Preprints of the 18th IFAC World Congress
Milano (Italy) August 28 - September 2, 2011

ℓ_0-System Gain and ℓ_1-Optimal Control

Simone Schuler, Christian Ebenbauer and Frank Allgöwer

Institute for Systems Theory and Automatic Control, University of Stuttgart, Germany, {schuler,ce,allgower}@ist.uni-stuttgart.de

Abstract: This paper introduces the concept of an ℓ_0-system gain for discrete-time LTI systems. It is shown that the ℓ_0-gain is characterized by the number of non-zero entries in the impulse response of the system and hence gives a natural extension of the notion of sparsity from signals to systems. With this newly introduced system gain, we give a system theoretic explanation of the sparse closed loop response of ℓ_1-optimal controlled systems by showing that the ℓ_1-optimal control problem is the best convex relaxation (in the sense of Lagrangian duality) of an appropriately defined ℓ_0-optimal control problem.

1. INTRODUCTION

Sparse signal recovery has received a lot of attention in the signal processing literature during the past five years. A discrete-time signal is said to be sparse, if most of its entries are zero, analog to a sparse vector, i.e. a vector where most of its entries are zero. The ℓ_0-norm quantifies sparsity by counting the number of non-zero entries in a vector or signal. Finding sparse vectors is important in many applications such as in parameter estimation or identification, signal processing or model reduction (Peeters and Westra, 2004; Candes et al., 2006b). However, finding sparse vectors, for example by minimizing the ℓ_0-norm, is a difficult non-convex problem. In compressive sensing (Candes et al., 2006b,a; Donoho, 2006) sparse signals are reconstructed from linear measurements by replacing the ℓ_0-minimization with an ℓ_1-minimization. Surprising about this replacement is the fact that it can be theoretically justified to be effective, i.e., under suitable assumptions, ℓ_1-minimization solves the ℓ_0-problem with high probability. The ℓ_1-norm can be considered as a convex relaxation of the ℓ_0-norm. In Fazel (2002) it was shown via conjugate functions that the ℓ_1-norm is the convex envelope of the ℓ_0-norm and therefore its best convex relaxation (in the sense of Lagrangian duality).

Motivated by the success of compressive sensing in signal processing, it is reasonable to ask if there is a meaningful notion of sparsity in the context of systems theory and how such a concept could look like. In this paper, a first step in this direction is made. The idea is to introduce an ℓ_0-system gain for single input/single output systems in the spirit of robust control for e.g. ℓ_2- or ℓ_∞-system gains. The ℓ_0-gain of a system is defined as the smallest ratio of the number of non-zero entries in the output signal and the input signals. Moreover, it is shown that the system gain is characterized by the number of non-zero entries of the impulse response. The second contribution of this paper is motivated by the fact that in ℓ_1-optimal control (Dahleh and Khammash, 1993; Dahleh and Diaz-Bobillo, 1995) it was observed that ℓ_1-optimal controllers produce sparse optimal closed loop responses. From the view of compressive sensing and ℓ_1-minimization, this behavior is reasonable but a systems theoretic explanation of this problem seems to be not available in literature. The paper establishes such a systems theoretic explanation of the sparse response of ℓ_1-optimal controllers by showing that the ℓ_1-optimal control problem is the best convex relaxation (in the sense of Lagrangian duality) of an appropriate ℓ_0-optimal control problem formulated with the help of the newly introduced ℓ_0-system gain. While our motivation to introduce a notion of sparsity for systems is primarily of theoretical interest, there are direct connections of the ℓ_0-gain to application areas, like the design of sparse finite impulse response filters, sparse channels and system identification via Markov parameters .

The paper is organized as follows: Section 2 states mathematical preliminaries. In Section 3 , the ℓ_0-system gain is introduced. In Section 4, the ℓ_0-optimal control problem is formulated and a convex relaxation as well as its relation to the ℓ_1-optimal control problem is established in Section 5. Section 6 gives a review of the ℓ_1-optimal control problem and shows the connection to the previously introduced ℓ_0-optimal control problem. The paper concludes with a summary and an outlook.

2. MATHEMATICAL PRELIMINARIES

A state-space realization of a transfer matrix $G(z)$ is written as

$$\left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] := C(zI - A)^{-1}B = G(z).$$

The sign(x) of $x \in \mathbb{R}$ is defined as

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

The ℓ_0-norm of a vector $x \in \mathbb{R}^n$ is defined as

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

The ℓ_0-norm counts the non-zero entries in a vector. Strictly speaking, it is not a norm, since homogeneity is not fulfilled ($\|\alpha x\|_0 = \|x\|_0 \neq |\alpha| \cdot \|x\|_0$), but nevertheless the term is commonly used in literature. A vector is called

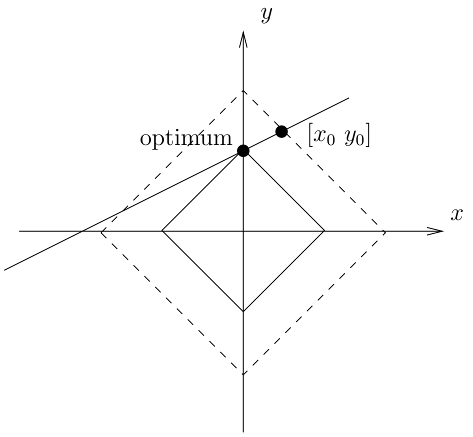

Fig. 1. The ℓ_1-ball for $\|[x\ y]\|_1 \leq 1$.

sparse, if it has a small ℓ_0-norm, i.e. if most of its entries are zero. The ℓ_1-norm of a vector $x \in \mathbb{R}^n$ is defined as

$$\|x\|_1 := \sum_{i=1}^{n} |x_i|.$$

As mentioned before, it was shown that the ℓ_1-norm is the convex envelope of the ℓ_0-norm (see Fazel, 2002). ℓ_1-minimization is known to lead to sparse solutions. An interpretation why this is the case is shown in Figure 1. If one minimizes the ℓ_1-norm of a vector subject to linear constraints, in most of the cases the solution lies on the corner of the ℓ_1-ball (Candes et al., 2008). In the following, we denote vector norms by $\|\cdot\|_i$ and signal norms by $\|\cdot\|_{\ell_i}$, $i = \{0, 1\}$. Notice that this distinction between ℓ_0-vector and ℓ_0-signal norm is not necessary but is done here because of conceptual clarity.

The convolution of two discrete time signals f and g is given by

$$(f * g)(k) = \sum_{i=0}^{\infty} f(i)g(k-i) = \sum_{i=0}^{\infty} f(k-i)g(i).$$

Consider an infinite dimensional sequence $x = \{x(1), x(2), \ldots, \} = \{x(k)\}$. Then the truncation operator $\mathcal{P}_N$ is defined by

$$\mathcal{P}_N(x) := \{x(1), x(2), \ldots, x(N), 0, 0, \ldots\}.$$

3. ℓ_0-SYSTEM GAIN

In this section, the concept of an ℓ_0-gain for discrete-time LTI systems is introduced. Moreover, a characterization of the ℓ_0-gain in terms of the number on non-zero entries in the impulse response of the system is derived.

3.1 Signal norm

We now consider $x = \{x(k)\} \in \ell_0$ being a discrete-time signal with the norm

$$\|x\|_{\ell_0} := \sum_{k=0}^{\infty} |\text{sign}(x(k))|.$$

The discrete-time impulse signal is given by

$$\delta(k) := \begin{cases} 1 & \text{for } k = 0 \\ 0 & \text{else .} \end{cases}$$

Note that the impulse signal δ fulfills $\|\delta\|_{\ell_0} = 1$.

3.2 Operator norm

Consider a discrete-time LTI system of the following form

$$x(k+1) = Ax(k) + Bw(k) \tag{1a}$$
$$z(k) = Cx(k) \tag{1b}$$

with matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times 1}$, and $C \in \mathbb{R}^{1\times n}$. For the sake of simplicity, we restrict ourself to SISO systems. It is expected that the multivariable case goes along similar lines and is part of our future work. The impulse response of this system is given in terms of its Markov parameters as

$$g(k) = \begin{cases} 0 & \text{for } k \leq 0 \\ CA^{k-1}B & \text{for } k > 0. \end{cases} \tag{2}$$

With the previously introduced ℓ_0-signal norm, we will now define the ℓ_0-gain of a system as the worst case ℓ_0 output signal for all possible input signals $w \in \ell_0$.

Definition 1. The induced ℓ_0-norm (or ℓ_0-gain) of an operator $G : \ell_0 \to \ell_0$ is defined as

$$\|G\|_{\ell_0-\text{ind}} := \sup_{w\neq 0} \frac{\|z\|_{\ell_0}}{\|w\|_{\ell_0}}, \quad w \in \ell_0$$

with $z = Gw$.

The above definition is exactly in the spirit of systems gains known from robust control theory.

A nice fact about the newly introduced ℓ_0-gain is that it is characterized by the sparsity of the impulse response of the system, as shown in the next theorem. Hence, this characterization justifies that the introduced ℓ_0-gain is a meaningful notion for sparsity.

Theorem 2. The ℓ_0-gain of the system (1) is the ℓ_0-norm of its impulse response, i.e.

$$\|G\|_{\ell_0-\text{ind}} = \|g\|_{\ell_0}.$$

Proof. Suppose w is an arbitrary input signal with $\|w\|_{\ell_0} = N$. Then two cases can be distinguished:

(1) If G has an infinite impulse response, then the ℓ_0-gain of (1) is infinity.
(2) If G has a finite impulse response, say $\|g\|_{\ell_0} := \|G\delta\|_{\ell_0} = M$, then

$$\frac{\|z\|_{\ell_0}}{\|w\|_{\ell_0}} \leq \frac{MN}{N} = \frac{\|G\delta\|_{\ell_0}}{\|\delta\|_{\ell_0}} = \|g\|_{\ell_0}.$$

The first inequality follows from the linearity of G, i.e. any output signal z is a superposition of scaled impulse responses (see Figure 2)

$$z(k) = \sum_{i=0}^{k} g(k-i)w(i)$$

with $g(k) = 0$, for $k > n$. Consequently, since w is arbitrary and $\|G\|_{\ell_0-\text{ind}} \leq M$, δ is an input signal such that $\|G\|_{\ell_0-\text{ind}} = M$. $\square$

Theorem 2 gives a direct link to finite impulse response filter, i.e. system (1) has a finite ℓ_0-gain if and only if it is a finite impulse response (FIR) filter. With the now defined ℓ_0-gain of a system, we want to design controllers/ filters K such that the closed loop system gain is minimal in terms of its ℓ_0-norm.

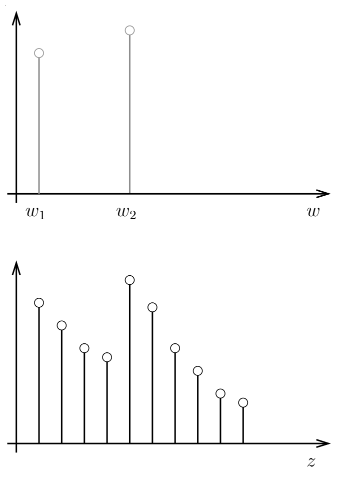

Fig. 2. System response as a superposition of impulse responses.

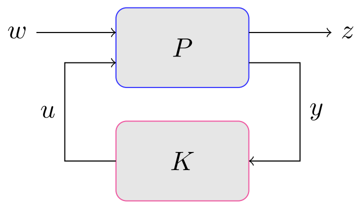

Fig. 3. Closed loop interconnection.

4. THE ℓ_0-OPTIMAL CONTROL PROBLEM

Consider the closed loop interconnection as depicted in Figure 3, where P is the generalized plant including all weighting functions and K is the controller to be designed, u is the controller output and y the measured output. The closed loop is given by $T(K) = \mathcal{F}_l(P, K)$, where $\mathcal{F}_l$ denotes the lower fractional transformation. We can reformulate this closed loop in terms of its Youla Parameterization (Youla et al., 1976)

$$T(Q) = H - UQV, \tag{3}$$

where H, U, V are transfer functions which are derived from the state space representation of the system and Q is a free, but stable transfer function. The controller K is then given by $K = \mathcal{F}_l(J, Q)$, where J can be computed from the state space description of the generalized plant P. We can now formulate the problem of finding a transfer function Q, such that the induced norm of the closed loop system is optimal in terms of the ℓ_0-gain:

$$\begin{aligned} \mu^0 &:= \inf_{Q} \|T(Q)\|_{\ell_0-\text{ind}} \\ &= \inf_{q\in\ell_0} \|h - u * q * v\|_{\ell_0}, \end{aligned} \tag{4}$$

where $*$ denotes the convolution operator. In other words, in the ℓ_0-optimal control problem, we search for an FIR filter Q such that the impulse response of the closed loop is sparse.

Following the ideas in Khammash (2000) we will reformulate the convolution terms. Let $p \in \ell_0$ be defined by $p := u * v$, and define p^k as

$$p^k := \{p(k), \ldots, p(0), 0, \ldots\}.$$

Since the convolution operation is associative and commutative it holds

$$\begin{aligned} u * q * v &= u * (q * v) \\ &= (u * v) * q \\ &= p * q. \end{aligned}$$

From this equation follows

$$\begin{aligned} (u * q * v)(k) =& (p * q)(k) \\ =& \langle p^k, q\rangle, \end{aligned}$$

where $\langle\cdot,\cdot\rangle$ denotes the standard inner product between vectors (ℓ_0-signals).

With this reformulation of the convolution terms, (4) can be written as

$$\mu^0 = \inf_{q\in\ell_0} \quad \|\{h(k) - \langle p^k, q\rangle\}\|_{\ell_0}. \tag{5}$$

This problem is an infinite dimensional optimization problem which is non-convex and difficult to solve.

If the ℓ_0-norm of a signal x is finite, then there exists an element $x(m)$ of the sequence which is the last non-zero element of x. It then holds

$$\|x\|_{\ell_0} = \|\mathcal{P}_N(x)\|_{\ell_0}, \quad \text{for } N \geq m.$$

Consequently, if (5) has a solution $q^* \in l_0$ (the infimum is attained), then it is equivalent to

$$\min_{q\in\ell_0} \quad \|\mathcal{P}_N(\{h(k) - \langle p^k, q\rangle\})\|_0 \tag{6}$$

with N sufficiently large. With this, we have derived an equivalent formulation of the optimization problem, which is now a finite-dimensional optimization problem. Compared to ℓ_1-optimal control, no relaxation is necessary to transform the infinite dimensional optimization problem into a finite one assuming that the infimum is attained (Khammash, 2000). In the next section, Lagrangian duality is applied to obtain a convex relaxation of the ℓ_0-optimal control problem.

5. LAGRANGIAN RELAXATION OF THE ℓ_0-OPTIMAL CONTROL PROBLEM

In this section, we want to show that the Lagrangian relaxation of the ℓ_0-problem is the ℓ_1-problem. We do this by showing that both, the ℓ_1-problem and the ℓ_0-problem have the same dual optimization problem. Since we know that the ℓ_1-problem is bi-dual, we can prove that the ℓ_1-problem is the bi-dual of the ℓ_0-problem and therefore its convex relaxation. Problem (6) can be written as a standard linear program of the form

$$\min_{x} \quad \|z\|_0 \tag{7a}$$

$$\text{s.t.} \quad h - Ax = z, \tag{7b}$$

where $h = [h(k)]$, $Ax = [\langle p^k, q\rangle]$ and $x = q$. Before we solve the problem stated in (7), we consider two auxiliary problems that will lead to the solution of the original problem.

The idea is to augment the original optimization problem by a new term and then show that the augmented problem leads to the same solution as the original problem. We will first consider a standard ℓ_1-minimization problem to show the ideas. Afterwards we will again consider the previous stated ℓ_1-optimal control problem. The dual optimization problem of the ℓ_1-minimization

$$\min_x \quad \|x\|_1 \tag{8a}$$
$$\text{s.t.} \quad Ax = b, \tag{8b}$$

is given by

$$\max_\nu \quad \nu^T b \tag{9a}$$
$$\text{s.t.} \quad \|A^T\nu\|_\infty \leq 1. \tag{9b}$$

A proof of this can be found for example in Boyd and Vandenberghe (2004). Consider first the optimization problem

$$\min_x \quad \|x\|_0 \tag{10a}$$
$$\text{s.t.} \quad Ax = b, \tag{10b}$$

for given $x \in \mathbb{R}^n$, $b \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m\times n}$, with $m < n$ and with optimal solution x^*. In Fazel (2002), it was shown that the ℓ_1-norm is the bi-conjugate of the ℓ_0-norm and therefore its convex envelope. However, this only holds when x is restricted to a bounded domain, i.e $C = \{x| \|x\|_\infty < 1\}$. Otherwise, the convex envelope is identically zero. Since we do not want to restrict the domain of x, we will present an alternative approach here. Consider the optimization problem

$$\min_x \quad c\|x\|_0 + \|x\|_1 \tag{11a}$$
$$\text{s.t.} \quad Ax = b \tag{11b}$$

with optimal solution $\tilde{x}^*$ and $c \in \mathbb{R}^+$ is a positive constant. This scaled optimization problem is easier to analyze and has the important property that the ℓ_0-norm of the optimal solution of (11) coincide with the solution of (10), for a sufficiently large constant c. The next lemma shows that both problems have the same optimal solution in terms of the ℓ_0-norm for sufficiently large constant c.

Lemma 3. Let x^*, $\tilde{x}^*$ be the solutions of (10) and (11), respectively. Then for sufficiently large c it holds

$$\|x^*\|_0 = \|\tilde{x}^*\|_0.$$

The proof is given in the appendix.

Lemma 3 allows us to search for sparse vectors by solving Problem (11) instead of Problem (10). Moreover, the advantage of (11) over (10) is, that its Lagrangian relaxation is not degenerated. To see this, we dualize Problem (11).

Lemma 4. The Lagrange dual of (11) is given by

$$\max_\nu \quad \nu^T b$$
$$\text{s.t.} \quad \|A^T\nu\|_\infty \leq 1.$$

The proof is given in the appendix.

Lemma 4 shows that the ℓ_1-minimization problem (8) and the augmented ℓ_0-optimization problem (11) have the same dual. Therefore the ℓ_1-minimization problem is the closest convex relaxation (in terms of Lagrangian duality) of the augmented ℓ_0-optimization problem.

After considering this two auxiliary problems, we can solve the problem originally stated in (6) and in its equivalent standard form (7). With

$$\xi = \begin{bmatrix} z \\ x \end{bmatrix}, \; \tilde{A} = [I \; A], \text{ and } D = [I \; 0],$$

the minimization problem (7) can be rewritten as

$$\min_\xi \quad \|D\xi\|_0$$
$$\text{s.t.} \quad \tilde{A}\xi = h.$$

This problem is similar to (10) and the same conclusions hold.

Corresponding to Lemma 4, we have the following theorem.

Theorem 5. The Lagrange relaxation of the optimization problem

$$\min_x \quad c\|z\|_0 + \|z\|_1 \tag{13a}$$
$$\text{s.t.} \quad h - Ax = z \tag{13b}$$

is given by

$$\max_\nu \quad \nu^T h$$
$$\text{s.t.} \quad \|\nu\|_\infty \leq 1$$
$$A^T\nu = 0.$$

Before we can prove this theorem, we state a Lemma similar to Lemma 3.

Lemma 6. Let ξ^*, $\tilde{\xi}^*$ be the solutions of (7) and (13), respectively. Then for sufficiently large c it holds

$$\|\xi^*\|_0 = \|\tilde{\xi}^*\|_0.$$

The proof of the theorem is similar to the proof of Lemma 3 and is omitted here.

Proof. [Theorem 5] The proof goes along the lines of the proof of Lemma 4. With ξ, $\tilde{A}$ and D as defined before, the minimization problem can be rewritten as

$$\min_\xi \quad c\|D\xi\|_0 + \|D\xi\|_1$$
$$\text{s.t.} \quad \tilde{A}\xi = h.$$

The Lagrangian of this problem is given by

$$\begin{aligned} L(\xi,\nu) &= c\|D\xi\|_0 + \|D\xi\|_1 + \nu^T(h - \tilde{A}\xi) \\ &= c\sum_{i=1}^{2n} d_{ii}|\text{sign}(\xi_i)| + \sum_{i=1}^{2n} d_{ii}|\xi_i| + \nu^T h - \nu^T\tilde{A}\xi \\ &= \sum_{i=1}^{2n} d_{ii}\left(c|\text{sign}(\xi_i)| + |\xi_i|\right) + \nu^T h - \nu^T\tilde{A}\xi \end{aligned}$$

and $d_{ii} = 0$ for $i > n$. With $\tilde{a}_i$ being the ith column of $\tilde{A}$, it follows

$$\begin{aligned} g(\nu) &= \min_\xi L(\xi,\nu) \\ &= \min_\xi\Big(\sum_{i=1}^{2n}\big(d_{ii}\left(c|\text{sign}(\xi_i)| + |\xi_i|\right) - (\nu^T\tilde{a}_i)\xi_i\big) \\ &\quad + \nu^T h\Big) \\ &= \begin{cases} -\infty & \text{for} \quad |\nu^T\tilde{a}_i| > d_{ii} \text{ for some } i \\ \nu^T h & \text{for} \quad |\nu^T\tilde{a}_i| \leq d_{ii} \text{ for all } i. \end{cases} \end{aligned}$$

This is due to the fact, that if $|\nu^T\tilde{a}_i| \leq d_{ii}$ then

$$\alpha(x_i) = d_{ii}\left(c|\text{sign}(x_i)| + |x_i|\right) - (\nu^T\tilde{a}_i)x_i$$

is positive definite ($\alpha(0) = 0$, $\alpha(x_i) \geq 0$), and if $|\nu^T\tilde{a}_i| > d_{ii}$, then $\alpha(\rho(\nu^T\tilde{a}_i)) \to -\infty$ for $\rho \to \infty$.

The dual problem can now be written as

$$\max_\nu \quad \nu^T h$$
$$\text{s.t.} \quad |\nu^T\tilde{a}_i| \leq d_{ii}$$

with $d_{ii} = 1$ for $i = 1 \ldots n$ and $d_{ii} = 0$ for $i = (n+1)\ldots 2n$. This can be rewritten as

$$\max_\nu \quad \nu^T h$$
$$\text{s.t.} \quad \|\nu^T\|_\infty \leq 1$$
$$A^T\nu = 0. \qquad \square$$

In the next section, we summarize the results obtained so far and establish the connection between the ℓ_0- and ℓ_1-optimal control problem.

6. THE ℓ_1-OPTIMAL CONTROL PROBLEM

In Khammash (2000) the ℓ_1-optimal control problem is formulated as follows

$$\mu^1 = \inf_{q \in \ell_1} \|h - u * q * v\|_1.$$

This problem has the same structure as Problem (4), with the only difference that the minimization takes place in ℓ_1. Reordering the convolution terms as shown in (Khammash, 2000) leads to the optimization problem

$$\mu^1 = \inf_{q \in \ell_1} \quad \|\{h(k) - \langle p^k, q\rangle\}\|_{\ell_1}, \tag{15}$$

which has the same structure as (5). In the same way as shown for the ℓ_0-optimal control problem, this problem can also be rewritten as

$$\min_x \quad \|z\|_1 \tag{16a}$$
$$\text{s.t.} \quad h - Ax = z, \tag{16b}$$

where $h = [h(k)]$, $Ax = [\langle p^k, q\rangle]$ and $x = q$. For finite k the results of Section 5 apply and consequently the proposed ℓ_0-optimal control problem is related to the ℓ_1-optimal control problem by Lagrangian duality. Summarizing, this explains why ℓ_1-optimal controlled systems have sparse impulse responses.

In contrast to the ℓ_0-optimal control problem, Problem (15) is a infinite dimensional problem and cannot be transformed into an equivalent finite dimensional problem. To tackle this numerically, the scaled q-method was introduced in Khammash (2000). Therefore, first the problem is rewritten as

$$\begin{aligned}\min_{q\in\ell_1} \quad & \max\{\|h - r\|_1, \alpha\|q\|_1\}\\ \text{s.t.}\ & r = u * q * v,\end{aligned}$$

and lower and upper bounds can be computed.
A lower bound for $\underline{\mu}_N(\alpha)$:

$$\begin{aligned}\underline{\mu}_N(\alpha) = \min_{q\in\ell_1} \ & \max\{\|h - r\|_1, \alpha\|q\|_1\}\\ \text{s.t.}\ & \mathcal{P}_N(r) = \mathcal{P}_N(u * q * v)\end{aligned}$$

A upper upper bound for $\bar{\mu}_N(\alpha)$:

$$\begin{aligned}\bar{\mu}_N(\alpha) = \min_{q\in\ell_1} & \max\{\|h - r\|_1, \alpha\|q\|_1\}\\ \text{s.t.}\ & r = u * \mathcal{P}_N(q) * v\end{aligned}$$

This allows to compute ℓ_1-optimal controllers and delivers the convex relaxation for suboptimal ℓ_0-controllers in terms of Lagrangian duality.

7. CONCLUSIONS

We introduced the ℓ_0-system gain for discrete-time LTI systems as a natural extension of the notion of sparsity from signals to systems. We could show that the ℓ_0-system gain is characterized by the number of non-zero entries of the impulse response. It was also shown that the ℓ_1-optimal control problem is the convex Lagrangian relaxation of the ℓ_0-optimal control problem. With this the observation of sparse optimal closed loop responses in ℓ_1-optimal control can be explained. While this work is motivated by the attempt to introduce a notion of sparsity for systems and by explaining the sparse response of ℓ_1-optimal controllers, it is hoped that an ℓ_0-system gain may also lead to new approaches in systems analysis and controller design. Ongoing research deals with the design of finite impulse response filters with minimum number of elements in the impulse response and the construction of sparse channels. Especially in network controlled systems ℓ_0-optimal control might be of interest since sparse signals often translate in low data rates. On the more theoretical side, future work should include the extension to multivariable systems as well as robustness issues.

Appendix A. PROOFS

Proof. [Lemma 3] The proof is given by contradiction. The solution set of (10) and the solution set of

$$\min_x \quad c\|x\|_0, \quad c > 0 \tag{A.1a}$$
$$s.t. \quad Ax = b. \tag{A.1b}$$

are identical. Let x^* be an optimal solution of (10) and $\tilde{x}^*$ be an optimal solution of (11) with $\|x^*\|_0 < \|\tilde{x}^*\|_0$. By optimality it follows

$$c\|x^*\|_0 + \|x^*\|_1 > c\|\tilde{x}^*\|_0 + \|\tilde{x}^*\|_1 \tag{A.2a}$$
$$\|x^*\|_1 > c(\underbrace{\|\tilde{x}^*\|_0 - \|x^*\|_0}_{\geq 1}) + \|\tilde{x}^*\|_1. \tag{A.2b}$$

Since x^* is finite and independent of c, i.e. x^* is a solution of (10) (and (A.1)) we obtain a contradiction in (A.2b) for c sufficiently large. Finally, $\|x^*\|_0 > \|\tilde{x}^*\|_0$ contradicts the assumption of optimality, since then, $\|\tilde{x}^*\|_0$ would be the optimal solution to (10). Therefore $\|x^*\|_0 = \|\tilde{x}^*\|_0$. □

Proof. [Lemma 4] Let a_i denote the ith column of A. The Lagrangian is given by

$$\begin{aligned}L(x,\nu) &= c\|x\|_0 + \|x\|_1 + \nu^T(b - Ax)\\ &= c\sum_{i=1}^n |\text{sign}(x_i)| + \sum_{i=1}^n |x_i| + \nu^T b - \sum_{i=1}^n (\nu^T a_i)x_i\\ &= \sum_{i=1}^n \big(c|\text{sign}(x_i)| + |x_i| - (\nu^T a_i)x_i\big) + \nu^T b.\end{aligned}$$

The Lagrange function is

$$\begin{aligned}g(\nu) &= \min_x L(x,\nu)\\ &= \begin{cases} -\infty & \text{for} \quad |\nu^T a_i| > 1 \text{ for some } i\\ \nu^T b & \text{for} \quad |\nu^T a_i| \leq 1 \text{ for all } i.\end{cases}\end{aligned}$$

This follows from the fact that if $|\nu^T a_i| \leq 1$ then

$$\alpha(x_i) = c|\text{sign}(x_i)| + |x_i| - (\nu^T a_i)x_i$$

is positive definite ($\alpha(0) = 0, \alpha(x_i) \geq 0$), and if $|\nu^T a_i| > 1$, then $\alpha(\rho(\nu^T a_i)) \to -\infty$ for $\rho \to \infty$. The dual problem is then as follows

$$\begin{aligned}\max_\nu \quad & \nu^T b\\ \text{s.t.} \quad & \|\nu^T A\|_\infty \leq 1.\end{aligned} \qquad \square$$

REFERENCES

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization.* Cambridge University Press.

Candes, E.J., Romberg, J.K., and Tao, T. (2006a). Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Information Theory*, 52(2), 489–509.

Candes, E.J., Romberg, J.K., and Tao, T. (2006b). Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8), 1207–1223.

Candes, E.J., Wakin, M.B., and Boyd, S. (2008). Enhancing sparsity by reweighted ℓ_1 minimization. *Journal of Fourier Analysis and Applications*, 14(5), 877–905.

Dahleh, M.A. and Diaz-Bobillo, I.J. (1995). *Control of Uncertain Systems - A Linear Programming Approach*. Prentice Hall, Englewood Cliffs, New Jersey.

Dahleh, M.A. and Khammash, M.H. (1993). Controller design for plants with structured uncertainties. *Automatica*, 29(1), 37–56.

Donoho, D.L. (2006). Compressed sensing. *IEEE Trans. Information Theory*, 52(4), 1289–1306.

Fazel, M. (2002). *Matrix Rank Minimization with Applications*. Ph.D. thesis, Department of Electrical Engineering, Stanford University.

Khammash, M.H. (2000). A new approach to the solution of the ℓ_1 control problem: The scaled-Q method. *IEEE Trans. Automat. Control*, 45(2), 180–187.

Peeters, R. and Westra, R. (2004). On the identification of sparse gene regulatory networks. In *Proc. of the 16th Int. Symp. on Math. Theo. of Netw.*, 1–14.

Youla, D.C., Jabr, H.A., and Bongiorno, J.J. (1976). Modern Wiener-Hopf design of optimal controllers - Part 2: The multivariable case. *IEEE Trans. Automat. Control*, 21(3), 319–338.