Faculty of Mathematical Sciences

University of Twente

University for Technical and Social Sciences

P.O. Box 217
7500 AE Enschede
The Netherlands
Phone: +31-53-4893400
Fax: +31-53-4893114
Email: memo@math.utwente.nl

Memorandum No. 1459

Inverse problems in 2-norm and infinity-norm controller synthesis

A.M. Holohan and G. Meinsma

August 1998

ISSN 0169-2690

Inverse Problems in 2-norm and infinity-norm Controller Synthesis.

Anthony M. Holohan∗[1] **and Gjerrit Meinsma**†

∗ School of Electronic Engineering, Dublin City University,
Dublin 9, Rep. of Ireland.

† Department of Systems, Signals and Control,
Faculty of Mathematical Sciences, University of Twente,
P.O. Box 217, 7500 AE, Enschede, The Netherlands.

Abstract

This paper studies certain inverse problems in the optimal frequency-domain synthesis of robust controllers, in both the 2-norm and the infinity-norm. These inverse problems identify the class of controllers which are optimal for some choice of weights. Their implications for loopshaping are discussed.

Keywords:

Inverse optimality, loopshaping, H_2 control, H_∞ control.

1991 Mathematics Subject Classification:

49N05, 49N50, 93B36, 93B52, 93D09, 93D25.

Contact Information:

∗Phone: +353-1-7045107, Fax: +353-1-7045508, E-mail: holohana@eeng.dcu.ie
†Phone: +31-53-4893451, Fax: +31-53-4340733, E-mail: g.meinsma@math.utwente.nl

[1]This work was completed while the first author was on sabbatical leave to the University of Twente.

1 Introduction

We begin by setting up the problem formulation, and giving some of our notation and terminology.

1.1 Robust Controller Synthesis - "Loopshaping"

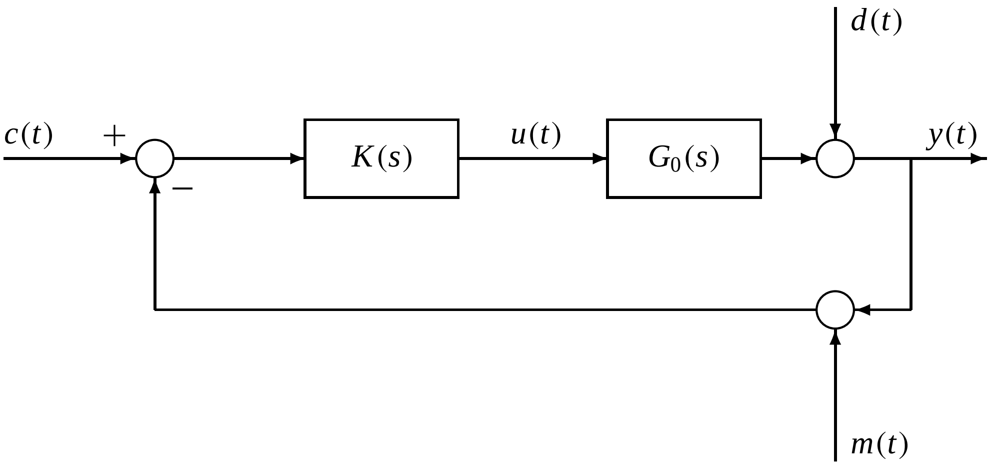

Figure 1. The standard feedback control loop.

Consider the familiar setup of Figure 1. The plant to be controlled will be denoted by $G_0(s)$. The feedback controller is $K(s)$. The loop gain is

$$L(s) = G_0(s)K(s)$$

The sensitivity function is

$$S(s) = \frac{1}{1+L(s)} = \frac{y(s)}{d(s)}$$

This is the transfer function from the disturbance input $d(s)$ to the plant's output $y(s)$. As is well known [1], a fundamental design objective is to obtain a $|S(j\omega)|$ which is sufficiently small. This objective is required for (i) sensitivity reduction, and (ii) disturbance attenuation. The complementary sensitivity function is

$$T(s) = \frac{L(s)}{1+L(s)} = -\frac{y(s)}{m(s)}$$

which is the transfer function from the measurement noise input $m(s)$ to minus the plant's output $-y(s)$. Another fundamental design objective is to obtain a $|T(j\omega)|$ which is sufficiently small. This objective is required for (i) robustness of stability, and (ii) measurement noise attenuation.

As is well known [1], the fact that

$$S + T = 1 \;\;\Rightarrow\;\; |S(j\omega)| + |T(j\omega)| \geq 1 \tag{1}$$

constitutes a fundamental limitation to what feedback can do. It means that there is an unavoidable tradeoff between keeping both $|S(j\omega)|$ and $|T(j\omega)|$ small. Moreover, another fundamental limitation is that the closed loop system be stable. This requirement means that the functions $S(s)$ and $T(s)$ must be analytic in the closed right half plane (CRHP) and must obey certain interpolation constraints [2]. Hence, it is usually the case that equality in eqn. (1) cannot be achieved at every frequency. This necessitates trading off one frequency against another.

A popular approach to practical frequency-domain design is to think of these objectives and limitations on a frequency-by-frequency basis, and to give $|S(j\omega)|$ and $|T(j\omega)|$ desirable "shapes", an approach nowadays termed "loopshaping". This is not always an easy task. Indeed, the two fundamental limitations make loopshaping a deep and subtle problem.

One approach to loopshaping is to use mathematical optimization to find a suitable controller. The following such optimal synthesis problems will be considered in this paper. Let $\mathcal{K}$ denote the set of all linear time-invariant (LTI) controllers that stabilize the given plant. We will deal with the 1-block 2-norm problems,

$$\inf_{K \in \mathcal{K}} \| W_1 S \|_2$$

$$\inf_{K \in \mathcal{K}} \| W_2 T \|_2$$

with the 1-block infinity-norm problems,

$$\inf_{K \in \mathcal{K}} \| W_1 S \|_\infty$$

$$\inf_{K \in \mathcal{K}} \| W_2 T \|_\infty$$

with the 2-block 2-norm problem,

$$\inf_{K \in \mathcal{K}} \| \; |W_1 S|^2 + |W_2 T|^2 \; \|_2$$

and with the 2-block infinity-norm problem,

$$\inf_{K \in \mathcal{K}} \| \; |W_1 S|^2 + |W_2 T|^2 \; \|_\infty$$

This paper studies the inverse optimality question for the above problems. That is, given a plant and a specific stabilizing controller, does there exist weights (W_1, W_2) for which the given controller is the optimal solution of one of these optimal synthesis problems. There is already a literature on these problems [3-5], and certain related problems [7], and we refer to these papers in more detail later. The remainder of this section gives some further notation which will be needed, and it states our assumptions. Section 2 settles the inverse problems for the 1-block cases. Section 3 describes some results from the literature. Section 4 gives some preliminary observations on the key role played by so-called positive real-axis crossovers. Section 5 contain the paper's main results, giving the solution to the 2-block 2-norm and infinity-norm inverse problems. Section 6 gives an example and Section 7 discusses the results. Finally, proofs are given in the appendix.

Some further notation will be needed. The dependence on the Laplace transform variable s is generally suppressed in our notation. For a given transfer function, $A(s)$ say, $P(A)$ will denote the number of open right half plane (ORHP) poles of A. Similarly, $Z(A)$ will denote the number of ORHP zeros of A. The clockwise winding number about the origin of the Nyquist curve of A will be denoted by $\mathrm{wno}(A)$. Given any stable biproper transfer function $A(s)$ which has no poles or zeros on the imaginary axis, then one can extract its ORHP zeros,

$$A(s) = \frac{a(-s)}{a(s)} A_{op}(s)$$

where the roots of the polynomial $a(-s)$ are precisely the ORHP zeros of $A(s)$. Then $A_{op}(s)$ is called the outer part or outer factor of $A(s)$, and $a(-s)/a(s)$ is called its inner part. This factorization is called an inner/outer factorization. We now adopt the notational convention that L, S and T inherit the subscript of the controller, that is, L_0, S_0, and T_0 are the loop gain, sensitivity function, and complementary sensitivity function, respectively, given by K_0, and similarly L_1, S_1 and T_1 correspond to controller K_1.

We now state our assumptions. Throughout this paper, the plant $G_0(s)$ and the controller $K_0(s)$ are viewed as fixed and given.

A1 $G_0(s)$ and $K_0(s)$ are real-rational transfer functions. So they represent finite dimensional LTI systems.

A2 K_0 stabilizes G_0.

A3 G_0 and K_0 are SISO.

A4 All weights are real-rational, stable and minimum phase. Weights are not identically zero. All H_∞ weights are biproper. All H_2 weights are strictly proper.

The above assumptions will be applied throughout this paper. For some results, we need in addition one or other of the following.

A5 Both G_0 and K_0 have no imaginary axis poles or zeros, infinity included. In particular, they are therefore biproper.

A6 The Nyquist curve of G_0K_0 does not meet the positive real-axis $(0, \infty)$ tangentially. That is, if $G_0K_0(j\omega_c) \in (0, \infty)$ at some frequency $s = j\omega_c$ then the slope of the Nyquist curve is not zero at this frequency.

Note that A5 is a very strong assumption. Removing it, where it is used, is the obvious next step in this line of research.

2 Inverse Optimality for One-Block Problems

This section treats inverse optimality for one-block problems.

Theorem 1 *Suppose that assumptions A1 to A5 are obeyed, and that $\mathcal{Z}(G_0) > 0$. Then, the following are equivalent.*

$$
\begin{array}{ll}
(a) & \exists\ W_1\ \ s.t.\ \ K_0 = \arg\inf_{K\in\mathcal{K}}\ \|\ W_1 S\ \|_\infty \\
(b) & \exists\ W_1\ \ s.t.\ \ K_0 = \arg\inf_{K\in\mathcal{K}}\ \|\ W_1 S\ \|_2 \\
(c) & \nexists\ K_1 \in \mathcal{K}\ \ s.t.\ \ |S_1(j\omega)| < |S_0(j\omega)|\ \ \forall\ \omega \in \mathbb{R}\cup\{\infty\} \\
(d) & \mathcal{P}(K_0) < \mathcal{Z}(G_0)
\end{array}
$$

Part (a) asks if K_0 is the infinity-norm optimal controller for some choice of weight. Part (b) asks if K_0 is the 2-norm optimal controller for some weight. Part (c) asks if $|S(j\omega)|$ can be strictly decreased at every frequency $s = j\omega$ (including infinity) by changing the controller. If so, K_0 would be a poor controller from the point of view of the sensitivity function. Part (d) gives the elegant answer to these questions.

The situation with the complementary sensitivity function is similar.

Theorem 2 *Suppose that assumptions A1 to A5 are obeyed, and that $\mathcal{P}(G_0) > 0$. Then, the following are equivalent.*

$$
\begin{array}{ll}
(a) & \exists\ W_2\ \ s.t.\ \ K_0 = \arg\inf_{K\in\mathcal{K}}\ \|\ W_2 T\ \|_\infty \\
(b) & \exists\ W_2\ \ s.t.\ \ K_0 = \arg\inf_{K\in\mathcal{K}}\ \|\ W_2 T\ \|_2 \\
(c) & \nexists\ K_1 \in \mathcal{K}\ \ s.t.\ \ |T_1(j\omega)| < |T_0(j\omega)|\ \ \forall\ \omega \in \mathbb{R}\cup\{\infty\} \\
(d) & \mathcal{Z}(K_0) < \mathcal{P}(G_0)
\end{array}
$$

The proofs are given in the appendix. We remark that these proofs are constructive in both directions. Thus, when K_0 is 2-norm (and so infinity-norm) optimal, the proofs allow the weights (in both the 2-norm and the infinity-norm cases) to be explicitely identified. In the other direction, when K_0 is not 2-norm or infinity-norm optimal, a simple formula explicitely identifies one $K_1 \in \mathcal{K}$ that strictly decreases $|S(j\omega)|$ (or $|T(j\omega)|$) at every frequency.

3 S&T-Optimality

We next describe briefly some results from the literature. One approach to the one-block infinity-norm inverse optimality problems is based on the following definitions [3-5].

Definition 1 *The controller K_0 is said to be S-optimal if*

$$K_0 = \arg \inf_{K\in K}\ \|\ W_1 S\ \|_\infty$$

for the weight $W_1 = [S_0]_{op}^{-1}$.

Definition 2 *The controller K_0 is said to be T-optimal if*

$$K_0 = \arg \inf_{K \in K} \| W_2 T \|_\infty$$

for the weight $W_2 = [T_0]_{op}^{-1}$.

The problem of finding conditions for a controller to be S-optimal and/or T-optimal was settled in [3-5]. In essence, this gives the equivalence between parts (a) and (d) in the previous two theorems, since the well known all-pass property of H_∞ optimal controllers means that there is only one candidate weight (up to rescaling by a constant) for which K_0 might be infinity-norm optimal. So when K_0 is infinity-norm optimal, the corresponding weight is essentially unique. We remark that when K_0 is 2-norm optimal, the corresponding weight may be unique or may be non-unique. However, the proof given in the appendix allows all such weights to be characterized.

Excepting parts (b), the above two theorems are a minor extension of the results in [3-5]. The proof in [3] was subsequently improved [4-5] and simplified by using a theorem due to Poreda [6]. Another different proof is given in the appendix. This proof is a further simplification, and uses only elementary arguments.

In [3-5], the 2-block infinity-norm inverse question is approached via the following definition.

Definition 3 *Let $W_1 = [S_0]_{op}^{-1}$ and $W_2 = [T_0]_{op}^{-1}$. Then K_0 is said to be S&T-optimal if*

$$K_0 = \arg \inf_{K \in K} \| \ |W_1 S|^2 + |W_2 T|^2 \|_\infty$$

Elegant conditions for S&T-optimality have been found [3]. The result is as follows.

Theorem 3 (Lenz et al.) *Suppose that assumptions A1 to A5 are obeyed. Then,*

K_0 *is S&T-optimal*

if and only if

$$P(K_0) + Z(K_0) < wno(T_0 - S_0)$$

Note that in this definition, the weights are chosen in a special way. Hence, it cannot be assumed that this result settles the "full" inverse problem, i.e. the 2-block infinity-norm inverse problem with arbitrary weights which obey A4. The difficulty is that although the above choice of weights makes $|W_1 S|^2 + |W_2 T|^2$ all-pass, there may be many other weights which do so too.

The above theory has been extended by Lenz [4,5], as follows.

Theorem 4 (Lenz) *Suppose that assumptions A1 to A5 are obeyed. Let* $W_1 = \alpha[S_0]_{op}^{-1}$ *and* $W_2 = \beta[T_0]_{op}^{-1}$ *where* α *and* β *are real numbers, not both zero. Then*

$$K_0 = \arg \inf_{K \in K} \| \ |W_1 S|^2 + |W_2 T|^2 \ \|_\infty$$

if and only if

$$P(K_0) + Z(K_0) < wno(\beta T_0 - \alpha S_0)$$

We will see later that this seemingly minor extension of the concept of S&T-optimality turns out to be interesting. We will say that K_0 is S&T-optimal for some (α, β) if the conditions of the above theorem hold for K_0 for some value of (α, β).

Feng and Smith [7] have studied a different inverse optimality question. They solved the inverse optimality question for McFarlane-Glover loopshaping [8], in the SISO case. Although this is a 2-block infinity-norm problem, it has a strong additional restriction on the weights. Specifically, when A5 holds, they require (W_1, W_2) to obey

$$W_1^* W_1 W_2^* W_2 = W_1^* W_1 + W_2^* W_2$$

So the class of 2-block infinity-norm problems considered in this paper is broader.

4 The Role of $\mathbb{R}_+$ Crossovers

Consider the Nyquist curve of the loop gain on the standard Nyquist D-contour, with small semi-circular indentations into the RHP around the finite imaginary axis poles and zeros. Let Γ denote this contour. Note that $L_0(s)$ is never zero and never infinity on this contour. The values of s for which

$$L_0(s) \ \in \ [0, \infty], \quad s \in \Gamma$$

will be called positive real-axis crossovers, or $\mathbb{R}_+$ crossovers. These are the points where the Nyquist curve intersects the positive real-axis. We will see shortly that these points play a central role in inverse optimality.

We introduce a classification of such points. A crossover $s = j\omega_c$ will be called a finite real-axis crossover if

$$L_0(j\omega_c) \ \in \ (0, \infty), \quad j\omega_c \in \Gamma$$

so that $L_0(j\omega_c)$ is neither zero nor infinity, and the point $s = j\omega_c$ lies on the imaginary axis and $\omega_c \neq \infty$. For obvious reasons, a finite crossover $j\omega_c$ will be called a downwards crossover if the derivative of the loop gain's phase, viz.

$$\left. \frac{\mathrm{d}}{\mathrm{d}\omega} \measuredangle L_0(j\omega) \right|_{\omega=\omega_c}$$

is strictly negative. It will be called an upwards crossover if this quantity is strictly positive. Let d_1 and u_1 denote the number of downwards and upwards finite real-axis crossovers respectively. Thus, these numbers do not count any points where the Nyquist curve of $L_0(j\omega)$ touches the positive real-axis tangentially, i.e. when the loop gain's phase and its derivative are both zero. Such points will be referred to as tangential touchings.

There may also be $\mathbb{R}_+$ crossovers due to the small semi-circular indentations in the contour about the finite imaginary axis poles and zeros of the loop gain. Recall that if L_0 has a pole of multiplicity n at $s = jp$, the small semi-circular part of the contour Γ around this pole is mapped under L_0 to n clockwise semi-circles of large radius centered about the origin. These may result in $\mathbb{R}_+$ crossovers. Note that they always cross the positive real axis in the downwards (clockwise) direction. Let d_2 be the number of such crossovers. Note that each single imaginary axis pole will contribute either zero or one such crossover, and that a double pole will contribute either zero, one or two such crossovers, depending on the loop gain's phase to the left of jp (i.e. $\measuredangle L_0(jp - j\epsilon)$). Similarly, if L_0 has a zero of multiplicity n at $s = jz$, the small semi-circular part of the contour Γ around this zero is mapped under L_0 to n counter-clockwise semi-circles of small radius. These may give rise to $\mathbb{R}_+$ crossovers, and, if so, they necessarily cross in the upwards (counter-clockwise) direction. Let u_2 be the total number of such crossovers.

The large semi-circular part of the contour Γ can also contribute $\mathbb{R}_+$ crossovers. If L_0 is biproper and $L_0(\infty) > 0$, then $L_0(\infty)$ gives a finite real-axis crossover which is either an upwards or a downwards crossover. If it is an upwards crossover, let $u_3 = 1$ and $d_3 = 0$, while if it is a downwards crossover, let $u_3 = 0$ and $d_3 = 1$. If L_0 is strictly proper, let $u_3 = 0 = d_3$. When L_0 is strictly proper, say with relative degree $n > 0$, the large semi-circular part of the contour gets mapped to n small semi-circles around the origin. These may result in crossovers close to the origin in the upwards (counterclockwise) direction. Let u_4 be the total number of such crossovers. Next, let

$$U(L_0) = u_1 - d_2 + u_3 \tag{2}$$

$$D(L_0) = d_1 - u_2 + d_3 - u_4 \tag{3}$$

Note that every $\mathbb{R}_+$ crossover appears in either $U(L_0)$ or $D(L_0)$, that the former gets the contributions from the finite upwards crossovers and those of arbitrarily large radius, and that the latter gets the contributions from the finite downwards crossovers and those of arbitrarily small radius.

The key role played by $\mathbb{R}_+$ crossovers comes from two easy observations. Firstly, the fact that

$$S_0 + T_0 = 1$$

immediately implies that

$$|S_0(j\omega)| + |T_0(j\omega)| \geq 1$$

Further, equality holds if and only if the loop gain is real and non-negative,

$$|S_0(j\omega)| + |T_0(j\omega)| = 1 \quad \Leftrightarrow \quad L_0(j\omega) \;\in\; [0, \infty]$$

Consider (any) one fixed frequency $s = j\omega_c$ for the moment. It is easily checked that the above observation shows that both $|S_0(j\omega_c)|$ and $|T_0(j\omega_c)|$ cannot be decreased simultaneously by changing the value of $K_0(j\omega_c)$ if and only if $L_0(j\omega_c) \in [0, \infty]$, i.e. if and only if $j\omega_c$ is a finite $\mathbb{R}_+$ crossover or is a pole, zero or tangential touching of L_0.

Secondly, considering again (any) one fixed frequency $s = j\omega_c$, it is easily verified that

$$\arg \min_{L(j\omega_c)} |W_1 S(j\omega_c)|^2 + |W_2 T(j\omega_c)|^2 \;=\; \left|\frac{W_1(j\omega_c)}{W_2(j\omega_c)}\right|^2$$

Clearly, this can be achieved by K_0 for some choice of infinity-norm weights obeying A4 if and only if $L_0(j\omega) \;\in\; (0, \infty)$, i.e. if and only if $j\omega_c$ is a finite $\mathbb{R}_+$ crossover (which is non-zero and non-infinite). It follows easily that if

$$L(j\omega_c) \;=\; \left|\frac{W_1(j\omega_c)}{W_2(j\omega_c)}\right|^2 \;\in\; [0, \infty] \tag{4}$$

then

$$|W_1 S(j\omega_c)|^2 + |W_2 T(j\omega_c)|^2$$

cannot be decreased at this $s = j\omega_c$, and conversely.

As shown in the appendix, the consequence is that a given controller K_0 can be H^∞ optimal because of a single point $s = j\omega_c$. It may or may not be true that there exists another stabilizing controller which decreases

$$|W_1 S(j\omega)|^2 + |W_2 T(j\omega)|^2$$

almost everywhere (a.e.), specifically, at all frequencies which are not $\mathbb{R}_+$ crossovers.

We need to dismiss some pathological cases. If L_0 is a negative constant, then it is trivial to verify that $K = -K_0$ is a stabilizing controller which reduces both $|S(j\omega)|$ and $|T(j\omega)|$ at all frequencies. So L_0 is neither 2-norm nor infinity-norm optimal. If L_0 is identically zero or a positive constant, the above observations show that equality in eqn. (1) holds at every frequency. It follows directly that L_0 is both 2-norm and infinity-norm optimal. It will prove convenient to accept the following convention in terminology. If L_0 is constant and non-negative, we take it that $U(L_0) = 1 = D(L_0)$, while if it is constant and negative, we take it that $U(L_0) = 0 = D(L_0)$, and A6 is not viewed as being violated. This terminology allows subsequent results to be stated more concisely, avoiding the need to mention this special case.

Assuming that $L_0(j\omega)$ is non-constant but real-valued everywhere implies that $L_0(s) = L_0(-s)$, so that L_0 is a non-constant real-rational function of s^2. This contradicts the stability assumption. So $L_0(j\omega)$ cannot be real-valued everywhere. Therefore its $\mathbb{R}_+$ crossovers are well defined. Thus, $U(L_0)$ and $D(L_0)$ are well defined, and in particular, they are finite integers.

5 Inverse Optimality for Two-Block Problems

We now turn our attention to inverse optimality for two-block problems, beginning with the 2-norm case.

Theorem 5 *Suppose that assumptions A1 to A4 and A6 are obeyed. Then, (a), (c) and (d) are equivalent, and they imply (b),*

$$
\begin{array}{ll}
(a) & \exists\, W_1, W_2 \;\; s.t. \;\; K_0 = \arg\inf_{K \in \mathcal{K}} \; \| \;\; |W_1 S|^2 + |W_2 T|^2 \; \|_2 \\
(b) & \nexists\, K_1 \in \mathcal{K} \;\; s.t. \; |S_1(j\omega)| \; < \; |S_0(j\omega)| \;\; and \\
 & |T_1(j\omega)| \; < \; |T_0(j\omega)| \;\; a.e. \; on \; j\mathbb{R} \cup \{\infty\} \\
(c) & \mathcal{D}(L_0) \; > \; \mathcal{Z}(K_0) - \mathcal{P}(G_0) \\
(d) & \mathcal{U}(L_0) \; > \; \mathcal{P}(K_0) - \mathcal{Z}(G_0)
\end{array}
$$

If, in addition, A5 is obeyed, then (a), (b), (c) and (d) are equivalent.

Part (a) deals with whether or not K_0 is the optimal solution of some 2-block 2-norm problem. Part (b) deals with whether or not both $|S(j\omega)|$ and $|T(j\omega)|$ can be decreased almost everywhere, i.e. for all but finitely many frequencies. Parts (c) and (d) give easily tested equivalent conditions. These involve the $\mathbb{R}_+$ crossovers of the Nyquist curve of $L_0(j\omega)$ and certain ORHP pole and zero counts.

This theorem solves the 2-block 2-norm inverse optimality question. Moreover, given A5, it says that a controller is *not* 2-norm optimal for any weights if and only if its $|S(j\omega)|$ and $|T(j\omega)|$ can be decreased almost everywhere. Clearly, such a controller would be a poor design.

It would be desirable to remove the need for assumption A5 for the (b) $\Rightarrow$ (a) part. This is the obvious next step in this line of research. The present author's have been unable to settle this part.

We now turn to the 2-block infinity-norm case.

Theorem 6 *Suppose that assumptions A1 to A4 are obeyed. Then, the following are equivalent.*

$$
\begin{array}{ll}
(a) & \exists\, W_1, W_2 \;\; s.t. \;\; K_0 = \arg\inf_{K \in \mathcal{K}} \; \| \;\; |W_1 S|^2 + |W_2 T|^2 \; \|_\infty \\
(b) & \nexists\, K_1 \in \mathcal{K} \;\; s.t. \; |S_1(j\omega)| \; < \; |S_0(j\omega)| \;\; and \\
 & |T_1(j\omega)| \; < \; |T_0(j\omega)| \;\; everywhere \; on \; j\mathbb{R} \cup \{\infty\} \\
(c) & \mathcal{P}(G_0) \; > \; \mathcal{Z}(K_0) \\
 & or \; L(j\omega_c) \in [0, \infty] \; for \; at \; least \; one \; \omega_c \in \mathbb{R} \cup \{\infty\} \\
(d) & \mathcal{Z}(G_0) \; > \; \mathcal{P}(K_0) \\
 & or \; L(j\omega_c) \in [0, \infty] \; for \; at \; least \; one \; \omega_c \in \mathbb{R} \cup \{\infty\}
\end{array}
$$

Part (a) is the infinity-norm inverse optimality question. Parts (a) and (b) show that a controller is not optimal for any 2-block infinity norm problem if and only if both $|S(j\omega)|$ and $|T(j\omega)|$ can be strictly decreased at every $s = j\omega$, including infinity. Parts (c) and (d) give testable conditions for this.

We note that there is a key difference between 2-norm and infinity-norm inverse optimality. These properties are intimately linked with, respectively, the questions of whether or not $|S(j\omega)|$ and $|T(j\omega)|$ can be decreased at almost every frequency, or at every frequency. The distinction between these questions is precisely the points where $L_0(j\omega_c) \in [0, \infty]$ for $\omega_c \in \mathbb{R} \cup \{\infty\}$. As discussed in Section 4, the presence of even one such point is enough to ensure that K_0 is infinity-norm optimal for some 2-block problem. Also, it is enough to ensure that $|S(j\omega)|$ and $|T(j\omega)|$ cannot both be decreased at some (such) frequencies. Given this observation, the only outstanding case is that of loop gains with no such points. But then A5 necessarily applies, and $U(L_0) = 0 = D(L_0)$. Because of this strong condition, the above theorem can be stated in quite a number of ways.

Theorem 7 *Suppose that assumptions A1 to A4 are obeyed. Then, (a), (e) and (f) are equivalent. If, in addition, A5 is obeyed. Then (a), (e), (f) and (g) are equivalent,*

$$
\begin{array}{ll}
(a) & \exists\, W_1, W_2 \;\; s.t. \;\; K_0 = \arg\inf_{K \in K} \; \| \; |W_1 S|^2 + |W_2 T|^2 \; \|_\infty \\
(e) & K_0 \textit{ is 2-norm optimal} \\
 & \textit{or } L(j\omega_c) \in [0, \infty] \textit{ for at least one } \omega_c \in \mathbb{R} \cup \{\infty\} \\
(f) & K_0 \textit{ is S-optimal} \\
 & \textit{or } K_0 \textit{ is T-optimal} \\
 & \textit{or } L(j\omega_c) \in [0, \infty] \textit{ for at least one } \omega_c \in \mathbb{R} \cup \{\infty\} \\
(g) & K_0 \textit{ is S\&T-optimal with } (\alpha, \beta)
\end{array}
$$

Part (e) emphasises the distinction between 2-norm and infinity-norm inverse optimality. Part (f) relates 2-block infinity-norm optimality to 1-block infinity-norm optimality, and part (g) relates it to the concept of S&T-optimality.

Two distinct notions of S&T-optimality were cited in Section 3. If the weights used in the definition of S&T-optimality are to obey A4, then A5 must be obeyed. When A5 does hold, then although the weaker version (with $\alpha = 1 = \beta$) is not equivalent to inverse infinity-norm optimality, the stronger version (with (α, β)) is.

6 Example

Suppose that the plant and the given controller are

$$G_0 = 1, \quad K_0 = \frac{-s+1}{2s+1}$$

The loop gain has only one $\mathbb{R}_+$ crossover. It occurs at $s = j0$. Choose

$$W_1 = [S_0]_{op}^{-1}, \quad W_2 = [T_0]_{op}^{-1}$$

Then,

$$|W_1 S_0(j\omega)|^2 + |W_2 T_0(j\omega)|^2 = 2 \;\; \forall \, \omega \in \mathbb{R} \cup \{\infty\}$$

No other controller can reduce this quantity at $s = j0$. Therefore, no other controller can reduce the infinity-norm of the above quantity. It follows that K_0 is infinity-norm optimal for the weights given. It is not 2-norm optimal. Indeed, as is trivially verified, the controller

$$K_1 = \frac{1}{s+1}$$

simultaneously reduces the modulus of both S and T at every $s = j\omega$ except $s = j0$. It is clear that K_0 is therefore a poor choice of controller. Thus, we have the (perhaps surprising) situation where K_0 is infinity-norm optimal, but it is a poor controller.

This example illustrates a number of points. First, infinity-norm optimality is not always a reliable indicator that the controller is a reasonable design. Second, there are certain choices of weights which, in infinity-norm loopshaping, are problematic. Third, it seems that dependable H_∞ loopshaping software should explicitely check for the condition of eqn. (4). Fourth, this poses the question of finding a definitive set of weight selection rules which ensure a good outcome. Fifth, it is tempting to suggest that 2-norm loopshaping may be more satisfactory in practice. Finally, the example suggests that it would be desirable for H_∞ loopshaping software to have the property that the controller returned is always 2-norm optimal for some problem.

7 Why Solve Inverse Problems?

Why solve inverse problems? Apart from the pursuit of knowledge for its own sake, the authors believe that there are convincing practical reasons for solving the inverse problems treated in this paper.

Let us agree to some terminology.

Definition 4 *A controller K will be called* reasonable *if its corresponding sensitivity and complementary sensitivity functions S and T cannot both be decreased at almost every frequency ω.*

The authors contend that every design procedure should ideally have two properties, as follows.

(a) The outcome of the design process (i.e. K_0) should always be reasonable.

(b) Every reasonable controller should be a possible outcome (i.e. be produced by some (W_1, W_2)).

The above analysis shows that H_∞ optimization does not always give a reasonable result. On the other hand, H_2 optimization always gives a reasonable result, and can in principle find all reasonable controllers. Clearly, this observation has implications for controller design. It also raises the issue of how H_∞ design software deals with (or should deal with) problems having positive real-axis crossovers and with weights which make these points tight. This may help to make H_∞ loopshaping software easier to use, and may reduce the number of design iterations needed.

The answer to an inverse problem depends on the rules for weight selection. Clearly, it would be desirable to know the complete, minimum set of rules for weight selection in H_∞ synthesis which ensure that the result is always reasonable.

All in all, the author's believe that solving such inverse problem's improves the theoretical foundations of loopshaping.

8 Appendix

This appendix is devoted to proving the theorems presented above.

First, some further notation will be needed. For a complex function, $f^*(s)$ denotes $\overline{f(-\bar{s})}$. Consequently, for any polynomial or rational function with real coefficients, $X^*(s) = X(-s)$. On the imaginary axis, $X^*(s)$ is the complex conjugate of $X(s)$. The degree of a polynomial $x(s)$ will be denoted by $\delta(x)$. The degrees of the numerator and denominator of a rational function X will be denoted by $\delta_n(X)$ and $\delta_d(X)$ respectively. The relative degree of X, denoted by $\delta_r(X)$, is defined to be

$$\delta_r(X) = \delta_d(X) - \delta_r(X)$$

Recall that for a given transfer function, A say, $\mathcal{P}(A)$ and $\mathcal{Z}(A)$ denote the number of ORHP poles and zeros of A respectively. Then, $\mathcal{P}(A^*)$ and $\mathcal{Z}(A^*)$ denote the number of open left half plane (OLHP) poles and zeros of A, repectively. The clockwise winding number about the points 0 and 1 of the Nyquist curve of A are denoted by $\text{wno}(A)$ and $\text{wno}(A, +1)$ respectively. Recall that the Principle of the Argument [12] says that

$$\text{wno}(A) = \mathcal{Z}(A) - \mathcal{P}(A)$$

We note that the engineering convention for the direction by which the Nyquist contour is traversed is different from the mathematics convention. Note also that, recalling the definitions of Section 4,

$$\text{wno}(L_0) = \mathcal{D}(L_0) - \mathcal{U}(L_0)$$

As is usual, H_∞ denotes the set of all stable transfer functions and RH_∞ denotes the set of all real-rational stable transfer functions.

Extensive use will be made of the Youla parameterization [9,10] of all stabilizing controllers. We use the rendition of Desoer *et al.* [11]. The latter treatment uses rational factorizations, while the former uses polynomial factorizations. We outline very briefly the Youla parameterization for the SISO case and our corresponding notation.

Any SISO rational transfer function G_0 can be written as the ratio of two transfer functions, $G_0 = N/D$ where both N and D are stable transfer functions which have no CRHP zeros in common. Such a factorization is said to be a stable coprime factorization.

Theorem 8 (Youla Parameterization) *Let $G_0 = N/D$ be a stable coprime factorization of G_0. Let K_0 be a stabilizing feedback controller. Then there exists stable transfer functions U and V such that*

$$NU + DV = 1 \tag{5}$$

and

$$K_0 = \frac{U}{V} \tag{6}$$

Furthermore, the set of all rational LTI controllers that stabilize the closed loop system is described by

$$K = \frac{U - QD}{V + QN} \tag{7}$$

as Q ranges through RH_∞, and the set of all LTI controllers that stabilize the closed loop system is also described by eqn (7) as Q ranges through H_∞.

This fundamental result describes precisely the set of *all* LTI stabilizing controllers for any given rational plant. This set is described in terms of the parameter Q, called here the Youla parameter. Eqn. (5) is called the Bezout identity. Simple algebra then confirms that for any Q, the corresponding sensitivity function and complementary sensitivity function are

$$S(Q) = DNQ + DV, \quad T(Q) = -NDQ + NU \tag{8}$$

respectively. The "given" controller K_0 then corresponds to $Q = 0$. We adopt the convention that Q_0 and Q_1 correspond respectively to the controllers K_0 and K_1, and then $Q_0 = 0$. The loop gain, the sensitivity function and the complementary sensitivity function which correspond to K_0 are then

$$L_0 = G_0 K_0, \quad S_0 = DV, \quad T_0 = NU$$

We will need the following inner/outer factorizations.

$$N = \frac{n^*}{n} N_{op}, \quad D = \frac{d^*}{d} D_{op}$$

$$U = \frac{u^*}{u} U_{op}, \quad V = \frac{v^*}{v} V_{op}$$

That is, given N, let $n(-s) = n^*(s)$ be the monic polynomial whose roots are precisely the ORHP zeros of N, and similarly for the others.

Lemma 1 *Suppose that assumptions A1 to A5 are obeyed. Then,*

$$\exists\ Q_1 \in H_\infty \text{ s.t. } |S_1(j\omega)| \ < \ |S_0(j\omega)| \ \forall\, \omega \in \mathbb{R} \cup \{\infty\}$$

if and only if

$$P(K_0) \geq Z(G_0)$$

For the proof, we will need the following classical theorem [12].

Theorem 9 (Rouché) *Suppose that the functions $f(s)$ and $g(s)$ are analytic inside and on the Jordan curve Γ. If*

$$|f(s) - g(s)| \ < \ |f(s)| \ \forall\, s \in \Gamma$$

then

$$f(s) \text{ and } g(s) \text{ have the same number of zeros inside } \Gamma$$

(and have no zeros on Γ).

Proof Suppose that there is some $Q_1 \in H_\infty$ which strictly reduces $|S(j\omega)|$ at every frequency including infinity. That is,

$$\exists\ Q_1 \in H_\infty \ \text{ s.t. } \ |S_1(j\omega)| \ < \ |S_0(j\omega)| \ \forall\, \omega \cup \{\infty\}$$

$$\Leftrightarrow \ |DNQ_1 + DV| < |DV| \ \forall\, \omega \cup \{\infty\}$$

$$\Leftrightarrow \ |Q_1 N + V| < |V| \ \forall\, \omega \cup \{\infty\}$$

Apply Rouche's Theorem,

$$\Rightarrow \ Z(V) = Z(QN)$$

$$\Leftrightarrow \ P(K_0) = Z(G_0) + Z(Q)$$

$$\Rightarrow \ P(K_0) \geq Z(G_0)$$

For the converse, suppose that $P(K_0) \geq Z(G_0)$. Then (and only then) n/v is proper. If it is biproper, let $a(s) = 1$. If it is strictly proper, let $a(s)$ be any strictly Hurwitz polynomial with degree $\delta(a) = \delta(v) - \delta(n)$. Consider

$$Q_1 = -\epsilon \left(\frac{n}{v}\right)^2 a^* a V_{op} N_{op}^{-1}$$

This particular choice of Q_1 is then in H_∞ as required. This Q_1 gives

$$\begin{aligned} |S_1(j\omega)| &= |DNQ_1 + DV| = |DV(Q_1 N V^{-1} + 1)| \\ &= |S_0(j\omega)| \left|1 - \epsilon \left(\frac{na}{v}\right)^* \left(\frac{na}{v}\right)\right| \end{aligned}$$

which clearly decreases $|S(j\omega)|$ everywhere, including infinity, for sufficiently small $\epsilon > 0$. ■

Lemma 2 *Suppose that both A and B are strictly proper, are stable, and are column vectors. Then,*

$$0 = \arg \inf_{Q \in H_\infty} \parallel AQ + B \parallel_2$$

if and only if $B^* A$ has no poles in the CRHP.

Proof Suppose that $B^* A$ has no poles in the CRHP. Then

$$\begin{aligned} \parallel AQ + B \parallel_2^2 &= \langle AQ + B, AQ + B \rangle \\ &= \parallel AQ \parallel_2^2 + \langle B, AQ \rangle + \langle AQ, B \rangle + \parallel B \parallel_2^2 \\ &= \parallel AQ \parallel_2^2 + \parallel B \parallel_2^2 + 2 \text{ Re} \int_{-\infty}^{+\infty} B^*(j\omega) A(j\omega) Q(j\omega) \mathrm{d}\omega \end{aligned}$$

and since $B^* AQ$ is analytic in the CRHP and is strictly proper, the above integral is zero, giving,

$$\begin{aligned} \parallel AQ + B \parallel_2^2 &= \parallel AQ \parallel_2^2 + \parallel B \parallel_2^2 \\ &\geq \parallel B \parallel_2 \end{aligned}$$

Therefore $Q = 0$ is optimal in this case.

Conversely, suppose that $B^* A$ has one or more unstable poles. Then BA^* has stable poles. It follows that Q_1 defined by

$$Q_1 = -\epsilon \pi_+ (A^* B)$$

is non-zero, where π_+ denotes projection onto H_2. Thus, to compute the optimal Q numerically, one can take the partial fraction expansion of $A^* B$ and let Q be minus the sum of the stable terms. The above Q_1 gives

$$\left\langle A^* B, Q_1 \right\rangle = -\epsilon \parallel \pi_+ (A^* B) \parallel_2^2 < 0$$

and so

$$\begin{aligned} \parallel AQ_1 + B \parallel_2^2 &= \parallel AQ_1 \parallel_2^2 + \langle B, AQ_1 \rangle + \langle AQ_1, B \rangle + \parallel B \parallel_2^2 \\ &= O(\epsilon^2) + \parallel B \parallel_2^2 - 2\epsilon \parallel \pi_+ (A^* B) \parallel_2^2 \\ &< \parallel B \parallel_2^2 \quad \text{for } \epsilon > 0 \text{ sufficiently small} \end{aligned}$$

So, if $B^* A$ has unstable poles, then $Q = 0$ is not optimal. ■

Lemma 3 *Suppose that assumptions A1 to A5 are obeyed. Then,*

$$\exists \, W_1 \text{ s.t. } K_0 = \arg \inf_{K \in K} \parallel W_1 S \parallel_2$$

if and only if

$$P(K_0) < Z(G_0)$$

Proof Using the Youla Parameterization,

$$W_1 S = W_1 D(QN + V) = W_1 DNQ + W_1 DV$$

From the previous lemma,

$$0 = \arg \inf_{Q \in H_\infty} \| W_1 S \|_2$$

$$\Leftrightarrow \ (W_1 DV)^*(W_1 DN) \text{ has no CRHP poles}$$

$$\Leftrightarrow \ (W_1 DV)^*(W_1 DN) \text{ has no ORHP poles}$$

When does this hold for some W_1 obeying A4? Clearly, it is necessary and sufficient that W_1 can be chosen to cancel all the ORHP poles of $(DV)^*(DN)$ while the ORHP poles of the factor W_1^* are cancelled by the ORHP zeros of $(DV)^*(DN)$. Counting the ORHP poles and zeros, this holds if and only if the number of ORHP poles to be cancelled is less than or equal to the number of ORHP zeros available to cancel them. That is,

$$P((W_1 DV)^*(W_1 DN)) \le Z((W_1 DV)^*(W_1 DN))$$

$$\Leftrightarrow \ \delta_d (W_1 DV)^* \le Z((W_1 DV)^*(W_1 DN))$$

$$\Leftrightarrow \ \delta_d (W_1 DV)^* \le \delta_n(W_1) + \delta_n(D) + Z(N) + \delta_n(V) - Z(V)$$

and since D and V are necessarily biproper, and W_1 must be strictly proper,

$$\Leftrightarrow \ Z(V) < Z(N)$$

$$\Leftrightarrow \ P(K_0) < Z(G_0)$$

as claimed. ∎

Lemma 4 *Suppose that assumptions A1 to A5 are obeyed. Then,*

$$\nexists \ W_1 \text{ s.t. } K_0 = \arg \inf_{K \in K} \| W_1 S \|_\infty$$

if and only if

$$\exists \ Q_1 \in H_\infty \text{ s.t. } |S_1(j\omega)| \ < |S_0(j\omega)| \ \forall \, \omega \cup \{\infty\}$$

Proof This equivalence is easy in the one-block case. Indeed, suppose that

$$\exists \ Q_1 \in H_\infty \text{ s.t. } |S_1(j\omega)| \ < |S_0(j\omega)| \ \forall \, \omega \cup \{\infty\}$$

If K_0 is H_∞ optimal for some W_1, then replacing $Q_0 = 0$ by Q_1 strictly reduces $|S(j\omega)|$ everywhere, which then reduces $\| W_1 S \|_\infty$ which is a contradiction. On the other hand, suppose that

$$\nexists \ W_1 \text{ s.t. } K_0 = \arg \inf_{K \in K} \| W_1 S \|_\infty$$

Choosing $W_1 = [S_0]_{op}^{-1}$ in particular gives

$$\| W_1 S_0 \|_\infty = |W_1 S_0(j\omega)| = 1 \ \ \forall\, \omega \cup \{\infty\}$$

and

$$K_0 \neq \arg\inf \| W_1 S \|_\infty$$

By the very meaning of infimum,

$$\exists\ Q_1 \in H_\infty \text{ s.t. } \ \| W_1 S_1 \|_\infty < \| W_1 S_0 \|_\infty = 1$$

and so

$$|W_1 S_1(j\omega)| \ \le \| W_1 S_1 \|_\infty < \ 1 = |W_1 S_0(j\omega)| \ \ \forall\ \ \omega \cup \{\infty\}$$

and the claim follows. Note that this proof relies on elementary arguments only, and does not depend on the (advanced) facts that H_∞ controllers always exist and are all-pass under our assumptions. ■

Remark Combining Lemma 3 with the contrapositives of Lemmas 1 and 4 establishes Theorem 1. The proof of Theorem 2 is almost identical. We turn now to Theorem 5, to the 2-block 2-norm case.

Lemma 5 *Suppose that assumptions A1 to A4 are obeyed. Then,*

$$\exists\ W_1, W_2 \ \ s.t. \ \ K_0 = \arg\inf_{K \in K} \left\| \begin{pmatrix} W_1 S \\ W_2 T \end{pmatrix} \right\|_2$$

if and only if

$$U(L_0) \ > \ P(K_0) - Z(G_0)$$

Proof Since

$$\begin{pmatrix} W_1 S \\ W_2 T \end{pmatrix} = \begin{pmatrix} W_1 \\ -W_2 \end{pmatrix} QDN + \begin{pmatrix} W_1 S_0 \\ W_2 T_0 \end{pmatrix}$$

then applying Lemma 2 shows that $Q = 0$ is optimal if and only if

$$ND\left(W_1^* W_1 S_0^* - W_2^* W_2 T_0^*\right) \text{ has no CRHP poles}$$

$$\Leftrightarrow \ N^* D^* \left(W_1^* W_1 S_0 - W_2^* W_2 T_0\right) \text{ has no CLHP poles}$$

Since this quantity cannot have $j\mathbb{R}$ poles, then

$$\Leftrightarrow \ N^* D^* \left(W_1^* W_1 S_0 - W_2^* W_2 T_0\right) \text{ has no OLHP poles}$$

Let $W_1 = Z P_1$ where P_1 is biproper and bistable. Clearly, this can always be done. We remark that for a given W_1 this factorization is highly non-unique. Next, define

P_2 to be $P_2 = W_2/Z$, so that $W_2 = ZP_2$, and let $A = P_2^* P_2 T_0 - P_1^* P_1 S_0$. Inserting these expressions gives

$$\Leftrightarrow \; Z^* Z N^* D^* A \;\text{ has no CLHP poles,}$$

Viewing P_1 and P_2 as fixed for the moment, when is this quantity anti-stable for some Z? Counting the OLHP poles and zeros as before, this holds if and only if the number of OLHP poles to be cancelled is less than or equal to the number of OLHP zeros available to cancel them. That is,

$$\Leftrightarrow \; \mathcal{P}(Z^*) + \mathcal{P}(A^*) \;\leq\; \mathcal{Z}(Z^*) + \mathcal{Z}(A^*) + \mathcal{Z}(G_0) + \mathcal{P}(G_0)$$

Suppose for the moment that A has no finite $j\mathbb{R}$ zeros or poles. Then,

$$\Leftrightarrow \; \delta_d(Z) + \delta_d(A) - \mathcal{P}(A) \;\leq\; \delta_n(Z) + \delta_n(A) - \mathcal{Z}(A) + \mathcal{Z}(G_0) + \mathcal{P}(G_0)$$

$$\Leftrightarrow \; \delta_r(Z) + \delta_r(A) + \mathcal{Z}(A) - \mathcal{P}(A) \;\leq\; \mathcal{Z}(G_0) + \mathcal{P}(G_0)$$

From the Principle of the Argument, applied to the the standard Nyquist D-contour with small semi-circular indentations into the RHP around the $j\mathbb{R}$ poles and zeros of L_0,

$$\Leftrightarrow \; \delta_r(Z) + \delta_r(A) + \text{wno}(A) \;\leq\; \mathcal{Z}(G_0) + \mathcal{P}(G_0)$$

$$\Leftrightarrow \; \text{wno}(P_2^* P_2 T_0 - P_1^* P_1 S_0) + \delta_r(Z) + \delta_r(A) \;\leq\; \mathcal{Z}(G_0) + \mathcal{P}(G_0)$$

$$\Leftrightarrow \; \text{wno}\left(P_1^* P_1 S_0 \left(\frac{P_2^* P_2 T_0}{P_1^* P_1 S_0} - 1\right)\right) + \delta_r(Z) + \delta_r(A) \;\leq\; \mathcal{Z}(G_0) + \mathcal{P}(G_0)$$

Letting $X = P_2/P_1$,

$$\Leftrightarrow \; \text{wno}\left(P_1^* P_1 S_0 \left(X^* X L_0 - 1\right)\right) + \delta_r(Z) + \delta_r(A) \;\leq\; \mathcal{Z}(G_0) + \mathcal{P}(G_0)$$

Next, we use some properties of winding numbers. Thus,

$$\Leftrightarrow \; \text{wno}\left(P_1^* P_1 S_0\right) + \text{wno}\left(X^* X L_0 - 1\right) + \delta_r(Z) + \delta_r(A) \;\leq\; \mathcal{Z}(G_0) + \mathcal{P}(G_0)$$

$$\Leftrightarrow \; \text{wno}\left(P_1^* P_1 S_0\right) + \text{wno}\left(X^* X L_0, +1\right) + \delta_r(Z) + \delta_r(A) \;\leq\; \mathcal{Z}(G_0) + \mathcal{P}(G_0)$$

Recall that P_1 is biproper and L_0 is proper. Then $P_1^* P_1 S_0$ is biproper. The term $\text{wno}(P_1^* P_1 S_0)$ can be evaluated by counting the positive real-axis crossovers of the Nyquist curve of $P_1^* P_1 S_0$. Since $\measuredangle P_1^* P_1(j\omega) = 0 \;\; \forall\, \omega$, the net number of crossovers of $P_1^* P_1 S_0$ is the same as that of S_0. Furthermore, by A4, P_1 has no poles or zeros on the imaginary axis or at infinity. Therefore, on the semi-circular indentations and on the large semi-circle of the contour, $P_1^* P_1$ resembles a fixed non-zero finite constant. Hence, the number of crossovers occurring on the semi-circular parts of the contour are the same for $P_1^* P_1 S_0$ and S_0. It then follows that

$$\text{wno}(P_1^* P_1 S_0) = \text{wno}(S_0) = \mathcal{P}(G_0) + \mathcal{P}(K_0)$$

Hence, $Q = 0$ is optimal for some weights if and only if $\exists\ W_1, W_2$ such that

$$P(G_0) + P(K_0) + \text{wno}(X^* X L_0, +1) + \delta_r(Z) + \delta_r(A) \ \le\ Z(G_0) + P(G_0)$$

Ranging over all weights is the same as ranging over all X and Z, (subject to A4), so the above is equivalent to $\exists\ X, Z$ such that

$$\text{wno}(X^* X L_0, +1) + \delta_r(Z) + \delta_r(A) \le\ Z(G_0) - P(K_0)$$

$$\Leftrightarrow \min_{X,Z} (\text{wno}(X^* X L_0, +1) + \delta_r(Z) + \delta_r(A)) \ \le\ Z(G_0) - P(K_0) \qquad (9)$$

Consider the issue of choosing X to minimize the term $\text{wno}(X^* X L_0, +1)$. This winding number can be determined from the positive real-axis crossovers of the Nyquist curve (on the indented contour) of $X^* X L_0$. We wish to relate the crossovers of $X^* X L_0$ to the crossovers of L_0 (on the same indented contour). Consider first the finite positive real-axis crossovers of $X^* X L_0$, that is, the points where $X^* X L_0(j\omega) \in (0, \infty)$. Since $\measuredangle X^* X(j\omega) = 0$ on the imaginary axis, i.e. on all of the Nyquist contour except the semi-circular parts, it follows that

$$X^* X L_0(j\omega) \in (0, \infty) \ \Leftrightarrow\ L_0(j\omega) \in (0, \infty)$$

so that the finite positive real-axis crossovers of $X^* X L_0(j\omega)$ and $L_0(j\omega)$ occur at precisely the same frequencies. We can choose X to locate these positive real-axis crossovers to the left or to the right of $+1$ as desired. To minimize the clockwise winding number of $X^* X L_0$ about $+1$, one locates the upward (counter clockwise) crossovers to the right of $+1$ and the downward (clockwise) crossovers to the left of $+1$. In this way, $\text{wno}(X^* X L_0, +1)$ gets a contribution of -1 from each upwards positive real-axis crossover of L_0, and avoids a contribution of $+1$ from every downwards positive real-axis crossover of L_0. Specifically, choose X so that it obeys the following conditions.

C1. For $L_0(j\omega_c) \in (0, +\infty)$ an upwards crossover, $X^* X L_0(j\omega_c) > 1$

C2. For $L_0(j\omega_c) \in (0, +\infty)$ a downwards crossover, $X^* X L_0(j\omega_c) < 1$

These conditions are easily realized. For instance, one can apply polynomial interpolation to the numerator of $X^* X$ for fixed denominator, while viewing it as a polynomial in s^2. It follows that A will have no $j\mathbb{R}$ zeros. Also, by assumption A6, there are no tangential touchings.

Next, consider the small semi-circular indentations of the contour about the (finite) $j\mathbb{R}$ zeros of L_0. In view of A4, $X^* X(j\omega)$ is finite and non-zero and L_0 is very small on these semi-circles, so that $X^* X L_0(s)$ is very small on the indentation, and so $X^* X L_0(s)$ lies to the left of $+1$. Hence, it makes no contribution to $\text{wno}(X^* X L_0, +1)$.

Next, consider the small semi-circular indentations of the contour about the $j\mathbb{R}$ poles of L_0. In view of A4, $X^* X(j\omega)$ is finite and non-zero, so on these semi-circular indentations, it resembles a fixed finite non-zero constant there. So these semi-circular indentations are mapped under $X^* X L_0(s)$ to very large semi-circles. These may or may not contribute to $\text{wno}(X^*XL_0, +1)$. However, such contributions do not depend on X, so the total contribution of this type to $\text{wno}(X^*XL_0, +1)$ is the same as the total contribution of this type to $\text{wno}(L_0, +1)$.

Next, consider the behaviour of the Nyquist curve of $X^*XL_0(s)$ on the large semi-circle of the contour, i.e. for s very large. If X^*XL_0 is strictly improper, this large semi-circle gets mapped to (possibly many) large semi-circles, thus contributing (possibly many) crossovers. The number of such crossovers depends on the relative degrees of X and L_0. Moreover, the relative degrees of X and L_0 are not independent of the terms $\delta_r(A)$ and $\delta_r(Z)$ in eqn. (9). To deal with this complication, it is best to consider various distinct cases separately, depending on the relative degrees of P_2 and X^*XL_0.

We will need some (easy) facts. Note that

$$\delta_r(P_2^*P_2T_0) \;=\; \delta_r(X^*XL_0)$$

Suppose that $P_2^*P_2T_0$ is biproper or strictly proper. Then, the term $\delta_r(A)$ is

$$\delta_r(A) = \delta_r(P_1^*P_1S_0 - P_2^*P_2T_0) = \delta_r(P_1^*P_1S_0) = 0 \tag{10}$$

provided

$$P_1^*P_1S_0(\infty) \;\neq\; P_2^*P_2T_0(\infty) \;\Leftrightarrow\; X^*XL_0(\infty) \;\neq\; +1$$

and this is assured by C1 and C2. Suppose on the other hand that $P_2^*P_2T_0$ is strictly improper. Then

$$\delta_r(A) = \delta_r(P_1^*P_1S_0 - P_2^*P_2T_0) = \delta_r(P_2^*P_2T_0) = \delta_r(X^*XL_0) < 0$$

so that

$$\delta_r(A) = \delta_r(X^*XL_0) < 0 \tag{11}$$

By A4, W_1 must be strictly proper, which holds if and only if

$$\delta_r(W_1) = \delta_r(Z) > 0 \;\Leftrightarrow\; \delta_r(Z) \geq 1 \tag{12}$$

Again by A4, W_2 must be strictly proper, so that

$$\delta_r(W_2) > 0$$

$$\Leftrightarrow\; \delta_r(W_2) = \delta_r(ZP_2) = \delta_r(Z) + \delta_r(P_2) = \delta_r(Z) + \delta_r(X) \geq 1$$

$$\Leftrightarrow\; \delta_r(Z) \geq 1 - \delta_r(X) \tag{13}$$

First, we deal with the case where P_2 is biproper. This will turn out to be the important case.

Case 1(a) Suppose that P_2 is biproper and L_0 is strictly proper. This case applies if and only if $\delta_r(W_1) = \delta_r(W_2)$ and $\delta_r(L_0) > 0$. It follows that $U(L_0)$ does not have a contribution from an upwards crossover at $L_0(\infty)$, and then $u_3 = 0$. Since X^*XL_0 is then also strictly proper, it cannot have a crossover to the right of +1 at $s = \infty$. Hence, choosing X so that C1 and C2 are obeyed gives

$$\min_X \text{wno}\,(X^*XL_0, +1) = -U(L_0)$$

Case 1(b) Suppose that both P_2 and L_0 are biproper. In this case, it is possible that L_0 has an upwards crossover at $s = \infty$, with $U(L_0)$ having a corresponding contribution of -1. This occurs if and only if

$$L_0(j\omega) > 0 \;\text{ and }\; \frac{\mathrm{d}}{\mathrm{d}\omega}\measuredangle L_0(j\omega) > 0 \;\text{ as }\; \omega \to \infty$$

Then, $U(L_0)$ includes a contribution of -1 from this crossover. By choosing X so that $X^*XL_0(\infty) > +1$, then $\text{wno}(X^*XL_0, +1)$ also gets a contribution of -1. If $L_0(\infty) > 0$, and it is a downwards crossover, then $X(\infty)$ may be chosen so $X^*XL_0(\infty) < +1$, thus avoiding a contribution of $+1$ to $\text{wno}(X^*XL_0, +1)$ from $L_0(\infty)$. Finally, if $L_0(\infty) < 0$, then neither $\text{wno}(X^*XL_0, +1)$ nor $U(L_0)$ can have a positive real-axis crossover at $s = \infty$. All in all, we see that $U(L_0)$ and $\text{wno}(X^*XL_0, +1)$ get the same contribution from $L_0(\infty)$. Hence, we again have that

$$\min_X \text{wno}\,(X^*XL_0, +1) = -U(L_0)$$

For both Cases 1(a) and 1(b), using eqns. (9) and (11) gives

$$\min_{X,Z} (\text{wno}\,(X^*XL_0, +1) + \delta_r(Z) + \delta_r(A)) \;\geq\; -U(L_0) + 1 + 0$$

Since Z must be strictly proper by A4, $\delta_r(Z) = 1$ is the best choice. So by choosing $\delta_r(Z) = 1$, this lower bound can be achieved, so that in fact

$$\min_{X,Z} (\text{wno}\,(X^*XL_0, +1) + \delta_r(Z) + \delta_r(A)) \;=\; -U(L_0) + 1$$

Combining Cases 1(a) and 1(b), and returning to eqn. (8), it follows that $Q = 0$ is optimal for some (W_1, W_2), subject to the additional constraint $\delta_r(P_2) = 0$, if and only if

$$\Leftrightarrow \; \min_{X,Z} (\text{wno}\,(X^*XL_0, +1) + \delta_r(Z) + \delta_r(A)) \;\leq\; Z(G_0) - P(K_0)$$

$$\Leftrightarrow \; -U(L_0) + 1 \;\leq Z(G_0) - P(K_0)$$

$$\Leftrightarrow \; -U(L_0) \;<\; Z(G_0) - P(K_0)$$

$$\Leftrightarrow \; U(L_0) \;>\; P(K_0) - Z(G_0)$$

The above case settles the inverse optimality problem with the additional constraint that

$$\delta_r(P_2) = 0 \;\Leftrightarrow\; \delta_r(W_1) = \delta_r(W_2)$$

The above proof shows that if the condition $U(L_0) > P(K_0) - Z(G_0)$ is obeyed, then we can construct weights for which $Q = 0$ is optimal. If this condition is not obeyed, then there are no weights obeying $\delta_r(W_1) = \delta_r(W_2)$ for which $Q = 0$ is optimal. It remains to show that the cases where P_2 is not biproper cannot achieve a strictly smaller minimum than Case 1.

Case 2(a) Suppose that X^*XL_0 is strictly proper with even relative degree, say $\delta_r(X^*XL_0) = 2n > 0$, and that $u_3 = 0$. The latter condition means that $U(L_0)$ does not have a contribution of -1 from $s = \infty$. Since X^*XL_0 is strictly proper, the large semi-circle gets mapped to very small semi-circle(s), and so cannot contribute to the winding number about $+1$. Hence, neither $U(L_0)$ nor $\mathrm{wno}(X^*XL_0, +1)$ gets a contribution from $s = \infty$, so that

$$\min_X \mathrm{wno}\,(X^*XL_0, +1) = -U(L_0)$$

Then, from eqns. (10) and (12),

$$\min_{X,Z} (\mathrm{wno}\,(X^*XL_0, +1) + \delta_r(A) + \delta_r(Z)) \;\geq\; -U(L_0) + 0 + 1$$

so this case cannot do strictly better than Case 1.

Case 2(b) Suppose that X^*XL_0 is strictly proper with even relative degree, say $\delta_r(X^*XL_0) = 2n > 0$, and that $u_3 = 1$. The latter condition means that $U(L_0)$ has a contribution of -1 from $s = \infty$. Since X^*XL_0 is strictly proper, the large semi-circle gets mapped to very small semi-circle(s), and so cannot contribute to the winding number about $+1$. Hence, the contribution of -1 from $u_3 = 1$ to $U(L_0)$ cannot be picked up by $\mathrm{wno}(X^*XL_0, +1)$ so that

$$\min_X \mathrm{wno}\,(X^*XL_0, +1) = -U(L_0) + 1$$

Then, from eqns. (10) and (12),

$$\min_{X,Z} (\mathrm{wno}\,(X^*XL_0, +1) + \delta_r(A) + \delta_r(Z))$$

$$\geq\; -U(L_0) + 2$$

Case 2(c) Suppose that $X^* X L_0$ is strictly proper with odd relative degree, say $\delta_r(X^* X L_0) = 2n - 1 > 0$. Then, $\delta_r(L_0)$ is necessarily odd and so is strictly proper, so that $u_3 = 0$. Since $X^* X L_0(\infty) = 0$, there can be no contribution from the large semi-circle to $\text{wno}(X^* X L_0, +1)$. Then,

$$\min_X \text{wno}\,(X^* X L_0, +1) = -U(L_0)$$

Then, from eqns. (10) and (12),

$$\min_{X,Z} \text{wno}\,(X^* X L_0, +1) + \delta_r(A) + \delta_r(Z) \;\geq\; -U(L_0) + 0 + 1$$

Next, we turn our attention to the case where $X^* X L_0$ is strictly improper. Here, the large semi-circle gets mapped to very large semi-circles, and so may contribute to the winding number about $+1$. Let us determine this contribution exactly. It depends on the relative degree of $X^* X L_0$ and on the asymptotic phase of the loop gain as the frequency tends to infinity. So there are several cases involved.

Case 3(a) Suppose that $X^* X L_0$ is strictly improper with even relative degree, say $\delta_r(X^* X L_0) = 2n < 0$, and that

$$\measuredangle L_0(j\omega) \in (-\epsilon + 180, +\epsilon + 180) \text{ for very large } \omega$$

i.e. $\measuredangle L_0(j\omega)$ tends to 180 (from above or below) as ω tends to infinity. Then $u_3 = 0$ necessarily and the large semi-circle contributes exactly $|n|$ clockwise crossovers, so that

$$\min_X \text{wno}\,(X^* X L_0, +1) = -U(L_0) + |n|$$

Then, from eqns. (11) and (13), and using $\delta_r(L_0) \geq 0$ gives

$$\min_{X,Z} \text{wno}\,(X^* X L_0, +1) + \delta_r(A) + \delta_r(Z)$$

$$\geq\; -U(L_0) + 1$$

Case 3(b) Suppose that $X^* X L_0$ is strictly improper with even relative degree, say $\delta_r(X^* X L_0) = 2n < 0$, and that

$$\measuredangle L_0(j\omega) \in (0, +\epsilon) \text{ for very large } \omega$$

i.e. $\measuredangle L_0(j\omega)$ decreases from above towards zero as ω tends to infinity. Then, $u_3 = 0$ necessarily and the large semi-circle contributes exactly $(|n| + 1)$ clockwise crossovers, so that

$$\min_X \text{wno}\,(X^* X L_0, +1) = -U(L_0) + |n| + 1$$

Then, from eqns. (11) and (13),

$$\min_{X,Z} \text{wno}\,(X^* X L_0, +1) + \delta_r(A) + \delta_r(Z)$$

$$\geq\ -U(L_0) + 2$$

Case 3(c) Suppose that $X^* X L_0$ is strictly improper with even relative degree, say $\delta_r(X^* X L_0) = 2n < 0$, that L_0 is strictly proper, and that

$$\measuredangle L_0(j\omega) \in (-\epsilon, 0) \ \text{ for very large } \omega$$

i.e. $\measuredangle L_0(j\omega)$ increases from below towards zero as ω tends to infinity. Then $u_3 = 0$ necessarily and the large semi-circle contributes exactly $(|n| - 1)$ clockwise crossovers, so that

$$\min_{X} \text{wno}\,(X^* X L_0, +1) = -U(L_0) + |n| - 1$$

Then, using eqns. (11) and (13), and since L_0 is strictly proper with even relative degree ($\Leftrightarrow \delta_r(L_0) \geq 2$),

$$\min_{X,Z} \text{wno}\,(X^* X L_0, +1) + \delta_r(A) + \delta_r(Z)$$

$$\geq\ -U(L_0) + 1$$

Case 3(d) Suppose that $X^* X L_0$ is strictly improper with even relative degree, say $\delta_r(X^* X L_0) = 2n < 0$, that L_0 is biproper, and that

$$\measuredangle L_0(j\omega) \in (-\epsilon, 0) \ \text{ for very large } \omega$$

i.e. $\measuredangle L_0(j\omega)$ increases from below towards zero as ω tends to infinity. These assumptions mean that $u_3 = 1$ necessarily. Since $X^* X L_0$ is strictly improper, this quantity cannot have a finite crossover at $s = \infty$, so that the minimum possible number of finite crossovers is $-U(L_0) + 1$. Also, the large semi-circle contributes exactly $(|n| - 1)$ clockwise crossovers, so that

$$\min_{X} \text{wno}\,(X^* X L_0, +1) = (-U(L_0) + 1) + (|n| - 1) = -U(L_0) + |n|$$

Then, using eqns. (11) and (13) and the fact that L_0 is biproper ($\Leftrightarrow \delta_r(L_0) = 0$),

$$\min_{X,Z} \text{wno}\,(X^* X L_0, +1) + \delta_r(A) + \delta_r(Z)$$

$$\geq\ -U(L_0) + 1$$

Case 3(e) Suppose that X^*XL_0 is strictly improper with odd relative degree, say $\delta_r(X^*XL_0) = 2n-1 < 0$, and that $\measuredangle L_0(j\omega)$ tends to -90. Then,

$$\min_X \text{wno}\,(X^*XL_0, +1) = -U(L_0) + |n|$$

Then, from eqns. (11) and (13), and the fact that L_0 has odd relative degree here ($\Rightarrow \delta_r(L_0) \geq 1$),

$$\min_{X,Z} \text{wno}\,(X^*XL_0, +1) + \delta_r(A) + \delta_r(Z)$$

$$\geq\ -U(L_0) + 1$$

Case 3(f) Suppose that X^*XL_0 is strictly improper with odd relative degree, say $\delta_r(X^*XL_0) = 2n-1 < 0$, and that $\measuredangle L_0(j\omega)$ tends to $+90$. Then,

$$\min_X \text{wno}\,(X^*XL_0, +1) = -U(L_0) + |n| + 1$$

Then, using eqns. (11) and (13) and $\delta_r(L_0) \geq 1$,

$$\min_{X,Z} \text{wno}\,(X^*XL_0, +1) + \delta_r(A) + \delta_r(Z)$$

$$\geq\ -U(L_0) + 2$$

Next, consider the case where X^*XL_0 is biproper.

Case 4(a) Suppose that X^*XL_0 is biproper, and that

$$X^*XL_0(\infty) > 0 \text{ and } \frac{\mathrm{d}}{\mathrm{d}\omega}\measuredangle X^*XL_0(\infty) > 0$$

If L_0 is biproper, then X must be biproper also. But this is Case 1, so we can assume here that L_0 is strictly proper, and then $u_3 = 0$. Since X^*XL_0 is biproper, the large semi-circle gets mapped to a (non-zero, non-infinite) point, namely $X^*XL_0(\infty)$. So X^*XL_0 has a real-axis crossover at infinity. Under the assumptions for this case, this crossover is both a positive real-axis crossover and an upwards crossover. So we must choose X to ensure that this contibution of -1 is picked up, so as to minimize the value of wno$(X^*XL_0, +1)$. To this end, we must choose X so that $X^*XL_0(\infty) > +1$. Then, the Nyquist curve of X^*XL_0 has an additional upwards crossover which was not counted in $U(L_0)$. Hence,

$$\min_X \text{wno}\,(X^*XL_0, +1) = -U(L_0) - 1$$

Then, from eqns. (10) and (13), and since X must be strictly improper ($\Leftrightarrow \delta_r(X) \leq -1$) here,

$$\min_{X,Z} \text{wno}\,(X^*XL_0, +1) + \delta_r(A) + \delta_r(Z)$$

$$\geq\ (-U(L_0) - 1) + 0 + 2\ =\ -U(L_0) + 1$$

Case 4(b) Suppose that X^*XL_0 is biproper, and that

$$\text{Either } X^*XL_0(\infty) < 0 \text{ or } \frac{\mathrm{d}}{\mathrm{d}\omega}\measuredangle X^*XL_0(\infty) < 0$$

For the same reason as in the previous case, we can assume that L_0 is strictly proper, so $u_3 = 0$. Since X^*XL_0 is biproper, it has a real-axis crossover at $s = \infty$. The assumptions for this case mean that this is either a negative real-axis crossover or a downwards positive real-axis crossover. If the latter holds, we must choose X so that $X^*XL_0(\infty) < +1$. to avoid the possible contribution of $+1$ from the crossover at $X^*XL_0(\infty)$. Then, the Nyquist curve of X^*XL_0 does not have an additional upwards crossover at $s = \infty$ to the right of $+1$. Also, since L_0 is strictly proper, $U(L_0)$ cannot have a contribution from $s = \infty$. Hence,

$$\min_X \text{wno}\,(X^*XL_0, +1) = -U(L_0)$$

Then, from eqns. (10) and (12),

$$\min_{X,Z} \text{wno}\,(X^*XL_0, +1) + \delta_r(A) + \delta_r(Z) \;\geq\; -U(L_0) + 0 + 1$$

Thus, we can conclude that none of Cases 2 to 4 can yield a value of

$$\min_{X,Z} \text{wno}\,((X^*XL_0, +1) + \delta_r(A) + \delta_r(Z))$$

which is strictly less than the value given by Case 1. ■

Remark It follows as a corollary that if $Q = 0$ is optimal for some (W_1, W_2), then it is also optimal for weights obeying $\delta_r(W_1) = 1 = \delta_r(W_2)$.

Lemma 6 *Suppose that assumptions A1 to A4 are obeyed. Then,*

$$D(L_0) + P(G_0) - Z(K_0) = U(L_0) + Z(G_0) - P(K_0)$$

Proof Since L_0 has no poles or zeros on the indented contour, wno(L_0) is well defined. Applying the Principle of the Argument to L_0 directly gives

$$\text{wno}(L_0) = Z(L_0) - P(L_0)$$

$$= Z(G_0) + Z(K_0) - P(G_0) - P(K_0)$$

Now, wno(L_0) can be evaluated by counting its positive real-axis crossovers. Thus,

$$\text{wno}(L_0) = D(L_0) - U(L_0)$$

This gives

$$D(L_0) - U(L_0) = Z(G_0) + Z(K_0) - P(G_0) - P(K_0)$$

$$\Leftrightarrow D(L_0) - Z(K_0) + P(G_0) = U(L_0) + Z(G_0) - P(K_0)$$

as claimed. ■

Lemma 7 *Suppose that assumptions A1 to A5 are obeyed. If*

$$U(L_0) \ \leq \ P(K_0) - Z(G_0)$$

then

$$\exists \ Q_1 \in H_\infty \ s.t. \ |S_1(j\omega)| \ < \ |S_0(j\omega)| \ \ and$$

$$|T_1(j\omega)| \ < \ |T_0(j\omega)| \ \ a.e. \ on \ j\mathbb{R} \cup \{\infty\}$$

Proof The proof is constructive. Suppose that $U(L_0) \leq P(K_0) - Z(G_0)$. Under this condition, we explicitely construct a $Q_1 \in H_\infty$ which decreases both $|S(j\omega)|$ and $|T(j\omega)|$ at almost all frequencies.

Another controller, corresponding to the Youla parameter Q, reduces $|S(j\omega)|$ and $|T(j\omega)|$ almost everywhere if and only if

$$|S(Q)| < |S_0|, \quad |T(Q)| < |T_0| \quad \text{a.e. on } j\mathbb{R}$$

Using eqn. (8) gives

$$|DV + DQN| < |DV|, \quad |NU - DQN| < |NU| \quad \text{a.e. on } j\mathbb{R}$$

that is, if and only if

$$|1 + QV^{-1}N| < 1, \quad |1 - QU^{-1}D| < 1 \quad \text{a.e. on } j\mathbb{R}$$

We restrict the search for such Q's to ones that have arbitrarily small norm. Then, the above reduce to the linear inequalities,

$$\text{Re}\,(-QV^{-1}N) > 0, \quad \text{Re}\,(QU^{-1}N) > 0 \quad \text{a.e. on } j\mathbb{R} \tag{14}$$

The proof proceeds as follows. First, we define a certain transfer function Q_1. Second, we show that if $U(L_0) \leq P(K_0) - Z(G_0)$ then this Q_1 is stable. Third, we show that $\sqrt{-L_{op}}\frac{dv}{nu}x$ has positive real part for almost all frequencies. Fourth, we show that the conditions of eqn. (14) are obeyed, so that this Q_1 reduces both $|S(j\omega)|$ and $|T(j\omega)|$ locally.

Let ω_k denote the frequencies for which $L_0(j\omega)$ crosses the positive real-axis. There are $U(L_0) + D(L_0)$ or $U(L_0) + D(L_0) - 1$ such crossovers (the possible crossover at $\omega = \infty$ is not needed here). Define

$$Q_1 = \sqrt{-L_{\text{op}}}N_{\text{op}}^{-1}V_{\text{op}}\frac{nd}{uv}x, \quad \text{with } x(s) = \epsilon \prod_k j(s - j\omega_k) \tag{15}$$

where ϵ is a real parameter which will be specified later, and where L_{op} denotes the outer part of the loop gain, and $\sqrt{-L_{\text{op}}}$ is that square root of $-L_{\text{op}}$ which is analytic in the CRHP. Such a square root exists because $-L_{\text{op}}$ and its inverse are analytic in the CRHP (see page 274 of [13]).

Next, we claim that this Q_1 is stable. Note that each of its factors $\sqrt{-L_{\text{op}}},\ N_{\text{op}}^{-1},\ V_{\text{op}}$ and $x(nd)/(uv)$ have all their poles in the open left half plane. Since $\sqrt{-L_{\text{op}}}N_{\text{op}}^{-1}V_{\text{op}}$ is biproper, it only remains to check that $\frac{nd}{uv}x$ is proper. Clearly,

$$\begin{aligned}\delta(ndx)-\delta(uv) &= Z(G_0)+P(G_0)+\delta(x)-Z(K_0)-P(K_0)\\ &\leq Z(G_0)+P(G_0)+(U(L_0)+D(L_0))-Z(K_0)-P(K_0)\\ &= (U(L_0)-P(K_0)+Z(G_0))+(D(L_0)-Z(K_0)+P(G_0))\\ &= 2(U(L_0)-P(K_0)+Z(G_0))\end{aligned}$$

where the last equality uses Lemma 6. Then,

$$U(L_0)\leq P(K_0)-Z(G_0)\ \Rightarrow\ \delta(ndx)\leq\delta(uv)$$

and so $\frac{nd}{uv}x$ is proper. Hence, Q_1 is stable.

Next, we claim that

$$\operatorname{Re}\frac{dv}{nu}\sqrt{-L_{op}}x>0\ \text{ a.e. on } j\mathbb{R} \tag{16}$$

It will be convenient to say that two functions f and g are *phase equivalent*, denoted by $f\overset{\measuredangle}{=}g$, if they have equal phase at almost all frequencies. Inequality (16) will be verified in a few steps. Since

$$L_0=\frac{NU}{DV}=\left(\frac{NU}{DV}\right)_{op}\left(\frac{nu}{dv}\right)^*\left(\frac{dv}{nu}\right)$$

$$=L_{op}\left(\frac{dv}{nu}\right)^2\left(\frac{nu}{dv}\right)^*\left(\frac{nu}{dv}\right)$$

and since

$$\measuredangle\left(\frac{nu}{dv}\right)^*\left(\frac{nu}{dv}\right)=0\ \text{ on } j\mathbb{R}$$

we see that

$$L_0\overset{\measuredangle}{=}L_{op}\left(\frac{dv}{nu}\right)^2\ \text{ on } j\mathbb{R} \tag{17}$$

This shows that L_0 is phase equivalent to a function which is analytic in the CRHP. Next, note that

$$L_0(j\omega)\in(0,\infty)$$

$$\Leftrightarrow\ L_{op}\left(\frac{dv}{nu}\right)^2(j\omega)\in(0,\infty)$$

$$\Leftrightarrow\ -L_{op}\left(\frac{dv}{nu}\right)^2(j\omega)\in(-\infty,0)$$

The points where this holds are the positive real-axis crossovers. Given the definition of x, these points are the zeros of x, and x is real valued on the imaginary axis. So multiplying by x gives

$$\Rightarrow \ -L_{op}\left(\frac{dv}{nu}\right)^2 x^2(j\omega) \notin (-\infty, 0) \ \ \forall\, \omega$$

$$\Rightarrow \ \sqrt{-L_{op}}\left(\frac{dv}{nu}\right) x \notin j\mathbb{R}/\{0\} \ \ \forall\omega$$

Thus, this quantity never crosses the imaginary axis. By choosing the factor ϵ in x as $+1$ or -1 as appropriate, it follows that the above quantity is confined to the CRHP, so that eqn. (16) holds, as required.

It needs to be checked that this quantity does not change sign as $s = j\omega$ goes through a zero. Certainly, $x(j\omega)$ is real-valued and it changes sign only at the ω_k's which are by definition the frequencies where L_0 crosses the positive real line. From (17) we see that these frequencies are exactly those frequencies where $\frac{dv}{nu}\sqrt{-L_{\mathrm{op}}}$ crosses the imaginary axis. As the frequency ω is increased towards ω_k, the term $\frac{dv}{nu}\sqrt{-L_{\mathrm{op}}}$ is about to cross the imaginary axis. Then the sign of $x(j\omega)$ changes at ω_k. This ensures that the product $\frac{dv}{nu}\sqrt{-L_{\mathrm{op}}}x$ remains on the same side of the imaginary axis. By fixing the sign of ϵ this side can always be taken to be the RHP.

This particular Q_1 gives

$$\begin{aligned} -V^{-1}NQ_1 &= \frac{v}{v^*}V_{\mathrm{op}}^{-1}\frac{n^*}{n}N_{\mathrm{op}}\sqrt{-L_{\mathrm{op}}}N_{\mathrm{op}}^{-1}V_{\mathrm{op}}\frac{nd}{uv}x \\ &= \frac{n^*n}{v^*v}\frac{dv}{nu}\sqrt{-L_{\mathrm{op}}}x \\ &\stackrel{\measuredangle}{=} \frac{dv}{nu}\sqrt{-L_{\mathrm{op}}}x \end{aligned}$$

and

$$\begin{aligned} U^{-1}DQ_1 &= \frac{1}{-L}(-V^{-1}NQ_1) \\ &\stackrel{\measuredangle}{=} \frac{1}{-\frac{d^2v^2}{n^2u^2}L_{\mathrm{op}}}\left(\frac{dv}{nu}\sqrt{-L_{\mathrm{op}}}x\right) \\ &= \frac{1}{\frac{dv}{nu}\sqrt{-L_{\mathrm{op}}}x}x^2 \\ &\stackrel{\measuredangle}{=} \frac{1}{\frac{dv}{nu}\sqrt{-L_{\mathrm{op}}}x} \end{aligned}$$

Here we used the fact that that $x^2(j\omega) \stackrel{\measuredangle}{=} 1$ which follows directly from the fact that $x(j\omega) \in \mathbb{R}$. It is trivial to verify that for any transfer function

$$A(j\omega) > 0 \ \text{ a.e.} \quad \Leftrightarrow \quad 1/A(j\omega) > 0 \ \text{ a.e.}$$

It follows that the two conditions in eqn. (14) hold if and only if the single condition of eqn. (16) holds. But this was verified above, completing the proof. ■

Remark Theorem 5 has now been established. Lemma 5 is (a) $\Leftrightarrow$ (d). Lemma 6 gives (c) $\Leftrightarrow$ (d) directly. Lemma 7, the only part which uses A5, is the contrapositive of (b) $\Rightarrow$ (d). For the converse of this part, note that if there exists a $K \in \mathcal{K}$ which reduces $|S(j\omega)|$ and $|T(j\omega)|$ at almost every frequency, then trivially K_0 could not be 2-norm optimal. So not (b) implies not (a), and the contrapositive gives (a) $\Rightarrow$ (b), and using Lemma 5 then gives (d) $\Rightarrow$ (b), and this part does not require A5. Next, we turn to the 2-block infinity-norm result.

Theorem 10 *Suppose that assumptions A1 to A4 are obeyed. Then, (a) through (e) are equivalent. If A5 is also obeyed, then (a) through (g) are equivalent.*

$$
\begin{array}{ll}
(a) & \exists\, W_1, W_2 \ \ s.t.\ \ K_0 = \arg\inf_{K\in\mathcal{K}} \ \| \ |W_1 S|^2 + |W_2 T|^2 \ \|_\infty \\
(b) & \nexists\, Q_1 \in \mathcal{H}_\infty \ s.t.\ |S_1(j\omega)| \ < \ |S_0(j\omega)| \ \ and \\
 & |T_1(j\omega)| \ < \ |T_0(j\omega)| \ \ everywhere\ on\ j\mathbb{R} \cup \{\infty\} \\
(c) & \mathcal{P}(G_0) \ > \ \mathcal{Z}(K_0) \\
 & or\ L(j\omega_c) \in [0, \infty]\ for\ at\ least\ one\ \omega_c \in \mathbb{R} \cup \{\infty\} \\
(d) & \mathcal{Z}(G_0) \ > \ \mathcal{P}(K_0) \\
 & or\ L(j\omega_c) \in [0, \infty]\ for\ at\ least\ one\ \omega_c \in \mathbb{R} \cup \{\infty\} \\
(e) & K_0\ is\ 2\text{-}norm\ optimal \\
 & or\ L(j\omega_c) \in [0, \infty]\ for\ at\ least\ one\ \omega_c \in \mathbb{R} \cup \{\infty\} \\
(f) & K_0\ is\ S\text{-}optimal \\
 & or\ K_0\ is\ T\text{-}optimal \\
 & or\ there\ is\ at\ least\ one\ \mathbb{R}_+\ crossover \\
(g) & K_0\ is\ S\&T\text{-}optimal\ with\ (\alpha, \beta)
\end{array}
$$

Proof We begin by considering the case where

$$L(j\omega_c) \in [0, \infty] \text{ for at least one } \omega_c \in \mathbb{R} \cup \{\infty\}$$

First, we claim that K_0 is infinity-norm optimal. Choose the real constants α and β so that

$$L(j\omega_c) = \left(\frac{\alpha}{\beta}\right)^2$$

where $s = j\omega_c$ is any one selected point where $L(j\omega_c) \in [0, \infty]$ for some $\omega_c \in \mathbb{R} \cup \{\infty\}$

Choose the weights to be

$$W_1 = \alpha Z^{-1}, \quad W_2 = \beta Z^{-1}$$

where Z is defined by the spectral factorization,

$$Z^* Z = \alpha^2 S_0^* S_0 + \beta^2 T_0^* T_0$$

Note that $|W_1 S(j\omega)|^2 + |W_2 T(j\omega)|^2$ is then all-pass. The remarks in Section 4 then show that no Q (stable or unstable) can reduce $|W_1 S(j\omega_c)|^2 + |W_2 T(j\omega_c)|^2$.

From the very definition of the infinity-norm, no Q can reduce $\| \ |W_1 S|^2 + |W_2 T|^2 \ \|_\infty$. This shows that $Q = 0$ is infinity-norm optimal, so that (a) holds.

Since

$$|S(j\omega_c)| + |T(j\omega_c)| = S(j\omega_c) + T(j\omega_c) = 1$$

it follows that no Q can reduce both $|S(j\omega_c)|$ and $|T(j\omega_c)|$, which shows that (b) holds. Since L_0 has at least one such $j\omega_c$ by assumption, it is trivial that (c), (d), (e) and (f) hold. Thus, when $L_0(j\omega_c) \in [0, \infty]$ (i.e. that L_0 has a finite $\mathbb{R}_+$ crossover or $L_0(j\omega_c) = 0$, $L_0(j\omega_c) = \infty$, $L_0(\infty) = 0$, or a tangential touchings), (a) to (f) are obeyed.

Now suppose that L_0 has no such $j\omega_c$. Then $U(L_0) = 0 = D(L_0)$ and A5 necessarily applies. Lemma 6 then gives

$$Z(K_0) - P(G_0) = P(K_0) - Z(G_0)$$

which shows that $(c) \Leftrightarrow (d)$. Lemma 7 shows that both $|S(j\omega)|$ and $|T(j\omega)|$ can be strictly decreased at every ω if $P(K_0) \geq Z(G_0)$, so that not $(d) \Rightarrow$ not (b), or $(b) \Rightarrow (d)$. The fact that not $(b) \Rightarrow$ not (a) is trivial, giving $(a) \Rightarrow (b)$. Under the present conditions, Lemma 5 shows that (c) implies 2-norm optimality for some W_1, W_2. But then it impossible that another controller could reduce $|ZW_1 S|^2 + |ZW_2 T_0|^2$ at every frequency, where Z here is chosen to make this expression all-pass. It follows that 2-norm optimality implies infinity-norm optimality. Hence, $(c) \Rightarrow (a)$.

For part (e), it suffices to note that when there are such $j\omega_c$'s, A5 holds and $U(L_0) = 0 = D(L_0)$, and then parts (c) and (d) of Theorem 6 become parts (c) and (d) of Theorem 5. For part (f), apply Theorems 1 and 2 when $U(L_0) = 0 = D(L_0)$. For part (g), choose α and β as above, $W_1 = \alpha[S_0]_{op}^{-1}$ and $W_2 = [T_0]_{op}^{-1}$. The claim then follows from the remarks in Section 4. ■

Acknowledgements

The authors wish to thank Prof. Huibert Kwakernaak of the University of Twente and Prof. Charles McCorkell of Dublin City University for making this work possible.

References

1. M. G. Safonov, A. J. Laub, and G. L. Hartmann, "Feedback Properties of Multivariable Systems: The Role and Use of the Return Difference Matrix," *IEEE Transactions on Automatic Control*, Vol. 26, pp. 47-65, 1981.

2. J. S. Freudenebrg and D. P. Looze, "Right Half Plane Poles and Zeros and Design Tradeoffs in Feedback Systems," *IEEE Transactions on Automatic Control*, Vol. 30, pp. 555-565, 1985.

3. K. E. Lenz, P. P. Khargonekar, and J. C. Doyle, "When is a Controller H^{∞}-Optimal?", *Mathematics of Control, Signals, and Systems*, Vol. 1, pp. 107-122, 1988.

4. K. E. Lenz, "Simple Mixed Sensitivity Optimal Controllers," *Systems and Control Letters*, Vol. 17, pp. 363-373, 1991.

5. K. E. Lenz, "Properties of Optimal Weighted Sensitivity Designs," *IEEE Transactions on Automatic Control*, Vol. 40, pp. 298-301, 1995.

6. S. J. Poreda, "A Characterization of Badly Approximable Functions," *Transactions of the American Mathematical Society*, Vol. 169, pp. 249-256, 1972.

7. J. Feng and M. C. Smith, "When is a Controller Optimal in the Sense of H_{∞} Loop-Shaping?", *IEEE Transactions on Automatic Control*, Vol. 40, pp. 2026-2039, 1995.

8. D. McFarlane and K. Glover, "A Loopshaping Design Procedure Using H_{∞} Synthesis," *IEEE Transactions on Automatic Control*, Vol. 37, pp. 759-769, 1992.

9. D. C. Youla, J. J. Bongiorno, and H. A. Jabr, "Modern Wiener-Hopf Design of Optimal Controllers. Part I," *IEEE Transactions on Automatic Control*, Vol. 21, pp. 3-13, 1976.

10. D. C. Youla, H. A. Jabr, and J. J. Bongiorno, "Modern Wiener-Hopf Design of Optimal Controllers. Part II," *IEEE Transactions on Automatic Control*, Vol. 21, pp. 319-338, 1976.

11. C. A. Desoer, R. Liu, J. Murray, and R. Saeks, "Feedback System Design: The Fractional Representation Approach to Analysis and Synthesis," *IEEE Transactions on Automatic Control*, Vol. 25, pp. 399-412, 1980.

12. L. Ahlfors, *Complex Analysis*, McGraw-Hill, 1953.

13. W. Rudin, *Real and Complex Analysis*, McGraw-Hill, 1987.