A Convex Approach to Sparse $\mathcal{H}_\infty$ Analysis & Synthesis

Seungil You and Nikolai Matni

***Abstract*— In this paper, we propose a new robust analysis tool motivated by large-scale systems. The $\mathcal{H}_\infty$ norm of a system measures its robustness by quantifying the worst-case behavior of a system perturbed by a unit-energy disturbance. However, the disturbance that induces such worst-case behavior requires perfect coordination among all disturbance channels. Given that many systems of interest, such as the power grid, the internet and automated vehicle platoons, are large-scale and spatially distributed, such coordination may not be possible, and hence the $\mathcal{H}_\infty$ norm, used as a measure of robustness, may be too conservative. We therefore propose a cardinality constrained variant of the $\mathcal{H}_\infty$ norm in which an adversarial disturbance can use only a limited number of channels. As this problem is inherently combinatorial, we present a semidefinite programming (SDP) relaxation based on the ℓ_1 norm that yields an upper bound on the cardinality constrained robustness problem. We further propose a simple rounding heuristic based on the optimal solution of our SDP relaxation, which provides a corresponding lower bound. Motivated by privacy in large-scale systems, we also extend these relaxations to computing the minimum gain of a system subject to a limited number of inputs. Finally, we also present a SDP based optimal controller synthesis method for minimizing the SDP relaxation of our novel robustness measure. The effectiveness of our semidefinite relaxation is demonstrated through numerical examples.**

I. Introduction

Structure, and in particular sparsity, has proven to be a powerful tool in the analysis and design of large-scale control systems. Lyapunov analysis [1], distributed performance certification [2], distributed optimal controller synthesis [3] and controller architecture design [4] all rely on and exploit structural properties of the underlying system to solve seemingly intractable problems in a computationally efficient manner. In contrast, in the context of robust control, adding additional structure to system uncertainty has traditionally made analysis and synthesis more difficult. For instance, linear matrix inequality (LMI) based necessary and sufficient conditions for the robust stability of a system subject to an unstructured delta block can be derived, but no such results exist if we restrict ourselves to highly structured delta blocks [5].

In this paper, we ask the following question, which we later interpret in terms of robustness to structured disturbances: given a large-scale system with p input channels, what $k \ll p$ input channels should be used to maximally (minimally) perturb the system using a unit energy input. We show that the answer can be obtained by suitably modifying the power semi-norm based definition of the $\mathcal{H}_\infty$ norm of a system to incorporate a cardinality constraint on the input; we therefore call the resulting performance metric the k-sparse $\mathcal{H}_\infty$ norm of the system.

We argue that questions pertaining to the maximal and minimal gains of a system restricted to a sparse subset of inputs arise naturally in the context of distributed system robustness analysis, consensus robustness analysis, privacy and system security. We further show that the resulting optimization problems are in fact a generalization of the maximal and minimal sparse eigenvalue problems, objects of central importance in certifying the performance of compressed sensing matrices [6] and in sparse PCA [7]. We also show touch upon how these restricted gains relate to similar conditions developed in the Regularization for Design (RFD) [4] framework that guarantee the recovery of optimal controller architectures.

Of course, the resulting optimization problems are combinatorial in nature, and are easily seen to be computationally difficult in general. Leveraging a novel *primal* formulation of the KYP lemma [8], we propose a semidefinite relaxation (akin to that proposed in [7]) for computing lower/upper bounds on the resulting minimal/maximal restricted gains of the system, and a simple rounding heuristic to obtain corresponding upper/lower bounds. We further derive the dual of the resulting semidefinite program (SDP) and show that it has similar structure to the traditional KYP LMI test, allowing for standard semidefinite programming based controller synthesis methods to be applied.

The paper is organized as follows: in Section II, we formally introduce the k-sparse $\mathcal{H}_\infty$ norm and the analogous k-sparse minimal gain of a system, and elucidate on several engineering applications. We also make connections to compressed sensing, restricted isometry constants and sparse PCA, as well as RFD. In Section III we present both a semidefinite relaxation and a rounding heuristic to obtain lower and upper bounds on the k-sparse $\mathcal{H}_\infty$ norm of a system. The dual to our semidefinite relaxation is derived in Section IV, and we show how it can be used to synthesize a centralized controller that minimizes the relaxed k-sparse $\mathcal{H}_\infty$ norm of the system. We present numerical examples in Section V, and end with a summary and discussion of future directions in Section VI.

A. Notation

We use $\mathcal{RH}_\infty$ to denote the space of stable real-rational proper transfer matrices. We use lower case Latin letters x to denote vectors, bold lower case Latin letters $\mathbf{x}$ to denote signals, upper case Latin letters X to denote matrices

This research was in part supported by NSF NetSE, AFOSR, the Institute for Collaborative Biotechnologies through grant W911NF-09-0001 from the U.S. Army Research Office

The authors are with the Control and Dynamical Systems, California Institute of Technology, Pasadena, CA 91125, USA {syou,nmatni}@caltech.edu

and upper case calligraphic letters $\mathcal{X}$ to denote elements of $\mathcal{RH}_\infty$.

We recall the definition of the power semi-norm, $\|\mathbf{x}\|_P^2 := \lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} x_k^* x_k$. For a matrix X, we denote the (i,j)th entry of X by X_{ij}, its conjugate transpose by X^*, its transpose by $X^\top$, the projection of X onto its diagonal elements by $\mathbf{diag}\,(X)$, and the range space and the null space of X by $\mathbf{Range}\,(X)$ and $\mathbf{Ker}\,(X)$, respectively. In addition, $|X|$ denotes the element-wise absolute value of X, and $\mathbf{1}$ denotes the all ones vector. We use $X \succeq 0$ to denote that X is positive semidefinite, and $X \succ 0$ to denote that X is positive definite.

II. k-SPARSE $\mathcal{H}_\infty$ ANALYSIS

We consider a discrete time linear time invariant system[1]

$$\mathcal{M}(z) = C(zI - A)^{-1}B + D \in \mathcal{RH}_\infty. \tag{1}$$

Recall that the $\mathcal{H}_\infty$ norm of $\mathcal{M}$ can be computed as the worst case gain in the output of the system induced by a disturbance of unit power semi-norm [9], [10], [8]:

$$\begin{aligned} \|\mathcal{M}\|_\infty^2 :=& \underset{\mathbf{w},\mathbf{x}}{\text{maximize}} && \|C\mathbf{x} + D\mathbf{w}\|_P^2 \\ & \text{subject to} && x_{k+1} = Ax_k + Bw_k,\ x_0 = 0, \\ &&& \|\mathbf{w}\|_P^2 \le 1. \end{aligned} \tag{2}$$

The $\mathcal{H}_\infty$ norm measures the worst-case behavior of a system subject to power semi-norm bounded disturbances, and it has well known implications on the robust stability of a system subject to unstructured uncertainty [5], as well as many practical interpretations [9].

One such interpretation is that an attacker seeks to maximize their disruption of the system using the disturbance $\mathbf{w}$ – in this case, the optimal disturbance $\mathbf{w}_\star$ to optimization problem (2) is precisely a disturbance that maximizes the attacker's impact on the system. Taking an opposite perspective, from the viewpoint of a system designer, the maximizing disturbance denotes a weak point of the system that may need to be addressed.

A seemingly innocuous assumption in the above analysis is that the attacker can simultaneously coordinate all of the disturbance channels: although reasonable in a centralized setting, this assumption may prove to be quite conservative when $\mathcal{M}$ is a distributed system. In particular, if there are many possible disturbances (B has many columns), and these disturbances enter through channels that are physically separated, it may be overly conservative to consider the response of the system to a centralized attack. In order to alleviate this conservativeness, we propose a *cardinality constrained* variation of optimization problem (2), in which we assume that at most k disturbance channels can have non-zero power semi-norms[2]:

$$\begin{aligned} \{\bar{\mu}_k(\mathcal{M})\}^2 :=& \underset{\mathbf{w},\mathbf{x}}{\text{maximize}} && \|C\mathbf{x} + D\mathbf{w}\|_P^2 \\ & \text{subject to} && x_{k+1} = Ax_k + Bw_k,\ x_0 = 0 \\ &&& \|\mathbf{w}\|_P^2 \le 1,\ \mathbf{Card}\,(\mathbf{w}) \le k. \end{aligned} \tag{3}$$

We refer to $\bar{\mu}_k(\mathcal{M})$ as the k-sparse $\mathcal{H}_\infty$ norm of system $\mathcal{M}$.

It should be clear that $\bar{\mu}_k(\mathcal{M}) \le \|\mathcal{M}\|_\infty$ for all k, but the size of the difference between these two quantities is unclear. If the gap is small, then this implies that the system exhibits near worst-case behavior even when perturbed by only a few carefully chosen disturbances, indicating a potential vulnerability that may need to be addressed. Conversely, if the gap is large, then considering the $\mathcal{H}_\infty$ norm of a system as a measure of robustness may be overly conservative. Regardless as to which situation occurs, the k-sparse $\mathcal{H}_\infty$ norm of a system provides valuable insight into its behavior and robustness properties.

Before elaborating on practical interpretations of the k-sparse $\mathcal{H}_\infty$ norm of a system, we show that the gap between $\bar{\mu}_k(\cdot)$ and $\|\cdot\|_\infty$ can indeed be made arbitrarily large for a fixed k by letting the state dimension of the underlying system tend to infinity. We defer illustrating the case for which the gap between the k-sparse and standard $\mathcal{H}_\infty$ norm of a system is small to Section V.A, where we present a power-grid motivated example for which $\bar{\mu}_3(\mathcal{M}) \approx \|\mathcal{M}\|_\infty$.

Example 1: *Consider a system $\mathcal{M}$, as in* (1)*, described by state-space parameters $(A, I_n, I_n, 0_{n,n})$, where $A = 0.99\frac{1}{n}\mathbf{1}\mathbf{1}^\top + 0.1(I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^\top) \in \mathbb{R}^{n\times n}$, I_n is the n-by-n identity matrix, and $0_{n,n}$ is the n-by-n zero matrix. Due to the special structure of the state-space parameters, optimization problems* (2) *and* (3) *can be solved analytically, and it can be shown that $\frac{\bar{\mu}_k(\mathcal{M})}{\|\mathcal{M}\|_\infty} \approx \sqrt{\frac{k}{n}}$. Thus for a fixed k the gap between $\bar{\mu}_k(\mathcal{M})$ and $\|\mathcal{M}\|_\infty$ can be made arbitrarily large by letting $n \to \infty$. Figure 1 shows $\frac{\bar{\mu}_k(\mathcal{M})}{\|\mathcal{M}\|_\infty}$ for $k = 5$ and $n = 5, \ldots, 30$.*

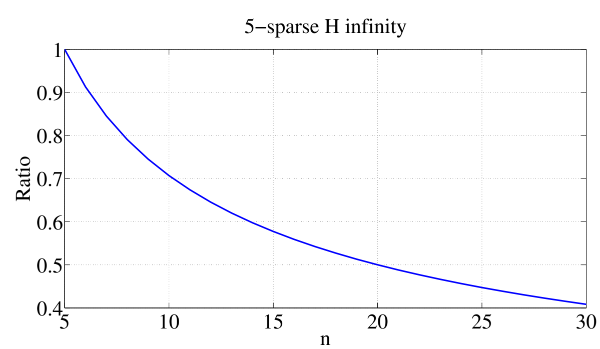

Fig. 1: The ratio, $\bar{\mu}_5(\mathcal{M})/\|\mathcal{M}\|_\infty$ for $n = 5, \cdots, 30$.

Robustness analysis for distributed system: Quantifying the robustness of a distributed system, such as the power grid,

[1] Although we present our analysis for discrete-time systems, analogous arguments and results hold for continuous time systems.

[2] We define the cardinality of a power signal, $\mathbf{Card}\,(\mathbf{w})$ as the number of indices i such that $\|\mathbf{w}_i\|_P > 0$. Notice that because we use the power semi-norm in this definition, all signals with finite ℓ_2 norm have a cardinality of 0, as their power semi-norm is 0.

allows the system designer to plan for and mitigate the worst case effects of un-modeled dynamics and disturbances. The need for robustness is increasingly important in the context of the power grid as it becomes more reliant on intermittent distributed energy resources, such as renewables. However, as mentioned, $\mathcal{H}_\infty$ analysis assumes that all such distributed energy resources coordinate with each other to destabilize the power network, which may be overly conservative and lead to loss of efficiency. Rather, we propose using the k-sparse $\mathcal{H}_\infty$ norm of the system to identify and quantify vulnerabilities of the system to potentially more realistic disturbances.

Robustness analysis for consensus network: The well studied problem of consensus (or synchronization) [11], [12], [13] is one in which a set of agents seek to converge to a common value using simple local averaging rules. When these local rules are linear and time invariant, the consensus protocol can be modeled as an LTI system. In this case, the A matrix defining the system is shown to satisfy the following properties [14]: $A\mathbf{1} = \mathbf{1}$, $A^\top \mathbf{1} = \mathbf{1}$, and $\rho(A - \frac{1}{n}\mathbf{1}\mathbf{1}^\top) < 1$, where n is the number of nodes in the network.

Although typically considered in a disturbance free setting, it is also natural to ask how much local disturbances applied to individual agents can affect the system's ability to reach consensus. Concretely, assume that each agent can be corrupted by a separate disturbance, i.e., that $B = I_n$, and we measure the effect of the disturbances on the deviation of each state x_k^j from the consensus value, as encoded by $z_k^j = x_k^j - \frac{1}{n}\sum_i x_k^i$, such that $C = I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^\top$, and $D = 0$. Note that the marginally stable mode of A is unobservable with respect to the measured output defined by C, and the system has a finite $\mathcal{H}_\infty$ norm and k-sparse $\mathcal{H}_\infty$ norm.

Whereas the $\mathcal{H}_\infty$ norm of the resulting system measures the effects of a worst-case attack on all agents, the k-sparse $\mathcal{H}_\infty$ norm measures the effects of worst-case attack on only k agents. From an attacker's perspective, this may result in a more realistically implementable strategy, and from a system designer's perspective, this provides valuable information as to which agents should be most closely monitored and protected from attack.

A. The k-sparse minimal gain of a system

We can also define the minimal k-sparse gain of system $\mathcal{M}$, which we denote by $\underline{\mu}_k(\mathcal{M})$ as

$$\begin{aligned}\{\underline{\mu}_k(\mathcal{M})\}^2 := \underset{\mathbf{w},\mathbf{x}}{\text{minimize}} \quad & \|C\mathbf{x} + D\mathbf{w}\|_P^2 \\ \text{subject to} \quad & x_{k+1} = Ax_k + Bw_k,\ x_0 = 0 \\ & \|\mathbf{w}\|_P^2 \geq 1,\ \mathbf{Card}(\mathbf{w}) \leq k. \end{aligned} \tag{4}$$

Privacy: An immediate interpretation of this optimization problem is in terms of *privacy*. Suppose that a publicly available variable is defined by $z_k = Cx_k$, and that a user wishes to transfer at least γ units of power to $y_k = Gx_k + Hw_k$ while minimizing their effect on the public variable. The optimal action for the user to take can be determined by solving optimization problem (4) with the added constraint

$$\|G\mathbf{x} + H\mathbf{w}\|_P^2 \geq \gamma^2 \tag{5}$$

System security: One can also view the user in the above scenario as an attacker, and the publicly available variable as a system monitor: in this case, the optimal input $\mathbf{w}_\star$ corresponds to the least detectable input that still disrupts the output $\mathbf{y}$ by γ units of power. Allowing for sparse optimal inputs $\mathbf{w}_\star$ makes for more realistically implementable actions by either a user or an attacker.

Connections to the Restricted Isometry Property and Regularization for Design: Our problem formulation seeks the minimal and maximal gains of a linear operator restricted to k-sparse subspaces. When the linear operator is a static matrix D, instead of a dynamical system ($A = B = C = 0$), then the cardinality constrained optimization problems (3) and (4) compute precisely the maximal and minimal *restricted eigenvalues* [15] of the matrix $D^\top D$, that is the maximal and minimal gains of D restricted to sparse subspaces. They are also closely linked to the Restricted Isometry Property (RIP) constant of the matrix, which can be used to state conditions for the recovery of sparse vectors [6] via convex optimization. The approach of computing restricted eigenvalues and their corresponding eigenvectors can also be used to perform sparse principal component analysis (sPCA) [7]. We can therefore view optimization problem (3) as a tool for bounding the restricted eigenvalues of an infinite dimensional LTI operator acting on signals in ℓ_2. Moreover, the k-sparse $\mathcal{H}_\infty$ norm and the k-sparse minimal gain of a system also have natural connections to the Regularization for Design (RFD) framework developed in [4]. In the RFD framework, atomic norms [16] are added as convex penalties to traditional model matching problems in order to design architecturally simple controllers. Control theoretic analogs to the recovery conditions found in the structured inference literature are stated in terms of restricted gains that are closely related to the k-sparse $\mathcal{H}_\infty$ norm and k-sparse minimal gain of a system – we are currently actively exploring the application of the computational methods developed in this paper to computing bounds on these restricted gains.

III. SDP RELAXATION OF k-SPARSE $\mathcal{H}_\infty$ ANALYSIS

As posed, optimization problems (3) and (4) are intractable: the optimization variables are infinite dimensional, and the cardinality constraint introduces a combinatorial aspect to the problem. In order to develop a computationally tractable framework, we propose an SDP based convex relaxation of the k-sparse $\mathcal{H}_\infty$ norm (3) and the k-sparse minimal gain (4). We begin by reviewing recent results on traditional $\mathcal{H}_\infty$ analysis [8], [17].

A. Review of $\mathcal{H}_\infty$ analysis

From previous work [8], [17], we know that using the gramian

$$V := \lim_{n\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \begin{bmatrix} x_k \\ w_k \end{bmatrix} \begin{bmatrix} x_k \\ w_k \end{bmatrix}^* \succeq 0,$$

optimization problem (2) can be transformed to the following equivalent finite dimensional semidefinite program:

$$\begin{array}{ll} \underset{V \succeq 0}{\text{maximize}} & \mathbf{Tr}\left(\begin{bmatrix} C^*C & C^*D \\ D^*C & D^*D \end{bmatrix} V\right) \\ \text{subject to} & \begin{bmatrix} I & 0 \end{bmatrix} V \begin{bmatrix} I \\ 0 \end{bmatrix} = \begin{bmatrix} A & B \end{bmatrix} V \begin{bmatrix} A^* \\ B^* \end{bmatrix} \\ & \sum_{i=n+1}^{n+m} V_{ii} \le 1, \end{array} \tag{6}$$

where n is the dimension of the state $\mathbf{x}$, m is the dimension of the disturbance $\mathbf{w}$, and V_{ii} is the ith diagonal component of V. The key idea of the proof is to construct the sinusoid $\mathbf{w}$ that achieves the $\mathcal{H}_\infty$ norm using a rank one solution of the semidefinite program (6). In the construction of $\mathbf{w}$, there is no prior structure imposed on $\mathbf{w}$. This means that, in general, all m disturbance channel are active and must therefore coordinate their actions.

B. SDP relaxation of k-sparse $\mathcal{H}_\infty$ analysis

Building on the result of the previous section, we propose and analyze a semidefinite relaxation of optimization problem (3) that can be used to compute an *upper bound* to the k-sparse $\mathcal{H}_\infty$ norm of a system. The relaxation to the k-sparse minimal gain of a system (4) is analogous, and stated without proof. To begin with, let us use the matrix

$$V := \lim_{n\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \begin{bmatrix} x_k \\ w_k \end{bmatrix} \begin{bmatrix} x_k \\ w_k \end{bmatrix}^* \succeq 0,$$

as in $\mathcal{H}_\infty$ analysis. For notational convenience, we partition $V = \begin{bmatrix} X & R \\ R^* & W \end{bmatrix}$ where $X \in \mathbb{C}^{n\times n}$, and $W \in \mathbb{C}^{m\times m}$.

We begin with a simple observation that lets us translate the cardinality constraint on $\mathbf{w}$ to one on the matrix W.

Proposition 1: $\mathbf{Card}\,(\mathbf{w}) \le k$ *if and only if* $\mathbf{Card}\,(\mathbf{diag}\,(W)) \le k$.

Proof: From the definition, $W_{ii} = \|\mathbf{w}_i\|_P^2$, where W_{ii} is (i,i)th entry of W and $\mathbf{w}_i$ is the ith component of the vector-valued signal $\mathbf{w}$. Therefore $\mathbf{Card}\,(\mathbf{w}) = \mathbf{Card}\,(\mathbf{diag}\,(W))$. ■

By applying the same procedure from [8] used to derive the SDP used for $\mathcal{H}_\infty$ analysis, we obtain the following optimization problem, which provides an upper bound of (3).

$$\begin{array}{ll} \underset{X,R,W}{\text{maximize}} & \mathbf{Tr}\left(\begin{bmatrix} C^*C & C^*D \\ D^*C & D^*D \end{bmatrix} \begin{bmatrix} X & R \\ R^* & W \end{bmatrix}\right) \\ \text{subject to} & X = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X & R \\ R^* & W \end{bmatrix} \begin{bmatrix} A^* \\ B^* \end{bmatrix} \\ & \mathbf{Tr}\,(W) \le 1 \\ & \mathbf{Card}\,(\mathbf{diag}\,(W)) \le k \\ & \begin{bmatrix} X & R \\ R^* & W \end{bmatrix} \succeq 0. \end{array} \tag{7}$$

For the standard $\mathcal{H}_\infty$ problem, this SDP relaxation is tight: the proof consists of constructing a disturbance $\mathbf{w}$ that achieves the optimal value of the SDP. Similarly, once a solution to optimization problem (7) is obtained, we can consider a system with disturbance inputs specified by the support of the optimal disturbance, and thus apply the methods of [8]. Thus, the cardinality constrained SDP (7) is in fact equivalent to k-sparse $\mathcal{H}_\infty$ optimization (3).

In applying the techniques from [8], we have reduced the optimization problem to a finite dimensional semidefinite program with an added cardinality constraint $\mathbf{Card}\,(\mathbf{diag}\,(W)) \le k$. In order to circumvent the intractability of this constraint, we propose using an ℓ_1 relaxation [18]. This approach is inspired by [7], in which the authors consider the ℓ_1 relaxation of an analogous cardinality constraint to obtain a semidefinite relaxation of the sparse PCA problem, in which one seeks the leading sparse singular vector of a matrix (as mentioned previously, this is closely related to the RIP constant of a matrix and to analogous quantities in RFD). In order to adapt this idea to our problem formulation, we need the following observation.

Proposition 2: *Consider* $W \in \mathbb{C}^{n\times n}$ *such that* $W \succeq 0$, $\mathbf{Tr}\,(W) \le 1$. *Then,* $\mathbf{1}^T|W|\mathbf{1} \le n$.

Proof: Consider a Hermitian matrix H where

$$H_{ij} = \begin{cases} 1 & \text{if } i = j \\ e^{\mathbf{i}\theta_{ij}} & \text{if } i \ne j, \end{cases}$$

for some θ_{ij}. If we construct H such that $H_{ij} = e^{\mathbf{i}\angle W_{ij}}$, then $\mathbf{1}^T|W|\mathbf{1} = \mathbf{Tr}\,(H^*W)$. This shows that $\mathbf{1}^T|W|\mathbf{1} \le \sup_H \mathbf{Tr}\,(H^*W)$, and from the Von Neumann's trace inequality [19], we have

$$\mathbf{Tr}\,(H^*W) \le \sum_i \sigma_i(W)\sigma_i(H),$$

where σ_i is the ith singular value of the matrix. Furthermore, by definition of H we have $\sigma_1(H) \le \sum_i \sigma_i(H) = \mathbf{Tr}\,(H) = n$. Therefore,

$$\begin{aligned} \mathbf{Tr}\,(H^*W) &\le \sum_i \sigma_i(W)\sigma_i(H) \le \sigma_1(H)\sum_i \sigma_i(W) \\ &\le n\,\mathbf{Tr}\,(W) \le n, \end{aligned}$$

and $\mathbf{1}^T|W|\mathbf{1} \le \sup_H \mathbf{Tr}\,(H^*W) \le n$. Notice that this upper bound is achieved by $W = \frac{1}{n}\mathbf{1}\mathbf{1}^T$, which shows the inequality is tight. ■

We can now connect the ℓ_1 norm bound to the cardinality constraint of optimization problem (3).

Proposition 3: *Consider a positive semidefinite matrix* W *with* $\mathbf{Tr}\,(W) \le 1$ *and* $\mathbf{Card}\,(\mathbf{diag}\,(W)) \le k$. *Then,* $\mathbf{1}^T|W|\mathbf{1} \le k$.

Proof: Without loss of generality, we can assume that $W_{11}, \cdots, W_{ii}$ for $1 \le i \le k$ are not zero, where W_{ii} is (i,i)th entry of W. Then from the Schur complement, we can easily check that W should have the form

$$W = \left[\begin{array}{c|c} \tilde{W} & 0 \\ \hline 0 & 0 \end{array}\right],$$

where $\tilde{W}$ is a $i \times i$ Hermitian matrix. Therefore, from Proposition 2, $\mathbf{1}^T|W|\mathbf{1} = \mathbf{1}^T|\tilde{W}|\mathbf{1} \leq i \leq k$, which concludes the proof. ∎

In the cardinality constrained problem (7), the W matrix satisfies the requirements of Proposition 3. This shows that if we replace the cardinality constraint on W in optimization problem (7) by a suitable ℓ_1 norm bound, then we have a larger feasible set. Although this procedure provides an upper bound to (7), the resulting optimization becomes a semidefinite program, which can be solved efficiently [20]. Therefore, we propose the following ℓ_1 based relaxation of (7), which is the main optimization problem in this paper.

$$\begin{aligned}
\{\bar{\mu}_k^{sdp}(\mathcal{M})\}^2 := \max_{X,R,W} \quad & \mathbf{Tr}\left(\begin{bmatrix} C^*C & C^*D \\ D^*C & D^*D \end{bmatrix}\begin{bmatrix} X & R \\ R^* & W \end{bmatrix}\right) \\
\text{s.t.} \quad & X = \begin{bmatrix} A & B \end{bmatrix}\begin{bmatrix} X & R \\ R^* & W \end{bmatrix}\begin{bmatrix} A^* \\ B^* \end{bmatrix} \\
& \mathbf{Tr}(W) \leq 1 \\
& \mathbf{1}^T|W|\mathbf{1} \leq k \\
& \begin{bmatrix} X & R \\ R^* & W \end{bmatrix} \succeq 0.
\end{aligned} \tag{8}$$

Although we omit the details, a similar argument for continuous time systems yields

$$\begin{aligned}
\underset{X,R,W}{\text{maximize}} \quad & \mathbf{Tr}\left(\begin{bmatrix} C^*C & C^*D \\ D^*C & D^*D \end{bmatrix}\begin{bmatrix} X & R \\ R^* & W \end{bmatrix}\right) \\
\text{subject to} \quad & XA^* + AX + R^*B^* + BR = 0 \\
& \mathbf{Tr}(W) \leq 1 \\
& \mathbf{1}^T|W|\mathbf{1} \leq k \\
& \begin{bmatrix} X & R \\ R^* & W \end{bmatrix} \succeq 0.
\end{aligned} \tag{9}$$

C. Extension to k-sparse minimal gain

In the previous section, we introduced a k-sparse minimal gain. A similar approach can be used to obtain the following SDP relaxation of (4).

$$\begin{aligned}
\{\underline{\mu}_k^{sdp}(\mathcal{M})\}^2 := \min_{X,R,W} \quad & \mathbf{Tr}\left(\begin{bmatrix} C^*C & C^*D \\ D^*C & D^*D \end{bmatrix}\begin{bmatrix} X & R \\ R^* & W \end{bmatrix}\right) \\
\text{s.t.} \quad & X = \begin{bmatrix} A & B \end{bmatrix}\begin{bmatrix} X & R \\ R^* & W \end{bmatrix}\begin{bmatrix} A^* \\ B^* \end{bmatrix} \\
& \mathbf{Tr}(W) \geq 1 \\
& \mathbf{1}^T|W|\mathbf{1} \leq k \\
& \begin{bmatrix} X & R \\ R^* & W \end{bmatrix} \succeq 0
\end{aligned}$$

Similarly, for the continuous time case, we have

$$\begin{aligned}
\underset{X,R,W}{\text{minimize}} \quad & \mathbf{Tr}\left(\begin{bmatrix} C^*C & C^*D \\ D^*C & D^*D \end{bmatrix}\begin{bmatrix} X & R \\ R^* & W \end{bmatrix}\right) \\
\text{subject to} \quad & XA^* + AX + R^*B^* + BR = 0 \\
& \mathbf{Tr}(W) \geq 1 \\
& \mathbf{1}^T|W|\mathbf{1} \leq k \\
& \begin{bmatrix} X & R \\ R^* & W \end{bmatrix} \succeq 0.
\end{aligned} \tag{10}$$

D. Rounding heuristic for solution refinement

Let $W^\star$ be the optimal solution to optimization problem (8). Since this matrix contains information about the worst-case disturbance, we can extract candidate worst case disturbance channels, and use those to obtain a corresponding lower bound to the value of optimization problem (3). The approach is simple: identify the k largest entries of $\mathbf{diag}(W)$, which we denote by $\{W_{i_1 i_1}, W_{i_2 i_2}, \cdots, W_{i_k i_k}\}$, and then restrict B and D to the column space corresponding to these disturbance channels. We can then compute the traditional $\mathcal{H}_\infty$ norm of the system defined by these restricted B and D matrices using classical methods. As we are choosing specific disturbance channels, this procedure yields a lower bound of the k-sparse $\mathcal{H}_\infty$ norm (3) of a system. The procedure can thus be summarized as follows:

Rounding heuristic:

1) Solve (8) to obtain $W^\star$.
2) Find the indices $\{i_1, \cdots, i_k\}$ such that $W^\star_{i_1 i_1} \geq \cdots \geq W^\star_{i_k i_k} \geq \cdots \geq W_{i_n i_n}$.
3) Construct $E := \begin{bmatrix} e_{i_1} & \cdots & e_{i_k} \end{bmatrix} \in \mathbb{R}^{m \times k}$ using a standard basis $\{e_i\}$ of $\mathbb{R}^m$.
4) Let $\tilde{B} := BE$, $\tilde{D} = DE$, and obtain $\bar{\mu}_k^{round}(\mathcal{M}) := \|\tilde{B}(e^{\mathbf{i}\theta}I - A)^{-1}C + \tilde{D}\|_\infty$.

Notice that step 3 chooses $i_1, \cdots, i_k$ to be the active disturbance channels. From this rounding procedure we obtain the inequality

$$\bar{\mu}_k^{round}(\mathcal{M}) \leq \bar{\mu}_k(\mathcal{M}) \leq \bar{\mu}_k^{sdp}(\mathcal{M})$$

Therefore, if the gap between $\bar{\mu}_k^{round}(\mathcal{M})$ and $\bar{\mu}_k^{sdp}(\mathcal{M})$ is not large, then $\bar{\mu}_k^{round}(\mathcal{M})$ effectively solves the k-sparse $\mathcal{H}_\infty$ problem and returns a candidate set of worst case disturbance channels. Notice that this heuristic can also be applied to the continuous time case and the minimal gain computation, but we omit these details.

IV. Dual problem & Controller Synthesis

As optimization problem (8) is an SDP, it is natural to consider its Lagrangian dual problem. To do this, let us begin with the following observation.

Proposition 4: *For $w \geq 0$, $\lambda \in \mathbb{C}$,*

$$\sup_{x \in \mathbb{C}}\{-w|x| + \mathbf{Re}(\lambda x)\} = \begin{cases} 0 & \text{if } |\lambda| \leq w \\ +\infty & \text{otherwise} \end{cases}.$$

Proof: Suppose $|\lambda| > w$. Let $x = \alpha\lambda^*$. Then

$$-w|x| + \mathbf{Re}(\lambda x) = \alpha|\lambda|(|\lambda| - w).$$

By taking $\alpha \to \infty$, we obtain the result.

Suppose $|\lambda| \leq w$. From Cauchy-Schwartz inequality,

$$-w|x| + \mathbf{Re}(\lambda x) \leq -w|x| + |\lambda||x| \leq (|\lambda| - w)|x| \leq 0,$$

for all $x \in \mathbb{C}$. Since the upper bound is achieved by $x = 0$, we can conclude the proof. ∎

With this technical tool in hand, we may proceed to derive the dual to optimization problem (8). First, we form the Lagrangian function in terms of $V = \begin{bmatrix} X & R \\ R^* & W \end{bmatrix}$.

$$
\begin{aligned}
&L(V,P,Q,\lambda,t) := \\
&\mathbf{Tr}\,(QV) + \mathbf{Tr}\left(\begin{bmatrix} C^*C & C^*D \\ D^*C & D^*D \end{bmatrix} V\right) \\
&+\mathbf{Tr}\left(P\left(\begin{bmatrix} A & B \end{bmatrix} V \begin{bmatrix} A^* \\ B^* \end{bmatrix} - \begin{bmatrix} I & 0 \end{bmatrix} V \begin{bmatrix} I \\ 0 \end{bmatrix}\right)\right) \\
&+\lambda\left(1 - \mathbf{Tr}\left(\begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} V\right)\right) \\
&+t\left(k - \mathbf{Tr}\left(\begin{bmatrix} 0 & 0 \\ 0 & \mathbf{11}^T \end{bmatrix} |V|\right)\right),
\end{aligned}
$$

where $P = P^*$, $Q \succeq 0$, $\lambda \geq 0$, $t \geq 0$.

Using cyclic property of the trace operator and from Proposition 4, we can obtain the dual function $d(Q,P,\lambda,t) := \sup_{V=V^*} L(V,P,\lambda,t)$ which becomes $\lambda + k \cdot t$ when,

$$
\begin{aligned}
&\left|Q + \begin{bmatrix} C^*C & C^*D \\ D^*C & D^*D - \lambda I \end{bmatrix} + \begin{bmatrix} A^*PA - P & A^*PB \\ B^*PA & B^*PB \end{bmatrix}\right| \\
&\leq \begin{bmatrix} 0 & 0 \\ 0 & t\mathbf{11}^T \end{bmatrix},
\end{aligned} \tag{11}
$$

where the inequality $\leq$ is a component-wise inequality. In addition, $d(P,\lambda,t) = +\infty$ if (Q,P,λ,t) does not satisfy (11). By defining $Y = Y^*$ to be the lower right bottom of (11), we obtain the following dual optimization problem to (8):

$$
\begin{aligned}
\underset{P,Y,\lambda,t}{\text{minimize}} \quad & \lambda + k \cdot t \\
\text{subject to} \quad & \begin{bmatrix} A^*PA - P & A^*PB \\ B^*PA & B^*PB - \lambda I - Y \end{bmatrix} \\
& + \begin{bmatrix} C^*C & C^*D \\ D^*C & D^*D \end{bmatrix} \preceq 0 \\
& |Y| \leq t\mathbf{11}^T \\
& P = P^*, Y = Y^*, t \geq 0, \lambda \geq 0.
\end{aligned} \tag{12}
$$

Notice that if we set $t = 0$, then we recover the SDP derived from the KYP lemma which computes the $\mathcal{H}_\infty$ norm of the system. It is clear that $t = 0$ is a suboptimal solution of (12), and therefore we can easily see that the $\mathcal{H}_\infty$ norm is an upper bound of (12) which is consistent with the definition of the k-sparse $\mathcal{H}_\infty$ norm.

In addition, it can be easily checked that the dual problem (12) is strictly feasible when A is stable by setting $Y = 0$, $t = 1$ and sufficiently large λ. In particular, for sufficiently large λ, only the upper left block of the LMI constraint in (12) is relevant, and if A is stable, one can construct a $P \succ 0$ such that $A^*PA - P + C^*C \prec 0$. Thus the dual problem is strictly feasible, and by Slater's condition, the duality gap is zero.

A similar derivation for a continuous time system, (9), gives us

$$
\begin{aligned}
\underset{P,Y,\lambda,t}{\text{minimize}} \quad & \lambda + k \cdot t \\
\text{subject to} \quad & \begin{bmatrix} A^*P + PA & PB \\ B^*P & -\lambda I - Y \end{bmatrix} \\
& + \begin{bmatrix} C^*C & C^*D \\ D^*C & D^*D \end{bmatrix} \preceq 0 \\
& |Y| \leq t\mathbf{11}^T \\
& P = P^*, Y = Y^*, t \geq 0, \lambda \geq 0,
\end{aligned} \tag{13}
$$

although we omit the detailed derivation.

A. k-sparse $\mathcal{H}_\infty$ synthesis

Here we modify the LMI approach to $\mathcal{H}_\infty$ controller synthesis presented in [21] so that the resulting controller minimizes the proposed semidefinite relaxation (12) of the k-sparse $\mathcal{H}_\infty$ measure. To begin with, consider the dynamical system

$$
\begin{aligned}
x_{k+1} &= Ax_k + B_1 w_k + B_2 u_k \\
z_k &= C_1 x_k + D_{11} w_k + D_{12} u_k \\
y_k &= C_2 x_k + D_{21} w_k,
\end{aligned}
$$

with dynamic controller

$$
\begin{aligned}
\zeta_{k+1} &= A_K \zeta_k + B_K y_k \\
u_k &= C_K \zeta_k + D_K y_k.
\end{aligned}
$$

The synthesis goal is to find a stabilizing controller (A_K, B_K, C_K, D_K) that minimizes our *SDP relaxation* (8) of the k-sparse $\mathcal{H}_\infty$ norm of the closed loop system[3] $(A_{cl}, B_{cl}, C_{cl}, D_{cl})$, subject to $\rho(A_{cl}) < 1$. From this stability requirement, together with the Lyapunov stability theorem, we can assume that $P \succ 0$ holds in optimization problem (12). In addition, we assume that (A, B_2, C_2) is stabilizable and detectable, thus ensuring feasibility.

Using $(A_{cl}, B_{cl}, C_{cl}, D_{cl})$, the matrix inequality constraint in the dual of k-sparse $\mathcal{H}_\infty$ analysis, (12), is given by

$$
\begin{aligned}
&\begin{bmatrix} A_{cl}^* P_{cl} A_{cl} - P_{cl} & A_{cl}^* P_{cl} B_{cl} \\ B_{cl}^* P_{cl} A_{cl} & B_{cl}^* P_{cl} B_{cl} - \lambda I - Y \end{bmatrix} \\
&\qquad + \begin{bmatrix} C_{cl}^* C_{cl} & C_{cl}^* D_{cl} \\ D_{cl}^* C_{cl} & D_{cl}^* D_{cl} \end{bmatrix} \prec 0
\end{aligned} \tag{14}
$$

We change the non-strict matrix inequality of optimization problem (12) to a strict inequality: as A_{cl} is required to be stable, a strictly feasible solution exists for a given $(A_{cl}, B_{cl}, C_{cl}, D_{cl})$, and therefore this variation does not change the optimal value of the following synthesis problem

$$
\begin{aligned}
\underset{A_K,B_K,C_K,D_K,P_{cl},Y,\lambda,t}{\text{minimize}} \quad & \lambda + k \cdot t \\
\text{subject to} \quad & (14), |Y| \leq t\mathbf{11}^\top, P_{cl} \succ 0 \\
& Y = Y^*, t \geq 0, \lambda \geq 0.
\end{aligned} \tag{15}
$$

[3]We refer the reader to [21] for details on how to express the closed loop system parameters $(A_{cl}, B_{cl}, C_{cl}, D_{cl})$ in terms of the open loop state-space parameters and (A_K, B_K, C_K, D_K).

Now let $T = \lambda I + Y$, then since $T - D_{cl}^* D_{cl} \succ 0$, T is positive definite. Conjugating constraint (14) by $\begin{bmatrix} I & 0 \\ 0 & T^{-1/2} \end{bmatrix}$, we see that (14) is true if and only if $\|(C_{cl}(zI - A_{cl})B_{cl} + D_{cl})T^{-1/2}\|_\infty < 1$. This representation of (14) allows for a simple modification of the classic LMI method for $\mathcal{H}_\infty$ synthesis [21], yielding the the following two SDPs:

$$\begin{aligned} &\underset{P,Q,Y,\lambda,t}{\text{minimize}} \quad \lambda + k \cdot t \\ &\text{subject to} \quad \begin{bmatrix} P & I_n \\ I_n & Q \end{bmatrix} \succ 0 \\ &\Pi_c^* \begin{bmatrix} AQA^* - A & AQC_1^* & B_1 \\ C_1QA^* & C_1QC_1^* - I & D_{11} \\ B_1^* & D_{11}^* & -\lambda I - Y \end{bmatrix} \Pi_c \prec 0 \\ &\Pi_o^* \begin{bmatrix} A^*PA - A & A^*PB_1 & C_1^* \\ B_1^*PA & B_1^*PB_1 - \lambda I - Y & D_{11}^* \\ C_1 & D_{11} & -I \end{bmatrix} \Pi_o \prec 0 \\ &|Y| \leq t\mathbf{1}\mathbf{1}^\top, Y = Y^*, t \geq 0, \lambda \geq 0 \end{aligned} \tag{16}$$

where

$$\Pi_c = \begin{bmatrix} N_c & 0 \\ 0 & I \end{bmatrix} \qquad \Pi_o = \begin{bmatrix} N_o & 0 \\ 0 & I \end{bmatrix}$$

$$\begin{aligned} \mathbf{Range}(N_c) &= \mathbf{Ker}\left(\begin{bmatrix} B_2^* & D_{12}^* \end{bmatrix}\right), \qquad N_c^* N_c = I \\ \mathbf{Range}(N_o) &= \mathbf{Ker}\left(\begin{bmatrix} C_2 & D_{21} \end{bmatrix}\right), \qquad N_o^* N_o = I \end{aligned}$$

After obtaining P, Q by solving the above SDP (16), we construct P_{cl} by $P_{cl} = \begin{bmatrix} P & P_2^* \\ P_2 & I \end{bmatrix}$, where P_2 is given by $P - Q^{-1} = P_2 P_2^*$. By applying the Schur complement to (14), we have

$$\begin{bmatrix} P_{cl}^{-1} & 0 & A_{cl} & B_{cl} \\ 0 & I & C_{cl} & D_{cl} \\ A_{cl}^* & B_{cl}^* & P_{cl} & 0 \\ C_{cl}^* & D_{cl}^* & 0 & \lambda I + Y \end{bmatrix} \succ 0 \tag{17}$$

which is clearly an LMI for a fixed P_{cl}. Finally, the following optimization returns the controller that achieves the optimal value of (16):

$$\begin{aligned} &\underset{A_K,B_K,C_K,D_K,Y,\lambda,t}{\text{minimize}} \quad \lambda + k \cdot t \\ &\text{subject to} \quad (17), |Y| \leq t\mathbf{1}\mathbf{1}^\top, Y = Y^* \\ &\qquad\qquad\quad t \geq 0, \lambda \geq 0. \end{aligned} \tag{18}$$

In summary, two SDPs (16) and (18) are needed to construct the controller. The first step is to solve (16) to find P, Q to construct P_{cl}, and then solve (18) to find the controller (A_K, B_K, C_K, D_K).

V. Numerical examples

A. A linearized swing dynamics

We consider the linearized swing dynamics of the New England power network [22], a widely-used benchmark example in the power control community. The system parameters can be found online [23]. We denote the nominal frequency by ω^0, the voltage magnitude at bus i by v_i, and the nominal phase angle bus i by θ_i^0. Going forward, all variables presented should be interpreted as deviations from this nominal steady state.

We consider the *linearized* swing dynamics [24].

$$\begin{aligned} \frac{d}{dt}\begin{bmatrix} \omega_G \\ P \end{bmatrix} &= \underbrace{\begin{bmatrix} -H^{-1}D_G & H^{-1}M_1 \\ -YM_1^\top & -YM_2^\top D_L^{-1} M_2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \omega_G \\ P \end{bmatrix}}_{x} \\ &+ \underbrace{\begin{bmatrix} H^{-1} & 0 \\ 0 & -YM_2^\top D_L^{-1} \end{bmatrix}}_{B_w} \underbrace{\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}}_{w}, \end{aligned}$$

where $\omega_G \in \mathbb{R}^{n_G}$ and $\omega_L \in \mathbb{R}^{n_L}$ are the frequencies at the generator bus and the load bus, respectively, P is the branch power flow vector, M is the signed incidence matrix of the topology of the power network, $M_1 = \begin{bmatrix} I_{n_G} & 0_{n_L \times n_L} \end{bmatrix} M$, $M_2 = \begin{bmatrix} 0_{n_G \times n_G} & I_{n_L} \end{bmatrix} M$, and H, D_G, D_L, and Y are diagonal matrices with entries specified by the inertia of the synchronous generator, the damping term of each bus, the damping term of each line and the admittance of the transmission line, respectively. We note that $n_G = 10$ and $n_L = 29$, and thus there are 39 potential disturbance channels that can affect the system.

We set the output to be

$$z = \underbrace{\begin{bmatrix} H^{1/2} & 0 \\ 0 & Y^{-1/2} \end{bmatrix}}_{C} \begin{bmatrix} \omega_G \\ P \end{bmatrix},$$

as $\frac{1}{2} z^* z$ corresponds to the total energy stored in the linearized swing dynamics.

Fig. 2 compares the k-sparse $\mathcal{H}_\infty$ norm obtained using the rounding heuristic described in Sec. III.D to the standard $\mathcal{H}_\infty$ norm of the system. As can be seen, the gap between 3-sparse $\mathcal{H}_\infty$ norm and the $\mathcal{H}_\infty$ norm is negligible, even though only 3 out of a possible 39 disturbances are used. This shows that the buses identified by the 3-sparse $\mathcal{H}_\infty$ analysis are a potential weak spot of the New England power network.

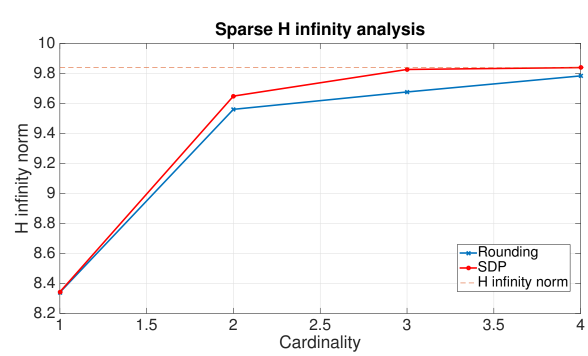

Fig. 2: The $\mathcal{H}_\infty$ norm and the lower bound of k-sparse $\mathcal{H}_\infty$ norm of New England power network. Bus 25, 30, 37 are identified in 3-sparse $\mathcal{H}_\infty$ analysis.

k-sparse controller	k-sparse $\mathcal{H}_\infty$ norm 1	2	3	$\mathcal{H}_\infty$
1	**1.1826**	1.3939	1.5078	1.7152
2	1.2289	**1.3340**	1.4116	1.5258
3	1.2509	1.3539	**1.4053**	1.5159
$\mathcal{H}_\infty$	1.3832	1.4172	1.4389	**1.5050**

TABLE I: k-sparse $\mathcal{H}_\infty$ norms of the controller obtained from our method. The minimum value in each column is emphasized.

B. *k-sparse $\mathcal{H}_\infty$ synthesis*

To illustrate the effectiveness of our synthesis approach, we apply our method to the following system:

$$A = \begin{bmatrix} 0.5 & 0.2 & 0 \\ 0.2 & 0.5 & 0.2 \\ 0 & 0.2 & 0.5 \end{bmatrix}, \quad B_1 = \begin{bmatrix} I_3 & 0_{3\times 3} \end{bmatrix}, \quad B_2 = I_3$$
$$C_1 = \begin{bmatrix} I_3 \\ 0_{3\times 3} \end{bmatrix}, \quad D_{11} = 0_{6\times 6}, \quad D_{12} = \begin{bmatrix} 0_{3\times 3} \\ I_3 \end{bmatrix}$$
$$C_2 = I_3, \quad D_{21} = \begin{bmatrix} 0_{3\times 3} & I_3 \end{bmatrix}, \quad D_{22} = 0_{3\times 3}.$$

We obtain the controller that minimizes the SDP relaxation of the k-sparse $\mathcal{H}_\infty$ norm using the convex optimization procedure described in Sec. IV.A. We then compute the true k-sparse $\mathcal{H}_\infty$ norm via exhaustive search – see Table I for the result of these numerical experiments.

Since our synthesis method is based on the SDP relaxation of the k-sparse $\mathcal{H}_\infty$ norm, the resulting controller may not be the true optimal controller. However, as we can see, the controllers computed with respect to relaxations of the k-sparse $\mathcal{H}_\infty$ norm exhibit better performance with respect to k disturbances than the general $\mathcal{H}_\infty$ optimal controller. In particular, if only k disturbances are allowed to coordinate their attack, then we see that if a controller is designed to mitigate the worst case effect of a larger number of disturbances, this can in fact lead to a degradation in the closed loop k-spare $\mathcal{H}_\infty$ norm of the system.

VI. Conclusion

Motivated by robustness properties of large-scale systems, we defined the k-sparse $\mathcal{H}_\infty$ norm and the k-sparse minimal gain of a system. As computing these objects involves solving a combinatorial optimization problem, we described semidefinite programming relaxations to these combinatorial optimization problems, as well as a simple rounding heuristic that provides a feasible, but possibility sub-optimal solution. We also developed a centralized controller synthesis method based on the KYP-like dual of our semidefinite relaxation, and confirmed its effectiveness through a numerical example. In future work, we aim to extend our synthesis methods to nested distributed systems [25], and to develop specialized solvers that exploit structure in the state-space parameters such that our methods can scale to systems with hundreds to thousands of states.

References

[1] R. Mason and A. Papachristodoulou, "Chordal sparsity, decomposing sdps and the Lyapunov equation," in *American Control Conference (ACC), 2014*, June 2014, pp. 531–537.

[2] C. Meissen, L. Lessard, and A. Packard, "Performance certification of interconnected systems using decomposition techniques," in *American Control Conference (ACC), 2014*, June 2014, pp. 5030–5036.

[3] 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, Feb 2006.

[4] N. Matni and V. Chandrasekaran, "Regularization for design," *arXiv preprint arXiv:1404.1972*, 2014.

[5] G. E. Dullerud and F. Paganini, *A course in robust control theory: A convex approach*. Springer New York, 2000.

[6] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9, pp. 589–592, 2008.

[7] A. d'Aspremont, L. El Ghaoui, M. I. Jordan, and G. R. Lanckriet, "A direct formulation for sparse PCA using semidefinite programming," *SIAM review*, vol. 49, no. 3, pp. 434–448, 2007.

[8] S. You and A. Gattami, "H-infinity analysis revisited," *arXiv preprint arXiv:1412.6160*, 2014.

[9] J. C. Doyle, B. A. Francis, and A. Tannenbaum, *Feedback control theory*. Macmillan Publishing Company New York, 1992.

[10] S. You and J. C. Doyle, "A Lagrangian dual approach to the Generalized KYP lemma," in *CDC*, 2013, pp. 2447–2452.

[11] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[12] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *Information Theory, IEEE Transactions on*, vol. 52, no. 6, pp. 2508–2530, 2006.

[13] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *Automatic Control, IEEE Transactions on*, vol. 48, no. 6, pp. 988–1001, 2003.

[14] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65–78, 2004.

[15] P. J. Bickel, Y. Ritov, and A. B. Tsybakov, "Simultaneous analysis of Lasso and Dantzig selector," *The Annals of Statistics*, pp. 1705–1732, 2009.

[16] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems," *Foundations of Computational mathematics*, vol. 12, no. 6, pp. 805–849, 2012.

[17] A. Gattami and B. Bamieh, "A simple approach to H_∞ analysis," in *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*. IEEE, 2013, pp. 2424–2428.

[18] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

[19] L. Mirsky, "A trace inequality of John von Neumann," *Monatshefte für Mathematik*, vol. 79, no. 4, pp. 303–306, 1975.

[20] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[21] P. Gahinet and P. Apkarian, "A linear matrix inequality approach to H_∞ control," *International journal of robust and nonlinear control*, vol. 4, no. 4, pp. 421–448, 1994.

[22] T. Athay, R. Podmore, and S. Virmani, "A practical method for the direct analysis of transient stability," *Power Apparatus and Systems, IEEE Transactions on*, no. 2, pp. 573–584, 1979.

[23] Ueda, "IEEE 10 Generator 39 Bus System," http://sys.elec.kitami-it.ac.jp/ueda/demo/WebPF/39-New-England.pdf.

[24] A. R. Bergen and V. Vittal, *Power Systems Analysis*, 2nd ed. Prentice Hall, 1999.

[25] C. W. Scherer, "Structured $\mathcal{H}_\infty$-optimal control for nested interconnections: A state-space solution," *Systems & Control Letters*, vol. 62, no. 12, pp. 1105–1113, 2013.