DYNAMIC PROGRAMMING AND THE MAXIMUM PRINCIPLE FOR CONTROL OF PIECEWISE DETERMINISTIC MARKOV PROCESSES

by

J.J. YE

DMS-617-IR September 1992
Dynamic Programming and the Maximum Principle for Control of Piecewise Deterministic Markov Processes

J.J. Ye*

Abstract

This paper studies optimal control of piecewise deterministic Markov processes. A stochastic maximum principle expressed in terms of adjoints given by deterministic differential inclusions is formulated. The relationship between dynamic programming optimality conditions and the stochastic maximum principle is given. Conditions are given under which the maximum principle can be stated in the normal form, the value function appeared in the maximum principle is eliminated, and some asymptotic transversality conditions hold in the case where the boundary hitting time is infinite. Implications of the maximum principle to computations are also discussed.

Key words: Optimal control, piecewise deterministic Markov process, dynamic programming, maximum principle, value function, asymptotic transversality conditions.

1991 AMS (MOS) subject classifications: Primary 49K45, secondary 49K15.

*Department of Mathematics and Statistics, P.O. Box 3045, University of Victoria, Victoria B.C., Canada V8W 3P4.
1 Introduction

Piecewise deterministic Markov processes (abbreviated as PDPs), introduced by Davis [5] in 1984, are continuous time homogeneous Markov processes consisting of a mixture of deterministic motion and random jumps. PDPs, which admit stochastic jump processes and deterministic dynamical systems as special cases, include virtually all of the stochastic models of applied probability except diffusions. PDPs provide a framework for studying optimization problems arising in queueing systems, inventory theory, resource allocation, capacity expansion problems (cf. [8]) and many other areas of operations research.

The optimal control theory of PDPs by dynamic programming approach has recently been developed by Vermes [16], Davis [6], Soner [15], Dempster and Ye [10] and Ye [18]. A stochastic maximum principle was given by Dempster and Ye [11]. A summary of results obtained by Dempster and Ye was given in Dempster [9]. The reader is also referred to Davis [7] for the complete theory of PDPs.

The stochastic maximum principle given in [11], however, is not satisfactory in the way that no nontriviality condition and transversality condition were given in the case where the boundary hitting time is infinite, and the value function was involved in the maximum principle. The main purpose of this paper is to resolve these problems and in the mean while, give the relationship between the maximum principle and dynamic programming.

A plan of the paper is as follows:

In section 2, we formulate the PDP optimal control problem precisely. Section 3 contains the reduction of the PDP optimal control problem to a family of deterministic control problems with a boundary condition. This control problem is nonstandard in that the terminal time is either $+\infty$ or the first time the problem hits the boundary of the state space. It is also an optimal control problem with nonsmooth problem data even in the case where all data of the PDP control problem is smooth. In section 4, we first develop a maximum principle for the deterministic control problem with a boundary condition. A stochastic maximum principle for the PDP control problem then follows by using the reduction result in Section 3. In section 5, we give the relationship between dynamic programming and the maximum principle. A by-product of the relation is a maximum principle for the PDP optimal control problem stated in
a strong form such that the nonnegative scalar in the maximum principle obtained in
section 4 can be taken as 1, the value function evaluated along the optimal trajectory is
replaced by the second adjoint variable, and some asymptotic transversality conditions
hold. Finally we conclude the paper by giving some remarks on the usefulness of the
maximum principle to compute the optimal controls in section 6.

2 The PDP optimal control problem

First we give a precise definition for a PDP. We shall take the interior of the state space
$E$ of our process to be a subset in $\mathbb{R}^n$ with interior $E^0 = \{ x \in \mathbb{R}^n : \psi(x) > 0 \}$ defined
by a boundary function $\psi \in C^1(\mathbb{R}^n)$ for which $\| \nabla \psi \| \neq 0$ for $x \in \partial E := \{ x \in \mathbb{R}^n : \psi = 0 \}$, the boundary of $E$. $E^0$ may or may not be connected and/or have compact
closure. $\mathcal{E}$ denotes the Borel sets of $E^0$. $C(E^0)$ denotes the set of all continuous and
bounded functions on $E^0$.

A piecewise deterministic process (PDP) taking values in $E$ is determined by its
three local characteristics :-

(i) A Lipschitz continuous vector field $f : E \rightarrow \mathbb{R}^n$ which determines a flow (or
integral curve) $\phi(t, x)$ in $E$ such that for $t \geq 0$
$$\frac{\partial}{\partial t} \phi(t, x) = f(\phi(t, x)) \quad \phi(0, x) = x \quad \text{for all } x \in E^0.$$ (i)

(ii) A jump rate $\lambda : E^0 \rightarrow \mathbb{R}_+ := [0, \infty)$ which is a measurable function such that
for each $x \in E^0$ there is an $\varepsilon(x) > 0$ such that
$$\int_0^{\varepsilon(x)} \lambda(\phi(t, x)) dt < \infty. \quad (1)$$

(iii) A transition measure $Q : E^0 \cup \partial E \rightarrow \mathcal{P}(E^0)$ where $\mathcal{P}(E^0)$ denotes the set of
probability measures on $E^0$ with the property that for each fixed $A \in \mathcal{E}$, the map
$x \mapsto Q(A; x)$ is measurable and $Q(\{x\}; x) = 0$.

From these characteristics a right-continuous sample path $x_t$ of the process $\{x_t : t > 0\}$
starting at $x \in E^0$ may be constructed as follows. Define $x_t := \phi(t, x)$ for $0 \leq t <
T_1$ where $T_1$ is the realization of the first jump time $T_1$ with generalized negative
exponential distribution determined by
$$P_x[T_1 > t] = \begin{cases} 
\exp[-\int_0^t \lambda(\phi(s, x))ds] & t < t_*(x) \\
0 & t \geq t_*(x).
\end{cases}$$
where
\[ t_*(x) := \inf\{ t > 0 : \phi(t, x) \in \partial E \} \] (2)
denotes the **boundary hitting time** (with the convention that \( \inf \emptyset = \infty \)).

Having realized \( T_1 = T_1 \) (possibly at \( T_1 = t_*(x) \)), we have \( x_{T_1} := \phi(T_1, x) \) and the **post jump state** \( x_{T_1} \) has distribution given by
\[ P_x [x_{T_1} \in A | T_1 = T_1] = Q(A; \phi(T_1, x)) \quad \forall A \in \mathcal{E}. \]

We may now restart the process at \( x_{T_1} = x_{T_1} \) according to the same recipe and proceeding recursively we obtain a sequence of jump time realizations \( T_1, T_2, \ldots \) between which \( x_t \) follows the integral curves of \( f \). Considering this construction as generic yields the process \( \{ x_t : t \geq 0, x_0 = x \} \) and the sequence of its jump times \( T_1, T_2, \ldots \).

Our jump rate assumption (1) implies that \( P_x [T_{k+1} > T_k] = 1 \) and we now further assume that \( P_x [T_n \uparrow \infty] = 1 \) for all \( x \in E^0 \).

As defined here the PDP \( \{ x_t \} \) is a temporally homogeneous strong Markov process with right continuous, left limited sample paths (**cf.** Davis [5] or [7]).

Now suppose that the local characteristics \( f, \lambda, Q \) of \( \{ x_t \} \) depend on a **control action** \( v \) from a compact set \( U \). The set of admissible controls may be different for interior and boundary states. We assume that \( v \in U_0 \subset \mathbb{R}^m \) if \( x \in E^0 \) and \( v \in U_\partial \subset \mathbb{R}^l \) if \( x \in \partial E \). Therefore, we shall distinguish transition measure \( Q_0(dy; x, v) \), for \( x \in E^0, v \in U_0 \), describing jumps from interior points, from \( Q_\partial(dy; x, v) \), for \( x \in \partial E, v \in U_\partial \), describing jumps from boundary points.

We assume that the following assumptions hold for the PDP optimal control problem defined in the end of this section:

**A1** The control sets \( U_0, U_\partial \) are compact.

**A2** The vector field \( f : E \times U_0 \rightarrow \mathbb{R}^n \) is bounded, continuous and Lipschitz continuous in \( x \in E^0 \) uniformly in \( v \in U_0 \).

**A3** The jump rate \( \lambda : E^0 \times U_0 \rightarrow \mathbb{R}_+ \) is bounded, continuous and Lipschitz continuous in \( x \in E^0 \) uniformly in \( v \in U_0 \).

**A4** As mentioned above, the transition measure \( Q \) may be expressed in terms of \( Q_0 := Q_{|E^0} : E^0 \times U_0 \rightarrow \mathcal{P}(E^0) \) and \( Q_\partial := Q_{|\partial E} : \partial E \times U_\partial \rightarrow \mathcal{P}(E^0) \).

\( Q_0 : E^0 \times U_0 \rightarrow \mathcal{P}(E^0) \) is bounded, continuous relative to the weak* topology.
on $P(E^0)$ and Lipschitz continuous in $x \in E^0$ (i.e. for all $\theta \in C(E^0)$ the map $x \mapsto \int_{E^0} \theta(y)Q_0(dy; x, v)$ is continuous and Lipschitz) uniformly in $v \in U_\theta$. $Q_\theta : \partial E \times U_\theta \rightarrow P(E^0)$ is bounded, continuous and Lipschitz continuous in $x \in E$ uniformly in $v \in U_\theta$.

(A5) The set of admissible controls $u := (u_0, u_\theta) \in C \subset C_0 \times C_\theta$ is defined in terms of the set of interjump open loop control functions

$$C_0 := \{u_0 \in B : u_0(\tau, z) : \mathbb{R}_+ \times E^0 \rightarrow U_0\},$$

where $\tau$ represents the time elapsed since the last jump and $z$ represents the post jump state, and the set of feedback boundary controls

$$C_\theta : \{u_\theta \in B : u_\theta : \partial E \rightarrow U_\theta\}$$

for which $P^u_\phi[\lim_n T_n = \infty] = 1$ for all $x \in E$, where for initial state $x$, $P^u_x(\cdot)$ is the probability measure (on path space) induced by $u$ and $\mathcal{L}$ denotes the set of all Borel measurable functions between a given domain and range.

(A6) The running cost $l_0 : E^0 \times U_\theta \rightarrow \mathbb{R}_+$ is bounded, continuous and Lipschitz continuous in $x \in E^0$ uniformly in $v \in U_\theta$. The boundary (jump) cost $l_\theta : \partial E \times U_\theta \rightarrow \mathbb{R}_+$ is bounded, continuous and Lipschitz continuous in $x \in E$ uniformly in $v \in U_\theta$.

The use of controls which are only measurable necessitates the open loop nature of the interjump control function $u_0(\cdot, z)$ (or $u_{0_\theta}(\cdot)$ in an obvious short notation) which is appropriate to the initial condition $z$ and which need only be specified from 0 to the controlled boundary hitting time $t_\ast(z) \leq \infty$ (cf. (2)) in case no random jump occurs up to this time elapsed from the last jump time. If a jump occurs before $t_\ast(z)$ elapses, to $w \in E^0$ say, the control function $u_0(\cdot, w)$ is used next and the corresponding flow $\phi^u(\cdot, w)$ (or $\phi^u_w(\cdot)$) is the unique absolutely continuous solution of the inhomogeneous dynamical system determined by the controlled vector field $f(\cdot, u_0(\cdot, w)) : E \times [0, t_\ast(w)) \rightarrow E$ as

$$\frac{\partial}{\partial \tau} \phi^u(\tau, w) = f(\phi^u(\tau, w), u_0(\tau, w)) \quad \phi^u(0, w) = w,$$

following from Caratheodory's existence and uniqueness theorem for first order differential equations.
The assumptions that $Q_\theta$ and $l_\theta$ have the Lipschitz continuous extension to $E \times U_\theta$ is only for the Lipschitz continuity of the value function and can be omitted if we assume that $E$ is star shape (cf. Dempster and Ye [10]).

The PDP optimal control problem is to find an admissible control $u = (u_0, u_\theta) \in \mathcal{C}$ so as to minimize the expected discounted total cost functional
\[
J^u_x(\cdot) := E^u_x[\int_0^\infty e^{-\delta t} l_0(x_t, u_0(\tau_t, z_t)) dt + \sum_{i} e^{-\delta T_i} I_\theta(x_{T_i}, u_\theta(x_{T_i})) 1_{\{x_{T_i} \in \delta E\}}],
\]
where $E^u_x$ denoted expectation with respect to $P^u_x$, $\delta > 0$ is the discount rate and $1_{\{\cdot\}}$ denote the indicator function of the event $\{\cdot\}$.

3 Reduction to deterministic control problems

We first formulate the deterministic optimal control problem with a boundary condition:

\begin{align*}
(P_z) \quad \text{minimize} \quad & J(z, u(\cdot)) := \int_0^{t^u_*(z)} e^{-\Lambda^u_*(z)} f_0(x(t), u(t)) dt + e^{-\Lambda^u_*(z)} F(x(t^u_*(z))) \\
\text{over the class} \; & \Omega \; \text{of all admissible pairs} \; (x(\cdot), u(\cdot)) \\
\text{such that} \; & u : [0, t^u_*(z)] \rightarrow \mathbb{R}^m \; \text{is measurable}, \notag \\
& u(t) \in U_0 \subset \mathbb{R}^m \; \forall t \in [0, t^u_*(z)], \notag \\
& \dot{x}(t) = f(x(t), u(t)) \; \text{a.e. } t \in [0, t^u_*(z)], \notag \\
& x(0) := z \in E^0,
\end{align*}

where $\Lambda^u_*(z) := \int_0^{t^u_*(z)} \lambda(x(s), u(s)) ds$ and $t^u_*(z)$ is called the boundary hitting time of the trajectory $x(t)$ corresponding to control $u$ for initial state $z$ defined by

\[
t^u_*(z) := \inf\{t > 0 : x(t) \in \partial E\}.
\]

In the case where the trajectory for initial state $z$ never reaches the boundary of $E$, $t^u_*(z) = \inf \emptyset = \infty$ by convention and $[0, t^u_*(z)]$ should be viewed as $[0, \infty)$. Where there is no confusion, we will simply use $t_*(z)$ even $t_*$ instead of $t^u_*(z)$.

To make sure that $J$ is well defined, we assume that $\lambda := \inf_{x \in E^0, u \in U_0} \lambda(x, u) > 0$. Thus even if $t^u_*(z)$ is $\infty$, the integral converges and in this case the term $e^{-\Lambda^u_*(z)} F(x(t^u_*(z)))$ vanishes (by virtue of the boundedness of the cost functions assumed below).

We also assume the following conditions hold for $(P_z)$:
(H1) the control set \( U_0 \) is compact in \( \mathbb{R}^m \),

(H2) \( f : E \times U_0 \to \mathbb{R}^n \) satisfies assumptions (A2),

(H3) \( f_0, \lambda : E^0 \times U_0 \to \mathbb{R}_+ \) are bounded, continuous and Lipschitz continuous in \( x \in E^0 \) uniformly in \( u \in U_0 \),

Now we discuss how one can relate the PDP optimal control problem with this deterministic optimal control problem. The idea of reduction comes from the observation that randomness of a PDP only appears at countable jump times. Consider the discrete time Markov process formed by the postjump state process \( \{ z_k := x_{T_k}, k = 0, 1, \cdots \} \). The PDP optimal control problem defined in section 1 can be reformulated as an infinite horizon discrete time stochastic control problem. Using the framework of dynamic programming in Bertsekas and Shreve [1], Dempster and Ye [10] showed that the PDP optimal control problem can be reduced to a family of deterministic control problem indexed by the postjump states. To use the framework of Bertsekas and Shreve, however, if the support of the transition measure is not countable, one needs to consider the PDP control problem with a set of admissible relaxed controls: \( u = (u_0, u_\theta) \in \mathcal{R} = \mathcal{R}_0 \times \mathcal{C}_\theta \) which is defined in terms of the set of relaxed open loop control functions

\[
\mathcal{R}_0 := \{ u_0 \in \mathcal{B} : u_\theta(\tau, z) : \mathbb{R}_+ \times E^0 \to IP(U_0) \},
\]

and the set of the feedback boundary controls \( \mathcal{C}_\theta \).

According to Dempster and Ye [10], we also need the following technical assumption. The assumption will also be needed in proving the maximum principle for \( (P_z) \).

(A7) There exists \( \alpha > 0 \) such that for all \( x \in \partial E \) and all \( v \in U_0 \)

\[
f(x, v) \cdot n(x) \geq \alpha > 0,
\]

where \( n(x) := -\nabla \psi(x)/\|\nabla \psi(x)\| \) is the unit outward normal to \( \partial E \in \mathbb{R}^n \) at the point \( x \in \partial E \) and \( \cdot \) denotes inner product.

Accordingly, we consider the problem \( (P_z) \) with the set of relaxed controls

\[
\tilde{C} = \{ u : [0, \infty) \to IP(U_0) \},
\]

and define the value function \( V : E \to \mathbb{R}_+ \) for problem \( (P_z) \) by

\[
V(z) := \inf_{(x(\cdot), u(\cdot)) \in \tilde{C}} J(z, u(\cdot)) \quad \forall z \in E^0
\]

\[
V(z) := F(z) \quad \forall z \in \partial E,
\]
where \( \tilde{\Omega} \) is the set of admissible pairs associated with the relaxed controls.

The following result obtained by Dempster and Ye [10] reduces the PDP optimal control problem to a family of deterministic control problems \((P_\theta)\).

**Proposition 1** Assume that assumptions (A1)-(A7) hold. Define the value function \( V : E \to \mathbb{R}_+ \) of the PDP optimal control problem by

\[
V(x) := \inf_{u \in \mathbb{R}} J_x(u) \quad \text{for all } x \in E. \tag{6}
\]

Then it coincides with the value function for problem \((P_\theta)\) with problem data defined by

\[
f_0(x, v) := l_0(x, v) + \int_{E^0} V(y) Q_0(dy;x,v) \lambda(x,v) \tag{7}
\]

\[
F(x) := \min_{v \in U_\theta} \{l_\theta(x,v) + \int_{E^0} V(y) Q_\theta(dy;x,v)\} \tag{8}
\]

\[
\tilde{\lambda}(z,v) := \lambda(z,v) + \delta. \tag{9}
\]

The PDP optimal control \( u = (u_0, u_\theta) \) is equivalent to choosing for each possible postjump state \( z \in E^0 \), an interjump control function \( u_{\theta z}(\cdot) \) which solves the deterministic control problem \((P_\theta)\) with problem data defined by (7), (8) and (9) and for each \( z \in \partial E \), a boundary control action \( u_\theta(z) \) which solves the following optimization problem:

\[
\min_{v \in U_\theta} \{l_\theta(z,v) + \int_{E^0} V(y) Q_\theta(dy;z,v)\}. \tag{10}
\]

Proposition 1 (consequently Theorem 2 and 3) is also true if we replace \( V(x) := \inf_{u \in \mathbb{R}} J_x(u) \) by \( V(x) = \inf_{u \in \mathbb{R}} J_x(u) \) in the case where the support of the transition measure is countable (e.g. the process with Markov disturbance) because in this case part 1 of Bertsesas and Shreve [1] will do the job and the result of proposition 1 holds without using the relaxed control set. From now on, by the PDP optimal control problem we mean the PDP optimal control problem with relaxed control. The values of the PDP optimal control problem with ordinary control and relaxed controls coincide under the following convexity assumptions:

- The set

\[
N_\theta(x) := \{ (f(x,v), \lambda(x,v), l) : l \geq l_0(x,v) + \lambda(x,v) \int_{E^0} \theta(y) Q_0(dy;x,v)\}, v \in U_0 \}
\]

is convex for all \( x \in E^0 \) and \( \theta \in C(E^0) \).
In the case the minimum value of the optimal control problem is attained by a relaxed control, the maximum principle obtained in the paper remain true provided that one explain a function $\phi(x(t), u(t))$ properly in terms of relaxed controls $u(t)$, i.e.

$$
\phi(x(t), u(t)) = \int_{U_0} \phi(x(t), u)u(t)(du).
$$

Note that $F(x)$ can be a nonsmooth function even when $l_0(\cdot, v)$ and $Q_0(A; \cdot, v)$ are smooth since it is a pointwise minima of some indexed family of functions. In fact, we can only assure that $F(x)$ is Lipschitz continuous. Therefore, instead of the usual gradient we should use the Clarke generalized gradient—a generalized calculus for Lipschitz continuous functions. To be more precise, let $\phi : \mathbb{R}^k \rightarrow \mathbb{R}$ be locally Lipschitz near a given point $x$. Then the generalized gradient of $\phi$ at $x$, denoted $\partial \phi(x)$, is the subset of $\mathbb{R}^k$ given by

$$
\partial \phi(x) := \co \{ \lim_i \nabla \phi(x_i) : x_i \rightarrow x, \text{ where } \nabla \phi(x_i) \text{ exists} \}.
$$

$\partial \phi(x)$ is in general a set-valued map and it coincides with the usual gradient $\{\nabla \phi(x)\}$ if only if $\phi$ is continuously differentiable. For a vector-valued function, the generalizedJacobians is similarly defined. The reader is referred to Clarke [3] for more details.

4 The Maximum Principle

In view of Proposition 1, a maximum principle for the PDP control problem will follow once the appropriate one for deterministic problem ($P_2$) is established. This deterministic control problem is however non-standard in that the terminal time $t_*$ is not fixed, but is instead either $+\infty$ or the first time the trajectory reaches the boundary of the state space. It is also a control problem with nonsmooth problem data even in the case where all problem data of the original PDP optimal control are smooth.

Given $x \in E^0$, $v \in U_0$, $p \in \mathbb{R}^n$, $q \in \mathbb{R}$, $r \in \mathbb{R}$, define the Hamiltonian function for ($P_2$) as

$$
H(x, p, q, v, r) := p \cdot f(x, v) - q\lambda(x, v) - rf_0(x, v).
$$

The following theorem provides a maximum principle for ($P_2$).

**Theorem 1** Let $(x^*(\cdot), u^*(\cdot))$ be an optimal solution for problem ($P_2$) and $t_*$ the corresponding boundary hitting time. Suppose the $f$ satisfies (A7), then there exist a scalar
\( r \in \{0, 1\} \) and absolutely continuous functions

\[
p : [0, t_*] \rightarrow \mathbb{R}^n \quad q : [0, t_*] \rightarrow \mathbb{R},
\]

such that:-

1. The adjoint variables \((p, q)\) satisfy the adjoint equations in the form of the differential inclusions:

\[
-\dot{p}(t) \in \partial_x f(x^*(t), u^*(t))^{\top} p(t) - q(t)\partial_x \lambda(x^*(t), u^*(t)) - r \partial_x f_0(x^*(t), u^*(t)) - \lambda(x^*(t), u^*(t))p(t) \quad a.e. \ t \in [0, t_*]
\]

\[
-\dot{q}(t) = -rf_0(x^*(t), u^*(t)) - q(t)\lambda(x^*(t), u^*(t)) \quad a.e. \ t \in [0, t_*].
\]

2. The optimal control \(u^*(t)\) maximizes the Hamiltonian pointwise, viz.

\[
\max_{v \in \mathcal{U}_0} H(x^*(t), p(t), q(t), v, r)
\]

\[
= H(x^*(t), p(t), q(t), u^*(t), r)
\]

\[
= 0 \quad a.e. \ t \in [0, t_*].
\]

3. The system is subject to the transversality condition: if \(t_* < \infty\), then

\[
-(p(t_*), q(t_*)) \in [r \partial F(x^*(t_*)) + \rho \partial d_E(x^*(t_*))] \times \{r(F(x^*(t_*)))\}.
\]

4. The adjoint variables satisfy the nontriviality condition

\[
|\langle p(0), q(0), r \rangle| > 0,
\]

where \(\partial\) denotes either the Clarke generalized gradient or the generalized Jacobian, \(\partial_x\) denotes either the partial generalized gradient or the partial generalized Jacobian with respect to \(x\), and \(d_C(x)\) is the distance from a point \(x\) to a set \(C\).

**Proof.** Problem \((P_e)\) can be equivalently posed as follows:

\[
(P_e) \quad \min \int_0^{t_*} x_0(t)f_0(x(t), u(t))dt + x_0(t_*)F(x(t_*))
\]

on the class \(\tilde{\Omega}\) of all pairs \((\tilde{x}(\cdot), u(\cdot))\) with

\[
\tilde{x}(\cdot) := (x(\cdot), x_0(\cdot))
\]

s.t.

\[
\tilde{x}(t) = (f(x(t), u(t)), -x_0(t)\lambda(x(t), u(t))) \quad a.e. \ t \in [0, t_*]
\]

\[
\tilde{x}(0) := (z, 1)
\]

\[
t_* := \inf \{t > 0 : x(t) \in \partial E\}
\]
For an optimal pair \((x^*(\cdot), u^*(\cdot))\) in \(\Omega\) we denote by \((\bar{x}^*(\cdot), u^*(\cdot))\) the corresponding solution for \((\bar{P}_\varepsilon)\) in the class \(\bar{\Omega}\).

Now we divide the analysis into two cases:

(a) the boundary hitting time of the optimal trajectory \(x^*(\cdot)\) is finite,

(b) the boundary hitting time of the optimal trajectory \(x^*(\cdot)\) is infinite.

Assumption (A7) postulates that when the trajectories get sufficiently closed to the boundary of \(E\), they must hit the boundary of \(E\) in such a way that the corresponding vector field element makes an acute angle with the outward pointing unit normal. As a result, we can find a tube about the optimal trajectory \(x^*(t)\) such that any trajectory in the tube hits the boundary at most once. Therefore, \((\bar{x}^*(\cdot), u^*(\cdot))\) is the optimal solution of the following problem:

\[
(P_C) \quad \text{minimize} \quad \int_0^{t_*} x_0(t)f_0(x(t), u(t))dt + x_0(t_*)F(x(t_*))
\]

on the class \(\bar{\Omega}\) of all pairs \((\bar{x}(\cdot), u(\cdot))\) with

\[
\bar{x}(\cdot) = (x(\cdot), x_0(\cdot))
\]

s.t.

\[
\dot{x}(t) = (f(x(t), u(t)), -x_0(t)\lambda(x(t), u(t))) \quad \text{a.e.} \quad t \in [0, t_*]
\]

\[
x(t) \in T(x^*; \varepsilon)
\]

\[
\bar{x}(0) := (z, 1)
\]

\[
(t_*, \bar{x}(t_*)) \in M,
\]

where \(M := [0, \infty) \times \partial E \times [0, 1]\) in case (a), \(M := \{\infty\} \times \mathbb{E} \times [0, 1]\) in case (b), \(T(x^*; \varepsilon)\) is the \(\varepsilon\)-tube about optimal trajectory \(x^*\) defined by

\[
T(x^*; \varepsilon) := \{v \in \mathbb{R}^n : |x^*(t) - v| < \varepsilon, t \geq 0\},
\]

and \(\varepsilon > 0\) is sufficiently small to ensure that \(T(x^*; \varepsilon) \subset \mathbb{E}\) for \(t \in [0, t_*]\).

Case (a). \(t_* < \infty\)

In this case the time interval is finite and the endpoint constraint set \([0, \infty) \times \partial E \times [0, 1]\) is closed in \(\mathbb{R}^{n+2}\), the nonsmooth deterministic maximum principle developed by Clarke in [3] is applicable.

The Hamiltonian function for the problem \((P_C)\) is defined as follows:

\[
\bar{H}(\bar{x}, \bar{p}, q, v, r) := \bar{p} \cdot f(x, v) - q x_0 \lambda(x, v) - r x_0 f_0(x, v)
\]
for \( \bar{x} := (x, x_0) \in \mathbb{R}^{n+1}, v \in U_0, \bar{p} \in \mathbb{R}^n, q \in \mathbb{R} \) and \( r \in \mathbb{R} \).

Identifying the data for \((P_C)\) with the corresponding data in Clarke [2, Theorem 5.2.3], we obtain the following maximum principle for the problem \((P_C)\):

There exist a scalar \( r \in \{0, 1\} \) and absolutely continuous functions

\[
\bar{p} : [0, t_*] \rightarrow \mathbb{R}^n \quad q : [0, t_*] \rightarrow \mathbb{R}
\]

such that:-

1. The adjoint variables \((\bar{p}, q)\) satisfies the adjoint equations in the form of the differential inclusions

\[
- \frac{d}{dt}(\bar{p}(t), q(t)) \in \partial_2 \bar{H}(\bar{x}^*(t), \bar{p}(t), q(t), u^*(t), r) \quad \text{a.e. } t \in [0, t_*].
\]

2. The optimal control function \( u^*(t) \) maximizes the Hamiltonian pointwise, viz.

\[
\max_{u \in U_0} \bar{H}(\bar{x}^*(t), \bar{p}, q, u, r) = \bar{H}(\bar{x}^*(t), \bar{p}, q, u^*(t), r) = 0 \quad \text{a.e. } t \in [0, t_*].
\]

3. The system is subject to the transversality condition

\[
-(\bar{p}(t_*), q(t_*)) \in r \partial \bar{F}(\bar{x}(t_*)) + \rho \partial d_{\partial E \times [0,1]}(x(t_*), x_0(t_*))
\]

where the function \( \bar{F}(\bar{x}) := x_0 F(x) \) and \( \rho \) is some nonnegative scalar.

4. The adjoint variables satisfy the nontriviality condition

\[
\| (\bar{p}, q) \|_{\infty} + \bar{r} > 0,
\]

where \( \| \cdot \|_{\infty} \) is the supremum norm for the space of appropriate functions on \([0, t_*]\).

Now we need to rearrange the expressions so that we have a maximum principle for the problem \((P_z)\).

Define \( p(.) := \bar{p}(\cdot)/x_0^*(\cdot) \), which is well defined since \( x_0^*(\cdot) > 0 \). It follows that (17) implies (13).
Let \( H_1(\tilde{x}, \tilde{p}, v) := \tilde{p} \cdot f(x, v) \) and \( H_2(\tilde{x}, q, v, r) = -x_0(q\bar{\lambda}(x, v) + rf_0(x, v)) \). Since \( H_1 \) as a function of \( \tilde{x} \) is independent of \( x_0 \) and \( H_2 \) can be written in a form \( x_0G(x) \), where \( x_0 > 0 \) and \( G(x) \) is continuous. It can be shown as in Ye [15, Proposition 1.8] that

\[
\partial_x H_1 = \partial_x H_1 \times \{0\}
\]
\[
\partial_x H_2 = \partial_x H_2 \times \{-q\lambda(x, u) - rf_0(x, u)\}.
\]

Consequently, we have

\[
\partial_x \tilde{H} := \partial_x[H_1 + H_2] \subset \partial_x H_1 + \partial_x H_2
\]
\[
= (\partial_x H_1 + \partial_x H_2) \times \{-q\lambda(x, u) - rf_0(x, u)\},
\]

where the inclusion (20) follows from the finite sums formula (see Clarke [2, Proposition 2.3.3]).

Therefore (16) implies

\[
-\frac{d}{dt} \tilde{p}(t) \in \partial_x H_1(\tilde{x}^*(t), \tilde{p}(t), u^*(t)) + \partial_x H_2(\tilde{x}^*(t), q(t), u^*(t), r)
\]
\[
= \partial_x f(x^*(t), u^*(t))^T \tilde{p}(t) - x_0^*(t)[q(t)\partial_x \bar{\lambda}(x^*(t), u^*(t)) + r \partial_x f_0(x^*(t), u^*(t))]
\]
\[
a.e. \; t \in [0, t_*] \tag{21}
\]
\[
-\frac{d}{dt} q(t) = -q(t)\bar{\lambda}(x^*(t), u^*(t)) - rf_0(x^*(t), u^*(t))
\]
\[
a.e. \; t \in [0, t_*]. \tag{22}
\]

Since \( \tilde{p}(t) := x_0^*(t)p(t) \) by definition, the left hand side of inclusion (21) is equal to

\[
\frac{d}{dt}[x_0^*(t)p(t)] = x_0^*(t)\frac{d}{dt} p(t) - x_0^*(t)\bar{\lambda}(x^*(t), u^*(t))p(t).
\]

Therefore inclusion (21) and the definition of \( \tilde{p} \) imply (11). Equation (12) is obtained from equation (22).

Since \( x_0^*(t_*) > 0 \) and \( F \) is continuous, we have

\[
\partial F(x^*(t_*)) = x_0^*(t_*) \partial F(x^*(t_*)) \times \{F(x^*(t_*))\}. \tag{23}
\]

Since \( \partial d_{C_1 \times C_2}(x_1, x_2) = \partial d_{C_1}(x_1) \times \partial d_{C_2}(x_2) \) (see Clarke [2, corollary of Theorem 2.4.5]), we have

\[
\partial d_{\beta F}(x^*(t_*), x_0^*(t_*)) = \partial d_{\beta F}(x^*(t_*)) \times \partial d_{[0,1]}[x_0^*(t_*)]
\]
\[
= \partial d_{\beta F}(x^*(t_*)) \times \{0\}, \tag{24}
\]

13
where the last equality follows from the fact that $x_0^*(t_*) \in (0, 1)$.

By substituting equalities (23) and (24) into inclusion (18), we have
\[-(\tilde{p}(t_*), \tilde{q}(t_*)) \in r[x_0^*(t_*) \partial F(x^*(t_*))] \times \{F(x^*(t_*))\} - \rho \partial d_{\partial} E(x^*(t_*)) \times \{0\} \]
\[= [rx_0^*(t_*) \partial F(x^*(t_*))] - \rho \partial d_{\partial} E(x^*(t_*))] \times \{rF(x^*(t_*))\},\]
from which we obtain the transversality condition (14).

If $r = 0$, it follows from (11) and (12) that the adjoint variables $(p, q)$ satisfy
\|[\dot{p}(t), \dot{q}(t)]\| \leq K|\dot{p}(t), q(t)|\|
for some constant $K$ by virtue of the boundedness and the uniform Lipschitz continuity of the problem data (cf. Clarke [2, Proposition 2.1.2]). This observation, together with Gronwall's lemma, implies that $(p, q)$ is either zero or else nonvanishing in $[0, t_*]$. But by (19), $(p, q)$ can not be zero, so $(p, q)$ is nonvanishing on $[0, t_*]$ as $r = 0$. This is the nontriviality condition (15).

Case (b). $t_* = \infty$. The situation of this case is not covered directly by the results employed in the previous case due to the fact that $t_* = \infty$. Applying the nonsmooth maximum principle for infinite horizon problems (see Ye [16, Theorem 2.1]) and rearranging the expressions as we did in case (a), we obtain the result.

Now we are ready to state a stochastic maximum principle for the PDP optimal control problem.

The Hamiltonian function for the PDP control problem is defined as follows:
\[H(x, p, q, v, r, \theta(\cdot)) := p \cdot f(x, v) - q(\lambda(x, v) + \delta) - r[l_0(x, v) + \lambda(x, v) \int_{E_0} \theta(y)Q_0(dy; x, v)]\]
for $x \in \mathbb{R}^n$, $u \in U_0$, $p \in \mathbb{R}^n$, $q, r \in \mathbb{R}$ and $\theta(\cdot) \in C(E^0)$.

By Proposition 1 and Theorem 1 we readily obtain the maximum principle for the PDP optimal control problem as follows:

**Theorem 2** Under assumptions (A1)-(A7), let $u^* = (u_0^*, u_2^*)$ be an ordinary control which solves the PDP optimal control problem. For each possible postjump state $z \in E^0$, let $u_{0z}(t)$ be the optimal interjump control function and let $x_2^*(t) := \phi_u^*(t)$ be the corresponding deterministic trajectory in $E^0$ on $[0, t_{u*}^*(z)]$ with initial point $z$. Then
there exist a scalar \( r \in \{0,1\} \) and absolutely continuous functions

\[
p_z : [0, t_\ast] \rightarrow \mathbb{R}^n \quad q_z : [0, t_\ast] \rightarrow \mathbb{R},
\]

such that:

1. The adjoint variables \((p_z, q_z)\) satisfy the adjoint equations in the form of the differential inclusions

\[
-\dot{p}_z(t) \in (\partial_x f(x^*_z(t), u^*_0(t)) - [\lambda(x^*_z(t), u^*_0(t)) + \delta]J_n) p_z(t) \\
-\dot{q}_z(t) \in -r \partial_x \lambda(x^*_z(t), u^*_0(t)) - r\partial_x l_0(x^*_z(t), u^*_0(t)) \\
+ \lambda(x^*_z(t), u^*_0(t)) \int_{E^0} V(y)Q_0(dy; x^*_z(t), u^*_0(t))
\]

\( \text{a.e. } t \in [0, t_\ast] \)  \hspace{1cm} (25)

\[
-\dot{q}_z(t) = -q_z(t)[\lambda(x^*_z(t), u^*_0(t)) + \delta] - r[l_0(x^*_z(t), u^*_0(t)) \\
+ \lambda(x^*_z(t), u^*_0(t)) \int_{E^0} V(y)Q_0(dy; x^*_z(t), u^*_0(t))]
\]

\( \text{a.e. } t \in [0, t_\ast] \).  \hspace{1cm} (26)

2. The optimal interior control function \( u^*_0(\cdot) \) maximizes the Hamiltonian pointwise, viz.

\[
\max_{v \in U_0} H(x^*_z(t), p_z(t), q_z(t), v, r, V(\cdot)) = H(x^*_z(t), p_z(t), q_z(t), u^*_0(t), r, V(\cdot)) = 0 \quad \text{a.e. } t \in [0, t_\ast]
\]  \hspace{1cm} (27)

3. The system is subject to the transversality condition: if \( t_\ast < \infty \), then

\[
-(p_z(t_\ast), q_z(t_\ast)) \in \{r\partial F(x^*_z(t_\ast)) + \rho \partial d_\partial(x^*_z(t_\ast))\} \times \{rF(x^*_z(t_\ast))\}
\]

where

\[
F(x) := l_\partial(x, u^*_\partial(x)) + \int_{E^0} V(y)Q_\partial(dy; x, u^*_\partial(x)).
\]

4. The adjoint variables satisfy nontriviality condition

\[
|(p_z(0), q_z(0), r)| > 0.
\]
Remark 1 For every \( z \in E^0 \), there is a multiplier function \((p_z, q_z)\) which depends Borel measurably on \( z \). Hence corresponding to the optimally controlled PDP \( \{x_t\} \) we may consider the multiplier process \((p, q)\) as a random process.

Since it is a necessary condition, the value function is equal to the expected cost functional associated with the given optimal control \( u^* \).

5 The relationship between dynamic programming and the maximum principle

Note that in Theorem 2, the possibility of \( r = 0 \) is not excluded, there is no transversality condition for the case \( t_* = \infty \), and the value function is involved. In this section, we discuss the relationship between Dynamic Programming and the Maximum Principle using nonsmooth analysis. A by-product is a maximum principle for the PDP control problem stated in a strong form such that the nonnegative scalar in Theorem 2 can be taken as 1, the value function appears in Theorem 2 is eliminated, and asymptotic transversality conditions

\[
\lim_{t \to \infty} q_z(t) = 0, \quad \lim_{t \to \infty} p_z(t) e^{-\int_0^t (\lambda(x^*_z(s), u^*_z(s)) + \delta)ds} = 0
\]

hold in the case \( t_*(z) = \infty \).

If the value function \( V(z) \) happens to be continuously differentiable, the dynamic programming approach characterizes the value function as a solution to the following Hamilton-Jacobi-Bellman (HJB) equation

\[
\min_{v \in U_0} \{ \nabla V(z) f(z, v) - \delta V(z) + l_0(z, v) + \lambda(z, v) \int_{E^0} (V(y) - V(z)) Q_0(dy; z, v) \} = 0 \quad \forall z \in E^0
\]

(29)

with the boundary condition

\[
V(z) := \min_{v \in U_0} \{ l_0(z, v) + \int_{E^0} V(y) Q_0(dy; z, v) \} \quad \forall z \in \partial E.
\]

(30)

Also the optimal control \((u^*_0, u^*_0)\) satisfies for each \( z \in \partial E \)

\[
\nabla V(x^*_z(t)) f(x^*_z(t), u^*_0(t)) - \delta V(x^*_z(t)) + l_0(x^*_z(t), u^*_0(t)) \\
+ \lambda(x^*_z(t), u^*_0(t)) \int_{E^0} (V(y) - V(x^*_z(t))) Q_0(dy; x^*_z(t), u^*_0(t)) = 0
\]

(31)
and if \( t_*(z) < \infty \)

\[
V(x_z^*(t_*)) = l_\theta(x_z^*(t_*), u_\theta^*(z)) + \int_{E_0} V(y)Q_\theta(dy; x_z^*(t_*), u_\theta^*(z)). \tag{32}
\]

As in the classical treatments of optimal control theory (cf. Clarke and Vinter [4], Fleming and Rishel [12]), if the value function \( V(z) \) happens to be twice continuously differentiable and all problem data are differentiable respected to \( x \), then by setting

\[
p_z(t) := -\nabla V(x_z^*(t)), \quad q_z(t) := -V(x_z^*(t)), \tag{33}
\]

we will be able to derive a maximum principle from the equation of dynamic programming.

Indeed, (29) and (31) implies the following: for each \( t \in [0, t_*] \),

\[
\nabla V(x_z^*(t))f(x_z^*(t), u_{0z}^*(t)) - \delta V(x_z^*(t)) + l_0(x_z^*(t), u_{0z}^*(t)) + \\
\lambda(x_z^*(t), u_{0z}^*(t)) \int_{E_0} (V(y) - V(x_z^*(t)))Q_0(dy; x_z^*(t), u_{0z}^*(t)) \\
\leq \nabla V(x) f(x, u_{0z}^*(t)) - \delta V(x) + l_0(x, u_{0z}^*(t)) \\
+ \lambda(x, u_{0z}^*(t)) \int_{E_0} (V(y) - V(x))Q_0(dy; x, u_{0z}^*(t)) \tag{34}
\]

and for each \( t \in [0, t_*] \) and \( u \in U_0 \),

\[
\nabla V(x_z^*(t))f(x_z^*(t), u_{0z}^*(t)) - \delta V(x_z^*(t)) + l_0(x_z^*(t), u_{0z}^*(t)) + \\
\lambda(x_z^*(t), u_{0z}^*(t)) \int_{E_0} (V(y) - V(x_z^*(t)))Q_0(dy; x_z^*(t), u_{0z}^*(t)) \\
\leq \nabla V(x_z^*(t))f(x_z^*(t), v) - \delta V(x_z^*(t)) + l_0(x_z^*(t), v) \\
+ \lambda(x_z^*(t), v) \int_{E_0} (V(y) - V(x_z^*(t)))Q_0(dy; x_z^*(t), v). \tag{35}
\]

From (35) we deduce the maximization of Hamiltonian condition (27). Differentiating the right hand side of (34) respect to \( x \) evaluated along \( x_z^*(t) \) and setting it to zero, we obtain the first adjoint equation (25).

Notice that

\[
q_z(t) = -\frac{d}{dt} V(x_z^*(t)) = -\nabla V(x_z^*(t))f(x_z^*(t), u_{0z}^*(t)).
\]

Using the HJB equation (31), we derive the second adjoint equation (26).

By virtue of (32), we have that for \( t_* < \infty \),

\[
q_z(t_*) := -V(x_z^*(t_*)) = -F(x_z^*(t_*)), \quad p_z(t_*) := -\nabla V(x_z^*(t_*)) = -\nabla F(x_z^*(t_*)),
\]

17
which is the transversality condition (28).

The hypothesis on the smoothness of the value function is however highly unrealistic (cf. [13]). As in Clarke and Vinter [4] and Vinter [17], we naturally hope to replace (33) by the following:

\[ p_z(t) \in -\partial V(x_z^*(t)), \quad q_z(t) = -V(x_z^*(t)). \]

For this purpose, we set out the following assumptions which ensures the Lipschitz continuity of the value function of the PDP optimal control problem:

(A8) \( Q_\theta \) is defined on \( \partial E \times U_\theta \) and has an extension to \( E \times U_\theta \) such that the extension \( Q_\theta : \partial E \times U_\theta \to \mathbb{H}^4(L^0) \) is bounded, continuous and Lipschitz continuous in \( x \in E \) uniformly in \( u \in U_\theta \). Similarly, \( l_\theta \) is defined on \( \partial E \times U_\theta \) and has an extension to \( E \times U_\theta \) such that the extension \( l_\theta : \partial E \times U_\theta \to \mathbb{R}_+ \) is bounded, continuous and Lipschitz continuous in \( x \in E \) uniformly in \( u \in U_\theta \).

(A9) The jump rate satisfies

\[ \inf_{x \in E^0, v \in U_0} \lambda(x, v) + \delta > \lambda^0_+, \quad (36) \]

where \( \lambda^0 := \sup_{x, y \in E^0, v \in U_0} \frac{(x - y) \cdot (f(x, v) - f(y, v))}{|x - y|^2} \).

**Remark 2** Assumption (A8) can be omitted if we assume that \( E \) is star shaped (see Dempster and Ye [10]).

It is obvious that condition (A9) is implied by the following condition:

\[ \exists \ r > 0 \ s.t. \ \inf_{x \in E^0, u \in U_0} \lambda(x, u) + \delta \geq L_f + r, \]

where \( L_f \) is the Lipschitz constant of \( f(\cdot, u) \). This is a reasonable assumption for a real "stochastic" problem (e.g. the capacity expansion problem (cf. [8]) satisfies this assumption).

Assumption (A8) and (A9) can be replaced with any condition under which the value function is Lipschitz continuous. For example if the boundary hitting time \( t_*(x) \) is known to be finite for all \( x \in E^0 \), then (A8) and (A9) can be dispensed with altogether, as it is then not needed in the proof.

The following result which ensures the Lipschitz continuity of the value function was proved in Ye [18] (see also Dempster and Ye [10]).
**Proposition 2** Under assumptions (A1)-(A9), the value function $V(z)$ is Lipschitz continuous.

We now give the relationship between the maximum principle and dynamic programming as in the following theorem.

**Theorem 3** Under assumptions (A1)-(A9), there exist adjoint variables $(p_z, q_z)$ which satisfy (1)-(3) of Theorem 2 with $r = 1$ such that

$$p_z(t) = -\partial V(x^*_z(t)), \quad q_z(t) = -V(x^*_z(t))$$ (37)

**Proof** Denote the value function for the deterministic problem $(P_C)$ by $\bar{V}(z, z_0)$. Then we have

$$\bar{V}(z, z_0) := \inf_{\tilde{\Omega}} \left\{ \int_0^{t_*} x_0(t) f_0(x(t), u(t)) dt + x_0(t_*) F(x(t_*)) \right\}$$

$$= \inf_{\tilde{\Omega}} \left\{ \int_0^{t_*} z_0 e^{-\int_0^t \lambda(x(s), u(s)) ds} f_0(x(t), u(t)) dt + z_0 e^{-\int_0^{t_*} \lambda(x(s), u(s)) ds} F(x(t_*)) \right\}$$

$$= z_0 V(z),$$

where $V(z)$ is the value function for the problem $(P_z)$.

Since $z_0 > 0$ and $V(z)$ is Lipschitz continuous, $\bar{V}$ is Lipschitz continuous and one has

$$\partial_z \bar{V}(z, z_0) = z_0 \partial V(z) \times \{V(z)\}.$$ (38)

By Proposition 1, for each possible postjump state $z \in E^0$, the interjump control function $u^*_{0z}(-)$ and the corresponding trajectory $x^*_{z}(-)$ solves problem $(P_z)$ with problem data defined by (7), (8) and (9). As in the proof of Theorem 1, $(P_z)$ can be equivalently posted as $(P_C)$. In the case where $t_* < \infty$, problem $(P_C)$ is a standard optimal control problem with end point constraints. Since the value function is Lipschitz continuous by Proposition 2, the problem is “calm” (cf. Clarke [3]), hence normal. Using the result in Clarke and Vinter [4] or Vinter [17], we conclude that there exists adjoint variables $(\bar{p}_z, q_z)$ for problem $(P_C)$ which satisfy (16)-(18) with $r = 1$ such that

$$- (\bar{p}_z(t), q_z(t)) \in \partial_z \bar{V}(x^*_z(t), x^*_{0z}(t)) \quad t \in [0, t_*].$$ (39)

where $x^*_{0z}(t) := e^{-\int_0^t (\lambda(x^*_z(s), u^*_{0z}(s)) + \delta) ds}$.

In the case where $t_* (z) = \infty$, problem $(P_C)$ is an infinite horizon optimal control problem with zero discount rate. Although assumption (A9) of Ye [16, Theorem 3.1]
is not satisfied due to the fact that the discount rate is zero, by using the fact that the value function $\bar{V}$ is Lipschitz continuous, the proof of [16, Theorem 3.1] goes through. Therefore we can apply [16, Theorem 3.1] and obtain (39) in this case.

From (39) and (38), we have

$$\tilde{p}_z(t) \in -x_{0z}^*(t)\partial V(x_z^*(t)), \quad q_z(t) = -V(x_z^*(t)).$$

As in the proof of Theorem 1, by defining $p_z(\cdot) := \tilde{p}_z(\cdot)/x_{0z}^*(\cdot)$, we recover the adjoint variables for problem $(P_z)$ which satisfy

$$p_z(t) \in -\partial V(x_z^*(t)), \quad q_z(t) = -V(x_z^*(t)).$$

The following corollary of Theorem 3 gives a maximum principle for the PDP optimal control problem in normal form which does not involve the value function and which provides some asymptotic transversality conditions.

**Corollary 1** Under assumptions (A1)–(A9), let $u^* = (u_0^*, u_0^*)$ be an optimal ordinary control of the PDP optimal control problem. For each possible postjump state $z \in E^0$, let $u_{0z}(t)$ be the optimal interjump control function, and let $x_z^*(t)$ be the corresponding deterministic trajectory in $E^0$ on $[0, t_*]$ with initial point $z$. Then there exist: absolutely continuous functions

$$p_z : [0, t_*] \longrightarrow \mathbb{R}^n \quad q_z : [0, t_*] \longrightarrow \mathbb{R},$$

such that:-

1. The adjoint variables $(p_z, q_z)$ satisfy the adjoint equations in the form of the differential inclusions

$$-\dot{p}_z(t) \in (\partial_x f(x_z^*(t), u_{0z}^*(t)) - [\lambda(x_z^*(t), u_{0z}^*(t)) + \delta]I_n)^T p_z(t)$$

$$-q_z(t)\partial_x \lambda(x_z^*(t), u_{0z}^*(t)) - \partial_x I_0(x_z^*(t), u_{0z}^*(t))$$

$$+\lambda(x_z^*(t), u_{0z}^*(t))\int_{E^0} q_y(0)Q_0(dy; x_z^*(t), u_{0z}^*(t))]$$

a.e. $t \in [0, t_*]$ \hspace{1cm} (40)

$$-\dot{q}_z(t) = -q_z(t)(\lambda(x_z^*(t), u_{0z}^*(t)) + \delta) - [I_0(x_z^*(t), u_{0z}^*(t))$$

$$+\lambda(x_z^*(t), u_{0z}^*(t))\int_{E^0} q_y(0)Q_0(dy; x_z^*(t), u_{0z}^*(t))]$$

a.e. $t \in [0, t_*].$
(2) The optimal interior control function $u^*_0(t)$ maximizes the Hamiltonian point-wise, viz.

$$
\max_{v \in U_0} H(x^*_z(t), p_z(t), q_z(t), v, 1, q(0)) = H(x^*_z(t), p_z(t), q_z(t), u^*_0(t), 1, q(0)) = 0 \quad \text{a.e. } t \in [0, t_*]
$$

(3) The system is subject to the transversality condition: If $t_* < \infty$, then

$$
-(p_z(t_*), q_z(t_*)) \in \{\partial F(x^*_z(t_*)) + \rho \partial \delta_E(x^*_z(t_*))\} \times \{F(x^*_z(t_*))\}
$$

where $\rho > 0$ is some constant and

$$
F(x) := l_0(x, u^*_0(x)) + \int_{E^0} q_y(0) Q_0(dy; x, u^*_0(x));
$$

If $t_* = \infty$, then

$$
\lim_{t \to \infty} q_z(t) = 0, \\
\lim_{t \to \infty} p_z(t) e^{-\int_0^t (\lambda(x^*_z(s), u^*_0(s)) + \delta)ds} = 0. \quad (42)
$$

**Proof.** The assertion follows from the fact that every element in $\partial V(x^*_z(t))$ is bounded by the Lipschitz constant of $V(\cdot)$ (which exists by virtue of Proposition 2).

**Remark 3** In the case where $l_0, Q_0, \lambda$ does not depend on controls, we may define

$$
\tilde{p}_z(t) := p_z(t) e^{-\int_0^t (\lambda(x^*_z(s), u^*_0(s)) + \delta)ds}
$$

as the first adjoint variable. Then (42) gives the strong transversality condition

$$
\lim_{t \to \infty} \tilde{p}_z(t) = 0.
$$

6 Concluding remarks and extensions

First, we discuss the usefulness of the maximum principle to compute optimal controls. Consider the case in which the process dies at the first jump time. Then (1)-(4) of
Corollary 1 reduce to the statement of the ordinary deterministic maximum principle because \( q_y(0) = 0 \).

Let us now consider the case in which the postjump state is the finite state Markov process with jump rate \( \lambda_{ij} \). The resulting stochastic process is a special case of PDP called "systems with jump Markov disturbances". More precisely, it is a PDP taking values in state space \( E \subset \bigcup_i \mathcal{H}^n \times \{i\} \) with controlled dynamics

\[
\dot{x}(t) = f^r(t)(x(t), u(t)), \quad \dot{r}(t) = 0,
\]

jump rate

\[
\lambda^i(x, u) = \lambda_{ii} = -\sum_{j \neq i} \lambda_{ij}, \quad \forall (x, i) \in E^0,
\]

and transition measure

\[
Q^i_0(A \times I; x, u) = -\delta_{(x)}(A) \sum_{j \in I} \lambda_{ij},
\]

where \( \delta_{(x)} \) denotes the 1-atom probability measure concentrated on \( x \).

Assume that (A1)–(A9) holds. Let the process dies at first time it reaches the boundary of the state space (or at the first time it reaches a terminal set). For each possible postjump state \((z, i) \in E^0\), let \( u^i_z(\cdot) \) denote the optimal interjump control function and \( x^i_z(\cdot) \) the corresponding trajectory. By Corollary 1, there exist absolutely continuous functions

\[
p^i_z : [0, t_*] \rightarrow \mathcal{H}^n, \quad q^i_z : [0, t_*] \rightarrow \mathcal{H}
\]

such that

\[
-p^i_z(t) \in \left[ \partial_x f^i(x^i_z(t), u^i_z(t)) - (\lambda_{ii} + \delta) I_n \right]^T p^i_z(t) - \partial_x \left[ l_0^i(x^i_z(t), u^i_z(t)) \right] - \sum_{j \neq i} q^j z(t) \lambda_{ij} \quad \text{a.e. } t \in [0, t_*] \tag{43}
\]

\[
-q^i_z(t) = -q^i z(t) (\lambda_{ii} + \delta) - l_0^i(x^i_z(t), u^i_z(t)) + \sum_{j \neq i} q^j z(t) \lambda_{ij} \quad \text{a.e. } t \in [0, t_*] \tag{44}
\]

\[
\max_{u \in U^0_0} \left\{ p^i_z(t) f^i(x^i_z(t), u) - l_0^i(x^i_z(t), u) \right\} = p^i_z(t) f^i(x^i_z(t), u^i_z(t)) - l_0^i(x^i_z(t), u^i_z(t)) \quad \text{a.e. } t \in [0, t_*]. \tag{45}
\]

If \( t_* < \infty \), then

\[
-p^i_z(t_*) \in \partial l_0^i(x^i_z(t_*)) + \rho \partial d_{\mathcal{E}}(x^i_z(t_*))
\]
\[-q^i_z(t_*) = l^i_0(x^i_z(t_*)).\]

If \(t_* = \infty\), then

\[\lim_{t \to \infty} q^i_z(t) = 0, \quad \lim_{t \to \infty} p^i_z(t)e^{-(\lambda_{ii} + \delta)t} = 0.\]

The statement of the maximum principle reduces to Rishel [11, Theorem 12] under certain conditions. As it was pointed out by Rishel in [14], consider the case in which the transition probabilities of the finite state Markov process are such that the state space of the process can be divided into an ordered set of disjoint classes and transitions are possible only from a state in a higher class to one in a lower class. Then the state of the lowest class are all absorbing. For these states, equations (43)–(45) reduces to the statement of the ordinary maximum principle because \(\lambda_{ij} = 0, \ j \neq i\). If the optimal interjump controls and adjoint variables \(q^i_y\) for these states can be obtained through use of the maximum principle, then \(q^i_y\) can be used in equations (43)–(45) to attempt to obtain the optimal interjump control functions and the adjoint variables corresponding states in the second class. Proceeding by induction in this manner, a control computation somewhat more complicated than the control computation using the maximum principle in the deterministic case is obtained.

The maximum principle obtained here may be used in approximating optimal controls. The idea is as follows: First, solve the optimal control with the current approximation of the value function. Then compute the value function associated with this optimal cost which can be used in the next approximation of the value function. Repeating the procedure until the value function is not sensibly modified any more. Algorithms of this type has been exploited by Boukas, Haurie and Ch. van Delft in [2].

Finally, we should mention that the case of interior jumps (i.e. jumps starting from \(E^0\) according to jump rate \(\lambda\) adds nothing conceptually new to the original PDP control problem. Instead of including a cost \(l(x, u)\) for the interior jumps one may simply add \(\lambda(x, u)l(x, u)\) to the running cost \(l_0(x, u)\). The results obtained in this paper also holds in the case where the state space is taken to be a union of sets in \(R^n\) or even manifolds, where the boundaries have suitable smoothness properties imposed on section 1.
7 Acknowledgements

The research of this paper was supported by NSERC grant WFA0123160 and a Centre de Recherches Mathématiques Postdoctoral Fellowship at the University de Montréal. The author is grateful to professor F.H. Clarke for Helpful suggestions.

References


