This is a repository copy of *Worst-case analysis of identification - BIBO robustness for closed loop data*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/697/

**Article:**

https://doi.org/10.1109/9.328804

**Reuse**
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher’s website.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
The following is a brief overview of the present work. In Section II we present some background material on coprime factorizations and the graph topology. In Section III we consider BIBO stable systems and the question of robustly convergent identification algorithms and input design. Section IV introduces ARX models for causal linear systems, and closed-loop identification is considered. Section V provides some error bounds for closed-loop identification in the framework of this paper. In Section VI we discuss briefly some generalizations to multivariable systems.

II. MATHEMATICAL PRELIMINARIES

Let \((I^p, \| \cdot \|_p)\) \((1 \leq p \leq \infty)\) denote the usual (real) sequence spaces. A linear discrete-time system is defined as a linear convolution operator \(G : I^p \rightarrow I^p\). As usual the linear system \(G\) is called \(I^p\) stable if

\[
\|G\|_{(p)} = \sup_{x \in I^p, x \neq 0} \frac{\|Gx\|_p}{\|x\|_p} < \infty.
\]

Here \(\|G\|_{(p)}\) is the induced operator norm, or the system gain, over \(I^p\). We shall often simplify the notation somewhat and write simply \(\|G\|\).

Let \(S^p\) denote the Banach space of linear shift-invariant causal \(I^p\) stable systems equipped with the operator norm \((1)\), and the known that \(S^\infty\) is isometrically isomorphic to \(l^1\). Thus, for a system \(G \in S^\infty\), we shall let \(G\) denote also the (unit) impulse response \((n)_n\geq 0(G)\). Then \(\|G\|_{(\infty)} = \|G\| = \sum_{n=0}^{\infty} |g_n| < \infty\).

A convenient way of representing both \(l^p\) stable and unstable systems \((p \in [1, \infty])\) is to consider the quotient field \(F(S^p)\). For \(S^p\) the \(l^p\) is isometrically isomorphic to \(l^1\). Thus, for a system \(G \in S^\infty\), we shall let \(G\) denote also the (unit) impulse response \((n)_n\geq 0(G)\). Then \(\|G\|_{(\infty)} = \|G\| = \sum_{n=0}^{\infty} |g_n| < \infty\).

A convenient way of representing both \(l^p\) stable and unstable systems \((p \in [1, \infty])\) is to consider the quotient field \(F(S^p)\). For \(S^p\) the \(l^p\) is isometrically isomorphic to \(l^1\). Thus, for a system \(G \in S^\infty\), we shall let \(G\) denote also the (unit) impulse response \((n)_n\geq 0(G)\). Then \(\|G\|_{(\infty)} = \|G\| = \sum_{n=0}^{\infty} |g_n| < \infty\).

A convenient way of representing both \(l^p\) stable and unstable systems \((p \in [1, \infty])\) is to consider the quotient field \(F(S^p)\). For \(S^p\) the \(l^p\) is isometrically isomorphic to \(l^1\). Thus, for a system \(G \in S^\infty\), we shall let \(G\) denote also the (unit) impulse response \((n)_n\geq 0(G)\). Then \(\|G\|_{(\infty)} = \|G\| = \sum_{n=0}^{\infty} |g_n| < \infty\).

A convenient way of representing both \(l^p\) stable and unstable systems \((p \in [1, \infty])\) is to consider the quotient field \(F(S^p)\). For \(S^p\) the \(l^p\) is isometrically isomorphic to \(l^1\). Thus, for a system \(G \in S^\infty\), we shall let \(G\) denote also the (unit) impulse response \((n)_n\geq 0(G)\). Then \(\|G\|_{(\infty)} = \|G\| = \sum_{n=0}^{\infty} |g_n| < \infty\).

A convenient way of representing both \(l^p\) stable and unstable systems \((p \in [1, \infty])\) is to consider the quotient field \(F(S^p)\). For \(S^p\) the \(l^p\) is isometrically isomorphic to \(l^1\). Thus, for a system \(G \in S^\infty\), we shall let \(G\) denote also the (unit) impulse response \((n)_n\geq 0(G)\). Then \(\|G\|_{(\infty)} = \|G\| = \sum_{n=0}^{\infty} |g_n| < \infty\).

A convenient way of representing both \(l^p\) stable and unstable systems \((p \in [1, \infty])\) is to consider the quotient field \(F(S^p)\). For \(S^p\) the \(l^p\) is isometrically isomorphic to \(l^1\). Thus, for a system \(G \in S^\infty\), we shall let \(G\) denote also the (unit) impulse response \((n)_n\geq 0(G)\). Then \(\|G\|_{(\infty)} = \|G\| = \sum_{n=0}^{\infty} |g_n| < \infty\).
consists of $l^\infty$ stable elements. (Closed-loop stability means that $H(G, K_0)$ is a BIBO stable operator.) $H(G, K_0)$ can be expressed as

$$H(G, K_0) = \begin{bmatrix} G_X & G_0 \\ \frac{1}{1+G} & 1 \\ \frac{1}{1+G} & \frac{1}{1+G} \end{bmatrix}.$$  

(3)

Now, there exist (by coprimeness of $X$ and $Y$) $E, F \in S^\infty$ such that $MX + EY = 1$. Using this identity and the expression for $H(G, K_0)$ we see that $1/(Y + XG), G/(Y + XG) \in S^\infty$, too. But then $(G/(Y + XG), X/(Y + XG))$ is a c.f. of $G$ over $S^\infty$. The only-if part follows by a result in [28, p. 364].

By this fact, the set of systems $CF(S^\infty)$ is really the 'largest' set of causal linear shift-invariant systems in which we can study identification of the unknown system from closed-loop time series.

It is possible to introduce several distance functions and metrics concerning with the stability to be a robust property. $CF(S^\infty)$ is a meterizable space, and it is possible to introduce several distance functions and metrics which induce the graph topology; we refer to [17] for details.

One further distance function may be defined as follows.

A c.f. $(N, D)$ of $G \in CF(S^\infty)$ over $S^\infty$ is said to be normalized (cf. [28]) if $N(z)/S^\infty + D(z)/S^\infty + 1 = 0$ for any $|z| = 1$. It is established in [17] that each $G \in CF(S^\infty)$ has a normalized c.f. over $S^\infty$ which is unique up to multiplication by $1$. Let $G_1, G_2 \in CF(S^\infty)$. Introduce the notation

$$A_i = \frac{D_i}{N_i}, \quad i = 1, 2$$

(4)

where $(N_i, D_i)$ is any coprime factorization (c.f.) of $G_i$ over $S^\infty$.

Define the quantity

$$\kappa(G_1, G_2) = \frac{1}{2\|A_1\|_1 + \|A_2\|_1} \|A_1 - A_2\|_1$$

(5)

where the infimum is thus taken over all BIBO stable c.f.'s of $G_1$ and $G_2$ with unit norm as indicated. We get the following result.

Theorem 2.2: Let $G \in CF(S^\infty)$, and let $(G_i)$ be a sequence in $CF(S^\infty)$. Then $\kappa(G, G_i) \to 0$ if and only if $G \to G$ in the $S^\infty$ graph topology.

Proof: Consider first the if part. Let $G \to G$ in the $S^\infty$ graph topology. Let $(N, D)$ be a c.f. of $G$ such that $\|D\|_N = 1$. Then there exists a sequence $\{(N_i, D_i)\}$, where $(N_i, D_i)$ is a c.f. of $G_i$, such that $N_i \to N, D_i \to D$. Let $a_i = \|D_i/N_i\|_1$. Denote

$$b_i = \left\| \begin{bmatrix} D_i \\ N_i \end{bmatrix} - \begin{bmatrix} D \\ N \end{bmatrix} \right\|_1$$

(6)

By the triangle inequality $|1 - b_i| \leq a_i \leq 1 + b_i$. Now, $b_i \to 0$, so that $a_i \to 1$. Again by the triangle inequality

$$\left\| \begin{bmatrix} D \\ N \end{bmatrix} - a_i^{-1} \begin{bmatrix} D_i \\ N_i \end{bmatrix} \right\|_1 \leq \left\| a_i^{-1} \right\|_1 \left\| \begin{bmatrix} D_i \\ N_i \end{bmatrix} - \begin{bmatrix} D \\ N \end{bmatrix} \right\|_1 \to 0$$

(7)

so that $\kappa(G, G_i) \to 0$ by definition (5). Consider now the only if part. Let $\kappa(G, G_i) \to 0$. For each $i, \epsilon_i > 0$, there exists, directly by definition (5), c.f.'s $(N_i, D_i)$ and $(N, D)$ of $G$ and $G_i$, respectively, satisfying $\|D_i/N_i\|_1 = 1, \|D/N\|_1 = 1$, such that

$$\left\| \begin{bmatrix} D \\ N \end{bmatrix} - \begin{bmatrix} D_i \\ N_i \end{bmatrix} \right\|_1 \leq 2\kappa(G, G_i) + \epsilon_i.$$  

(8)

Take $\epsilon_i = 1/i$. The above inequality then gives $N_i \to N$ and $D_i \to D$. This completes the proof of the theorem.

Thus, $\kappa$, like all the distance functions presented in [17], can be used to quantify the identification error in a way compatible with the needs of robust control design.

We end this section with a striking example illustrating some of the differences between the $S^\infty$ and $S^2$ settings.

Example 2.1: Consider the system $G$ with transfer function $G(z) = \exp\{-\alpha(1+\gamma)/|1-\gamma|\}$, where $\alpha > 0$ is a constant. The induced $l^2$ norm of the system $G$ is equal to $\sup_{\gamma} |G(e^{i\theta})| = 1$, so that $G$ does not amplify the energy of any $l^2$ signal. We see directly, however, that $G$ is not $l^\infty$ stable.

Now $G \in CF(S^\infty)$ as $G$ is $l^\infty$ stable but $G \notin CF(S^2)$. We see as follows. Assume to the contrary that $G \notin CF(S^2)$. Then there exists a coprime factorization $(N, D)$ of $G$ over $S^\infty$. But then $D(z)/G(z) = S(z)$ and both $N(z)$ and $D(z)$ are continuous on the unit circle while $G(z)$ is discontinuous at $z = 1$, necessarily $D(1) = 0$. But then $N(1) = 0$ as $G(z)$ is bounded on the unit circle.

Thus $N$ and $D$ can not be coprime. This contradiction establishes the claim.

In fact, by an analogous reasoning it is seen that no system whose transfer function is bounded but not continuous on the unit circle can have a coprime factorization over $S^\infty$.

Note that the system $G$ above is in the quotient field $F(S^\infty)$ of $S^\infty$ (e.g., $G = z(1-z)^{-1}$, $S = (1-z)^{-1}$, $D = (1-z)^{-1}$, where both the numerator and denominator are in $S^\infty$). Now $G \notin CF(S^\infty)$, it then follows by Theorem 2.1 that $G$ can not be stabilized in the $l^\infty$ sense by any controller in $CF(S^\infty)$ (and yet $G$ is $l^2$ stable).

III. WORST-CASE IDENTIFICATION OF STABLE SYSTEMS

Worst-case identification of BIBO stabilizable feedback systems is closely related to the theory of identification of BIBO stable systems. Therefore we begin by considering the question of robust identification of BIBO stable systems from the point of view of input design.

We suppose that we are given the stable model $y = h \ast u + \nu$, with $h \in l^1$, unknown, $\nu \in l^\infty$, comprising the noise, with $u \in l^\infty$ being the input, and with $\nu$ the measured output. (Here $\ast$ denotes the convolution product.) It is desired to choose $u(t)$ for $t \geq 0$ such that given $y(0), \ldots, y(n)$ we may construct an identified model $h_n$ such that the following robust convergence condition is satisfied

$$\lim_{n \to \infty} \inf_{\|u\|_1 \leq 1} \sup_{t \geq 0} \|h_n - h\|_1 = 0,$$

(9)

It is known [21], [11] that with $u$ chosen to be an impulse or step no such algorithm can exist. On the other hand it is known [15], [27] that some input designs (e.g., Galois sequences) do guarantee the existence of such an algorithm. In [4] a sufficient condition is given. The following result gives a necessary and sufficient condition. It has the further consequence that for a large class of inputs one cannot identify $h \in l^1$ even asymptotically, unless one is given further prior information. We state the result in a general case when the input $u(t)$ can be nonzero (but is unknown) for $t < 0$. This corresponds to input over which we have no control, and we make only the assumption that $u(t)$ is bounded for $t < 0$. Let us write $u = u_+ + u_-$, where $u_+(t) = 0$ for $t < 0$ and $u_-(t) = 0$ for $t \geq 0$. In fact the question whether $u = 0$ does not affect the existence or otherwise of an algorithm.

Theorem 3.1: Given $u_+ \in l^\infty$ and output measurements $y_0, y_1, \ldots$, where $y = h \ast u + \nu$ and $\nu$ is noise, as above, then there is a robustly convergent identification algorithm using $y$ if and
only if \( u_+ \) satisfies
\[
3C > 0 \quad \text{such that for all } k \in \mathbb{N}, \quad ||k \ast u_+||_\infty \geq C||k||_1, \tag{10}
\]

**Proof (Sufficiency):** Note that, by compactness, condition (10) is equivalent to the following: there exists \( C > 0 \) such that for all \( p \geq 0 \) there is a number \( n_p \) such that if \( k \in \mathbb{N} \) and \( k = (k_0, k_1, \ldots, k_p, 0, 0, \ldots) \), then \( ||k \ast u_+(j)||_\infty \geq C||k||_1 \) for some \( 0 \leq j \leq n_p \) (although the value of \( C \) here is not the same). Without loss of generality we can assume that \( n_p > 0 \). It is also true that we may assume that \( j > p \) if we wish; for suppose that \( k = (k_0, k_1, \ldots, k_p, 0, 0, \ldots) \) is such that \( ||k \ast u_+(j)||_\infty < \epsilon ||k||_1 \) for \( j > p \); now consider the vector \( l_n \) with \( n \) repetitions of \( k \). We find that
\[
||l_n \ast u_+||_\infty < \epsilon ||l_n||_1 + ||u_+||_\infty ||k||_1 < 2\epsilon ||l_n||_1 \tag{11}
\]
for sufficiently large \( n \).

Given data \( (y(0), \ldots, y(n)) \), \( n = n_p \) choose \( \hat{h}_n = (h_0, \ldots, h_p, 0, 0, \ldots) \) to minimize \( \alpha = \max_{y \in \mathbb{C}^n} ||(\hat{h}_n \ast u_+) - y)||_1 \). Write \( P_0 h = (h(0), \ldots, h(p), 0, 0, \ldots) \).

Clearly
\[
\alpha \leq ||l_n||_1 + ||u_+||_\infty ||h - P_0 h||_1 \tag{13}
\]

But
\[
||l_n \ast u_+(j)||_1 \leq \alpha + ||y(j)||_1 + ||h - P_0 h||_1 ||u_+||_\infty \leq \alpha + ||l_n||_1 + ||h - P_0 h||_1 ||u_+||_\infty \tag{14}
\]

Combining these equations, we obtain
\[
||h - P_0 h||_1 \leq (1/C)(2||v||_\infty + 3||l_n||_1 ||u_+||_\infty) \tag{16}
\]
which shows that \( ||h - h_n||_1 \to 0 \) as \( p \to \infty \), i.e., \( n \to \infty \) and \( ||v||_\infty \to 0 \) (if \( n \) is not an \( n_p \) for any \( p \) we can if we wish ignore all data after the largest \( n_p \) \( \leq n \)).

**Necessity:** Since we have no control over \( u_+ \), to obtain the necessity condition we need only prove necessary in the case \( u_+ = 0 \). Robust convergence implies that for all \( n \) sufficiently large and \( \epsilon \) sufficiently small, given any \( (y(0), \ldots, y(n)) \) with \( ||y(j)||_1 \leq \epsilon \) for all \( j \), the model \( h_n \) satisfies \( ||h_n||_1 \leq 1 \), since the output \( y \) could be produced by taking \( h = 0 \) and \( y = (h \ast u_+) + v = v \) and must converge in norm to zero.

If (10) is not satisfied, choose \( h \in \mathbb{N} \) with \( ||h||_1 = 2 \) and with \( ||h \ast u_+||_\infty < \epsilon \). Now given \( y(j) = (h \ast u_+) + v \) for \( j = 0, \ldots, n \), \( n \) large, the identified model \( h_n \) must converge to \( h \) in \( \mathbb{N} \) since \( v = 0 \). But since \( ||h||_1 = 2 \) and \( ||h_n||_1 \leq 1 \) this is a contradiction.

As a result we see that no sequence \( u \in \mathbb{C}^n \), that is a sequence tending to zero, can be used as the input if we require a robustly convergent identification algorithm.

**Corollary 3.1:** If \( u_+ \in \mathbb{C}^n \) then there is no robustly convergent algorithm using measurements of the output \( y = h \ast u + v \).

**Proof:** Given any \( \epsilon > 0 \) there exists an index \( n \) such that \( ||u(t)||_1 < \epsilon/2 \) if \( t \geq n \). Write \( u' = P_0 u = (u(0), \ldots, u(n), 0, \ldots) \).

Let
\[
h = \frac{1}{X}(1/N, 1/X, \ldots, 1/X, 0, 0, \ldots) \tag{17}
\]
with \( N \) nonzero terms. Then \( ||h||_1 = 1 \), but \( ||h \ast u'||_\infty \leq (n + 1)||u'||_\infty/N \) since \( u' \) has only \( n + 1 \) nonzero terms. Hence
\[
||h \ast u'||_\infty \leq \epsilon/(2 + (n + 1)||u'||_\infty/N < \epsilon \tag{18}
\]
for sufficiently large \( N \), which means that condition (10) fails to hold and so there is no robustly convergent algorithm available.

**IV. COPRIME ARX MODELS**

A convenient way to represent the input/output signal dependency of causal linear shift-invariant systems in \( CF(S^\infty) \) is to use ARX models [13]. Thus consider
\[
A(q^{-1})y(t) = B(q^{-1})u(t) + v(t) \tag{19}
\]
where \( q^{-1} \) is the backward shift operator (i.e., \( q^{-1}y(t) = y(t - 1) \) etc.), \( y \) is the output, \( u \) is the input, \( v \) is a bounded disturbance, and \( A(\cdot), B(\cdot), \) interpreted as complex-valued functions of the complex variable \( \cdot \), are functions analytic in the open unit disk with absolutely convergent Fourier series. Here it is usual to take \( A(0) \neq 0 \).

We write (19) now as a convolution operator equation
\[
Ay = Bu + v \tag{20}
\]
where the meaning of the symbols is obvious from (19). \( A \) and \( B \) are \( L^\infty \) stable operators.

A problem with the ARX representation as above is that care must be taken to avoid problems caused by possible appearance of common factors, as these can make the estimation of the unknown system unreliable in the graph topology (common stable minimum phase factors, however, are allowed).

This problem can be avoided if knowledge of some stabilizing controller for the unknown system is available. This is usually a realistic assumption in applications. Thus let \( K_0 \in CF(S^\infty) \) denote a stabilizing controller for the unknown system \( G \in CF(S^\infty) \). Let \( (X_0, Y_0) \) denote a c.f. of \( K_0 \) over \( S^\infty \). Let \( X_0, D_0 \in S^\infty \) be such that the Bezout identity \( X_0H_0 + D_0Y_0 = 1 \) is satisfied. By the Youla parameterization of all plants in \( CF(S^\infty) \) stabilizable by \( K_0 \), we can express (20) as
\[
D_0y - N_0u = R(X_0y + Y_0u) + v \tag{21}
\]
where \( R \in S^\infty \).

The unknown system \( G \) is given by the Youla parameterization
\[
G = (X_0 + RY_0)(D_0 - RX_0)^{-1} \tag{22}
\]
This can be thought of as a reversal of the roles of plant and controller.

Thus we see that “all” we need to do is to identify \( R \) accurately enough in the \( S^\infty \) norm (which is the same as to identify accurately the impulse response of \( R \) in the \( L^1 \) norm), and we are then guaranteed to get a good approximation to the unknown system \( G \) in the \( S^\infty \) graph topology.

The \( R \)-scheme as described above becomes particularly transparent when the input is chosen as follows [9]. Define
\[
u(t) = [K_0(\cdot) - y] + r_2(t) \tag{23}
\]
where \( r_1 \) and \( r_2 \) are bounded reference inputs. The reference inputs \( r_1 \) and \( r_2 \) are act as probing signals to generate sufficient information about the c.f. \( (N, D) \) of the unknown plant, \( N = X_0 + RY_0 \), and \( D = D_0 - RX_0 \). With the above choice of \( u \), we see that
\[
D_0y - N_0u = R(X_0r_1 + Y_0r_2) + v \tag{24}
\]
For our purposes it is useful to make the following choices of \( r_1 \) and \( r_2 \) (cf. [23]). Take \( r_1 = N_0u \) and \( r_2 = D_0w \), where \( w \) is a bounded signal. Then
\[
D_0y - N_0u = Rw + v \tag{25}
\]
Note that the common reference signal generator \( w \) acts now as a simple bounded input in (25). In the sequel we shall refer to the above scheme as the R-scheme.

Let \( A \) denote a closed-loop identification algorithm mapping the assumed experimental information \( Y_n = \{(D_0y)(t) - (N_0u)(t)\}_{t=0}^{n-1} \), \( W_n = \{w(t)\}_{t=0}^{n-1} \) into a model \( A(Y_n, W_n) \in CF(S^\infty) \) of the unknown system \( G \). Here \( n \) is the number of available input/output pairs. We shall say that \( A \) is untuned if it only uses the \emph{a priori} information about the unknown system that it is stabilizable by a known controller \( K_0 \).

**Theorem 4.1:** There exists a bounded reference input \( w \) and an untuned closed-loop identification algorithm \( A \) such that

\[
\lim_{n \to \infty} \sup_{|w(t)|, |\gamma(t)| < \infty} m(G - A(Y_n, W_n)) = 0
\]

for any \( G \in CF(S^\infty) \) stabilized by \( K_0 \in CF(S^\infty) \), and for any \( \alpha \geq 0 \). Here \( m = d, \rho, \beta, \gamma \) or \( \kappa \) (see Section VII).

This result follows directly from the above description of the R-scheme, using the results of [17] and an explicit construction for robustly convergent \( l^1 \) identification algorithms given in Section III; see also [15] and [26], [27].

It is desirable to have some additional prior information about the unknown system \( G \) to derive useful worst-case identification error bounds after a given finite number \( n \) of data points—it suffices to have information which defines a relatively compact set of \( R \) operators; see [15] or [11].

**V. ERROR BOUNDS**

In this section we shall derive error estimates in the various distance functions given in [17] based on 1) \( l^1 \) error bounds for the operator \( R \) in the R-scheme of Section IV and 2) \( l^1 \) modeling error bounds for the factors of a c.f. \((N, D)\) of the unknown system. Combining the derived error bounds with \( l^1 \) identification error bounds for the respective stable operators (which have been considered in [26], [27], [10], [15], [11] as well as in Section III) gives then a general approach for deriving identification error bounds compatible with the notion of robustness of BIBO stability.

Let us suppose that

\[
G = (N_0 + RY_0)/(D_0 - RX_0)
\]

is the true system and that

\[
\hat{G} = (N_0 + \hat{R}Y_0)/(D_0 - \hat{R}X_0)
\]

is the identified model, an approximation to \( G \) produced by taking an estimate \( \hat{R} \) for \( R \). Let \( \nu(G, \hat{G}), \gamma(G, \hat{G}), \rho(G, \hat{G}), \delta(G, \hat{G}) \) denote the subspace gap, the projection gap, the rho function, and the graph metric distance, respectively, for \( G \) and \( \hat{G} \). Detailed definitions of these quantities can be found in [17].

**Theorem 5.1:** Let \( P \) be the projection

\[
P = \begin{bmatrix} D_0 - RX_0 \\ N_0 + RY_0 \end{bmatrix} \begin{bmatrix} w \\ r \end{bmatrix}
\]

and \( \hat{P} \) the corresponding projection using \( \hat{R} \) in place of \( R \). Then

\[
\gamma(G, \hat{G}) \leq \gamma(P - \hat{P})
\]

\[
\leq \| R - \hat{R} \| 1/2 \| X_0 \| 1 + \| Y_0 \| 1 \|^2.
\]

By the triangle inequality, that is

\[
\| XU - NU \| \leq \| NUU' - NU \| + \| U' - I \| + \| X - N \| \| U \| 1 \|
\]

(33)

for any \( U' \) as defined above and \( NUU' - NU \| 1 < 1/2 \). If \( XN - XN \) is small enough this is bounded by \( K \max(A/(\| DA \| 1), \| DA \| 1) \) for some constant \( K \) depending only on \( N \) and \( D \).

**Proof:** This follows immediately from the definitions of the projections \( P \) and \( \hat{P} \), together with the properties of \( \gamma \) and \( \delta \) from [17], on noting that

\[
\| XU \| 1 \left( \frac{2}{\| X \| 1} \| DA \| 1 + \| DA \| 1 \|^2 + (2/\| D \| 1) \| DA \| 1 + \| DA \| 1 \|^2 \right)
\]

(32)

It is also convenient to obtain a bound for the graph metric \( d \), and to do this we require a quantitative version of the theorem in [17] that guarantees the existence of normalized coprime factors. Note that we shall use the notation \( \| H \| 1 \equiv \sum_{\infty}^{\infty} | h_n(\omega) | \) for any \( H(\omega) = \sum_{\infty}^{\infty} h_n(\omega) \) having an a.c.F.s. on the unit circle, i.e., for any \( H \) with \( \| H \| 1 < \infty \).

We recall (cf. [17]) that if \( G = N/D \) is a coprime factorization over \( S^\infty \) and we write \( F(\omega) = 1/(|N(\omega)|^2 + |D(\omega)|^2) \), then \( 1/F \) has an a.c.F.s. If we then write \( \ln F(\omega) = \sum_{\infty}^{\infty} a_n(\omega) e^{i\omega n} \), \( V(\omega) = a_0/2 + \sum_{\infty}^{\infty} a_n(\omega) e^{i\omega n} \), and \( U(\omega) = \exp(V(\omega)) \), then \( G = N/U/D \) is a normalized coprime factorization.

**Theorem 5.2:** Let \( G = N/D \) and \( \hat{G} = \hat{N}/\hat{D} \) be two not necessarily normalized coprime factorizations over \( S^\infty \). Let \( \hat{F} \) (and analogously \( F \)) be defined as above and let \( NU/\hat{D}, \hat{N}/D \hat{U} \) be corresponding normalized coprime factorizations. Let \( XN - XN \) and \( D \) be small enough over \( \| DA \| 1 \leq 1/2 \), if \( XN - XN \) is small enough this is bounded by \( K \max(A/(\| DA \| 1), \| DA \| 1) \) for some constant \( K \) depending only on \( N \) and \( D \).

**Proof:** Note that \( \| DA \| 1 \) is well defined since \( 1/F \) has an a.c.F.s. We have that \( F = \ln F + (1 + D/A/F) \) and

\[
\| I + D/A/F \| 1 \leq 2\| I \| 1 + 2\| D/A/F \| 1 \| \| D/A/F \| 1 < 1/2, \text{ as may be seen by considering the Taylor expansion of } \ln(1 + x).
\]

Now \( U' = UU' \), where \( U' = \exp(V - V) \) and hence, since \( \| V - V \| 1 \leq 2\| I \| 1 + 2\| D/A/F \| 1 \), also we have \( XN - XN \| 1 \leq \exp(\| V - V \| 1) \leq C(\| V - V \| 1) \), where \( C \) is the maximum value of the increasing function \( \exp(x - 1/2) \) in the range \( [0, \| V - V \| 1] \). But if \( \| DA \| 1 \| 1 \leq 1/2 \) then \( \| V - V \| 1 \leq 1/2 \) and so \( C \leq 1/2 \). Thus \( \| U' - I \| 1 \leq 2\| D/A/F \| 1 \).

\[
\| XU - NU \| = \| NUU' - NU \| \leq \| NU \| 1 \| U' - I \| 1 + \| X - N \| \| U \| 1 \|
\]

(34)

by the triangle inequality, that is

\[
\| XU - NU \| \leq \| NU \| 1 \| U \| 1 + \| DA \| 1 \| U \| 1 \|
\]

(35)

Clearly a similar estimate holds for \( \| D \| 1 \), and this implies the required bound.

Finally a bound for \( \| DA \| 1 \) can be obtained showing that it is less than a constant (depending on \( N \) and \( D \)) times the maximum of \( \| XN \| 1 \) and \( \| DA \| 1 \), by noting that

\[
\Delta F = \frac{2N(XN + (\| DA \| 1)^2 + 2D(\| D \| 1 + (\| D \| 1)^2)}{\| X \|^2 + 2N(XN + (\| DA \| 1)^2 + (\| DA \| 1)^2}
\]

(36)

This gives the estimate for \( d(G, \hat{G}) \) in the form required.
Note that as \(\rho(G, \hat{G}) \leq d(G, \hat{G})\) the above result provides also an upper bound for the rho function evaluated for \(G\) and \(\hat{G}\).

An application of this result to the \(R\)-scheme gives

\[
\rho(G, \hat{G}) \leq d(G, \hat{G}) \leq K ||R - \hat{R}||, \text{ max } \{||X_0||, ||Y_0||\} \quad (36)
\]

for some constant \(K > 0\) and for \(||R - \hat{R}||\), small enough.

Consider now the kappa function. We get the following result.

**Theorem 5.6:** Let \(G = XD^{-1}, \hat{G} = \hat{X}D^{-1} \in CF(S^\infty)\), where \((X, D)\) and \((\hat{X}, \hat{D})\) are coprime factorizations over \(S^\infty\). Let \(\Delta X = \hat{X} - X\) and \(\Delta D = \hat{D} - D\) satisfy

\[
\left\| \Delta D \right\| < \left\| \Delta X \right\|. \quad (37)
\]

Then

\[
2\kappa(G, \hat{G}) < \left\| \Delta D \right\| \left( \left\| \Delta X \right\| - \left\| \Delta D \right\| \right)^{-1}. \quad (38)
\]

**Proof:** Denote \(a^{-1} = \left\| \Delta X \right\|\) and \(\hat{a}^{-1} = \left\| \Delta X \right\|\). Then

\[
\left\| D \right\| - \hat{D} \left\| \Delta D \right\| \leq \left| 1 - \hat{a}^{-1} \right| + \hat{a} \left\| \Delta D \right\|. \quad (39)
\]

Now as (the triangle inequality)

\[
\hat{a} < 1/\left( a^{-1} - \left\| \Delta D \right\| \right) \quad (40)
\]

\[
\left| 1 - \hat{a}^{-1} \right| < \left\| \Delta D \right\| \left( a^{-1} - \left\| \Delta D \right\| \right) \quad (41)
\]

the result follows by the definition of the kappa function.

Applying this result to the \(R\)-scheme gives the estimate

\[
2\kappa(G, \hat{G}) < ||R - \hat{R}|| \left( \frac{Y_0}{-X_0} \right) \left( \frac{X_0 + RY_0}{D0 - RX_0} \right)^{-1} \quad (42)
\]

for \(||R - \hat{R}||\), small enough.

**VI. GENERALIZATIONS**

We have so far considered only single-input single-output (SISO) systems. The purpose of this section is to indicate generalizations to multi-input multi-output (MIMO) systems.

Let \(G = XD^{-1}\) be a matrix operator of size \(m \times n\) with elements in \(S^\infty\), and \((X, D)\) a smooth right coprime factorization of \(G\), i.e., \(X\) and \(D\) are BIBO stable matrix operators (so that their elements are in \(S^\infty\)), \(D \neq 0\) such that there exist BIBO stable matrix operators \(X, Y\) so that the right Bezout identity (see e.g., [28]) \(X + YD = I\) is satisfied. We shall use the superscript \(H\) to denote complex conjugate transpose, i.e., \(X^H(e^{j\theta}) = X^\dagger(e^{-j\theta})\) where the superscript \(\dagger\) denotes matrix transpose.

**Theorem 6.1:** Let \(G = XD^{-1}\), where \((X, D)\) is a smooth right coprime factorization of \(G\) in the sense defined above. Then there exists a smooth normalized coprime factorization \((M, E)\) of \(G\), i.e., a coprime factorization of \(G = ME^{-1}\) such that \(M, E\) have elements in \(S^\infty\) and

\[
M^H(e^{j\theta})M(e^{j\theta}) + E^H(e^{j\theta})E(e^{j\theta}) = I \quad (43)
\]

for any \(\theta \in [0, 2\pi]\).

**Proof:** Note that \(\tilde{F}^{-1} = X^H\hat{X} + D^H\hat{D}\) has a matrix-valued a.c.F.s, and that \(det(\tilde{F}^{-1})\) is bounded away from zero on the unit circle. It follows by Wiener’s famous result [12, p. 202] that \(\tilde{F}\) has an a.c.F.s. Furthermore, \(\tilde{F}\) is positive definite. Thus we can define the matrix logarithm \(\ln \tilde{F}\) of \(\tilde{F}\). (The matrix exponential of \(\ln \tilde{F}\) satisfies \(exp(\ln \tilde{F}) = \tilde{F}\).) The logarithm \(\ln \tilde{F}\) has an a.c.F.s by the Wiener-Levy theorem [12, p. 210]. Denote \(\ln \tilde{F}(e^{j\theta}) = \sum_{k=-\infty}^{\infty} A_k e^{jk\theta}\).

Define

\[
\hat{V}(z) = A_0 + \sum_{k \neq 0} A_k z^k \quad (44)
\]

The matrix-valued function \(\hat{V}(z)\) is analytic in the open unit disk and has an a.c.F.s. Define the matrix exponential \(\hat{U} = \exp \hat{V}(z)\). Note that \(\hat{U}^{-1}\) is analytic in the open unit disk and has a matrix-valued a.c.F.s. Set \(M = NU\) and \(E = DU\). Now

\[
\ln \tilde{F}(e^{j\theta}) = \hat{V}(e^{j\theta}) + \hat{V}^H(e^{j\theta}) \quad (45)
\]

so that

\[
(\hat{U}^{-1})^X(e^{j\theta}) = \exp [\hat{V} + \hat{V}^H](e^{j\theta}) = \hat{F}(e^{j\theta}) \quad (46)
\]

Thus

\[
M^H\hat{M} + \hat{E}^H\hat{E} = \hat{U}^H(X^H\hat{X} + D^H\hat{D})\hat{U} = \hat{U}^H\hat{F}^{-1}\hat{U} = I \quad (47)
\]

so that \((M, E)\) is the required smooth normalized coprime factorization of \(G\). This completes the proof of the Theorem.

It is then possible to generalize the various distance functions and to obtain error estimates for MIMO systems in a way rather similar to those given in Section V.

**VII. CONCLUSIONS**

A fundamental characterization of inputs which guarantee the existence of robustly convergent \(\tilde{l}\) identification algorithms for BIBO stable systems has been given. We have applied this to worst-case analysis of identification of feedback systems from closed-loop time series measuring identification error with distance functions which lead to the weakest convergence notions in which feedback stability is a robust property. Bounds on the worst-case identification error in these distance functions have been obtained through \(\tilde{l}\) identification of certain stable operators.

**REFERENCES**


