

Suboptimal coverings for continuous spaces of control tasks

James A. Preiss

JAPREISS@USC.EDU and Gaurav S. Sukhatme

GAURAV@USC.EDU

University of Southern California, Los Angeles, USA *

Abstract

We propose the α -suboptimal covering number to characterize multi-task control problems where the set of dynamical systems and/or cost functions is infinite, analogous to the cardinality of finite task sets. This notion may help quantify the function class expressiveness needed to represent a good multi-task policy, which is important for learning-based control methods that use parameterized function approximation. We study suboptimal covering numbers for linear dynamical systems with quadratic cost (LQR problems) and construct a class of multi-task LQR problems amenable to analysis. For the scalar case, we show logarithmic dependence on the “breadth” of the space. For the matrix case, we present experiments 1) measuring the efficiency of a particular constructive cover, and 2) visualizing the behavior of two candidate systems for the lower bound. ¹

1. Introduction

An advanced control system such as a mobile robot may be required to perform many different tasks. If the task set is finite, like selecting between “map an environment” and “deliver a package”, then its size is naturally quantified by the number of tasks. If the task set is infinite, like delivering packages with arbitrary mass and inertial properties, then its size is not so easily quantified. Even if the task space is equipped with a metric or measure, these structures may be only weakly linked to the diversity of behavior required for good performance on all tasks.

Our interest in this issue is motivated by multi-task paradigms in learning-based control, where the policy is selected from a parameterized family of functions that map state and task parameters directly to actions. As the task space expands from a singleton set, we expect to need a more expressive class of functions to represent a good multi-task policy. In this work, we propose the α -suboptimal covering number to capture this idea. For a task space Φ and a suboptimality ratio $\alpha > 1$, we define $N_\alpha(\Phi)$ as the size of the smallest set of single-task policies \mathcal{C} such that for every $\phi \in \Phi$, at least one $\pi \in \mathcal{C}$ has a cost ratio no greater than α relative to the optimal policy for ϕ . If the policies in \mathcal{C} are parameterized functions, then \mathcal{C} provides an upper bound on the number of parameters needed to represent an α -suboptimal multi-task policy. In switching-based adaptive control, where ϕ is unknown, a smaller \mathcal{C} implies a faster convergence time.

To study suboptimal covering numbers in a concrete setting, we consider linear dynamical systems with quadratic cost functions, or LQR problems. LQR problems are a common setting to analyze learning algorithms because detailed properties are known (Fazel et al., 2018, for example). This has led to new inquiries into their fundamental properties (Bu et al., 2019). We construct a family of well-behaved multi-task LQR problems where Φ is controlled by a “breadth” parameter

* Gaurav S. Sukhatme holds concurrent appointments as a Professor at USC and as an Amazon Scholar. This paper describes work performed at USC and is not associated with Amazon.

1. An extended version of this paper with source code is available at <https://arxiv.org/abs/2104.11865>.

$\theta \in [1, \infty)$, and for which $N_\alpha(\Phi_\theta)$ is finite and increasing in θ . For the special case of a scalar LQR problem, we derive matching logarithmic upper and lower bounds on $N_\alpha(\Phi_\theta)$ as a function of θ . As an effort towards analogous bounds for the matrix case, we present empirical results intended to shed light on the problem structure. For the upper bound, we analyze properties of a logical extension of our scalar cover. For the lower bound, we visualize suboptimal neighborhoods for two choices of “extremal” systems, revealing surprising topological behavior for one choice.

This paper is an initial step towards a comprehensive theory. In addition to a more complete picture of deterministic LQR systems, ideas of α -suboptimal coverings could be applied to a wide range of multi-task problems. We also hope they will lead to insights about function class expressiveness in learning-based multi-task control.

2. Problem setting

In this section we define α -suboptimal covering numbers with respect to an abstract multi-task control problem, independent of distinctions such as continuous vs. discrete time and stochastic vs. deterministic dynamics. We then instantiate these definitions for a particular class of LQR problems.

Notation The set of all functions $\mathcal{X} \mapsto \mathcal{Y}$ is denoted by $\mathcal{Y}^\mathcal{X}$. The relation $A \succeq B$ (resp. $A \succ B$) denotes that $A - B$ is positive semidefinite (resp. definite). Matrices of zeros and ones, with dimension implied by context, are denoted by $\mathbf{0}$ and $\mathbf{1}$. The integers $\{1, \dots, N\}$ are denoted by $[N]$.

Definition 1 A multi-task optimal control problem is defined by an arbitrary state space \mathcal{X} , action space \mathcal{U} , and task space Φ ; a class of reference policies $\Pi_{\text{ref}} \subseteq \mathcal{U}^\mathcal{X}$, and a strictly positive objective function $J : \Phi \times \mathcal{U}^\mathcal{X} \mapsto \mathbb{R}_{>0}$. The partial application of J for $\phi \in \Phi$ is denoted by $J_\phi : \mathcal{U}^\mathcal{X} \mapsto \mathbb{R}$. The optimal reference cost for an task is denoted by $J_\phi^* = \inf_{\pi \in \Pi_{\text{ref}}} J_\phi(\pi)$.

Definition 2 Consider a multi-task optimal control problem $(\mathcal{X}, \mathcal{U}, \Phi, \Pi_{\text{ref}})$ and a suboptimality ratio $\alpha > 1$. Given a policy $\pi : \mathcal{X} \mapsto \mathcal{U}$, we define its α -suboptimal neighborhood as $\mathcal{N}_\alpha(\pi) = \left\{ \phi \in \Phi : J_\phi(\pi) / J_\phi^* \leq \alpha \right\}$. A set of policies $\mathcal{C} \subseteq \mathcal{U}^\mathcal{X}$ is an α -suboptimal cover of Φ if $\bigcup_{\pi \in \mathcal{C}} \mathcal{N}_\alpha(\pi) = \Phi$. The α -suboptimal covering number of Φ , denoted $N_\alpha(\Phi)$, is the size of the smallest finite α -suboptimal cover of Φ if one exists, or ∞ otherwise.

Standard LQR problem A continuous-time, deterministic, infinite-horizon, time-invariant LQR problem with full-state feedback is defined by state space $\mathcal{X} = \mathbb{R}^n$, action space $\mathcal{U} = \mathbb{R}^m$, linear dynamics $\dot{x} = Ax + Bu$, where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, and quadratic cost

$$J(\pi) = \mathbb{E}_{x(0) \sim \mathcal{N}(\mathbf{0}, I)} \int_0^\infty (x^\top Qx + u^\top Ru) dt,$$

where $Q \succeq \mathbf{0}$ and $R \succ \mathbf{0}$ are cost matrices of appropriate dimensions and $\mathcal{N}(\mathbf{0}, I)$ is the unit Gaussian distribution. For the purposes of this paper, the pair (A, B) is *controllable* if $J(\pi) < \infty$ for some policy π . If (A, B) is controllable, then the optimal policy is the linear $u = K^*x$, where $K^* \in \mathbb{R}^{m \times n}$ can be computed by finding the unique maximal positive semidefinite solution P of the algebraic Riccati equation $A^\top P + PA - PBB^{-1}B^\top P + Q = \mathbf{0}$ (henceforth called the *maximal solution*) and letting $K^* = -R^{-1}B^\top P$ (Kalman, 1960). Additionally, $J(K^*) = \text{Tr}[P]$. An arbitrary controller $K \in \mathbb{R}^{m \times n}$ is *stabilizing* if $J(K) < \infty$, in which case $J(K)$ satisfies

$$J(K) = \text{Tr} \left[(Q + K^\top RK)W \right], \text{ where } W = \int_0^\infty e^{t(A+BK)^\top} e^{t(A+BK)} dt. \quad (1)$$

W can be computed by solving the Lyapunov equation $(A + BK)^\top W + W(A + BK) + I = \mathbf{0}$.

Multi-dynamics LQR A fully general formulation of multi-task LQR would allow variations in each of (A, B, Q, R) , but this creates redundancy. Any LQR problem where $Q \succ 0$ is equivalent via change of coordinates to another LQR problem where $Q = I$ and $R = I$. To reduce redundancy, we consider only *multi-dynamics* LQR problems where $Q = I_{n \times n}$ and $R = I_{m \times m}$ in this work. The reference policy class is linear: $\Pi_{\text{ref}} = \mathbb{R}^{m \times n}$.

A multi-dynamics LQR problem can be defined by $\Phi = \mathbf{A} \times \mathbf{B}$ for some sets $\mathbf{A} \subseteq \mathbb{R}^{n \times n}$ and $\mathbf{B} \subseteq \mathbb{R}^{n \times m}$, but it is not obvious how to design \mathbf{A} and \mathbf{B} . To support an asymptotic analysis of $N_\alpha(\Phi)$, the task space Φ should have a real-valued ‘‘breadth’’ parameter θ that sweeps from a single task to sets with arbitrarily large, but finite, covering numbers. Matrix norm balls are a popular representation of dynamics uncertainty in the robust control literature, but they can easily contain uncontrollable pairs, and removing the uncontrollable pairs can lead to an infinite covering number. For example, in the scalar problem $\mathbf{A} = \{a\}$, $\mathbf{B} = [-\theta, 0) \cup (0, \theta]$, where $a \neq 0$, it can be shown that no α -suboptimal cover is finite.

These properties are worrying, but the example \mathbf{B} is pathological. The zero crossing is analogous to reversing the direction of force applied by an actuator in a physical system. A more relevant multi-dynamics problem is variations in mass or actuator strength, whose signs are fixed. We formalize this idea with the following definition.

Definition 3 Fix $A \in \mathbb{R}^{n \times n}$ and a breadth parameter $\theta \geq 1$. Let $\mathbf{A} = \{A\}$ and $\mathbf{B} = \{U\Sigma V^\top : \Sigma \in \Sigma\}$, where $\Sigma = \{\text{diag}(\sigma) : \sigma \in [\frac{1}{\theta}, 1]^d\}$. The matrices $U \in \mathbb{R}^{n \times d}$ and $V \in \mathbb{R}^{m \times d}$ each have rank d , where $0 < d \leq \min\{n, m\}$. The tuple (A, U, V, θ) fully defines a multi-task LQR problem in decomposed dynamics form, or DDF problem for brevity.

We will abuse notation and associate Φ with both $\mathbf{A} \times \mathbf{B}$ and Σ when the meaning is clear from context. The continuity of the LQR cost (1) with respect to B and the compactness of Φ for any θ imply that $N_\alpha(\Phi_\theta)$ is always finite. Variations in A are redundant in the scalar case where we focus our theoretical work in this paper. The definition can be extended to include them in future work.

Linearized quadrotor example As an example of a realistic DDF problem, we consider the quadrotor helicopter illustrated in Figure 1. Near the hover state, its full nonlinear dynamics are well approximated by a linearization. The state is given by $x = (\mathbf{x}, \mathbf{v}, \mathbf{r}, \boldsymbol{\omega})$, where $\mathbf{x} \in \mathbb{R}^3$ is position, $\mathbf{v} \in \mathbb{R}^3$ is linear velocity, $\mathbf{r} \in \mathbb{R}^3$ is attitude Euler angles, and $\boldsymbol{\omega} \in \mathbb{R}^3$ is angular velocity. The inputs $u \in \mathbb{R}_{\geq 0}^4$ are the squared angular velocities of the propellers.

Many factors influence the response to inputs, including geometry, mass, moments of inertia, motor properties, and propeller aerodynamics. These can be combined and partially nondimensionalized into four control authority parameters to form $\phi \in \Phi$. The hover state occurs at $x = \mathbf{0}$, $u \propto \mathbf{1}$, where the constant input counteracts gravity. The linearized dynamics are given by

$$\dot{x} = \underbrace{\begin{bmatrix} 0 & I & 0 & 0 \\ 0 & 0 & G & 0 \\ 0 & 0 & 0 & I \\ 0 & 0 & 0 & 0 \end{bmatrix}}_A x + \underbrace{\begin{bmatrix} 0 & 0 \\ \hat{e}_z & 0 \\ 0 & 0 \\ 0 & I \end{bmatrix}}_U \underbrace{\begin{bmatrix} \sigma_z & & & \\ & \sigma_\phi & & \\ & & \sigma_\theta & \\ & & & \sigma_\psi \end{bmatrix}}_\Sigma \underbrace{\begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}}_{V^\top} u, \quad G = \begin{bmatrix} 0 & g & 0 \\ -g & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where g is the gravitational constant and $\hat{e}_z = [0 \ 0 \ 1]^\top$. The parameters $(\sigma_z, \sigma_\phi, \sigma_\theta, \sigma_\psi)$ denote the thrust, roll, pitch, and yaw authority constants respectively. Since we use the convention $\sigma \in [\frac{1}{\theta}, 1]$, the maximum value of each constant can be varied by scaling the columns of U .

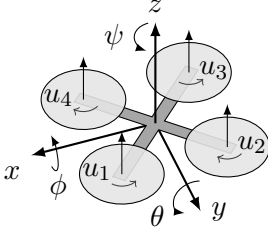


Figure 1: Quadrotor helicopter with position states x, y, z , attitude states ϕ, θ, ψ , and propeller speed inputs u_1, u_2, u_3, u_4 . The linearized dynamics at hover, subject to variations in mass, geometry, etc., can be expressed in decomposed dynamics form—see Section 2.

3. Theoretical results

In this section we show logarithmic upper and lower bounds on the growth of $N_\alpha(\Phi_\theta)$ in θ for scalar DDF problems. We present several intermediate results in matrix form because they are needed for our empirical results later. We begin with a key lemma in the framework of *guaranteed cost control* (GCC) from [Petersen and McFarlane \(1994\)](#), simplified for our use case.

Lemma 4 (GCC synthesis, [Petersen and McFarlane \(1994\)](#)) *Given the multi-task LQR problem $\mathbf{A} = \{A\}$, $\mathbf{B} = \{B_1\Delta + B_2 : \|\Delta\| \leq 1\}$, where $B_1, B_2 \in \mathbb{R}^{m \times p}$ are arbitrary for arbitrary p , and the state cost matrix is $Q \succ \mathbf{0}$, if there exists $\tau > 0$ such that $P \succ \mathbf{0}$ solves the Riccati equation*

$$A^\top P + PA + P \left(\frac{1}{\tau} B_1 B_1^\top - \frac{1}{1+\tau} B_2 B_2^\top \right) P + Q = \mathbf{0}, \quad (2)$$

then the controller $K = -\frac{1}{1+\tau} B_2^\top P$ has cost $J_B(K) \leq \text{Tr}[P]$ for all $B \in \mathbf{B}$. Also, $\text{Tr}[P]$ is a convex function of τ .

We use the notation $P, \tau, K = \text{GCC}(A, B_1, B_2, Q)$ to indicate that P, τ solve (2) and K is the corresponding controller. It is straightforward to show that any DDF problem can be expressed in the form required by Lemma 4 with additional constraints on Δ .

In the original presentation, [Petersen and McFarlane \(1994\)](#) treat B_1 as given, so they accept that (2) may have no solution (for example, when $A = 2I, B_1 = I, B_2 = \mathbf{0}$). Our application requires constructing values of B_1, B_2 that guarantee a solution, motivating the following lemma. We abbreviate the reference text [Lancaster and Rodman \(1995\)](#) as [Lan95](#).

Lemma 5 (existence of α -suboptimal GCC) *For the DDF problem (A, U, V, θ) , if $B \in \mathbf{B}$ and $\alpha > 1$, then there exists $B_1 \neq \mathbf{0} \in \mathbb{R}^{m \times n}$ such that the GCC Riccati equation (2) with $B_2 = B$ has a solution (P, τ) satisfying $\text{Tr}[P] \leq \alpha J_B^*$.*

Proof For this proof, it will be more convenient to write the algebraic Riccati equation as

$$A^\top P + PA - PDP + Q = \mathbf{0}, \quad (3)$$

where $D \succeq \mathbf{0}$. Let $\mathcal{D} = \{D \succeq \mathbf{0} : (A, D) \text{ is controllable}\}$. Controllability of (A, B) implies that $BB^\top \in \mathcal{D}$ ([Lan95](#), Corollary 4.1.3). Let Ric_+ denote the map from \mathcal{D} to the maximal solution of (3), which is continuous ([Lan95](#), Theorem 11.2.1), and let $\mathcal{D}_\alpha = \{D \in \mathcal{D} : \text{Tr}[\text{Ric}_+(D)] < \alpha J_B^*\}$. The set \mathcal{D}_α is open in \mathcal{D} by continuity and is nonempty because it contains BB^\top . Now define $B_1(\tau) = \tau B$ for $\tau \in (0, \frac{1}{2})$. The equivalent of D in the GCC Riccati equation (2) becomes

$$D(\tau) = -\frac{1}{\tau} B_1(\tau) B_1(\tau)^\top + \frac{1}{1+\tau} B_2 B_2^\top = \frac{1-\tau-\tau^2}{1+\tau} BB^\top.$$

As a positive multiple of BB^\top , we know $D(\tau) \in \mathcal{D}$, and because $\lim_{\tau \rightarrow 0} D(\tau) = BB^\top$, the set of τ for which $D(\tau) \in \mathcal{D}_\alpha$ is nonempty. Any such τ and $B_1(\tau)$ provide a solution. \blacksquare

Finally, the following comparison result will be useful in several places.

Lemma 6 (Lan95, Corollary 9.1.6) *Given two algebraic Riccati equations*

$$A^\top P + PA - PBB^\top P + Q = \mathbf{0} \quad \text{and} \quad \tilde{A}^\top P + P\tilde{A} - P\tilde{B}\tilde{B}^\top P + \tilde{Q} = \mathbf{0},$$

with maximal solutions P and \tilde{P} , let $X = \begin{bmatrix} Q & A^\top \\ A & -BB^\top \end{bmatrix}$ and $\tilde{X} = \begin{bmatrix} \tilde{Q} & \tilde{A}^\top \\ \tilde{A} & -\tilde{B}\tilde{B}^\top \end{bmatrix}$. If $X \succeq \tilde{X}$, then $P \succeq \tilde{P}$.

3.1. Scalar upper bound

We are now ready to bound the covering number for scalar systems. The first lemma bounding $J_{a,b}^*$ will be useful for the lower bound also. We then construct a cover inductively.

Lemma 7 *In a scalar LQR problem, if $a > 0$ and $0 < b \leq 1$, then the optimal scalar LQR cost satisfies the bounds $2a/b^2 < J_{a,b}^* < (2a+1)/b^2$.*

Proof The lower bound is visible from the closed-form solution for the scalar Riccati equation, which is $J_{a,b}^* = \frac{a + \sqrt{a^2 + b^2}}{b^2}$. The upper bound is obtained by substituting $a^2 + b^2 \leq (a+1)^2$. \blacksquare

Lemma 8 *If $p, \tau, k = \text{GCC}(a, b_1, b_2, q)$, then for any $\beta \in (0, 1)$, there exists $k' \in \mathbb{R}$ such that $p', \tau, k' = \text{GCC}(a, \beta b_1, \beta b_2, \beta^{-2}q)$, where $p' = \beta^{-2}p$.*

Proof In the scalar system, the GCC matrix Riccati equation (2) reduces to the quadratic equation

$$\left(\frac{1}{\tau} b_1^2 - \frac{1}{1+\tau} b_2^2 \right) p^2 + 2ap + q = 0. \quad (4)$$

Substituting $p' = \beta^{-2}p$ into (4) and multiplying by β^{-2} yields a new instance of (4) with the parameters $b'_1 = \beta b_1$, $b'_2 = \beta b_2$, $q' = \beta^{-2}q$, for which p' is a solution with τ unchanged. \blacksquare

Theorem 9 *For the scalar DDF problem defined by $\mathbf{A} = \{a\}$, where $a > 0$, and $\mathbf{B} = [\frac{1}{\theta}, 1]$, if $\alpha \geq \frac{2a+1}{2a}$, then $N_\alpha(\mathbf{B}) = O(\log \theta)$.*

Proof We construct a cover from the upper end of \mathbf{B} . By Lemma 7, the condition $\alpha \geq \frac{2a+1}{2a}$ implies that $J_{b=1}^* < \alpha 2a < \alpha J_{b=1}^*$. Therefore, by Lemmas 4 and 5, there exists $\beta \in (0, 1)$ and p, τ, k such that $p, \tau, k = \text{GCC}(a, (1-\beta)/2, (1+\beta)/2, 1)$ and $p \leq \alpha 2a$.

Proceeding inductively, suppose that for $N \geq 1$, we have covered $[\beta^N, 1]$ by the intervals $\mathbf{B}_n = [\beta^{n+1}, \beta^n]$ for $n \in \{0, \dots, N-1\}$, and each \mathbf{B}_n has a controller k_n such that

$$\beta^{-2n} p, \tau, k_n = \text{GCC}(a, (\beta^n - \beta^{n+1})/2, (\beta^n + \beta^{n+1})/2, \beta^{-2n}).$$

Then the existence of the desired \mathbf{B}_N, k_N follows immediately from Lemma 8.

By Lemma 6, for each \mathbf{B}_n the GCC state cost $q_n = \beta^{-2n} \geq 1$ is an upper bound on the cost if we replace q_n with 1 to match the DDF problem. Therefore, for each interval \mathbf{B}_n , for all $b \in \mathbf{B}_n$,

$$\alpha J_b^* \geq \alpha J_{\beta^n}^* > \beta^{-2n} \alpha 2a \geq \beta^{-2n} p \geq J_b(k_n),$$

where first inequality is due to Lemma 6, the second is due to Lemma 7, the third is by construction of p , and last is due to the GCC guarantee of k_n . Hence, $\mathbf{B}_n \subseteq \mathcal{N}_\alpha(k_n)$. We cover the full \mathbf{B} when $\beta^N \leq \frac{1}{\theta}$, which is satisfied by $N \geq -\log \theta / \log \beta$. \blacksquare

3.2. Scalar lower bound

For the matching lower bound, we begin by deriving a simplified overestimate of $\mathcal{N}_\alpha(k)$. We then show that the true $\mathcal{N}_\alpha(k)$ is still a closed interval moving monotonically with k . Finally, we argue that the gaps between consecutive elements of a cover grow at most geometrically, while the range of k values in a cover must grow linearly with θ .

Lemma 10 *For a scalar DDF problem with $a \geq 1$, $\mathbf{B} = [\frac{1}{\theta}, 1]$, for any $k < 0$, if $\alpha \geq 3/2$, then $\mathcal{N}_\alpha(k) \subseteq \frac{1}{|k|}[c_1 - c_2, c_1 + c_2]$, where c_1 and c_2 are constants depending on α and a .*

Proof Beginning with the closed-form solution for $J_b(k)$, which can be derived from (1), we define

$$J_b(k) = \frac{1 + k^2}{-2(a + bk)} \geq \frac{k^2}{-2(a + bk)} \triangleq \underline{J}_b(k). \quad (5)$$

By Lemma 7, we have $J_b^* < 3a/b^2 \triangleq \overline{J}_b^*$, so $\tilde{r} = \underline{J}_b(k)/\overline{J}_b^*$ is a lower bound on the suboptimality of k . Computing $\partial^2 \tilde{r}/\partial b^2$ shows that \tilde{r} is strictly convex in b on the domain $a + bk < 0$, so the α -sublevel set of \tilde{r} is the closed interval with boundaries where $\tilde{r} = \alpha$. This equation is quadratic in b with the solutions $b = -a(3\alpha \pm \sqrt{9\alpha^2 - 6\alpha})/k$. The resulting interval contains $\mathcal{N}_\alpha(k)$. ■

Lemma 11 *For a scalar DDF problem, if $\alpha > 1$ and $k < -1$, then $\mathcal{N}_\alpha(k)$ is either empty or a closed interval $[b_1, b_2]$, with b_1 and b_2 positive and nondecreasing in k .*

Proof The result follows from the quasiconvexity of both the suboptimality ratio $J_b(k)/J_b^*$ and the cost $J_b(k)$. Showing these requires some tedious calculations. For details, see the extended version of this paper. ■

Theorem 12 *For a scalar DDF problem defined by $a = 1$ and $\mathbf{B} = [\frac{1}{\theta}, 1]$, if $\alpha \geq 3/2$, then $N_\alpha(\mathbf{B}) = \Omega(\log \theta)$.*

Proof From the closed-form solution $k_{a,b}^* = -(a + \sqrt{a^2 + b^2})/b$, we observe that $k_b^* < -1$ for all $b \in \mathbf{B}$. This, along with the quasiconvexity of $J_b(k)$ in k , implies that there exists a minimal α -suboptimal cover \mathcal{C} for which all $k_i < -1$. Suppose $\mathcal{C} = k_1, \dots, k_N$ is such a cover, ordered such that $k_i < k_{i+1}$. Then by Lemma 11, $\mathcal{N}_\alpha(k_i)$ and $\mathcal{N}_\alpha(k_{i+1})$ must intersect, so their overestimates according to Lemma 10 certainly intersect, therefore satisfying

$$\frac{c_1 + c_2}{-k_{i+1}} \geq \frac{c_1 - c_2}{-k_i} \implies \frac{k_{i+1}}{k_i} \leq \frac{c_1 + c_2}{c_1 - c_2} \implies \frac{k_N}{k_1} \leq \left(\frac{c_1 + c_2}{c_1 - c_2} \right)^{N-1}.$$

By Lemma 10, to cover $b = 1$ controller k_1 must satisfy $k_1 \geq -(c_1 + c_2)$, and to cover $b = \frac{1}{\theta}$, controller k_N must satisfy $k_N \leq -\theta(c_1 - c_2)$. Along with the previous result, this implies

$$\left(\frac{c_1 + c_2}{c_1 - c_2} \right)^{N-1} \geq \theta \frac{c_1 - c_2}{c_1 + c_2} \implies N \geq \frac{\log \theta}{\log \frac{c_1 + c_2}{c_1 - c_2}}.$$

Recalling that c_1 and c_2 only depend on a and α , the $\Omega(\log \theta)$ dependence on θ is established. ■

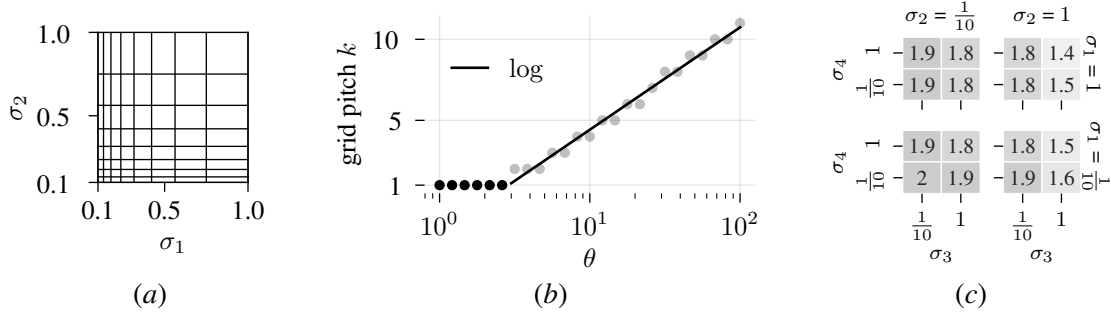


Figure 2: Application of geometric grid cover to linearized quadrotor. (a) Illustration of geometric grid partition. (b) Empirical upper bound on grid pitch k . (c) Suboptimality ratios for corner cells in empirical cover. Discussion in Section 4.

Remarks

- For the upper bound, it may be possible to compute or bound β in the scalar case as a function of a and α , but the analogous result will likely be harder to obtain in the matrix case.
- These results impose a lower bounds on α greater than 1. We believe this is a mild condition in control applications: if the application demands a suboptimality ratio very close to 1, then the size of the suboptimal cover is likely to become impractical for storage. However, further theoretical results building upon suboptimal coverings may require eliminating the bound.

4. Empirical results

For matrix DDF problems, we present empirical results as a first step towards covering number bounds. We begin by testing a cover construction. If the construction fails to achieve a conjectured upper bound in a numerical experiment, then either the conjecture is false, or the construction is not efficient. A natural idea is to extend the geometrically spaced sequence of b values from Lemma 7 to multiple dimensions. We now make this notion, illustrated in Figure 2(a), precise.

Definition 13 (Geometric grid partition) *Given a DDF problem with $\Sigma = [\frac{1}{\theta}, 1]^d$, and a grid pitch $k \in \mathbb{N}_+$, select s_1, \dots, s_{k+1} such that $s_1 = \frac{1}{\theta}$, $s_{k+1} = 1$, and $\frac{s_{i+1}}{s_i} > 0$ is constant. For each $j \in [k]^d$, define the grid cell $\Sigma(j) = \prod_{i=1}^d [s_{j(i)}, s_{j(i)+1}]$, where $j(i)$ is the i^{th} component of j . The cells satisfy $\Sigma = \bigcup_{j \in [k]^d} \Sigma(j)$, thus forming an partition (up to boundaries) of Σ into k^d cells.*

Empirical upper bound on $N_\alpha(\Phi)$. In this experiment, we construct 2-suboptimal covers of the linearized quadrotor for varying θ using geometric grids. We begin with the guess $k = 1$. For each grid cell $\Sigma(j)$, we compute a controller $K(j)$ using GCC synthesis and check if $\Sigma(j) \subseteq \mathcal{A}_2(K(j))$. (This requires evaluating only one Lyapunov equation due to Lemma 6.) If not, we increment k and try again. Termination is guaranteed by continuity. Results for this experiment with $\theta \in [1, 100]$ are shown in Figure 2(b). The required grid pitch k follows roughly logarithmic growth, as indicated by the linear least-squares best-fit curve in black. Small values of θ are excluded from the fit (indicated by black markers), as we do not expect the asymptotic behavior to appear yet.

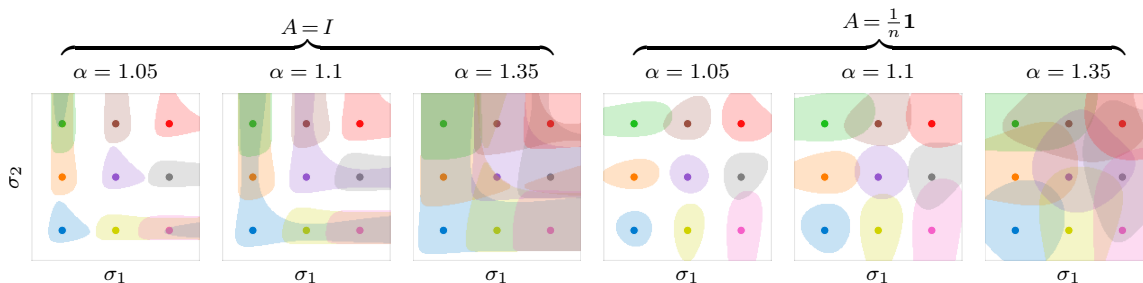


Figure 3: α -suboptimal neighborhoods for geometric grid partition in 2D system. *Left*: minimum coupling; $A = I$. *Right*: maximum coupling; $A = \frac{1}{n}\mathbf{1}$. *Columns*: varying suboptimality threshold α . All axes are logarithmic. Colors have no meaning. Discussion in Section 4.1.

These results do not rule out the $\log(\theta)^d$ growth suggested by the geometric grid construction. Testing larger values of θ is computationally difficult because the number of grid cells becomes huge and the GCC Riccati equation (2) becomes numerically unstable for very small Σ .

Efficiency of geometric grid partition. Given an α -suboptimal geometric grid cover, we examine a measurable quantity that may reflect the “efficiency” of the cover. Intuitively, in a good cover we expect the worst-case suboptimality ratio of each controller $K(j)$ relative to its grid cell $\Sigma(j)$ to be close to α . If it close to α for some cells but significantly less than α for others, then the grid pitch around the latter cells is finer than necessary. We visualize results for this computation on the linearized quadrotor with $\theta = 10$, $k = 4$ in Figure 2(c) — only the corners of the $4 \times 4 \times 4 \times 4$ grid are shown. The suboptimality ratio is close to $\alpha = 2$ for cells with low control authority (near $\Sigma = \frac{1}{\theta}I$), but drops to around 1.4 for cells with high control authority (near $\Sigma = I$). The difference suggests that the geometric grid cover could be more efficient in the high-authority regime.

Efficiency of GCC synthesis. One possible source of conservativeness is that Lemma 4 applies to the affine image of a $m \times n$ -dimensional matrix norm ball, but we only require guaranteed cost on a d -dimensional affine subspace of matrices. In other words, we ask GCC synthesis to ensure α -suboptimality on systems that are not actually part of Φ . If this is negatively affecting the result, then we should observe that the worst-case cost of $K(j)$ on $\Sigma(j)$ is less than the trace of the solution P for the GCC Riccati equation (2). The worst-case cost always occurs at the minimal $\Sigma \in \Sigma(j)$ by Lemma 6; we evaluate it with (1). For the quadrotor, a mismatch sometimes occurs for smaller values of θ , but it does not occur for the large values of θ .

4.1. Suboptimal neighborhood visualizations

We now present intuition-building experiments towards a covering number lower bound for matrix DDF problems. A lower bound requires a class of DDF problem that can be instantiated for any dimensionality d . Two “extremal” systems come to mind: *minimum coupling*, where $A = I$, and *maximum coupling*, where $A = \frac{1}{n}\mathbf{1}$. Note that for minimum coupling, an α -suboptimal policy is not necessarily α -suboptimal on each scalar subsystem—if it were, the lower bound $\log(\theta)^d$ would trivially follow from the results in Section 3.



Figure 4: α -suboptimal neighborhoods for the three-dimensional decomposed dynamics system with minimal coupling ($A = U = V^\top = I_{3 \times 3}$) and breadth $\theta = 100$. Neighborhoods shown for α ranging from 1.04 to 1.2 with a fixed controller.

We show approximate suboptimal neighborhoods for a two-dimensional system in Figure 3. We select a geometric grid of Σ values (indicated by the circular markers) and synthesize their LQR-optimal controllers. Then, we evaluate the suboptimality ratio of each controller on a finer grid of Σ values to get approximate neighborhoods, indicated by the semi-transparent regions. We repeat this experiment with three values of α for both choices of A .

Interestingly, the neighborhoods for $A = I$ are not always connected. In the plot for $\alpha = 1.05$ (far left), the neighborhood for the minimal Σ has another component that overlaps other neighborhoods to its top and right. If we increase to $\alpha = 1.1$, the components join into an “L”-shaped region. In contrast, the neighborhoods for $A = \frac{1}{n}\mathbf{1}$ seem more well-behaved. For both choices of A , the neighborhoods are of comparable size.

To verify that this behavior is not an artifact of the two-dimensional case only, we repeat the experiment in three dimensions. Figure 4 shows neighborhoods of one controller $K = K_{(2/\theta)I}^*$ for α ranging from 1.04 to 1.2. As α grows, $\mathcal{N}_\alpha(K)$ shows similar topological phases as the 2D case. In the simply-connected phase (large α), the neighborhood appears to include any Σ where at least one σ_i is sufficiently small. If this property holds in higher dimensions, then it would be possible to construct a cover using only controllers of uniform gain in all dimensions for large α .

5. Related work

Suboptimal coverings are closely related to several topics in control theory. Robust control synthesis under parametric uncertainty (Dullerud and Paganini, 2000) can be interpreted as seeking a policy that performs well on all of Φ without observing the particular $\phi \in \Phi$. Most problem statements in robust synthesis admit problem instances with no solution; the goal is to find a robust policy *if* one exists. Gain-scheduled control considers a multi-task setup identical to ours, while adaptive control control adds the complication that ϕ is not known to the policy.

Adaptive and gain-scheduled policies of the *self-tuning* type synthesize a single-task policy after estimating ϕ , but this relies on the assumption that control synthesis can be computed quickly (Åström and Wittenmark, 2013). In contrast, methods of the *multi-model* type use a precomputed set of policies (Murray-Smith and Johansen, 1997). Although all multi-model methods impose some kind of coverage condition on the policy set, researchers have focused more on the switching rule than the policy set. For example, Fu and Barmish (1986); Stilwell and Rugh (1999); Yoon et al. (2007) non-constructively assert the existence of a finite cover by continuity and compactness arguments. To address the need for small covers, Anderson et al. (2000); McNichols and Fadali (2003); Tan et al. (2004); Fekri et al. (2006); Du et al. (2012) propose constructive algorithms for

various classes of Φ , sometimes with arguments for minimality but without bounds on the covering number. [Jalali and Golmohammad \(2012\)](#) bound covering numbers in terms of frequency-domain properties of Φ as opposed to state-space parameters like mass and geometry. The most closely related work to ours is from [Fu \(1996\)](#), who shows a tight bound of 2^n for the stability covering number of a relatively broad Φ . This result is complementary to ours: suboptimality is a stronger criterion than stability, but our class of Φ is more restrictive. We are not aware of prior work that bounds covering numbers in a setup based on local suboptimality, as opposed to a single global performance measure.

Multi-task control is also a popular topic in deep learning research, where it is often motivated by ideas of lifelong skill acquisition in robotics. Domain randomization methods follow the spirit of robust control ([Peng et al., 2017](#)), but usually optimize for the average case instead of a worst-case guarantee. Many methods where the policy observes ϕ use architectural constructs that can only be applied to finite task sets ([Yang et al., 2017](#); [Parisotto et al., 2016](#); [Devin et al., 2017](#)). A common approach for infinite task spaces is to treat ϕ as a vector input alongside the system state. [Yu et al. \(2017\)](#); [Chen et al. \(2018\)](#) use this approach for dynamics parameters; [Schaul et al. \(2015\)](#) use it for navigation goals. There is evidence that policy class influences these methods: in a recent benchmark ([Yu et al., 2019](#)), the concatenated-input architecture trails the finite-task architecture. Other investigations into the difficulty of learning policies for multi-task control include methods to condition the multi-task optimization landscape ([Yu et al., 2020](#)) or balance disparate cost ranges ([van Hasselt et al., 2016](#)).

6. Conclusion and future work

In this paper, we introduced and motivated the α -suboptimal covering number to quantify infinite task spaces for multi-task control problems. We defined a particular class of multi-task linear-quadratic regulator problems amenable to analysis of the α -suboptimal covering number, and showed logarithmic dependency on the problem “breadth” parameter θ in the scalar case. Towards analogous results for the matrix case, we presented empirical studies intended to shed light on possible proof techniques. For the upper bound, we considered a natural covering construction that would preserve logarithmic dependence on θ but give exponential dependence on dimensionality. Experiments did not rule out its validity. For the lower bound, we visualized suboptimal neighborhoods for two possible system classes and observed interesting topological behavior for the minimal-coupling class.

After extending our current results to the matrix case, in future work the analysis can be applied to other classes of multi-task LQR problems including variations in A, Q, R , discrete time, and stochastic dynamics. It will be interesting to see if there are major differences between LQR variants. We also hope that suboptimal covers and covering numbers will be a useful tool for analyzing how the size of the task space affects the required expressiveness of function classes used in practice as multi-task policies, such as neural networks.

References

- Brian D. O. Anderson, Thomas S. Brinsmead, Franky De Bruyne, Joao Hespanha, Daniel Liberzon, and A. Stephen Morse. Multiple model adaptive control. part 1: Finite controller coverings. *International Journal of Robust and Nonlinear Control*, 10(11-12):909–929, 2000.
- Karl J Åström and Björn Wittenmark. *Adaptive Control*. Courier Corporation, 2013.
- Jingjing Bu, Afshin Mesbahi, and Mehran Mesbahi. On topological and metrical properties of stabilizing feedback gains: the MIMO case. *CoRR*, abs/1904.02737, 2019.
- Tao Chen, Adithyavairavan Murali, and Abhinav Gupta. Hardware conditioned policies for multi-robot transfer learning. In *NeurIPS*, pages 9355–9366, 2018.
- Coline Devin, Abhishek Gupta, Trevor Darrell, Pieter Abbeel, and Sergey Levine. Learning modular neural network policies for multi-task and multi-robot transfer. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 2169–2176, 2017.
- Jingjing Du, Chunyue Song, and Ping Li. Multimodel control of nonlinear systems: An integrated design procedure based on gap metric and H_∞ loop shaping. *Industrial & Engineering Chemistry Research*, 51(9):3722–3731, 2012.
- Geir E. Dullerud and Fernando Paganini. *A Course in Robust Control Theory: A Convex Approach*. Springer-Verlag New York, 2000.
- Maryam Fazel, Rong Ge, Sham M. Kakade, and Mehran Mesbahi. Global convergence of policy gradient methods for linearized control problems. *CoRR*, abs/1801.05039, 2018.
- Sajjad Fekri, Michael Athans, and Antonio Pascoal. Issues, progress and new results in robust adaptive control. *International Journal of Adaptive Control and Signal Processing*, 20(10):519–579, 2006.
- Minyue Fu. Minimum switching control for adaptive tracking. In *Proceedings of 35th IEEE Conference on Decision and Control*, volume 4, pages 3749–3754, 1996.
- Minyue Fu and B. Ross Barmish. Adaptive stabilization of linear systems via switching control. *IEEE Transactions on Automatic Control*, 31(12):1097–1103, 1986.
- Ali Akbar Jalali and Hassan Golmohammad. An optimal multiple-model strategy to design a controller for nonlinear processes: A boiler-turbine unit. *Computers & Chemical Engineering*, 46: 48–58, 2012.
- R. E. Kalman. Contributions to the theory of optimal control. *Boletín de la Sociedad Matemática Mexicana*, 5:102–199, 1960.
- Peter Lancaster and Leiba Rodman. *Algebraic Riccati Equations*. Clarendon Press, 1995.
- Kenneth H. McNichols and M. Sami Fadali. Selecting operating points for discrete-time gain scheduling. *Computers & Electrical Engineering*, 29(2):289–301, 2003.

- Roderick Murray-Smith and Tor Arne Johansen. *Multiple Model Approaches to Modelling and Control*. Taylor and Francis, London, 1997.
- Emilio Parisotto, Lei Jimmy Ba, and Ruslan Salakhutdinov. Actor-mimic: Deep multitask and transfer reinforcement learning. In *International Conference on Learning Representations (ICLR)*, 2016.
- Xue Bin Peng, Marcin Andrychowicz, Wojciech Zaremba, and Pieter Abbeel. Sim-to-real transfer of robotic control with dynamics randomization. *CoRR*, abs/1710.06537, 2017.
- Ian R. Petersen and Duncan C. McFarlane. Optimal guaranteed cost control and filtering for uncertain linear systems. *IEEE Transactions on Automatic Control*, 39(9):1971–1977, 1994.
- Tom Schaul, Daniel Horgan, Karol Gregor, and David Silver. Universal value function approximators. volume 37 of *Proceedings of Machine Learning Research*, pages 1312–1320, Lille, France, Jul 2015.
- Daniel J. Stilwell and Wilson J. Rugh. Interpolation of observer state feedback controllers for gain scheduling. *IEEE Transactions on Automatic Control*, 44(6):1225–1229, 1999.
- Wen Tan, Horacio J. Marquez, and Tongwen Chen. Operating point selection in multimodel controller design. In *Proceedings of the 2004 American Control Conference*, volume 4, pages 3652–3657, 2004.
- Hado P van Hasselt, Arthur Guez, Matteo Hessel, Volodymyr Mnih, and David Silver. Learning values across many orders of magnitude. In *Advances in Neural Information Processing Systems*, pages 4287–4295, 2016.
- Zhaoyang Yang, Kathryn E Merrick, Hussein A Abbass, and Lianwen Jin. Multi-task deep reinforcement learning for continuous action control. In *IJCAI*, pages 3301–3307, 2017.
- Myung-Gon Yoon, Valery A. Ugrinovskii, and Marek Pszczel. Gain-scheduling of minimax optimal state-feedback controllers for uncertain LPV systems. *IEEE Transactions on Automatic Control*, 52(2):311–317, 2007.
- Tianhe Yu, Deirdre Quillen, Zhanpeng He, Ryan Julian, Karol Hausman, Chelsea Finn, and Sergey Levine. Meta-world: A benchmark and evaluation for multi-task and meta reinforcement learning. In *CoRL*, volume 100 of *Proceedings of Machine Learning Research*, pages 1094–1100. PMLR, 2019.
- Tianhe Yu, Saurabh Kumar, Abhishek Gupta, Sergey Levine, Karol Hausman, and Chelsea Finn. Gradient surgery for multi-task learning. *CoRR*, abs/2001.06782, 2020.
- Wenhao Yu, Jie Tan, C. Karen Liu, and Greg Turk. Preparing for the unknown: Learning a universal policy with online system identification. In *Robotics: Science and Systems (RSS)*, 2017.