On a Relaxation of Time-Varying Actuator Placement

Alex Olshevsky, Member, IEEE

Abstract—We consider the time-varying actuator placement problem in continuous time, where the goal is to maximize the trace of the controllability Grammian. A natural relaxation of the problem is to allow the binary \((0, 1)\) variable indicating whether an actuator is used at a given time to take on values in the closed interval \([0, 1]\). We show that all optimal solutions of both the original and the relaxed problems can be given via an explicit formula, and that, as long as the input matrix has no zero columns, the solutions sets of the original and relaxed problem coincide.

Index Terms—Control of networks, network analysis and control, optimization.

I. INTRODUCTION

We consider the time-varying actuator placement problem: informally, given a differential equation with input, we would like to optimize some controllability-related objective while using few nonzero inputs per time step. This is motivated by scenarios where setting an input to something nonzero at a given time carries a fixed cost that can be much larger than the cost of synthesizing the input itself. Our variation of the problem is “time-varying” in the sense that we allow different inputs to be nonzero at different times; this is in contrast to “fixed” actuator placement problems, where one has to select the same set of actuators to be nonzero across all time.

Formally, our goal is to choose a diagonal matrix \(V(t)\) whose entries lie in the binary set \([0, 1]\) optimizing some controllability-related properties of the resulting differential equation

\[
\dot{x}(t) = Ax(t) + BV(t)u(t),
\]

where we assume that \(A \in \mathbb{R}^{n \times n}\) and \(B \in \mathbb{R}^{n \times m}\) is a matrix with no zero columns.\(^1\) The multiplication of the input \(u(t)\) by the diagonal matrix \(V(t)\) can be thought of as choosing to use only certain actuators. Indeed, if \(V_i(t) = 0\), then \(u_i(t)\) has no effect on \(x(t)\), and the \(i\)'th entry of the input is ignored at time \(t\).

Typical controllability-related objectives are usually formulated in terms of the controllability Grammian, which we define\(^2\)

\[
W_V = \int_0^T e^{At}BV(t)V(t)^TB^Te^{A^Tt} \, dt.
\]

The most natural objective is perhaps to minimize \(\text{Tr}(W_V^{-1})\), which is proportional to the average energy to move from the origin to a uniformly random point on the unit sphere (see discussion in [17]). However, this function is often challenging to reason about. For example, as a consequence of [16], [23] a number of optimization problems involving \(\text{Tr}(W_V^{-1})\) are NP-hard.

We follow several recent papers which instead consider maximization of \(\text{Tr}(W_V)\). This is because \(\text{Tr}(W_V)\) is easier to reason about and can be used to construct bounds on \(\text{Tr}(W_V^{-1})\) (see discussion in [10], [14], [15]). Furthermore, we will seek to do so in the presence of an upper bound on the number of actuators used per unit time step. More formally, denoting \(V(t) = \text{diag}(v(t))\), it is typically assumed that the diagonal entries \(v_i(t) \in [0, 1]\) satisfy the constraint

\[
\int_{0}^{T} \sum_{i=1}^{m} |v_i(t)| \, dt \leq \alpha,
\]

for some \(\alpha\). We will refer to functions \(v_i(t)\) satisfying these constraints as feasible. Note that, because we have constrained \(v_i(t) \in [0, 1]\), this is the same as requiring that

\[
\sum_{i=1}^{m} \mu(\{t : v_i(t) = 1\}) \leq \alpha,
\]

where \(\mu(\cdot)\) denotes the Lebesgue measure. We will naturally assume that \(\alpha \in (0, mT]\), as otherwise the problem is trivial.

A natural relaxation of the problem is to allow each \(v_i(t)\) to lie in \([0, 1]\) instead of requiring it to take on the binary values \([0, 1]\). We will refer to this as the relaxed time-varying actuator placement problem, and the version where \(v_i(t)\) are

\(^1\)If \(B\) does have zero columns, then the corresponding entry of \(u(t)\) does not affect \(x(t)\). Consequently, we can simply delete the nonzero columns of \(B\) and reindex the vector \(u(t)\).

\(^2\)It would be more standard to replace \(t\) by \(T - t\) in the definition of the controllability Grammian, but since that definition is equivalent to the one we give with a “flipped” \(V(t)\), we prefer to avoid dealing with \(T - t\)'s throughout this letter.
required be in \([0, 1]\) will be referred to as the **original time-varying actuator placement problem**. These definitions lead to the main question which is the concern of this letter, namely **understanding when the optimal solutions sets of the original and relaxed problems coincide**.

Compelling motivation for both problems also from the dual formulation, where one considers instead the trace of the observability Grammian. Since the trace of the inverse of the observability Grammian measures the performance of the least-squares estimate of the initial condition over a fixed time-horizon (e.g., [5, p. 21]), our problem formulation can be motivated as “sensor scheduling” where one seeks to achieve the best performance while using as few sensors as possible.

### A. Previous Work

This letter is most closely related to the recent work [10], where the same question was considered. We next give a statement of the main results of [10].

Let us adopt the notation \(b_j\) for the \(j\)th column of the matrix \(B\). Further, for \(i = 1, \ldots, m\), we consider the functions

\[
f_i(t) = b_i^T e^{A t} e^{B t} b_i.
\]

It is then possible to give a condition in terms of the functions \(f_i(t)\) for the relationship between the original and relaxed time-varying actuator placement problems.

**Theorem 1** [10]: The optimal solution set of the relaxed problem is non-empty. Further, assuming that \(f_i(t)\) is not constant for all \(i \in \{1, \ldots, m\}\), we have that:

1. All optimal solutions of the relaxed problem take \([0, 1]\) values almost everywhere.
2. The optimal solution of the solution sets of the original and relaxed actuator scheduling problems coincide.

This theorem is an amalgamation of [10, Ths. 1–3]. It provides an answer to the motivating concern of the present paper. Related theorems in more general settings were also proved in [9] and [11]. However, the condition that \(f_i(t)\) are not constant is only shown to be sufficient in this theorem.

It is natural to observe that in discrete-time the optimal schedule for the original problem can be found by a greedy method (see [7, Th. 5]). This suggests it may be possible to give a characterization of the optimal schedule for the original & relaxed problems in the continuous-time model studied here.

This letter is also related to the works [14], [15] which studied combinatorial implications of maximization of the trace of the controllability Grammian, relating them to quantities like centrality and communicability in graphs. Also related is [3] which studied combinatorial aspects of the smallest eigenvalue of the controllability Grammian, which is a measure of the maximum control energy to go from the origin to a point on the unit sphere. Moreover, [2] studied a convex relaxation of the problem of maximizing the smallest eigenvalue of the controllability Grammian. Finally, [9] studied the optimization aspects of finding good actuator schedules with a reformulation of the cost function.

Beyond that, the time-varying actuator placement problem is quite old; for example, a version of it dates back to a paper of Athens in 1972 [1]. There is quite a bit of recent work on understanding efficient algorithms as well as fundamental limitations for this problem. For example, fundamental limitations in terms of unavoidably large control energy have been studied in [17], [19], [26] among others. Algorithms for actuator placement, in either the fixed or time-varying regime, based on randomized sampling [4], [8], [12], [20], convex relaxation [22], [24], or greedy methods [6], [21], [25], [27] were studied in recent works. Given the relatively large amount of work done on different versions of the problem which are not directly related to our motivating concern, we refer the reader to the above papers for a broader overview of the field.

### B. Our Contribution

We show that, under our assumption that \(B\) has no zero columns, the optimal solution sets of the original and relaxed problems always coincide. This comes out as a byproduct of an explicit formula for the solution of the relaxed problem. In turn, this is done by drawing a connection to (a modification) of the classical notion of a rearrangement of a function.

### II. STATEMENT OF THE MAIN RESULT

#### A. The (Asymmetric) Rearrangement

We need to introduce several concepts and notations to state our main result. We adopt the standard notation that the indicator function \(1_{\mathcal{X}}(x)\) equals one if the point \(x\) belongs to the set \(\mathcal{X}\) and zero otherwise. Given a Lebesgue measurable subset \(\mathcal{X} \subset \mathbb{R}\) of the real line of finite measure, we define its rearrangement \(\mathcal{X}^*\) to be the interval \([0, l]\) whose length \(l\) is the same as the Lebesgue measure of \(\mathcal{X}\). As already mentioned, we adopt the convention of using \(\mu(\mathcal{X})\) to denote the Lebesgue measure of the set \(\mathcal{X}\).

Given a measurable nonnegative function \(f : [0, a] \to \mathbb{R}\) with bounded range, its rearrangement \(f^* : [0, a] \to \mathbb{R} \cup \{+\infty\}\) is defined as

\[
f^*(x) = \int_0^x 1_{\{y \in [0,a] \mid f(y) > t\}^*}(x)\ dt
\]

Intuitively, the rearrangement is that it corresponds to “sorting” the function \(f(x)\). In particular [13]:

**Proposition 1**: The rearrangement \(f^*(x)\) is measurable, non-increasing and its level sets have the same measure as the level sets of \(f(x)\), i.e., for all \(a \in \mathbb{R}\),

\[
\mu(f(x) > a) = \mu(f^*(x) > a)
\]

\[
\mu(f(x) = a) = \mu(f^*(x) = a).
\]

This proposition provides intuition for the notion of a rearrangement: because \(f^*(x)\) has the same level sets as \(f(x)\) while being non-increasing, it can be thought of as a “sorted” version of \(f(x)\). We illustrate this with an example.

**Example 1**: Consider \(f(x) = x^2\) defined on the domain \([0, 1]\). In that case, \(y \in [0, 1] : f(y) > t\) is the set \(y \in [0, 1] : y^2 > t\) which, for \(t \in [0, 1]\), equals \((\sqrt{t}, 1]\); when \(t > 1\), the set \(y \in [0, 1] : f(y) > t\) is empty. The rearrangement of \((\sqrt{t}, 1]\) is \([0, 1 - \sqrt{t}]\). Thus, for any \(x \in [0, 1]\) we have that,

\[
f^*(x) = \int_0^1 1_{[0, 1-\sqrt{t}]}(x)\ dt = \int_0^{(1-x)^2} 1\ dt = (1-x)^2.
\]
It is, of course, immediate that \( f(x) = x^2 \) and \( f^*(x) = (1-x)^2 \), when defined over the domain \([0, 1]\), have level sets of the same measure.

### B. Statement of the Main Result

Our first step is to take the functions \( f_i : [0, T] \rightarrow \mathbb{R}, i = 1, \ldots, m \), defined in Eq. (1) and define their “concatenation” \( F(t) : [0, mt] \rightarrow \mathbb{R} \) which consists on putting these functions “side by side” on the interval \([0, mt]\). Formally,

\[
F(t) = f_i(t) \text{ when } t \in ((i - 1)T, iT).
\]

Since the functions \( f_i(t) \) are clearly continuous, they have bounded range over \([0, T]\). Further, these functions are clearly nonnegative. As a result, \( F(t) \) is also nonnegative and also has bounded range, and therefore the rearrangement \( F^*(t) \) is well-defined.

**Definition 1:** We will say that a function \( g : [0, a] \rightarrow \mathbb{R} \cup \{+\infty\} \) is strictly decreasing to the right at \( x_0 \in (0, a) \) if \( g(x_0) > g(x_0 + \epsilon) \) for all \( \epsilon \) small enough. Similarly, we will say that \( g \) is strictly decreasing from the left at \( x_0 \in (0, a) \) if \( g(x_0 - \epsilon) > g(x_0) \) for all \( \epsilon \) small enough.

The main result of this letter is the following theorem.

**Theorem 2:**

1) Suppose \( F^*(x) \) is strictly decreasing to the right at \( x = a \). Then the unique\(^3\) optimal solution to the relaxed time-varying actuator placement problem is

\[
v^\text{opt, r}_i(t) = 1 \{ \tau : f_i(\tau) > F^*(a) \}(t).
\]

2) Suppose \( F^*(x) \) is strictly decreasing to the right at \( x = a \). Then the unique optimal solution to the relaxed time-varying actuator placement problem is

\[
v^\text{opt, l}_i(t) = 1 \{ \tau : f_i(\tau) > F^*(a) \}(t).
\]

3) Suppose \( F^*(x) \) is not strictly decreasing at \( x = a \) either from the left or to the right. Then the set of optimal solutions to the relaxed time-varying actuator placement problem has more than one element. However, all optimal solutions of the relaxed problem can be parametrized as

\[
v^\text{opt, nl}_i(t) = 1 \{ \tau : f_i(\tau) > F^*(a) \}(t) + \sum_{l=1}^{m} \mu(S_l(t)),
\]

for some sets

\[
S_l \subset \{ \tau : f_i(\tau) = F^*(a) \},
\]

with

\[
\sum_{l=1}^{m} \mu(S_l) = \alpha - \mu(\{ \tau : F(\tau) > F^*(a) \})
\]

Since all the solutions exhibited in this theorem are binary, the immediate implication is that the solutions of the original and relaxed problem are always the same. In particular, the condition that the functions \( f_i(t) \) not be constant in Theorem 1 is not necessary, once the pathological cases where \( B \) has a zero column are ruled out (to see the two conditions are not equivalent, consider \( A = 0, B = I \)).

This theorem shows that we can write down the optimal solution(s) of the relaxed problem via an explicit formula. Moreover, once the notion of the rearrangement has been introduced, the theorem is quite intuitive: informally, it says that we have to take the “top slice” of the functions \( f_1(t), \ldots, f_m(t) \) after sorting. This makes sense if one thinks of the variable \( i \) at time \( T - t \).

### III. PROOF OF THE MAIN RESULT

In this section, we will prove our main result, Theorem 2. We begin by stating several properties of the rearrangement which we will find useful as propositions.

The propositions below hold for all functions \( f(x) \) and \( g(x) \) such that their rearrangements can be defined, i.e., these functions must be nonnegative and have bounded range. Their proofs are fairly standard. Indeed, it is common in the literature to deal with the “symmetric non-increasing rearrangement” in which \( f^*(x) \) is further constructed to be symmetric. For such a notion of rearrangement, the proofs of these facts appear in many places; a standard reference is [13, Ch. 3], which provides hints for many of these, and which can be consulted for general background. We thus do not provide the proofs of these propositions; for completeness, they may be found in the version of this letter on the arxiv \[18\].

**Proposition 2 (Conservation of the \( L^1 \) Norm):**

\[
\int_0^a f(x) \, dx = \int_0^a f^*(x) \, dx.
\]

**Proposition 3 (Hardy-Littlewood Inequality):**

\[
\int_0^a f(x) g(x) \, dx \leq \int_0^a f^*(x) g^*(x) \, dx
\]

Moreover, we have equality if and only if for almost all \( s, t \), we have that

\[
\mu(\{ x : f(x) \geq t \} \cap \{ x : g(x) \geq s \}) = \min(\mu(\{ x : f(x) \geq t \}), \mu(\{ x : g(x) \geq s \})).
\]

**Proposition 4 (Monotonicity):** If \( f(x) \leq g(x) \) for all \( x \), then \( f^*(x) \leq g^*(x) \) for all \( x \). In particular, since a constant function is the rearrangement of itself, if \( f(x) \leq 1 \) for all \( x \in [0, a] \), then \( f^*(x) \leq 1 \) for all \( x \in [0, a] \).

**Proposition 5 (Integral Identity):** Suppose \( b < a \). Then, if \( f^*(x) \) is strictly decreasing to the right at \( x = b \), we have that

\[
\int_0^a f(x) 1_{\{ \tau : f(\tau) \geq F^*(b) \}}(x) \, dx = \int_0^b f^*(x) \, dx
\]

On the other hand, if \( f^*(x) \) is strictly decreasing from the left at \( x = b \), then

\[
\int_0^a f(x) 1_{\{ \tau : f(\tau) > F^*(b) \}}(x) \, dx = \int_0^b f^*(x) \, dx
\]

**Proposition 6 (Integral Identity):** Let \( (b_1, b_2) \) be the largest open interval containing \( b \) on which the function \( f^*(x) \) is constant. Then if \( S \) is a subset of the set \( \{ x : f(x) = f^*(b) \} \),

\[
\int_0^a f(x) 1_{\{ \tau : f(\tau) > F^*(b) \}}(x) \, dx = \int_0^b f^*(x) \, dx.
\]
then
\[
\int_0^a f(x) \left( \mathbb{1}_{\{x \in f(t) > f^*(b)\}}(x) + \mathbb{1}_2(x) \right) \, dx = f^*(b) \mu(S) + \int_0^b f^*(x) \, dx.
\]

We now turn to our first lemma, which introduces some notation pertaining to functions \( f^*(x) \) which are not strictly decreasing from either direction at a point.

**Lemma 1:** Suppose \( f : [0, a] \to \mathbb{R} \) and \( f^*(x) \) is not decreasing from either the left or the right at the point \( x = b \) with \( b \in (0, a) \). Then there exists an open interval containing \( b \), which we denote by \( (b', b') \), such that

1) \( f^*(x) \) is constant on this interval.
2) \( (b', b') \) is the largest open interval containing \( b \) with this property.
3) \[
\mu(\{ x : f(x) > f^*(b) \}) = b'
\]
\[
\mu(\{ x : f(x) = f^*(b) \}) = b' - b
\]
\[
\mu(\{ x : f(x) \geq f^*(b) \}) = b'
\]

**Proof:** Since \( f^*(x) \) is nonincreasing, it is immediate that if it is not strictly decreasing from the left or to the right at \( x = b \), then it must be constant on some interval \((l, r)\) containing \( b \). We can then define \( a' \) to be infimum of all \( l \) such that \((l, r)\) is an interval containing \( b \) on which \( f^*(x) \) is constant; defining \( b' \) similarly, we obtain that \( (b', b') \) is the largest open interval containing \( b \) where \( f^*(x) \) is constant. This proves parts (1) and (2).

For part (3), we have that by Proposition 3,
\[
\mu(\{ x : f(x) > f^*(b) \}) = \mu(\{ x : f^*(x) > f^*(b) \}) = \mu(\{ 0, b' \}) \text{ or } \mu(\{ 0, b' \}) = b'
\]
and the proof of the second and third identities of item (3) proceed similarly.

Our next lemma is a straightforward generalization to the continuous space, of the fact that the largest convex combination of a set of numbers (subject to constraint on how big the weights can be) puts as much weight as possible on the largest numbers. We present it without proof.

**Lemma 2:** Let \( g(t) : [0, a] \to \mathbb{R} \) be a nonincreasing function and suppose \( b \in (0, a) \). Then,
\[
\max_{\gamma(t) \in [0,1], \int_0^a \gamma(t) \, dt = b} \int_0^a \gamma(t) g(t) \, dt = \int_0^b g(t) \, dt.
\]

Moreover,

1) If \( g(x) \) is either decreasing to the right or from the left at \( x = b \), then the maximum is uniquely achieved by the function \( \gamma(t) = 1_{[0,b]}(t) \).
2) If \( g(x) \) is not decreasing from the right or the left at \( t = b \), let \( (b', b') \) be the open interval guaranteed by Lemma 1. Then the functions which achieve the maximum are
\[
\gamma_{\text{opt}}(t) = 1_{[0,b']}(t) + \lambda(t) 1_{Q}(t)
\]
where \( Q \subset [b', b'] \), and \( \int_Q \lambda(t) \, dt = b - b' \).

We next exploit Lemma 2 by applying it to the rearrangement of a function \( f^*(x) \), which of course is nonincreasing. The result is stated as the following lemma.

**Lemma 3:** Let \( a, b \) be scalars satisfying \( b \leq a \) and let us define \( B \) as the set of functions \( \beta(t) : [0, a] \to [0, 1] \) with
\[
\int_0^a \beta(t) \, dt \leq b.
\]

Then
\[
\max_{\beta \in B} \int_0^a \beta(t) f(t) \, dt = \int_0^b f^*(t) \, dt.
\]

Moreover:

1) If \( f^*(t) \) is strictly decreasing to the right at \( t = b \), then the unique \( \beta(t) \) which achieves this maximum is \( \beta(t) = 1_{[f(t) \leq f^*(b)]}(t) \).
2) If \( f^*(t) \) is strictly decreasing from the left at \( t = b \), then the unique \( \beta(t) \) which achieves this maximum is \( \beta(t) = 1_{[f(t) > f^*(b)]}(t) \).
3) If \( f^*(t) \) is neither strictly decreasing from the left nor to the right at \( t = b \), then the \( \beta(t) \) which take values in \([0, 1] \) almost everywhere which achieve this maximum are
\[
\beta(t) = 1_{[f(t) \leq f^*(b)]}(t) + 1_{Q}(t),
\]
where \( Q \subset \{ t : f(t) = f^*(b) \} \) with \( \mu(Q) = b - \mu(\{ t : f(t) > f^*(b) \}) \).

**Proof:** We prove the lemma in the case \( f^*(t) \) is strictly decreasing to the right at \( t = b \); the other cases are similar. Observe we only need to prove Eq. (3) with an inequality rather than equality, since by Proposition 5, the function \( \beta(t) = 1_{[f(t) \leq f^*(b)]} \) makes the left-hand side of Eq. (3) equal to the right-hand side.

Using Proposition 3 we have that for any \( \beta(t) \in B \),
\[
\int_0^a \beta(t) f(t) \, dt \leq \int_0^a \beta^*(t) f^*(t) \, dt.
\]

By Proposition 3 we have that \( \beta^* \) integrates to \( b \) over \([0, a] \) just like \( \beta \). By Proposition 4, we have that \( \beta^*(t) \leq 1 \). Moreover, since the rearrangement of a function is always nonnegative as an immediate consequence of Eq. (2), we also have that \( \beta^*(t) \) is nonnegative. Thus:
\[
\int_0^a \beta(t) f(t) \, dt \leq \max_{\gamma(t) \in [0,1], \int_0^a \gamma(t) \leq b} \int_0^a \gamma(t) f^*(t) \, dt.
\]

However, since \( f^*(t) \) is nonincreasing by Proposition 4, by Lemma 2 we must have that \( \gamma_{\text{opt}}(t) = 1_{[0,b]}(t) \).

Thus
\[
\int_0^a \beta(t) f(t) \, dt \leq \int_0^a 1_{[0,b]}(t) f^*(t) \, dt.
\]

Since this holds for all \( \beta \in B \), we have proved Eq. (3).

It remains to characterize what \( \beta(t) \) achieve equality in Eq. (3). For this, all the inequalities in the proof we’ve just given must be satisfied with equality. We next go through several of these inequalities and discuss how having equality in them constrains \( \beta(t) \).

First, by Lemma 2 the optimal \( \gamma(t) \) is unique and equals \( 1_{[0,b]}(t) \), so we must have \( \beta^*(t) = 1_{[0,b]}(t) \). In particular, this implies that \( \beta(t) \) is the indicator function of a set of Lebesgue measure \( b \). Let \( S \) denote that set.
Second, applying the equality conditions of Proposition 3 to Eq. (4), the set $S$ is such that for almost all $s, t$,
\[ \mu(\{x : f(x) \geq t\} \cap \{x : l_s(x) \geq s\}) \]
and
\[ \min(\mu(\{x : f(x) \geq t\}), \mu(\{x : l_s(x) \geq s\})) \]
are the same. But this implies that
\[ \mu(\{x : f(x) \geq t\} \cap S) = \min(\mu(\{x : f(x) \geq t\}), \mu(S)) \]
for almost all $t$. The last statement implies that, for almost all $t$, the level set $\{x : f(x) \geq t\}$ either contains or is contained in $S$ (up to a set of zero measure).

It is tempting to argue that since $\mu(S) = b$ and $\mu(\{x : f(x) \geq f^*(b)\}) = b$ (because $f^*(x)$ is decreasing to the right at $x = b$), and one of these two sets contains the other, they must be equal (again up to a set of measure zero). Unfortunately, it might be that the choice $t = f^*(b)$ is not included in the “almost all” $t$ above. Thus we proceed as follows. We consider the level sets $L_\epsilon = \{x : f(x) \geq f^*(b) - \epsilon\}$ with $\epsilon > 0$ which satisfy $\mu(L_\epsilon) \geq b$. Since the measure of $S$ is exactly $b$, it follows that $S \subseteq L_\epsilon$, up to a set of measure zero, almost all $\epsilon$. Since the sets $L_\epsilon$ are nested, $S$ is in fact contained in every $L_\epsilon$ for $\epsilon > 0$, and $S$ is contained in the intersection of $\{L_\epsilon, \epsilon > 0\}$, which is $\{x : f(x) \geq f^*(b)\}$. Since, as mentioned above, the last set has the same measure as $S$, we conclude $S$ equals $S$ up to a set of measure zero.

With this last lemma in place, we can now turn to the proof of our main result.

Proof of Theorem 2: We first argue that the solutions we have proposed are feasible. First, suppose that $F^*(x)$ is strictly decreasing to the right at $x = \alpha$. Then, using Proposition 3, we have
\[ \sum_{i=1}^m \int_0^T |v^*_i(t)| dt = \sum_{i=1}^m \mu(\{\tau : f_i(\tau) \geq F^*(\alpha)\}) \]
\[ = \mu(\{\tau : F(\tau) \geq F^*(\alpha)\}) \]
\[ = \mu(\{\tau : F^*(\tau) \geq F^*(\alpha)\}) \]
\[ = \alpha, \]
where the last step follows from the assumption that $F^*$ is decreasing to the right at $\alpha$. The case when $F^*(x)$ is strictly decreasing from the right at $x = \alpha$ follows by an almost identical argument.

Next, suppose $F^*(x)$ is not strictly decreasing either from the left or to the right at $x = \alpha$. Then
\[ \sum_{i=1}^m \int_0^T |v^*_i(t)| dt = \sum_{i=1}^m \mu(\{\tau : f_i(\tau) > F^*(\alpha)\}) \]
\[ + \mu(S) \]
\[ = \mu(\{\tau : F^*(\tau) > F^*(\alpha)\}) \]
\[ + (\alpha - \mu(\{\tau : F^*(\tau) > F^*(\alpha)\})) \]
\[ = \alpha, \]
where we relied on Proposition 3 in the second step. To summarize, we have shown that under the appropriate assumptions, all three of $v^*_i(t)$, $v_{i, l}(t)$ and $v^*_i, nlr(t)$ achieve the cost of $f^*_0 F^*(t)$.

We next show that, for any choice of functions $v_i(t)$, we will attain a cost that is upper bounded by $f^*_0 F^*(t)$.

Let us adopt the notation $J_{\text{opt}, t}$ for the cost corresponding to the functions $v^*_i(t)$. Let us compute this cost under the assumption that $F^*(t)$ is decreasing from the right at $t = \alpha$. We have that
\[ J_{\text{opt}, t} = \text{Tr} \int_0^T \sum_{i=1}^m e^{A t} b_i \mathbb{1}_{\{f_i(\tau) \geq F^*(\alpha)\}}(t) b_i^T e^{A^T t} dt \]
\[ = \int_0^T \sum_{i=1}^m 1_{\{f_i(\tau) \geq F^*(\alpha)\}}(t) b_i^T e^{A^T t} e^{A t} b_i dt \]
\[ = \int_0^T \sum_{i=1}^m 1_{\{f_i(\tau) \geq F^*(\alpha)\}}(t) f_i(t) dt \]
\[ = \int_0^\alpha F^*(t) dt, \]
and the last step used Proposition 5 and the assumption that $F^*(t)$ is strictly decreasing to the right at $t = \alpha$. Thus we have shown that the choice of functions $v^*_i(t)$ achieves a cost of $f^*_0 F^*(t) dt$. The case when $F^*(t)$ is strictly decreasing from the left at $t = \alpha$ follows by an almost identical argument.

We next argue that, under the assumption that $F^*(t)$ is decreasing neither from the left nor the right, the functions $v^*_i, nlr$ achieve the same cost. Indeed, let $(\alpha', \alpha')$ be the largest open interval containing $\alpha$ on which $F^*(t)$ is non-decreasing whose existence was guaranteed by Lemma 1. Note that lemma tells us that $\mu(\{\tau : F^*(\tau) > F^*(\alpha)\}) = \alpha'$.

Define $S' = S_i + (i-1)T$, where $S_i$ comes from the theorem statement and we translate it by $(i-1)T$ to make it so that $S' \subset ((i-1)T, iT]$. Further defining $S = \cup_{i=1}^m S'_i$ and proceeding similarly as before,
\[ J_{\text{opt}, \text{nlr}} = \int_0^{mT} F(t) \mathbb{1}_{\{f_i(\tau) > F^*(\alpha)\}}(t) + \mathbb{1}_S(t) dt \]
\[ = \int_0^\alpha F^*(t) dt + \mu(S) F^*(\alpha) \]
\[ = \int_0^\alpha F^*(t) dt + (\alpha - \mu(F^*(\tau) > F^*(\alpha))) F^*(\alpha) \]
\[ = \int_0^\alpha F^*(t) dt + (\alpha - \alpha') F^*(\alpha) \]
\[ = \int_0^\alpha F^*(t) dt, \]
where we have that $\mu(\{\tau : F^*(\tau) > F^*(\alpha)\}) = \alpha'$.
Indeed, for any feasible functions \( v_i(t) \) we have that \( v_i(t) \in [0, 1] \) which means that

\[
J_r = \text{Tr} \int_0^T \sum_{i=1}^m v_i(t) e^{At} b_i b_i^T e^{AT} dt 
\]

\[
\leq \text{Tr} \int_0^T \sum_{i=1}^m v_i(t) e^{At} b_i b_i^T e^{AT} dt
\]

\[
= \sum_{i=1}^m \int_0^T v_i(t) f_i(t) dt \int_0^T q(t) F(t) dt
\]

where we define

\[
q(t) = \sum_{i=1}^m v_i(t - (i - 1)T) \mathbb{1}_{[t-(i-1)T,T]}(t)
\]

Observing that

\[
\int_0^{mT} |q(t)| dt = \int_0^T \sum_{i=1}^m |v_i(t)| dt \leq \alpha,
\]

by feasibility of \( v_i(t) \), we can now apply Lemma 3 to Eq. (8) and, by Eq. (5), obtain \( J_r \leq \int_0^T F^*(t) dt \).

Putting it all together, we have thus shown that, under the appropriate assumptions, each of \( v_{opt,r}(t), v_{opt,s}(t), v_{opt,nr}(t) \) are optimal. It remains to prove that these are the only optimal choices. For this, we must analyze the cases of equality in the above bounds.

Observe that to achieve equality we need to have equality starting from the Eq. (6) through the end of the proof. In particular, we must have equality in the application of Lemma 3. But that lemma spells out conditions for equality. In particular, Lemma 3 forces \( q(t) = \mathbb{1}_{\{t:F(t) \geq F^*(t)\}}(t) \) when \( F^*(t) \) is decreasing from the right at \( \alpha \) and \( q(t) = \mathbb{1}_{\{t:F(t) > F^*(t)\}}(t) \) when it is decreasing from the left. By definition of \( q(t) \), this is the same as having \( v_i(t) = \mathbb{1}_{\{t:F(t) \geq F^*(t)\}}(t) \) and \( v_i(t) = 1 - \mathbb{1}_{\{t:F(t) > F^*(t)\}}(t) \). This concludes the proof for the case where \( F^*(t) \) is either strictly decreasing to the right or from the left at \( \alpha \).

It only remains to analyze the cases of equality in the case where \( F^*(t) \) is not strictly decreasing either from the left or to the right at \( t = \alpha \). First, observe that because \( B \) has no zero columns and \( e^{AT} \) is always nonsingular, we have that \( \text{Tr}(e^{AT} b_i b_i^T e^{AT}) > 0 \), for all \( i \in \{1, \ldots, m\} \) and \( t \in [0, T] \). In particular, because \( v_i(t) \in [0, 1] \), the implication of this is that to achieve equality going from Eq. (6) to Eq. (7), we must have that each \( v_i(t) \in \{0, 1\} \) almost everywhere. This implies the function \( q(t) \) must be binary as well.

Having established that, we apply the conditions for equality in the last item of Lemma 3. That lemma tells us that we must have \( q(t) = \mathbb{1}_{\{t:F(t) \geq F^*(t)\}}(t) + 1_S \) for a subset \( S \) of the set \( \{t:F(t) = F^*(t)\} \) with \( \mu(S) = \alpha - \mu(\{t:F(t) > F^*(t)\}) \). This concludes the proof.

\[\text{REFERENCES}\]

[1] M. Athans, “On the determination of optimal costly measurement strategies for linear stochastic systems,” *Automatica*, vol. 8, no. 4, pp. 397–412, 1972.

[2] G. Baggio, S. Zampieri, and C. W. Scherer, “Gramian optimization with input-power constraints,” in *Proc. IEEE 58th Conf. Decis. Control (CDC)*, 2019, pp. 5686–5691.

[3] N. Bof, G. Baggio, and S. Zampieri, “On the role of network centrality in the controllability of complex networks,” *IEEE Trans. Control Netw. Syst.*, vol. 4, no. 3, pp. 643–653, Sep. 2017.

[4] S. D. Bopardikar, “Sensor selection via randomized sampling,” 2017. [Online]. Available: arXiv:1712.06511.

[5] S. Boyd, “Introduction to linear dynamical systems,” Lecture Notes EEE263, Stanford University, Stanford, CA, USA, 2008.

[6] P. V. Chanekar, N. Chopra, and S. Azarm, “Optimal actuator placement for linear systems with limited number of actuators,” in *Proc. Amer. Control Conf. (ACC)*, 2017, pp. 334–339.

[7] A. S. A. Dilip, “The controllability Gramian, the hadamard product, and the optimal actuator/leader and sensor selection problem,” *IEEE Control Syst. Lett.*, vol. 3, no. 4, pp. 883–888, Oct. 2019.

[8] A. Hashemi, M. Ghasemi, H. Vikalo, and U. Topcu, “A randomized greedy algorithm for near-optimal sensor scheduling in large-scale sensor networks,” in *Proc. Amer. Control Conf. (ACC)*, 2018, pp. 1027–1032.

[9] T. Ikeda and K. Kashima, “On sparse optimal control for general linear systems,” *IEEE Trans. Autom. Control*, vol. 64, no. 5, pp. 2077–2083, May 2019.

[10] T. Ikeda and K. Kashima, “Sparsity-constrained controllability maximization with application to time-varying control node selection,” *IEEE Control Syst. Lett.*, vol. 2, no. 3, pp. 321–326, Jul. 2018.

[11] T. Ikeda and K. Kashima, “Sparse optimal feedback control for continuous-time systems,” in *Proc. 18th Eur. Control Conf. (ECC)*, 2019, pp. 3728–3733.

[12] A. Jadbabaie, A. Olshevsky, and M. Siami, “Deterministic and randomized actuator scheduling with guaranteed performance bounds,” 2018. [Online]. Available: arXiv:1805.00606.

[13] E. H. Lieb and M. Loss, “Analysis,” in *Graduate Studies in Mathematics*, vol. 14. Providence, RI, USA: American Mathematical Society, 2001.

[14] E. Nozari, F. Pasqualetti, and J. Cortés, “Time-invariant versus time-varying actuator scheduling in complex networks,” in *Proc. Amer. Control Conf. (ACC)*, 2017, pp. 4995–5000.

[15] E. Nozari, F. Pasqualetti, and J. Cortés, “Heterogeneity of central nodes explains the benefits of time-varying control scheduling in complex dynamical networks,” *J. Complex Netw.*, vol. 7, no. 5, pp. 659–701, 2019.

[16] A. Olshevsky, “Minimal controllability problems,” *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 3, pp. 249–258, Sep. 2014.

[17] A. Olshevsky, “Eigenvalue clustering, control energy, and logarithmic capacity,” *Syst. Control Lett.*, vol. 96, pp. 45–50, Oct. 2016.

[18] A. Olshevsky, “On a relaxation of time-varying actuator placement,” 2019. [Online]. Available: arXiv:1912.00454.

[19] F. Pasqualetti, S. Zampieri, and F. Bullo, “Controllability metrics, limitations and algorithms for complex networks,” *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 1, pp. 40–52, Mar. 2014.

[20] M. Siami and A. Jadbabaie, “Deterministic polynomial-time actuator scheduling with guaranteed performance,” in *Proc. Eur. Control Conf. (ECC)*, 2018, pp. 113–118.

[21] T. Summers and M. Kamgarpour, “Performance guarantees for greedy maximization of non-submodular controllability metrics,” in *Proc. 18th Eur. Control Conf. (ECC)*, 2019, pp. 2796–2801.

[22] T. Summers and I. Shames, “Convex relaxations and gramian rank constraints for sensor and actuator selection in networks,” in *Proc. Int. Symp. Intell. Control*, 2016, pp. 1–6.

[23] V. Tsioumas, M. A. Rahimian, G. J. Pappas, and A. Jadbabaie, “Minimal actuator placement with bounds on control effort,” *IEEE Trans. Control Netw. Syst.*, vol. 3, no. 1, pp. 67–78, Mar. 2016.

[24] A. Zare, H. Mohammad, N. K. Dhingra, T. T. Georgiou, and M. R. Jovanovic, “Proximal algorithms for large-scale statistical modeling and sensor/actuator selection,” *IEEE Trans. Autom. Control*, early access, Oct. 18, 2019, doi: 10.1109/TAC.2019.2948268.

[25] H. Zhang, R. Ayoub, and S. Sundaram, “Sensor selection for Kalman filtering of linear dynamical systems: Complexity, limitations and greedy algorithms,” *Automatica*, vol. 78, pp. 202–210, Apr. 2017.

[26] Y. Zhao and J. Cortés, “Gramian-based reachability metrics for bilinear networks,” *IEEE Trans. Control Netw. Syst.*, vol. 4, no. 3, pp. 620–631, Sep. 2017.

[27] Y. Zhao, F. Pasqualetti, and J. Cortés, “Scheduling of control nodes for improved network controllability,” in *Proc. IEEE 55th Conf. Decis. Control (CDC)*, 2016, pp. 1859–1864.