and

$$\begin{aligned} P(\tau^{M,n}\leq T) &= P\left(\widehat{\Gamma}^n(\widehat{E}\times[0,T])\geq M\right)\\ &\leq M^{-1}E\left[\widehat{\Gamma}^n(\widehat{E}\times[0,T])\right]\\ &= M^{-1}T\mu_1^E(E), \end{aligned}$$

so $\tau^{M,n}\to\infty$ $a.s.$ as $M\to\infty$.

Since if M_k^n is a martingale, $M_k^n(\cdot\wedge\tau^{M,n})$ is also a martingale,

$$\begin{aligned} E\Big[\Big\{ & Z_k^n(t_{m+1}\wedge\tau^{M,n}) - Z_k^n(t_m\wedge\tau^{M,n})\\ & -\int_{t_m\wedge\tau^{M,n}}^{t_{m+1}\wedge\tau^{M,n}}\int_U \frac{d\mu_0^E}{d\mu_n^E}(G^{-1}(Z^n(s))Ag_k(G^{-1}(Z^n(s)),u)\eta_0(G^{-1}(Z^n(s)),du)\,ds\\ & -\int_{\widehat{E}\times(t_m\wedge\tau^{M,n},t_{m+1}\wedge\tau^{M,n}]} Bg_k(G^{-1}(Z^n(s)),u)\eta_1((G^{-1}(Z^n(s)),du)\widehat{\Gamma}^n(dz\times ds)\Big\}\\ & \cdot\prod_{i=1}^m h_i(Z^n(t_i),\widehat{\Gamma}^n_{t_i})\Big] = 0 \end{aligned} \tag{4.13}$$

holds for each k, $m\geq 1$, $0\leq t_1\leq t_2\leq\cdots\leq t_m<t_{m+1}$, and $h_i\in\overline{C}(\widehat{E}\times\mathcal{L}_{t_i}(\widehat{E}))$, $i=1,\ldots,m$.
Now for each M and n, define the random measure $\widehat{\Gamma}^{M,n}$ by

$$\int_{\widehat{E}\times[0,t]} h(z,s)\,\widehat{\Gamma}^{M,n}(dz\times ds) = \int_{\widehat{E}\times[0,t\wedge\tau^{M,n}]} h(z,s)\widehat{\Gamma}^n(dz\times ds),$$

for all bounded, continuous functions h. The following observations should be made about $\widehat{\Gamma}^{M,n}$. First, $\widehat{\Gamma}^{M,n}$ is monotone in M in that for every nonnegative function h,

$$\int h\,d\widehat{\Gamma}^{M_1,n}\leq\int h\,d\widehat{\Gamma}^{M_2,n},$$

whenever $M_1\leq M_2$. Second, as $M\to\infty$, $\widehat{\Gamma}^{M,n}\to\widehat{\Gamma}^n$ and, moreover, $\widehat{\Gamma}^{M,n}_t=\widehat{\Gamma}^n_t$ on the set $\{\tau^{M,n}\geq t\}$, where, as above, $\widehat{\Gamma}^{M,n}_t$ and $\widehat{\Gamma}^n_t$ denote the restrictions of $\widehat{\Gamma}^{M,n}$ and $\widehat{\Gamma}^n$, respectively, to random measures on $\widehat{E}\times[0,t]$. Finally, $\{\widehat{\Gamma}^{M,n}\}$ also satisfies

$$E[\widehat{\Gamma}^{M,n}(H\times[0,t])]\leq E[\widehat{\Gamma}^n(H\times[0,t])]=\nu_1^{\widehat{E}}(H),$$

and so is relatively compact.

Using a diagonal argument and the Skorohod representation theorem, if necessary, for each M there exist τ^M, Γ^M and some subsequence (to simplify notation, we assume the entire sequence) such that $\Gamma^{M,n} \to \widehat{\Gamma}^M$ and $\tau^{M,n} \to \tau^M$ a.s. and for all but countably many t, $Z^n(t) \to Z(t)$ and $\Gamma_t^{M,n} \to \widehat{\Gamma}_t^M$ a.s.

τ^M and $\widehat{\Gamma}^M$ inherit a number of properties from $\tau^{M,n}$ and $\widehat{\Gamma}^{M,n}$. First, $\tau^M \to \infty$ *a.s.* as $M \to \infty$. Second, $\widehat{\Gamma}^M$ is monotone in M, and since $\tau^M \geq \lim_{\epsilon \searrow 0} \inf\{t > 0 : \widehat{\Gamma}(\widehat{E} \times [0,t]) > M - \epsilon\}$, it follows that $\widehat{\Gamma}_t^M = \widehat{\Gamma}_t$ *a.s.* on the set $\{\widehat{\Gamma}(\widehat{E} \times [0,t]) < M\}$ and hence,

$$\widehat{\Gamma}_t^M \nearrow \widehat{\Gamma}_t \ a.s. \quad \forall t \geq 0.$$

Now choose $t_1, \ldots, t_m, t_{m+1} \in \mathcal{T} := \{t : (Z^n(t), \widehat{\Gamma}_t^n) \to (Z(t), \widehat{\Gamma}_t) \ a.s.\}$ in (4.13). Since $|Z_k| \leq a_k$,

$$E\left[Z_k^n(t) \prod_{i=1}^m h_i(Z^n(t_i), \widehat{\Gamma}_{t_i}^n)\right] \to E\left[Z_k(t) \prod_{i=1}^m h_i(Z(t_i), \widehat{\Gamma}_{t_i})\right], \quad t \in \mathcal{T}. \tag{4.14}$$

Note also that

$$\left| E\left[Z_k^n(t) \prod_{i=1}^m h_i(Z^n(t_i), \widehat{\Gamma}_{t_i}^n)\right] - E\left[Z_k^n(t \wedge \tau^{M,n}) \prod_{i=1}^m h_i(Z^n(t_i), \widehat{\Gamma}_{t_i}^n)\right]\right| \leq 2a_k \prod_{i=1}^m \|h_i\| P\{\tau^{M,n} \leq t\}$$

and

$$\left| E\left[Z_k(t) \prod_{i=1}^m h_i(Z(t_i), \widehat{\Gamma}_{t_i})\right] - E\left[Z_k(t \wedge \tau^M) \prod_{i=1}^m h_i(Z(t_i), \widehat{\Gamma}_{t_i})\right]\right| \leq 2a_k \prod_{i=1}^m \|h_i\| P\{\tau^M \leq t\}.$$

Now recall the definitions of the measures $\nu_0^{\widehat{E}}$, $\mu_1^{\widehat{E}}$, and $\nu_n^{\widehat{E}}$ on $\widehat{E}$ (from the first paragraph of the proof). Observe that the measures $\nu_0^{\widehat{E}}$ and $\nu_1^{\widehat{E}}$ are absolutely continuous with respect to $\nu_n^{\widehat{E}}$ with Radon-Nikodym derivatives $\dfrac{d\nu_i^{\widehat{E}}}{d\nu_n^{\widehat{E}}}(z) = \dfrac{d\mu_i^E}{d\mu_n^E}(G^{-1}(z))$.

We claim that for each $g \in \mathcal{L}^1(\nu_0^{\widehat{E}})$,

$$E\left[\int_{t_m}^{t_{m+1}} g(Z^n(s)) \frac{d\nu_0^{\widehat{E}}}{d\nu_n^{\widehat{E}}}(Z^n(s)) ds \prod_{i=1}^m h_i(Z^n(t_i), \widehat{\Gamma}_{t_i}^n)\right] \to E\left[\int_{t_m}^{t_{m+1}} g(Z(s)) ds \prod_{i=1}^m h_i(Z(t_i), \widehat{\Gamma}_{t_i})\right], \tag{4.15}$$

as $n \to \infty$. To see this, fix $g \in \mathcal{L}^1(\nu_0^{\widehat{E}})$ and let $\epsilon > 0$ be given and select $g_\epsilon \in C(\widehat{E})$ (recall, $\widehat{E}$ is compact so g_ϵ is bounded) such that

$$\int_{\widehat{E}} |g(z) - g_\epsilon(z)| \nu_0^{\widehat{E}}(dz) < \epsilon.$$

Then, recalling that $Z^n(s)$ has distribution $\nu_n^{\widehat{E}}$ and the definition of λ_0^n in (4.4),

$$\left| E\left[\int_{t_m}^{t_{m+1}} (g(Z^n(s)) - g_\epsilon(Z^n(s))) \frac{d\nu_0^{\widehat{E}}}{d\nu_n^{\widehat{E}}}(Z^n(s))\, ds \prod_{i=1}^m h_i(Z^n(t_i), \widehat{\Gamma}_{t_i}^n) \right] \right|$$

$$\le E\left[\int_{t_m}^{t_{m+1}} |g(Z^n(s)) - g_\epsilon(Z^n(s))| \cdot \frac{d\nu_0^{\widehat{E}}}{d\nu_n^{\widehat{E}}}(Z^n(s))\, ds \right] \prod_{i=1}^m ||h_i||$$

$$\le \epsilon(t_{m+1} - t_m) \prod_{i=1}^m ||h_i||.$$

Similarly, $Z(t)$ will have distribution $\nu_0^{\widehat{E}}$ and so

$$\left| E\left[\int_{t_m}^{t_{m+1}} (g(Z(s)) - g_\epsilon(Z(s)))ds \prod_{i=1}^m h_i(Z(t_i), \widehat{\Gamma}_{t_i}) \right] \right|$$

$$\le E\left[\int_{t_m}^{t_{m+1}} |g(Z(s)) - g_\epsilon(Z(s))|\, ds \right] \prod_{i=1}^m ||h_i||$$

$$\le \epsilon(t_{m+1} - t_m) \prod_{i=1}^m ||h_i||.$$

We now consider the convergence of

$$E\left[\int_{t_m}^{t_{m+1}} g_\epsilon(Z^n(s)) \frac{d\nu_0^{\widehat{E}}}{d\nu_n^{\widehat{E}}}(Z^n(s))\, ds \prod_{i=1}^m h_i(Z^n(t_i), \widehat{\Gamma}_{t_i}^n) \right].$$

Since for each i, $h_i(Z^n(t_i), \widehat{\Gamma}_{t_i}^n) \to h_i(Z(t_i), \widehat{\Gamma}_{t_i})$ *a.s.* as $n \to \infty$, and for almost all $s \in [t_m, t_{m+1}]$, $g_\epsilon(Z^n(s)) \to g_\epsilon(Z(s))$ *a.s.* and $\frac{d\nu_0^{\widehat{E}}}{d\nu_n^{\widehat{E}}}(Z^n(s)) \to 1$ in probability, as $n \to \infty$, (4.15) follows. A similar argument establishes

$$E\left[\int_{t_m \wedge \tau^{M,n}}^{t_{m+1} \wedge \tau^{M,n}} g(Z^n(s)) \frac{d\nu_0^{\widehat{E}}}{d\nu_n^{\widehat{E}}}(Z^n(s))ds \prod_{i=1}^m h_i(Z^n(t_i), \widehat{\Gamma}_{t_i}^n) \right]$$

$$\to \; E\left[\int_{t_m \wedge \tau^M}^{t_{m+1} \wedge \tau^M} g(Z(s))ds \prod_{i=1}^m h_i(Z(t_i), \widehat{\Gamma}_{t_i}) \right]. \tag{4.16}$$

Turning to the convergence of the terms involving the random measures, observe that

$$\int_{\widehat{E} \times (t_m, t_{m+1}]} g(z) \widehat{\Gamma}^{M,n}(dz \times ds) \to \int_{\widehat{E} \times (t_m, t_{m+1}]} g(z) \widehat{\Gamma}^M(dz \times ds) \; a.s. \tag{4.17}$$

for all $g \in C(\widehat{E})$. Since these random variables are bounded by $\|g\|M$, the bounded convergence theorem implies

$$E\left[\int_{\widehat{E} \times (t_m, t_{m+1}]} g(z) \widehat{\Gamma}^{M,n}(dz \times ds) \prod_{i=1}^m h_i(Z^n(t_i), \Gamma_{t_i}^n) \right] \tag{4.18}$$

$$\to E\left[\int_{\widehat{E}\times(t_m,t_{m+1}]} g(z)\widehat{\Gamma}^M(dz\times ds)\prod_{i=1}^m h_i(Z(t_i),\widehat{\Gamma}_{t_i})\right].$$

Noting that for $g\geq 0$,

$$E[\int_{\widehat{E}\times(t_m,t_{m+1}]} g(z)\widehat{\Gamma}^{M,n}(dz\times ds)]\leq E[\int_{\widehat{E}\times(t_m,t_{m+1}]} g(z)\widehat{\Gamma}^n(dz\times ds)]=(t_{m+1}-t_m)\int_{\widehat{E}} gd\nu_1^{\widehat{E}},$$

this convergence can be extended to all $g\in\mathcal{L}^1(\nu_1^{\widehat{E}})$ by approximating g by $g_\epsilon\in C(\widehat{E})$ as above. In particular, $g\in\mathcal{L}^1(\nu_1^{\widehat{E}})$ if $|g(z)|\leq C_g\widehat{\psi}_B(z)\equiv C_g\int_U\psi_B(G^{-1}(z)),u)\eta_1(G^{-1}(z),du)$.
Taking g in (4.16) to be $g(z)=\int_U Ag_k(G^{-1}(z),u)\eta_0(G^{-1}(z),du)$ and g in (4.18) to be $g(z)=\int_U Bg_k(G^{-1}(z),u)\eta_1(g^{-1}(z),du)$, along with (4.14), we have

$$\begin{aligned}
&\Bigg|E\Bigg[\Bigg\{Z_k(t_{m+1}\wedge\tau^M)-Z_k(t_m\wedge\tau^M)\\
&\qquad-\int_{t_m}^{t_{m+1}} I_{\{s\leq\tau^M\}}Ag_k(G^{-1}(Z(s)),u)\eta_0(G^{-1}(Z(s)),du)ds \qquad (4.19)\\
&\qquad-\int_{\widehat{E}\times(t_m,t_{m+1}]} Bg_k(G^{-1}(z),u)\eta_1(G^{-1}(z),du)\widehat{\Gamma}^M(dz\times ds)\Bigg\}\prod_{i=1}^m h_i(Z(t_i),\widehat{\Gamma}_{t_i})\Bigg]\Bigg|\\
&\qquad\leq 4a_k\prod_{i=1}^m\|h_i\|P\{\tau_M\leq t\}.
\end{aligned}$$

Defining $\widehat{\psi}_A(z)=\int_U\psi(G^{-1}(z),u)\eta_0(G^{-1}(z),du)$, the expression in the expectation is dominated by

$$R_M=\left(2a_k+\int_{t_m}^{t_{m+1}} a_{g_k}\widehat{\psi}_A(Z(s))ds+\int_{\widehat{E}\times(t_m,t_m]} b_{g_k}\widehat{\psi}_B(z)\widehat{\Gamma}^M(dz\times ds)\right)\prod_{i=1}^m\|h_i\|. \qquad (4.20)$$

Let R be defined as R_M with $\widehat{\Gamma}^M$ replaced by $\widehat{\Gamma}$. Noting that $R_M\nearrow R$ and $E[R]<\infty$, the dominated convergence theorem implies that as $M\to\infty$, the expectation on the left side of (4.19) converges to the left side of (4.8) while the right side converges to zero. Consequently, (4.8) holds for $t_1,\ldots,t_m\in\mathcal{T}$. Right continuity of Z and $\widehat{\Gamma}$ then implies (4.8) holds for all t, and thus (4.9) is a martingale.

The random measure $\widehat{\Gamma}$ has stationary increments but need not be adapted to the filtration generated by Z. Without loss of generality, assume the process Z is defined for all t, not just for $t\geq 0$, and assume that $\widehat{\Gamma}$ takes values in measures on $\widehat{E}\times\mathbb{R}$. Define $\overline{\mathcal{F}}_t^Z=\sigma(Z(s):-\infty<s\leq t)\vee\mathcal{N}$, where $\mathcal{N}$ denotes the null sets, so that $\{\overline{\mathcal{F}}_t^Z\}$ is the completion of the filtration generated by the process Z. Let $\mathcal{F}_t=\overline{\mathcal{F}}_{t+}^Z$. Then by Lemma 6.1, using the space $\widehat{E}$ and taking $H(x,s)=e^{-|s|}$, there exists a predictable random measure $\tilde{\Gamma}$ satisfying (6.16). As a result, (4.9) will be an $\overline{\mathcal{F}}_{t+}^Z$-martingale with $\tilde{\Gamma}$ replacing $\widehat{\Gamma}$. Note that $\tilde{\Gamma}$ has stationary increments.

Define $X=G^{-1}(Z)$ and the random measure Γ on $E\times\mathbb{R}$ by

$$\int_{E\times R} h(x,s)\Gamma(dx\times ds)=\int_{\widehat{E}\times R} h(G^{-1}(z),s)\tilde{\Gamma}(dz\times ds).$$

By working with the completions, $\overline{\mathcal{F}}_{t+}^X=\overline{\mathcal{F}}_{t+}^Z$, which implies (1.18) is an $\{\overline{\mathcal{F}}_{t+}^X\}$-martingale for each g_k and hence for each $f\in\mathcal{D}$. □

5 Proof of Theorem 1.11

Proof. We essentially follow the proof of Theorem 4.1 of Kurtz and Stockbridge [21]. Let α be chosen to satisfy (1.21). Define the operators $\tilde{A}$ and $\tilde{B}$ by

$$\tilde{A}(\phi f)(x,\theta,s)=\phi(\theta,s)\widehat{A}f(x)+\phi'(\theta,s)f(x)+\alpha\left[\phi(-\theta,0)\int_E f(y)\nu_0(dy)-\phi(\theta,s)f(x)\right] \quad (5.1)$$

and

$$\tilde{B}(\phi f)(x,\theta,s)=\phi(\theta,s)\widehat{B}f(x), \quad (5.2)$$

for $f\in\mathcal{D}$ and $\phi\in\mathcal{D}_1=\{\phi:\phi,\phi'\in\overline{C}(\{-1,1\}\times[0,\infty))\}$, where ϕ' denotes differentiation with respect to the second variable. Taking $\tilde{\mathcal{D}}=\{f\phi: f\in\mathcal{D},\phi\in\mathcal{D}_1\}$, $\tilde{A}$, $\tilde{B}$ and $\tilde{\mathcal{D}}$ satisfy Condition 1.2 with $\psi_{\tilde{A}}=\int_U\psi_A(\cdot,u)\eta_0(\cdot,du)$ and $\psi_{\tilde{B}}=\int_U\psi_B(\cdot,u)\eta_1(\cdot,du)$. Define the measures $\tilde{\mu}_0\in\mathcal{P}(E\times\{-1,1\}\times[0,\infty))$ and $\tilde{\mu}_1\in\mathcal{M}(E\times\{-1,1\}\times[0,\infty))$ by

$$\begin{aligned}\int_{E\times\{-1,1\}\times[0,\infty)} & h(x,\theta,s)\tilde{\mu}_0(dx\times d\theta\times ds) && (5.3)\\ &=\alpha\int_0^\infty e^{-\alpha s}\int_E\frac{h(x,-1,s)+h(x,1,s)}{2}\nu_s(dx)ds\\ \int_{E\times\{-1,1\}\times[0,\infty)} & h(x,\theta,s)\tilde{\mu}_1(dx\times d\theta\times ds) && (5.4)\\ &=\alpha\int_{E\times[0,\infty)} e^{-\alpha s}\frac{h(x,-1,s)+h(x,1,\theta)}{2}\mu(dx\times ds),\end{aligned}$$

for $h\in B(E\times\{-1,1\}\times[0,\infty))$. The following computation verifies that $(\tilde{A},\tilde{B},\tilde{\mu}_0,\tilde{\mu}_1)$ satisfy (1.19). For $f\in\mathcal{D}$ and $\phi\in\mathcal{D}_1$, and setting $\overline{\phi}(s)=(\phi(-1,s)+\phi(1,s))/2$,

$$\begin{aligned}&\int\tilde{A}(\phi f)d\tilde{\mu}_0+\int\tilde{B}(\phi f)d\tilde{\mu}_1\\ &=\alpha\int_0^\infty\int_E e^{-\alpha s}\left[\overline{\phi}(s)\widehat{A}f+\overline{\phi}'(s)f+\alpha\left(\overline{\phi}(0)\int fd\nu_0-\overline{\phi}(s)f\right)\right]d\nu_s ds\\ &\quad+\alpha\int_{E\times[0,\infty)}e^{-\alpha s}\overline{\phi}(s)\widehat{B}fd\mu\\ &=\alpha\int_0^\infty\left(e^{-\alpha s}\overline{\phi}'(s)-\alpha e^{-\alpha s}\overline{\phi}(s)\right)\left(\int fd\nu_s\right)ds\\ &\quad+\alpha\int_0^\infty e^{-\alpha s}\overline{\phi}(s)\left(\int\widehat{A}fd\nu_s\right)ds+\alpha\overline{\phi}(0)\int fd\nu_0+\alpha\int e^{-\alpha s}\overline{\phi}(s)\widehat{B}fd\mu\\ &=\alpha\int_0^\infty\left(e^{-\alpha s}\overline{\phi}'(s)-\alpha e^{-\alpha s}\overline{\phi}(s)\right)\left(\int_E fd\nu_0+\int_0^s\int_E\widehat{A}fd\nu_r dr+\int_{E\times[0,s]}\widehat{B}fd\mu\right)ds\\ &\quad+\alpha\int_0^\infty e^{-\alpha s}\overline{\phi}(s)\left(\int\widehat{A}fd\nu_s\right)ds+\alpha\overline{\phi}(0)\int fd\nu_0+\alpha\int e^{-\alpha s}\overline{\phi}(s)\widehat{B}fd\mu\\ &=\alpha\int_0^\infty\left(\frac{d}{ds}(e^{-\alpha s}\overline{\phi}(s))\right)\left(\int_E fd\nu_0+\int_{E\times[0,\infty)}I_{[0,s]}(r)\widehat{A}fd\nu_r dr\right.\end{aligned}$$

$$
\begin{aligned}
&\qquad + \int_{E\times[0,\infty)} I_{[0,s]}(r)\widehat{B} f d\mu \Big) ds \\
&+\alpha\int_0^\infty e^{-\alpha s}\overline{\phi}(s)\left(\int \widehat{A} f d\nu_s\right) ds + \alpha\overline{\phi}(0)\int f d\nu_0 + \alpha\int e^{-\alpha s}\overline{\phi}(s)\widehat{B} f d\mu \\
= 0,
\end{aligned}
$$

where the last equality follows by interchanging the order of the integrals with respect to r and s and observing that all the terms cancel.

At this point we could apply Corollary 1.9 to obtain the existence of a stationary space-time process (Y, Θ, S) and boundary measure Γ with stationary distribution given by $\tilde{\mu}_0$; however, we need to specify a more explicit form for the random measure. As in the proof of Theorem 1.7, for each n, define the operators $\tilde{B}_n = n\tilde{B}$ and the measures $\tilde{\mu}_1^n = \frac{1}{n}\tilde{\mu}_1$ and $\tilde{\mu}_n = K_n^{-1}(\tilde{\mu}_0 + \tilde{\mu}_1^n)$. Apply Lemma 4.1 to get the stationary processes Z^n, Θ^n, and S^n, and processes λ_0^n and λ_1^n satisfying (4.4) such that

$$
\begin{aligned}
\phi(\Theta^n(t), S^n(t)) f(G^{-1}(Z^n(t))) &- \int_0^t K_n\tilde{A}(\phi f)(G^{-1}(Z^n(s)), \Theta^n(s), S^n(s)) d\lambda_0^n(s) \qquad (5.5)\\
&- \int_0^t K_n\tilde{B}_n(\phi f)(G^{-1}(Z^n(s)), \Theta^n(s), S^n(s)) d\lambda_1^n(s)
\end{aligned}
$$

is an $\{\mathcal{F}_t^{Z^n,\Theta^n,S^n}\}$-martingale.

Existence of limiting processes Z, Θ, and S follow as in the proof of Theorem 1.7, and in the current setting, $(\Theta^n, S^n) \Rightarrow (\Theta, S)$ in the Skorohod topology. This stronger convergence of (Θ^n, S^n) allows us to be more explicit in describing a boundary measure $\widehat{\Gamma}$.

For each n, let $\widehat{\Gamma}^n \in \mathcal{L}(\widehat{E})$ be the random measure satisfying

$$
\begin{aligned}
\widehat{\Gamma}^n(H_1\times[0,T]) &= \int_0^T nK_n I_{H_1}(Z^n(s)) d\lambda_1^n(s) \qquad (5.6)\\
&= \int_0^T I_{H_1}(Z^n(s))\frac{d\tilde{\mu}_1}{d\tilde{\mu}_n}(G^{-1}(Z^n(s)), \Theta^n(s), S^n(s))\, ds.
\end{aligned}
$$

Note that

$$
nK_n\int_0^t h(G^{-1}(Z^n(s)), \Theta^n(s), S^n(s)) d\lambda_1^n(s) = \int_{\widehat{E}\times[0,t]} h(G^{-1}(z), \Theta^n(s), S^n(s))\widehat{\Gamma}^n(dz\times ds),
$$

and hence

$$
t\int_{E\times[0,\infty)} h(x,\theta,s)\tilde{\mu}_1(dx\times d\theta\times ds) = E\Big[\int_{\widehat{E}\times[0,t]} h(G^{-1}(z), \Theta^n(s), S^n(s))\widehat{\Gamma}^n(dz\times ds)\Big]. \qquad (5.7)
$$

In terms of $\widehat{\Gamma}^n$, (5.5) becomes

$$
\begin{aligned}
\phi(\Theta^n(t), S^n(t)) f(G^{-1}(Z^n(t))) &- \int_0^t K_n\tilde{A}(\phi f)(G^{-1}(Z^n(s)), \Theta^n(s), S^n(s)) d\lambda_0^n(s) \qquad (5.8)\\
&- \int_{\widehat{E}\times[0,t]} \tilde{B}(\phi f)(G^{-1}(z), \Theta^n(s), S^n(s))\widehat{\Gamma}^n(dz\times ds).
\end{aligned}
$$

Since

$$\begin{aligned} E[\widehat{\Gamma}^n(\widehat{E}\times[0,T])] &= E\left[\int_0^T I_{\widehat{E}}(Z^n(s))\frac{d\tilde{\mu}_1}{d\tilde{\mu}_n}(G^{-1}(Z^n(s)),\Theta^n(s),S^n(s))ds\right] \\ &= T\tilde{\mu}_1(E\times\{-1,1\}\times[0,\infty)), \end{aligned}$$

the argument in the proof of Theorem 1.7 shows that the sequence $\{\widehat{\Gamma}^n\}$ is relatively compact, and the existence of the limit $(Z,\Theta,S,\widehat{\Gamma})$, at least along a subsequence, follows as before. The $\widehat{E}\times\{-1,1\}\times[0,\infty)$-valued process (Z,Θ,S) (which we may take to be defined for $-\infty<t<\infty$) is stationary and the random measure $\widehat{\Gamma}$ (which we may take to be defined on $\widehat{E}\times(-\infty,\infty)$) has stationary time-increments. The convergence of (5.8) to a martingale follows as before, except for the last term. Applying the Skorohod representation theorem, we will assume that the convergence is almost sure.

Taking $f\equiv 1$ in (5.8), we see that

$$\phi(\Theta^n(t),S^n(t))-\int_0^t[\phi'(\Theta^n(s),S^n(s))+\alpha(\phi(-\Theta^n(s),0)-\phi(\Theta^n(s),S^n(s)))]d\lambda_0^n(s)$$

is a martingale, and it follows that (Θ^n,S^n) can be written as $(\Theta^n(t),S^n(t))=(\tilde{\Theta}^n(\lambda_0^n(t)),\tilde{S}^n(\lambda_0^n(t)))$, where $(\tilde{\Theta}^n,\tilde{S}^n)$ is a solution of the martingale problem for C given by $C\phi(\theta,r)=\phi'(\theta,r)+\alpha(\phi(-\theta,0)-\phi(\theta,r))$, for $\phi\in\mathcal{D}_1$. Uniqueness of this martingale problem (cf. [11, Theorem 4.4.1]) and the fact that $\lambda_0(t)\to t$ implies that (Θ,S) is a stationary solution of the martingale problem for C. It follows (see [21], page 624) that S is exponentially distributed at each time t, increases linearly at rate 1 up to a random time that is exponentially distributed with parameter α at which time it jumps to 0, and the cycle repeats. Similarly, $\Theta(t)=\Theta(0)(-1)^{N(t)}$, where $N(t)$ is the number of returns to zero made by S in the time interval $(0,t]$. Note also that (Θ^n,S^n) converges to (Θ,S) in the Skorohod topology.

Some care needs to be taken in analyzing the convergence of the last term in (5.8). We can approximate $\tilde{B}(\phi f)(G^{-1}(z),\Theta^n(s),S^n(s))$ by $h(z,\Theta^n(s),S^n(s))$ with $h\in\overline{C}(\widehat{E}\times\{-1,1\}\times[0,\infty))$ (that is, select h so that $\int_{\widehat{E}\times\{-1,1\}\times[0,\infty)}|\tilde{B}(\phi f)(G^{-1}(z),\theta,r)-h(z,\theta,r)|\tilde{\mu}_1(dz\times d\theta\times dr)$ is small), but we cannot rule out the possibility that $\widehat{\Gamma}$ has a discontinuity at the jump times of (Θ,S). In particular, $\mu(E\times\{0\})$ may not be zero. For instance, in the transaction cost models (see Example 1.5 and [8, 25]), if the support of ν_0 is not a subset of the control region, then the optimal solution may instantaneously jump to the boundary of the control region at time zero. In this situation, $\{\nu_t\}$ will not be right continuous at zero, but will satisfy

$$\int_E f\nu_{0+}=\int_E f d\nu_0+\int_E\widehat{B}f(x)\mu(dx\times\{0\}).$$

Since in the process we are constructing, $Y=G^{-1}(Z)$ "starts over" with distribution ν_0 at each time τ_k that S jumps back to zero, in this situation Y must take an instantaneous jump governed by $\widehat{\Gamma}$ so that $Y(\tau_k+)$ has distribution ν_{0+}.

Let $\tau_1=\inf\{t>0:S(t)=0\}$, and for $k\geq 1$, let $\tau_{k+1}=\inf\{t>\tau_k:S(t)=0\}$. Then we can write $\widehat{\Gamma}=\widehat{\Gamma}_0+\widehat{\Gamma}_1$ so that

$$\phi(\Theta(t),S(t))f(G^{-1}(Z(t+)))\quad-\quad\int_0^t\tilde{A}(\phi f)(G^{-1}(Z(s)),\Theta(s),S(s))ds \tag{5.9}$$

$$
\begin{aligned}
&- \int_{\widehat{E}\times[0,t]} \tilde{B}(\phi f)(G^{-1}(z),\Theta(s-),S(s-))\widehat{\Gamma}_0(dz\times ds) \\
&- \int_{\widehat{E}\times[0,t]} \tilde{B}(\phi f)(G^{-1}(z),\Theta(s),0)\widehat{\Gamma}_1(dz\times ds)
\end{aligned}
$$

is an $\{\overline{\mathcal{F}}_{t+}^{Z,\Theta,S,\widehat{\Gamma}_0,\widehat{\Gamma}_1}\}$-martingale for $f\in\mathcal{D}$ and $\phi\in\mathcal{D}_1$, where the support of the measure $\widehat{\Gamma}_1$ is on $\widehat{E}\times\{\tau_k, k\geq 1\}$.

Taking $\phi\equiv 1$, we see that

$$
\begin{aligned}
f(G^{-1}(Z(t+))) \quad &- \int_0^t (\widehat{A}f(G^{-1}(Z(s)) + \alpha\left[\int_E f(y)\nu_0(dy) - f(G^{-1}(Z(s))\right])ds \qquad (5.10) \\
&- \int_{\widehat{E}\times[0,t]} \widehat{B}f(G^{-1}(z))\widehat{\Gamma}(dz\times ds)
\end{aligned}
$$

is an $\{\overline{\mathcal{F}}_{t+}^{Z,\Theta,S,\widehat{\Gamma}}\}$-martingale.

Define the E-valued process $Y=G^{-1}(Z)$ and the random measure $\tilde{\Gamma}$ on $E\times[0,\infty)$ by

$$
\int_{E\times[0,t]} h(x,s)\tilde{\Gamma}(dx\times ds) = \int_{\widehat{E}\times[0,t]} h(G^{-1}(z),s)\widehat{\Gamma}(dz\times ds).
$$

We may assume that $\tilde{\Gamma}$ is adapted to the filtration $\{\overline{\mathcal{F}}_{t+}^{Y,\Theta,S}\}$, by applying Lemma 6.1, if necessary. We can rewrite (5.10) as

$$
f(Y(t+)) - \int_0^t (\widehat{A}f(Y(s)) + \alpha\left[\int_E f(y)\nu_0(dy) - f(Y(s)\right])ds - \int_{E\times[0,t]} \widehat{B}f(y)\tilde{\Gamma}(dy\times ds), \tag{5.11}
$$

and (5.11) is an $\{\overline{\mathcal{F}}_{t+}^{Y,S}\}$-martingale.

Now let $\tau_0=\sup\{t\leq 0: S(t)=0\}$. Define the process X by $X(t)=Y(\tau_1+t)$, $t\geq 0$, the random measure Γ on $E\times[0,\infty)$ by $\Gamma(H\times[t_1,t_2])=\tilde{\Gamma}(H\times[\tau_1+t_1,\tau_1+t_2])$, for $H\in\mathcal{B}(E)$ and $0\leq t_1<t_2$, and the filtration $\mathcal{G}_t=\overline{\mathcal{F}}_{(\tau_1+t)+}^{Y,\Theta,S}$, $t\geq 0$. Then, an application of the optional sampling theorem (cf. [11, Theorem 2.2.13]) implies

$$
f(X(t)) - \int_0^t \left[\widehat{A}f(X(r)) + \alpha\left(\int f d\nu_0 - f(X(r))\right)\right]dr - \int_{E\times[0,t]} \widehat{B}f(x)\Gamma(dx\times ds)
$$

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

Define

$$
L(t) = [\alpha(\tau_1-\tau_0)]^{-1}e^{\alpha t}I_{[0,\tau_2-\tau_1)}(t) = [\alpha(\tau_1-\tau_0)]^{-1}e^{\alpha(\tau_2\wedge(\tau_1+t)-\tau_1)}\frac{1+\Theta(\tau_1)\Theta((\tau_1+t)\wedge\tau_2)}{2},
$$

and observe that L is a $\{\mathcal{G}_t\}$-martingale with $E[L(t)]=1$. Let $\widehat{P}$ be a new probability measure having Radon-Nikodym derivative $L(t)$ on $\mathcal{G}_t$ with respect to the original probability measure P, and denote expectation with respect to $\widehat{P}$ by $E^{\widehat{P}}[\cdot]$.

Observe that, under P,

$$[\alpha(\tau_1-\tau_0)]^{-1}e^{\alpha t}I_{[0,\tau_2-\tau_1)}(t)f(X(t))-\int_0^t[\alpha(\tau_1-\tau_0)]^{-1}e^{\alpha r}I_{[0,\tau_2-\tau_1)}(r)\widehat{A}f(X(r))\,dr$$
$$-\int_{E\times[0,t]}[\alpha(\tau_1-\tau_0)]^{-1}e^{\alpha r}I_{[0,\tau_2-\tau_1)}(r)\widehat{B}f(x)\Gamma(dx\times dr)$$

is a $\{\mathcal{G}_t\}$-martingale which implies that for each $m\geq 1$, $0\leq t_1\leq\cdots\leq t_m<t_{m+1}$, and $h_i\in\overline{C}(E\times\mathcal{L}_{t_i}(E))$, $i=1,\ldots,m$

$$\begin{aligned}
0 &= E\Bigg[[\alpha(\tau_1-\tau_0)]^{-1}\Bigg\{e^{\alpha t_{m+1}}I_{[0,\tau_2-\tau_1)}(t_{m+1})f(X(t_{m+1}))-e^{\alpha t_m}I_{[0,\tau_2-\tau_1)}(t_m)f(X(t_m))\\
&\qquad -\int_{t_m}^{t_{m+1}}e^{\alpha r}I_{[0,\tau_2-\tau_1)}(r)\widehat{A}f(X(r))dr\\
&\qquad -\int_{E\times(t_m,t_{m+1}]}e^{\alpha r}I_{[0,\tau_2-\tau_1)}(r)\widehat{B}f(x)\Gamma(dx\times dr)\Bigg\}\prod_{i=1}^m h_i(X(t_i),\Gamma_{t_i})\Bigg]\\
&= E\Bigg[[\alpha(\tau_1-\tau_0)]^{-1}\Bigg\{e^{\alpha t_{m+1}}I_{[0,\tau_2-\tau_1)}(t_{m+1})f(X(t_{m+1}))\\
&\quad -E[e^{\alpha t_{m+1}}I_{[0,\tau_2-\tau_1)}(t_{m+1})|\mathcal{G}_{t_m}]f(X(t_m))\\
&\quad -\int_{t_m}^{t_{m+1}}E[e^{\alpha t_{m+1}}I_{[0,\tau_2-\tau_1)}(t_{m+1})|\mathcal{G}_r]\widehat{A}f(X(r))dr\\
&\quad -\int_{E\times(t_m,t_{m+1}]}E[e^{\alpha t_{m+1}}I_{[0,\tau_2-\tau_1)}(t_{m+1})|\mathcal{G}_r]\widehat{B}f(x)\Gamma(dx\times dr)\Bigg\}\prod_{i=1}^m h_i(X(t_i),\Gamma_{t_i})\Bigg]\\
&= E^{\widehat{P}}\Bigg[\Bigg\{f(X(t_{m+1}))-f(X(t_m))-\int_{t_m}^{t_{m+1}}\widehat{A}f(X(r))dr\\
&\qquad -\int_{E\times(t_m,t_{m+1}]}\widehat{B}f(x)\Gamma(dx\times dr)\Bigg\}\prod_{i=1}^m h_i(X(t_i),\Gamma_{t_i})\Bigg],
\end{aligned}$$

where the elimination of the conditioning in the last term in the braces follows by Lemma 6.2. It follows that

$$f(X(t))-f(X(0))-\int_0^t\widehat{A}f(X(r))dr-\int_{E\times[0,t]}\widehat{B}f(x)\Gamma(dx\times dr)$$

is an $\{\mathcal{F}_t^{X,\Gamma}\}$-martingale under $\widehat{P}$.

We now derive the distribution of $X(t)$. First for each $h\in\overline{C}(E\times[0,\infty))$,

$$\begin{aligned}
&E^{\widehat{P}}\left[\alpha\int_0^T e^{-\alpha t}h(X(t),t)dt\right]\\
&\qquad = E\left[[\alpha(\tau_1-\tau_0)]^{-1}\alpha e^{\alpha T}I_{[0,\tau_2-\tau_1)}(T)\int_0^T e^{-\alpha t}h(X(t),t)dt\right]\\
&\qquad = E\left[(\tau_1-\tau_0)^{-1}\int_0^T E[e^{\alpha T}I_{[0,\tau_2-\tau_1)}(T)|\mathcal{G}_t]e^{-\alpha t}h(X(t),t)dt\right]
\end{aligned}$$

$$= E\left[(\tau_1 - \tau_0)^{-1}\int_0^T I_{[0,\tau_2-\tau_1)}(t)h(X(t),t)dt\right]$$
$$= E\left[(\tau_1 - \tau_0)^{-1}\int_{\tau_1}^{\tau_2\wedge(\tau_1+T)} h(Y(t),S(t))dt\right],$$

and letting $T \to \infty$ yields

$$E^{\widehat{P}}\left[\alpha\int_0^\infty e^{-\alpha t}h(X(t),t)dt\right] = E\left[(\tau_1-\tau_0)^{-1}\int_{\tau_1}^{\tau_2} h(Y(t),S(t))dt\right]. \tag{5.12}$$

For $t \geq 0$, define $\tau_0^t = \sup\{r \leq t : S(r) = 0\}$, $\tau_1^t = \inf\{r > t : S(r) = 0\}$, and $\tau_2^t = \inf\{r > \tau_1^t : S(r) = 0\}$. Note that $\tau_i^0 = \tau_i$ for $i = 0, 1, 2$. The quantity

$$(\tau_1^t - \tau_0^t)^{-1}\int_{\tau_1^t}^{\tau_2^t} h(Y(r),S(r))dr$$

is stationary in t and for $t \in [\tau_k, \tau_{k+1})$,

$$(\tau_1^t - \tau_0^t)^{-1}\int_{\tau_1^t}^{\tau_2^t} h(Y(r),S(r))dr = (\tau_{k+1}-\tau_k)^{-1}\int_{\tau_{k+1}}^{\tau_{k+2}} h(Y(r),S(r))dr.$$

Let $N(t)$ denote the number of jumps of S in the interval $(0,t]$. Then by stationarity,

$$\begin{aligned}
E&\left[(\tau_1-\tau_0)^{-1}\int_{\tau_1}^{\tau_2} h(Y(r),S(r))dr\right]\\
&= E\left[T^{-1}\int_0^T (\tau_1^t-\tau_0^t)^{-1}\int_{\tau_1^t}^{\tau_2^t} h(Y(r),S(r))drdt\right]\\
&= T^{-1}E\left[\sum_{i=1}^{N(T)+1}\frac{T\wedge\tau_i - \tau_{i-1}\vee 0}{\tau_i-\tau_{i-1}}\left(\int_{\tau_i}^{\tau_{i+1}} h(Y(r),S(r))dr\right)\right]\\
&= T^{-1}E\left[\int_0^T h(Y(r),S(r))dr\right]\\
&\quad -T^{-1}E\left[\int_0^{\tau_2\wedge T} h(Y(r),S(r))dr\right]\\
&\quad +T^{-1}E\left[\int_{\tau_1}^{\tau_2}\frac{T\wedge\tau_1}{\tau_1-\tau_0}h(Y(r),S(r))dr\right]\\
&\quad +T^{-1}E\left[I_{\{N(T)>1\}}\int_T^{\tau_{N(T)+1}} h(Y(r),S(r))dr\right]\\
&\quad +T^{-1}E\left[I_{\{N(T)>0\}}\frac{T-\tau_{N(T)}}{\tau_{N(T)+1}-\tau_{N(T)}}\int_{\tau_{N(T)+1}}^{\tau_{N(T)+2}} h(Y(r),S(r))dr\right].
\end{aligned}$$

The first term of the right hand side equals $\int h(x,s)\tilde{\mu}_0(dx\times\{-1,1\}\times ds)$ by the stationarity of (Y,S), and the other terms converge to 0 as $T\to\infty$. By (5.3) and (5.12), we obtain

$$E^{\widehat{P}}\left[\alpha\int_0^\infty e^{-\alpha t}h(X(t),t)dt\right] = \alpha\int_0^\infty e^{-\alpha t}\int_E h(x,t)\nu_t(dx)dt. \tag{5.13}$$

Let $\{h_k\} \subset \overline{C}(E)$ be a countable collection which is separating (see [11, p. 112]). Taking $h(x,t)$ in (5.13) to be of the form $\phi(t)h_k(x)$, we see that

$$E^{\widehat{P}}[h_k(X(t))] = \int_E h_k(x)\nu_t(dx), \qquad a.e.\ t,$$

and since $\{h_k\}$ is separating, it follows that $X(t)$ has distribution ν_t for $a.e.\ t$.

Following a similar argument, we determine $E[\Gamma]$. For $h \in \overline{C}(E \times [0,\infty))$, we have

$$\begin{aligned}
&E^{\widehat{P}}\left[\int_{E\times[0,T)} \alpha e^{-\alpha r} h(x,r)\Gamma(dx\times dr)\right] \\
&\quad = E\left[[\alpha(\tau_1-\tau_0)]^{-1}\alpha e^{\alpha T} I_{[0,\tau_2-\tau_1)}(T)\int_{E\times[0,T)} e^{-\alpha r}h(x,r)\Gamma(dx\times dr)\right] \\
&\quad = E\left[(\tau_1-\tau_0)^{-1}\int_{E\times[0,T)} E[e^{\alpha T}I_{[0,\tau_2-\tau_1)}(T)|\mathcal{G}_r]e^{-\alpha r}h(x,r)\Gamma(dx\times dr)\right] \\
&\quad = E\left[(\tau_1-\tau_0)^{-1}\int_{E\times[0,T)} I_{[0,\tau_2-\tau_1)}(r)h(x,r)\Gamma(dx\times dr)\right] \\
&\quad = E\left[(\tau_1-\tau_0)^{-1}\int_{E\times[\tau_1,\tau_2\wedge(\tau_1+T))} h(x,r-\tau_1)\tilde{\Gamma}(dx\times dr)\right],
\end{aligned}$$

where the second equality follows by Lemma 6.2. Letting $T\to\infty$, we obtain

$$\begin{aligned}
&E^{\widehat{P}}\left[\int_{E\times[0,\infty)} \alpha e^{-\alpha r}h(x,r)\Gamma(dx\times dr)\right] \\
&\quad = E\left[(\tau_1-\tau_0)^{-1}\int_{E\times[\tau_1,\tau_2)} h(x,r-\tau_1)\tilde{\Gamma}(dx\times dr)\right].
\end{aligned} \tag{5.14}$$

Recalling the definitions of τ_i^t, for $i=0,1,2$, and τ_k, for $k\geq 0$, the quantity $(\tau_1^t-\tau_0^t)^{-1}\int_{E\times[\tau_1^t,\tau_2^t)} h(x,r-\tau_1^t)\tilde{\Gamma}(dx\times dr)$ is stationary in t, and for $t\in[\tau_k,\tau_{k+1})$,

$$\begin{aligned}
&(\tau_1^t-\tau_0^t)^{-1}\int_{E\times[\tau_1^t,\tau_2^t)} h(x,r-\tau_1^t)\tilde{\Gamma}(dx\times dr) \\
&\quad = (\tau_{k+1}-\tau_k)^{-1}\int_{E\times[\tau_{k+1},\tau_{k+2})} h(x,r-\tau_{k+1})\tilde{\Gamma}(dx\times dr).
\end{aligned}$$

Proceeding exactly as before we have

$$\begin{aligned}
&E\left[(\tau_1-\tau_0)^{-1}\int_{E\times[\tau_1,\tau_2)} h(x,r-\tau_1)\tilde{\Gamma}(dx\times dr)\right] \\
&\quad = E\left[T^{-1}\int_0^T (\tau_1^t-\tau_0^t)^{-1}\int_{E\times[\tau_1^t,\tau_2^t)} h(x,r-\tau_1^t)\tilde{\Gamma}(dx\times dr)dt\right]
\end{aligned}$$

$$
\begin{aligned}
&= T^{-1}E\left[\int_{E\times[0,T)} h(x,r-\tau_0^r)\tilde{\Gamma}(dx\times dr)\right] \\
&\quad -T^{-1}E\left[\int_{E\times[0,\tau_2\wedge T)} h(x,r-\tau_0^r)\tilde{\Gamma}(dx\times dr)\right] \\
&\quad +T^{-1}E\left[\frac{\tau_1\wedge T}{\tau_1-\tau_0}\int_{E\times[\tau_1,\tau_2)} h(x,r-\tau_1)\tilde{\Gamma}(dx\times dr)\right] \\
&\quad +T^{-1}E\left[I_{\{N(T)>1\}}\int_{E\times[T,\tau_{N(T)+1})} h(x,r-\tau_{N(T)})\tilde{\Gamma}(dx\times dr)\right] \\
&\quad +T^{-1}E\left[I_{\{N(T)>0\}}\frac{T-\tau_{N(T)}}{\tau_{N(T)+1}-\tau_{N(T)}}\int_{E\times[\tau_{N(T)+1},\tau_{N(T)+2})} h(x,r-\tau_{N(T)+1})\tilde{\Gamma}(dx\times dr)\right].
\end{aligned}
$$

Since $S(r)=r-\tau_0^r$, by (5.7) and the definition of $\tilde{\Gamma}$, the first term equals $\int_{E\times[0,\infty)} h(x,r)\tilde{\mu}_1(dx \times\{-1,1\}\times dr)$ and the other terms converge to 0 as $T\to\infty$. Thus combining (5.14), the above limit, and (5.5), we have established that

$$
E^{\widehat{P}}\left[\int_{E\times[0,\infty)} \alpha e^{-\alpha r}h(x,r)\Gamma(dx\times dr)\right] = \int_{E\times[0,\infty)} \alpha e^{-\alpha s}h(x,s)\mu(dx\times ds). \tag{5.15}
$$

Taking h of the form $h(x,r)=\alpha^{-1}e^{\alpha r}I_{H_1\times H_2}(x,r)$ for $H_1\in\mathcal{B}(E)$ and $H_2\in\mathcal{B}([0,\infty))$ with H_2 bounded, we have $E^{\widehat{P}}[\Gamma(H_1\times H_2)]=\mu(H_1\times H_2)$, and hence the result. $\square$

6 Appendix

6.1 Existence of an adapted compensator for a random measure

Let $(\Omega,\mathcal{F},P)$ be a probability space with a filtration $\{\mathcal{F}_t\}$. Let $\mathcal{P}$ be the predictable σ-algebra, that is, the smallest σ-algebra of sets in $[0,\infty)\times\Omega$ such that for each $\{\mathcal{F}_t\}$-adapted, left-continuous process X, the mapping $(t,\omega)\to X(t,\omega)$ is $\mathcal{P}$-measurable. A process X is $\{\mathcal{F}_t\}$-*predictable* (or simply *predictable* if the filtration is clear from context) if the mapping $(t,\omega)\to X(t,\omega)$ is $\mathcal{P}$-measurable.

Let (E,r) be a complete separable metric space, and let $\mathcal{P}_E=\mathcal{B}(E)\times\mathcal{P}$. A process Z with values in $M(E)$ (the space of $\mathcal{B}(E)$-measurable functions) is predictable if the mapping $(x,t,\omega)\to Z(x,t,\omega)$ is $\mathcal{P}_E$-measurable. A random measure Π on $E\times[0,\infty)$ is adapted if for each $D\in\mathcal{B}(E)$, the process $\Pi(D\times[0,t])$ is adapted, and Π is predictable if for each $D\in\mathcal{B}(E)$, the process $\Pi(D\times[0,t])$ is predictable. The following result is essentially the existence of the dual predictable projection of a random measure. (See, for example, Jacod and Shiryaev [12], Theorem II.1.8).

Lemma 6.1 *Let $\{\mathcal{F}_t\}$ be a complete, right-continuous filtration. Let Γ be a random measure on $E\times[0,\infty)$ (not necessarily adapted). Suppose that there exists a strictly positive predictable process H such that*

$$
E\left[\int_{E\times[0,\infty)} H(x,s)\Gamma(dx\times ds)\right]<\infty.
$$

Then there exists a predictable random measure $\widehat{\Gamma}$ such that for each $B(E)$-valued predictable Z satisfying $|Z| < K$ for some constant $K < \infty$,

$$M_Z(t) = E\left[\int_{E\times[0,t]} Z(x,s)H(x,s)\Gamma(dx \times ds)\middle|\mathcal{F}_t\right] - \int_{E\times[0,t]} Z(x,s)H(x,s)\widehat{\Gamma}(dx\times ds), \quad (6.16)$$

is an $\{\mathcal{F}_t\}$-martingale. In addition, there exist a kernel γ from $((0,\infty)\times\Omega,\mathcal{P})$ to E and a nondecreasing, right-continuous, predictable process A such that

$$\widehat{\Gamma}(dx \times ds,\omega) = \gamma(s,\omega,dx)dA(s,\omega) + \delta_{\{0\}}(ds)E[\Gamma(dx\times\{0\})|\mathcal{F}_0](\omega). \quad (6.17)$$

Proof. We separate out the atom of Γ at time 0 (which may or may not exist) by defining $\widehat{\Gamma}(\cdot\times\{0\}) = E[\Gamma(\cdot\times\{0\})|\mathcal{F}_0]$. This explains the second term in (6.17).

For $D \in \mathcal{P}_E$, define $\nu(D) = E[\int_{E\times(0,\infty)} I_D(x,s)H(x,s)\Gamma(dx\times ds)]$ and for $C\in\mathcal{P}$, define $\nu_0(C) = \nu(E\times C)$. Since E is Polish, there exists a transition function γ_0 from $((0,\infty)\times\Omega,\mathcal{P})$ into E such that

$$\nu(D) = \int_{(0,\infty)\times\Omega} \gamma_0(s,\omega,D_{(s,\omega)})\nu_0(ds\times d\omega),$$

where $D_{(s,\omega)} = \{x : (x,s,\omega)\in D\}$. In particular, for each $G\in\mathcal{B}(E)$, $\gamma_0(\cdot,\cdot,G)$ is $\mathcal{P}$-measurable. Let $A_0(t) = \int_{E\times(0,t]} H(x,s)\Gamma(dx\times ds)$, and note that

$$\nu_0(C) = E\left[\int_{E\times(0,\infty)} I_C(s)H(x,s)\Gamma(dx\times ds)\right] = E\left[\int_0^\infty I_C(s)dA_0(s)\right], \quad C\in\mathcal{P},$$

so

$$\nu(D) = \int_\Omega\int_0^\infty \gamma_0(s,\omega,D_{(s,\omega)})dA_0(s,\omega)P(d\omega) = E\left[\int_0^\infty \gamma_0(s,\cdot,D_{(s,\cdot)})dA_0(s)\right],$$

for every $D\in\mathcal{P}_E$ which implies

$$E\left[\int_{E\times(0,\infty)} Z(x,s)H(x,s)\Gamma(dx\times ds)\right] = E\left[\int_0^\infty\int_E Z(x,s)\gamma_0(s,\cdot,dx)dA_0(s)\right]$$

for every bounded, predictable Z.

There exists a nondecreasing, right-continuous, predictable process A, such that $A_0 - A$ is a martingale and hence

$$\nu_0(C) = E\left[\int_0^t I_C(s)dA(s)\right], \quad C\in\mathcal{P},$$

and

$$\nu(D) = E\left[\int_0^\infty \gamma_0(s,\omega,D_{(s,\omega)})dA(s)\right].$$

For $G\in\mathcal{B}(E\times(0,\infty))$, define

$$\widehat{\Gamma}(G,\omega) = \int_0^\infty\int_G \frac{1}{H(x,s)}\gamma_0(s,\omega,dx)dA(s),$$

and observe that

$$
\begin{aligned}
E\left[\int_{E\times(0,t]} Z(x,s)H(x,s)\widehat{\Gamma}(dx\times ds)\right] &= E\left[\int_0^t\int_E Z(x,s)\gamma_0(s,dx)dA(s)\right] \\
&= E\left[\int_0^t\int_E Z(x,s)\gamma_0(s,dx)dA_0(s)\right] \\
&= E\left[\int_{(0,t]\times E} Z(x,s)H(x,s)\Gamma(dx\times ds)\right].
\end{aligned}
$$

Let $t \geq 0$ and $r > 0$, and let R be bounded and $\mathcal{F}_t$-measurable. Note that if Z is predictable, then $\tilde{Z}(x,s) = RI_{(t,t+r]}(s)Z(x,s)$ is predictable. It follows that

$$
\begin{aligned}
&E[(M_Z(t+r) - M_Z(t))R] \\
&\quad= E\left[\left(E\left[\int_{E\times[0,t+r]} Z(x,s)H(x,s)\Gamma(dx\times ds)\middle|\mathcal{F}_{t+r}\right]\right.\right. \\
&\qquad\qquad \left.\left. - E\left[\int_{E\times[0,t]} Z(x,s)H(x,s)\Gamma(dx\times ds)\middle|\mathcal{F}_t\right]\right)R\right] \\
&\qquad\quad -E\left[\int_{E\times(t,t+r]} Z(x,s)H(x,s)\widehat{\Gamma}(dx\times ds)R\right] \\
&\quad= E\left[\left(\int_{E\times[0,t+r]} Z(x,s)H(x,s)\Gamma(dx\times ds)\right.\right. \\
&\qquad\qquad \left.\left. - \int_{E\times[0,t]} Z(x,s)H(x,s)\Gamma(dx\times ds)\right)R\right] \\
&\qquad\quad -E\left[\int_{E\times(t,t+r]} Z(x,s)H(x,s)\widehat{\Gamma}(dx\times ds)R\right] \\
&\quad= E\left[\int_{E\times(t,t+r]} RZ(x,s)H(x,s)\Gamma(dx\times ds)\right] \\
&\qquad\quad -E\left[\int_{E\times(t,t+r]} RZ(x,s)H(x,s)\widehat{\Gamma}(dx\times ds)\right] \\
&\quad= 0,
\end{aligned}
$$

so M_Z is a martingale. □

6.2 Conditioning and random measures

Let Γ be an adapted random measure on $E\times[0,\infty)$ defined on $(\Omega,\mathcal{F},P)$ and satisfying $\Gamma(E\times[0,t]) < \infty$ a.s., for each $t > 0$. Let V be a stochastic process defined on $(\Omega,\mathcal{F},P)$, indexed by $E\times[0,\infty)$. V is *measurable* if the the mapping $(x,t,\omega) \in E\times[0,\infty)\times\Omega \to V(x,t,\omega) \in \mathbb{R}$ satisfies $V^{-1}(C) \in \mathcal{B}(E)\times\mathcal{B}[0,\infty)\times\mathcal{F}$ for each $C \in \mathcal{B}(\mathbb{R})$. Let $\{\mathcal{F}_t\}$ be a filtration in $\mathcal{F}$. Γ is *adapted* to $\{\mathcal{F}_t\}$ if $\Gamma(G\times[0,s])$ is $\mathcal{F}_t$-measurable for all $G \in \mathcal{B}(E)$ and $0 \leq s \leq t$.

Let $\mathcal{O}$ be the optional σ-algebra, that is, the smallest σ-algebra of sets in $[0,\infty)\times\Omega$ such that for each $\{\mathcal{F}_t\}$-adapted, right continuous process X, the mapping $(t,\omega)\to X(t,\omega)$ is $\mathcal{O}$-measurable. A variant of the optional projection theorem (see, for example, [11], Corollary 2.4.5) ensures that if V is nonnegative, then there exists a $\mathcal{B}(E)\times\mathcal{O}$-measurable function $\widehat{V}$ such that

$$E[V(x,\tau)|\mathcal{F}_\tau]=\widehat{V}(x,\tau)\quad a.s.,$$

for every finite $\{\mathcal{F}_t\}$-stopping time τ. We will simply write $E[V(x,t)|\mathcal{F}_t]$ for $\widehat{V}(x,t)$.

Lemma 6.2 *Let $\{\mathcal{F}_t\}$ be a complete, right-continuous filtration. Let Γ be an $\{\mathcal{F}_t\}$-adapted random measure on $E\times[0,\infty)$, satisfying $E[\Gamma(E\times[0,t])]<\infty$, for every $t>0$. Let V be a nonnegative, measurable process on $E\times[0,\infty)$. Suppose $E[\int_{E\times[0,t]}V(x,s)\Gamma(dx\times ds)]<\infty$, for all $t\geq 0$. Then*

$$M_V(t)=E[\int_{E\times[0,t]}V(x,s)\Gamma(dx\times ds)|\mathcal{F}_t]-\int_{E\times[0,t]}E[V(x,s)|\mathcal{F}_s]\Gamma(dx\times ds)$$

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

$$E\left[\int_{E\times[0,t]}V(x,s)\Gamma(dx\times ds)\right]=E\left[\int_{E\times[0,t]}E[V(x,s)|\mathcal{F}_s]\Gamma(dx\times ds)\right].\tag{6.18}$$

Proof. The collection of bounded V for which (6.18) holds is a linear space that is closed under bounded, pointwise convergence. Let ξ be an $\mathbb{R}$-valued random variable, $C\in\mathcal{B}(E)$, and $0\leq a<b$. Define $V(x,s)=\xi I_C(x)I_{(a,b]}(s)$. Then, letting $\pi=\{s_k\}$ be a partition of $(a\wedge t,b\wedge t]$ and $\gamma(s)=\min\{s_k\in\pi:s_k\geq s\}$

$$\begin{aligned}
E\left[\int_{E\times[0,t]}V(x,s)\Gamma(dx\times ds)\right] &= E[\xi\Gamma(C\times(a\wedge t,b\wedge t])]\\
&= \sum E[\xi\Gamma(C\times(s_k,s_{k+1}])]\\
&= \sum E[E[\xi|\mathcal{F}_{s_{k+1}}]\Gamma(C\times(s_k,s_{k+1}])]\\
&= E\left[\int_{E\times[0,t]}E[\xi|\mathcal{F}_{\gamma(s)}]I_C(x)I_{(a,b]}(s)\Gamma(dx\times ds)\right]\\
&= E\left[\int_{E\times[0,t]}E[V(x,s)|\mathcal{F}_{\gamma(s)}]\Gamma(dx\times ds)\right].
\end{aligned}$$

Letting $\max(s_{k+1}-s_k)\to 0$, $\gamma(s)\to s+$, by the right continuity of $E[\xi|\mathcal{F}_s]$, we have (6.18). A monotone class argument (see, for example, Corollary A.4.4 in [11]) then implies (6.18) for all bounded, measurable processes V, and the extension to all positive V follows by the monotone convergence theorem.

Let Z be bounded and $\mathcal{F}_t$-measurable. Then

$$\begin{aligned}
&E[Z(M_V(t+h)-M_V(t))]\\
&\quad= E\Bigg[Z\Big(\int_{E\times[0,t+h]}V(x,s)\Gamma(dx\times ds)-\int_{E\times[0,t]}V(x,s)\Gamma(dx\times ds)
\end{aligned}$$

$$
\begin{aligned}
&\qquad - \int_{E\times(t,t+h]} E[V(x,s)|\mathcal{F}_s]\Gamma(dx\times ds)\Big)\Bigg] \\
&= E\left[\int_{E\times(t,t+h]} ZV(x,s)\Gamma(dx\times ds)\right. \\
&\qquad \left. - \int_{E\times(t,t+h]} E[ZV(x,s)|\mathcal{F}_s]\Gamma(dx\times ds)\right] \\
&= 0,
\end{aligned}
$$

where the last equality follows by (6.18). This result implies that M_V is a martingale. □

References

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

[2] A.G. Bhatt and R.L. Karandikar, Invariant measures and evolution equations for Markov processes, *Ann. Probab.* **21** (1993), 2246–2268.

[3] A.G. Bhatt, G. Kallianpur and R.L. Karandikar, Uniqueness and robustness of solutions of measure-valued equations of nonlinear filtering, *Ann. Probab.* **23** (1995), 1895–1938.

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

[5] A.G. Bhatt and R.L. Karandikar, Characterization of the optimal filter: the non-Markov case, *Stochastics Stochastics Rep.,* **66** (1999), 177–204.

[6] V.S. Borkar and M.K. Ghosh, Ergodic control of multidimensional diffusions I: The existence results, *SIAM J. Control Optim.* **26** (1988), 112–126.

[7] J. G. Dai and J. M. Harrison, Steady-State Analysis of RBM in a Rectangle: Numerical Methods and a Queueing Application, *Ann. Appl. Probab.* **1** (1991), 15-35.

[8] M. H. A. Davis and A. R. Norman, Portfolio selection with transaction costs. *Math. Oper. Res.* **15** (1990), 676–713.

[9] P. Donnelly and T.G. Kurtz. Particle representations for measure-valued population models. *Ann. Probab.* **27** (1999), 166–205.

[10] P.E. Echeverria, A criterion for invariant measures of Markov processes, *Z. Wahrsch. verw. Gebiete* **61** (1982), 1-16.

[11] S.N. Ethier and T.G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley, New York (1986).

[12] J. Jacod and A.N. Shiryaev, *Limit Theorems for Stochastic Processes,* Springer-Verlag, Berlin (1987).

[13] A. Jakubowski, A non-Skorohod topology on the Skorohod space. *Electron. J. Probab.* 2 (1997), paper no. 4, 21 pp.

[14] I. Karatzas, A class of singular stochastic control problems. *Adv. in Appl. Probab.* **15** (1983), 225–254.

[15] T.G. Kurtz, Martingale problems for constrained Markov processes, in *Recent Advances in Stochastic Calculus*, J.S. Baras, V. Mirelli, eds., Springer–Verlag, New York, 1990.

[16] T.G. Kurtz, A control formulation for constrained Markov processes, in *Mathematics of Random Media*, Lectures in Applied Mathematics, **27**, 1991.

[17] T.G. Kurtz, Random time changes and convergence in distribution under the Meyer-Zheng conditions, *Ann. Probab.*, **19** (1991), 1010–1034.

[18] T.G. Kurtz, Averaging for martingale problems and stochastic approximation, *Proceedings of the Joint U.S.–French Workshop on Applied Stochastic Analysis. Springer Lecture Notes in Control and Inf. Sci.* **177** (1992), 186-209.

[19] T.G. Kurtz. Martingale problems for conditional distributions of Markov processes. *Electron. J. Probab.* **3** (1998), Paper 9, 29 pp.

[20] T.G. Kurtz and D.L. Ocone. Unique characterization of conditional distributions in nonlinear filtering. *Ann. Probab.* **16** (1988), no. 1, 80–107.

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

[22] T.G. Kurtz and R.H. Stockbridge, Martingale problems and linear programs for singular control, *Thirty-seventh annual Allerton conference on communication, control, and computing (Monticello, Ill., 1999)*, 11-20, Univ. Illinois, Urbana-Champaign, Ill.

[23] H.L. Royden, *Real Analysis, 2nd ed.,* MacMillan, New York, 1968.

[24] S.E. Shreve, An introduction to singular stochastic control. *Stochastic differential systems, stochastic control theory and applications (Minneapolis, Minn., 1986), 513–528, IMA Vol. Math. Appl.*, **10**, Springer, New York-Berlin, 1988.

[25] S.E. Shreve and H.M. Soner, Optimal investment and consumption with transaction costs. *Ann. Appl. Probab.* **4** (1994), 609–692.

[26] H.M. Soner and S.E. Shreve, Regularity of the value function for a two-dimensional singular stochastic control problem. *SIAM J. Control Optim.* **27** (1989), 876–907.

[27] R.H. Stockbridge, Time-average control of martingale problems: Existence of a stationary solution, *Ann. Probab.* **18** (1990), 190–205.

[28] R.H. Stockbridge, Time-average control of martingale problems: A linear programming formulation, *Ann. Probab.* **18** (1990), 206–217.

[29] A. Weiss, Invariant measures of diffusions in bounded domains, Ph.D. dissertation, New York University, 1981.