Tutorial Paper

Beyond Singular Values and Loop Shapes

Gunter Stein
Honeywell Systems and Research Center, Minneapolis, Minnesota 55418
and
John C. Doyle
California Institute of Technology, Pasadena, California 91125

This paper provides a tutorial look at the status of singular value loop shaping as a paradigm for multivariable feedback design. It shows that this paradigm is effective whenever a design problem's specifications are spatially round, but that it can be arbitrarily conservative otherwise. This happens because singular value conditions for robust performance are not tight (sufficient but not necessary) and can severely overstate actual requirements. An alternate paradigm is discussed that promises to overcome these limitations. This alternative includes a more general problem formulation, a new matrix function *mu*, and tight conditions for both robust stability and robust performance. The current state of development of this paradigm supports analysis and comparisons of existing feedback designs. To a limited extent, it also supports formal mathematical synthesis of new optimal designs, although much research remains to be done in the synthesis area.

I. Introduction

EVER since the basic work of Nyquist,[1] Bode,[2] and others, the classical approach to feedback design has followed a frequency domain perspective—given a plant described by rational transfer function $G(s)$, design a rational compensator $K(s)$, such that the closed-loop feedback system meets three basic design requirements—stability, performance, and robustness.

As is well known, the stability requirement imposes functional constraints on certain transfer functions of the closed-loop system, e.g., a Nyquist encirclement count for the function $\det(I + GK)$.[3] Likewise, the performance requirement imposes magnitude constraints on certain other transfer functions. For example, one of the most frequently encountered performance requirements is to keep output errors small in the face of commands and (output) disturbances. This requires the output sensitivity function

$$S(s) \triangleq [I + G(s)K(s)]^{-1} \tag{1}$$

to be small for all frequencies, $s = j\omega$, where the disturbances and/or reference commands are large.

The third feedback design requirement, robustness, calls for the first two requirements to be achieved not only for the nominal plant but also for an entire set of neighboring plants that arise from the inevitable presence of modeling errors and plant uncertainties. This imposes still other magnitude constraints on transfer functions. For example, a commonly used model of plant uncertainties is the so-called unstructured multiplicative perturbation at the output. Stability for this uncertainty model requires that the complementary output sensitivity function

$$T(s) \triangleq G(s)K(s)[I + G(s)K(s)]^{-1} \tag{2}$$

be small for all frequencies where the uncertainties are large.[4] However, since $S(s) + T(s) = I$, both S and T cannot be small simultaneously. Thus, specifications on Eqs. (1) and (2) lead to one of the most basic tradeoffs in feedback design.

For classical single-input/single-output (SISO) systems, the meanings of "small" and "large" in this basic tradeoff are, of course, understood in terms of the absolute values of the respective complex-valued functions at each frequency. Hence, SISO designers working in the frequency domain have viewed the design problem as one of altering the shapes of (Bode) magnitude plots of sensitivity and complementary sensitivity functions to meet design specifications. Indeed, because $|S(s)| \approx 1/|GK(s)|$ whenever the loop transfer function $GK(s)$ is large, and $|T(s)| \approx |GK(s)|$ whenever $GK(s)$ is small, the shapes of these magnitude functions are intimately tied to the shape of $GK(s)$, and the entire design process is often referred to simply as "loop shaping."

Over the last few years, the loop-shaping process has been successfully formalized and generalized to multi-input/multi-output (MIMO) design problems.[4] Key ingredients of the generalization include 1) the use of singular values as appropriate measures of magnitude for matrix-valued transfer functions, 2) the development of formal mathematical conditions that guarantee robust stability and robust performance of MIMO feedback systems in terms of these magnitude measures, and 3) the development of design procedures (e.g., LQG/LTR[5] and H_∞^{6-8}) that help to synthesize desired multivariable loop shapes.

Design experience with these new results shows that singular value loop shaping is an effective MIMO design paradigm only for problems whose objectives can be reduced to spatially round specifications on $S(s)$ and $T(s)$ alone. Unfortunately, many design requirements that arise in MIMO problems cannot be usefully expressed in this form. This paper identifies some of these latter requirements and describes a design framework and certain recently developed tools that promise to deal with them more effectively.

The paper begins in Sec. II with a brief review of the singular value loop-shaping paradigm. It then discusses skewed-specification design problems that are not easily handled by this process in Sec. III. Two dubious solutions for such problems—plant inversion and round controls—are examined in detail. Finally, a more general design framework and associated research results that promise to address skewed problems more effectively are presented in Sec. IV. Section V provides concluding comments. The paper presents no new theoretical results and should be viewed as a tutorial look at the current status of this branch of frequency domain MIMO design.

Received Sept. 26, 1988.

*Corporate Fellow; also, Adjunct Professor, EECS, Massachusetts Institute of Technology, Cambridge, MA 02139.
†Associate Professor, Electrical Engineering.

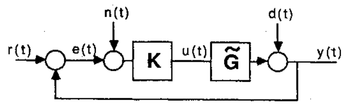

Fig. 1 Generic feedback design problem.

II. Singular Value Loop-Shaping Paradigm

Our generic multivariable feedback design problem is illustrated in Fig. 1. The loop consists of a plant and a compensator in a unity feedback arrangement. The plant can be any element from a set of plants designated by $\tilde{G}$. Each system in $\tilde{G}$ is assumed to be linear, finite-dimensional, time invariant, and can thus be modeled by a rational transfer function matrix $\tilde{G}(s)$. Similarly, the compensator K is modeled by a rational $K(s)$. The design problem is to find a K that makes the feedback loop internally stable for all possible plants $\tilde{G}$ and causes it to respond well to various external signs. "Responding well" typically means that the plant's outputs accurately track external commands $r(t)$, even in the face of disturbances, $d(t)$ and sensor noise $n(t)$ and that the plant's inputs and/or other selected signals remain within applicable limits. (Throughout the paper, $x(t)$ and $x(s)$ will designate time functions and their Laplace transforms, respectively.)

A. Formal Design Problem

This MIMO design problem can be formalized by specifying precise mathematical statements for the aforementioned qualitative performance objectives and by specifying the set of plants $\tilde{G}$ over which those objectives must be achieved. We will start with some very simple specifications for these elements of the design problem. More complex situations are treated later.

Performance Objectives

As formal performance objectives, we will require selected transfer functions of the closed-loop system to be sufficiently small in a weighted H_∞-norm sense. To begin with, the selected functions will consist only of the sensitivity function, and our mathematical performance specification will be the following:

$$\sup_\omega \sigma_{\max}[w(j\omega)\tilde{S}(j\omega)] < 1 \tag{3}$$

Here $\tilde{S}(s)$ denotes the sensitivity function produced by any plant from a plant set to be defined, $w(j\omega)$ is the (scalar) transfer function of a specified stable, minimum phase, invertible system wI, and $\sigma_{\max}[\,]$ denotes the largest singular value of its matrix-valued argument.

Qualitatively, specification (3) expresses a desire to keep the sensitivity function of the closed-loop system smaller than the function $1/w(j\omega)$ for all plants and all frequencies. This desire can arise from various sources. For example, we may need to suppress sinusoidal disturbances with known magnitudes, $\|d(j\omega)\| = D(\omega)$, down to specified residual error levels, $\|y(j\omega)\| < \delta$. Then, because $y = \tilde{S}d$ in the absence of other signals, $w(j\omega)$ would be chosen such that $|w(j\omega)| = D(\omega)/\delta$. Alternatively, if we need to track commands, $\|r(j\omega)\| = R(\omega)$ with tracking errors $\|y(j\omega) - r(j\omega)\| < \delta$, then $w(j\omega)$ would be chosen as $|w(j\omega)| = R(\omega)/\delta$, and more generally if disturbances and commands exist simultaneously, $w(j\omega)$ would be chosen as $|w(j\omega)| = [R(\omega) + D(\omega)]/\delta$.

Specification (3) can also be motivated more formally for signals other than sinusoids. For example, if disturbance $d(t)$ is assumed to be generated by any function from the L_2-unit ball, passed through linear system wI, then Eq. (3) is a necessary and sufficient condition that the residual error $y(t)$ again belongs to the L_2-unit ball.[9] Thus, Eq. (3) makes L_2 errors small for sets of L_2 disturbances, and analogously for sets of L_2 commands.

Similar formal arguments can also be made for disturbances and commands derived from sets of constant power signals and from stochastic signals whose power spectra must be reduced to spectra below specified bounds. The implication is that H_∞-norm specifications such as Eq. (3) provide a useful way to express a variety of disparate performance objectives. Since several such objectives typically arise simultaneously in specific engineering situations, the actual transfer functions selected in Eq. (3) and the specified weighting functions $w(j\omega)$ will generally represent a composite of requirements with no single formal mathematical origin.

Set of Plants

Given that we are using the size of transfer functions to describe performance requirements, it is convenient to use the same concept to describe sets of plants. As a starting point, consider $\tilde{G}$ to be generated by the unstructured multiplicative perturbations mentioned in the introduction. That is,

$$\tilde{G} = [I + w_\Delta \Delta']G \tag{4}$$

where G is a nominal plant value, Δ' is any stable perturbation system whose transfer function satisfies

$$\sup_\omega \sigma_{\max}[\Delta'(j\omega)] \le 1 \tag{5}$$

and w_Δ is a specified stable minimum phase invertible system, WI, scalar times identity. (The stability assumption on Δ' can be relaxed to the condition that G and $(I + \Delta')G$ have the same number of unstable modes[4] or, more generally, that they are path connected in an appropriate graph topology.[10] Throughout the paper, the symbol $(\;)'$ will be used to indicate normalized quantities, e.g., unit-norm systems or signals whose actual or desired norms are bounded by unity.)

For SISO systems, this characterization of $\tilde{G}$ includes all systems whose frequency responses, at each frequency ω, fall into a circle with radius $|w_\Delta(j\omega)G(j\omega)|$ centered about a nominal model, $G(j\omega)$. The magnitude $|w_\Delta(j\omega)|$ is the normalized error magnitude of the nominal model. In engineering situations, this plant set proves useful for describing those model uncertainties that remain after systematic parameter variations (structured errors) are accounted for. Such remaining uncertainties are generally due to unknown and/or neglected dynamics, resonances, time delays, nonlinearities, etc. They are usually small at low frequencies ($|w(j\omega)| < 1$), but invariably grow much larger than unity ($|w(j\omega)| > 1$, errors greater than 100%) as frequency increases.

B. Loop-Shape Specifications

The preceding performance objectives and plant sets imply that the shapes of certain singular value magnitudes, as functions of frequency, must be appropriately constrained. In particular, suppose we are dealing only with the nominal plant, i.e., $\tilde{G} \equiv G$. Then, according to Eq. (3), the Bode plot of $\sigma_{\max}[S(j\omega)]$ must fall below the Bode plot of $1/|w(j\omega)|$ over the entire frequency range. That is,

$$\sigma_{\max}[S(j\omega)] < \frac{1}{|w(j\omega)|} \quad \text{for all } \omega \tag{6}$$

Moreover, since $S = (I + GK)^{-1}$, the loop GK must be shaped such that

$$\sigma_{\min}[1 + GK(j\omega)] > |w(j\omega)| \quad \text{for all } \omega \tag{7}$$

and that

$$\sigma_{\min}[GK(j\omega)] > |w(j\omega)| \quad \text{whenever } |w(j\omega)| \gg 1 \tag{8}$$

where $\sigma_{\min}[\,]$ denotes the smallest singular value of its argument.

Equations (6–8) define loop-shape specifications that any stabilizing compensator $K(s)$ must satisfy in order to achieve nominal performance. Of course, since $\tilde{G}$ will not always equal G, $K(s)$ must also satisfy two other requirements. It must maintain stability for all elements in the defined set of plants (this will be called "robust stability") and it must satisfy the performance objective for all plants (this will be called "robust performance"). These additional requirements also impose specifications on loop shapes, as summarized by the following analysis theorem.

Theorem 1—Robust Stability and Performance for Equations (3–5)

Suppose that the nominal feedback system in Fig. 1 is stable (i.e., it is stable with $\tilde{G} \equiv G$). Then 1) the perturbed system is stable for all plants defined by Eqs. (4) and (5) if and only if

$$\sigma_{\max}[w_\Delta(j\omega)T(j\omega)] < 1 \qquad \text{for all } \omega \tag{9}$$

where $T(s)$ is the complementary sensitivity function defined by Eq. (2), and 2) the perturbed system satisfies specification (3) for all plants defined by Eqs. (4) and (5)

$$\sigma_{\max}[w(j\omega)S(j\omega)] < [1 - \sigma_{\max}[w_\Delta(j\omega)T(j\omega)]] \qquad \text{for all } \omega \tag{10}$$

and it satisfies the objective for all plants only if

$$\sigma_{\max}[w(j\omega)S(j\omega)] < [1 - \sigma_{\min}[w_\Delta(j\omega)T(j\omega)]] \qquad \text{for all } \omega \tag{11}$$

The first result follows from the fact that the feedback loop remains stable if and only if the function

$$\det[I + (I + w_\Delta\Delta')GK] = \det[I + GK]\,\det[I + w_\Delta\Delta'T] \tag{12}$$

remains nonzero along the $j\omega$ axis (and therefore in the right half plane) for all Δ'. The second result can be derived directly from the following expression for the sensitivity function produced by individual plants in the plant set

$$\tilde{S}(s) \triangleq \{I + [I + w_\Delta(s)\Delta'(s)]G(s)K(s)\}^{-1}$$
$$= S(s)[I + w_\Delta(s)\Delta'(s)T(s)]^{-1} \tag{13}$$

This expression for $\tilde{S}(s)$ shows that performance is maintained in the face of $\Delta'(s)$ whenever the nominal performance requirement on $S(s)$ is tightened sufficiently to offset any amplification from the factor $[I + w_\Delta\Delta'T]^{-1}$. Equations (10) and (11) simply reflect the worst- and best-case values this amplification can take when maximized over Δ'. Note that these equations reduce to a single necessary and sufficient condition for SISO systems and also for MIMO systems whose complementary sensitivity functions have approximately equal maximum and minimum singular values, i.e., whose condition numbers $\kappa[T]$ satisfy

$$\kappa[T] \triangleq \sigma_{\max}[T]/\sigma_{\min}[T] \approx 1 \tag{14}$$

Matrices that satisfy the property of Eq. (14) will be called "spatially round." This term is motivated by interpretations of matrices as mappings that, in general, transform unit spheres into ellipsoidal objects. Spatially round matrices map unit spheres into other spheres, scaled in magnitude but not altered in shape. Such matrices include scalars times identity and also scalars times any product of unitary matrices.

Even without spatially round $T(s)$, however, the sufficient condition (10) alone is not unduly conservative. This is so because the factor $1 - \sigma_{\max}[w_\Delta T]$ in Eq. (10) must be positive for robust stability and is typically designed to be 0.5 or greater to provide some design margin. Also, the factor $1 - \sigma_{\min}[w_\Delta T]$ in Eq. (11) is never greater than unity. Hence, the true robust performance requirement is typically overstated by less than a factor of 2.

The significance of theorem 1 is that it defines robust stability and robust performance solely in terms of acceptable shapes for the nominal functions $\sigma_{\max}[S(j\omega)]$, $\sigma_{\max}[T(j\omega)]$, and $\sigma_{\max}[GK(j\omega)]$. In particular, from Eq. (9), robust stability is guaranteed if and only if

$$\sigma_{\max}[T(j\omega)] < \frac{1}{|w_\Delta(j\omega)|} \qquad \text{for all } \omega \tag{15a}$$

and, because $T = GK(I + GK)^{-1}$, this is assured if

$$\sigma_{\max}[GK(j\omega)] < \frac{1}{|w_\Delta(j\omega)| - 1} \approx \frac{1}{|w_\Delta(j\omega)|} \qquad \text{whenever } |w_\Delta(j\omega)| \gg 1 \tag{15b}$$

Similarly, from Eqs. (10) and (11), robust performance is guaranteed if and "almost only if" (to within a factor of 2) a linear combination of $\sigma_{\max}[S]$ and $\sigma_{\max}[T]$ is small enough, namely,

$$\sigma_{\max}[S(j\omega)] < \frac{1 - |w_\Delta(j\omega)|\sigma_{\max}[T(j\omega)]}{|w(j\omega)|} \qquad \text{for all } \omega \tag{16}$$

Note that this requirement has the same form as the nominal performance specification [Eq. (6)] but is tightened up sufficiently to account for the presence of $\Delta'(s)$. Thus, to design for robust performance under the loop-shaping paradigm, we simply design for nominal performance against a tightened specification. We must, of course, also design for stability robustness, using specification (15).

C. Synthesis Methods

In terms of the preceding paradigm, feedback synthesis amounts to finding compensators $K(s)$ that stabilize the nominal plant and shape the feedback loop's magnitude functions such that applicable loop-shape specifications are satisfied. Several classical textbooks describe semiformal, but effective, procedures for accomplishing these loop-shaping goals for SISO systems (for example, Ref. 11). Various attempts have been made over the years to generalize these procedures to MIMO problems with varying degrees of success (e.g., inverse Nyquist methods applied to diagonally dominant systems,[12] direct-Nyquist and Bode methods applied to characteristic loci,[13] root-locus methods applied to multivariable functions,[14] etc.).

It turns out that some of the most effective methods for shaping MIMO loops use modern optimization-based synthesis tools. For example, H_∞ methods[6–8] have recently been developed to compute compensators that directly minimize weighted sensitivity criteria,

$$J = \sup_\omega \sigma_{\max}[w(j\omega)S(j\omega)]$$

or weighted complementary sensitivity criteria,

$$J = \sup_\omega \sigma_{\max}[w_\Delta(j\omega)S(j\omega)]$$

or so-called "mixed sensitivity" criteria constructed by augmenting S and T,

$$J = \sup_\omega \sigma_{\max}[w(j\omega)S(j\omega) \quad w_\Delta(j\omega)T(j\omega)] \tag{17}$$

Similarly, modified versions of the LQG problem (i.e., LQG/LTR) are available to minimize mixed sensitivity criteria in an H_2 sense[5]:

$$J = \frac{1}{\pi}\int_0^\infty \mathrm{Tr}[|w|^2 SS^* + TT^*]\,\mathrm{d}\omega \tag{18}$$

We note that none of these criteria precisely matches the robust performance specification of Eq. (16). Certainly, sensitivities or complementary sensitivities alone do not do so, and even mixed sensitivites do not match the linear combination of $\sigma_{\max}[S]$ and $\sigma_{\max}[T]$ required by Eq. (16). Nevertheless, the mixed sensitivity problems have proven quite useful in design applications.

The various synthesis options will not be discussed further here. Rather, our objective is to identify design requirements that do not lend themselves readily to the overall MIMO loop-shaping paradigm.

III. More Complex Problems

The design problem described previously can be generalized and made more versatile for design purposes by including more complex performance requirements and/or plant sets. Some of these generalizations still fit nicely under the loop-shaping paradigm. Examples include performance specifications on $T(s)$ due to sensor noise, bandwidth limits, or control signal saturations. There are other generalizations, however, that do not fit as well. This section describes some of them.

A. Problems for which Loop-Shape Specifications Fail

Many engineering design situations give rise to performance objectives and/or plant sets that are not spatially round. Consider the following prototype.

Peformance Requirements

Let the weighted H_∞-norm performance objective in Eq. (3) be generalized to

$$\sup_{\omega} \sigma_{\max}[W_e(j\omega)\tilde{S}(j\omega)W_d(j\omega)^{-1}] < 1 \tag{19}$$

where $W_e(s)$ and $W_d(s)^{-1}$ are specified transfer function matrices. $W_e(s)$ shapes the frequency content and weights spatial directions of the errors we wish to keep small, and $W_d(s)^{-1}$ shapes frequency content and weights spatial directions of external disturbances and/or commands. As before, both W_e and W_d^{-1} correspond to stable, minimum phase, invertible dynamic systems. Examples will be given shortly.

Set of Plants

Let the plants $\tilde{G}$ correspond to a generalization of the unstructured multiplicative perturbations in Eq. (4), namely,

$$\tilde{G} = [I + W_v^{-1}\Delta' W_z]G \tag{20}$$

where Δ' is again any stable perturbation operator with norm less than or equal to unity, and W_v^{-1} and W_z are stable, minimum phase, invertible systems which shape spatial directions and frequency content of inputs and outputs of the unit-norm perturbation.

The principal MIMO loop-shape requirements for this new problem description are given by the following theorem.

Theorem 2—Loop-Shape Specifications for Equations (19) and (20)

1) Nominal performance is satisfied if and only if

$$\sigma_{\max}[W_e(j\omega)S(j\omega)W_d(j\omega)^{-1}] < 1 \quad \text{for all } \omega \tag{21}$$

2) Stability is robust if and only if

$$\sigma_{\max}[W_z(j\omega)T(j\omega)W_v(j\omega)^{-1}] < 1 \quad \text{for all } \omega \tag{22}$$

3) Performance is robust if (not only if)

$$\sigma_{\max}[W_e S W_d^{-1}] < \frac{1 - \sigma_{\max}[W_z T W_v^{-1}]}{\kappa[W_v W_d^{-1}]} \quad \text{for all } \omega \tag{23}$$

These results are generalizations of results in Sec. II and reduce back to Sec. II when the weights are specialized to be spatially round (e.g., $W_e = wI$, $W_d = I$, $W_z = w_\Delta I$, and $W_v = I$). In particular, the first result corresponds to Eq. (6) and, like Eq. (6), is true by definition [e.g., a consequence of Eq. (19)]. The second result corresponds to Eq. (9) and is derived via stability arguments analogous to Eq. (12). The last result corresponds to Eqs. (10) and (11) and is based on the following weighted generalization of the sensitivity function in Eq. (13)

$$\begin{aligned}\tilde{M}(s) &= W_e(s)\tilde{S}(s)W_d(s)^{-1}\\ &= W_e S[I + W_v^{-1}\Delta' W_z T]^{-1}W_d^{-1}\\ &= W_e S W_d^{-1} W_d[I + W_v^{-1}\Delta' W_z T]^{-1}W_d^{-1}\\ &= (W_e S W_d^{-1})(W_v W_d^{-1})^{-1}[I + \Delta'(W_z T W_v^{-1})]^{-1}(W_v W_d^{-1})\end{aligned} \tag{24}$$

Equation (24) shows that the tightest singular value bound that we can place on $\tilde{M}(s)$ for all $\Delta'(s)$ is $\sigma_{\max}[W_e S W_d^{-1}]$ $\kappa[W_v W_d^{-1}]/(1 - \sigma_{\max}[W_z T W_v^{-1}])$. This bound leads directly to condition (23), and thus Eq. (23) is the least conservative sufficient condition for robust performance that can be established via singular values.

Given the preceding requirements, the loop-shaping design process presumably proceeds just as in Sec. II. We design for robust stability using Eq. (22) and for robust performance using nominal performance specification (21), but tightened to Eq. (23) to account for uncertainties. Unfortunately, the last step runs into difficulties. Even though Eq. (23) is the least conservative singular value bound, it can be arbitrarily poor for some design situations—requiring much tighter (often unachievably tight) nominal performance than is actually necessary to assure robust performance. On the other hand, we will see as follows that there are also design situations for which Eq. (23) is not only sufficient but nearly necessary as well. Because Theorem 2 provides no way to distinguish one situation from the other, designers are forced to assume the worst and, therefore, to settle for potentially very conservative outcomes. This predicament exists whenever the weighted external signals and/or plant perturbations produce large condition numbers $\kappa[W_v W_d^{-1}]$, i.e., whenever the problem specifications are not spatially round. The predicament is illustrated with a simple example below.

B. Example with Skewed Specifications

As mentioned earlier, problem specifications that are not spatially round occur frequently and naturally in many engineering situations. As an example, consider the class of basic feedback design problems illustrated in Fig. 2. We want to keep output errors small ($\sup_\omega \sigma_{\max}[wS]$ less than unity) for all plants in a set generated by unstructured multiplicative perturbations occurring at the plant input, not at the output as in Sec. II. This plant set is described by

$$\tilde{G} = G[I + w_\Delta \bar{\Delta}'] \tag{25}$$

where $\bar{\Delta}'$ again satisfies the assumption of Eqs. (4) and (5). A few manipulations show that the corresponding multiplicative perturbation at the output is given by the following.

$$\Delta' = G\bar{\Delta}' G^{-1} \tag{26}$$

which leads to the following weights for Eqs. (19) and (20).

$$W_e(s) = w(s)I; \qquad W_d(s)^{-1} = I \tag{27}$$

$$W_z(s) = G(s)^{-1}; \qquad W_v(s)^{-1} = w_\Delta(s)G(s) \tag{28}$$

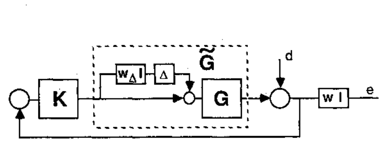

Fig. 2 A design problem with skewed specifications.

Evidently, these specifications are not spatially round whenever the condition number of G is large. (For simplicity, assume that G is stable, minimum phase, and without singularities on the $j\omega$ axis.)

C. Dubious Control Solution for Skewed Problems

The preceding skewed design problem arises because round performance requirements and round uncertainty requirements are levied at different points of the feedback system—performance required at the plant's output, uncertainties applied at its input and visa versa. The problem derives its skewness from the skewness of $G(s)$. It if often tempting to solve this class of skewed problems with plant-inverting controllers that achieve good (identical) loop shapes at both the input and output simultaneously. This is a dubious solution, however, because identical loop shapes typically satisfy only two of the specifications in theorem 2, not all three.

The two specifications that can be readily satisfied by plant inversion are nominal performance and robust stability. This is easy to verify for the example. Consider the following compensator:

$$K(s) = G(s)^{-1}\lambda(s) \tag{29}$$

where $\lambda(s)$ is a scalar loop-transfer function that makes $K(s)$ proper and stabilizes the closed loop. This compensator produces diagonal sensitivity and complementary sensitivity functions with identical diagonal elements, namely,

$$S = \frac{1}{(1+\lambda)}I, \qquad T = \frac{\lambda}{(1+\lambda)}I \tag{30}$$

Substituting these expressions into Eqs. (21) and (22) shows how nominal performance and robust stability can indeed be achieved for the example. We simply choose $\lambda(s)$ to satisfy

$$\sigma_{\max}[W_e S W_d^{-1}] = \sigma_{\max}[wS] = \frac{|w(j\omega)|}{|1+\lambda(j\omega)|} < 1 \quad \text{for all } \omega \tag{31}$$

$$\sigma_{\max}[W_z T W_v^{-1}] = \sigma_{\max}[G^{-1}Tw_\Delta G] = \frac{|w_\Delta(j\omega)\lambda(j\omega)|}{|1+\lambda(j\omega)|} < 1 \quad \text{for all } \omega \tag{32}$$

Note that choosing $K(s)$ in this way amounts to solving SISO loop-shaping problems. Solutions exist whenever the weights $w(j\omega)$ and $w_\Delta(j\omega)$ are sized to be compatible with the basic $S+T=I$ transfer function constraint.

Next, consider the robust performance specification of Eq. (23). For the example, this specification is tighter than Eq. (21) by at least the factor

$$\kappa[W_v W_d^{-1}] = \kappa[G] \tag{33}$$

and requires $|\lambda(j\omega)|$ in Eq. (31) to be increased accordingly. However, at frequencies where $|w_\Delta(j\omega)|$ exceeds unity, large increases of $|\lambda(j\omega)|$ will surely violate robust stability, e.g., Eq. (32). Hence, for reasonably tight requirements $w(j\omega)$ and $w_\Delta(j\omega)$ and large condition numbers $\kappa[G(j\omega)]$, we will be unable to find plant-inverting compensators to satisfy the sufficient condition. Moreover, when this situation arises, we cannot appeal to excessive conservatism of Eq. (23) to save the design. Such appeals fail because Eq. (23), when restricted to plant-inverting compensators, is not only sufficient for robust performance but almost necessary as well. This fact follows from Eq. (24), which takes a much simpler form for our example with plant-inverting controls, namely,

$$\tilde{M}(s) = wsG[I + \Delta' w_\Delta t]^{-1}G^{-1} \tag{34}$$

where ws and $w_\Delta t$ are the scalar-weighted sensitivity and complementary sensitivity functions achieved by the controller [i.e., as given by Eqs. (30–32)]. Equation (34) shows that plant perturbations can cause performance deteriorations from the nominal $|ws|$ by a factor as large as

$$F = \sup_{\Delta} \sigma_{\max}[G(I+\Delta' w_\Delta t)^{-1}G^{-1}] = \left[\inf_{\Delta'} \sigma_{\min}[G(I+\Delta' w_\Delta t)G^{-1}]\right]^{-1} \tag{35}$$

This factor is bounded from above and below by the following limits:

$$\frac{|w_\Delta t|}{1+|w_\Delta t|}\frac{\kappa[G]}{1-|w_\Delta t|} \le F \le \frac{\kappa[G]}{1-|w_\Delta t|} \tag{36}$$

The upper limit can be derived directly from singular value calculations, as a special case of Eq. (23), and the lower limit is derived later in Sec. IV using the structured singular value. (The lower limit is valid only for $\kappa[G]$ sufficiently large. It can also be derived directly from Eq. (35) by finding the smallest amplification of $G(I + w_\Delta \Delta' t)G^{-1}$ over input vectors of the form $x = \alpha v_1 + \beta v_m$, $\alpha^2 + \beta^2 = 1$, where v_1 and v_m are the right singular vectors of G^{-1} corresponding to the largest and smallest singular values, respectively.) Note that the lower limit differs from the upper limit only by the scalar $|w_\Delta t|/(1+|w_\Delta t|)$ whose value can be as large as 0.5 for small stability margins (e.g., $|w_\Delta t| \approx 1$). More typical values are 0.33 for $|w_\Delta t| \approx 0.5$, and really small values do not occur unless w_Δ itself is very small. Hence, the worst-case performance deterioration anticipated by Eq. (23) is nearly realized by plant-inverting controls.

D. Another Dubious Control Solution

Plant inversion may be viewed as one extreme of possible solutions for skewed problems, namely, the extreme that removes all skewness of $G(s)$ with exactly opposite skewness in the controller. Another extreme of possible solutions would be to remove none of the skewness of $G(s)$ by constraining the compensator to satisfy $\kappa[K(s)] \approx 1$ at all s. This *spatially round control* option is also a dubious solution because it is rarely able to satisfy the nominal performance specification of Eq. (21) and the stability robustness specification of Eq. (22) simultaneously. However, round controllers do have the significant advantage that, in the event Eqs. (21) and (22) can be satisfied, the robust performance specification (23) no longer needs to be considered because it is known to be excessively conservative for this design situation.

The conservatism of Eq. (23) for round controllers is again easy to establish for our example. Starting with Eqs. (24) and (28), $\tilde{M}(s)$ reduces to the simpler form

$$\tilde{M}(s) = (W_e S W_d^{-1})G[1 + \Delta' G^{-1}Tw_\Delta G]^{-1}G^{-1} = (W_e S W_d^{-1})[I + \Delta' T w_\Delta]^{-1} \tag{37}$$

It can therefore be bounded by

$$\sigma_{\max}[\tilde{M}] \le \frac{\sigma_{\max}[W_e S W_d^{-1}]}{1-\sigma_{\max}[w_\Delta T]} = \frac{\sigma_{\max}[W_e S W_d^{-1}]}{1-\sigma_{\max}[w_\Delta G^{-1}TG]} = \frac{\sigma_{\max}[W_e S W_d^{-1}]}{1-\sigma_{\max}[W_z T W_v^{-1}]} \tag{38}$$

Here the second to last step uses the facts that $\sigma_{\max}[T] = \sigma_{\max}[KTK^{-1}]$ for spatially round K and that $KTK^{-1} = G^{-1}TG$. The last step merely resubstitutes Eq. (28).

Note that the relation $KTK^{-1} = G^{-1}TG$ makes Eq. (38) valid whenever either K or G is spatially round. In either case, therefore, $\sigma_{\max}[M]$ can be kept below unity by a more reason-

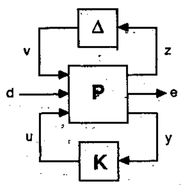

Fig. 3 General problem description.

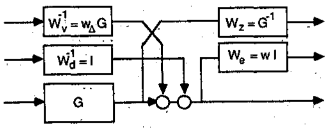

Fig. 4 *P*-system for example in Sec. III.

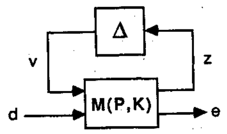

Fig. 5 Closed loop *M*-system.

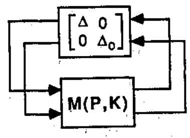

Fig. 6 Stability condition for robust performance.

able sufficient condition not involving the condition number of G:

$$\sigma_{\max}[W_e S W_d^{-1}] < 1 - \sigma_{\max}[W_z T W_v^{-1}]$$

E. Implications

The preceding discussions illustrate that theorem 2, with Eq. (23), provides only a limited tool for MIMO feedback design. For some controllers, Eq. (23) is necessary but difficult to satisfy. For other controllers, it is unnecessary, but other conditions [Eqs. (21) and (22)] are difficult to satisfy. The theorem provides no direct way to distinguish these cases or to identify a middle ground between them where effective control solutions might be found.

Overall, this implies that singular value loop shaping is inadequate as a paradigm for general MIMO feedback design and that we must seek more powerful alternatives. One such alternative, currently in the research stage, is described in the next section.

IV. Alternate MIMO Design Framework

The limitations previously described can be overcome in part with an alternate design framework that has been developed over the last few years.[15,16] This alternate framework consists of a more general problem description, a more suitable measure of magnitude for matrix transfer functions, and certain key analysis and synthesis results. These various elements are described briefly in this section.

A. Problem Description

The generalized problem is illustrated in Fig. 3. It consists of a general system $\boldsymbol{P}$ with three pairs of input/output variables. The first pair consists of measured outputs $y(t)$ and control inputs $u(t)$. The second pair consists of performance variables $e'(t)$ and external input signals $d'(t)$; and the third pair consists of output signals $z'(t)$, and inputs $v'(t)$, through which unit-norm perturbations Δ' are connected back into the system. (As before, primed quantities are normalized to unit size.) The design problem is to find a compensator $\boldsymbol{K}$ which stabilizes $\boldsymbol{P}$ and keeps the transfer matrix between d' and e' appropriately small for all stable Δ'.

This problem description is very general because the internal structure of $\boldsymbol{P}$ can be chosen to represent many different problem specifications. One example of internal structure is shown in Fig. 4, which corresponds to the class of skewed design problems described in Sec. III.A. We note that $\boldsymbol{P}$ includes the usual input-output description of the plant $\boldsymbol{G}$ and also the weighting functions $\boldsymbol{W}_e$ and $\boldsymbol{W}_d^{-1}$, which shape performance variables and external signals, as well as $\boldsymbol{W}_z$ and $\boldsymbol{W}_v^{-1}$, which shape the plant perturbation. The types of external signals [whether $r(t)$, $d(t)$, $n(t)$ and/or others] are also defined by the internal structure, as are the locations of perturbations (whether at outputs, inputs, and/or elsewhere). Various examples of internal structures of $\boldsymbol{P}$ for other problem specifications can be found in other references.[15,16]

B. Analysis Results

Beyond its generality, Fig. 3 is important because it comes equipped with a nonconservative necessary and sufficient condition for robust performance. In order to describe this new condition, we first close the compensator feedback loop in Fig. 3 to get the closed-loop system in Fig. 5. The system $\boldsymbol{M(P,K)}$ in this figure has a 2×2 block-structured transfer matrix $M(s)$, whose blocks are defined in terms of the original 3×3 partition of $P(s)$ as follows:

$$M_{ij}(s) = P_{ij}(s) + P_{i3}(s)[I - K(s)P_{33}(s)]^{-1}K(s)P_{3j}(s) \quad i,j = 1,2 \tag{39}$$

Suppose that this system is stable. Then the following results apply.

Theorem 3[16]

1) Nominal performance is satisfied if and only if

$$\sigma_{\max}[M_{22}(j\omega)] < 1 \qquad \text{for all } \omega \tag{40}$$

2) Stability is robust if and only if

$$\sigma_{\max}[M_{11}(j\omega)] < 1 \qquad \text{for all } \omega \tag{41}$$

3) Performance is robust if and only if

$$\mu[M(j\omega)] < 1 \qquad \text{for all } \omega \tag{42}$$

where $\mu[\,]$ is a function to be defined shortly.

The first of these results is a generalization of Eq. (21). It is again true by definition—performance is satisfied for $\Delta'(s) \equiv 0$ if $\sup_\omega \sigma_{\max}[M_{22}(j\omega)]$ is less than or equal to unity. Note that $M_{22}(s)$ can consist of any transfer matrix we may wish to keep small, not just sensitivity $S(s)$ as in Eq. (21). It also includes any desired weighting matrices. These features of the new design framework are important because they provide flexibility to handle many engineering situations where performance is not expressed in terms of sensitivity alone.

The second result is a generalization of Eq. (22). It follows from stability conditions with the Δ' loop closed—namely, that $\det[I - \Delta'(s)M_{11}(s)]$ must remain nonzero on the $j\omega$ axis for all $\Delta'(s)$. Again, $M_{11}(s)$ can consist of any transfer matrix, not just $T(s)$ as in Eq. (22) and it also includes any desired weights. These features provide the important engineering flexibility to represent uncertainties at their known locations in a plant, rather than forcing them to be reflected entirely to the plant's input or output.

The third result is the most significant one. It is a generalization of Eq. (23) that provides necessary and sufficient conditions for robust performance, not just sufficient conditions. It can be established by starting with the definition that performance is robust if and only if the e'/d' transfer matrix with the Δ' loop closed remains H_∞ norm bounded by unity—that is, if and only if

$$\sigma_{\max}[M_{22} + M_{21}(I - \Delta' M_{11})^{-1}\Delta' M_{12}] < 1 \quad \text{for all } \omega \text{ and all } \Delta' \tag{43}$$

Notice that this last norm bound is also a necessary and sufficient condition for system M in Fig. 6 to continue to remain stable even if we choose to connect a second norm-bounded perturbation, say $\Delta_0'(s)$, across the e' and d' terminals [to see this, compare the form of Eq. (43) with our other stability conditions, Eqs. (9), (22), and (41)]. Robust performance, therefore, is exactly equivalent to robust stability in the face of two perturbations, Δ' and Δ_0', connected around system M in the block-diagonal structured arrangement shown in Fig. 6. The latter stability is assured, of course, if and only if the function $\det[I - \mathrm{diag}(\Delta',\Delta_0')M(j\omega)]$ remains nonzero for all ω.

These observations bring us to the function $\mu[\,]$ in Eq. (42). This function was defined specifically to test the kind of determinant conditions already identified. Its full definition for complex matrices is the following[17]:

$$\mu[M] \triangleq \left[\min \left\{ \varepsilon \,\middle|\, \begin{array}{l} \det[I - \varepsilon X M] = 0 \\ \text{for some } X = \mathrm{diag}[\Delta_1,\Delta_2,\ldots,\Delta_m] \\ \text{with } \sigma_{\max}[\Delta_i] \le 1 \quad \text{for all } i \end{array} \right\} \right]^{-1} \tag{44}$$

In words, μ is the reciprocal of the smallest value of scalar ε that makes the matrix $I - \varepsilon XM$ singular for some X in a block-diagonal perturbation set. If no such ε exists, μ is taken to be zero. Notice that this definition reduces to the conventional singular value in the absence of structure (i.e., when the number of blocks, m, in X is one). For this reason, μ has been called the structured singular value. Note also that the value of μ depends on the number of blocks in the structure as well as on the block dimensions. Technically, therefore, μ arguments should include not only matrix M but also a multi-index describing the structure. By convention and for sake of notational simplicity, this latter dependence is suppressed.

It is clear from Eq. (44) that μ can be applied to the transfer matrices of Fig. 6 to test whether $\det[I - \mathrm{diag}(\Delta',\Delta_0')M]$ remains nonzero along the $j\omega$ axis. In fact, the determinant remains nonzero if and only if $\mu[M] < 1$ on the axis. This is a tight condition for robust stability with respect to the two perturbation blocks, and equivalently, a tight condition for robust performance [Eq. (42)]. More rigorous details of this argument are given in Ref. 17.

It is also important to note that Eq. (44) is not limited to 2×2 block structures. It can be used to test stability with respect to any number of diagonal blocks. This permits us to establish robust stability with respect to plant sets characterized by several unstructured perturbations and simultaneously to establish robust performance. Equation (44) also extends readily to real-valued perturbations and so reduces many parametric system analysis problems to μ calculations. More generally still, the structured singular value concept also extends to time varying and nonlinear systems.[18] The calculations required for these extended cases, however, continue to impose substantial challenges.

C. Numerics for the Structured Singular Value

Like singular values, μ is useful for practical numerical analyses as well as for theoretical ones. First-generation computer algorithms have been developed to evaluate the function for fixed complex matrices with complex perturbations. When used repeatedly, these algorithms can generate plots of $\mu[M(j\omega)]$ across frequency and thus provide practical Bode-like analysis tests for robust stability/performance of any given candidate feedback design.

To date, μ algorithms are based on the following inequalities, proven in Ref. 17:

$$\sup_U |\lambda[UM]| \le \mu[M] \le \inf_D \sigma_{\max}[DMD^{-1}] \tag{45}$$

where $|\lambda[M]|$ denotes the largest eigenvalue magnitude (spectral radius) of M and where

$$U \triangleq \mathrm{diag}[U_1,U_2,\ldots,U_m] \qquad U_i \text{ unitary}$$

$$D \triangleq \mathrm{diag}[d_1 I_1, d_2 I_2,\ldots,d_m I_m] \qquad d_i \text{ real-valued scalar}$$

with block dimensions of U_i and I_i matching those of X in Eq. (44). While these inequalities look formidable at first glance, they can be readily established by combining the following three facts.

1) $\mu[UM] = \mu[DMD^{-1}] = \mu[M]$ for all block structured U and D defined previously. This follows directly from Eq. (44).

2) $\mu[M] \le \sigma_{\max}[M]$. This follows by interpreting Eq. (44) as a definition of $\sigma_{\max}$, i.e., when X has one block and is therefore least constrained.

3) $|\lambda[M]| \le \mu[M]$. This also follows from Eq. (44) by interpreting it as a definition of $|\lambda|$ when X has m repeated scalar blocks and is therefore most constrained.

Reference 17 shows that the left-hand side of Eq. (45) is actually an equality and thus provides a potential way to compute μ. Unfortunately, the implied maximization over block-structured unitary matrices U can have many local maxima. The right-hand side of Eq. (45) is also an equality, at least for structures with three or fewer blocks. Its implied minimization over block-structured matrices D has convexity properties, which imply that all local minima are global. Thus, local search methods can provide numerically tractable approaches to finding the infimum.

For block structures with four or more blocks, experience has shown that the upper bound in Eq. (45) still provides an accurate estimate for μ. Indeed, despite extensive search, no example has been found where the upper bound exceeds μ by more than about 15%. Although this evidence is suggestive, no guarantee on the quality of the upper bound has been proven. For more complicated block structures involving repeated, possibly real, perturbations, the upper bound in Eq. (45) can be much poorer. Fortunately, alternative upper bounds and other algorithms are under development and appear to be very promising.[19–22]

D. Some Analytical Applications of μ

As a way to illustrate the alternate design framework and the utility of structured singular values, we revisit the example from Sec. III. Our objectives will be to verify some of the conclusions from Secs. II and III and, more important, to establish the lower bound on performance deterioration of plant-inverting controls [Eq. (36)]. As noted earlier, the general system description P for the example is given in Fig. 4. After closing the $K(s)$-feedback loop in the figure, the matrix $M(s)$ for this description is the following:

$$M(s) = \begin{bmatrix} W_z T W_v^{-1} & W_z T W_d^{-1} \\ W_e S W_v^{-1} & W_e S W_d^{-1} \end{bmatrix} = \begin{bmatrix} W_z T \\ W_e S \end{bmatrix} [W_v^{-1} \quad W_d^{-1}] \tag{46}$$

Then, by Eq. (44), the structured singular value is the reciprocal of the smallest ε, which will make the following determinant vanish for some Δ' and some Δ'_0:

$$\det[I - \varepsilon\, \mathrm{diag}(\Delta',\Delta_0')M] = \det[I - \varepsilon(W_v^{-1}\Delta' W_z T + W_d^{-1}\Delta_0' W_e S)] \tag{47}$$

Round Weights

If all the weights W_x in Eq. (47) are scalars times identity as in Sec. II, then the determinant in Eq. (47) becomes $\det[I - \varepsilon(w_\Delta \Delta' T + w\Delta_0' S)]$, and the smallest ε that can make it vanish is bounded by

$$\varepsilon_{\min} \geq 1/(\sigma_{\max}[w_\Delta T] + \sigma_{\max}[wS])$$

Its reciprocal μ is thus bounded by

$$\mu[M] \leq \sigma_{\max}[w_\Delta T] + \sigma_{\max}[wS] \tag{48}$$

Moreover, since $\mu[M]$ must be less than unity for robust performance, the condition

$$\sigma_{\max}[wS] < 1 - \sigma_{\max}[w_\Delta T]$$

is sufficient for robust performance, as was shown previously in Sec. II, Eq. (12).

Round Controllers/Round Plants

On the other hand, if the weights are skewed as described in the example of Sec. III [Eqs. (27) and (28)], then the applicable determinant becomes

$$\det[I - \varepsilon\,\mathrm{diag}(\Delta', \Delta_0')M] = \det[I - \varepsilon(w_\Delta G\Delta' G^{-1} T + w\Delta_0' S)] \tag{49a}$$

$$= \det[I - \varepsilon(w_\Delta T K^{-1}\Delta' K + wS\Delta_0')] \tag{49b}$$

where Eq. (49b) is obtained from Eq. (49a) by factoring out S and S^{-1} on the right and left, respectively. These expressions imply that Eq. (48) still holds whenever either $K(s)$ or $G(s)$ is spatially round, and consequently, that Eq. (38) in Sec. III is sufficient for robust performance in these cases.

Plant-Inverting Controllers

Next, consider what happens when the plant-inversion approach is used to "solve" the example problem. The matrix $M(s)$ in Eq. (46) then becomes

$$M(s) = \begin{bmatrix} w_\Delta tI & wtG^{-1} \\ w_\Delta sG & wsI \end{bmatrix} \tag{50}$$

Using the upper bound from Eq. (45) (an equality because there are only two blocks), its structured singular value is given by

$$\mu[M] = \inf_d \sigma_{\max}\begin{bmatrix} w_\Delta tI & w_\Delta t(dG)^{-1} \\ wsdG & wsI \end{bmatrix}$$

$$= \inf_d \sigma_{\max}\begin{bmatrix} w_\Delta tI & w_\Delta t(d\Sigma)^{-1} \\ wsd\Sigma & wsI \end{bmatrix} \tag{51a}$$

$$= \inf_d \max_i \sigma_{\max}\begin{bmatrix} w_\Delta t & w_\Delta t/(d\sigma_i) \\ wsd\sigma_i & ws \end{bmatrix} \tag{51b}$$

$$= \inf_d \max_i \left[|ws|^2 + |w_\Delta t|^2 + |wsd\sigma_i|^2 + \frac{|w_\Delta t|^2}{|d\sigma_i|^2} \right]^{1/2} \tag{51c}$$

In Eq. (51a), the scaling variables d_1 and d_2 from Eq. (45) are selected as $d_1 = 1$ and $d_2 = (w_\Delta/w)d$ for algebraic convenience. The matrix $\Sigma = \mathrm{diag}(\sigma_1 \cdots \sigma_m)$ is obtained by substituting the singular value decomposition $U\Sigma V^*$ for G and factoring out its right and left unitary matrices. Equation (51b) is obtained by elementary row/column permutations to isolate two-dimensional blocks, and Eq. (51c) is an explicit singular value calculation based on the fact that each such block is an outer product of two vectors. An examination of Eq. (51c) shows the scaling variable d must be chosen to reduce the term $|wsd\sigma_i|^2 + |w_\Delta t/d\sigma_i|^2$ for the largest singular value of G while at the same time not excessively increasing the same term for the smallest singular value (or visa versa). Indeed, the best choice of d is

$$|d|^2 = |w_\Delta t| \,/\, |ws\sigma_{\max}[G]\sigma_{\min}[G]| \tag{52}$$

which makes the two terms equal, and when substituted back into Eq. (51c), gives the following explicit expression for the μ values achieved by plant inversion:

$$\mu^2[M] = |ws|^2 + |w_\Delta t|^2 + |ws|\,|w_\Delta t|\left[\kappa[G] + \frac{1}{\kappa[G]}\right] \tag{53}$$

Finally, bounding this μ value by unity gives the following necessary and sufficient condition for robust performance whenever $\kappa[G]$ is large:

$$|ws| < \frac{1 - |w_\Delta t|^2}{|w_\Delta t|\kappa[G]} = \frac{1 - |w_\Delta t|}{\kappa[G]}\,\frac{1 + |w_\Delta t|}{|w_\Delta t|} \tag{54}$$

This condition confirms the lower bound in Eq. (36).

Beyond confirming Eq. (36), Eq. (53) shows that μ values achieved by inversion of highly skewed plants are proportional to the square root of the plant's condition number, with proportionality constant equal to the geometric mean of the nominal performance and robust stability test values, i.e.,

$$\mu[M] \approx \sqrt{|ws|\,|w_\Delta t|}\,\sqrt{\kappa[G]} \tag{55}$$

This should be contrasted with Eq. (48), which shows that μ is bounded by the sum of these test values whenever weights, controllers, or plants are spatially round.

For plant inverting controllers, therefore, robust performance requires that one of the nominal test values, $|ws|$ or $|w_\Delta t|$, be very small (proportional to $1/\kappa[G]$) whenever the other one takes on typical SISO design values (e.g., ≈ 0.5). Equations (54) and (23) result when the small one is chosen to be $|ws|$. As an alternative, we could attempt to make both test values small and equal. They would then need to be proportional only to $1/\sqrt{\kappa[G]}$. Unfortunately, the relation $S(s) + T(s) = I$ precludes this alternative unless the weights w and w_Δ themselves are very small.

E. Numerical Example

The preceding analyses confirm our basic conclusions from Sec. III that both round controls and plant inversion are dubious solutions for highly skewed design problems. The one is dubious because nominal performance and robust stability test values often cannot both be made small enough, and the other is dubious because the multiplier $\kappa[G]$, or even $\sqrt{\kappa[G]}$, relating μ to these test values is simply too large. Fortunately, there are other control alternatives that do better. The existence of these alternatives as well as the dubious properties of round controls and plant inversion are next illustrated with a simple numerical example.

Consider the following plant with condition number $\kappa = 625$:

$$G(s) = \begin{bmatrix} \dfrac{25}{s} & 0 \\ 0 & \dfrac{0.04}{s} \end{bmatrix} \tag{56}$$

Let the performance specification be round levied at the output, and let the plant uncertainty specification be round levied at the input (and thus skewed at the output as in the example of Sec. III) with the following weights:

$$w(s) = \frac{0.2(s+1)}{(s+0.001)}I \qquad w_\Delta(s) = \frac{200(s+1)}{(s+1000)}I \tag{57}$$

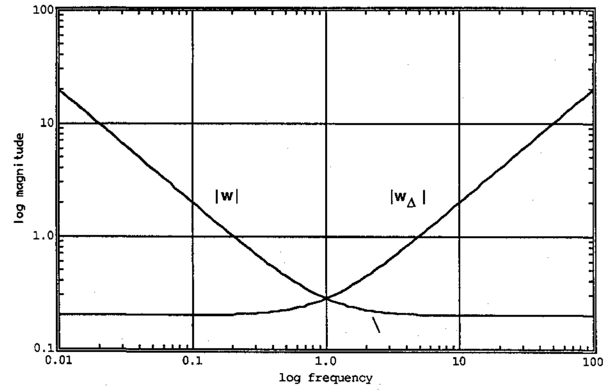

Fig. 7 Weights for example in Sec. III.

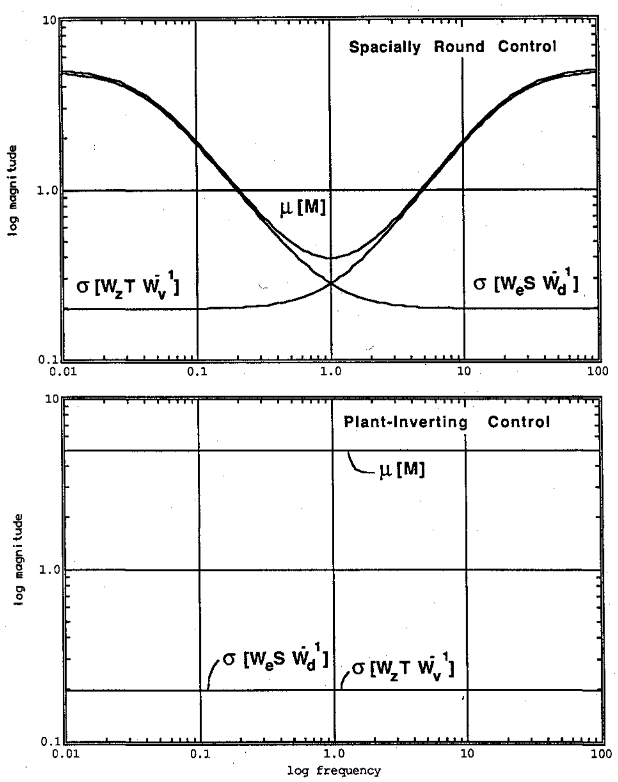

Fig. 8 Performance of dubious controls.

Bode plots of these weights are shown in Fig. 7. They call for the feedback system's sensitivity function to be less than 5ω for $0.001 \le \omega \le 1$ rad/s and to be no greater than 5.0 for higher frequencies. The plant set for which this must be achieved consists of the nominal plant with unstructured multiplicative input perturbations whose magnitude is 0.2 (20%) below 1 rad/s but rises to 100% at 5 rad/s and continues to increase steadily beyond that frequency. (These specifications are mirror images, reflected about 1 rad/s for symmetry. They should be treated only as illustrations, not serious specifications. Sensitivities as high as 5.0 are rarely tolerable in practice, and model errors of 20% are not always possible, even at low frequencies.) The general problem description again corresponds to Fig. 4 with Eqs. (56) and (57) replacing the generic symbols in the figure.

Our two dubious control alternatives for this design problem are evaluated in Fig. 8. Figure 8a corresponds to a round controller, and Fig. 8b corresponds to a plant inverter, i.e.,

$$K_a(s) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad K_b(s) = \begin{bmatrix} 0.04 & 0 \\ 0 & 25 \end{bmatrix} \tag{58}$$

Each figure shows Bode plots for the nominal performance test, $\sigma_{\max}[W_e S W_d^{-1}]$, for the robust stability test $\sigma_{\max}[W_z T W_v^{-1}]$, and for the $\mu[M]$ function achieved by the controller.

As predicted by Eq. (53), the plant inverter's μ values exceed specifications, even though its nominal performance and robust stability test values are adequate. The $\sqrt{\kappa[G]} = 25$ amplification factor is simply too high. Note that the nominal test values and μ values are identical across frequency as a result of our (convenient) choice of weight, plant, and controller parameters.

The round controller, on the other hand, fails to meet even the nominal specifications. Nominal performance is violated at low frequencies and robust stability is violated at high frequencies. Note, however, that the μ values are not amplified nearly as much over the nominal test values because the sum bound in Eq. (48) applies.

It is important to recognize that μ values greater than unity, such as those exhibited by these controllers, have *doubly severe negative effects*. For instance, the $\mu = 5$ values of the round controller at low frequencies not only mean that performance will be five times worse than specified at those frequencies. They also mean that this will happen for plant perturbations that are five times smaller than the perturbations specified at those frequencies. [This follows from Eq. (44) and the fact (not proven here) that the determinant vanishes for points on the boundary of each of the blocks in the perturbation structure X.] For the case at hand, this means that sensitivities will be as large as 25ω for plant errors no larger than 0.04 (4%)—well short of the 5ω at 20% specification.

Next, consider the following family of alternate controllers for the design example:

$$K_\eta = \begin{bmatrix} 25^{-\eta} & 0 \\ 0 & 25^{\eta} \end{bmatrix} \tag{59}$$

Heuristically, this family performs various degrees of partial plant inversion uniformly across frequency. It interpolates between one extreme, the round controller at $\eta = 0$, to the other extreme, the full plant inverter at $\eta = 1$. The family's performance is summarized in Fig. 9a, which shows the worst case μ values (i.e., $\sup_\omega \mu[M(j\omega)]$) achieved for various values of parameter η. Note that every value $0 < \eta < 1$ performs better than the two extremes and that there is a well-defined optimum near $\eta = 0.4$ that is more than twice as good (four times as good, counting doubly severe effects) than the extremes. Bode plots of $\mu[M]$, $\sigma_{\max}[W_e S W_d^{-1}]$, and $\sigma_{\max}[W_z T W_v^{-1}]$ for this optimum value are shown in Fig. 9b.

F. Mu Synthesis

Figure 9 illustrates that controllers exist for highly skewed design problems that are superior to both round controls and to plant inversion. Of course, K_η at $\eta = 0.4$ still falls well short of specification. We are left to wonder whether other alternatives exist that do even better and, more fundamentally, whether there is an ultimate best robust performance level and what controllers achieve that level.

These questions motivate a need for more formal design procedures to solve highly skewed problems. We call such procedures "mu synthesis"—meaning formal theories and associated numerical tools to find compensators $\boldsymbol{K}$, which stabilize system $\boldsymbol{P}$ in Fig. 3 and that minimize the resulting worst-case value of $\mu[M(\boldsymbol{P},\boldsymbol{K})(j\omega)]$ in Fig. 5. While complete solutions of this design problem are not yet available, a so-called D-K iteration scheme has been invented that yields useful answers.[16] D-K iterations are based on the upper bound in Eq. (45) that equates μ optimization to the following two-parameter minimization problem

$$\inf_{\boldsymbol{K}} \sup_{\omega} \mu[M(\boldsymbol{P},\boldsymbol{K})(j\omega)] = \inf_{\boldsymbol{K}} \sup_{\omega} \inf_{D(\omega)} \sigma_{\max}[D(\omega)M(\boldsymbol{P},\boldsymbol{K})(j\omega)D(\omega)^{-1}] \tag{60}$$

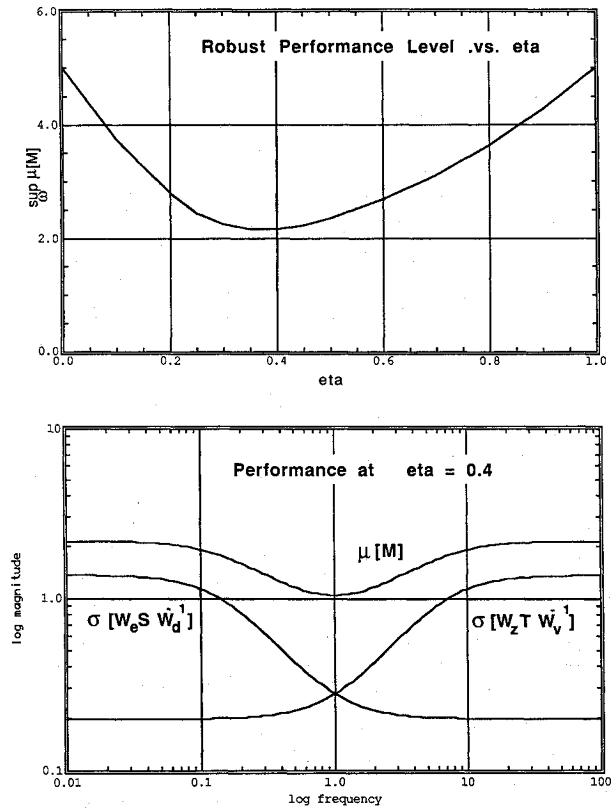

Fig. 9 Performance of partial-plant-inverting controls.

where $D(\omega)$ is a block-diagonal scaling matrix applied pointwise across frequency to the frequency response matrix $M(\boldsymbol{P},\boldsymbol{K})(j\omega)$.

D-K iterations proceed by performing this two-parameter minimization in sequential fashion—first minimizing over $\boldsymbol{K}$ with $D(\omega)$ fixed, then minimizing pointwise over $D(\omega)$ with $\boldsymbol{K}$ fixed, then again over $\boldsymbol{K}$, and again over $D(\omega)$, and so forth. Details of this process are summarized in the following steps:

1) Fix an initial estimate of the scaling matrix $D(\omega)$ pointwise across frequency. When applied to the 2×2 block structure of $\boldsymbol{M}$ in Fig. 5, this matrix has the form $\mathrm{diag}[d_1(\omega)I_1, d_2(\omega)I_2]$. Alternatively, the scaling matrix can be applied to the 3×3 block structure of $\boldsymbol{P}$ in Fig. 3, in which case it has the form $\mathrm{diag}[d_1(\omega)I_1, d_2(\omega)I_2, I_3]$. In either form, the first scale factor d_1 can be set equal to unity, leaving only a single scalar function of frequency to be selected.

2) Find a state-space realization for $D(\omega)$, and construct a state-space model for system $\boldsymbol{DPD}^{-1}$. Again for the structure in Fig. 3, this step calls for a state-space realization of a SISO system $\boldsymbol{d}$, whose frequency response magnitude matches $d(\omega)$. This realization must then be incorporated into an overall state-space model for the scaled system:

$$\boldsymbol{DPD}^{-1} = \begin{bmatrix} I_1 & 0 & 0 \\ 0 & \boldsymbol{d}I_2 & 0 \\ 0 & 0 & I_3 \end{bmatrix} \boldsymbol{P} \begin{bmatrix} I_1 & 0 & 0 \\ 0 & \boldsymbol{d}^{-1}I_2 & 0 \\ 0 & 0 & I_3 \end{bmatrix}$$

3) Solve an H_∞-optimization problem to minimize $\sup_\omega \sigma_{\max}[M(\boldsymbol{DPD}^{-1},\boldsymbol{K})(j\omega)]$ over all stabilizing $\boldsymbol{K}$. Note that this optimization problem uses the scaled version of $\boldsymbol{P}$. Let its minimizing controller be denoted by $\boldsymbol{K}^*$.

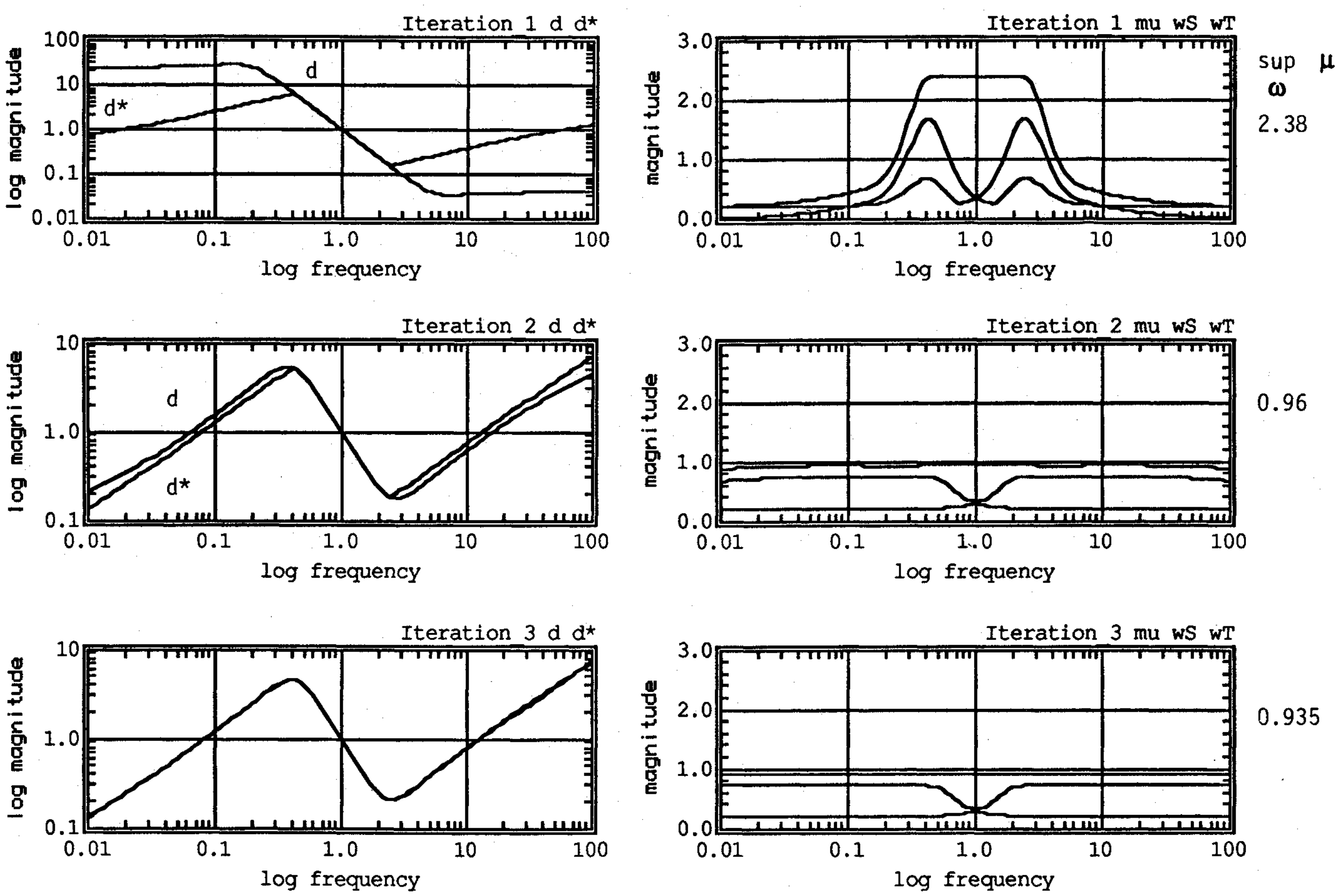

Fig. 10 D-K-iteration convergence to "global minimum."

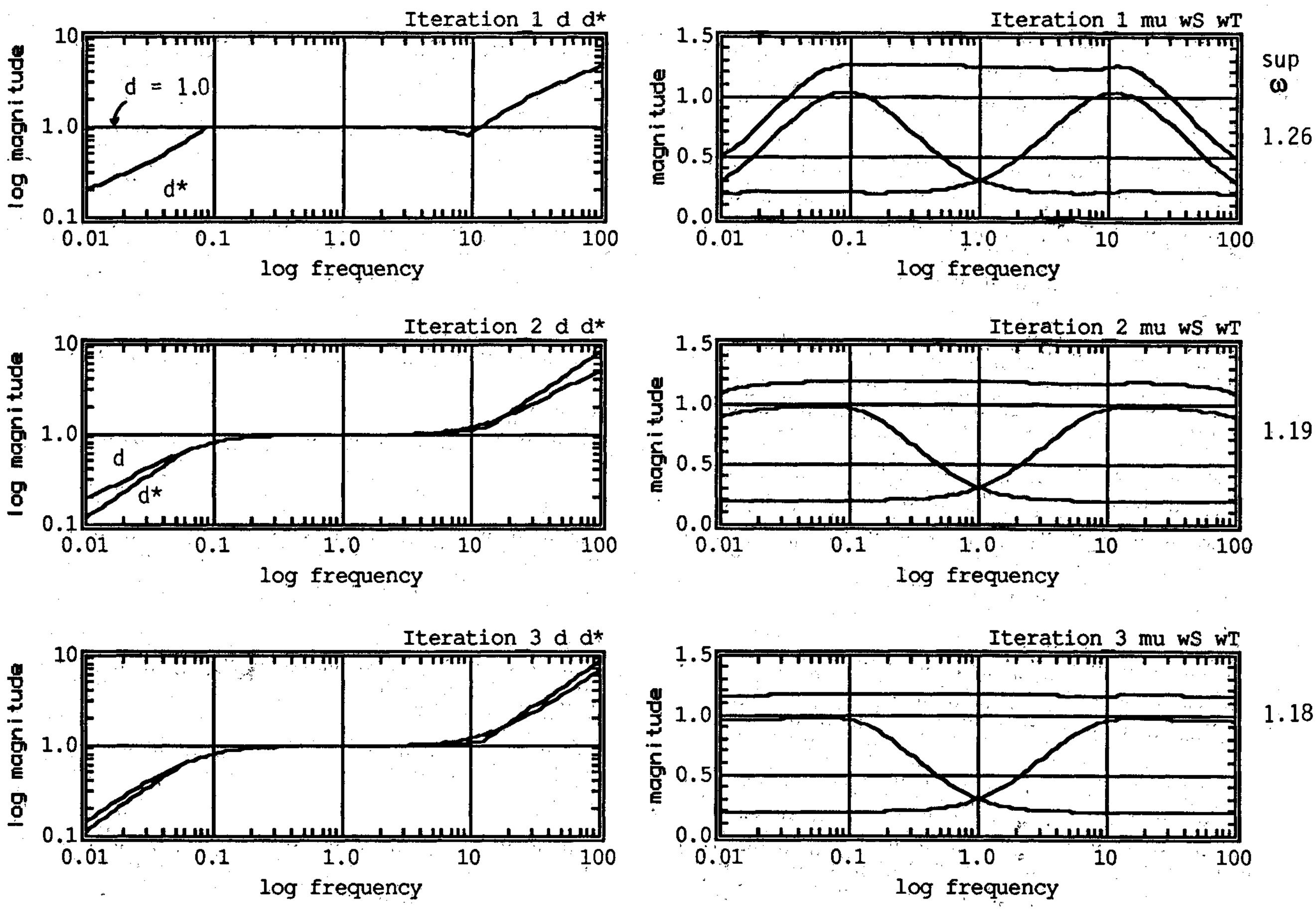

Fig. 11 ***D-K*-iteration convergence to local minimum.**

4) Minimize $\sigma_{\max}[D(\omega)M(P,K^*)(j\omega)D(\omega)^{-1}]$ over $D(\omega)$, pointwise across frequency. Note that this evaluation uses the minimizing K^* from step 3, but P is unscaled. The minimization itself produces a new scaling function. Let this new function be denoted by $D^*(\omega)$.

5) Compare $D^*(\omega)$ with the previous estimate $D(\omega)$. Stop if they are close. Else replace $D(\omega)$ with $D^*(\omega)$ and return to step 2.

These iterations are practical because the H_∞-optimization problem in step 2 has a numerically tractable solution discovered only recently.[8] This H_∞ solution is itself a significant step forward in MIMO design. It encompasses the various special H_∞ problems mentioned in Sec. II as special cases and provides state-space-based computational algorithms, whose complexity is comparable to LQG algorithms. Hence, they can handle design problems of significant engineering size. Detailed descriptions of the algorithms are left to Ref. 23.

Although *D-K* iterations are practical, they are incomplete solutions of mu-synthesis problems for two reasons. First, the iterations require a frequency response approximation process in step 2. With the current state of the art, this process relies more on engineering skill than on formal mathematics. Second, while the iterations are monotonically decreasing [with perfect state-space realizations of $D(\omega)$], they converge to local minima of Eq. (60), not necessarily to global minima.

Both the practicality of *D-K* iterations and their limitations are illustrated in Figs. 10 and 11. These figures show two *D-K* iteration trials applied to our example problem in Sec. IV.E. Figure 10 shows a sequence of iterations that, we believe, converge to a point very close to the global minimum of the example, while Fig. 11 shows a sequence that clearly converges to a local minimum. The alleged global minimum corresponds to $\sup_\omega \mu[M(j\omega)] \approx 0.935$. This value barely meets the original design specifications. However, it is five times better (25 times better, counting double severe effects) than either round control or plant-inversion solutions, thus making *D-K* iterations well worth the effort.

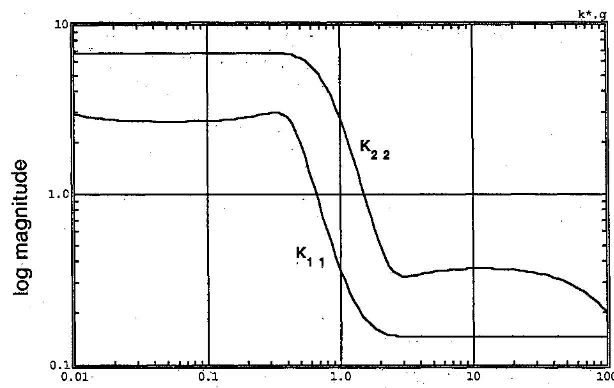

Fig. 12 **Frequency response of *K****.**

Frequency responses of the alleged globally minimizing controller K^* are shown in Fig. 12. Simple interpretations explaining why this controller produces good results remain to be discovered and are left as a challenge for serious readers. We also encourage explorations of other controllers so as to refute or verify our claim that K^* is close to μ optimal.

V. Conclusion

In this paper, we have taken a tutorial look at singular value loop shaping as a paradigm for multivariable feedback design, and have attempted to motivate the need to go beyond this paradigm in general design situations.

Our principal conclusions are that singular value loop shaping is effective for design problems with spatially round specifications but that it can be excessively conservative (often impossibly conservative) when the specifications are highly skewed. Unfortunately, skewed problems arise frequently and naturally in MIMO design situations. A simple example with performance requirements levied at the output and uncertainties occurring at the input of a poorly conditioned plant was used to illustrate this point. Neither plant inversion nor round controls are able to satisfy robust performance specifications for this example.

An alternate design paradigm, which promises to handle skewed problems more effectively, was discussed. This paradigm involves a very general and flexible design setup, coupled with tight necessary and sufficient conditions for robust performance. The latter are provided by a new matrix function, the structured singular value *mu*. This function is already being widely used to analyze and compare performance properties of existing controllers. Progress is also being made on formal mu-synthesis methods that find new controllers to minimize the structured singular value directly. Although incomplete at this time, one such method is *D-K* iteration. This method was applied to our example to yield substantial robust performance improvements over other control options.

Much research remains to be done to make mu synthesis as routine as, say, H_2 synthesis with LQG algorithms or H_∞ synthesis with the newly discovered LQG-like algorithms. This research work is underway.

Acknowledgment

This work was supported in part by Honeywell and by research grants from NASA, ONR, and AFOSR.

References

[1]Nyquist, H., "Regeneration Theory," *Bell System Technical Journal*, Jan. 1932.

[2]Bode, H. W., *Network Analysis and Feedback Amplifier Design*, D. Van Nostrand, Princeton, NJ, 1945.

[3]Rosenbrock, H. H., *State Space Methods and Multivariable Theory*, Nelson, London, 1970.

[4]Doyle, J. C., and Stein, G., "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis," *IEEE Transactions Auto. Control*, Feb., 1981.

[5]Stein, G., and Athans, M., "The LQG/LTR Procedure for Multivariable Feedback Control Design," *IEEE Transactions Auto. Control*, Feb. 1987.

[6]Zames, G., and Francis, B. A., "Feedback, Minimax Sensitivity, and Optimal Robustness," *IEEE Transactions Auto. Control*, Vol. AC-28, 1983, pp. 585–601.

[7]Francis, B. A., *A Course in H_∞ Control Theory*, Springer-Verlag, Berlin, Heidelberg, 1987.

[8]Doyle, J. C., Glover, K., Khargonekar, P., and Francis, B. A., "State-space Solutions to Standard H_2 and H_∞ Control Problems," *Proceedings of the American Control Conference*, Atlanta, GA, June 1988.

[9]Desoer, C. A., and Vidyasagar, M., *Feedback Systems: Input-Output Properties*, Academic, New York, 1975.

[10]Vidyasagar, M., "The Graph Metric for Unstable Plants and Robustness Estimates for Feedback Stability," *IEEE Transactions Auto. Control*, Vol. AC-29, 1984, pp. 403–418.

[11]Horowitz, I. M., *Synthesis of Feedback Systems*, Academic, New York, 1963.

[12]Rosenbrock, H. H., *Computer-Aided Control System Design*, Academic, London, 1974.

[13]MacFarlane, A. G. J., and Kouvaritakis, B., "A Design Technique for Linear Multivariable Feedback-Systems," *International Journal Control*, Vol. 25, 1977, pp. 837–879.

[14]Postlethwaite, I., and MacFarlane, A. G. J., *A Complex Variable Approach to the Analysis of Linear Multivariable Feedback Systems*, Springer-Verlag, Berlin, 1979.

[15]Doyle, J. C., Wall, J. E., and Stein, G., "Performance and Robustness Analysis for Structured Uncertainty," *Proceedings of the 21st IEEE Conference on Decision and Control*, Orlando, FL, Dec. 1982, pp. 629–636.

[16]Doyle, J. C., "Lecture Notes," ONR/Honeywell Workshop on Advances in Multivariable Control, Minneapolis, MN, Oct. 1984.

[17]Doyle, J. C., "Analysis of Feedback Systems with Structured Uncertainty," *Proceedings of IEE*, Nov. 1982.

[18]Krause, J., et al., "A General Statement of Structured Singular Value Concepts," *Proceedings of the American Control Conference*, Atlanta, GA, June 1988.

[19]Fan, M. K. H., and Tits, A., "Characterization and Efficient Computation of the Structured Singular Value," *IEEE Transactions Auto. Control*, Aug. 1986.

[20]de Gaston, R. E., and Safonov, M. G., "Calculations of the Multiloop Stability Margin," *Proceedings of the American Control Conference*, Seattle, WA, June 1986.

[21]Fan, M. K. H., Tits, A., and Doyle, J. C., "Robustness in the Presence of Joint Parametric Uncertainty and Unmodelled Dynamics," *Proceedings of the American Control Conference*, Atlanta, GA, June 1988.

[22]Packard, A., and Doyle, J. C., "Structured Singular Value with Repeated Scalar Blocks," *Proceedings of the American Control Conference*, Atlanta, GA, June 1988.

[23]Glover, K., and Doyle, J. C., "State-space Formulae for all Stabilizing Controllers that satisfy an H_∞-norm Bound," *System and Control Letters* (to be published).