On Subspace Decompositions of Finite Horizon DP Problems with Switched Linear Dynamics

Manolis C. Tsakiris and Danielle C. Tarraf

Abstract—We consider finite horizon dynamic programming problems for switched linear dynamics over a finite dimensional but otherwise arbitrary state-space, where the cost function depends only on the state. We start from a decomposition of the underlying state-space as the sum of invariant subspaces under all linear transformations induced by the input, and pose the following question: Assuming the cost function has an additive structure compatible with the state-space decomposition, under what conditions does the cost-to-go function also exhibit an additive structure? We begin by deriving a necessary and sufficient condition for the existence of this additive structure. We then propose a sufficient condition, expressed in terms of the cardinality of the input set: While not necessary, this condition has the advantage of being readily verifiable. Finally, we characterize the resulting reduction in complexity for the case of systems with finite state-spaces, and we conclude with a simple illustrative example.

I. INTRODUCTION

Dynamic programming (DP), pioneered by Bellman [1], has found wide-ranging applications in areas as diverse as operations research [3], bioinformatics [13], biology [12] and coding theory and cryptography [8]. Dynamic programming ideas are also extensively used in various control settings, in particular when solving optimal control problems for switched linear systems, whether using exact models [2], [14] or approximate finite state models [20], [19], [17].

The well-known Principle of Optimality results in a fairly intuitive solution approach for DP problems. However, it also suffers from the curse of dimensionality, namely that the computational complexity of the DP algorithm increases exponentially with the dimension of the underlying state-space as well as that of the underlying input space. As the complexity of the problem becomes prohibitive in obtaining an optimal solution, various methods have been developed where the cost-to-go function, also known as the value function, is approximated and a suboptimal solution is obtained in an efficient manner [5],[14]. The importance of the structure of the cost-to-go function has also been noted in the literature [10].

In this paper, we consider finite horizon dynamic programming problems of switched linear dynamics over a finite dimensional, but otherwise arbitrary, state-space, with the cost function depending only on the state. Building on ideas developed in our previous work, [23], [22], we start from a decomposition of the underlying state-space as the sum of invariant subspaces under all linear transformations induced by the input, we assume that the cost function of the state exhibits an additive structure compatible with the underlying state-space decomposition, and we pose the question: Under what conditions does the cost-to-go function likewise exhibit an additive structure? Such a structure would allow for significant reduction in the complexity of computing the cost-to-go function, particularly for finite state systems long of particular interest to us [18], [21], and would suggest a desirable property of a parametric family to be used in approximating the value function [10].

Specifically, we present three contributions: First, we derive a necessary and sufficient condition for the cost-to-go function to exhibit an additive structure, assuming that the cost function is additive. While this characterization is of theoretical relevance, its practical value is limited as checking this condition is as expensive as solving the original DP problem. Thus our second contribution consists of a readily verifiable sufficient condition: Perhaps surprisingly, this sufficient condition is expressed in terms of the cardinality of the input set. Our third contribution characterizes the resulting reduction in complexity for a class of finite memory systems, specifically switched linear systems defined over finite fields.

Even though the developments presented in this work were mainly the result of theoretical interest, they become relevant in applications involving optimal control of switched systems, where the state of each system is determined by the states of a set of similarly behaved subsystems. A prominent example is the emerging field of control of quantum systems, in which the state is a linear combination of the eigenstates of the Hamiltonian operator that describes the system. Optimal control is of great importance in quantum systems, e.g. quantum memories, and dynamic programming has been used in that context [4], [7]. Furthermore it has been shown that in certain instances stabilization of a quantum system can only be achieved by means of switching controls [6]. A second example where our results become relevant is the emerging field of multi-agent systems, where it has been shown that in some important cases a multi-agent system can be modelled as a switched linear system [9] or as a linear system over a finite field [16].

The paper is organized as follows: We present the problem setup and problem statement in Section II. We state the main results in Section III and present their derivation in Section IV. We conclude with illustrative examples in Section V.

A word on notation: \( \mathbb{R}^+ \) denotes the set of non-negative reals. For sets \( \mathcal{X} \) and \( \mathcal{Y} \), \( \mathcal{X}^\mathcal{Y} \) denotes the set of all maps from \( \mathcal{Y} \) to \( \mathcal{X} \). For \( f : \mathcal{X} \to \mathbb{R}^+ \), \( \arg\min_{x \in \mathcal{X}} f(x) \) denotes
the set of arguments $x$ in $\mathcal{X}$ that minimize function $f$. For maps $f : \mathcal{X} \to \mathcal{Y}$ and $g : \mathcal{Y} \to \mathcal{W}$, $g \circ f$ denotes the composite map from $\mathcal{X}$ to $\mathcal{W}$ defined by $g \circ f(x) = g(f(x))$. The symbol $\rightarrow$ is used to denote an injective mapping of sets. We denote by $\text{diag}(A_1, \ldots, A_k)$ a square block diagonal matrix where square matrix $A_i$ appears as the $i^{\text{th}}$ block diagonal entry. We use $0_k$ to denote a column vector of size $k$ containing zeros in its entries and similarly $0_{k \times k}$ a $k \times k$ zero matrix. If $x_1, \ldots, x_k$ are elements of a vector space $\mathcal{X}$ over a field $\mathcal{F}$, we denote by $< x_1, \ldots, x_k >$ the subspace spanned by $x_1, \ldots, x_k$ over $\mathcal{F}$. We use $\oplus$ to denote direct sum of subspaces. We use $\mathcal{L}(\mathcal{X})$ to denote the set of all linear transformations $\mathcal{X} \to \mathcal{X}$. If $\mathcal{X}$ is a vector space, $\mathcal{S}, \mathcal{T}$ subspaces such that $\mathcal{X} = \mathcal{S} \oplus \mathcal{T}$ and $\sigma \in \mathcal{L}(\mathcal{X})$, we denote by $\sigma|_{\mathcal{S}}$ the linear transformation $\sigma|_{\mathcal{S}} : \mathcal{S} \to \mathcal{X}$ defined by $\sigma|_{\mathcal{S}}(s) = \sigma(s)$, for any $s \in \mathcal{S}$.

II. PRELIMINARIES

A. The Dynamic Programming Algorithm

Let $\mathcal{X}$ be a finite dimensional vector space of dimension $n$ over a field $\mathcal{F}$, and let $\mathcal{U}$ be a finite set. Associate with every $u \in \mathcal{U}$ a distinct element $A_u$ of $\mathcal{L}(\mathcal{X})$.

Consider the discrete-time dynamical system defined by the state transition equation

$$x_{t+1} = A_u x_t \tag{1}$$

where $x_t \in \mathcal{X}$, $u_t \in \mathcal{U}$, $t \in \mathcal{T} = \{0, 1, \ldots, T - 1\}$ and $T$ is a given positive integer. Consider also a non-negative cost function

$$g : \mathcal{X} \to \mathbb{R}^+ \tag{2}$$

Given any initial state $x_0 \in \mathcal{X}$, the optimal control problem consists of finding an optimal policy $\pi^*(x_0) \in \mathcal{U}^T$ that minimizes the additive cost

$$J(x_0, \pi) = \sum_{t=0}^{T} g(x_t) \tag{3}$$

over all control policies $\pi \in \mathcal{U}^T$, when $x_t$ evolves according to (1). We will denote this DP problem defined by equation (1), cost function (2) and objective (3), by $\langle \{A_u\}, g \rangle$ for short.

The solution of this problem hinges on the Principle of Optimality [1]. Indeed, let $\mathcal{T}_i = \{t, t+1, \ldots, T - 1\}$ and define the cost-to-go function

$$J^*_t(x) = \begin{cases} \min_{u \in \mathcal{U}} \sum_{t=1}^{T} g(x_r) & , \ t \in \mathcal{T} \\ g(x) & , \ t = T \end{cases} \tag{4}$$

for the system evolving according to (1) with $x_t = x$. The principle of optimality can then be stated as

$$J^*_t(x) = g(x) + \min_{u \in \mathcal{U}} J^*_{t+1}(A_u x) \tag{5}$$

for any $t \in \mathcal{T}$ and $x \in \mathcal{X}$. This observation is used as the basis of the recursive DP algorithm, which is applied backwards in time [3] to compute the cost-to-go function and a corresponding optimal controller $u^* : \mathcal{X} \times \mathcal{T} \to \mathcal{U}$ satisfying

$$u^*(x, t) \in \arg \min_{u \in \mathcal{U}} J^*_{t+1}(A_u x). \tag{6}$$

For notational simplicity, we will be writing $u^*_t(x)$ to denote $u^*(x, t)$ and $J^*(x, t)$ to denote $J^*_{t+1}(x, t)$, the optimal value of $J(x, t)$.

B. Problem Statement

Assume that there exists a finite index set $\mathcal{I}$ with cardinality $r > 1$ and subspaces $\mathcal{X}_i$ of $\mathcal{X}$ such that

$$\mathcal{X} = \bigoplus_{i \in \mathcal{I}} \mathcal{X}_i \tag{7}$$

and

$$A_u(\mathcal{X}_i) \subseteq \mathcal{X}_i, \ \forall u \in \mathcal{U}, \ \forall i \in \mathcal{I}. \tag{8}$$

**Remark 1:** Given linear transformation $A : \mathcal{X} \to \mathcal{X}$ of the vector space $\mathcal{X}$, there is always a natural decomposition of $\mathcal{X}$ as the direct sum of $A$-invariant subspaces, irrespectively of the underlying field $\mathcal{F}$. Indeed, in the standard setting where $\mathcal{X} = \mathbb{R}^n$ and the eigenvalues of $A$ are real, the invariant subspaces are the generalized eigenspaces of the Jordan canonical form of matrix $A$. The more general case is addressed in the literature (see Chapter 7 in [15] and Chapter XIV.2 in [11]). For examples of switched systems where the structure of (5) and (6) arises, the reader is referred to Section V.

Let $f : \mathcal{X} \to \mathbb{R}^+$. We say $f \in \mathcal{G}_n$, if for any $x \in \mathcal{X}$ with components $x_i$ in $\mathcal{X}_i$, we have

$$f(x) = \sum_{i \in \mathcal{I}} f(x_i). \tag{9}$$

$\mathcal{G}_n$ is thus the set of all real-valued functions of the state that additively decompose in a manner compatible with the underlying decomposition of the state-space.

The problem we are interested in addressing in this paper is then the following:

**Problem 1:** Consider given a DP problem $\langle \{A_u\}, g \rangle$ and a decomposition (5), (6) of the state-space. Assuming that $g \in \mathcal{G}_n$, find conditions under which the cost-to-go function is additive along the subspaces, that is

$$J^*_t(x) \in \mathcal{G}_n \ \forall t \in \mathcal{T}. \tag{10}$$

**Remark 2:** Problem 1 admits an intuitive interpretation. Indeed, under conditions (5) and (6), we can think of $r$ independent subsystems, driven in parallel by a common input. Each subsystem is a switched linear system with state-space $\mathcal{X}_i$, dynamics $x_{i,t+1} = A_u|_{\mathcal{X}_i} x_{i,t}$, cost function $g(x_i)$ and cost-to-go function $J^*_t(x_i)$. Note that in general we have

$$J^*(x_0) \geq \sum_{i \in \mathcal{I}} J^*(x_{i,0}). \tag{11}$$

In other words, the control policy that is optimal for the overall interconnected system is not optimal for each individual system, in general. Problem 1 can thus be interpreted...
as simply asking: When is equality achieved in (7), or equivalently, when do all the subsystems admit a common optimal control policy?

III. MAIN RESULTS

We state and briefly discuss our main results in this Section. We begin with the following proposition, which formalizes the intuitive interpretation of Problem 1 given in Remark 2.

**Proposition 1:** Let \( g \in \mathcal{G}_s \). Then the following are equivalent:

1. \( J^*_t \in \mathcal{G}_s \), \( \forall t \in T \)

2. For any \( x \in \mathcal{X} \), \( t \in T \), there exists \( u^*_t(x) \in \cap_{i \in I} \arg \min_{u \in \mathcal{U}} J^*_{t+1}(A_u x_i) \)

where \( x_1 + \ldots + x_r, x_i \in \mathcal{X}_i \), is the unique decomposition of \( x \).

**Theorem 1:** Let \( g \in \mathcal{G}_s \) and for any \( i \in I \) define \( m_i = \text{card}((A_u)_{x_i} : u \in \mathcal{U}) \). If \( \text{card}(\mathcal{U}) = \prod_{i \in I} m_i \), then \( J^*_t \in \mathcal{G}_s \), \( \forall t \in T \).

Theorem 1 gives a sufficient criterion that is simple to check, under which Problem 1 has a solution, i.e. \( J^*_t \in \mathcal{G}_s \). Note that in this case, we have a significant reduction in computing the cost-to-go function, and thus an optimal controller. More specifically, in the case where the field \( F \) is finite, we have an exponential reduction in complexity, as seen in the following statement:

**Proposition 2:** Suppose the conditions of Theorem 1 are true and that the underlying field \( F \) is finite containing \( q \) elements. Then we have a reduction in the computational complexity from \( g^n \cdot T \cdot \text{card}(\mathcal{U}) \) to \((\sum_{i=1}^n q^m_i) \cdot T \cdot \text{card}(\mathcal{U})\) function evaluations, where \( n_i = \text{dim}(\mathcal{X}_i) \).\(^1\)

IV. DERIVATION OF MAIN RESULTS

In this section we provide the derivations of the main results stated in the previous section. We begin with Proposition 1.

**Proof:** [Proposition 1] (\( \Rightarrow \)) Suppose \( J^*_t \in \mathcal{G}_s \), \( \forall t \in T \). Then for any \( t \in T \) and for any \( x \in \mathcal{X} \) with component \( x_i \in \mathcal{X}_i \), we have that

\[
J^*_t(x) = \sum_{i \in I} J^*_t(x_i)
\]

or equivalently by the principle of optimality (4)

\[
g(x) + \min_{u \in \mathcal{U}} J^*_{t+1}(A_u x) = \sum_{i \in I} \left[ g(x_i) + \min_{u \in \mathcal{U}} J^*_{t+1}(A_u x_i) \right]
\]

or equivalently since \( g \in \mathcal{G}_s \)

\[
\min_{u \in \mathcal{U}} J^*_{t+1}(A_u x) = \sum_{i \in I} \min_{u \in \mathcal{U}} J^*_{t+1}(A_u x_i).
\]

\(^1\)This is indeed a reduction for reasonably large values of \( n \); e.g. for \( q = 2 \) we need \( n \geq 3 \), for \( q = 3 \) we need \( n \geq 2 \).

For \( u^*_t(x_i) \in \arg \min_{u \in \mathcal{U}} J^*_{t+1}(A_u x_i) \) and \( u^*_t(x) \in \arg \min_{u \in \mathcal{U}} J^*_{t+1}(A_u x) \) we rewrite (8) as

\[
J^*_{t+1}(A_u^* x) = \sum_{i \in I} J^*_{t+1}(A_u^* x_i) .
\]

Now, by hypothesis \( J^*_{t+1} \in \mathcal{G}_s \) and so

\[
J^*_{t+1}(A_u^* x) = \sum_{i \in I} J^*_{t+1}(A_u^* x_i) .
\]

From equations (9) and (10) we have

\[
\sum_{i \in I} J^*_{t+1}(A_u^* x_i) = \sum_{i \in I} J^*_{t+1}(A_u^* x_i) .
\]

Now, by definition of \( u^*_t(x_i) \) we have that

\[
J^*_{t+1}(A_u x_i) \leq J^*_{t+1}(A_u x_i), \forall u \in \mathcal{U}, \forall i \in I.
\]

Consequently, we have that

\[
J^*_{t+1}(A_u^* x_i) \leq J^*_{t+1}(A_u^* x_i) .
\]

and summing (13) over all \( i \in I \) gives

\[
\sum_{i \in I} J^*_{t+1}(A_u^* x_i) \leq \sum_{i \in I} J^*_{t+1}(A_u^* x_i) .
\]

If any of the inequalities in (13) is strict, then the inequality in (14) will also be strict, in violation of (11). Hence

\[
J^*_{t+1}(A_u^* x_i) = J^*_{t+1}(A_u^* x_i)
\]

holds for every \( i \in I \), and so

\[
u^*_t(x) \in \cap_{i \in I} \arg \min_{u \in \mathcal{U}} J^*_{t+1}(A_u x_i).
\]

(\( \Leftarrow \)) Conversely, suppose that for any \( x \in \mathcal{X} \), \( t \in T \), there exists

\[
u^*_t(x) \in \cap_{i \in I} \arg \min_{u \in \mathcal{U}} J^*_{t+1}(A_u x_i).
\]

We will show by backward induction on \( t \) that this implies \( J^*_t \in \mathcal{G}_s \). We know this is true by hypothesis for \( t = T \), since \( J^*_T = g \). For \( t = T - 1 \) we have for any \( x \in \mathcal{X} \) with component \( x_i \in \mathcal{X}_i \) that

\[
\min_{u \in \mathcal{U}} J^*_T(A_u x) = \min_{u \in \mathcal{U}} \left[ \sum_{i \in I} J^*_T(A_u x_i) \right]
\]

and so

\[
\min_{u \in \mathcal{U}} J^*_T(A_u x) \geq \sum_{i \in I} \min_{u \in \mathcal{U}} J^*_T(A_u x_i) .
\]

Now by hypothesis there exists some \( u_0 \in \cap_{i \in I} \arg \min_{u \in \mathcal{U}} J^*_T(A_u x_i) \). Thus the left hand term of (16) satisfies

\[
\min_{u \in \mathcal{U}} J^*_T(A_u x) = J^*_T(A_{u_0} x),
\]

while the right hand term satisfies

\[
\sum_{i \in I} \min_{u \in \mathcal{U}} J^*_T(A_{u_0} x_i) = \sum_{i \in I} J^*_T(A_{u_0} x_i) = J^*_T(A_{u_0} x)
\]
with the second equality following from \( J^*_T = g \in G_s \). It follows from (18) and (17) that equality is achieved in (16), and we can write
\[
\min_{u \in \mathcal{U}} J^*_T(A_u x) = \sum_{i \in \mathcal{I}} \left[ \min_{u \in \mathcal{U}} J^*_T(A_u x_i) \right].
\] (19)

Adding \( g(x) \) to both sides of (19) and noting that \( g \in G_s \), we can write
\[
g(x) + \min_{u \in \mathcal{U}} J^*_T(A_u x) = \sum_{i \in \mathcal{I}} \left[ g(x_i) + \min_{u \in \mathcal{U}} J^*_T(A_u x_i) \right].
\]

Now, invoking the principle of optimality (4), the last equation says that \( J^*_T(A_u x) \) achieved in (18) and (17) that equality is achieved in \( J^*_T \in G_s \). The induction step can now be performed in a similar manner to show that \( J^*_T \in G_s \) for all \( t \in T \).

To prove Theorem 1 we need the following Lemma:

**Lemma 1:** Suppose that \( g \in G_s \). If for any choice of elements \( u_1, \ldots, u_r \) in \( \mathcal{U} \) there exists a corresponding choice of \( u_0 \in \mathcal{U} \) such that
\[
A_{u_0} | x_i = A_{u_i} | x_i, \quad 1 \leq i \leq r,
\]
then \( J^*_T \in G_s \), \( \forall t \in T \).

**Proof:** By backward induction on \( t \). Since \( J^*_T = g \in G_s \), we have
\[
\min_{u \in \mathcal{U}} J^*_T(A_u x) \geq \sum_{i \in \mathcal{I}} \left[ \min_{u \in \mathcal{U}} J^*_T(A_u x_i) \right].
\] (20)

Letting \( u_{T-1}^*(x_i) \in \arg \min_{u \in \mathcal{U}} J^*_T(A_u x_i) \), \( i \in \mathcal{I} \), we can write the right hand side of (20) as
\[
\sum_{i \in \mathcal{I}} \left[ \min_{u \in \mathcal{U}} J^*_T(A_u x_i) \right] = \sum_{i \in \mathcal{I}} J^*_T(A_{u_{T-1}^*} x_i).
\] (21)

Now by hypothesis there exists some \( u_0 \in \mathcal{U} \) such that
\[
A_{u_0} | x_i = A_{u_{T-1}^*} | x_i, \quad \forall i \in \mathcal{I}
\]
and consequently
\[
A_{u_0} x_i = A_{u_{T-1}^*} x_i, \quad \forall i \in \mathcal{I}.
\]

Thus equation (21) can be written as
\[
\sum_{i \in \mathcal{I}} \left[ \min_{u \in \mathcal{U}} J^*_T(A_u x_i) \right] = \sum_{i \in \mathcal{I}} J^*_T(A_{u_0} x_i).
\]

Since \( J^*_T = g \in G_s \)
\[
J^*_T(A_{u_0} x) = \sum_{i \in \mathcal{I}} J^*_T(A_{u_0} x_i) = \sum_{i \in \mathcal{I}} \left[ \min_{u \in \mathcal{U}} J^*_T(A_u x_i) \right].
\] (22)

By equation (20) \( J^*_T(A_u x) \) is bounded below by \( \sum_{i \in \mathcal{I}} \left[ \min_{u \in \mathcal{U}} J^*_T(A_u x_i) \right] \). By equation (22) this lower bound is achieved by \( u_0 \), hence we have
\[
J^*_T(A_{u_0} x) = \min_{u \in \mathcal{U}} J^*_T(A_u x).
\]

This, together with (22) gives
\[
\min_{u \in \mathcal{U}} J^*_T(A_u x) = \sum_{i \in \mathcal{I}} \left[ \min_{u \in \mathcal{U}} J^*_T(A_u x_i) \right].
\] (23)

Adding \( g(x) \) to both sides and invoking the principle of optimality (4), yields that \( J^*_T(A_u x) \in G_s \). The induction step can now be performed in a similar manner to show that \( J^*_T \in G_s \), \( \forall t \in T \).

**Proof:** [Theorem 1] Define the subset \( A_i \) of \( \mathcal{L}(X_i) \) as
\[
A_i = \{ (A_u) | x_i : u \in \mathcal{U} \}
\]
and let \( m_i = \text{card}(A_i) \). Define \( A \subseteq \mathcal{L}(X) \) by
\[
A = \{ A \in \mathcal{L}(X) : A | x_i \in A_i, \forall i \in \mathcal{I} \}.
\]

Obviously \( A_n | x_i \in A_i, \forall u \in \mathcal{U}, \forall i \in \mathcal{I} \) and thus \( A_n \subseteq A \), \( \forall n \in \mathcal{U} \). Hence the mapping \( u \mapsto A_u \), denoted by \( \phi(u) = A_u \) is of the form
\[
\phi : \mathcal{U} \rightarrow A.
\]

Note that \( \phi \) is injective because by hypothesis \( u \neq u' \Rightarrow A_u \neq A_{u'} \). By construction we have that
\[
\text{card}(A) = \prod_{i \in \mathcal{I}} m_i.
\]

Now suppose that \( \text{card}(\mathcal{U}) = \text{card}(A) \). Since \( \phi \) is an injective mapping of the finite set \( \mathcal{U} \) into the finite set \( A \) of equal cardinality, then \( \phi \) will also be surjective. Given any \( u_i \in \mathcal{U}, i \in \mathcal{I} \), let \( A \) be the linear transformation of \( X \) defined by \( A | x_i = A_{u_i} | x_i \). Then \( A \in A \) and by the surjectivity of \( \phi \), there will exist \( u \in \mathcal{U} \) such that \( \phi(u) = A_u = A \). Then the statement follows by invoking Lemma 1.

**Proof:** [Proposition 2] The computation of \( J^*_T \) is done recursively, using the principle of optimality (4). The operations involved in the computation are function evaluations, comparisons, matrix vector multiplications and additions, with the first two consisting of the computational burden. We prove the argument for the function evaluations (the argument for the comparisons is similar). In general, to compute \( J^*_T \) from \( J^*_{T+1} \) at \( x \in X \), we need to compute \( \min_{u \in \mathcal{U}} J^*_T(A_u x) \), which requires \( \text{card}(\mathcal{U}) \) function evaluations. Hence computation of \( J^*_T \) requires \( \text{card}(X) \cdot \text{card}(\mathcal{U}) = q^n \cdot \text{card}(\mathcal{U}) \) function evaluations. This must be done for \( T \) time steps giving a total complexity of \( q^n \cdot T \cdot \text{card}(\mathcal{U}) \) function evaluations. If however \( J^*_T \in G_s \), then we only need to compute \( J^*_T | x_i, \forall i \in \mathcal{I} \). Following a similar argument as above, the complexity of computing \( J^*_T | x_i \) for all time steps is \( q^n \cdot T \cdot \text{card}(\mathcal{U}) \), where \( n_i = \text{dim}(X_i) \). Hence the total complexity is \( (\sum_{i=1}^r q^n) \cdot T \cdot \text{card}(\mathcal{U}) \).

**V. EXAMPLES**

**Example 1:** We begin with an example where the structure of (5) and (6) arises. Let \( X = \mathbb{R}^n \) and let \( A \) be an \( n \times n \) real matrix with real eigenvalues and with Jordan decomposition \( A = SJS^{-1} \). Let the input set be \( \mathcal{U} = \{ 1, 2, \ldots, k \} \), where \( \kappa \) is a positive integer equal to the cardinality of \( \mathcal{U} \). Now let
Then the generalized eigenspaces of $A$ are invariant under $A^u = S(JA)^u S^{-1}$ for any $u$.

**Example 2:** As an example where the theory developed in this paper applies, consider the case where the underlying field is the reals $\mathbb{R}$ and the state-space $\mathcal{X}$ is $\mathbb{R}^n$, $n \geq 3$. Define the input set to be all the binary strings of length 3, i.e.

$$\mathcal{U} = \{000, 001, 010, 011, 100, 101, 110, 111\}.$$ 

Let $n_1, n_2, n_3$ be positive integers such that $n = n_1 + n_2 + n_3$. Now, for each $i = 1, 2, 3$, let $A^{(i)}_0$, $A^{(i)}_1$ be two distinct $n_i \times n_i$ real valued matrices and let

$$u = (ijk) \mapsto A_u = \begin{bmatrix} A^{(1)}_i & 0 & 0 \\ 0 & A^{(2)}_i & 0 \\ 0 & 0 & A^{(3)}_i \end{bmatrix}, \quad i, j, k \in \{0, 1\}.$$ 

Then the common invariant subspaces of the state-space under the system transformations induced by the input are

$$\mathcal{X}_1 = \langle e_1, \ldots, e_{n_1} \rangle_\mathbb{R},$$

$$\mathcal{X}_2 = \langle e_{n_1+1}, \ldots, e_{n_1+n_2} \rangle_\mathbb{R},$$

$$\mathcal{X}_3 = \langle e_{n_1+n_2+1}, \ldots, e_n \rangle_\mathbb{R}$$

where $e_i$ is a $n$-tuple with a 1 in the $i^{th}$ entry and zeros everywhere else. Define the cost function $g : \mathbb{R}^n \to \mathbb{R}^+$ to be a positive-definite quadratic form represented by the matrix

$$Q_g = \begin{bmatrix} Q_1 & 0 & 0 \\ 0 & Q_2 & 0 \\ 0 & 0 & Q_3 \end{bmatrix}$$

where $Q_1, Q_2, Q_3$ are $n_i \times n_i$, $i = 1, 2, 3$, positive-definite real matrices. Clearly $g$ satisfies the splitting property $g \in \mathcal{G}_s$. Finally, define the finite horizon to be equal to a positive integer $T$.

Now, for any $u \in \mathcal{U}$, $A_u | \mathcal{X}_i$ can take only two values: either $\text{diag} \left( A^{(1)}_0, 0_{n_2 \times n_2}, 0_{n_3 \times n_3} \right)$ or $\text{diag} \left( A^{(1)}_1, 0_{n_2 \times n_2}, 0_{n_3 \times n_3} \right)$. Similarly, $A_u | \mathcal{X}_2$ can only take the values $\text{diag} \left( 0_{n_1 \times n_1}, A^{(2)}_0, 0_{n_3 \times n_3} \right)$ or $\text{diag} \left( 0_{n_1 \times n_1}, A^{(2)}_1, 0_{n_3 \times n_3} \right)$, and $A_u | \mathcal{X}_3$ can only take the values $\text{diag} \left( 0_{n_1 \times n_1}, 0_{n_2 \times n_2}, A^{(3)}_0 \right)$ or $\text{diag} \left( 0_{n_1 \times n_1}, 0_{n_2 \times n_2}, A^{(3)}_1 \right)$. So, referring to the notation of Theorem 1

$$m_i = 2, \quad i = 1, 2, 3.$$ 

Thus

$$\prod_{i=1}^{3} m_i = 8 = \text{card}(\mathcal{U})$$

and so by Theorem 1 we have that for any $x_i \in \mathbb{R}^{n_i}, i \in \{1, 2, 3\}$ and for any $t \in \{0, 1, \ldots, T\}$

$$J^*_t \left( \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) = \text{diag} \left( \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right).$$

**VI. Conclusions**

We considered finite horizon DP problems with underlying switched linear dynamics over a finite dimensional, but otherwise arbitrary, state-space, and where the cost function depends only on the state. Assuming the existence of a decomposition of the state-space as the sum of invariant subspaces under all linear transformations induced by the input, and assuming that the cost function decomposes additively along these subspaces, we showed that the cost-to-go function is also additive if and only if there exists an input symbol that is optimal with respect to all of the components of the state vector at every time instant. Next, we proposed and derived a readily verifiable sufficient condition for this decomposition to occur: Interestingly, this condition is written in terms of the cardinality of the input set. Finally, we characterized the resulting complexity reduction for the case of finite memory systems, of particular interest to us.

**VII. Acknowledgments**

This research was supported by AFOSR YIP award FA9550-11-1-0118 and NSF CAREER award ECCS 0954601.

**References**


