On Computing the Worst-Case Peak Gain of Linear Systems

V. Balakrishnan and S. Boyd*

(To Appear in Systems and Control Letters, 1992)

Abstract

Based on the bounds due to Doyle and Boyd, we present simple upper and lower bounds for the ℓ^1-norm of the 'tail' of the impulse response of finite-dimensional discrete-time linear time-invariant systems. Using these bounds, we may in turn compute the ℓ^∞-gain of these systems to any desired accuracy. By combining these bounds with results due to Khammash and Pearson, we derive upper and lower bounds for the worst-case ℓ^∞-gain of discrete-time systems with diagonal perturbations.

Keywords: SISO discrete-time LTI systems, computation of ℓ^∞-gain, discrete-time systems with diagonal perturbations, worst-case ℓ^∞-gain.

1 Notation

$\mathbf{Z}_+$, $\mathbf{R}$, $\mathbf{R}_+$ and $\mathbf{C}$ denote the set of nonnegative integers, real numbers, nonnegative real numbers and complex numbers respectively. All the sequences in this note are defined over $\mathbf{Z}_+$. The ℓ^∞-norm of a complex-valued sequence u is defined as

$$\|u\|_\infty \stackrel{\Delta}{=} \sup_{k\geq 0} |u(k)|.$$

Thus, the ℓ^∞-norm of a sequence is its peak value. The ℓ^1-norm of a complex-valued sequence u is defined as

$$\|u\|_1 \stackrel{\Delta}{=} \sum_{k\geq 0} |u(k)|.$$

For a matrix $P \in \mathbf{R}^{n\times n}$, P^T stands for the transpose. $\sigma_1(P), \sigma_2(P), \ldots, \sigma_n(P)$ are the singular values of P in decreasing order. $\rho(P)$ denotes the spectral radius, which is the maximum magnitude of the eigenvalues of P. I stands for the identity matrix, with size determined from context.

*Research supported in part by AFOSR under contract F49620-92-J-0013.

2 Bounds for the ℓ^∞-gain

Consider a stable, finite-dimensional discrete-time linear time-invariant (LTI) system described by the state equations

$$\begin{aligned} x(k+1) &= Ax(k) + bu(k), \quad x(0) = 0, \\ y(k) &= cx(k) + du(k), \end{aligned} \tag{1}$$

where the input $u(k) \in \mathbf{R}$, the output $y(k) \in \mathbf{R}$ and the state $x(k) \in \mathbf{R}^n$. We assume that $\{A, b, c, d\}$ is minimal. The *impulse response* of system (1) is the real sequence given by

$$h(k) \triangleq \begin{cases} d, & k = 0, \\ cA^{k-1}b, & k > 0. \end{cases}$$

The ℓ^∞-gain of system (1), which is the largest possible peak value of the output y over all possible inputs u with a peak value of at most one, is just $\|h\|_1$:

$$\|h\|_1 = \sup_{\|u\|_\infty > 0} \frac{\|y\|_\infty}{\|u\|_\infty}.$$

$\|h\|_1$ is usually approximated by summing only a finite, typically large (say N) number of terms:

$$S_N = \sum_{k=0}^{N} |h(k)| \leq \|h\|_1.$$

Obviously, S_N is a lower bound for $\|h\|_1$, and increases monotonically to $\|h\|_1$ with increasing N. The 'error' $\|h\|_1 - S_N$ is just the ℓ^1 norm of the tail, $\sum_{k>N} |h(k)|$. Many simple bounds on this error are possible; for instance, if the poles of the system (1) are distinct, we may write down a residue expansion for the impulse response $h(k)$:

$$h(k) = \begin{cases} d, & k = 0, \\ \sum_{i=1}^{n} r_i {p_i}^{k-1}, & k > 0. \end{cases}$$

where $p_1, p_2, \ldots, p_n$ are the distinct poles of the system and r_i are the residues (see for example, [7], Chapter 2). Then,

$$\sum_{k>N} |h(k)| \leq \sum_{i=1}^{n} |r_i| \frac{|p_i|^N}{1 - |p_i|}. \tag{2}$$

Similar bounds are possible when the poles are not distinct.

The first purpose of this note is to present more sophisticated, and in many cases, substantially better bounds for the ℓ^1-norm of the tail. These bounds are based on Theorem 2 of [2], which states that for the system (1),

$$|d| + \sigma_1(W_o^{\frac{1}{2}} W_c^{\frac{1}{2}}) \leq \|h\|_1 \leq |d| + 2\sum_{i=1}^{n} \sigma_i(W_o^{\frac{1}{2}} W_c^{\frac{1}{2}}), \tag{3}$$

where

$$W_o = \sum_{k=0}^{\infty} (A^T)^k c^T c A^k \quad \text{and} \quad W_c = \sum_{k=0}^{\infty} A^k b b^T (A^T)^k$$

are the observability and controllability Gramians respectively [4]. $\sigma_i(W_o^{\frac{1}{2}} W_c^{\frac{1}{2}})$ are just the Hankel singular values of the system (1).

We now observe that $\{0, h(N+1), h(N+2), \ldots\}$, the tail of the impulse response of system (1), is just the impulse response of the system $\{A, A^N b, c, 0\}$. Applying bounds (3) to this system, we have for any $N \geq 0$,

$$\sigma_1(W_o^{\frac{1}{2}} A^N W_c^{\frac{1}{2}}) \leq \sum_{k>N} |h(k)| \leq 2\sum_{i=1}^{n} \sigma_i(W_o^{\frac{1}{2}} A^N W_c^{\frac{1}{2}}). \tag{4}$$

Thus, we have upper and lower bounds for $\|h\|_1$:

$$S_N + \sigma_1(W_o^{\frac{1}{2}} A^N W_c^{\frac{1}{2}}) \leq \|h\|_1 \leq S_N + 2\sum_{i=1}^{n} \sigma_i(W_o^{\frac{1}{2}} A^N W_c^{\frac{1}{2}}), \quad \forall N \geq 0. \tag{5}$$

The ratio between the upper and lower bounds for $\|h\|_1$ in (4) is at most $2n$, whereas the ratio between the residue-expansion based upper bound (2) and any lower bound can be arbitrarily large.

We next show that with increasing N, the difference between the upper and lower bounds converges monotonically to zero. W_o satisfies the Lyapunov equation

$$A^T W_o A - W_o + c^T c = 0,$$

which implies that

$$(A^T)^k W_o A^k - (A^T)^{k-1} W_o A^{k-1} + (A^T)^{k-1} c^T c A^{k-1} = 0$$

for $k = 1, 2, \ldots$ Therefore,

$$(W_o^{\frac{1}{2}} A^k W_c^{\frac{1}{2}})^T (W_o^{\frac{1}{2}} A^k W_c^{\frac{1}{2}}) \leq (W_o^{\frac{1}{2}} A^{k-1} W_c^{\frac{1}{2}})^T (W_o^{\frac{1}{2}} A^{k-1} W_c^{\frac{1}{2}}), \quad k = 1, 2, \ldots$$

This immediately means

$$\sigma_i(W_o^{\frac{1}{2}} A^k W_c^{\frac{1}{2}}) \leq \sigma_i(W_o^{\frac{1}{2}} A^{k-1} W_c^{\frac{1}{2}}), \quad i = 1, 2, \ldots, n \text{ and } k = 1, 2, \ldots,$$

from which it follows that the difference between the upper and lower bounds in (5) converges monotonically to zero with increasing N.

The above argument shows that all of the Hankel singular values of the impulse response of the 'tail' system $\{A, A^N b, c, 0\}$ decrease monotonically (to zero, since the system is stable) as $N \to \infty$. In fact, we can say more: If we normalize the Hankel singular values by dividing them by the first one, the number of 'normalized' Hankel singular values that converge to nonzero values as $N \to \infty$ equals the number of 'dominant' Jordan blocks of A, that is, the number of Jordan blocks of A which

- correspond to an eigenvalue of A with maximum magnitude, and
- which have the largest size among all Jordan blocks corresponding to an eigenvalue with maximum magnitude.

Thus, for large N, the number of significant terms in the sum $\sum_{i=1}^{n} \sigma_i(W_o^{\frac{1}{2}} A^N W_c^{\frac{1}{2}})$ is just the 'effective order' of the tail system $\{A, A^N b, c, 0\}$.

Finally, we discuss informally a scheme for finding

$$N_{\min} = \min \left\{ N \;\middle|\; 2 \sum_{i=1}^{n} \sigma_i(W_o^{\frac{1}{2}} A^N W_c^{\frac{1}{2}}) - \sigma_1(W_o^{\frac{1}{2}} A^N W_c^{\frac{1}{2}}) < \epsilon \right\},$$

which is the smallest value of N for which the difference between the upper and lower bounds in (5) is less than ϵ. As a preliminary step, $W_c^{\frac{1}{2}}$ and $W_o^{\frac{1}{2}}$ are computed. Then:

1. *We find the smallest positive integer M such that $N_{\min} \leq 2^M$.*

   This is done iteratively where at the kth iteration, we form the matrix A^{2^k} by squaring $A^{2^{k-1}}$ and check if

   $$\sigma_1(W_o^{\frac{1}{2}} A^{2^k} W_c^{\frac{1}{2}}) + 2 \sum_{i=2}^{n} \sigma_i(W_o^{\frac{1}{2}} A^{2^k} W_c^{\frac{1}{2}}) < \epsilon,$$

and stop if the condition is satisfied. Clearly, M iterations are needed. Each iteration involves three $n \times n$ matrix multiplies and one computation of singular values. For use in part (2), we store the matrices $\{A, A^2, \ldots, A^{2^M}\}$.

2. *By a simple bisection,* $N_{\min}$ *is then located in the set* $\{2^{M-1}, 2^{M-1}+1, \ldots, 2^M\}$.

   We assume that $M \geq 2$, since computing $N_{\min}$ is trivial otherwise. We start by forming $\tilde{A} = A^{(2^{M-1}+2^{M-2})}$ and checking if

   $$\sigma_1(W_o^{\frac{1}{2}}\tilde{A}W_c^{\frac{1}{2}}) + 2\sum_{i=2}^{n}\sigma_i(W_o^{\frac{1}{2}}\tilde{A}W_c^{\frac{1}{2}}) < \epsilon.$$

   (Note that since $A^{2^{M-1}}$ and $A^{2^{M-2}}$ are both already available from step (1), and therefore this involves three $n \times n$ matrix multiplies and one computation of singular values.) If the answer is yes, then N lies in the set $\{2^{M-1}, 2^{M-1}+1, \ldots, 2^{M-1}+2^{M-2}\}$. Otherwise, N lies in the set $\{2^{M-1}+2^{M-2}, \ldots, 2^M\}$. By continuing this process (at most $M-1$ times) of halving the set where N lies, we may compute $N_{\min}$ exactly.

Once $N_{\min}$ is found, $S_{N_{\min}}$ can be computed to give $\|h\|_1$ to within an absolute accuracy of ϵ (assuming infinite precision arithmetic; we have not considered the effects of data rounding here).

The exact determination of $N_{\min}$ takes approximately $6M$ matrix multiplies and $2M$ computations of singular values. Forming $S_{N_{\min}}$ takes about $N_{\min}$ matrix-vector multiplies and $N_{\min}$ vector-vector inner products. (Recall that $2^{M-1} < N_{\min} \leq 2^M$.) Since computing singular values is by far the most expensive of the above calculations, it might prove advantageous to not compute $N_{\min}$ exactly, but to instead use an upper bound obtained by terminating the bisection in step (2) earlier. Computation may be further reduced by first balancing system (1), so that the Gramians W_c and W_o are diagonal and equal.

We note that for calculating the $\mathbf{H}_\infty$-norm of system (1) to within a relative accuracy ϵ, there exist methods (see [1]) where the computational effort involved depends only on ϵ and the state dimension n. However for determining $\|h\|_1$ using the bounds in (5) to within an accuracy of ϵ

(relative or absolute), the number of computations depends on the system matrices A, b, c and d as well. We know of no way to overcome this deficiency.

3 Bounds for the worst-case ℓ^∞-gain

We now combine the results of the previous section with results from [5] to derive bounds for the worst-case ℓ^∞-gain of discrete-time LTI systems with diagonal uncertainty. We consider the system shown in Figure 1: H is a stable discrete-time LTI plant. $\Delta_1, \Delta_2, \ldots, \Delta_m$ are scalar LTI perturbations that act on the system. Now, for some notation (indices $i, j = 1, 2, \ldots, m$):

δ_i	:	Impulse response of perturbation Δ_i.
h_{00}	:	Open-loop ($\Delta = 0$) impulse response from w to z.
h_{i0}	:	Open-loop ($\Delta = 0$) impulse response from w to y_i.
h_{0i}	:	Open-loop ($\Delta = 0$) impulse response from u_i to z.
$h_{\mathrm{cl}}(\Delta)$	:	Closed-loop impulse response from w to z.

We assume that $\|\delta_i\|_1 \leq 1$ and denote by Ω the corresponding set of all possible perturbations Δ.

The quantity of interest is the worst-case (*i.e.* maximum possible) ℓ^∞-gain from w to z, which we define as

$$L_{\mathrm{wc}} = \sup_{\Delta \in \Omega} \|h_{\mathrm{cl}}(\Delta)\|_1.$$

In [5], Khammash and Pearson show that the $L_{\mathrm{wc}} \geq 1$ if and only if the following condition holds:

There exists some nonzero $x = [x_0, ..., x_m]$ with $x_i \geq 0$ such that

$$x_i \leq \sum_{j=0}^{m} \|h_{ij}\|_1 x_j \quad i = 0, 1, ..., m. \qquad \text{(COND)}$$

Condition (COND) may be expressed simply in terms of a matrix whose (i,j)-entry is $\|h_{ij}\|_1$, $i, j = 0, 1, \ldots, m$.

Fact 1 *Condition* COND *holds if and only if the spectral radius of the matrix*

$$M = \begin{bmatrix} \|h_{00}\|_1 & \|h_{01}\|_1 & \cdots & \|h_{0m}\|_1 \\ \|h_{10}\|_1 & \|h_{11}\|_1 & \cdots & \|h_{1m}\|_1 \\ \vdots & \vdots & \ddots & \vdots \\ \|h_{m0}\|_1 & \|h_{m1}\|_1 & \cdots & \|h_{mm}\|_1 \end{bmatrix}$$

is at least one.

This fact, stated without proof in Theorem 1 of [6], is immediate from the following characterization of the spectral radius of a nonnegative matrix (a matrix with nonnegative entries) M (see, for example, page 504, corollary 8.3.3 of [3]):

$$\rho(M) = \max_{x \geq 0,\, x \neq 0} \; \min_{0 \leq i \leq m,\, x_i \neq 0} \frac{1}{x_i} \sum_{j=0}^{m} M_{ij} x_j .$$

(M_{ij} refers to the (i,j)-entry of M.)

By simply scaling w and z by $1/\sqrt{\gamma}$ ($\gamma > 0$) as in Figure 2, and applying Fact 1, we conclude that

$$L_{\mathrm{wc}} = \sup\{\gamma \mid \rho\left(D_\gamma M D_\gamma\right) \geq 1\}, \tag{6}$$

where

$$D_\gamma = \begin{bmatrix} 1/\sqrt{\gamma} & 0 \\ 0 & I \end{bmatrix}.$$

For convenience, we partition M as

$$M = \begin{bmatrix} M^{(11)} & M^{(12)} \\ M^{(21)} & M^{(22)} \end{bmatrix}, \tag{7}$$

where $M^{(11)} \in \mathbf{R}_+$, $M^{(12)} \in \mathbf{R}_+^{1\times m}$, $M^{(21)} \in \mathbf{R}_+^{m\times 1}$ and $M^{(22)} \in \mathbf{R}_+^{m\times m}$.

If $\rho\left(D_\gamma M D_\gamma\right) \geq 1$ for all $\gamma > 0$, then we define $L_{\mathrm{wc}} = \infty$. This corresponds to the case when $\rho(M^{(22)}) \geq 1$, and the system is not ℓ^∞-stable (see [5]). On the other hand, if $\rho\left(D_\gamma M D_\gamma\right) < 1$ for all $\gamma > 0$, we define $L_{\mathrm{wc}} = 0$. This corresponds to the case when either the first row (or the first column) of M is identically zero (with $\rho(M^{(22)}) < 1$). Then $h_{\mathrm{cl}}(\Delta) = 0$ for all Δ.

Of course, every entry of M is the ℓ^∞-gain of some LTI system; therefore, the remarks made in Section 2 about computing ℓ^∞-gains apply here as well. We may however use the fact that M is

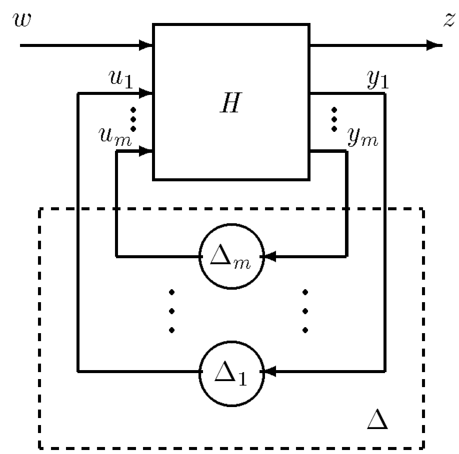

Figure 1: Linear system with diagonal uncertainty Δ.

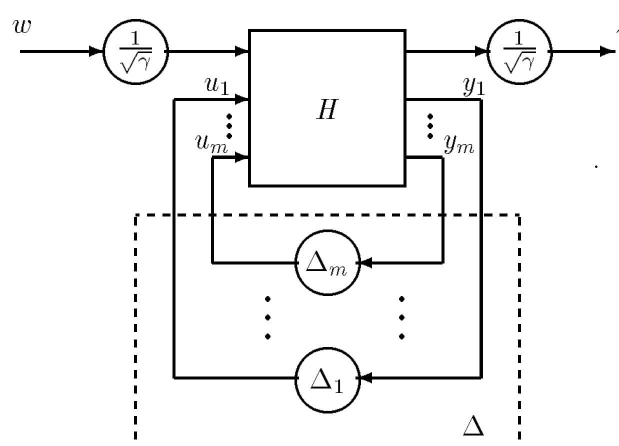

Figure 2: Uncertain linear system with the impulse response from w to z scaled by $1/\gamma$.

nonnegative to derive bounds on L_{wc} based on the bounds for the entries of M. We start with the following fact.

Fact 2 *The spectral radius of a nonnegative matrix is a nondecreasing function of its entries.*

(See Corollary 8.1.19 on page 491 of [3].)

Fact 2 implies that $\rho(D_\gamma P D_\gamma)$ is a nondecreasing function of the entries of the nonnegative matrix P and a nonincreasing function of $\gamma > 0$. These, in turn, mean that the function $\Phi(P)$ of a nonnegative matrix P defined by

$$\Phi(P) = \sup\{\gamma \mid \rho\left(D_\gamma P D_\gamma\right) \geq 1\}$$

is nondecreasing with the entries of P. We then have the following bounds for L_{wc}:

Theorem 1 *Let α_{ij}^N and β_{ij}^N be lower and upper bounds for $\|h_{ij}\|_1$ computed via equation (5) for some $N > 0$. Let M_{lb}^N and M_{ub}^N be matrices with (i,j)-entry α_{ij}^N and β_{ij}^N respectively ($i,j = 0, 1, \ldots, m$). Then*

$$L_{\mathrm{lb}}^N = \Phi(M_{\mathrm{lb}}^N) = \sup\{\gamma \mid \rho\left(D_\gamma M_{\mathrm{lb}}^N D_\gamma\right) \geq 1\},$$

and

$$L_{\mathrm{ub}}^N = \Phi(M_{\mathrm{ub}}^N) = \sup\{\gamma \mid \rho\left(D_\gamma M_{\mathrm{ub}}^N D_\gamma\right) \geq 1\},$$

are lower and upper bounds respectively for L_{wc}, i.e. $L_{\mathrm{lb}}^N \leq L_{\mathrm{wc}} \leq L_{\mathrm{ub}}^N$.

Computation of L_{lb}^N and L_{ub}^N is straightforward, once we make the following observation:

Fact 3 *The spectral radius of a nonnegative matrix is also an eigenvalue.*

(See Theorem 8.3.1 on page 503 of [3].)

Given a $(m+1) \times (m+1)$ matrix P, we first partition conformally as with M in equation (7) as

$$P = \begin{bmatrix} P^{(11)} & P^{(12)} \\ P^{(21)} & P^{(22)} \end{bmatrix}.$$

Then, $\Phi(P) = \infty$ if $\rho(P^{(22)}) \geq 1$. Otherwise, we note that $\rho\left(D_\gamma P D_\gamma\right) = \rho\left(D_\gamma^2 P\right)$, and solve for $D_\gamma^2 P x = x$ for some nonzero $(m+1)$-vector x to obtain

$$\Phi(P) = P^{(11)} + P^{(12)}(I - P^{(22)})^{-1} P^{(21)}.$$

The above formula shows that if $\rho(P^{(22)}) < 1$, $\Phi(P)$ is just the unique solution to the equation $\rho(D_\gamma P D_\gamma) = 1$.

4 Conclusion

We have presented simple bounds on the ℓ^∞-gain of single-input single-output linear discrete time systems. We have shown how to combine these bounds with recent results from [5] to compute guaranteed bounds for the worst-case ℓ^∞ gain of discrete-time LTI systems with diagonal uncertainty. The bounds may be easily extended to block diagonal uncertainties as well as to continuous time systems.

References

[1] S. Boyd and V. Balakrishnan, A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its $\mathbf{L}_\infty$-norm, *Systems Control Lett.* **15** (1990) 1–7.

[2] S. Boyd and J. Doyle, Comparison of peak and RMS gains for discrete-time systems, *Systems Control Lett.* **9** (1987) 1–6.

[3] R. A. Horn and C. A. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, United Kingdom, 1985).

[4] T. Kailath, *Linear Systems* (Prentice-Hall, New Jersey, 1980).

[5] M. Khammash and J. B. Pearson, Performance robustness of discrete-time systems with structured uncertainty, *IEEE Trans. Automat. Control* **36** (4) (1991) 398–412.

[6] M. Khammash and J. B. Pearson, Robustness synthesis for discrete-time systems with structured uncertainty, *Proceedings of the 1991 American Control Conference*, (1991) 2720–2724.

[7] K. Ogata, *Discrete-Time Control Systems* (Prentice-Hall, Englewood Cliffs, New Jersey, 1987).