Temporal Difference Methods for the Maximal Solution of Discrete-Time Coupled Algebraic Riccati Equations

Oswaldo L. V. Costa *,† Julio C.C. Aya†
Departamento de Engenharia de Telecomunicações e Controle
Escola Politécnica da Universidade de São Paulo
CEP: 05508 900 São Paulo SP Brazil
oswaldo@lac.usp.br, julio@lac.usp.br

15th February 2000

Abstract

In this paper we present an iterative technique for deriving the maximal solution of a set of discrete-time coupled algebraic Riccati equations (CARE), based on temporal difference methods. CARE are related to the optimal control of Markovian jump linear systems and have been extensively studied over the last few years. We trace a parallel with the theory of temporal difference algorithms for Markovian decision processes to develop a λ-policy iteration like algorithm for the maximal solution of these equations. For the special cases in which $\lambda = 0$ and $\lambda = 1$ we have the situation in which the algorithm reduces to the iterations of the Riccati difference equations (value iteration) and quasi-linearization method (policy iteration) respectively. The advantage of the proposed method is that an appropriate choice of λ between 0 and 1 can speed up the convergence of the policy evaluation step of the policy iteration method by using value iteration.

Keywords: temporal difference methods, jump systems, Markov parameters, optimal control.

This work was suported in part by:
*FAPESP (Research Council of the State of São Paulo)
†CNPq(Brazilian National Research Council)

1 Introduction

In this paper we consider the following class of models in an appropriate probabilistic space $(\Omega, P, \{\mathcal{F}_k\}, \mathcal{F})$, known in the international literature as discrete-time Markovian jump linear systems (cf. [14]):

$$\begin{aligned} x(k+1) &= A_{\theta(k)}x(k) + B_{\theta(k)}u(k) \\ x(0) &= x_0, \theta(0) = \theta_0. \end{aligned} \tag{1}$$

Here $\theta(k)$ is a Markov chain taking values in $\{1, \ldots, N\}$ with transition probability matrix $\mathcal{P} = (p_{ij})$. Let $\Sigma = \{u = (u(0), \ldots); u(k)$ is $\mathcal{F}_k$-measurable for each $k\}$. For $u \in \Sigma$, one considers the following quadratic cost for system (1)

$$J_{(x_0,\theta_0)}(u) = E_{(x_0,\theta_0)}\left(\sum_{k=0}^{\infty}\left(\left\|C_{\theta(k)}x(k)\right\|^2 + \left\|D_{\theta(k)}u(k)\right\|^2\right)\right) \tag{2}$$

and it is desired to minimize (2) over $u \in \Sigma$. It has been shown in the literature (cf. [5], [11],[17]) that the solution of this problem is associated to the existence of a solution $P = (P_1, \ldots, P_N)$, $P_i \geq 0$, $i = 1, \ldots, N$, to the following set of coupled algebraic Riccati equations (CARE), for $i = 1, \ldots, N$

$$X_i = A_i'\mathcal{E}_i(X)A_i + C_i'C_i - A_i'\mathcal{E}_i(X)B_i\left(B_i'\mathcal{E}_i(X)B_i + D_i'D_i\right)B_i'\mathcal{E}_i(X)A_i \tag{3}$$

where $\mathcal{E}(X) = (\mathcal{E}_1(X), \ldots, \mathcal{E}_N(X))$ is defined, for $X = (X_1, \ldots, X_N)$ as

$$\mathcal{E}_i(X) = \sum_{j=1}^{N} p_{ij}X_j. \tag{4}$$

If such a solution P exists, it can be shown (see [5]) that the optimal control law for the problem posed by equations (1), (2) and (3) is given by the feedback control law

$$u(k) = F_{\theta(k)}x(k) \tag{5}$$

where $F = (F_1, \ldots, F_N)$ is given by

$$F_i = -\left(B_i'\mathcal{E}_i(P)B_i + D_i'D_i\right)^{-1}B_i'\mathcal{E}_i(P)A_i. \tag{6}$$

Related to the infinite horizon problem in (1), (2), we have the finite horizon problem (7) in which, for some $S = (S_1, \ldots, S_N) \in \mathbb{H}^{n+}$, it is desired to find $u(0), \ldots, u(\kappa)$ such that minimizes the functional

$$\begin{aligned} J_{(x_0,\theta_0)}(u) = E_{(x_0,\theta_0)}\Bigg(&\sum_{k=0}^{\kappa}\left(\left\|C_{\theta(k)}x(k)\right\|^2 + \left\|D_{\theta(k)}u(k)\right\|^2\right) \\ &+ x(\kappa+1)'S_{\theta(\kappa+1)}x(\kappa+1)\Bigg). \end{aligned} \tag{7}$$

As proved in [5], the solution of this problem can be obtained from the following set of coupled Riccati difference equations for $k = \kappa, \ldots, 0$,

$$\begin{aligned} P_i(k) &= A_i'\mathcal{E}_i(P(k+1))A_i + C_i'C_i \\ &\quad - A_i'\mathcal{E}_i(P(k+1))B_i\left(B_i'\mathcal{E}_i(P(k+1))B_i + D_i'D_i\right)B_i'\mathcal{E}_i(P(k+1))A_i \\ &= (A_i + B_iF_i(k))'\mathcal{E}_i(P(k+1))(A_i + B_iF_i(k)) + C_i'C_i + F_i(k)'D_i'DF_i(k) \end{aligned} \tag{8}$$

where $P(\kappa+1) = S$ and

$$F_i(k) = -\left(B_i'\mathcal{E}_i(P(k+1))B_i + D_i'D_i\right)^{-1} B_i'\mathcal{E}_i(P(k))A_i. \tag{9}$$

The optimal control is given by

$$u(k) = F_{\theta(k)}(k)x(k)$$

and under some appropriate conditions (see [5]), $P(k)$ converges to a positive semi-definite solution P of (3).

Tracing a parallel between the above method and the theory of Markovian decision processes (MDP for short), we could relate the iterations (8),(9) with those in the so called value iteration technique in MDP. On the other hand, as will be presented in Lemma 4 below (see remark 1), the so-called quasi-linearization method for deriving the solution of (3) could be seen as the policy iteration technique for MDP, involving a policy evaluation part, and a policy improvement part.

The goal of this paper is to apply temporal difference methods for obtaining the optimal solution P of the CARE (3), (4). Some numerical algorithms based on Riccati equations, Lyapunov equations, and convex programming have been proposed in the current literature (cf. [9], [1], [7], [16], [10], [8], [3]). However if the number of operating modes (N) of the system is too large, an specific method which uses only value iteration techniques with a fast rate of convergence could be advantageous.

Temporal difference (TD for short) methods have been applied to solve problems related to MDP (see for instance [2],[15]). Usually it is preferable to carry out the policy evaluation step of the policy iteration method by using value iteration and a drawback of this approach is that the value iteration algorithm may converge very slow for discount factor close to 1. The idea of TD methods is to asymptotically induce a reduction on the discount factor without altering the cost-to-go of a given policy. This reduction on the discount factor depends on a parameter usually named λ, given origin to the TD λ-policy iteration methods. For $\lambda = 1$, the method reduces to the policy iteration technique, while that for $\lambda = 0$, we have the value iteration technique.

In this paper we shall apply these ideas to obtain an iterative method for the maximal solution of (3), (4), to be defined in the next section. As we shall see below, this iterative method traces a close parallel with the TD λ-policy iteration methods in MDP. For $\lambda = 0$ our technique reduces to that presented in equations (8), (9) (value iteration), while for $\lambda = 1$ it reduces to the quasi-linearization method presented in Lemma 4 below (policy iteration).

The paper is presented in the following way. Section 2 deals with the notation and some preliminary results. Section 3 presents the temporal difference method for solving the CARE (3). The main result is Theorem 1, in which the proof of convergence is established. Section 4 presents a numerical example and section 5 concludes the paper with some final comments.

2 Notation and Preliminary Results

For $\mathbb{X}$ and $\mathbb{Y}$ complex Banach spaces we set $\mathbb{B}(\mathbb{X},\mathbb{Y})$ the Banach space of all bounded linear operators of $\mathbb{X}$ into $\mathbb{Y}$, with the uniform induced norm represented by $\|\cdot\|$. For simplicity we shall set $\mathbb{B}(\mathbb{X}) := \mathbb{B}(\mathbb{X},\mathbb{X})$. The spectral radius of an operator $\mathcal{G} \in \mathbb{B}(\mathbb{X})$ will be denoted by $r_\sigma(\mathcal{G})$. If $\mathbb{X}$ is a Hilbert space then the inner product will be denoted by $\langle .;. \rangle$, and for $\mathcal{G} \in \mathbb{B}(\mathbb{X})$, $\mathcal{G}^*$ will denote the adjoint operator of $\mathcal{G}$. As usual, $\mathcal{G} \geq 0$ ($\mathcal{G} > 0$ respectively) will denote that the operator $\mathcal{G} \in \mathbb{B}(\mathbb{X})$ will be positive semi-definite (positive definite). In particular we shall denote by $\mathbb{C}^n$ the n dimensional complex Euclidean spaces and by $\mathbb{B}(\mathbb{C}^n,\mathbb{C}^m)$ the normed bounded linear space of all $m \times n$ complex matrices, with $\mathbb{B}(\mathbb{C}^n) := \mathbb{B}(\mathbb{C}^n,\mathbb{C}^n)$.

Set $\mathbb{H}^{n,m}$ the linear space made up of all N-sequences of complex matrices $V = (V_1,\ldots,V_N)$ with $V_i \in \mathbb{B}(\mathbb{C}^n,\mathbb{C}^m)$, $i = 1,\ldots,N$ and, for simplicity, set $\mathbb{H}^n := \mathbb{H}^{n,n}$. For $V = (V_1,\ldots,V_N) \in \mathbb{H}^{n,m}$, we consider the following norm in $\mathbb{H}^{n,m}$

$$\|V\| := \sum_{i=1}^{N} \|V_i\| . \tag{10}$$

It is easy to verify that $\mathbb{H}^{n,m}$ equipped with the above norm is a Banach space. We set

$$\mathbb{H}^{n+} := \{V = (V_1,\ldots,V_N) \in \mathbb{H}^n ; V_i \geq 0, \quad i = 1,\ldots,N\}$$

and shall write, for $V = (V_1,\ldots,V_N) \in \mathbb{H}^n$ and $S = (S_1,\ldots,S_N) \in \mathbb{H}^n$, that $V \geq S$ if

$$V - S = (V_1 - S_1,\ldots,V_N - S_N) \in \mathbb{H}^{n+},$$

and that $V > S$ if $V_i - S_i > 0$ for $i = 1,\ldots,N$. For $\Gamma = (\Gamma_1,\ldots,\Gamma_N) \in \mathbb{H}^n$ we define the following operator $\mathcal{L}(\cdot) = (\mathcal{L}_1(\cdot),\ldots,\mathcal{L}_N(\cdot)) \in \mathbb{B}(\mathbb{H}^n)$ for $V = (V_1,\ldots,V_N) \in \mathbb{H}^n$ and $i,j = 1,\ldots,N$,

$$\mathcal{L}_i(V) := \Gamma_i' \mathcal{E}_i(\cdot)\Gamma_i \tag{11}$$

$$\mathcal{G}_j(V) := \sum_{i=1}^{N} p_{ij}\Gamma_i V_i \Gamma_i' \tag{12}$$

where the operator $\mathcal{E}(\cdot) = (\mathcal{E}_1(\cdot),\ldots,\mathcal{E}_N(\cdot)) \in \mathbb{B}(\mathbb{H}^n)$ is defined as in (4). It is easy to verify that the operators $\mathcal{E}$, $\mathcal{L}$, and $\mathcal{G}$ map $\mathbb{H}^{n+}$ into $\mathbb{H}^{n+}$, and that $r_\sigma(\mathcal{L}) = r_\sigma(\mathcal{G})$.

We assume in model (1) and cost function (2) that $A = (A_1,\ldots,A_N) \in \mathbb{H}^n$, $B = (B_1,\ldots,B_N) \in \mathbb{H}^{m,n}$, $C = (C_1,\ldots,C_N) \in \mathbb{H}^{n,p}$ and $D = (D_1,\ldots,D_N) \in \mathbb{H}^{m,p}$. It has been shown in [4], [13], that model (1) with $u(k) = F_{\theta(k)}x(k)$, and $V_i(k) = E\left(x(k)x(k)^* 1_{\{\theta(k)=i\}}\right)$, $V(k) = (V_1(k),\ldots,V_N(k)) \in \mathbb{H}^{n+}$ leads to

$$V(k+1) = \mathcal{G}\big(V(k)\big), \qquad k = 0,1,\ldots,$$

where $\Gamma_i = A_i + B_i F_i$ in (12). Next we define the stability concept that we shall consider in the following sections.

Definition 1 *We say that* $F = (F_1, \dots, F_N) \in \mathbb{H}^{n,m}$ *stabilizes* (A, B) *in the mean square sense if, when we make* $u(k) = F_{\theta(k)}x(k)$ *in system (1), we have that* $E\left(\|x(k)\|^2\right) \to 0$ *as* $k \to \infty$ *for any initial condition* $x(0)$ *and* $\theta(0)$. *We say that* (A, B) *is mean square stabilizable if for some* $F = (F_1, \dots, F_N) \in \mathbb{H}^{n,m}$, *we have that* F *stabilizes* (A, B) *in the mean square sense.*

The following result, proved in [4], shows that $F = (F_1, \dots, F_N)$ stabilizes system (1) in the mean square sense if and only if the expectral radius of the operator (12) (or (11)) in closed loop is less than one.

Lemma 1 $F = (F_1, \dots, F_N) \in \mathbb{H}^{n,m}$ *stabilizes* (A, B) *in the mean square sense if and only if* $r_\sigma(\mathcal{G}) < 1$, *where* $\mathcal{G}$ *is as in (12) with* $\Gamma_i = A_i + B_iF_i$.

The next lemma, proved in [6], will be crucial for the development of the paper. Set $\Gamma_i = A_i + B_iF_i$, $i = 1, \dots, N$, for some $F = (F_1, \dots, F_N) \in \mathbb{H}^{n,m}$. Consider also the operator $\bar{\mathcal{L}}(\cdot) = (\bar{\mathcal{L}}_1(\cdot), \dots, \bar{\mathcal{L}}_N(\cdot)) \in \mathbb{B}(\mathbb{H}^n)$ defined as

$$\bar{\mathcal{L}}_i(\cdot) := \Lambda_i'\mathcal{E}_i(\cdot)\Lambda_i, \qquad i = 1, \dots, N \tag{13}$$

where $\Lambda_i = A_i + B_iG_i$, $i = 1, \dots, N$, for some $G = (G_1, \dots, G_N) \in \mathbb{H}^{n,m}$.

Lemma 2 *Let* $\mathcal{L}$ *and* $\bar{\mathcal{L}}$ *be as defined in (11) and (13) above, with* $\Gamma_i = A_i + B_iF_i$ *and* $\Lambda_i = A_i + B_iG_i$, $i = 1, \dots, N$. *Suppose that* $r_\sigma(\mathcal{L}) < 1$ *and for some* $P = (P_1, \dots, P_N) \geq 0$ *and* $\delta > 0$,

$$P_i - \Lambda_i'\mathcal{E}_i(P)\Lambda_i \geq \delta(G_i - F_i)'(G_i - F_i), \qquad i = 1, \dots, N. \tag{14}$$

Then $r_\sigma(\bar{\mathcal{L}}) < 1$.

We make the following definitions:

Definition 2 *We define* $\mathcal{F}(\cdot) = (\mathcal{F}_1(\cdot), \dots, \mathcal{F}_N(\cdot)) : \mathbb{H}^{n+} \to \mathbb{H}^{n,m}$, $\mathcal{V}(\cdot) = (\mathcal{V}_1(\cdot), \dots, \mathcal{V}_N(\cdot)) : \mathbb{H}^{n,m} \to \mathbb{H}^n$ *and* $\mathcal{R} = (\mathcal{R}_1(\cdot), \dots, R_N(\cdot)) : \mathbb{H}^{n+} \to \mathbb{H}^{n+}$ *as*

$$\begin{aligned}
\mathcal{F}_i(X) &:= -\left(B_i'\mathcal{E}_i(X)B_i + D_i'D_i\right)^{-1} B_i'\mathcal{E}_i(X)A_i \\
\mathcal{V}_i(F) &:= C_i'C_i + F_i'D_i'D_iF_i \\
\mathcal{R}_i(X) &:= A_i'\mathcal{E}_i(X)A_i + C_i'C_i - A_i'\mathcal{E}_i(X)B_i\left(B_i'\mathcal{E}_i(X)B_i + D_i'D_i\right)^{-1} B_i'\mathcal{E}_i(X)A_i
\end{aligned} \tag{15}$$

where $X = (X_1, \dots, X_N) \in \mathbb{H}^{n+}$, $F = (F_1, \dots, F_N) \in \mathbb{H}^{n,m}$.

We have also the following result (see proof in [4]):

Lemma 3 *Suppose that* $F = (F_1, \dots, F_N) \in \mathbb{H}^{n,m}$ *stabilizes* (A, B) *in the mean square sense. Then there exists a unique solution* $S(F)$ *for the equation in* Y

$$Y_i - (A_i + B_iF_i)'\,\mathcal{E}_i(Y)\,(A_i + B_iF_i) - (C_i'C_i + F_i'D_i'D_iF_i) = 0, \quad i = 1, \dots, N. \tag{16}$$

Moreover,

$$S(F) = \sum_{k=0}^{\infty} \mathcal{L}^k\left(\mathcal{V}(F)\right). \tag{17}$$

The following identity will be useful in the sequel: for any $F = (F_1, \dots, F_N) \in \mathbb{H}^{n,m}$, we have that

$$
\begin{aligned}
&(A_i + B_i F_i)' \, \mathcal{E}_i(S) \, (A_i + B_i F_i) + F_i' D_i' D_i F_i \\
&= (A_i + B_i \mathcal{F}_i(S))' \, \mathcal{E}_i(S) \, (A_i + B_i \mathcal{F}_i(S)) + \mathcal{F}_i(S)' D_i' D_i \mathcal{F}_i(S) \\
&+ (F_i - \mathcal{F}_i(S))' \, (B_i' \mathcal{E}_i(S) B_i + D_i' D_i) \, (F_i - \mathcal{F}_i(S)) \, . \qquad (18)
\end{aligned}
$$

The next Lemma provides the existence of the maximal solution for (3) whenever (A, B) is mean square stabilizable. It is based on a quasi-linearization technique for the CARE (3), and parallels the policy iteration technique for MDP (see Remark 1 below).

Lemma 4 *Suppose that (A, B) is mean square stabilizable and considers $F^0 = \left(F_1^0, \dots, F_N^0\right) \in \mathbb{H}^{n,m}$ such that stabilizes (A, B) in the mean square sense. Then for $l = 0, 1, 2, \dots,$ there exists $X^l = \left(X_1^l, \dots, X_N^l\right)$ which satisfies the following properties:*

a) $X^0 \geq X^1 \geq \dots \geq X^l \geq X$, for arbitrary $X \in \mathbb{H}^{n+}$ such that $X \geq \mathcal{R}(X)$.

b) $r_\sigma(\mathcal{L}^l) < 1$, where $\mathcal{L}^l(\cdot) = (\mathcal{L}_1^l(\cdot), \dots, \mathcal{L}_N^l(\cdot))$ and for $i = 1, \dots, N$,

$$
\begin{aligned}
\mathcal{L}_i^l(\cdot) &:= A_i^{l\prime} \mathcal{E}_i(\cdot) A_i^l, \\
A_i^l &:= A_i + B_i F_i^l, \\
F_i^l &:= \mathcal{F}_i\left(X^{l-1}\right) \text{ for } l = 1, 2, \dots .
\end{aligned}
$$

c) X^l satisfies $X^l = \mathcal{L}^l\left(X^l\right) + \mathcal{V}\left(F^l\right)$ and is given by $X^l = \sum_{k=0}^{\infty} \left(\mathcal{L}^l\right)^k \left(\mathcal{V}\left(F^l\right)\right)$.

Moreover there exists $X^+ = \left(X_1^+, \dots, X_N^+\right) \in \mathbb{H}^{n+}$ such that $X^+ = \mathcal{R}\left(X^+\right)$, $X^+ \geq X$ for any $X \in \mathbb{H}^{n+}$ such that $X \geq \mathcal{R}(X)$, and $X^l \to X^+$ as $l \to \infty$. Furthermore $r_\sigma(\mathcal{L}^+) \leq 1$, where $\mathcal{L}^+(\cdot) = \left(\mathcal{L}_1^+(\cdot), \dots, \mathcal{L}_N^+(\cdot)\right)$ is defined as $\mathcal{L}_i^+(\cdot) = A_i^{+\prime} \mathcal{E}_i(\cdot) A_i^+$, for $i = 1, \dots, N$, and

$$
\begin{aligned}
F_i^+ &= \mathcal{F}_i\left(X^+\right) \\
A_i^+ &= A_i + B_i F_i^+ .
\end{aligned}
$$

Remark 1 *Step c) in Lemma 4, which corresponds the calculation of the solution of the linear system*

$$
X^l = \mathcal{L}^l\left(X^l\right) + \mathcal{V}\left(F^l\right),
$$

can be seen as the policy evaluation step in the policy iteration technique for MDP, while from identity (18),

$$
\begin{aligned}
&(A_i + B_i F_i)' \mathcal{E}_i(X^l)(A_i + B_i F_i) + C_i' C_i + F_i' D_i' D_i F_i = \\
&\quad X_i^l - \left(F_i^l - \mathcal{F}_i\left(X^l\right)\right)' \left(B_i' \mathcal{E}_i(X^l) B_i + D_i' D_i\right) \left(F_i^l - \mathcal{F}_i\left(X^l\right)\right) \\
&\quad + \left(F_i - \mathcal{F}_i\left(X^l\right)\right)' \left(B_i' \mathcal{E}_i(X^l) B_i + D_i' D_i\right) \left(F_i - \mathcal{F}_i\left(X^l\right)\right)
\end{aligned}
$$

and the right hand side is minimized in F_i by choosing $F_i = F_i^{l+1} = \mathcal{F}_i\left(X^l\right)$, which can be seen as the policy improvement step.

3 Temporal Difference Methods

The policy evaluation step of the policy iteration method presented in Lemma 4 could be performed by value iteration since, as seen in c) of Lemma 4,

$$X^l = \sum_{k=0}^{\infty} \left(\mathcal{L}^l\right)^k \left(\mathcal{V}(F^l)\right)$$

and from the fact that $r_\sigma\left(\mathcal{L}^l\right) < 1$, the above summation could be truncated for a sufficiently large integer κ, and

$$X^l \simeq \sum_{k=0}^{\kappa} \left(\mathcal{L}^l\right)^k \left(\mathcal{V}(F^l)\right).$$

However if $r_\sigma\left(\mathcal{L}^l\right)$ is close to one, the convergence above could be too slow, and thus we are motivated to use an approach in which the spectral radius of the associated operator is reduced in order to accelerate the policy iteration method. We present below this policy-iteration like algorithm, which parallels the $TD(\lambda)$ methods in Neuro-dynamic programming (see, for instance, [2],[15]).

Consider $\lambda \in (0,1)$. Suppose that at iteration ℓ, we are given $\left(S^{\ell-1}, F^\ell\right)$ such that $\sqrt{\lambda}F^\ell$ stabilizes $\sqrt{\lambda}(A,B)$ in the mean square sense, that is, $r_\sigma(\lambda\mathcal{L}^\ell) < 1$, where $\mathcal{L}^\ell(\cdot) = \left(\mathcal{L}_1^\ell(\cdot), \ldots, \mathcal{L}_N^\ell(\cdot)\right)$ and

$$\mathcal{L}_i^\ell(\cdot) := \left(A_i + B_i F^\ell\right)' \mathcal{E}_i(\cdot) \left(A_i + B_i F^\ell\right).$$

For $k = 0, 1, \ldots$, we define the TD associated to a transition from $\theta(k)$ to $\theta(k+1)$

$$\begin{aligned}
\Gamma^\ell(k) &= \Phi^\ell(k-1)' \Big(C'_{\theta(k)} C_{\theta(k)} + F^{\ell'}_{\theta(k)} D'_{\theta(k)} D_{\theta(k)} F^\ell_{\theta(k)} \\
&\quad + \left(A_{\theta(k)} + B_{\theta(k)} F^\ell_{\theta(k)}\right)' S^{\ell-1}_{\theta(k+1)} \left(A_{\theta(k)} + B_{\theta(k)} F^\ell_{\theta(k)}\right) - S^{\ell-1}_{\theta(k)} \Big) \Phi^\ell(k-1) \\
\Phi^\ell(k) &= \left(A_{\theta(k)} + B_{\theta(k)} F^\ell_{\theta(k)}\right) \Phi^\ell(k-1) \\
\Phi^\ell(-1) &= I.
\end{aligned}$$

We can view $\Gamma^\ell(k)$ as the one stage cost to go matrix at time k, for a discrete-time Markovian jump system (1). The cost-to-go matrix with discount factor λ is given by

$$\Delta_i^\ell = \sum_{k=0}^{\infty} E\left(\lambda^k \Gamma^\ell(k) | \theta(0) = i\right). \tag{19}$$

The next cost matrix is then given by

$$S_i^\ell = S_i^{\ell-1} + \Delta_i^\ell \tag{20}$$

and we shall refer to the above method as λ-policy iteration method. Note that

$$E\left(\Gamma^\ell(k) | \theta(0) = i\right) = \left(\mathcal{L}_i^\ell\right)^k \left(\mathcal{V}\left(F^\ell\right)\right) + \left(\mathcal{L}_i^\ell\right)^{k+1} \left(S^{\ell-1}\right) - \left(\mathcal{L}_i^\ell\right)^k \left(S^{\ell-1}\right) \tag{21}$$

and since by hypothesis $r_\sigma(\lambda\mathcal{L}^\ell) < 1$, there exists $a \geq 0$, $0 < b < 1$ such that

$$\left\| \left(\lambda\mathcal{L}^\ell\right)^k \right\| \leq ab^k \tag{22}$$

and it follows from (21) and (22) that the summation in (19) is well defined.

Remark 2 *For the special case in which $\lambda = 1$, we have that the above method coincides with the quasi-linearization method presented in Lemma 4, which could be seen also as the policy iteration technique known for the MDP. Indeed, from (19) and (21), and recalling that by assumption $F^\ell = \left(F_1^\ell, \ldots, F_N^\ell\right)$ stabilizes (A, B) in the mean square sense, we have that*

$$\begin{aligned}\Delta_i^\ell &= \sum_{k=0}^{\infty} \left(\left(\mathcal{L}_i^\ell\right)^k (\mathcal{V}(F^\ell)) + (\mathcal{L}_i^\ell)^{k+1}(S^{\ell-1}) - (\mathcal{L}_i^\ell)^k(S^{\ell-1}) \right) \\ &= S_i\left(F^\ell\right) - S_i^{\ell-1}.\end{aligned}$$

Therefore,

$$S_i^\ell = S_i^{\ell-1} + \Delta_i^\ell = S_i\left(F^\ell\right).$$

Remark 3 *For the special case in which $\lambda = 0$, we have that the above method coincides with the iterations of the coupled discrete-time Riccati difference equations in equations (8) and (9). Indeed, in this case,*

$$\begin{aligned}\Delta_i^\ell &= E\left(\Gamma^\ell(0) \,|\theta(0) = i\right) \\ &= C_i'C_i + F_i^{\ell'} D_i'D_iF_i^\ell + \left(A_i + B_iF_i^\ell\right)' \mathcal{E}_i\left(S^{\ell-1}\right)\left(A_i + B_iF_i^\ell\right) - S_i^{\ell-1}\end{aligned}$$

and thus

$$S_i^\ell = C_i'C_i + F_i^{\ell'} D_i'D_iF_i^\ell + \left(A_i + B_iF_i^\ell\right)' \mathcal{E}_i\left(S^{\ell-1}\right)\left(A_i + B_iF_i^\ell\right)$$

which coincides with (8).

Define $\mathcal{M}^\ell(S) = \left(\mathcal{M}_1^\ell(S), \ldots, \mathcal{M}_N^\ell\right)(S)$ as

$$\begin{aligned}\mathcal{M}_i^\ell(S) &= \lambda\left(A_i + B_iF_i^\ell\right)'\mathcal{E}_i(S)\left(A_i + B_iF_i{}^\ell\right) + (1-\lambda)\left(A_i + B_iF_i{}^\ell\right)'\mathcal{E}_i\left(S^{\ell-1}\right)\left(A_i + B_iF_i^\ell\right) \\ &\quad + C_i'C_i + F_i^{\ell'} D_i'D_iF_i^\ell \\ &= \lambda\mathcal{T}_i^\ell(S) + (1-\lambda)\mathcal{T}_i{}^\ell\left(S^{\ell-1}\right)\end{aligned}$$

where $\mathcal{T}^\ell(S) = \left(\mathcal{T}_1^\ell(S), \ldots, \mathcal{T}_N^\ell(S)\right)$ is defined as

$$\begin{aligned}\mathcal{T}_i^\ell(S) &= \left(A_i + B_iF_i^\ell\right)'\mathcal{E}_i(S)\left(A_i + B_iF_i^\ell\right) + C_i'C_i + F_i^{\ell'} D_i'D_iF_i{}^\ell \\ &= \mathcal{L}_i^\ell(S) + \mathcal{V}\left(F^\ell\right).\end{aligned}$$

Lemma 5 *For $k = 0, 1, \ldots$*

$$\begin{aligned}\left(\mathcal{M}^\ell\right)^k(S) &= (1-\lambda)\left[\mathcal{T}^\ell\left(S^{\ell-1}\right) + \lambda\left(\mathcal{T}^\ell\right)^2\left(S^{\ell-1}\right) + \cdots + \lambda^{k-1}\left(\mathcal{T}^\ell\right)^k\left(S^{\ell-1}\right)\right] \\ &\quad + \lambda^k\left(\mathcal{T}^\ell\right)^k(S).\end{aligned}$$

Proof: For simplicity, we shall drop the superscript ℓ on $\mathcal{M}^\ell$, $\mathcal{T}^\ell$, $\mathcal{L}^\ell$ and F^ℓ. We have that

$$\mathcal{M}(S) = \lambda \mathcal{T}(S) + (1-\lambda)\mathcal{T}\left(S^{\ell-1}\right).$$

Let us apply induction on k to show that

$$\mathcal{M}^k(S) = (1-\lambda)\left[\mathcal{T}\left(S^{\ell-1}\right) + \lambda\mathcal{T}^2\left(S^{\ell-1}\right) + \cdots + \lambda^{k-1}\mathcal{T}^k\left(S^{\ell-1}\right)\right] + \lambda^k\mathcal{T}^k(S).$$

For $k=1$, the result holds by definition. Suppose it holds for k. Then

$$\begin{aligned}
\mathcal{M}^{k+1}(S) &= \mathcal{M}\left(\mathcal{M}^k(S)\right) = \lambda\mathcal{T}\left(\mathcal{M}^k(S)\right) + (1-\lambda)\mathcal{T}\left(S^{\ell-1}\right) \\
&= \lambda\mathcal{T}\left((1-\lambda)\left[\mathcal{T}^\ell\left(S^{\ell-1}\right) + \lambda\left(\mathcal{T}^\ell\right)^2\left(S^{\ell-1}\right) + \ldots + \lambda^{k-1}\left(\mathcal{T}^\ell\right)^k\left(S^{\ell-1}\right)\right] + \lambda^k\mathcal{T}^k(S)\right) \\
&+ (1-\lambda)\mathcal{T}\left(S^{\ell-1}\right) \\
&= \lambda\mathcal{L}\left((1-\lambda)\left[\mathcal{T}^\ell\left(S^{\ell-1}\right) + \cdots + \lambda^{k-1}\left(\mathcal{T}^\ell\right)^k\left(S^{\ell-1}\right)\right] + \lambda^k\mathcal{T}^k(S)\right) + \lambda\mathcal{V}(F) \\
&+ (1-\lambda)\mathcal{T}\left(S^{\ell-1}\right) \\
&= \lambda(1-\lambda)\mathcal{L}\left(\mathcal{T}^\ell\left(S^{\ell-1}\right)\right) + \cdots + \lambda(1-\lambda)\lambda^{k-1}\mathcal{L}\left(\mathcal{T}^\ell\right)^k\left(S^{\ell-1}\right) + \lambda^{k-1}\mathcal{L}\left(\mathcal{T}^k(S)\right) \\
&+ \lambda\mathcal{V}(F) + (1-\lambda)\mathcal{T}\left(S^{\ell-1}\right).
\end{aligned}$$

Note now that

$$\begin{aligned}
&\lambda(1-\lambda) + \lambda^2(1-\lambda) + \cdots + \lambda^k(1-\lambda) + \lambda^{k+1} = \\
&\lambda\left((1-\lambda) + \lambda(1-\lambda) + \cdots + \lambda^{k-1}(1-\lambda) + \lambda^k\right) = \\
&\qquad \lambda\left(1-\lambda\left(1+\cdots+\lambda^{k-1}\right) + \lambda^k\right) = \lambda\left((1-\lambda)\frac{(1-\lambda^k)}{(1-\lambda)} + \lambda^k\right) = \lambda
\end{aligned}$$

and thus

$$\begin{aligned}
\mathcal{M}^{k+1}(S) &= \lambda(1-\lambda)\left(\mathcal{L}\left(\mathcal{T}\left(S^{\ell-1}\right)\right) + \mathcal{V}(F)\right) + \cdots + \lambda^k(1-\lambda)\left(\mathcal{L}\left(\mathcal{T}^k\left(S^{\ell-1}\right)\right) + \mathcal{V}(F)\right) \\
&+ \lambda^{k+1}\left(\mathcal{L}\left(\mathcal{T}^k(S)\right) + \mathcal{V}(F)\right) + (1-\lambda)\mathcal{T}\left(S^{\ell-1}\right) \\
&= (1-\lambda)\mathcal{T}\left(S^{\ell-1}\right) + \lambda(1-\lambda)\mathcal{T}^2\left(S^{\ell-1}\right) + \cdots + \lambda^k(1-\lambda)\mathcal{T}^{k+1}\left(S^{\ell-1}\right) \\
&+ \lambda^{k+1}\mathcal{T}^{k+1}(S),
\end{aligned}$$

completing the proof.

Lemma 6 *Suppose that $\sqrt{\lambda}F^\ell$ stabilizes $\sqrt{\lambda}(A,B)$ in the mean square sense. Then*

i) There exists a unique R satisfying

$$R = \mathcal{M}^\ell(R).$$

ii) $R = S^\ell$ where S^ℓ is as in eq. (20).

iii) For $0 \le \lambda < 1$,

$$S^\ell = (1-\lambda)\sum_{k=0}^{\infty}\lambda^k\left(\mathcal{T}^\ell\right)^{k+1}\left(S^{\ell-1}\right).$$

Proof:

i) *Since* $r_\sigma(\lambda\mathcal{L}^\ell) < 1$, *we have that there exists a unique* R *such that*

$$R = \lambda\mathcal{L}^\ell(R) + (1-\lambda)\mathcal{L}^\ell\left(S^{\ell-1}\right) + \mathcal{V}\left(F^\ell\right)$$

which is given by (see eq. (17) in Lemma 3)

$$R = \sum_{k=0}^{\infty} \left(\lambda\mathcal{L}^\ell\right)^k \left((1-\lambda)\mathcal{L}^\ell\left(S^{\ell-1}\right) + \mathcal{V}\left(F^\ell\right)\right).$$

ii) *From i), (19), and (21),*

$$\begin{aligned} R &= \sum_{k=0}^{\infty} \lambda^k \left(\mathcal{L}^\ell\right)^{k+1}\left(S^{\ell-1}\right) - \sum_{k=1}^{\infty} \lambda^k \left(\mathcal{L}^\ell\right)^{k}\left(S^{\ell-1}\right) + S^{\ell-1} + \sum_{k=0}^{\infty} \lambda^k \left(\mathcal{L}^\ell\right)^{k}\left(\mathcal{V}\left(F^\ell\right)\right) \\ &= \sum_{k=0}^{\infty} \lambda^k \left(\mathcal{L}^\ell\right)^{k+1}\left(S^{\ell-1}\right) - \sum_{k=1}^{\infty} \lambda^k \left(\mathcal{L}^\ell\right)^{k}\left(S^{\ell-1}\right) + S^{\ell-1} + \sum_{k=0}^{\infty} \lambda^k \left(\mathcal{L}^\ell\right)^{k}\left(\mathcal{V}\left(F^\ell\right)\right) \\ &= \sum_{k=0}^{\infty} \lambda^k \left(\left(\mathcal{L}^\ell\right)^{k}\left(\mathcal{V}\left(F^\ell\right)\right) + \left(\mathcal{L}^\ell\right)^{k+1}\left(S^{\ell-1}\right) - \left(\mathcal{L}^\ell\right)^{k}\left(S^{\ell-1}\right)\right) + S^{\ell-1} \\ &= \Delta^\ell + S^{\ell-1} = S^l \end{aligned}$$

iii) *We have that*

$$\mathcal{T}^\ell(S) = \mathcal{L}^\ell(S) + \mathcal{V}\left(F^\ell\right)$$

and thus

$$\left(\lambda\mathcal{T}^\ell\right)^k(S) = \left(\lambda\mathcal{L}^\ell\right)^k(S) + \sum_{r=0}^{k-1}\left(\lambda\mathcal{L}^\ell\right)^{k-r-1}\left(\lambda^{r+1}\mathcal{V}\left(F^\ell\right)\right).$$

Thus, since $r_\sigma\left(\lambda\mathcal{L}^\ell\right) < 1$, *we have for some constant* $a \geq 0$, $\lambda < b < 1$ *(see (22)) that*

$$\left\|\left(\lambda\mathcal{L}^\ell\right)^k\right\| \leq ab^k.$$

Thus, recalling the definition in (10), we have

$$\begin{aligned} \left\|\left(\lambda\mathcal{T}^\ell\right)^k(S)\right\| &\leq \left\|\left(\lambda\mathcal{L}^\ell\right)^k\right\|\|S\| + \left(\sum_{r=0}^{k-1}\left\|\left(\lambda\mathcal{L}^\ell\right)^{k-r-1}\right\|\lambda^{r+1}\right)\|\mathcal{V}(F^\ell)\| \\ &\leq ab^k\|S\| + ab^k\left\|\mathcal{V}\left(F^\ell\right)\right\|\left(\sum_{r=0}^{k-1}\left(\frac{\lambda}{b}\right)^{r+1}\right) \\ &\leq ab^k\|S\| + ab^k\left\|\mathcal{V}\left(F^\ell\right)\right\|\left(\frac{\lambda}{b}\right)\frac{1}{1-\left(\frac{\lambda}{b}\right)} \end{aligned}$$

and taking the limit as k *goes to infinity we conclude that*

$$\lambda^k\left(\mathcal{T}^\ell\right)^k(S) \to 0 \text{ as } k \to \infty,$$

and from Lemma 5,

$$\left(\mathcal{M}^{\ell}\right)^{r}(S)\uparrow(1-\lambda)\sum_{k=0}^{\infty}\lambda^{k}\left(\mathcal{T}^{\ell}\right)^{k+1}\left(S^{\ell-1}\right)\ \ \textit{as } r\rightarrow\infty,$$

for any S. On the other hand,

$$S^{\ell}=\mathcal{M}^{\ell}\left(S^{\ell}\right)\ \ \textit{and thus } S^{\ell}=\left(\mathcal{M}^{\ell}\right)^{r}\left(S^{\ell}\right)$$

which shows that

$$S^{\ell}=(1-\lambda)\sum_{k=0}^{\infty}\lambda^{k}\left(\mathcal{T}^{\ell}\right)^{k+1}\left(S^{\ell-1}\right).$$

¿From Lemma 6,

$$S^{\ell}=\lambda\mathcal{L}^{\ell}\left(S^{\ell}\right)+(1-\lambda)\mathcal{L}^{\ell}\left(S^{\ell-1}\right)+\mathcal{V}\left(F^{\ell}\right). \tag{23}$$

Consider $F^{0}=\left(F_{1}^{0},\ldots,F_{N}^{0}\right)\in\mathbb{H}^{n,m}$, $0\leq\lambda<1$, and $S^{-1}\in\mathbb{H}^{n+}$, $V\in\mathbb{H}^{n+}$ such that

A1) $r_{\sigma}\left(\lambda\mathcal{L}^{0}\right)<1$
A2) $V\geq P$ where P is the maximal solution of CARE (3)
A3) $\lambda\mathcal{T}^{0}\left(S^{-1}\right)+(1-\lambda)\mathcal{T}^{0}(V)\leq S^{-1}$.

Remark 4 *Suppose that F^{0} stabilizes (A,B) in the mean square sense. Then from Lemma 3, there exists a unique solution $Y\geq0$ satisfying $Y=\mathcal{T}^{0}(Y)$ and $Y\geq P$. Indeed, from (18), for any $F=(F_{1},\ldots,F_{N})\in\mathbb{H}^{n,m}$, we have*

$$\begin{aligned}P_{i}&=(A_{i}+B_{i}F_{i})'\,\mathcal{E}_{i}(P)\,(A_{i}+B_{i}F_{i})+F_{i}'D_{i}'D_{i}F_{i}+C_{i}'C_{i}\\&\quad-(F_{i}-\mathcal{F}_{i}(P))'\,(B_{i}'\mathcal{E}_{i}(P)B_{i}+D_{i}'D_{i})\,(F_{i}-\mathcal{F}_{i}(P))\end{aligned} \tag{24}$$

and since

$$Y_{i}=\left(A_{i}+B_{i}F_{i}^{0}\right)'\mathcal{E}_{i}(P)\left(A_{i}+B_{i}F_{i}^{0}\right)+F_{i}^{0'}D_{i}'D_{i}F_{i}^{0}+C_{i}'C_{i}$$

it follows from (24) with $F=F^{0}$ that

$$(Y_{i}-P_{i})-\left(A_{i}+B_{i}F_{i}^{0}\right)'\mathcal{E}_{i}(Y-P)\left(A_{i}+B_{i}F_{i}^{0}\right)\geq0$$

and since $r_{\sigma}(\mathcal{L}^{0})<1$ we have from Lemma 3 that $Y\geq P$. Therefore for this F^{0} and whit $S^{-1}=Y$, $V=Y$, A1),A2),A3) will be satisfied for any $\lambda=[0,1)$.

We define the λ-policy iteration algorithm like for $\ell=0,1,\ldots$ in the following way:

i) $F_{i}^{\ell}=\mathcal{F}_{i}\left(S^{\ell-1}\right)$ (except for $\ell=0$).

ii) S^{ℓ} the solution of equation (20).

Theorem 1 *Suppose that (A,B) is mean square stabilizable and A1), A2), A3) are satisfied. Then for each $\ell = 0, 1, \ldots,$ we have that $\sqrt{\lambda}F^{\ell}$ stabilizes $\sqrt{\lambda}(A,B)$ in the mean square sense, and*

$$S^{\ell} \downarrow P \text{ as } \ell \uparrow \infty$$

where P is the maximal solution for the CARE given by (3).

Proof: First of all, notice from A3) that

$$\mathcal{M}^0\left(S^{-1}\right) = \lambda \mathcal{T}^0\left(S^{-1}\right) + (1-\lambda)\mathcal{T}^0(V) \leq S^{-1}$$

and thus,

$$\left(\mathcal{M}^0\right)^k\left(S^{-1}\right) \leq \left(\mathcal{M}^0\right)^{k-1}\left(S^{-1}\right) \leq S^{-1}.$$

Taking the limit as $k \to \infty$, we get from A1), and Lemmas 5, and 6, that

$$S^0 = \lim_{k\to\infty} \left(\mathcal{M}^0\right)^k\left(S^{-1}\right) \leq S^{-1}.$$

Moreover,

$$\begin{aligned} S_i^0 = &\Big(\lambda\left(A_i + B_iF_i^1\right)'\mathcal{E}_i\left(S^0\right)\left(A_i + B_iF_i^1\right) + C_i'C_i + F_i^{1'}D_i'D_iF_i^1 \\ &+ \left(F_i^1 - F_i^0\right)'\left(B_i'E_i\left(S^0\right)B_i + D_i'D_i\right)\left(A_i + B_iF_i^0\right)\left(F_i^1 - F_i^0\right)\Big) \\ &+ (1-\lambda)\Big(\left(A_i + B_iF_i^0\right)'\mathcal{E}_i\left(S^{-1}\right)\left(A_i + B_iF_i^0\right) + C_i'C_i + F_i^{0'}D_i'D_iF_i^0\Big) \end{aligned}$$

and thus, for some $\delta^0 > 0$

$$S_i^0 - \lambda\left(A_i + B_iF_i^1\right)'\mathcal{E}_i\left(S^0\right)\left(A_i + B_iF_i^1\right) \geqslant \delta^0\lambda\left(F_i^1 - F_i^0\right)'\left(F_i^1 - F_i^0\right)$$

showing, according to equation (14) in Lemma 2, that $\sqrt{\lambda}F^1$ stabilizes $\sqrt{\lambda}(A,B)$ in the mean square sense. By making $F = F^0$ in (24) and recalling from A2), A3) that

$$\lambda\mathcal{T}^0\left(S^{-1}\right) + (1-\lambda)\mathcal{T}^0(P) \leq \lambda\mathcal{T}^0\left(S^{-1}\right) + (1-\lambda)\mathcal{T}^0(V) \leq S^{-1}$$

we have,

$$\begin{aligned} &\left(S_i^{-1} - P_i\right) - \lambda\left(A_i + B_iF_i^0\right)'\mathcal{E}_i\left(S^{-1} - P\right)\left(A_i + B_iF_i^0\right) \geq \\ &\quad \left(F_i^0 - \mathcal{F}_i\left(P\right)\right)'\left(B_i'\mathcal{E}_i\left(P\right)B_i + D_i'D_i\right)\left(F_i^0 - \mathcal{F}_i(P)\right) \geq 0 \end{aligned}$$

and since from A1), $r_\sigma(\lambda\mathcal{L}^0) < 1$, we have from Lemma 3, equation (17), that $S^{-1} \geq P$. This implies that

$$\begin{aligned} S_i^0 \geq &\lambda\left(A_i + B_iF_i^0\right)'\mathcal{E}_i\left(S^0\right)\left(A_i + B_iF_i^0\right) \\ &+ (1-\lambda)\left(A_i + B_iF_i^0\right)'\mathcal{E}_i(P)\left(A_i + B_iF_i^0\right) \\ &+ C_i'C_i + F_i^{0'}D_i'D_iF_i^0 \end{aligned}$$

and from equation (24) again, we have,

$$\begin{aligned} &\left(S_i^0 - P_i\right) - \lambda\left(A_i + B_iF_i^0\right)'\mathcal{E}_i\left(S^0 - P\right)\left(A_i + B_iF_i^0\right) \geq \\ &\quad \left(F_i^0 - \mathcal{F}_i(P)\right)'\left(B_i'\mathcal{E}_i(P)B_i + D_i'D_i\right)\left(F_i^0 - \mathcal{F}_i(P)\right) \geq 0 \end{aligned}$$

which implies that $S^0 \geq P$. *Suppose now that*

$$\begin{aligned}&P \leq S^{\ell} \leq S^{\ell-1}\\ &\sqrt{\lambda}F^{\ell} \text{ stabilizes } \sqrt{\lambda}(A,B) \text{ in the mean square sense.}\end{aligned}$$

We want to show that

$$\begin{aligned}&P \leq S^{\ell+1} \leq S^{\ell}\\ &\sqrt{\lambda}F^{\ell+1} \text{ stabilizes } \sqrt{\lambda}(A,B) \text{ in the mean square sense.}\end{aligned}$$

Indeed, from equation (18),

$$\begin{aligned}S_i^{\ell} = {}& \lambda\Big(\left(A_i + B_iF_i^{\ell+1}\right)' \mathcal{E}_i\left(S^{\ell}\right)\left(A_i + B_iF_i^{\ell+1}\right) + C_i'C_i + F_i^{\ell+1'}D_i'D_iF_i^{\ell+1}\\ &+ \left(F_i^{\ell+1} - F_i^{\ell}\right)'\left(B_i'\mathcal{E}_i\left(S^{\ell}\right)B_i + D_i'D_i\right)\left(F_i^{\ell+1} - F_i^{\ell}\right)\Big)\\ &+ (1-\lambda)\left(\left(A_i + B_iF_i^{\ell}\right)'\mathcal{E}_i\left(S^{\ell-1}\right)\left(A_i + B_iF_i^{\ell}\right) + C_i'C_i + F_i^{\ell'}D_i'D_iF_i^{\ell}\right)\end{aligned}$$

and thus, for some $\delta^{\ell} > 0$

$$S_i^{\ell} - \lambda\left(A_i + B_iF_i^{\ell+1}\right)'\mathcal{E}_i\left(S^{\ell}\right)\left(A_i + B_iF_i^{\ell+1}\right) \geq \delta^{\ell}\lambda\left(F_i^{\ell+1} - F_i^{\ell}\right)'\left(F_i^{\ell+1} - F_i^{\ell}\right)$$

which shows, from Lemma 2, that $\sqrt{\lambda}F^{\ell+1}$ *stabilizes* $\sqrt{\lambda}(A,B)$ *in the mean square sense. Notice that from the identity in (18) we have*

$$\begin{aligned}S_i^{\ell} = {}& \lambda\left(A_i + B_iF_i^{\ell}\right)'\mathcal{E}_i\left(S^{\ell}\right)\left(A_i + B_iF_i^{\ell}\right) + (1-\lambda)\left(A_i + B_iF_i^{\ell}\right)'\mathcal{E}_i\left(S^{\ell-1}\right)\left(A_i + B_iF_i^{\ell}\right)\\ &+ C_i'C_i + F_i^{\ell'}D_i'D_iF_i^{\ell}\\ &\geqslant \lambda\left(A_i + B_iF_i^{\ell}\right)'\mathcal{E}_i\left(S^{\ell}\right)\left(A_i + B_iF_i^{\ell}\right) + (1-\lambda)\left(A_i + B_iF_i^{\ell}\right)'\mathcal{E}_i\left(S^{\ell}\right)\left(A_i + B_iF_i^{\ell}\right)\\ &+ C_i'C_i + F_i^{\ell'}D_i'D_iF_i^{\ell}\\ &= \left(A_i + B_iF_i^{\ell}\right)'\mathcal{E}_i\left(S^{\ell}\right)\left(A_i + B_iF_i^{\ell}\right) + C_i'C_i + F_i^{\ell'}D_i'D_iF_i^{\ell}\\ &= \left(A_i + B_iF_i^{\ell+1}\right)'\mathcal{E}_i\left(S^{\ell}\right)\left(A_i + B_iF_i^{\ell+1}\right) + C_i'C_i + F_i^{\ell+1'}D_i'D_iF_i^{\ell+1}\\ &+ \left(F_i^{\ell+1} - F_i^{\ell}\right)'\left(B_i'\mathcal{E}_i\left(S^{\ell}\right)B_i + D_i'D_i\right)\left(F_i^{\ell+1} - F_i^{\ell}\right)\end{aligned}$$

and

$$\begin{aligned}S_i^{\ell+1} = {}& \lambda\left(A_i + B_iF_i^{\ell+1}\right)'\mathcal{E}_i\left(S^{\ell+1}\right)\left(A_i + B_iF_i^{\ell+1}\right)\\ &+ (1-\lambda)\left(A_i + B_iF_i^{\ell+1}\right)'\mathcal{E}_i\left(S^{\ell}\right)\left(A_i + B_iF_i^{\ell+1}\right)\\ &+ C_i'C_i + F_i^{\ell+1'}D_i'D_iF_i^{\ell+1}\end{aligned}$$

and therefore,

$$\begin{aligned}\left(S_i^{\ell} - S_i^{\ell+1}\right) - \lambda\left(A_i + B_iF_i^{\ell+1}\right)'\mathcal{E}_i\left(S^{\ell} - S^{\ell+1}\right)\left(A_i + B_iF_i^{\ell+1}\right) \geq&\\ \left(F_i^{\ell+1} - F_i^{\ell}\right)'\left(B_i'\mathcal{E}_i\left(S^{\ell}\right)B_i + D_i'D_i\right)\left(F_i^{\ell+1} - F_i^{\ell}\right) \geq 0.&\end{aligned}$$

Thus, since $\sqrt{\lambda}F^{\ell+1}$ *stabilizes* $\sqrt{\lambda}(A,B)$ *in the mean square sense, we get that*

$$S^{\ell} - S^{\ell+1} \geq 0.$$

Remains to show that $S^{\ell+1} \geq P$. We have that

$$\begin{aligned} S_i^{\ell+1} &\geq \lambda \left(A_i + B_i F_i^{\ell+1}\right)' \mathcal{E}_i \left(S^{\ell+1}\right) \left(A_i + B_i F_i^{\ell+1}\right) \\ &+ (1-\lambda) \left(A_i + B_i F_i^{\ell+1}\right)' \mathcal{E}_i(P) \left(A_i + B_i F_i^{\ell+1}\right) \\ &+ C_i' C_i + F_i^{\ell+1\prime} D_i' D_i F_i^{\ell+1} \end{aligned} \tag{25}$$

and from (24) and (25) again we have that

$$\begin{aligned} &\left(S_i^{\ell+1} - P_i\right) - \lambda \left(A_i + B_i F_i^{\ell+1}\right)' \mathcal{E}_i \left(S^{\ell+1} - P\right) \left(A_i + B_i F_i^{\ell+1}\right) \geq \\ &\qquad \left(F_i^{\ell+1} - \mathcal{F}_i(P)\right)' \left(B_i' \mathcal{E}_i(P) B_i + D_i' D_i\right) \left(F_i^{\ell+1} - \mathcal{F}_i(P)\right) \geq 0 \end{aligned}$$

and since $r_\sigma\left(\lambda \mathcal{L}^{\ell+1}\right) < 1$, we have from Lemma 3, equation (17), that indeed $S^{\ell+1} \geq P$. Thus we have that $S^\ell \downarrow S$ as $\ell \uparrow \infty$ for some $S \in \mathbb{H}^{n+}$, and $S \geq P$. Moreover,

$$\begin{aligned} S_i &= \lambda \left(A_i + B_i \mathcal{F}_i(S)\right)' \mathcal{E}_i(S) \left(A_i + B_i \mathcal{F}_i(S)\right) + (1-\lambda) \left(A_i + B_i \mathcal{F}_i(S)\right)' \mathcal{E}_i(S) \left(A_i + B_i \mathcal{F}_i(S)\right) \\ &\quad + C_i' C_i + \mathcal{F}_i(S)' D_i' D_i \mathcal{F}_i(S) \\ &= \left(A_i + B_i \mathcal{F}_i(S)\right)' \mathcal{E}_i(S) \left(A_i + B_i \mathcal{F}_i(S)\right) + C_i' C_i + \mathcal{F}_i(S)' D_i' D_i \mathcal{F}_i(S) \end{aligned}$$

that is, S is a solution for the CARE. Since P is maximal (Lemma 4), $P \geq S$, and therefore, $P = S$.

4 Numerical Example

To illustrate the use of the results developed in the previous sections, we have chosen a simple economic system based on Samuelson's multiplier-accelerator model [12] which appears in state equation form:

$$x(k+1) = A_{\theta(k)} x(k) + B_{\theta(k)} u(k).$$

We considered the following three-operating mode discrete time jump linear:

Table 1: Dates for the multiplier-accelerator model

	Operating Modes		
Parameters	1	2	3
matrix A	$A_1 = \begin{bmatrix} 0 & 1 \\ -2.5 & 3.2 \end{bmatrix}$	$A_2 = \begin{bmatrix} 0 & 1 \\ -4.3 & 4.5 \end{bmatrix}$	$A_3 = \begin{bmatrix} 0 & 1 \\ 5.3 & -5.2 \end{bmatrix}$
matrix B	$B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$B_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$B_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$
matrix $Q_i = C_i' C_i$	$Q_1 = \begin{bmatrix} 3.6 & -3.8 \\ -3.8 & 4.87 \end{bmatrix}$	$Q_2 = \begin{bmatrix} 10 & -3 \\ -3 & 8 \end{bmatrix}$	$Q_3 = \begin{bmatrix} 5 & -4.5 \\ -4.5 & 4.5 \end{bmatrix}$
matrix $R_i = D_i' D_i$	$R_1 = 2.6$	$R_2 = 1.165$	$R_3 = 1.111$

The discrete time state transition probability matrix is assumed to be as

$$P = \begin{bmatrix} 0.67 & 0.17 & 0.16 \\ 0.30 & 0.47 & 0.23 \\ 0.26 & 0.10 & 0.64 \end{bmatrix}.$$

Therefore we must have $F^0 = \left(F_1^0, F_2^0, F_3^0\right) \in \mathbb{H}^{2,1}$, $0 \leq \lambda < 1$, and $S^{-1} \in \mathbb{H}^{2+}$, $V \in \mathbb{H}^{2+}$. If we choose F^0 such that the $r_\sigma(\mathcal{L}^0) < 1$, $S^{-1} = \mathcal{T}^0(S^{-1})$ and $V = S^{-1}$ then, as stated in Remark 4, the conditions $A1), A2)$ and $A3)$ are satisfied for any value of $0 \leq \lambda < 1$, and therefore according to Theorem 1, the algorithm will converge to the maximal solution P.

We have considered three values of λ; $\lambda = 0.1$, $\lambda = 0.5$ and $\lambda = 0.9$. For each of these values we have solved S^{ℓ} by value iteration, and we have taken the mean number of iterations, denoted by ς, needed until convergence of S^{ℓ}. We have also denoted by ℓ_f the number of iterations needed until convergence of S^{ℓ} to P. For $\lambda = 0.1$ we obtained $\varsigma = 3.375$ and $\ell_f = 16$; for $\lambda = 0.5$ we got $\varsigma = 5.08$ and $\ell_f = 12$; and for $\lambda = 0.9$ we obtained $\varsigma = 6.5$ and $\ell_f = 8$. As expected, as λ increases, so does ς and ℓ_f decreases. In all the three cases, the optimal solution was:

Table 2: Optimal solution

	Solution of CARE	Optimal Control Law
1	$P_1 = \begin{bmatrix} 18.6616 & -18.9560 \\ -18.9560 & 28.1086 \end{bmatrix}$	$F_1 = \begin{bmatrix} 2,3172 & -2,3317 \end{bmatrix}$
2	$P_2 = \begin{bmatrix} 30.8818 & -21.6010 \\ -21.6010 & 36,2739 \end{bmatrix}$	$F_2 = \begin{bmatrix} 4,1684 & -3,7131 \end{bmatrix}$
3	$P_3 = \begin{bmatrix} 35,4175 & -38,6129 \\ -38,6129 & 49,7079 \end{bmatrix}$	$F_3 = \begin{bmatrix} -5,1657 & 5,7933 \end{bmatrix}$

5 Conclusion

In this paper we have traced a parallel between the temporal difference methods (TD for short) for Markovian decision processes (MDP for short) with a λ-policy iteration like algorithm for obtaining the maximal solution of the set of coupled algebraic Riccati equations (CARE for short) for the optimal control of discrete-time Markovian jump linear systems. We related the iterations (8),(9) with those in the so called value iteration technique in MDP and the so-called quasi-linearization method presented in Lemma 4 with the policy iteration technique for MDP, involving a policy evaluation part, and a policy improvement part (see Remark 1).

TD have been applied to solve problems related to MDP (see for instance [2], [15]). Usually it is preferable to carry out the policy evaluation step of the policy iteration method by using value iteration and a drawback of this approach is that the value iteration algorithm may converge very slow for discount factor close to 1. The idea of TD is speed up the convergence of the policy evaluation step of the policy iteration method by using value iteration. This is done by using a parameter, usually named λ, which gives origin to the TD λ-policy iteration methods. For $\lambda = 1$, the method reduces to the policy iteration technique, while that for $\lambda = 0$, we have the value iteration technique.

By applying these ideas to the maximal solution of (3), (4), we obtained an iterative method that traces a close parallel with the TD λ-policy iteration methods in MDP. For $\lambda = 0$ our technique reduces to that presented in equations (8), (9) (value iteration), while for $\lambda = 1$ it reduces to the quasi-linearization method presented in Lemma 4 below (policy iteration). As for the MDP, the advantage of the method is that with an appropriate choice of λ between 0 and 1, we can speed up the convergence of the policy evaluation step of the policy iteration method by using value iteration.

References

[1] H. Abou-Kandil, G. Freiling, and G. Jank. On the solution of discrete-time markovian jump linear quadratic control problems. *Automatica*, 31(5):765–768, 1995.

[2] M. Ait-Rami and L. E. Ghaoui. LMI optimization for nonstandard riccati equations arising in stochastic control. *IEEE Trans. Automatic Control*, 41(11):1666–1671, 1996.

[3] D. P. Bertesekas and J.N. Tsitsilklis. *Neuro-Dynamic Programming*. Athena Scientific, 1996.

[4] O.L.V. Costa and E.K. Boukas. Necessary and sufficient condition for robust stability and stabilizability of continuous-time linear systems with markovian jumps. *J. Optimization Theory and Applications*, 99:359–379, 1998.

[5] O.L.V. Costa and M.D. Fragoso. Stability results for discrete-time linear systems with markovian jumping parameters. *J. Math. Analysis and Applic*, 179:154–178, 1993.

[6] O.L.V. Costa and M.D. Fragoso. Discrete-time LQ-optimal control problems for infinite markov jump parameter systems. *IEEE Trans. Automat. Control*, 40:2076–2088, 1995.

[7] O.L.V. Costa and R.P. Marques. Maximal and stabilizing hermitian solutions for discrete-time coupled algebraic riccati equations. *Mathematics of Control, Signals and Systems*, 12:167–195, 1999.

[8] O.L.V. Costa, J.B.R. Do Val, and J.C. Geromel. A convex programming approach to $\mathcal{H}_2$-control of discrete-time markovian jump linear systems. *Int. J. Control*, 66:557–579, 1997.

[9] Z. Gajic and I. Borno. Lyapunov iterations for optimal control of jump linear systems at steady state. *IEEE Trans. Automat. Control*, 40(11):481–498, 1995.

[10] Y. Ji and H.J. Chizeck. Controllability, observability and discrete-time markovian jump linear quadratic control. *Int. J. Control*, 48:481–498, 1988.

[11] Y. Ji, H.J. Chizeck, X. Feng, and K.A. Loparo. Stability and control of discrete-time jump linear systems. *Control Th. and Adv. Tech*, 7:247–270, 1991.

[12] W.P. Blair Jr and D.D. Sworder. Feedback control of a class of linear discrete system with jump parameters and quadratic cost criteria. *Int. J. Control*, 21:833–841, 1975.

[13] M. Mariton. Almost sure and moments stability of jump linear systems. *Systems and Control Letters*, 11:393–397, 1988.

[14] M. Mariton. *Jump Linear Systems in Automatic Control.* Marcel Dekker, 1990.

[15] R. S. Sutton and A. G. Barto. *Reinforcement Learning - An Introduction.* MIT Press, 1998.

[16] J.B.R. Do Val, J.C. Geromel, and O.L.V. Costa. Uncoupled riccati iterations for the linear quadratic control problem of discrete-time markov jump linear systems. *IEEE Trans. Automat. Control*, 43:1727–1733, 1998.

[17] J.B.R. Do Val, J.C. Geromel, and O.L.V. Costa. Solutions for the linear quadratic control problem of markov jump linear systems. *J. Optimization Theory and Applications*, 103:283–311, 1999.