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

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

Abstract—This paper is concerned with the optimal 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, a necessary and sufficient condition is derived to guarantee the existence of a distributed controller that generates the same input and state trajectories as the optimal centralized controller for a given deterministic system. This condition is then translated into a convex optimization problem. Subsequently, a regularization term is incorporated into the objective of the proposed optimization problem to indirectly account for the stability of the distributed control system. The designed optimization problem has a closed-form solution (explicit formula) and can be efficiently solved for large-scale systems that require low computational efforts. Furthermore, strong theoretical lower bounds are derived on the optimality guarantee of the designed distributed controller in the case where the proposed conditions do not hold. 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. The results are then extended to stochastic systems that are subject to input disturbance and measurement noise. By building upon the developed methodology, we partially address some long-standing problems, such as finding the minimum number of free elements required in the distributed controller under design to achieve a performance close to the optimal centralized one. Numerical results on a power network and several random systems are reported to demonstrate the efficacy of the proposed method.

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]–[6], dynamically decoupled systems [7], [8], strongly connected systems [9], and optimal static distributed systems [10], [11].

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]–[19]. This has been carried out in the seminal work [20] by deriving a sufficient condition named quadratic invariance, which has been specialized in [21] by deploying the concept of partially ordered sets. These conditions have been further investigated in several other papers [22]–[24]. A different approach is taken in the recent papers [25] and [26], 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 [27], [28], we have shown in [29]–[31] 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, we have proved in [32] that either a large input weighting matrix or a large noise covariance can convexify the optimal distributed control problem for stable systems, and hence one can use a variety of iterative algorithms to find globally optimal distributed controllers. Since SDPs and iterative algorithms are often computationally expensive for large-scale problems, it is desirable to develop a computationally-cheap method for designing suboptimal distributed controllers.

A. Contributions

Consider the gap between the optimal costs of the optimal centralized and distributed control problems. This gap 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 an optimal centralized controller, is it possible to design a stabilizing static distributed controller with a given structure whose performance is close to that of the best 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 optimal centralized controller are identical for a given initial state. We translate the condition into an optimization problem, where the closeness of the optimal 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. This problem has a closed-form solution, which depends on the given sparsity pattern of the to-be-designed controller. Subsequently, a lower bound is obtained to guarantee the performance of the designed static distributed controller. This lower bound determines the distance between the performances of the designed controller and the optimal centralized one. We show that the proposed convex program indirectly maximizes the derived lower bound while striving to achieve a closed-loop stability. In the extreme case, if the optimal value of the optimization problem is zero, the distributed and centralized controller are equivalent in terms of their performances. 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.

The proposed technique could be used to quantify how the sparsity pattern of the unknown controller affects the performance of the optimal distributed control system. We show that if the number of free elements in each row of the unknown static distributed controller gain is higher than the rank of some Lyapunov matrix, then there exists a distributed controller with the same performance as the optimal centralized one. However, this controller may not be stabilizing and therefore having a higher number of free elements in each row of the controller gain increases the likelihood of finding both a stabilizing and a high-performance controller. Based on this observation, the subspace of high-performance distributed controllers is explicitly characterized, and an optimization problem is designed to seek a stabilizing controller in the subspace of high-performance and structurally constrained controllers. The efficacy of the developed mathematical framework is demonstrated on a power network and random systems.

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. Numerical examples are provided in Section V. Concluding remarks are drawn in Section VI. Some of the proofs are given in the appendix.

Notations: The space of real numbers is denoted by \( \mathbb{R} \). The symbols \( \text{trace}\{W\} \) and \( \text{null}\{W\} \) denote the trace and the null space of a matrix \( W \), respectively. \( I_m \) denotes the identity matrix of dimension \( m \). The symbols \( (.)^T \) and \( (.)^* \) are used for transpose and Hermitian transpose, respectively. The symbols \( |W|_2 \) and \( |W|_F \) denote the 2-norm and Frobenius norm of \( W \), respectively. The \( (i,j) \)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 Frobenius inner product of the matrices \( W_1 \) and \( W_2 \) is denoted as \( \langle W_1,W_2 \rangle \). The expected value of a random variable \( x \) is shown as \( \mathbb{E}\{x\} \). The symbol \( \lambda_{\max}(W) \) is used to show the maximum eigenvalue of a symmetric matrix \( W \). The spectral radius of \( W \) is the maximum absolute value of the eigenvalues of matrix \( W \) and is denoted by \( \rho(W) \).

II. Problem Formulation

In this paper, the optimal distributed control (ODC) problem is studied. 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 \( K \subseteq \mathbb{R}^{n \times n} \) as a linear subspace with some pre-specified sparsity pattern (enforced zeros in certain entries). A feedback gain belonging to \( K \) is called a distributed (decentralized) controller with its sparsity pattern captured by \( K \). In the case of \( K = \mathbb{R}^{n \times n} \), 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

\[
    x[\tau + 1] = Ax[\tau] + Bu[\tau], \quad \tau = 0, 1, \ldots, \infty
\]

with the known matrices \( A \in \mathbb{R}^{n \times n} \), \( B \in \mathbb{R}^{n \times m} \) and \( x[0] \in \mathbb{R}^n \). 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:

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

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

\[
    \min_{K \in \mathbb{R}^{n \times n}} J(K)
\]

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 $K$, the additional constraint $K \in K$ should be added to the optimal centralized control problem, and it is well-known that Riccati equations cannot be used to find an optimal distributed controller in general. We refer to this problem as the infinite-horizon stochastic ODC problem.

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

$$\begin{cases}
y[\tau] = x[\tau] + Fv[\tau]
\end{cases}, \quad \tau = 0, 1, 2, \ldots \quad (4)$$

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 $\tau$. Associated with the system (4) under an arbitrary controller $u[\tau] = KY[\tau]$, consider the cost functional

$$J(K) = \lim_{\tau \to +\infty} \mathbb{E} \left\{ x[\tau]^TQx[\tau] + u[\tau]^TRu[\tau] \right\} \quad (5)$$

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 $K$ (note that the operator $\lim_{\tau \to +\infty}$ in the definition of $J(K)$ can be alternatively changed to $\lim_{\tau \to +\infty}$ without affecting the solution, due to the closed-loop stability). Finding an optimal distributed controller with a pre-defined structure is NP-hard and intractable in its worst case. Therefore, we seek to find a near-globally optimal 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 deterministic system (1) with the performance index (2) or the stochastic system (4) with the performance index (5). Given a matrix $K_d \in K$ and a percentage number $\mu \in [0, 100]$, it is said that the distributed controller $u[\tau] = K_d x[\tau]$ has the global optimality guarantee of $\mu$ if

$$\frac{J(K_c)}{J(K_d)} \times 100 \geq \mu \quad (6)$$

where $K_c$ denotes an optimal centralized controller gain for the corresponding deterministic or stochastic ODC problem.

To understand Definition 2 if $\mu$ is equal to 90% for instance, it means that the distributed controller $u[\tau] = K_d x[\tau]$ is at most 10% worse than the best (static) centralized controller with respect to the cost function (2) or (5). It also implies that if there exists a better static distributed controller, it outperforms $K_d$ by at most a factor of 0.1. This paper aims to address two problems.

**Objective 1) Distributed Controller Design:** 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 $u[\tau] = K_d x[\tau]$.

**Objective 2) Minimum number of required communications:** Given the optimal centralized controller $K_c$, find the minimum number of free (nonzero) elements in the sparsity patterns imposed by $K$ that is required to ensure the existence of a stabilizing controller $K_d$ with a high optimality guarantee.

### III. Distributed Controller Design: Deterministic Systems

In this section, we study the design of static distributed controllers for deterministic systems. 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. In what follows, we will investigate these criteria.

**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. The next lemma derives a necessary and sufficient condition under which the centralized and distributed controllers generate the same state trajectory for the system (1).

**Lemma 1.** Given the optimal centralized gain $K_c$, an arbitrary distributed controller gain $K_d \in K$, and the initial state $x[0]$, the relation

$$x_c[\tau] = x_d[\tau], \quad \tau = 0, 1, 2, \ldots \quad (7)$$

holds if and only if

$$B(K_c - K_d)(A + BK_c)x[0] = 0, \quad \tau = 0, 1, 2, \ldots \quad (8)$$

**Proof.** The proof is provided in the appendix.

**Lemma 2** investigates the equivalence of the centralized and distributed controllers from the perspective of the closeness of the state trajectories. The next lemma studies the analogy of the input trajectories for the centralized and distributed control systems.

**Lemma 2.** Given the optimal centralized gain $K_c$, an arbitrary distributed controller gain $K_d \in K$, and the initial state $x[0]$, the relation

$$u_c[\tau] = u_d[\tau], \quad \tau = 0, 1, 2, \ldots \quad (9)$$

holds if and only if

$$(K_c - K_d)(A + BK_c)x[0] = 0, \quad \tau = 0, 1, 2, \ldots \quad (10)$$

**Proof.** The proof is provided in the appendix.
Using Lemmas 1 and 2, we aim to derive necessary and sufficient conditions for the equivalence of the state and input trajectories generated by the centralized and distributed controllers.

**Theorem 1.** Given the optimal centralized gain $K_c$, an arbitrary gain $K_d \in \mathcal{K}$, and the initial state $x[0]$, the relations

$$
\begin{align*}
    u_c[\tau] &= u_d[\tau], & \tau = 0, 1, 2, \ldots \\
    x_c[\tau] &= x_d[\tau], & \tau = 0, 1, 2, \ldots
\end{align*}
$$

(11a)

hold if and only if

$$
(K_c - K_d)(A + BK_c)^T x[0] = 0, \quad \tau = 0, 1, 2, \ldots
$$

(12)

**Proof.** This theorem is an immediate consequence of Lemmas 1 and 2. \( \square \)

Theorem 1 derives a necessary and sufficient condition in order for a distributed control system to perform identically to its centralized counterpart. To flourish this condition, an optimization problem will be introduced below.

**Optimization A.** This problem is defined as

$$
\begin{align*}
    \min_K \frac{1}{2} \text{trace} \left\{ (K_c - K)P(K_c - K)^T \right\} \\
    \quad \text{s.t. } K \in \mathcal{K}
\end{align*}
$$

(13a)

(13b)

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

$$
(A + BK_c)P(A + BK_c)^T - P + x[0]x[0]^T = 0
$$

(14)

Since $P$ 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 $x[0]$, the relations

$$
\begin{align*}
    u_c[\tau] &= u_d[\tau], & \tau = 0, 1, 2, \ldots \\
    x_c[\tau] &= x_d[\tau], & \tau = 0, 1, 2, \ldots
\end{align*}
$$

(15a)

(15b)

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 (12) is equivalent to the optimal objective value of Optimization A being equal to 0. To this end, define the semi-infinite matrix

$$
X = \begin{bmatrix}
    x[0] & (A + BK_c)x[0] & (A + BK_c)^2 x[0] & \cdots
\end{bmatrix}
$$

(16)

Now, observe that (12) is satisfied if and only if the Frobenius norm of $(K_c - K_d)X$ is equal to 0 or equivalently

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

(17)

On the other hand, if $P$ is defined as $XX^T$, then it is the unique solution of (14). This completes the proof. \( \square \)

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 $K_c$ 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.

**B. Stability Criterion**

In the preceding subsection, we have derived conditions to guarantee that the centralized and distributed control systems possess the same input and state trajectories for a given initial state. However, these conditions do not necessarily ensure the stability of the distributed closed-loop system. To elaborate on this statement, assume that condition (12) is satisfied, implying that $x_c[\tau] = x_d[\tau]$ and $u_c[\tau] = u_d[\tau]$ for every nonnegative integer $\tau$. Assume also that $A + BK_d$ is diagonalizable as $A + BK_D = VDV^{-1}$, where $V$ is a matrix consisting of the eigenvectors of $A + BK_d$ and $D$ is a diagonal matrix containing the eigenvalues of $A + BK_d$. One can write

$$
x_d[\tau] = (A + BK_d)^T x[0] = VDV^{-1} x[0]
$$

(18)

Moreover, due to the stability of the centralized closed-loop system, one can write

$$
0 = \lim_{\tau \to \infty} \| x_c[\tau] \|_2 = \lim_{\tau \to \infty} \| x_d[\tau] \|_2 = \lim_{\tau \to \infty} \| VDV^{-1} x[0] \|_2
$$

(19)

The above equation does not imply that all diagonal entries of $D$ have norms less than 1 (to guarantee stability). Instead, it implies that $x[0]$ is orthogonal to every eigenvector whose corresponding eigenvalue is an unstable mode.

It follows from the above discussion that whenever the centralized and distributed control systems have the same input and state trajectories, $x[0]$ 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, we introduce an optimization problem next.

**Optimization B.** This problem is defined as

$$
\begin{align*}
    \min_K \frac{1}{2} \text{trace} \left\{ (K_c - K)^T B^T B(K_c - K) \right\} \\
    \quad \text{s.t. } K \in \mathcal{K}
\end{align*}
$$

(19a)

(19b)

**Lemma 3.** There exists a strictly positive number $\epsilon$ such that an arbitrary distributed controller $u_c[\tau] = K_d x_d[\tau]$ with a gain $K_d \in \mathcal{K}$ stabilizes the system (1) if the objective value of Optimization B at the point $K_d$ is less than $\epsilon$.

**Proof.** 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.,

$$
A + BK_d = A + BK_c + B(K_d - K_c)
$$

(20)

The proof follows from the above equation. \( \square \)

Note that there are several techniques in matrix perturbation and robust control to maximize or find a sub-optimal value $\epsilon$. Note also that the stability criterion (19) 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.

**Optimization C.** Given a constant number \( \alpha \in [0, 1] \), this problem is defined as the minimization of the function

\[
C(K) = \alpha \times C_1(K) + \left(1 - \alpha \right) \times C_2(K)
\]  

with respect to the matrix variable \( K \in \mathcal{K} \), where

\[
C_1(K) = \text{trace} \left\{ (K_c - K)P(K_c - K)^T \right\} \quad (22a)
\]

\[
C_2(K) = \text{trace} \left\{ (K_c - K)^T B^T B(K_c - K) \right\} \quad (22b)
\]

Note that \( C_1(K) \) accounts for the performance of the distributed controller and \( C_2(K) \) indirectly enforces a closed-loop stability. Assume that each matrix in the space \( \mathcal{K} \) has \( l \) free entries to be designed. Denote these unknown parameters as \( h_1, h_2, ..., h_l \). Furthermore, let \( M_1, ..., M_l \in \mathbb{R}^{m \times n} \) be constant 0-1 matrices such that \( M_i(i, j) \) is equal to 1 if the pair \((i, j)\) is the location of the free entry \( h_i \) in \( K \in \mathcal{K} \) and is zero otherwise, for every \( t \in \{1, 2, ..., l\} \).

In the next subsection, we will connect the optimal objective value of Optimization C to the stability and the performance guarantee of the distributed controller under design. Before developing that result, we aim to show that the solution of Optimization C can be found via an explicit formula.

**Theorem 3.** Consider the matrix \( X \in \mathbb{R}^{l \times l} \) and vector \( y \in \mathbb{R}^l \) with the entries

\[
X(i, j) = \alpha \times \text{trace} \left\{ M_i P M_j^T \right\} + \left(1 - \alpha \right) \text{trace} \left\{ M_i^T B^T B M_j \right\}
\]

\[
y(i) = \alpha \times \text{trace} \left\{ M_i P K_c^T \right\} + \left(1 - \alpha \right) \text{trace} \left\{ M_i^T B^T B K_c \right\}
\]

for every \( i, j \in \{1, 2, ..., l\} \). 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

\[
\mathcal{K} = \left\{ \sum_{i=1}^{l} M_i h_i \left| h \in \mathbb{R}^l \right. \right\}
\]

for \( M_1, ..., M_l \in \mathbb{R}^{m \times n} \) (note that \( h_i \)'s are the entries of \( h \)). Substituting \( K_d = \sum_{i=1}^{l} M_i h_i \) into \( (21) \) and taking its gradient with respect to \( h \) lead to the optimality condition

\[
\sum_{j=1}^{l} \alpha \times \text{trace} \left\{ M_i P M_j^T \right\} h_j + \sum_{j=1}^{l} \left(1 - \alpha \right) \times \text{trace} \left\{ M_i^T B^T B M_j \right\} h_j = \alpha \times \text{trace} \left\{ M_i P K_c^T \right\} + \left(1 - \alpha \right) \times \text{trace} \left\{ M_i^T B^T B K_c \right\}
\]

The above equation can be written in a compact form as \( Xh = y \). Note that since \( (21) \) 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 \( h \).

**Remark 1.** Depending on the ranks of the matrices \( P \) and \( B \) and the positions of the free elements in each matrix \( K \in \mathcal{K} \), the equation \( Xh = y \) may have more than one solution. In other words, the null space of matrix \( X \) can have a dimension higher than 0. Due to Theorem 3, each feasible solution of \( Xh = y \) yields a solution for Optimization C. This degree of freedom enables us to obtain a set of candidate distributed controllers.

To illustrate Theorem 3, consider the case where \( m = n \) and \( \mathcal{K} \) consists of diagonal matrices. One naive strategy to design \( K_d \) is to simply remove the off-diagonal entries of \( K_c \) and keep its diagonal. However, Optimization C proposes that the diagonal elements of the distributed controller be equal to

\[
h_k = \frac{\alpha(P^k, K_c^k) + \left(1 - \alpha \right)(B^T B)_{k,k} K_c}{\alpha P(k,k) + \left(1 - \alpha \right)\|B_k\|^2_2}
\]

for \( k \in \{1, 2, ..., m\} \), where

- \( P^k \) and \( K_c^k \) denote the \( k^{th} \) rows of the matrices \( P \) and \( K_c \), respectively.
- \( (B^T B)_k \), \( K_c^k \) and \( B_k \) denote the \( k^{th} \) columns of the matrices \( B^T B, K_c \), and \( B \), respectively.

The diagonal distributed controller gain \( K_d \) proposed by the above equation has the property that the \( (k, k) \)th entry of \( K_d \) is a weighted sum of the elements of the \( k \)th row and \( k \)th column of \( K_c \), where the weights come from the Lyapunov matrix \( P \) and \( B \). The performance of the diagonal controller designed using this simple formula will be evaluated later in this paper.

D. Lower Bound on Optimality Guarantee

So far, a convex optimization problem has been designed whose explicit solution produces a distributed controller with the right sparsity pattern such that it yields the same performance as the optimal centralized controller if the optimal objective value of this optimization problem is zero. Then, it has been argued that if the objective value is not zero but small enough at optimality, then its corresponding distributed controller has a high optimality guarantee. In this section, this statement will be formalized by finding a lower bound on the global optimality guarantee of the designed distributed controller. In particular, it is aimed to show that this lower bound is in terms of the value of \( C_1(K_d) \) in \( (21) \), and that a small \( C_1(K_d) \) translates into a high optimality guarantee (where \( K_d \) is a solution of Optimization C that can be explicitly found in terms of \( K_c \) using Theorem 3). 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 4.** Given the optimal centralized gain \( K_c \) and an
arbitrary stabilizing gain $K_d \in \mathcal{K}$, the relations
\begin{align}
\sum_{\tau=0}^{\infty} \|x_d[\tau] - x_c[\tau]\|_2^2 &\leq \left( \frac{\kappa(V)\|B\|_2}{1 - \rho(A + BK_d)} \right)^2 C_1(K_d) \tag{27a} \\
\sum_{\tau=0}^{\infty} \|u_d[\tau] - u_c[\tau]\|_2^2 &\leq \left( \frac{\kappa(V)\|K_d\|_2\|B\|_2}{1 - \rho(A + BK_d)} \right)^2 C_1(K_d) \tag{27b}
\end{align}
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. \hfill \Box

Notice that, according to the statement of Lemma 4, the upper bounds in (27a) and (27b) are valid if the distributed controller gain $K_d$ makes the system stable. According to (16) and (17), one can verify that
\begin{equation}
C_1(K_d) = \sum_{\tau=0}^{\infty} \| (K_d - K_c)(A + BK_c)^\tau x(0) \|_2^2 \tag{28}
\end{equation}

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 $C_1(K)$ and $C_2(K)$. While the second term indirectly accounts for stability, the first term $C_1(K)$ directly appears in the upper bounds in (27a) and (27b). Hence, Optimization C aims to minimize the deviation between the trajectories of the distributed and centralized control systems.

**Theorem 4.** Assume that $Q = I_n$ and $R = I_m$. Given the optimal centralized gain $K_c$ and an arbitrary stabilizing gain $K_d \in \mathcal{K}$, the relations
\begin{equation}
\left( 1 + \mu \sqrt{C_1(K_d)} \right)^2 J(K_c) \geq J(K_d) \geq J(K_c) \tag{29}
\end{equation}
hold, where
\begin{equation}
\mu = \max \left\{ \frac{\kappa(V)\|B\|_2}{\sqrt{\sum_{\tau=0}^{\infty} \|x_c[\tau]\|_2^2}}, \frac{1 - \rho(A + BK_d)}{1 - \rho(A + BK_d) + \kappa(V)\|K_d\|_2\|B\|_2} \right\} \tag{30}
\end{equation}

**Proof.** The proof is provided in the appendix. \hfill \Box

Notice that whenever the optimal solution of Optimization C does not satisfy the equation $C_1(K_d) = 0$, Theorem 2 cannot be used to show the equivalence of the distributed and centralized controllers. Instead, Theorem 4 quantifies the similarity between the two control systems. It also states that one may find a distributed controller with a high performance guarantee by minimizing the objective of Optimization C. More precisely, it follows from (29) that
\begin{equation}
\frac{J(K_c)}{J(K_d)} \geq \frac{1}{(1 + \mu \sqrt{C_1(K_d)})^2} \tag{31}
\end{equation}

Since a small $C_1(K_d)$ in (31) results in a high optimality guarantee for the designed distributed controller, this theorem justifies why it is beneficial to minimize (21), which in turn minimizes $C_1(K_d)$ while striving to find a stabilizing controller. Another implication of Theorem 4 is as follows:

if there exists a better linear static distributed controller with the given structure, it outperforms $K_d$ by at most a factor of $(1 + \mu \sqrt{C_1(K_d)})^2$.

**Remark 2.** One may speculate that there is no guarantee that the parameter $\mu$ remains small if $C_1(K)$ is minimized via Optimization C. In particular, it could theoretically occur that $\rho(A + BK_d)$ approaches 1 and $\mu$ goes to infinity. However, note that $\mu$ is implicitly controlled by the term $C_2(K)$ in the objective function of Optimization C. In fact, by minimizing $\|B(K_d - K_c)\|_F$ in the objective function, it is attempted to make that eigenvalues and eigenvectors of $A + BK_d$ close to those of $A + BK_c$. This implies that $\kappa(V)$ and $\rho(A + BK_d)$ are forced to be close to $\kappa(V')$ and $\rho(A + BK_c)$, where $V'$ is the eigenvector matrix of $A + BK_c$.

**Remark 3.** The bounds in Theorem 4 are derived to substantiate the reason behind the minimization of $C(K_d)$ in Optimization C for the controller design. However, these bounds are rather conservative compared to the actual performance of the designed distributed controller $K_d$. It will be shown through simulations that while the lower bound in (31) may not be satisfactory for the optimality guarantee, the actual optimality guarantee is high and close to 100% in several examples. Finding tighter bounds on the optimality guarantee is left as future work.

**Remark 4.** Notice that Theorem 4 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 in order to use the bounds in Theorem 4. Define $Q_d$ and $R_d$ as $Q = Q_d^TQ_d$ and $R = R_d^TR_d$, respectively. The ODC problem with the tuple $(A, B, x_d[\cdot], u_d[\cdot])$ can be reformulated with respect to a new tuple $(\tilde{A}, \tilde{B}, \tilde{x}[\cdot], \tilde{u}[\cdot])$ defined as
\begin{align*}
\tilde{A} &= Q_d A Q_d^{-1}, & \tilde{B} &= Q_d B R_d^{-1}, \\
\tilde{x}[\tau] &= Q_d x_d[\tau], & \tilde{u}[\tau] &= R_d u[\tau].
\end{align*}

Furthermore, in order to extend the result of Theorem 4 to general positive definite matrices $Q$ and $R$, the following mapping for the basis matrices $M_1, ..., M_l$ is required
\begin{equation}
M_i \triangleq R_d M_i Q_d^{-1}, \quad i \in \{1, 2, ..., l\}
\end{equation}

While this transformation is indeed useful to obtain a good lower bound, one can alternatively derive a lower bound on the optimality guarantee for arbitrary positive-definite matrices $Q$ and $R$ by multiplying the first and second terms in (30) by $\lambda_{\max}(Q)$ and $\lambda_{\max}(R)$, respectively.

### E. Sparsity Pattern

Consider a general discrete Lyapunov equation
\begin{equation}
M P M^T - P + H H^T = 0 \tag{32}
\end{equation}
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 (32) whenever
approximation of Optimization A below.

Approximate Optimization A. This problem is defined as

\[
\min_K \text{trace } \{(K_c - K) \hat{P}(K_c - K)^T\} \tag{33a}
\]

s.t. \(K \in \mathcal{K}\) 

(33b)

Let \(W \in \mathbb{R}^{n \times r}\) be a matrix whose columns are those eigenvectors of \(\hat{P}\) associated with the \(r\) nonzero eigenvalues of this matrix. Under the mild (generic) condition that every \(r\) rows of \(W\) are linearly independent, it is aimed to prove that if the number of free elements in each row of every matrix \(K \in \mathcal{K}\) is greater than or equal to \(r\), the optimal value of Approximate Optimization A is zero.

**Theorem 5.** The optimal objective value of Approximate Optimization A is 0 if the spark of \(W\) is \(r + 1\) and each row of every matrix \(K \in \mathcal{K}\) has at least \(r\) free elements.

**Proof.** The optimal objective value of Approximate Optimization A is 0 if there is a controller \(K_d\) such that \((K_c - K_d)W = 0\) or equivalently

\[
K_d^TW = K_c^TW, \quad j = 1, 2, \ldots, n \tag{34}
\]

where \(K_j^c\) and \(K_j^d\) denote the \(j\)th rows of \(K_c\) and \(K_d\), respectively. Note that the rows of \(K_d\) \(\in \mathcal{K}\) can be designed independently. On the other hand, (34) has a solution \(K_j^d\) with the right sparsity pattern because it has at least \(r\) free elements to be designed and the corresponding rows of \(W\) are linearly independent by assumption. This completes the proof.

**Corollary 1.** Given a natural number \(r\), assume that the rank of \(P\) is \(r\) and that each row of every matrix \(K \in \mathcal{K}\) has at least \(r\) free elements. Then, there exists a controller \(K_d \in \mathcal{K}\) whose global optimality degree is 100%.

**Proof.** The proof follows from Theorems 2 and 5.

It is desirable to show that the difference between the objective values of Optimization A and Approximate Optimization A can be upper bounded in terms of the maximum eigenvalue of \(P - \hat{P}\).

**Theorem 6.** Let \(K_d\) denote an optimal solution of Approximate Optimization A. The difference between the optimal objective values of Optimization A and Approximate Optimization A is upper bounded by the expression \(\sqrt{n} \times \lambda_{\max}(P - \hat{P}) \times \| \hat{P}(K_c - K_d)^T(K_c - K_d) \|_F \).

**Proof.** Since the matrix \(P - \hat{P}\) is positive semidefinite, the objective function of Approximate Optimization A is lower bounded on that of Optimization A. This implies that the difference between the objective functions of these two problems evaluated at the solution of Approximate Optimization A, i.e.,

\[
\text{trace } \{(K_c - K_d)(P - \hat{P})(K_c - K_d)^T\} \tag{35}
\]

which is equal to

\[
\text{trace } \{(K_c - K_d)^T(K_c - K_d)(P - \hat{P})\} \tag{36}
\]

Using the Cauchy-Schwarz inequality, it can be verified that

\[
\text{trace } \{(K_c - K_d)^T(K_c - K_d)(P - \hat{P})\} \leq \| (K_c - K_d)^T(K_c - K_d) \|_F \| P - \hat{P} \|_F \tag{37}
\]

The proof follows from a property of the Frobenius norm.

**Remark 5.** Due to Corollary 1, there is a controller \(K_d\) whose global optimality degree is close to 100% if the number of free elements in each row of every matrix \(K \in \mathcal{K}\) is greater than or equal to the approximate rank of \(P\) (i.e., the number of clearly dominant eigenvalues). If the degree of freedom of \(K_d\) in each row is higher than \(r\), then there are infinitely many distributed controllers with a high optimality degree, and then the chance of existence of a stabilizing controller among those candidates would be higher.

Motivated by Remark 5 and Corollary 1, we introduce an optimization problem next.

**Approximate Optimization D.** This problem is defined as

\[
\min_K \text{trace } \{(K_c - K)B^TB(K_c - K)^T\} \tag{38a}
\]

s.t. \(K \in \mathcal{K}\)

\[(K_c - K)W = 0 \tag{38c}\]

Approximate Optimization D aims to make the closed-loop system stable while imposing a constraint on the performance of the distributed controller. In particular, the designed optimization problem indirectly searches for a stabilizing distributed controller in the subspace of high-performance controllers with the prescribed sparsity pattern. As before, Approximate Optimization D has a closed-form solution. More precisely, analogous to Theorem 3, one can derive an explicit formula for \(K_d\) through a system of linear equations.

For each row \(i\) of \(K_d\), let \(r_i\) denote the number of free elements at row \(i\). Since the rank of \(\hat{P}\) is equal to \(r\), in order to assure that the system of equations (38) is not overdetermined, assume that \(r_i \geq r\) for every \(i \in \{1, 2, \ldots, m\}\). Furthermore, define \(l_i^r\) as a 0-1 row vector of size \(n\) such
that \( l_i^j(k) = 1 \) if the \( j \)th free element of row \( i \) in \( K_d \) resides in column \( k \) of \( K_d \) and \( l_i^j(k) = 0 \) otherwise, for every \( j \in \{1, 2, ..., r_i\} \). Define

\[
  l_i = [l_i^1T, l_i^2T, ..., l_i^r_iT]T
\]

The set of all permissible vectors for the \( j \)th row of \( K_d \) is characterized in terms of the left null space of \( l_iW \) and an initial vector as

\[
  K_d^i = K_0^i + \beta_i \text{null}(l_iW) \quad l_i
\]

where \( \beta_i \) is an arbitrary row vector with size equal to the number of rows in \( \text{null}(l_iW) \) and \( K_0^i \) is equal to

\[
  K_0^i = K_0^iW(\bar{I}_W)^{-1}l_i
\]

where

\[
  \bar{I}_W = [I_1T, I_2T, ..., I_{\tau_m}T]T
\]

Therefore, the set of permissible distributed controllers with the structure imposed by (38) can be characterized as

\[
  \tilde{K} = \{ K_0 + \beta N | \beta \in \mathcal{B} \}
\]

where

\[
  K_0 = [K_0^1T, K_0^2T, ..., K_0^mT]T
\]

and \( \mathcal{B} \) is the set of all matrices in the form of

\[
  \begin{bmatrix}
    \beta_1 & 0 & \cdots & 0 \\
    0 & \beta_2 & \cdots & 0 \\
    \vdots & \ddots & \ddots & \vdots \\
    0 & \cdots & \cdots & \beta_m
  \end{bmatrix}
\]

for arbitrary vectors \( \beta_i \) with size equal to the number of rows in \( \text{null}(l_iW) \). Similar to the basis matrices used in Theorem 3 denote \( M_1, ..., M_t \) as 0-1 matrices such that, for every \( t \in \{1, 2, ..., l\} \), \( M_t(i, j) \) is equal to 1 if \( (i, j) \) is the location of the \( \tau_m^j(k) \)th element of the vector \( [\beta_1, \beta_2, ..., \beta_m]T \) in \( \mathcal{B} \) and is 0 otherwise.

**Theorem 7.** Consider the matrix \( \tilde{X} \in \mathbb{R}^{l \times l} \) and vector \( \tilde{y} \in \mathbb{R}^l \) with the entries

\[
  \tilde{X}(i, j) = \text{trace}\left\{ N^T \tilde{M}_i B^T B \dot{M}_j N \right\}
\]

\[
  \tilde{y}(i) = \text{trace}\left\{ N^T \tilde{M}_i B^T B(K_c - K_0) \right\}
\]

for every \( i, j \in \{1, 2, ..., l\} \). The matrix \( K_d \) is an optimal solution of Approximate Optimization D if and only if it can be expressed as \( K_d = K_0 + \beta N \), where \( \beta \) is defined as the matrix in (45) and the parameters \( \beta_1, \beta_2, ..., \beta_m \) satisfy the linear equation \( \tilde{X}[\beta_1, \beta_2, ..., \beta_m]T = \tilde{y} \).

**Proof.** The method used in the proof of Theorem 3 can be adopted to prove this theorem after noting that the set of permissible distributed controllers with the structure imposed by (38) is equal to the set \( \tilde{K} \) in (43).

**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

\[
  \Sigma_d = \mathcal{E}\{Ed[\tau]d(\tau)^T E^T\}, \quad \Sigma_v = \mathcal{E}\{Fv[\tau]v(\tau)^T F^T\}
\]

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 \( \tau \). Let \( K_c \) denote the gain of the optimal static distributed 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 \( u[\tau] = K_c y[\tau] \) with a high global optimality guarantee such that \( K_d \in \tilde{K} \). For an arbitrary discrete-time random process \( y[\tau] \) 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). Hence, it is not necessary for the centralized and distributed control systems to have similar trajectories for states and inputs at all times as long as they have similar limiting behaviors.

Before proceeding with the main results for stochastic systems, it is worthwhile to mention that whenever the states and inputs of the distributed control system are compared to their counterparts for the centralized one on different times, it is implicitly assumed that the measurement noise and input disturbance are equal for both systems at each time step. The reason behind this assumption is that the system under study undergoes the same input disturbance and measurement noise in both distributed and centralized cases. Therefore, at a given time, the measured states contain the same input disturbance and measurement noise for both centralized and distributed control systems.

**Lemma 5.** Given the optimal centralized gain \( K_c \) and an arbitrary stabilizing distributed controller gain \( K_d \in \tilde{K} \), the relations

\[
  \mathcal{E}\{u_c[\infty]\} = \mathcal{E}\{u_d[\infty]\} = \mathcal{E}\{x_c[\infty]\} = \mathcal{E}\{x_d[\infty]\} = 0
\]

hold.

**Proof.** The proof is a direct consequence of the stability of the closed-loop systems and the Gaussian distributions of the measurement noise and disturbance. The details are omitted for brevity. \( \square \)

Lemma 5 states that any stabilizing controller, regardless of its structure, generates states and inputs whose expected values attenuate to zero as \( \tau \to \infty \). However, in order to guarantee that the objective value (5) for the distributed control system is close to that of the centralized control system, the second moments of \( x_c[\infty] \) and \( x_d[\infty] \) should be similar. To this end, we propose an optimization problem to indirectly minimize...
Lemma 6. Given the optimal centralized gain $K_c$ and an arbitrary stabilizing distributed controller gain $K_d \in \mathcal{K}$, the relation
\[
\mathcal{E} \left\{ \| x_c[\infty] - x_d[\infty] \|_2^2 \right\} = \text{trace} \{ P_1 + P_2 - P_3 - P_4 \} \tag{49}
\]
holds, where $P_1, P_2, P_3$ and $P_4$ are the unique solutions of the equations
\[
(A+BK_c)P_1(A+BK_c)^T - P_1 + \Sigma_d + (BK_d) \Sigma_v (BK_d)^T = 0 \tag{50a}
\]
\[
(A+BK_c)P_2(A+BK_c)^T - P_2 + \Sigma_d + (BK_c) \Sigma_v (BK_c)^T = 0 \tag{50b}
\]
\[
(A+BK_d)P_3(A+BK_c)^T - P_3 + \Sigma_d + (BK_c) \Sigma_v (BK_c)^T = 0 \tag{50c}
\]
\[
(A+BK_c)P_4(A+BK_d)^T - P_4 + \Sigma_d + (BK_c) \Sigma_v (BK_c)^T = 0 \tag{50d}
\]

Proof. The proof is provided in the appendix. \hfill \Box

Note that (50a) and (50c) are Lyapunov equations, whereas (50b) and (50d) are Stein equations. These equations all have unique solutions if $A+BK_d$ and $A+BK_c$ are stable. Lemma 6 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 (50a) and $K_d \in \mathcal{K}$. However, this is a hard problem in general. In particular, the minimization of the singleton $\text{trace} \{ P_1 \}$ subject to (50a) and $K_d \in \mathcal{K}$ 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).

In what follows, we will propose an optimization problem as the counterpart of Optimization C for stochastic systems.

**Stochastic Optimization C**: Given a constant number $\alpha \in [0,1]$, this problem is defined as the minimization of the function
\[
C_\alpha(K) = \alpha \times C_\alpha^*(K) + (1 - \alpha) \times C_\alpha^*(K) \tag{51}
\]
with respect to the matrix variable $K \in \mathcal{K}$, where $P_s$ is the unique solution to
\[
(A+BK_c)P_s(A+BK_c)^T - P_s + \Sigma_d + (BK_c) \Sigma_v (BK_c)^T = 0 \tag{52}
\]
and
\[
C_1^*(K) = \text{trace} \left\{ (K_c - K) (\Sigma_v + P_s) (K_c - K)^T \right\} \tag{53a}
\]
\[
C_2^*(K) = \text{trace} \left\{ (K_c - K)^T B^T B (K_c - K) \right\} \tag{53b}
\]

It is straightforward to observe that Theorem 3 can be adopted to find an explicit formula for all solutions of Stochastic Optimization C by replacing $P$ with $\Sigma_v + P_s$ in the definitions of $X$ and $y$.

**Lemma 7.** $C_1^*(K_d)$ in Stochastic Optimization C is equal to
\[
\mathcal{E} \left\{ \left( (K_c - K_d) (Fv[\infty] + x_c[\infty]) \right)^2 \right\} \tag{54}
\]

Proof. One can verify that
\[
\| (K_c - K_d) (Fv[\infty] + x_c[\infty]) \|^2_2 = \text{trace} \left\{ \left( (K_c - K_d) (Fv[\infty] + x_c[\infty]) \right)^T (K_c - K_d)^T \right\} \tag{55}
\]

Moreover,
\[
x[\infty] = \lim_{\tau \to \infty} \sum_{i=0}^{\tau-1} (A+BK)^{\tau-1-i} (Ed[i] + BK Fv[i]) \tag{56}
\]

The proof is completed by substituting (56) into (55) and taking its expected value. \hfill \Box

The next lemma proves that the covariance matrix of the random vector $\Delta x[\tau]$ converges to a finite and constant matrix under a stability condition.

**Lemma 8.** Given the optimal centralized gain $K_c$ and an arbitrary stabilizing distributed controller gain $K_d \in \mathcal{K}$, the term $\mathcal{E} \left\{ \Delta x[\tau] \Delta x[\tau]^T \right\}$ converges to a finite constant matrix as $\tau$ goes to infinity.

Proof. The proof is immediate using Lemma 6. \hfill \Box

Lemmas 7 and 8 can be combined to prove a main result stated below.

**Lemma 9.** Given the optimal centralized gain $K_c$ and an arbitrary stabilizing gain $K_d \in \mathcal{K}$, the relations
\[
\mathcal{E} \left\{ \| x_c[\infty] - x_d[\infty] \|_2^2 \right\} \leq \left( \frac{\kappa(V) \| B \|^2_2}{1 - \rho(A+BK_d)} \right)^2 C_1^*(K_d) \tag{57a}
\]
\[
\mathcal{E} \left\{ \| u_c[\infty] - u_d[\infty] \|_2^2 \right\} \leq \left( 1 + \frac{\kappa(V) \| K_d \|^2_2 \| B \|^2_2}{1 - \rho(A+BK_d)} \right)^2 C_1^*(K_d) \tag{57b}
\]
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. \hfill \Box

In what follows, the counterpart of Theorem 4 will be presented for stochastic systems.

**Theorem 8.** Assume that $Q = I_n$ and $R = I_m$. Given the optimal centralized gain $K_c$ and an arbitrary stabilizing gain $K_d \in \mathcal{K}$, the relations
\[
\left( 1 + \mu_s \sqrt{C_1^*(K_d)} \right)^2 J(K_c) \geq J(K_d) \geq J(K_c) \tag{58}
\]
hold, where
\[
\mu_s = \max \left\{ \frac{\kappa(V) \| B \|^2_2}{(1 - \rho(A+BK_d)) \sqrt{\mathcal{E} \left\{ \| x_c[\infty] \|_2^2 \right\}}} , \frac{1 - \rho(A+BK_d) + \kappa(V) \| K_d \|^2_2 \| B \|^2_2}{(1 - \rho(A+BK_d)) \sqrt{\mathcal{E} \left\{ \| u_c[\infty] \|_2^2 \right\}}} \right\} \tag{59}
\]

Proof. The proof is a consequence of Lemma 9 and the argument made in the proof of Theorem 4. \hfill \Box
It can be inferred from Theorem [8] that Stochastic Optimization C aims to indirectly maximize the optimality guarantee and assure stability. Similar to the deterministic case, we can extend the results of Theorem [8] to stochastic systems with general positive-definite matrices $Q$ and $R$, using Remark [4] after two additional changes of parameters

$$\hat{E} = Q_d E, \quad \hat{F} = Q_d F$$

**Remark 6.** It is well-known that finding the optimal centralized controller in the presence of measurement noise would be a difficult problem in general. If the optimal controller $K_c$ is not available, one can use a near-globally optimal controller that performs similarly to the near-globally optimal centralized controller. 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 [31]).

V. NUMERICAL RESULTS

Three examples will be offered in this section to demonstrate the efficacy of the proposed controller design technique.

A. Example 1: Power Networks

In this part, we study the distributed frequency control problem for electrical power systems. The goal is to design a distributed controller that controls the frequency of a system consisting of a number of generators and loads that are connected together via an underlying transmission network. As a case study, we consider the IEEE 39-Bus New England Power System. The single-line diagram of this system is provided in Figure 1. The distributed controller constrained by a user-defined structure is used to optimally adjust the mechanical power input to each generator. This pre-determined communication topology specifies which generators exchange their rotor angle and frequency measurements with one another. To formulate the problem, we consider the widely used per-unit swing equation

$$M_i \dot{\theta}_i + D_i \dot{\theta}_i = P_M - P_{EI}$$

where $\theta_i$ denotes the voltage (or rotor) angle at a generator bus $i$ (in rad), $P_M$ is the mechanical power input to the generator at bus $i$ (in per unit), $P_{EI}$ is the electrical active power injection at bus $i$ (in per unit), $M_i$ is the inertia of the generator at bus $i$ (in pu-sec²/rad), and $D_i$ is the damping coefficient of the generator at bus $i$ (in pu-sec/rad) [39]. The electrical real power $P_{EI}$ in (60) can be found using the nonlinear AC power flow equation

$$P_{EI} = \sum_{j=1}^{n} V_i^* V_j [ G_{ij} \cos(\theta_i - \theta_j) + B_{ij} \sin(\theta_i - \theta_j) ]$$

where $n$ denotes the number of buses in the system, $V_i$ is the voltage phasor at bus $i$, $G_{ij}$ is the line conductance, and $B_{ij}$ is the line susceptance. To simplify the formulation, a commonly-used technique is to use the DC power flow equation corresponding to (61) in which all the voltage magnitudes are assumed to be 1 per unit, each branch is modeled as a series inductor, and the angle differences across each line are assumed to be relatively small:

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

It is possible to rewrite (62) into the matrix format $P_E = L \theta$, where $P_E$ and $\theta$ are the vectors of real power injections and voltage (or rotor) angles at all the generator buses. In this equation, $L$ denotes the Laplacian matrix and can be found as follows [40]:

$$L_{ii} = \sum_{j=1, j \neq i}^{n} B_{ij}^{Kron} \quad \text{if } i = j$$

$$L_{ij} = -B_{ij}^{Kron} \quad \text{if } i \neq j$$

where $B_{ij}^{Kron}$ is the susceptance of the Kron reduced admittance matrix $Y_{Kron}$ defined as

$$Y_{Kron}^{ij} = Y_{ij} - \frac{Y_{ik} Y_{kj}}{Y_{kk}} \quad (i, j = 1, 2, \ldots, n \text{ and } i, j \neq k)$$

where $k$ is the index of the non-generator bus to be eliminated from the admittance matrix and $n$ is the number of generator buses. Note that the Kron reduction method aims to eliminate the static buses of the network because the dynamics and interactions of only the generator buses are of interest [41].

Using $\theta = [\theta_1, \ldots, \theta_n]^T$ as rotor angle state vector and $w = [w_1, \ldots, w_n]^T$ as the frequency state vector and substituting the matrix format of $P_E$ into (60), the state space model of the swing equation used for frequency control in power systems could be written as

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

Fig. 1: Single-line diagram of IEEE 39-Bus New England Power System.
Localized Deterministic Case: In this experiment, we generate the entries of the initial state $x$ what generators are allowed to communicate. We will study where each node represents a generator and each line specifies structure of the controller: distributed, localized, star and ring. $R$ and $G$ optimality. The 39-bus system has 10 generators, labeled as $G_0, G_1, ..., G_{10}$. 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 2, where each node represents a generator and each line specifies what generators are allowed to communicate. We will study both deterministic and stochastic cases below, by choosing the entries of the initial state $x(0)$ uniformly from the interval $[0, 1]$. 

Deterministic Case: In this experiment, we generate the weighting matrices $Q$ and $R$ in a random fashion as $Q = Q Q^T$ and $R = 0.1 \times R R^T$, where $Q$ and $R$ are normal random matrices. We then design distributed controllers using the explicit formula given in Theorem 3 and with the reformulation introduced in Remark 4. Figure 3a shows the global optimality guarantee for the four aforementioned control topologies for different values of the parameter $\alpha$. The maximum absolute eigenvalue of the closed-loop matrix $A + BK_d$ is depicted in Figure 3b. A number of observations can be made:

- For those values of $\alpha$ that make the system stable, the optimality guarantee is at least 80% for all topologies. Note that a distributed controller with a significantly higher optimality guarantee may not exist due to its low degree of freedom.
- For those values of $\alpha$ making all topologies stabilizing, the ring structure has the best optimality guarantee (near 100%), while the fully distributed one has the lowest guarantee (near 80%).
- According to Figure 3b, the choice of $\alpha$ is critical for the ring and star structures. The choices $0.92 \leq \alpha \leq 1$ and $0.77 \leq \alpha \leq 1$ make the system unstable for the ring and star topologies, respectively. Moreover, all values of $\alpha$ in the interval $[0.02, 1]$ lead to high-performance controllers for the fully distributed and localized topologies.

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 $\Sigma_d = I_n$ and $\Sigma_v = \sigma I_n$, with $\sigma$ varying from 0 to 5. The matrices $Q$ and $R$ are selected as $I_n$ and $0.1 \times I_m$, respectively, and $\alpha$ is chosen as 0.4. First, a near-globally optimal static centralized controller $K_c$ is designed using the ODC solver that is based on an SDP relaxation of the problem (see [31]). This static centralized controller is then used in Stochastic Optimization C to obtain the distributed controller $K_d$, as explained in Remark 6. The simulation results are provided in Figure 4. 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 decentralized controller is relatively low. This may be due to the potential large gap between the globally optimal costs of the optimal centralized and fully decentralized control problems. Furthermore, notice that the designed distributed controllers are all found using the explicit formulas provided in this paper, as opposed to solving optimization problems.

B. Example 2: Stable Random Systems

The objective is to demonstrate the performance of the proposed design technique on stable systems. Consider $n = m = 40$, $Q = I_n$ and $R = I_m$. We generate random continuous-time systems and then discretize them according to following rules:

- The entries of $A_c$ are chosen randomly from a Gaussian distribution with the mean 0 and variance 25.
- The entries of $B_c$ are chosen randomly from a normal distribution.
- After constructing $A_c$, the matrix is rescaled by a real number so that its maximum absolute eigenvalue becomes equal to 0.8.
- $K$ is assumed to be diagonal (off-diagonal elements are forced to be zero).
- $A$ and $B$ are obtained by discretizing ($A_c, B_c$) using the zero-order hold method with the sampling time of 0.1 second.

Notice that $K = 0$ is a trivial stabilizing distributed controller for the above system. However, this choice of the controller may not result in a high optimality guarantee. We design a diagonal controller $K_d$ using the explicit formula given in [26]
for 100 random systems generated as above, with $\alpha = 0.98$. We arrange the resulting optimality guarantees in ascending order and label their corresponding trials as 1, 2, ..., 100. Furthermore, we calculate the optimality guarantee for the trivial controller $K = 0$ for each random system. The results of this experiment are provided in Figure 5. It can be observed that the designed diagonal controller always has a better optimality guarantee than the trivial solution $K = 0$. Note that our formula designs the $(i, i)^{th}$ entry of $K_d$ using a linear combination of the entries in the $i^{th}$ row and $i^{th}$ column of the optimal centralized controller $K_c$. Another approach is to simply use the diagonal of $K_c$ as a candidate diagonal controller $K_d$. To assess this difference, we compute the cross-correlation between the controller designed using Theorem 5 and a truncated version of $K_c$ for these 100 trials. The cross-correlation for these trials has the mean 0.4892 and standard deviation 0.1217. This shows that the designed controller is completely different from the naive method of truncating the centralized controller.

C. Example 3: Highly-Unstable Random Systems

In Example 2, it was shown that the proposed explicit formula could simply design high-performance distributed controllers for stable systems. This example investigates highly unstable systems for which the design of a stabilizing distributed controller is challenging itself without even imposing any performance criterion. Consider Example 2, but assume that the randomly generated matrices $A_c$'s are not scaled down to make the maximum absolute eigenvalue equal to 0.8 (i.e., their unstable eigenvalues are untouched). Suppose that each off-diagonal entry of $K$ is forced to be zero with probability $p$. We consider three scenarios associated with $p$ equal to 0.1, 0.2 and 0.3. The diagonal entries of $K$ are all set to be free elements by assuming that each subsystem can measure its own state.

Note that the above class of random systems is highly unstable with the maximum absolute eigenvalue of $A$ as high as 9. The optimality guarantees and maximum absolute eigenvalues of the designed distributed controllers $K_d$ are given in Figure 6 for a sample random system with a varying parameter $\alpha$. The following observations can be made:

- As long as $\alpha$ is not too close to 1, the designed controllers are stabilizing. This implies that it is crucial to incorporate both stability and performance terms in the objective of Optimization C.
- As the probability of forced zeros in the controller increases, the optimality guarantee of the designed controller becomes more dependent on the value of $\alpha$. Moreover, the optimality guarantee for the controller with $p = 0.1$ is higher than those of the other controllers, and it is almost 100% for a wide range of $\alpha$. This is due to the fact that the controller has many free elements for
controllers for the three scenarios of \( p \) equal to 0.1, 0.2 and 0.3. We arrange the obtained maximum absolute eigenvalues in ascending order and subsequently label their corresponding trials as 1, 2, ..., 100. Figure 7 shows the optimality guarantees and maximum absolute eigenvalues of the designed distributed controllers. For \( p = 0.1 \), the proposed method always yields stabilizing controllers with optimality guarantees near to 100%. For \( p = 0.2 \), 99 control systems are stable with optimality guarantees near to 100%. For \( p = 0.3 \), 54 control systems are stable with high optimality guarantees. Note that the designed controllers are different from truncated versions of \( K_c \) (by simply discarding 10%-30% entries of \( K_c \)). More precisely, the cross-correlation between the controller designed using Theorem 3 and a truncated version of \( K_c \) for the above 100 trials with \( p = 0.1 \) has the mean 0.6245 and standard deviation 0.0677.

As mentioned earlier, the eigenvalues of \( P \) in the Lyapunov equation \((14)\) may decay rapidly. To support this statement, we compute the eigenvalues of \( P \) for previously generated 100 random unstable systems. Subsequently, we arrange the absolute eigenvalues of \( P \) for each system in ascending order and label them as \( \lambda_1, \lambda_2, ..., \lambda_{40} \). For every \( i \in \{1, 2, ..., 40\} \), the mean of \( \lambda_i \) for these 100 independent random systems is drawn in Figure 8a (the variance is very low). It can be observed that only 15% of the eigenvalues are dominant and \( P \) can be well approximated by a low-rank matrix. Due to Corollary 1 in order to ensure a high optimality guarantee, each row of \( K_d \) should have at least 6 free elements. However, as discussed earlier, more free elements may be required to make the system stable via Approximate Optimization D. To study the minimum number of free elements needed to achieve a closed-loop stability and a high optimality guarantee, we find distributed controllers using Approximate Optimization D for different numbers of free elements at each row of \( K_d \). In all of these random systems, \( P \) is approximated by a rank-6 matrix \( \hat{P} \). For each sparsity level, we then calculate the percentage of closed-loop systems that become stable using Approximate Optimization D. Moreover, their average global optimality guarantee using the designed \( K_d \) is also obtained. The results are provided in Figures 8b and 8c. The number of stable closed-loop systems increases quickly as the number of free elements in each row of the distributed controller gain exceeds 25. As mentioned earlier, there could be a non-trivial gap between the minimum number of free elements satisfying the performance criterion and the minimum number of free elements required to make the closed-loop system stable. Furthermore, it can be observed in Figure 8c that the designed distributed controller has an optimality guarantee close to 100% for all stable closed-loop systems. This optimality guarantee is ensured via constraint (3c) in Approximate Optimization D.

VI. Conclusions

This paper studies the optimal 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 best (static) centralized controller. To this end, we derive a necessary and sufficient condition under which there exists a distributed controller that produces the same input and state trajectories as the optimal centralized controller for a given deterministic system. 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. 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. The derived explicit formulas for the controller design may help partially answer some open problems, such as finding the minimum number of free elements needed in the distributed controller to attain a performance close to the optimal centralized one. The proposed method is evaluated on a power network and several random systems.

**REFERENCES**


Combining (67) and (68) leads to (10).

To prove that (10) implies (9), suppose that the equation (10) is satisfied. By pre-multiply the left side of (10) with $B$, it follows from Lemma 1 that $x_c[\tau] = x_d[\tau]$. Therefore,

$$u_c[\tau] - u_d[\tau] = K_c x_c[\tau] - K_d x_d[\tau] = K_c x_c[\tau] - K_d x_c[\tau] = (K_c - K_d) (A + BK_d)^\tau x[0] = 0$$

This yields the equation (9), and completes the proof. □

Proof of Lemma 4
First, we prove the inequality (7/3). It is straightforward to verify that

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

(70)

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 (70) by $V$ yields that

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

(71)

Taking the 2-norm from both sides of (71) leads to

$$\| \Delta \tilde{x}[\tau + 1] \|_2 \leq \| V B (K_d - K_c) x_c[\tau] \|_2 + \| D \|_2 \| \Delta \tilde{x}[\tau] \|_2$$

(72)

(note that $\| D \|_2 \leq \| D \|_2 \leq \rho(A + BK_d)$). It can be concluded from (72) that

$$\| \Delta \tilde{x}[\tau + 1] \|_2 - \rho(A + BK_d) \| \Delta \tilde{x}[\tau] \|_2$$

or equivalently

$$\| \Delta \tilde{x}[\tau + 1] \|_2 + \rho(A + BK_d)^2 \| \Delta \tilde{x}[\tau] \|_2$$

$$- 2\rho(A + BK_d) \| \Delta \tilde{x}[\tau + 1] \|_2 \| \Delta \tilde{x}[\tau] \|_2$$

$$\leq \| V B (K_d - K_c) x_c[\tau] \|_2$$

(73)

Summing up both sides of (74) over all values of $\tau$ gives rise to the inequality

$$\sum_{\tau = 0}^{\infty} \| \Delta \tilde{x}[\tau + 1] \|_2^2 + \rho(A + BK_d)^2 \sum_{\tau = 0}^{\infty} \| \Delta \tilde{x}[\tau] \|_2^2$$

$$- 2\rho(A + BK_d) \sum_{\tau = 0}^{\infty} \| \Delta \tilde{x}[\tau + 1] \|_2 \| \Delta \tilde{x}[\tau] \|_2$$

$$\leq \sum_{\tau = 0}^{\infty} \| V B (K_d - K_c) x_c[\tau] \|_2^2$$

(75)

Using the Cauchy-Schwarz inequality, one can write

$$\sum_{\tau = 0}^{\infty} \| \Delta \tilde{x}[\tau + 1] \|_2^2 + \rho(A + BK_d)^2 \sum_{\tau = 0}^{\infty} \| \Delta \tilde{x}[\tau] \|_2^2$$

$$- 2\rho(A + BK_d) \sqrt{\left( \sum_{\tau = 0}^{\infty} \| \Delta \tilde{x}[\tau + 1] \|_2^2 \right) \left( \sum_{\tau = 0}^{\infty} \| \Delta \tilde{x}[\tau] \|_2^2 \right)}$$

$$\leq \sum_{\tau = 0}^{\infty} \| V B (K_d - K_c) x_c[\tau] \|_2^2$$

(76)
Note that $\sum_{\tau=0}^{\infty} |\Delta \hat{x}[\tau + 1]|^2 = \sum_{\tau=0}^{\infty} |\Delta \hat{x}[\tau]|^2$. Hence, it can be inferred from (76) and (28) that
\[
\sum_{\tau=0}^{\infty} |\Delta \hat{x}[\tau]|^2 \leq \left( \frac{|V_2|}{2} \right)^2 \frac{1}{1 - \rho(A + BK_d)} C_1(K_d) \tag{77}
\]
Since $\Delta x[\tau] = V^{-1} \Delta \hat{x}[\tau]$, one can write
\[
\sum_{\tau=0}^{\infty} |\Delta x[\tau]|^2 \leq \left( \frac{|V_2|}{2} \right)^2 \frac{1}{1 - \rho(A + BK_d)} C_1(K_d) \tag{78}
\]
This proves the inequality (27a). The above argument can be adopted to prove (27b) after noting that
\[
\Delta u[\tau] = (K_d - K_c) x[\tau] + K_d \Delta x[\tau] \tag{79}
\]

**Proof of Theorem 4** According to Lemma 3 one can write
\[
\sum_{\tau=0}^{\infty} |x_d[\tau]|^2 + \sum_{\tau=0}^{\infty} |x_c[\tau]|^2 \leq \left( \frac{\kappa(V)}{|B|} \right)^2 \frac{1}{1 - \rho(A + BK_d)} C_1(K_d) \tag{80}
\]
Dividing both sides of (80) by $\sum_{\tau=0}^{\infty} |x[\tau]|^2$ and using the Cauchy-Schwarz inequality for $\sum_{\tau=0}^{\infty} |x[\tau]|^2 |x_d[\tau]|^2$ yield that
\[
\frac{\sum_{\tau=0}^{\infty} |x_d[\tau]|^2}{\sum_{\tau=0}^{\infty} |x[\tau]|^2} \leq \left( 1 + \frac{\kappa(V)}{|B|} \right)^2 \frac{1}{1 - \rho(A + BK_d)} C_1(K_d) \tag{81}
\]
Likewise,
\[
\frac{\sum_{\tau=0}^{\infty} |u_d[\tau]|^2}{\sum_{\tau=0}^{\infty} |u[\tau]|^2} \leq \left( 1 + \frac{\kappa(V)}{|B|} \right)^2 \frac{1}{1 - \rho(A + BK_d)} C_1(K_d) \tag{82}
\]
Combining (81) and (82) leads to
\[
\frac{J(K_d)}{J(K_c)} = \sum_{\tau=0}^{\infty} |x_d[\tau]|^2 + \sum_{\tau=0}^{\infty} |u_d[\tau]|^2 \leq (1 + \mu) \frac{C_1(K_d)}{C_1(K_d)} \tag{83}
\]
This completes the proof. □

**Proof of Lemma 9** It is straightforward to verify that
\[
\mathcal{E} \left\{ |x_c[\infty] - x_d[\infty]|^2 \right\} = \text{trace} \left\{ \mathcal{E} \left\{ x_d[\infty] x_d[\infty]^T \right\} \right\}
+ \text{trace} \left\{ \mathcal{E} \left\{ x_c[\infty] x_c[\infty]^T \right\} \right\}
- \text{trace} \left\{ \mathcal{E} \left\{ x_d[\infty] x_c[\infty]^T \right\} \right\}
- \text{trace} \left\{ \mathcal{E} \left\{ x_c[\infty] x_d[\infty]^T \right\} \right\} \tag{84}
\]
On the other hand, since $d[\cdot]$ and $v[\cdot]$ are independent and identically distributed random vectors, the equation
\[
\mathcal{E} \left\{ d_1[\tau_1] d_2[\tau_2] \right\} = \mathcal{E} \left\{ v_1[\tau_1] v_2[\tau_2] \right\} = 0 \text{ holds for every two different indices } \tau_1 \text{ and } \tau_2. \text{ In addition, we have } \mathcal{E} \left\{ v_1[\tau_1] d_2[\tau_2] \right\} = \mathcal{E} \left\{ d_1[\tau_1] v_2[\tau_2] \right\} = 0, \text{ for all nonnative integers } \tau_1 \text{ and } \tau_2. \text{ It follows from these relations and the equations (84) and (59) that}
\]
\[
\mathcal{E} \left\{ x_d[\tau] x_d[\tau]^T \right\} = P_1, \quad \mathcal{E} \left\{ x_c[\tau] x_c[\tau]^T \right\} = P_2, \quad \mathcal{E} \left\{ x_c[\tau] x_d[\tau]^T \right\} = P_3, \quad \mathcal{E} \left\{ x_c[\tau] x_d[\tau]^T \right\} = P_4, \tag{85}
\]
where $P_1, P_2, P_3$ and $P_4$ satisfy (59). Moreover, note that (50a) and (50b) are Lyapunov equations, whereas (50c) and (50d) are Stein equations. These equations all have unique solutions since $A + BK_d$ and $A + BK_c$ are stable. This completes the proof. □

**Proof of Lemma 9** The proof is similar to that of Lemma 4. First, we prove the inequality (57a). It is straightforward to verify that
\[
\Delta x[\tau + 1] = B(K_d - K_c)(Fv[\tau] + x_c[\tau]) + (A + BK_d) \Delta x[\tau] \tag{86}
\]
Using the eigen-decomposition of $A + BK_d$ as $V^{-1}DV$ and pursuing the method developed in the proof of Lemma 4 one can write
\[
\mathcal{E} \left\{ (\Delta \hat{x}[\tau + 1]|_2 - \rho(A + BK_d) \Delta x[\tau]|_2 \right\} \leq \mathcal{E} \left\{ (V(B(K_d - K_c)(Fv[\tau] + x_c[\tau])|_2 \right\} \tag{87}
\]
where $\Delta \hat{x}[\tau] = V \Delta x[\tau]$. It follows from the Cauchy-Schwarz inequality and (59) that
\[
\mathcal{E} \left\{ (\Delta \hat{x}[\tau + 1]|_2 + \rho(A + BK_d)^2 \mathcal{E} \left\{ (\Delta \hat{x}[\tau]|_2 \right\} \right\} - 2 \rho(A + BK_d) \mathcal{E} \left\{ (\Delta \hat{x}[\tau + 1]|_2 \right\} \mathcal{E} \left\{ (\Delta \hat{x}[\tau]|_2 \right\} \leq \mathcal{E} \left\{ (V(B(K_d - K_c)(Fv[\tau] + x_c[\tau])|_2 \right\} \tag{88}
\]
According to Lemma 8, $\mathcal{E} \left\{ \Delta x[\tau] \Delta x[\tau]^T \right\}$ (and hence, $\mathcal{E} \left\{ \Delta \hat{x}[\tau] \Delta \hat{x}[\tau]^T \right\}$) converges to a constant matrix as $\tau$ approaches infinity. This implies that $\mathcal{E} \left\{ (\Delta \hat{x}[\tau + 1]|_2 \right\}$ and $\mathcal{E} \left\{ (\Delta \hat{x}[\tau]|_2 \right\}$ converge to the same number. Taking the limits from both sides of (88) and interchanging the limit operator with the expected value operator result in
\[
\mathcal{E} \left\{ |\Delta \hat{x}[\infty]|^2 \right\} \leq \left( \frac{|V_2|}{2} \right)^2 \frac{1}{1 - \rho(A + BK_d)} C_1(K_d) \tag{89}
\]
Substituting $\Delta x[\infty] = V^{-1} \Delta \hat{x}[\infty]$ into the above inequality leads to
\[
\mathcal{E} \left\{ |\Delta x[\infty]|^2 \right\} \leq |V^{-1}|^2 \mathcal{E} \left\{ |\Delta \hat{x}[\infty]|^2 \right\} \leq \left( \frac{|V^{-1}|^2 |V_2|^2}{2} \right)^2 \frac{1}{1 - \rho(A + BK_d)} C_1(K_d) \tag{90}
\]
The inequality (57b) can be proved similarly. □