Transformation of Optimal Centralized Controllers Into Near-Globally Optimal Static Distributed Controllers

Salar Fattahi*, Ghazal Fazelnia+, Javad Lavaei*, and Murat Arcak†
*Department of Industrial Engineering and Operations Research, University of California, Berkeley
+Department of Electrical Engineering, Columbia University
†Department of Electrical Engineering and Computer Science, University of California, Berkeley

***Abstract*—This paper is concerned with the optimal static distributed control problem for linear discrete-time deterministic and stochastic systems. The objective is to design a stabilizing static distributed controller whose performance is close to that of the optimal centralized controller. To this end, we fist consider deterministic systems, where the initial state is either given or belongs to a known bounded region. Given an arbitrary centralized controller, we derive a condition under which there exists a distributed controller that generates input and state trajectories close to their counterparts in the centralized closed-loop system. This condition for the design of a distributed controller is translated into an optimization problem, where the optimal objective value of this problem quantifies the closeness of the designed distributed and given centralized control systems. The results are then extended to stochastic systems that are subject to input disturbance and measurement noise. The proposed optimization problem has a closed-form solution (explicit formula) and can be efficiently solved for large-scale systems. The mathematical framework developed in this paper is utilized to design a near-globally optimal distributed controller based on the optimal centralized controller, and strong theoretical lower bounds on the global optimality guarantee of the obtained distributed controller are derived. We show that if the optimal objective value of the proposed convex program is sufficiently small, the designed controller is stabilizing and nearly globally optimal. To illustrate the effectiveness of the proposed method, case studies on aircraft formation and frequency control of power systems are offered.**

I. INTRODUCTION

The area of distributed control has been created to address computation and communication challenges in the control of large-scale real-world systems. The main objective is to design a controller with a prescribed structure, as opposed to the traditional centralized controller, for an interconnected system consisting of an arbitrary number of interacting local subsystems. This structurally constrained controller is composed of a set of local controllers associated with different subsystems, which are allowed to interact with one another according to the given control structure. The names "decentralized" and "distributed" are interchangeably used in the literature to refer to structurally constrained controllers (the latter term is often used for geographically distributed systems). It has been known that solving the long-standing optimal decentralized control problem is a daunting task due to its NP-hardness [1], [2]. Since it is not possible to design an efficient algorithm to solve this complex problem in its general form unless $P = NP$, several methods have been devoted to solving the optimal distributed control problem for special structures, such as spatially distributed systems [3], [4], localizable systems [5], [6], strongly connected systems [7], optimal static distributed systems [8], decentralized systems over graphs [9], [10], and quadratically-invariant systems [11].

Email: fattahi@berkeley.edu, gf2293@columbia.edu, lavaei@berkeley.edu and arcak@berkeley.edu

This work was supported by the ONR YIP Award, DARPA Young Faculty Award, AFOSR YIP Award, NSF CAREER Award, and NSF EPCN Award.

Due to the evolving role of convex optimization in solving complex problems, more recent approaches for the optimal distributed control problem have shifted toward a convex reformulation of the problem [12]–[20]. This has been carried out in the seminal work [21] by deriving a sufficient condition named quadratic invariance, which has been specialized in [22] by deploying the concept of partially order sets. These conditions have been further investigated in several other papers [23]–[25]. A different approach is taken in the recent papers [26] and [27], where it has been shown that the distributed control problem can be cast as a convex program for positive systems. Using the graph-theoretic analysis developed in [28], [29], it is shown in [30]–[32] that a semidefinite programming (SDP) relaxation of the distributed control problem has a low-rank solution for finite- and infinite-time cost functions in both deterministic and stochastic settings. The low-rank SDP solution may be used to find a near-globally optimal distributed controller. Moreover, it is proved in [33] that either a large input weighting matrix or a large noise covariance can convexify the optimal static distributed control problem for stable systems, and hence one can use a variety of iterative algorithms to find globally optimal static distributed controllers. Since SDPs and iterative algorithms are often computationally prohibitive for large-scale problems, it is desirable to develop a computationally-cheap method for designing suboptimal distributed controllers.

A. Contributions

The gap between the optimal costs of the optimal centralized and distributed control problems could be arbitrarily large in practice (as there may not exist a stabilizing controller with the prescribed structure). This paper is focused on systems for which this gap is expected to be relatively small. The main problem to be addressed is the following: given a centralized controller, is it possible to design a stabilizing static distributed controller with a given structure whose performance is close to that of the centralized one? The primary objective of this paper is to propose a candidate distributed controller via an explicit formula, which is indeed a solution to a system of linear equations.

In this work, we first study deterministic systems and derive a necessary and sufficient condition under which the states and inputs produced by a candidate static distributed controller and the given optimal centralized controller are identical for a given initial state. We translate the condition into an optimization problem, where the closeness of the given centralized and distributed control systems are captured by the smallness of the optimal objective value of this convex program. We then incorporate a regularization term into the objective function of the optimization problem to account for the stability of the closed-loop system. In real-world problems, it is often the case that the initial state is not known precisely. Therefore, in addition to closed-loop stability guarantee, the designed distributed controller should offer some optimality guarantee for every initial state belonging to a given uncertainty region. To address this issue, the proposed optimization problem is generalized to handle uncertain initial states as well. This optimization problem has a closed-form solution, which depends on the prescribed sparsity pattern of the to-be-designed controller as well as the properties of the uncertainty region of the initial state. A lower bound is obtained to guarantee the performance of the designed static distributed controller. This lower bound quantifies the distance between the performances of the designed controller and the given centralized one in the worst case. We show that the proposed convex program indirectly maximizes the derived lower bound while striving to achieve closed-loop stability. By building upon the derived results for deterministic systems, the proposed method is extended to stochastic systems that are subject to disturbance and measurement noises. We show that these systems benefit from similar lower bounds on optimality guarantee.

In this paper, we design a suboptimal distributed controller based on the optimal centralized controller. However, finding the best centralized controller could be computationally expensive for large-scale systems with tens of thousands of states. Under such circumstances, one may need to find an approximate solution of the Ricatti equations corresponding to the optimal centralized controller as delineated in [34]–[37]. The memory and computational complexities of these iterative methods are almost linearly proportional to the size of the problem, and they benefit from a quadratic convergence. This would lead to a near-global centralized controller, which can then be used to design a suboptimal distributed controller using the optimization problem proposed in this work. More precisely, the developed mathematical framework can be applied to any arbitrary centralized controller to obtain a distributed controller with a given sparsity pattern such that the centralized and distributed control systems perform similarly.

To demonstrate the efficacy of the developed mathematical framework, we consider two case studies in this paper. The first one is an instance of the control problem for multi-agent systems. In particular, we consider the aircraft formation problem in which each aircraft should make decisions solely based on its relative distance from the neighboring agents [38], [39]. In the second case study, we consider the frequency control problem for power systems with different topological restrictions on the distributed controller [32]. We will show that the synthesized distributed controllers offer high performance and closed-loop stability guarantees in both of the case studies.

The rest of this paper is organized as follows. The problem is formulated in Section II. Deterministic systems are studied in Section III, followed by an extension to stochastic systems in Section IV. The complexity analysis of the developed method is discussed in Section V. Case studies are provided in Section VI. Concluding remarks are drawn in Section VII. Some of the proofs are given in the appendix.

Notations: The space of real numbers is denoted by $\mathbb{R}$. The symbol $\mathrm{trace}\{W\}$ denotes the trace of a matrix W. I_m denotes the identity matrix of dimension m, where the subscript m is dropped when the dimension is implied by the context. The symbol $(\cdot)^T$ is used for transpose. The symbols $\|W\|_2$ and $\|W\|_F$ denote the 2-norm and Frobenius norm of W, respectively. The $(i,j)^{\text{th}}$ entry of a matrix W is shown as $W(i,j)$ or W_{ij}, whereas the i^{th} entry of a vector w is shown as $w(i)$ or w_i. The symbols $\lambda_W^{\max}$ or $\lambda_{\max}(W)$ refer to the maximum eigenvalue of a symmetric matrix W. The maximum absolute value of the eigenvalues of W is denoted by $\rho(W)$, and is called the spectral radius of the matrix W. The notation $W \succeq 0$ means that the symmetric matrix W is positive semidefinite. For a real number y, the notation $(y)_+$ denotes the maximum of 0 and y. The expected value of a random variable x is shown as $\mathcal{E}\{x\}$.

II. Problem Formulation

In this paper, the optimal static distributed control problem for systems with quadratic cost functions is studied. For simplicity of notation, this problem is referred to as *optimal distributed control (ODC)* henceforth (note that the term "ODC" is used only for static controllers and does not imply dynamic controllers in this paper). The objective is to develop a cheap, fast and scalable algorithm for the design of distributed controllers for large-scale systems. It is aimed to obtain a static distributed controller with a pre-determined structure that achieves a high performance compared to the optimal centralized controller. We implicitly assume that the gap between the optimal values of the optimal centralized and distributed control problems is not too large (otherwise, our method cannot produce a high-quality static distributed controller since there is no such controller). The mathematical framework to be developed here is particularly well-suited for mechanical and electrical systems such as power networks that are not highly unstable, for which it is empirically known that the above-mentioned gap is relatively small (note that the design problem is still hard even if the gap is small).

Definition 1. *Define $\mathcal{K} \subseteq \mathbb{R}^{m\times n}$ as a linear subspace with some pre-specified sparsity pattern (enforced zeros in certain entries). A feedback gain belonging to $\mathcal{K}$ is called a* **distributed (decentralized) controller** *with its sparsity pattern captured by $\mathcal{K}$. In the case of $\mathcal{K} = \mathbb{R}^{m\times n}$, there is no structural constraint imposed on the controller, which is referred to as a* **centralized controller**. *Throughout this paper, we use the notations K_c, K_d, and K to show an optimal centralized controller gain, a designed (near-globally optimal) distributed* *controller gain, and a variable controller gain (serving as a variable of an optimization problem), respectively.*

In this work, we will study two versions of the ODC problem, which are stated below.

Infinite-horizon deterministic ODC problem: Consider the discrete-time system

$$x[\tau+1] = Ax[\tau] + Bu[\tau], \qquad \tau = 0,1,...,\infty \tag{1}$$

with the known matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, where the initial state $x[0] \in \mathbb{R}^n$ may or may not be known *a priori*. The objective is to design a stabilizing static controller $u[\tau] = Kx[\tau]$ to satisfy certain optimality and structural constraints. Associated with the system (1) under an arbitrary controller $u[\tau] = Kx[\tau]$, we define the following cost function for the closed-loop system:

$$J(K) = \sum_{\tau=0}^{\infty} \left(x[\tau]^T Q x[\tau] + u[\tau]^T R u[\tau] \right) \tag{2}$$

where Q and R are constant positive-definite matrices of appropriate dimensions. Assume that the pair (A, B) is stabilizable. The minimization problem of

$$\min_{K\in\mathbb{R}^{m\times n}} J(K) \tag{3}$$

subject to (1) and the closed-loop stability condition is an optimal centralized control problem and the optimal controller gain can be obtained from the Riccati equation. However, if there is an enforced sparsity pattern on the controller via the linear subspace $\mathcal{K}$, the additional constraint $K \in \mathcal{K}$ should be added to the optimal centralized control problem, and it is well-known that Riccati equations cannot be used to find an optimal static distributed controller in general. We refer to this problem as the infinite-horizon deterministic ODC problem.

Infinite-horizon stochastic ODC problem: Consider the discrete-time system

$$\begin{cases} x[\tau+1] = Ax[\tau] + Bu[\tau] + Ed[\tau] \\ \quad\;\; y[\tau] = x[\tau] + Fv[\tau] \end{cases} \qquad \tau = 0,1,2,... \tag{4}$$

where A, B, E, F are constant matrices, and $d[\tau]$ and $v[\tau]$ denote the input disturbance and measurement noise, respectively. Furthermore, $y[\tau]$ is the noisy state measured at time τ. Associated with the system (4) under an arbitrary controller $u[\tau] = Ky[\tau]$, consider the cost functional

$$J(K) = \lim_{\tau\to+\infty} \mathcal{E}\left\{ x[\tau]^T Q x[\tau] + u[\tau]^T R u[\tau] \right\} \tag{5}$$

The infinite-horizon stochastic ODC problem aims to minimize the above objective function for the system (4) with respect to a stabilizing distributed controller K belonging to $\mathcal{K}$ (note that the operator $\lim_{\tau\to+\infty}$ in the definition of $J(K)$ can be alternatively changed to $\lim_{\tau'\to+\infty} \frac{1}{\tau'} \sum_{\tau=0}^{\tau'}$ without affecting the solution, due to the closed-loop stability).

Finding an optimal static distributed controller with a predefined structure is NP-hard and intractable in its worst case. Therefore, we seek to find a near-globally optimal static distributed controller. To measure the performance of the designed suboptimal distributed controller, the value of the objective function evaluated at the designed distributed controller is compared to that of the optimal centralized controller.

Definition 2. Consider the system (1) or (4) with the cost function (2) or (5), respectively. Given $K_d \in \mathcal{K}$ and a number $\mu \in [0,1]$, it is said that the distributed controller has the global optimality guarantee μ if it satisfies the inequality

$$\frac{J(K_c)}{J(K_d)} \geq \mu \tag{6}$$

We interchangeably denote the global optimality guarantee as a number μ between 0 and 1 or in percentage as $100 \times \mu\%$. For example, if $\mu = 0.95$, then the inequality (6) implies that the underlying distributed controller K_d is at most 5% worse than the optimal centralized controller K_c. This means that if there exists a better distributed controller, it would outperform K_d by at most 5%.

The problem under investigation in this paper is as follows: Given the deterministic system (1) or the stochastic system (4), find a distributed controller $u[\tau] = K_d x[\tau]$ such that

i) The design procedure for obtaining K_d is based on a simple formula with respect to K_c, rather than solving an optimization problem.
ii) The controller $u[\tau] = K_d x[\tau]$ has a high global optimality guarantee.
iii) The system (1) is stable under the controller $u[\tau] = K_d x[\tau]$.

III. Distributed Controller Design: Deterministic Systems

In this section, we study the design of static distributed controllers for deterministic systems. First, we assume that the initial state of the system is known *a priori*. For the sake of simplicity of notations, the initial state $x[0]$ is denoted as x henceforth. We consider two criteria in order to design a distributed controller. The first criterion is about the performance of the to-be-designed controller. The second criterion is concerned with the stability of the system under the designed controller. Next, we generalize the results to the case where the initial state is not known, but belongs to an uncertainty region.

A. Performance Criterion

Consider the optimal centralized controller $u[\tau] = K_c x[\tau]$ and an arbitrary distributed controller $u[\tau] = K_d x[\tau]$. Let $x_c[\tau]$ and $u_c[\tau]$ denote the state and input of the system (1) under the centralized controller. Likewise, define $x_d[\tau]$ and $u_d[\tau]$ as the state and input of the system (1) under the distributed controller. Given the initial state of the system, the next theorem derives a necessary and sufficient condition under which the centralized and distributed controllers generate the same state and input trajectories for the system (1).

Theorem 1. *Given the optimal centralized gain K_c, an arbitrary gain $K_d \in \mathcal{K}$, and the initial state x, the relations*

$$u_c[\tau] = u_d[\tau], \qquad \tau = 0,1,2,... \tag{7a}$$

$$x_c[\tau] = x_d[\tau], \qquad \tau = 0,1,2,... \tag{7b}$$

hold if and only if

$$(K_c - K_d)(A + BK_c)^\tau x = 0, \qquad \tau = 0, 1, 2, \ldots \tag{8}$$

Proof. The proof is provided in the appendix. □

To exploit the condition introduced in Theorem 1, an optimization problem will be introduced below.

Optimization A. This problem is defined as

$$\min_{K} \ \text{trace}\left\{(K_c - K)P_x(K_c - K)^T\right\} \tag{9a}$$

$$\text{s.t.} \ \ K \in \mathcal{K} \tag{9b}$$

where the symmetric positive-semidefinite matrix $P_x \in \mathbb{R}^{n\times n}$ is the unique solution of the Lyapunov equation

$$(A + BK_c)P_x(A + BK_c)^T - P_x + xx^T = 0 \tag{10}$$

Since P_x is positive semidefinite and the feasible set $\mathcal{K}$ is linear, Optimization A is convex. The next theorem explains how this optimization problem can be used to study the analogy of the centralized and distributed control systems.

Theorem 2. *Given the optimal centralized gain K_c, an arbitrary gain $K_d \in \mathcal{K}$, and the initial state x, the relations*

$$u_c[\tau] = u_d[\tau], \qquad \tau = 0, 1, 2, \ldots \tag{11a}$$

$$x_c[\tau] = x_d[\tau], \qquad \tau = 0, 1, 2, \ldots \tag{11b}$$

hold if and only if the optimal objective value of Optimization A is zero and K_d is a minimizer of this problem.

Proof. In light of Theorem 1, we need to show that condition (8) is equivalent to the optimal objective value of Optimization A being equal to 0. To this end, define the semi-infinite matrix

$$X = \begin{bmatrix} x & (A + BK_c)x & (A + BK_c)^2x & \cdots \end{bmatrix} \tag{12}$$

Now, observe that (8) is satisfied if and only if the Frobenius norm of $(K_c - K_d)X$ is equal to 0 or equivalently

$$\text{trace}\left\{(K_c - K_d)XX^T(K_c - K_d)^T\right\} = 0 \tag{13}$$

On the other hand, if P_x is defined as XX^T, then it is the unique solution of (10). This completes the proof. □

Theorem 2 states that if the optimal objective value of Optimization A is 0, then there exists a distributed controller $u_d[\tau] = K_d x_d[\tau]$ with the structure induced by $\mathcal{K}$ whose global optimality guarantee is 100%. Roughly speaking, a small optimal value for Optimization A implies that the centralized and distributed control systems can become close to each other. This statement will be formalized later in the paper. One may speculate that since the optimal centralized controller is unique, these does not exist a different (distributed) controller with the same performance and therefore the optimality guarantee of 100% is not reachable. However, note that the designed distributed controller and its optimal centralized counterpart would perform identically at the given initial state, but could be completely different in another initial state. In fact, it will later be shown that if the exact value of the initial state is not known, the optimality guarantee of precisely 100% is not achievable in general, but there is a way to modify Theorem 2 to make the optimality guarantee close to 100%.

Consider a general discrete Lyapunov equation

$$MPM^T - P + HH^T = 0 \tag{14}$$

for constant matrices M and H. It is well known that if M is stable, the above equation has a unique positive semidefinite solution P. Extensive amount of work has been devoted to the behavior of the eigenvalues of the solution of (14) whenever HH^T is low rank. [35], [36], [40]–[42]. Those papers show that if HH^T possesses a small rank compared to the size of P, the eigenvalues of P tend to decay quickly. As a result, one can notice that since xx^T has rank 1 in the Lyapunov equation (10), the matrix P_x tends to have a small number of dominant eigenvalues. In the extreme case, if the closed-loop matrix $A + BK_c$ is 0 (the most stable discrete system) or alternatively if x is chosen to be one of the eigenvectors of $A + BK_c$, the matrix P_x becomes rank-1. If there exists a distributed controller $K_d \in \mathcal{K}$ such that $K_c - K_d$ belongs to the subspace spanned by those eigenvectors corresponding to the insignificant eigenvalues of P_x, the optimal objective value of Optimization A would be small. In this case, although achieving the optimality guarantee of 100% is not possible, the guarantee would be close to 100% (this will be proven in this paper). From another perspective, this implies that although the infinite-horizon deterministic ODC problem with a known initial state may have a unique globally optimal solution in the form of a centralized controller, it may also possess several near-optimal controllers in the form of distributed controllers.

B. Stability Criterion

In the preceding subsection, we obtained a condition to guarantee an identical behavior for the centralized and distributed control systems associated with a given initial state. However, the condition does not necessarily ensure the stability of the distributed closed-loop system. In fact, whenever the centralized and distributed control systems have identical trajectories, the initial state x resides in the stable manifold of the system $x[\tau + 1] = (A + BK_d)x[\tau]$, but the closed-loop system is not necessarily stable. To address this issue, notice that $A + BK_d$ could be interpreted as a structured additive perturbation of the closed-loop system matrix corresponding to the centralized controller K_c, i.e.,

$$A + BK_d = A + BK_c + B(K_d - K_c) \tag{15}$$

Lemma 1. *There exists a strictly positive number ε such that an arbitrary distributed controller $u[\tau] = K_d x[\tau]$ with a gain $K_d \in \mathcal{K}$ stabilizes the system* (1) *if the norm of $B(K_d - K_c)$ at the point K_d is less than ε.*

Proof. The proof follows from (15). □

Motivated by Lemma 1, the aim of the next optimization problem is to minimize the Frobenius norm of $B(K_d - K_c)$.

Optimization B. This problem is defined as

$$\min_{K} \ \text{trace}\left\{(K_c - K)^T B^T B(K_c - K)\right\} \tag{16a}$$

$$\text{s.t.} \ \ K \in \mathcal{K} \tag{16b}$$

Note that there are several techniques in matrix perturbation and robust control to maximize or find a sub-optimal value ε [43]. Note also that the stability criterion (16a) is conservative, and can be improved by exploiting any possible structure in the matrices A and B together with the set $\mathcal{K}$.

C. Candidate Distributed Controller

Optimization A and Optimization B were introduced earlier to separately guarantee a high performance and closed-loop stability for a to-be-designed controller K_d. To benefit from both approaches, they will be merged into a single convex program below. To this end, define the functions

$$C_1(K,P) = \text{trace}\left\{(K_c-K)P(K_c-K)^T\right\} \tag{17a}$$
$$C_2(K) = \text{trace}\left\{(K_c-K)^T B^T B(K_c-K)\right\} \tag{17b}$$
$$C(K,P,\omega) = \omega \times C_1(K,P) + (1-\omega)\times C_2(K) \tag{17c}$$

where the parameters P, K and ω belong to $\mathbb{R}^{m\times n}$, $\mathbb{R}^{n\times n}$ and $\mathbb{R}$, respectively.

Optimization C. Given a constant number $\omega \in [0,1]$ and an $n\times n$ positive semidefinite matrix P, this problem is defined as the minimization of the function $C(K,P,\omega)$ with respect to the matrix variable $K \in \mathcal{K}$.

Note that $C(K,P_x,\omega)$ has two terms, where $C_1(K,P_x)$ accounts for the performance of the distributed controller while $C_2(K)$ indirectly enforces closed-loop stability. Assume that each matrix in the space $\mathcal{K}$ has l free entries to be designed. Denote these unknown parameters as $h_1, h_2, \dots h_l$. Furthermore, let $M_1,\dots,M_l \in \mathbb{R}^{m\times n}$ be constant 0-1 matrices such that $M_t(i,j)$ is equal to 1 if the pair (i,j) is the location of the free entry h_t in $K\in\mathcal{K}$ and is zero otherwise, for every $t\in\{1,2,\dots,l\}$. It is desirable to prove that the solution of Optimization C can be found via an explicit formula.

Theorem 3. *Given a constant number $\omega\in[0,1]$ and an $n\times n$ positive semidefinite matrix P, consider the matrix $X\in\mathbb{R}^{l\times l}$ and the vector $y\in\mathbb{R}^l$ with the entries*

$$\begin{aligned} X(i,j) &= \omega\ \text{trace}\left\{M_i P M_j^T\right\} \\ &\quad + (1-\omega)\ \text{trace}\left\{M_i^T B^T B M_j\right\} \end{aligned} \tag{18a}$$
$$\begin{aligned} y(i) &= \omega\ \text{trace}\left\{M_i P K_c^T\right\} \\ &\quad + (1-\omega)\ \text{trace}\left\{M_i^T B^T B K_c\right\} \end{aligned} \tag{18b}$$

for every $i,j\in\{1,2,\dots,l\}$. A matrix K_d is an optimal solution of Optimization C if and only if it can be expressed as $K_d = \sum_{i=1}^l M_i h_i$ such that the vector h defined as $[h_1 \cdots h_l]^T$ is a solution to the linear equation $Xh=y$.

Proof. The space of permissible controllers can be characterized as

$$\mathcal{K} \triangleq \left\{ \sum_{i=1}^l M_i h_i \;\middle|\; h\in\mathbb{R}^l \right\} \tag{19}$$

for $M_1,\dots,M_l\in\mathbb{R}^{m\times n}$ (note that h_i's are the entries of h). Substituting $K_d=\sum_{i=1}^l M_i h_i$ into (17) and taking its gradient with respect to h lead to the optimality condition

$$\begin{aligned} &\sum_{j=1}^l \omega\ \text{trace}\left\{M_i P M_j^T\right\} h_j \\ &+ \sum_{j=1}^l (1-\omega)\ \text{trace}\left\{M_i^T B^T B M_j\right\} h_j \\ &= \omega\ \text{trace}\left\{M_i P K_c^T\right\} + (1-\omega)\ \text{trace}\left\{M_i^T B^T B K_c\right\} \end{aligned} \tag{20}$$

The above equation can be written in a compact form as $Xh=y$. Note that since (17) is convex with respect to h and the constraint $K_d\in\mathcal{K}$ is linear, the above optimality condition is necessary and sufficient for the optimality of h. □

D. Unknown Initial State

Consider the case where the initial state of the system is not known precisely a priori. In this case, the equation (10) cannot be used because it depends on the unknown initial state x. In this subsection, the objective is to modify (10) to accommodate uncertainty on the initial state in the framework proposed earlier for designing a high-performance distributed controller. Assume that the initial state belongs to the uncertainty region $\mathcal{E}$. Throughout the rest of this paper, we assume that the uncertainty region is defined as

$$\mathcal{E} = \{a + Mu : u\in\mathbb{R}^{n\times n}, \|u\|_2 \le 1\} \tag{21}$$

for some $a\in\mathbb{R}^{n\times 1}$ and $M\in\mathbb{R}^{n\times n}$. If $\mathcal{E}$ does not have the above ellipsoidal expression, one may use an outer ellipsoidal approximation of this region at the expense of designing distributed controllers with a lower performance (see [44] for more details). Define

$$\mathcal{L}(P,x) = (A+BK_c)P(A+BK_c)^\top - P + xx^\top \tag{22}$$

Normally, there does not exist a common matrix P satisfying $\mathcal{L}(P,x)=0$ for every initial state in $\mathcal{E}$. To bypass this issue, we introduce a new optimization problem to design a matrix P such that $\mathcal{L}(P,x)$ is maintained close to zero for every initial state in the uncertainty region. This problem is defined as

$$\min_{\alpha,P} \quad \alpha \tag{23a}$$
$$\text{s.t.} \quad -\alpha I \le \mathcal{L}(P,x) \le \alpha I, \qquad \forall x\in\mathcal{E} \tag{23b}$$
$$P \ge 0 \tag{23c}$$

Notice that (23) is a semidefinite programming (SDP) with an infinite number of constraints. There are two potential issues. First, using the optimal solution of (23) as a surrogate for P_x in (10) may not necessarily lead to a high performance distributed controller (since the optimality condition introduced in Theorem 2 may no longer hold for the distributed controller designed based on the optimal solution of (23)). Furthermore, (23) is an infinite-dimensional optimization problem and cannot be solved efficiently unless it is formulated as a finite-dimensional problem. In the sequel, we remedy both of the above-mentioned problems. First, we contrive an explicit solution that is nearly optimal for (23). Second, we derive a lower bound on the performance of the distributed controller designed based on the obtained explicit solution to guarantee the near global optimality of the decontroller for all initial states belonging to the uncertainty region. Define

$$s(\mathcal{E}) = \max_{\|y\|_2=1} \left\{ |a^\top y| \times \|My\|_2 \right\} \tag{24}$$

The following theorem studies the solution of the optimization problem (23).

Theorem 4. *Suppose that α^* is the optimal objective value of* (23). *Furthermore, define P^* as the unique solution of the Lyapunov equation*

$$(A+BK_c)P^*(A+BK_c)^\top - P^* + aa^\top + M^2 = 0 \tag{25}$$

Then, the following statements hold:

1. (β, P^*) *is a feasible solution for* (23)*, where*
$$\beta = 2\max\left\{s(\mathcal{E}), (\lambda_M^{\max})^2\right\} \tag{26}$$
2. $0 \le \beta - \alpha^* \le \left((\lambda_M^{\max})^2 - s(\mathcal{E})\right)_+$

Proof. The proof is provided in appendix. □

Remark 1. Notice that if $s(\mathcal{E}) \ge (\lambda_M^{\max})^2$, the tuple (β, P^*) is an optimal solution of (23). One sufficient condition for the satisfaction of $s(\mathcal{E}) \ge (\lambda_M^{\max})^2$ is the inequality $\|Ma\|_2 \ge (\lambda_M^{\max})^2$. Roughly speaking, this condition holds for those ellipsoids with the properties that the ratio of the largest to smallest diameters is not large and that the center of the ellipsoid is sufficiently far from the origin. For example, the condition is automatically satisfied whenever the uncertainty region is equal to a norm-2 ball that does not include the origin.

In the next subsection, we will use Optimization C with the input $P = P^*$ to design a high-performance controller. Note that although Theorem 4 does not offer closed-form formulas for α^* and β, only the matrix P^* is needed in Optimization C which can be found efficiently by solving the Lyapunov equation (25). In summary, two steps should be taken to design a distributed controller,:

1. Solve the Lyapunov equation (25) in order to find P^*.
2. Solve Optimization C with the input $P = P^*$ to obtain K_d.

Note that when $\mathcal{E}$ consists of a single initial state x, P^* coincides with P_x.

E. Lower Bound on Optimality Guarantee

Consider Optimization C with the input $P = P^*$. A distributed controller can be designed by solving this problem, which has an explicit solution due to Theorem 3. In this subsection, the objective is to derive a lower bound on the global optimality guarantee of the designed distributed controller. In particular, it is desirable to show that the optimality guarantee depends on how small the optimal value of $C_1(K, P^*)$ is. To this end, we first derive an upper bound on the deviation of the state and input trajectories generated by the distributed controller from those of the centralized controller.

Lemma 2. *Given the optimal centralized gain K_c, an arbitrary stabilizing gain $K_d \in \mathcal{K}$ and an initial state x, the relations*

$$\sum_{\tau=0}^{\infty} \|x_d[\tau] - x_c[\tau]\|_2^2 \le \left(\frac{\kappa(V)\|B\|_2}{1-\rho(A+BK_d)}\right)^2 C_1(K_d, P_x) \tag{27a}$$

$$\sum_{\tau=0}^{\infty} \|u_d[\tau] - u_c[\tau]\|_2^2 \le \left(1 + \frac{\kappa(V)\|K_d\|_2\|B\|_2}{1-\rho(A+BK_d)}\right)^2 C_1(K_d, P_x) \tag{27b}$$

hold, where $\kappa(V)$ is the condition number in 2-norm of the eigenvector matrix V of $A + BK_d$.

Proof. The proof is provided in the appendix. □

Notice that, according to the statement of Lemma 2, the upper bounds in (27a) and (27b) are valid if the distributed controller gain K_d makes the system stable. According to (12) and (13), one can verify that

$$C_1(K_d, P_x) = \sum_{\tau=0}^{\infty} \|(K_d - K_c)(A+BK_c)^\tau x\|_2^2 \tag{28}$$

An important observation can be made on the connection between Optimization C and the upper bounds in (27a) and (27b). Note that Optimization C minimizes a combination of $C_1(K, P)$ and $C_2(K)$. While the second term indirectly accounts for stability, the first term $C_1(K, P)$ directly appears in the upper bounds in (27a) and (27b) (recall that $P^* = P_x$ if $\mathcal{E}$ consists of the single initial state x). Hence, Optimization C aims at minimizing the deviation between the trajectories of the distributed and centralized control systems.

Definition 3 ([45])**.** For a stable matrix X, define the radius of stability as $r(X) = \inf_{0\le\theta\le 2\pi} \|(e^{\mathbf{i}\theta} - X)^{-1}\|^{-1}$.

It can be verified that $r(X) > 0$ and $r(X) + \rho(X) \le 1$.

Lemma 3. *Assume that R satisfies the Lyapunov equation $XRX^T - R + Y = 0$ for a stable matrix X. Then, we have*

$$\|R\|_2 \le \frac{\|Y\|_2}{r(X)^2} \tag{29}$$

Proof. See [46] and [45]. □

Lemma 4. *The relation*

$$\|P - P_x\|_2 \le \frac{\alpha}{r(A+BK_c)^2} \tag{30}$$

holds for every $x \in \mathcal{E}$ and every feasible solution (α, P) of the optimization problem (23).

Proof. One can write

$$-\alpha I \le \mathcal{L}(P, x) \le \alpha I \tag{31}$$

for every $x \in \mathcal{E}$. Subtracting $\mathcal{L}(P_x, x) = 0$ from (31) yields

$$-\alpha I \le \mathcal{L}(P - P_x, 0) \le \alpha I \tag{32}$$

According to (32) and Lemma 3, one can write

$$\|P - P_x\|_2 \le \frac{\|\mathcal{L}(P-P_x, 0)\|_2}{r(A+BK_c)^2} \le \frac{\alpha}{r(A+BK_c)^2} \tag{33}$$

□

Theorem 5. *Assume that $Q = I_n$ and $R = I_m$. Consider the optimal centralized gain K_c, an arbitrary stabilizing gain $K_d \in \mathcal{K}$, and a feasible solution (α, P) for the optimization problem* (23). *Suppose that the matrix $A + BK_d$ is diagonalizable. The controller $u[\tau] = K_d x[\tau]$ has the global optimality guarantee μ for every initial state $x \in \mathcal{E}$, where*

$$\mu = \frac{1}{\left(1 + \zeta\sqrt{C_1(K_d, P) + \eta\alpha}\right)^2} \tag{34}$$

and

$$\zeta = \max\left\{\frac{\kappa(V)\|B\|_2}{(1-\rho(A+BK_d))\sqrt{\sum_{\tau=0}^{\infty}\|x_c[\tau]\|_2^2}}\,,\right.$$
$$\left.\frac{1-\rho(A+BK_d)+\kappa(V)\|K_d\|_2\|B\|_2}{(1-\rho(A+BK_d))\sqrt{\sum_{\tau=0}^{\infty}\|u_c[\tau]\|_2^2}}\right\} \tag{35a}$$

$$\eta = \frac{\|K_c - K_d\|_F^2}{r(A+BK_c)^2} \tag{35b}$$

Proof. The proof is provided in appendix □

Theorem 5 quantifies the similarity between the behaviors of the system under the optimal centralized controller and an arbitrary distributed controller. The developed bound depends on $C_1(K_d, P)$ and $\rho(A+BK_d)$, which are both taken care of by Optimization C with the tradeoff coefficient ω. Hence, this paper proposes designing the distributed controller K_d based on Optimization C with the input $P = P^*$. Since this controller can be explicitly characterized in terms of the centralized controller K_c and the sparsity space $\mathcal{K}$ using Theorem 3, the global optimality guarantee of this controller depends on K_c and $\mathcal{K}$ via a closed-form formula. This formula could be used to study how changing the sparsity pattern of the controller space affects the performance loss with respect to the optimal centralized controller.

Remark 2. Notice that Theorem 5 is developed for the case of $Q = I_n$ and $R = I_m$. However, its proof can be adopted to derive similar bounds for the general case. Alternatively, for arbitrary positive-definite matrices Q and R, one can transform them into identity matrices through a reformulation of the ODC problem. Define Q_d and R_d as $Q = Q_d^T Q_d$ and $R = R_d^T R_d$, respectively. The ODC problem with the tuple $(A, B, x[\cdot], u[\cdot])$ can be reformulated with respect to a new tuple $(\bar{A}, \bar{B}, \bar{x}[\cdot], \bar{u}[\cdot])$ defined as

$$\bar{A} \triangleq Q_d A Q_d^{-1}, \quad \bar{B} \triangleq Q_d B R_d^{-1},$$
$$\bar{x}[\tau] \triangleq Q_d x[\tau], \quad \bar{u}[\tau] \triangleq R_d u[\tau],$$

Furthermore, in order to extend the result of Theorem 3 to general positive definite Q and R, the following mapping for the basis matrices $M_1, ..., M_l$ is required

$$\bar{M}_i \triangleq R_d M_i Q_d^{-1}, \quad i \in \{1, 2, ..., l\}$$

Remark 3. Although K_c is assumed to be the optimal centralized controller, none of the results developed in this paper uses the optimality property of K_c. This implies that the proposed mathematical framework works for every arbitrary stabilizing controller K_c, which produces a distributed controller K_d based on K_c together with a lower bound on the ratio $\frac{J(K_c)}{J(K_d)}$. As a by-product of this result, since finding the best centralized controller could be computationally expensive for large-scale systems with tens of thousands of states, one could consider K_c as a suboptimal centralized controller that is easy to obtain using the existing heuristic or approximation methods, and then design a distributed controller based on K_c. Then, one can conclude that the global optimality guarantee of the designed distributed controller K_d is equal to the product of μ given in Theorem 5 and the global optimality guarantee of K_c.

IV. Distributed Controller Design: Stochastic Systems

In this section, the results developed earlier are generalized to stochastic systems. For input disturbance and measurement noise, define the covariance matrices

$$\Sigma_d = \mathcal{E}\left\{Ed[\tau]d[\tau]^T E^T\right\}, \ \Sigma_v = \mathcal{E}\left\{Fv[\tau]v[\tau]^T F^T\right\} \tag{36}$$

for all $\tau \in \{0, 1, ..., \infty\}$. It is assumed that $d[\tau]$ and $v[\tau]$ are identically distributed and independent random vectors with Gaussian distribution and zero mean for all times τ. Let K_c denote the gain of the optimal static centralized controller $u[\tau] = K_c y[\tau]$ minimizing (5) for the stochastic system (4). Note that if $F = 0$, the matrix K_c can be found using the Riccati equation. The goal is to design a stabilizing distributed controller $u[\tau] = K_d y[\tau]$ with a high global optimality guarantee such that $K_d \in \mathcal{K}$. For an arbitrary discrete-time random process $a[\tau]$ with $\tau \in \{0, 1, ..., \infty\}$, denote the random variable $\lim_{\tau\to+\infty} a[\tau]$ as $a[\infty]$ if the limit exists. Note that the closeness of the random tuples $(u_c[\infty], x_c[\infty])$ and $(u_d[\infty], x_d[\infty])$ is sufficient to guarantee that the centralized and distributed controllers lead to similar performances. This is due to the fact that only the limiting behaviors of the states and inputs determine the objective value of the optimal control problem in (5).

Lemma 5. *Given the optimal centralized gain K_c and an arbitrary stabilizing distributed controller gain $K_d \in \mathcal{K}$, the relation*

$$\mathcal{E}\left\{\|x_c[\infty] - x_d[\infty]\|_2^2\right\} = \operatorname{trace}\{P_1 + P_2 - P_3 - P_4\} \tag{37}$$

holds, where P_1, P_2, P_3 and P_4 are the unique solutions of the equations

$$(A+BK_d)P_1(A+BK_d)^T - P_1 + \Sigma_d + (BK_d)\Sigma_v(BK_d)^T = 0 \tag{38a}$$

$$(A+BK_c)P_2(A+BK_c)^T - P_2 + \Sigma_d + (BK_c)\Sigma_v(BK_c)^T = 0 \tag{38b}$$

$$(A+BK_d)P_3(A+BK_c)^T - P_3 + \Sigma_d + (BK_d)\Sigma_v(BK_c)^T = 0 \tag{38c}$$

$$(A+BK_c)P_4(A+BK_d)^T - P_4 + \Sigma_d + (BK_c)\Sigma_v(BK_d)^T = 0 \tag{38d}$$

Proof. The proof is provided in the appendix. □

Lemma 5 implies that in order to minimize $\mathcal{E}\left\{\|x_c[\infty] - x_d[\infty]\|_2^2\right\}$, the trace of $P_1 + P_2 - P_3 - P_4$

should be minimized subject to (38) (by substituting K with K_d) and $K \in \mathcal{K}$. However, this is a hard problem in general. In particular, the minimization of the singleton $\mathrm{trace}\{P_1\}$ subject to (38a) and $K \in \mathcal{K}$ is equivalent to the ODC problem under study (if Q and R are identity matrices). Due to the possible intractability of the minimization of $\mathcal{E}\left\{\|x_c[\infty]-x_d[\infty]\|_2^2\right\}$, we aim to minimize an upper bound on this function (similar to the deterministic case).

Define P_s as the unique solution of the matrix equation

$$(A+BK_c)P_s(A+BK_c)^T - P_s + \Sigma_d + (BK_c)\Sigma_v(BK_c)^T = 0 \tag{39}$$

Lemma 2 will be generalized to stochastic systems next.

Lemma 6. *Given the optimal centralized gain K_c and an arbitrary stabilizing gain $K_d \in \mathcal{K}$, the relations*

$$\mathcal{E}\left\{\|x_c[\infty]-x_d[\infty]\|_2^2\right\} \leq \left(\frac{\kappa(V)\|B\|_2}{1-\rho(A+BK_d)}\right)^2 \times C_1(K_d, P_s+\Sigma_v) \tag{40a}$$

$$\mathcal{E}\left\{\|u_c[\infty]-u_d[\infty]\|_2^2\right\} \leq \left(1+\frac{\kappa(V)\|K_d\|_2\|B\|_2}{1-\rho(A+BK_d)}\right)^2 \times C_1(K_d, P_s+\Sigma_v) \tag{40b}$$

hold, where $\kappa(V)$ is the condition number in 2-norm of the eigenvector matrix V of $A+BK_d$.

Proof. The proof is omitted due to the similarity to the proof of Lemma 2, and is moved to [47]. □

In what follows, the counterpart of Theorem 5 will be presented for stochastic systems.

Theorem 6. *Assume that $Q = I_n$ and $R = I_m$. Consider the optimal centralized gain K_c and an arbitrary stabilizing gain $K_d \in \mathcal{K}$ for which $A+BK_d$ is diagonalizable. The controller $u[\tau] = K_d x[\tau]$ has the global optimality guarantee μ_s, where*

$$\mu_s = \frac{1}{\left(1+\zeta_s\sqrt{C_1(K_d, P_s+\Sigma_v)}\right)^2} \tag{41}$$

and

$$\zeta_s = \max\left\{\frac{\kappa(V)\|B\|_2}{(1-\rho(A+BK_d))\sqrt{\mathcal{E}\{\|x_c[\infty]\|_2^2\}}},\ \frac{1-\rho(A+BK_d)+\kappa(V)\|K_d\|_2\|B\|_2}{(1-\rho(A+BK_d))\sqrt{\mathcal{E}\{\|u_c[\infty]\|_2^2\}}}\right\} \tag{42}$$

Proof. The proof is a consequence of Lemma 6 and the argument made in the proof of Theorem 5. □

Similar to the deterministic case, one can extend the results of Theorem 6 to stochastic systems with general positive-definite matrices Q and R, using Remark 2 after two additional changes of parameters

$$\bar{E} \triangleq Q_d E, \quad \bar{F} \triangleq Q_d F$$

Theorem 6 quantifies the similarity between the behaviors of the system under the optimal centralized controller and an arbitrary distributed controller, which depends on $C_1(K_d, P_s+\Sigma_v)$ and $\rho(A+BK_d)$. Since both of these terms are incorporated into Optimization C with the tradeoff coefficient ω, this paper proposes designing the distributed controller K_d based on Optimization C with the input $P = P_s + \Sigma_v$. This controller has a closed-form solution due to Theorem 3, which enables expressing its global optimality guarantee as a function of the optimal centralized controller K_c and the sparsity space $\mathcal{K}$.

Remark 4. It is well-known that finding the optimal centralized controller in the presence of measurement noise is a difficult problem in general. If the optimal controller K_c is not available, one can use a near-globally optimal feedback gain as a substitute for K_c in Optimization C (in order to design a distributed controller that performs similarly to the near-globally optimal centralized controller, as described in Remark 3). Such a controller could be designed using a convex relaxation or the Riccati equation for the LQG problem. To evaluate the optimality guarantee of the designed distributed controller, one can compare $J(K_d)$ against a lower bound on $J(K_c)$ (e.g., using the SDP relaxation proposed in [32]).

V. Complexity Analysis

So far, we have proposed a unified framework to design a near-globally optimal distributed controller in three cases of deterministic systems with known initial states, deterministic systems with unknown initial states, and stochastic systems. The controller synthesis procedures are all based on Optimization C, each with a different matrix input P. On the other hand, Theorem 3 states that Optimization C has a closed-form solution. It is desirable to study the computational complexity of the proposed designed procedures, which amounts to the same complexity as finding the closed-form solution of Optimization C.

Notice that in order to find the elements of X and y in (18), one should first obtain B^TB and P. The complexity of finding B^TB is $\mathcal{O}(m^2n)$, using ordinary matrix multiplication. Furthermore, P can be found by solving the Lyapunov equation (10), (25), or (39), depending on the type of the distributed controller problem under study. The complexity of solving the Lyapunov equation is $\mathcal{O}(n^3)$ in the worst case [35]. However, in case where the optimal solution benefits from a low-rank approximation (which is the case in (10) or (25) with a small uncertainty region), there are some iterative algorithms that can be used to find the solution of (10) more efficiently. For instance, [48] shows that the low-rank property of the optimal solution can be used to design an iterative algorithm with the complexity of $\mathcal{O}(n)$ per iteration and a quadratic convergence. Next, we need to construct the matrix X and the vector y. A naive way of finding the entries of X and y is to resort to the matrix multiplications shown in (18). However, X and y can be found more efficiently. First, consider the term $M_iPM_j^T$ in (18a). If we denote the respective locations of the entries h_i and h_j in K_d as (r_i, k_i) and (r_j, k_j), one can easily verify that $\mathrm{trace}\{M_iPM_j^T\}$ is equal to $P(k_i, k_j)$ if $r_i = r_j$ and is zero otherwise. Similarly, it can be shown that $\mathrm{trace}\{M_i^TB^TBM_j\}$ is equal to the $(r_i, r_j)^{\text{th}}$ element of B^TB if $k_i = k_j$ and is zero otherwise. Therefore, instead of using matrix multiplications to find X, one can determine the entries of this matrix based on the locations of the free elements in

K_d and their corresponding entries in P and $B^T B$, which can be performed with the complexity of $\mathcal{O}(l^2)$. Likewise, finding the vector y has the complexity of $\mathcal{O}((n+m)l)$.

Finally, consider the complexity of obtaining h in the equation $Xh = y$. Using the LU factorization, this can be performed in $\mathcal{O}(l^3)$. However, notice that the matrix X is sparse in many cases where a sparse distributed controller is sought. This is due to the fact that, based on the above explanation, $X(i,j)$ is equal to zero for all pairs of h_i and h_j that do not belong to the same row or column of K_d. For instance, if the subspace of admissible distributed controllers $\mathcal{K}$ has a diagonal structure, X shares the same structure and therefore the complexity of finding h reduces to $\mathcal{O}(l)$. Inspired by this observation, one can utilize the general sparsity structure of X and find the solution of $Xh = y$ using a sparse LU factorization. The complexity analysis of solving a system of sparse linear equations using LU factorization is beyond the scope of this paper and can be found in [49].

The above analysis implies that in the natural case where $m \le n$ and $l = O(n)$, the complexity of the proposed controller design procedure is almost the same as that of the inversion of an $n \times n$ matrix. This analysis assumes that the centralized controller K_c is provided as an input. It is often the case that the complexity of finding the optimal centralized controller using Riccati equations dominates the complexity of finding its corresponding distributed controller based on the method developed in this paper.

VI. Numerical Results

Two case studies will be offered in this section to demonstrate the efficacy of the proposed controller design technique. The simulations are run on a laptop computer with an Intel Core i7 quad-core 2.50 GHz CPU and 16GB RAM. The results are reported based on a serial implementation in MATLAB.

A. Multi-Agent Systems

To illustrate the performance of the method developed in this paper on multi-agent systems, consider the planar vertical takeoff and landing (PVTOL), where the model for each aircraft (agent) is given as [50]:

$$\ddot{X}^i(t) = v(t), \quad \ddot{\theta}^i(t) = \frac{1}{\delta}\left(\sin\theta^i(t) + v^i(t)\cos\theta^i(t)\right) \tag{43}$$

Note that X and θ are the horizontal position and angle of aircraft i, respectively, and δ depends on the coupling between the rolling moment and lateral acceleration of the aircraft. Assuming that all agents are stabilized vertically, we only consider their horizontal position in this problem. Consider the feedback rule

$$v^i(t) = \alpha\dot{X}^i(t) + \beta\theta^i(t) + \gamma\dot{\theta}^i(t) + u^i(t) \tag{44}$$

for the control of aircraft i. The first three terms in (44) are used for the internal stability of the horizontal speed and angle of each aircraft. The last term needs to be designed using a controller with the structure to be delineated so that the agents maintain their relative positions. In particular, the goal is to design a distributed controller K_d that minimizes the objective function

$$J_c(K_d) = \int_0^\infty \left(y(t)^\top Q y(t) + u(t)^\top R u(t)\right) dt \tag{45}$$

while ensuring the stability of the closed-loop system. The structure of the distributed controller is captured via an undirected graph. If there exists an edge between agents i and j, it means that both agents have access to their relative distance. To illustrate our technique, consider a system consisting of 4 aircraft whose communication structure is in the form of a path graph. This communication structure also defines the sparsity of the to-be-designed distributed controller. Assume that the desired distance between adjacent agents is equal to d. By defining the state of aircraft $i \in \{1,2,3,4\}$ as

$$x^i(t) = [\dot{X}^i(t),\ \theta^i(t),\ \dot{\theta}^i(t)]^\top \tag{46}$$

the linearized model of each agent can be described as (1) where

$$A = \begin{bmatrix} \alpha & \beta & \gamma \\ 0 & 0 & 1 \\ \frac{\alpha}{\delta} & \frac{\beta+1}{\delta} & \frac{\gamma}{\delta} \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \\ \frac{1}{\delta} \end{bmatrix} \tag{47}$$

(please refer to [51] for more details). Note that the parameters α, β, and γ are used to guarantee the internal stability of each aircraft. As explained in [51], for $\delta = 0.1$, the values $\alpha = 90.62$, $\beta = -42.15$ and $\gamma = -13.22$ for the state feedback controller of each agent ensure their internal stability. Since the agents have access to their relative distance from their neighbors, we define

$$z^i(t) = [X^i(t) - X^{i+1}(t) - d,\ x^i(t)^\top]^\top \tag{48}$$

for $1 \le i \le 3$ and $z^4(t) = x^4(t)$. Based on the above definition, the state-space model of the whole system can be described as

$$\dot{z}(t) = \begin{bmatrix} \tilde{A} & H_4 & 0 & 0 \\ 0 & \tilde{A} & H_4 & 0 \\ 0 & 0 & \tilde{A} & H_3 \\ 0 & 0 & 0 & A \end{bmatrix} z(t) + \begin{bmatrix} \tilde{B} & 0 & 0 & 0 \\ 0 & \tilde{B} & 0 & 0 \\ 0 & 0 & \tilde{B} & 0 \\ 0 & 0 & 0 & B \end{bmatrix} u(t) \tag{49}$$

Note that the vectors $z(t)$ and $u(t)$ are the concatenations of $z^i(t)$ and $u^i(t)$ for all agents, respectively. Moreover, H_4 is a 4×4 matrix whose $(i,j)^{\text{th}}$ entry is equal to -1 if $(i,j) = (1,2)$ and is zero otherwise. Similarly, H_3 is a 3×3 matrix whose $(i,j)^{\text{th}}$ entry is equal to -1 if $(i,j) = (1,1)$ and is zero otherwise. Finally, $\tilde{A}$ and $\tilde{B}$ are defined as

$$\tilde{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & \alpha & \beta & \gamma \\ 0 & 0 & 0 & 1 \\ 0 & \frac{\alpha}{\delta} & \frac{\beta+1}{\delta} & \frac{\gamma}{\delta} \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \frac{1}{\delta} \end{bmatrix} \tag{50}$$

The structure of the distributed controller for the system described in (49) can be viewed as

$$K_d = \left[\begin{array}{cccc|cccc|cccc|ccc} \star & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline \star & 0 & 0 & 0 & \star & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \star & 0 & 0 & 0 & \star & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \star & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{51}$$

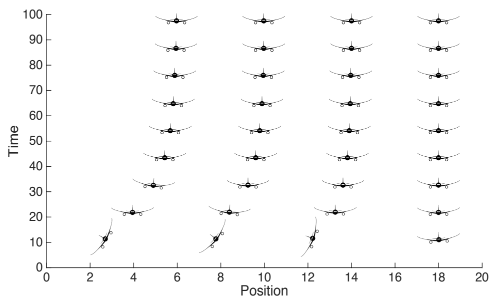


Fig. 1: Formation of 4 aircraft over the horizontal axis with $d = 4$.

where each "$\star$" corresponds to a to-be-designed free element of the distributed controller. Since the goal is to bring every aircraft to its pre-specified relative location as quickly as possible with the least amount of effort, we define the weighting matrix Q to be a diagonal matrix with $Q(k,k)$ equal to 100 if the k^{th} element of $x(t)$ corresponds to the relative positions of neighboring agents and equal to one otherwise. Furthermore, we choose R to be the identity matrix with appropriate dimension. Suppose that our estimate of the initial state of the whole system is equal to a vector a whose k^{th} element is uniformly drawn from the interval $[-2\ 2]$. Furthermore, assume that due to the uncertainty in our estimation, we consider a maximum amount of $0.2 \times |a|$ as our estimation error, where $|\cdot|$ is the entry-wise absolute value operator. This means that the initial state of the system can reside anywhere between $a - 0.2 \times |a|$ and $a + 0.2 \times |a|$. It is easy to observe that the smallest-volume outer ellipsoidal approximation of this uncertainty region can be described as $\mathcal{E} = \{a + Mu : u \in \mathbb{R}^{15\times 1}, \|u\|_2 \le 1\}$, where M is a diagonal matrix with the k^{th} diagonal entry equal to $0.2 \times |a_k| \times \sqrt{15}$.

We discretize the system using the zero-order hold method with the sampling time equal to 0.01 and then find the distributed controller via the method presented in this paper. The regularization coefficient ω in Optimization C is chosen to be 0.9. The free entries of the designed distributed controller are obtained as:

$$K_d(1,1) = -8.84,\ K_d(2,1) = 4.72,\ K_d(2,5) = -7.30 \quad (52)$$

$$K_d(3,5) = 6.18,\ K_d(3,9) = -4.90,\ K_d(4,9) = 9.67 \quad (53)$$

This controller makes the closed-loop system stable. We also find the optimal centralized LQR controller for the continuous system in order to measure the optimality guarantee of the designed distributed controller. For 100 uniformly and independently chosen initial states from the uncertainty region, the average cost function using the optimal centralized LQR controller is 7793.49, whereas the average cost function for the designed distributed controller is 8178.40 (both costs correspond to the original continuous-time system). Moreover, the average optimality guarantee of these trials is 95.28% with the standard deviation of 0.32. Figure 1 shows a snapshot of the coordination of the four aircraft for one of these trials.

To assess the performance of the proposed technique for the design of the distributed controller in larger-scale problems, consider an extended instance of the described multi-agent problem with 100 agents. This system has 399 states and 100 inputs. The initial state has an estimate of a whose elements are drawn from standard uniform distribution. Furthermore, we consider a maximum amount of $0.1 \times |a|$ as our estimation error. Similar to the previous case, one can verify that the outer ellipsoidal approximation of this uncertainty region is equal to $\mathcal{E} = \{a + Mu : u \in \mathbb{R}^{15\times 1}, \|u\|_2 \le 1\}$, where M is a diagonal matrix with the k^{th} diagonal entry equal to $0.1 \times |a_k| \times \sqrt{399}$. The design of the optimal LQR controller using the function `dlqr` in MATLAB has a runtime of 4.82 seconds. The second step is to find the unique solution of the Lyapunov equation (25) . The average elapsed time for solving this equation is 0.15 seconds using the function `dlyap`. The last step is to solve Optimization C. We use Theorem 3 to cast Optimization C as a system of linear equations. Using the function `mldivide` in MATLAB that automatically exploits the sparsity structures of X and $\mathbf{y}$, this system of linear equations can be solved in 0.18 seconds. The designed distributed controller makes the closed-loop system stable and provides an optimality guarantee of 79%. An important observation can be made based on this case study as follows: the design of a high performance and sparse controller using the proposed method can be carried out with a negligible overhead compared to the design of the optimal LQR controller. Finally, we compare our method to the ODC solver in [52], which is based on the SDP relaxation of the optimal distributed control problem. Within a time limit of 30 minutes, the solver was terminated because it did not converge to a meaningful solution. This is due to the fact that the computational complexity of solving SDPs makes them prohibitive to use in medium- and large-sized problems.

B. Power Networks

In this case study, we consider the frequency control problem for power systems. The aim is to control the frequency of a power system with a distributed controller that respects a certain sparsity structure. This sparsity structure determines which generators can share their rotor angle and frequency with each other. We consider the IEEE 39-bus New England Power System whose single line diagram is given in Figure 2. The relationship between the rotor angle of different generators and their frequency can be described by the per-unit swing equation

$$M_i \ddot{\theta}_i + D_i \dot{\theta}_i = P_{Mi} - P_{Ei} \quad (54)$$

where θ_i is the voltage (or rotor) angle at a generator bus i (in rad), P_{M_i} denotes the mechanical power input to the generator at bus i (in per unit), P_{E_i} shows the electrical active power injection at bus i (in per unit), M_i is the inertia coefficient of the generator at bus i (in pu-sec^2/rad), and D_i is the damping coefficient of the generator at bus i (in pu-sec/rad) [53]. The relationship between the electrical active power injection P_{E_i} and the voltage angles can be described by a nonlinear equation, known as AC power flow equation. In order to simplify these equations and to linearize the representation of the system, a widely-used method is to utilize the following DC power flow equations as an approximation of the nonlinear relationship between the active power injection and voltage

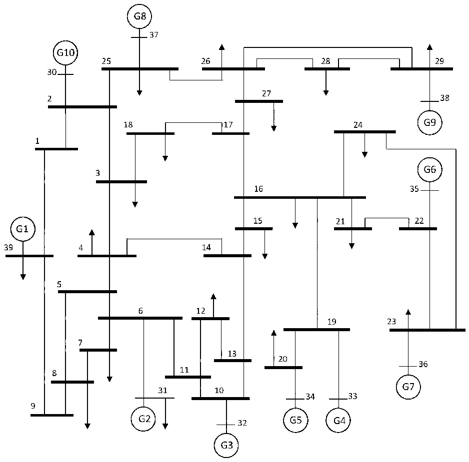


Fig. 2: Single-line diagram of IEEE 39-Bus New England Power System.

angles:

$$P_{Ei} = \sum_{j=1}^{n} B_{ij}(\theta_i - \theta_j) \tag{55}$$

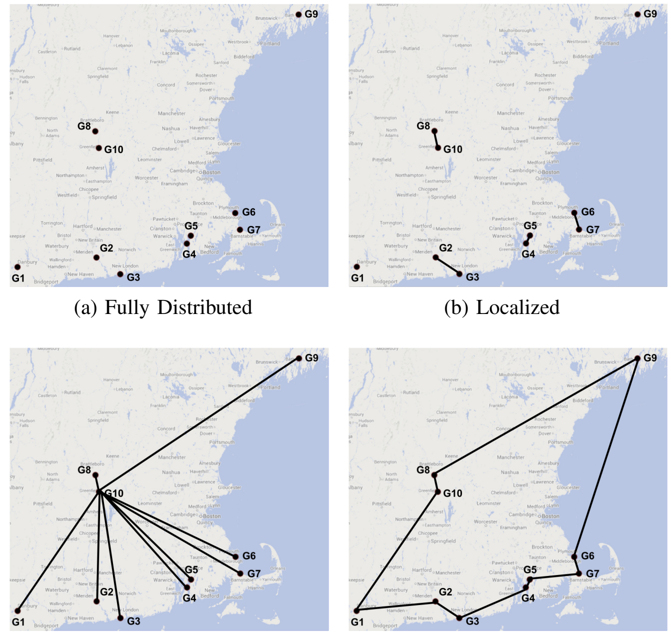
where n is the number of buses in the system and B_{ij} is the susceptance of the line (i, j). Writing (55) in a matrix form gives rise to the following state space representation of frequency control problem:

$$\begin{bmatrix} \dot{\theta}(t) \\ \dot{w}(t) \end{bmatrix} = \begin{bmatrix} 0_{n\times n} & I_n \\ -M^{-1}L & -M^{-1}D \end{bmatrix} \begin{bmatrix} \theta(t) \\ w(t) \end{bmatrix} + \begin{bmatrix} 0_{n\times n} \\ M^{-1} \end{bmatrix} P_M(t) \tag{56}$$

where $\theta(t) = [\theta_1(t), \ldots, \theta_n(t)]^T$ and $w(t) = [w_1(t), \ldots, w_n(t)]^T$ represent the state of the rotor angles and the frequency of generators at time t, respectively. Furthermore, $M = \text{diag}(M_1, \ldots, M_n)$ and $D = \text{diag}(D_1, \ldots, D_n)$.

The goal is to first discretize the system with the sampling time of 0.2 second, and then design a distributed controller to stabilize the system while achieving a high degree of optimality. The 39-bus system has 10 generators, labeled as $G_1, G_2, ..., G_{10}$. We consider four different topologies for the structure of the controller: distributed, localized, star and ring. A visual illustration of these topologies is provided in Figure 3, where each node represents a generator and each line specifies what generators are allowed to communicate. The state and input weighting matrices are chosen to be I and $0.1 \times I$, respectively.

Deterministic Case: In this experiment, the uncertainty region is considered as an sphere centered at $[1, 1, ..., 1]^T$ with the radius ψ to be specified later. In order to evaluate the performance of the proposed method, we analyze the global optimality guarantee of the designed distributed controller for different topologies with respect to the radius of the uncertainty region varied from 0.1 to 6. Furthermore, the regularization coefficient ω in Optimization C is set to 0.5. For each radius and topology, we consider 1000 independent trials with initial states uniformly chosen from the spherical uncertainty region. The results are provided in Figure 4. It can be observed that the ring topology has the best performance for different radii. The maximum and minimum optimality



(a) Fully Distributed (b) Localized

(c) Star Topology (G10 in center) (d) Ring

Fig. 3: Communication structures studied in Example 1 for the IEEE 39-Bus test System (borrowed from [32]).

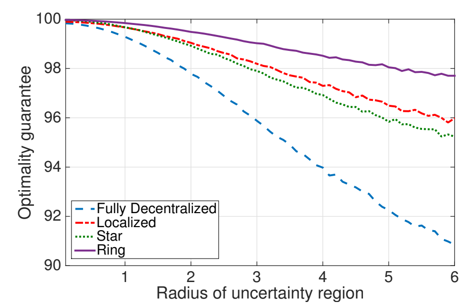


Fig. 4: Global optimality guarantee of the designed distributed controller for different topologies.

guarantees for this structure are equal to 99.97% and 97.70% (corresponding to the radii 0.1 and 6), respectively. Moreover, the worst performance corresponds to the fully distributed controller with the maximum and minimum optimality guarantees equal to 99.83% and 90.85%, respectively. Finally, it can be observed that the star and localized structures have relatively similar performances with respect to the radius of the uncertainty region.

Stochastic Case: Suppose that the power system is subject to input disturbance and measurement noise. The disturbance may be caused by certain non-dispatchable supplies and fluctuating loads. The measurement noise could arise from the inaccuracy of the rotor angle and frequency measurements. We assume that $\Sigma_d = I$ and $\Sigma_v = \sigma I$, with σ varying from 0 to 5. The ω is chosen as 0.5. First, a near-globally optimal static centralized controller K_c is designed using the SDP relaxation of the problem given in [32]. This static centralized controller is then used in Optimization C to obtain the distributed controller K_d, as explained in Remark 4. The simulation results are provided in Figure 5. It can be observed that the designed controllers are all stabilizing with no exceptions. Moreover, the global optimality guarantees for the ring, star, localized, and fully distributed topologies are above 95%, 91.8%, 88%, and 61.7%, respectively. Note that the optimality guarantee of the designed fully distributed controller is relatively low. This may be due to the potential large gap between the globally optimal costs of the optimal centralized and fully distributed control problems.

VII. Conclusion

This paper studies the optimal static distributed control problem for linear discrete-time systems. The goal is to design a stabilizing static distributed controller with a pre-defined structure, whose performance is close to that of the given (static) centralized controller. To this end, we derive a necessary and sufficient condition under which there exists a distributed controller that produces similar input and state trajectories as the optimal centralized controller for a given deterministic system with a known initial state. We then convert this condition into a convex optimization problem. We also add a regularization term into the objective of the proposed optimization problem to account for the stability of the distributed control system indirectly. This optimization problem is extended to deterministic systems with an unknown initial state belonging to a bounded uncertainty region. We derive a theoretical lower bound on the optimality guarantee of the designed distributed control, and prove that a small optimal objective value for this optimization problem brings about a high optimality guarantee for the designed distributed controller. The proposed optimization problem has a closed-form solution, which depends on the optimal centralized controller as well as the prescribed sparsity pattern for the unknown distributed controller. The results are then extended to stochastic systems that are subject to input disturbance and measurement noise. To demonstrate the performance of the developed design method, extensive simulations are performed on two real-world problems, namely aircraft formation as a multi-agent system and the frequency control problem for power systems.

References

[1] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal of Control*, vol. 6, no. 1, 1968.

[2] J. N. Tsitsiklis and M. Athans, "On the complexity of decentralized decision making and detection problems," *IEEE Conference on Decision and Control*, 1984.

[3] B. Bamieh, F. Paganini, and M. A. Dahleh, "Distributed control of spatially invariant systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 7, pp. 1091–1107, 2002.

[4] N. Motee and A. Jadbabaie, "Optimal control of spatially distributed systems," *IEEE Transactions on Automatic Control*, vol. 53, no. 7, pp. 1616–1629, 2008.

[5] Y.-S. Wang, N. Matni, and J. C. Doyle, "Localized LQR optimal control," *53rd IEEE Conference on Decision and Control*, pp. 1661–1668, 2014.

[6] Y.-S. Wang, N. Matni, and J. C. Doyle, "Localized LQR control with actuator regularization," *American Control Conference*, pp. 5205–5212, 2016.

[7] J. Lavaei, "Decentralized implementation of centralized controllers for interconnected systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 7, pp. 1860–1865, 2012.

[8] F. Lin, M. Fardad, and M. R. Jovanovic, "Augmented lagrangian approach to design of structured optimal state feedback gains," *IEEE Transactions on Automatic Control*, vol. 56, no. 12, pp. 2923–2929, 2011.

[9] J. Swigart and S. Lall, "Optimal controller synthesis for decentralized systems over graphs via spectral factorization," *IEEE Transactions on Automatic Control*, vol. 59, no. 9, pp. 2311–2323, 2014.

[10] A. Lamperski and L. Lessard, "Optimal decentralized state-feedback control with sparsity and delays," *Automatica*, vol. 58, pp. 143–151, 2015.

[11] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *IEEE Transactions on Automatic Control*, vol. 51, no. 2, pp. 274–286, 2006.

[12] G. A. de Castro and F. Paganini, "Convex synthesis of localized controllers for spatially invariant systems," *Automatica*, vol. 38, no. 3, pp. 445 – 456, 2002.

[13] 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.

[14] X. Qi, M. Salapaka, P. Voulgaris, and M. Khammash, "Structured optimal and robust control with multiple criteria: a convex solution," *IEEE Transactions on Automatic Control*, vol. 49, no. 10, pp. 1623–1640, 2004.

[15] K. Dvijotham, E. Theodorou, E. Todorov, and M. Fazel, "Convexity of optimal linear controller design," *Conference on Decision and Control*, 2013.

[16] N. Matni and J. C. Doyle, "A dual problem in H_2 decentralized control subject to delays," *American Control Conference*, 2013.

[17] Y.-S. Wang, N. Matni, and J. C. Doyle, "Localized LQR optimal control," *arXiv preprint arXiv:1409.6404*, 2014.

[18] N. Matni, "Distributed control subject to delays satisfying an H_1 norm bound," in *Conference Decision and Control*, 2014.

[19] Y. Wang, J. Lopez, and M. Sznaier, "A convex optimization approach to synthesizing sparse dynamic output feedback controllers," *IFAC-PapersOnLine*, vol. 48, no. 14, pp. 95–100, 2015.

[20] R. Arastoo, M. Bahavarnia, M. V. Kothare, and N. Motee, "Closed-loop feedback sparsification under parametric uncertainties," *IEEE 55th Conference on Decision and Control (CDC)*, pp. 123–128, 2016.

[21] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *IEEE Transactions on Automatic Control*, vol. 51, no. 2, pp. 274–286, 2006.

[22] P. Shah and P. A. Parrilo, "H_2-optimal decentralized control over posets: a state-space solution for state-feedback," *http://arxiv.org/abs/1111.1498*, 2011.

[23] L. Lessard and S. Lall, "Optimal controller synthesis for the decentralized two-player problem with output feedback," *American Control Conference*, 2012.

[24] A. Lamperski and J. C. Doyle, "Output feedback H_2 model matching for decentralized systems with delays," *American Control Conference*, 2013.

[25] M. Rotkowitz and N. Martins, "On the nearest quadratically invariant information constraint," *IEEE Transactions on Automatic Control*, vol. 57, no. 5, pp. 1314–1319, 2012.

[26] T. Tanaka and C. Langbort, "The bounded real lemma for internally positive systems and H-infinity structured static state feedback," *IEEE Transactions on Automatic Control*, vol. 56, no. 9, pp. 2218–2223, 2011.

[27] A. Rantzer, "Distributed control of positive systems," *http://arxiv.org/abs/1203.0047*, 2012.

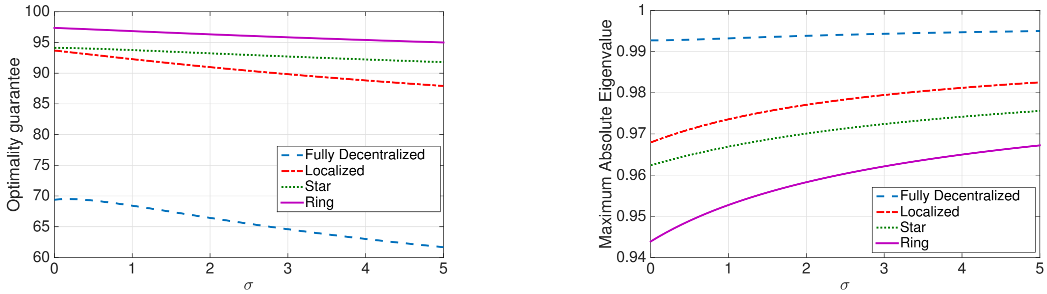
[28] S. Sojoudi and J. Lavaei, "On the exactness of semidefinite relaxation for nonlinear optimization over graphs: Part I," *IEEE Conference on Decision and Control*, 2013.

[29] S. Sojoudi and J. Lavaei, "On the exactness of semidefinite relaxation for nonlinear optimization over graphs: Part II," *IEEE Conference on Decision and Control*, 2013.

[30] J. Lavaei, "Optimal decentralized control problem as a rank-constrained optimization," *Allerton*, 2013.

[31] A. Kalbat, R. Madani, G. Fazelnia, and J. Lavaei, "Efficient convex relaxation for stochastic optimal distributed control problem," *Allerton*, 2014.

[32] G. Fazelnia, R. Madani, A. Kalbat, and J. Lavaei, "Convex relaxation for optimal distributed control problems," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 206–221, 2017.



(a) Optimality Guarantee

(b) Maximum Absolute Eigenvalues

Fig. 5: Global optimality guarantee and maximum absolute eigenvalue of $A + BK_d$ for four different topologies and different values of σ, under the assumptions $\Sigma_d = I$ and $\Sigma_v = \sigma I$ (stochastic case).

[33] S. Fattahi and J. Lavaei, "On the convexity of optimal decentralized control problem and sparsity path," *To appear in IEEE American Control Conference*, 2017.

[34] P. Benner and H. Mena, "BDF methods for large-scale differential Riccati equations," *Proc. of Mathematical Theory of Network and Systems*, vol. 10, 2004.

[35] P. Benner, J.-R. Li, and T. Penzl, "Numerical solution of large-scale Lyapunov equations, Riccati equations, and linear-quadratic optimal control problems," *Numerical Linear Algebra with Applications*, vol. 15, no. 9, pp. 755–777, 2008.

[36] L. Grasedyck, W. Hackbusch, and B. N. Khoromskij, "Solution of large scale algebraic matrix Riccati equations by use of hierarchical matrices," *Computing*, vol. 70, no. 2, pp. 121–165, 2003.

[37] E. K.-w. Chu and P. C.-Y. Weng, "Large-scale discrete-time algebraic Riccati equations—doubling algorithm and error analysis," *Journal of Computational and Applied Mathematics*, vol. 277, pp. 115–126, 2015.

[38] J. Lavaei, A. Momeni, and A. G. Aghdam, "A model predictive decentralized control scheme with reduced communication requirement for spacecraft formation," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 2, pp. 268–278, 2008.

[39] D. M. Stipanović, G. Inalhan, R. Teo, and C. J. Tomlin, "Decentralized overlapping control of a formation of unmanned aerial vehicles," *Automatica*, pp. 1285–1296, 2004.

[40] L. Grubišić and D. Kressner, "On the eigenvalue decay of solutions to operator Lyapunov equations," *Systems and control letters*, vol. 73, pp. 42–47, 2014.

[41] J. Baker, M. Embree, and J. Sabino, "Fast singular value decay for Lyapunov solutions with nonnormal coefficients," *SIAM Journal on Matrix Analysis and Applications*, vol. 36, no. 2, pp. 656–668, 2015.

[42] T. Penzl, "Eigenvalue decay bounds for solutions of Lyapunov equations: the symmetric case," *Systems and Control Letters*, vol. 40, no. 2, pp. 139–144, 2000.

[43] G. Dullerud and F. Paganini, *Course in Robust Control Theory*. Springer-Verlag New York, 2000.

[44] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, 2004.

[45] T. Mori, "On the relationship between the spectral radius and stability radius for discrete systems," *IEEE Transactions on Automatic Control*, vol. 35, no. 7, 1990.

[46] M. K. Tippett and D. Marchesin, "Upper bounds for the solution of the discrete algebraic Lyapunov equation," *Automatica*, vol. 35, no. 8, pp. 1485–1489, 1999.

[47] S. Fattahi, G. Fazelnia, and J. Lavaei, "Transformation of optimal centralized controllers into near-globally optimal static distributed controllers," *http://www.ieor.berkeley.edu/~lavaei/ODC_Journal_2016.pdf*, 2016.

[48] T. Penzl, "A cyclic low-rank smith method for large sparse Lyapunov equations," *SIAM Journal on Scientific Computing*, vol. 21, no. 4, pp. 1401–1418, 1999.

[49] P. E. Gill, W. Murray, M. A. Saunders, and M. H. Wright, "Maintaining LU factors of a general sparse matrix," *Linear Algebra and its Applications*, vol. 88, pp. 239–270, 1987.

[50] J. Hauser, S. Sastry, and G. Meyer, "Nonlinear control design for slightly non-minimum phase systems: Application to v/stol aircraft," *Automatica*, pp. 665–679, 1992.

[51] M. Arcak, "Synchronization and pattern formation in diffusively coupled systems," *IEEE Conference on Decision and Control*, pp. 7184–7192, 2012.

[52] G. Fazelnia and J. Lavaei, "ODC Solver," *http://www.ieor.berkeley.edu/~lavaei/Software.html*.

[53] M. A. Pai, *Energy Function Analysis for Power System Stability*. Kluwer Academic Publishers, Boston, 1989.

[54] S.-D. Wang, T.-S. Kuo, and C.-F. Hsu, "Trace bounds on the solution of the algebraic matrix Riccati and Lyapunov equation," *IEEE Transactions on Automatic Control*, vol. 31, no. 7, pp. 654–656, 1986.

Appendix

Proof of Theorem 1: Note that

$$x_d[\tau] = (A + BK_d)^\tau x, \quad x_c[\tau] = (A + BK_c)^\tau x \tag{57}$$

First, we prove that (7b) holds if and only if

$$B(K_c - K_d)(A + BK_c)^\tau x = 0, \qquad \tau = 0, 1, 2, \ldots \tag{58}$$

To prove the necessity part, suppose that $x_d[\tau] = x_c[\tau]$ and $x_d[\tau+1] = x_c[\tau+1]$. One can write

$$\begin{aligned} 0 &= (A + BK_d)^{\tau+1} x - (A + BK_c)^{\tau+1} x \\ &= (A + BK_d)(A + BK_c)^\tau x - (A + BK_c)^{\tau+1} x \\ &= B(K_c - K_d)(A + BK_c)^\tau x \end{aligned}$$

To prove the sufficiency part, we use a mathematical induction. The validity of the base case can be easily verified. Assume that $x_d[k] = x_c[k]$ for $\tau = k$, and consider the case $\tau = k + 1$. It follows from the equality $BK_c(A + BK_c)^\tau x = BK_d(A + BK_d)^\tau x$ of the induction step that

$$\begin{aligned} (A + BK_c)^{k+1} x &= A(A + BK_d)^k x + BK_d(A + BK_d)^k x \\ &= (A + BK_d)^{k+1} x \end{aligned}$$

Next, we show that (7a) implies (8). To this end, assume that the equation (7a) is satisfied. Since $x_c[0] = x_d[0] = x$, the relation $x_c[\tau] = x_d[\tau]$ holds for every nonnegative integer τ (note that the system (1) generates identical state signals under two identical input signals $u_c[\tau]$ and $u_d[\tau]$). Now, one can write

$$u_c[\tau] = K_c x_c[\tau] = K_c(A + BK_c)^\tau x \tag{59a}$$

$$u_d[\tau] = K_d x_d[\tau] = K_d(A + BK_d)^\tau x \tag{59b}$$

On the other hand, the relation $x_c[\tau] = x_d[\tau]$ can be expressed as

$$(A + BK_c)^\tau x = (A + BK_d)^\tau x \tag{60}$$

Combining (59) and (60) leads to (8).

To prove that (8) implies (7a), suppose that the equation (8) is satisfied. By pre-multiplying the left side of (8) with B, it follows from (58) that $x_c[\tau] = x_d[\tau]$. Therefore,

$$\begin{aligned} u_c[\tau] - u_d[\tau] &= K_c x_c[\tau] - K_d x_d[\tau] \\ &= K_c x_c[\tau] - K_d x_c[\tau] \\ &= (K_c - K_d)(A + BK_c)^\tau x = 0 \end{aligned} \tag{61}$$

This yields the equation (8), and completes the proof □

Proof of Theorem 4: The optimization problem (23) can be written in the form of

$$\min \quad \alpha \tag{62a}$$
$$\text{s.t.} \quad \mathcal{L}(P, x) \leq \alpha I, \qquad \forall x \in \mathcal{E} \tag{62b}$$
$$-\alpha I \leq \mathcal{L}(P, x), \quad \forall x \in \mathcal{E} \tag{62c}$$
$$P \geq 0 \tag{62d}$$

Notice that (62b) is equivalent to $y^\top \mathcal{L}(P, x) y \leq \alpha$ for every $x \in \mathcal{E}$ and y such that $\|y\|_2 = 1$. This implies that (62b) is equivalent to

$$y^\top \mathcal{L}(P, 0) y + \max_{\substack{x = a + Mu \\ \|u\|_2 = 1}} (y^\top x)^2 \leq \alpha \tag{63}$$

for every y such that $\|y\|_2 = 1$. Now, consider

$$\max_{\substack{x = a + Mu \\ \|u\|_2 \leq 1}} \{(y^\top x)^2\} \tag{64}$$

Using S-procedure, it can be easily verified that (64) is equal to $(|a^\top y| + \|My\|_2)^2$. Therefore, (63) can be reduced to

$$y^\top \mathcal{L}(P, 0) y + (|a^\top y| + \|My\|_2)^2 \leq \alpha \tag{65}$$

Now, consider (62c). Similar to the previous case, this constraint is equivalent to

$$-\alpha \leq y^\top \mathcal{L}(P, 0) y + \min_{\substack{x = a + Mu \\ \|u\|_2 \leq 1}} (y^\top x)^2 \tag{66}$$

for every y such that $\|y\|_2 = 1$. As before, one can use strong duality to show that

$$\min_{\substack{x = a + Mu \\ \|u\|_2 \leq 1}} \{(y^\top x)^2\} = \left((|a^\top y| - \|My\|_2)_+\right)^2 \tag{67}$$

Therefore, (66) is equivalent to

$$-\alpha \leq y^\top \mathcal{L}(P, 0) y + \left((|a^\top y| - \|My\|_2)_+\right)^2 \tag{68}$$

Now, it follows from (65) and (68) that the inequalities

$$0 \leq \alpha + y^\top \mathcal{L}(P, 0) y + \left(|a^\top y| - \|My\|_2\right)^2 \tag{69a}$$
$$0 \leq \alpha - y^\top \mathcal{L}(P, 0) y - (|a^\top y| + \|My\|_2)^2 \tag{69b}$$

should be satisfied for every y such that $\|y\|_2 = 1$ (note that we did not use the "+" operator in the above equations). Combining (69a) and (69b), one can verify that the inequality

$$2(|a^\top y|)\|My\|_2 \leq \alpha \tag{70}$$

is satisfied for every y such that $\|y\|_2 = 1$. This implies that

$$s(\mathcal{E}) \leq \alpha \tag{71}$$

Next, we will show that the defined pair of (β, P^*) is indeed feasible for (62). First, notice that since P^* satisfies (25) and $aa^T + M^2 \geq 0$, we have $P^* \geq 0$. Next, we show the feasibility of (62b) and (62c) via their equivalence to (65) and (68), respectively. Combining the definitions of β and P^* with (65) yields that

$$0 \leq \beta + (a^\top y)^2 + \|My\|_2^2 - (|a^\top y| + \|My\|_2)^2 \tag{72}$$

This is equivalent to

$$2|a^\top y|\|My\|_2 \leq \beta \tag{73}$$

which holds for every y such that $\|y\|_2 = 1$, due to the definition of β. This implies that (β, P^*) satisfies (62b). Similarly, one can substitute β and P^* in (68) to derive the inequality

$$0 \leq \beta - (a^\top y)^2 - \|My\|_2^2 + ((|a^\top y| - \|My\|_2)_+)^2 \tag{74}$$

If $|a^\top y| \geq \|My\|_2$, the above inequality holds due to (73). Now, assume that $|a^\top y| < \|My\|_2$. We need to show that

$$(a^\top y)^2 + \|My\|_2^2 \leq \beta \tag{75}$$

for every y such that $\|y\|_2 = 1$. However, note that

$$(a^\top y)^2 + \|My\|_2^2 \leq 2\|My\|_2^2 \leq 2(\lambda_M^{\max})^2 \leq \beta \tag{76}$$

which certifies that (75) holds for every feasible y. This implies that (β, P^*) satisfies (62c) and, hence, it is feasible for (62). The second part of the theorem follows from the definition of β and the fact that $s(\mathcal{E}) \leq \alpha^*$ (due to (71)). □

Proof of Lemma 2: Define $\Delta x[\tau] = x_d[\tau] - x_c[\tau]$ for $\tau = 0, 1, 2, \ldots$. First, we prove the inequality (27a). It is straightforward to verify that

$$\Delta x[\tau + 1] = B(K_d - K_c)x_c[\tau] + (A + BK_d)\Delta x[\tau] \tag{77}$$

Consider the eigen-decomposition of $A + BK_d$ as $V^{-1}DV$. Define $\Delta\tilde{x}[\tau] = V\Delta x[\tau]$. Multiplying both sides of (77) by V yields that

$$\Delta\tilde{x}[\tau + 1] = VB(K_d - K_c)x_c[\tau] + D\Delta\tilde{x}[\tau] \tag{78}$$

Taking the 2-norm from both sides of (78) leads to

$$\begin{aligned} \|\Delta\tilde{x}[\tau + 1]\|_2 &\leq \|VB(K_d - K_c)x_c[\tau]\|_2 + \|D\|_2 \times \|\Delta\tilde{x}[\tau]\|_2 \\ &\leq \|VB(K_d - K_c)x_c[\tau]\|_2 + \rho(A + BK_d) \times \|\Delta\tilde{x}[\tau]\|_2 \end{aligned} \tag{79}$$

(note that $\|D\Delta\tilde{x}[\tau]\|_2 \leq \|D\|_2\|\Delta\tilde{x}[\tau]\|_2$ and $\|D\|_2 \leq \rho(A + BK_d)$). It can be concluded from (79) that

$$\begin{aligned} (\|\Delta\tilde{x}[\tau + 1]\|_2 - \rho(A + BK_d) \times \|\Delta\tilde{x}[\tau]\|_2)^2 \\ \leq \|VB(K_d - K_c)x_c[\tau]\|_2^2 \end{aligned} \tag{80}$$

or equivalently

$$\begin{aligned} \|\Delta\tilde{x}[\tau + 1]\|_2^2 + \rho(A + BK_d)^2\|\Delta\tilde{x}[\tau]\|_2^2 \\ - 2\rho(A + BK_d)\|\Delta\tilde{x}[\tau + 1]\|_2\|\Delta\tilde{x}[\tau]\|_2 \\ \leq \|VB(K_d - K_c)x_c[\tau]\|_2^2 \end{aligned} \tag{81}$$

Summing up both sides of (81) over all values of τ gives rise to the inequality

$$\sum_{\tau=0}^{\infty}\|\Delta\tilde{x}[\tau+1]\|_2^2+\rho(A+BK_d)^2\sum_{\tau=0}^{\infty}\|\Delta\tilde{x}[\tau]\|_2^2 - 2\rho(A+BK_d)\sum_{\tau=0}^{\infty}\|\Delta\tilde{x}[\tau+1]\|_2\|\Delta\tilde{x}[\tau]\|_2 \leq \sum_{\tau=0}^{\infty}\|VB(K_d-K_c)(x_c[\tau])\|_2^2 \quad (82)$$

Using the Cauchy-Schwarz inequality, one can write

$$\sum_{\tau=0}^{\infty}\|\Delta\tilde{x}[\tau+1]\|_2^2+\rho(A+BK_d)^2\sum_{\tau=0}^{\infty}\|\Delta\tilde{x}[\tau]\|_2^2 - 2\rho(A+BK_d)\sqrt{(\sum_{\tau=0}^{\infty}\|\Delta\tilde{x}[\tau+1]\|_2^2)(\sum_{\tau=0}^{\infty}\|\Delta\tilde{x}[\tau]\|_2^2)} \leq \sum_{\tau=0}^{\infty}\|VB(K_d-K_c)(x_c[\tau])\|_2^2 \quad (83)$$

Note that $\sum_{\tau=0}^{\infty}\|\Delta\tilde{x}[\tau+1]\|_2^2=\sum_{\tau=0}^{\infty}\|\Delta\tilde{x}[\tau]\|_2^2$. Hence, it can be inferred from (83) and (28) that

$$\sum_{\tau=0}^{\infty}\|\Delta\tilde{x}[\tau]\|_2^2 \leq \left(\frac{\|V\|_2\|B\|_2}{1-\rho(A+BK_d)}\right)^2 C_1(K_d,P_x) \quad (84)$$

Since $\Delta x[\tau]=V^{-1}\Delta\tilde{x}[\tau]$, one can write

$$\begin{aligned}\sum_{\tau=0}^{\infty}\|\Delta x[\tau]\|_2^2 &\leq \|V^{-1}\|_2^2\sum_{\tau=0}^{\infty}\|\Delta\tilde{x}[\tau]\|_2^2\\ &\leq \left(\frac{\|V^{-1}\|_2\|V\|_2\|B\|_2}{1-\rho(A+BK_d)}\right)^2 C_1(K_d,P_x)\end{aligned} \quad (85)$$

This proves the inequality (27a). The above argument can be adopted to prove (27b) after noting that

$$\Delta u[\tau]=(K_d-K_c)x_c[\tau]+K_d\Delta x[\tau] \quad (86)$$

where $\Delta u[\tau]=u_d[\tau]-u_c[\tau]$. □

Proof of Theorem 5: According to Lemma 2, one can write

$$\sum_{\tau=0}^{\infty}\|x_d[\tau]\|_2^2+\sum_{\tau=0}^{\infty}\|x_c[\tau]\|_2^2 \leq \left(\frac{\kappa(V)\|B\|_2}{1-\rho(A+BK_d)}\right)^2 C_1(K_d,P_x) + 2\sum_{\tau=0}^{\infty}\|x_c[\tau]\|_2\|x_d[\tau]\|_2 \quad (87)$$

Dividing both sides of (87) by $\sum_{\tau=0}^{\infty}\|x_c[\tau]\|_2^2$ and using the Cauchy-Schwarz inequality for $\sum_{\tau=0}^{\infty}\|x_c[\tau]\|_2\|x_d[\tau]\|_2$ yield that

$$\frac{\sum_{\tau=0}^{\infty}\|x_d[\tau]\|_2^2}{\sum_{\tau=0}^{\infty}\|x_c[\tau]\|_2^2} \leq \left(1+\frac{\kappa(V)\|B\|_2}{(1-\rho(A+BK_d))\sqrt{\sum_{\tau=0}^{\infty}\|x_c[\tau]\|_2^2}}\sqrt{C_1(K_d,P_x)}\right)^2 \quad (88)$$

Likewise,

$$\frac{\sum_{\tau=0}^{\infty}\|u_d[\tau]\|_2^2}{\sum_{\tau=0}^{\infty}\|u_c[\tau]\|_2^2} \leq \left(1+\frac{1-\rho(A+BK_d)+\kappa(V)\|K_d\|_2\|B\|_2}{(1-\rho(A+BK_d))\sqrt{\sum_{\tau=0}^{\infty}\|u_c[\tau]\|_2^2}}\sqrt{C_1(K_d,P_x)}\right)^2 \quad (89)$$

Combining (88) and (89) leads to

$$\begin{aligned}\frac{J(K_d)}{J(K_c)} &= \frac{\sum_{\tau=0}^{\infty}\|x_d[\tau]\|_2^2+\sum_{\tau=0}^{\infty}\|u_d[\tau]\|_2^2}{\sum_{\tau=0}^{\infty}\|x_c[\tau]\|_2^2+\sum_{\tau=0}^{\infty}\|u_c[\tau]\|_2^2}\\ &\leq (1+\zeta\sqrt{C_1(K_d,P_x)})^2\end{aligned} \quad (90)$$

On the other hand, we have

$$\begin{aligned}&\text{trace}\{(K_c-K_d)P_x(K_c-K_d)^T\}\\ &=\text{trace}\{(K_c-K_d)P(K_c-K_d)^T\}\\ &\quad+\text{trace}\{(K_c-K_d)(P_x-P)(K_c-K_d)^T\}\end{aligned} \quad (91)$$

According to Lemma 4, one can verify that

$$\begin{aligned}&\text{trace}\{(K_c-K_d)(P_x-P)(K_c-K_d)^T\}\\ &=\text{trace}\{(K_c-K_d)^T(K_c-K_d)(P_x-P)\}\\ &\overset{(a)}{\leq} \lambda_{\max}(P_x-P)\text{trace}\{(K_c-K_d)^T(K_c-K_d)\}\\ &\leq \|P-P_x\|_2\|K_c-K_d\|_F^2\\ &\leq \frac{\|K_c-K_d\|_F^2}{r(A+BK_c)^2}\alpha\end{aligned} \quad (92)$$

where the third line (a) is due to the inequality

$$\text{trace}\{XY\}\leq\lambda_{\max}(X)\text{trace}\{Y\} \quad (93)$$

for a symmetric matrix X and a positive semi-definite matrix Y (please refer to [54]). Hence, the relation

$$C_1(K_d,P_x)\leq C_1(K_d,P)+\frac{\|K_c-K_d\|_F^2}{r(A+BK_c)^2}\alpha \quad (94)$$

holds for every $x\in\mathcal{E}$. The proof is completed after combining (94) with (89).

Proof of Lemma 5: It is straightforward to verify that

$$\begin{aligned}\mathcal{E}\left\{\|x_c[\infty]-x_d[\infty]\|_2^2\right\} &= \text{trace}\left\{\mathcal{E}\left\{x_d[\infty]x_d[\infty]^T\right\}\right\}\\ &+\text{trace}\left\{\mathcal{E}\left\{x_c[\infty]x_c[\infty]^T\right\}\right\}\\ &-\text{trace}\left\{\mathcal{E}\left\{x_d[\infty]x_c[\infty]^T\right\}\right\}\\ &-\text{trace}\left\{\mathcal{E}\left\{x_c[\infty]x_d[\infty]^T\right\}\right\}\end{aligned} \quad (95)$$

On the other hand, since $d[\cdot]$ and $v[\cdot]$ are independent and identically distributed random vectors, the equation $\mathcal{E}\left\{d[\tau_1]d[\tau_2]^T\right\}=\mathcal{E}\left\{v[\tau_1]v[\tau_2]^T\right\}=0$ holds for every two different indices τ_1 and τ_2. In addition, we have $\mathcal{E}\left\{v[\tau_1]d[\tau_2]^T\right\}=\mathcal{E}\left\{d[\tau_1]v[\tau_2]^T\right\}=0$, for all nonnative integers τ_1 and τ_2. Therefore,

$$x[\infty]=\lim_{\tau\to\infty}\sum_{i=0}^{\tau-1}(A+BK)^{\tau-1-i}(Ed[i]+BKFv[i]) \quad (96)$$

yields

$$\begin{aligned}&\mathcal{E}\left\{x_d[\tau]x_d[\tau]^T\right\}=P_1, \quad \mathcal{E}\left\{x_c[\tau]x_c[\tau]^T\right\}=P_2,\\ &\mathcal{E}\left\{x_d[\tau]x_c[\tau]^T\right\}=P_3, \quad \mathcal{E}\left\{x_c[\tau]x_d[\tau]^T\right\}=P_4,\end{aligned} \quad (97)$$

where P_1, P_2, P_3 and P_4 satisfy (38). Note that (38a) and (38b) are Lyapunov equations, whereas (38c) and (38d) are Stein equations, which all have unique solutions since $A+BK_d$ and $A+BK_c$ are stable. This completes the proof. $\square$

Murat Arcak received the B.S. degree in electrical engineering from the Bogazici University, Istanbul, Turkey, in 1996 and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of California, Santa Barbara, CA, USA, in 1997 and 2000, respectively. Currently, he is a Professor at University of California, Berkeley, CA, USA, in the Electrical Engineering and Computer Sciences Department. Prior to joining the University of California, Berkeley, in 2008, he was a faculty member at the Rensselaer Polytechnic Institute, Troy, NY, USA. His research is in dynamical systems and control theory with applications to synthetic biology, multi-agent systems, and transportation. Prof. Arcak received a CAREER Award from the National Science Foundation in 2003, the Donald P. Eckman Award from the American Automatic Control Council in 2006, the Control and Systems Theory Prize from the Society for Industrial and Applied Mathematics (SIAM) in 2007, and the Antonio Ruberti Young Researcher Prize from the IEEE Control Systems Society in 2014. He is a member of SIAM.

Salar Fattahi is a Ph.D. student in Industrial Engineering and Operations Research at University of California, Berkeley. He received the B.Sc. in Electrical Engineering from Sharif University of Technology, Tehran, Iran in 2014 and the M.Sc. degree in Electrical Engineering from Columbia University, New York, NY, USA. He is the recipient of IEEE Transactions of Smart Grid best reviewer award in 2017, Outstanding Graduate Student Instructor award from University of California, Berkeley in 2017, and Armstrong fellowship award from Columbia University in 2014.

Ghazal Fazelnia is a Ph.D. student in Electrical Engineering at Columbia University. She received the B.Sc. in Electrical Engineering from Sharif University of Technology, Tehran, Iran in 2013. She is a Bronze Medalist in National Mathematics Olympiad, Iran in 2008 and recipient of Microsoft Research Graduate Womens Scholarship in 2015.

Javad Lavaei is an Assistant Professor in the Department of Industrial Engineering and Operations Research at University of California, Berkeley. He obtained his Ph.D. degree in Computing and Mathematical Sciences from California Institute of Technology in 2011. Prof. Lavaei has won multiple awards, including NSF CAREER Award, Office of Naval Research Young Investigator Award, Google Faculty Research Award, the Canadian Governor General's Gold Medal, Northeastern Association of Graduate Schools Master's Thesis Award, and Silver Medal in the 1999 International Mathematical Olympiad. He is an Associate Editor of the IEEE Transactions on Smart Grid and serves on the conference editorial board for both the IEEE Control Systems Society and the European Control Association. He is a co-recipient of the 2015 INFORMS Optimization Society Prize for Young Researchers, and the recipient of the 2016 Donald P. Eckman Award from the American Automatic Control Council.