Decentralized State Feedback Control for Interconnected Process Systems

Simone Schuler,∗ Ulrich Münz∗∗,⋆ and Frank Allgöwer∗

∗ Institute for Systems Theory and Automatic Control, University of Stuttgart, Germany, {schuler,allgower}@ist.uni-stuttgart.de
** Siemens Corporate Technology, Munich, Germany, ulrich.muenz@siemens.com

Abstract: We consider the problem of constructing decentralized state feedback controllers for linear continuous-time systems. Different from existing approaches, where the topology of the controller is fixed a-priori, the topology of the controller is part of the optimization problem. Structure optimization is done in terms of a minimization of the required feedback and subject to a predefined bound on the tolerable loss of the achieved $\mathcal{H}_\infty$-performance of the decentralized controller compared to an $\mathcal{H}_\infty$-optimal centralized controller. We develop a computationally efficient formulation of the decentralized control problem by convex relaxations which makes it attractive for practical applications.

Keywords: decentralized control, $\ell_p$-minimization, multivariable systems, controller structure design, $\mathcal{H}_\infty$-control

1. INTRODUCTION

The analysis and control of interconnected systems is one of the big challenges of modern engineering science (see Murray, 2002). Examples of such interconnected systems include highly coupled chemical plants, heat exchangers, chemical reaction networks and distributed power generation networks (smart grids). The constituent parts of interconnected systems are the individual dynamical subsystems, the interconnection topology, and the dynamics between the subsystems and the controller architecture. In a decentralized framework, the controllers are spatially distributed, and each controller has access to a different subset of local measurements. Decentralized control is often preferred to one centralized controller, since less measurement links are necessary and less information needs to be transmitted (ˇSiljak, 1991). In general, decentralized controller design consists of two different problems: First, the structure of the decentralized controller has to be designed, and second, the controller itself has to be designed. While all this represents a general paradigm for the control of distributed systems and decentralized controller design, there has been a long history within the process control community focusing on the design of controller architecture and decentralized control. For example the first work on preferable loop pairings (Bristol, 1966) was motivated by typical problems in multivariable process control, an alternative approach was for example presented by Niederlinski (1971). Generalizations and other interaction measures can among others be found in Gagnepain and Niederlinski (1971). Interaction measures analyze the underlying structure of an interconnected system and answer therefore the question, how a good structure of the decentralized controller might look like. The design of the decentralized controller itself is independent thereof.

In decentralized controller design, different levels of decentralization can be considered. Figure 1(a) shows exemplary a network of three interconnected subsystems $\Sigma_i$, $i = 1, 2, 3$ with a centralized controller $K$ that has access to and can influence all the subsystems. In Figure 1(c), the network is alternatively controlled by three individual controllers $K_{ii}$, where each controller has only access to its own associated output. This is a completely decentralized control structure. As an example for an intermediate structure, in Figure 1(b) additional measured outputs from other subsystems are used to improve the performance of the control loop. This leads to a partially decentralized control structure. Interaction measures as described in the previous paragraph often try to identify the loop pairings, such that a completely decentralized controller as depicted in Figure 1(c) can be implemented while stability and performance requirements are achieved. However, achievable performance generally decreases for decentralized controllers when compared to centralized ones and sometimes not even the stability of the interconnected system can be guaranteed. It is the goal of this paper to design partially decentralized controllers, with only small $\mathcal{H}_\infty$-performance degradation compared to an optimal centralized controller.

With the advent of convex optimization and the efficient computation of $\mathcal{H}_\infty$- and $\mathcal{H}_2$-optimal controllers, decentralized controller design came again into focus and was considered in Rotkowitz and Lall (2006) and Shah and

© IFAC, 2012. All rights reserved.
Parrilo (2008), as well as in Qi et al. (2004) and Scherer (2002). These approaches focus on the design of the $\mathcal{H}_\infty$-or $\mathcal{H}_2$-optimal controller itself and not on the design of the controller architecture. They have the following common features: (i) the structure of the controller to be designed is fixed and restricted to special cases, and network and controller have to share a common structure; (ii) the structure of the controller has to be specified in advance, that is, they do not consider the problem of structure design but only the design of the controller itself. Considering the problem of controller architecture and controller design itself, a natural question one may raise is whether it is possible to design both the controller structure and the controller itself simultaneously. This is especially of interest in the fast growing field of networked control systems with highly interacting subsystems, where it is often not clear, how the controller topology should look like to achieve good performance.

The above discussions motivate the main contribution of the current paper. We search for a tradeoff between the number of measurement links, i.e. the degree of decentralization and the achievable $\mathcal{H}_\infty$-performance of the system. We combine control theoretic insight with results from compressed sensing (see e.g. Candès et al., 2006a; Donoho, 2006) to systematically achieve decentralization with guaranteed system theoretic properties by computationally efficient algorithms. We define a pattern operator to represent the structure of the controller. In this sense, the controller topology is not specified in advance, but considered as an optimization variable. We design state feedback controllers for linear continuous-time interconnected systems, where the number of measurement links is minimized subject to an $\mathcal{H}_\infty$-performance constraint. By means of a system augmentation approach, we first present a novel characterization of the $\mathcal{H}_\infty$-performance of the closed-loop system. Then, the non-convex structure optimization of the pattern operator is relaxed by a convex weighted $\ell_1$-minimization, and the resulting problem can be tackled by finding a solution of iterative convex optimization problems. Moreover, an algorithm is proposed to optimize the initial values such that the solvability of the original problem can be improved. Similar results for discrete-time systems can be found in Schuler et al. (2010, 2011). Simultaneous design of controller structure and controller itself for interacting single integrator systems was also independently reported in Fardad et al. (2011).

The remainder of the article is organized as follows. After introducing the mathematical preliminaries in Section 2, the formulation of the decentralized control problem is presented in Section 3. Section 4 is devoted to the design of a decentralized controller for networks of interconnected subsystems. The paper concludes with an illustrative example in Section 5 and a summary in Section 6.

2. MATHEMATICAL PRELIMINARIES

The $p$-norm of a vector $x \in \mathbb{R}^n$ for $1 \leq p < \infty$ is defined as

$$
\|x\|_p := \left( \sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}.
$$

For $p \in \mathbb{R}^+$, the function $|x|_0 := \lim_{p \to 0} \|x\|_p$, is often referred as the 0-norm for convenience, despite not being a true norm. $|x|_0$ corresponds exactly to the number of non-zero entries in $x$. Alternatively, it is also common to define

$$
|\|x\|_0 := \sum_{i=1}^{n} |\text{sign}(x_i)|
$$

with

$$
\text{sign}(x_i) = \begin{cases} -1 & \text{for } x_i < 0 \\ 0 & \text{for } x_i = 0 \\ 1 & \text{for } x_i > 0. \end{cases}
$$

Both definitions are equivalent. A vector is called sparse if its 0-norm is small compared to the dimension of the vector, i.e., if most of its entries are zero. The 0-norm is used in the present context to achieve sparse controller structures. A brief introduction into sparsity measures in general and sparsity promoting optimal control is given in the Appendix.

Performance is specified in this paper by the $\mathcal{H}_\infty$-norm. The $\mathcal{L}_2$-induced norm (or $\mathcal{L}_2$-gain) of a dynamical system $\mathcal{H} : \mathcal{L}_2^m \to \mathcal{L}_2^n$ is defined as

$$
\|\mathcal{H}\|_{\mathcal{L}_2} := \sup_{u \in \mathcal{L}_2^m/\{0\}} \frac{\|Hw\|_{\mathcal{L}_2^n}}{\|w\|_{\mathcal{L}_2^m}},
$$

and corresponds for a linear system $\mathcal{H}$ to the $\mathcal{H}_\infty$-norm $\|H(s)\|_\infty = \sup_{\omega \in \mathbb{R}} \sigma(H(j\omega))$, where $H(s) = C(sI - A)^{-1}B + D$ is a transfer function of the dynamical system $\mathcal{H}$ and $\sigma(H)$ denotes the largest singular value of $H$.

Notational specifications as used in the paper are given next. Given a matrix $M = [m_1 \ldots m_n]$ with $m_i \in \mathbb{R}^n$ being its ith column, we define a vector vec($M$) $\in \mathbb{R}^{nm \times 1}$ by

$$
\text{vec}(M) = [m_1^T \ldots m_n^T]^T.
$$

A symmetric and positive definite (resp. positive semi-definite) matrix $M$ is written as $M > 0$ (resp. $M \geq 0$), $M^T$ and $M^{-1}$ denote the transpose and inverse of a matrix.
A matrix \( M \) with elements \( m_{ij} \) is also written as \( M = [m_{ij}] \). For a given real matrix \( M \in \mathbb{R}^{n \times m} \) with \( n < m \), the orthogonal complement \( M^\perp \) is defined as the matrix with \( n \times (m-n) \) columns that satisfies \( MM^\perp = 0 \). A (block-) diagonal matrix with elements \( m_1, \ldots, m_n \) on the diagonal is abbreviated as \( \text{diag}(m_1, \ldots, m_n) \) or simply \( \text{diag}(m_n) \). If \( A \in \mathbb{R}^{m \times n} \) and \( B \in \mathbb{R}^{m \times n} \), \( A \otimes B \) denotes the Hadamard product of \( A \) and \( B \).

3. FORMULATION OF DECENTRALIZED CONTROL PROBLEM

Starting with classical centralized control, the problem of finding a decentralized controller is formulated as a minimization problem over the number of measurement links between subsystems and controllers. This is done subject to an upper bound on the \( H_\infty \)-performance degradation between the decentralized control loop and the centralized one. This leads to a block structure in the decentralized controller, where we achieve decentralization by setting as many blocks as possible to zero. This approach is formalized in the following.

3.1 Interconnected Systems

We consider a network consisting of \( N \) interconnected subsystems, where the \( i \)th subsystem is given by

\[
\begin{align}
\dot{x}_i &= A^{ij}x_i + \sum_{j=1}^{N} A^{ij} x_j + \sum_{j=1}^{N} B^{ij}_1 w_j + B^{ij}_2 u_i \quad (1a) \\
\dot{z}_i &= C^{ij} x_i + \sum_{j=1}^{N} D^{ij}_1 w_j + D^{ij}_2 u_i \quad (1b)
\end{align}
\]

where \( x_i \in \mathbb{R}^n \) is the state of subsystem \( i \) and \( A^{ij} \neq 0 \) if and only if subsystem \( j \) influences subsystem \( i \) in a direct way. The exogenous input \( w_i \in \mathbb{R}^q \) represents disturbance signals, \( u_i \in \mathbb{R}^q \) is the control input to each subsystem and \( z_i \in \mathbb{R}^p \) is the controlled output. Here, for simplicity of notation and clarity of presentation, it is assumed that all subsystems have the same number of states and (exogenous) inputs. This assumption can be dropped and these numbers can be different for different subsystems without changing the presented results.

Let \( x = [x_1^T, \ldots, x_N^T]^T \), then the entire model of the network composed of \( N \) subsystems in the form of (1) can be written as

\[
\Sigma_p : \begin{bmatrix} \dot{x} \\ z \end{bmatrix} = \begin{bmatrix} A & B_1 & B_2 \\ C & D_1 & D_2 \end{bmatrix} \begin{bmatrix} x \\ w \\ u \end{bmatrix} \quad (2)
\]

with \( A = [A^{ij}] \), \( B_1 = [B^{ij}_1] \), \( D_1 = [D^{ij}_1] \) and \( B_2 = \text{diag}(B_2^{ij}) \), \( C = \text{diag}(C^{ij}) \), \( D_2 = \text{diag}(D_2^{ij}) \), \( w = [w_1^T, \ldots, w_N^T]^T \), \( u = [u_1^T, \ldots, u_N^T]^T \), \( z = [z_1^T, \ldots, z_N^T]^T \), \( i = 1, \ldots, N \).

Assumptions on the entire model are stabilizability of \((A,B_2)\) and detectability of \((A,C_2)\).

3.2 Centralized Controller

In a centralized framework the controller has access to all states. A state feedback controller is then given by

\[
u = K x, \quad K \in \mathbb{R}^{q_2 N \times N}, \quad \text{and} \quad K = [K^{ij}] \quad (3)
\]

In general, the matrix \( K \) has no structure and the controller uses all possible measurement links between sensors and controller. The closed loop is given by

\[
\Sigma_{cf}(K) : \begin{bmatrix} \dot{x} \\ z \end{bmatrix} = \begin{bmatrix} A + B_2 K & B_1 \\ C + D_2 K & D_1 \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} \quad (4)
\]

We search for a controller that minimizes the influence of the exogenous input \( w \) on the performance output \( z \) in terms of the \( H_\infty \)-norm of the closed loop. This centralized controller uses all possible degrees of freedom and can be designed via convex optimization (Gahinet and Apkarian, 1994) using standard LMI solvers (Sturm, 1999; Löfberg, 2004). We have the following assumptions on \( K \).

Assumption 1. The centralized controller \( K \) achieves

(i) stability of the closed loop and
(ii) optimality in terms of the \( H_\infty \)-norm of the closed loop.

3.3 Decentralized Controller

Next, we describe a decentralized controller such that less measurement links between sensors and controllers are necessary. We especially consider that each controller does not have access to all subsystems but only knows the states of a few subsystems, i.e. we want to remove measurement links between subsystems and controllers. Therefore we search for controllers

\[
u_i = \sum_{j=1}^{N} \hat{K}^{ij} x_j, \quad \hat{K}^{ij} \in \mathbb{R}^{q_2 \times n}
\]

where \( \hat{K}^{ij} = 0 \) for as many pairs \((i,j)\) as possible. If \( \hat{K}^{ij} = 0 \) no link from subsystem \( j \) to controller \( i \) is necessary. Combining all individual controllers to one state feedback controller similar to (3), leads to a decentralized controller

\[
u = \hat{K} x, \quad \hat{K} \in \mathbb{R}^{q_2 N \times N}, \quad \text{and} \quad \hat{K} = [\hat{K}^{ij}] \quad (5)
\]

Similar to (4) the closed loop is given by

\[
\hat{\Sigma}_{cf}(\hat{K}) : \begin{bmatrix} \dot{x} \\ z \end{bmatrix} = \begin{bmatrix} A + B_2 \hat{K} & B_1 \\ C + D_2 \hat{K} & D_1 \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} \quad (6)
\]

Note that \( \hat{K} \) has a decentralized structure if several submatrices \( \hat{K}^{ij} \) are zero. In fact, decentralization of the controller is often modeled with a sparsity structure on the centralized controller, e.g., a completely decentralized control structure (see Figure 1(c)) would be represented by a block diagonal structure in the centralized controller \( K \), that is \( K_{ij} = 0 \) for \( i \neq j \). Since \( \hat{K}^{ij} \in \mathbb{R}^{q_2 \times n} \), decentralization corresponds to zero matrices in the \((i,j)\) block of the decentralized controller. We want to find decentralized controllers with as many blocks of zeros as possible.

To represent the structure of the decentralized controller, we introduce the pattern operator \( \mathcal{Z}(\hat{K}, p) \in \mathbb{R}^{m \times n} \), such that each element \( z^{ij} \) of \( \mathcal{Z}(\hat{K}, p) \) represents one block \( \hat{K}^{ij} \) of the decentralized controller \( \hat{K} \) and \( z^{ij} = 0 \) if and only if
\( K^{ij} = 0 \). Different realizations of the pattern operator are possible, e.g.

\[
Z(\hat{K}, p) := \begin{bmatrix}
\| \hat{K}_{11} \|^p & \cdots & \| \hat{K}_{1N} \|^p \\
\vdots & \ddots & \vdots \\
\| \hat{K}_{N1} \|^p & \cdots & \| \hat{K}_{NN} \|^p
\end{bmatrix},
\]

with \( p \in \{0, 1, 2, \infty\} \) for example. Now, each subcontroller \( \hat{K}^{ij} \) is represented by a single element in the pattern operator and it holds that if \( z^{ij} = 0 \) no link from subsystem \( j \) to controller \( i \) is necessary. This pattern operator will enable us later to optimize the structure of the decentralized controller, i.e. maximize the number of zeros in this operator.

### 3.4 Error System

As said before, the centralized controller can be computed via convex optimization and the global optimum can be found. Any decentralized controller will have a worse performance. In the present context, we allow a small performance degradation in terms of the \( H_\infty \)-performance of the closed loop for the decentralized controller and want to use as few measurement links as possible to achieve this performance. Performance degradation due to the decentralization of the controller can now be investigated by the analysis of the error system \( \Sigma_e = \Sigma_{cl} - \hat{\Sigma}_{cl} \)

\[
\dot{\Sigma}_e(\hat{K}, \hat{K}) : \begin{bmatrix} \dot{x}_e \\ \dot{e} \end{bmatrix} = \begin{bmatrix} A_e & B_e \\ C_e & 0 \end{bmatrix} \begin{bmatrix} x_e \\ w \end{bmatrix},
\]

(7)

where

\[
\begin{bmatrix} A_e & B_e \\ C_e & D_e \end{bmatrix} = \begin{bmatrix} A + B_2 \hat{K} & 0 \\ 0 & A + B_2 \hat{K} \\ C + D_2 \hat{K} & -(C + D_2 \hat{K}) \end{bmatrix} 0
\]

with \( x_e = [x \ \dot{x}]^T \) and \( e = z - \dot{z} \). Note that the closed loop can also be written in the following way:

\[
A_e = A_0 + F \hat{K} R, \quad C_e = C_0 + G \hat{K} R.
\]

(8)

with

\[
A_0 = \begin{bmatrix} A + B_2 \hat{K} & 0 \\ 0 & A \\ C + D_2 \hat{K} & -C \end{bmatrix}, \quad F = \begin{bmatrix} 0 \\ B_2 \\ 0 \end{bmatrix}, \quad R = [0 \ I]
\]

\[
C_0 = [C + D_2 \hat{K} & -C], \quad G = -D_2.
\]

This representation of the error system will simplify the reformulation of the \( H_\infty \)-performance constraint in the next section. Since the state variables of the error system \( \Sigma_e \) in (7) are composed of \( x \) and \( \dot{x} \), and the system \( \Sigma_{cl} \) in (4) is asymptotically stable by Assumption 1, the asymptotic stability of \( \hat{\Sigma}_{cl} \) is then equivalent to the asymptotic stability of \( \Sigma_e \). Thus, the requirement of the closed-loop system with the decentralized controller to be stable is achieved by guaranteeing the stability of the error system.

#### 3.5 Decentralized Control Problem

In the following, we specify an acceptable performance degradation and minimize the 0-norm of the pattern operator \( Z(\hat{K}, p) \) in order to derive the optimally structured corresponding feedback. This can be formulated in the following decentralized control problem.

**Problem 2. (Decentralized Control Problem)**

Given the linear system in (2) and the centralized controller in (3), determine a decentralized controller (5), such that

(i) the closed-loop system \( \hat{\Sigma}_{cl} \) in (6) is asymptotically stable and

(ii) the 0-norm of the pattern operator \( Z(\hat{K}, p) \) is minimized subject to a given maximal \( H_\infty \)-performance degradation \( \gamma \), that is

\[
\min_{\hat{K}} \|\text{vec}(Z(\hat{K}, p))\|_0
\]

\[
\text{subject to } \|\Sigma_{cl}(\hat{K}) - \hat{\Sigma}_{cl}(\hat{K})\|_\infty < \gamma
\]

and \( \gamma > 0 \).

The problem formulation implies that the performance degradation of the closed loop controlled with the decentralized controller does not exceed \( \gamma \). Recall that the 0-norm of a vector is a measure of its sparsity. In this way, minimizing \( \|Z(\hat{K}, p)\|_0 \) attempts to maximize the number of zero-elements of the pattern operator and therefore minimizes the number of measurement links. In the presented problem set-up controller design and topology design are combined into one optimization problem and solved simultaneously.

However, the minimization of the 0-norm is a non-convex problem and requires a combinatorial search (see Candès et al., 2006a). Furthermore, the \( H_\infty \)-performance constraint is also non-convex and the well known convex reformulations cannot be applied to the presented set-up as will be seen in the next section. Therefore, Problem 2 is non-convex both in the objective function and in the constraint. In the following section, we show how this problem can be approximated by a numerically tractable convex optimization problem.

### 4. Convex Relaxation of the Decentralized Control Problem

This section focuses on the design of decentralized \( H_\infty \)-controllers as defined in Problem 2. We first introduce a system augmentation approach to reformulate the characterization of the \( H_\infty \)-performance of the closed loop system. This enables us to incorporate the decentralization constraint into the controller design. Then, we use a weighted \( \ell_1 \)-minimization to relax the numerically exhaustive combinatorial exact solution of the \( \ell_0 \) objective function and show how a decentralized controller can be found by the iterative solution of convex optimization problems.

#### 4.1 Reformulation of the \( H_\infty \)-performance Constraint

For the characterization of the \( H_\infty \)-norm, the classical Bounded Real Lemma formulation is used:

**Lemma 3.** (Gahinet and Apkarian, 1994) Consider a continuous-time transfer function \( H(s) \) with realization \( H(s) = D_e + C_e(sI - A_e)B_e \) . The following statements are equivalent

(i) \( \|D_e + C_e(sI - A_e)B_e\|_\infty < \gamma \) and \( A_e \) is asymptotically stable;

(ii) there exists a symmetric positive definite solution \( P_e > 0 \) to the LMI
\[
\begin{bmatrix}
A^T P + P A_e & PB_e & C^T \\
B^T P & -\gamma I & 0 \\
C_e & D_e & -\gamma I
\end{bmatrix} \begin{bmatrix}
P \\
\end{bmatrix} + \begin{bmatrix}
P A_e + A^T P & P B_e & C^T \\
B^T P & -\gamma I & 0 \\
C_e & D_e & -\gamma I
\end{bmatrix} \begin{bmatrix}
P \end{bmatrix} < 0. \quad (10)
\]

Note that if matrix inequality (10) is used for controller synthesis, it is a bilinear matrix inequality due to the multiplication between the Lyapunov matrix \(P\) and the controller matrix \(\hat{K}\) embedded in the closed loop matrices \(A_e\) and \(C_e\) and therefore it is non-convex. In existing convex approaches to \(\mathcal{H}_\infty\)-state feedback, the controller matrix does not appear explicitly in the synthesis conditions. This does not facilitate optimization with structural constraints as imposed on the controller in this article, where we are additionally interested in the minimization of \(\|Z(\hat{K}, p)\|_0\).

For these reasons, the approaches in Gahinet and Apkarian (1994), Sanei et al. (1990) or Iwasaki and Skelton (1994) are not directly applicable to the controller design problem discussed here. To overcome these difficulties, we introduce a system augmentation approach as initially developed in Shu and Lam (2009) for static output feedback and further adopted to positive filtering in Li et al. (2011) and decentralized control in Schuler et al. (2011). Such a formulation enables us to deal with the \(\mathcal{H}_\infty\) decentralized control problem considered in this article. Instead of the system augmentation approach, a simple \((P, \hat{K})\)-iteration would also be possible, but classical \((P, \hat{K})\)-iteration has in general not very good convergence properties.

Define an auxiliary variable
\[
v = \hat{K} R x_e
\]
and consider \(x^*_e = [x^*_e \ v^*_e]^T\) as an augmented variable, then we can write system (7) as a singular system
\[
\Sigma_s : \begin{bmatrix}
E_s \ x_s \ e \\
\end{bmatrix} = \begin{bmatrix}
A_s & B_s \\
C_s & 0
\end{bmatrix} \begin{bmatrix}
x_s \\
w
\end{bmatrix},
\]
with
\[
E_s = \begin{bmatrix}
I & 0 \\
0 & 0
\end{bmatrix}, \quad A_s = \begin{bmatrix}
A_0 & F \\
\hat{K} R & -I
\end{bmatrix}, \quad B_s = \begin{bmatrix}
B_e \\
0
\end{bmatrix},
\]
\[
C_s = \begin{bmatrix}
C_0 & G
\end{bmatrix}.
\]

**Theorem 4.** Given the decentralized controller \(\hat{K}\), the following statements are equivalent

(i) the error system in (7) is asymptotically stable and satisfies \(\|\Sigma_s\|_\infty < \gamma\),

(ii) there exist a symmetric matrix \(P_1 > 0\) and diagonal matrix \(P_2 > 0\) such that
\[
\Omega := \begin{bmatrix}
P^T A_e + A^T P_1 & P_1 B_e & C^T \\
B^T P_1 & -\gamma I & 0 \\
C_e & 0 & -\gamma I
\end{bmatrix} < 0, \quad (13)
\]

with
\[
P_s = \begin{bmatrix}
P_1 & 0 \\
-\frac{1}{\alpha} P_1 \hat{K} R & \frac{1}{\alpha}
\end{bmatrix}.
\]

**Proof.** (ii) \(\Rightarrow\) (i). Suppose there exist \(P_1 > 0\) and diagonal \(P_2 > 0\) such that (13) holds. Define a non-singular matrix
\[
T = \begin{bmatrix}
I & 0 & 0 \\
\hat{K} R & 0 & 0 \\
0 & I & 0 \\
0 & 0 & I
\end{bmatrix},
\]
Pre- and post-multiplying (13) from left and right by \(T^T\) and \(T\), respectively, we have
\[
\Omega := T^T \Omega T = \begin{bmatrix}
P_1 A_e + A^T P_1 & P_1 B_e & C^T \\
B^T P_1 & -\gamma I & 0 \\
C_e & 0 & -\gamma I
\end{bmatrix} < 0.
\]

The third leading principal submatrix of \(\Omega\) is identical to (10). Based on Lemma 3 this indicates that the error system in (7) is asymptotically stable, and satisfies \(\|\Sigma_s\|_\infty < \gamma\).

(i) \(\Rightarrow\) (ii). If the error system in (7) is asymptotically stable, and satisfies \(\|\Sigma_s\|_\infty < \gamma\), then it follows from Lemma 3 that there exists a matrix \(P_1\), such that (10) holds. Given any diagonal matrix \(X > 0\), there must exist a scalar \(\alpha > 0\) satisfying
\[
-\alpha X = \begin{bmatrix}
P_1 F & 0 \\
0 & 0
\end{bmatrix} \begin{bmatrix}
A^T P_1 + P_1 A_e & P_1 B_e & C^T \\
B^T P_1 & -\gamma I & 0 \\
C_e & 0 & -\gamma I
\end{bmatrix}^{-1} \begin{bmatrix}
P_1 F & 0 \\
0 & 0
\end{bmatrix} < 0.
\]

By choosing \(P_2 = \alpha X\) and applying Schur complement equivalence to (14), we have \(\Omega = T^{-T} \Omega T^{-1} < 0\), which completes the proof. \(\square\)

**Remark 5.** The characterization of the \(\mathcal{H}_\infty\)-norm in Theorem 4 is necessary and sufficient and therefore equivalent to the formulation of Lemma 3. The difference is the separation of the controller matrix \(\hat{K}\) and the Lyapunov matrix \(P_1\). Additionally, the controller matrix can be further parameterized by the diagonal matrix \(P_2\). For the necessity part of the proof, the diagonal structure of \(P_2\) is not needed, but it will enable us later on to design the decentralized controller \(\hat{K}\).

**Theorem 6.** There exists a decentralized controller \(\hat{K}\) to solve Problem 2, if and only if there exist matrices \(P_1 > 0\), diagonal \(P_2 > 0\), and \(L\) and \(U\), such that
\[
\Pi(U) := \begin{bmatrix}
F^T P_1 + L R & -P_2 & 0 & G^T \\
B^T P_1 & 0 & -\gamma I & 0 \\
C_0 & G & 0 & -\gamma I
\end{bmatrix} < 0,
\]
with
\[
\Delta = \begin{bmatrix}
P_1 F & R^T L^T & P_1 B_e & C^T \\
B^T P_1 & -\gamma I & 0 \\
C_e & 0 & -\gamma I
\end{bmatrix}^{-1} \begin{bmatrix}
P_1 F & 0 \\
0 & 0
\end{bmatrix} < 0.
\]

In this case, the decentralized controller is given by
\[
\hat{K} = P_2^{-1} L
\]

**Proof.** **Sufficiency:** From (16), we have \(L = P_2 \hat{K}\). Substituting this into (15), it is identical to (13) except for the first leading principle. For the first leading principle, we observe that for any \(U\) it holds that
\[
-\hat{K} R \hat{R}^T P_2 \hat{K} R \leq -R^T \hat{K}^T P_2 \hat{K} R
\]
\[
+ (U - \hat{K} R)^T P_2 (U - \hat{K} R)
\]
\[
=! U^T P_2 U - U^T P_2 \hat{K} R - R^T \hat{K}^T P_2 U. \quad (17)
\]

According to Theorem 4 this completes the sufficiency of the proof.

**Necessity:** Assume that the error system (7) is asymptotically stable with \(\|\Sigma_s\|_\infty < \gamma\). Then, according to Theorem 4, there exist a matrix \(P_1 > 0\) and a diagonal
matrix $P_2 > 0$ such that (13) holds. By selecting $U = \hat{K}R$ we have
\[
-\hat{R}^T \hat{K}^T P_2 \hat{K}R = -\hat{R}^T \hat{K}^T P_2 (U - \hat{K}R) = U^T P_2 U - U^T P_2 \hat{K}R - \hat{R}^T \hat{K}^T P_2 U.
\] (18)
Substitute (18) into (13), and let $L = P_2 \hat{K}$, it is equivalent to (15). This completes the proof. □

**Remark 7.** Theorem 6 presents a necessary and sufficient condition for the existence of a decentralized controller. Although inequality (15) in Theorem 6 is still non-linear due to the multiplication of $U$ and $L$, one can see that the controller matrix $\hat{K}$ is only coupled with the diagonal matrix $P_2$, which is independent of the Lyapunov matrix $P_1$. Due to the fact that $P_2$ is diagonal, the sparsity structure of $\hat{K}$ is preserved by specifying the sparsity of $L$, hence no conservatism is added.

Even though, Theorem 6 is still non-linear, the advantage lies in the fact that the structural constraints added to $\hat{K}$ can now be imposed on $L$ explicitly and it holds $Z(\hat{K}, p) = Z(L, p)$. However, it is well known that the 0-norm is non-convex, which further makes Problem 2 difficult to solve. Stimulated by the field of compressed sensing (see Candès et al., 2006b; Donoho, 2006; Baraniuk, 2007), we use a convex relaxation of the 0-norm and further develop an iterative algorithm to solve the decentralized control problem as shown next.

### 4.2 Convex Relaxation of the $\ell_0$-Constraint

Minimizing $\|Z(L, p)\|_0$ attempts to maximize the zero-elements in the pattern operator. However, as stated before this approach is of little practical use, since the optimization problem is non-convex and NP-hard as its solution requires a combinatorial search which grows faster than polynomial as $n$ grows (see Candès et al., 2006a).

Similar to the convex relaxation for rank minimization in Fazel et al. (2001), we will use the convex envelope of the function $\|Z(L, p)\|_0$ defined next.

Let the map $f$ be defined as $f : \mathbb{X} \to \mathbb{R}$, where $\mathbb{X} \subseteq \mathbb{R}^n$. The convex envelope of $f$ (on $\mathbb{X}$), denoted $f_{\text{env}}$, is defined as the point-wise largest convex function $g$ such that $g(x) \leq f(x)$ for all $x \in \mathbb{X}$.

**Lemma 8.** (Fazel (2002)). The convex envelope of the function $f = \|x\|_0 = \sum_{i=1}^{n} |\text{sign}(x_i)|$ on $\mathbb{X} = \{x \in \mathbb{R}^n | \|x\|_1 \leq 1\}$ is $f_{\text{env}}(x) = \|x\|_1 = \sum_{i=1}^{n} |x_i|$. Note that the minimization of $\|x\|_1$ can be solved using linear programming. Additionally, this is the best possible convex relaxation since the 1-norm is the convex envelope of the 0-norm.

As described in Candès et al. (2008), reweighted $\ell_1$-minimization can be used to improve the results of the minimization. In this direction, weights $m_{ij}^{\gamma} \geq 0$ can be assigned to each variable $L_{ij}$ as

\[
\min \|M \circ Z(L, p)\|_1, \quad \text{subject to } \Gamma, \quad \text{subject to } \Pi(U_k) < 0.
\] (19)

where $M = [m_{ij}^{\gamma}]$ are non-negative weights. For the described design problem, the weights are free parameters. They counteract the influence of the signal magnitude on the $\ell_1$-penalty function. If $m_{ij}^{\gamma} = 1$ for all $(i, j)$, the weighted 1-norm reduces to the regular 1-norm. If the weights $m_{ij}^{\gamma}$ are chosen to be inversely proportional to the magnitude of $z_{ij}$

\[
\left\{ \begin{array}{l}
m_{ij}^{\gamma} = 1/|z_{ij}^{\gamma}|, \quad z_{ij}^{\gamma} \neq 0, \\
m_{ij}^{\gamma} = \infty, \quad z_{ij}^{\gamma} = 0,
\end{array} \right.
\] (20)

then the weighted 1-norm and the 0-norm coincide. The above weighting scheme cannot be implemented, since the weights depend on the solution of the optimization problem. In Candès et al. (2008) an iterative scheme is proposed, where the weights are the solution of the previous iteration. In the context of Problem 2, these weights can also be used to include system and control theoretic insight into the optimization problem to improve the result. For example, results from interaction measures (see e.g. Bristol, 1966; Grosdidier and Morari, 1986) can be used to choose appropriate initial weights. If other knowledge about the system is available, e.g. some measurement links are very unattractive since they are related to high implementation costs or just impossible to implement, those measurement links can also be penalized by a large initial weight in the reweighted $\ell_1$-minimization. Furthermore, the diagonal weights $m_{ii}$ can always be set to zero, since local measurement does not increase the cost.

### 4.3 Relaxed Decentralized Controller Design

Based on the discussion of the previous two sections, we can relax Problem 2 with a convex objective function and the new characterization of the $\mathcal{H}_\infty$-performance.

**Problem 9.** (Relaxed Decentralized Control Problem). Given the linear system (2), find $P_1 > 0$, diagonal $P_2 > 0$, $L$ and $U$ satisfying

\[
\inf \|M \circ Z(L, p)\|_1, \quad \text{subject to (15)}.
\]

As already stated in Remark 7, the matrix inequality (15) is generally not linear with respect to the parameters $P_1$, $P_2$, $L$ and $U$. However, if we fix the matrix $U$, then it becomes an LMI in the other parameters. Thus, the remaining question is how to choose $U$ properly. Based on the above discussion, we propose the following algorithm to solve Problem 9.

**Algorithm 1** Relaxed decentralized controller design algorithm

1. Set $k = 1$. For fixed $\gamma$, find $U_1$ such that the following system

\[
\Sigma_{\text{inst}} : \begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} A_0 + F U_1 & B_1 \\ C_0 + G U_1 & 0 \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix}.
\] (21)

achieves an $\mathcal{H}_\infty$-performance $\|\Sigma_{\text{inst}}\|_\infty < \gamma$. Choose a sufficiently small $\gamma$. If no $U$ can be found, exit. No solution can be found.

2. For fixed $U_k$, solve the following convex optimization problem for the parameters $\Gamma = \{P_1 > 0$, diagonal $P_2 > 0, L\}$:

\[
\inf_{\Gamma} \|M_k \circ Z(L, p)\|_1 \quad \text{subject to } \Pi(U_k) < 0.
\] (22)

Denote $P_{2k}$, and $L_k$ as the solution of $P_{2k}$, and $L_k$.

3. Terminate on convergence. Otherwise update $U_{k+1}$ as $U_{k+1} = (P_{2k})^{-1}L_k$ and $m_{ij}^{k+1} = (|z_{ij}^{k}| + \varepsilon)^{-1}$. Set $k = k + 1$, then go to Step 2.
Remark 10. The small positive number ε is introduced to ensure that all weights are well-defined when \( z_{ij}^k = 0 \).

Remark 11. While the \( \ell_1 \)-relaxation leads to a semidefinite program that can be solved in polynomial time, there are still limitations to the solvable problem sizes. The symmetric matrix \( P_1 \) has \( nN(nN + 1) \) decision variables, the diagonal matrix \( P_2 \) has \( q_2 N \) decision variables and the matrix \( L \) has \( q_2 n N^2 \) decision variables. In practice, the algorithm performs well for problems with up to 1000 variables. Larger problems would require special solvers that take advantage of the structure of the individual problem.

Note that for Algorithm 1 convergence to a global optimum cannot be guaranteed. However, since the error systems consider the difference between the optimal centralized controller and the decentralized controller, the optimization starts already close to the optimum, which makes convergence more likely. Furthermore, it is well-known that the performance of an iterative algorithm usually depends on the choice of initial values, because poor selection of a starting point often results in the iteration being trapped and no solution can be found. A good selection may lead to a feasible solution.

4.4 Initial Value Optimization

Althought \( U_1 \) can be directly computed as the solution of

\[
\begin{bmatrix}
\bar{P}A_0^T + A_0P + WT^TW + B_0PC_0^T + WT^TG & B_0 \gamma I & 0 \\
0 & -\gamma I & 0 \\
0 & 0 & -\gamma I
\end{bmatrix} < 0
\]

with \( U_1 = W_1 \bar{P}^{-1} \). We want to show how this solution can be improved for a better feasibility of Algorithm 1.

It follows from (17) that, if for some \( \hat{K}^* \), \( P_1^* > 0 \), diagonal \( P_2^* > 0 \) inequality (13) holds, then (15) will also be feasible, that is \( \Pi(U) < 0 \), provided that \( (U - \hat{K}^*R)^T P_2^*(U - \hat{K}^*R) \) is ‘small’ enough. Thus, the solvability of the optimization problem can be further improved by choosing the initial value \( U_1 \) such that \( \| U_1 - \hat{K}^*R \|^2 \) is small enough. As \( P_2^* \) should be relatively large (because of (14)), \( \| U_1 - \hat{K}^*R \| \) should be sufficiently small. The following proposition provides an equivalent characterization of how \( \| U_1 - \hat{K}^*R \| \) can be made as small as possible.

Theorem 12. Given \( U_1 \) and a sufficiently small scalar \( \epsilon \), the following statements are equivalent:

(i) There exists \( \hat{K}^* \) such that the error system (7) is asymptotically stable with closed loop performance \( \| \Sigma_e \|_\infty < \gamma \), and \( \| U_1 - \hat{K}^*R \| \leq v_1 \epsilon \), where \( v_1 \) is a scalar.

(ii) The system (21) is asymptotically stable with \( \| \Sigma_{\text{intr}} \|_\infty < \gamma \) and \( \| U_1 R^2 \| \leq v_2 \epsilon \), where \( v_2 \) is a scalar.

Proof. (i) \( \Rightarrow \) (ii): We have

\[
\| U_1 R^2 \| = \| U_1 R^1 - KRR^-1 \| \\
\leq \| U_1 - \hat{K} R \| R^{-1} \| R \| \leq v_1 \epsilon \| R \| : = v_2 \epsilon
\]

In the following, we will prove that system (21) is asymptotically stable with \( \| \Sigma_{\text{intr}} \|_\infty < \gamma \). It follows from Lemma 3 that, if system (7) is asymptotically stable with \( \| \Sigma_e \|_\infty < \gamma \), there exists \( R > 0 \), such that

\[
(\hat{A} + \hat{B}R\hat{C})^T \hat{P} + \hat{P}(\hat{A} + \hat{B}R\hat{C}) < 0.
\]

Furthermore, it holds

\[
(\hat{A} + BU\hat{C})^T \hat{P} + \hat{P}(\hat{A} + BU\hat{C}) = \\
(\hat{A} + \hat{B}R\hat{C})^T \hat{P} + \hat{P}(\hat{A} + \hat{B}R\hat{C}) + \\
(\hat{B}(U - \hat{K}R)\hat{C})^T \hat{P} + \hat{P}(\hat{B}(U - \hat{K}R)).
\]

This, if \( (U - \hat{K}R) \) is sufficiently small, based on (25) and (26), we have that

\[
(\hat{A} + BU\hat{C})^T \hat{P} + \hat{P}(\hat{A} + BU\hat{C}) < 0.
\]

Hence, (21) is asymptotically stable with \( \| \Sigma_{\text{intr}} \| < \gamma \), which completes the first part of the proof.

(ii) \( \Rightarrow \) (i): We have that \( RR^T = I \) and \( R^2 = [I \ 0]^T \). If we select \( \hat{K} = UR^{-1} \), then

\[
(U - \hat{K}R)[R^T \ R^2] = [0 \ UR^2].
\]

Note, that \( [R^T \ R^2] \) is a permutation matrix and consider that

\[
U - \hat{K}R = [0 \ UR^2] [R^T \ R^2]^{-1}.
\]

Then, the following holds

\[
\| U - \hat{K}R \| \leq \| 0 \ UR^2 \| \| [R^T \ R^2]^{-1} \| \leq \epsilon
\]

In addition, if system (21) is asymptotically stable with \( H_\infty\)-performance less than \( \gamma \), then

\[
(\hat{A} + BU\hat{C})^T \hat{P} + \hat{P}(\hat{A} + BU\hat{C}) < 0.
\]

With a similar proof line as in (i) \( \Rightarrow \) (ii) we have that (25) holds, which further indicates that \( \hat{K} \) exists, such that \( \| \Sigma_e \|_\infty < \gamma \). This together with (27) gives that (i) holds, which completes the whole proof.

Based on Theorem 12, we propose the following algorithm to optimize the initial value of \( U \).

Algorithm 2 Initial Value Optimization

1. Set \( j = 1 \). Solve the LMI (24) to obtain \( U_1 \).
2. For fixed \( U_j \), find \( P_j \) such that

\[
(\hat{A} + BU_j\hat{C})^T P_j + P_j(\hat{A} + BU_j\hat{C}) < 0.
\]

3. For fixed \( P_j \), minimize \( \epsilon_j \) subject to

\[
(\hat{A} + BU_j\hat{C})^T P_j + P_j(\hat{A} + BU_j\hat{C}) < 0
\]

\[
\begin{bmatrix}
-\epsilon_j & (U_j R^2)^T \\
U_j R^2 & -I
\end{bmatrix} < 0
\]

Denote \( \epsilon_j^* \) and \( U_j^* \) as the optimized value of \( \epsilon_j \) and \( U_j \).

4. If \( |\epsilon_j^* - \epsilon_{j-1}^*| < \delta_2 \), where \( \delta_2 \) is a prescribed tolerance, STOP. An initial choice of \( U_1 \) is obtained. If not, set \( U_{j+1} = U_j^* \) and \( j = j + 1 \) and go to Step 2.

Remark 13. Note that the stopping criteria suggested in Step 4 is heuristic, and convergence of the initial optimization algorithm is naturally guaranteed, since \( \epsilon_j^* \) is
monotonically decreasing with respect to \( j \) with a lower bound of zero.

We now have a computationally tractable solution to the decentralized controller design problem as defined in Problem 2. The exhaustive combinatorial search of the 0-norm to achieve a sparse structure of the controller was relaxed by the computationally attractive weighted \( \ell_1 \)-minimization. The non-convex \( H_{\infty} \)-performance constraint was reformulated by means of a system augmentation approach. Both relaxations lead to the relaxed decentralized controller problem as stated in Problem 9. This relaxed decentralized controller design problem can then be solved using Algorithm 1. To improve the initial value of the iteration, Algorithm 2 was introduced. In the next section, we will apply our results to an illustrative example.

5. EXAMPLE

The example is a decentralized interconnected system taken from \( \text{COMPIE} \) (Leibfritz, 2004, Example 'Decentralized Interconnected Systems 4') and consists of three subsystems with two states each

\[
\dot{x} = \begin{bmatrix} 0 & 10.5 & 10.6 & 0 \\ -0.1 & 20.5 & 0 & 0.5 \\ -1 & 3 & 0.5 & 0.5 \\ -3 & -4 & 0.5 & 0.5 & 0 \end{bmatrix} x + I w + \begin{bmatrix} 10 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{bmatrix} u \\
z = I x
\]

The system is unstable and fully coupled. While the subsystems have all the same size, the size of the individual control inputs is different. Therefore decentralized controller design for this system is more general than the presented set-up. A centralized controller designed using convex optimization techniques admits a closed loop performance of \( \| \Sigma_c \|_\infty = 0.8271 \). Figure 2 shows the \( H_{\infty} \)-performance level of the closed loop with different decentralized controllers computed with Algorithm 1 and Algorithm 2 with \( p = 1 \) for the objective function and \( p = \infty \) for the computation of the weights of the \( \ell_1 \)-minimization. The initial weight was chosen as the element-wise inverse of \( Z(K, \infty) \). As can be seen, the \( H_{\infty} \)-norm of the closed loop is increasing with increasing sparsity of the controller. That is, there is a trade-off between \( H_{\infty} \)-performance level and decentralization of the controller. Nevertheless, almost no performance degradation is visible between the central controller and a decentralized controller up to \( \| Z(K, \infty) \|_0 = 5 \). Figure 3 shows the non-zero blocks of the 3-by-3 block matrix \( \hat{K} \), i.e. the input of sub-system 1 depends on the state of subsystem 3 if there is a dot at \((3,1)\). Thus, it shows the structure of the controllers in Figure 2. As can be seen, there is an ordering of the subcontrollers in their influence on the achieved \( H_{\infty} \)-performance. Especially, the lower left element is important for the closed loop performance. Only a large performance degradation allows to cancel this element to achieve \( \| Z(K, \infty) \|_0 = 4 \), whereas the first four elements can be omitted with almost no performance degradation.

6. CONCLUSION

We presented a design method for \( H_{\infty} \) decentralized control of multivariable interconnected subsystems which are omnipresent in process control, e.g., highly coupled chemical plants, chemical reaction networks or distributed power generation networks. We formulated the decentralized controller design problem by minimizing the 0-norm of a pattern operator representing the structure of the controller to optimize the controller topology. In contrast to existing design methods for decentralized control, the structure of the decentralized controller and the controller itself are designed simultaneously and the minimization of the pattern operator is the objective function of the resulting optimization problem. For the resulting combinatorial optimization problem, computationally tractable convex relaxations have been provided. More specifically, by means of a system augmentation approach and a reweighted \( \ell_1 \)-relaxation, an iterative algorithm was developed to deal with the relaxed decentralized control problem. To further improve the results of the approach, an algorithm to optimize the initial values was presented. An illustrative example shows the trade-off between closed loop performance and achieved decentralization. Future research will deal with extensions to the more general case of static and dynamic output feedback as well as the problem of optimal sensor placement which can be seen as the dual problem to the presented decentralized controller design.

Appendix A. INTRODUCTION TO SPARSITY MEASURES AND SPARSITY PROMOTING OPTIMIZATION

In the following, we want to give a brief introduction into different sparsity measures and how sparsity promoting optimization can be used for structure and topology design.
in control. Sparse representation of signals is of fundamental importance in many fields such as blind source separation, compression, sampling and signal analysis (Hurley and Rickard, 2009). More recently, the concept of sparse controller and filter design was also considered in the control system community (Baran et al., 2010; Schuler et al., 2011; Fardad et al., 2011; Lin et al., 2011). Sparse controllers are especially of interest in networked control systems and decentralized control, where the controllers are spatially distributed and each controller has only access to a subset of the available information. Sparse filters are of interest since they offer the opportunity to omit arithmetic operations and the elimination and deactivation of circuits components. Generally speaking, in a sparse representation, a small number of coefficients or elements contain a large proportion of the energy. This heuristic interpretation leads to several possible sparsity measures (see Hurley and Rickard (2009) for an overview). Which sparsity measure should be used depends heavily on the considered application. Typically used sparsity measures are listed in Table A.1. We will describe their properties in the following with a special emphasis on how they can be used for optimization. This discussion is based on Hurley and Rickard (2009); Lin et al. (2011) and Boyd and Vandenberghe (2004, Chapter 6).

The \( \ell_0 \)-measure or 0-norm as introduced before is equal to the number of non-zero coefficients of a vector \( x \in \mathbb{R}^n \). This is the traditional sparsity measure in many mathematical settings. The derivative of the 0-norm contains no information, since it is always zero when defined. Therefore it is not commonly used in optimization problems. Exhaustive combinatorial search is the only way to find the sparsest solution when the 0-norm is considered (Candès et al., 2006a). Furthermore, the 0-norm is very sensitive to noise and numerical accuracy. In noisy settings, the 0-norm is often replaced by the \( \ell_{0,\epsilon} \)-measure, where only elements with absolute value larger than a threshold \( \epsilon \) are counted. Clearly, the solution then depends heavily on the choice of \( \epsilon \), which is different for each application and might not be known. An optimization with \( \ell_{0,\epsilon} \) as penalty function is also difficult, since the gradient contains again no information. Instead, the \( \ell_0 \)-measure is often used for \( 0 < p < 1 \). The level sets have a star-like shape and are depicted in Figure A.1(d). As can be seen, the level-sets of the \( p \)-norm and the feasible set intersect on one of the axes, i.e. the achieved solution is sparse. The log-sum measure has also a star-like shape and enforces sparsity outside some range. Both, the \( \ell_p \) and the log-sum measures promote sparse solutions, but the drawback of these measures is, that when used in optimization, the optimization problem has a non-convex objective function and is difficult to solve.

The 1-norm, approximates the 0-norm and is easily calculated. When used as penalty function in an optimization problem, this leads to a convex objective function and does not increase the complexity of a convex optimization problem. Furthermore, it was shown in Fazel (2002) that the 1-norm is the convex envelope of the 0-norm and therefore the best convex relaxation. The 1-norm is depicted in Figure A.1(b). As can be seen in the figure, not always a sparse solution is achieved. To overcome this problem, the weighted \( \ell_1 \)-minimization was introduced (Candès et al., 2008). In the weighted \( \ell_1 \)-minimization, a weight on each element counteracts the magnitude of this element and enforces the sparsity of the achieved solution (see also Figure A.1(c)). When the weight is chosen inversely proportional to the magnitude of this element, the 0-norm and the 1-norm coincide. This two important properties, namely the convex formulation and the coincidence with the 0-norm for properly chosen weights make the weighted \( \ell_1 \)-optimization very attractive. As can be seen in Figure A.1(a) this property only holds for \( p = 1 \). Minimizations over \( \ell_p \)-measure with \( p > 1 \) (e.g. \( p = 2 \) or \( p = \infty \)) are also convex problems but will in general not lead to sparse solutions. The previous discussion also motivates, why the weighted \( \ell_1 \)-minimization was used in this paper to design controllers with decentralized structure.

### Table A.1. Commonly used sparsity measures.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \ell_0 )</td>
<td>( \sum_{i=1}^{n}</td>
</tr>
<tr>
<td>( \ell_{0,\epsilon} )</td>
<td>( \sum_{i=1}^{n}</td>
</tr>
<tr>
<td>( \ell_1 )</td>
<td>( \sum_{i=1}^{n}</td>
</tr>
<tr>
<td>weighted ( \ell_1 )</td>
<td>( \sum_{i=1}^{n} m_i</td>
</tr>
<tr>
<td>( \ell_p )</td>
<td>( \left( \sum_{i=1}^{n}</td>
</tr>
<tr>
<td>log-sum</td>
<td>( \sum_{i=1}^{n} \log(1 + \frac{</td>
</tr>
</tbody>
</table>

### References


Fig. A.1. The solution $F^*$ of a constraint optimization problem is the intersection of the feasible set with the smallest sub-level set of the penalty function.


