

# Efficient Convex Relaxation for Stochastic Optimal Distributed Control Problem

Abdulrahman Kalbat, Ramtin Madani, Ghazal Fazelnia, and Javad Lavaei  
Department of Electrical Engineering, Columbia University

**Abstract**—This paper is concerned with the design of an efficient convex relaxation for the notorious problem of stochastic optimal distributed control (SODC) problem. The objective is to find an optimal structured controller for a dynamical system subject to input disturbance and measurement noise. With no loss of generality, this paper focuses on the design of a static controller for a discrete-time system. First, it is shown that there is a semidefinite programming (SDP) relaxation for this problem with the property that its SDP matrix solution is guaranteed to have rank at most 3. This result is due to the extreme sparsity of the SODC problem. Since this SDP relaxation is computationally expensive, an efficient two-stage algorithm is proposed. A computationally-cheap SDP relaxation is solved in the first stage. The solution is then fed into a second SDP problem to recover a near-global controller with an enforced sparsity pattern. The proposed technique is always exact for the classical  $H_2$  optimal control problem (i.e., in the centralized case). The efficacy of our technique is demonstrated on the IEEE 39-bus New England power network, a mass-spring system, and highly-unstable random systems, for which near-optimal stabilizing controllers with optimality degrees above 90% are designed under a wide range of noise levels.

## I. INTRODUCTION

The area of decentralized control is created to address the challenges arising in the control of real-world systems with many interconnected subsystems. The objective is to design a structurally constrained controller—a set of partially interacting local controllers—with the aim of reducing the computation or communication complexity of the overall controller. The local controllers of a decentralized controller may not be allowed to exchange information. The term *distributed control* is often used in lieu of decentralized control in the case where there is some information exchange between the local controllers. It has been long known that the design of an optimal decentralized (distributed) controller is a daunting task because it amounts to an NP-hard optimization problem [1], [2]. Great effort has been devoted to investigating this highly complex problem for special types of systems, including spatially distributed systems [3], [4], [5], [6], [7], dynamically decoupled systems [8], [9], weakly coupled systems [10], and strongly connected systems [11].

Due to the recent advances in the area of convex optimization, the focus of the existing research efforts has shifted from deriving a closed-form solution for the above control synthesis problem to finding a convex formulation of the problem that can be efficiently solved numerically [12], [13], [14], [15], [16]. This has been carried out in the seminal work [17] by deriving a sufficient condition named quadratic invariance,

which has been generalized in [18] by deploying the concept of partially order sets. These conditions have been further investigated in several other papers [19], [20], [21]. A different approach is taken in the recent papers [22] and [23], where it has been shown that the distributed control problem can be cast as a convex optimization for positive systems. A sparsity-promoting technique is developed in [24].

There is no surprise that the decentralized control problem is computationally hard to solve. This is a consequence of the fact that several classes of optimization problems, including polynomial optimization and quadratically-constrained quadratic program as a special case, are NP-hard in the worst case. Due to the complexity of such problems, various convex relaxation methods based on linear matrix inequality (LMI), semidefinite programming (SDP), and second-order cone programming (SOCP) have gained popularity [25], [26]. These techniques enlarge the possibly non-convex feasible set into a convex set characterizable via convex functions, and then provide the exact or a lower bound on the optimal objective value. The SDP relaxation usually converts an optimization with a vector variable to a convex optimization with a matrix variable, via a lifting technique. The exactness of the relaxation can then be interpreted as the existence of a low-rank (e.g., rank-1) solution for the SDP relaxation. Several papers have studied the existence of a low-rank solution to matrix optimizations with linear or nonlinear (e.g., LMI) constraints.

Consider the design of a structured controller for a deterministic system. It is shown in [27] that this problem admits an SDP relaxation in the finite-horizon case with the property that the rank of its SDP is upper bounded by 4. This result is extended to an infinite-horizon control problem in [28]. The current paper significantly generalizes our previous results in multiple directions. The main focus of this work is on the stochastic optimal distributed control (SODC) problem, where the system is subject to input disturbance and measurement noise. First, we show that the SODC problem has an SDP relaxation such that the rank of its SDP solution is at most 3. This solution can be converted to a near-global controller, but the SDP problem is computationally expensive. The main contribution of this work is to develop an efficient computational method to find a high-performance controller. To this end, we offer a two-stage convex program, where a computationally-cheap SDP relaxation is solved in the first stage and its solution is then passed to a second convex optimization to retrieve a solution. The superiority of this technique is discussed in detail. In the extreme case, this technique correctly solves the classical (centralized)  $H_2$  optimal control problem, as a substitute for the Riccati equations. The proposed algorithm is applied to the frequency control of a New England power network, a mass-spring system, and several highly-unstable

Emails: ak3369@columbia.edu, madani@ee.columbia.edu, gf2293@columbia.edu, and lavaei@ee.columbia.edu.

This work was supported by NSF CAREER and ONR YIP.

random systems.

This paper is organized as follows. The problem is formulated in Section II. The main results of the paper, consisting of an SDP relaxation as well as a two-stage convex program, is developed in Section III. Two case studies on power networks and mass-spring systems are provided in Section IV, followed by simulations on random systems. Finally, some concluding remarks are drawn in Section V.

**Notations:**  $\mathbb{R}$  and  $\mathbb{S}^n$  denote the sets of real numbers and  $n \times n$  symmetric matrices, respectively.  $\text{rank}\{W\}$  and  $\text{trace}\{W\}$  denote the rank and trace of a matrix  $W$ . The notation  $W \succeq 0$  means that  $W$  is symmetric and positive semidefinite. Given a matrix  $W$ , its  $(l, m)$  entry is denoted as  $W_{lm}$ . The superscript  $(\cdot)^{\text{opt}}$  is used to show the globally optimal value of an optimization parameter. The symbols  $(\cdot)^T$  and  $\|\cdot\|$  denote the transpose and 2-norm operators, respectively. The notation  $|x|$  shows the size of a vector  $x$ . The expected value of a random variable  $x$  is shown as  $\mathcal{E}\{x\}$ .

## II. PROBLEM FORMULATION

Consider the discrete-time system

$$\begin{cases} x[\tau + 1] = Ax[\tau] + Bu[\tau] + Ed[\tau] \\ y[\tau] = Cx[\tau] + Fv[\tau] \end{cases} \quad \tau = 0, 1, 2, \dots$$

with the known matrices  $A$ ,  $B$ ,  $C$ ,  $E$ , and  $F$ , where

- $x[\tau] \in \mathbb{R}^n$ ,  $u[\tau] \in \mathbb{R}^m$  and  $y[\tau] \in \mathbb{R}^r$  denote the state, input and output of the system.
- $d[\tau]$  and  $v[\tau]$  denote the input disturbance and measurement noise, which are assumed to be zero-mean white-noise random processes.

The goal is to design an optimal distributed controller. With no loss of generality and in order to simplify the presentation, we focus on the static case where the objective is to design a static controller of the form  $u[\tau] = Ky[\tau]$  under the constraint that the controller gain  $K$  must belong to a given linear subspace  $\mathcal{K} \subseteq \mathbb{R}^{m \times r}$ . The set  $\mathcal{K}$  captures the sparsity structure of the unknown constrained controller  $u[\tau] = Ky[\tau]$  and, more specifically, it contains all  $m \times r$  real-valued matrices with forced zeros in certain entries. This paper is mainly concerned with the following problem.

**Stochastic Optimal Distributed Control (SODC) problem:** Design a stabilizing static controller  $u[\tau] = Ky[\tau]$  to minimize the cost function

$$\lim_{\tau \rightarrow +\infty} \mathcal{E} \left( x[\tau]^T Q x[\tau] + u[\tau]^T R u[\tau] \right) \quad (1)$$

subject to the system dynamics and the controller requirement  $K \in \mathcal{K}$ , given positive-definite matrices  $Q$  and  $R$ .

## III. MAIN RESULTS

Define two covariance matrices as below:

$$\Sigma_d = \mathcal{E}\{Ed[\tau]d[\tau]^T E^T\}, \quad \Sigma_v = \mathcal{E}\{Fv[\tau]v[\tau]^T F^T\} \quad (2)$$

for every  $\tau \in \{0, 1, 2, \dots\}$ . In what follows, the SODC problem will be formulated as a nonlinear optimization program.

**Theorem 1.** *The SODC problem is equivalent to finding a controller  $K \in \mathcal{K}$ , a symmetric Lyapunov matrix  $P \in \mathbb{S}^n$ , and auxiliary matrices  $G \in \mathbb{S}^n$ ,  $L \in \mathbb{R}^{n \times r}$  and  $M \in \mathbb{S}^r$  to minimize the objective function*

$$\text{trace}\{P\Sigma_d + M\Sigma_v + K^T R K \Sigma_v\} \quad (3)$$

subject to the constraints

$$\begin{bmatrix} G & G & (AG + BL)^T & L^T \\ G & Q^{-1} & 0 & 0 \\ AG + BL & 0 & G & 0 \\ L & 0 & 0 & R^{-1} \end{bmatrix} \succeq 0, \quad (4a)$$

$$\begin{bmatrix} P & I \\ I & G \end{bmatrix} \succeq 0, \quad (4b)$$

$$\begin{bmatrix} M & (BK)^T \\ BK & G \end{bmatrix} \succeq 0, \quad (4c)$$

$$L = KCG \quad (4d)$$

*Proof:* A sketch of the proof will be provided in this extended abstract. It is straightforward to verify that

$$\begin{aligned} x[\tau] &= (A + BKC)^\tau x[0] \\ &+ \sum_{t=0}^{\tau-1} (A + BKC)^t Ed[\tau - t - 1] \\ &+ \sum_{t=0}^{\tau-1} (A + BKC)^t BK F v[\tau - t - 1] \end{aligned} \quad (5)$$

for  $\tau = 1, 2, \dots$ . On the other hand, since the controller under design must be stabilizing,  $(A + BKC)^\tau$  approaches zero as  $\tau$  goes to  $+\infty$ . In light of the above equation, it can be verified that

$$\begin{aligned} \mathcal{E} \left\{ \lim_{\tau \rightarrow +\infty} (x[\tau]^T Q x[\tau] + u[\tau]^T R u[\tau]) \right\} &= \\ = \mathcal{E} \left\{ \lim_{\tau \rightarrow +\infty} (x[\tau]^T (Q + C^T K^T R K C) x[\tau]) \right\} &= \\ = \text{trace}\{P\Sigma_d + (BK)^T P (BK) \Sigma_v + K^T R K \Sigma_v\} \end{aligned} \quad (6)$$

where

$$P = \sum_{t=0}^{\infty} ((A + BKC)^t)^T (Q + C^T K^T R K C) (A + BKC)^t \quad (7)$$

or equivalently

$$(A + BKC)^T P (A + BKC) - P + Q + (KC)^T R (KC) = 0, \quad (8a)$$

$$P \succeq 0 \quad (8b)$$

After pre- and post-multiplying (8a) with  $P^{-1}$  and using the Schur complement formula, the constraints (8a) and (8b) can be combined as

$$\begin{bmatrix} P^{-1} & P^{-1} & S^T & P^{-1}(KC)^T \\ P^{-1} & Q^{-1} & 0 & 0 \\ S & 0 & P^{-1} & 0 \\ (KC)P^{-1} & 0 & 0 & R^{-1} \end{bmatrix} \succeq 0 \quad (9)$$

where  $S = (A + BKC)P^{-1}$  (note that 0's in the above matrix are zero matrices of appropriate dimensions). By changing the

variable  $P^{-1}$  to  $G$  and defining  $L$  as  $KCP^{-1}$ , the following observations can be made:

- The condition (9) is identical to the set of constraints (4a) and (4d).
- The cost function (6) can be expressed as

$$\text{trace}\{G^{-1}\Sigma_d + (BK)^T G^{-1}(BK)\Sigma_v + K^T R K \Sigma_v\} \quad (10)$$

- Since  $P$  appears only once in the constraints of the optimization problem (3)-(4) (i.e., the condition (4b)) and the objective function of this optimization includes the term  $\text{trace}\{P\Sigma_d\}$ , the optimal value of  $P$  is equal to  $G^{-1}$ .
- Similarly, the optimal value of  $M$  is equal to  $(BK)^T G^{-1}(BK)$ .

The proof follows from the above facts.  $\blacksquare$

The SODC problem is cast as a (deterministic) nonlinear program in Theorem 1. This optimization problem is non-convex due only to the complicating constraint (4d). More precisely, the removal of this nonlinear constraint makes the optimization problem a semidefinite program (note that the term  $K^T R K$  in the objective function is convex due to the assumption  $R \succ 0$ ).

The traditional  $H_2$  optimal control problem (i.e., in the centralized case) can be solved using Riccati equations. It will be shown in the next theorem that the abovementioned semidefinite program correctly solves the centralized  $H_2$  optimal control problem.

**Theorem 2.** *Consider the special case where  $C = I$ ,  $\Sigma_v = 0$  and  $\mathcal{K}$  contains the set of all unstructured controllers. Then, the SODC problem has the same solution as the convex optimization problem obtained from the nonlinear optimization (3)-(4) by removing its non-convex constraint (4d).*

*Proof:* It is easy to verify that a solution  $(K^{\text{opt}}, P^{\text{opt}}, G^{\text{opt}}, L^{\text{opt}}, M^{\text{opt}})$  of the convex problem stated in the theorem can be mapped to the solution  $(L^{\text{opt}}(G^{\text{opt}})^{-1}, P^{\text{opt}}, G^{\text{opt}}, L^{\text{opt}}, M^{\text{opt}})$  of the non-convex problem (3)-(4) and vice versa (recall that  $C = I$  and  $\Sigma_v = 0$  by assumption). This completes the proof.  $\blacksquare$

Theorem 2 states that a classical optimal control problem can be precisely solved via a convex relaxation of the nonlinear optimization (3)-(4) by eliminating its constraint (4d). However, this simple convex relaxation does not work satisfactorily for a general control structure  $\mathcal{K}$ . To design a better relaxation, define

$$w := [1 \quad h \quad \text{vec}\{CG\}]^T \quad (11)$$

where  $h$  is a row vector containing the variables (free parameters) of  $K$ , and  $\text{vec}\{CG\}$  is a row vector containing all entries of  $G$ . It is possible to write the bilinear matrix term  $KCG$  as a linear function of the entries of the parametric matrix  $ww^T$ . Hence, by introducing a new matrix variable  $W$  playing the role of  $ww^T$ , the nonlinear constraint (4d) can be rewritten as a linear constraint in term of  $W$ . Now, one can relax the mapping constraint  $W = ww^T$  to  $W \succeq 0$  and another constraint stating that the first column of  $W$  is equal to  $w$ . This convex problem is referred to as **SDP relaxation of the SODC** in this work. In the case where the relaxation

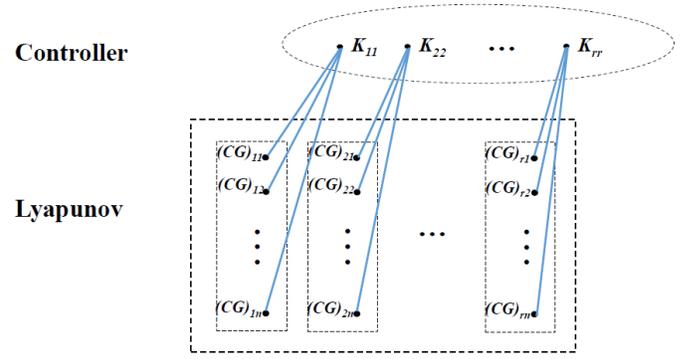


Fig. 1: The sparsity graph for the SDP relaxation of the SODC problem in the case where  $\mathcal{K}$  consists of diagonal matrices (vertex 1 is ignored to improve the legibility of the paper). Note that  $(CG)_{ij}$ 's denote the entries of  $CG$ .

has the same solution as SODC, the relaxation is said to be exact.

**Theorem 3.** *Consider the case where  $\mathcal{K}$  contains only diagonal matrices. The following statements hold regarding the SDP relaxation of the SODC problem:*

- The relaxation is exact if it has a solution  $(K^{\text{opt}}, P^{\text{opt}}, G^{\text{opt}}, L^{\text{opt}}, M^{\text{opt}}, W^{\text{opt}})$  such that  $\text{rank}\{W^{\text{opt}}\} = 1$ .*
- The relaxation always has a solution  $(K^{\text{opt}}, P^{\text{opt}}, G^{\text{opt}}, L^{\text{opt}}, M^{\text{opt}}, W^{\text{opt}})$  such that  $\text{rank}\{W^{\text{opt}}\} \leq 3$ .*

*Proof:* To study the SDP relaxation of the aforementioned control problem, we need to define a sparsity graph  $\mathcal{G}$ . Let  $\eta$  denote the number of rows of  $W$ . The graph  $\mathcal{G}$  has  $\eta$  vertices with the property that two arbitrary disparate vertices  $i, j \in \{1, 2, \dots, \eta\}$  are connected in the graph if  $W_{ij}$  appears in at least one of the constraints of the SDP relaxation other than the global constraint  $W \succeq 0$ . For example, vertex 1 is connected to all remaining vertices of the graph. The graph  $\mathcal{G}$  with its vertex 1 removed is depicted in Figure 1. This graph is acyclic and therefore the treewidth of the graph  $\mathcal{G}$  is at most 2. Hence, the SDP relaxation has a matrix solution with rank at most 3 [29].  $\blacksquare$

Theorem 3 states that the SDP relaxation of the SODC problem has a low-rank solution. However, it does not imply that every solution of the relaxation is low-rank. It follows from our recent work [29] that there are infinitely many convex programs that are able to turn a high-rank solution of the SDP relaxation into a matrix solution with rank at most 3 for the design of a diagonal controller. Theorem 3 will be generalized below.

**Proposition 1.** *The SODC problem has a convex relaxation with the property that its exactness amounts to the existence of a rank-1 matrix solution  $W^{\text{opt}}$ . Moreover, it is always guaranteed that this relaxation has a solution such that  $\text{rank}\{W^{\text{opt}}\} \leq 3$ .*

*Proof:* As discussed in our recent paper [28], there exist two binary matrices  $\Phi_1$  and  $\Phi_2$  such that  $K = \Phi_1 \overline{\text{diag}}\{k\} \Phi_2$  for every  $K \in \mathcal{K}$ , where  $\overline{\text{diag}}\{k\}$  denotes a diagonal matrix whose diagonal contains the free (variable) entries of  $K$ . Hence, the design of a structured control gain  $K$  for the system

( $A, B, C, E, F$ ) amounts to the design of a diagonal control gain  $\text{diag}\{k\}$  for the system ( $A, B\Phi_1, \Phi_2C, E, \Phi_2F$ ) (after updating the matrices  $Q$  and  $R$  accordingly). It follows from Theorem 3 that the SDP relaxation of the SODC problem equivalently formulated for the new system satisfies the properties of this theorem. ■

In this section, it has been shown that the SODC problem has an SDP relaxation with a low-rank solution. Nevertheless, there are many SDP relaxations with this property and it is desirable to find the one offering the highest lower bound on the optimal solution of the SODC problem. To this end, the abovementioned SDP relaxation should be reformulated in such a way that the diagonal entries of the matrix  $W$  are incorporated into as many constraints of the problem as possible in order to indirectly penalize the rank of the matrix  $W$ . This idea will be flourished in the next section, but for a computationally-cheap relaxation of the SOCP problem.

#### A. Computationally-Cheap SDP Relaxation

The proposed SDP relaxation has a high dimension for a large-scale system, which makes it less interesting for computational purposes. Moreover, the quality of its optimal objective value can be improved using some indirect penalty technique. The objective of this section is to offer a computationally-cheap SDP relaxation for the SOCP problem, whose solution outperforms that of the previous SDP relaxation. For this purpose, Consider an invertible matrix  $\Phi$  such that

$$C\Phi = \begin{bmatrix} \Lambda & 0 \end{bmatrix} \quad (12)$$

where  $\Lambda$  is a diagonal matrix and “0” is an  $r \times (n - r)$  zero matrix. Define also

$$\mathcal{K}^2 = \{KK^T \mid K \in \mathcal{K}\} \quad (13)$$

Indeed,  $\mathcal{K}^2$  captures the sparsity pattern of the matrix  $KK^T$ . For example, if  $\mathcal{K}$  consists of block-diagonal (rectangular) matrix,  $\mathcal{K}^2$  will also include block-diagonal (square) matrices. Let  $\mu_1 \in \mathbb{R}$  and  $\mu_2 \in \mathbb{R}$  be two positive numbers such that

$$Q \succ \mu_1 \times \Phi^{-T}\Phi^{-1}, \quad R \succ \mu_2 \times \Lambda^2 \quad (14)$$

where  $\Phi^{-T}$  denotes the transpose of the inverse of  $\Phi$ . Define  $\hat{Q} := Q - \mu_1 \times \Phi^{-T}\Phi^{-1}$  and  $\hat{R} := R - \mu_2 \times \Lambda^2$ .

**Computationally-Cheap SDP Relaxation:** This optimization problem is defined as the minimization of

$$\text{trace}\{P\Sigma_d + M\Sigma_v + \mu_2\mathbf{W}_{33}\Sigma_v + K^T\hat{R}K\Sigma_v\} \quad (15)$$

subject to the constraints

$$\begin{bmatrix} G - \mathbf{W}_{22} & G & (AG + BL)^T & L^T \\ G & \hat{Q}^{-1} & 0 & 0 \\ AG + BL & 0 & G & 0 \\ L & 0 & 0 & R^{-1} \end{bmatrix} \succeq 0, \quad (16a)$$

$$\begin{bmatrix} P & I \\ I & G \end{bmatrix} \succeq 0, \quad (16b)$$

$$\begin{bmatrix} M & (BK)^T \\ BK & G \end{bmatrix} \succeq 0, \quad (16c)$$

$$\mathbf{W} := \begin{bmatrix} I_n & \vdots & \Phi^{-1}G & \begin{bmatrix} \Lambda K^T \\ 0 \end{bmatrix} \\ \vdots & \hat{Q}\Phi^{-T} & \vdots & \vdots \\ \vdots & \vdots & \mathbf{W}_{22} & \vdots \\ \vdots & \begin{bmatrix} K\Lambda & 0 \end{bmatrix} & \vdots & \mathbf{W}_{33} \end{bmatrix} \succeq 0, \quad (16d)$$

$$K \in \mathcal{K}, \quad (16e)$$

$$\mathbf{W}_{33} \in \mathcal{K}^2, \quad (16f)$$

with the parameter set  $\{K, L, G, P, M, \mathbf{W}\}$ , where the dependent variables  $\mathbf{W}_{22}$  and  $\mathbf{W}_{33}$  represent two blocks of  $\mathbf{W}$ .

The following remarks can be made regarding the computationally-cheap SDP relaxation:

- The constraint (16a) corresponds to the Lyapunov inequality, where  $\mathbf{W}_{22}$  in its first block aims to play the role of  $P^{-2}$ .
- The constraint (16b) ensures that the relation  $P = G^{-1}$  occurs at optimality (at least for one of the solution of the problem).
- The constraint (16c) ensures that the relation  $M = (BK)^T G^{-1}(BK)$  occurs at optimality
- The constraint (16d) is a surrogate for the only complicating constraint of the SODC problem, i.e.,  $L = KCG$ .
- Since no non-convex rank constraint is imposed in order to maintain the convexity of the relaxation, a possible rank constraint is compensated in various ways. More precisely, the entries of  $\mathbf{W}$  are constrained in the objective function (15) through the term  $\text{trace}\{\mu_2\mathbf{W}_{33}\Sigma_v\}$ , in the first block of the constraint (16a) through the term  $G - \mathbf{W}_{22}$ , and also via the constraints (16e) and (16f).
- The proposed relaxation takes advantage of the sparsity of not only  $K$ , but also  $KK^T$  (through the constraint (16f)).

**Theorem 4.** *The computationally-cheap SDP relaxation is a convex relaxation of the SODC problem, whose optimal cost is greater than or equal to that of the (regular) SDP relaxation with the variable set  $(K, L, P, G, M, W)$ . Furthermore, the relaxation is exact if it possesses a solution  $(K^{\text{opt}}, L^{\text{opt}}, P^{\text{opt}}, G^{\text{opt}}, M^{\text{opt}}, \mathbf{W}^{\text{opt}})$  such that  $\text{rank}\{\mathbf{W}^{\text{opt}}\} = n$ .*

*Proof:* Since the proof is long, it is omitted in this extended abstract. ■

Once the computationally-cheap SDP relaxation is solved, a controller  $K$  must be recovered. This can be achieved in two ways as explained below.

**Direct Recovery Method:** A near-optimal controller  $\hat{K}$  for the SODC problem is chosen to be equal to the optimal controller  $K^{\text{opt}}$  obtained from the computationally-cheap SDP relaxation.

**Indirect Recovery Method:** Let  $(K^{\text{opt}}, L^{\text{opt}}, P^{\text{opt}}, G^{\text{opt}}, M^{\text{opt}},$

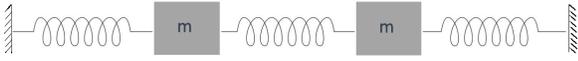


Fig. 2: A mass-spring system with two masses

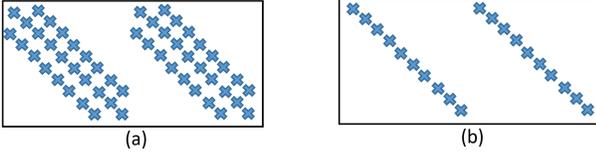


Fig. 3: Two control topologies studied for the mass-spring system: (a) localized, (b) decentralized.

$\mathbf{W}^{\text{opt}}$ ) denote a solution of the computationally-cheap SDP relaxation. A near-optimal controller  $\hat{K}$  for the SODC problem is recovered by solving a convex program with the variables  $K \in \mathcal{K}$  and  $\gamma \in \mathbb{R}$  to minimize the cost function

$$\varepsilon \times \gamma + \text{trace}\{(BK)^T (G^{\text{opt}})^{-1} (BK) \Sigma_v + K^T R K \Sigma_v\} \quad (17)$$

subject to the constraint

$$\begin{bmatrix} (G^{\text{opt}})^{-1} - Q + \gamma I_n & (A + BKC)^T & (KC)^T \\ (A + BKC) & G^{\text{opt}} & 0 \\ (KC) & 0 & R^{-1} \end{bmatrix} \succ 0 \quad (18)$$

where  $\varepsilon$  is a pre-specified nonnegative number.

The direct recovery method assumes that the controller  $K^{\text{opt}}$  obtained from the computationally-cheap SDP relaxation is near-optimal, whereas the indirect method assumes that the controller  $K^{\text{opt}}$  might be quite wrong while the inverse of the Lyapunov matrix is near-optimal. The indirect method is built on the SDP relaxation by fixing  $G$  at its optimal value and then perturbing  $Q$  as  $Q - \gamma I_n$  to facilitate the recovery of a stabilizing controller (it may rarely happen that a stabilizing controller can be recovered from a solution  $G^{\text{opt}}$  if  $\gamma$  is set to zero). Although none of the proposed recovery methods is universally better than the other one, we have verified in numerous simulation that the indirect recovery method significantly outperforms the direct recovery method with a high probability.

#### IV. SIMULATIONS RESULTS

The objective of this section is to elucidate the results of this work on two physical systems and several random systems. Note that the computation time for each SDP relaxation studied below has been less than 2 seconds on a desktop computer with an Intel Core i7 quad-core 3.4 GHz CPU and 16 GB RAM (using CVX and MOSEK).

##### A. Case Study I: Mass-Spring System

In this part, the aim is to evaluate the performance of the proposed controller design technique on the *Mass-Spring* system, as a classical physical system. Consider a mass-spring system consisting of  $N$  masses. This system is exemplified in Figure 2 for  $N = 2$ . The system can be modeled in the continuous-time domain as

$$\dot{x}_c(t) = A_c x_c(t) + B_c u_c(t) \quad (19)$$

where the state vector  $x_c(t)$  can be partitioned as  $[o_1(t)^T \ o_2(t)^T]^T$  with  $o_1(t) \in \mathbb{R}^n$  equal to the vector of positions and  $o_2(t) \in \mathbb{R}^n$  equal to the vector of velocities of the  $N$  masses. In this example, we assume that  $N = 10$  and adopt the values of  $A_c$  and  $B_c$  from [24]. The goal is to design a static sampled-data controller with a pre-specified structure (the controller is composed of a sampler, a static discrete-time controller and a zero-order holder). To this end, we first discretize the system with the sampling time of 0.1 seconds. Two control structures of ‘‘decentralized’’ and ‘‘localized’’ are to be studied for the matrix  $K \in \mathbb{R}^{10 \times 20}$ , which are shown in Figure 3. The entries marked as ‘‘\*’’ are the free entries of the rectangular matrix  $K$  under design. We assume that the system is subject to both input disturbance and measurement noise. Consider the case where  $\Sigma_d = I$  and  $\Sigma_v$  is a scalar multiple of the identity matrix such that the value of the identical diagonal entries of  $\Sigma_v$  could vary from 0 to 5. Using the computationally-cheap SDP relaxation in conjunction with the indirect recovery method, a near-optimal controller was designed for each of aforementioned control structures under various noise levels. The results are reported in Figure 4, with ‘‘optimality degree’’ defined as:

$$\text{Optimality degree (\%)} = 100 - \frac{\text{upper bound} - \text{lower bound}}{\text{upper bound}} \times 100$$

where ‘‘upper bound’’ and ‘‘lower bound’’ denote the cost of the recovered near-optimal controller and the optimal objective of the SDP relaxation, respectively. The structured controllers designed using the SDP relaxation were all stable with optimality degrees higher than 95% in the worst case and close to 99% in many cases.

##### B. Case Study II: Frequency Control in Power Systems

In this part, the performance of the computationally-cheap SDP relaxation combined with the indirect recovery method will be evaluated on the problem of designing an optimal distributed frequency control for IEEE 39-Bus New England Power System. The one-line diagram of this system is shown in Figure 5. The main objective of the unknown controller is to optimally adjust the mechanical power input to each generator as well as being structurally constrained by a user-defined communication topology. This pre-determined communication topology specifies which generators exchange their rotor angle and frequency measurements with one another.

In this example, we stick with a simple classical model of the power system. However, our result can be deployed for a complicated high-order model with nonlinear terms (our SDP relaxation may be revised to handle possible nonlinear terms in the dynamics). To derive a simple state-space model of the power system, we start with the widely used per-unit swing equation

$$M_i \ddot{\theta}_i + D_i \dot{\theta}_i = P_{Mi} - P_{Ei} \quad (20)$$

where  $\theta_i$  denotes the voltage (or rotor) angle at bus  $i$  (in rad),  $P_{Mi}$  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 coefficient of the generator at bus  $i$  (in pu-sec<sup>2</sup>/rad), and  $D_i$  is the damping coefficient of

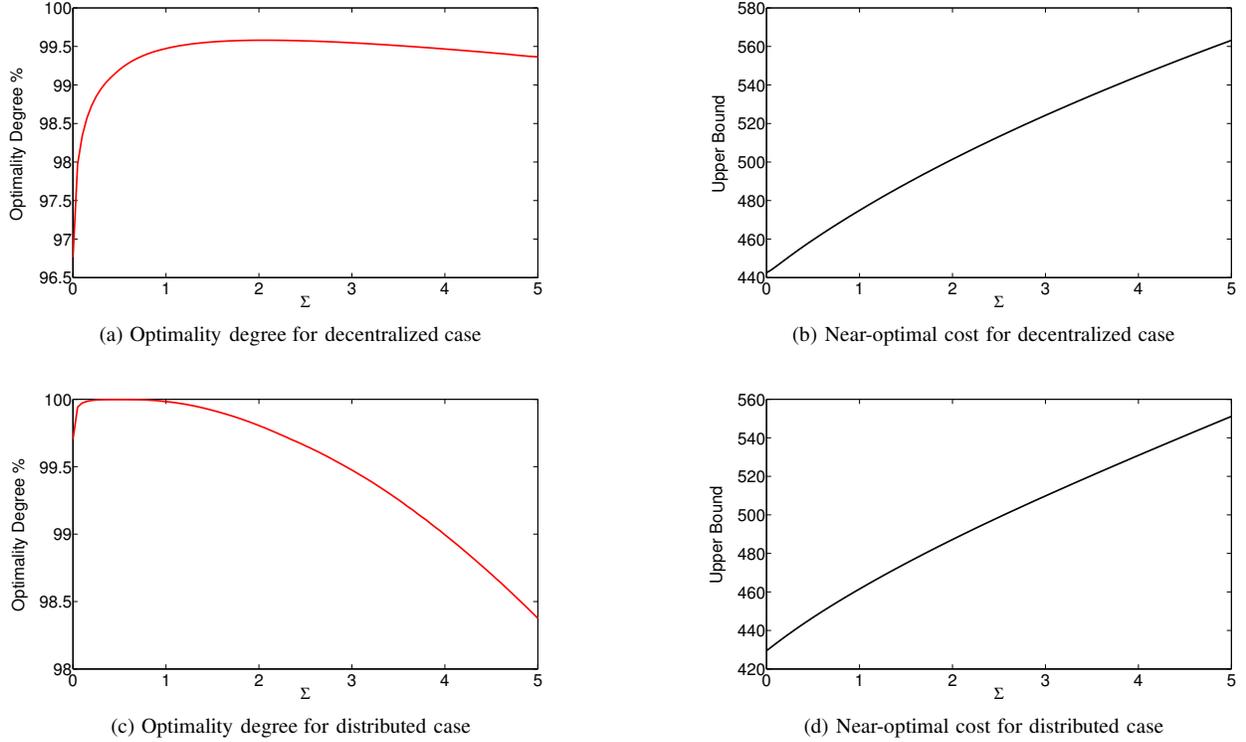


Fig. 4: The optimality degree and the optimal cost of the near-optimal controller designed for the mass-spring system for two different control structures. The noise covariance matrix  $\Sigma_v$  is assumed to be a multiple of the identity matrix and the x-axis in each figure shows the value of the identical diagonal entries of  $\Sigma_v$ .

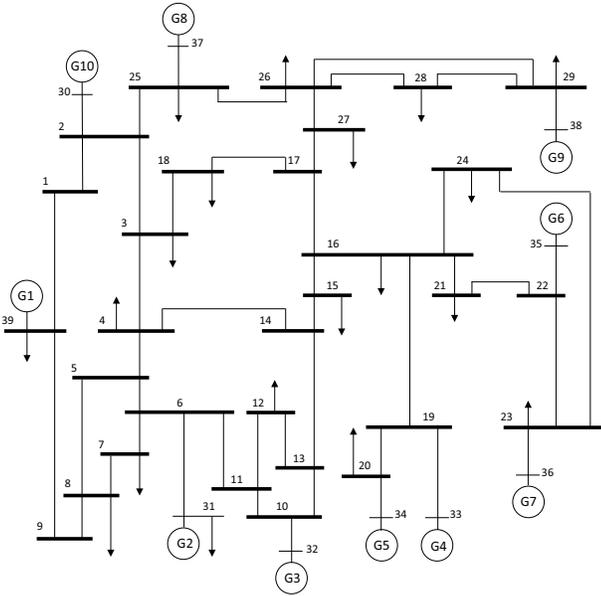


Fig. 5: Single line diagram of IEEE 39-Bus New England Power System.

equations:

$$P_{Ei} = \sum_{j=1}^n |V_i||V_j| [ G_{ij} \cos(\theta_i - \theta_j) + B_{ij} \sin(\theta_i - \theta_j) ] \quad (21)$$

where  $n$  denotes the number of buses in the system,  $V_i$  is the voltage phasor at bus  $i$ ,  $\theta_i$  is the voltage (or rotor) angle 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 approximate equation (21) by its corresponding DC power flow equation stated below:

$$P_{Ei} = \sum_{j=1}^n B_{ij}(\theta_i - \theta_j) \quad (22)$$

The approximation error is often small in practice due to the common practice of power engineering, which rests upon the following assumptions:

- For most networks,  $G \ll B \rightarrow G = 0$
- For most neighbouring buses,  $|\theta_i - \theta_j| \leq (10^\circ \text{ to } 15^\circ)$   
 $\rightarrow \sin(\theta_i - \theta_j) \approx \theta_i - \theta_j$   
 $\rightarrow \cos(\theta_i - \theta_j) \approx 1$
- In per unit,  $|V_i|$  is close to 1 (0.95 to 1.05)  
 $\rightarrow |V_i||V_j| \approx 1$

It is possible to rewrite (22) 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 generator buses (by removing the load buses and the intermediate zero buses). In this

the generator at bus  $i$  (in pu-sec/rad) [30]. The electrical real power  $P_{Ei}$  in (20) comes from the nonlinear AC power flow

equation,  $L$  denotes the Laplacian matrix and can be found as follows [31]:

$$\begin{aligned} L_{ii} &= \sum_{j=1, j \neq i}^{\bar{n}} B_{ij}^{Kron} & \text{if } i = j \\ L_{ij} &= -B_{ij}^{Kron} & \text{if } i \neq j \end{aligned} \quad (23)$$

where  $B^{Kron}$  is the susceptance of the Kron reduced admittance matrix  $Y^{Kron}$  defined as

$$Y_{ij}^{Kron} = Y_{ij} - \frac{Y_{ik}Y_{kj}}{Y_{kk}} \quad (i, j = 1, 2, \dots, n \text{ and } i, j \neq k) \quad (24)$$

where  $k$  is the index of the non-generator bus to be eliminated from the admittance matrix and  $\bar{n}$  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 [32].

By defining the rotor angle state vector as  $\theta = [\theta_1, \dots, \theta_{\bar{n}}]^T$  and the frequency state vector as  $w = [w_1, \dots, w_{\bar{n}}]^T$  and by substituting the matrix format of  $P_E$  into (20), the state space model of the swing equation used for frequency control in power systems could be written as

$$\begin{bmatrix} \dot{\theta} \\ \dot{w} \end{bmatrix} = \begin{bmatrix} 0_{\bar{n} \times \bar{n}} & I_{\bar{n}} \\ -M^{-1}L & -M^{-1}D \end{bmatrix} \begin{bmatrix} \theta \\ w \end{bmatrix} + \begin{bmatrix} 0_{\bar{n} \times \bar{n}} \\ M^{-1} \end{bmatrix} P_M \quad (25)$$

$$y = \begin{bmatrix} \theta \\ w \end{bmatrix}$$

where  $M = \text{diag}(M_1, \dots, M_{\bar{n}})$  and  $D = \text{diag}(D_1, \dots, D_{\bar{n}})$ . It is assumed that both rotor angle and frequency are available for measurement at each generator (implying that  $C = I_{2\bar{n}}$ ). This is a reasonable assumption with the recent advances in Phasor Measurement Unit (PMU) technology [33].

By substituting the per-unit inertia (M) and damping (D) coefficients for the 10 generators of IEEE 39-Bus system [34] based on the data in Table I, the continuous-time state space model matrices  $A_c$ ,  $B_c$  and  $C_c$  can be found. The system is then discretized to the discrete-time model matrices  $A$ ,  $B$  and  $C$  with the sampling time of 0.2 second. The initial values of the rotor angle ( $\theta_0$ ) were calculated by solving power (or load) flow problem for the system using MATPOWER [35]. In practice, the rotor speed does not vary significantly from synchronous speed and thus the initial frequency ( $w_0$ ) was assumed to be 1.0 per unit. Both  $\theta_0$  and  $w_0$  are reported for each generator in Table I.

The 39-bus system has 10 generators, labeled as  $G_1, G_2, \dots, G_{10}$ . Four communication topologies are considered in this work: decentralized, localized, star, and ring. In order to better understand how the interactions among the 10 generators in the system are related to the communication structures, the Kron reduced network of the system is visualized by the weighted graph shown in Figure 6. In a fully decentralized structure, none of the generators communicate with each other. In a localized communication structure, the generators may only communicate with their close neighbors. In a star topology, a single generator is able to communicate with all other

Bus	Gen	M	D	$\theta_0$	$w_0$
30	G10	4	5	-0.0839	1.0
31	G2	3	4	0.0000	1.0
32	G3	2.5	4	0.0325	1.0
33	G4	4	6	0.0451	1.0
34	G5	2	3.5	0.0194	1.0
35	G6	3.5	3	-0.0073	1.0
36	G7	3	7.5	0.1304	1.0
37	G8	2.5	4	0.0211	1.0
38	G9	2	6.5	0.127	1.0
39	G1	6	5	-0.2074	1.0

TABLE I: Generators' data and initial values (in per unit) for IEEE 39-Bus New England Power System.

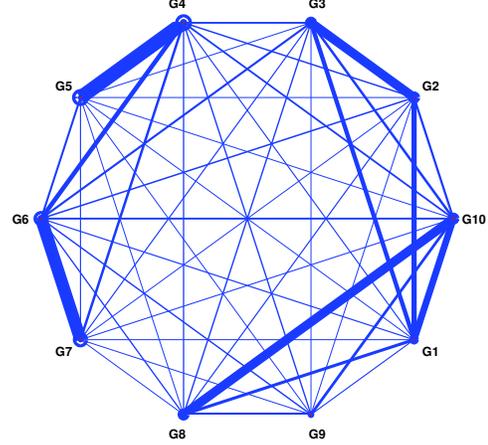


Fig. 6: Weighted graph of the Kron reduced network of IEEE 39-Bus New England Power System. Weights (thicknesses) of all edges are normalized to the minimum off-diagonal entry of the susceptance  $B^{Kron}$ .

generators in the system. The ring communication structure—forming a closed path—aims to provide communications between neighbors. These topologies are visualized in Figure 7. The locations of the generators in the figure are based on the exact coordinates of the power plants named in [36]. Note that  $G_1$  represents a group of generators, but it is considered as a single node near the border between New York and Connecticut in this map.  $G_4$  and  $G_5$  are very close in distance, but  $G_4$  was somewhat shifted from its real coordinates to make the communication link between them visible in this map.

To conduct some experiments, assume that  $Q = I$ ,  $R = 0.1I$ ,  $\Sigma_d = I$ , and  $\Sigma_v$  is a multiple of the identity matrix where the value of its identical diagonal entries could vary from 0 to 15. Using the proposed SDP relaxation, near-optimal controllers were designed for each of the four foregoing communication topologies under various levels of noise. The designed structured controllers were always stable with optimality degrees always higher than 90%. The findings are reported in Figures 8 and 9

### C. Random Systems

The goal of this example is to test the efficiency of the computationally-cheap SDP relaxation combined with the indirect recovery method on 100 highly-unstable random systems. Assume that  $n = m = r = 25$ , and that  $C, Q, R$  are identity matrices of appropriate dimensions. Suppose that

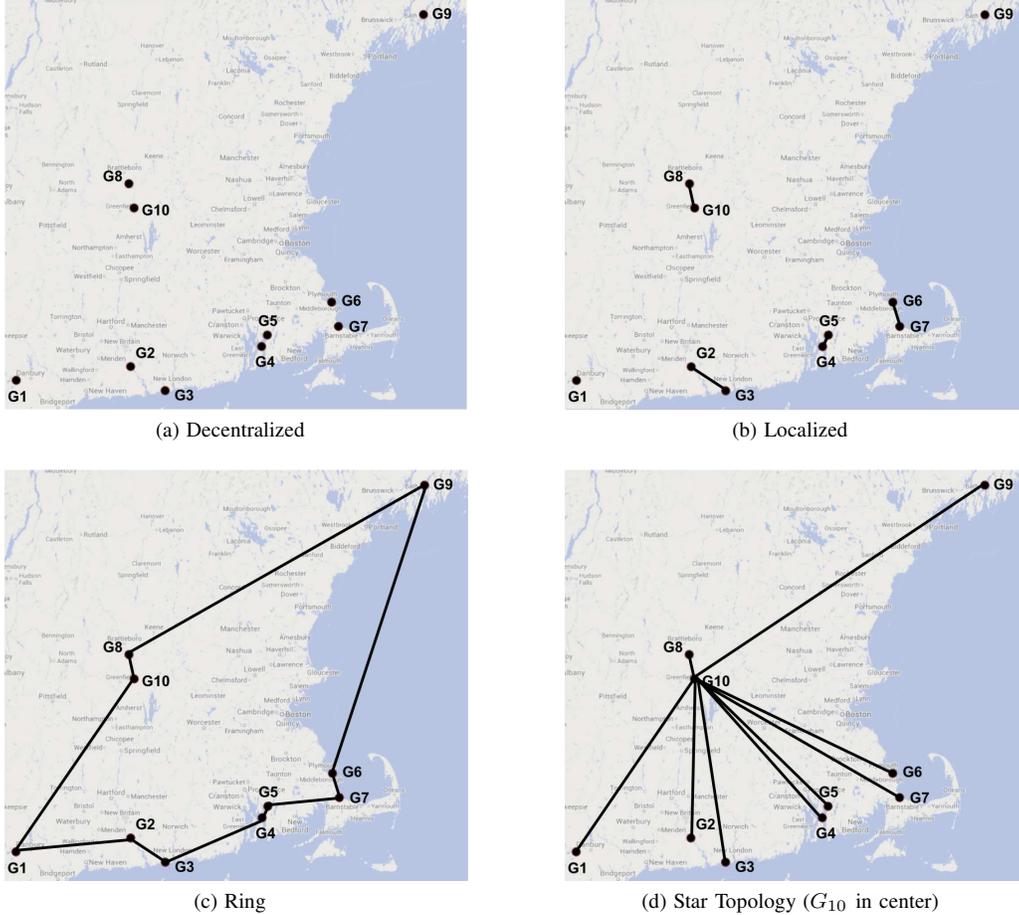


Fig. 7: Four communication topologies studied for IEEE 39-bus system.

$\Sigma_d = I$  and  $\Sigma_v = 0$ . To make the problem harder, assume that the controller under design must satisfy the hard constraint  $\text{trace}\{KK^T\} \leq 2$  (to avoid a high gain  $K$ ). We generated hundred random tuples  $(A, B, \mathcal{K})$  according to the following rules:

- The entries of  $A$  were uniformly chosen from the interval  $[0, 0.5]$  at random.
- The entries of  $B$  were uniformly chosen from the interval  $[0, 1]$  at random.
- Each entry of the matrix  $K$  was enforced to be zero with the probability of 70%.

Note that although the matrices  $A$  and  $B$  are nonnegative, the matrix  $K$  under design can have both positive and negative entries. The randomly generated systems are highly unstable with the maximum absolute eigenvalue as high as 6 (instability for discrete-time systems requires a maximum magnitude less than 1). Although the control of such systems was not easy and the control structure was enforced to be 70% sparse with an enforced sparsity pattern, the proposed technique was always able to design a “stabilizing” near-optimal controller with an optimality degree between 50% and 75%. The results are reported in Figure 10.

## V. CONCLUSIONS

This paper studies the stochastic optimal distributed control (SODC) problem for discrete-time systems. The objective is to design a (static) distributed controller with a pre-determined structure to minimize the expected value of a quadratic cost functional. We cast the SODC problem as a rank-constrained optimization with only one non-convex constraint requiring the rank of a variable matrix to be 1. Dropping the rank constraint leads to a semidefinite programming (SDP) relaxation of the SODC problem. It is shown that the nonlinearity of the SODC problem appears in such a sparse way that its SDP relaxation has a matrix solution with rank at most 3. This low-rank solution helps to find a near-global solution of the SODC problem. Since the derived SDP relaxation is computationally expensive, an efficient two-stage convex program is developed. A computationally-cheap SDP relaxation is first solved, and then part of the SDP solution is fed into a second convex program to retrieve a near-global controller. This two-stage optimization problem works as well as the Riccati equations in the centralized case. The proposed technique is tested on two physical networks and several random systems to show the possibility of designing structured controllers with global optimality degrees as high as 90%.

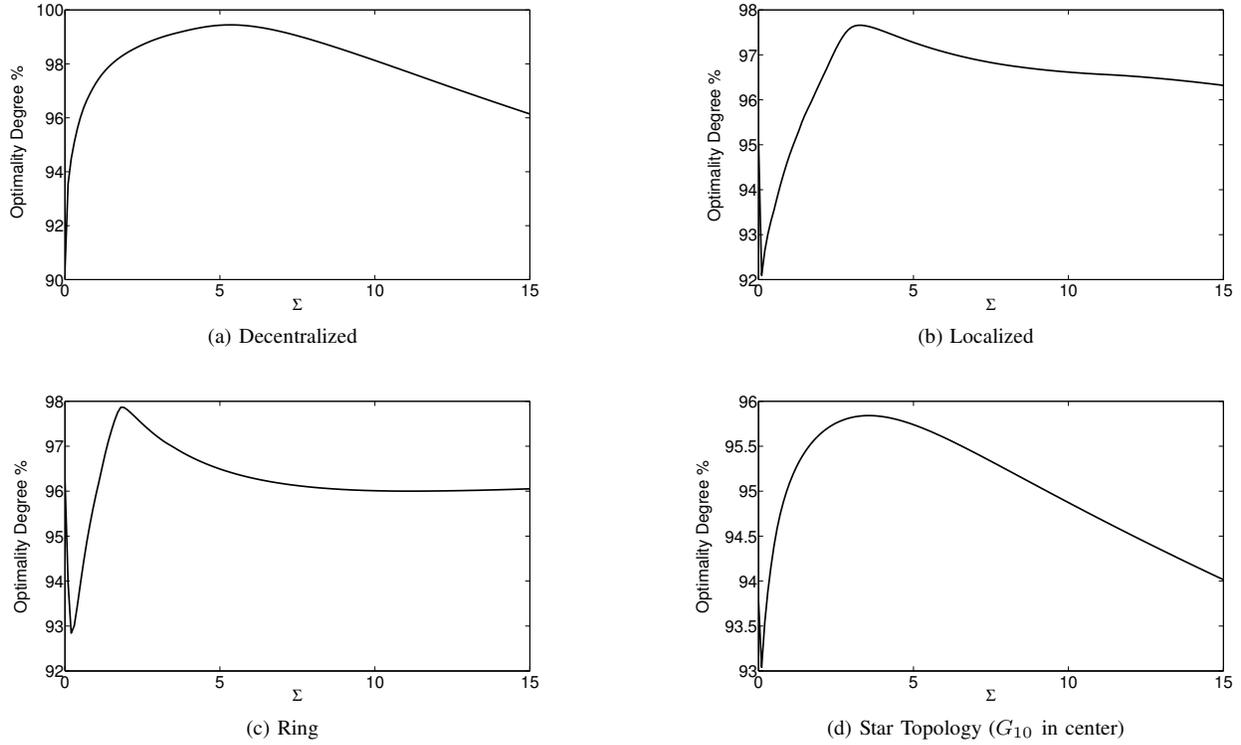


Fig. 8: The optimality degree of the near-optimal controller designed for the power system for four different control structures. The noise covariance matrix  $\Sigma_v$  is assumed to be a multiple of the identity matrix and the x-axis in each figure shows the value of the identical diagonal entries of  $\Sigma_v$ .

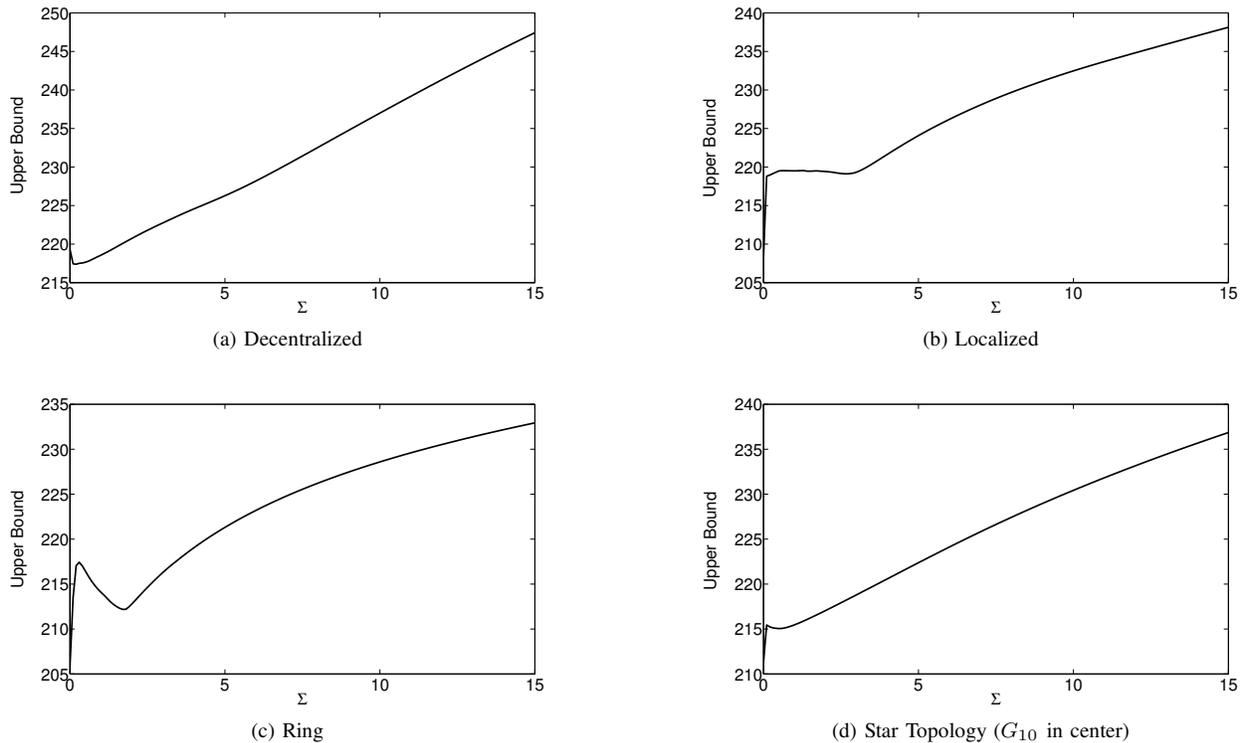


Fig. 9: The optimal cost of the near-optimal controller designed for the power system for four different control structures. The noise covariance matrix  $\Sigma_v$  is assumed to be a multiple of the identity matrix and the x-axis in each figure shows the value of the identical diagonal entries of  $\Sigma_v$ .

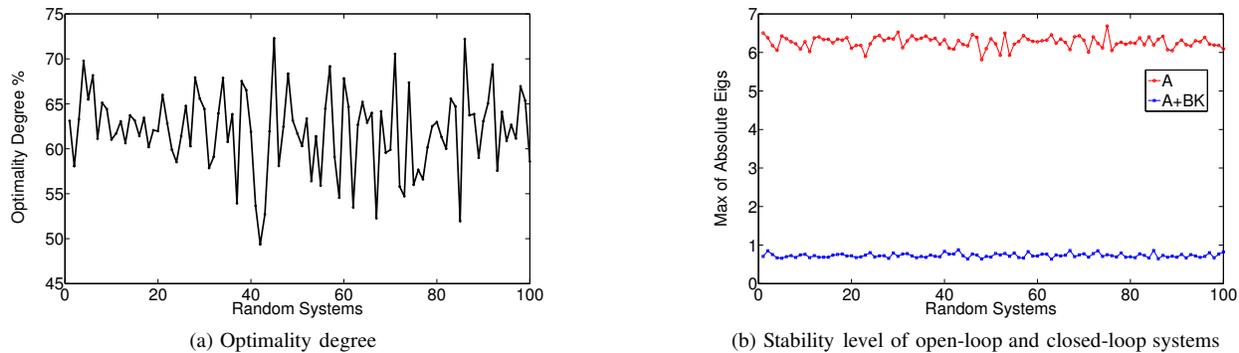


Fig. 10: The optimality degree and the stability level (maximum of the absolute eigenvalues) associated with 100 near-optimal sparse controllers designed for 100 highly-unstable random systems.

## 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] R. D'Andrea and G. Dullerud, "Distributed control design for spatially interconnected systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 9, pp. 1478–1495, 2003.
- [4] 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.
- [5] C. Langbort, R. Chandra, and R. D'Andrea, "Distributed control design for systems interconnected over an arbitrary graph," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1502–1519, 2004.
- [6] N. Motee and A. Jadbabaie, "Optimal control of spatially distributed systems," *Automatic Control, IEEE Transactions on*, vol. 53, no. 7, pp. 1616–1629, 2008.
- [7] G. Dullerud and R. D'Andrea, "Distributed control of heterogeneous systems," *IEEE Transactions on Automatic Control*, vol. 49, no. 12, pp. 2113–2128, 2004.
- [8] T. Keviczky, F. Borrelli, and G. J. Balas, "Decentralized receding horizon control for large scale dynamically decoupled systems," *Automatica*, vol. 42, no. 12, pp. 2105–2115, 2006.
- [9] F. Borrelli and T. Keviczky, "Distributed LQR design for identical dynamically decoupled systems," *Automatic Control, IEEE Transactions on*, vol. 53, no. 8, pp. 1901–1912, 2008.
- [10] D. D. Siljak, "Decentralized control and computations: status and prospects," *Annual Reviews in Control*, vol. 20, pp. 131–141, 1996.
- [11] J. Lavaei, "Decentralized implementation of centralized controllers for interconnected systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 7, pp. 1860–1865, 2012.
- [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," *Automatic Control, IEEE Transactions on*, 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," in *American Control Conference*, 2013.
- [17] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *Automatic Control, IEEE Transactions on*, vol. 51, no. 2, pp. 274–286, 2006.
- [18] 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.
- [19] L. Lessard and S. Lall, "Optimal controller synthesis for the decentralized two-player problem with output feedback," *American Control Conference*, 2012.
- [20] A. Lamperski and J. C. Doyle, "Output feedback  $H_2$  model matching for decentralized systems with delays," *American Control Conference*, 2013.
- [21] M. Rotkowitz and N. Martins, "On the nearest quadratically invariant information constraint," *Automatic Control, IEEE Transactions on*, vol. 57, no. 5, pp. 1314–1319, 2012.
- [22] 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.
- [23] A. Rantzer, "Distributed control of positive systems," <http://arxiv.org/abs/1203.0047>, 2012.
- [24] F. Lin, M. Fardad, and M. R. Jovanovi, "Design of optimal sparse feedback gains via the alternating direction method of multipliers," *IEEE Transactions on Automatic Control*, vol. 58, no. 9, 2013.
- [25] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, 1996.
- [26] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, 2004.
- [27] J. Lavaei, "Optimal decentralized control problem as a rank-constrained optimization," *52th Annual Allerton Conference on Communication, Control and Computing*, 2013.
- [28] G. Fazelnia, R. Madani, and J. Lavaei, "Convex relaxation for optimal distributed control problem," Submitted to *IEEE Conference on Decision and Control*, 2014.
- [29] R. Madani, G. Fazelnia, S. Sojoudi, and J. Lavaei, "Finding low-rank solutions of sparse linear matrix inequalities using convex optimization," *Technical Report*. [Online]. Available: [http://www.ee.columbia.edu/~lavaei/LMI\\_Low\\_Rank.pdf](http://www.ee.columbia.edu/~lavaei/LMI_Low_Rank.pdf)
- [30] M. A. Pai, *Energy Function Analysis for Power System Stability*. Kluwer Academic Publishers, Boston, 1989.
- [31] F. Dorfler and F. Bullo, "Novel insights into lossless ac and dc power flow," in *2013 IEEE Power and Energy Society General Meeting (PES)*, July 2013, pp. 1–5.
- [32] A. R. Bergen and V. Vittal, *Power Systems Analysis*. Prentice Hall, 1999, vol. 2.
- [33] M. Andreasson, D. Dimarogonas, H. Sandberg, and K. Johansson, "Distributed control of networked dynamical systems: Static feedback, integral action and consensus," *IEEE Transactions on Automatic Control*, vol. PP, no. 99, pp. 1–1, 2014.
- [34] I. E. Atawi, "An advance distributed control design for wide-area power system stability," Ph.D. dissertation, Swanson School Of Engineering,, University of Pittsburgh, Pittsburgh, Pennsylvania, 2013.
- [35] R. Zimmerman, C. Murillo-Sanchez, and R. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb 2011.
- [36] M. Ilic, J. Lacalle-Melero, F. Nishimura, W. Schenler, D. Shirmohammadi, A. Crough, and A. Catelli, "Short-term economic energy management in a competitive utility environment," *IEEE Transactions on Power Systems*, vol. 8, no. 1, pp. 198–206, Feb 1993.