Unconstrained and Constrained Optimal Control of Piecewise Deterministic Markov Processes

Oswaldo Costa, François Dufour, Alexey Piunovskiy

Universidade de Sao Paulo
Institut de Mathématiques de Bordeaux
INRIA Bordeaux Sud-Ouest
University of Liverpool

This study has been carried out with financial support from the French State, managed by the French National Research Agency (ANR) in the frame of the "Investments for the future" Programme IdEx Bordeaux - CPU (ANR-10-IDEX-03-02)
Outline

1. Piecewise deterministic Markov processes
   ▶ Introduction
   ▶ Parameters of the model
   ▶ Construction of the controlled process
   ▶ Admissible strategies
2. Optimization problems
   ▶ Unconstrained and constrained problems
   ▶ Assumptions
3. Non explosion
4. The unconstrained problem and the dynamic programming approach
5. The constrained problem and the linear programming approach
Introduction

Davis (80’s)

General class of non-diffusion stochastic hybrid models: deterministic trajectory punctuated by random jumps.

Applications

Engineering systems, biology, operations research, management science, economics, dependability and safety, ...
Parameters of the model

- the state space: $X$ open subset of $\mathbb{R}^d$ (boundary $\partial X$).
- the flow: $\phi(x, t) : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ satisfying $\phi(x, t + s) = \phi(\phi(x, s), t)$ for all $x \in \mathbb{R}^d$ and $(t, s) \in \mathbb{R}^2$.
- active boundary:
  $\Delta = \{z \in \partial X : z = \phi(x, t) \text{ for some } x \in X \text{ and } t \in \mathbb{R}_+^* \}$.
  For $x \in \bar{X} = X \cup \Delta$, $t^*(x) = \inf\{t \in \mathbb{R}_+ : \phi(x, t) \in \Delta\}$.

- $A$ is the action space, assumed to be a Borel space.
  $A^g \in \mathcal{B}(A)$ (respectively $A^i \in \mathcal{B}(A)$) is the set of gradual or continuous (respectively impulsive) actions satisfying $A = A^i + A^g$. 
Parameters of the model

- The set of *feasible* actions in state $x \in X$ is $A(x) \subset A$. Let us introduce the following sets $K = K^i \cup K^g$ with

$$K^g = \{(x, a) \in X \times A^g : a \in A(x)\}$$

$$K^i = \{(x, a) \in \Delta \times A^i : a \in A(x)\}$$

- The jumps intensity $\lambda$ which is a $\mathbb{R}_+^+$-valued measurable function defined on $K^g$.

- The stochastic kernel $Q$ on $X$ given $K$ satisfying $Q(X \setminus \{x\}|x, a) = 1$ for any $(x, a) \in K^g$. It describes the state of the process after any jump.
Uncontrolled process

Definition of a PDMP

Parameters: flow $\phi$, intensity of the jumps $\lambda$, transition kernel $Q$
Uncontrolled process

Definition of a PDMP

Parameters: flow \( \phi \), intensity of the jumps \( \lambda \), transition kernel \( Q \)
Uncontrolled process

Definition of a PDMP

Parameters: flow $\phi$, intensity of the jumps $\lambda$, transition kernel $Q$
Uncontrolled process

Definition of a PDMP

Parameters: flow \( \phi \), intensity of the jumps \( \lambda \), transition kernel \( Q \)
Uncontrolled process

Definition of a PDMP
Parameters: flow \( \phi \), intensity of the jumps \( \lambda \), transition kernel \( Q \)
Construction of the controlled process

The canonical space $\Omega = \bigcup_{n=1}^{\infty} \Omega_n \cup (X \times (\mathbb{R}_+^* \times X)^\infty)$ with $\Omega_n = X \times (\mathbb{R}_+^* \times X)^n \times (\{\infty\} \times \{x_\infty\})^\infty$.

Introduce the mappings $X_n : \Omega \rightarrow X_\infty = X \cup \{x_\infty\}$ by $X_n(\omega) = x_n$ and $\Theta_n : \Omega \rightarrow \mathbb{R}_+^*$ by $\Theta_n(\omega) = \theta_n$; $\Theta_0(\omega) = 0$ where

$$\omega = (x_0, \theta_1, x_1, \theta_2, x_2, \ldots) \in \Omega.$$

In addition $T_n(\omega) = \sum_{i=1}^{n} \Theta_i(\omega) = \sum_{i=1}^{n} \theta_i$ with $T_\infty(\omega) = \lim_{n \rightarrow \infty} T_n(\omega)$.

$H_n$ is the set of path up to $n$.

$H_n = (X_0, \Theta_1, X_1, \ldots, \Theta_n, X_n)$ is the history of the process up to $n$. 
Construction of the process

The controlled process $\{\xi_t\}_{t \in \mathbb{R}^+}$:

$$\xi_t(\omega) = \begin{cases} 
\phi(X_n, t - T_n) & \text{if } T_n \leq t < T_{n+1} \text{ for } n \in \mathbb{N}; \\
x_\infty & \text{if } T_\infty \leq t.
\end{cases}$$

The flow is not controlled.
Admissible strategies and conditional distribution

An admissible control strategy is a sequence \( u = (\pi_n, \gamma_n)_{n \in \mathbb{N}} \) such that, for any \( n \in \mathbb{N} \),

- \( \pi_n \) is a stochastic kernel on \( A^g \) given \( H_n \times \mathbb{R}^*_+ \):
  \[
  \pi_n(da|h_n, t) = 1 \text{ for } t \in ]0, t^*(x_n)[,
  \]
- \( \gamma_n \) is a stochastic kernel on \( A^i \) given \( H_n \):
  \[
  \gamma_n(da|h_n) = 1
  \]

where \( h_n = (x_0, \theta_1, x_1, \ldots \theta_n, x_n) \in H_n \).

The set of admissible control strategies is denoted by \( \mathcal{U} \).
Admissible strategies and conditional distribution

For an admissible control strategy $u = (\pi_n, \gamma_n)_{n \in \mathbb{N}}$, we can equivalently consider the random processes with values in $\mathcal{P}(A^g)$ and $\mathcal{P}(A^i)$ respectively as

$$\pi(da|t) = \sum_{n \in \mathbb{N}} l_{\{T_n < t \leq T_{n+1}\}} \pi_n(da|H_n, t - T_n)$$

and

$$\gamma(da|t) = \sum_{n \in \mathbb{N}} l_{\{T_n < t \leq T_{n+1}\}} \gamma_n(da|H_n),$$

for $t \in \mathbb{R}_+^*$. 


Admissible strategies and conditional distribution

Interaction of \( u = (\pi_n, \gamma_n)_{n \in \mathbb{N}} \) and the parameters of the model:

- the intensity of jumps
  \[
  \lambda^u_n(h_n, t) = \int_{A^g} \lambda(\phi(x_n, t), a)\pi_n(da|h_n, t),
  \]

and the corresponding rate of jumps

\[
\Lambda^u_n(h_n, t) = \int_{[0,t]} \lambda^u_n(h_n, s)ds,
\]

- the distribution of the state after a (stochastic) jump
  \[
  Q^{g,u}_n(dx|h_n, t) = \frac{1}{\lambda^u_n(h_n, t)} \int_{A^g} Q(dx|\phi(x_n, t), a)\lambda(\phi(x_n, t), a)\pi_n(da|h_n, t)
  \]

- the distribution of the state after a (boundary) jump
  \[
  Q^{i,u}_n(dx|h_n) = \int_{A^i} Q(dx|\phi(x_n, t^*(x_n)), a)\gamma_n(da|h_n).
  \]
Admissible strategies and conditional distribution

We want the joint distribution of the next sojourn time and state to be given by $G_n$

\[ G_n(\Gamma_1 \times \Gamma_2|h_n) = \left[ I\{x_n=x_\infty\} + e^{-\Lambda_n^u(h_n,+\infty)} I\{x_n\in\mathbb{X}\} I\{t^*(x_n)=\infty\} \right] \delta(+\infty,x_\infty)(\Gamma_1 \times \Gamma_2) \]

\[ + I\{x_n\in\mathbb{X}\} \left[ \delta_{t^*(x_n)}(\Gamma_1) Q_n^{i^*,u}(\Gamma_2|h_n) e^{-\Lambda_n^u(h_n,t^*(x_n))} I\{t^*(x_n)<\infty\} \right] \]

\[ + \int_{0,t^*(x_n)[\cap \Gamma_1} Q_n^{g^*,u}(\Gamma_2|h_n, t) \lambda_n^u(h_n, t) e^{-\Lambda_n^u(h_n,t)} dt \],

where $\Gamma_1 \in \mathcal{B}(\bar{\mathbb{R}}^*_+)$, $\Gamma_2 \in \mathcal{B}(\mathbb{X}_\infty)$ and $h_n = (x_0, \theta_1, x_1, \ldots, \theta_n, x_n) \in \mathcal{H}_n$. 
Consider an admissible strategy $u \in \mathcal{U}$ and an initial state $x_0 \in \mathcal{X}$

$$\mathbb{P}^u_{x_0}\left((\Theta_{n+1}, X_{n+1}) \in \Gamma_1 \times \Gamma_2 | \mathcal{F}_{T_n}\right) = G_n(\Gamma_1 \times \Gamma_2 | H_n)$$

→ the conditional distribution of $(\Theta_{n+1}, X_{n+1})$ given $\mathcal{F}_{T_n}$ under $\mathbb{P}^u_{x_0}$ is $G_n(\cdot | H_n)$ ($\{\mathcal{F}_t\}$ is the natural filtration of the process).
Admissible strategies and conditional distribution

Consider an admissible strategy \( u \in \mathcal{U} \) and an initial state \( x_0 \in X \). There exists a probability \( \mathbb{P}^u_{x_0} \) on \( (\Omega, \mathcal{F}) \) such that

\[
\mathbb{P}^u_{x_0}(\{X_0 = x_0\}) = 1
\]

and the positive random measure \( \nu \) defined on \( \mathbb{R}^*_+ \times X \) by

\[
\nu(dt, dx) = \sum_{n \in \mathbb{N}} \frac{G_n(dt - T_n, dx | H_n)}{G_n([t - T_n, +\infty] \times X_\infty | H_n)} I\{T_n < t \leq T_{n+1}\}
\]

is the compensator of

\[
\mu(dt, dx) = \sum_{n \geq 1} I\{T_n(\omega) < \infty\} \delta(T_n(\omega), X_n(\omega))(dt, dx).
\]

with respect to \( \mathbb{P}^u_{x_0} \) (Jacod, *Multivariate point processes*, 1975).
Outline

1. Piecewise deterministic Markov processes
   - Introduction
   - Parameters of the model
   - Construction of the controlled process
   - Admissible strategies

2. Optimization problems
   - Unconstrained and constrained problems
   - Assumptions

3. Non explosion

4. The unconstrained problem and the dynamic programming approach

5. The constrained problem and the linear programming approach
Unconstrained and constrained problems

Cost functions

- \((C^g_j)_{j \in \{0, 1, \ldots, p\}}\) associated with a continuous action. Real-valued mapping defined on \(K^g\).

- \((C^i_j)_{j \in \{0, 1, \ldots, p\}}\) associated with an impulsive action on the boundary. Real-valued mapping defined on \(K^i\).

The associated infinite-horizon discounted criteria corresponding to an admissible control strategy \(u \in \mathcal{U}\) are given by

\[
\mathcal{V}_j(u, x_0) = \mathbb{E}^u_{x_0} \left[ \int_{0, +\infty} e^{-\alpha s} \int_{A(\xi_s)} C^g_j(\xi_s, a) \pi(da|s) ds \right]
\]

\[
+ \mathbb{E}^u_{x_0} \left[ \int_{0, +\infty} e^{-\alpha s} I\{\xi_s- \in \Delta\} \int_{A(\xi_s-)} C^i_j(\xi_s-, a) \gamma(da|s) \mu(ds, X) \right]
\]

for any \(j \in \{0, 1, \ldots, p\}\).
Unconstrained and constrained problems

- The optimization problem without constraint consists in minimizing the performance criterion

$$\inf_{u \in \mathcal{U}} \mathcal{V}_0(u, x_0).$$

- The optimization problem with $p$ constraints consists in minimizing the performance criterion

$$\inf_{u \in \mathcal{U}} \mathcal{V}_0(u, x_0)$$

such that the constraint criteria

$$\mathcal{V}_j(u, x_0) \leq B_j$$

are satisfied for any $j \in \mathbb{N}_p^*$, where $(B_j)_{j \in \mathbb{N}_p^*}$ are real numbers representing the constraint bounds.
Different classes of strategies

- **feasible**, if \( u \in \mathcal{U} \) and \( \forall j(u, x_0) \leq B_j \), for \( j \geq 1 \).

- **stationary**, if for some \((\pi, \gamma) \in \mathcal{P}^g \times \mathcal{P}^i\) the control strategy \( u = (\pi_n, \gamma_n)_{n \in \mathbb{N}} \) is given by \( \pi_n(da|h_n, t) = \pi(da|\phi(x_n, t)) \) and \( \gamma_n(db|h_n) = \gamma(db|\phi(x_n, t^*(x_n))) \).

- **non-randomized stationary**, if \( \pi_n(\cdot|h_n, t) = \delta_{\varphi^s(\phi(x_n,t))}(\cdot) \) and \( \gamma_n(\cdot|h_n) = \delta_{\varphi^s(\phi(x_n,t))}(\cdot) \), where \( \varphi^s : \overline{X} \to \mathcal{A} \) is a measurable mapping satisfying \( \varphi^s(y) \in \mathcal{A}(y) \) for any \( y \in \overline{X} \).
Hypotheses

Assumption A. There are constants $K \geq 0, \varepsilon_1 > 0$ and $\varepsilon_2 \in [0, 1[$ such that

(A1) For any $(x, a) \in K^g$, $\lambda(x, a) \leq K$

(A2) $\inf_{(z, b) \in K^i \cap K} Q(A_{\varepsilon_1}|z, b) \geq 1 - \varepsilon_2$, with

$$A_{\varepsilon_1} = \{x \in X : t^*(x) > \varepsilon_1\}.$$

Assumption B.

(B1) The set $A(y)$ is compact for every $y \in \overline{X}$.

(B2) The kernel $Q$ is weakly continuous.

(B3) The function $\lambda$ is continuous on $K^g$.

(B4) The flow $\phi$ is continuous on $\mathbb{R}_+ \times \mathbb{R}^p$.

(B5) The function $t^*$ is continuous on $\overline{X}$.
Assumption C.

(C1) The multifunction $\Psi^g$ from $X$ to $A$ defined by $\Psi(x) = A(x)$ is upper semicontinuous. The multifunction $\Psi$ from $\Delta$ to $A$ defined by $\Psi^i(z) = A(z)$ is upper semicontinuous.

(C2) The cost function $C^g_0$ (respectively, $C^i_0$) is bounded and lower semicontinuous on $K^g$ (respectively, $K^i$).
Outline

1. Controlled piecewise deterministic Markov processes
   - Introduction
   - Parameters of the model
   - Construction of the process
   - Admissible strategies

2. Optimization problems
   - Unconstrained and constrained problems
   - Different classes of strategies
   - Hypotheses

3. Non explosion

4. The unconstrained problem and the dynamic programming approach

5. The constrained problem and the linear programming approach
Lemma

Suppose Assumption A is satisfied. Then there exists $M < \infty$ such that, for any control strategy $u \in \mathcal{U}$ and for any $x_0 \in \mathcal{X}$

$$
\mathbb{E}^u_{x_0} \left[ \sum_{n \in \mathbb{N}^*} e^{-\alpha T_n} \right] \leq M \quad \text{and} \quad \mathbb{P}^u_{x_0}(T_\infty < +\infty) = 0.
$$
Elements of proof:

- For any control strategy $u$, $x_0 \in X$ we have for any $j \in \mathbb{N}$

$$\mathbb{P}^u_{x_0}(\Theta_{j+2} + \Theta_{j+1} > \varepsilon_1 | H_j) \geq e^{-2K\varepsilon_1}(1 - \varepsilon_2).$$

- Now,

$$\mathbb{E}^u_{x_0} \left[ e^{-\alpha(\Theta_{j+1} + \Theta_{j+2})} | H_j \right]$$

$$\leq \mathbb{P}^u_{x_0}(\Theta_{j+1} + \Theta_{j+2} \leq \varepsilon_1 | H_j)$$

$$+ e^{-\alpha \varepsilon_1} \mathbb{P}^u_{x_0}(\Theta_{j+1} + \Theta_{j+2} > \varepsilon_1 | H_j)$$

$$= 1 + [e^{-\alpha \varepsilon_1} - 1] \mathbb{P}^u_{x_0}(\Theta_{j+1} + \Theta_{j+2} > \varepsilon_1 | H_j)$$

$$\leq 1 + [e^{-\alpha \varepsilon_1} - 1][1 - \varepsilon_2]e^{-2K\varepsilon_1} = \kappa < 1.$$
Elements of proof:

- For any $j \in \mathbb{N}^*$,

$$
\mathbb{E}_x^u \left[ e^{-\alpha T_{2j+1}} \right] = \mathbb{E}_x^u \left[ e^{-\alpha T_{2j-1}} \mathbb{E}_x^u \left[ e^{-\alpha (\Theta_{2j} + \Theta_{2j+1})} | H_{2j-1} \right] \right]
\leq \kappa \mathbb{E}_x^u \left[ e^{-\alpha T_{2j-1}} \right],
$$

and so

$$
\mathbb{E}_x^u \left[ e^{-\alpha T_{2j+1}} \right] \leq \kappa^j \mathbb{E}_x^u \left[ e^{-\alpha T_1} \right] \leq \kappa^j.
$$

Similarly,

$$
\mathbb{E}_x^u \left[ e^{-\alpha T_{2j+2}} \right] \leq \kappa^j \mathbb{E}_x^u \left[ e^{-\alpha T_2} \right] \leq \kappa^j.
$$

for any $j \in \mathbb{N}$.

- Therefore,

$$
\mathbb{E}_x^u \left[ \sum_{n \in \mathbb{N}^*} e^{-\alpha T_n} \right] \leq \frac{2}{1 - \kappa}.
$$
Outline

1. Controlled piecewise deterministic Markov processes
   ▶ Introduction
   ▶ Parameters of the model
   ▶ Construction of the process
   ▶ Admissible strategies

2. Optimization problems
   ▶ Unconstrained and constrained problems
   ▶ Different classes of strategies
   ▶ Hypotheses

3. Non explosion

4. The unconstrained problem and the dynamic programming approach

5. The constrained problem and the linear programming approach
There are two approaches to deal with such problems:

- **the associated discrete-stage Markov decision model:**
There are two approaches to deal with such problems:

- **the infinitesimal approach (HJB equation):**
Notation and preliminary results:

▶ \( \mathbb{A}(\mathbb{X}) \) is the set of functions \( g \in \mathbb{B}(\mathbb{X}) \) such that for any \( x \in \mathbb{X} \), the function \( g(\phi(x, \cdot)) \) is absolutely continuous on \( [0, t^*(x)] \cap \mathbb{R}^+ \).

▶ Let \( g \in \mathbb{A}(\mathbb{X}) \), there exists a real-valued measurable function \( \lambda g \) defined on \( \mathbb{X} \) satisfying for any \( t \in [0, t^*(x)] \)

\[
g(\phi(x, t)) = g(x) + \int_{[0,t]} \lambda g(\phi(x, s)) \, ds.
\]

▶ Let \( R \in \mathcal{P}(\mathbb{X} | \mathbb{Y}) \). Then \( Rf(y) \overset{\text{def}}{=} \int_{\mathbb{X}} f(x) R(dx | y) \) for any \( y \in \mathbb{Y} \) and measurable function \( f \). For any measure \( \eta \) on \( (\mathbb{Y}, \mathcal{B}(\mathbb{Y})) \), \( \eta R(\cdot) \overset{\text{def}}{=} \int_{\mathbb{Y}} R(\cdot | y) \eta(dy) \).

▶ \( q(dy | x, a) \overset{\text{def}}{=} \lambda(x, a)[Q(dy | x, a) - \delta_x(dy)] \)
Sufficient conditions for the existence of a solution for the HJB equation associated with the optimization problem.

**Theorem**

*Suppose assumptions A, B and C hold. Then there exist $W \in \mathbb{A}(X)$ and $W' \in \mathbb{B}(X)$ satisfying*

\[-\alpha W(x) + x W(x) + \inf_{a \in A^g(x)} \left\{ C^g_0(x, a) + q W(x, a) \right\} = 0,

*for any $x \in X$, and*

\[W(z) = \inf_{b \in A^i(z)} \left\{ C^i_0(z, b) + Q W(z, b) \right\},

*for any $z \in \Delta$. Moreover, for any $x \in X$*

\[W(x) = \inf_{u \in \mathcal{U}} V_0(u, x).

Sufficient conditions for the existence of an optimal strategy.

**Theorem**

Suppose assumptions A, B and C hold. There exists a measurable mapping \( \hat{\varphi} : \overline{X} \rightarrow A \) such that \( \hat{\varphi}(y) \in A(y) \) for any \( y \in \overline{X} \) and satisfying

\[
C^g_0(x, \hat{\varphi}(x)) + qW(x, \hat{\varphi}(x)) = \inf_{a \in A(x)} \{ C^g_0(x, a) + qW(x, a) \}
\]

for any \( x \in X \), and

\[
C^i_0(z, \hat{\varphi}(z)) + QW(z, \hat{\varphi}(z)) = \inf_{b \in A(z)} \{ C^i_0(z, b) + QW(z, b) \}.
\]

for any \( z \in \Delta \). Moreover, the stationary non-randomized strategy \( \hat{\varphi} \) is optimal.
Elements of proof:

- Define recursively \( \{W_i\}_{i \in \mathbb{N}} \) as

\[
W_{i+1}(y) = \mathcal{B} W_i(y),
\]

with \( W_0(y) = -K_A l_{A_{\varepsilon_1}}(y) - (K_A + K_B) l_{A_{\varepsilon_1}}(y) \) and

\[
\mathcal{B} V(y) = \int_{0\, t^*(y)[} e^{-(K+\alpha)t} \mathcal{R} V(\phi(y, t))dt
\]

\[+ e^{-(K+\alpha)t^*(y)} \mathcal{I} V(\phi(y, t^*(y))),\]

where

\[
\mathcal{R} V(x) = \inf_{a \in A(x)} \left\{ C_g^0(x, a) + qV(x, a) + KV(x) \right\},
\]

and

\[
\mathcal{I} V(z) = \inf_{b \in A(z)} \left\{ C_i^0(z, b) + QV(z, b) \right\}.
\]
\begin{itemize}
  \item $W_i$ is lower semicontinuous and
  \[ |W_i(y)| \leq K_A I_{A_{\epsilon_1}}(y) + (K_A + K_B) I_{A_{\epsilon_1}}(y). \]
  \item $\mathcal{B}$ is monotone ($V_1 \leq V_2 \Rightarrow \mathcal{B} V_1 \leq \mathcal{B} V_2$), $\{W_i\}_{i \in \mathbb{N}}$ is increasing and $W_i \rightarrow W$ and $W$ is bounded and lower semicontinuous.
  \item $\lim_{i \rightarrow \infty} \mathcal{R} W_i(x) = \mathcal{R} W(x)$, for any $x \in X$
  $\lim_{i \rightarrow \infty} \mathcal{T} W_i(z) = \mathcal{T} W(z)$ for any $z \in \Delta$. 
\end{itemize}
By using the bounded convergence Theorem,

\[ W(y) = \mathcal{B} W(y) = \int_{[0,t^*(y)[} e^{-(K+\alpha)t} \mathcal{R} W(\phi(y, t)) dt \]

\[ + e^{-(K+\alpha)t^*(y)} \mathcal{T} W(\phi(y, t^*(y)))) , \]

where \( y \in \overline{X} \).

Then \( W \in \mathbb{A}(\overline{X}) \) and there exists \( \mathcal{X} W \in \mathbb{B}(X) \)

\[-\alpha W(x) + \mathcal{X} W(x) + \inf_{a \in \mathcal{A}(x)} \left\{ C_{0}^{g}(x, a) + q W(x, a) \right\} = 0, \]

for any \( x \in X \), and

\[ W(z) = \inf_{b \in \mathcal{A}(z)} \left\{ C_{0}^{i}(z, b) + Q W(z, b) \right\} , \]

for any \( z \in \Delta \).
Outline

1. Controlled piecewise deterministic Markov processes
   - Introduction
   - Parameters of the model
   - Construction of the process
   - Admissible strategies

2. Optimization problems
   - Unconstrained and constrained problems
   - Different classes of strategies
   - Hypotheses

3. Non explosion

4. The unconstrained problem and the dynamic programming approach

5. The constrained problem and the linear programming approach
The linear programming approach

The method has been extensively studied in the literature

- **Continuous and discrete time MDP:**

- **Controlled martingale problems:**
Occupation measure

For any admissible control strategy $u \in \mathcal{U}$, the occupation measure $\eta_u \in \mathcal{M}(K)$ associated with $u$ is defined as follows

$$\eta_u(\Gamma) = \mathbb{E}_{x_0}^u \left[ \int_{\Gamma \cap K^g} \int_{]0,\infty[} e^{-\alpha s} \delta_{\xi_s}(dx) \pi(da \mid s) ds \right]$$

$$+ \mathbb{E}_{x_0}^u \left[ \int_{\Gamma \cap K^i} \sum_{n \in \mathbb{N}^*} e^{-\alpha T_n} \delta_{\xi_{T_n-}}(dz) \gamma(db \mid T_n-) \right].$$

for any $\Gamma \in \mathcal{B}(K)$. 
Linear programming approach

The infinite-horizon discounted criteria can be rewritten as

\[
\mathcal{V}_j(u, x_0) = \mathbb{E}^u_{x_0} \left[ \int_{0, +\infty} e^{-\alpha s} \int_{A(\xi_s)} C^g_j(\xi_s, a) \pi(da|s)ds \right] \\
+ \mathbb{E}^u_{x_0} \left[ \int_{0, +\infty} e^{-\alpha s} \mathbb{I}_{\{\xi_s \in \Delta\}} \int_{A(\xi_{s-})} C^i_j(\xi_{s-}, a) \gamma(da|s)\mu(ds, X) \right]
= \eta^g_u(C^g_j) + \eta^i_u(C^i_j)
\]

where \(\eta^g_u\) (resp. \(\eta^i_u\)) denotes the restriction of \(\eta_u\) to \(K^g\) (resp. \(K^i\)).
A finite measure $\eta \in \mathcal{M}(\mathbf{K})$ is called admissible if, for any $(W, \mathcal{X}W) \in \mathbb{A}(\mathbf{X}) \times \mathbb{B}(\mathbf{X})$, the following equality holds

$$
\int_{\mathbf{X}} \left[ \alpha W(x) - \mathcal{X}W(x) \right] \hat{\eta}^g(dx) + \int_{\Delta} W(z)\hat{\eta}^i(dz)
$$

$$
= W(x_0) + \int_{\mathbf{K}^g} qW(x, a)\eta^g(dx, da) + \int_{\mathbf{K}^i} QW(z, b)\eta^i(dz, db).
$$

with $\hat{\eta}^g$ (resp. $\hat{\eta}^i$) denotes the marginal of $\eta^g$ (resp. $\eta^i$) w.r.t. to $\mathbf{X}$. 

Admissible measure
Occupation and admissible measures

The next important result shows the link between the set of admissible measures and the set of occupation measures.

**Theorem**

*Suppose Assumption A is satisfied. Then the following assertions hold.*

i) For any control strategy $u \in \mathcal{U}$, the occupation measure $\eta_u$ is admissible.

ii) Suppose that the measure $\eta$ is admissible. Then there exist stochastic kernels $\pi \in \mathcal{P}^g$ and $\gamma \in \mathcal{P}^i$ for which the stationary control strategy $u = (\pi, \gamma) \in \mathcal{U}_s$ satisfies $\eta = \eta_u$. 
Linear programming approach

The constrained linear program, labeled $\text{LP}$, is defined as

$$\inf_{(\eta^g, \eta^i) \in \mathcal{M}} \eta^g(C_0^g) + \eta^i(C_0^i)$$

where $\mathcal{M}$ is the set of measures $(\eta^g, \eta^i)$ in $\mathcal{M}(K^i) \times \mathcal{M}(K^g)$ such that $\eta^g + \eta^i$ is admissible and satisfies

$$\eta^g(C_j^g) + \eta^i(C_j^i) \leq B_j.$$
Linear programming approach

**Theorem**

*Suppose Assumption A holds and the cost functions $C^g_j$ and $C^i_j$ are bounded from below for any $j \in \mathbb{N}_p$. Then the values of the constrained control problem and the linear program $\mathbb{LP}$ are equivalent:*

$$\inf_{(\eta^g, \eta^i) \in \mathbb{M}} \eta^g(C^g_0) + \eta^i(C^i_0) = \inf_{u \in \mathbb{U}^f} \mathcal{V}_0(u, x_0).$$
Theorem

Suppose Assumptions A, B and (C1) are satisfied. Assume the cost functions $C^g_j$ (resp. $C^i_j$) are bounded from below and lower semicontinuous on $K^g_j$ (resp. $K^i_j$) for any $j \in \mathbb{N}_p$.

If the set of feasible strategies is non empty then the $\text{LP}$ is solvable and there exists a stationary feasible strategy $u^*$ satisfying

$$
\eta^g_{u^*}(C^g_0) + \eta^i_{u^*}(C^i_0) = \inf_{(\eta^g, \eta^i) \in \mathcal{M}} \eta^g(C^g_0) + \eta^i(C^i_0)
$$

$$
= \inf_{u \in \mathcal{U}^f} \mathcal{V}_0(u, x_0) = \mathcal{V}_0(u^*, x_0).
$$