Advanced Control Law Tuning and Performance Assessment

Michael J. Grimble and Pawel Majecki
Industrial Control Centre
University of Strathclyde
Graham Hills Building, 50 George Street
Glasgow G1 1QE
UNITED KINGDOM

m.grimble@eee.strath.ac.uk

ABSTRACT

It is evident that military applications for 21st century will be highly complex, multivariable systems. Designing optimal controllers for such applications will require the use of mathematical models which describe the complexity of the underlying processes as accurately as possible. For robustness it is essential that these models include the uncertainties in the estimated process dynamics and trajectories. Controllers resulting from optimal control strategies (like LQG, H_2, H_∞) will usually be of very high order and that can cause implementation and computational problems. Some of these problems can be overcome by using controllers that are of lower order and restricted structure.

The subject of restricted structure controller design and performance assessment is relatively new and enables the expected performance to be assessed against a much more realistic criterion. That is, the performance figures take into account the limitations of the existing control system structure, and hence provide a more accurate measure of the possible performance improvement.

However, the design and/or tuning of restricted-structure controllers in order to provide the performance comparable with full-order solutions is still a very contentious issue. Added to this is the need to achieve robust properties and performance specifications required by military applications. Very few methods for designing restricted-structure controllers exist that allow the performance and robustness objectives to be combined into one relatively simple optimisation problem. This lecture presents an LQG/H2-based method that tackles the above mentioned issues.

This work also provides the exciting possibility of enabling the multivariable structures of systems to be assessed. Thus for example, it is possible to check whether a diagonal multivariable controller, upper triangular, lower triangular or sparser structures might be almost as good as a full multivariable control law. This technique provides advantages over the commonly used technique of so called relative gain array for judging the best structure for a multivariable system.

All these techniques are applicable to continuous or discrete-time linear systems, although the nonlinear behaviour can readily be accommodated in the design through the use of multiple-model strategy. This lecture, however, will also introduce a control law that is derived for the control of nonlinear, possibly time-varying systems. The solution for this Nonlinear Generalized Minimum Variance control law is original and was obtained using a simple operator representation of the process. The quadratic cost index involves both error and control signal costing terms. The controller obtained is simple to implement and includes an internal model of the process. In one form it might be considered a nonlinear version of the Smith Predictor.

1.0 INTRODUCTION

The design and performance assessment of optimal controllers (Kwakernaak and Sivan, 1972 [1], Desborough and Harris 1993 [2], Uduehi and Grimble 2001 [3]) is considered as well as the ways in

Grimble, M.J.; Majecki, P. (2006) Advanced Control Law Tuning and Performance Assessment. In *Achieving Successful Robust Integrated Control System Designs for 21st Century Military Applications – Part II* (pp. 2-1 – 2-42). Educational Notes RTO-EN-SCI-166, Paper 2. Neuilly-sur-Seine, France: RTO. Available from: http://www.rto.nato.int/abstracts.asp.

which robustness properties can be modified and assessed. The minimization of quadratic cost functions for stochastic linear systems in such a way that the robustness margins are improved is of interest. There are of course well known guaranteed robustness properties for systems with state feedback (Anderson and Moore 1971 [4], Safonov and Athans 1986 [5], Doyle 1978 [6]) and in some problems these properties can be recovered using a loop transfer recovery approach ([7],[8],[9],[10],[11],[12]). However, more direct methods of tuning robustness properties would be valuable (Doyle and Stein, 1981 [13]). This is explored for both full-order optimal output feedback controllers and for low order (restricted structure) optimal solutions.

Two aspects of the design of full-order and *restricted structure* (*RS*) optimal controllers are considered in the following. The first involves robustness improvement and the second is concerned with the noise rejection properties. An LQG criterion is to be optimized but the usual robustness and noise rejection properties will not hold if the controller structure is limited to say a PID or a low order lead-lag form. Grimble (1999 [14], 2000 [15]) introduced a polynomial systems approach to restricted structure optimal control design and this is the philosophy followed here. However, the design of such controllers previously focussed on performance issues and the robustness/noise-rejection aspects were not considered in any detail. The strategy for robustness improvement is to add a fictitious signal and a sensitivity costing term in the criterion. This enables the penalty on sensitivity to be directed at modifying the robustness properties. The normal LQG cost-index does of course include error and control signals that depend upon the sensitivity functions. However, this does not enable these sensitivity terms to be costed in a particular way. The proposed robustness weighting term gives free choice of the weighting function and enables the H_2 norm of a weighted sensitivity function to be minimized. This relates to the definition of a so-called Dual Criterion (Grimble, 1986 [16]), but the results here are focussed on the design issues and they use what might be termed a Kucera polynomial systems approach, (1980 [17]). The frequency-domain polynomial systems approach is particularly helpful when determining the frequency response behaviour to noise and disturbance signals (Grimble, 1994 [18]).

The above control algorithms and methodology are based on the linearized description of the process and are valid in a particular operating point. The aim in the following is to introduce a controller for nonlinear multivariable, possibly time-varying, processes. This *Nonlinear GMV* (NGMV) control law is related to the family of LQG designs but is based more on the rich heritage of minimum variance controls. Åström introduced the *Minimum Variance* (MV) controller assuming the linear plant was minimum phase and later derived the MV controller for processes that could be non-minimum phase (Åström 1979 [19]). The latter was guaranteed to be stable on non-minimum phase processes, whereas the former was unstable. Hastings-James (1970 [20]) and later Clarke and Hastings-James (1971, [21]), modified the first of these control laws by adding a control costing term. This was termed a *Generalized Minimum Variance* (GMV) control law and enabled non-minimum phase processes to be stabilized, although when the control weighting tended to zero the control law reverted to the initial algorithm of Åström, which was unstable. However, the control law had similar characteristics to LQG design in some cases and was much simpler to implement. This simplicity was exploited very successfully in the so-called generalized MV self-tuning controller introduced by Clarke and Gawthrop (1975, [22]). The control of nonlinear non-minimum phase linear systems using a GMV type algorithm was considered by Grimble (1981 [23]). The use of GMV control laws for linear systems designs was reviewed in Grimble (1988 [24]). The use of dynamic cost weightings in the GMV cost index (1994 [18]) provided additional flexibility and the dynamic costing solution was exploited to obtain a Generalized H_∞ controller (Grimble 1993 [25]). All of these results were applicable to linear discrete-time stochastic processes.

The NGMV control law uses a similar stochastic framework but generalizes and extends these results to account for nonlinear system behaviour. The structure of the system was defined so that a simple controller structure and solution are obtained. When the system is linear the results revert to those for the GMV controller referred to above (Grimble 2001 [26]). There is some loss of generality in assuming the reference and disturbance models are represented by linear subsystems. However, the plant model can be in a very general nonlinear operator form, which might involve state-space, transfer operators or even nonlinear function look up tables. That is, the input sub-system to the plant might include valves or a servo-system that has no traditional equation based model. The input nonlinear subsystem can be a black box. No state space or operator structure is needed. The optimal solution reveals all that is needed is a method of computing the output from such subsystem given a control input. If on the other hand an equation based model is available it may be used directly. For this reason the nonlinear part of the plant is represented in operator (unstructured form) for most of the analysis. The nonlinear dynamic terms, in the plant, only need to be open loop stable and can be very complicated. The ability to introduce very general plant structures, without formal models, is a major advantage of the method.

For linear systems stability is ensured when the combination of a control weighting function and an error weighted plant model is strictly minimum phase. For nonlinear systems a related operator equation must have a stable inverse. It is shown that if there exists say a PID controller that will stabilize the nonlinear system, without transport delay elements, then a set of cost weightings can easily be defined to guarantee the existence of this inverse and thereby ensure the stability of the closed loop. If the plant is open-loop stable the solution can be realized in a particularly simple form which relates to the well known Smith Predictor for systems with significant transport delays. This has the advantage of providing some confidence in the practical utility of the solution and also introduces what might be termed an extension of these *Smith* controllers for nonlinear plants. A so-called *Nonlinear Smith Predictor* will therefore be introduced. The main advantage over other nonlinear control design methods (Isidori 1995 [27]) is the simplicity of the solution.

2.0 SYSTEM MODEL

The system shown in Fig. 1 is for the time being assumed to be linear, continuous-time and single-input, single-output. The external white noise sources drive colouring filters which represent the reference $W_r(s)$, measurement noise $W_n(s)$, robustness modification $W_p(s)$ and disturbance $W_d(s)$ subsystems. The robustness model does not exist physically but is introduced for design modification. The system equations become:

Input disturbance: $$d(s) = W_d(s)\xi(s) \tag{2.1}$$

Robustness signal: $$p(s) = W_p(s)\eta(s) \tag{2.2}$$

Output: $$y(s) = d(s) + p(s) + W(s)u(s) \tag{2.3}$$

Reference: $$r(s) = W_r(s)\zeta(s) \tag{2.4}$$

Tracking error: $$e(s) = r(s) - y(s) \tag{2.5}$$

Observations: $$z(s) = y(s) + n(s) \tag{2.6}$$

Measurement noise: $$n(s) = W_n(s)\omega(s) \tag{2.7}$$

Control signal: $$u(s) = C_0(s)(r(s) - z(s)) \tag{2.8}$$

The system transfer functions are all assumed to be functions of the Laplace transform complex number in the complex frequency domain. For notational simplicity the arguments in $W(s)$ and the other models are omitted.

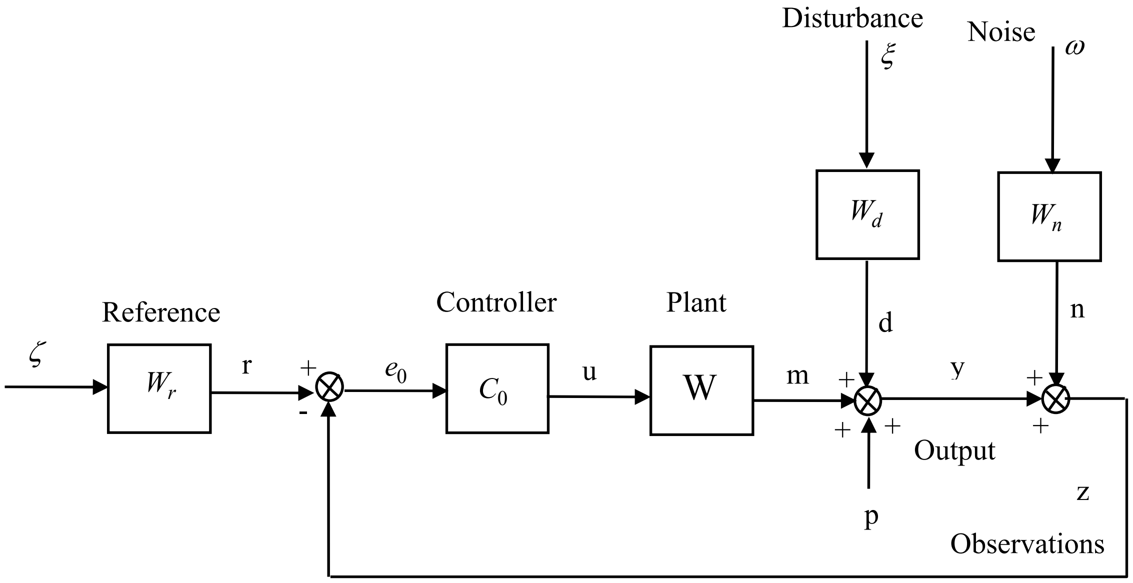

Figure 1: Single Degree of Freedom Unity Feedback Control System with Measurement Noise, Reference and Disturbance Models Assumptions

Assumptions:

- The white noise sources, ξ, ω, η *and* ζ are zero-mean and mutually statistically independent. The intensities of these signals are without loss of generality taken to be of value unity.
- The system W is assumed free of unstable hidden modes and the reference W_r, noise W_n, robustness W_p and disturbance W_d subsystems are asymptotically stable.

The following expressions may easily be derived for the output, error, observations, controller input, control and sensitivity costing signals:

Output *:* $$y(s) = WC_0(1+WC_0)^{-1}(r(s)-n(s)) + (1+WC_0)^{-1}(d(s)+p(s)) \quad (2.9)$$

Error *:* $$e(s) = r(s) - y(s) = (1+WC_0)^{-1}(r(s)-d(s)-p(s)) + WC_0(1+WC_0)^{-1}n(s) \quad (2.10)$$

Observations*:* $$z(s) = WC_0(1+WC_0)^{-1}r(s) + (1+WC_0)^{-1}(d(s)+n(s)+p(s)) \quad (2.11)$$

Controller input *:* $$e_0(s) = r(s) - z(s) = (1+WC_0)^{-1}(r(s)-d(s)-n(s)-p(s)) \quad (2.12)$$

Control signal*:* $$u(s) = (1+WC_0)^{-1}C_0(r(s)-d(s)-n(s)-p(s)) \quad (2.13)$$

Robustness signal*:* $$h(s) = (1+WC_0)^{-1}p(s) \quad (2.14)$$

These equations include the following sensitivity operators:

Sensitivity: $$S = (1 + WC_0)^{-1} \tag{2.15}$$

Complementary sensitivity : $$T = 1 - S = WC_0 S \tag{2.16}$$

Control sensitivity: $$M = C_0 S = C_0(1 + WC_0)^{-1} \tag{2.17}$$

The system shown in Fig. 1 may be represented in polynomial form (*Kailath* 1980 [28]), where the system transfer functions are written as:

System : $$W = A^{-1}B \tag{2.18}$$

Reference generator: $$W_r = A^{-1}E \tag{2.19}$$

Input disturbance: $$W_d = A^{-1}C_d \tag{2.20}$$

Measurement noise: $$W_n = A^{-1}C_n \tag{2.21}$$

Robustness signal model: $$W_p = A^{-1}C_p \tag{2.22}$$

There is no loss of generality in assuming these models have a common denominator A polynomial. The various polynomials are not necessarily coprime but the system transfer function is assumed to be free of unstable hidden modes. The coprime representation of the system is denoted by $A_0^{-1}B_0$, where $B=B_0U_0$ and $A=A_0U_0$.

The spectrum of the signal $r(s)$-$d(s)$-$n(s)$-$p(s)$ in equations (2.12) and (2.13) is denoted by $\Phi_{ff}(s)$ and a generalised spectral-factor Y_f may be defined from this spectrum, using:

$$Y_f Y_f^* = \Phi_{ff} = \Phi_{rr} + \Phi_{dd} + \Phi_{nn} + \Phi_{pp} \tag{2.23}$$

In polynomial form $Y_f = A^{-1}D_f$. The disturbance model is assumed to be such that D_f is strictly Hurwitz and satisfies:

$$D_f D_f^* = EE^* + C_d C_d^* + C_n C_n^* + C_p C_p^* \tag{2.24}$$

The role of the robustness model W_p may now be explained since it is one of the components in the combined signal spectrum Φ_{ff} defined in (2.23). Clearly, if Φ_{pp} is dominant then the spectrum $\Phi_{ff} \to \Phi_{pp}$. This implies that $\Phi_{e_0 e_0} \to S\Phi_{pp}S^*$ and this spectrum will provide the robustness term (suitably weighted) introduced in the cost function, in the next section.

3.0 LQG CRITERION AND RESTRICTED STRUCTURE CONTROL

The LQG cost-function to be minimized (*Youla et al* 1976 [29]) is defined as:

$$J = \frac{1}{2\pi j} \oint_D \{Q_c(s)\Phi_{ee}(s) + R_c(s)\Phi_{uu}(s) + P_c(s)\Phi_{e_0e_0}(s)\}ds \tag{3.1}$$

where Q_c, R_c, P_c represent dynamic weighting elements, acting on the spectra of the error *e(t),* feedback control *u(t)* and controller input $e_0(t)$ signals. The R_c weighting term is assumed to be positive definite and Q_c , P_c are assumed to be positive-semidefinite on the D contour of the s-plane. The robustness weighting term can be motivated as in the dual criterion results of *Grimble* (1986 [30]). The error, control and robustness weightings can be written in polynomial form as:

$$Q_c = \frac{Q_{cn}}{A_q^* A_q} = \frac{B_q^* B_q}{A_q^* A_q}, \qquad R_c = \frac{R_{cn}}{A_r^* A_r} = \frac{B_r^* B_r}{A_r^* A_r} \quad \text{and} \qquad P_c = \frac{P_{cn}}{A_q^* A_q} = \frac{B_p^* B_p}{A_q^* A_q} \tag{3.2}$$

where A_q is a Hurwitz polynomial and A_r is a strictly Hurwitz polynomial (*Grimble and Johnson* 1988 [31]). The problem will be to minimize the above criterion with the controller chosen to have a specified structure [15]. The P_c weighting term represents a robustness weighting to be explained later. For later use let the weightings be rewritten using the common denominator, so that:

$$\tilde{Q}_c = Q_c + P_c = \frac{Q_{cn} + P_{cn}}{A_q A_q^*} = \frac{\overline{Q}_{cn}}{A_q A_q^*} \text{ or } \tilde{Q}_c = \tilde{Q}_{cn} / (A_q A_q^* A_r A_r^*) = \tilde{Q}_{cn} / (A_w A_w^*)$$

where $\tilde{Q}_{cn} = \overline{Q}_{cn} A_r A_r^*$ *and* $A_w = A_q A_r$. Similarly for the control weighting:

$$\tilde{R}_c = \frac{R_{cn}}{A_r A_r^*} = \tilde{R}_{cn} / (A_q A_q^* A_r A_r^*) = \tilde{R}_{cn} / (A_w A_w^*)$$

where $\tilde{R}_{cn} = R_{cn} A_q A_q^*$.

Theorem 3.1: Restricted Structure Single Degree of Freedom LQG Control Problem

Consider the LQG error and control weighted criterion defined in (3.1), and the system introduced in §2. The conditions that determine the LQG controller of restricted structure are derived below. The derivation includes the robustness/sensitivity costing and coloured measurement noise model. These terms were not considered in the previous polynomial approaches to the restricted structure control design problem. The cost-function to be minimised was defined in as,

$$J = \frac{1}{2\pi j} \oint_D \{Q_c \Phi_{ee} + R_c \Phi_{uu} + P_c \Phi_{e_0e_0}\}ds$$

Noting the independence of the noise sources and recalling (2.10) to (2.13) and the definitions for sensitivity in equations (2.15) to (2.17) obtain by substituting in (3.1) :

$$J = \frac{1}{2\pi j} \oint_D \left\{ \begin{aligned} &Q_c(1 - WM)\Phi_{ff}(1 - W^*M^*) + P_c(1 - WM)\Phi_{ff}(1 - M^*W^*) \\ &+ R_c M \Phi_{ff} M^* - Q_c(1 - WM)\Phi_{nn} - \Phi_{nn}(1 - M^*W^*)Q_c + Q_c \Phi_{nn} \end{aligned} \right\} ds$$

$$J = \frac{1}{2\pi j}\oint_D \left\{ \begin{matrix} (W^*\tilde{Q}_c W + R_c)M\Phi_{ff}M^* - M^*W^*(\tilde{Q}_c\Phi_{ff} - Q_c\Phi_{nn}) \\ -(\tilde{Q}_c\Phi_{ff} - Q_c\Phi_{nn})WM + \tilde{Q}_c\Phi_{ff} - Q_c\Phi_{nn} \end{matrix} \right\} ds \tag{3.3}$$

where $\tilde{Q}_c = Q_c + P_c$ and this may be written in the alternative forms $\tilde{Q}_c = \frac{\bar{Q}_{cn}}{A_q A_q^*} = \frac{Q_{cn} + P_{cn}}{A_q A_q^*}$ or using a *common denominator* for the spectral factor:

$$\tilde{Q}_c = \tilde{Q}_{cn} / (A_q A_q^* A_r A_r^*) = \tilde{Q}_{cn} / (A_w A_w^*) \tag{3.4}$$

where $\tilde{Q}_{cn} = \bar{Q}_{cn} A_r A_r^*$. Then define,

$$\Phi_{fp} = W^*(\tilde{Q}_c\Phi_{ff} - Q_c\Phi_{nn}) \tag{3.5}$$

and $\Phi_{ff} = \Phi_{rr} + \Phi_{dd} + \Phi_{nn} + \Phi_{pp}$.

The *generalised spectral factors* Y_c and Y_f due to *Shaked* (1976 [32]) may now be defined, using:

$$Y_c^* Y_c = W^*\tilde{Q}_c W + R_c \tag{3.6}$$

$$Y_f Y_f^* = \Phi_{ff} = \Phi_{rr} + \Phi_{dd} + \Phi_{nn} + \Phi_{pp} \tag{3.7}$$

Completing the squares in equation (3.3) obtain:

$$J = \frac{1}{2\pi j}\oint_D \{(Y_c M Y_f - \frac{\Phi_{fp}}{Y_c^* Y_f^*})(Y_c M Y_f - \frac{\Phi_{fp}}{Y_c^* Y_f^*})^* + \Phi_0\} ds \tag{3.8}$$

where

$$\Phi_0 = \tilde{Q}_c\Phi_{ff} - Q_c\Phi_{nn} - \frac{\Phi_{fp}\Phi_{fp}^*}{Y_c Y_c^* Y_f Y_f^*} \tag{3.9}$$

Substituting in the spectral-factor expressions (3.6) and (3.7), using the polynomial system models in equations (2.18) to (2.22), obtain

$$Y_f Y_f^* = (EE^* + C_d C_d^* + C_n C_n^* + C_p C_p^*)/(AA^*) \tag{3.10}$$

$$Y_c^* Y_c = (B^*\tilde{Q}_{cn} B + A^*\tilde{R}_{cn} A)/(A^* A_w^* A_w A) \tag{3.11}$$

where

$$\tilde{Q}_{cn} = (Q_{cn} + P_{cn})A_r A_r^* \quad \text{and} \quad \tilde{R}_{cn} = R_{cn} A_q A_q^* \tag{3.12}$$

To obtain the polynomial spectral-factors define the filter D_f and control D_c spectral factors from (3.6) and (3.7) respectively, as:

$$D_f D_f^* = EE^* + C_d C_d^* + C_n C_n^* + C_p C_p^* \tag{3.13}$$

$$D_c^* D_c = B^*\tilde{Q}_{cn} B + A^*\tilde{R}_{cn} A \tag{3.14}$$

Recalling that A_q and A_r are normally chosen to be coprime, the generalized spectral factors may be written in the form:

$$Y_f = A^{-1}D_f \quad \text{and} \quad Y_c = A_c^{-1}D_c \quad \text{where} \quad A_c = AA_w \quad \text{and} \quad A_w = A_q A_r$$

The various terms in the criterion (3.1) may now be simplified by substituting from the polynomial system models in and the spectral factor results given above.

$$\Phi_{fp} = W^*(\tilde{Q}_c\Phi_{ff} - Q_c\Phi_{nn}) = \frac{B^*}{A^*}\left(\frac{\bar{Q}_{cn}D_f D_f^* - Q_{cn}C_n C_n^*}{A_q A_q^* AA^*}\right)$$

$$\Phi_{fp}/(Y_c^* Y_f^*) = B^*\left(\frac{\bar{Q}_{cn}D_f D_f^*}{AA_q} - \frac{Q_{cn}C_n C_n^*}{AA_q}\right)\frac{A_r^*}{D_c^* D_f^*} \tag{3.15}$$

The measurement noise subsystem must be asymptotically stable and after cancellation of common terms may be written as: $A^{-1}C_n = A_{n0}^{-1}C_{n0}$, where $A = A_{n0}A_0$. The following diophantine equations must be introduced:

Feedback Diophantine equations: Calculate (G_0, H_0, F_0) with F_0 of minimum degree:

$$D_c^* G_0 + F_0 AA_q = B^* \bar{Q}_{cn} A_r^* D_f \tag{3.16}$$

$$D_c^* H_0 - F_0 BA_r = A^* R_{cn} A_q^* D_f \tag{3.17}$$

Implied equation: Multiplying (3.16) by BA_r and (3.17) by AA_q and adding the equations obtain:

$$D_c^*(G_0 BA_r + H_0 AA_q) = (B^* \tilde{Q}_{cn} B + A^* \tilde{R}_{cn} A)D_f$$

and after division by D_c^* obtain:

$$G_0 BA_r + H_0 AA_q = D_c D_f \tag{3.18}$$

Measurement noise equation:

$$D_c^* D_f^* X_0 + Y_0 A_{n0} A_q = B^* Q_{cn} C_{n0} C_n^* A_r^* \tag{3.19}$$

The terms in the cost-optimization problem may now be considered, substituting from the above polynomial equations. Substituting from the Diophantine equations (3.16) and (3.17):

$$\frac{\Phi_{fp}}{Y_c^* Y_f^*} = \left(\frac{G_0}{AA_q} + \frac{F_0}{D_c^*}\right) - \left(\frac{X_0}{A_{n0}A_q} + \frac{Y_0}{D_c^* D_f^*}\right) \tag{3.20}$$

Considering now the term $Y_c M Y_f$, writing $C_0 = C_{0d}^{-1}C_{0n}$, obtain :

$$Y_c M Y_f = \frac{D_c D_f C_{0n}}{A A_w (A C_{0d} + B C_{0n})} \tag{3.21}$$

The first squared term in (3.8), using (3.21) now becomes:

$$Y_c M Y_f - \frac{\Phi_{fp}}{Y_c^* Y_f^*} = \frac{[D_c D_f C_{0n} - (G_0 - X_0 A_0) A_r (A C_{0d} + B C_{0n})]}{A A_w (A C_{0d} + B C_{0n})} + \frac{(Y_0 - F_0 D_f^*)}{D_c^* D_f^*} \tag{3.22}$$

Substituting from the *implied Diophantine* equation:

$$Y_c M Y_f - \frac{\Phi_{fp}}{Y_c^* Y_f^*} = \frac{[(H_0 A_{n0} A_q + X_0 A_r B) C_{0n} - (G_0 A_{n0} - X_0 A) A_r C_{0d}]}{A_{n0} A_w (A C_{0d} + B C_{0n})} + \frac{(Y_0 - F_0 D_f^*)}{D_c^* D_f^*} \tag{3.23}$$

This cost term expression may be written in the form:

$$Y_c M Y_f - \frac{\Phi_{fp}}{Y_c^* Y_f^*} = T_1^+ + T_1^- \tag{3.24}$$

where the term within the square brackets in (3.23) denoted by T_I^+. This term is stable, since A_w is Hurwitz and the closed-loop characteristic polynomial $\rho_c = (A C_{0d} + B C_{0n})$ is required to be strictly Hurwitz for $J_{min} < \infty$. The final term in (3.23) is strictly unstable since D_c^* is strictly non-Hurwitz.

4.0 COST FUNCTION MINIMIZATION AND PARAMETRIC OPTIMIZATION

Given the simplification of terms in the cost-function presented above the cost minimization procedure may be followed (*Grimble and Johnson*, 1988 [31]). Note that the cost- function (3.8) may be written, using (3.24) as:

$$J = \frac{1}{2\pi j} \oint_D \{(T_I^+ + T_I^-)(T_I^+ + T_I^-)^* + \Phi_0\} ds \tag{4.1}$$

From the Residue theorem the integrals of the cross-terms $T_I^+ T_I^{-*}, T_I^- T_I^{+*}$ can be shown to be zero. This result follows because $\oint T_1^+ T_1^{-*} ds = -\oint T_1^- T_1^{+*} ds$ but the term $T_1^- T_1^{+*}$ is analytic for *all* s in the left half plane so that the sum of the residues obtained in calculating $\oint T_1^- T_1^{+*} ds$ is zero.

Note that this result still applies if the function $T_1^- T_1^{+*}$ contains poles on the $j\omega$ axis, since they can be avoided by the D contour, using small semi-circular detours in the left-half plane. These semi-circles are centred on these poles and do not contribute in the limiting case as the radius tends to zero. Also observe that the term containing $T_1^- T_1^{+*}$ could lead to an infinite cost should such terms be present. However, these may not be present, since the optimal control may be chosen so that they cancel. The practical case when this arises is when the error weighting includes an integrator $A_q(s) = s$. When the controller denominator $C_{0d}(s)$ includes integral action the A_q polynomial cancels throughout the term. The consequence is that the criterion can have a finite minimum, even though certain cost function terms include j axis poles. The cost-function therefore simplifies as:

$$J = \frac{1}{2\pi j} \oint_D \{(T_I^+ T_I^{+*} + T_I^- T_I^{-*}) + \Phi_0\} ds \tag{4.2}$$

Since the terms T_j^- and Φ_0 are independent of the controller, the criterion J is minimised when the first term involving T_I^+ is minimized. However, if the feedback controller C_0 has a restricted structure then it is unlikely that T_I^+ can be set to zero. It follows that to minimize the cost-function the first term in (4.2) should be minimised, through the choice of C_0, namely: $J_0 = \dfrac{1}{2\pi j} \oint_D \{T_1^+ T_1^{+*}\} ds$ (4.3)

For a finite solution to this cost minimization problem to exist the T_I^+ term must be asymptotically stable. Inspection of this term:

$$T_1^+ = \frac{[(H_0 A_{n0} A_q + X_0 A_r B) C_{0n} - (G_0 A_{n0} - X_0 A) A_r C_{0d}]}{A_{n0} A_w (A C_{0d} + B C_{0n})} \tag{4.4}$$

reveals that all terms are asymptotically stable but the weighting A_q could include a j axis zero (A_q is only assumed to be Hurwitz). However, it is assumed that although the structure of the controller $C_0 = C_{0n} C_{0d}^{-1}$ is limited, $C_{0\mathrm{d}}$ will have zeros at the j axis zeros of the chosen weighting A_q. Thus, such a zero will cancel and under the given assumptions T_I^+ is asymptotically stable.

Then the LQG controller of restricted structure may be calculated from a simple direct optimization problem. First compute the filtering and control spectral factors D_f and D_c (strictly Hurwitz due to the system description) using:

$$D_f D_f^* = EE^* + C_d C_d^* + C_n C_n^* + C_p C_p^* \tag{4.5}$$

$$D_c^* D_c = B^* \tilde{Q}_{cn} B + A^* \tilde{R}_{cn} A \tag{4.6}$$

The following *regulating Diophantine equations* must then be solved for (G_0, H_0, F_0), with F_0 of minimum degree:

$$D_c^* G_0 + F_0 A A_q = B^* \bar{Q}_{cn} A_r^* D_f \tag{4.7}$$

$$D_c^* H_0 - F_0 B A_r = A^* R_{cn} A_q^* D_f \tag{4.8}$$

and the following *measurement noise Diophantine equation* must be solved for (X_0, Y_0), with Y_0 of smallest degree:

$$D_c^* D_f^* X_0 + Y_0 A_{n0} A_q = B^* Q_{cn} C_{n0} C_n^* A_r^* \tag{4.9}$$

The optimal controller $C_0 = C_{0n} C_{0d}^{-1}$ must then be found to minimize the following component in the cost-function term:

$$J_0 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \{T_1^+(j\omega) T_1^+(-j\omega)\} d\omega \tag{4.10}$$

where $$T_1^+ = \frac{[(H_0 A_{n0} A_q + X_0 A_r B) C_{0n} - (G_0 A_{n0} - X_0 A) A_r C_{0d}]}{A_{n0} A_q A_r (A C_{0d} + B C_{0n})}$$

If the controller has a specified limited structure the minimum of the cost term J_{0min} will be non-zero. For an unconstrained solution the minimum is achieved when $T_1^+ = 0$ and the minimum of J_0 (denoted J_{0min}) is zero. It has been explained in Theorem 3.1, that the computation of the optimal feedback controller C_0 reduces to minimization of the term J_0.

It is clear from (4.4) that T_1^+ can be written in the form: $T_1^+ = \left(C_{0n} L_1 - C_{0d} L_2\right) / \left(C_{0n} L_3 + C_{0d} L_4\right)$ (4.11)

where $C_0 = C_{0n} / C_{0d}$ has a specified structure which can be as expressed below;

***Reduced order**:* $$C_0(s) = \frac{c_{n0} + c_{n1}s + \ldots + c_{np}s^p}{c_{d0} + c_{d1}s + \ldots + c_{dv}s^v}$$

where $v \geq p$ is less than the order of the system (plus weightings)

Lead lag: $$C_0(s) = \frac{(c_{n0} + c_{n1}s)(c_{n2} + c_{n3}s)}{(c_{d0} + c_{d1}s)(c_{d2} + c_{d3}s)}$$

PID: $$C_0(s) = k_0 + k_1 / s + k_2 s$$

The assumption must be made that a stabilising control law exists for the assumed controller structure. Note that the controller structure should be consistent with the choice of error weighting, in the sense that, if A_q includes a j axis zero, then the controller denominator $C_{0d}(s)$ should also include such a zero. The solution of this optimisation problem may be obtained using the following results. Assume, for example, that C_0 has a modified PID structure of the form:

$$C_0 = k_0 + (k_1 / s) + (k_2 s / (1 + s\tau)) \tag{4.12}$$

so that the numerator : $$C_{0n} = k_0(1 + s\tau)s + k_1(1 + s\tau) + k_2 s^2 \tag{4.13}$$

and the denominator : $$C_{0d} = s(1 + s\tau) \tag{4.14}$$

Let the superscripts r and i denote the real and imaginary parts of a complex function, so that $C_{0n} = C_{0n}^r + jC_{0n}^i$ and $C_{0d} = C_{0d}^r + jC_{0d}^i$. The controller numerator term may be split into frequency dependent components, through comparison with (4.13):

$$C_{0n}(j\omega) = -k_0\omega^2\tau + k_1 - k_2\omega^2 + j(k_0\omega + k_1\omega\tau) \tag{4.15}$$

and $$C_{0n}^r(j\omega) = -k_0\omega^2\tau + k_1 - k_2\omega^2 \quad \text{and} \quad C_{0n}^i(j\omega) = k_0\omega + k_1\omega\tau \tag{4.16}$$

Similarly, for the denominator term: $$C_{0d}(j\omega) = -\omega^2\tau + j\omega \tag{4.17}$$

and hence $$C^r_{0d}(j\omega) = -\omega^2\tau \quad \text{and} \quad C^i_{0d}(j\omega) = \omega \tag{4.18}$$

If the solution of the optimization problem is to be found by iteration, the denominator term in T_1^+ can be assumed to be known and the minimisation can then be performed on the numerator (linear terms). Thus, to set up this problem let, $T_1^+ = C_{0n} L_{n1} - C_{0d} L_{n2}$

where $$L_{n1} = L_1 / (C_{0n} L_3 + C_{0d} L_4) \quad \text{and} \quad L_{n2} = L_2 / (C_{0n} L_3 + C_{0d} L_4) \tag{4.19}$$

Substituting from (4.11) and (4.19).

$$T_1^+ = C^r_{0n} L^r_{n1} - C^i_{0n} L^i_{n1} - C^r_{0d} L^r_{n2} + C^i_{0d} L^i_{n2} + j(C^i_{0n} L^r_{n1} + C^r_{0n} L^i_{n1} - C^r_{0d} L^i_{n2} - C^i_{0d} L^r_{n2})$$

and after substitution from (4.16) and (4.18) obtain

$$T_1^+ = \begin{cases} k_0\left(-\omega^2\tau L^r_{n1} - \omega L^i_{n1} + j(\omega L^r_{n1} - \omega^2\tau L^i_{n1})\right) + k_1\left(L^r_{n1} - \omega\tau L^i_{n1} + j(\omega\tau L^r_{n1} + L^i_{n1})\right) \\ +k_2\left(-\omega^2 L^r_{n1} - j\omega^2 L^i_{n1}\right) + \omega^2\tau L^r_{n2} + \omega L^i_{n2} + j(\omega^2\tau L^i_{n2} - \omega L^r_{n2}) \end{cases}$$

The real and imaginary part of T_1^+ may therefore be written as: $T_1^+ = T_1^{+r} + jT_1^{+i}$ and it follows that, $\left|T_1^+\right|^2 = \left(T_1^{+r}\right)^2 + \left(T_1^{+i}\right)^2$. Write a vector form of the above equations as:

$$\begin{bmatrix} T_1^{+r} \\ T_1^{+i} \end{bmatrix} = F \begin{bmatrix} k_0 \\ k_1 \\ k_2 \end{bmatrix} - L = Fx - L,$$

where

$$F(\omega) = \begin{bmatrix} -\omega(\omega\tau L^r_{n1} + L^i_{n1}) & (L^r_{n1} - \omega\tau L^i_{n1}) & -\omega^2 L^r_{n1} \\ \omega(L^r_{n1} - \omega\tau L^i_{n1}) & (\omega\tau L^r_{n1} + L^i_{n1}) & -\omega^2 L^i_{n1} \end{bmatrix} \text{ and } L(\omega) = \begin{bmatrix} -\omega(L^i_{n2} + \omega\tau L^r_{n2}) \\ \omega L^r_{n2} - \omega^2\tau L^i_{n2} \end{bmatrix}$$

The cost-function can be optimised directly but a simple iterative solution can be obtained if the integral is approximated (*Yukitomo et al* 1998 [33]) by a summation with a sufficient number of frequency points $\{\omega_1, \omega_2, ..., \omega_N\}$. The optimisation can then be performed by minimising the sum of squares at each of the frequency points. The minimization of the cost term J_0 is therefore required where,

$$J_0 = \sum_{k=1}^{N} (Fx - L)^T (Fx - L) = (b - Ax)^T (b - Ax) \tag{4.20}$$

where $$A = \begin{bmatrix} F(\omega_1) \\ \vdots \\ F(\omega_N) \end{bmatrix}, \quad b = \begin{bmatrix} L(\omega_1) \\ \vdots \\ L(\omega_N) \end{bmatrix}, \quad x = \begin{bmatrix} k_0 \\ k_1 \\ k_2 \end{bmatrix} \tag{4.21}$$

Assuming the matrix $A^T A$ is non-singular the least squares optimal solution (*Noble* 1969 [34]) follows as:

$$x = (A^T A)^{-1} A^T b \tag{4.22}$$

Lemma 4.1 : Restricted Structure LQG Controller Solution Properties

The *characteristic polynomial* which determines stability and the *implied equation* are given as:

$$\rho_c = AC_{0d} + BC_{0n}$$

$$AA_q H_0 + BA_r G_0 = D_c D_f \tag{4.23}$$

The minimum value of the cost-function, with controller of restricted structure, is given as:

$$J_{min} = \frac{1}{2\pi j} \oint_D \{T_1^+ T_1^{+*} + T_1^- T_1^{-*} + \Phi_0\} ds$$

$$J_{\min} = \frac{1}{2\pi j} \oint_D \{T_1^+ T_1^{+*} + \frac{(Y_0 - F_0 D_f^*)^* (Y_0 - F_0 D_f^*)}{D_c^* D_c D_f D_f^*} + \Phi_0\} ds \tag{4.24}$$

Proof: The implied equation (4.23) follows from (4.7) and (4.8) by multiplying (4.7) by BA_r and (4.8) by AA_q, and adding.

4.1 DESIGN AND ROBUSTNESS IMPROVEMENT

The LQG controller should be designed in such a way that it is consistent with the restricted controller structure of interest. For example, A_q should approximate a differentiator if near integral action is required. In fact the assumption made in deriving Theorem 3.1 was that the controller structure is compatible with the choice of error weighting and if $1/A_q$ includes a j axis pole then this will be included in the chosen controller. In fact the usual situation will be that the designer decides the controller should include integral action and the weighting ($1/A_q$) will be chosen as an integrator. The control weighting $1/A_r$ is not so critical but if for example, a PID structure is to be used, then the point at which the differential (lead term) comes in can help to determine the A_r weighting. Clearly, there is no point in designing an LQG controller which has an ideal response, in some sense, but cannot be approximated by the chosen controller structure. Thus, the weightings should be selected so that the closed-loop properties are satisfactory but taking into consideration the limitations of the controller structure required. The basic concept proposed is straightforward. That is, in the region of the unity gain crossover frequency for the open loop system, or the phase margin frequency, the distance $|1+WC_0|$ should normally be maximised. This requires the sensitivity to be minimized, particularly in this sensitive region. By costing the sensitivity directly a mechanism is provided to improve robustness (*Horowitz* 1979 [35]) but there are some subtleties to address:

- The weighting P_c needs to be increased from zero where performance is presumably maximized (the LQG cost is optimized) up to a level where robustness is adequate and performance still acceptable.
- If pure sensitivity costing is required P_c could cancel the combined noise dynamics $\Phi_{ff} = Y_f Y_f^*$ which is unrealistic. The alternative is to make the model $W_p = A^{-1} C_p$ large, relative to the other noise terms and also a constant ($C_p = \rho A$ say) and this will introduce a fictitious stochastic term affecting

the noise and disturbance rejection, and reference tracking, properties. The size of the scalar ρ also therefore involves a compromise.

- The weighting and frequency shaping effects are only important in the decade above and a little below the crossover frequency referred to. The shaping might therefore be introduced by a weighting that approximates an ideal window function but this increases the order of the weighting term.

The above design choices and trade-offs detract from the approach but the prize is quite important and worthy of the effort. That is, the provision of a tuning variable, or variables, in a cost index where the robustness properties of an optimal controller can be manipulated and traded against performance/stochastic properties.

4.2 APPLICATIONS OF RS TECHNIQUE FOR MIMO SYSTEMS

Although the theory and the derivation of the control law presented in the previous sections assumed for simplicity a single-input, single-output system, similar results can also be obtained in the multivariable case. The restricted-structure controller design problem then becomes particularly interesting since it can be considered from a number of viewpoints.

First and most obvious application is optimal tuning / benchmarking of the existing controller. In the case of an r×m multivariable plant, the controller transfer-function matrix is of size m×r, however only a few possible feedback connections between outputs and controls are probably used. In fact, the most common situation is a square system with as many inputs as outputs, controlled by a multi-loop (decentralized) controller – which is often a set of SISO PID or PI controllers. By specifying the required control structure to be the actual structure, it is possible to obtain "optimal" parameters for the existing controller, in terms of the given cost function.

Another use of the technique may be in I/O pairing. For example, in a 3x3 MIMO system, the following multi-loop PID control configurations are possible:

$$A)\begin{bmatrix} PID & & \\ & PID & \\ & & PID \end{bmatrix} \quad B)\begin{bmatrix} PID & & \\ & & PID \\ & PID & \end{bmatrix} \quad C)\begin{bmatrix} & PID & \\ PID & & \\ & & PID \end{bmatrix}$$

$$D)\begin{bmatrix} & & PID \\ PID & & \\ & PID & \end{bmatrix} \quad E)\begin{bmatrix} & & PID \\ & PID & \\ PID & & \end{bmatrix} \quad F)\begin{bmatrix} & PID & \\ & & PID \\ PID & & \end{bmatrix}$$

The problem is the "best" choice of input/output pairs for a multi-loop control system. Methods such as Relative Gain Array are widely used in industry, however they take into account only steady-state information. On the other hand, pre-calculating the values of the cost function for all possible multi-loop configurations will determine the best "dynamic" pairing.

Finally, by computing benchmark figures corresponding to all possible controller configurations, with off-diagonal elements, it is possible to determine the potential benefits resulting from the introduction of additional loops to the system. If this turns out to be greater than the cost of installation, wiring, maintenance etc., the optimal tuning parameters are then available.

5.0 ROBUST CONTROL DESIGN EXAMPLE

To illustrate the effect of the robustness weighting element and the fictitious robustness signal *{p(t)}* a simple example is considered. The system models may be listed as:

Coprime system model:

$$W = \frac{B_0}{A_0} = \frac{1000(s+2)(s+6)}{(s+0.7)(s+3.9)(s+100)(s^2+2\xi\omega_0 s+\omega_0^2)}$$

where $\xi = 0.1$ and $\omega_0 = 10$. Also write

$$W = \frac{B}{A} = \frac{B_0 U_0}{A_0 U_0} \quad \text{where} \quad U_0 = (s+3.2)s$$

Disturbance model: $$W_d = \frac{C_d}{A} = \frac{1000}{(s+100)(s^2+2\xi\omega_0 s+\omega_0^2)(s+3.2)s}$$

where, $C_d = 1000(s+0.7)(s+3.9)$

Reference model : $$W_r = \frac{E}{A} = \frac{1}{(s^2+2\xi\omega_0 s+\omega_0^2)s}$$

where $E = (s+0.7)(s+3.9)(s+3.2)(s+100)$

Noise model: $$W_n = \frac{C_n}{A} = \frac{0.1}{(s+100)}$$

where $C_n = 0.1(s+0.7)(s+3.9)(s^2+2\xi\omega_0 s+\omega_0^2)(s+3.2)s$

Fictitious robustness signal: $$W_p = \frac{C_p}{A} = \frac{\rho}{(s+100)}$$

Cost Function Weightings

The cost function weightings may be defined as: $Q_c = \dfrac{(0.01s+1)(-0.01s+1)}{(s+10^{-6})(-s+10^{-6})}$

$$R_c = (10s+1)(-10s+1) \text{ and } P_c = \rho_1 \frac{(0.01s+1)(-0.01s+1)}{(s+10^{-6})(-s+10^{-6})}$$

Results

The frequency responses of the different system models are shown in Fig. 2. The system is low pass with a resonant subsystem and the disturbance model includes an integrator. The measurement noise model only rolls off at high frequencies. Consider first the full order optimal case and the use of a large $\rho = 1000$, then as ρ_1 varies the unit step responses of the closed loop system are as shown in Fig. 3. This represents the case where there is a large fictitious disturbance model W_p but where the robustness weighting ρ varies between 0 to 1000. The faster responses occur as ρ increases, since the effect is related to that when Q_c increases. The corresponding closed-loop frequency responses are shown in Fig. 3.

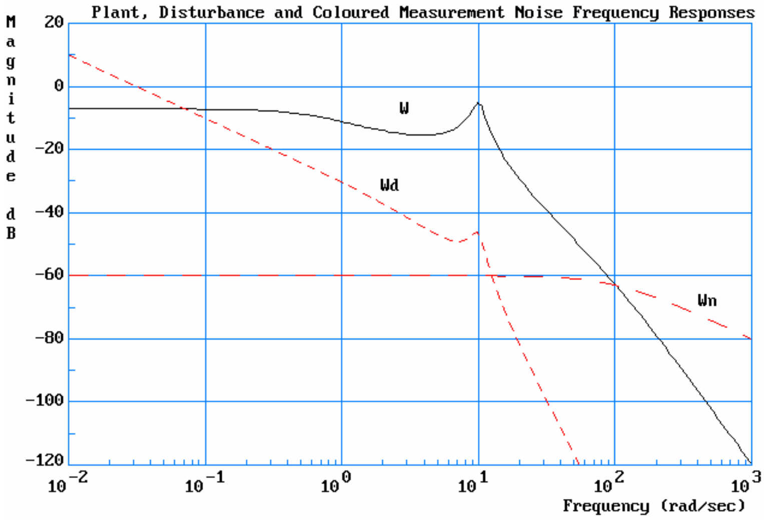

Figure 2: Bode Frequency Responses for the System, Disturbance and Noise

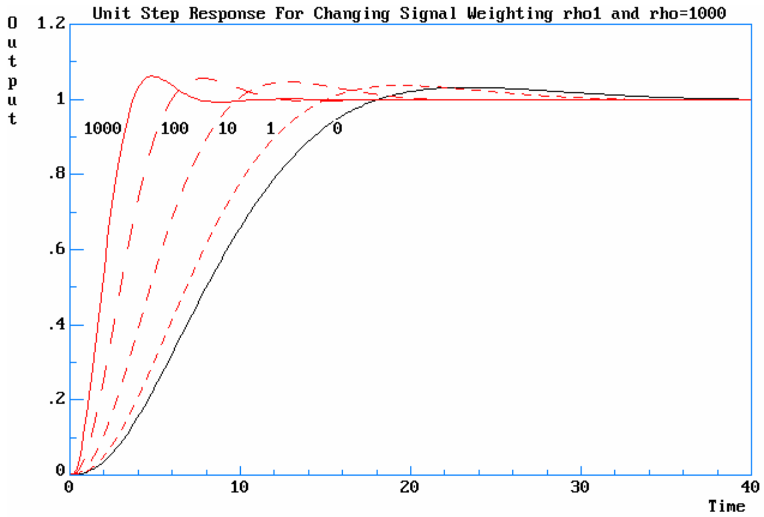

Figure 3: Closed Loop System Unit Step Responses for Fixed ρ = 1000 and Varying Weighting ρ_1

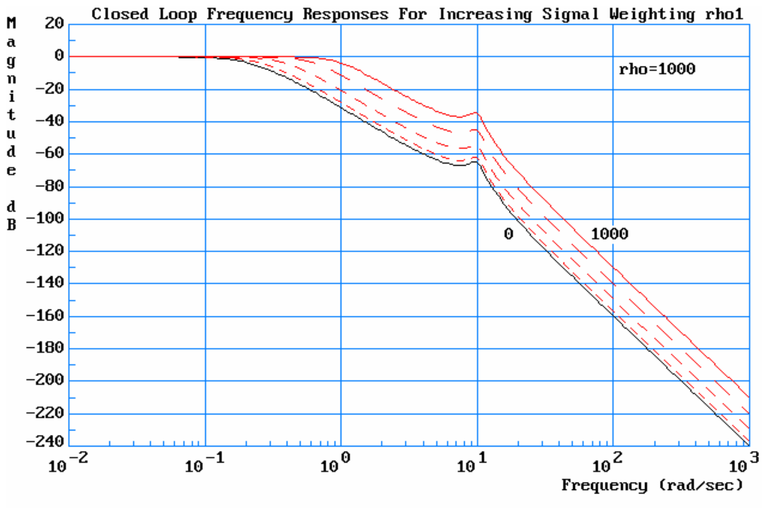

Figure 4: Bode Diagram of Closed Loop Responses for ρ = 1000 and Changing Value of ρ1

Restricted structure controller

The unit step responses are compared in Fig. 5 for the full-order and restricted controller designs. The case $\rho = 0$ and $\rho_1 = 10$ was considered and the corresponding controller and sensitivity-function frequency responses are shown in Figs 6 and 7. The restricted structure control is particularly good from an overshoot perspective. However, one reason is the higher controller gains at high frequencies for the restricted structure control law. The computed controllers were obtained as:

Optimal Full-Order Controller:

$$C_0(s) = \frac{0.493426(s+0.2205695)(s+0.7)(s+3.199631)(s+3.9)(s^2+1.999939s+99.99969)(s+100)}{(s+0.000001)(s+1.09342\times10^{-6})(s+1.008371)(s+3.866217)(s^2+14.81793s+65.05414)} \times (s^2+7.520861s+166.0202)$$

Optimal Restricted Structure:

$$C_0(s) = \frac{-0.0641248s^2 - 0.3587621s + 1.208584}{s(0.2s+1)}$$

The step responses shown in Fig. 8 are for the case $\rho = \rho_1 = 100$. The results are much faster and the restricted structure design is again good, relative to the full-order solution.

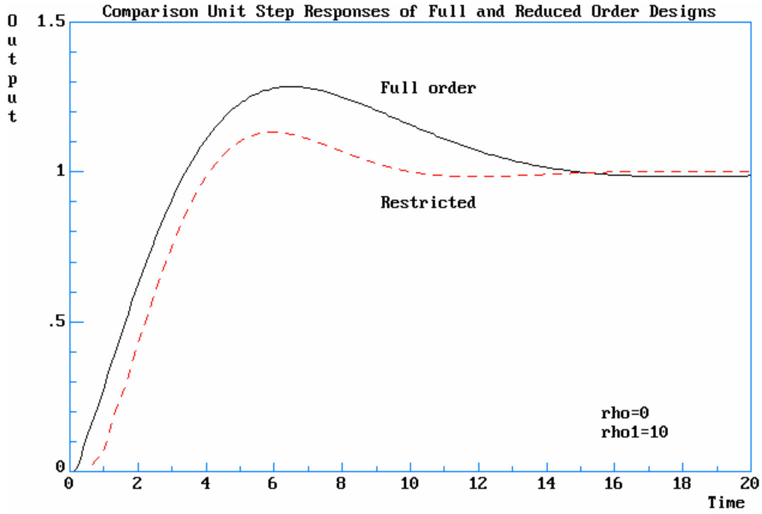

Figure 5: Comparison of Closed-Loop Unit Step Responses of Full and RS Control Designs

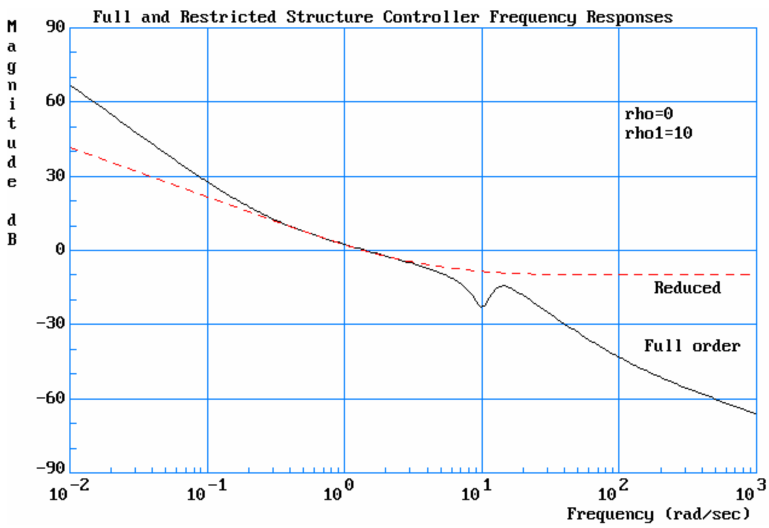

Figure 6: Bode Comparison of Controller Frequency Responses

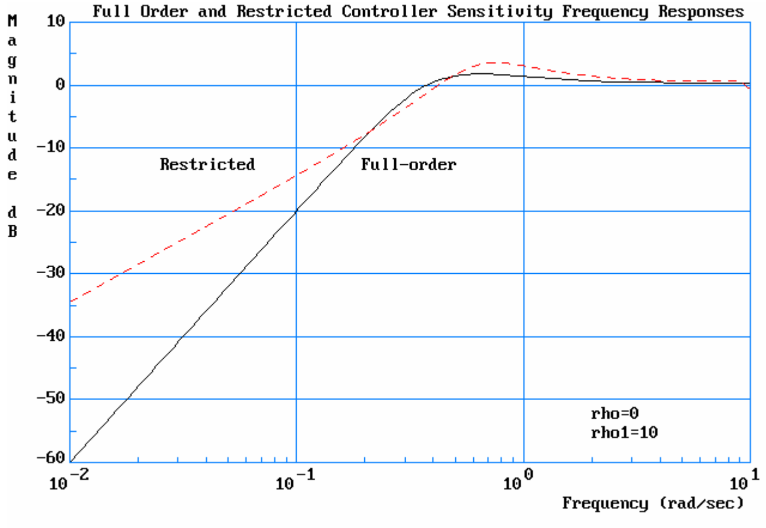

Figure 7: Bode Comparison of Sensitivity Function Frequency Responses

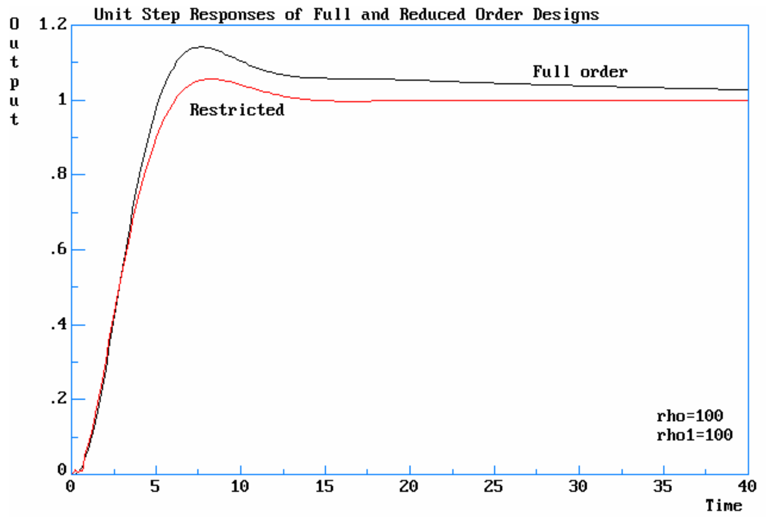

Figure 8: Comparison of Unit Step Responses for Full-Order and RS Control Designs

Example Conclusions

The example reveals that the robustness weighting terms and the fictitious robust costing signal *{p(t)}* certainly affects the overshoots which represent a measure of robustness, both on closed-loop frequency and time responses. The tuning variables ρ and ρ_1 affect the robustness of this minimum-phase open loop stable system in much the way expected. However, the alteration of robustness properties is not a straightforward matter, since any values of ρ and ρ_1 above zero will cause a measure of sub-optimality in stochastic (LQG cost) terms. The most surprising results were the very good results obtained for the restricted structure control designs. The explanation was the higher high frequency gains employed that reduced the peaks on the sensitivity function frequency responses. In this problem changes in the measurement noise model again did not have a large effect. However, results that were not shown were obtained for a coloured measurement noise model with a peak in the low frequency range. In this case the controller gains are significantly reduced and this slows the speed of response of the system and the overshoot increases markedly.

6.0 NONLINEAR GENERALIZED MINIMUM VARIANCE CONTROL

The system description is of restricted generality and is carefully chosen so that simple results are obtained. The plant itself is nonlinear and may be time-varying and have quite a general form. However, the reference and disturbance signals are assumed to have linear time-invariant model representations. This is not very restrictive, since in many applications the models for the disturbance and reference signals are only LTI approximations.

The system is shown in Fig. 9 and includes the *nonlinear* plant model and the linear reference/disturbance models. There is no loss of generality in assuming that the zero mean white noise sources $\{\zeta(t)\}$ and $\{\xi(t)\}$ have identity covariance matrices (Atherton 1982 [36]). There is no requirement to specify the distribution of the noise sources, since it will be shown that the special structure of the system leads to a prediction equation, which is dependent upon the *linear* disturbance and reference models.

6.1 System Models and Signals

The polynomial matrix system models, for the ($r \times m$) multivariable system, shown in Fig. 9, may now be introduced. Part of the system is represented by linear and part by nonlinear models. The linear disturbance, reference and plant *output* subsystem models have the left-coprime polynomial matrix representation:

$$[W_{d0}(z^{-1}),\ W_r(z^{-1}),\ W_{0k}(z^{-1})] = A^{-1}(z^{-1})[C_d(z^{-1}),\ E_r(z^{-1}),\ B_{0k}(z^{-1})] \tag{5.1}$$

The polynomial matrix system models, for the system, may be listed as follows:

Disturbance model: $$W_d(z^{-1}) = A^{-1}(z^{-1})C_d(z^{-1})$$

Reference model: $$W_r(z^{-1}) = A^{-1}(z^{-1})E_r(z^{-1}) \tag{5.2}$$

Without loss of generality these models have the common denominator polynomial matrix $A(z^{-1})$. Note that the arguments of the polynomial matrices are often omitted for simplicity. The subsystem associated with the plant inputs is assumed to be unstructured (need not have known equations) and of the form:

Nonlinear time-varying plant model: $$(\mathcal{W}u)(t) = D_k(\mathcal{W}_k u)(t) \tag{5.3}$$

where D_k denotes a diagonal matrix: $D_k = \text{diag}\left\{z^{-k_1}, z^{-k_2}, ..., z^{-k_r}\right\}$ of the common delay elements in the respective output signal paths.

One of the main strengths of the method is that no model is required for the nonlinear subsystem: $(\mathcal{W}_k u)(t)$. It is necessary to assume this is stable and to have some means of computing the output from this block but a traditional equation based model is not essential. That is, look-up tables may be employed, old Fortran code may be available that enables the output to be computed for a given input, or as in current research, a fuzzy neural model, may be fitted to real plant data. These methods, which do not involve a conventional model, can provide all that is needed to compute the control law. Most of the results do not need a more detailed breakdown of the plant model structure. However, in the later sections, to show the system can be stabilised, it will be assumed that any unstable modes of the plant are included in a stable/unstable linear time invariant block of polynomial matrix form: $W_{0k} = A^{-1}B_{0k}$. Thus, the delay free plant model term $(\mathcal{W}_k u)(t) = W_{0k}(\mathcal{W}_{1k} u)(t)$ and hence the total plant model may be written as:

$$(\mathcal{W}u)(t) = D_k W_{0k}(\mathcal{W}_{1k} u)(t) \tag{5.4}$$

The signals shown in the system model of Fig. 9 may be listed as follows:

Error signal: $$e(t) = r(t) - y(t) \tag{5.5}$$

Plant output: $$y(t) = d(t) + (\mathcal{W}u)(t) \tag{5.6}$$

Reference: $$r(t) = W_r \omega(t) \tag{5.7}$$

Disturbance signal: $$d(t) = W_d \xi(t) \tag{5.8}$$

Combined signal: $$f(t) = r(t) - d(t) \tag{5.9}$$

The *power spectrum* for the combined reference and disturbance model can be computed, noting these are linear subsystems, using:

$$\Phi_{ff} = \Phi_{rr} + \Phi_{dd} = W_r W_r^* + W_d W_d^* \tag{5.10}$$

and the *generalized spectral-factor* Y_f may be computed using:

$$Y_f Y_f^* = \Phi_{ff} \qquad ((5.11)$$

where the system models ensure Y_f is strictly minimum phase. Note that a measurement noise model has not been included to simplify the equations. This is appropriate so long as the control cost-function weighting, introduced in the next section, ensures controller roll-off at high frequencies.

6.2 Optimal Nonlinear Generalized Minimum Variance (NGMV) Problem

The optimal NGMV control problem involves the minimisation of the variance of the signal $\{\phi_0(t)\}$ in Fig. 9. This signal involves a $(r \times m)$ *dynamic cost function weighting* matrix: $P_c(z^{-1})$ on the error signal, represented by linear polynomial matrices as: $P_c = P_{cd}^{-1}P_{cn}$. It also includes an m-square, *nonlinear dynamic control signal costing* operator term: $(\mathcal{F}_c u)(t)$. The choice of dynamic weightings is critical to the design and typically P_c is low-pass and $\mathcal{F}_c$ is a high-pass transfer. The signal:

$$\phi_0(t) = P_c e(t) + (\mathcal{F}_c u)(t) \tag{5.12}$$

is to be minimized in a variance sense, so that the cost index to be minimised:

$$J = E\{\phi_0^T(t)\phi_0(t)\} = E\{trace\{\phi_0(t)\phi_0^T(t)\}\} \tag{5.13}$$

where $E\{\cdot\}$ denotes the expectation operator. Note that in some applications the signal $\phi_0(t)$ may represent an inferred output. That is, this signal can represent the output from a subsystem that cannot be measured directly.

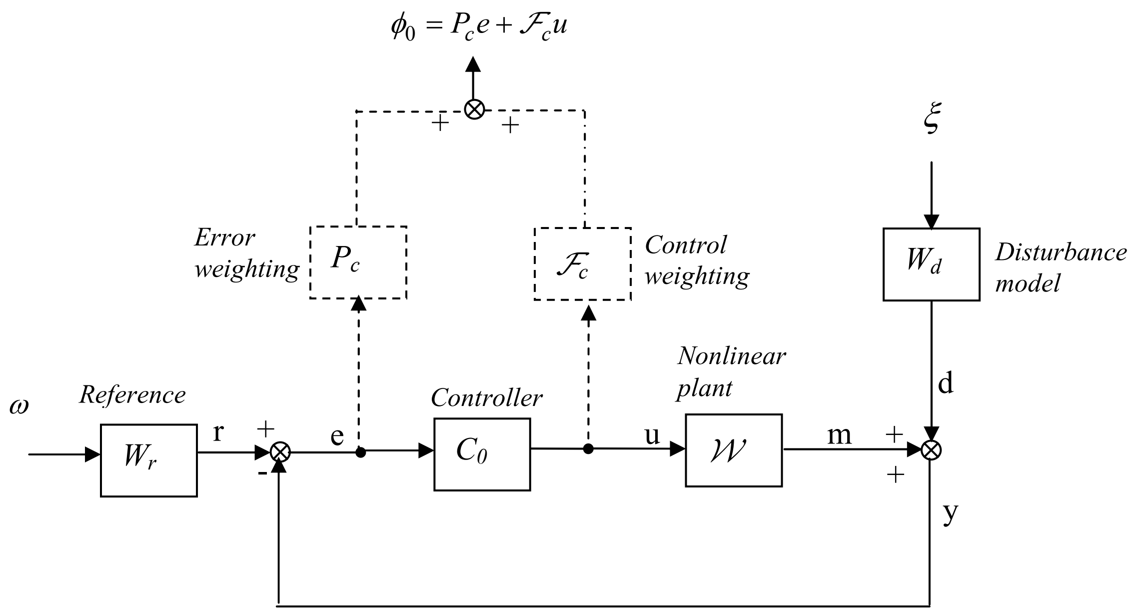

Figure 9: Single Degree of Freedom Closed Loop Feedback Control System for the Nonlinear Plant (signal ϕ_0 is dependent on the weightings shown dotted)

If the smallest of the delays in each output channel of the plant are of magnitudes: $\{k_1, k_2, ..., k_r\}$, respectively, this implies the control at time *t* affects the j^{th} output at least k_j steps later. For this reason the control signal costing can be defined to have the form:

$$(\mathcal{F}_c u)(t) = D_k (\mathcal{F}_{ck} u)(t) \tag{5.14}$$

Typically this will be a linear operator but it may also be chosen to be nonlinear to cancel the plant input nonlinearities in appropriate cases. The control weighting operator $\mathcal{F}_{ck}$ is assumed to be full rank and invertible.

Theorem 6.1: NGMV Optimal Controller

The NGMV optimal controller to minimize the variance of the weighted error and control signals may be computed from the following equations. The assumption is made that the nonlinear possibly time-varying operator $(P_c \mathcal{W}_k - \mathcal{F}_{ck})$ has a stable causal inverse, due to the choice of weighting operators P_c and $\mathcal{F}_c$. The smallest degree solution (G_0, F_0), with respect to F_0, must be computed from the polynomial matrix equation:

$$A_p P_{cd} F_0 + D_k G_0 = P_{cf} D_f \tag{5.15}$$

where the left coprime polynomial matrices A_{pf} and P_{cf} satisfy: $A_p^{-1} P_{cf} = P_{cn} A^{-1}$ (5.16) and the spectral factor Y_f is written in the polynomial matrix form: $Y_f = A^{-1}D_f$.

***Optimal control signal*:** The optimal NGMV control action can be computed as:

$$u(t) = \left(F_0 Y_f^{-1}\mathcal{W}_k - \mathcal{F}_{ck}\right)^{-1}\left(\left(A_p P_{cd}\right)^{-1} G_0 Y_f^{-1} e\right)(t) \tag{5.17}$$

Proof: The proof involves collecting results in the next section. ■

Remarks:

- The following solution is simplified if D_k and the weighting P_c and spectral factor Y_f commute. This assumption is certainly valid if the delay elements are the same in each channel $D_k = z^{-k}I$ or if P_c and Y_f are diagonal transfers, which is also reasonable for many applications.
- The class of problems considered are those for which a solution to the Diophantine equation can be found where the row degrees of $\mathrm{F}_0(z^{-1})$ are less than the delay path magnitudes $\{k_1, k_2, ..., k_r\}$ and this is ensured under the conditions listed in the previous remark.
- A further consequence of the above assumption is that the matrices: $F_0 Y_f^{-1}$, $G_0 Y_f^{-1}$ and D_k commute, which is a property employed later in the proof.

6.3 Solution of the Nonlinear Optimal Control Problem

A simple optimisation argument is used in the following. The signal to be minimised is shown to consist of both linear and nonlinear terms. However, the stochastic part of the problem involves linear models so that a prediction equation may easily be derived. This enables the signal to be written in terms of *future* and *past* white noise related terms. The optimal *causal* solution is therefore that which sets the past terms to zero. This will include some of the nonlinear control input dependent terms and the optimal control follows.

Consider the minimisation of the signal $\{\phi_0(t)\}$, where $\phi_0(t)$ represents the weighted sum of error and control signals and is the same dimension as the input signal. This fictitious or inferred output is defined as: $\phi_0(t) = P_c e(t) + (\mathcal{F}_c u)(t)$, where P_c is assumed to be a linear and $\mathcal{F}_c$ can be a linear or nonlinear operator. Now from the equations in §6.1: $e = r - y = r - d - \mathcal{W}u$ and hence,

$$\phi_0(t) = P_c(r - d - \mathcal{W}u) + \mathcal{F}_c u = P_c(r-d) - (P_c\mathcal{W} - \mathcal{F}_c)u \tag{5.18}$$

Notice this last term involves non-linear operators and the notation implies that the signal:

$$(P_c\mathcal{W} - \mathcal{F}_c)u = P_c(\mathcal{W}u)(t) - (\mathcal{F}_c u)(t)$$

Assumption: An important assumption will now be recalled that does not affect stability properties but may cause a degree of sub-optimality in disturbance rejection. That is, the model for the signal $f = r - d$ is assumed to be linear. In fact, disturbance models are often determined by nonlinear power spectrum models but are approximated well by a linear system driven by white noise. A typical example arises in ship positioning applications where the *Pierson-Moskowitz* spectrum is used to model wave motion but where this is normally approximated by a lightly damped second-order linear system.

Spectral Factor: Recall the signal: $f = Y_f \varepsilon$, where Y_f is a linear transfer and $\varepsilon(t)$ denotes a zero mean white noise signal of identity covariance matrix. The operator Y_f follows from a standard spectral-factor computation, given the disturbance W_d and the reference W_r signal models (Grimble, 1994 [18]). The Y_f may be assumed to have the following polynomial matrix form: $Y_f = A^{-1}D_f$ where from the system description D_f is strictly Schur. Thence, from the first term in (5.18):

$$P_c(r-d) = P_c f = P_c Y_f \varepsilon = P_{cd}^{-1} P_{cn} A^{-1} D_f \varepsilon$$

Now introduce the left coprime polynomial matrices A_p and P_{cf} satisfying: $P_{cn}A^{-1} = A_p^{-1}P_{cf}$

then $P_c f = \left(A_p P_{cd}\right)^{-1} P_{cf} D_f \varepsilon$ and from equation (5.15) the weighted error and control signals have the form:

$$\phi_0 = \left(A_p P_{cd}\right)^{-1} P_{cf} D_f \varepsilon - \left(P_c \mathcal{W} - \mathcal{F}_c\right) u \tag{5.19}$$

6.3.1 Diophantine equation

Introduce the following linear Diophantine equation, to expand the *combined disturbance and reference model* into two groups of terms:

$$A_p P_{cd} F_0 + D_k G_0 = P_{cf} D_f \tag{5.20}$$

where the solution for (F_0, G_0) satisfies the row j degree of $F_0 < k_j$. Hence,

$$P_c Y_f = \left(A_p P_{cd}\right)^{-1} P_{cf} D_f = F_0 + \left(A_p P_{cd}\right)^{-1} D_k G_0 \tag{5.21}$$

The first polynomial matrix includes delay elements in the j[th] channel, up to and including z^{-k_j+1} and the last term involves delay elements greater than or equal to k_j in each output channel. Substituting into (5.19), obtain the inferred output signal:

$$\phi_0 = F_0 \varepsilon + \left(A_p P_{cd}\right)^{-1} D_k G_0 \varepsilon - \left(P_c \mathcal{W} - \mathcal{F}_c\right) u \tag{5.22}$$

but $\varepsilon = Y_f^{-1} f = Y_f^{-1}\left(e + \mathcal{W}u\right) = Y_f^{-1}(r-d)$ and hence substituting in (5.22):

$$\phi_0 = F_0 \varepsilon + \left(A_p P_{cd}\right)^{-1} D_k G_0 Y_f^{-1} e - \left(-\left(A_p P_{cd}\right)^{-1} D_k G_0 Y_f^{-1} \mathcal{W} + P_c \mathcal{W} - \mathcal{F}_c\right) u$$

$$= F_0 \varepsilon + \left(A_p P_{cd}\right)^{-1} D_k G_0 Y_f^{-1} e + \left(\left(A_p P_{cd}\right)^{-1} \left(D_k G_0 - A_p P_{cd} P_c Y_f\right) Y_f^{-1} \mathcal{W} + \mathcal{F}_c\right) u \tag{5.23}$$

but $A_p P_{cd} P_c Y_f = A_p P_{cn} A^{-1} D_f = P_{cf} D_f$ and hence (5.21) gives: $D_k G_0 - P_{cf} D_f = -A_p P_{cd} F_0$

These last two equations then give the desired weighted error and control signal as:

$$\phi_0 = F_0\varepsilon + \left(A_p P_{cd}\right)^{-1} D_k G_0 Y_f^{-1} e + \left(\mathcal{F}_c - F_0 Y_f^{-1} \mathcal{W}\right) u \qquad (5.24)$$

The control signal at time t affects the j[th] system output at time $t+k_j$ and hence the control signal costing term $\mathcal{F}_c$ should include a delay of k_j steps, so that $\mathcal{F}_c = D_k \mathcal{F}_{ck}$. Moreover, since in general the control signal costing is required on each signal channel, the $\mathcal{F}_{ck}$ weighting may be defined to be of full rank and invertible. Equation (5.24) may be simplified further if $A_p P_{cd}$ and D_k and $F_0 Y_f^{-1}$ and D_k commute, which is certainly the case under the assumptions on P_c and Y_f discussed after the Theorem at the end of the last section. From (5.24) the inferred or fictitious output may be written as:

$$\phi_0(t) = F_0\varepsilon(t) + D_k\left((\mathcal{F}_{ck}u)(t) - F_0 Y_f^{-1}(\mathcal{W}_k u)(t) + \left(A_p P_{cd}\right)^{-1} G_0 Y_f^{-1} e(t)\right) \qquad (5.25)$$

6.3.2 Optimisation

To compute the optimal control signal inspect the form of the weighted error and control signals in equation (5.25). Since the row degrees of F_0 are required to be less than k_j (the magnitude of the delay in the j[th] channel) the j[th] row of the first term is dependent upon the values of the white noise signal components: $\varepsilon(t),\ldots,\varepsilon(t-k_j+1)$. The remaining terms in the expression for the j[th] row are all delayed by at least k_j steps and therefore depend upon the earlier values: $\varepsilon(t-k)$, $\varepsilon(t-k_j-1)$, ….it follows that the first and the remaining terms are statistically independent. The first term on the right of (5.25) is independent of the control action and the smallest variance is achieved when the remaining terms are set to zero. The optimal control signal must therefore satisfy:

$$u(t) = \mathcal{F}_{ck}^{-1}\left(F_0 Y_f^{-1}(\mathcal{W}_k u)(t) - \left(A_p P_{cd}\right)^{-1} G_0 Y_f^{-1} e(t)\right) \qquad (5.26)$$

and this may be represented in the block diagram of Fig. 10.

Recall the signal $\phi_0 = P_c e + \mathcal{F}_c u = P_c(r-y) + \mathcal{F}_c u$ involves a weighting $\mathcal{F}_c$ that normally has a negative sign to ensure ϕ_0 is minimized by a signal u having negative feedback. The forward path gain of the controller block is therefore usually positive.

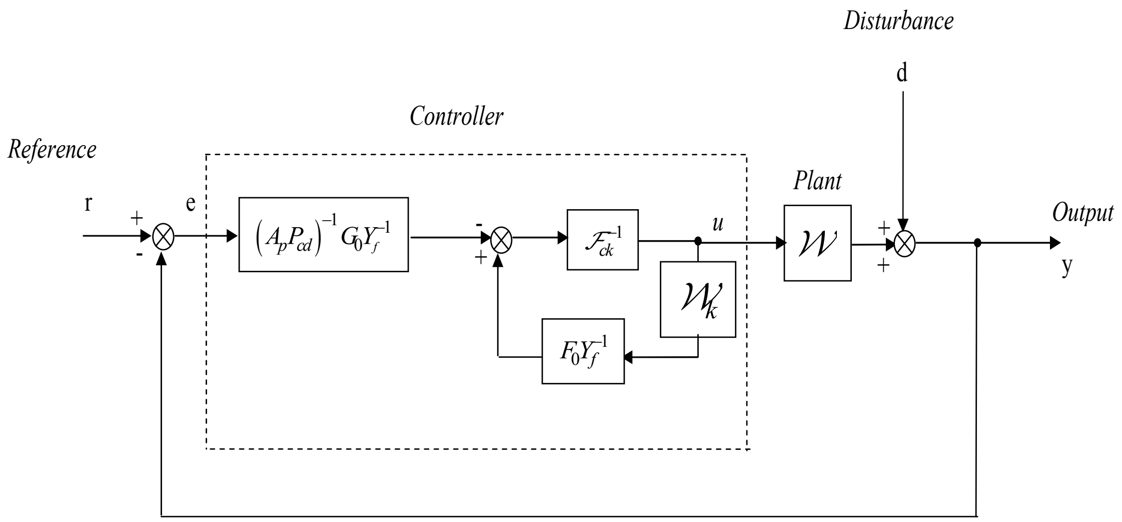

Figure 10: Control Signal Generation and Controller Modules

6.3.3 Alternative Expression for the Control Signal

An alternative expression for the control signal may be found that is useful for stability analysis. From equation (5.26), recalling that $F_0 Y_f^{-1}$ and D_k, are assumed to commute:

$$u(t) = \mathcal{F}_{ck}^{-1}\left(F_0 Y_f^{-1}(\mathcal{W}_k' u)(t) - (A_p P_{cd})^{-1} G_0 Y_f^{-1}\left(r(t) - d(t) - (\mathcal{W}u)(t)\right)\right)$$
$$= \mathcal{F}_{ck}^{-1}\left(\left(F_0 Y_f^{-1} + (A_p P_{cd})^{-1} D_k G_0 Y_f^{-1}\right)(\mathcal{W}_k' u)(t) - (A_p P_{cd})^{-1} G_0 Y_f^{-1}\left(r(t) - d(t)\right)\right)$$

Note from the Diophantine equation (5.20) the first term:

$$F_0 Y_f^{-1} + (A_p P_{cd})^{-1} D_k G_0 Y_f^{-1} = (A_p P_{cd})^{-1}(A_p P_{cd} F_0 + D_k G_0) Y_f^{-1} = (A_p P_{cd})^{-1} P_{cf} A \ = P_{cd}^{-1} P_{cn}$$

Thence, the expression for the control signal simplifies as:

$$u(t) = \mathcal{F}_{ck}^{-1}\left(P_c(\mathcal{W}_k' u)(t) - (A_p P_{cd})^{-1} G_0 Y_f^{-1}\left(r(t) - d(t)\right)\right)$$

or since:

$$(\mathcal{F}_{ck} - P_c \mathcal{W}_k')u = -(A_p P_{cd})^{-1} G_0 Y_f^{-1}\left(r(t) - d(t)\right)$$

$$u(t) = (P_c \mathcal{W}_k' - \mathcal{F}_{ck})^{-1}(A_p P_{cd})^{-1} G_0 Y_f^{-1}\left(r(t) - d(t)\right) \qquad (5.27)$$

where the existence of a stable causal inverse of the nonlinear operator $(P_c \mathcal{W}_k' - \mathcal{F}_{ck})$ is assumed. This latter result may also be confirmed by applying the optimisation argument directly to (5.24).

Note that the nonlinear operator may be computed, assuming the existence of the inverse of the control weighting $\mathcal{F}_{ck}$. For example, if the signal $\psi(t)$ is defined as:

$$\psi(t) = \left(P_c \mathcal{W}_k - \mathcal{F}_{ck}\right)u = P_c\left(\mathcal{W}_k u\right)(t) - \left(\mathcal{F}_{ck}\, u\right)(t)$$

then the control signal: $$u(t) = \mathcal{F}_{ck}^{-1}\left(P_c\left(\mathcal{W}_k u\right)(t) - \psi(t)\right)$$

but this cannot be used for computations, since the right-hand side also includes u at time t.

To avoid this problem in the computation of $u(t)$ the operator $\left(P_c \mathcal{W}_k - \mathcal{F}_{ck}\right)$ must be split into two parts involving a term without a delay $\mathcal{N}_0$ and a term that depends upon past values of the control action $\mathcal{N}_1$. That is,

$$\psi(t) = (P_c \mathcal{W}_k - \mathcal{F}_{ck})u = (\mathcal{N}_0 u)(t) + (\mathcal{N}_1 u)(t) \tag{5.28}$$

and the control signal can then be computed as:

$$u(t) = \mathcal{N}_0^{-1}\left(\psi(t) - (\mathcal{N}_1 u)(t)\right) \tag{5.29}$$

Since the weightings may be chosen freely the pulse response operator $\mathcal{N}_o$ can be assumed to be full rank and the inverse exists. This may be computed from $\mathcal{N}u = (P_c \mathcal{W}_k - \mathcal{F}_{ck})u$, if a model is available, by setting $z^{-1} = 0$, so that: $\mathcal{N}_0 = \mathcal{N}\big|_{z^{-1}=0} = (P_c \mathcal{W}_k - \mathcal{F}_{ck})\big|_{z^{-1}=0}$ and $\mathcal{N}_1 = \mathcal{N} - \mathcal{N}_0$. The results suggest the method of implementing the inverse operator shown in Fig. **11**.

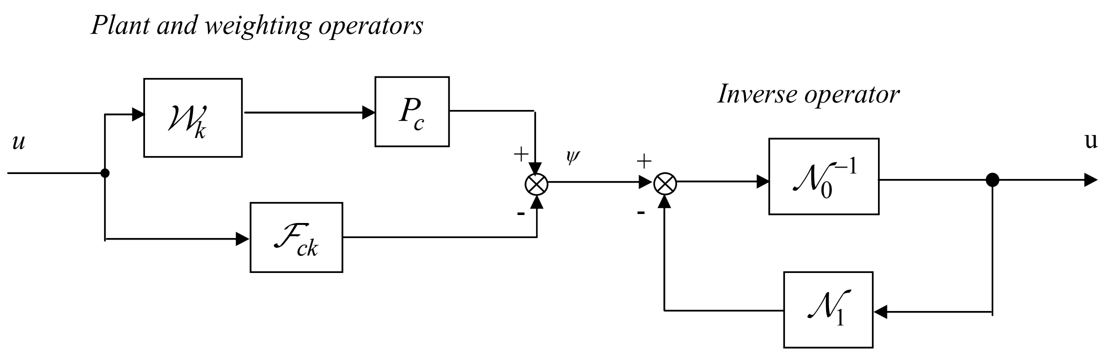

Figure 11: Nonlinear Time-Varying Operator $\left(P_c \mathcal{W}_k - \mathcal{F}_{ck}\right)$ and its Inverse

Limiting Case: Note that in the limiting case, for a square system, when $\mathcal{F}_{ck} \to 0$ the optimal control signal (5.27) becomes: $u_{\min\text{var}}(t) = \left(P_c \mathcal{W}_k\right)^{-1}\left(A_p P_{cd}\right)^{-1} G_0 Y_f^{-1}\left(r(t) - d(t)\right)$ and clearly the *minimum variance* control for the nonlinear system includes the stable inverse of the plant model, when one exists. The nonlinear terms are cancelled in this somewhat unrealistic case. In practice control costing must be employed to ensure unrealistic high gains do not occur, and hence pure *minimum variance* control is not normally an option. However, the link to this well known controller is valuable to provide confidence in the solution.

6.3.4 Existence of the Stable Nonlinear Operator Inverse

The above result (5.27) indicates a necessary condition for optimality is that the operator $\left(P_c\mathcal{W}_k - \mathcal{F}_{ck}\right)$ must have a stable inverse. For the case of linear systems the requirement is that the operator is strictly minimum-phase. This reveals that one of the restrictions on the choice of cost weightings is that this stability condition be fulfilled.

The simple solution obtained to this problem depends upon the existence of the stable inverse of the operator $\left(P_c\mathcal{W}_k - \mathcal{F}_{ck}\right)$. An important question is whether sensible choices of the weightings will lead to this condition. To show this covers a very wide class of systems consider the case where $\mathcal{F}_{ck}$ is linear and negative so that: $\mathcal{F}_{ck} = -F_k$. Then obtain: $\left(P_c\mathcal{W}_k + F_k\right)u = F_k\left(F_k^{-1}P_c\mathcal{W}_k + I\right)u$ and note that the term $\left(I + F_k^{-1}P_c\mathcal{W}_k\right)$ represents the return-difference operator for a system with feedback controller $K_c = F_k^{-1}P_c$. This interesting equivalence is important, since the stability of the inverse operator is clearly related to the stability of the feedback loop. It will be shown later that this feedback loop actually arises in a nonlinear version of the *Smith Predictor*.

***Starting Point for Weighting Selection**:* Consider the delay free plant $\mathcal{W}_k$ and assume a PID controller exists K_c to stabilize the closed loop system. Then a starting point for weighting choice, that will ensure $\left(P_c\mathcal{W}_k + F_k\right)$ is stably invertible, is $F_k^{-1}P_c = K_c$. That is, when the optimal controller is to be applied to an existing system it is very likely that a PID controller that stabilizes the plant is available, and will therefore provide a starting point for weighting selection. To demonstrate that such a weighting selection

is reasonable consider the scalar case and let a controller K_c have the form:

$$K_c = k_0 + \frac{k_1}{1-z^{-1}} + k_2\left(1-z^{-1}\right) = \left(\left(k_0+k_1+k_2\right) - \left(k_0+2k_2\right)z^{-1} + k_2z^{-2}\right)\Big/\left(1-z^{-1}\right) \tag{5.30}$$

and assume the PID gains are positive numbers, with small derivative gain. Then it is simple to confirm that if $F_k = 1$ the P_{cn} term is minimum phase and has real zeros. The $P_c = P_{cd}^{-1}P_{cn}$ term then includes integral action at low frequency and a lead term at high frequency, which implies the resulting optimal controller should have similar characteristics to the PID controller involving high gain at both low and high frequencies. This suggests that some modification will be necessary to limit the gain at high frequencies. Using the equivalence to a filtered PID controller leads to weightings of similar characteristics but with a limited gain at high frequencies. Recall the above argument was aimed at showing there are likely to be weighting choices that are reasonable and lead to a stable inverse for the operator $\left(P_c\mathcal{W}_k - \mathcal{F}_{ck}\right)$, under the above very reasonable assumption.

Realising the Controller and Minimum Cost

The controller expression may also be expressed using the inverse of the nonlinear operator (from (5.26)) as:

$$u(t) = \left(F_0Y_f^{-1}\mathcal{W}_k - \mathcal{F}_{ck}\right)^{-1}\left(\left(A_pP_{cd}\right)^{-1}G_0Y_f^{-1}e\right)(t) \tag{5.31}$$

The controller has the structure shown in Fig. 12, or equivalently, expanding the nonlinear compensator block, it corresponds with the controller structure of Fig. 10.

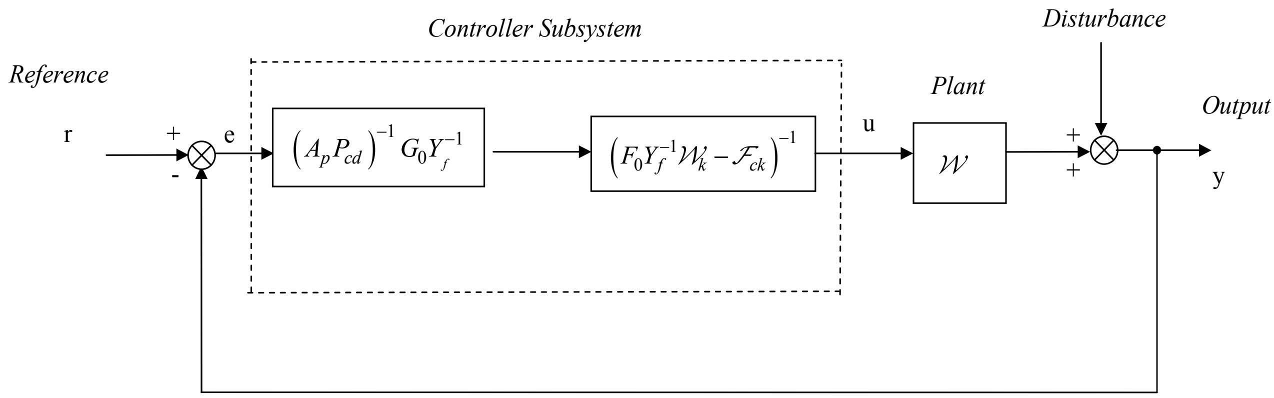

Figure 12: Equivalent Single Degree of Freedom Nonlinear Controller Structure

***Minimum Cost**:*

The *minimum cost* is clearly due to the first (time-invariant linear) term in (5.24). That is, using Parseval's theorem:

$$J_{\min} = E\left\{ \left(F_0 \varepsilon(t)\right)^T \left(F_0 \varepsilon(t)\right) \right\} = \frac{1}{2\pi j} \oint_{|z|=1} \text{trace} \left\{ F_0\left(z^{-1}\right) F_0\left(z^{-1}\right)^* \right\} \frac{dz}{z} \qquad (5.32)$$

It is interesting that this expression for the minimum cost, which can provide a benchmark cost for nonlinear controller design, depends only on the reference and disturbance signal models that are LTI. This arises because the control action effectively removes the nonlinear plant model from the prediction of the signal $\{\phi(t)\}$, whose variance is being minimised.

Minimum Variance and Special Cases

To verify the solution in the special case when the control weighting $\mathcal{F}_c$ tends to zero, a modification to the proof is required. From equation (5.26) the optimal control signal must satisfy:

$$F_0 Y_f^{-1}\left(\mathcal{W}_k u\right)(t) = \left(A_p P_{cd}\right)^{-1} G_0 Y_f^{-1} e(t)$$

For a square system, where the inverse of $\mathcal{W}_k$ exists, the *minimum variance* control, for a nonlinear, possibly time-varying, process: $u(t) = \left(\mathcal{W}_k^{-1}\left(Y_f F_0^{-1} P_{cd}^{-1} A_p^{-1} G_0 Y_f^{-1} e\right)\right)(t)$ (5.33)

and if the system is *scalar*, then for nonlinear systems: $u(t) = \left(\mathcal{W}_k^{-1}\left(F_0^{-1} P_{cd}^{-1} A^{-1} G_0 e\right)\right)(t)$

If now the plant is assumed linear then the plant operator $\mathcal{W}_k$ may be written in polynomial form as: $W_k = A^{-1} B_{0k}$ and the *scalar* minimum variance control: $u(t) = \dfrac{G_o}{B_{0k} F_0 P_{cd}} e(t)$ (5.34)

This is the stabilizing control if the process is minimum phase. This is a reminder that the *GMV control* *law*, with control costing, is only stabilizing when the control weighting is chosen appropriately. There are also many applications where the inclusion of the inverse of the plant model in the control law would suggest unrealistically high gains and poor behaviour to noise and uncertainties. However, this special case is relatively simple and there are some applications like the winding mechanisms for coal mines, where inverse static characteristics are employed routinely.

Relationship to the Smith Predictor

The optimal controller can be expressed in a similar form to that of a Smith Predictor. This provides a new nonlinear version of the Smith Predictor. Moreover, it provides an optimal method of tuning and provides optimal stochastic disturbance rejection and tracking properties. However, the introduction of this structure also limits the application of the solution on open-loop unstable systems. That is, although the structure illustrates a useful link between the new solution and the *Smith* time delay compensator, it also has the same disadvantage, that it may only be used on open-loop stable systems. Nevertheless the structure is intuitively reasonable and should be valuable in applications. This *Nonlinear Smith Predictor* will now be derived.

Observe that the system in Fig. 10 may be redrawn as in Fig. 13. The changes are made to the linear subsystems by adding and subtracting equivalent terms.

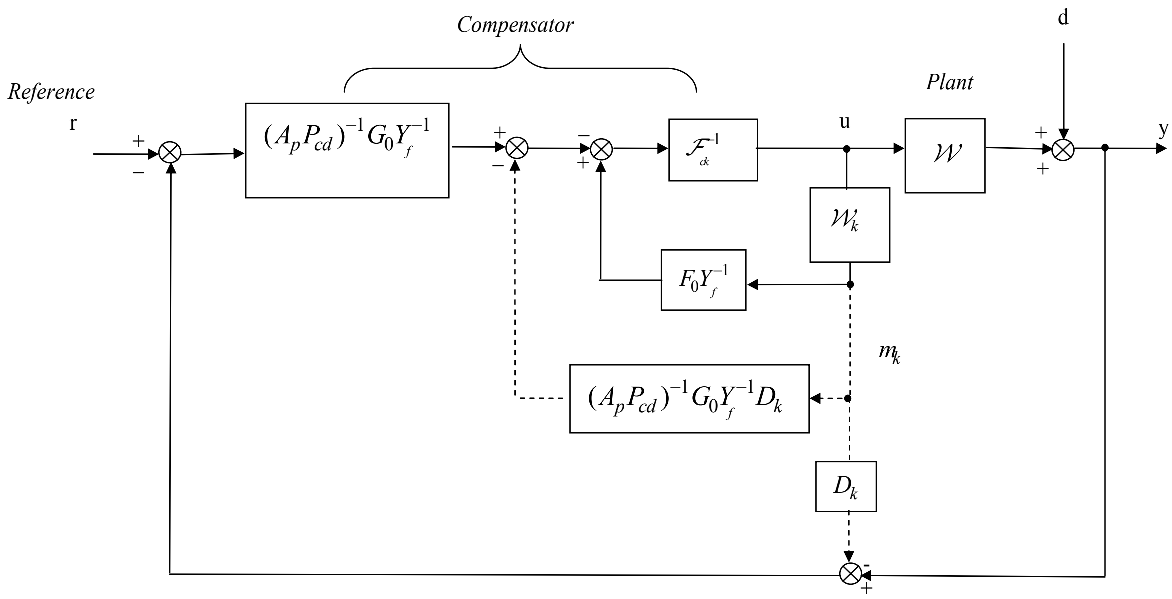

Figure 13: Modification to the Controller Structure Shown Dotted

Now combine the two linear inner loop blocks, by first defining the signal: $m_k(t)=(\mathcal{W}_k'u)(t)$ as follows: $\left(F_0Y_f^{-1}+\left(A_pP_{cd}\right)^{-1}G_0Y_f^{-1}D_k\right)m_k=\left(A_pP_{cd}\right)^{-1}\left(A_pP_{cd}F_0+D_kG_0\right)Y_f^{-1}m_k$

but substituting from (5.20), assuming D_k and G_0 commute:

$$\left(F_0Y_f^{-1}+\left(A_pP_{cd}\right)^{-1}G_0Y_f^{-1}D_k\right)m_k=P_cm_k \tag{5.35}$$

The system may therefore be redrawn as shown in Fig. 14 where the control action clearly satisfies equation (5.27). Now observe that the compensator may be rearranged, as shown in Fig. 15. This latter structure is essential if P_c includes an integrator, which introduces integral action. That is, P_{cd}^{-1} must be placed in the inner error channel, rather than in individual blocks as in Fig. 14.

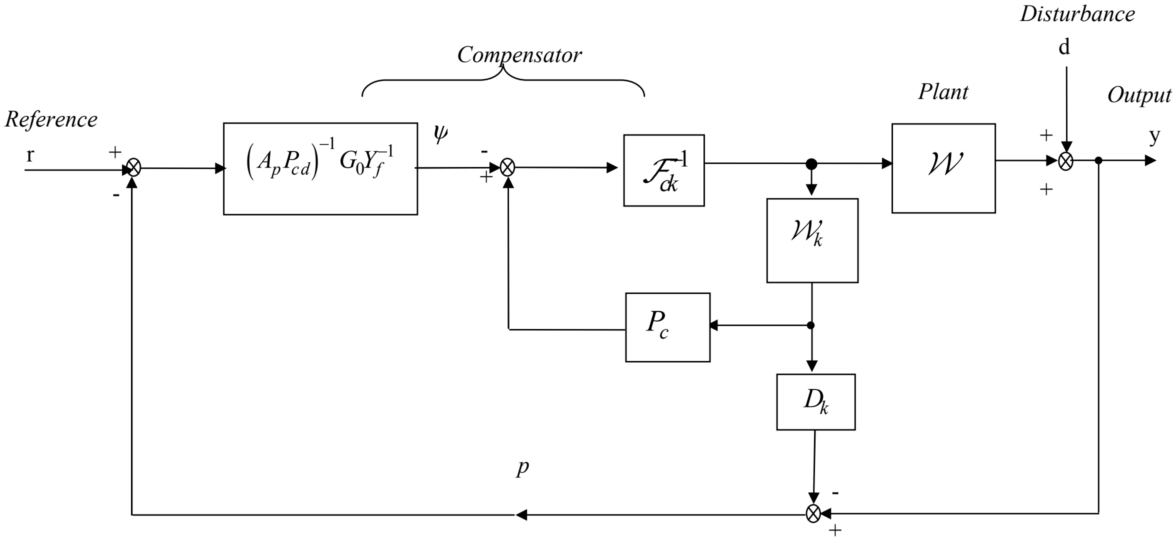

Figure 14: Nonlinear Smith Predictor Compensator and Internal Model Structure

The structure in Fig. 14 is intuitively reasonable and easy to explain. Note from the control signal u to the feedback signal p that the transfer is null when the model $D_k \mathcal{W}_k$ matches the plant model. It follows that the control action, due to reference signal r changes, is not due to feedback but involves the open-loop stable compensator involving the block $A_p^{-1} G_0 Y_f^{-1}$ and the *inner* nonlinear feedback loop. This inner-loop has the weightings $\mathcal{F}_{ck}^{-1} P_c$ acting like an inner-loop controller. If these weightings are chosen to be of the usual form it will represent a filtered PID controller. Thus, the control action due to the reference changes will be due to the cascade of these two blocks.

It should be emphasized that the choice of the weightings to be equal to a PID control law is only a starting point for design, since stability is easier to achieve. However, the control weighting can have an additional lead term (or alternatively a high frequency lag term may be added to the error P_c weighting. The high frequency characteristics of the optimal controller will then have more realistic roll off. Under the given assumptions the resulting system is stable. This follows because the plant is stable, the inner-loop is stable (due to choice of weightings) and there are only stable terms in the input block.

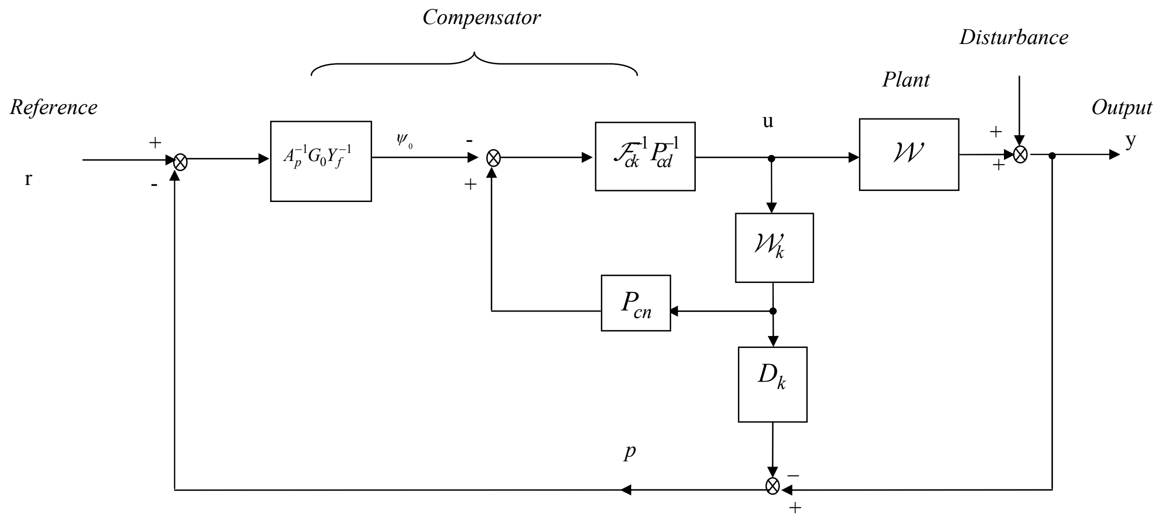

Figure 15: Nonlinear Smith Predictor Compensator and Internal Model Structure ($\psi_0 = P_{cd}\psi$)

Design Issues

In general, it seems relatively easy to obtain an NGMV design very close (and normally better) than the existing PID performance, and then use the proposed parameterization (which is only one of a number of possible choices) to achieve further improvement. The approach is not in competition with PID of course. There is every reason to use the simplest possible controller that will do the job. The NGMV was only compared with the PID design for this low order problem. It has the advantage that if the plant is high order, a stabilising PID control law may not exist

As the controller includes the nonlinear model of the plant, it should be robust against any changes of the operating point, whereas any linear controller may have problems regulating across the whole operating range. The above results do, of course, correspond to no (or little) plant/model mismatch. The choice of cost weightings to optimise robustness will be the subject of future research.

Some of the other areas where further investigation is warranted, that relate to design and applications, may be listed as:

- A slight generalization is to define a completely nonlinear objective function, so that the error weighting is nonlinear $\mathcal{P}_c$.
- Constraints on input actuators, like mechanical bending limits, can be allowed for using *barrier* functions, which may be absorbed into the plant model as a further nonlinearity.
- Integral action can be included by defining the error weighting to include an integrator. In this case the way in which the optimal control deals with windup should be analyzed.
- The inclusion of different model structures/types for the plant model $\mathcal{W}$, like a neural network, to provide a learning capability.
- A low order solution may be obtained using a r*estricted structure* control design philosophy.

7.0 NONLINEAR GMV CONTROL PROBLEM

The computation of a NGMV controller is illustrated below in the design of a scalar nonlinear discrete-time dynamic system, given in the following nonlinear state-space form:

$$x_1(t+1) = \frac{x_1(t)\cdot x_2(t)}{1+x_1^2(t)} + u(t)$$

$$x_2(t+1) = e^{-x_1^2(t)x_2^2(t)} + u(t)$$

$$y(t) = x_1(t-4)$$

Let the initial state *x(0) = 0* (the stable equilibrium point of the autonomous system). Observe that the output *y(t)* includes a transport delay of *k = 4* samples. The open-loop system response to a series of steps is shown in Fig. 16, and the nonlinearity present in the system is clearly evident from the wide range of responses.

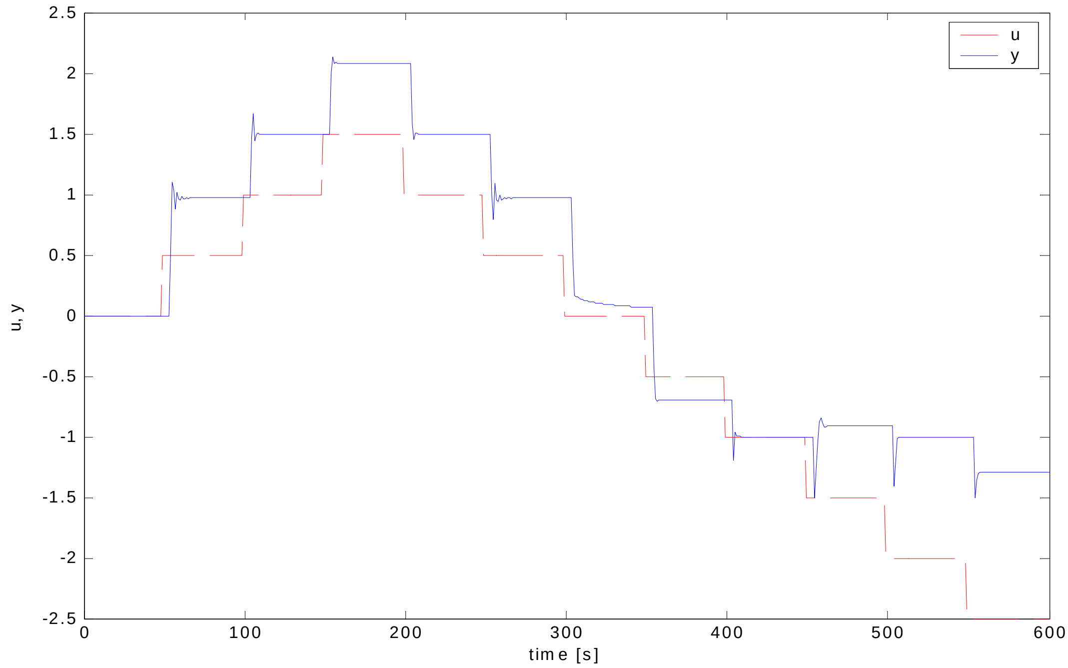

Figure 16: Open-loop Plant Responses for Operating Regions

For the nonlinear GMV controller design, the linear reference model has been defined as: $W_r = 0.05\big/\left(1-0.99z^{-1}\right)$, and is the stochastic analogue of a near step reference changes. The model of the additive linear disturbance acting on the system output was chosen as: $W_d = 0.05\big/\left(1-0.8z^{-1}\right)$. Assume the plant is controlled by the nominal stabilizing PID controller, denoted $C_1(z^{-1})$, with filtered derivative term: $C_1 = K\left(1+\dfrac{1}{T_i(1-z^{-1})}+\dfrac{T_d(1-z^{-1})}{(1-\tau_d z^{-1})}\right)$ and with the tuning parameters: K=0.1, T_i=4s, T_d=1s and τ_d=0.5. As explained in section 4.4, the nominal dynamic weightings for the NGMV design may be defined in terms of this controller as: $P_c = C_1,\quad F_c = -z^{-4}$. The Bode plots of these weightings are shown in Fig. 17.

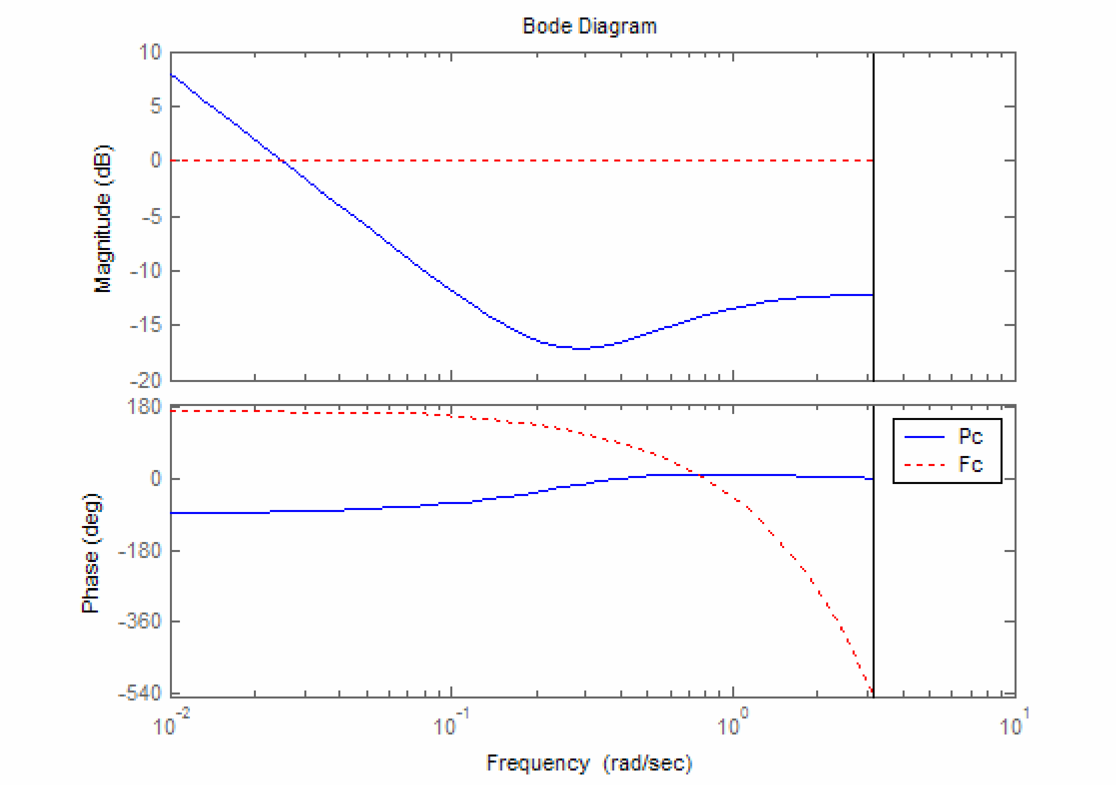

Figure 17: Frequency Responses of the Dynamic Weightings (nominal design)

The reference tracking of a sequence of steps for the two nominal controllers is shown in Fig. 18, and the corresponding output and control signal variances are collected in Table 1. Note that the nominal PID tuning parameters have only been found to stabilize the delay-free plant and are not ‘optimized’ in any sense. However, this controller is useful in that it can provide initial design parameters for the NGMV controller that will stabilize the plant, i.e. make the nonlinear operator stable and invertible. As can be seen from Fig. 18 and Table 1, the performance of the initial nonlinear controller design is close to that of the original PID, although it is normally more robust to the changes of the operating point (this can be seen for the set-point equal to zero). The stochastic performance of the nonlinear controller is also slightly better. The importance of this result is not the controller produced but that it provides a painless way to obtain an initial choice of cost weightings.

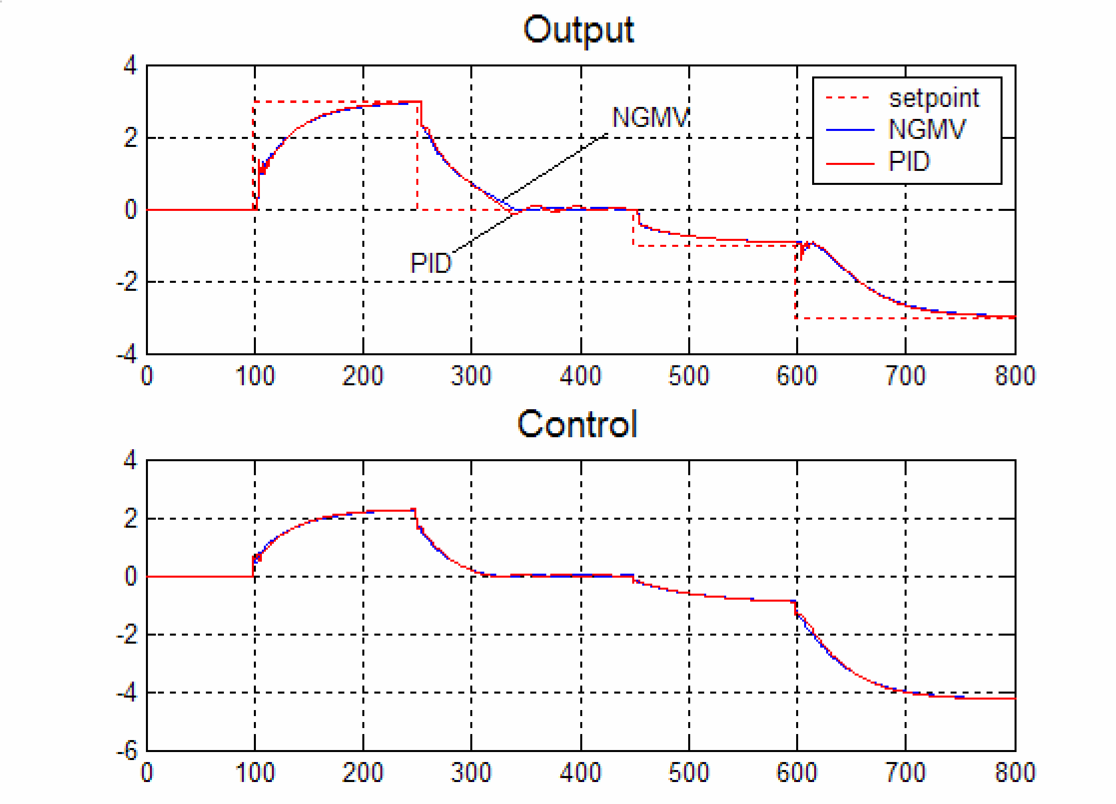

Figure 18: Time Responses: Nominal PID and NGMV Controllers

Table 1: Stochastic Performance: Nominal PID and NGMV Controllers

Op. point	Controller	Var[e]	Var[u]	Var[phi]
3	PID	0.01662	0.00117	0.00082
	NGMV	0.01672	0.00104	0.00082
0	PID	0.01568	0.00046	0.00054
	NGMV	0.01180	0.00025	0.00038
-1	PID	0.01122	0.00168	0.00051
	NGMV	0.01117	0.00166	0.00052
-3	PID	0.00729	0.00115	0.00035
	NGMV	0.00726	0.00115	0.00033

The nominal NGMV design will now be modified by changing the control weighting. Parameterize the control weighting as: $F_{ck} = -\rho(1-\gamma z^{-1})$, where ρ is a positive scalar and γ is a value from 0 to 1, to introduce a lead term to the weighting. Note that a lead term on the control signal weighting is normally useful to reduce the high frequency gain of the controller. For the nominal design: ρ=1 and γ=0. The Bode plots for some combinations of these two parameters are shown in Fig. 19.

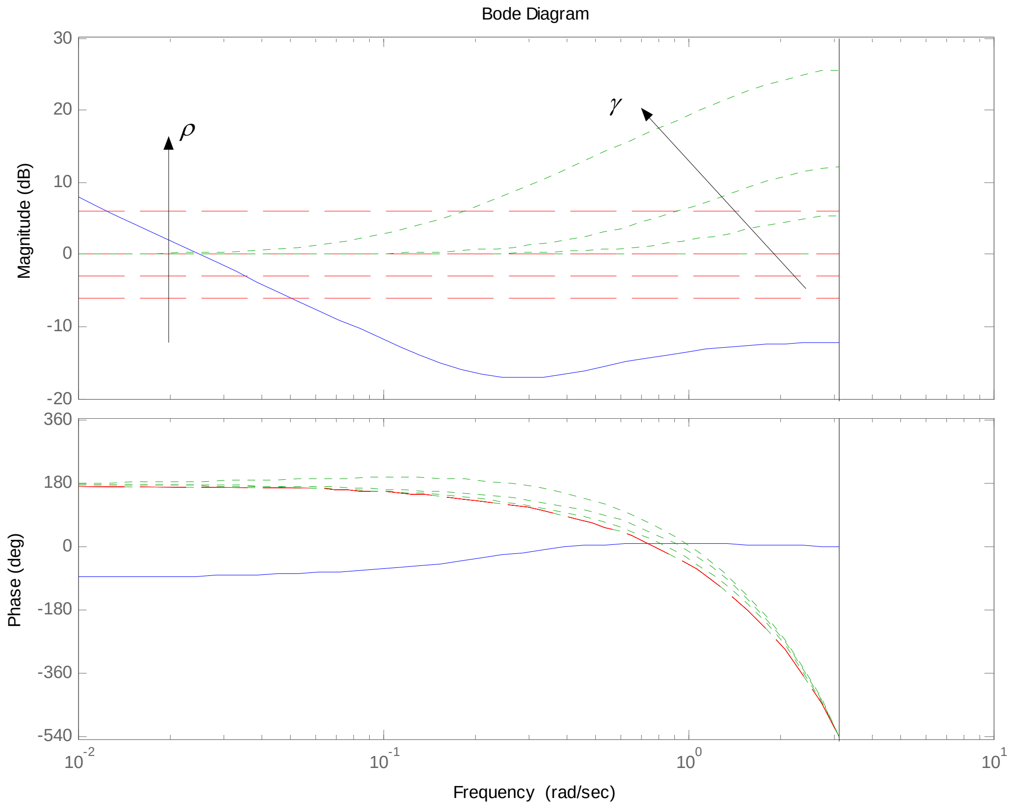

Figure 19: Frequency Responses of the Dynamic Cost Function Weightings Pc (solid), Fc with γ=0 (dashed), Fc with ρ=1 (dotted)

The parameterization involves two tuning parameters and is meant to simplify the design task. For example, decreasing the value of ρ (reducing the control weighting) leads to a faster response and a more violent control action. This can also be seen from the stochastic performance results (with added disturbance noise). Interestingly, there is very little change in the error variance. On the other hand, while decreasing ρ to a value of 0.35 leads to some undesirable oscillatory behaviour, adding a lead term helps resolve this problem. The Figure 12 and Table 2 present the simulation results for different values of ρ (γ=0). As expected increasing ρ results in a slower step response characteristic, providing a simple tuning mechanism.

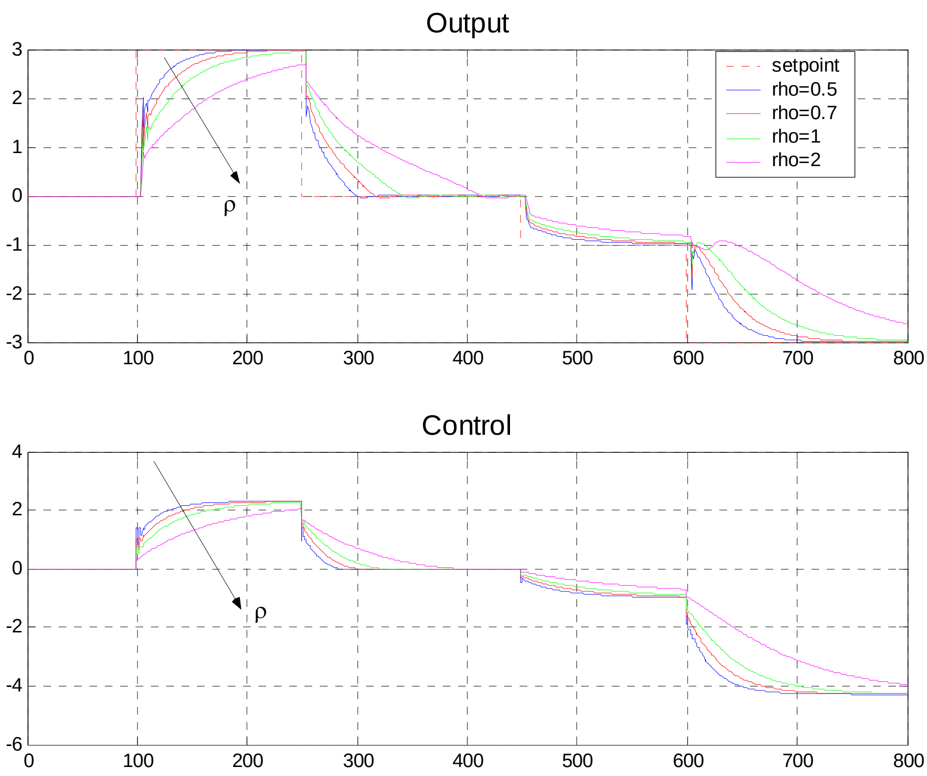

Figure 20: Time Responses for Weighting Parameters: ρ = 0.5, 0.7, 1, 2; γ = 0

Table 2: Stochastic Performance for Weighting Parameters: ρ = 0.5, 0.7, 1, 2; γ = 0

Op. point	rho	Var[e]	Var[u]	Var[phi]
3	0.5	0.01528	0.00334	0.00263
	0.7	0.01515	0.00199	0.00162
	1	0.01512	0.00120	0.00131
	2	0.01518	0.00049	0.00169
0	0.5	0.01330	0.00095	0.00233
	0.7	0.01324	0.00059	0.00126
	1	0.01322	0.00035	0.00077
	2	0.01312	0.00013	0.00055
-1	0.5	0.01042	0.00345	0.00352
	0.7	0.01017	0.00201	0.00171
	1	0.01015	0.00117	0.00127
	2	0.01023	0.00063	0.00214
-3	0.5	0.00604	0.00182	0.00258
	0.7	0.00593	0.00129	0.00140
	1	0.00595	0.00089	0.00104
	2	0.00617	0.00042	0.00141

For comparison, the nominal PID controller has been retuned and its performance compared with that of the NGMV controller with design parameters ρ=0.5 and γ=0.3. A set of PID parameters were obtained that were close to the NGMV design in terms of the speed of response, but the plant nonlinearity still caused some oscillatory behaviour in the PID control design responses.

In the last experiment, the plant time delay was increased from 4 to 10 samples. For the controller design, the same weightings were used as before but the NGMV controller obtained was of course different, reflecting the change in the time-delay. Then the Nonlinear Control Design Blockset of Matlab was used to find the optimal PID parameters, given the desired response. The boundary constraints were relaxed until a feasible set of parameters was found. However, it was not possible to tune the PID controller for satisfactory responses, across the whole operating range.

The Fig. 21 shows the response of the NGMV controller and of two of the PID controllers obtained. The dynamic response of the NGMV controller is very close to the original one, despite the significant increase in the time delay. It was not possible to obtain, for the PID controllers, both fast transient responses at the operating point = 3 and no oscillatory behaviour at the operating point = 0. The PID controller did not have time delay compensation, so it might be argued that it is not a fair comparison, nevertheless it demonstrates a potential of the NGMV controller to control highly nonlinear plants with significant time delays. Moreover, although the link to a Smith Predictor time delay compensator was made, the approach has the significant advantage over the Smith Predictor, that it provides a stochastic control design procedure, whereas the Smith Predictor only provides a structure (there is no guidance how to design the controller for say disturbance rejection).

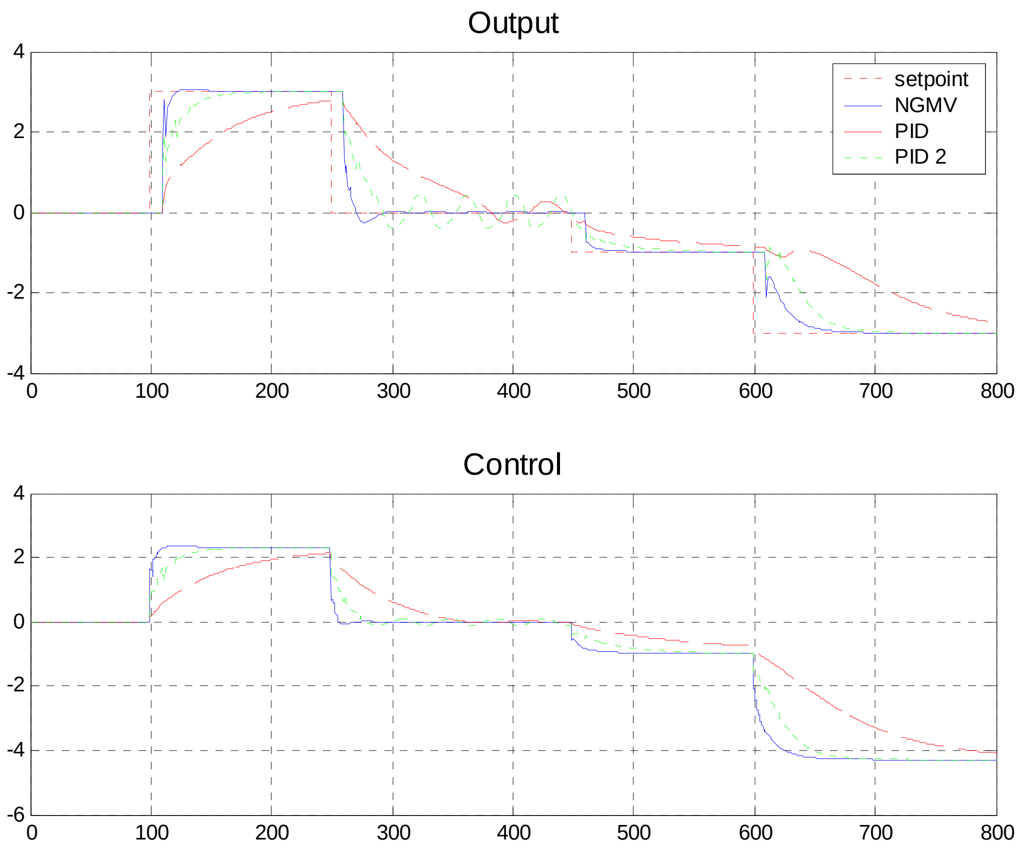

Figure 21: Time Responses for Time Delay Increased from 4 to 10 Samples

8.0 CONCLUSIONS

The robustness of full-order and restricted structure optimal control problems was considered for continuous-time linear systems. The emphasis was on the improvement of robustness by adding a sensitivity costing term in the cost index and by introducing a fictitious disturbance model W_p. The effect of measurement noise was also investigated and this was introduced in the feedback system model. The robustness weighting acts directly on the sensitivity function and may improve robustness margins but other properties will probably deteriorate like the measurement noise rejection properties. If robustness is more important than stochastic properties then attention would turn to H_∞ cost minimization and many of the ideas presented above would apply (Grimble, 1986 [30]). However, such an approach is readily embedded in the usual mixed sensitivity H_∞ design problem.

A relatively simple controller for nonlinear multivariable and possibly time-varying systems was also introduced. The closed loop stability of the system was shown to depend upon the existence of a stable inverse for a particular loop operator. This operator depended upon the cost weighting definitions. It was shown that a possible starting point for weighting selection was through the relationship to a PID controller. That is, if it is assumed that PID controller exists, to stabilize the delay free plant model $\mathcal{W}_k^{\prime}$, then this guaranteed the existence of at least one set of control weightings that would ensure closed-loop stability.

A major advantage of the NGMV solution is that the only knowledge of the nonlinear plant model that is required, is the ability to compute an output for a given control input sequence. Such a model, assumed stable, could be in Fortran or C code, or might even include look up tables or a neural network. The remaining computations concern the linear disturbance and reference signal models and knowledge of the transport delay element of length k. These are representative linear approximations and experience suggests they will be adequate so long as they capture the dominant frequency response behaviour. It follows that such a controller can be calculated without the usual model information required in traditional model based control law design.

The relationship to the Smith Predictor was discussed for two reasons. Firstly the extension of *Smith's* ideas to the nonlinear problem is interesting and provides a practical method of implementing these controllers, when the plant is open-loop stable. Secondly the physical structure is useful to provide an intuitive understanding of the operation and properties of the proposed Nonlinear GMV controller. The structure in Fig. 14, that describes the *Nonlinear Smith Predictor*, is particularly illuminating.

9.0 REFERENCES

[1] Kwakernaak, H., and R. Sivan, 1972, Linear optimal control systems, Wiley, New York.

[2] Desborough L. D. and Harris T.J. 1993, Performance assessment measures for univariate feedback/feedforward control, Canadian Journal of Chemical Engineering, Vol. 71, pp. 605-615.

[3] D. Uduehi, M. Grimble, 2001, Process Control Loop Benchmarking and Revenue Optimisation, Workshop Tutorial Paper, Proceedings of the American Control Conference, Arlington, USA, pp. 4313-27

[4] Anderson, B., and J. Moore, 1971, Linear optimal control, Prentice Hall, Englewood Cliffs.

[5] Safonov, M.G., and M. Athans, 1977, Gain and phase margin for multiloop LQG regulators, IEEE Trans. Aut. Cont., AC-22, 2, pp. 173-179.

[6] Doyle, J.C., 1978, Guaranteed margins for LQG regulators, IEEE Trans Auto. Control. AC-23, 4, pp. 756-757.

[7] Athans, M., 1986, A tutorial on the LQG/LTR method, Proc. American Control Conference, Seattle, WA.

[8] Martin, R., L. Valavani, and M. Athans, 1986, Multivariable control of a submersible using the LQG/LTR design methodology, Proc. American Control Conference, San Diego, pp. 1191-1200.

[9] Moore, J.B., and J.D. Blight, 1982, Performance and robustness trades in LQG regulator design, IEEE CDC Conf., San Diego, pp 1191-1200.

[10] Pfeil, W.H., M. Athans, and H.A. Spang III, 1986, Multivariable control of the G1700 engine using the LQG/LTR design methodology, Proc. American CDC, Seattle.

[11] Rodriguez, A., and M. Athans, 1986, Multivariable control of a twin-lift helicopter system using the LQG/LTR design methodology, Proc. American Control Conference, Seattle, WA.

[12] Stein, G., 1986, A critique of the LQG/LTR design method, Proc. American Control Conference, Seattle.

[13] Doyle, J.C., and G. Stein, Multivariable feedback design: concepts for a classical/modern synthesis, IEEE Trans. Auto. Control, AC-23, 4, pp. 607-611

[14] Grimble, M.J., 1999, Restricted structure feedforward stochastic optimal control, CDC 99, Phoenix, Arizona.

[15] Grimble, M.J., 2000, Restricted structure LQG optimal control for continuous time systems, IEE Proc. Control. Theory Appl. Vol. 147, No. 2.

[16] Grimble, M.J., 1986, Dual criterion stochastic optimal control problem for robustness improvement, IEEE Transactions on Automatic Control, Vol. AC 31. No. 2.

[17] Kucera, V., 1980, Stochastic multivariable control: a polynomial equation approach, IEEE Trans. On Auto. Contr., AC-25, 5, pp. 913-919.

[18] Grimble, M.J., 1994, Robust industrial control, Prentice Hall, Hemel Hempstead.

[19] Åström, K J, 1979, Introduction to stochastic control theory, Academic Press, London.

[20] Hastings-James, R, 1970, A linear stochastic controller for regulation of systems with pure time delay, University of Cambridge, Department of Engineering, Research Report No CN/70/3.

[21] Clarke, D W, and Hastings-James, R, 1971, Design of digital controllers for randomly disturbed systems, Proc. IEE, Vol. 118, No. 10, 1502-1506.

[22] Clarke, D W, and Gawthrop, P J, 1975, Self-tuning controllers, Proc IEE., Vol. 122, No 9, pp 929-934.

[23] Grimble, M J, 1981, A control weighted minimum-variance controller for non-minimum phase systems, Int J Control, Vol 33, No 4, pp 751-762.

[24] Grimble, M J, 1988, Generalized minimum-variance control law revisited, Optimal Control

Applications and Methods, Vol. 9, pp 63-77.

[25] Grimble, M J, 1993, H_∞ Multivariable control law synthesis, IEE Proceedings, Control Theory and Applications, Vol. 140, No 5, pp 353-363.

[26] Grimble, M J, 2001, Industrial control systems design, John Wiley, Chichester.

[27] Isidori, A, 1995, Nonlinear Control Systems, Springer Verlag, London.

[28] Kailath, T., 1980, Linear systems, Prentice Hall, Englewood, Cliffs, NJ

[29] Youla, D.C., H.A. Jabr and J.J. Bongiorno, 1976, Modern Wiener-Hopf design of optimal controllers, part 11 : the multivariable case, IEEE Trans. On Auto. Control, AC-21, No. 3, pp. 319-338.

[30] Grimble, M.J., 1986, Optimal H_∞ robustness and the relationship to LQG design problems, IJC, Vol. 43, No. 2, pp. 351-372.

[31] Grimble, M.J. and Johnson, M.A., 1988, Optimal multivariable control and estimation theory : theory and applications, Vols. I and II, John Wiley, London.

[32] Shaked, U., 1976, A general transfer function approach to linear stationary filtering and steady-state optimal control problems, Int. J. Control, 24, 6, pp. 741-770.

[33] Yukitomo, M., Yutaka Iino, Shiro Hino, Kazuki Takahashi and Kazue Nagata, 1998, A new PID controller tuning system and its application to a flue gas temperature control in a gas turbine power plant, IEEE Conference on Control Applications, Trieste, Italy.

[34] Noble, B., 1969, Applied linear algebra, Prentice Hall, Engelwood Cliffs.

[35] Horowitz, I.M., 1979, Quantitative synthesis of uncertain multiple-input, multiple-output feedback systems, Int. J. Control, Vol. 30, No. 1, pp. 81-106.

[36] Atherton D P, 1982, Nonlinear Control Engineering, Van Nostrand Reinhold, New York.