

## A DYNAMIC PROGRAMMING ALGORITHM FOR THE OPTIMAL CONTROL OF PIECEWISE DETERMINISTIC MARKOV PROCESSES\*

ANTHONY ALMUDEVAR†

**Abstract.** A piecewise deterministic Markov process (PDP) is a continuous time Markov process consisting of continuous, deterministic trajectories interrupted by random jumps. The trajectories may be controlled with the object of minimizing the expected costs associated with the process. A method of representing this controlled PDP as a discrete time decision process is presented, allowing the value function for the problem to be expressed as the fixed point of a dynamic programming operator. Decisions take the form of trajectory segments. The expected costs may then be minimized through a dynamic programming algorithm, rather than through the solution of the Bellman–Hamilton–Jacobi equation, assuming the trajectory segments are numerically tractable. The technique is applied to the optimal capacity expansion problem, that is, the problem of planning the construction of new production facilities to meet rising demand.

**Key words.** piecewise deterministic Markov process, dynamic programming, capacity expansion

**AMS subject classifications.** 93E20, 60J75

**PII.** S0363012999364474

**1. Introduction.** In this paper a technique for minimizing expected costs associated with piecewise deterministic Markov processes (PDPs) is developed. Such processes may be described as continuous time Markov processes consisting of continuous, deterministic trajectories interrupted by random jumps. A comprehensive definition and theoretical development of these processes can be found in Davis [4]. Many problems in operations research can be naturally expressed in this framework; hence there is a great deal of interest in optimization problems associated with these processes.

A PDP is usually defined on a state space  $E \subset \mathbb{R}^p$  partitioned into a boundary  $E_\delta$  and interior  $E_o$ , although the state space definition in [4] is somewhat more general. We let  $\mathcal{E}$  denote the Borel subsets of  $E$ , and we will let  $\mathcal{P}(E)$  be the space of probability measures on the measurable space  $(E, \mathcal{E})$ , endowed with the topology of weak convergence. Under suitable regularity conditions a PDP can be uniquely determined by a vector field  $f : E \rightarrow \mathbb{R}^p$ , an intensity function  $\lambda : E \rightarrow \mathbb{R}^+$ , and stochastic kernels  $q_o : E_o \rightarrow \mathcal{P}(E)$  and  $q_\delta : E_\delta \rightarrow \mathcal{P}(E)$ . Between jumps the PDP  $\hat{x}(t)$  obeys  $d\hat{x}(t)/dt = f(\hat{x}(t))$ , and jumps occur at rate  $\lambda(x)$  when the process is at state  $x$ , independently of the process history. If a jump occurs at  $x \in E_o$ , the process is transferred immediately to a new state given randomly by probability measure  $q_o(dx | x)$ . If the process reaches the boundary at  $x \in E_\delta$ , the process is transferred immediately to a new state given randomly by probability measure  $q_\delta(dx | x)$ . We will always assume that  $q_o(E_o | x) = 1$  and  $q_\delta(E_o | x) = 1$ .

A controlled PDP is defined when the quadruple  $(f, \lambda, q_o, q_\delta)$  is allowed to depend on a control parameter  $u$ . In addition, cost is assumed at a rate  $l_o(x, u)$  when the process is at  $x \in E_o$  and control  $u$  is applied, and a discrete cost  $l_\delta(x, u)$  is assumed

---

\*Received by the editors November 10, 1999; accepted for publication (in revised form) February 19, 2001; published electronically July 25, 2001. This research was supported by the Natural Sciences and Engineering Research Council of Canada.

<http://www.siam.org/journals/sicon/40-2/36447.html>

†Department of Mathematics and Computing Science, Saint Mary's University, Halifax, NS, Canada, B3H 3C3 ([anthony.almudevar@stmarys.ca](mailto:anthony.almudevar@stmarys.ca)).

when the process reaches the boundary at  $x \in E_\delta$  and control  $u$  is applied. A *control policy*  $\Phi$  is equivalent to a specification for each  $x \in E$  of an open loop continuous time control function to be applied from  $x$  until the next jump (Vermes [11]).

If we define  $J_\Phi(x)$  to be the expected cost under control policy  $\Phi$  from initial state  $x$ , possibly under geometric discounting, the *value function* is then defined as

$$J^*(x) = \inf_{\Phi} J_\Phi(x),$$

where the infimum is taken over all admissible control policies. The object is to find, if it exists, a control policy whose expected cost achieves this infimum.

In the existing literature the value function for this problem is typically given as a solution to a Bellman–Hamilton–Jacobi (BHJ) equation. In [11] a limiting form of the BHJ equation is given as a necessary and sufficient optimality condition. In Dempster and Ye [9] a generalized BHJ equation, expressed in terms of the generalized Clarke gradient (Clarke [3]), is given as a necessary and sufficient optimality condition. In Soner [10] a viscosity solution approach to the BHJ equation is proposed, and, more recently, the viscosity solution to the BHJ equation has been developed in Davis and Farid [8], which has advantages with respect to the availability of numerical methods for solution.

In this paper we use an approach similar to that introduced by Davis [5] and developed in [9] and Davis [6], in which the problem is reformulated in terms of an imbedded discrete time process, in which a stage consists of the intrajump deterministic portion of the process. The principal difference is that in the approach proposed in this article the problem remains in the discrete time domain up to and including the solution algorithm. The concept of a continuously applied control parameter will play no role. Instead, a discrete time decision process is defined in which a decision consists of the selection of a trajectory segment, in this way constructing the deterministic trajectory in a piecewise fashion. This means that the BHJ equation plays no role. Ultimately, the value function is calculable as a fixed point of a dynamic programming operator in discrete time. Here we do not admit direct control over the cost function and the jump rate, unlike the other models cited in the above literature, although in principle the methodology could be extended to incorporate the cost function and the jump rate into the action space.

Apart from the more limited control, this allows a more uniform approach to the calculation of optimal policies and a weakening of regularity conditions. For the generalized BHJ equation in [9], conditions are imposed which guarantee that the value function is Lipschitz, which excludes many problems of practical importance (see [6]). The viscosity solution approach allows milder assumptions. In [8] the state space is required to be bounded, and the cost rate and jump rate are assumed to be bounded and uniformly continuous. In comparison, in this article the state space need not be bounded, the jump rate is bounded but not necessarily uniformly continuous, and the cost rate may be lower semicontinuous and need not be bounded. In fact, conditions are placed only on suitably defined integrals of the cost rate (see section 2). As for the trajectory, in the context of the BHJ equation the vector field  $f$  is generally assumed to be Lipschitz. In the approach presented here there is no explicit vector field and no other restrictions other than that the path can be constructed in a piecewise manner from trajectories selected from a compact set. This admits a wider variety of control structures, including certain types of impulse controls.

In section 2, we define a discrete time decision process imbedded in the PDP and obtain conditions under which the resulting transition measure will be continuous

on the state-action product space. In section 3, we discuss some results for discrete time decision processes from Bertsekas and Shreve [1] which may be applied to the problem under consideration here. In addition, with some additional assumptions we show that the dynamic programming operator is a contraction mapping. In section 4 we show how this may be applied to the capacity expansion problem considered in Davis et al. [7]. Section 5 contains some concluding remarks and possible extensions of this work.

**2. Reduction of a PDP to a discrete time process.** Let  $E \subset \mathfrak{R}^p$  be a state space containing a boundary  $E_\delta$ . Let  $E_o = E - E_\delta$  be the interior of  $E$ . Possibly,  $E_\delta = \emptyset$ , the empty set. We also have the intensity function and stochastic kernels  $(\lambda, q_o, q_\delta)$  as defined in section 1, all assumed to be Borel measurable mappings.

Then let  $I_T$  be a time scale interval  $[0, T]$  if  $T < \infty$  and  $[0, \infty)$  if  $T = \infty$ . Let  $A$  be an action space consisting of a family of continuous trajectories  $\alpha : I_T \rightarrow \mathfrak{R}^p$  with  $\alpha(0) = 0$ . It will be assumed that  $A$  is a compact metric space in which convergence implies pointwise convergence. We define

$$B(x, \alpha) = \inf\{t \in I_T : x + \alpha(t) \in E_\delta\},$$

which is the time taken for the trajectory  $x + \alpha(t)$  to reach the boundary, and let

$$t_f(x, \alpha) = \min\{B(x, \alpha), T\},$$

adopting the convention that  $\inf \emptyset = \infty$ . For each  $x \in E$  let  $A_x \subset A$  be a subset of trajectories available at state  $x$ , which gives the *state-action space*

$$\Gamma = \{(x, \alpha) \in E \times A : \alpha \in A_x\}.$$

We assume that  $x + \alpha(t) \in E$  when  $t \leq t_f(x, \alpha)$  for all  $(x, \alpha) \in \Gamma$ . Generally, the following condition will be satisfied:

$$(A.1) \quad x + \alpha(t) = x + \alpha(B(x, \alpha)) \quad \forall t \geq B(x, \alpha), t \in I_T, \text{ when } (x, \alpha) \in \Gamma \text{ and } B(x, \alpha) < \infty,$$

which implies that a trajectory comes to rest upon reaching the boundary. In addition, (A.1) implies  $x + \alpha(B(x, \alpha)) \in E_\delta$  for all  $(x, \alpha) \in \Gamma$  and that the only admissible action when  $x \in E_\delta$  is  $\alpha \equiv 0$ .

We can define iteratively the continuous time process  $\{\hat{x}(t) \in E : t \geq 0\}$  and the imbedded discrete time decision process  $\{(\hat{x}_n, \hat{\alpha}_n) \in \Gamma : n \geq 0\}$  with the associated event time process  $\{\hat{t}_n \geq 0 : n \geq 0\}$ . Suppose we have state  $\hat{x}_n \in E_o$ , decision  $\hat{\alpha}_n \in A_{x_n}$ , and time  $\hat{t}_n$ . The process then follows the trajectory

$$(2.1) \quad \hat{x}(t) = \hat{x}_n + \hat{\alpha}_n(t - \hat{t}_n), \quad t \geq \hat{t}_n,$$

until time  $\hat{t}_n + t_f(\hat{x}_n, \hat{\alpha}_n)$ , unless a random jump occurs along the trajectory before then, say, at time  $t' \in (\hat{t}_n, \hat{t}_n + t_f(\hat{x}_n, \hat{\alpha}_n))$ , in which case (2.1) holds until  $t'$ . These jumps occur at rate  $\lambda(x)$  when the process is in state  $x \in E_o$ , independently of the process history. If such a jump occurs at state  $x'$ , then the new state  $\hat{x}_{n+1} \in E_o$  is given randomly by the distribution  $q_o(dx | x')$ , and we set  $\hat{t}_{n+1} = t'$ . If no jump occurs before  $\hat{t}_n + t_f(\hat{x}_n, \hat{\alpha}_n)$  then set  $\hat{t}_{n+1} = \hat{t}_n + t_f(\hat{x}_n, \hat{\alpha}_n)$ . In this case, if the process has reached the boundary at state  $x' \in E_\delta$  (i.e.,  $B(\hat{x}_n, \hat{\alpha}_n) < \infty$ ), then the new state  $\hat{x}_{n+1} \in E_o$  is given randomly by the distribution  $q_\delta(dx | x')$ . Otherwise, if the end of the trajectory segment  $\hat{\alpha}_n$  has been reached before the boundary (i.e.,

$B(\hat{x}_n, \hat{\alpha}_n) = \infty, T < \infty$ ), then set  $\hat{x}_{n+1} = \hat{x}_n + \hat{\alpha}_n(\hat{t}_{n+1} - \hat{t}_n)$ . A new decision  $\hat{\alpha}_{n+1} \in A_{\hat{x}_{n+1}}$  is then made. An initial state and decision  $(\hat{x}_0, \hat{\alpha}_0) \in \Gamma$  is specified, with  $\hat{t}_0 = 0$ . If  $\hat{x}_0 \in E_\delta$ , we will set  $\hat{t}_1 = \hat{t}_0 = 0, \hat{\alpha}_0 \equiv 0$ , and  $\hat{x}_1$  will be determined by  $q_\delta(dx | \hat{x}_0)$ . Then  $\hat{x}_k \in E_o$  for  $k \geq 1$ .

This defines the transition measure  $Q : \Gamma \rightarrow \mathcal{P}(E)$  for the process  $(\hat{x}_n, \hat{\alpha}_n)$ , where  $Q(K | x, \alpha)$  is the probability that  $\hat{x}_{n+1} \in K$  given that trajectory  $\hat{\alpha}_n = \alpha$  is selected at state  $\hat{x}_n = x$ . Assuming (A.1) holds, this is given explicitly by

$$\begin{aligned}
 Q(K | x, \alpha) &= \int_0^{t_f(x, \alpha)} q_o(K | x + \alpha(t)) \lambda(x + \alpha(t)) \exp(-\Lambda(t, x, \alpha)) dt \\
 &+ q_\delta(K | x + \alpha(B(x, \alpha))) \exp(-\Lambda(B(x, \alpha), x, \alpha)) I\{B(x, \alpha) < \infty\} \\
 (2.2) \quad &+ I\{x + \alpha(T) \in K\} \exp(-\Lambda(T, x, \alpha)) I\{B(x, \alpha) = \infty, T < \infty\},
 \end{aligned}$$

where

$$\Lambda(t, x, \alpha) = \int_0^t \lambda(x + \alpha(w)) dw.$$

Here,  $I\{S\}$  is the indicator function of set  $S$ . Since we assume  $q_o(E_o | x) = 1$  and  $q_\delta(E_o | x) = 1$ , we necessarily have  $Q(E_o | x, \alpha) = 1$  for  $(x, \alpha) \in \Gamma$ . Note also that if  $x \in E_\delta$ , we have  $\alpha \equiv 0, t_f(x, \alpha) = B(x, \alpha) = 0$ , and  $Q(K | x, \alpha) = q_\delta(K | x)$ . It will be useful to know when the transition measure is continuous with respect to weak convergence on  $\Gamma$  (with  $E \times A$  assuming the product topology). We prove below that  $Q$  will be continuous under the following assumptions:

- (B.1)  $q_o(dx | x)$  is continuous on  $E_o$  with respect to weak convergence.
- (B.2)  $q_\delta(dx | x)$  is continuous on  $E_\delta$  with respect to weak convergence.
- (B.3)  $\lambda$  is continuous on  $E, \lambda \leq M_\lambda$  for some  $M_\lambda < \infty$ .
- (B.4) The sets  $B_1 = \{(x, \alpha) \in \Gamma : B(x, \alpha) < \infty\}$  and  $B_2 = \{(x, \alpha) \in \Gamma : B(x, \alpha) = \infty\}$  are both closed.
- (B.5)  $B(x, \alpha)$  is continuous on  $B_1$ .

*Remark.* If a nontrivial boundary is present, a special condition is typically necessary for the continuity of  $Q$  to hold. Generally, some assumption which governs the behavior of the trajectory near the boundary is required. Informally, these assumptions typically require that if a trajectory approaches the boundary, it does so in some direct manner. In [11] the minimum velocity in the direction normal to the boundary is bounded away from 0. This condition is weakened in [8] to require only that where the boundary is approachable it is approachable nontangentially. Assumptions (B.4) and (B.5) are used here to govern trajectory behavior near the boundary. They will not be natural to many problems and are not satisfied by the capacity expansion problem considered in [7]. However, we show in section 4 how a reasonable redefinition of the problem can force (B.4) and (B.5) to hold.

**THEOREM 2.1.** *If assumptions (A.1) and (B.1)–(B.5) hold, then  $Q(dx | x, \alpha)$  is continuous on  $\Gamma$  with respect to weak convergence.*

*Proof.* Let  $\gamma = \{(x_n, \alpha_n) : n \geq 1\}$  be a convergent sequence in  $\Gamma$  with limit  $(x_0, \alpha_0)$ . We then have

$$\lim_{n \rightarrow \infty} x_n + \alpha_n(t) = x_0 + \alpha_0(t) \quad \forall t \in I_T,$$

and hence by (B.3)

$$(2.3) \quad \lim_{n \rightarrow \infty} \lambda(x_n + \alpha_n(t)) = \lambda(x_0 + \alpha_0(t)) \quad \forall t \in I_T.$$

By (B.3)  $\lambda$  is bounded, so applying the dominated convergence theorem gives

$$(2.4) \quad \lim_{n \rightarrow \infty} \Lambda(t, x_n, \alpha_n) = \Lambda(t, x_0, \alpha_0) \quad \forall t \in I_T.$$

Next, recall that if  $\{\mu_n : n \geq 1\}$  is any sequence of probability measures in  $\mathcal{P}(E)$ , an equivalent definition of weak convergence of the sequence to a probability measure  $\mu_0$  is

$$\liminf_{n \rightarrow \infty} \mu_n(K) \geq \mu_0(K) \quad \forall \text{ open sets } K$$

(see, for example, Theorem 29.1 in Billingsley [2]), so it suffices to show that

$$(2.5) \quad \liminf_{n \rightarrow \infty} Q(K | x_n, \alpha_n) \geq Q(K | x_0, \alpha_0) \quad \forall \text{ open sets } K \in \mathcal{E},$$

for each convergent sequence  $\gamma$ . Since  $B_1$  and  $B_2$  are closed, we may assume that  $\gamma \subset B_1$  or  $\gamma \subset B_2$ . We now examine separately the three following cases.

*Case 1:*  $T = \infty$ ,  $\gamma \subset B_2$ . In this case we have

$$Q(K | x_n, \alpha_n) = \int_0^\infty q_o(K | x_n + \alpha_n(t)) \lambda(x_n + \alpha_n(t)) \exp(-\Lambda(t, x_n, \alpha_n)) dt.$$

By (B.1), (2.3), and (2.4) we may assert for open  $K$

$$(2.6) \quad \begin{aligned} & \liminf_{n \rightarrow \infty} q_o(K | x_n + \alpha_n(t)) \lambda(x_n + \alpha_n(t)) \exp(-\Lambda(t, x_n, \alpha_n)) \\ & \geq q_o(K | x_0 + \alpha_0(t)) \lambda(x_0 + \alpha_0(t)) \exp(-\Lambda(t, x_0, \alpha_0)); \end{aligned}$$

hence (2.5) holds by Fatou's lemma.

*Case 2:*  $T < \infty$ ,  $\gamma \subset B_2$ . In this case we have

$$(2.7) \quad \begin{aligned} Q(K | x_n, \alpha_n) &= \int_0^T q_o(K | x_n + \alpha_n(t)) \lambda(x_n + \alpha_n(t)) \exp(-\Lambda(t, x_n, \alpha_n)) dt \\ &+ I\{x_n + \alpha_n(T) \in K\} \exp(-\Lambda(T, x_n, \alpha_n)) \end{aligned}$$

for all  $n \geq 0$ . Using an argument similar to that used for Case 1, we have

$$(2.8) \quad \begin{aligned} & \liminf_{n \rightarrow \infty} \int_0^T q_o(K | x_n + \alpha_n(t)) \lambda(x_n + \alpha_n(t)) \exp(-\Lambda(t, x_n, \alpha_n)) dt \\ & \geq \int_0^T q_o(K | x_0 + \alpha_0(t)) \lambda(x_0 + \alpha_0(t)) \exp(-\Lambda(t, x_0, \alpha_0)) dt \end{aligned}$$

for all open sets  $K \in \mathcal{E}$ . Then

$$\liminf_{n \rightarrow \infty} I\{x_n + \alpha_n(T) \in K\} \geq I\{x_0 + \alpha_0(T) \in K\}$$

for open  $K \in \mathcal{E}$ , which, when combined with (2.4), (2.7), and (2.8), gives (2.5) for Case 2.

Case 3:  $\gamma \subset B_1$ . In this case we necessarily have  $t_f(x_n, \alpha_n) = B(x_n, \alpha_n)$ ,  $n \geq 0$ , so that

$$Q(K | x_n, \alpha_n) = \int_0^{B(x_n, \alpha_n)} q_o(K | x_n + \alpha_n(t)) \lambda(x_n + \alpha_n(t)) \exp(-\Lambda(t, x_n, \alpha_n)) dt \\ + q_\delta(K | x_n + \alpha_n(B(x_n, \alpha_n))) \exp(-\Lambda(B(x_n, \alpha_n), x_n, \alpha_n))$$

for all  $n \geq 0$ . By assumption (B.5)  $B(x_n, \alpha_n) \rightarrow_n B(x_0, \alpha_0)$ . Then, using (2.6), we have

$$\liminf_{n \rightarrow \infty} q_o(K | x_n + \alpha_n(t)) \lambda(x_n + \alpha_n(t)) \exp(-\Lambda(t, x_n, \alpha_n)) I\{t \leq B(x_n, \alpha_n)\} \\ \geq q_o(K | x_0 + \alpha_0(t)) \lambda(x_0 + \alpha_0(t)) \exp(-\Lambda(t, x_0, \alpha_0)) I\{t < B(x_0, \alpha_0)\}$$

for open  $K \in \mathcal{E}$ , so by Fatou's lemma

$$\liminf_{n \rightarrow \infty} \int_0^{B(x_n, \alpha_n)} q_o(K | x_n + \alpha_n(t)) \lambda(x_n + \alpha_n(t)) \exp(-\Lambda(t, x_n, \alpha_n)) dt \\ (2.9) \quad \geq \int_0^{B(x_0, \alpha_0)} q_o(K | x_0 + \alpha_0(t)) \lambda(x_0 + \alpha_0(t)) \exp(-\Lambda(t, x_0, \alpha_0)) dt$$

for all open sets  $K \in \mathcal{E}$ . By assumption (A.1) we must have

$$\lim_{n \rightarrow \infty} x_n + \alpha_n(B(x_n, \alpha_n)) = x_0 + \alpha_0(B(x_0, \alpha_0)),$$

and by assumption (B.2)

$$\liminf_{n \rightarrow \infty} q_\delta(K | x_n + \alpha_n(B(x_n, \alpha_n))) \geq q_\delta(K | x_0 + \alpha_0(B(x_0, \alpha_0)))$$

for all open sets  $K \in \mathcal{E}$ . We then have

$$\lim_{n \rightarrow \infty} \lambda(x_n + \alpha_n(t)) I\{t \leq B(x_n, \alpha_n)\} = \lambda(x_0 + \alpha_0(t)) I\{t \leq B(x_0, \alpha_0)\},$$

except possibly at  $t = B(x_0, \alpha_0)$ . The sequence  $\{B(x_n, \alpha_n) : n \geq 1\}$  is bounded since  $B_1$  is closed. Then with assumption (B.3) the dominated convergence theorem applies, giving

$$\lim_{n \rightarrow \infty} \Lambda(B(x_n, \alpha_n), x_n, \alpha_n) = \Lambda(B(x_0, \alpha_0), x_0, \alpha_0),$$

so that

$$\liminf_{n \rightarrow \infty} q_\delta(K | x_n + \alpha_n(B(x_n, \alpha_n))) \exp(-\Lambda(B(x_n, \alpha_n), x_n, \alpha_n)) \\ \geq q_\delta(K | x_0 + \alpha_0(B(x_0, \alpha_0))) \exp(-\Lambda(B(x_0, \alpha_0), x_0, \alpha_0))$$

for all open sets  $K \in \mathcal{E}$ , which with (2.9) gives (2.5) for Case 3, which completes the proof.  $\square$

With respect to assumption (B.5), assumption (A.1) is sufficient to guarantee the lower semicontinuity of  $B(x, \alpha)$  on  $B_1$ , as shown in Lemma 2.2 below, but upper semicontinuity must be verified separately.

LEMMA 2.2. *If assumption (A.1) holds,  $B(x, \alpha)$  is lower semicontinuous on  $B_1$ .*

*Proof.* Suppose that  $\gamma = \{(x_n, \alpha_n) : n \geq 1\}$  is a convergent sequence in  $B_1$  with limit  $(x_0, \alpha_0) \in B_1$ . We show that

$$(2.10) \quad \liminf_{n \rightarrow \infty} B(x_n, \alpha_n) \geq B(x_0, \alpha_0)$$

for any such sequence. Suppose there exists an infinite subsequence  $\{(x_{n_k}, \alpha_{n_k}) : k \geq 1\}$  and a  $\beta < B(x_0, \alpha_0)$  such that  $B(x_{n_k}, \alpha_{n_k}) \leq \beta$  for all  $k \geq 1$ . Then

$$(2.11) \quad \lim_{k \rightarrow \infty} x_{n_k} + \alpha_{n_k}(B(x_0, \alpha_0)) = x_0 + \alpha_0(B(x_0, \alpha_0)),$$

and by assumption (A.1) and the fact that  $B(x_{n_k}, \alpha_{n_k}) \leq \beta < B(x_0, \alpha_0)$ ,  $k \geq 1$ , we must have

$$(2.12) \quad \begin{aligned} \lim_{k \rightarrow \infty} x_{n_k} + \alpha_{n_k}(B(x_0, \alpha_0)) &= \lim_{k \rightarrow \infty} x_{n_k} + \alpha_{n_k}(\beta) \\ &= x_0 + \alpha_0(\beta). \end{aligned}$$

However, (2.11) and (2.12) are contradictory since  $x_0 + \alpha_0(B(x_0, \alpha_0)) \in E_\delta$ , but  $x_0 + \alpha_0(\beta) \in E_o$ ; hence any convergent sequence must satisfy (2.10).  $\square$

Finally, we assume there is a nonnegative expected cost  $g : \Gamma \rightarrow \mathfrak{R}^+$  associated with each stage. This cost may be specified by letting  $C_T$  be the family of measurable functions  $c : I_T \rightarrow \mathfrak{R}^+$ . The cost of a stage is then determined by a mapping  $h_o : \Gamma \rightarrow C_T$  which represents the rate at which cost is assumed at a time  $t$  after decision  $\alpha$  is made from state  $x$ . We may also have a boundary cost  $h_\delta(x)$ ,  $x \in E_\delta$ , assumed when the process reaches the boundary at  $x$ . Then if  $W_{(x,\alpha)}$  is the random time spent in the stage, the cost assumed in the stage given  $W_{(x,\alpha)} = w$  is

$$H_{(x,\alpha)}(w) = \int_0^w h_o(t | x, \alpha) dt + h_\delta(x + \alpha(B(x, \alpha)))I\{w = B(x, \alpha), B(x, \alpha) < \infty\}.$$

Then  $g$  is given by

$$(2.13) \quad \begin{aligned} g(x, \alpha) &= E[H_{(x,\alpha)}(W_{(x,\alpha)})] \\ &= \int_0^{t_f(x,\alpha)} h_o(t | x, \alpha) \exp(-\Lambda(t, x, \alpha)) dt \\ (2.14) \quad &+ h_\delta(x + \alpha(B(x, \alpha))) \exp(-\Lambda(B(x, \alpha), x, \alpha))I\{B(x, \alpha) < \infty\}. \end{aligned}$$

In the following discussion any regularity condition will be placed on  $g$  directly.

**3. Optimization for lower semicontinuous costs.** We give a general definition (following [1]) of a stochastic discrete time decision process  $\{(\hat{x}_n, \hat{\alpha}_n) : n \geq 0\}$ , where  $\hat{x}_n$  and  $\hat{\alpha}_n$  are elements of a state space and action space  $E$  and  $A$ , both assumed to be Borel spaces. Let  $\mathcal{P}(E)$  and  $\mathcal{P}(A)$  be the space of all probability measures on the Borel sets of  $E$  and  $A$ , respectively, endowed with the topology of weak convergence. For each  $x \in E$  we assume that there is a set of available actions  $A_x \subset A$ . We then have state-action space

$$\Gamma = \{(x, \alpha) \in E \times A : \alpha \in A_x\},$$

where  $E \times A$  is endowed with the product topology (and is also a Borel space). We assume there is a stochastic kernel  $Q(dx | x, \alpha)$  which is a Borel measurable mapping

from  $\Gamma$  to  $\mathcal{P}(E)$ . Finally, we have a lower semianalytic cost function  $g : \Gamma \rightarrow \mathfrak{R}^+$ . Define a policy

$$\Phi = \{\tilde{\phi}_n : n \geq 0\}$$

as a sequence of stochastic kernels  $\tilde{\phi}_n(dy | x_0, \alpha_0, \dots, x_{n-1}, \alpha_{n-1}, x_n)$  which are universally measurable mappings from  $(\times^n \Gamma) \times E$  to  $\mathcal{P}(A)$  satisfying

$$\tilde{\phi}_n(A_{x_n} | x_0, \alpha_0, \dots, x_{n-1}, \alpha_{n-1}, x_n) = 1,$$

and let  $\Pi$  be the class of all such policies. For a given policy  $\Phi \in \Pi$  the process  $(\hat{x}_n, \hat{\alpha}_n)$  can then be defined iteratively by considering a current state  $\hat{x}_n$  and the process history  $\{(\hat{x}_k, \hat{\alpha}_k) : k = 0, \dots, n-1\}$ . Decision  $\hat{\alpha}_n$  is then given randomly by the distribution  $\tilde{\phi}_n(dy | \hat{x}_0, \hat{\alpha}_0, \dots, \hat{x}_{n-1}, \hat{\alpha}_{n-1}, \hat{x}_n)$ , and then state  $\hat{x}_{n+1}$  is given randomly by the distribution  $Q(dx | \hat{x}_n, \hat{\alpha}_n)$ . We are given an initial state  $\hat{x}_0$ . Then a cost of  $\sum_n g(\hat{x}_n, \hat{\alpha}_n)$  is assumed. (We do not consider at this point geometric discounting.) Define

$$J_\Phi(x) = E \left[ \sum_{n=0}^{\infty} g(\hat{x}_n, \hat{\alpha}_n) \mid \hat{x}_0 = x \right],$$

which denotes the expected cost assumed by the process under policy  $\Phi$  with initial state  $\hat{x}_0 = x$ . If  $\tilde{\phi}_n$  is parametrized by  $x_n$  only, then  $\Phi$  is a Markov policy. Let  $\Pi_1$  be the class of all mappings  $\phi : E \rightarrow A$  with  $\phi(x) \in A_x$  for all  $x \in E$ . We will be interested primarily in *nonrandomized stationary Markov policies*, that is, policies for which there is some  $\phi \in \Pi_1$  such that for all  $n \geq 0$ ,  $\tilde{\phi}_n(dy | x_n)$  is a point mass at  $\phi(x_n)$ . (In this case we will simply write  $\tilde{\phi}_n(dy | x_n) = \phi(x_n) \cdot$ )

We then define the problem:

(P) minimize  $J_\Phi(x)$  over all policies  $\Phi \in \Pi$  for each  $x \in E$ .

Define the value function

$$J^*(x) = \inf_{\Phi \in \Pi} J_\Phi(x), \quad x \in E.$$

For universally measurable  $J : E \rightarrow \mathfrak{R}^+$  define the operator  $T$  mapping  $J$  to  $TJ : E \rightarrow \mathfrak{R}^+$  by

$$(3.1) \quad (TJ)(x) = \inf_{\alpha \in A_x} \left( g(x, \alpha) + \int_E J(x') Q(dx' | x, \alpha) \right)$$

for all  $x \in E$ . For  $\phi \in \Pi_1$ , define the operator  $T_\phi$  mapping universally measurable  $J : E \rightarrow \mathfrak{R}^+$  to  $T_\phi J : E \rightarrow \mathfrak{R}^+$  by

$$(T_\phi J)(x) = g(x, \phi(x)) + \int_E J(x') Q(dx' | x, \phi(x))$$

for all  $x \in E$ . Letting  $J_0 \equiv 0$ , define the sequence

$$(3.2) \quad J_{k+1} = TJ_k, \quad k \geq 0.$$

(The sequence is well defined, since if  $J$  is lower semianalytic, so is  $TJ$ . See [1, Section 8.2].)



It is easy to verify that  $g \geq 0$  implies that  $T$  is monotone in the sense that  $TJ_2 \geq TJ_1$  if  $J_2 \geq J_1$ . Then  $J_1 \geq J_0$ , and hence  $J_2 = TJ_1 \geq TJ_0 = J_1$ . By extending this argument we conclude that  $\{J_k\}$  is increasing, so that the limit

$$(3.3) \quad J_\infty = \lim_{k \rightarrow \infty} J_k$$

exists.

The model defined in this section is a *lower semicontinuous model with positive cost* according to the definition given in [1, Definition 8.7, p. 208] if the following conditions hold:

- (C.1)  $A$  is compact.
- (C.2)  $\Gamma$  is a closed subset of  $E \times A$ .
- (C.3)  $g(x, \alpha)$  is lower semicontinuous on  $\Gamma$ .
- (C.4) The transition measure  $Q(dx | x, \alpha)$  is weakly continuous on  $\Gamma$ .

We summarize some results from [1, Proposition 8.6, Corollary 9.4.1, Proposition 9.8, Corollary 9.17.2, Proposition 9.18, pp. 209, 221, 225, 235, 236] in the following theorem.

**THEOREM 3.1.** *Under assumptions (C.1)–(C.4), the following hold.*

- (i) *If  $J \in \mathcal{J}$  is lower semicontinuous, then so is  $TJ$  (from proof of Proposition 8.6).*
- (ii)  *$J^*$  is lower semianalytic, and  $J^* = TJ^*$  (Corollary 9.4.1, Proposition 9.8).*
- (iii) *There exists a Borel measurable nonrandomized stationary Markov policy  $\Phi^*$  such that  $J_{\Phi^*} = J^*$  (Corollary 9.17.2).*
- (iv)  *$J^* = J_\infty$ , where  $J_\infty$  is lower semicontinuous (Corollary 9.17.2).*
- (v) *There exists a sequence  $\{\phi_k \in \Pi_1 : k \geq 0\}$ , where  $\phi_k$  is universally measurable, such that  $T_{\phi_k} J_k = TJ_k$ ,  $k \geq 0$ . Each sequence  $\{\phi_k(x)\}$ ,  $x \in E$ , has an accumulation point. If  $\phi^* \in \Pi_1$  is universally measurable and  $\phi^*(x)$  is an accumulation point of  $\{\phi_k(x)\}$  when  $J^*(x) < \infty$ , then  $\Phi^* = (\phi^*, \phi^*, \dots)$  is an optimal policy (Proposition 9.18).*

Suppose the state space  $E$  contains a measurable set  $E_K$  such that once the process enters  $E_K$  it does not leave and it assumes no further cost. Let  $\mathcal{J}_K$  be the set of all  $J : E \rightarrow \mathfrak{R}^+$  with  $J(x) = 0$  for all  $x \in E_K$ . We must then have  $J_\Phi \in \mathcal{J}_K$  for any policy  $\Phi$ . Furthermore, suppose there is some  $r > 0$  such that from any state the probability of subsequently entering  $E_K$  is at least  $r$  for all  $(x, \alpha) \in \Gamma$ . Under these assumptions, it is shown below that  $T$  is a contraction mapping on  $\mathcal{J}_K$ ; hence there is at most one fixed point of  $T$  in  $\mathcal{J}_K$ . This can be summarized by the following assumptions:

- (D.1)  $\exists r > 0$  such that  $Q(E_K | x, \alpha) \geq r$  for all  $(x, \alpha) \in \Gamma$ .
- (D.2)  $Q(E_K | x, \alpha) = 1$  for all  $x \in E_K, \alpha \in A$ .
- (D.3)  $g(x, \alpha) = 0$  for all  $x \in E_K, \alpha \in A$ .

**THEOREM 3.2.** *Suppose (D.1)–(D.3) hold. Then  $T$  is a contraction mapping of universally measurable  $J \in \mathcal{J}_K$  to  $\mathcal{J}_K$  with contraction constant  $1 - r$ .*

*Proof.* By (D.2), (D.3), and the definition of  $T$ , if  $J(x) = 0$  on  $E_K$  for universally measurable  $J$ , then  $TJ \in \mathcal{J}_K$ .

If  $J_1, J_2 \in \mathcal{J}_K$  are universally measurable, then for any  $\phi \in \Pi_1$  we have

$$\begin{aligned} \|T_\phi J_2 - T_\phi J_1\| &= \sup_{x \in E} \left| \int_E J_2(x') Q(dx' | x, \phi(x)) - \int_E J_1(x') Q(dx' | x, \phi(x)) \right| \\ &\leq \sup_{x \in E} \int_E |J_2(x') - J_1(x')| Q(dx' | x, \phi(x)) \end{aligned}$$

$$\begin{aligned}
&\leq \sup_{x \in E} \int_{E-E_K} |J_2(x') - J_1(x')| Q(dx' | x, \phi(x)) \\
&\leq \sup_{x \in E} \|J_2 - J_1\| (1 - Q(E_K | x, \phi(x))) \\
&\leq \|J_2 - J_1\| (1 - r)
\end{aligned}$$

since  $|J_2(x) - J_1(x)| = 0$  when  $x \in E_K$ . For  $\epsilon > 0$  we may select  $\phi$  so that

$$T_\phi J_1 \leq T J_1 + \epsilon,$$

and then

$$\begin{aligned}
T J_2 - T J_1 &\leq T J_2 - T_\phi J_1 + \epsilon \\
&\leq T_\phi J_2 - T_\phi J_1 + \epsilon \\
&\leq \|J_2 - J_1\| (1 - r) + \epsilon.
\end{aligned}$$

This holds for all  $\epsilon > 0$ , so  $T J_2 - T J_1 \leq \|J_2 - J_1\| (1 - r)$ . A similar argument gives  $T J_1 - T J_2 \leq \|J_2 - J_1\| (1 - r)$ , completing the proof.  $\square$

The model discussed in this section is directly applicable to the imbedded discrete time decision process introduced in section 2. Using the notation of that section, if  $E$  and  $E_o$  are measurable subsets of  $\mathfrak{R}^p$  and if  $A$  can be defined as a compact metric space, then  $E$  and  $A$  are both Borel spaces; then it remains to verify that  $\Gamma$  is closed. It must then be verified that the transition measure (2.2) is continuous on  $\Gamma$ , possibly through Theorem 2.1. Then the cost  $g$  must be lower semicontinuous on  $\Gamma$ . Under these conditions, assumptions (C.1)–(C.4) hold and Theorem 3.1 applies, and the optimum expected cost may be calculated through the dynamic programming algorithm (3.2)–(3.3). An optimal policy may be obtained as the limit defined in Theorem 3.1(v).

With respect to the process of section 2, assumption (D.1) will hold under various circumstances. Geometric discounting may be introduced by adding to  $E$  a kill state  $\Delta$  and assuming that the process jumps to  $\Delta$  at some fixed rate. If  $\lambda$  is bounded and  $B(x, \alpha)$  is bounded away from 0, then assumption (D.1) will be satisfied. Alternatively, there may be some target set which the state-action space is constrained to reach in one stage within some bounded time, barring a jump. If the process remains in this set with no further costs, then assumption (D.1) will be satisfied.

**4. The capacity expansion problem.** We now consider the optimal capacity expansion problem considered by Davis et al. [7]. We suppose that for a certain commodity there is a demand rate  $d$  which increases in time according to a compound Poisson process with constant rate  $\lambda > 0$ . Suppose there are enough plants to supply the commodity at rate  $s$ . At any time a decision to build a new plant may be made, which requires a total cost of  $C$ . Let  $y$  be the amount already invested in the plant being currently built. If no plant is currently being built, then  $y = 0$ . The rate of investment will then be  $\dot{y} \in [0, c]$ , where  $c$  represents the maximum possible investment rate. Once a plant is completed, capacity  $s$  is increased by  $L$  units. We let  $z = s - d$ . If  $z > 0$ , then there is overcapacity, and if  $z < 0$ , there is undercapacity. Let  $h : \mathfrak{R} \rightarrow \mathfrak{R}^+$  represent the rate at which cost is assumed due to overcapacity or undercapacity  $z$  with  $h(0) = 0$ .

The problem is to derive a policy, giving the investment rate at any state, which minimizes the total expected cost under geometric discounting. In [7] a technique for solving the BHJ equation for this problem is given. It should be noted that an

optimal solution does not necessarily exist. Examples are given in [7] of a problem in which for certain values of  $z$  it is  $\epsilon$ -optimal to build the current plant to within a small amount  $\beta$  of completion, with the expected cost function improving as  $\beta$  approaches 0, but not optimal to complete it. This suggests introducing as a control constraint the requirement that a plant be completed if it is within some fixed amount of completion. It will be shown below that this constraint forces assumptions (B.4) and (B.5) to hold. It is also shown in [7] that any optimal policy will specify either maximum investment rate  $c$  or minimum investment rate 0.

We need to specify  $E_\delta, E_o, A, \Gamma, \lambda, q_o, q_\delta, g$  as defined in section 2. The state space will be

$$\begin{aligned} E &= [0, C] \times \mathfrak{R}, \\ E_o &= [0, C] \times \mathfrak{R}, \\ E_\delta &= E - E_o. \end{aligned}$$

Then we interpret  $(y, z) \in E$  as the state at which the current plant has  $y$  currently invested and  $z = s - d$ . As in [7], we will suppose that the investment rate is either 0 or  $c$ . Hence from a starting point  $(y, z)$  the decision will consist of determining how much to invest in the current plant at rate  $c$ . The action space  $A$  is then the family of parametric curves  $\alpha : [0, \infty) \rightarrow \mathfrak{R}^2$  of the form

$$(4.1) \quad \alpha(t) = \begin{cases} (ct, 0), & 0 \leq t < a/c, \\ (a, 0), & a/c \leq t, \end{cases}$$

for  $a \in [0, C]$ . Thus we have  $T = \infty$ . The trajectories in (4.1) are homeomorphic to the interval  $[0, C]$ , so that  $A$  will subsequently be represented by  $[0, C]$ . We let  $\beta$  be a positive constant less than  $C$ . The action space will be constrained so that if  $y \in (C - \beta, C]$ , the project must be completed at rate  $c$ , and if  $y \in [0, C - \beta]$ , then an amount  $a \leq C - \beta - y$  or  $a = C - y$  may be invested. Hence we have state-action space

$$\begin{aligned} \Gamma &= \Gamma_1 \cup \Gamma_2, \\ \Gamma_1 &= \{(y, z, a) \in E \times [0, C] : y + a = C\}, \\ \Gamma_2 &= \{(y, z, a) \in E \times [0, C] : y + a \leq C - \beta\}. \end{aligned}$$

We introduce geometric discounting by adding to  $E$  a kill state  $\Delta$  to which the process jumps at a rate  $\eta > 0$ . At this state no further costs are assumed. We can then define the overall jump intensity as  $\lambda(y, z) \equiv \lambda + \eta$ . If the magnitude of any demand jump equals in distribution some nonnegative random variable  $Z$ , let  $P_Z(\cdot | y, z)$  be the probability measure of the random vector equal in distribution to  $(y, z - Z) \in E$ . Then

$$q_o(K | y, z) = \frac{\lambda}{\lambda + \eta} P_Z(K | y, z) + \frac{\eta}{\lambda + \eta} I\{\Delta \in K\}, \quad (y, z) \in E_o,$$

and

$$q_\delta(K | y, z) = I\{(0, z + L) \in K\}, \quad (y, z) \in E_\delta,$$

for any  $K \in \mathcal{E}$ , the Borel subsets of  $E$ . Then  $q_o$  and  $q_\delta$  are continuous on  $E_o$  and  $E_\delta$ , respectively. We also have  $B(y, z, a) < \infty$  in  $\Gamma_1$  and  $B(y, z, a) = \infty$  in  $\Gamma_2$ , which

are both closed sets. Assumptions (B.1)–(B.4) of Theorem 2.1 are then satisfied. Furthermore,  $B(y, z, a) = a/c$  on  $\Gamma_1$  so that assumption (B.5) of Theorem 2.1 is satisfied. Since assumption (A.1) also holds, we may conclude that the transition measure  $Q$  defined in (2.2) is continuous on  $\Gamma$ .

For state  $(y, z)$  and decision  $a$  we can then calculate the immediate stage cost  $g$ ,

$$g(y, z, a) = \frac{h(z) + c}{\lambda + \eta} (1 - \exp(-(\lambda + \eta)(a/c)))$$

for  $(y, z, a) \in \Gamma_1$  and

$$g(y, z, a) = \frac{h(z)}{\lambda + \eta} + \frac{c}{\lambda + \eta} (1 - \exp(-(\lambda + \eta)(a/c)))$$

for  $(y, z, a) \in \Gamma_2$ . If  $h(z)$  is lower semicontinuous, then so is  $g$  on  $\Gamma$ . Then assumptions (C.1)–(C.4) are satisfied so that Theorem 3.1 applies and algorithm (3.2)–(3.3) becomes

$$(4.2) \quad J_0(y, z) \equiv 0,$$

$$(4.3) \quad J_{k+1}(y, z) = \inf_a \left( g(y, z, a) + \int_E J_k(y', z') Q(dy', dz' | y, z, a) \right)$$

for all  $(y, z) \in E$ ,  $k \geq 0$ , where the infimum is taken over  $a \in [0, C - \beta - y] \cup \{C - y\}$  if  $y \leq C - \beta$ , and over the singleton  $\{C - y\}$  if  $y > C - \beta$ . Since at any state  $(C, z)$  the process transfers immediately to  $(0, z + L)$ , we may set  $J_k(C, z) = J_k(0, z + L)$  for all  $k \geq 1$ ,  $z \in \mathfrak{R}$ . Note that  $J_k(\Delta) = 0$ .

As an example, we will apply this algorithm to a case considered in [7], in which jumps in demand consist of one unit with probability one. The integral in (4.3) becomes

$$\int_E J_k(y', z') Q(dy', dz' | y, z, a) = \int_0^{a/c} J_k(y + ct, z - 1) \lambda \exp(-(\lambda + \eta)t) dt + J_k(0, z + L) \exp(-(\lambda + \eta)(a/c))$$

for  $(y, z, a) \in \Gamma_1$ , and

$$\int_E J_k(y', z') Q(dy', dz' | y, z, a) = \int_0^{a/c} J_k(y + ct, z - 1) \lambda \exp(-(\lambda + \eta)t) dt + J_k(y + a, z - 1) \frac{\lambda}{\lambda + \eta} \exp(-(\lambda + \eta)(a/c))$$

for  $(y, z, a) \in \Gamma_2$ . Then let

$$J_{k+1}^a(y, z) = g(y, z, a) + \int_E J_k(y', z') Q(dy', dz' | y, z, a)$$

for all  $(y, z, a) \in \Gamma$ ,  $k \geq 0$ . If  $L$  is an integer, then we may confine attention to a semigrd on  $E$  by constraining  $z$  to be an integer. We will discretize the problem by considering only states  $\{(Ci/n, z) : i = 0, 1, \dots, n\}$  for some large  $n$ . Choose  $\beta = C(i^*/n)$  for some positive integer  $i^* < n$ . Then (4.3) can be calculated for a

given  $J_k$  numerically. To reduce the number of calculations necessary we can evaluate the discretized version of (4.3) using backwards recursion by setting

$$\begin{aligned} J_{k+1}(C - \beta, z) &= \min\{J_{k+1}^0(C - \beta, z), J_{k+1}^\beta(C - \beta, z)\}, \\ J_{k+1}(C(1 - i/n), z) &= \min\{V_{wait}, V_{go}\}, \quad i = i^* + 1, \dots, n, \end{aligned}$$

where

$$\begin{aligned} (4.4) \quad V_{wait} &= \frac{h(z) + \lambda J_k(C(1 - i/n), z - 1)}{\lambda + \eta}, \\ V_{go} &= \left( C \frac{h(z) + c}{nc} + J_{k+1}(C(1 - (i - 1)/n), z) \right) \exp(-(\lambda + \eta)C/(nc)) \\ (4.5) \quad &+ \lambda/(\lambda + \eta) J_k(C(1 - i/n), z - 1) (1 - \exp(-(\lambda + \eta)C/(nc))). \end{aligned}$$

Intuitively, when the process is in state  $(C - \beta, z)$  there are only two options available: completing the project or waiting. So we calculate the expected cost for each option and set  $J_{k+1}(C - \beta, z)$  to be the smaller value. Then consider state  $(C - \beta - C/n, z)$ . Again, there are two choices: either proceeding to point  $(C - \beta, z)$  or waiting. Then  $V_{wait}$  in (4.4) with  $i = i^* + 1$  represents the expected cost of waiting. If the choice is to proceed, the assumption is that the process reaches state  $(C - \beta, z)$  with probability  $\exp(-(\lambda + \eta)C/(nc))$  and then assumes the optimal choice there. Otherwise, the process jumps to point  $(C - \beta - C/n, z - 1)$  or  $\Delta$ , with probabilities  $\lambda/(\lambda + \eta)$  and  $\eta/(\lambda + \eta)$ , respectively. For this choice  $V_{go}$  in (4.5) with  $i = i^* + 1$  represents the expected cost. Then set  $J_{k+1}(C - \beta - C/n, z)$  to be the smaller of these two values. Continue in this manner, decreasing  $y$  by  $C/n$ , until  $J_{k+1}$  is calculated for state  $(0, z)$ , and then repeat this algorithm for all values of  $z$ . Then  $J_{k+1}$  is used to calculate  $J_{k+2}$  in the same manner, continuing in this way until convergence is achieved.

This algorithm was applied to a set of parameters  $L = 1$ ,  $C = 1$ ,  $c = 1$ ,  $\lambda = 0.8$ ,  $\eta = 0.05$ , and  $h(z) = 1.5|z|$  on the range  $10 \leq z \leq 10$  with  $n = 50$  and  $\beta = 3/50$ . Note that to calculate  $J_{k+1}(y, z)$  the values of  $J_k(y, z - 1)$  and  $J_k(0, z + 1)$  are required; hence the range over which  $J_k$  can be calculated will decrease by one unit in each direction of  $z$  with each iteration. In [7] this is dealt with by setting appropriate boundary conditions. We do the same here with the constraint  $J_k(y, -10) = 250$ ,  $k \geq 0$ . This quantity is roughly the expected cost when construction continues indefinitely from state  $(0, -10)$ . We also assumed that it will be optimal to wait at all states  $(y, 10)$ . These constraints allow the calculation of  $J_k$  on the entire range of interest.

It was found that the optimal policy could be expressed by the quantities  $w(z)$ ,  $z = -10, -9, \dots, 10$ , where it will be optimal to construct as long as  $y < 1 - w(z)$ . The quantities found were  $w(z) = 0$  for  $z = -10, \dots, -1$ ;  $w(z) = 1$  for  $z = 4, \dots, 10$ ; and  $w(0) = 0.06$ ,  $w(1) = 0.06$ ,  $w(2) = 0.22$ , and  $w(3) = 0.62$ . Note that  $0.06 = \beta$ . The same example calculated in [7] gives  $w(1) = 0.0158$ ,  $w(2) = 0.2225$ , and  $w(3) = 0.6612$ . Also in [7], for  $z = 0$  it was found that the expected cost improved as  $w(0) \rightarrow 0$ , but it was not optimal to set  $w(0) = 0$ . Accordingly, the algorithm proposed here calculated  $w(0) = \beta$ . Similarly, where  $w(1) = 0.0158 \leq \beta$  in [7],  $w(1)$  by the above algorithm was found to be  $\beta$ . The other values were the same using both methods. Convergence was achieved by 50 iterations.

It should be noted that the solution techniques used in [7] require some prior assumption about the form of the optimal policy. Two classes of policy are considered: the “invest until complete” (IUC) policy and the “follow realized demand” (FRD) policy. For an IUC policy there is nonincreasing  $w(z) \in [0, C]$  such that construction

takes place when  $y \geq 1-w(z)$ . Essentially, a plant is completed once started under this policy. For an FRD policy there is nondecreasing  $w(z) \in [0, C]$  such that construction takes place as long as  $y < 1-w(z)$ . (The optimal policy calculated in this section is an FRD policy.) Necessary and sufficient optimality conditions are developed separately for each class and are then investigated separately. No such distinction has to be made in the technique presented in this paper.

**5. Concluding remarks.** The problem of minimum cost piecewise deterministic processes under a broad class of controls was considered with the objective of verifying the existence of an optimal control and with proposing a unified approach to a numerical solution. The approach is fundamentally different from other discussions of this problem in the literature in that the control problem is presented as a discrete time decision process in which a decision consists of the selection of a trajectory segment from a compact space. The BHJ equation plays no role. If the action space is numerically tractable, a straightforward fixed point algorithm based on a dynamic programming operator can be used to calculate the optimal control.

In the BHJ equation method the velocity field is commonly assumed to be Lipschitz-continuous. This means that a solution to the BHJ equation could also be constructed from a sequence of trajectory segments taken from a suitably defined compact space, making the theory presented here applicable also to control models treated in the literature cited above (although one would need to establish some smoothness conditions on an optimal trajectory as a necessary condition). The solution methodology, however, is more natural for problems in which the trajectory segments are parametrizable in finite dimensions, although the infinite dimension control could be approximated with splines. It is important to note that the discrete time decision process also admits more coarse varieties of control. For example, we may define piecewise linear control policies, which would have the effect of allowing control to be exerted only at regular time intervals.

It is anticipated that further work in this area will result in an expansion of the definition of the action space to include some control over jump rate and cost function. This would make the range of applicable models similar to that of methods based on the BHJ equation. It would also be of some value to allow trajectory time lengths to vary and hence be subject to control. This would significantly expand the classes of admissible control structures. A more complete treatment of, for example, impulse-type controls would then be possible.

#### REFERENCES

- [1] D. BERTSEKAS AND S. E. SHREVE, *Stochastic Optimal Control: The Discrete Time Case*, Academic Press, New York, 1978.
- [2] P. BILLINGSLEY, *Probability and Measure*, 3rd ed., John Wiley and Sons, New York, 1995.
- [3] F. H. CLARKE, *Optimization and Nonsmooth Analysis*, John Wiley and Sons, New York, 1983.
- [4] M. H. A. DAVIS, *Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models*, J. R. Stat. Soc., Ser. B Stat. Methodol., 46 (1984), pp. 353–388.
- [5] M. H. A. DAVIS, *Control of piecewise-deterministic processes via discrete time dynamic programming*, in Proceedings of the 3rd Bad Honnef Conference on Stochastic Differential Systems, Lecture Notes in Control and Inform. Sci. 16, Springer-Verlag, Berlin, 1986, pp. 140–150.
- [6] M. H. A. DAVIS, *Markov Models and Optimization*, Monogr. Statist. Appl. Probab. 49, Chapman and Hall, London, 1993.
- [7] M. H. A. DAVIS, M. A. H. DEMPSTER, S. P. SETHI, AND D. VERMES, *Optimal capacity expansion under uncertainty*, Adv. Appl. Probab., 19 (1987), pp. 156–176.

- [8] M. H. A. DAVIS AND M. FARID, *Piecewise-deterministic processes and viscosity solutions*, in Stochastic Analysis, Control, Optimization and Applications. A Volume in Honor of W. H. Fleming, Systems Control Found. Appl., W. M. McEneaney, G. Yin, and Q. Zhang, eds., Birkhäuser Boston, Boston, 1999, pp. 249–268.
- [9] M. A. H. DEMPSTER AND J. J. YE, *Necessary and sufficient optimality conditions for control of piecewise deterministic Markov processes*, Stochastics Stochastics Rep., 40 (1992), pp. 125–145.
- [10] H. M. SONER, *Optimal control with a state space constraint II*, SIAM J. Control Optim., 24 (1986), pp. 1110–1122.
- [11] D. VERMES, *Optimal control of piecewise deterministic Markov processes*, Stochastics, 14 (1985), pp. 165–208.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.