

On the convexity of a class of structured optimal control problems for positive systems

Neil K. Dhingra, Marcello Colombino, and Mihailo R. Jovanović

Abstract—We study a class of structured optimal control problems for positive systems in which the design variable modifies the main diagonal of the dynamic matrix. For this class of systems, we establish convexity of both the \mathcal{H}_2 and \mathcal{H}_∞ optimal control formulations. In contrast to previous approaches, our formulation allows for arbitrary convex constraints and regularization of the design parameter. We provide expressions for the gradient and subgradient of the \mathcal{H}_2 and \mathcal{H}_∞ norms and establish graph-theoretic conditions under which the \mathcal{H}_∞ norm is continuously differentiable. Finally, we develop a customized proximal algorithm for computing the solution to the regularized optimal control problems and apply our results for HIV combination drug therapy design.

I. INTRODUCTION

Even when standard optimal control problems admit convex formulations, the introduction of structural constraints or regularizers often makes the problem intractable. Consequently, significant effort has been devoted to identifying classes of convex problems. These include funnel causal and quadratically invariant systems [1], [2], positive systems [3], structured and sparse consensus and synchronization networks [4]–[6], optimal sensor/actuator selection [7], [8], and symmetric modifications to symmetric linear systems [9].

Positive systems have received much attention in recent years because of convenient properties that arise from their structure. A system is called positive if, for every nonnegative initial condition and input signal, its state and output remain nonnegative [10]. Such systems appear in the models of heat transfer, chemical networks, and probabilistic networks. In [11], the authors show that the KYP lemma greatly simplifies for positive systems, thereby allowing for decentralized \mathcal{H}_∞ synthesis via Semidefinite Programming (SDP). In [3], it is shown that static output-feedback can be solved via Linear Programming (LP) for a class of positive systems. In [12], [13], the authors develop necessary and sufficient conditions for robust stability of positive systems with respect to induced \mathcal{L}_1 – \mathcal{L}_∞ norm-bounded perturbations. In [14], [15], it is shown that the structured singular value is equal to its convex upper bound for positive systems. Thus, assessing robust stability with respect to induced \mathcal{L}_2 norm-

bounded perturbation is also tractable.

Most of the recent literature focuses on control design for positive systems with respect to induced \mathcal{L}_1 – \mathcal{L}_∞ norms or induced \mathcal{L}_2 norms [11], [13], [16]. In this paper, we show that a class of \mathcal{H}_2 control problems, which is not tractable for general systems, is convex for positive systems. We also show that the \mathcal{H}_∞ performance metric is convex in the original controller variables. This (i) allows us to formulate convex optimization problems where the control parameter is kept as an optimization variable; and (ii) facilitates a straightforward implementation of constraints or regularizers on the control parameter in the optimal control formulation.

The paper is organized as follows. In Section II, we formulate the regularized optimal control problem for a class positive systems. In Section III, we establish convexity of both \mathcal{H}_2 and \mathcal{H}_∞ optimal control formulations. In Section IV, we provide an example from combination drug therapy design. Finally, in Section V, we conclude the paper and summarize the ongoing research directions.

Notation

The set of real numbers is denoted by \mathbb{R} . \mathbb{R}_+ (\mathbb{R}_{++}) denote the set of nonnegative (positive) reals. The set of $n \times n$ Metzler matrices (matrices with nonnegative off diagonal elements) is denoted by \mathbb{M}^n . The set of $n \times n$ nonnegative (positive) diagonal matrices is denoted by \mathbb{D}_+^n (\mathbb{D}_{++}^n). Given a matrix A , A^T denotes its transpose. We use $\bar{\sigma}(A)$ to indicate the largest singular value of A , $\text{trace}(A)$ to denote its trace, and $\|A\|_F^2 := \text{trace}(A^T A)$ to denote its Frobenius norm. We write $A \geq 0$ ($A > 0$) if A has nonnegative (positive) entries and $A \succcurlyeq 0$ ($A \succ 0$) to denote that A is symmetric and positive semidefinite (definite). The vector inner product is given by $\langle x, y \rangle := x^T y$ and the matricial inner product is given by $\langle X, Y \rangle := \text{trace}(X^T Y)$. Given a set \mathcal{C} we define the indicator function

$$I_{\mathcal{C}}(x) := \begin{cases} 0 & \text{if } x \in \mathcal{C} \\ +\infty & \text{otherwise.} \end{cases}$$

We define the sparsity pattern $\text{sp}(u)$ of a vector u to be the set of indices for which u_i is nonzero. Finally, the ℓ_1 norm of the vector u is given by, $\|u\|_1 := \sum_i |u_i|$.

II. PROBLEM FORMULATION

A. Background material

We begin with a definition of a linear positive system.

Neil K. Dhingra and Mihailo R. Jovanović are with the University of Minnesota. E-mail addresses: {dhin0008, mihailo}@umn.edu. Marcello Colombino is with the Automatic Control Laboratory, ETH Zurich, Switzerland. E-mail address: mcolombi@control.ee.ethz.ch.

Financial support from the University of Minnesota Doctoral Dissertation Fellowship, the University of Minnesota Informatics Institute Transdisciplinary Faculty Fellowship, the National Science Foundation under award ECCS-1407958, and the Swiss National Science Foundation under grant 2-773337-12 is gratefully acknowledged.

Definition 1: A linear system described by the state-space representation,

$$\begin{aligned}\dot{x} &= Ax + Bd \\ y &= Cx + Dd,\end{aligned}$$

is *positive* if and only if A is Metzler and B , C , and D are nonnegative matrices.

We also recall several useful results related to positive systems that are easily derived.

Lemma 1: Let $A \in \mathbb{M}^n$. Then $e^A \geq 0$.

Proof: The Metzler matrix A can be written as $\tilde{A} - \alpha I$ with $\tilde{A} \geq 0$ and $\alpha > 0$. Then $e^A = e^{-\alpha} e^{\tilde{A}} \geq 0$ because $e^{\tilde{A}} \geq 0$. ■

Lemma 2: Let $A \in \mathbb{M}^n$ be Hurwitz and $Q \succ 0$ be a nonnegative matrix. Then the solution X to the Lyapunov equation

$$AX + XA^T + Q = 0$$

is elementwise nonnegative.

Proof: This follows directly from Lemma 1 and the fact that $X = \int_0^\infty e^{At} Q e^{A^T t} dt$. ■

Lemma 3: Let $A \in \mathbb{R}_+^{n \times n}$. Then the left and right singular vectors, w and v , associated with the largest singular value of A are nonnegative. If $A \in \mathbb{R}_{++}^{n \times n}$, then w and v are positive and unique.

Proof: This follows from the application of the Perron theorem [17, Theorem 8.2.11] to AA^T and $A^T A$. ■

B. \mathcal{H}_2 and \mathcal{H}_∞ optimal control problems

We consider a class of positive systems,

$$\begin{aligned}\dot{x} &= (A + K(u))x + Bd \\ z &= Cx\end{aligned}\tag{1}$$

where A is a Metzler matrix, B and C are positive matrices, $K: \mathbb{R}^m \rightarrow \mathbb{D}^n$ is a linear function of the control parameter $u \in \mathbb{R}^m$, $x(t) \in \mathbb{R}^n$ is the state vector, $z(t) \in \mathbb{R}^q$ is the performance output, and $d(t) \in \mathbb{R}^p$ is the disturbance input.

Our objective is to design a stabilizing *diagonal* matrix $K(u)$ that minimizes a measure of the amplification from d to z . In the case of white stochastic disturbances, we consider the steady-state variance of the output z ,

$$J_2(u) := \lim_{t \rightarrow \infty} \mathbf{E}(z^T(t)z(t))$$

where \mathbf{E} is the expectation operator. This performance metric quantifies the \mathcal{H}_2 norm of system (1). Another measure of the input-output amplification is the \mathcal{H}_∞ norm of the closed-loop system is given by,

$$J_\infty(u) := \sup_{\omega \in \mathbb{R}} \bar{\sigma}(C(j\omega I - A - K(u))^{-1}B)$$

where $\bar{\sigma}(\cdot)$ is the largest singular value of a given matrix. Minimization of J_2 or J_∞ over u may be ill-posed because there is no penalty on the control effort in the performance

output z . Thus, both J_2 and J_∞ might have infimums which are not attainable by a finite u .

To penalize the control effort and impose additional structural requirements on u , we consider a class of regularized optimal control problems

$$\begin{aligned}\text{minimize}_u & J(u) + g(u) \\ \text{subject to} & A + K(u) \text{ Hurwitz.}\end{aligned}\tag{2}$$

Although the regularization term $g(u)$ can be any convex function of u , we restrict our attention to the following regularizers

$$\begin{aligned}g_1(u) &= u^T R u \\ g_2(u) &= \gamma \|u\|_1 \\ g_3(u) &= I_{\mathcal{C}}(u),\end{aligned}\tag{3}$$

with $R \succ 0$ and $\gamma > 0$. The quadratic penalty limits the magnitude of u , the ℓ_1 norm promotes sparsity of u , and the indicator function enforces that u belongs to a convex set \mathcal{C} . We refer the reader to [7], [18], [19] for information about some recent uses of regularization in the optimal control problems.

It is well known that the maximum eigenvalue of $A + K(u)$ is a convex function of u [20]. Recently, it has been shown that the weighted \mathcal{L}_1 norm of the response of system (1) at time T from a nonnegative initial condition $x_0 \geq 0$,

$$\int_0^T c^T x(t) dt$$

is a convex function of u for every nonnegative vector $c \in \mathbb{R}^m$ [16], [21]. Furthermore, decentralized \mathcal{H}_∞ control of positive systems can be cast as a semidefinite program (SDP) using a suitable change of coordinates [11]. Recent work has provided a characterization of this problem via a linear program (LP) [22]. However, since both the SDP and the LP formulations require a change of variables that does not preserve the structure of $K(u)$, it is difficult to explicitly impose structural constraints or penalties on u .

In this paper, we show that both the \mathcal{H}_2 and the \mathcal{H}_∞ norms are convex functions of the original optimization variable u and provide explicit expressions for the (sub)gradients. This allows us to develop efficient descent algorithms that solve regularized optimal control problems of the form (2).

C. Applications

The class of systems that we consider is encountered in a variety of applications ranging from network theory to the control of biological systems.

1) *Combination drug therapy design for HIV treatment:* As shown in [23], [24], the problem of designing drug dosages for treating HIV can be expressed as

$$\dot{x} = \left(-L + \text{diag}(p) - \sum_{k=1}^m u_k D_k \right) x + d.$$

Here, the i th component of the state vector x represents the population of the i th HIV mutant. The diagonal matrix $\text{diag}(p)$ specifies the replication rate of each mutant, and

L is a directed graph Laplacian which quantifies the rate of mutation from one mutant to another. The control input u_k is the dose of drug k and D_k is a given diagonal matrix which specifies how efficiently drug k kills each HIV mutant. Imposing constraints on drug dosages is challenging using existing positive systems tools for \mathcal{H}_2 and \mathcal{H}_∞ design because the drugs u do not appear explicitly as optimization variables [21], [24], [25], making regularization difficult.

2) *Leader selection in directed networks:* We consider directed networks with n nodes. Each node x_i updates its state by averaging with its set of neighbors \mathcal{N}_i ,

$$\dot{x}_i = \sum_{j \in \mathcal{N}_i} w_{ij}(x_j - x_i) + d_i,$$

where w_{ij} is the weight of the edge from j to i , and d_i is a disturbance affecting node i . In contrast to this relative information exchange, some nodes – ‘leaders’ – are equipped with absolute information about their state,

$$\dot{x}_i = \sum_{j \in \mathcal{N}_i} w_{ij}(x_j - x_i) - u_i(x_i - r_i),$$

where u_i is a control gain and r_i a reference signal. The aggregate dynamics can be compactly written as,

$$\dot{x} = -(L + K(u))x + d$$

where L is a graph Laplacian [26] that reflects the network structure and $K(u) = \text{diag}(u)$ is a diagonal matrix whose nonzero entries identify the leaders.

One application arises in vehicular formations where the ‘leaders’ are equipped with absolute measurements such as GPS data and the regular nodes can only measure their distances from neighboring vehicles. Another application is seen in opinion dynamics in social networks where it is desired to identify the ‘leaders’ for targeted advertising.

III. OPTIMAL CONTROL

In this section we develop the main technical results of the paper. We show that the \mathcal{H}_2 and \mathcal{H}_∞ norms of the positive system (1) are convex in the decision variable u . We also provide explicit characterizations of the gradient of J_2 and the subgradient set of J_∞ and present a condition for J_∞ to be continuously differentiable over u .

A. \mathcal{H}_2 optimal control

We begin with the convexity result for the \mathcal{H}_2 norm.

Proposition 4: Let A be Metzler, B and C be positive matrices, and $K(u)$ be a diagonal matrix that depends linearly on u . Then, the square of the \mathcal{H}_2 norm of the closed-loop system (1) is a convex function of u .

Proof: The square of the \mathcal{H}_2 norm is given by,

$$J_2(u) = \langle C^T C, X_c \rangle := \text{trace}(C^T C X_c)$$

where X_c is the controllability Gramian of the closed-loop system determined by the solution to the Lyapunov equation

$$(A + K(u))X_c + X_c(A + K(u))^T + BB^T = 0.$$

For Hurwitz $A + K(u)$, X_c can be expressed as,

$$X_c = \int_0^\infty e^{(A+K(u))t} BB^T e^{(A+K(u))^T t} dt.$$

Substituting this expression into the trace and rearranging terms yields,

$$\begin{aligned} J_2(u) &= \int_0^\infty \|C e^{(A+K(u))t} B\|_F^2 dt \\ &= \int_0^\infty \sum_{i,j} \left(c_i^T e^{(A+K(u))t} b_j \right)^2 dt \end{aligned}$$

where c_i^T is the i th row of C and b_j is the j th column of B . From [16, Lemma 3] we know that

$$c^T e^{(A+K(u))t} b$$

is a convex function of u when c and b are nonnegative. Since the range of this function is \mathbb{R}_+ and $(\cdot)^2$ is nondecreasing over \mathbb{R}_+ , the composition rules for convex functions [27] imply that the expression $(c_i^T e^{(A+K(u))t} b_j)^2$ is convex. Finally, convexity of J_2 follows from the linearity of the sum and integral operators. ■

Remark 1: Using [16, Lemma 4], convexity of the quadratic cost

$$\int_0^T x^T(t) C^T C x(t) dt$$

also holds over a finite or infinite time horizon for a piecewise constant u . Even though this allows for the design of time-varying controllers using model predictive control, in this paper we restrict our attention to a constant u .

Before we provide an explicit characterization for ∇J_2 , we recall the definition of the adjoint of a linear operator.

Definition 2: The adjoint of a linear operator $K: \mathbb{R}^m \rightarrow \mathbb{R}^{n \times n}$ is the linear operator $K^\dagger: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^m$ which satisfies

$$\langle X, K(u) \rangle = \langle K^\dagger(X), u \rangle.$$

Proposition 5: Let $A + K(u)$ be Hurwitz and let K be a linear operator. Then the gradient of J_2 is given by,

$$\nabla_u J_2(u) = 2K^\dagger(X_c X_o) \quad (4)$$

where X_c and X_o are the controllability and observability Gramians of the closed-loop system (1),

$$\begin{aligned} (A + K(u))X_c + X_c(A + K(u))^T + BB^T &= 0 \\ (A + K(u))^T X_o + X_o(A + K(u)) + C^T C &= 0. \end{aligned}$$

Proof: The Lagrangian associated with J_2 is

$$\begin{aligned} \mathcal{L}(u, X_c, X_o) &= \langle C^T C, X_c \rangle + \\ &\quad \langle X_o, (A + K(u))X_c + X_c(A + K(u))^T + BB^T \rangle \end{aligned}$$

Variation with respect to X_c and X_o yields

$$\begin{aligned} \nabla_{X_o} \mathcal{L} &= (A + K(u))X_c + X_c(A + K(u))^T + BB^T \\ \nabla_{X_c} \mathcal{L} &= (A + K(u))^T X_o + X_o(A + K(u)) + C^T C. \end{aligned}$$

By rewriting \mathcal{L} as

$$\mathcal{L}(u, X_c, X_o) = 2 \langle K^\dagger(X_c X_o), u \rangle + \langle C^T C, X_c \rangle + \langle X_o, A X_c + X_c A^T + B B^T \rangle$$

it follows that the gradient of \mathcal{L} with respect to u is,

$$\nabla_u \mathcal{L} = 2K^\dagger(X_c X_o) \quad (5)$$

where the controllability and observability Gramians of the closed-loop system solve the stationarity conditions $\nabla_{X_o} \mathcal{L} = 0$ and $\nabla_{X_c} \mathcal{L} = 0$, respectively. ■

Remark 2: The expression for the gradient provided in Proposition 5 applies to any linear system that can be stabilized with a linear function $K(u)$. However, when the system is not positive and K is not linear, the convexity of $J_2(u)$ is not guaranteed.

Remark 3: When $K(u) := \sum_i u_i K_i$ with $K_i \in \mathbb{D}^n$, each component of the gradient can be expressed as

$$\nabla_{u_i} J_2 = 2 \langle X_c X_o, K_i \rangle.$$

Since K is a diagonal operator, each element of the gradient is a weighted sum of the diagonal elements of the matrix $X_c X_o$. Lemma 2 implies that, for a positive system, the Gramians X_c and X_o are nonnegative matrices. Thus, the diagonal of $X_c X_o$ is positive and J_2 is a monotone function of its diagonal elements.

B. \mathcal{H}_∞ optimal control

We first briefly summarize the standard \mathcal{H}_∞ optimal control formulation for positive systems applied to (1). The KYP lemma for positive systems [11] states that $J_\infty(u) \leq \gamma$ if and only if there exist $P \in \mathbb{D}_{++}^n$ such that,

$$\begin{bmatrix} C^T C + (A + K(u))^T P + P(A + K(u)) & P B \\ B^T P & -\gamma^2 \end{bmatrix} \prec 0.$$

The change of variables $K(u)P = Y \in \mathbb{D}^n$ yields,

$$\begin{bmatrix} C^T C + AP + Y + PA + Y & P B \\ B^T P & -\gamma^2 \end{bmatrix} \prec 0.$$

from which $K(u)$ can be recovered as $K(u) = YP^{-1}$. Using the results in [22, Theorem 1], this LMI can be cast as an LP. However, the change of variables used in both the LMI and the LP does not allow us to penalize or constrain u directly. This is because regularizers $g(u)$ given by (3) are different from the standard quadratic penalty on the control effort.

The sparsity pattern of the diagonal matrix $K(u)$ is the same as the sparsity pattern of the diagonal matrix Y . Thus, promoting the sparsity of Y is equivalent to promoting the sparsity of $K(u)$. In particular, sparsity in u can be readily enforced in a special case when $K(u) = \text{diag}(u)$. However, in general u is a bilinear function of P^{-1} and Y and it is much more challenging to promote sparsity of u for $K(u) := \sum_k u_k D_k$ with diagonal D_k using the above formulation. It is therefore desirable to provide a convex characterization for the \mathcal{H}_∞ problem in the original set of coordinates.

Proposition 6: Let A be Metzler, B and C be positive matrices, and $K(u)$ be a diagonal matrix that depends linearly on u . Then, the \mathcal{H}_∞ norm of the closed-loop system (1) is a convex function of u .

Proof: The \mathcal{H}_∞ norm is defined as

$$J_\infty(u) = \sup_\omega \bar{\sigma} \left(C(j\omega I - (A + K(u)))^{-1} B \right).$$

For positive systems, the supremum is achieved at $\omega = 0$ [11]. Since $A + K(u)$ is Hurwitz,

$$-(A + K(u))^{-1} = \int_0^\infty e^{(A+K(u))t} dt \quad (6)$$

and therefore,

$$J_\infty = \bar{\sigma} \left(C \int_0^\infty e^{(A+K(u))t} dt B \right).$$

Using [16, Lemma 3], we conclude that each element in the argument of $\bar{\sigma}$ is a nonnegative convex function of u .

The maximum singular value $\bar{\sigma}(X)$ can be expressed as a convex function of the entries of X [27],

$$\bar{\sigma}(X) = \max_{\|w\|=1, \|v\|=1} w^T X v. \quad (7)$$

When X is a nonnegative matrix, v and w are nonnegative by Lemma 3. Since

$$w^T (X + \alpha e_i e_j^T) v \geq w^T X v$$

for any nonnegative α , $\bar{\sigma}(X)$ is nondecreasing in each element of X . Thus, the composition rules for convex functions [27] imply that J_∞ is a convex function of u . ■

Proposition 7: Let K be a linear operator and $A_{cl} := A + K(u)$ be Hurwitz. Then,

$$\begin{aligned} \partial J_\infty(u) = & \left\{ \sum_i \alpha_i K^\dagger (A_{cl}^{-1} B v_i w_i^T C A_{cl}^{-1}) \right. \\ & \left. \mid -w_i^T (C A_{cl}^{-1} B) v_i = J_\infty(u), \alpha_i \in \mathcal{P} \right\} \end{aligned} \quad (8)$$

where $K^\dagger(u)$ is the adjoint of $K(u)$ and \mathcal{P} is defined by

$$\mathcal{P} := \left\{ \alpha \mid \alpha_j \geq 0, \sum_j \alpha_j = 1 \right\}.$$

Proof: The largest singular value of the matrix X is determined by (7). The subdifferential set of the supremum over a set of differentiable functions,

$$f(x) = \sup_{i \in \mathcal{I}} f_i(x)$$

is the convex hull of the subgradients of each function that achieves the maximum [28, Theorem 1.13],

$$\partial f(x) = \sum_{j | f_j(x) = f(x)} \alpha_j \nabla f_j(x)$$

where $\alpha \in \mathcal{P}$. Therefore, the subgradient of $\bar{\sigma}$ is given by

$$\partial \bar{\sigma}(X) = \left\{ \sum_j \alpha_j w_j v_j^T \mid w_j^T X v_j = \bar{\sigma}(X), \alpha \in \mathcal{P} \right\}.$$

The matricial derivative of X^{-1} and the application of the chain rule yield (8). ■

Remark 4: The expression for the gradient provided in Proposition 7 applies to any linear system that can be stabilized with a linear function $K(u)$. However, when the system is not positive and K is not diagonal, the convexity of $J_\infty(u)$ is not guaranteed.

It is well known that, in general, the \mathcal{H}_∞ norm is not differentiable. We next provide graph theoretic conditions for J_∞ to be continuously differentiable.

Definition 3 (Graph associated to a matrix): Given $A \in \mathbb{R}^{n \times n}$ we denote the graph associated to A as $\mathcal{G}(A) = (\mathcal{V}, \mathcal{E})$, with the set of vertices $\mathcal{V} = \{1, \dots, n\}$ and the set of edges $\mathcal{E} := \{(i, j) \mid A_{ji} \neq 0\}$.

Definition 4 (Strongly connected graph): A graph $(\mathcal{V}, \mathcal{E})$ is strongly connected if there is a directed path linking every two distinct nodes in \mathcal{V} .

Lemma 8: [29, Proposition 1.34] Let $A \in \mathbb{R}^{n \times n}$ be such that $\mathcal{G}(A)$ has self-edges at each node. The following statements are equivalent:

- $\mathcal{G}(A)$ is strongly connected.
- A^{n-1} is positive.

Proposition 9: Let A be Metzler, B and C be nonnegative matrices, and $K(u)$ be a diagonal linear operator such that $A_{cl} := A + K(u)$ is Hurwitz. If the graph associated with A is strongly connected, J_∞ is a continuously differentiable function of u .

Proof: Since A is Metzler, we can always find an $\alpha > 0$ sufficiently large so that $A_{cl} = \tilde{A} - \alpha I$ with \tilde{A} nonnegative and $\tilde{A}_{ii} > 0$ for every $i \in \{1, \dots, n\}$. The edge set of $\mathcal{G}(\tilde{A})$ contains all edges of $\mathcal{G}(A)$ plus all self-edges. Since self-edges play no role in strong connectivity, if $\mathcal{G}(A)$ is strongly connected so is $\mathcal{G}(\tilde{A})$. From Lemma 8 we conclude that $\tilde{A}^{n-1} > 0$ and from the definition of the matrix exponential we have

$$e^{A_{cl}} = e^{-\alpha} e^{\tilde{A}} = e^{-\alpha} \sum_{k=0}^{\infty} \frac{\tilde{A}^k}{k!} > 0 \quad (9)$$

From (6) and (9), we conclude that if the graph associated with A is strongly connected, $-A_{cl}^{-1}$, and therefore $-CA_{cl}^{-1}B$ are positive matrices. By Lemma 3, w and v are positive and unique. This implies that (8) is unique for each stabilizing u and, thus, J_∞ is continuously differentiable. ■

IV. EXAMPLE

Consider the HIV combination drug therapy problem described in Section II-C.1. Following [25], [30], [31], we study a system with 35 mutants (x) and 5 drugs (u). The sparsity pattern of the A matrix, shown in Fig. 1, corresponds to the mutation pattern and replication rates of the HIV mutants and $K(u)$ specifies the effect of drug therapy.

Several clinically relevant properties, such as maximum dose or budget constraints, may be directly enforced in this formulation. Other conditions can be promoted via convex penalties, such as drug j requiring drug i via $u_i \geq u_j$ or

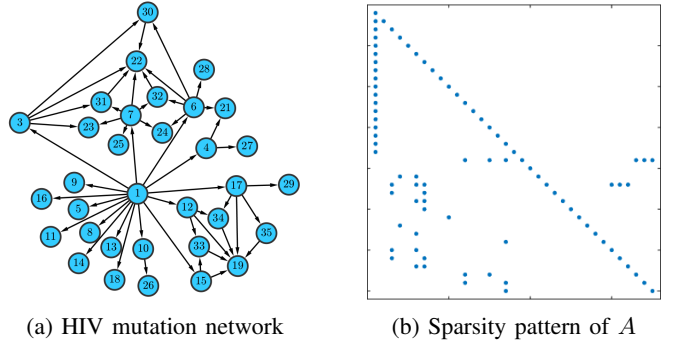


Fig. 1: Mutation pattern in the HIV model.

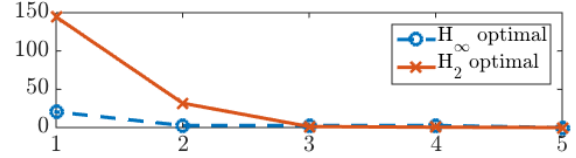


Fig. 2: Percent \mathcal{H}_2 and \mathcal{H}_∞ performance degradation as a function of the number of drugs N .

mutual exclusivity of drugs i and j via $u_i + u_j \leq 1$. We design \mathcal{H}_2 and \mathcal{H}_∞ -optimal drug doses using two types of convex regularizer g .

A. Budget constraint

We first impose a unit budget constraint on the drug doses and solve the \mathcal{H}_2 and \mathcal{H}_∞ -optimal problems,

$$\begin{aligned} & \text{minimize} && J(u) \\ & \text{subject to} && \sum_i u_i = 1, \quad u_i \geq 0. \end{aligned}$$

using proximal gradient and proximal subgradient methods [32], [33]. Table I contains the optimal doses and illustrates the tradeoff between \mathcal{H}_2 and \mathcal{H}_∞ performance.

Antibody	$u_{\mathcal{H}_2}$	$u_{\mathcal{H}_\infty}$	$J_2(u_{\mathcal{H}_2})$	$J_2(u_{\mathcal{H}_\infty})$	$J_\infty(u_{\mathcal{H}_2})$	$J_\infty(u_{\mathcal{H}_\infty})$
3BC176	0.5952	0.9875	0.6017	1.1947	0.1857	0.1084
PG16	0	0				
45-46G54W	0.2484	0.0125				
PGT128	0.1564	0				
10-1074	0	0				

TABLE I: Optimal budgeted doses and $\mathcal{H}_2/\mathcal{H}_\infty$ norms.

B. Sparsity-promoting framework

Although the budget constraint above is naturally sparsity-promoting, we introduce an algorithm to select sparser sets of drugs for a quadratically regularized problem. In Algorithm 1, we use a reweighted ℓ_1 penalty function [34] to select a few drugs and then perform a ‘polishing’ step to design the optimal doses of those drugs. Fig. 2 shows the percent performance degradation relative to the optimal dose using all 5 drugs with $B = C = I$, $R = I$ and γ varying from 0.01 to 10 in 50 logarithmically spaced increments.

Algorithm 1: Sparsity-promoting algorithm for N drugs

Set $\gamma > 0$, $R \succ 0$, $w = \mathbf{1}$, $\varepsilon > 0$;
while $\text{card}(u_\gamma) > N$ **do**
 $u_\gamma = \underset{u}{\text{argmin}} \quad J(u) + u^T R u + \sum w_i |u_i|$
 increase γ , $w_i = (u_i + \varepsilon)^{-1}$;
end
 $u^* = \underset{u}{\text{argmin}} \quad J(u) + u^T R u$
 subject to $\text{sp}(u) \subseteq \text{sp}(u_\gamma)$.

V. CONCLUDING REMARKS

In this paper, we establish convexity of the closed loop \mathcal{H}_2 and \mathcal{H}_∞ norms for a class of structured optimal control design problems for positive systems. We provide the (sub)gradients of these performance measures and give graph theoretic conditions for the \mathcal{H}_∞ norm to be continuously differentiable. Our formulation is amenable to arbitrary convex constraints and regularizers and can be used to address challenges in several applications, including leader selection in directed networks and combination drug therapy.

Our ongoing work focuses on extending these results to other classes of linear operators and time-varying control inputs. We also plan to study the leader selection and combination drug therapy problems in more detail.

ACKNOWLEDGMENTS

We would like to thank Anders Rantzer for useful discussion on positive systems and Vanessa Jonsson for help with the HIV model.

REFERENCES

- [1] B. Bamieh and P. G. Voulgaris, "A convex characterization of distributed control problems in spatially invariant systems with communication constraints," *Systems & Control Letters*, vol. 54, no. 6, pp. 575–583, 2005.
- [2] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *IEEE Trans. Automat. Control*, vol. 51, pp. 274–286, Feb 2006.
- [3] A. Rantzer, "Distributed control of positive systems," in *Proceedings of 50th IEEE Conference on Decision and Control*, pp. 6608–6611, Dec 2011.
- [4] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed average consensus with least-mean-square deviation," *J. Parallel Distrib. Comput.*, vol. 67, no. 1, pp. 33–46, 2007.
- [5] F. Lin, M. Fardad, and M. R. Jovanović, "Identification of sparse communication graphs in consensus networks," in *Proceedings of the 50th Annual Allerton Conference on Communication, Control, and Computing*, (Monticello, IL), pp. 85–89, 2012.
- [6] M. Fardad, F. Lin, and M. R. Jovanović, "Design of optimal sparse interconnection graphs for synchronization of oscillator networks," *IEEE Trans. Automat. Control*, vol. 59, no. 9, pp. 2457–2462, 2014.
- [7] B. Polyak, M. Khlebnikov, and P. Shcherbakov, "An LMI approach to structured sparse feedback design in linear control systems," in *Proc. of the 2013 European Control Conference*, pp. 833–838, 2013.
- [8] N. K. Dhingra, M. R. Jovanović, and Z. Q. Luo, "An ADMM algorithm for optimal sensor and actuator selection," in *Proceedings of the 53rd IEEE Conference on Decision and Control*, (Los Angeles, CA), pp. 4039–4044, 2014.
- [9] N. K. Dhingra and M. R. Jovanović, "Convex synthesis of symmetric modifications to linear systems," in *Proceedings of the 2015 American Control Conference*, (Chicago, IL), pp. 3583–3588, 2015.
- [10] L. Farina and S. Rinaldi, *Positive Linear Systems: Theory and Applications*. John Wiley & Sons, 2011.
- [11] T. Tanaka and C. Langbort, "The bounded real lemma for internally positive systems and H-infinity structured static state feedback," *IEEE Trans. Automat. Control*, vol. 56, no. 9, pp. 2218–2223, 2011.
- [12] C. Briat, "Robust stability and stabilization of uncertain linear positive systems via integral linear constraints: \mathcal{L}_1 gain and \mathcal{L}_∞ gain characterization," *International Journal of Robust and Nonlinear Control*, vol. 23, no. 17, pp. 1932–1954, 2013.
- [13] Y. Ebihara, D. Peaucelle, and D. Arzelier, " \mathcal{L}_1 gain analysis of linear positive systems and its application," in *Proceedings of 50th IEEE Conference on Decision and Control*, pp. 4029–4034, 2011.
- [14] M. Colombino and R. Smith, "Convex characterization of robust stability analysis and control synthesis for positive linear systems," in *Proceedings of the 53rd IEEE Conference on Decision and Control*, (Los Angeles, California), 2014.
- [15] M. Colombino and R. Smith, "A Convex Characterization of Robust Stability for Positive and Positively Dominated Linear Systems," *IEEE Trans. Automat. Control*. To appear. Available online at <http://arxiv.org/pdf/1508.04983v1.pdf>, June 2016.
- [16] P. Colaneri, R. H. Middleton, Z. Chen, D. Caporale, and F. Blanchini, "Convexity of the cost functional in an optimal control problem for a class of positive switched systems," *Automatica*, vol. 4, no. 50, pp. 1227–1234, 2014.
- [17] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge university press, 1985.
- [18] F. Lin, M. Fardad, and M. R. Jovanović, "Design of optimal sparse feedback gains via the alternating direction method of multipliers," *IEEE Trans. Automat. Control*, vol. 58, no. 9, pp. 2426–2431, 2013.
- [19] N. Matni and V. Chandrasekaran, "Regularization for design," in *Proceedings of the 53rd IEEE Conference on Decision and Control*, (Los Angeles, CA), pp. 1111–1118, 2014.
- [20] J. E. Cohen, "Convexity of the dominant eigenvalue of an essentially nonnegative matrix," *Proceedings of the American Mathematical Society*, vol. 81, no. 4, pp. 657–658, 1981.
- [21] A. Rantzer and B. Bernhardsson, "Control of convex monotone systems," in *Proceedings of the 53rd IEEE Conference on Decision and Control*, 2014.
- [22] A. Rantzer, "An Extended Kalman-Yakubovich-Popov Lemma for Positive Systems," *IFAC-PapersOnLine*, vol. 48, no. 11, pp. 242–245, 2015.
- [23] E. Hernandez-Vargas, P. Colaneri, R. Middleton, and F. Blanchini, "Discrete-time control for switched positive systems with application to mitigating viral escape," *International Journal of Robust and Nonlinear Control*, vol. 21, no. 10, pp. 1093–1111, 2011.
- [24] V. Jonsson, A. Rantzer, and R. M. Murray, "A scalable formulation for engineering combination therapies for evolutionary dynamics of disease," in *Proceedings of the 2014 American Control Conference*, pp. 2771–2778, IEEE, 2014.
- [25] V. Jonsson, N. Matni, and R. M. Murray, "Reverse engineering combination therapies for evolutionary dynamics of disease: An H_∞ approach," in *Proceedings of the 52nd IEEE Conference on Decision and Control*, pp. 2060–2065, 2013.
- [26] D. M. Cvetković, M. Doob, and H. Sachs, *Spectra of graphs: theory and application*, vol. 87. Academic Pr, 1980.
- [27] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [28] N. Z. Shor, *Minimization methods for non-differentiable functions*, vol. 3. Springer Science & Business Media, 2012.
- [29] F. Bullo, J. Cortés, and S. Martinez, *Distributed Control of Robotic Networks: A Mathematical Approach to Motion Coordination Algorithms*. Princeton University Press, 2009.
- [30] F. Klein, A. Halper-Stromberg, J. A. Horowitz, H. Gruell, J. F. Scheid, S. Bournazos, H. Mouquet, L. A. Spatz, R. Diskin, A. Abadir, *et al.*, "HIV therapy by a combination of broadly neutralizing antibodies in humanized mice," *Nature*, vol. 492, no. 7427, pp. 118–122, 2012.
- [31] V. Jonsson, A. Rantzer, and R. M. Murray, "A scalable formulation for engineering combination therapies for evolutionary dynamics of disease," *Proceedings of the 2014 American Control Conference*, 2014.
- [32] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.
- [33] D. P. Bertsekas, *Nonlinear programming*. Athena Scientific, 1999.
- [34] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, pp. 877–905, 2008.