Multivariable Control of a Nonlinear Process

Output Prioritisation by Error Redistribution

Warren D. Carew

This dissertation is presented in fulfilment of the requirements for the degree of
Doctor of Philosophy in Electrical Engineering
at the University of Cape Town

Thesis Supervisor
Professor M. Braae

December 2001

Abstract

All real systems experience input saturation. Since the first proportional controller was implemented, systems have been susceptible to saturation. The introduction of the proportional-integral (PI) controller complicates matters further, since it reduces the steady state error to zero for zero-order setpoint tracking. This encompasses a fair portion of the implemented control strategies in the world.

It is clear that saturation, by limiting the input state for the process, limits the possible output region. This is true in an absolute as well as a relative sense, due to limited rate of change of the physical actuator. Hence not only is the absolute output range limited, but the nonlinear effects also limit the output value at the next instant.

In single variable digital systems the performance degradation can be crippling. Performance and stability are easily lost in systems with either high gains or small control ranges. In multivariable systems, the condition is aggravated by the transference of the saturation effects across all interacting loops. Not only is the loop experiencing saturation affected, but also the loops with which it is interacting. This results in a situation where these loops also lose their ability to track their setpoint. Moreover, in high gain systems, stability is jeopardised.

This dissertation presents a general solution for the practical problem that given the physical constraint of saturation, a multivariable system's control engineer would be able to prioritise the process outputs, and be given a choice over which outputs are maintained during periods of nonlinear operation.

A review of the work done in anti-windup bumpless transfer (AWBT), using Kothare *et al.'s* (1994) "*Unified framework for analysis of anti-windup bumpless transfer techniques*" as a reference, it will be shown how these techniques are employed to ensure system stability and linear performance recovery. Anti-windup (AW) Techniques that are designed specifically for multivariable systems will also be reviewed. The performance of these techniques, while the requested operating points are outside of the realisable operating region, will be of particular interest. AWBT does not in general meet the objective of allowing the engineer to prioritise the nonlinear mode's operation during these periods.

The novel concept of error redistribution (ER) is introduced to meet this objective. Error redistribution is a multivariable technique which allows the redistribution of error in the saturating loops to the non-saturating, or rather, the correcting loops, to maintain the setpoint of prioritised outputs; thus exploiting all degrees of freedom to reach the nonlinear mode objective. The formalisation of this concept, along with a stability proof and design guidelines are presented. The guidelines include being able to measure the suitability of using ER on a process. The nature of the technique allows the designer a controlled use, applying it only to the loops where it will be effective and with minimal disruption to the process.

Throughout this thesis, a simulated example of a multivariable thermal process (MVTP) is used to illustrate the principles and results relevant to dealing with saturation using AW and ER compensation. An application chapter includes the results of applying ER to a laboratory version of the MVTP; a simulated distillation column and a typical gold mine milling circuit. Comparisons are done for different pairings of ER compensation and AW techniques, including the implementation of the Hanus conditioning technique and artificial-nonlinearity (AN) described by Peng *et al.* (1998).

The result of this thesis is that error redistribution has been proven to be a viable technique for the optimisation of process outputs during nonlinear operation. The general structure of the system can easily implemented on industrial control platforms.

Acknowledgements

Thank-you to all the people who have helped me in finishing this thesis. My family and friends, in particular: my wife, Monique; the "Control Lab" postgraduates: Karl Prince, Ignatious Thithi, Jacek Naronzy; and, Mr Malcolm Attfield and Professor Martin Braae of the Electrical Engineering Department, University of Cape Town.

Thanks also to the Measurement and Control Division at Mintek, for sponsoring this thesis and in particular Dr Dave Hulbert for directing the work to areas of practical interest.

Table of Contents

List of Figures

List of Tables

Nomenclature

Process and General State-Space Form

$\mathbf{G}(s)$ is the $m \times n$ process being considered. In general, the systems will be dealt with in the state-space form, where the transfer function matrix $\mathbf{G}(s)$ corresponds to the minimal state-space realisation as:

$$\mathbf{G}(s) = \mathbf{C}_G(s\mathbf{I} - \mathbf{A}_G)^{-1}\mathbf{B}_G + \mathbf{D}_G = \left[\begin{array}{c|c} \mathbf{A}_G & \mathbf{B}_G \\ \hline \mathbf{C}_G & \mathbf{D}_G \end{array}\right]$$

The partitioned matrix will be used to represent the state-space form. Where necessary, subscripts will indicate to which system the matrix belongs.

In general, indices i and j will be used in reference with the process output and the error vector.

Controllers and Compensation Indicators

$\mathbf{K}(s)$ represents the $n \times m$ linear control strategy.

$\hat{\mathbf{K}}(s)$ the AWBT compensated and,

$\tilde{\mathbf{K}}(s)$ the ER compensated control strategy.

In general, indices k and l will be used in reference to the control action.

Saturation and Correcting Loop Indicators

Of particular interest to the discussion in this thesis are *saturating* and *correcting* output loops, indices for a particular case will include a superscript "s" or "c" to indicate which case is been referred to (e.g. i^s or i^c). In general, the index i^c will be used for correcting loops and j^s for the saturating loops.

A superscript * will indicate the actual plant input in the case of control action, $\underline{\mathbf{u}}$, and the actual or effective error redistribution matrix in the case of the ER matrix $\boldsymbol{\delta}$.

The saturating minimum and maximum for a particular control action will be represented by $\underline{u}_i$ and $\bar{u}_i$ respectively.

Glossary

The following is a collection of terms and abbreviations as used in the context of this document.

AW	Anti-windup compensation
AWBT	Anti-Windup Bumpless Transfer
CAW	Conventional Anti-windup
CLSR	Closed Loop Step Response
ER	Error Redistribution
δ	Error Redistribution Matrix
$\varepsilon_{i^c j^s}$	Error Redistribution Function from saturating loop j^s to correcting loop i^c
GCT	Generalised Conditioning Technique
HCT	Hanus Conditioning Technique
ITAE / O	Integral-Time-Absolute-Error / Operating Point
ISE	Integral-Squared-Error
MAW	Modified Anti-windup Technique
MIMO	Multi-Input Multi-Output
MVT(P)	Multivariable Thermal Process
ε_u , ε_e	Nonlinear sensitivity parameters
OPx	Operating Point set x – a particular set of setpoint changes for a case study
$\underline{\mathbf{P}}$	Prioritisation Vector
Settling time	Time in which the output reaches within +/-2% of its final value after a setpoint change or disturbance to the system. In the case if a multi-output system, it is taken as the maximum time, across all outputs, to reach this band.
SISO	Single-Input Single-Output
Steady State	When transients from setpoint changes or disturbances have decayed to zero.

1 Introduction

All real systems experience input saturation and in multi-input multi-output (MIMO) systems, all outputs are not of equal importance.

The physical world is usually limited by the rules of physics. This means that process inputs exhibit nonlinear effects that are not included in the classical modelling of the world. The scope of this work will be limited to the most common and usually most significant nonlinearity that all real systems that are regarded to be linear exhibit: input saturation. No matter what form the actuator may take, it will physically be limited at least by saturation and a maximum rate of change. In many cases, the rate of change limitation is swamped by the slower response times of the process itself. However, if the closed loop responses are to be in the same order as this limit, the response will be adversely affected.

Saturation, as will be demonstrated with examples in the text, has the ability to destroy the efforts of the control engineer in the most baffling of ways. In the single-input single-output (SISO) case, the effects of saturation are relatively easy to identify and once accepted as an issue, steps can be implemented to compensate for these effects. Anti-windup bumpless transfer (AWBT) techniques have received much attention over the last twenty years as tougher performance requirements have pushed systems into nonlinear regions.

For MIMO systems, the effect of saturation is often subtle. If there is a high degree of interaction between outputs and even one of many inputs to the system saturate, the entire system could be destabilized. The instability that can be introduced through saturation and the lack of understanding of this phenomenon has, in some industries, led engineers to shun classical multivariable techniques as theoretical tools not applicable in the real world.

Nonlinearities are not the problem; it is the failure of our analytical tools which do not account for the nonlinearities that cause inconsistency between our models and the real world. In analogue circuits (electronic or mechanical), saturation occurring in the controller is limited by the physics of the situation. This results in the higher inputs required by the process not being delivered and the systems response time is degraded. For a reasonably robust system, stability should not be threatened. The important aspect is that the controllers' states are restricted to values that reflect the realisable controller output.

In digital systems, where the value of a variable is only restricted by the numerical representation being used, a vast state-space is available to the controller, far exceeding that which produces realisable control actions. The internal values of the control states are not limited in classical linear control theory. A fundamental corner stone of classical control theory is that it applies to linear-time-invariant (LTI) systems and that the Theory of Superposition applies. In the analogue case, reality resolves some issues. However, the virtual universe that has been created in the digital computer has pushed back the boundaries of the nonlinearities that plague reality. This creates a universe in which the states of the controller are able to wander outside of the state-space region that produces realisable values at the actuator. This phenomenon is generally referred to as windup. Doyle *et al.*(1987) have promoted the concept that the anti-windup effort should not be limited to saturation, but rather to the greater problem of resolving controller output and plant input discrepancies caused by any nonlinear effect.

Having accepted that saturation will degrade performance and that AWBT techniques will salvage stability and some degree of performance, Industry has asked, *if performance is to sacrificed, is it not at least possible to choose which setpoints are maintained, because some are more important than others?*

Hence the objective of this work is:

1. to highlight the effects of saturation on multivariable systems;
2. to show how the appropriate use of AWBT can restore stability to the system; and
3. to introduce error redistribution (ER) as a novel means by which the control engineer can optimise performance during periods of nonlinear operation.

The techniques presented and developed in this thesis are discussed in terms of industrial process control applications. The results are not limited to this particular industry, but the examples and general discussions will focus on the process industry.

Throughout this work, a simulated example of a multivariable thermal process (MVTP) will be used to highlight the points being discussed. In Section 5.4 a laboratory MVTP system is used as an implemented example. Other application examples in Chapter 5 will demonstrate the use of ER compensation in a simulated distillation column and a milling circuit. The scope will be limited to the study of modern multivariable techniques, but not extending into model predictive control (MPC); although references to this technology will be made where appropriate.

1.1 Why Error Redistribution?

In the case where the requested operating point exists outside the physically realisable region, the control actions saturate. By implementing anti-windup (AW) compensation, system stability and some performance can be salvaged. AW techniques work to restrict the effects of saturation from propagating through the entire control system, thereby maintaining stability and, where possible, the non-saturating loops at their respective setpoints.

From a system performance point of view, it is quite possible that the outputs have a particular priority and that if a setpoint is to be lost, the designer would like to choose which setpoints are to be maintained at the cost of those of lower priority.

Error redistribution (ER) is a method by which a realisable operating point, closest to the minimum of some design performance index, is obtained while the requested operating point is outside of the realisable output space.

1.2 The Error Redistribution Principle

Error redistribution (ER) is a means to utilize the process interactions to minimise the performance index by exploiting the available (non-saturating) control loops when saturation is occurring in a particular loop. This is done by redistributing the error in control loops experiencing actuator saturation, the *saturating loops*, to non-saturating control loops, the *correcting loops*, that have high interaction terms with the *saturating loops*, thus using all degrees of freedom available to meet the performance requirements.

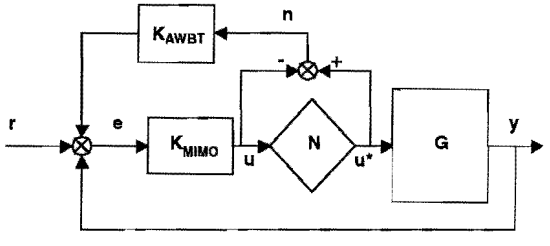


Fig.1 2x2 MIMO Control System, with AWBT compensation

A 2x2 system as in Fig.1 will be used to illustrate the principle of ER. The control loop being considered is a typical MIMO system including actuator saturation, ***N***, and AWBT

compensation, K_{AWBT}. ER compensation is not yet included. Fig.2 shows the available control space, u_1 *vs.* u_2, and in Fig.3 the corresponding output space y_1 *vs.* y_2., where the realisable steady state operating space, as a function of $\underline{u}^*(t)$, is indicated. The bounds of this region, which can be established by examining the mapping from the realisable control action space, $\underline{u}^*(t)$, to the output space, $\underline{y}(t)$, are assumed not to be known.

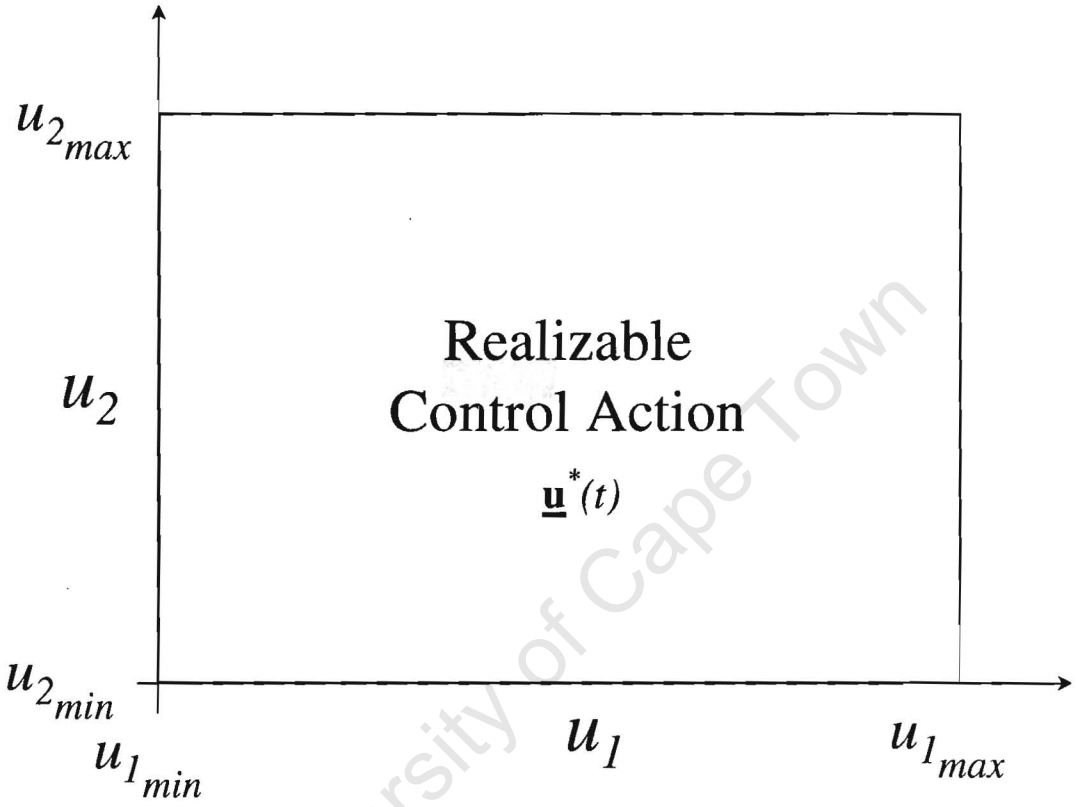


Fig.2 Realisable Control Action Space, $\underline{u}^*(t)$

At time t_0, the system is at operating point $\underline{y}(t_0) = \mathbf{A}$. At time t_1, the setpoints are changed to request the operating point $\mathbf{B}$, which exists outside the realisable region. Operating in a non-linear mode, AWBT techniques act to restrict the control action to the realisable region. Some AWBT techniques have the effect of when not being able to satisfy $y_1 = r_1(t_1)$, will act so as to maintain $y_2 = r_2(t_0) = r_2(t_1)$ and the system will come to steady state at operating point $\mathbf{B}'$ (see Section 3.3.1). It is obvious that should the priority of y_1 be greater than that of y_2, then by moving the operating point to $\mathbf{B}''$, y_1 would reach it's setpoint at the cost of y_2. At t_2 the demanded setpoints are again in the realisable region. The control system will then return to its linear mode of operation and operating point $\mathbf{C}$ will be obtained.

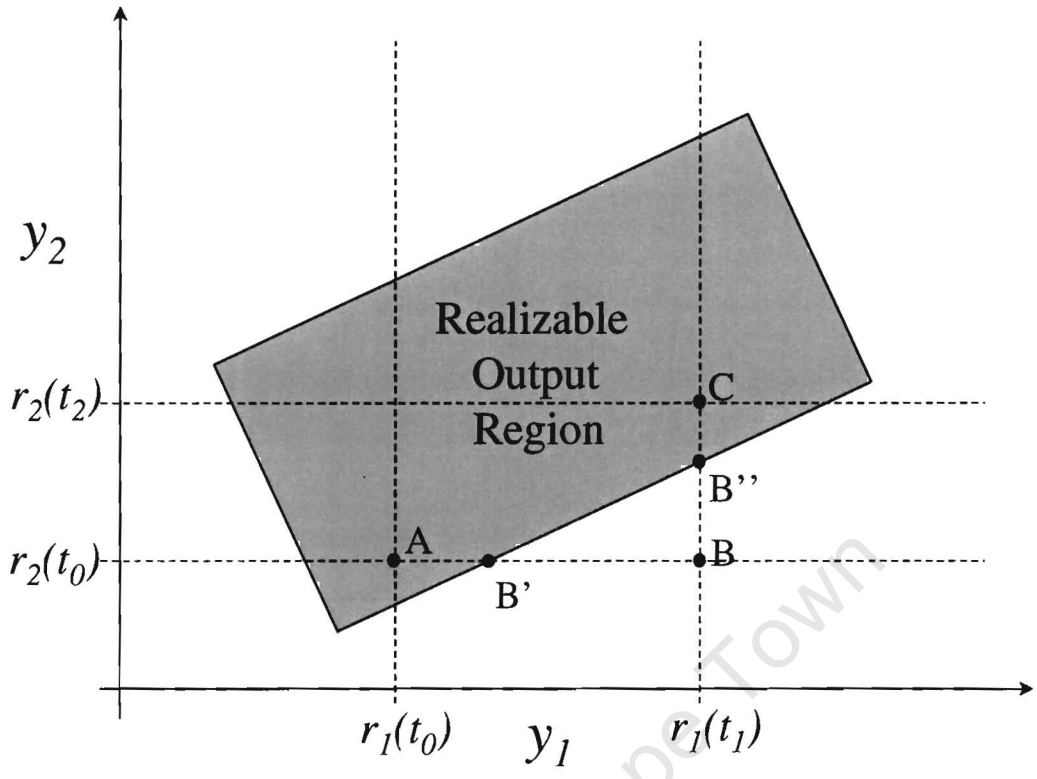


Fig.3 Realisable System Output Space

The action of AWBT is to restrict the controller's states to those that represent realisable control actions and thus the process is restricted to a realisable and stable output operating point. From the discussion above, exploiting the available degrees of freedom, the possibility exists to redistribute the error in y_1, the *saturating loop*, to y_2, the *correcting loop*, ensuring the performance of y_1 at the calculated expense of y_2. This is the principle of ER that will be formalised in the following sections.

The main work resulting in this thesis was undertaken during the period of 1995 through 1997 with the sponsorship of the Measurement and Control Division, Mintek. This projects progress was documented in the progress reports, Carew (1996-97), to Mintek.

2 Real World Multivariable Systems

To highlight the issues brought about by input saturation, and to have a frame of reference for the discussion of anti-windup bumpless transfer (AWBT) and error redistribution (ER) techniques, a simulation of a 2x2 example based on the multivariable thermal process (MVTP) will be considered. The MVTP is described in detail in Appendix E, and in Chapter 5.4, an implementation of AWBT and ER compensation will be presented.

A typical MVTP process model is found to be:

$$\mathbf{G}(s)=\begin{bmatrix} \dfrac{1.38}{7.5s+1} & \dfrac{1.97}{8.75s+1} \\ \dfrac{0.97}{6.25s+1} & \dfrac{2.05}{8.75s+1} \end{bmatrix} \tag{2-1}$$

This is a typical model obtained from the analysis of step test data. The data was fitted to a first order response for the effect of the step change on each input to each output. The units are in degrees Celsius per volt $\left({}^{0}C/_{V}\right)$ for all transfer functions.

A Nyquist analysis using a Nyquist array with Gershgorin bands for the system, $\mathbf{G}(s)$, would show that the system is diagonally dominant. (Maciejowski, 1989). A diagonalising controller can be designed using classical methods. Let $\mathbf{K}_d(s)$ be that controller:

$$\mathbf{K}_d(s)=\begin{bmatrix} 1 & -\dfrac{10.69s+1.426}{8.75s+1} \\ -\dfrac{4.156s+0.475}{6.25s+1} & 1 \end{bmatrix} \tag{2-2}$$

Designing a SISO PI controller for each diagonalised loop:

$$\mathbf{K}_{PI}(s)=\begin{bmatrix} \dfrac{3.00s+1}{s} & 0 \\ 0 & \dfrac{1.00s+0.25}{s} \end{bmatrix} \tag{2-3}$$

So that the final linear controller is given by:

$$\mathbf{K}(s)=\begin{bmatrix} \dfrac{3.00s+1}{s} & -\dfrac{10.69s^2+4.10s+0.3565}{8.75s^2+1.00s} \\ -\dfrac{12.47s^2+5.581s+0.475}{6.25s^2+1.00s} & \dfrac{1.00s+0.25}{s} \end{bmatrix} \quad (2\text{-}4)$$

2.1 Evaluation of performance

An objective means of measuring system performance is required for the comparison of techniques. Three performance indices will be used in the performance evaluation:

- Integral-Square Error (ISE)
- Integral of Time multiplied Absolute Error (ITAE)
- Integral of Time multiplied Absolute Operating Point Error (ITAO)

All indices are error based; therefore a smaller value indicates improved performance in the context of the index. Since the performance evaluation will be in relation to the ideal situation, where there are no nonlinear effects, all the indices will be normalised by the ideal response value for that index under similar conditions.

The equation defining the index and brief summation of the properties of performance that each will expose follows. It is important to realise that no one index summarises all aspects of the system's performance, so these indices should be referred to with engineering insight into the objective being considered. In particular, most indices will be considered on a per step per loop basis so as to highlight the relative performance between loops.

2.1.1 Integral-Square Error (ISE)

The ISE index is calculated for each output loop, i, as:

$$ISE_i = \int_{t_0(op)}^{t_1(op)} e_i^2(t)dt \quad (2\text{-}5)$$

Where $t_0(op) \rightarrow t_1(op)$ is the time interval of each step that is being considered. The sum of ISE_i over all loops gives the overall performance.

This is a common performance index used to reflect on the rise times of process.

2.1.2 *Integral of Time multiplied Absolute Error (ITAE)*

The ITAE index is calculated for each loop, i, and for each change in operating point:

$$ITAE_{i,op} = \int_{t_0(op)}^{t_1(op)} t\left|e_i(t)\right|dt \tag{2-6}$$

This index weights later errors, such as oscillation, lingering transients, and steady state errors, more harshly. A lower value indicates a more rapid return to steady-state setpoint tracking.

2.1.3 *Integral of Time multiplied Absolute Operating Point Error (ITAO)*

The ITAO index is calculated for each change in operating point:

$$ITAO_{op} = \int_{t_0(op)}^{t_1(op)} t\left\|\underline{\mathbf{e}}(t)\right\|dt \tag{2-7}$$

This is the integral of time multiplied error norm, the distance of the output trajectory from operating point. This is a generalisation of the ITAE index to the more usual multivariable case.

2.2 *Ideal system responses*

Having designed the compensator and controller, simulations are done to verify system performance, so generating the ideal response indices. A two-step change in the requested operating point will be considered for this case study:

$$\underline{\mathbf{r}}(t) = \begin{cases} r_1 = 0 & t < 10 \\ r_1 = 5 & t \geq 10 \\ r_2 = 1 & t < 160 \\ r_2 = 4 & t \geq 160 \end{cases} \tag{2-8}$$

For performance analysis, the time response (including control action), the control trajectory and output trajectory, will be given for each case.

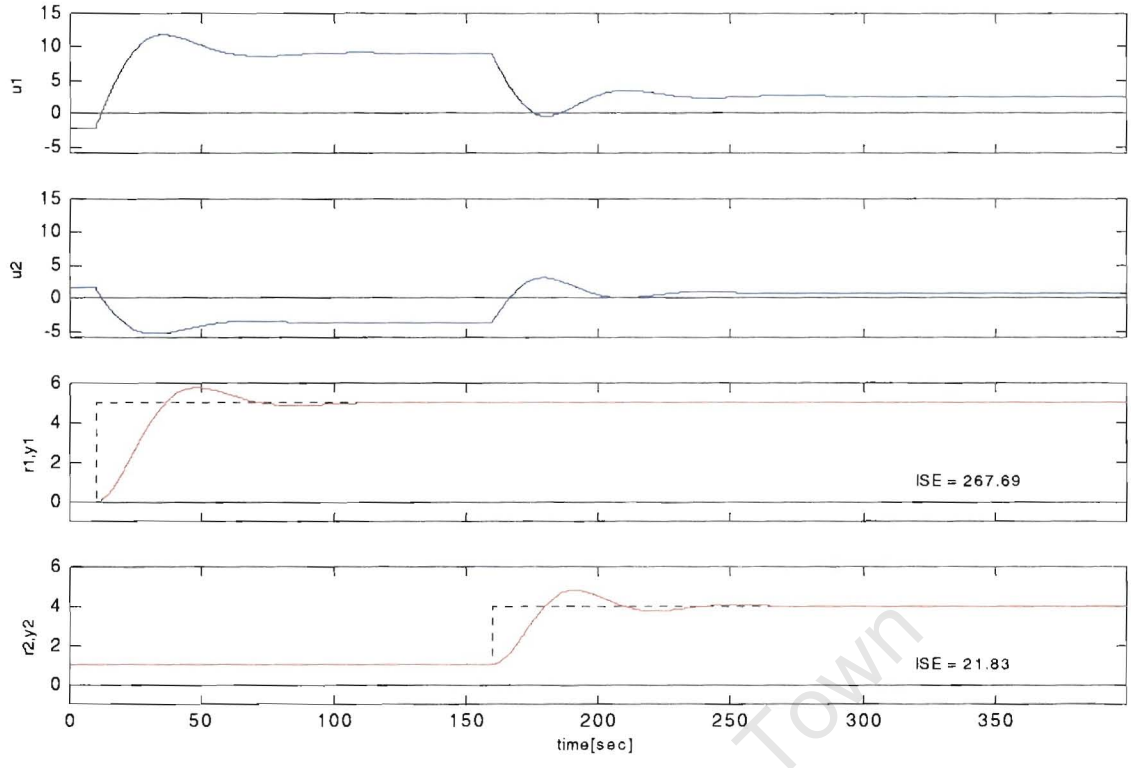


Fig.4 Ideal MVT Process Closed Loop Step Response (CLSR)

Since enough control action is available, the controller is able to compensate for the interaction terms within the process when a step change occurs. This results in only a small disturbance in the output not being stepped and the system moves quickly to its new steady state.

The settling time for each loop, being the maximum time across all loops, for the system to reach within +/-2% of its steady-state value after a setpoint change, are tabulated, Table 1, from the time response curves in Fig.4:

Table 1 Ideal MVT Settling Times

Step time	y_1	y_2
$r_1 = 10$	59	-
$r_2 = 160$	-	75

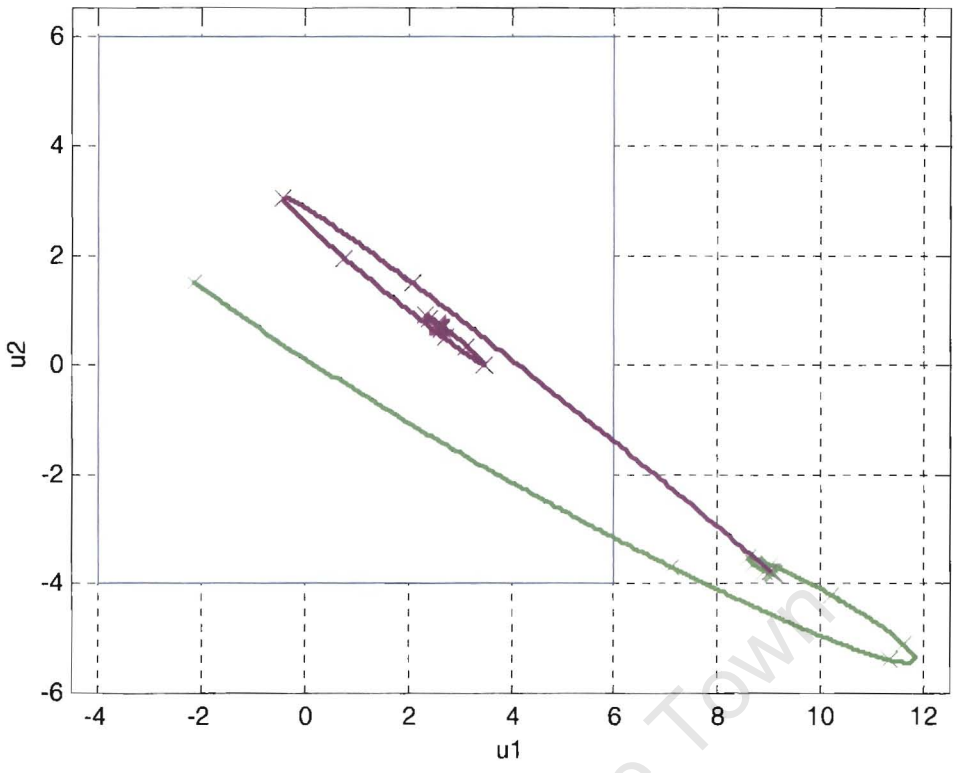


Fig.5 Ideal MVT Process Control Trajectory

In the ideal case, $\underline{\mathbf{u}}$ is not bounded and an infinite output space is available. The control trajectory is shown in Fig.5. The bounded region represents the realisable control region that will be introduced by saturation. The green and purple curves correspond to the trajectory for the first and second steps respectively. The X-markers on the trajectories are spaced at 10 seconds intervals so as to indicate the speed at which the control actions changed. The trace properties are consistent across all the following trajectory plots.

Examining the trajectory for the first step, the system moves quickly (< 10 seconds) out of the region of available control action and, within another 50 seconds, has reached within +/-2% of steady-state. With the second step, the system moves again within 10 seconds into the realisable control region and settles to within +/-2% of steady-state within 70 seconds.

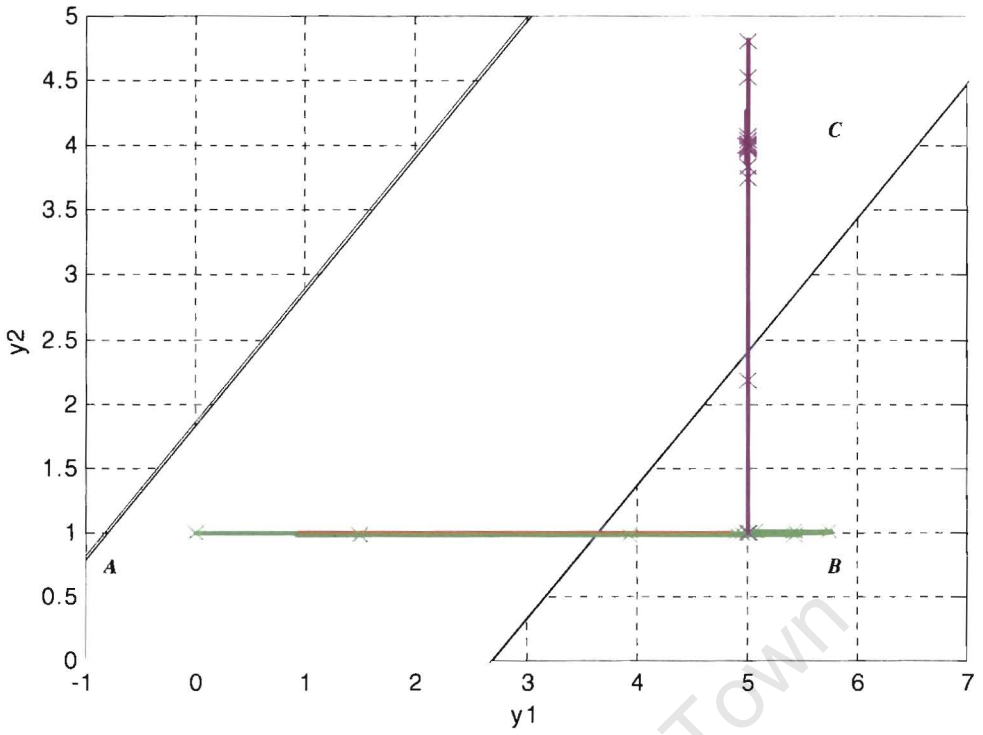


Fig.6 Ideal MVT Process Output Trajectory

The output trajectory plot, Fig.6, shows the path the system takes to obtain the operating point. Again, the bounded region relates to the realisable output region that will be imposed through the inclusion of saturation in the next section. The red trajectory indicates the requested operating points; in this case it is mostly covered by the actual trajectory.

2.2.1 Performance Indices

As can be seen, in Table 2, the outputs perform satisfactorily, and the following values are obtained for the various performance indices:

Table 2 Ideal MVT Process Performance Indices

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	268	0.0	1550	37	1551
2	0	87.3	13	1233	1233
Loop Totals	268	87.3	1563	1270	
Overall	355		2833		2784

Notice the low or zero values of the indices where the step takes place in the other loop. This indicates the successful decoupling of the MIMO process. In the following analysis of the responses, these indices will form a baseline for comparisons and all indices will be normalised according to them.

2.3 Saturating system responses

Saturation will be defined as the vector function:

$$\underline{\mathbf{u}}^* = \mathbf{N}(\underline{\mathbf{u}}) = \mathbf{sat}(\underline{\mathbf{u}}) = \begin{cases} \overline{u}_i & u_i > \overline{u}_i \\ u_i & \text{else} \\ \underline{u}_i & u_i < \underline{u}_i \end{cases} \tag{2-9}$$

For the MVTP being considered, the range of the control action for both loops is:

$$\underline{\mathbf{u}}^* = \mathbf{N}(\underline{\mathbf{u}}) = \mathbf{sat}(\underline{\mathbf{u}}) = \begin{cases} +6 & u_i > +6 \\ u_i & \text{else} \\ -4 & u_i < -4 \end{cases} \tag{2-10}$$

Having restricted the available control action space, the realisable output region has been restricted to approximately the following region in Fig.7.

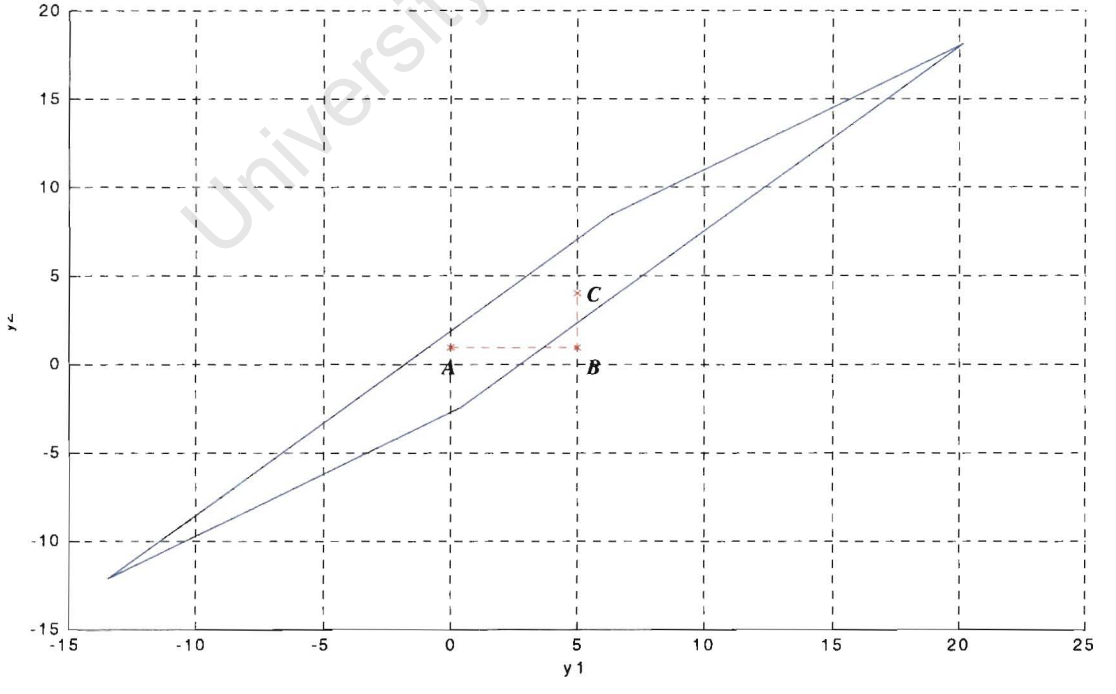


Fig.7 MVT Process Realisable Region, with requested Operating Points

The requested operating points are also included on Fig.7, where it is seen that operating point after the first step change, B, is outside of the realisable region. It is important to note that the setpoints of $y_1 = 5$ and $y_2 = 1$ are valid. If the system was comprised of two single loops, each would be obtainable at any point. However, the off-diagonal terms in the process model exclude these from being obtained concurrently. The realisable region is obtained by considering all combinations of the input extremes. The realisable control action as defined by Equ.2-10, is:

$$u_1^* = \begin{bmatrix} -4 \\ -4 \end{bmatrix}, u_2^* = \begin{bmatrix} 6 \\ -4 \end{bmatrix}, u_3^* = \begin{bmatrix} 6 \\ 6 \end{bmatrix} \text{ and } u_4^* = \begin{bmatrix} -4 \\ 6 \end{bmatrix} \tag{2-11}$$

The rhomboid given here is described by multiplying these input values by the steady state gains of the process matrix, from Equ.2-1,

$$\mathbf{G}_\infty = \begin{bmatrix} 1.38 & 1.97 \\ 0.97 & 2.05 \end{bmatrix} \tag{2-12}$$

Giving:

$$\mathbf{G}_\infty u_1^* = \begin{bmatrix} -13.4 \\ -12.1 \end{bmatrix}, \mathbf{G}_\infty u_2^* = \begin{bmatrix} 0.40 \\ -2.38 \end{bmatrix}, \mathbf{G}_\infty u_3^* = \begin{bmatrix} 20.1 \\ 18.1 \end{bmatrix} \text{ and } \mathbf{G}_\infty u_4^* = \begin{bmatrix} 6.30 \\ 8.42 \end{bmatrix} \tag{2-13}$$

This realisable region is the steady-state realisable region. Due to transient states within higher order systems, the excitation of those states could result in the system's transients reaching operating points outside this steady-state realisable region. Hence, higher order process models would have a transient region that would encompass the steady-state region. In general, references to the "*realisable output region*" are to the "*steady-state realisable region*" as described here, unless otherwise stated.

It is apparent that the realisable output region can only be approximated for the physical system, as the model for the system is surely only valid within a certain range of the operating point used to create the model and fringe effects will limit the accuracy. This will be a key concern in implementing a stabilising or optimising strategy.

The following responses are generated by the saturating system:

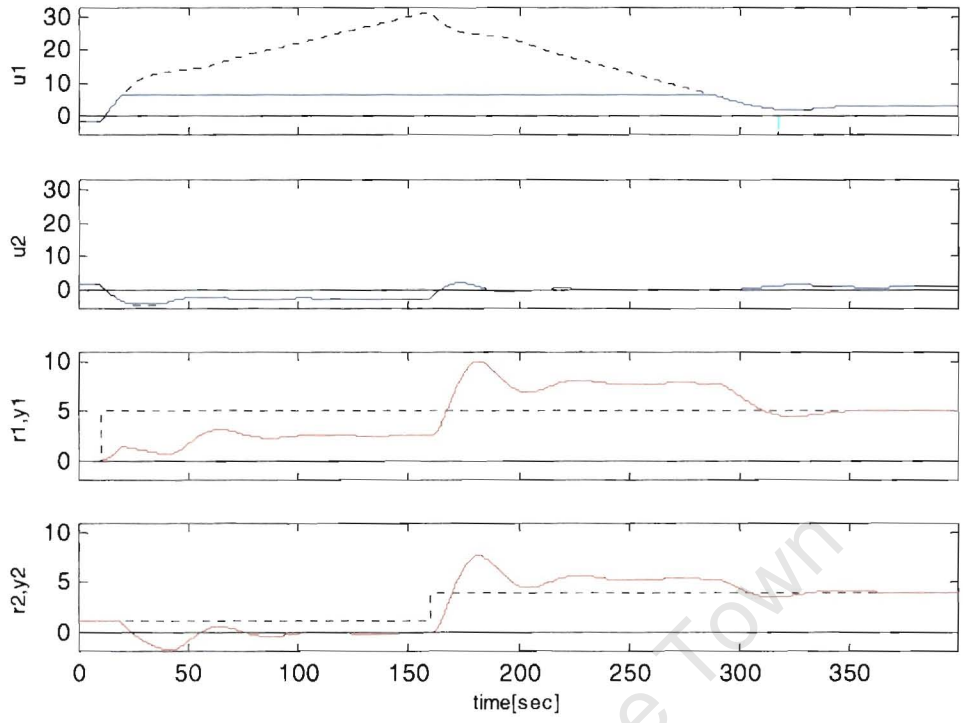


Fig.8 Saturating MVT Process CLSR

The time responses in Fig.8 show that the controller output winds up considerably during the period of saturation. The settling time is not particularly relevant. The system only returns to linear operation after 350 seconds. The table below indicates the "settling times" taken to reach the nonlinear mode final value.

Table 3 Saturating MVT Process Settling Times

Step time	y_1	y_2
$r_1 = 10$	115	-
$r_2 = 160$	-	173

Three effects of note are:

1. The controller in loop 1 never reaches steady state. Its' states continue changing even though the process inputs have saturated. This is the cause of continuing the integration of the error and results in the phenomenon of "integrator windup".

2. The system takes considerable time to re-enter the linear region. Because of integrator windup, the error in the loop must first go negative to wind down the internal controller states before the control action and controller outputs are again equal.
3. The second output loses its setpoint although there is still available control action in this loop; the controller's off-diagonal terms are contributing to the second control signal as though the first loop was actually obtaining the calculated, wound-up, value.

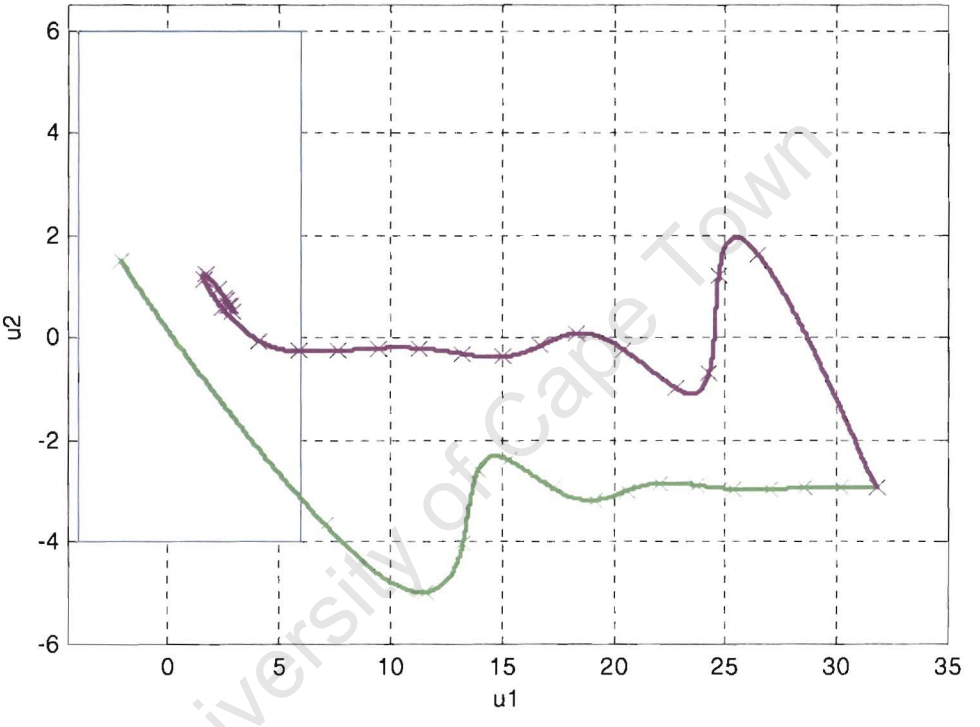


Fig.9 Saturating MVT Process Requested Control Trajectory

From the time responses in Fig.8 and the control trajectories in Fig.9, it is evident that although the error derivate has approached zero and the system is in a steady state, the calculated control in u_1 is continuing to increase. This indicates that the integrator in the controller is continuing to integrate the error, but due to the saturation, there is no effect on the process outputs.

From the output trajectory in Fig.10, it is evident that the plants "steady-state" operating point, SAT, after the first step change is not satisfactory. The system is operating in an unstable mode and the controller's states are moving along a divergent path. Without

external intervention, the states would continue to windup and the setpoints would not ever be reached.

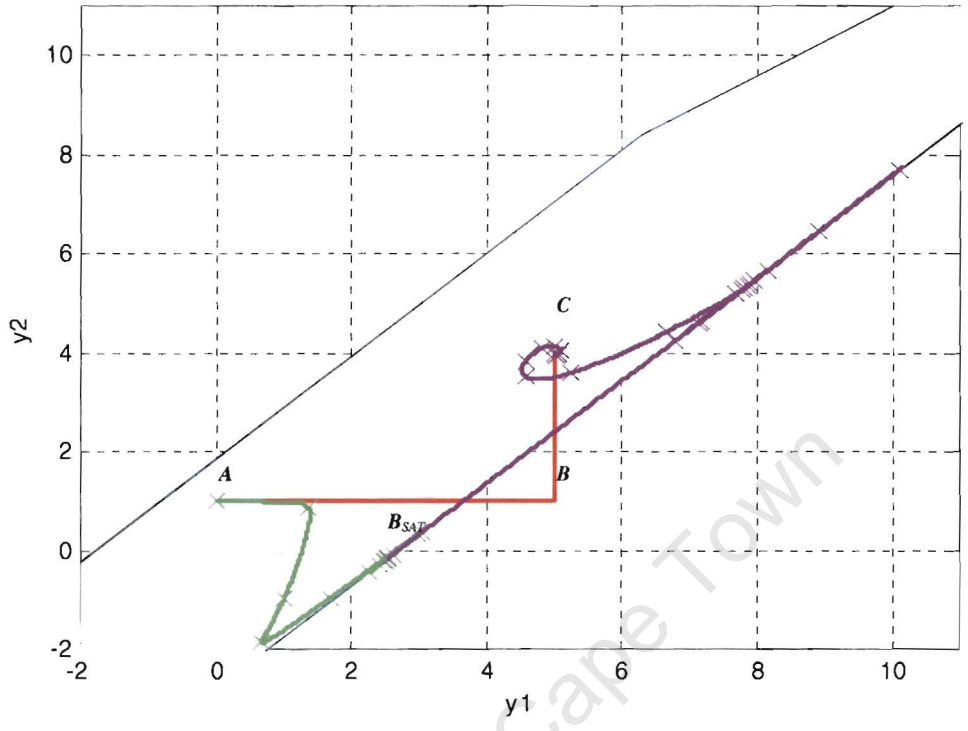


Fig.10 Saturating MVT Process Output Trajectory

The second step, returning to the realisable region, is such an intervention. After a considerable time period, the controller state re-enters the region of realisable control and the system returns to linear performance. During the return to linear operation, the process only sees the changes in u_2. Again, the controller is trying to compensate for the non-existent effect of u_1 and the transition to linear performance is less than ideal. These effects indicate that even if the system returns to linear region, the wound-up states can still have a negative impact on the system's performance.

2.3.1 Performance Indices

The indices will be normalised to the ideal case, which would be represented by the indices, across the steps, producing a net value of 1.00. For comparison the normalised ideal table is given in Table 4. The performance index values, Table 5, for the saturating case show a poor response.

Comparing Table 4 and Table 5, the saturating system performs approximately 9 times worse than the ideal system in the ISE sense and 20-40 times worse in the ITAE and ITAO sense. The ITAE and ITAO indices are noticeably worse because these indices penalise the steady state error. As noted earlier, even though the second loop's control action has not saturated, this loop has also experienced performance degradation.

Table 4 Normalised Ideal MVT Process Performance Indices

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.00	0.00	0.992	0.029	0.557
2	0.00	1.00	0.008	0.971	0.443
Loop Totals	1.00	1.00	1.00	1.00	
Overall	2.0		2.0		1.00

Table 5 Saturating MVT Process Performance Indices

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	5.01	3.37	18.5	10.9	11.6
2	4.66	5.11	19.7	12.2	12.6
Loop Totals	9.67	8.48	38.2	23.0	
Overall	18.15		61.2		24.2

In the context of output optimisation, where the performance of y_1 may be more significant than that of y_2, this system has obviously failed.

2.3.2 Analytical Discussion

To summarise these effects analytically for this controller design, a detailed analysis of the interactions within the controller, the process and their interactions is required:

- The changes in control action are caused by changes in the error inputs into the controller.

- The nature of the controller is that u_1 includes the term $k_{12}e_2$ that compensates for the effect that changes in u_2 has on y_1 because of the term $g_{12}u_2$.

- Similarly, u_2 includes the term $k_{21}e_1$ that compensates for the effect that changes in u_1 has on y_2 because of the term $g_{21}u_1$.

In effect, applying an open loop compensator that reduces the control action, u_i by the amount that the increase in the other control action u_j will affect the output y_i through process interaction term g_{ij}.

This effect is visible in the ideal case. Considering Fig.4, the Ideal MVT Process' time response's, during the step in r_1, u_1 changes to produce the desired effect of $y_1 = r_1$. Because of this step, u_2 is compensating for the impact this change in u_1 would have on y_2, because of the term $g_{21}u_1$. This compensation is successful and no disturbances are experienced by y_2. Examination of the Ideal MVT Process Control Trajectory, Fig.5, shows that until steady state was reached, u_1 and u_2 acted together to move y_1 to r_1 and to prevent disturbances affecting y_2.

In the saturating case, examination of the requested Control Trajectory in Fig.9 shows that after approximately t=100 seconds, u_2 approaches its final value until the next step occurs at t=160seconds. It is also noticed that u_1 appears to be increasing at a constant rate, again, until the next step in the operating point occurs. Returning to the Saturating MVT Process time responses in Fig.8, it is apparent that y_1 does not reach the new setpoint but does settle to a final value; and that y_2 has been disturbed by the step even through u_2 has available control action. Lastly, after the second step, the system takes a considerable amount of time to re-enter its linear mode of operation.

Obviously, these phenomena are a result of the saturation being included. By considering the component terms within the system, an explanation for these are given:

- Each process input is limited to the range: $\underline{u}_i < u_i^* < \bar{u}_i$

- Should the error, e_1, drive the integral action in k_{11} so large that: $k_{11}e_1 > \bar{u}_1 \pm k_{12}e_2$ or $k_{11}e_1 < \underline{u}_1 \pm k_{12}e_2$ then $u_1^* = \underline{u}_1$ or $u_1^* = \bar{u}_1$ respectively.
- In the saturating mode, the effect of u_1 on:
    - the output y_1 approaches the limit $g_{\infty 11}\bar{u}_1$,
    - and, the output y_2 approaches the limit $g_{\infty 21}\bar{u}_1$.
- Hence, the final output values are:
    - $y_{\infty 1} = g_{\infty 11}\bar{u}_1 + \lim_{s \to 0} g_{12}u_2 s$
    - $y_{\infty 2} = g_{\infty 21}\bar{u}_1 + \lim_{s \to 0} g_{22}u_2 s$

    From Fig.9, we see the final value for $u_2 \to 3.00$, giving the "steady-state" output values, $\underline{y} \approx \begin{bmatrix} 2.37 & -0.33 \end{bmatrix}^T$ as expected from Fig.10.

This discussion shows that given the saturating state of u_1 and the final value of u_2, the final output is as expected. This does not explain why u_2 has this particular steady-state value.

- u_2 should be compensating for the effect of u_1 on y_2 through $k_{21}e_1$
- Since k_{21} contains an integrator, after the transients have decayed, this term increases linearly. Note this term depends on e_1 and not u_1 which continues to windup.
- u_2 is also dependant on the term $k_{22}e_2$, k_{22} also contains an integrator and any error is integrated linearly after the transients have decayed.

Since $e_2 > 0$, it is expected that the integrator in k_{22} would continue to integrate until this error is reduced. However, the point at which the rate of integration of e_1 is equivalent, but opposite to the rate of integration of e_2, the final value for u_2 is an equilibrium point and holds this value. If the respective rates vary sufficiently, u_2 would also have been driven to saturation.

Thus, both setpoints are lost, even though there is control action available in one of the loops.

2.4 *Effects of Saturation*

As illustrated in the example, saturation of the process input has severe degradation effects on the performance of a system. These typically include:

- Loss of setpoints – due to compensating for non-existing control action, the controller pushes the process off the desired operating point. This is true even in the loops which are not experiencing saturation.
- Reduction in disturbance rejection ability – due to the limited control action available to the controller, large disturbances, while the control action is near saturation, will be poorly dealt with.
- Instability – in high gain controllers, the effects of windup can result in limit cycling and in the extreme, the frequency will be limited only by the maximum rate of change of the physical system.

In all cases, the issues arising due to saturation are caused by the discrepancy between controller states and the actual controller input to the process. In the chapters that follow, anti-windup techniques and error redistribution will be discussed. From their design, it is apparent that these techniques will be useful in dealing with the wider problem of nonlinearieties that result in state discrepancies within controllers.

3 Anti-windup - Dealing with Saturation

In chapter 2 the MVTP example was introduced and the affects of saturation discussed. Since the 1960's control engineering researchers have been studying methods to compensate for these affects (Fertik and Ross, 1967). However, as most reviewed literature indicates a sound theoretical basis for dealing with saturation and the windup it causes within controllers has been missing. Several attempts have been made to consolidate the mostly *ad hoc* techniques that make up the Anti-windup Bumpless Transfer (AWBT) literature. One of the most recent such attempts was made by Kothare *et al.* (1994) and will be used as the foundation for error redistribution (ER) stability. In the following sections a discussion of popular techniques and an overview of Kothare's Unified Framework for AWBT given. Anti-windup (AW) compensation will be applied to the MVTP system followed by a discussion of the benefits and shortcomings of AWBT techniques.

Initially the anti-windup problems was thought of in the context of integrators in the controller winding up, i.e. having the value of the integral component of the controller become so large that the composite output exceeded the output value that was obtainable by the input actuator. This effect was immediately apparent when implementing digital PI controllers. Since the 1960's, researchers like Fertik and Ross (1967) and Kramer and Jenkins (1971) have been working on the issues around digital controller windup.

It's interesting to note that the windup affect would not necessarily have had a severe impact on the performance of a controller implemented in a continuous (analog electronics or mechanical) system. The nature of the controller itself was such that on reaching saturation the integrators would be forced to stop integrating and hold their value at the saturation limit. With the event of digital computers where the variable value representing the output is limited only by the bit-resolution of the number system, the integrators would continue to integrate the error term until it overflows the register.

During the 1970's and 1980's the availability of digital computers and their ability to realise complex mathematical algorithms has led to their exploitation in all areas of automation. With this the "nuisance" of windup, Hanus (1980), increased and many researchers have looked for anti-windup techniques appropriate for their particular favourite control strategy. In one of the tome's of control theory, Åström and Wittenmark (1984) presents the anti-reset windup techniques for digital PI controllers and shows how state-space observer techniques can be used to limit controller states. This method broadened the scope of

anti-windup techniques and placed the focus on the states of controller, rather than just the integration term.

Hanus *et al.* (1987) presents the "conditioning technique" as a generalised anti-windup and bumpless transfer method. This was an extension of his previous work (Hanus, 1980) taking the concept of realisable reference and realisable control action to the MIMO case. This technique is the generalisation of the back-calculation method proposed by others, such as Fertik and Ross (1967).

The anti-reset techniques was generalised by Doyle *et al.* (1987a,b) in the formulation of conventional anti-windup (CAW) and nonlinear modified anti-windup (MAW) for the MIMO case. In particular, Doyle and Packard (1987) makes the point that in MIMO systems the gain of the system is dependant on the direction of the control vector $\underline{\mathbf{u}}$. Doyle *et al.*(1987), again, point to the windup phenomenon as being a special case of problem where the states of the controller are not correlating to the actual process inputs.

Many have discussed the possibilities of exploiting internal model control (IMC) for anti-windup compensation (Doyle *et al.*(1987), Campo *et al.*(1989), Campo and Morari (1990) and Zheng *et al.*(1994) amongst others), but the consensus is that these systems perform sluggishly and other forms of AW compensation are, in general, better. As such, this work will not focus on the IMC based techniques.

Kothare *et al.*(1994) also discuss the extended Kalman filter as a observer-based AWBT technique. Hanus's work was built on by Walgama *et al.* (1992) to deal with the "short sightedness problem". While other "independent" algorithms continue to be presented, for example Larsson (1994). But most of these are been shown to be a special case of an existing class.

Hence, the most general interpretation of AWBT techniques has come to encompass all methods whereby the states of the controller are restricted so as to be consistent between the controller's calculated outputs and the actual process inputs.

The foundations for the unified framework presented by Kothare *et al.*(1994) were presented by Campo and Morari(1990). Kothare *et al.*(1994) shows how a wide range of AWBT compensation techniques can be realised in the framework.

The primary objectives of AWBT techniques, as summarised by Kothare *et al.*(1994), are quoted as:

- Ensure stability of the closed loop system during nonlinear operation.
- Linear performance recovery when the system returns to the linear mode.
- And graceful transition to and from the nonlinear modes.

Since publication, the framework has being able to assist in the synthesis of new AWBT techniques, Saeki *et al.*(1996), and has formed a platform for discussion and comparison of methods, Edwards and Postlethwaite (1998). It must be noted that this framework is just another step towards understanding and formalising AW techniques, but there are techniques just as MAW which do not fit the framework in all senses and others have suggested improvements to or alternative paramertisations of the framework, Peng *et al.*(1998). This has allowed the field to move onto issues beyond the basic need of stabilising saturating control loops to performance related issues. In the second part of this chapter newer techniques, mostly developed over and after the period when ER compensation was been developed (Carew 1996-1997) which deal with performance in MIMO systems, Peng *et al.*(1998), Kapoor *et al.*(1998).

Attention has been given to the similarity between AWBT and model predictive control (MPC) and the possibilities of producing the AWBT functionality with MPC (Cherukuri, 1998).

3.1 Presentation of Unified Framework for AWBT

Since this framework is central to the development of the error redistribution (ER) techniques presented in Chapter 4, a presentation of Kothare *et al.*'s (1994) framework is presented here.

The linear controller, $\mathbf{K}(s)$, and the open loop process model, $\mathbf{G}(s)$, are represented by the state space forms:

$$\mathbf{K} = \left[\begin{array}{c|c} \mathbf{A_K} & \mathbf{B_K} \\ \hline \mathbf{C_K} & \mathbf{D_K} \end{array}\right] \text{ and } \mathbf{G} = \left[\begin{array}{c|c} \mathbf{A_G} & \mathbf{B_G} \\ \hline \mathbf{C_G} & \mathbf{D_G} \end{array}\right], \tag{3-1}$$

Fig.11 represents Kothare's Framework, where the system inputs are $\underline{\mathbf{w}} = [\underline{\mathbf{r}} \quad \underline{\mathbf{d}}]^T$, the setpoints, $\underline{\mathbf{r}}$ and disturbances, $\underline{\mathbf{d}}$, and the actual control action/inputs to the process $\underline{\mathbf{u}}^*$.

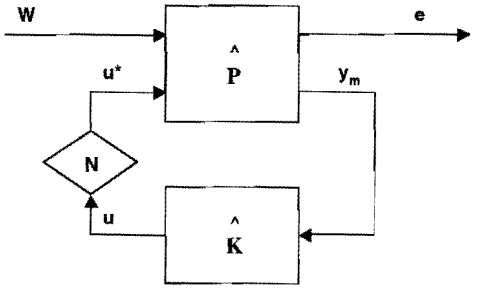


Fig.11 Unified Framework for AWBT

The interconnection matrix is given as:

$$\hat{\mathbf{P}} = \left[\begin{array}{cc|c} \mathbf{I} & -\mathbf{I} & -\mathbf{G} \\ \mathbf{I} & -\mathbf{I} & -\mathbf{G} \\ \hline \mathbf{0} & \mathbf{I} & \mathbf{M} \end{array}\right] \tag{3-2}$$

The process outputs are the errors, $\underline{\mathbf{e}}$, and the feedback term $\underline{\mathbf{y}}_m = [\underline{\mathbf{e}} \quad \underline{\mathbf{u}}_m]^T$, including the error and the approximated or measured value of $\underline{\mathbf{u}}^*$, called $\underline{\mathbf{u}}_m$. It is not always possible to know the exact value of $\underline{\mathbf{u}}^*$, modern intelligent instrumentation provided some assistance in reporting the realised value of the requested control action. For a generic formulation, Kothare *et al.* (1994) use:

$$\underline{\mathbf{u}}_m = \mathbf{I}\underline{\mathbf{d}} + \mathbf{M}\underline{\mathbf{u}}^* \tag{3-3}$$

Where $\mathbf{M}$ is ideally $\mathbf{I}$, but may include measurement noise, filters and other related factors. In practice, where the limitations of $\underline{\mathbf{u}}^*$ are well defined, or are purposefully used to contain the control action, a nonlinear function is employed to calculate $\underline{\mathbf{u}}_m$. For simplicity, the assumption of $\underline{\mathbf{u}}_m \approx \underline{\mathbf{u}}^*$ will be made.

$\underline{\mathbf{y}}_m$ is fed back to the AWBT compensated controller, $\hat{\mathbf{K}}(s)$. Kothare's framework is derived from defining $\hat{\mathbf{K}}$ as the following state-space system:

$$\begin{aligned}\dot{\underline{\mathbf{x}}} &= \mathbf{A}_{\hat{\mathbf{K}}}\underline{\mathbf{x}}_{\hat{\mathbf{K}}} + \mathbf{B}_{\hat{\mathbf{K}}}\underline{\mathbf{y}}_m + \underline{\boldsymbol{\varepsilon}}_S \\ \underline{\mathbf{u}} &= \mathbf{C}_{\hat{\mathbf{K}}}\underline{\mathbf{x}}_{\hat{\mathbf{K}}} + \mathbf{D}_{\hat{\mathbf{K}}}\underline{\mathbf{y}}_m + \underline{\boldsymbol{\varepsilon}}_O\end{aligned} \tag{3-4}$$

Where, the AWBT correction term, $\underline{\boldsymbol{\varepsilon}}$, as a function of the AWBT operator, Λ, and the difference between the actual process input, $\underline{\mathbf{u}}^*$, and the control output, $\underline{\mathbf{u}}$, is:

$$\underline{\boldsymbol{\varepsilon}} = \begin{bmatrix}\underline{\boldsymbol{\varepsilon}}_S \\ \underline{\boldsymbol{\varepsilon}}_O\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Lambda}_S \\ \boldsymbol{\Lambda}_O\end{bmatrix}\left(\underline{\mathbf{u}}^* - \underline{\mathbf{u}}\right) \tag{3-5}$$

The AWBT correction term is broken into state correction, $\underline{\boldsymbol{\varepsilon}}_S$, and output correction, $\underline{\boldsymbol{\varepsilon}}_O$, components, thus allowing full access to the states and outputs of the controller.

Where the AWBT operator, $\boldsymbol{\Lambda} = \begin{bmatrix}\Lambda_S & \Lambda_O\end{bmatrix}^T$, for linear performance recovery, should be memoryless and therefore a constant matrix. As Kothare *et al.* (1994) note, that although restrictive, most AWBT systems fit this model and so it seems reasonable.

The formulation of the AWBT compensated controller comes from considering equations (3-4 and 3-5). The decomposition of the $\hat{\mathbf{K}}(s)$ supplies the following model:

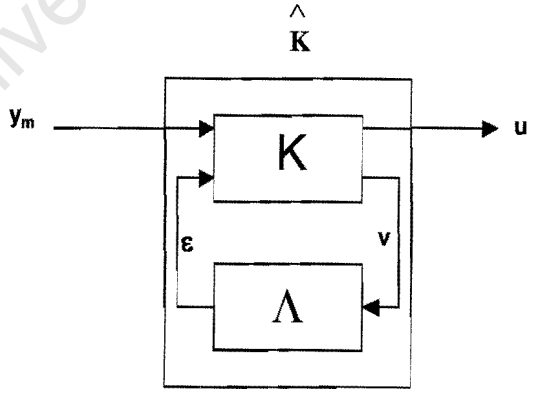


Fig.12 Decomposition of $\hat{\mathbf{K}}(s)$

Where the decomposed terms are:

$$K = \left[\begin{array}{c|cccc} \mathbf{A_K} & \mathbf{B_K} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \hline \mathbf{C_K} & \mathbf{D_K} & \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{array}\right] \text{ and } \underline{\mathbf{v}} = \begin{bmatrix} \underline{\mathbf{x}} \\ \underline{\mathbf{e}} \\ \underline{\mathbf{u}}_m \\ \underline{\boldsymbol{\varepsilon}}_S \\ \underline{\boldsymbol{\varepsilon}}_O \end{bmatrix} \tag{3-6}$$

From the definition of the AWBT correction term given in equation (3-5), the AWBT operator will have the following structure:

$$\underline{\boldsymbol{\varepsilon}} = \begin{bmatrix} \underline{\boldsymbol{\varepsilon}}_S \\ \underline{\boldsymbol{\varepsilon}}_O \end{bmatrix} = \Lambda \underline{\mathbf{v}} = \begin{bmatrix} \mathbf{\Lambda}_1 \\ \mathbf{\Lambda}_2 \end{bmatrix} \begin{bmatrix} -\mathbf{C_K} & -\mathbf{D_K} & \mathbf{I} & \mathbf{0} & -\mathbf{I} \end{bmatrix} \underline{\mathbf{v}} \tag{3-7}$$

As mentioned earlier, the AWBT compensated controller is derived from combining the AWBT operator and the linear controller. Giving the following as the AWBT compensated controller:

$$\hat{\mathbf{K}} = \begin{bmatrix} \mathbf{U} & \mathbf{I} - \mathbf{V} \end{bmatrix} \tag{3-8}$$

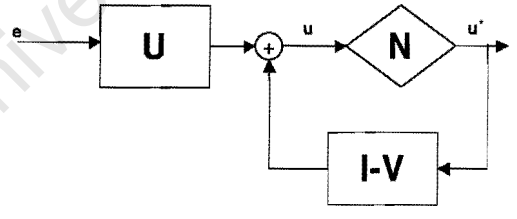


Fig.13 Controller implementation

where:

$$\mathbf{U} = \left[\begin{array}{c|c} \mathbf{A_U} & \mathbf{B_U} \\ \hline \mathbf{C_U} & \mathbf{D_U} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{A_K} - \mathbf{H_1 C_K} & \mathbf{B_K} - \mathbf{H_1 D_K} \\ \hline \mathbf{H_2 C_K} & \mathbf{H_2 D_K} \end{array}\right] \tag{3-9}$$

$$\mathbf{V} = \left[\begin{array}{c|c} \mathbf{A_V} & \mathbf{B_V} \\ \hline \mathbf{C_V} & \mathbf{D_V} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{A_K} - \mathbf{H_1 C_K} & -\mathbf{H_1} \\ \hline \mathbf{H_2 C_K} & \mathbf{H_2} \end{array}\right] \tag{3-10}$$

Kothare *et al.* (1994) show that most existing AWBT schemes fit this model, where:

$$\mathbf{H}_1 = \mathbf{\Lambda}_1(\mathbf{I}+\mathbf{\Lambda}_2)^{-1} \text{ and } \mathbf{H}_2 = (\mathbf{I}+\mathbf{\Lambda}_2)^{-1} \tag{3-11}$$

Provided that $\mathbf{H}_2$ is invertible, $\mathbf{V}(s)$ and $\mathbf{U}(s)$ form the left co-prime factors of $\mathbf{K}(s)$:

$$\mathbf{K}(s) = \mathbf{V}(s)^{-1}\mathbf{U}(s) \tag{3-12}$$

3.1.1 Stability in the framework

As discussed by Kothare *et al.*(1994) and Campo and Morari(1990), when a system is saturating it is in effect open loop. This may happen at the decision of the operator too. Hence, the system is required to be globally stable, and the individual components should themselves be asymptotically open loop stable.

3.1.2 Framework Techniques

The techniques which have been shown to be special cases of the framework are listed, for a detailed discussion see Kothare et al.(1994) and the respective references.

3.1.2.1 Anti-reset windup (ARW)

The concept of introducing an additional feedback loop to restrict windup of integral terms has been the subject of much study since the 1960's (Fertik and Ross, 1967). This is a standard SISO technique. In the most general form:

$$\hat{\mathbf{K}}(s) = \left[\begin{array}{c|cc} -\dfrac{1}{\tau_r} & \dfrac{k}{\tau_I\tau_r}(\tau_r-\tau_I) & \dfrac{k}{\tau_r} \\ \hline 1 & k & 0 \end{array}\right] \tag{3-13}$$

Within the parameterisation of the framework equations (3-11), this translates to:

$$\mathbf{H}_1 = \frac{1}{\tau_r} \text{ and } \mathbf{H}_2 = 1 \tag{3-14}$$

It is useful to note that a favoured heuristic choice is $\tau_r = \tau_I$

3.1.2.2 Conventional anti-windup (CAW)

CAW (Doyle *et al.*, 1987) is a generalisation of anti-reset windup and is easily extended from SISO to MIMO problems. It involves the high gain feedback of the error between the controller output and process inputs, the saturation error, to drive the states of the controller to those being represented by the process inputs. CAW maps into the framework equations (3-11) using the following parameterisation:

$$\mathbf{H}_1 = \mathbf{BX}(\mathbf{I}+\mathbf{DX})^{-1} \text{ and } \mathbf{H}_2 = (\mathbf{I}+\mathbf{DX})^{-1} \tag{3-15}$$

X is the design parameter that allows the designer to optimise the performance of the AW compensation.

3.1.2.3 Conditioning Techniques

Hanus (1980) presents the concept of the realisable reference, $\underline{\mathbf{w}}^r$, and present realisability. In this sense, the realisable reference is the reference that should be inputed into the system so that the control action is equal to the process input. Extending back calculation, by introduction of $\underline{\mathbf{w}}^r$ into the state-space equations and solving for the control output, $\underline{\mathbf{u}}$, in terms of the process input, $\underline{\mathbf{u}}^*$, produces the following form of the Hanus conditioned controller (HCT):

$$\hat{\mathbf{K}}(s) = \left[\begin{array}{c|cc} \mathbf{A}_K - \mathbf{B}_K\mathbf{D}_K^{-1}\mathbf{C}_K & \mathbf{0} & \mathbf{B}_K\mathbf{D}_K^{-1} \\ \hline \mathbf{C}_K & \mathbf{D}_K & \mathbf{0} \end{array}\right] \tag{3-16}$$

Which is of the same form as the preferred ARW implementation. The parameterisation as per equations (3-11) is:

$$\mathbf{H}_1 = \mathbf{BD}^{-1} \text{ and } \mathbf{H}_2 = \mathbf{I} \tag{3-17}$$

This method has no tuneable parameters. Its advocates, such as Vrancic and Peng *et al.* (1996) and Peng *et al.*(1996) would claim that this is because this represents an optimal design.

Walgama *et al.* (1992) and Hanus and Peng (1992) show that some of the "short sightedness" of the HCT can be overcome by modifying the realisable reference by a user parameter, ρ, or a filtered form of the saturation error. This variation on the conditioning technique is the generalised conditioning technique GCT, one form given by:

$$\mathbf{H}_1 = \mathbf{B}\left(\mathbf{D}_\mathbf{K} + \rho\mathbf{I}\right)^{-1} \text{ and } \mathbf{H}_2 = \mathbf{I} \tag{3-18}$$

3.1.2.4 Observer based Anti-windup (OAW)

Åström and Hägglund (1988) and Åström and Rundqwist (1989) suggested that an observer, driven by the saturation error, could be used to corrected the states of the controller. In the simplest form:

$$\hat{\mathbf{K}}(s) = \left[\begin{array}{c|cc} \mathbf{A}_\mathbf{K} - \mathbf{L}\mathbf{C}_\mathbf{K} & \mathbf{B}_\mathbf{K} - \mathbf{L}\mathbf{D}_\mathbf{K} & \mathbf{L} \\ \hline \mathbf{C}_\mathbf{K} & \mathbf{D}_\mathbf{K} & \mathbf{0} \end{array}\right] \tag{3-19}$$

Again, this is a similar form to HCT and ARW, with the parameterisation in terms of equation (3-11) is:

$$\mathbf{H}_1 = \mathbf{L} \text{ and } \mathbf{H}_2 = \mathbf{I} \tag{3-20}$$

A modification of the process matrix, $\hat{\mathbf{P}}$ allows the extended Kalman Filter anti-windup strategy, which is an observer-based technique, to be to expressed in terms of the framework.

A paper by Kapoor *et al.*(1998) presents a synthesis method for implementing an observer based AW compensation in MIMO systems which produces good setpoint tracking in the examples presented.

3.1.2.5 Internal Model Control Anti-windup (IMCAW)

The nature of IMC control lends intuitively to the possiblilities of anti-windup compensation, since the controller states should always be in correlation to the process inputs if the process model and process are driven from the same process inputs. However, these control strategies tend to have slow dynamic response and this effects the AW compensation. Kothare *et al.*(1994) show that a two degree of freedom model fits the framework and Zheng *et al.* (1994) show a IMC design that has been modified to improve it's AW characteristics.

3.2 MIMO Specific techniques

These techniques are aware of the performance issues related to MIMO systems. These techniques have similar objectives in mind to those driving ER compensation. The two main areas of consideration come from: Doyle and Packard (1987) and Doyle *et al.*(1987) who indicated that the concept of directionality of the control action is significant to the performance of MIMO systems; and Hanus *et al.*(1987) who had proposed the use of the conditioning technique to MIMO systems.

3.2.1 Modified Anti-windup

Modified anti-windup, MAW, is an attempt by Doyle and Packard (1987) and Doyle *et al.*(1987) to improve the performance of MIMO systems by maintaining the directionality in the input vector $\underline{\mathbf{u}}^*$.

The concept of MAW is to scale the control action so that the largest signal is within the realisable control space. This is done by directly modifying the output and state equations of the controller. The implementation is in equations (3-21 to 3-22) and Fig.14.

$$\begin{aligned}\dot{\underline{\mathbf{x}}} &= \left(\mathbf{A}_K + \beta(\alpha(t)-1)\mathbf{I}\right)\underline{\mathbf{x}}_K + \mathbf{B}_K\underline{e} \\ \underline{\mathbf{u}} &= \mathbf{C}_K\underline{\mathbf{x}}_K + \alpha(t)\mathbf{D}_K\underline{e}\end{aligned} \tag{3-21}$$

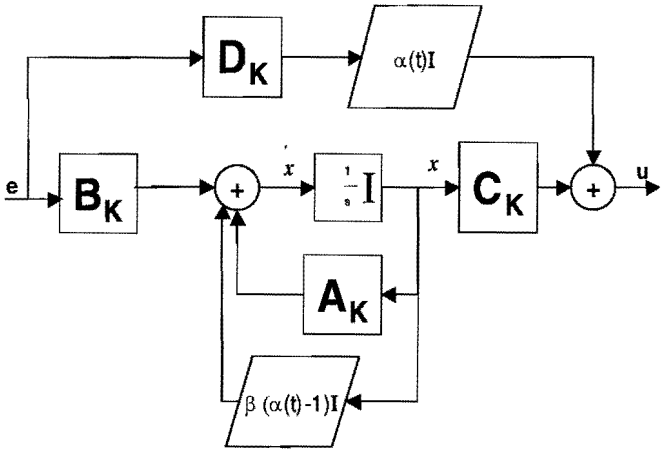


Fig.14 MAW Implementation Diagram

The MAW scaling parameter, $\alpha(t)$, is calculate as:

$$\alpha(t) = \begin{cases} 1 & \forall i, u_i = u_i^* \\ \min\left(u_i^* / u_i\right), \text{where } u_i \neq 0 & \text{else} \end{cases} \tag{3-22}$$

Moreover, the corresponding AWBT operator would be:

$$\Lambda = \begin{bmatrix} \beta(\alpha-1)\mathbf{I} & 0 & 0 & 0 & 0 \\ 0 & (\alpha-1)\mathbf{D}_K & 0 & 0 & 0 \end{bmatrix} \tag{3-23}$$

β is a design parameter setting the effectiveness of containing the states. Stability of MAW compensated systems is ensured since the compensation only modifies the real part of the eigenvalues. Unfortunately, this does not factor into the $\mathbf{H}_1$ and $\mathbf{H}_2$ parameters of the framework.

Independent tests by Mattern (1993) and Marcopoli and Phillips (1994) show that although neither CAW nor MAW are absolutely superior, MAW properties showed promise in the MIMO case.

3.2.2 Artificial Non-linearity

The Artificial Non-linearity (AN) concept is that to aid performance a non-linear operation is carried out on the output of the controller and has the general form, as in Fig.15, of:

$$\underline{\mathbf{u}}^{AN}(t) = \mathbf{AN}(\underline{\mathbf{u}}(t)) \tag{3-24}$$

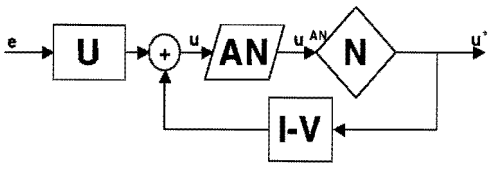


Fig.15 AN Implementation Diagram

In 1998, Peng *et al.* (1996) present a comparison of AN designs that where previously suggested by Campo and Morari (1990) and what the authors present as an optimal design by Hanus and Kinnaert (1989), which is reworked for implementation by the authors.

The Campo and Morari (1990) AN is a simple scaling of $\underline{\mathbf{u}}(t)$ as used in MAW. The Hanus and Kinnaert version is derived from manipulation of the realisable reference, $\underline{\mathbf{w}}^r$,in terms of a constrained control space. A brief description in terms of the Kothare *et al.*(1994) framework, as opposed to the modified framework of Peng *et al.*(1998), is presented here. For more detailed derivation of this method refer to Hanus and Kinnaert (1989) and Peng *et al.*(1998).

The control action being generated is:

$$\underline{\mathbf{u}} = \mathbf{U}(s)(\underline{\mathbf{w}} - \underline{\mathbf{y}}) + (\mathbf{I} - \mathbf{V}(s))\underline{\mathbf{u}}^* \tag{3-25}$$

Let $\underline{\mathbf{w}}^r$ be the realisable reference that would satisfy:

$$\underline{\mathbf{u}}^* = \mathbf{N}(\underline{\mathbf{u}}) = \mathbf{U}(s)(\underline{\mathbf{w}}^r - \underline{\mathbf{y}}) + (\mathbf{I} - \mathbf{V}(s))\underline{\mathbf{u}}^* \tag{3-26}$$

It is easy to see that the realisable reference can be found in terms of the control action signals:

$$\underline{\mathbf{w}}^r = \underline{\mathbf{w}} - \mathbf{U}^{-1}(s)(\underline{\mathbf{u}}^* - \underline{\mathbf{u}}) \tag{3-27}$$

The use of quadratic optimisation of the difference between the realisable reference and the demanded reference:

$$J\left(\underline{\mathbf{w}}^{r}-\underline{\mathbf{w}}\right)=\left(\underline{\mathbf{w}}^{r}-\underline{\mathbf{w}}\right)^{T}\boldsymbol{\Lambda}\left(\underline{\mathbf{w}}^{r}-\underline{\mathbf{w}}\right) \tag{3-28}$$

Can be used to form an optimal AN function.

Solving for the HCT case, equation (3-27) can be found as:

$$\underline{\mathbf{w}}^{r}=\underline{\mathbf{w}}-\mathbf{D}_{\mathbf{K}}^{-1}(s)(\underline{\mathbf{u}}^{*}-\underline{\mathbf{u}}) \tag{3-29}$$

Describing the saturation element, equation (2-9) by the following constraint inequality:

$$\mathbf{H}\underline{\mathbf{u}}^{*}+\underline{\mathbf{b}}\leq 0 \tag{3-30}$$

Where $\mathbf{H}=\left[\mathbf{I}_{m\times m}\quad -\mathbf{I}_{m\times m}\right]^{T}$ and $\underline{\mathbf{b}}=\left[-\bar{u}_{1}\ \ldots\ -\bar{u}_{m}\quad \underline{u}_{1}\ \ldots\ \underline{u}_{m}\right]^{T}$.

Using HCT solve equation (3-29) for $\underline{\mathbf{u}}^{*}$ and substitute into equation (3-30). Hence the constraint defining the control action region for the HCT case is:

$$\mathbf{H}\mathbf{D}_{\mathbf{K}}\left(\underline{\mathbf{w}}^{r}-\underline{\mathbf{w}}\right)+\mathbf{H}\underline{\mathbf{u}}+\underline{\mathbf{b}}\leq 0 \tag{3-31}$$

Application of the Kuhn-Tucker multipliers and reducing the system to active constraints represented by, $\mathbf{H}_{0}$ and $\underline{\mathbf{b}}_{0}$, Hanus and Kinnaert (1989) have formulated the close solution of this optimisation problem when used with HCT:

$$\underline{\mathbf{u}}^{AN}(t)=-\mathbf{D}_{\mathbf{K}}\boldsymbol{\Lambda}^{-1}\mathbf{D}_{\mathbf{K}}^{T}\mathbf{H}_{0}^{T}\left(\mathbf{H}_{0}\mathbf{D}_{\mathbf{K}}\boldsymbol{\Lambda}^{-1}\mathbf{D}_{\mathbf{K}}^{T}\mathbf{H}_{0}^{T}\right)^{-1}\left(\mathbf{H}_{0}\underline{\mathbf{u}}(t)+\underline{\mathbf{b}}_{0}\right)+\underline{\mathbf{u}}(t) \tag{3-32}$$

Peng *et al.*(1996) show that using an example presented by Zheng *et al.*(1994) that the Hanus and Kinnaert artificial non-linearity technique out performs the Campo and Morari (1990) AN.

3.2.3 *Model Predictive Control*

Model Predictive Control (MPC) is an optimal control technique where the best control action is calculated by simulating the process response over the event horizon and then applying the optimal control action to output at the next control instant.

The MPC algorithm has long been seen as an optimisation rather than a stabilisation technique due to the slowness of generating the resultant control action. With the increase in online computing power MPC has gained industrial acceptance (Qin and Bagewell, 1996). The usefulness of MPC is that once the process has been modelled, a designer is able to apply any mathematical strategies to iterate for the optimal control action. The structure of the solution is that all available degree's of freedom can be utilised.

De Prada and Valentin (1996) have demonstrated the usefulness of MPC in this arena, however, this is outside the scope of this thesis.

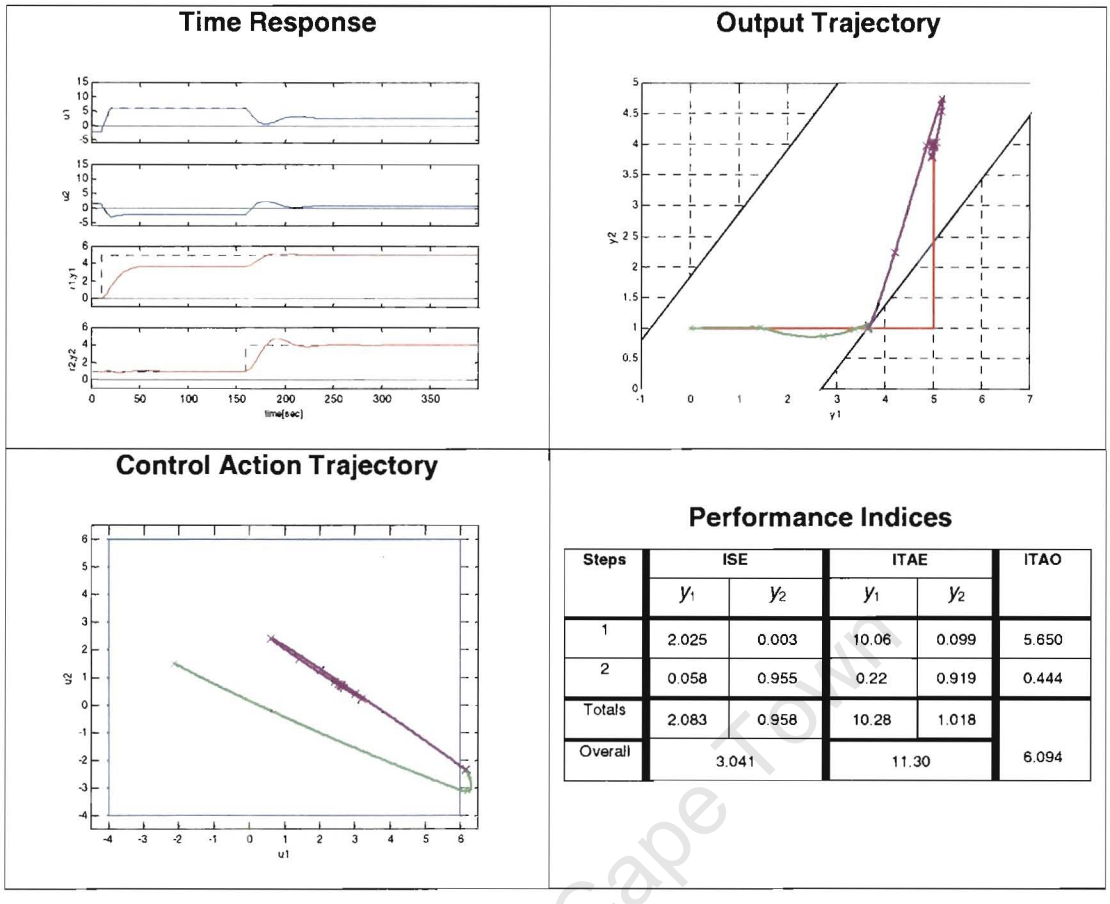
3.3 *Including Anti-windup Compensation*

The various AW techniques work to maintain the controller states with different net effects. A number of significant implementations of different techniques will be presented and discussed in the final section of this chapter.

3.3.1 *Conventional Anti-windup*

Examination of a number of simulations indicated that a value of $\mathbf{X} = 10\mathbf{I}$, for the CAW tuning parameter in equation (3-15), would perform adequately in terms of the performance indices as well as the primary objective of restricting the requested control action to the available control space.

Table 6 below indicates the times taken to reach steady-state, as defined in the glossary. The time response trends in Fig.16:Time Response show greatly improved performance over the saturating case. The system maintains the setpoint in the second loop and although the performance of the first loop is unsatisfactory, the system has remained stable and performed gracefully.



Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	2.025	0.003	10.06	0.099	5.650
2	0.058	0.955	0.22	0.919	0.444
Totals	2.083	0.958	10.28	1.018	
Overall	3.041		11.30		6.094

Fig.16 CAW MVT Compensation Results

Comparing the performance indices of the CAW compensated case, Fig.16, to the normalised ideal response given in Table 4 and the saturating response given in Table 5, the overall performance of the system has been greatly improved. Loop 2 performed better than the ideal case, while the first loops performance suffered due to the limited control action. In an ISE sense the performance was twice that of the ideal case, since the transients were well behaved. In the ITAE sense, the performance of the first loop was ten times that of the ideal case. This is due to the large steady-state error that the system experienced during the first step.

Table 6 CAW MVT Process Settling Times

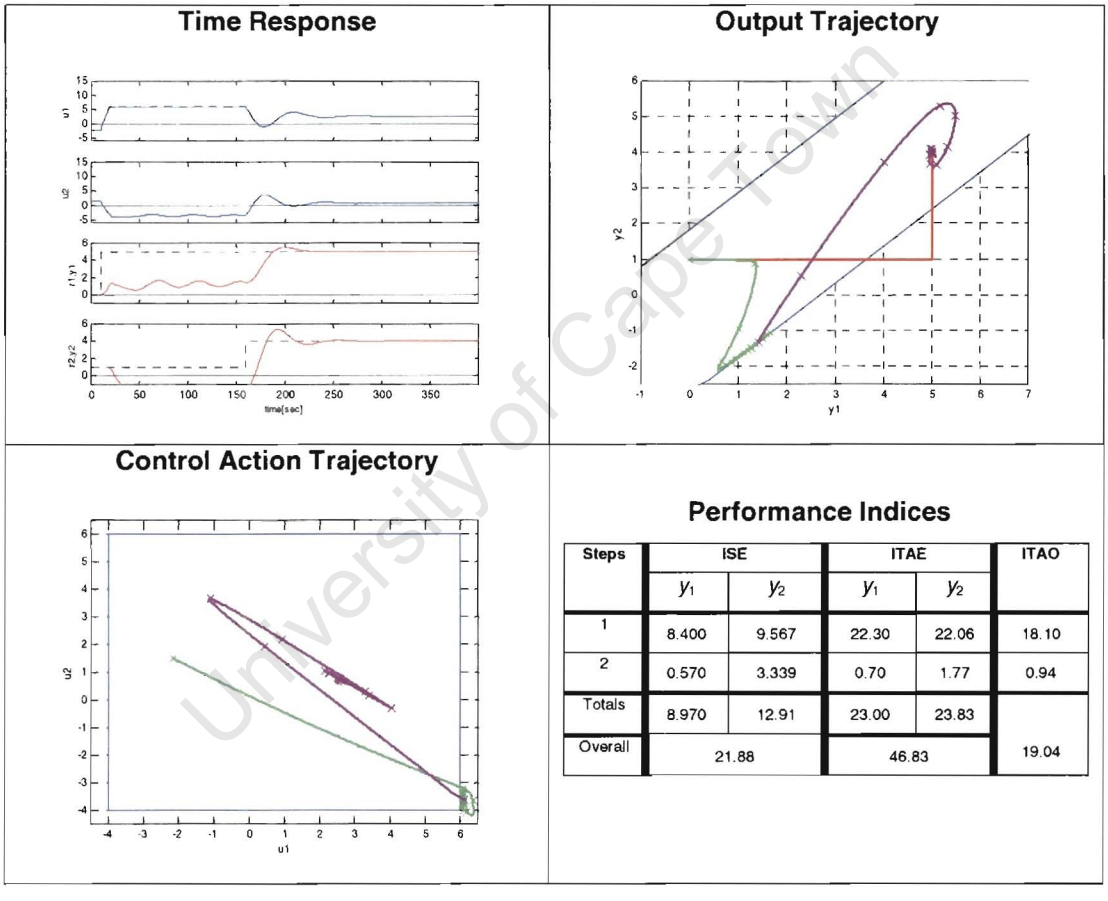
Step time	y_1	y_2
$r_1 = 10$	42	-
$r_2 = 160$	-	67

The control action trajectory in Fig.16 shows that the requested control action was better contained to the available control action.

After the first step, the steady-state position of the process is shown to be maintaining the second loop setpoint and the system moves smoothly to operating point C after the second step.

3.3.2 Hanus Conditioning Technique

HCT compensation has no tuneable parameters. It is implemented according to equations (3-17). From the results in Fig.17 it is clear that this method has also contained the controller states, but the performance is significantly poorer than the CAW case in that both setpoints have been lost.



Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	8.400	9.567	22.30	22.06	18.10
2	0.570	3.339	0.70	1.77	0.94
Totals	8.970	12.91	23.00	23.83	
Overall	21.88		46.83		19.04

Fig.17 HCT MVT Compensation Results

Compared to the saturating case discussed in Section 2.3, the system performs better with the performance returning to the linear mode soon after the second step. The settling times are given in Table 7.

Table 7 HCT MVT Process Settling Times

Step time	y_1	y_2
$r_1 = 10$	-	-
$r_2 = 160$	-	85

3.3.3 Modified Anti-windup

This technique is the one of the first to make allowances for maintaining the directionality of the MIMO control vector. It directly modifies both the outputs and states of the controller.

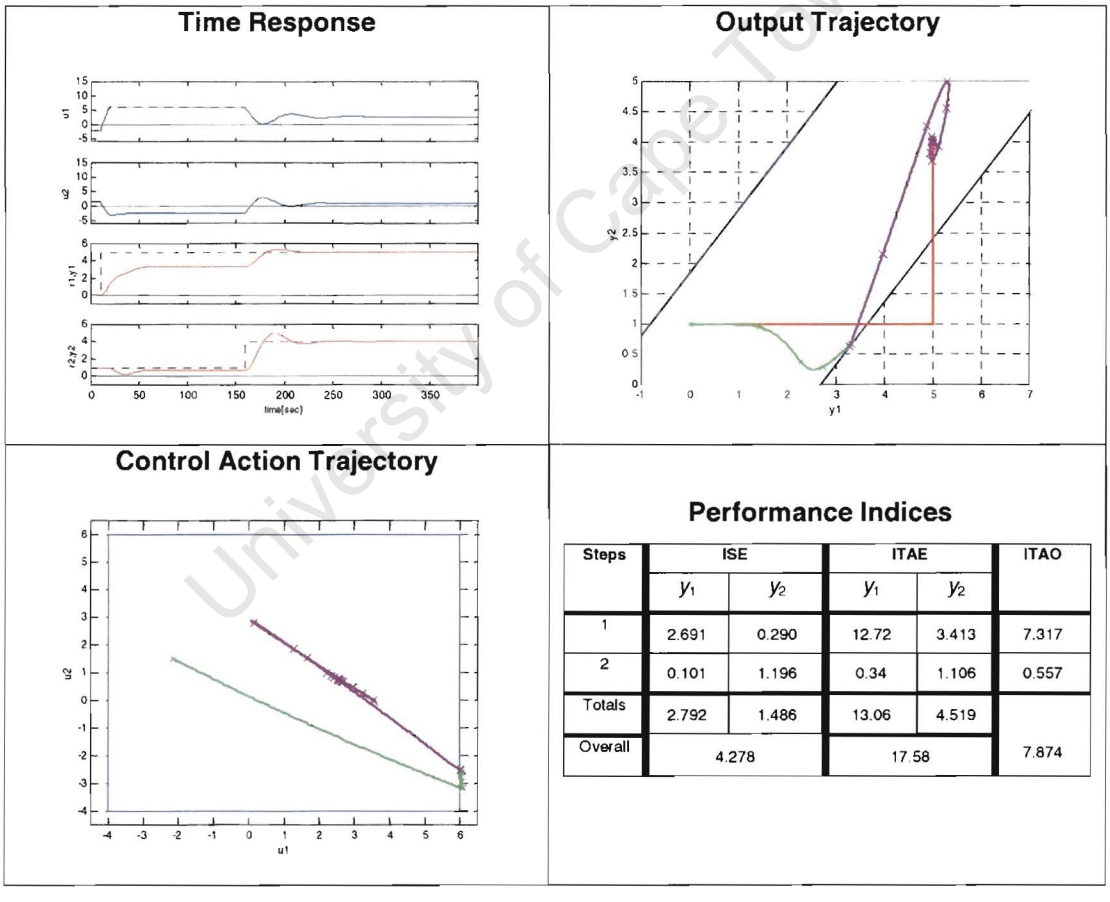


Fig.18 MAW MVT Compensation Results

This implementation is modeled on that suggested by Mattern (1993). For the simulation in Fig.18 a tuning of $\beta = 10$ is used in equation (3-22), additional results are in Appendix B. By introducing a nonlinear element to scale the control vector when an input is saturating, this compensation's performance is significantly better than that reported in the HCT case.

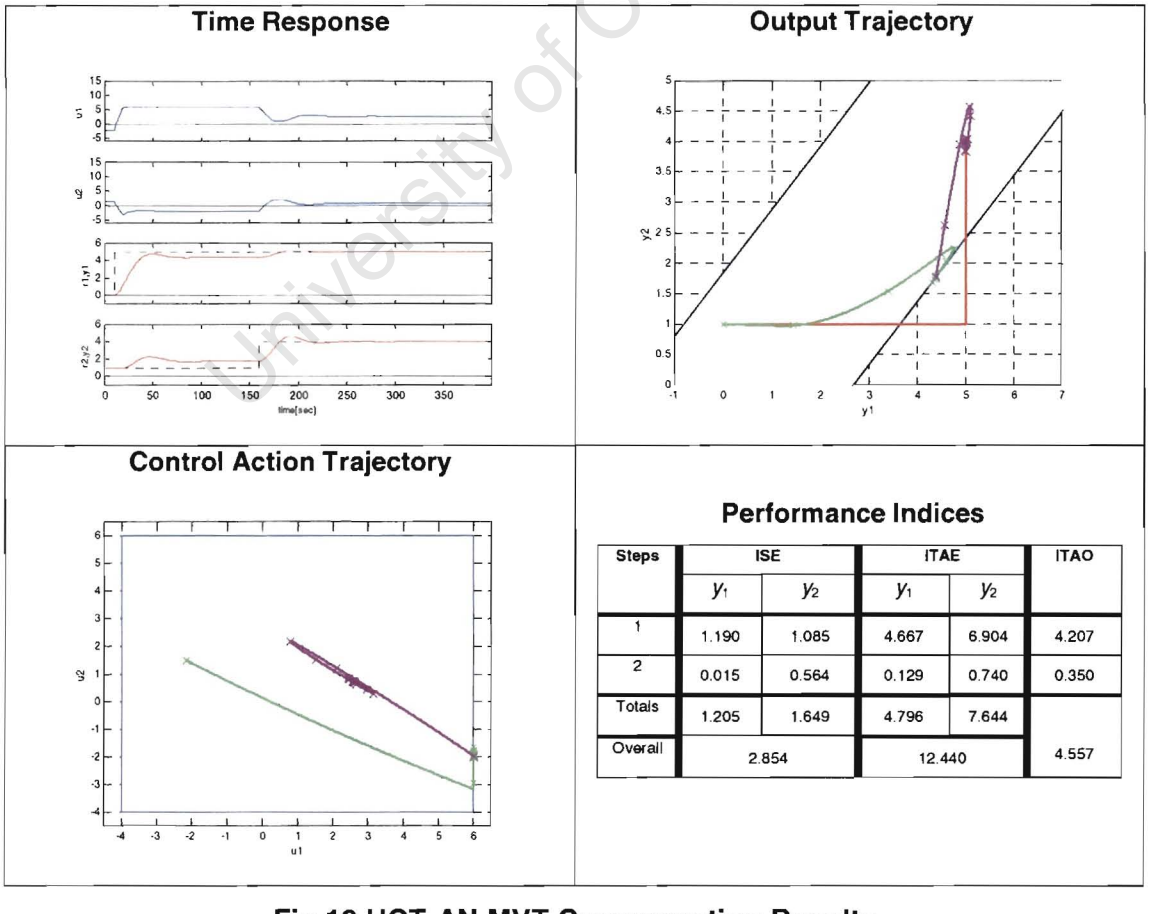
Increasing the tuning parameter, and thereby placing a greater feedback on the controller states, pulls the control action closer to the realisable region; but for $\beta = 10$, the performance is still less than that obtained by the CAW example. Settling times are given in Table 8.

Table 8 MAW MVT Process Settling Times

Step time	y_1	y_2
$r_1 = 10$	40	-
$r_2 = 160$	-	80

3.3.4 Artificial Non-linearity

This method introduces a closed form optimisation to the output from a HCT compensated system, so as to produce a realisable control action that is constrained and optimised to limit the impact on the system performance. This implementation is modelled on that suggested by Vrancic and Peng (1996). For this simulation, a value of $\Lambda = \mathbf{I}$ was used in equation (3-32). Additional results will be discussed in Chapter 5.



Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.190	1.085	4.667	6.904	4.207
2	0.015	0.564	0.129	0.740	0.350
Totals	1.205	1.649	4.796	7.644	
Overall	2.854		12.440		4.557

Fig.19 HCT-AN MVT Compensation Results

The results in Fig.19 are clearly better than the previous results in terms of the containment of the control action. This is because the nature of the AN, which has being designed to operate with the HCT, is such that when active the saturating signal should be set to the specified limit. Both loops have left their setpoints, but from the output trajectory in Fig.19, clearly some equilibrium between the two setpoints has been found.

The settling times are given in Table 9.

Table 9 HCT-AN MVT Process Settling Times

Step time	y_1	y_2
$r_1 = 10$	50	-
$r_2 = 160$	-	50

3.3.5 Anti-windup Performance Evaluation

Typically, windup is most likely during the transient from one operating point to the next. The example operating point changes used here deliberately place the operating point outside the realisable operating region.

In terms of the performance indices, the overall rating would be, from worst to best performing: HCT, MAW, CAW and HCT-AN. All made significant improvements over the uncompensated case by containing the controller states. The return to linear performance and setpoint tracking was a function of how close to the operating point the systems found themselves at the step change.

Clearly, the combination of HCT and it's AN implementation, being the only formulation to include more than the control boundaries in it's design criteria, has being the only compensation considered here that indicates the possibility of optimising the operation during this nonlinear mode of operation.

In Chapter 5 further comparisons between the error redistribution compensation and AN compensation are made.

3.4 Discussion of Anti-windup Techniques

By far the most significant feature of AWBT is the maintenance of stability in the system during nonlinear operation. However, as the examples presented here as well as those by Doyle *et al.*(1987), Zheng *et al.*(1994) and Peng *et al.*(1998), amongst others, show different AW techniques do not perform equally well in all cases. This is why a framework is so significant, as it allows researchers to analyse the different techniques in a common structure. In general, the synthesis of AWBT techniques is still based on a heuristic approach which is then fitted into the framework.

The paper by Peng *et al.*(1996) giving advise of designing AW compensated controllers, shows the formulation of conditioning techniques for the PID controller. This particular formulation shows that the conditioning technique is a special case of the conventional anti-windup (CAW) in that the structure of the solution is that of CAW, but the derivation of the AWBT feedback element is derived from heuristic knowledge of the process.

As many such as Campo and Morari (1990), Kothare *et al.*(1994), Peng *et al.*(1998) have said, anti-windup compensation can be added after the design of the linear controller. The results from the increasing experience, including the examples presented here, shows that the relationship between controller and compensator play a larger roll in the performance than at first appeared, and performance gains are possible if the controller is designed with AW in mind, Vrancic and Peng (1996) and Kapoor *et al.*(1998).

The factors that determine the successful performance of AW compensation have being elusive. Hanus *et al.*(1987) quote a list of successful application of the conditioning technique. Campo and Morari (1990) in presenting a framework prior to Kothare *et al.*(1994) attempts to specify the performance requirements and thereby the synthesis techniques for generating an optimal AW compensator. As commented, the means to solve this H_∞ problem are not at hand (Edwards and Postlethwaite, 1998).

Currently, in the literature there appear to be two main focuses: firstly, the application of the techniques that have been developed and refined over the last 20 year to meet higher performance requirements (Kapoor *et al.*1998 , Turner and Walker 2000); and secondly, the bridging of the gap between classical MIMO and modern techniques in the union of AWBT and MPC technologies (Cherukuri and Nikolaou 2001).

4 Error Redistribution – Output Optimisation

From the literature reviewed in Chapter 3 it is evident that a large amount of effort is going into resolving the saturation problem. In all of these cases, the objective has been to ensure stability and robustness during nonlinear operation. Error Redistribution (ER) is introduced as a novel means to allow prioritised responses in the nonlinear mode. By prioritised response, cognisance is taken of the fact that all outputs are not necessarily of equal consequence.

Using the MVTP example, the necessity of including this objective will be highlighted. To begin with, the generic ER principle will be discussed and then MVTP responses with ER compensation will be presented. The formulation of the ER compensation technique will follow.

4.1 The Error Redistribution Principle

Multi-input Multi-output (MIMO) systems are classified as such due to the interaction between an input and several outputs. In the case were there is no interaction between loops, the transfer function model for the process is diagonal; each loop can simply be treated as a separate SISO problem. This does not mean that the effects of saturation can be ignored, but that AWBT techniques should be applied to ensure stability and robust performance.

ER is a MIMO technique that aims to exploit the interaction between inputs and outputs to meet designed performance requirements. In any real process the inputs, or rather the control action actuators, will be bounded by physical constraints. These include at least saturation and rate of change limits, possibly deadband, hystersis and backlash to name but a few of the common physical limits. Hence it is possible to state that the absolute realisable control action space is limited to a finite space and that, depending on the nature of nonlinearities, the path that the control action can take at any point will be limited; resulting in the currently realisable space being a subset of the absolute.

Given that the input space to a process is limited, the realisable output space of a stable process would too be limited to a realisable output space. The steady state realisable output region would be a subset of the absolute region as more complex systems can experience overshoot and thus extend their outputs passed the steady state limits.

To summarise, the input to any process is a bounded region of control action space which can be mapped to a corresponding output space region. At any point in time, the currently attainable control action could be a subset of the absolute, while the output space can extend past the steady state region.

Having ascertained that a limited output space exists; it is not always possible to define this space accurately. Firstly, it would require an exceptional model of the process and in particular the nonlinearities in the actuators. Secondly, drift or failure within the system would result in a time variance of the model that would have to be compensated for. Thirdly, disturbances, either at the actuator or within the process itself, can have the effect of changing this output space. Therefore, if any useful calculation of this space were to be done, it would have to be conservative at the best. Hence ER, while accepting and working within the conditions that the realisable output region imposes, does not make direct use of this.

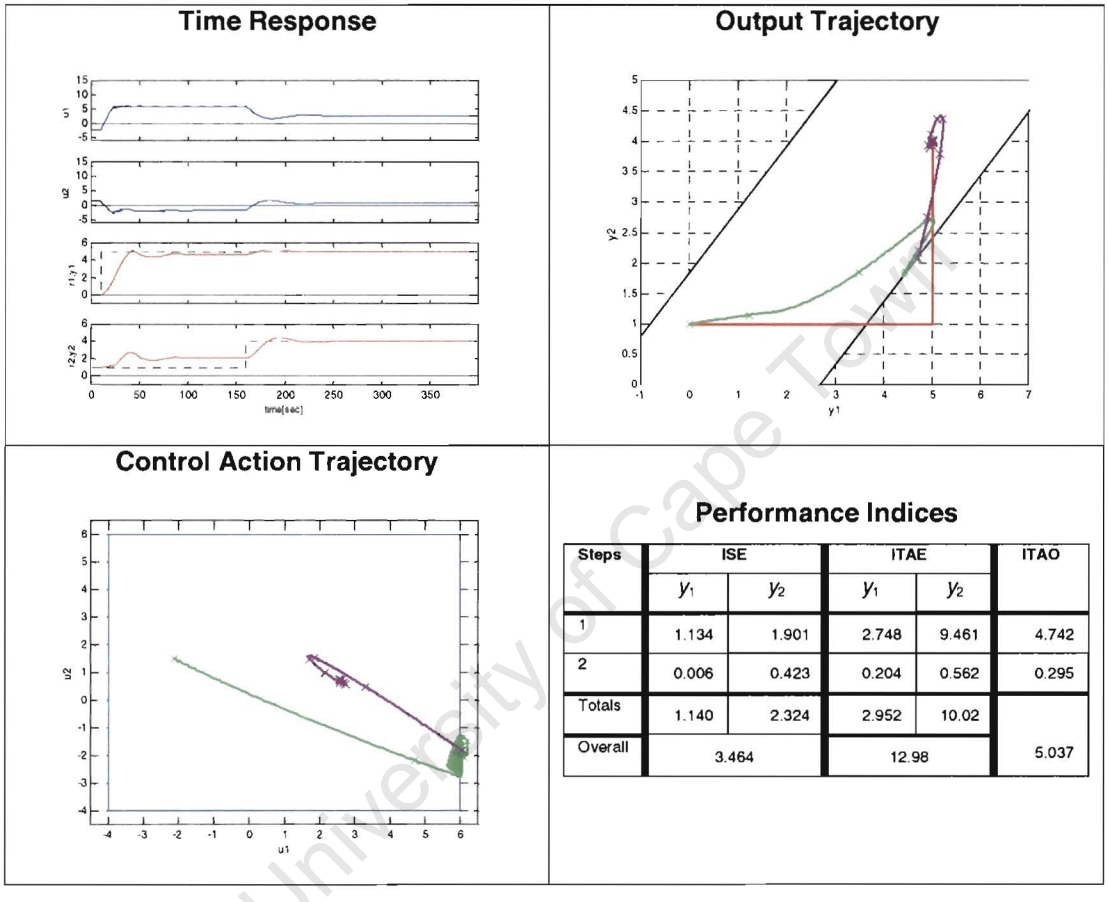
Consider Fig.3, this shows a typical control action region mapped to a corresponding output space region for a 2x2 1st order system. It is evident that operating points can exist where either setpoints would be obtainable, but not concurrently. In fact, the system would find that at least one input had saturated in attempting to reach this unrealisable operating point. Analysis of such conditions where CAW compensation has been employed, indicate that this is indeed the case. CAW acts to prevent the effects of saturating control action permeating throughout the entire controller and process by limiting the controller states to the currently realised control action. This condition results in the controller maintaining outputs at setpoint on the non-saturating loops; while the saturating loop experiences a steady state error during this nonlinear mode of operation.

Where the saturating loop is of higher priority than the non-saturating loop, ER compensation allows the designer to specify this and will in effect move the operating point to minimise the error in the high priority loop. The means by which this is realised is by redistributing the error in the saturating loop to the non-saturating loop, which will be referred to as the correcting loop. This redistribution of error will effect a change in the process input/controller output of the correcting loop, which, through the off diagonal terms in the process, will be used to minimise the error in the saturating loop.

This technique was developed independently, Carew (1996-1997), to the HCT-AN technique, presented by Peng *et al.* (1998), that was discussed in the previous chapter. Comparisons of the two techniques will be presented in Chapter 5.

4.2 *MVTP ER Compensated Responses*

In Chapter 3 the response of this system with various AW compensation implementations was presented. Now consider that if loop of y_1 is of greater significance than y_2. ER compensation is added to the system, giving y_1 a priority of $P_1 = 1$ and y_2 a priority of $P_2 = 2$.



Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.134	1.901	2.748	9.461	4.742
2	0.006	0.423	0.204	0.562	0.295
Totals	1.140	2.324	2.952	10.02	
Overall	3.464		12.98		5.037

Fig.20 ER MVT Compensation results

Fig.20 shows the response of the system with ER compensation. It is clear that the first loop, the saturating loop, has improved significantly and is again comparable to the ideal response (see Fig.4). As a consequence of this the second loops' performance has been degraded. This loop is the correcting loop and, as explained earlier, the available control action in the correcting loop is used to move to an operating point so as to minimise the error in the saturating loop. Since this is a dynamic system, the correcting loop is seen to search for the optimal operating point where the error in both loops is minimised. The table below indicates the times taken to reach steady-state.

Table 10 ER compensated MVT Process Settling Times

Step time	y_1	y_2
$r_1 = 10$	62	-
$r_2 = 160$	-	53

The exact implementation of the ER compensation for this case will be discussed in section 4.6. The control action in Fig.20 is, in general, held within the available control space, but experiences more "jitter" in the process of optimising the output.

4.3 Considerations for ER Compensation

To allow the formulisation of ER compensation, the concepts of output prioritisation and nonlinear operation have to be defined. A suggested prioritisation and performance evaluation strategy are discussed next. These are given as guidelines for the most general of cases.

4.3.1 Optimisation - Prioritisation and Performance

The ER function needs to take the relative priority of the saturating loop and correcting loop(s) into consideration. The prioritisation vector, $\underline{\mathbf{P}}$, is defined as:

$$\underline{\mathbf{P}} = [P_1 \quad P_2 \quad \ldots \quad P_m]^T \tag{4-1}$$

where m = the number of outputs and,

P_i = Priority of the output y_i,

where the highest priority value is 1 and the lowest possible is m.

Since it is possible that several outputs are of the same priority, the method of assigning priority is:

- The most significant output is assigned the priority $P_i = 1$,
- A $P_i = 0$ would indicate that output i is not to be used in ER,
- If output y_i and y_j are of the same priority, $P_i = P_j$,
- Since this vector is of relative priority, fractional increments are possible $P_i = P_j + 0.2$

- Priorities are numbered sequentially but not uniquely, i.e:

$$\underline{\mathbf{P}} = \begin{bmatrix} 3 & 1 & 2 & 3 & 2 & 0 & 4 \end{bmatrix}^T.$$

An $m \times m$ error redistribution matrix, $\boldsymbol{\delta}$, based on a priority vector $\underline{\mathbf{P}}$ is used to select the corrective action of the ER function. In principle, the ER matrix maps the error to the redistributed error vector, $\underline{\mathbf{e}}^{ER}$:

$$\underline{\mathbf{e}}^{ER} = \boldsymbol{\delta}\underline{\mathbf{e}} \tag{4-2}$$

Where any off-diagonal term, $\delta_{i^c j^s} \neq 0$ would indicate that should the j^s th loop be operating in a nonlinear mode, i.e. if the dominant input to y_{j^s}, u_{k^s}, is saturating, the i^c th loop would be used to exercise correcting action, in principle $e_{ER_{i^c}} = e_{i^c} + \delta_{i^c j^s} e_{j^s}$, on the former hence shifting the operating point of y_{i^c}. To generate the error redistribution matrix:

- First calculate the relative priorities:

$$\delta_{i^c j^s} = \begin{cases} P_{i^c} - P_{j^s} & \text{where } P_{i^c} - P_{j^s} > 0 \text{ and } P_{j^s} \neq 0 \\ 0 & \text{else} \end{cases} \tag{4-3}$$

  Obviously if there is no process interaction between outputs *i* and *j*, δ_{ij} should be set to zero.

- Secondly, obtain the total effective ER across all correcting loops for output y_{j^s}:

$$\delta^{\Sigma}_{j^s} = \delta^{\Sigma}\left(j^s\right) = \sum_{i^c}^{i^c \neq j^s} \delta_{i^c j^s} \tag{4-4}$$

- Finally, add the diagonal (this indicates normal linear operation) and normalise:

$$\delta_{i^c j^s} = \begin{cases} \delta_{i^c j^s} \Big/ \delta^{\Sigma}_{j^s} & \text{where } i^c \neq j^s \\ 1 & i^c = j^s \end{cases} \tag{4-5}$$

Where the performance index is chosen as the weighted integral-square-error, $wISE$:

$$wISE = \sum_i w_i \int_{t_1}^{t_2} e_i^2(t)dt \tag{4-6}$$

it can easily be related to economic factors dictating a preference to a particular system output. Based on the priority vector, $\underline{\mathbf{P}}$, a weighting function is suggested:

$$w_i = \begin{cases} \frac{(m+1)-P_i}{m} & \text{where } P_i \neq 0 \\ 1/m & \text{else} \end{cases} \tag{4-7}$$

This weighting function gives the output with priority of $P_i = 1$ a weighting of one and all others a fraction based on their relative priorities. For example, consider the priority vectors and their resultant weighting functions:

$$\begin{aligned} \underline{\mathbf{P}} &= [1 \quad 2 \quad 3]^T \Rightarrow \underline{\mathbf{w}} = [1 \quad 2/3 \quad 1/3]^T \\ \underline{\mathbf{P}} &= [1 \quad 2 \quad 2 \quad 3]^T \Rightarrow \underline{\mathbf{w}} = [1 \quad 3/4 \quad 3/4 \quad 1/2]^T \end{aligned} \tag{4-7}$$

4.3.2 Suitable conditions for ER

ER can be applied to multivariable systems where significant interaction exists within the process model, $\mathbf{G}(s)$, and AW has been applied. ER is particularly suitable when the process and linear controller are diagonally dominant, but not necessarily in both row and column sense.

Even in cases where there is no difference in the priority of outputs, the designer can still use ER to ensure optimal overall system performance. This will be illustrated in the application example. Linear performance recovery is key to the success of ER. This makes it necessary that ER operates primarily when the system is experiencing nonlinear operation:

The nonlinear condition is detected by:

$$\left|u_{k^s}^* - u_{k^s}\right| > \varepsilon_u \text{ and } \left|e_{i^s}\right| > \varepsilon_e \tag{4-8}$$

being true. Where the nonlinear sensitivity parameters, ε_u and ε_e, are fractions of the range of the control action and the output. e_{i^s} is the error of the saturating loop and $u^*_{k^s} - u_{k^s}$ is the difference between the dominant actuator input to the plant, $u^*_{k^s}$, and the controller output u_{k^s}. To ensure linear performance recovery, the ER function, E, should only exhibit influence during this condition. Or more realistically $\lim_{t \to t_l} \mathrm{E} = 0$, where t_l is some finite time after the system has returned to a linear mode.

The nonlinear sensitivity parameters, ε_u and ε_e, need to be selected with the units of the respective signals in mind. If the system is being designed with the units of % of full scale of the control action or output, a scalar percentage often serves. However, if the loops are being designed using their natural units, vector forms of the parameters will be required. In general, in the discussion presented here, a scalar form will be used.

A stronger constraint of $\left|\dot{e}_i\right| < \varepsilon_{\dot{e}}$ could also be imposed where the system will naturally reach the desired operating point. However, simulations have shown that the effects of ER during these periods improve the rate at which saturating loops reach their operating point.

Notice, that most AWBT techniques and ER are designed as functions of the difference between the controller output and the process input, and in the case of ER, the process outputs. In any situation where a difference occurs, i.e. not only at saturation due to limits, but also at saturation due to equipment failure, AWBT and ER will work to stabilise and optimise the process.

4.4 Formulation of ER Compensation

Working from Kothare *et al.*(1994) AWBT framework, ER compensation is included. Firstly, it is necessary to detect the nonlinear mode of operation and then what, if any, corrective action is to taken. This function is carried out by the ER State Selector (ERSS) that produces the effective ER matrix, $\boldsymbol{\delta}^*$, which is derived from the ER matrix $\boldsymbol{\delta}$, equations (4-2 through 4-5), and the system signals. Secondly, a suitable ER transfer function is to be developed in section 4.4.1. Fig.21 shows the placement of the ER components in the controller structure to the ER compensated controller, $\tilde{\mathbf{K}}(s)$.

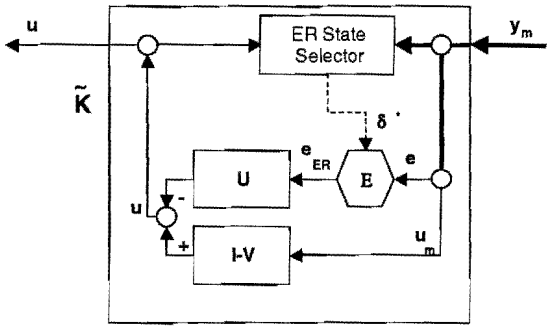


Fig.21 Error Redistribution Compensation

The effective ER matrix is therefore defined as:

$$\boldsymbol{\delta}^*_{i^c j^s} = \begin{cases} 1 & i^c = j^s \\ \delta_{i^c j^s} & \left|u_{m_{k^s}} - u_{k^s}\right| > \varepsilon_u \text{ and } \left|u_{m_{l^c}} - u_{l^c}\right| < \varepsilon_u \text{ and } \left|e_{j^s}\right| > \varepsilon_e = \mathrm{ERSS}(\underline{\mathbf{e}}, \underline{\mathbf{u}}, \underline{\mathbf{u}}_m, \boldsymbol{\delta}) \\ 0 & \text{otherwise} \end{cases} \quad (4\text{-}9)$$

In the linear mode of operation, $\boldsymbol{\delta}^* = \mathbf{I}$. Where the dominant process input to loop j^s, $u_{m_{k^s}}$, is not equal to the controller output, u_{k^s}, and the output y_{j^s} has not reached setpoint, the $j^{s\,\text{th}}$ column's off-diagonal elements are replaced with the values from the ER matrix where $\delta_{i^c j^s} \neq 0$, except when the control action on the correcting loop, u_{l^c}, is itself saturating. Thus indicating that the $i^{c\,\text{th}}$ loops will be used as correcting loops for the saturating $j^{s\,\text{th}}$ loop.

4.4.1 Design of the ER function

The ER function is designed to allow the controller to utilize all the available degrees of freedom (non-saturated loops) to optimise the process output. Therefore, the ER compensation error vector, $\underline{\mathbf{e}}^{ER}$, is generated and passed onto the controller:

$$e^{ER}_{i^c}(s) = \sum_{j^s}^{m} \mathrm{E}_{i^c j^s}\left(\delta^*_{i^c j^s}, \underline{\mathbf{e}}, s\right) \quad (4\text{-}10)$$

The structure of the ER function, $\mathrm{E}\left(\boldsymbol{\delta}^*, \underline{\mathbf{e}}, s\right)$, is:

$$\mathrm{E}(\boldsymbol{\delta}^*, \underline{\mathbf{e}}, s) = \begin{cases} \varepsilon_{ij}(s) e_i & i = j \text{ then } \varepsilon_{ij}(s) \equiv 1 \\ \varepsilon_{ij}(s) \delta_{ij}^* e_j & i \neq j \text{ and } \delta_{ij}^* \neq 0 \\ \varepsilon_{ij}(s) & i \neq j \text{ and } \delta_{ij}^* = 0 \end{cases} \tag{4-11}$$

The only non-zero off diagonal terms in $\boldsymbol{\delta}^*$, $\delta_{ij}^* \neq 0$ will be those where the priority of y_j exceeds that of y_i. The designer can set $\varepsilon_{ij}(s)$ to zero if, even though the priorities may justify the taking of corrective action, none is deemed necessary or more likely only a limited number of interacting loops are chosen to effect ER compensation. To ensure linear performance recovery, $\varepsilon_{ij}(s)$ must have zero steady state value, i.e. must not be an integrator. Thus ensuring that the ER function reduces to an identity in the linear mode, $\mathrm{E}(\underline{\mathbf{I}}, \underline{\mathbf{e}}, s) \rightarrow \underline{\mathbf{e}}$, after some time.

4.4.2 Control Loop Interaction

Examining the interaction of the controller, process and ER function, the ε_{ij} terms have the following effect on any controller output u_i:

$$u_j = \sum_k^m k_{jk} e_k^{ER} = \sum_k^m k_{jk} \sum_l^m \varepsilon_{kl} \delta_{kl}^* e_l \tag{4-12}$$

And, therefore, the following affect on the process output:

$$y_i = \sum_j^n g_{ij} u_j = \sum_j^n g_{ij} \sum_k^m k_{jk} \sum_l^m \varepsilon_{kl} \delta_{kl}^* e_l \tag{4-13}$$

Hence the effect of ER, Δ^{ER}, on the saturating and correcting outputs is:

$$\Delta^{ER} y_{i^s} = \sum_{j^c} g_{i^s j^c} k_{j^c k^c} \sum_{l^s} \varepsilon_{k^c l^s} \delta_{k^c l^s}^* e_{l^s} \tag{4-14}$$

where $\delta_{k^c l^s}^* \neq 0$ and $k^c \neq l^c$

$$\Delta^{ER} y_{i^c} = g_{i^c j^c} k_{j^c k^c} \sum_{l^s} \varepsilon_{k^c l^s} \delta_{k^c l^s}^* e_{l^s} \tag{4-15}$$

where $\delta^*_{k^c l^s} \neq 0$ and $k^c \neq l^c$.

Where more than one loop, i^c, is to be employed in correcting for saturating loop, j^s, the inclusion of the ER matrix term $\delta_{i^c j^s}$ in equation (4-5) distributes the correcting action across the available loops so as to minimise the deviation from the requested operating point.

From studying these interactions, two main points of concern are raised:

(i) The effect of off-diagonal non-zero terms in $\boldsymbol{\delta}^*$ will propagate throughout the controller as shown in equation (4-13). This should not adversely affect the performance of loops not engaged in ER compensation.

(ii) In an $m \times n$ system, it is possible that at a particular point in time more than one loop is saturating, and therefore more than one set of loops are trying to affect correcting action. Again, this means that a number of controller outputs will be affected by the correcting action. The question that must be asked is: if these correcting action adversely affect the performance of the non-saturating loops and given the interactions, will ER in fact be able to achieve the desired affect of modifying the saturating loop output, y_{i^s}, through the off-diagonal process elements $g_{i^s j^c}(s)$.

Indeed, in decentralized control strategies, it is the purpose of the controller elements k_{kl} to compensate the dominant control action, u_k, for output y_i for the effect of $g_{il}u_l$, where $l \neq k$. Hence, if the control actions of the loops not being compensated are in a linear mode, the controller will naturally compensate for the additional error terms in the vector $\underline{\mathbf{e}}^{ER}$, which addresses the concern in (i). However, the concern in (ii) relates to what extent this effect can be exploited. If multiple loops were affecting compensation, it would simplify matters to use as few loops as possible to compensate for a particular output, thus reducing the chance of consuming the available control action in linear loops.

If u_k is saturating, the loop is in effect open, and the controller is not able to compensate. However, in general the action of $g_{i^s j^c}u_{j^c}$ will act to return the system to a linear mode. Because of these two opposing affects, the system searches for the operating point closest to meeting the prioritised operating point. In the non-saturating loops the additional

interaction allows the controller to compensate for the effects of ER in the correcting loops, hence not indiscriminately and adversely affecting the system as a whole. Special care does need to be taken that the gains of the ER functions are not so high as to consume all the available control action in the other loops.

4.4.2.1 Ideal ER function

Consider the loop equations in equations (4-14 through 4-15) and the objective of ER compensation, which is to affect a change equal to the error in the loop while the loop is in a non-linear mode of operation. For ease of discussion, consider the case of loop i^s saturating and loop j^c effecting correcting action on the square process $G(s)$:

$$\Delta^{ER} y_{i^s} = g_{i^s j^c} k_{j^c j^c} \varepsilon_{j^c i^s} \delta^*_{j^c i^s} e_{i^s} \approx \delta^*_{j^c i^s} e_{i^s} \tag{4-16}$$

Solving for $\varepsilon_{j^c i^s}(s)$, indicating the theoretical ideal ER function:

$$\varepsilon_{j^c i^s}(s) = \left(g_{i^s j^c} k_{j^c j^c}\right)^{-1} \tag{4-17}$$

And substituting into equation (4-15), shows the expected effect on the correcting loop:

$$\Delta^{ER} y_{j^c} \approx \left(g_{j^c j^c} \Big/ g_{i^s j^c} \right) \delta^*_{j^c i^s} e_{i^s} \tag{4-18}$$

The ER function in equation (4-18) is usually not causal and results in pole-zero cancellation which is not good for internal stability. If $k_{j^c j^c}$ includes an integrator, it is not possible to approximate $\varepsilon_{j^c i^s}$ with the steady-state gain, as this goes to infinity.

4.4.2.2 ER Hold

The discussion has dealt with ER in terms of continuous time systems, and as such if the loop time of the system, $dt \rightarrow 0$, then it would present a smooth effect on the system. However, realising that the effective ER matrix can introduce ER terms as impulses to the system and this is usually true when implementing the controller as a continuous controller in a computer algorithm. Application to the discrete realisation of the controller and process model is justified, and will be reformulated in discrete terms as part of future work.

In the simulation of continuous systems, two paths exist: either the simulation step should $dt \rightarrow 0$ or, at least, for the process to have a smaller step than the control strategy. Alternatively, the use of a hold on the effective ER matrix, $\boldsymbol{\delta}^*$, similar to a zero-order hold, would prevent excessive jitter of the control action.

The simple state machine to implement such a hold is:

```
If δ*_{n-1} = I and δ*_n <> I then
        δ*_n = δ*_n                 % Activate ER matrix
        LastTime = CurrentTime
Else
        If LastTime + HoldPeriod > CurrentTime then
                δ*_n = δ*_{n-1}     % Keep effective ER matrix constant
        Else
                δ*_n = δ*_n         % Enough time has elapsed,
                                    % so update ER matrix
                LastTime = CurrentTime
        End If
End If
```

Since the effect of ER would be held constant for the hold period, the choice of an ER function's gain should be inversely proportional to the hold period.

4.4.2.3 *Realistic Gains*

By considering an approximation of the available control action to the next control instance, an upper bound to the ER gains can be found. Notice that this approximation is not necessary for performance or the stability presented later, but rather helps to indicate the ideal range for the ER gains.

Defining the following terms as the estimated full range values for:

- The control action available to each input:

$$U_{j^c} = \left| \overline{u}_{j^c} - \underline{u}_{j^c} \right| \tag{4-19}$$

- The steady state output range for each output:

$$Y_{i^s} = \left| \overline{y}_{i^s} - \underline{y}_{i^s} \right| \tag{4-20}$$

Hence, the maximum impulse values for the error redistribution vector, $\underline{\mathbf{E}}^{ER}$, would be where the error was the full output range, $e_{i^s} = Y_{i^s}$, so:

$$\Delta E_{i^c}^{ER}(\mathbf{e}) = \sum_{j^s}^{i^c \neq j^s} \varepsilon_{i^c j^s} \delta_{i^c j^s}^{*} e_{j^s} = \sum_{j^s}^{i^c \neq j^s} \varepsilon_{i^c j^s} \delta_{i^c j^s}^{*} Y_{j^s} \tag{4-21}$$

Realistically, ER is only effective when e_{i^s} is within the over the following range:

$$\varepsilon_e \leq e_{i^s} \leq Y_{i^s} \tag{4-22}$$

The initial value response to this impulse should be limited to the available control action:

$$U_{j^c} \geq \left| \sum_{l} \lim_{s \to \infty} sk_{j^c l}(s) \Delta E_l^{ER} \right| \tag{4-23}$$

Defining the effective initial values as

$$K_{j^c l} = \lim_{s \to \infty} s k_{j^c l}(s) \tag{4-24}$$

Gives

$$\begin{aligned} U_{j^c} &\geq \left| \sum_l K_{j^c l} \Delta E_l^{ER} \right| \\ U_{j^c} &\geq \left| \sum_l K_{j^c l} \left[\sum_i^{i \neq l} \varepsilon_{li} \delta_{li}^* e_i \right] \right| \end{aligned} \tag{4-25}$$

Since all terms excepting ε_{li} are known, a linear matrix equation in terms of the non-zero ER functions can be written and solved giving the upper limits for these terms.

Typically the range of error, e_i, given in equation (4-22) represents the worst case, and the maximum value used when considering ε_{li} should be guided by the largest typical step performed by the system. As a rule of thumb, a third of the realisable output span seems reasonable:

$$\varepsilon_{e_i} \leq e_i \leq Y_i / 3 \tag{4-26}$$

Note the chatter in the control action is a function of magnitude of ε_{li}, so the smallest effective values are preferred.

Drawing on the issues raised in the above discussions to form, the following guidelines to find a suitable ER function are given.

4.4.3 ER Guidelines

It is evident that excessive use of ER would have a severe impact on the systems performance. Therefore, the following guidelines should be applied in formulating the ER function $\varepsilon_{ij}(s)$:

- The steady-state gain matrix $G_0 = G(s)\big|_{s=0}$ is a good indicator of which loops to pair in ER compensation:

- If $\left| g_{j^s i^c} / g_{i^c i^c} \right| << 1$, then the effect of interaction between the saturating loop j^s and the correcting loop i^c is small. To use such a loop for compensation would require large gains in $\varepsilon_{i^c j^s}(s)$ and the resulting large changes in the control action and correcting output. Therefore, it would be preferable to set $\delta_{i^c j^s} = 0$ and use a loop with a larger interaction ratio as the correcting loop.
- $\varepsilon_{i^c j^s}(s)$ should convert the units of e_{j^s} to those of e_{i^c} to account for scaling differences between loops, the scaling term, from considering the interaction equations (4-14,4-15 and 4-17), is $g_{i^c i^c} / g_{j^s i^c}$.

This information should be used to modify the ER matrix, $\boldsymbol{\delta}$, at the first step mentioned in section 4.3.1, equation (4-3), to reduce the number of non-zero terms to the minimum required.

- From equation (4-17), to compensate directional sense differences:

$$\operatorname{sgn}\left(\varepsilon_{i^c j^s}\right) = \operatorname{sgn}\left(g_{j^s i^c} k_{i^c i^c}\right) \tag{4-27}$$

- Experimentation has shown that a simple gain or in systems with large dead-times, a first-order system is sufficient for effective ER compensation. Where possible, the steady-state value of $\left(g_{i^s j^c} k_{j^c j^c}\right)^{-1}$ from equation (4-17) is a good initial value from which tuning to improve transient performance can take place.
- The gain of $\varepsilon_{ij}(s)$ should not be too large, to avoid the introduction of limit cycles. To reduce the impact on the on non-saturating loops, $\varepsilon_{ij}(s)$ should be as small as possible, to avoid consuming all the control action available for compensation and so doing introducing limit cycles. The limiting value can be derived from equation (4-25).
- All $\varepsilon_{ij}(s)$ are to return to zero after some time t_l after returning to the linear mode, so that linear performance recovery is achieved. This implies that $\varepsilon_{ij}(s)$ should not include an integrator. In addition t_l should be comparable to the linear response time of the saturating loop y_i.

4.5 Stability Analysis

A key consideration is that the system should remain stable in the linear and nonlinear modes; and that the transition between modes should not affect stability. As already discussed, all control system components are chosen to be asymptotically stable. It is apparent that the ER state selection mechanism, equation (4-9), can result in a combination of ER correcting loops becoming active.

The stability analysis must therefore be executed for each possible combination. If each individual mode is found to be globally asymptotically stable, the ER compensated system will at least be globally Lyapunov stable. This is because even when switching between stable modes, the system can become trapped in a limit cycle, oscillating between modes. A closer examination of the stability analysis and the mathematical justification follows.

4.5.1 Modes of Operation

Since the ER state selection would fire a maximum of m rules, each rule would change one row of E. Where the number of rows, were there no off-diagonal non-zero terms, is m_0, there are:

$$S = 2^{m-m_0} \tag{4-28}$$

possible modes of operation, one corresponding to the linear mode and the rest corresponding to a particular ER function, E_s for each.

4.5.2 Restructuring the Control Loop

Given that the controller can for any particular mode be represented in the following block diagram:

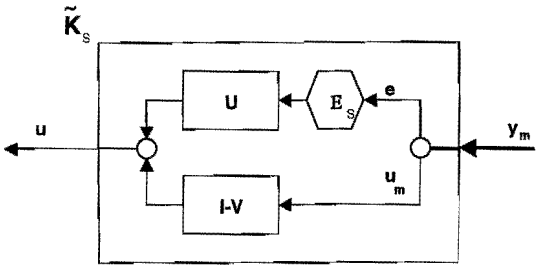


Fig.22 ER Compensation for mode S

Block-diagram manipulations are used to find an open loop model, $\mathbf{G}^S_{O/L}$, for the system in each mode. Combining the ER function with the $\mathbf{U}$ controller Fig.22 to form:

$$\mathbf{U}^S = \mathbf{U}\mathrm{E}_S \tag{4-29}$$

the controller for each nonlinear mode. An open loop system block, $\mathbf{Q}^S$ as in Fig.23 can represent each nonlinear mode.

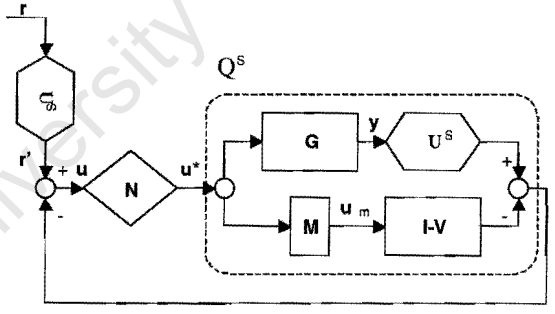


Fig.23 Re-organized feedback path for mode S

Where:

$$\begin{aligned}\mathbf{r}' &= \mathbf{U}^S\mathbf{r} \\ \mathbf{V}' &= (\mathbf{I}-\mathbf{V})\mathbf{M} \\ \mathbf{Q}^S &= \mathbf{U}^S\mathbf{G}-\mathbf{V}'\end{aligned} \tag{4-30}$$

And the output of the reorganised system is therefore: $\underline{\mathbf{y}}_{Q^S}(s) = \mathbf{Q}^S(s)\underline{\mathbf{u}}^*(s)$:

4.5.3 Stability Proof

Given that the components systems of each $\mathbf{Q}^S$ are asymptotically stable, as required by Kothare *et al.*(1994). For closed loop stability the combined system, $\mathbf{U}^S$, must be open loop stable. In addition, for a globally stable system, $\mathbf{Q}^S$, must also be open loop stable. This can be verified by considering the eigenvalues of $\mathbf{A}_{\mathbf{U}^S}$ and $\mathbf{A}_{\mathbf{Q}^S}$ or in the case where dead-times exist in the process, by multivariable Nyquist techniques for each of the S modes.

Since $\underline{\mathbf{u}}^*$ is bounded, if $\mathbf{A}_{\mathbf{Q}^S} < \underline{\mathbf{0}}$, i.e. the eigenvalues of $\mathbf{A}_{\mathbf{Q}^S}$ are found to be in the LHP, a Lyapunov function could be drawn around the maximum of the extent of the system across all operating modes. Extending the principles of circle criterion discussed by Kosut (1983): since the nonlinear element, $\mathbf{N}$, is saturation, the nonlinear output is bounded, and since $\mathbf{Q}^S$ is asymptotically stable, it's state's set would be closed hence, by the circle criterion, a positive constant, k can be found such that:

$$\left\| \mathbf{N}\left(-\underline{\mathbf{y}}_{\mathbf{Q}^S}(j\omega)\right) \right\| \leq k \left\| \underline{\mathbf{y}}_{\mathbf{Q}^S}(j\omega) \right\| \tag{4-31}$$

Then if

$$k\left\| \underline{\mathbf{y}}_{\mathbf{Q}^S}(j\omega) \right\| < 1, \quad \text{for } 0 \leq \omega \leq \infty \tag{4-31}$$

This is sufficient for the closed-loop system to be globally asymptotic stable.

In addition, a more tedious stability analysis can carried out by finding the closed loop system, $\mathbf{H}^S$, in terms of $\mathbf{Q}^S$:

$$\mathbf{H}^S = \left[\begin{array}{c|c} \mathbf{A}_{\mathbf{Q}^S} - \mathbf{B}_{\mathbf{Q}^S}\left(1+\mathbf{D}_{\mathbf{Q}^S}\right)^{-1}\mathbf{C}_{\mathbf{Q}^S} & \mathbf{B}_{\mathbf{Q}^S} - \mathbf{B}_{\mathbf{Q}^S}\left(1+\mathbf{D}_{\mathbf{Q}^S}\right)^{-1}\mathbf{D}_{\mathbf{Q}^S} \\ \hline \left(1+\mathbf{D}_{\mathbf{Q}^S}\right)^{-1}\mathbf{C}_{\mathbf{Q}^S} & \left(1+\mathbf{D}_{\mathbf{Q}^S}\right)^{-1}\mathbf{D}_{\mathbf{Q}^S} \end{array}\right] \tag{4-32}$$

Again, if $\mathbf{A}_{\mathbf{H}^S}$ is negative definite, asymptotic stability would be ensured. The Lyapunov consideration, to encompass all states, continues as before. For completeness, a brief

statement of Lyaponuv's second method is presented, followed by a suggested means of finding an encompassing Lyapunov function.

4.5.3.1 Lyaponuv's Second Method [Barnett, Section 5.4]:

Where a system is defined by:

$$\dot{\underline{\mathbf{x}}} = f(\underline{\mathbf{x}}) \tag{4-33}$$

If a Lyaponuv function, $V(\underline{\mathbf{x}})$, exists such that:

- $V(\underline{\mathbf{x}})$ and all $\partial V / \partial x_i$ are continuous,
- $V(\underline{\mathbf{x}})$ is positive definite and
- $\dot{V}(\underline{\mathbf{x}})$ is negative semi-definite.

Then:

1. The origin of equation (4-33) is stable.
2. The origin of equation (4-33) is asymptotically stable if, in addition, the derivative is negative definite.

It is important to note that the Lyapnuov method is not necessary, i.e. if a Lyapunov function say $V_1(\underline{\mathbf{x}})$ is chosen and fails the tests, this is not proof of instability as $V_2(\underline{\mathbf{x}})$ may meet the stability requirements.

In application to a linear model,

$$\dot{\underline{\mathbf{x}}} = \mathbf{A}\underline{\mathbf{x}} \tag{4-34}$$

a quadratic form of a Lyapunov function is assumed as:

$$V(\underline{\mathbf{x}}) = \underline{\mathbf{x}}^T \mathbf{P}_L \underline{\mathbf{x}} \tag{4-35}$$

Where $\mathbf{P}_L$, which is usually referred to as $\mathbf{P}$ but is not so as not to be confused with the prioritisation vector, is symmetric positive definite. Considering the derivative of $V(\underline{\mathbf{x}})$:

$$\begin{aligned}\dot{V}(\underline{\mathbf{x}}) &= \underline{\mathbf{x}}^T\left(\mathbf{A}^T\mathbf{P}_L + \mathbf{P}_L\mathbf{A}\right)\underline{\mathbf{x}} \\ -\mathbf{Q} &= \mathbf{A}^T\mathbf{P}_L + \mathbf{P}_L\mathbf{A}\end{aligned} \tag{4-36}$$

Hence, if and only if $\mathbf{Q}$ is a real symmetric positive definite matrix and $\mathbf{P}_L$ is also positive definite, the real matrix $\mathbf{A}$ produces a stable system.

4.5.3.2 Finding an all-encompassing Lyapunov Function

Confirming that for each possible ER compensation combination, the resulting systems, $\mathbf{Q}^{S_1}, \mathbf{Q}^{S_2}, ..., \mathbf{Q}^{S_n}$, are asymptotically stable, let $\mathbf{Q}$ in equation (4-36) be:

$$\mathbf{Q} = \mathbf{A}_{\mathbf{Q}^S}^T\mathbf{A}_{\mathbf{Q}^S} \quad \text{or} \quad \mathbf{Q} = \mathbf{A}_{\mathbf{H}^S}^T\mathbf{A}_{\mathbf{H}^S} \tag{4-37}$$

Solve for $\mathbf{P}_L$ in equation (4-36) (using MATLAB $\mathbf{P}_L$=lyap(A,Q) function). If $\mathbf{P}_L$ is found to be positive definite, equation (4-37) is a valid Lyapunov function. One can be found for each mode, since we know the mode to be stable. A Lyapunov function $V_S(\underline{\mathbf{x}})$ (with the largest coefficients) can be evaluated for some constant k such that this one encompasses the Lyapunov functions for all other modes.

Hence, it is sufficient that if the separate modes are asymptotically stable, the system as a whole will be Lyapunov stable. The significance of this is that typically, Lyapunov stability criteria tend to produce very conservative stability limits. Since it is not necessary to actually find the Lyapunov function in each case a more accurate stability is possible.

A discrete time form of the Lyapunov function can be used to produce the same result for a discrete time state-space form. On a practical note, continuous time systems are often used in the process control industry to model and design the control strategy; but are implemented on digital computers, and are in fact then discrete implementations. When dealing with the nonlinear effects the sample time of the "continuous system" has to be far greater (by at least a factor of ten) than under normal conditions. If the effort was made to

transform the model to the discrete time, the stability ranges would be far closer to realistic gains indicated in Section 4.4.2.2.

4.6 Implementing ER compensation

A system with suitable AWBT compensation, can have ER compensation applied. The parameters and their selection, in terms of the MVT example are presented here as a summary and application of the theoretical discussion above.

4.6.1 Prioritisation

The objective of ER compensation is to improve the performance of high priority loops during periods of nonlinear operation. To accomplish this each loop is given a priority number, 1 being the highest. For the MVT example, the first loop will be given the higher priority:

$$\mathbf{P} = \begin{bmatrix} 1 & 2 \end{bmatrix}^T \tag{4-38}$$

Equations (4-2 through 4-5) then produce the ER distribution matrix.

$$\delta = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \tag{4-39}$$

The off diagonal terms indicate which loops should be used to compensate for the higher priority loops. These terms will be considered for ER functions.

For evaluation purposes, performance indices with weighting functions can be used. A suggested weighting function is given in equations 4-6 and 4-7. This allows fair consideration of the systems performance including the priorities given to the different loops. However, for ease of comparison, no weighting will be used in this example.

4.6.2 Nonlinear Sensitivity Parameters

The nonlinear sensitivity parameters, ε_e and ε_u, as defined in section 4.3.2, allow the designer to specify tolerances after which ER compensation becomes effective. These

values should be in terms of % error or the units of the output, in which case a vector form would be used.

For the MVT example, a value of $\varepsilon_e = 0.1$ will result in a tolerance of less than 1% of the full range and will be used in our example. In addition, ε_u will be taken as $\varepsilon_u = 0.001$, the more effect the AWBT strategy, the smaller this number, which is used to sense that the control action is saturating, and therefore that the process is in a nonlinear mode.

4.6.3 ER Function Matrix

The structure of the ER Function matrix is derived from the ER distribution matrix. For the MVT example, from equations 4-39 and 4-11, the form is:

$$\mathbf{ER} = \begin{bmatrix} 1.0 & 0.0 \\ \varepsilon_{21} & 1.0 \end{bmatrix} \tag{4-40}$$

Only one ER function is required. To design this function, ε_{21}, consider the guidelines in Section 4.4.3:

- The term ε_{21}'s function is to redistribute the errors in loop y_1 to loop y_2.
- Considering $\mathbf{G}_0$:
  - $\left|g_{12}/g_{22}\right| = \left|1.97/2.05\right| \approx 0.96$ which is ≈ 1, thus indicating that changes in the input u_2, through changes in the effective error for y_2, are capable of producing comparable changes in y_1.
  - To convert from units of e_1 to those of e_2: $g_{22}/g_{12} \approx 1\left[{}^{o}C/{}_{o}C\right]$
- $\text{sgn}(g_{12}k_{22}) = +1$
- The upper limit for ε_{21} is found from considering equations (4-19 through 4-26). The estimated full ranges, from Fig.7, are:

$$\begin{aligned} U_2 &= \left|\overline{u}_2 - \underline{u}_2\right| \approx \left|6.0 - (-4.0)\right| \approx 10 \\ Y_1 &= \left|\overline{y}_1 - \underline{y}_1\right| \approx \left|20 - (-14)\right| \approx 34 \\ Y_2 &= \left|\overline{y}_2 - \underline{y}_2\right| \approx \left|18 - (-12)\right| \approx 30 \end{aligned} \tag{4-41}$$

From the composite controller:

$$K_{22} = \lim_{s \to \infty} sk_{22}(s) = 0.25 \tag{4-42}$$

Forming the linear matrix equation from (4-25):

$$U_2 \geq \left| K_{22} \varepsilon_{21} e_1 \right| \tag{4-43}$$

The range of e_1 is:

$$\begin{aligned} \varepsilon_e \leq e_1 \leq Y \\ 0.1 \leq e_1 \leq 34 \end{aligned} \tag{4-44}$$

Which from equation (4-43) implies a maximum value for ε_{21} of ranging from 1.17 to 400. Thus the suggested limit for the e_1, taken at $Y_1/3$, so as not to be too conservative, gives the follow ranges for e_1 and ε_{21}:

$$\begin{aligned} 0.1 \leq e_1 \leq 11.33 \\ 0 \leq \varepsilon_{21} \leq 3.53 \end{aligned} \tag{4-45}$$

The following engineering diagram represented the region of consideration for ε_{21}. It is formed by considering the limited amount of control action and the stability limit.

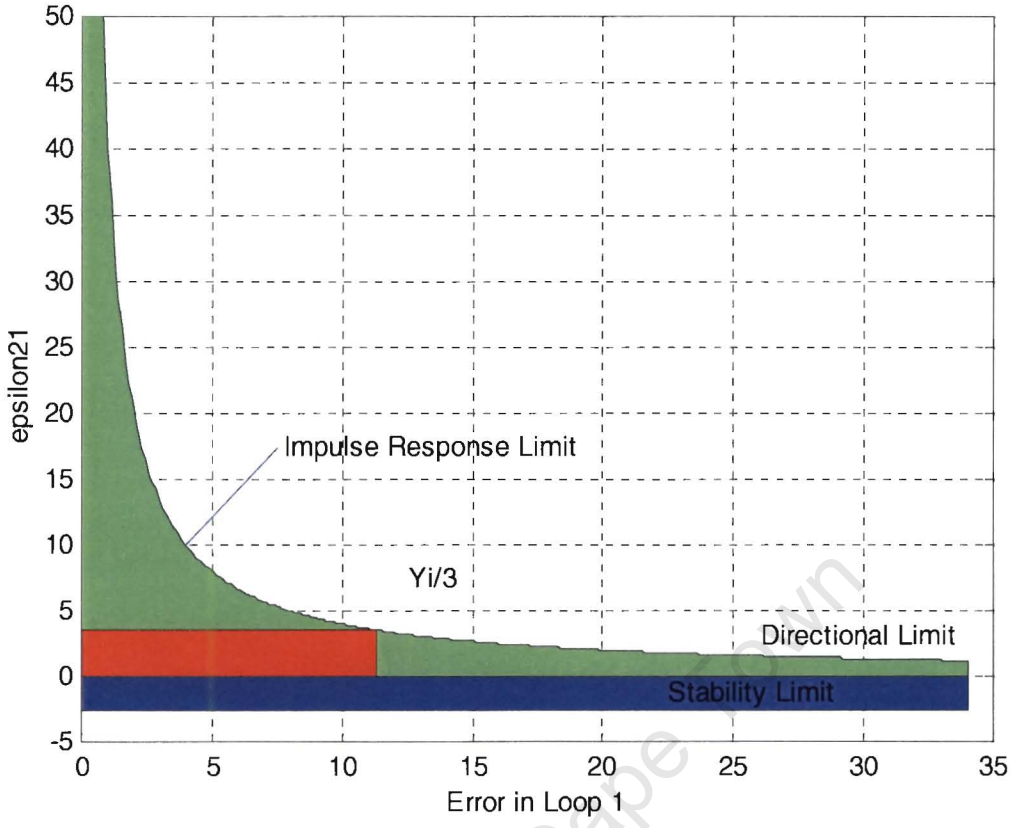


Fig.24 ER Function Design Diagram

Given these guidelines, it is expected that y_1 could be suitably compensated by redistributing its error to y_2. The ratio $\left|g_{12}/g_{22}\right|$ been close to unity, large values (>1) are expected for the ER function. From the design diagram, where the $Y_1/3$ point is $\varepsilon_{21} \approx 3.529$, the value for ε_{21} is taken as $\varepsilon_{21} = 3.5$. The resulting ER function matrix is thus:

$$\mathbf{ER} = \begin{bmatrix} 1.0 & 0.0 \\ 3.5 & 1.0 \end{bmatrix} \tag{4-46}$$

4.6.4 Performance Evaluation

The performance indices are given in Fig.20. Of particular interest is that the performance indices for the first loop have been reduced while the second loop's performance deteriorated. This is the desired effect of ER, to improve the performance of the higher

priority loop. The ITAO, which is based on the combined loop performance, has in this unweighted example in fact improved over the AWBT compensated case. This will not always be the case, but does show that ER can improve the overall system performance in this case even if loops were equally important.

Table 11 Resultant Percentage improvements

Loop/Index	SAT→AWBT	AWBT→ER
1/ISE	464%	199%
2/ISE	883%	41%
1/ITAE	371%	370%
2/ITAE	2253%	10%
ITAO	395%	122%

The percentage improvements listed in Table 11 show that the first loop, in terms of ISE, experiences an improvement of 364% by the addition of AWBT, and another 99% performance improvement over the AWBT compensated system by the inclusion of ER compensation. In all cases, as expected, there is a degradation of the second loop's performance. Again the ITAO, shows an overall improvement of 22% in the ER compensated case.

The sub-sections that follow demonstrate the effect of varying ER based parameters on the performance indices. The key points are summarised here:

The value of the ER function has the greatest overall effect on the performance indices with larger gains producing smaller performance numbers. For this example:

Smaller ε_{21} values resulted in:

- Limited jitter in the control action, and
- Greater steady state error.

Larger ε_{21} values result in:

- increased jitter in the control action,
- the improvement in the higher priority loop increases at a lower rate than the degradation of the lower priority loops, and
- the nonlinear sensitivity parameters, in particular ε_e, must be chosen with care so as not to have an exaggerated effect during nonlinear transients in the case of realisable operating points.

The effect of the ER Hold was to reduce the jitter of the control signals while still delivering significant performance improvements.

4.6.4.1 Varying of Nonlinear Sensitivity Parameters

To illustrate the impact of these parameters, the following tables present the performance total loop ISE indices for operating point 1 (OP1) which exists in the realisable region and operating point 3 (OP3) which is outside of the realisable region. The OP1 case produces nonlinear transients, i.e. the control action saturates momentarily, but the system is able to realise the operating point.

$$\underline{\mathbf{r}}^1(t)=\begin{cases} r_1=0 & t<10 \\ r_1=3 & t\geq 10 \\ r_2=1 & t<160 \\ r_2=2 & t\geq 160 \end{cases} \text{ and } \underline{\mathbf{r}}^3(t)=\begin{cases} r_1=0 & t<10 \\ r_1=5 & t\geq 10 \\ r_2=1 & t<160 \\ r_2=4 & t\geq 160 \end{cases} \tag{4-47}$$

As indicated, when the AW compensation is well tuned, ε_u must be small so that the system can detect that the control actions are saturating. In the case where instruments can feedback this information, a direct check of whether u_i is at a limit can be made. The difference technique to sense saturation allows for the compensation of drift and failure.

Table 12 ISE for varying Nonlinear Sensitivity Parameters

OP = 1, y1

ε_e						
	0.001	1.0198	1.0198	1.0198	0.9955	0.9955
	0.010	1.0198	1.0198	1.0198	0.9955	0.9955
	0.100	1.0198	1.0198	1.0198	1.0201	1.0201
		1	**0.1**	**0.01**	**0.001**	**0.0001**
AWBT=1.0198		ε_u				

OP = 3, y1

ε_e						
	0.001	2.0862	1.2733	1.0537	1.0467	1.0466
	0.010	2.0862	1.2733	1.0537	1.0467	1.0467
	0.100	2.0862	1.2733	1.0540	1.0472	1.0471
		1	**0.1**	**0.01**	**0.001**	**0.0001**
AWBT=2.0862		ε_u				

OP = 1, y2

ε_e						
	0.001	1.0198	1.0198	1.0198	1.4838	1.4842
	0.010	1.0198	1.0198	1.0198	1.4845	1.4744
	0.100	1.0198	1.0198	1.0198	1.0198	1.0198
		1	**0.1**	**0.01**	**0.001**	**0.0001**
AWBT=1.0198		ε_u				

OP = 3, y2

ε_e						
	0.001	0.9614	1.5416	2.2824	2.3262	2.3271
	0.010	0.9614	1.5416	2.2825	2.3161	2.3269
	0.100	0.9614	1.5416	2.2778	2.3238	2.3250
		1	**0.1**	**0.01**	**0.001**	**0.0001**
AWBT=0.9614		ε_u				

Once saturation has been detected, ε_e specifies the size of error for which ER compensation will be applied. When the requested operating point is outside of the

realisable region, ε_u dominates, but in the case of realisable operating points that have nonlinear transients, ε_e can be used to de-sensitise the compensator so as not to interfere with the linear performance.

Table 13 ITAE for varying Nonlinear Sensitivity Parameters

OP = 1, y1

ε_e						
	0.001	1.0710	1.0710	1.0710	0.7159	0.7163
	0.010	1.0710	1.0710	1.0710	0.7159	0.7160
	0.100	1.0710	1.0710	1.0710	1.0734	1.0738
AWBT=1.0710		**1**	**0.1**	**0.01**	**0.001**	**0.0001**
		ε_u				

OP = 3, y1

ε_e						
	0.001	10.387	5.8678	2.8438	2.8025	2.8022
	0.010	10.387	5.8678	2.8436	2.8027	2.8025
	0.100	10.387	5.8678	2.8455	2.8066	2.8057
AWBT=10.387		**1**	**0.1**	**0.01**	**0.001**	**0.0001**
		ε_u				

OP = 1, y2

ε_e						
	0.001	1.1008	1.1008	1.1008	2.9918	2.9910
	0.010	1.1008	1.1008	1.1008	2.9923	2.9927
	0.100	1.1008	1.1008	1.1008	1.1017	1.1020
AWBT=1.1008		**1**	**0.1**	**0.01**	**0.001**	**0.0001**
		ε_u				

OP = 3, y2

ε_e						
	0.001	1.0396	6.5174	10.212	10.273	10.274
	0.010	1.0396	6.5174	10.212	10.273	10.273
	0.100	1.0396	6.5174	10.208	10.267	10.269
AWBT=1.0396		**1**	**0.1**	**0.01**	**0.001**	**0.0001**
		ε_u				

4.6.4.2 Varying of ER function

Larger values of the ER function will help to reduce the steady state error, having this closer to the desired setpoint. However, as discussed, there are a number of factors that act to limit the ER functions gain. Performance indices for a variation of ER function values are given in Table 14. The response trends for these simulations are given in Appendix C.

4.6.4.3 Varying of ER Hold Time

ER compensation should act at every loop iteration. However, this can introduce a large amount of jitter onto the control action – in particular when working with continuous systems. By freezing the effective ER matrix for a minimum time period, which models the quantatisation that takes place in implementing a control system on a digital platform, a less active control signal is obtained while still performing well in terms of performance optimisation. Table 15 has the results of the MVT example using a 1 second hold on the effective ER matrix. The response trends for this simulation is given in Appendix C.

Table 14 Performance Indices for Varied ER Function

$\varepsilon_{21} = 1.20$

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.17	1.11	4.81	6.93	4.25
2	0.01	0.55	0.12	0.71	0.33
Loop Totals	1.18	1.66	4.93	7.64	
Overall	2.84		13.1		4.58

$\varepsilon_{21} = 3.5 - 10\% = 3.15$

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.05	1.85	2.85	9.48	4.72
2	0.01	0.42	0.09	0.62	0.29
Loop Totals	1.06	2.27	3.94	10.1	
Overall	3.32		14.0		5.01

$\varepsilon_{21} = 3.5$

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.04	1.92	2.72	9.66	4.78
2	0.01	0.41	0.09	0.61	0.28
Loop Totals	1.05	2.33	2.81	10.3	
Overall	3.38		13.1		5.06

$\varepsilon_{21} = 3.5 + 10\% = 3.85$

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.04	1.96	2.56	9.84	4.84
2	0.00	0.40	0.09	0.60	0.28
Loop Totals	1.04	2.36	2.67	10.4	
Overall	3.40		13.1		5.12

$\varepsilon_{21} = 10.0$

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.01	2.46	1.51	11.3	5.34
2	0.00	0.33	0.08	0.5	0.26
Loop Totals	1.01	2.79	1.59	11.8	
Overall	3.81		13.4		5.60

Table 15 ER on MVT with 1 sec hold Performance Indices

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.333	1.473	3.777	8.634	4.670
2	0.010	0.491	0.153	0.706	0.330
Totals	1.343	1.964	3.930	9.340	
Overall	3.307		13.27		5.000

4.6.5 Practical Realisation

A key feature of any control strategy is the ease with which it can be implemented. Theory aside, the algorithm for this example can be implemented with the psuedo code in Table 16.

Table 16 AWBT and ER compensation algorithm for MVT

$\Delta\underline{\mathbf{u}} = \underline{\mathbf{u}}^* - \underline{\mathbf{u}}$	Find the difference between the controller output and the process input
$\underline{\mathbf{e}} = \underline{\mathbf{r}} - \underline{\mathbf{y}} + \mathbf{X}_{CAW} \cdot \Delta\underline{\mathbf{u}}$	Apply CAW compensation
if	If...
$\lvert\Delta u_1\rvert > \varepsilon_u$ and	The high priority loop's control action is saturating...
$\lvert\Delta u_2\rvert < \varepsilon_u$ and	And the correcting loop is not saturating...
$\lvert\underline{e}_1\rvert > \varepsilon_e$ then	And there is an error in the high priority loop...
$e_2 = e_2 + \varepsilon_{21} e_1$	Then redistribute the error in the saturating high priority loop to the correcting loop according to the ER function
end if	End If

In Section 5.4, a practical implementation on the PlantStar industrial control platform for the laboratory MVT process will be presented.

4.7 ER Application Summary

ER compensation has been formulated to optimise prioritised outputs during periods of actuator saturation. It is applicable to MIMO processes where the process and controller are open loop stable and the interaction between process inputs and outputs is sufficient to allow inputs other than the dominant one to effect a change on an output.

The application of ER compensation to a control system follows the following design steps:

1. Design the linear controller.

   This should be designed to meet the linear performance requirements for the closed loop system.
2. Apply a suitable AWBT technique. (Chapter 3)

   This may require the iteration through a number of possible compensation strategies and its tuning parameters before acceptable performance is obtained. By acceptable performance it is meant that the controller states are constrained to a state-space that produces realizable outputs. Or, more likely, does not excessively overshoot the realizable values for any actuator.

3. Consider prioritization of outputs. (Section 4.3.1)

   It is clear that any number of prioritisations can be given to process outputs and an ER function matrix developed for each. This would allow the process operator to choose an ER optimisation strategy based on external influences.
4. Use the prioritisation vector to create the ER matrix. (Section 4.3.1)
5. Consider the nonlinear sensitivity parameters. (Section 4.3.2)
6. Design the ER Function Matrix:
   a. work through the ER Guidelines to develop each ER Function (Section 4.3.3). These will highlight the suitability of applying ER compensation for a particular ER Function.
   b. Given the initial ER function values, consider the S resulting effective ER function matrices and therefore modes of operation (Section 4.5.1).
   c. Use the restructured control loop in Section 4.5.2, and confirm that each ER mode is stable. Based on the Stability Proof discussed in Section 4.5.3, this is a sufficient condition for the overall performance to Lyapunov stable.
7. Simulate the compensated system and consider performance indices (Section 2.1) with some weighting in terms of the allocated prioritisation, (Section 4.3.1).

Based on the resulting change in performance steps 6 and 7 may need a number of iterations.

4.8 Discussion on Error Redistribution

ER compensation has, in the MVT process case study, been able to optimise high priority loops by the successful redistribution of error in these loops to lower priority loops. By exploiting the degrees of freedom available through the process interaction terms, an optimal nonlinear operating point is obtained.

In the case of poorly chosen nonlinear sensitivity parameters, ER can have a slight degradation effect on systems that experience transient saturation.

Since ER is formulated to be independent of the nature of the AWBT compensation, but does require that the controller states are contained by suitable AW compensation; it can be applied with all forms of AWBT that can be represented in Kothare's Unified Framework. This framework supplies the infrastructure which this thesis expands to develop the Lyapunov stability test for ER compensated systems.

A key feature of ER compensation is that it is an online feedback strategy as opposed to a back calculation or off-line optimisation strategy. Since ER compensation is driven by the process error and the detection of the nonlinear mode, it does not require a well-defined knowledge of the realisable control region or realisable operating region. In particular, detecting the nonlinear modes through differences between the requested control action and the realised (measured) process inputs will extend the optimisation effects of AWBT and ER compensation to actuator failure and process drift. ER can also be used where soft-limits are required to maintain the controller/process integrity or in the implementation of safety-loops.

Another key issue from a practical point of view is the ease with which this compensation can be implemented within a control system.

5 Applications and Case Studies

This chapter presents four case studies of the application of ER compensation. The first is a continuation of the MVT simulated example where the application of ER with different AW techniques and the AN technique is considered. The next two studies are carried out on simulations of real world processes. The first is 2x2 Distillation Column with dead times. The second is a 3x3 Milling Circuit with largely varying dead times which adversely affect the systems performance. All simulations are done in MATLAB, using the ER Toolbox that was developed through the course of this work, see Appendix A.

The fourth case study will be the application of ER to the real multivariable thermal process on the industrial control platform PlantStar.

5.1 Simulated Multivariable Thermal Process

The implementation of ER had the desired effect of optimising the process outputs during nonlinear operation. This section considers other possible combinations of the AW and ER and the application of HCT-AN optimisation to this example. Changes to the realisable control region are also considered.

Where only the performance indices are presented here, the full results are given in Appendix C.

5.1.1 Combinations of AW and ER Compensation

Compared to the respective AW compensated responses in Section 3.3, there is significant performance improvement and, as desired, the first loops perfromance is the most affected. Note that the performance indices are not weighted in Table 17 and Table 18.

Table 17 HCT-ER Performance Indices on MVT Process

ER hold = 0 seconds

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.575	0.557	6.640	5.245	4.454
2	0.030	0.696	0.207	0.827	0.391
Totals	1.605	1.253	6.847	6.072	
Overall	2.858		12.92		4.845

ER hold = 1 second

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.331	1.478	3.821	8.853	4.677
2	0.011	0.507	0.158	0.732	0.341
Totals	1.342	1.985	3.439	9.585	
Overall	3.327		13.02		5.018

Table 18 MAW-ER Performance Indices on MVT Process

ER hold = 0 seconds

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.144	1.229	4.471	7.555	4.348
2	0.008	0.453	0.096	0.668	0.312
Totals	1.152	1.682	4.567	8.223	
Overall	2.834		12.79		4.660

ER hold = 1 second

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.120	2.070	2.646	9.682	4.806
2	0.005	0.415	0.202	0.557	0.292
Totals	1.125	2.485	2.848	10.24	
Overall	3.610		13.09		5.098

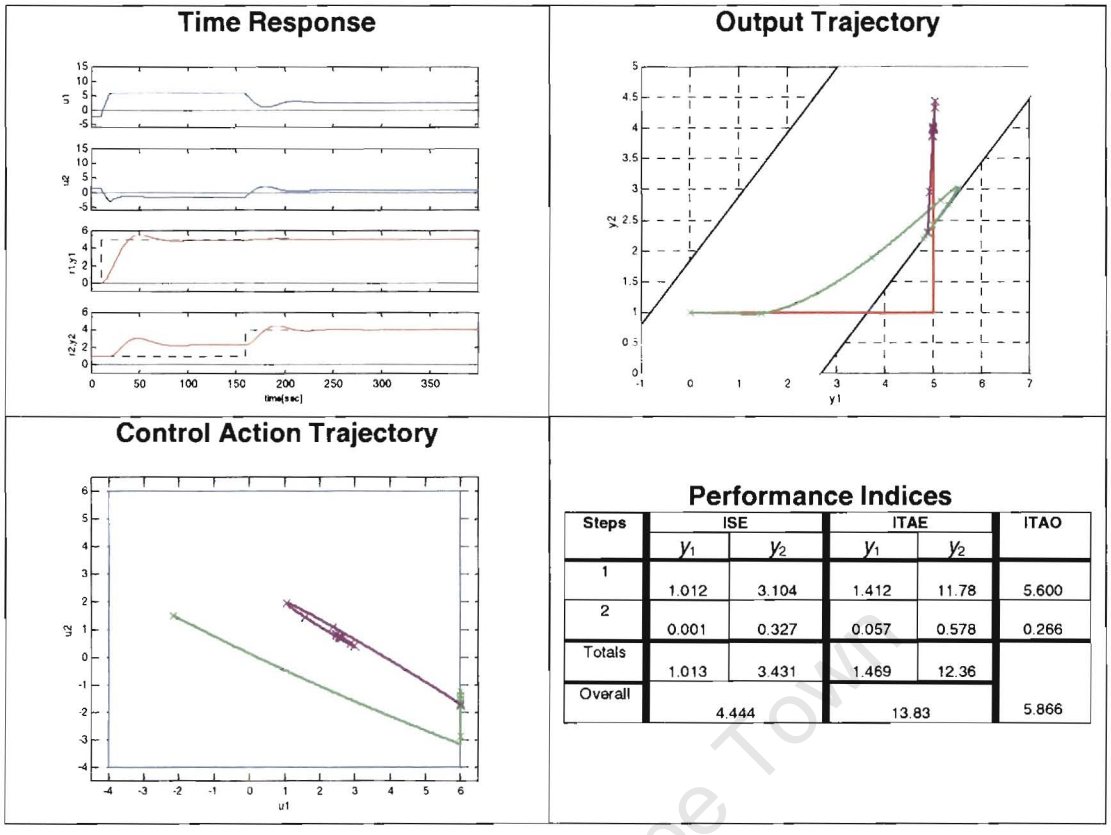
For HCT-ER, Table 17, the improvement in ISE for the first loop, y_1, produces indices 1/5.6 and 1/6.7 of those of the HCT AW only compensated systems, Section 3.3.2 , for ER hold being 0 seconds and 1 second respectively. Similar comparisons for the MAW-ER case, Table 18, indicate improvement of the order of 1/2.4 times that of the MAW AW comepensated system, Section 3.3.3.

The ITAE improvements for HCT-ER y_1 are 1/3.9 and 1/8.1 of the HCT system, for ER hold of 0 and 1 seconds. In this case y_2 was also improved with reductions of 1/3.9 and 1/2.5 over the HCT compensated system. For MAW-ER the improvement is 1/2.8 and 1/4.6 for ER hold of 0 and 1 seconds for y_1 with only a reducation by 1.8 and 2.7 times the performance in y_2. These performance numbers indicate that the steady state error for y_1 during the nonlinear mode has been reduced. Interestlying in these cases the nett effect of improvements on y_1 and the improvements or degradation of y_2 has been to improve the overall perfromance and the ITAO numbers are seen to be reduced.

In Section 3.3.4, it was evident that AN has the potential to produce similar results. Fig.25 shows the results of the application of HCT-AN with weighting of:

$$\Lambda = \begin{bmatrix} 10 & 0 \\ 0 & 1 \end{bmatrix} \tag{5-3}$$

Even though the performance of HCT, in Section 3.3.2, performed poorly, the combination of HCT and AN, Section 3.3.4, is most effective in optimising the operating point. The performance indices are given in Fig.25. A concern with the implementation of the AN in equation (3-32) is that if the control region was to vary with external operating conditions it would revert to the performance of the HCT compensated system, see Section 5.1.2.



Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.012	3.104	1.412	11.78	5.600
2	0.001	0.327	0.057	0.578	0.266
Totals	1.013	3.431	1.469	12.36	
Overall	4.444		13.83		5.866

Fig.25 Optimised HCT-AN Compensated Results

All the performance indices quoted in earlier chapters and in this section have not included the priority weighting discussed in Section 4.3.1 and the totals for ISE and ITAE indices between the MAW-ER and HCT-ER examples are reasonably similar. However, to easily compare these totals in an optimisation sense, the prioritisation weighting, from equation (4-7), of $\underline{\mathbf{w}} = \begin{bmatrix} 1 & 1/2 \end{bmatrix}^T$ will be applied producing Table 19 and Table 20 for ER hold = 0 seconds.

Table 19 MVT Process Weighed ISE Performance Index Totals

HCT-ER	MAW-ER	CAW-ER	HCT-AN
2.232	1.993	2.302	2.729

Table 20 MVT Process Weighed ITAE Performance Index Totals

HCT-ER	MAW-ER	CAW-ER	HCT-AN
9.883	8.679	7.962	7.649

The weighted performance indices totals in Table 19 and Table 20, show that though the transients performance, ISE Performance Indices, were worse than the HCT-ER and MAW-ER cases, the CAW-ER and HCT-AN cases have clearly performed better in an

ITAE sense, thus indicating the reduction of steady state error in y_1 during nonlinear operation.

5.1.2 Reduced Realisable Control Region

If the control region was to shrink from having an upper limit of +6 to +5, the control strategy should be able to perform robustly. Fig.26 shows the results for an ER compensated system under these conditions. Results for the saturation, various AW, HCT-ER and HCT-AN compensated systems are given in Appendix C.

For a 10% reduction of the control region in each signal, the percentage degradation from the ER results given in Section 4.2 is given in Table 21.

Table 21 Percentge Degradation with reduced Control Region for CAW-ER

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	2.3	73.7	24.2	33.5	31.1
2	0.0	-35.0	-19.6	-18.2	20.0

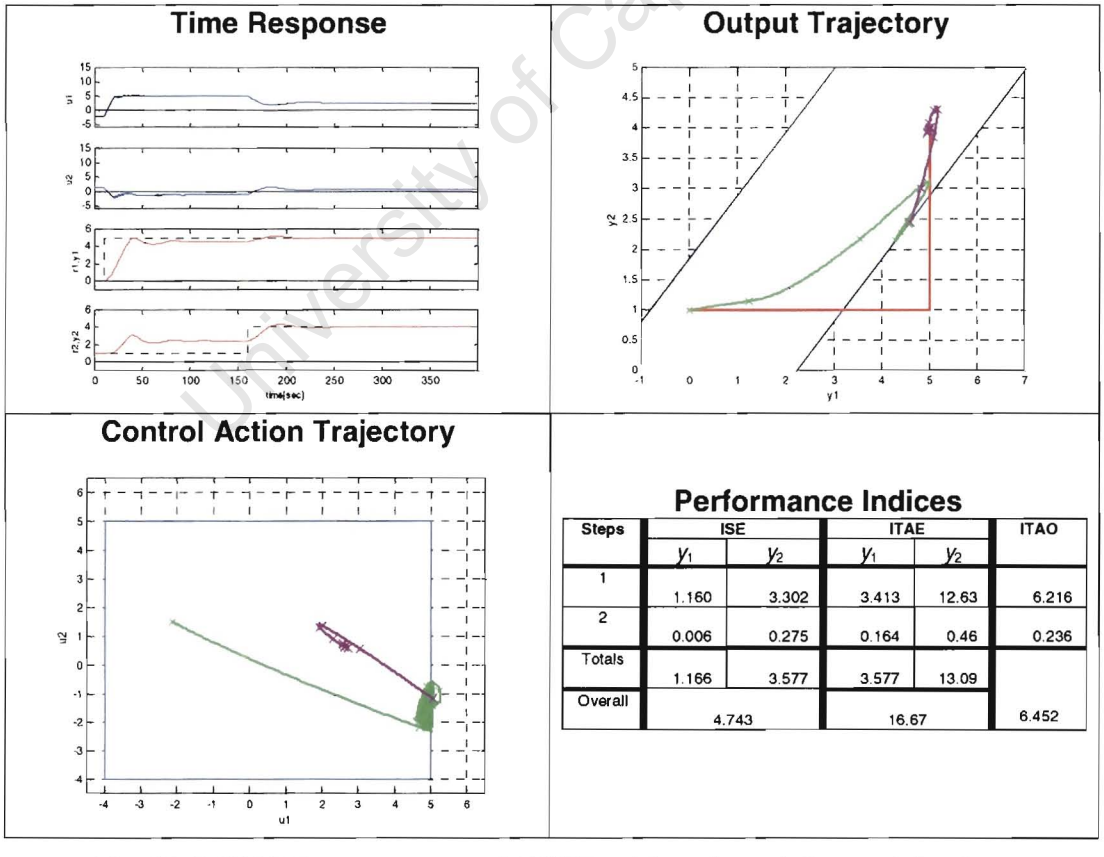


Fig.26 CAW-ER compensated MVTP with reduced Control Region

The transient response is very similar, with only a 2.3% ISE increase, but the larger steady state error increase the ITAE by 24.2%. A large increase in the performance indices for y_2 during step one indicate that the steady state error for y_2 is greater. The negative percentages indicate an improvement over the previous situation. This is because the system is operating closer to the final operating point. The performance of this system has proven to be adequate under the variation of the realisable control region.

The optimised compensators, HCT-ER, HCT-AN and CAW-ER performance indices are given in Table 22 and Table 23 for comparison.

Table 22 MVT Process with Reduced Control Region Comparitive ISE

Steps	y_1			y_2		
	CAW-ER	HCT-ER	HCT-AN	CAW-ER	HCT-ER	HCT-AN
1	1.160	1.840	14.02	3.302	0.944	17.23
2	0.006	0.054	1.06	0.275	0.591	4.60
Total	1.166	1.894	15.08	3.577	1.535	21.83

Table 23 MVT Process with Reduced Control Region Comparitive ITAE

Steps	y_1			y_2		
	CAW-ER	HCT-ER	HCT-AN	CAW-ER	HCT-ER	HCT-AN
1	3.413	8.559	35.82	12.63	6.736	28.94
2	0.164	0.214	0.94	0.46	0.700	2.00
Total	3.577	8.773	36.76	13.09	7.436	30.94

In this example, the CAW-ER has best optimised y_1 both in ISE and ITAE senses. The HCT-AN case performed the worst and both outputs have lost their setpoints. ER is more robust in this case because it is a feedback system. It is not dependent on absolute knowledge of the realizable output space (either due to the system model, or the actuator limitations).

5.2 Distillation Column

Distillation Columns are common chemical processes. The Wood and Berry Distillation Column (Wood and Berry, 1973) is a well-documented example. It is represented by the following transfer function matrix. Of particular interest is that it is a full structure 2x2 system with dead times in each term.

$$\mathbf{G}(s)=\begin{bmatrix} \dfrac{12.8}{16.7s+1}e^{-s} & \dfrac{-18.9}{21s+1}e^{-3s} \\ \dfrac{6.6}{10.9s+1}e^{-7s} & \dfrac{-19.4}{14.4s+1}e^{-3s} \end{bmatrix} \tag{5-1}$$

5.2.1 Linear Controller

An INA compensator for the system is suggested by Singh (1999). The decoupling technique has introduced non-minimal phase zeros into the controller design.

$$\mathbf{K}_d(s)=\begin{bmatrix} \dfrac{1.391s^4+1.360s^3+7.751s^2+3.612s+8.547}{1s^4+2.473s^3+5.385s^2+6.233s+6.283} & \dfrac{1.236s^2-2.858s+3.081}{1s^2+2.312s+2.49} \\ \dfrac{0.3165s^2-0.1688s+0.735}{1s^2+0.161s+2.523} & 1 \end{bmatrix} \tag{5-2}$$

Designing a SISO PI controller for each diagonalised loop:

$$\mathbf{K}_{PI}(s)=\begin{bmatrix} \dfrac{1.598s+0.18}{8.88s} & 0 \\ 0 & \dfrac{-0.324s+0.036}{9s} \end{bmatrix} \tag{5-3}$$

So that the final linear controller is given by:

$$\mathbf{K}(s)=\begin{bmatrix} \dfrac{2.223s^5+2.424s^4+12.63s^3+.169s^2+14.31s+1.538}{8.88s^5+21.96s^4+47.82s^3+55.35s^2+55.79s} & -\dfrac{0.4005s^3-0.8815s^2+0.8954s+0.1109}{9s^3+20.81s^2+22.41s} \\ \dfrac{0.5059s^3-0.2128s^2+1.144s+0.1323}{8.88s^3+1.43s^2+22.4s} & -\dfrac{0.324s+0.036}{9s} \end{bmatrix} \tag{5-4}$$

The actuator input range is limited to $-0.10<u_1<+0.31$ and $-0.10<u_1<+0.25$.

5.2.2 AW and ER Compensation

AW compensation in the form of HCT (see Section 3.1.2.3) will be used to maintain the controller states during saturation. For purposes of this example, the relative output priority is assigned as:

$$\mathbf{P}=\begin{bmatrix}1 & 2\end{bmatrix}^T \tag{5-6}$$

Thus, loop y_2 will be used to compensate for errors in loop y_1. The prioritisation vector generates the ER distribution matrix, as per equations (4-2 through 4-5):

$$\delta = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \tag{5-7}$$

Thus indicating that the some function of error, e_1, will be added to e_2. The nonlinear sensitivity parameters, that are used to "select" the mode of operation (see Section 4.3.2), are taken as:

$$\varepsilon_u = 0.0001 \text{ and } \varepsilon_e = 0.0001 \tag{5-8}$$

From the prioritisation and resulting ER matrix, the form of the ER function matrix is:

$$\mathbf{ER} = \begin{bmatrix} 1.0 & 0.0 \\ \varepsilon_{21} & 1.0 \end{bmatrix} \tag{5-9}$$

Only one ER function is required. To design this function, ε_{21}, consider the guidelines in Section 4.4.3:

- The term ε_{21}'s function is to redistribute the errors in loop y_1 to loop y_2.
- Considering $\mathbf{G}_0$:
  - $\left|g_{12}/g_{22}\right| = \left|-18.9/-19.4\right| \approx 0.97$ which is ≈ 1, thus indicating that changes in the input u_2, through changes in the effective error for y_2, are capable of producing comparable changes in y_1.
  - To convert from units of e_1 to those of e_2: $g_{22}/g_{12} \approx 1$
- $\text{sgn}(g_{12}k_{22}) = +1$
- The upper limit for ε_{21} is found from considering equations (4-19 through 4-25). The estimated full ranges are:

$$U_2 = \left|\bar{u}_2 - \underline{u}_2\right| \approx \left|0.25 - (-0.10)\right| \approx 0.35$$
$$Y_1 = \left|\bar{y}_1 - \underline{y}_1\right| \approx \left|6.0 - (-6.0)\right| \approx 12.0 \qquad (5\text{-}10)$$
$$Y_2 = \left|\bar{y}_2 - \underline{y}_2\right| \approx \left|4.0 - (-5.5)\right| \approx 9.5.0$$

From the composite controller:

$$K_{21} = \lim_{s \to \infty} sk_{21}(s) = {}^{0.1109}\!/_{22.41} = 0.0049$$
$$K_{22} = \lim_{s \to \infty} sk_{22}(s) = {}^{0.036}\!/_{9.00} = 0.0040 \qquad (5\text{-}11)$$

Forming the linear matrix equation from (4-25):

$$U_2 \geq \left|K_{21}Y_1 + K_{22}(\varepsilon_{21}Y_1 + Y_2)\right|$$
$$5.26 \geq \left|\varepsilon_{21}\right| \qquad (5\text{-}12)$$

Given these guidelines, it is expected that y_1 could be suitably compensated by redistributing its error to y_2. The ratio $\left|{}^{g_{12}}\!/_{g_{22}}\right|$ being close to unity, large values (>1) are expected for the ER function. A stability check indicates a range of $-36.36 < \varepsilon_{21} < 51.51$ produces stable systems. The initial value for ε_{21} is taken as $\varepsilon_{21} \approx 1.0$. Iterating through a range of values for ε_{21}, $1 < \varepsilon_{21} < 5$. After some simulations, a value of $\varepsilon_{21} = 5$ was chosen. The resulting ER function matrix is thus:

$$\mathbf{ER} = \begin{bmatrix} 1.0 & 0.0 \\ 5.0 & 1.0 \end{bmatrix} \qquad (5\text{-}13)$$

5.2.3 Response Set

The first set of results, Fig.27 and Table 24 through Table 26, present the ideal, saturating and HCT compensated results are supplied for comparison. For clarity, the performance index numbers have not been weighted.

Table 24 Distillation Column Performance Index ISE (Part 1)

Steps	y_1			y_2		
	IDEAL	SAT	HCT	IDEAL	SAT	HCT
1	0.999	51.21	52.99	0.474	33.87	35.03
2	0.001	64.07	64.05	0.526	78.94	78.91
Total	1.000	115.3	117.0	1.000	112.8	113.9

Table 25 Distillation Column Performance Index ITAE (Part 1)

Steps	y_1			y_2		
	IDEAL	SAT	HCT	IDEAL	SAT	HCT
1	0.728	143.3	144.7	0.513	29.83	30.12
2	0.272	133.1	133.0	0.487	37.23	37.20
Total	1.000	276.4	277.7	1.000	67.06	67.32

Table 26 Distillation Column Performance Index ITAO (Part 1)

Steps			
	IDEAL	SAT	HCT
1	0.538	42.21	42.63
2	0.462	45.48	45.45
Total	1.000	87.69	88.08

In the second set, the results for optimised system are presented in Fig.28 and Table 27 through Table 29. The first two results are for ER compensation and the third and fourth are for AN compensated systems.

The two ER results, ER 1 and ER 2, are for a zero and one second hold on the effective ER matrix. The two AN results are for a weighting of $\Lambda_{11} = 10$ and $\Lambda_{11} = 15$ respectively.

The first set of results, Fig.27 and Table 24 through Table 26, indicate poor performance of HCT where the internal windup has not allowed the return to linear performance by the end of the simulation. The compensation is active since the bounded control actions indicate that the controller states are being contained.

ER and AN systems are presented in the second part of the results,Fig.28 and Table 27 through Table 29. All these responses resemble the ideal closer than the HCT compensated system. In all cases, the controller states are better contained, and the output performance is acceptable.

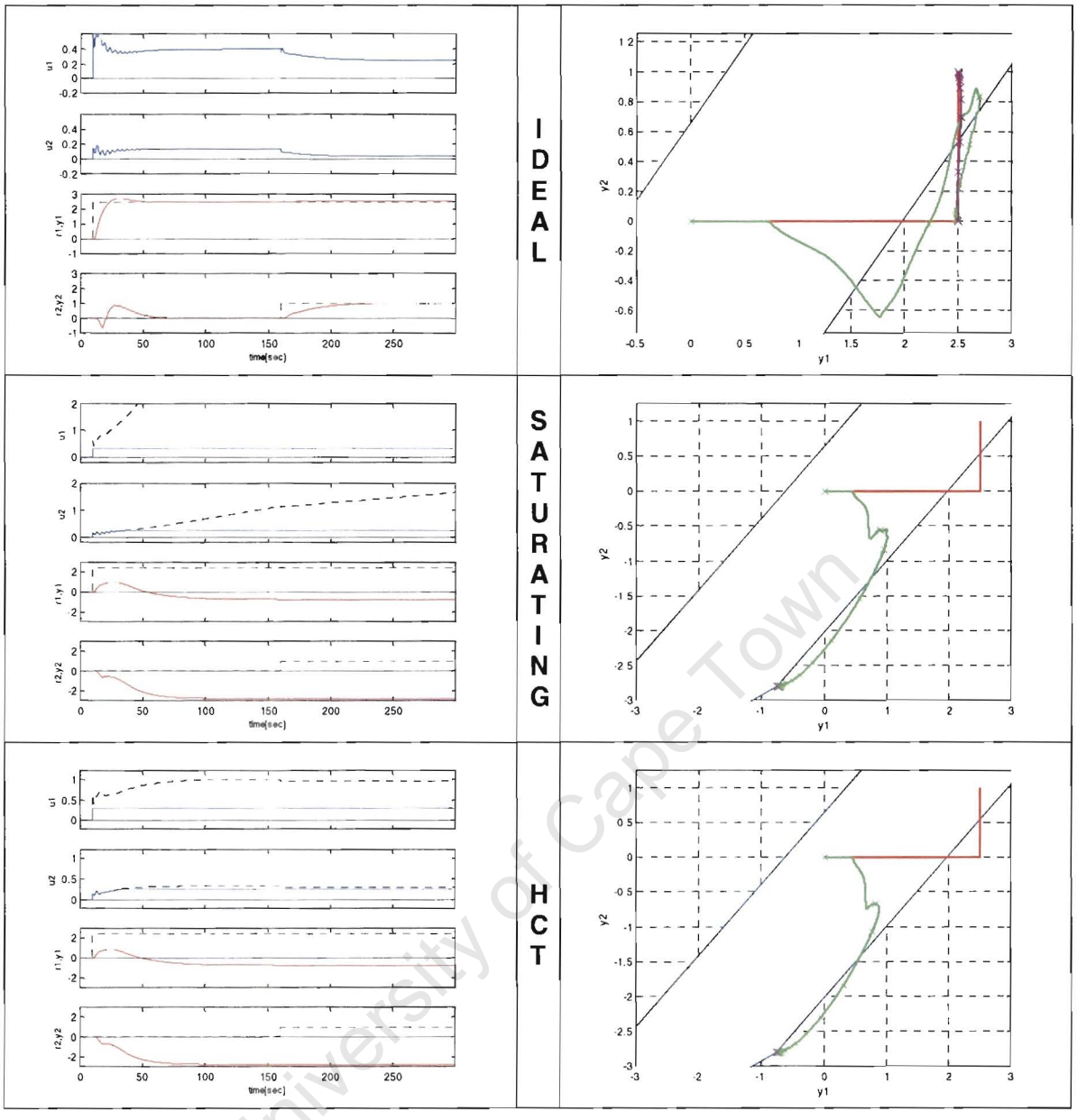


Fig.27 Distillation Column Ideal, Saturating and HCT Compensated Response

In comparing these results, the ER compensation supplies a large and rapidly changing control action to the process during the first 20-30 seconds, and thereafter settles down. The AN compensated results show reasonable control action during this period. However, there is noticeable oscillation of control action u_2 during the nonlinear operation. The increased Λ in AN 2 was an attempt to improve on the output performance of the system.

The AN systems out-performed the ER systems in the ISE index by approximately 30% in y_1, indicating that the AN systems where quicker to respond to the nonlinear situation. The performance of ER 2 in y_2 was significantly worse; this is a combination of the

relatively high ER function of $\varepsilon_{21} = 5$ and the hold time of one second resulting in a longer excursion from the setpoint for y_2. This is reflected in the higher ITAE for y_2 in ER 2 as compared to ER 1.

Table 27 Distillation Column Performance Index ISE (Part 2)

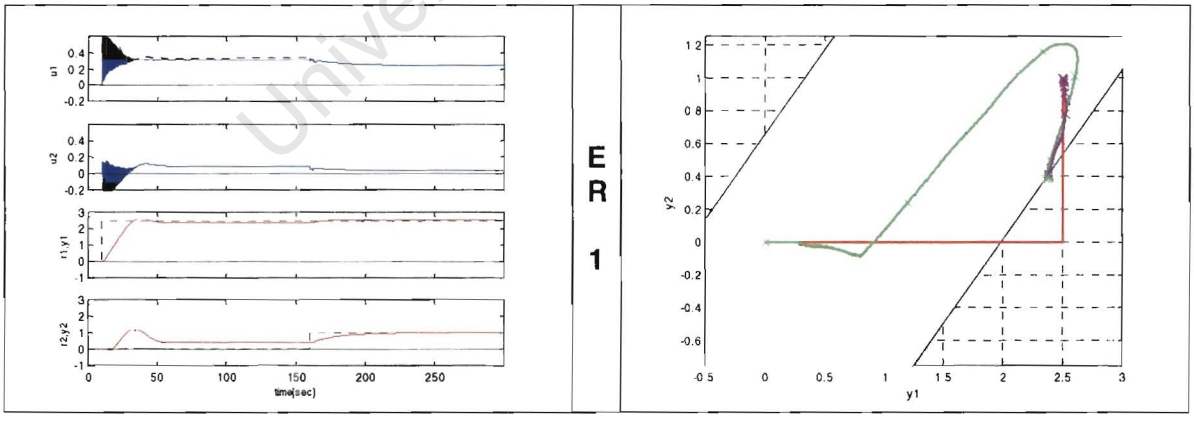
Steps	y_1				y_2			
	ER 1	ER 2	AN 1	AN 2	ER 1	ER 2	AN 1	AN 2
1	2.054	2.134	1.671	1.531	1.696	3.625	1.381	1.732
2	0.004	0.004	0.010	0.006	0.154	0.133	0.194	0.170
Total	2.058	2.138	1.681	1.537	1.850	3.758	1.575	1.902

Table 28 Distillation Column Performance Index ITAE (Part 2)

Steps	y_1				y_2			
	ER 1	ER 2	AN 1	AN 2	ER 1	ER 2	AN 1	AN 2
1	6.271	6.110	8.947	7.021	4.913	5.520	4.110	4.651
2	0.179	0.172	0.213	0.191	0.245	0.209	0.265	0.254
Total	6.450	6.282	9.160	7.212	5.158	5.729	4.375	4.905

Table 29 Distillation Column Performance Index ITAO (Part 2)

Steps	ER 1	ER 2	AN 1	AN 2
1	4.854	5.381	4.404	4.662
2	0.233	0.200	0.252	0.242
Total	5.087	5.581	4.656	4.904



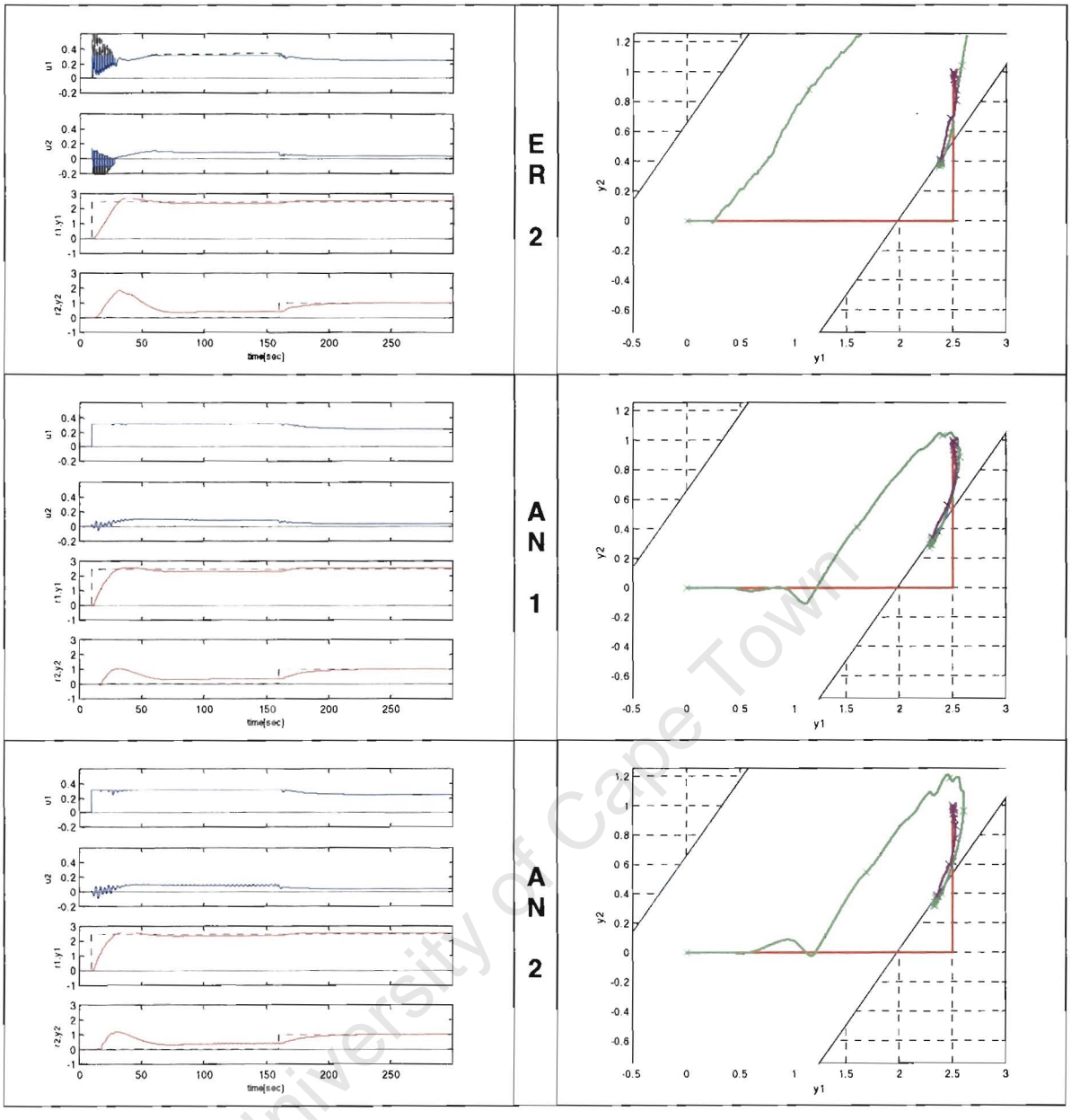


Fig.28 Distillation Column ER and AN Compensated Responses

The ITAE indices show the ER systems out-performing the AN compensated systems in y_1. This indicates that the reduction in steady state error was better. For ER 1 approximately 37% improvement over the steady state error in AN 1, and approximately 14% improvement over AN 2, was achieved.

The ITAO indicates that the best steady state performance, if all outputs where of equal significance, would have been achieved by AN 1. This is easily observed in Fig.28 were the effect on y_2 was limited, this meant that y_1's performance suffered.

The continuing increase of Λ would allow the AN system to produce similar results to the ER, but this introduced an oscillatory control action and output for most of the duration of the first step.

5.3 *Milling Circuit*

In another application to an industrial process, a gold milling circuit was used to perform a simulated case study. This example is based on a typical milling circuit which is modelled and an INA compensated control strategy is applied (Hulbert and Braae, 1981). The principal model, $\mathbf{G}(s)$, is a 3x3 transfer function matrix that describes the mapping between change in inputs, $\Delta\underline{\mathbf{u}}$, to changes in outputs, $\Delta\underline{\mathbf{y}}$, between the following pairs:

$$\Delta\underline{\mathbf{y}} = \mathbf{G}(s)\Delta\underline{\mathbf{u}}$$

$$\begin{bmatrix} \Delta PCF \\ \Delta PCD \\ \Delta PSM \end{bmatrix} = \mathbf{G}(s) \begin{bmatrix} \Delta RMFD \\ \Delta PD \\ \Delta SD \end{bmatrix} \tag{5-16}$$

where the plant inputs are:

$$\Delta\underline{\mathbf{u}} = \begin{bmatrix} \Delta RMFD = \text{Feed rate of water to rod mill} \left[kg \cdot s^{-1}\right] \\ \Delta PD = \text{Primary sump Diluation} \left[kg \cdot s^{-1}\right] \\ \Delta SD = \text{Secondary sump Diluation} \left[kg \cdot s^{-1}\right] \end{bmatrix} \tag{5-17}$$

and the plant outputs are:

$$\Delta\underline{\mathbf{y}} = \begin{bmatrix} \Delta PCF = \text{Flow of feed to primary cyclone} \left[l \cdot s^{-1}\right] \\ \Delta PCD = \text{Density of feed to Primary Cyclone} \left[kg \cdot m^{-3}\right] \\ \Delta PSM = \text{Particle Size} \left[\% < 75\mu m\right] \end{bmatrix} \tag{5-18}$$

It is important to note that the outputs are of different units and therefore of different scales. For a fair comparison of loop performance, the error term should be taken as a fraction of the total range of the error for each loop.

The following transfer function matrix describes the milling circuit:

$$\mathbf{G}(s)=\begin{bmatrix} \dfrac{6.72}{1766s+1} & \dfrac{1.257}{61.3s+1} & \dfrac{0.1866}{573s+1} \\ \dfrac{80.9}{1978s+1} & \dfrac{-3.61}{73.7s+1} & \dfrac{0.854}{654s+1} \\ \dfrac{-5.25}{1059s+1}e^{-942s} & \dfrac{176.8s+0.255}{20900s^2+302s+1}e^{-197s} & \dfrac{122.4s+0.0657}{26300s^2+329s+1}e^{-76s} \end{bmatrix} \tag{5-19}$$

This model was found about the operation point of (estimated from Fig 12 in Hulbert and Braae (1981)):

$$\underline{\mathbf{y}}_0 \approx \begin{bmatrix} PCF = 110 \pm 30 \\ PCD = 1425 \pm 75 \\ PSM = 80 \pm 10 \end{bmatrix} \tag{5-20}$$

Since we are interested in the dynamic response about this point, the simulations origin will be about this point. To highlight the effects of constrained inputs, the input range was limited to $-4.5 < RMFD < 1.5$, $-7.0 < PD < 7.0$ and $-8.0 < SD < 8.0$.

A point to note is that due to the spread of the dead-time terms across the Particle Size output, ΔPSM, the term with the greatest steady-state effect is not been used as the dominant control action. This is not a "best-practice" approach, but was effective due to the significant dead times between inputs and the *PSM* output.

5.3.1 Linear Controller

The INA decoupling controller, $\mathbf{K}_d(s)$:

$$\mathbf{K}_d(s) = \left[K_d^1(s) \,\vdots\, K_d^2(s) \,\vdots\, K_d^3(s) \right] \tag{5-21}$$

is composed of the following column vectors:

$$\mathbf{K}_d^1(s) = \begin{bmatrix} \dfrac{92.0e9s^5 + 2.23e10s^4 + 14.8e6s^3 + 43.6e3s^2 + 48.2s + 0.30}{642e12s^6 + 18.8e12s^5 + 176e9s^4 + 702e6s^3 + 1.33e6s^2 + 1.5e3s + 1} \\ \dfrac{2.26e12s^5 + 50e9s^4 + 311e6s^3 + 743e3s^2 + 904s + 0.650}{18.9e15s^6 + 265e12s^5 + 1.31e9s^4 + 2.98e9s^3 + 4.09e6s^2 + 3.44e3s + 1} \\ \dfrac{473s + 0.0967}{16.4e9s^4 + 194e6s^3 + 695e6s^2 + 877s + 1} \end{bmatrix}^T \tag{5-22}$$

$$\mathbf{K}_d^2(s)=\begin{bmatrix} \dfrac{308e9s^5+5.84e9s^4+16.8e6s^3+26.2e3s^3+14.9s+0.00255}{2.26e12s^6+8.47e12s^5+106e9s^4+503e6s^3+1.03e6s^2+1.35e3s+1} \\ -\dfrac{1.68e12s^5+32.8e9s^4+96.2e6s^3+169e3s^3+254s+92.3e-3}{6.65e15s^6+148e12s^5+911e9s^4+2.24e9s^3+3.29e6s^2+3.23e3s+1} \\ -\dfrac{0.116s+0.539}{81.3e6s^3+486e3s^2+599s+1} \end{bmatrix}^T \tag{5-23}$$

$$\mathbf{K}_d^3(s)=\begin{bmatrix} -\dfrac{2.55e6s^3+56.7e3s^2+251s+0.116}{20.8e6s^3+406e3s^2+1.21e3s+1} \\ -\dfrac{7.47e3s^2+153e3s+0.603}{310e3s^2+1.11e3s+1} \\ \dfrac{1.57e3s+8.33}{533s+1} \end{bmatrix}^T \tag{5-24}$$

The decoupling compensator allows the design of three PI controllers for the de-coupled loops. These are chosen to enhance the transient response and to eliminate steady state error:

$$K_{PI}(s)=\begin{bmatrix} \dfrac{1000s+10}{1000s} & 0 & 0 \\ 0 & \dfrac{54.6s+0.6}{91s} & 0 \\ 0 & 0 & \dfrac{23.45s+0.35}{67s} \end{bmatrix} \tag{5-25}$$

The resultant composite controller is given here for completeness:

$$\mathbf{K}(s)=\left[K^1(s)\ \vdots\ K^2(s)\ \vdots\ K^3(s)\right] \tag{5-26}$$

$$K^1(s)=\begin{bmatrix} \dfrac{920e12s^6+23.2e12s^5+170e9s^4+584e6s^2+919e3s^2+782s+0.30}{642e15s^7+18.8e15s^6+176e12s^5+702e9s^4+1.33e9s^3+1.56e6s^2+1e3s} \\ \dfrac{22.6e15s^6+523e12s^5+3.61e12s^4+10.5e9s^3+16.5e6s^2+15.5e3s+6.50}{18.9e18s^7+265e15s^6+1.31e15s^5+2.98e12s^4+4.09e9s^2+3.44e6s+1e3s} \\ -\dfrac{4.73e6s^2+5.70e3s+0.967}{16.4e12s^5+194e9s^4+695e6s^3+877e3s+1e3s} \end{bmatrix} \tag{5-27}$$

$$\mathbf{K}^2(s)=\begin{bmatrix} \dfrac{16.8e12s^6+504e9s^5+4.42e9s^4+11.5e6s^3+16.5e3s^2+9.08s+1.53e-3}{20.6e15s^7+77.1e12s^6+9.65e12s^5+45.8e9s^4+93.7e6s^3+123e3s^2+91s} \\ -\dfrac{91.7e12s^6+2.80e12s^5+24.9e9s^4+67.0e6s^3+115e3s^2+157s-55.4e-3}{605e15s^7+13.5e15s^6+82.9e12s^5+204e9s^4+309e6s^3+294e3s^2+91s} \\ -\dfrac{6.33s^2+29.4s+0.323}{7.34e9s^4+4.42e6s^3+54.5e3s^2+91s} \end{bmatrix} \tag{5-28}$$

$$\mathbf{K}^3(s)=\begin{bmatrix} -\dfrac{59.8e6s^4+2.22e6s^3+25.7s^2+90.6s+40.6e-3}{1.40e9s^4+27.0e6s^3+81.1e3s^2+67s} \\ -\dfrac{175e3s^3+6.20e3s^2+67.7s+0.211}{20.8e6s^3+74.4e3s^2+67s} \\ \dfrac{36.8e3s^2+745s+2.92}{35.7e3s^2+67s} \end{bmatrix} \quad (5\text{-}29)$$

5.3.2 AW and ER Compensation

Having implemented a suitable linear solution, AW compensation and ER compensation are added to the controller. Since the controller already is of a high order, the simplest AW compensation is applied. In this case, CAW with parameters:

$$CAW=\begin{bmatrix} 20 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 10 \end{bmatrix} \quad (5\text{-}30)$$

The CAW constants are entered into the ER Toolbox to check that the AWBT compensation produces a stable system.

To apply ER compensation, the relative priority of each output is considered. The physical design of the milling circuit dictates the following prioritisation:

$$\mathbf{P}=\begin{bmatrix}3 & 2 & 1\end{bmatrix}^T \quad (5\text{-}31)$$

where the Particle Size, ΔPSM, is the primary output function. It will be shown that this prioritisation is not ideal for ER, but an attempt at optimisation is made to highlight the issues discussed in the formulation of ER. The prioritisation vector is used to generate the ER distribution matrix as per equations (4-1 through 4-5)

The nonlinear sensitivity parameters are taken as:

$$\varepsilon_u = 0.0001 \text{ and } \varepsilon_e = 0.0001 \quad (5\text{-}32)$$

Considering the prioritisation, the following is the form of the ER function matrix:

$$\mathbf{ER} = \begin{bmatrix} 1.0 & \varepsilon_{12} & \varepsilon_{13} \\ 0 & 1.0 & \varepsilon_{23} \\ 0 & 0 & 1.0 \end{bmatrix} \tag{5-33}$$

There is a maximum of three ER functions, $\varepsilon_{12}, \varepsilon_{13}$ and ε_{23}, that are required to be designed. To design ε_{12}, the following steps are considered:

- The term ε_{12}'s function is to redistribute errors in loop y_2 to loop y_1.
- Considering $\mathbf{G}_0$:
  - $\left|g_{21}/g_{11}\right| = \left|80.9/6.72\right| \approx 12$ is $>>1$, thus indicating that for small changes in the input u_1, relatively large changes can be expected in y_2.
  - To convert from units of e_2 to those of e_1: $g_{11}/g_{21} = 0.08\left[l.m^3/kg.s\right]$

Given these guidelines, it is expected that y_2 can be suitably compensated by redistributing its error to y_1. The initial value for ε_{12} is taken as $\varepsilon_{12} \approx 0.08$. Iterating through a range of values for ε_{12}, the range of values that produced stable systems is limited to $0 < \varepsilon_{12} < 0.808$. After some simulation, the value of $\varepsilon_{12} \approx 0.8$ was chosen.

The terms ε_{13} and ε_{23} are considered next:

- Both ε_{13} and ε_{23} are to redistribute errors in loop y_3 to loop y_1 and y_2 respectively.
- Considering $\mathbf{G}_0$:
  - $\left|g_{31}/g_{11}\right| = \left|-5.25/6.72\right| \approx 0.78$ is not $>>1$, thus indicating that large changes in the input u_1, will only result in relatively small changes in y_3.
  - $\left|g_{32}/g_{22}\right| = \left|0.255/-3.61\right| \approx 0.07$ is not $>>1$, thus indicating that large changes in the input u_2, will only result in very small changes in y_3.
- To convert from units of e_3 to those of e_1: $g_{11}/g_{31} = -1.28\left[l/\%<75\mu m.s\right]$

- To convert from units of e_3 to those of e_2: $g_{22}/g_{32} = -14.2\left[kg/\%<75\mu m.s.m^3\right]$

Given these guidelines, it is expected that ER compensation will not perform well in compensating for saturation in loop y_3. The initial values for ε_{13} and ε_{23} are taken as $\varepsilon_{13} \approx -1.2$ and $\varepsilon_{23} \approx -14$. Iterating through a range of values for ε_{13} and ε_{23}, the range of values that produced a stable system is limited to $-1.73 < \varepsilon_{13} < 1.01$, $-1.15 < \varepsilon_{23} < 2.38$. After some simulation, the values of $\varepsilon_{13} \approx -1.0$ and $\varepsilon_{23} \approx -1.0$ were decided upon.

Having considered all of the above, the following is the resultant ER function matrix:

$$\mathbf{ER} = \begin{bmatrix} 1.0 & 0.8 & -1.0 \\ 0 & 1.0 & -1.0 \\ 0 & 0 & 1.0 \end{bmatrix} \tag{5-34}$$

In this case study a number of operating point sets are considered. Only one full result will be presented here and the remainder will be in Appendix D. All operating points are defined by the vector function:

$$\underline{\mathbf{r}}^x(t) = \left[r_1^x(t) \quad r_2^x(t) \quad r_3^x(t)\right]^T \tag{5-35}$$

the separate setpoint functions will be specified at the beginning of each example.

5.3.3 Response Set for Operating Point 1

Table 30 Milling Circuit Performance Index ISE (OP1)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.333	0.323	0.346	0.332	0.000	2.26	1.49	0.180	0.129	5.40	1.11	1.12
2	0.000	0.021	0.001	0.019	0.667	1.29	5.00	0.789	0.062	0.25	0.13	0.28
3	0.000	0.037	0.001	0.005	0.000	2.58	0.20	0.007	0.808	1.57	1.65	1.83
Total	0.333	0.381	0.348	0.357	0.667	6.13	6.69	0.976	1.00	7.21	2.89	3.23

Table 31 Milling Circuit Performance Index ITAE (OP1)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.327	0.85	0.467	0.24	0.062	10.2	4.55	0.92	0.295	3.45	1.56	1.29
2	0.002	1.22	0.208	1.23	0.532	13.8	34.9	5.16	0.220	1.64	0.67	0.90
3	0.005	3.43	0.063	1.34	0.072	52.5	0.423	2.09	0.485	6.30	7.85	7.97
Total	0.333	5.50	0.738	2.81	0.666	76.6	39.8	8.17	1.00	11.4	10.1	10.2

Table 32 Milling Circuit Performance Index ITAO (OP1)

Steps	IDEAL	SAT	AWBT	ER
1	0.384	3.20	1.49	1.16
2	0.200	1.79	1.91	1.31
3	0.416	6.72	6.73	6.90
Total	1.00	11.7	10.13	9.38

$$r_1^1(t)=\begin{cases}0.0 & t<100\\ 10 & t\geq 100\end{cases},\ r_2^1(t)=\begin{cases}0.0 & t<5000\\ -10 & t\geq 5000\end{cases}\ \text{and}\ r_3^1(t)=\begin{cases}0.0 & t<10000\\ 2 & t\geq 10000\end{cases} \qquad (5\text{-}36)$$

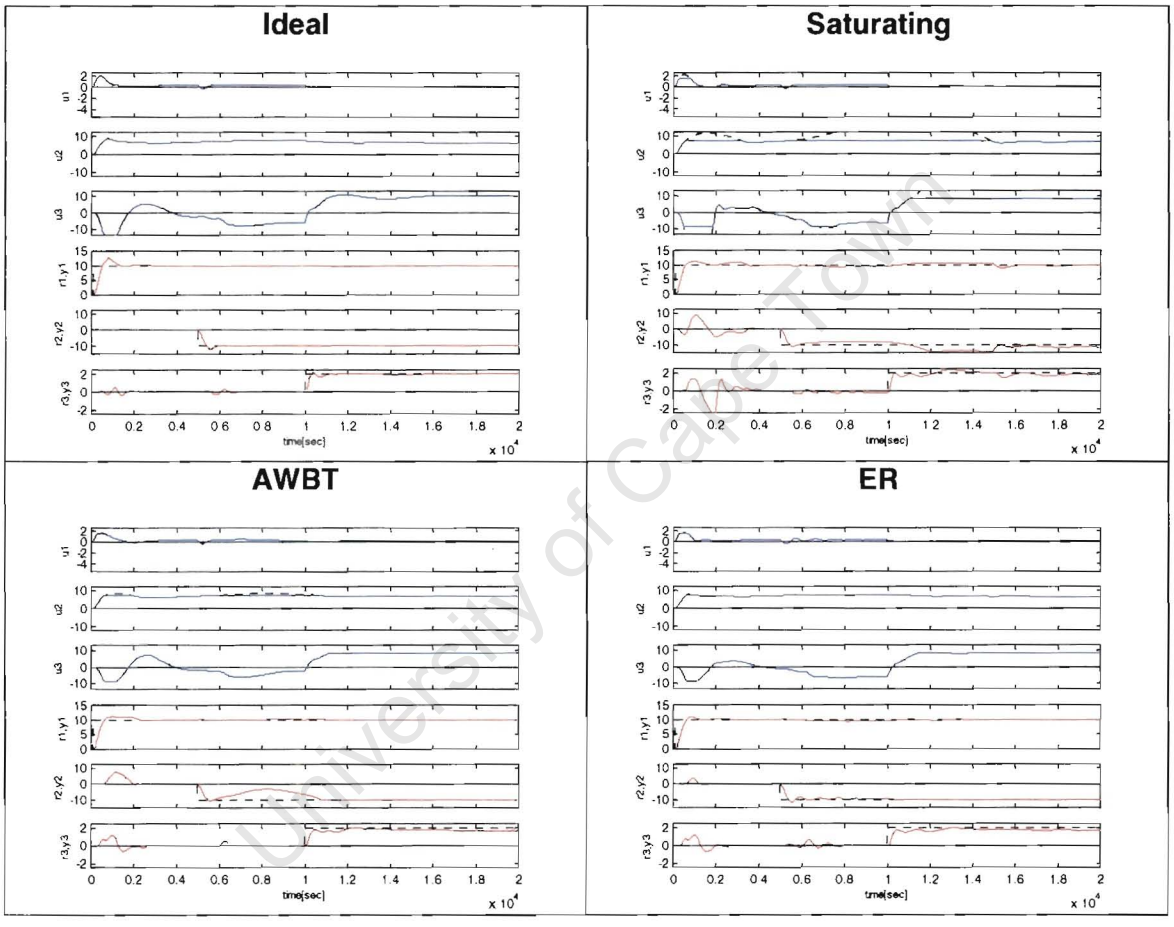


Fig.29 Milling Circuit Time Responses: Operating Point Set 1

5.3.4 Comments

AW compensation has a positive effect on the system, stability is maintained and performance is improved over the non-compensated saturating case.

As expected, good ER compensation by loop y_1 of loop y_2 result in excellent improvements in y_2. However, as predicted from the analysis in 5.3.2, ER compensation of y_3 with loops

y_1 and y_2 have little impact on the performance. Considering the results of the simulated responses to the other operating point sets, they indicate that ER compensation reduces the performance in loops y_1 and y_2 with little or no benefit in y_3 performance and should be removed, i.e. set $\varepsilon_{13} = 0.0$ and $\varepsilon_{23} = 0.0$. This would improve the performance of ER compensation of y_1 on y_2.

5.4 Laboratory Multivariable Thermal Process

This case study was implemented on a laboratory multivariable thermal process, MVTP, using the industrial control platform PlantStar. A more detailed discussion of the properties of the MVTP is given in Appendix E.

Fig.30 Photograph of the Multivariable Thermal Process

The photograph in Fig.30 shows the MVTP. On both the left and right, the control panel for each thermal source and sensor can be seen. The sensors are shown near the centre of the measurement area, which is a metallic grid. The sensors are held in place with magnetic contacts on their feet and are easily moved and placed. The experiment unit top has a matrix of fitment holes for the repeatable placement of the thermal sources.

The following block diagram, Fig.31, depicts the control system about the MVTP. The MVTP is connected to the PlantStar control platform, running on Windows 2000, through either a PLC or PC interface card. PlantStar supports OPC I/O interfaces and users can create PlantStar modules to interface with custom PC interface cards.

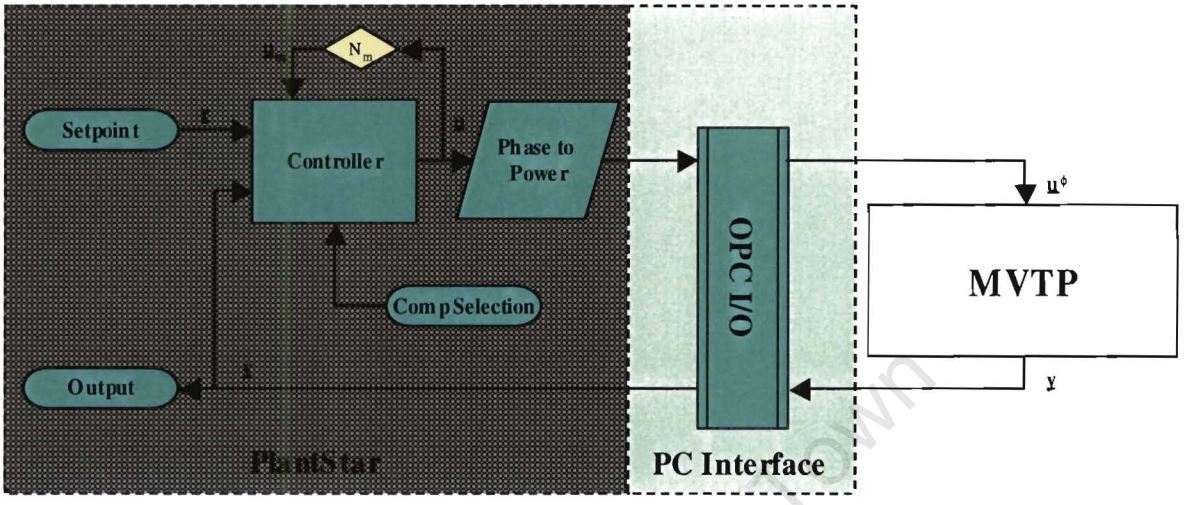


Fig.31 MVTP Control System

PlantStar provides the interface for changing operating points, *Setpoint*, and compensation techniques, *Compensation Selection* as well as viewing and data gathering functions, *Output*. A PlantStar module was written to do the Power to Phase conversion (see Appendix E) to linearise the effect of the PWM used to drive the thermal source. The control action, $\underline{\mathbf{u}}$, is in terms of power; while the actual process input is a phase signal $\underline{\mathbf{u}}^{\phi}$, which corresponds to the requested power level. Since the saturation limits of the system are known, N_m is used to derive $\underline{\mathbf{u}}_m$.

5.4.1 Process Model and Linear Controller

The *Controller* is implemented using standard PlantStar models and PlantStar Language to control the logic. The process outputs are scaled 0 to 100 degrees and the process inputs are from 0 to 10, representing 0 to 100% of the maximum possible power of 1200W to each of the thermal sources. The PlantStar control strategy includes the power to phase conversion.

For this study the system will be modelled as a 2x2 transfer matrix function. The model will be obtained but carrying out step tests about the operating point. Stepping the process

inputs, which correspond to the power supplied to the thermal sources, from 30 to 60% the step responses were analysed and an approximate first order model found:

$$\mathbf{G}(s)=\begin{bmatrix} \dfrac{3.00}{11s+1} & \dfrac{2.99}{12s+1} \\ \dfrac{1.19}{10s+1} & \dfrac{3.46}{11s+1} \end{bmatrix} \tag{5-37}$$

The units are in degrees Celsius per volt $\left({}^{0}C/_{V}\right)$ for all transfer functions. Following the example in Chapter 2, a decoupling controller and SISO PI controllers are designed for each loop. The resultant linear controller being:

$$\mathbf{K}(s)=\begin{bmatrix} \dfrac{1.50s+0.50}{3s} & -\dfrac{49.28s^2+2.091s+1.493}{108.1s^2+9.01s} \\ -\dfrac{19.63s^2+8.33s+0.595}{103.7s^2+10.37s} & \dfrac{1.50s+0.50}{3s} \end{bmatrix} \tag{5-38}$$

The simulated closed loop response for a set of operating points that do not saturate the control signals is given in Fig.32. This shows both reasonable performance with good decoupling of the off-diagonal terms.

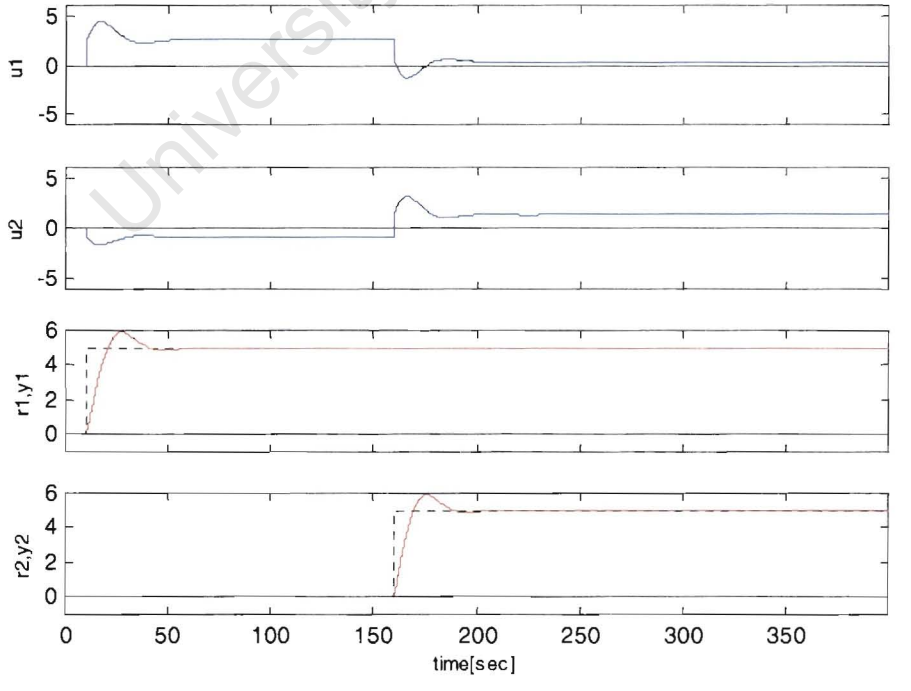


Fig.32 Simulated Ideal Response for Laboratory MVTP

To confirm the design, non-saturating closed loop step tests were performed, Fig.33, with a control step time of 250 milliseconds. The results showed reasonable setpoint tracking and fair decoupling of the g_{21} term during steps on y_1. However, y_1 showed significant disturbance during the stepping of y_2 thus indicating that the controller has not optimally decoupled g_{12}. The error in design could reside in the method of synthesizing $\mathbf{K}(s)$ and/ or the tuning of the PI controllers or in the model $\mathbf{G}(s)$ of the system. In either case, the compensation tests will be run on this sub-optimally controlled system. This will give an indication of the robustness of the compensation to deal with a sub-optimal and real-world situation, where drift in the process tends to shift the control system away from the optimal.

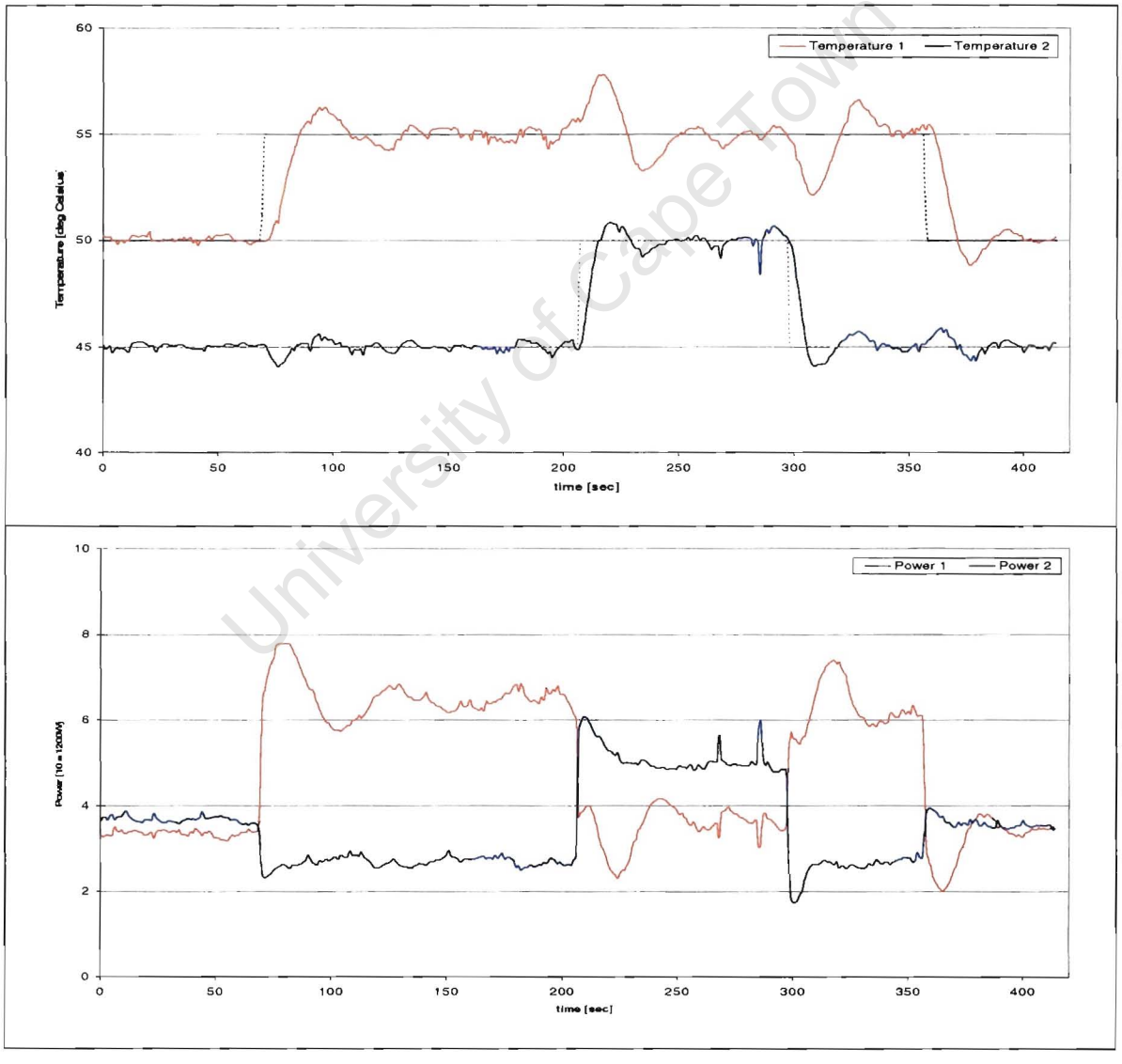


Fig.33 Temperature and Power Traces For Ideal Response of Laboratory MVTP

For this example the performance index numbers have been normalized over the duration of each step. The results for the non-saturating step, in Table 33, are not directly comparable to the saturating case and compensated cases, but serves as a reference for the order of magnitude of indices that can be expected. In this example, y_1 and y_2 correspond to *Temperatue 1* and *Temperature 2* in degrees Celsius; and u_1 and u_2 correspond to *Power 1* and *Power 2* in tens of percent of the total power, 1200Watts.

Table 33 Performance Indices for the Non-saturating Laboratory MVT Process

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	0.51	0.02	7.99	3.64	9.14
2	0.27	0.24	4.18	2.82	5.57
3	0.25	0.23	2.71	1.18	3.06
4	0.52	0.02	1.51	0.68	1.73
Totals	1.56	0.51	16.41	8.33	
Overall	2.07		24.74		19.50

5.4.2 Saturation

To experience the full nonlinear affect, the control region was not truncated to known linear responses, but allowed to approach the maximum limits that the system could safelly produce. Hence the saturation function is:

$$\underline{\mathbf{u}}^* = \mathbf{N}(\underline{\mathbf{u}}) = \mathbf{sat}(\underline{\mathbf{u}}) = \begin{cases} +9.9 & u_i > +9.9 \\ u_i & \text{else} \\ 0.2 & u_i < +0.2 \end{cases} \tag{5-39}$$

From running at a steady-state operating point of $\underline{\mathbf{r}}^0 = [45 \quad 45]^T$, the system goes through the following sequence of operating points:

$$\underline{\mathbf{r}}^1 = \begin{bmatrix} 65 \\ 45 \end{bmatrix} \text{ and } \underline{\mathbf{r}}^2 = \begin{bmatrix} 65 \\ 50 \end{bmatrix} \tag{5-40}$$

The results of the uncompensated system are dire. In Fig.34 the requested control action for the first loop experiences severe windup while the second loops control action is not saturating, this effect is discussed in Section 2.3.2.

5.4.3 AW and ER Compensation

As a first attempt at containing the controller states, CAW compensation will be added to the control system. Since the control time steps are larger than those used in the simulated case study, a smaller value of the CAW parameters of equation (3-15) will be used: $\mathbf{X} = \mathbf{I}$.

As in the earlier example, the relative output priority is assigned as:

$$\mathbf{P} = \begin{bmatrix} 1 & 2 \end{bmatrix}^T \tag{5-41}$$

Thus, loop y_2 will be used to compensate for errors in loop y_1. The prioritisation vector generates the ER distribution matrix, as per equations (4-2 through 4-5):

$$\delta = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \tag{5-42}$$

Thus indicating that the some function of error, e_1, will be added to e_2. The nonlinear sensitivity parameters, that are used to "select" the mode of operation (see Section 4.3.2), are taken as:

$$\varepsilon_u = 0.001 \text{ and } \varepsilon_e = 0.5 \tag{5-43}$$

The value for ε_e is chosen from considering the ideal response where the temperature was roughly held within half a degree of the setpoint during steady-state operation. ε_u was simply chosen to be small enough so that any saturation while $|e_1| > \varepsilon_e$ would enable compensation.

From the prioritisation and resulting ER matrix, the form of the ER function matrix is:

$$\mathbf{ER} = \begin{bmatrix} 1.0 & 0.0 \\ \varepsilon_{21} & 1.0 \end{bmatrix} \tag{5-45}$$

Only one ER function is required. To design this function, ε_{21}, consider the guidelines in Section 4.4.3:

- The term ε_{21}'s function is to redistribute the errors in loop y_1 to loop y_2.
- Considering $\mathbf{G}_0$ from equation (5-37):
  - $\left|g_{12}/g_{22}\right| = \left|2.98/3.457\right| \approx 0.86$ which less than but close to unity, thus indicating that changes in the input u_2, through changes in the effective error for y_2, are capable of producing comparable changes in y_1.
  - To convert from units of e_1 to those of e_2: $g_{22}/g_{12} \approx 1.16$
- $\text{sgn}(g_{12}k_{22}) = +1$

So as not to consume all the available control action in one iteration, a realistic upper limit for ε_{21} can be found from considering equations (4-19 through 4-25).

Given these guidelines, it is expected that y_1 could be suitably compensated by redistributing its error to y_2. The ratio $\left|g_{12}/g_{22}\right|$ being close to unity, large values (>1) are expected for the ER function. The initial value for ε_{21} is taken as suggested above as $\varepsilon_{21} \approx 1.16$. The stability check indicates that the values for CAW and ER produce a stable system. The first attempt ER function matrix is thus:

$$\mathbf{ER} = \begin{bmatrix} 1.00 & 0.00 \\ 1.16 & 1.00 \end{bmatrix} \tag{5-46}$$

5.4.4 Response Set

The following are the time responses for the saturating (Fig.34), CAW compensated (Fig.35) and CAW-ER compensated (Fig.36) systems when passed through the operating points in equation (5-40). The performance indices are tabulated in Table 34 through Table 36 with time normalised values for each case. Clear performance improvements for y_1 are obtained for acceptable degradation in y_2.

Table 34 Laboratory MVT Process Performance Index ISE

Steps	y_1			y_2		
	SAT	CAW	ER	SAT	CAW	ER
1	21.94	18.78	9.55	1.88	0.42	2.59
2	0.88	0.34	0.10	0.49	0.34	0.08
Total	22.86	19.12	9.65	2.37	0.76	2.67

Table 35 Laboratory MVT Process Performance Index ITAE

Steps	y_1			y_2		
	SAT	CAW	ER	SAT	CAW	ER
1	86.66	57.62	25.22	28.76	3.99	28.29
2	11.96	5.77	4.47	4.75	3.10	1.77
Total	98.62	63.39	29.69	33.51	7.09	30.06

Table 36 Laboratory MVT Process Performance Index ITAO

Steps			
	SAT	CAW	ER
1	91.37	57.86	38.92
2	12.98	6.94	5.26
Total	104.4	64.80	44.18

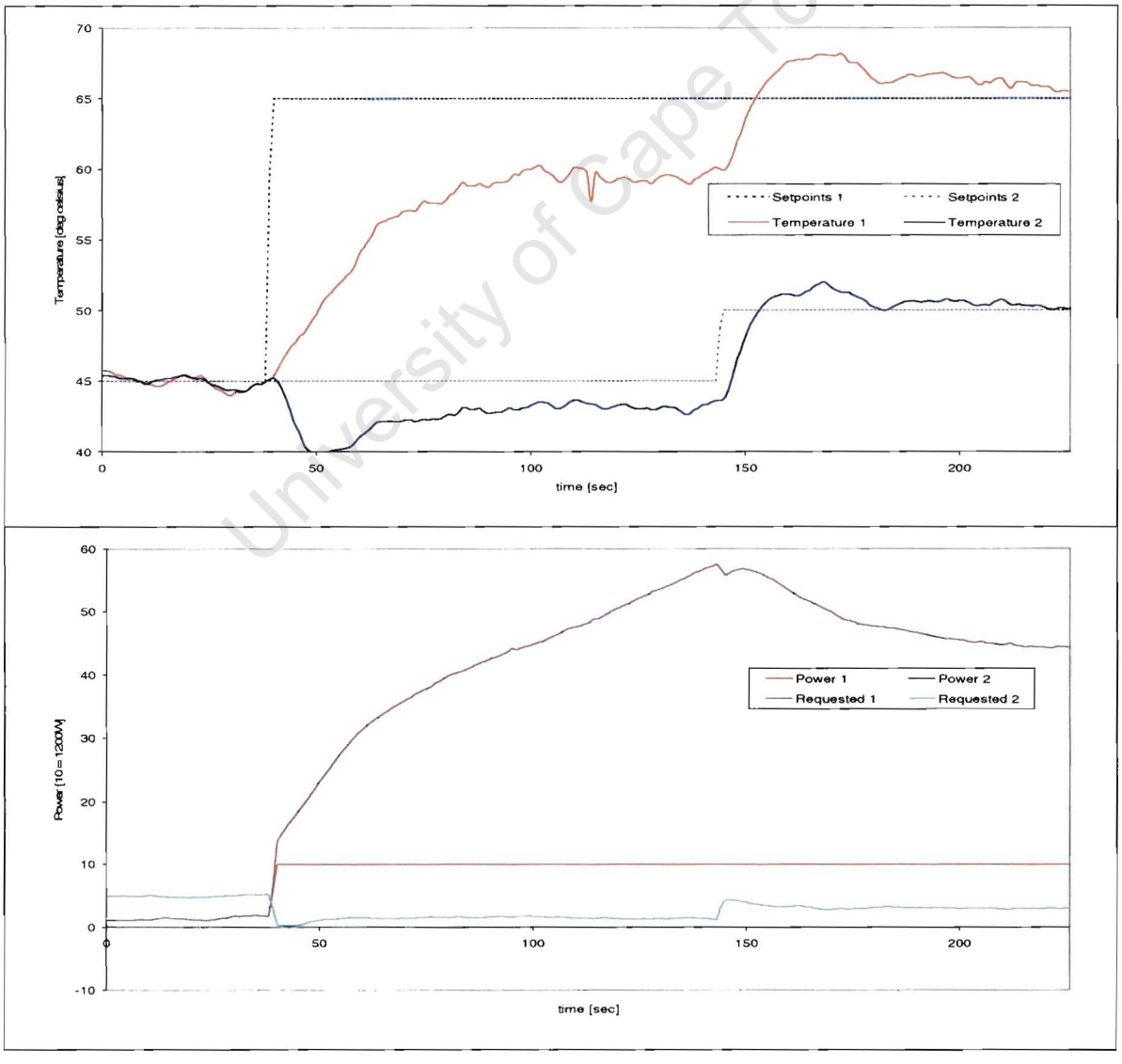


Fig.34 Temperature and Power Traces For Saturating Response of Laboratory MVTP

This uncompensated system clearly exhibits windup of u_1, while u_2 is not saturating. The windup also affects the return to linear performance which only accurs after 360 seconds.

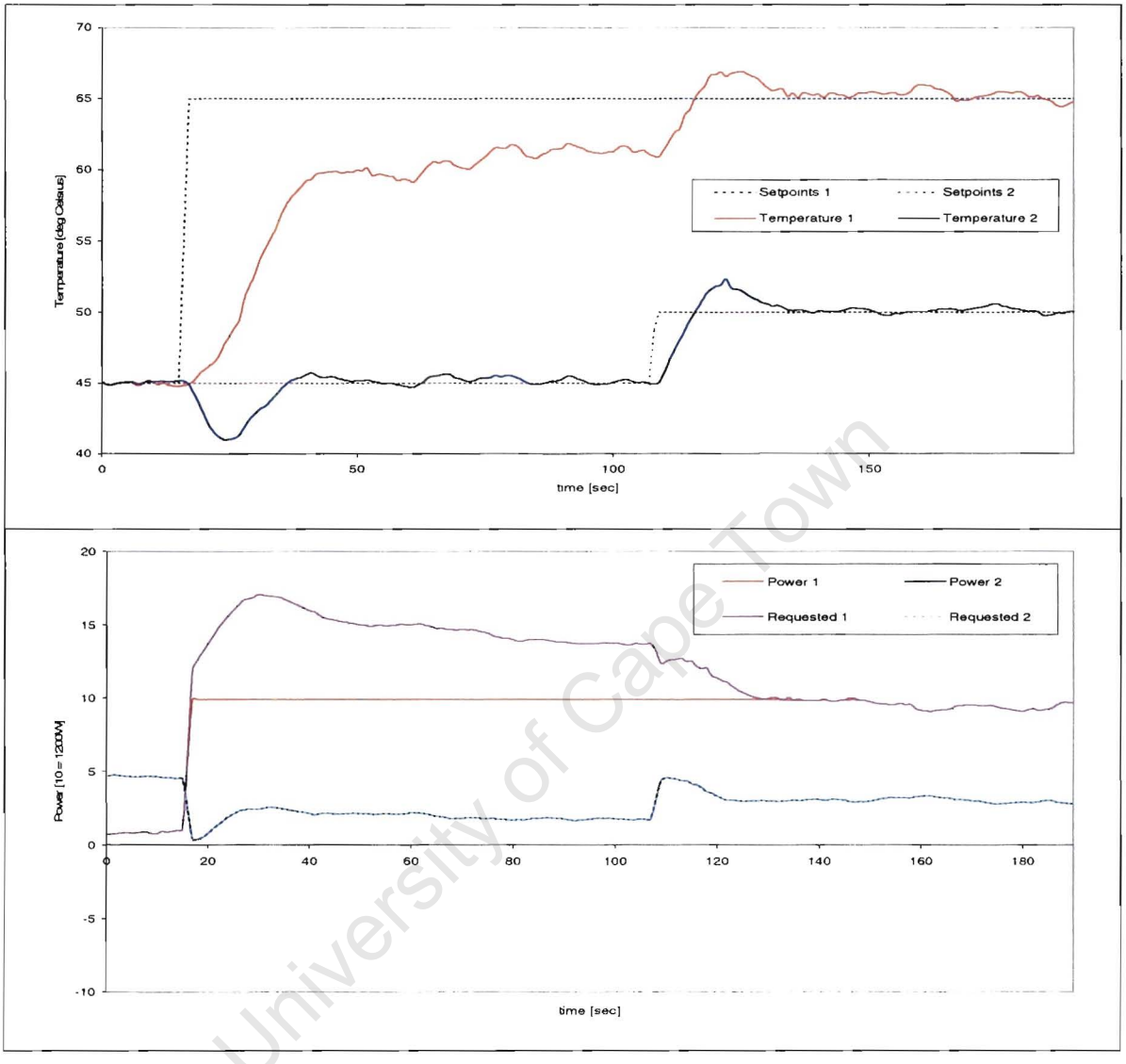


Fig.35 Temperature and Power Traces For CAW Response of Laboratory MVTP

The inclusion of CAW compensation in Fig.35 has clearly bounded the controller states when compared with the saturating response in Fig.34. The output performance has been greatly enhanced, with y_2 maintaining its setpoint. During $\underline{\mathbf{r}}^2$, u_1 quickly returns to realisable region and both y_1 and y_2 obtain their setpoints.

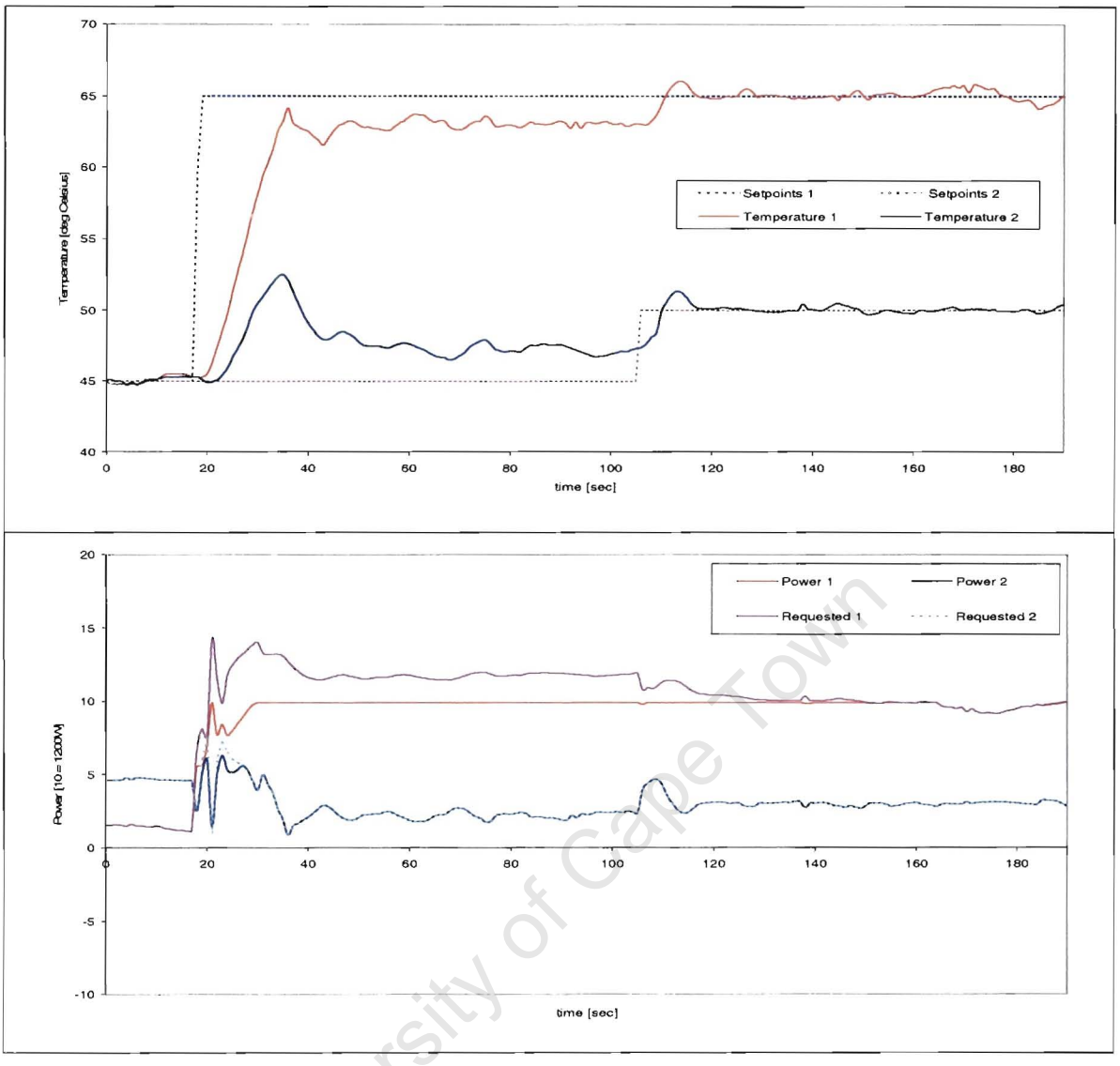


Fig.36 Temperature and Power Traces For CAW-ER Response of Laboratory MVTP

In Fig.36, ER compensation has been added to the CAW compensation. Here with an acceptable increase in movement of the control signals, the performance of y_1 has been greatly improved over that seen in Fig.35.

5.5 Discussions and Recommendations from Case Studies

The application of AW, AN and ER compensation techniques have been successful in the stabilisation of the four examples discussed in this chapter. Both HCT-AN and ER compensation have shown good abilities at optimising processes during nonlinear operation. Comparison of the weighted performance indices gave a fair indication of the improvement of performance in terms of an optimal operating point during nonlinear operation.

For the simulated MVTP example HCT-AN produced better results; and in the Distillation Column example, HCT-ER out-performed HCT-AN. The difference between AN and ER techniques can be summarised as:

- ER shifts the demanded reference to an optimal operating point by feedback. This generates a hunting response as the system searches for the optimal operating point. ER exploits the interaction present in the process as well as the effects of the controller to improve performance. Being a regulator around the process, which is trying to minimise the error in high priority loops, it is not critically dependent on the process and boundary conditions. However, the examples have shown that the performance improvement does depend on the AW compensation being used. Improving the guidelines to include issues related to the AW technique should be considered in future work.

- AN uses the back calculation of the realisable setpoint by reversing the effects of the controller. When the controller is formulated such as with the HCT compensation, there exists a closed form solution to the nonlinear optimisation. In other cases the nonlinear optimisation, equation (3-28), involving the inverse of the controller, $\mathbf{U}(S)$, needs to be solved online. The implementation used here, as suggested by Vrancic (1997), depends on the definition of the boundary conditions and should be reformulated to use the sensed control action boundaries.

From the guidelines for ER implementation given in Chapter 4, the easy application of this technique was possible in all the examples. In the case of the Milling Circuit, limited effectiveness of ER was predicted and this was found to be the case.

6 Further Study

The scope of this thesis was to present error redistribution (ER) as a viable means of performance optimisation during nonlinear operation. Being the initial formulation of ER compensation there are number of issues that require further study and integration into work that has been ongoing in the field of AWBT research. References referred to in this Chapter are included in a separate appendix: Further Study References.

In particular, the ER guidelines used to form the ER functions should be the result of some optimisation problem. A more rigorous formulation may be possible by exploiting the realisable output region defined by (3-27) upon which the HCT-AN is based. Since ER compensation uses the ER State Selector, equation (4-9), to detect the nonlinear mode and the resulting ER function matrix operates on the error vector, $\underline{e}$, it would not be as sensitive to changes in the realisable control space as HCT-AN in Section 5.1.2. On the practical side, the application and comparison of HCT-AN to the laboratory MVT would be of interest.

A number of researchers have been looking into the stability and synthesis techniques for AWBT (Kothare and Morari, 1999; Mulder *et al.*,2000, 2001; Grimm *et al.*, 2001; Mulder and Kothare, 2002) as well as the application of MPC to AWBT (Morari and Lee, 1997).

In particular, work by Mulder *et al.* (2001) present a synthesis technique for the AWBT compensation including an optimisation component using linear matrix inequalities. The stability analysis presented follows similar lines as that used in the ER compensation analysis. An investigation into whether ER is a special case would be of interest. This would also facilitate a more rigorous derivation of the ER function matrix.

Any investigation into the development of the ER function matrix should include the consideration of the class of problems to which the formulation can be applied. In particular, the application to non-minimal phase processes and unstable systems. Also systems with different tracking requirements, such as ramped or oscillatory systems.

A control structure technique, called selectors, should also be considered and a comparison drawn between their functionality and that of ER. Selectors are used in Single-Input-Multiple-Output (SIMO) systems where there are separate controllers, each measuring a separate output and having independent setpoints, the selector is used to

choose the highest, lowest or median of the control actions to output to the shared actuator (Åström, 2001). The concept of 'override selector' is that when process outputs exceed limits, the selector switches to a specified controller. The application appears to be of an ad hoc nature where each implementation would depend on an in depth understanding of process (Protuner Application Manual, 2002).

ER compensation employs a Multiple-Input-Single-Output (MISO) approach in that a number of inputs can be employed to optimise a single output within a MIMO system.

7 Conclusions

The purpose of this thesis was to devise an intuitive, but theoretically justifiable, method for the optimisation of output responses during nonlinear operation, the method being easily applied to industrial situations. Error redistribution (ER) has been shown to meet this objective.

The approach taken was to first examine the effects of saturation in MIMO systems in Chapter 2 and to explain phenomenon relating to them. In particular, the condition under which a single control action in a 2x2 system saturates but both setpoints are lost was discussed in Section 2.3.2. The clear significance of the correlation of states within the controller and those that were represented by the realised process input was illustrated.

Over the last two decades a number of anti-windup bumpless transfer (AWBT) techniques have been devised to deal with saturation within control systems. As the tendency is to solve simpler problems first, most of the techniques (Anti-reset Windup, Conventional Anti-windup, Hanus Conditioning Technique) were evolved from SISO solutions to MIMO implementations. Chapter 3 presented an overview of one of the most recent unifying frameworks for AWBT. A number of newer techniques that have taken greater cognisance of MIMO issues (Modified Anti-windup, Artificial Nonlinearity) were also included. Five of these compensation techniques were implemented in the multivariable thermal (MVT) process simulation that accompanies the discussion throughout the text.

The AW compensation techniques were all improvements on the uncompensated system but with varying degrees of success. Containing the states of the controller, and therefore its ouputs, ensured stability and improved linear recovery but only the implementation of HCT-AN included MIMO performance criteria in its design. Current literature also shows examples of the varying effectiveness of techniques. This is driving the design method from one that precluded AW issues from the linear design to a more holistic approach where the control strategy is designed with AW in mind. There are two significant paths of current research, one looking at optimisations to the heuristic and traditional AW compensation design and implementations; and the more generic approach of model predictive control (MPC).

In Chapter 4, the novel concept of error redistribution (ER) was introduced as a means by which the designer could make use of the available control action in non-saturating loops

to optimise the performance of higher prioritised outputs during modes of nonlinear operation. The results of applying CAW and ER compensation to the MVT example were shown, along with a design procedure and stability proof.

The application and case studies in Chapter 5 discussed:

1. the use of ER with 2 other AW compensations (HCT, MAW) and the comparison with the HCT-AN technique;
2. a Distillation Column, a 2x2 system with dead-time and non-minimal phase introduced by the decoupling controller;
3. a Milling Circuit, a 3x3 industrially implemented control system with significant dead-time; and
4. a real application to a laboratory model of the MVT Process with the industrial control platform PlantStar.

Throughout this thesis, simulations and comparisons of the results for ideal, saturating and compensated systems were produced using MATLAB. An ER toolbox was developed to aid designers check stability and performance of AW and ER compensated systems. Analysis of the performance of the systems was completed using integral-square-error (ISE), integral-time-multiplied-absolute-error (ITAE) and the integral-time-multiplied-absolute-operating-point-error (ITAO) performance indices. Using these allowed an assessment of the performance considering the rise-times, steady-state and overall steady-state performance on a per step basis. A weighting system was also introduced to measure the optimisation obtained by the applied compensation.

The focus of this thesis has been the performance of a process during nonlinear operation that is not due to controller dynamics, but are a result of the operating points been placed outside those that are physically realisable by the system.

The application of AW techniques indicated that the containment of controller states was critical to stable and linear performance. No method presented a clear optimal implementation. Techniques based around the Hanus conditioning technique, HCT and GCT, which have a good theoretical basis for their claim to an optimal solution, were clearly out-performed in examples presented both here and in the literature. Other techniques that work to maintain the direction of the control vector during saturation, such as MAW, also under-performed compared to simpler methods such as CAW.

To the question posed by industry in the introduction, ER and HCT-AN have both been proven successful methods of implementing optimisation during nonlinear operation. Examples have shown that although both ER and HCT-AN systems offer significant improvements over standard AW compensation, neither are guaranteed to produce the optimal performance.

An artificial-nonlinearity based around HCT has a closed form solution, and of the techniques considered, this was the only one to produce comparable results to ER compensation. However, a general solution requires the online solving of a nonlinear optimisation problem.

The application of ER compensation has proven effective. ER guidelines allowed the easy design of, and implementation on both simulated and industrial systems. The performance of the systems considered was consistent with what the theoretical analysis predicted. Although the formulation and stability analysis allow for the introduction of complex ER functions, only simple gains where used in the examples presented. Additional design guidelines should be developed around the formation of these functions, including the effects of the AW. Chapter 6 Further Study, discusses avenues for further research in this area to follow.

To summarise, the contribution being made by this thesis is:

- The evaluation of the effect of anti-windup compensation techniques on processes whilst operating in nonlinear modes.

- The concept of error-redistribution for performance optimisations during nonlinear operation has been presented.

- A formalisation of the technique with theoretical analysis and inclusion of the ER compensation technique into Kothare's unified framework for AWBT techniques for the stability analysis.

- A set of practical implementation guidelines for designers to exploit ER. These include the suitability of applying ER to the process been considered.

- Comparison of the performance of ER with an artificial-nonlinear optimisation, HCT-AN.

- Extensive simulated application case studies along with a MATLAB toolbox with which to undertake such studies.

- Implementation of ER compensation on a laboratory Multivariable Thermal Process (MVTP), designed during this work, on an industrial control platform - Mintek's PlantStar.

A. MATLAB Error Redistribution Toolbox

To facilitate the analysis of AWBT and ER compensation a collection of MATLAB functions and reports have been developed. These allow the user to define the process and linear controller; include saturation, AWBT and ER compensation; and execute a simulation followed by a performance evaluation. They provide a number of graphs and performance indices in a report form for comparison purposes. All the simulations presented throughout this thesis were done using this toolbox. This appendix is given as a brief user guide for the toolbox.

Overview

The first step is to define the ideal system; the user specifies the Process and the Linear Controller. The saturation parameters for the process inputs define the non-linear element.

The AWBT framework allows a certain amount of automation of the conversion of the simpler AWBT techniques to the framework's parameterisation. The user can specify the type of AWBT to be used as well as the particular parameters required for the compensation. This operation also carries out the necessary stability check on the models.

The user supplies the ER parameters and ER function; again, the toolbox has stability checks.

Once the problem is defined, necessary operating points are specified; the case study is then ready for processing with the *ProcessERStudy* function. This function uses Simulink to simulate the four cases, and a number of other functions to process the data and produce the ER Study report.

The results can be saved as a MATLAB workspace, thus allowing additional analysis without undergoing the simulations again. The reports are in HTML format for ease of distribution.

It is important that the maximum time step is limited to one that produces equivalent performance indices for all four cases when the simulation has no periods of saturation.

Defining the Case Study

The toolbox is designed to only process one ER case study at a time. The toolbox defines the global variables that will be given specific values to be used to in the study.

These parameters are required for the analysis, they are defined in the file *ERGLOBALS.m* , and include the following:

Variable	Description
ERStudyName	Name by which to identify study
ssK	Minimal state-space representation of Linear Controller, **K**
ssG	Minimal state-space representation of Process Model, **G**
ssGint	Column vector of initial states for Process Model
AWBTTYPE	String parameter selecting the AWBT Technique `'ARW'` - Anti-reset windup (single variable only) `'CAW'` - Conventional Anti-windup `'HCT'` - Hanus Conditioning Technique `'GCT'` - Generalised Conditioning Technique `'OAW'` - Observer-based Anti-windup `'IMC-AW'`- Internal model controller Anti-windup Derived from Kothare *et al.* (1994)
AWBTPARAMETERS	Parameters required for selected AWBT technique (defined in MATLAB help)
Umin	Column vector of Minimum Input Control Action
Umax	Column vector of Maximum Input Control Action
Ee	ER Error non-linear sensitivity parameter, ε_e
Eu	ER Control action non-linear sensitivity parameter, ε_u
PRIORITISATION	ER Prioritisation Vector, $\underline{\mathbf{P}}$
NOOUTPUTS	Number of outputs
ssER	Minimal state-space representation of ER function
MAXSTEPSIZE	Maximum step size for simulation, in seconds
STOPTIME	Length of simulation, from 0 seconds to STOPTIME, in seconds
TOLERANCE	Tolerance used during Simulink simulations
STEPTIMES	Column vector of the time at which to step this particular row's setpoint from the initial to final value, additional columns are for additional operating points
STEPINITIAL	Column vector of the initial setpoint values, additional columns are for additional operating points

STEPFINAL	Column vector of the final value for each setpoint, additional columns are for additional operating points
RESULTS	An empty matrix used to store the case study results.

To correctly parameterise and check the system's stability, two functions are provided: *ConvertToAWBTFramework* will transform the linear controller into the AWBT framework; *ERStateSpaceStabilityCheck* will confirm the stability of the ER compensated system.

[ssKint, I_V, I_Vint, U, Uint, StableAWBT] =
ConvertToAWBTFramework (ssK, AWBTTYPE, AWBTPARAMETERS)

The arguments are the state-space representation of the linear controller, K, and the AWBT variables described above.

Resultant	Description
ssKint	Column vector of initial states for Linear Controller
I_V	AWBT framework system I – V
I_Vint	Column vector of initial states for I_V
U	AWBT framework system U
Uint	Column vector of initial states for U
StableAWBT	Returns 1 if AWBT systems, I_V and U are open loop stable

[StableER,StableOLER,StableG,StableU,StableI_V,StableERFunc,StableUER] =
ERStateSpaceStabilityCheck (G_,U_,I_V_,ERFunc_)

The arguments are the state-space representations of the process,G, and the AWBT systems, U and I-V, as described above.

Resultant	Description
StableG	1 if the process G is open loop stable
StableU	1 if the AWBT U is open loop stable
StableI_V	1 if the AWBT I-V is open loop stable
StableERFunc	1 if the ER compensation is open loop stable
StableUER	1 if the ER formulation Us is open loop stable
StableOLER	1 if the ER formulation Qs is open loop stable
StableER	1 if the ER formulation is closed loop stable

Processing the Case Study

Having defined the study, the *ProcessERStudy* function will execute the simulations. This function requires that the following arguments describe which case study and operating point are being considered and the various properties of the resulting report:

```
[IdealPIstr,ISE,ISEstr,IAE,IAEstr,ITAE,ITAEstr,ITAO,ITAOstr,ISU,ISUstr] =
        ProcessERStudy (studyname_,filename_,...
                        op_,Urange_,timerange_,traj_range_, ...
                        range_,simulate_,show_,erroranalysis_,normalise_,report_)
```

The arguments are used to specify the type of processing and reporting that is required based on the case study that has been defined in global variables.

Argument	Description
studyname_	Name that will be use to identify the study, usually ERStudyName
filename_	Specific of Simulink model for this case study, usually 'Model'
op_	Particular operating point info for Study. Relating to the number of the column of the STEPxxxx parameters
Urange_	U range: `[Umin, Umax]`

timerange_	Axes information for the time plots, for all the u_i and r_i, y_i axes:
	`[umin, rymin;`
	`umax, rymax]`

traj_range_	Axes information for the trajectory plots, one column for each r_i, y_i and u_i. The last column has the time marker interval.
	`[ry1min, ry2min,… u1min, u2min,… ,ryinterval;`
	`ry1max, ry2max,… u2max, u2max,… ,uinterval]`

range_	The portion of the simulated time (0 – STOPTIME) that is to be used in the analysis: `[start; stop]`
simulate_	This function can either just process the existing RESULTS matrix, or do the simulations and populate the RESULTS before processing depending on whether this parameter if 0 or 1

show_	A number of graph outputs are possible, bitwise-or these values together to generate the required graphs:
	`1 - Time responses`
	`2 - Output trajectories`
	`4 - Control trajectories`

	8 - To truncate control trajectories
erroranalysis_	Options in the interpretation of the error can be specified here: 1 - Convert the error to a % of the steady-state realisable region 2 - Use prioritised weighting
normalise_	1 to normalise indices by the ideal response index, else 0
report_	1 to generate HTML reports, else 0

The function assumes that the following global variables are defined (their descriptions are given in the preceding section):

ssG	SsK	I_V	U	SsER
ssGint	SsKint	I_Vint	Uint	SsQ
Umin	Umax	Ee	Eu	PRIORITISATION
STEPTIMES	STEPINITIAL	STEPFINAL	STOPTIME	MAXSTEPSIZE
TOLERANCE	NOOUTPUTS			

The simulation of the case study is achieved using a Simulink model. The toolbox comes with a generic multivariable system, 'Modelxxxxx.mdl'. The file name is in the form of the specific model with the prefix indicating which case: IDEAL, SATurating, AWBT compensated or ER compensated. Due to the way delays are implemented in Simulink, another set of files for a 2x2 and 3x3 systems with delays are given in the examples.

During processing the command window will display the progress and current operations being executed. If any graphs were selected, figures will be opened to display the graphs; and if reporting was selected, a browser window will be opened presenting the report. The format of the report is discussed later.

The resultants of this function give a raw and formatted (XXXXstr) version of the performance indices. The IdealPIstr always gives the absolute values for the Ideal case as a reference for when the indices are normalised to the ideal case. Typically the report is the best way to view these results.

Custom Automations

The definition of the Case Study involves a large number of parameters. The three examples discussed throughout this text are included on the CD as case studies and examples of how to exploit the functionality in the toolbox. A brief description of the four typical custom files used follows:

SimXXXXX.m

This file defines the process, controller, compensation and operating points. It also runs the conversions and stability checks.

The compensation frameworks are state-space formulations. MATLAB provides functions to transform transfer function matrix forms into minimal state-space realisations.

When working from a transfer function matrix problem, the following variables are generally used to define the process:

Variable	Description
G	Definition of Process
Kd	Definition of Precompensator
Kpi	Definition of Diagonal loop controllers
Ks	Complete Linear Controller

Where state-space forms are to be derived from the transfer function models, the Matlab function, *ss(XX,'min')* is used to produce a minimal state-space form.

ProcessXXXX.m

This file encapsulates the properties and parameters for the *ProcessERStudy* function call for this case study.

ProcessAllXXXX.m

This file loops though all the defined operating points.

XXXX.Example.m

This file runs through all operating points, but has internal parameters that control its behaviour, including whether or not to do the simulations. In particular, this file copies the resultant reports and case study results to separate directories for easy analysis.

Reports and Outputs

All the simulation data is stored in the RESULTS matrix and if the workspace is saved with the case study defining global variables, the processing can be done later with greater speed as the simulations take the longest to do.

The main HTML report has the following chapters:

1. Process Parameters

This chapter summarise the control strategy and compensation parameters.

2. Time and Trajectory Responses

2.1. Time Trends

The control action and setpoint/output response versus time.

2.2. Control Trajectories

The control action trajectories in the control action space.

2.3. Output Trajectories

The setpoint and output trajectories in the output space.

On the trajectory graphs, the shaped area represents the realisable control space or steady-state realisable output space, respectively.

3. Performance Indices Analysis

The tabulated performance indices for the case study

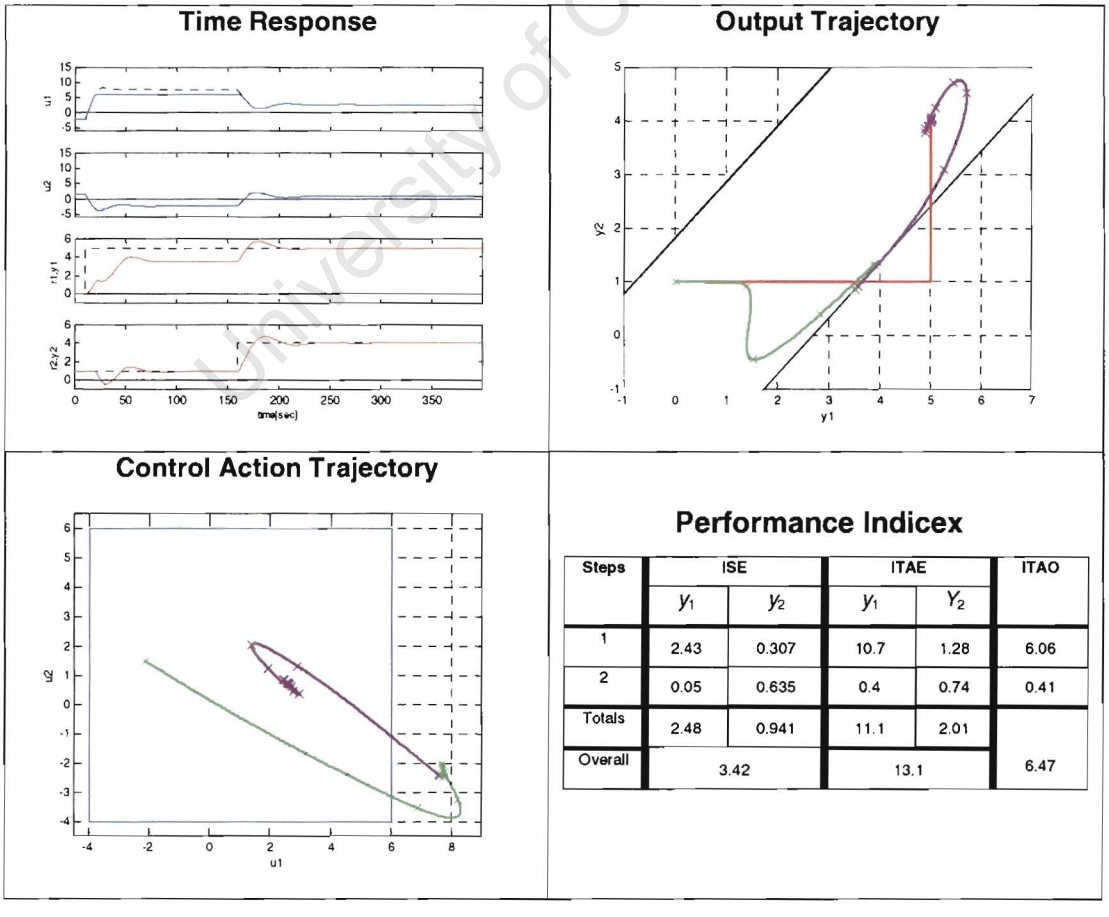
B. Other AWBT Compensation Techniques

Using Kothare's framework several other AWBT compensation techniques were applied to the simulated MVTP problem. Those results are included in this appendix for completeness. Each result set is for the same operating point change. The time response, the output and control trajectories and the performance index table are presented.

All techniques are significant improvements over the uncompensated and saturating system. Of the presented techniques, CAW used in the main text produced the best results. It is also clear that the CAW compensation, in the main text, performed better in that y_2 held it's setpoint during the nonlinear operation.

Generalised Conditioning Technique

This technique was an evolution of the Hanus conditioning technique and is discussed in Section 3.1.2.3. From equations (3-18) For this implementation, $\rho = 1$.

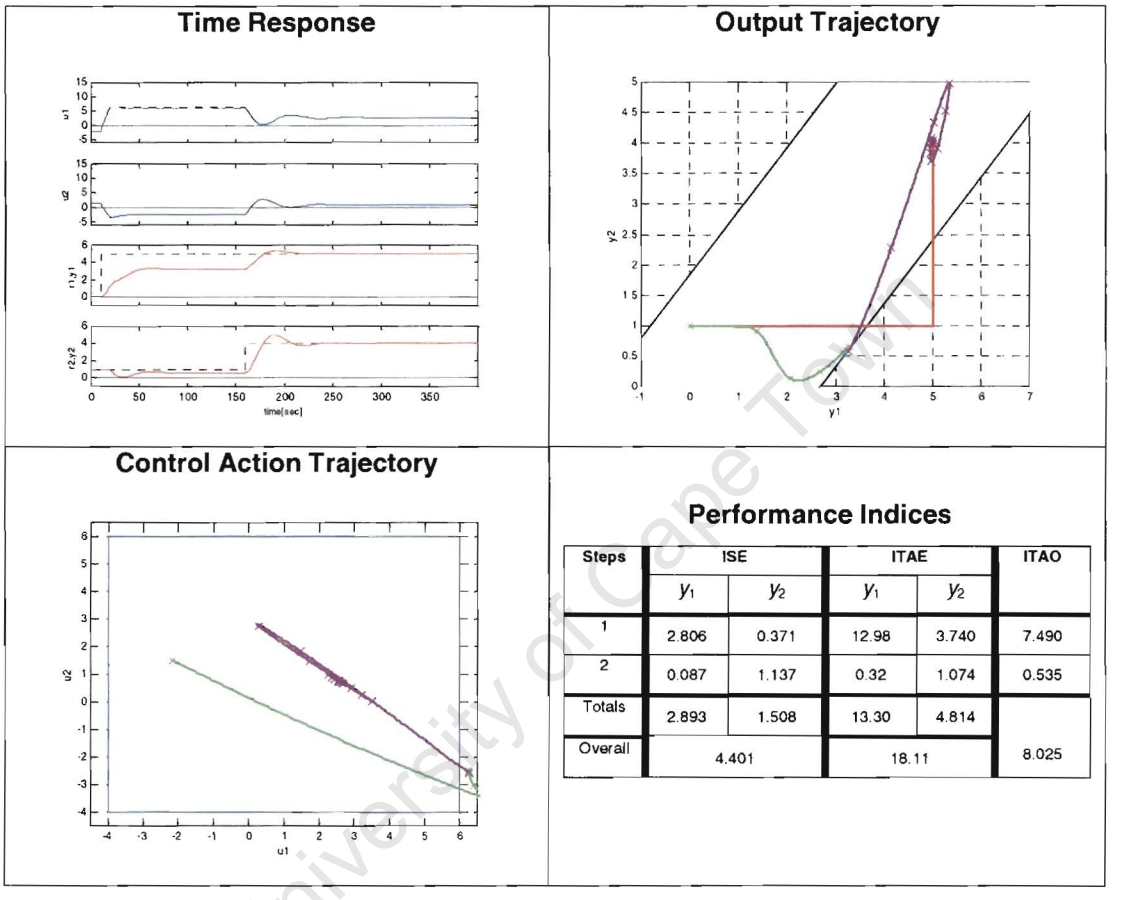


Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	Y_2	
1	2.43	0.307	10.7	1.28	6.06
2	0.05	0.635	0.4	0.74	0.41
Totals	2.48	0.941	11.1	2.01	
Overall	3.42		13.1		6.47

Fig.37 GCT Compensated MVT Process Results

From the control action trajectory it is seen that the compensation does not maintain as tight a correlation between control states and process input; but the time response is significantly improved over HCT, see Section 3.3.2.

Modified Anti-windup with β=1



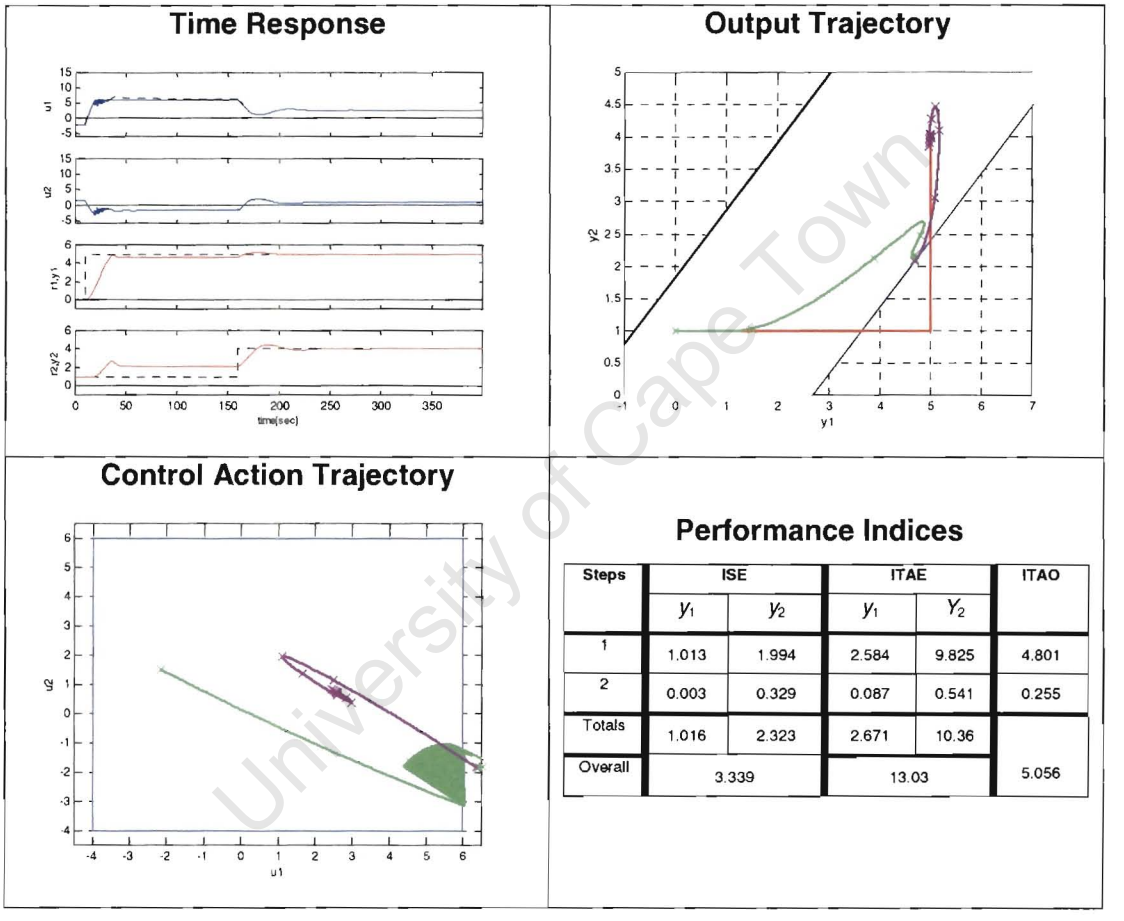
Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	2.806	0.371	12.98	3.740	7.490
2	0.087	1.137	0.32	1.074	0.535
Totals	2.893	1.508	13.30	4.814	
Overall	4.401		18.11		8.025

Fig.38 MAW $\beta = 1$ Compensated MVT Process Results

C. MVTP ER Compensated Simulations

The results are for using the same ER implementation as discussed in Section 4.6. Various combinations of AW techniques and ER variations are given here for completeness and to compliment the discussion in Chapters 4 and Section 5.1. Note that none of the performance indices are weighted.

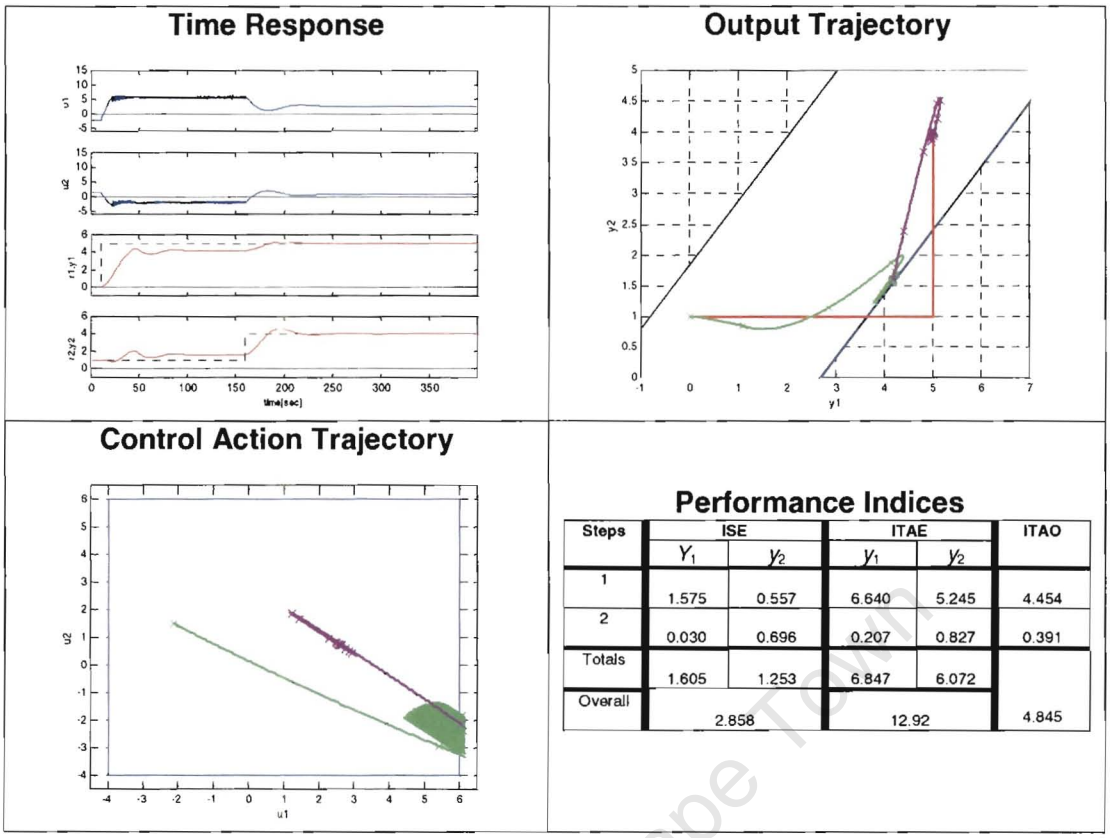
Pairing of AW techniques with ER



Performance Indices

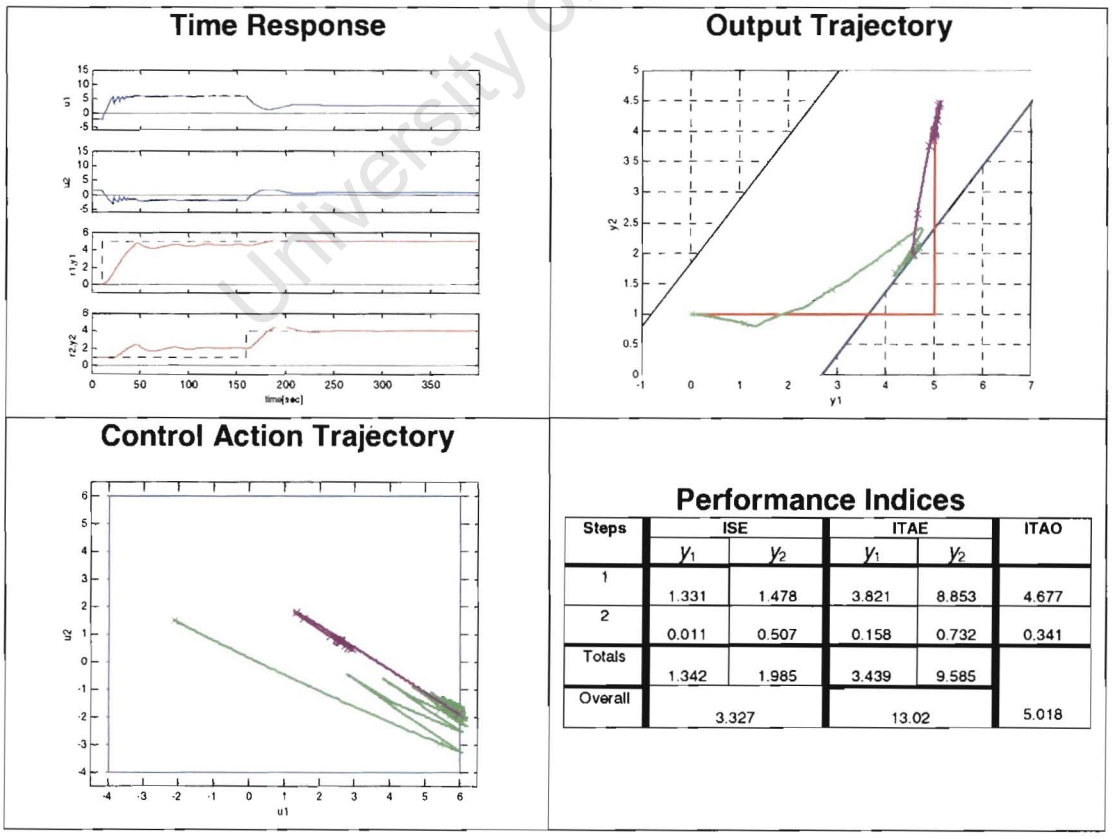
Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	Y_2	
1	1.013	1.994	2.584	9.825	4.801
2	0.003	0.329	0.087	0.541	0.255
Totals	1.016	2.323	2.671	10.36	
Overall	3.339		13.03		5.056

Fig.39 GCT-ER Compensated MVT Process Results



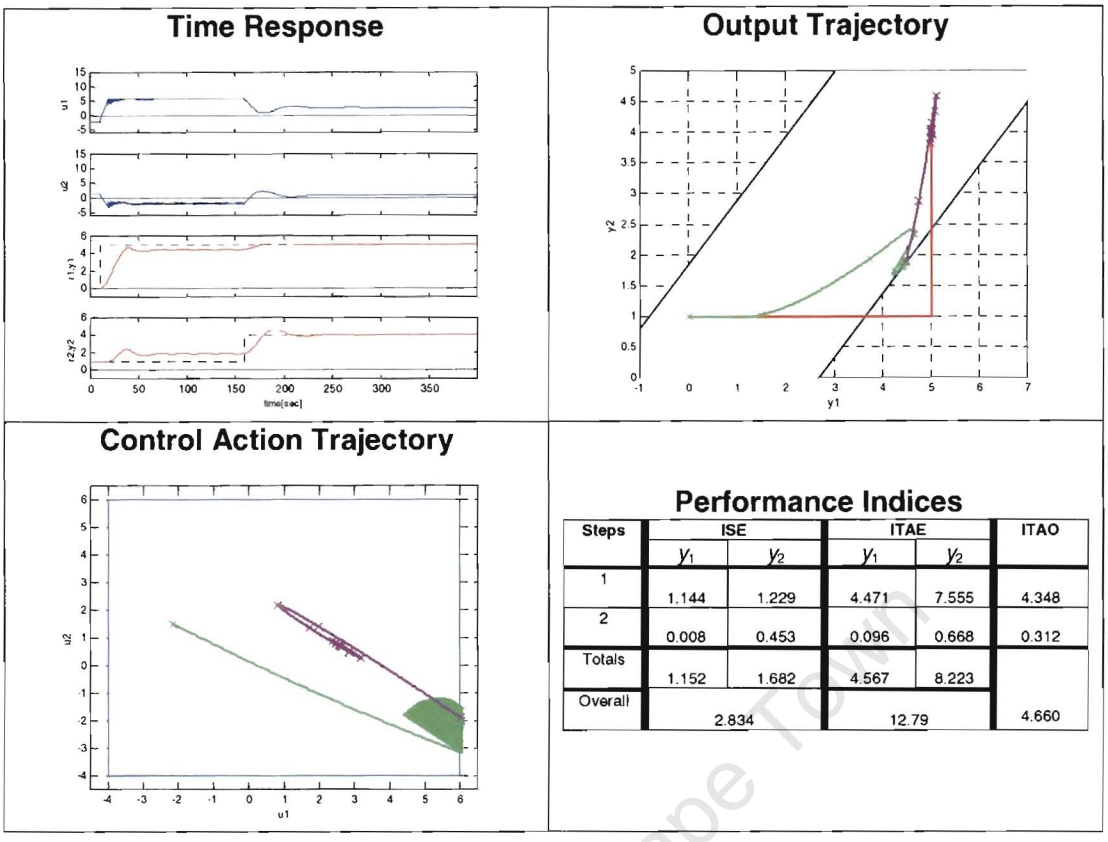
Steps	ISE		ITAE		ITAO
	Y_1	y_2	y_1	y_2	
1	1.575	0.557	6.640	5.245	4.454
2	0.030	0.696	0.207	0.827	0.391
Totals	1.605	1.253	6.847	6.072	
Overall	2.858		12.92		4.845

Fig.40 HCT-ER Compensated MVT Process Results



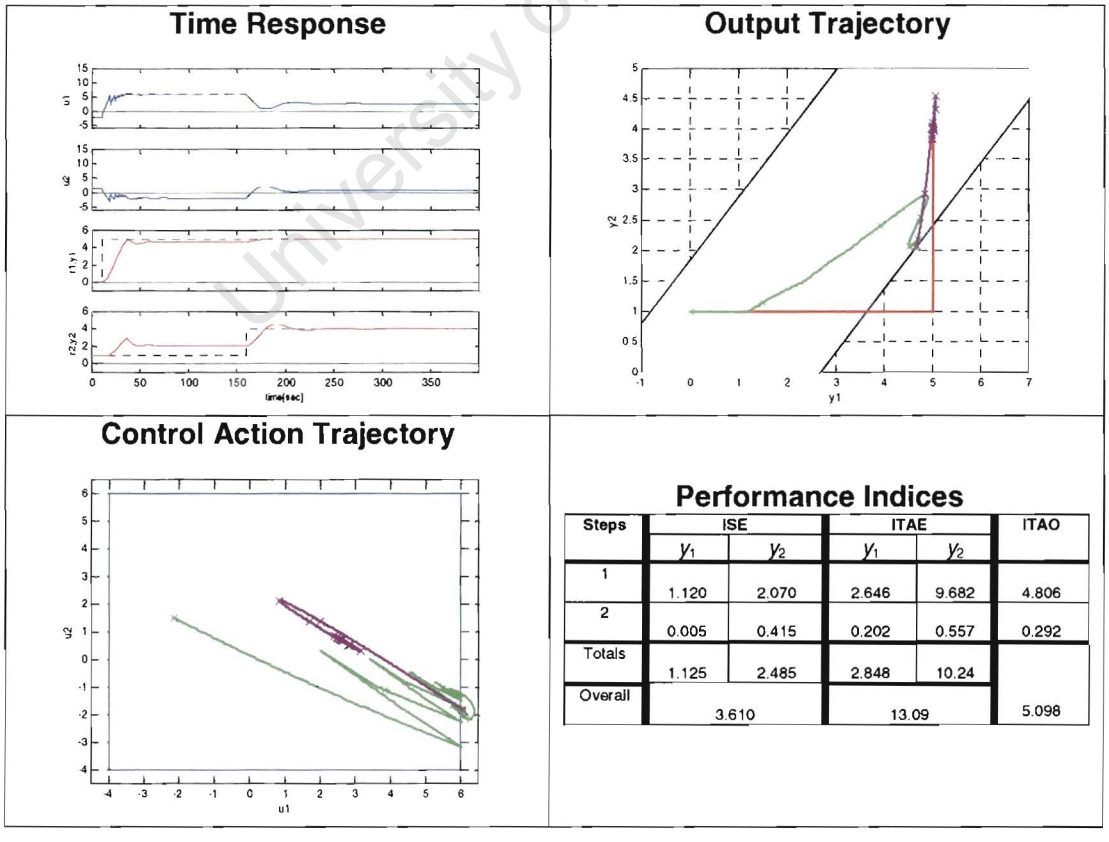
Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.331	1.478	3.821	8.853	4.677
2	0.011	0.507	0.158	0.732	0.341
Totals	1.342	1.985	3.439	9.585	
Overall	3.327		13.02		5.018

Fig.41 HCT-ER (with 1 second hold) Compensated MVT Process Results



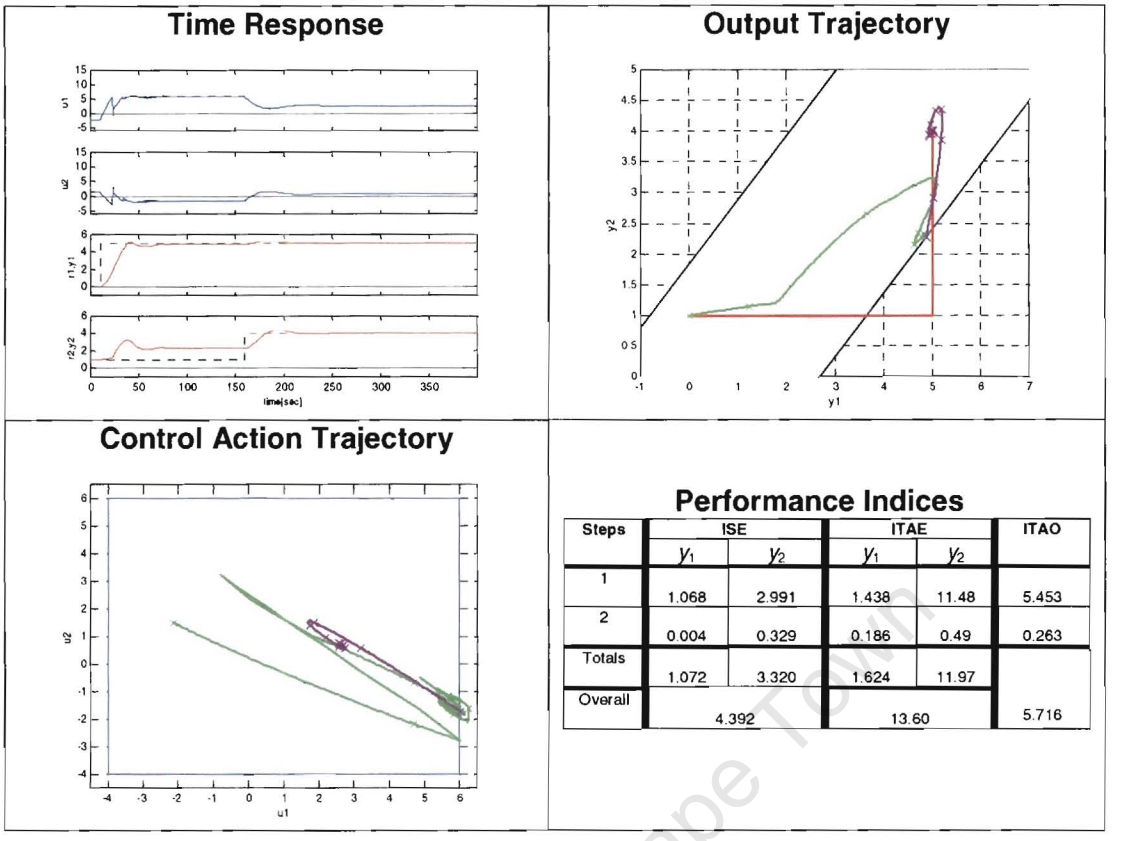
Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.144	1.229	4.471	7.555	4.348
2	0.008	0.453	0.096	0.668	0.312
Totals	1.152	1.682	4.567	8.223	
Overall	2.834		12.79		4.660

Fig.42 MAW-ER Compensated MVT Process Results



Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.120	2.070	2.646	9.682	4.806
2	0.005	0.415	0.202	0.557	0.292
Totals	1.125	2.485	2.848	10.24	
Overall	3.610		13.09		5.098

Fig.43 MAW-ER (with 1 second hold) Compensated MVT Process Results



Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.068	2.991	1.438	11.48	5.453
2	0.004	0.329	0.186	0.49	0.263
Totals	1.072	3.320	1.624	11.97	
Overall	4.392		13.60		5.716

Fig.44 CAW-ER (with 1 second hold) Compensated MVT Process Results

Varying of ER Function

These figures are for a varying ER function value with CAW compensation. The performance indices are in

Table 14 with the discussion of the results in Section 4.6.4.2.

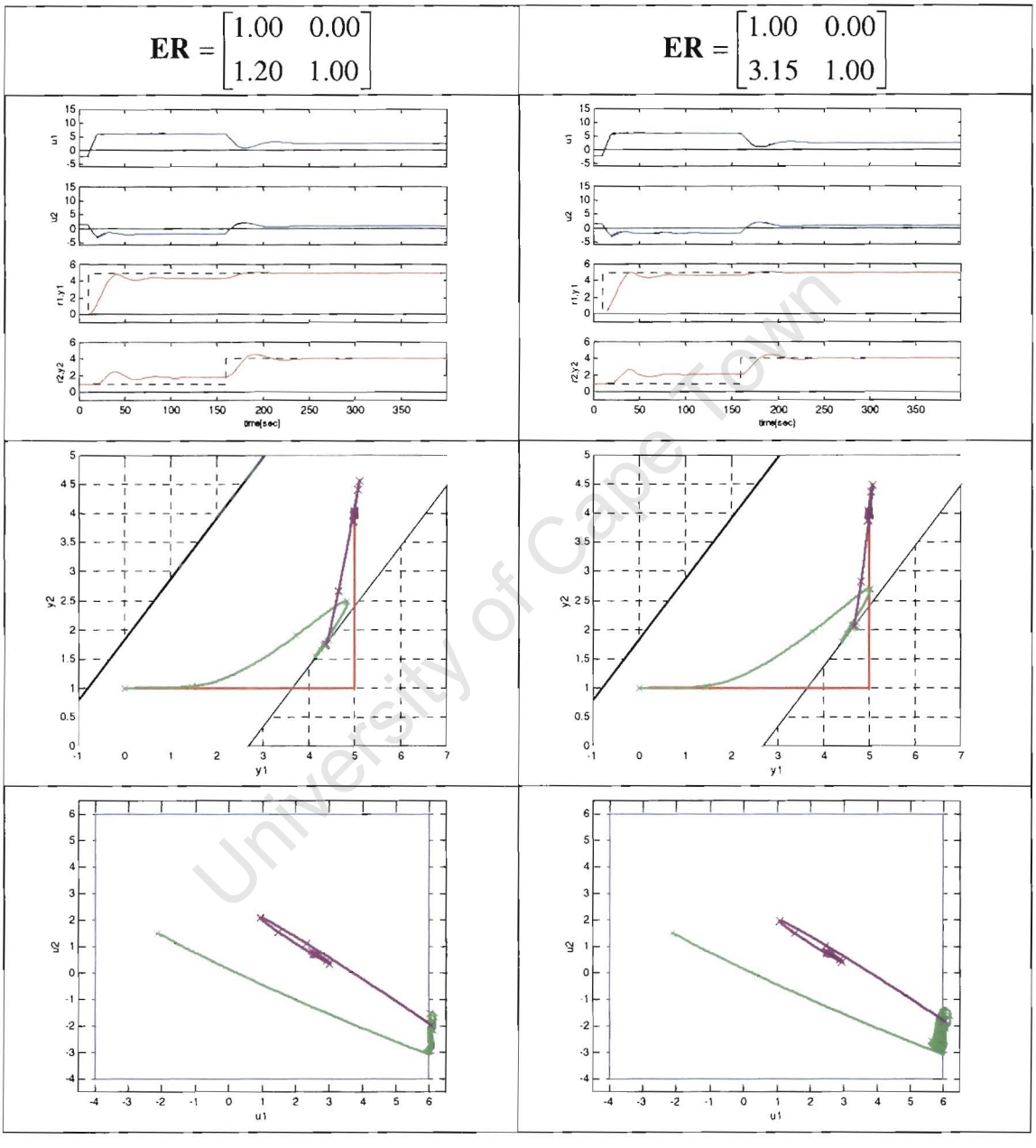


Fig.45 CAW-ER (with $\varepsilon_{21} = 1.20$ and 3.15) Compensated MVT Process Results

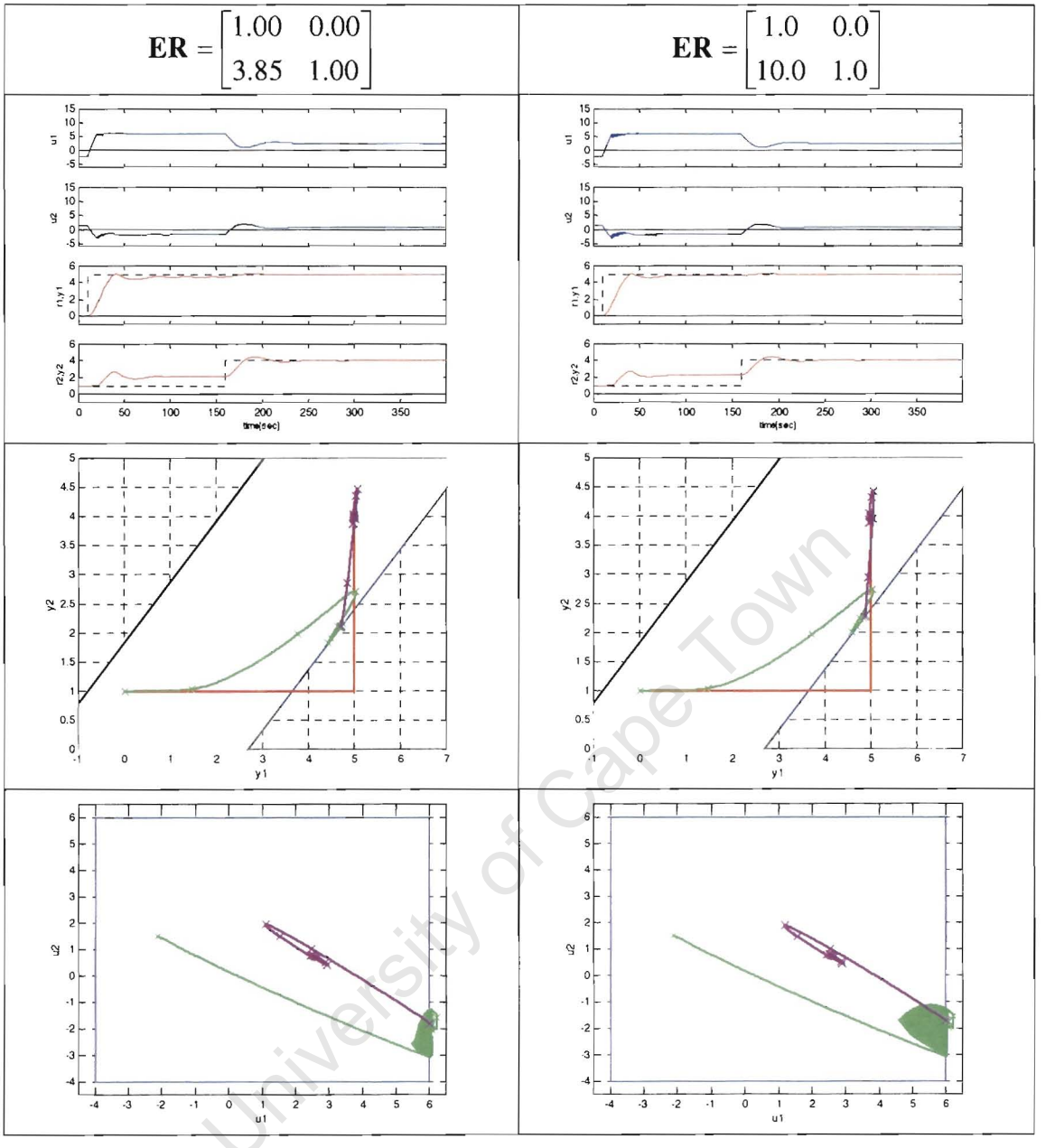
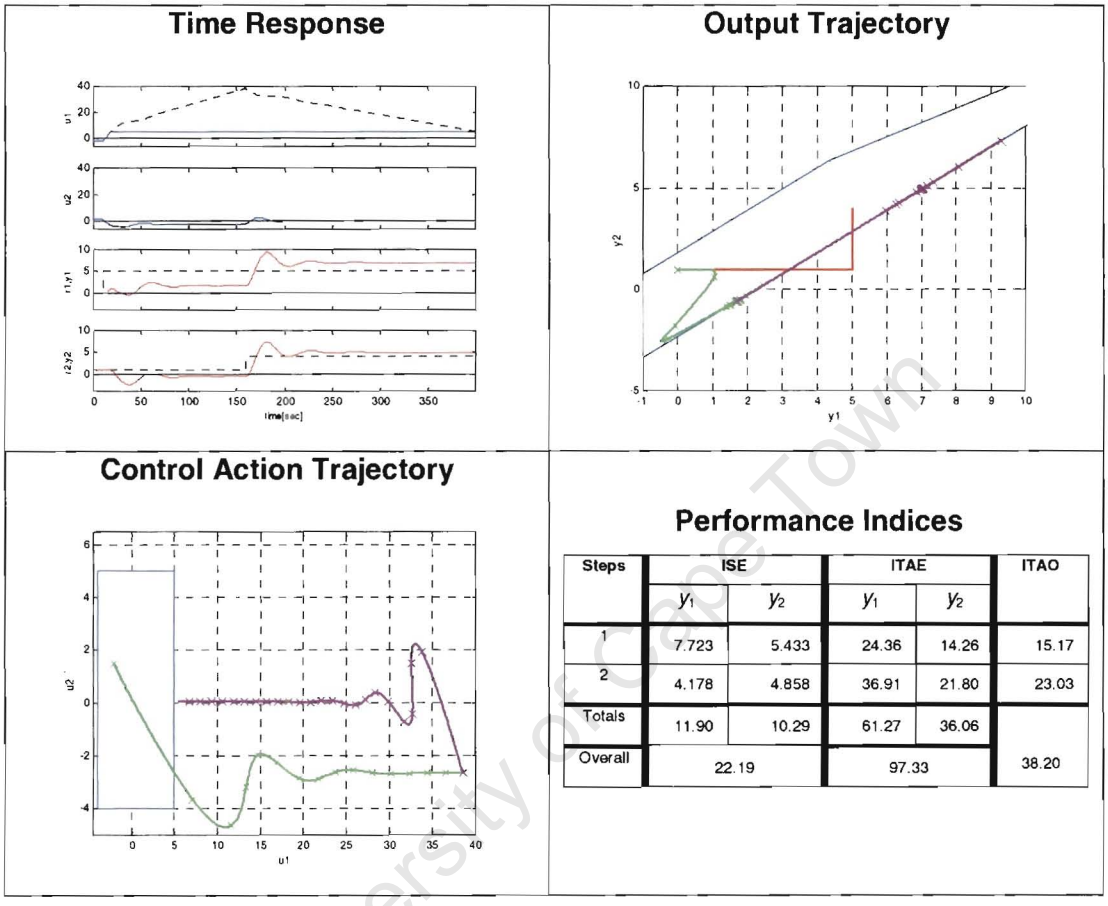


Fig.46 CAW-ER (with $\varepsilon_{21} = 3.85$ and 10.0) Compensated MVT Process Results

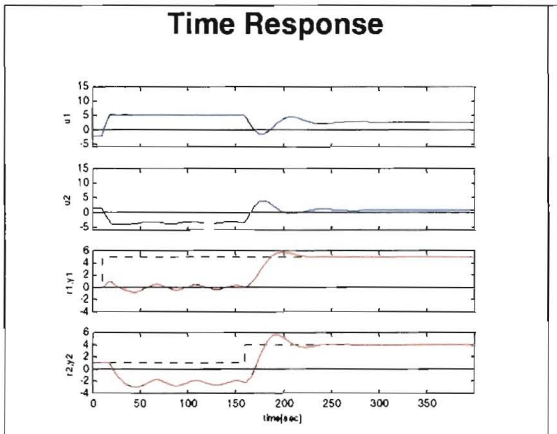
Varying Control Space

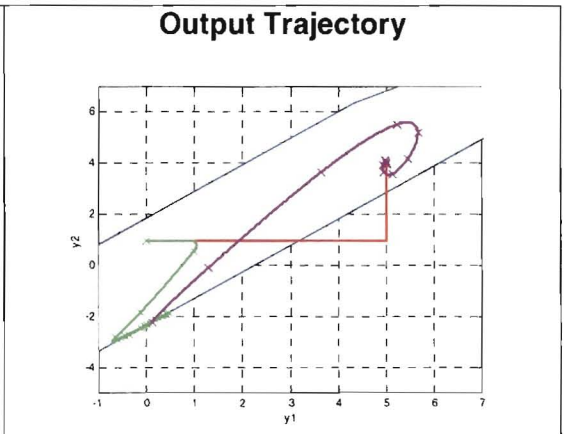
For the MVTP example being discussed in the text, various simulations are presented where the realisable control region has been reduced by 10% in each control signal. This discussion is in Section 5.1.2.

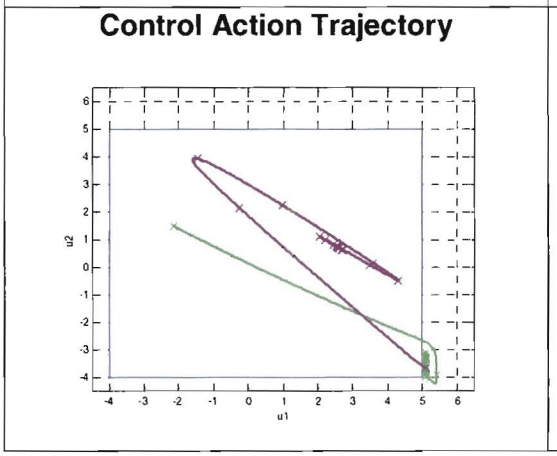


Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	7.723	5.433	24.36	14.26	15.17
2	4.178	4.858	36.91	21.80	23.03
Totals	11.90	10.29	61.27	36.06	
Overall	22.19		97.33		38.20

Fig.47 Saturating MVT Process Results with Reduced Control Region



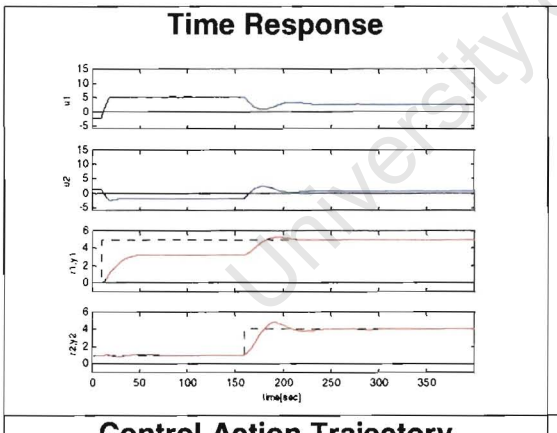


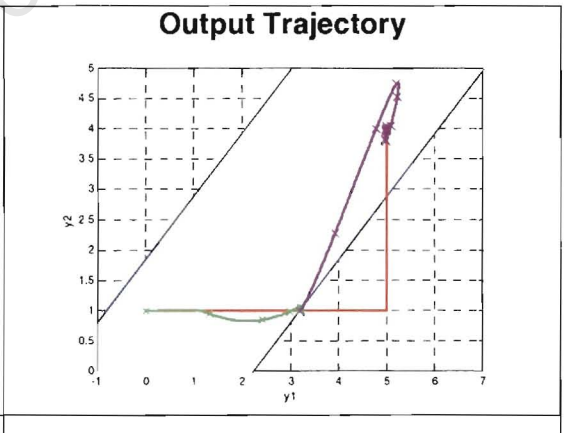


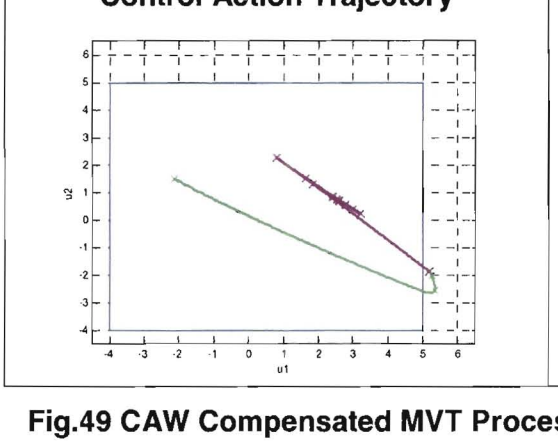
Performance Indices

Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	14.02	17.23	35.82	28.94	24.08
2	1.06	4.60	0.94	2.00	1.13
Totals	15.08	21.83	36.76	30.94	
Overall	36.91		67.70		25.21

Fig.48 HCT Compensated MVT Process Results with Reduced Control Region



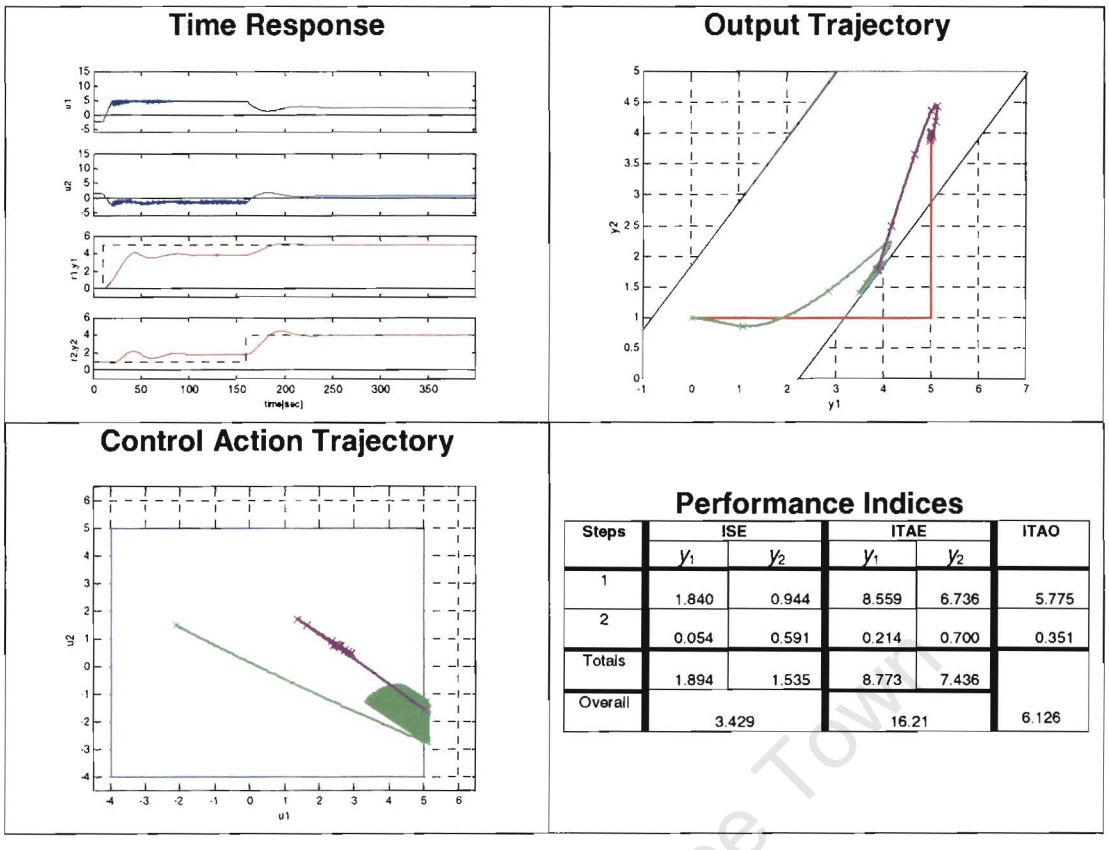




Performance Indices

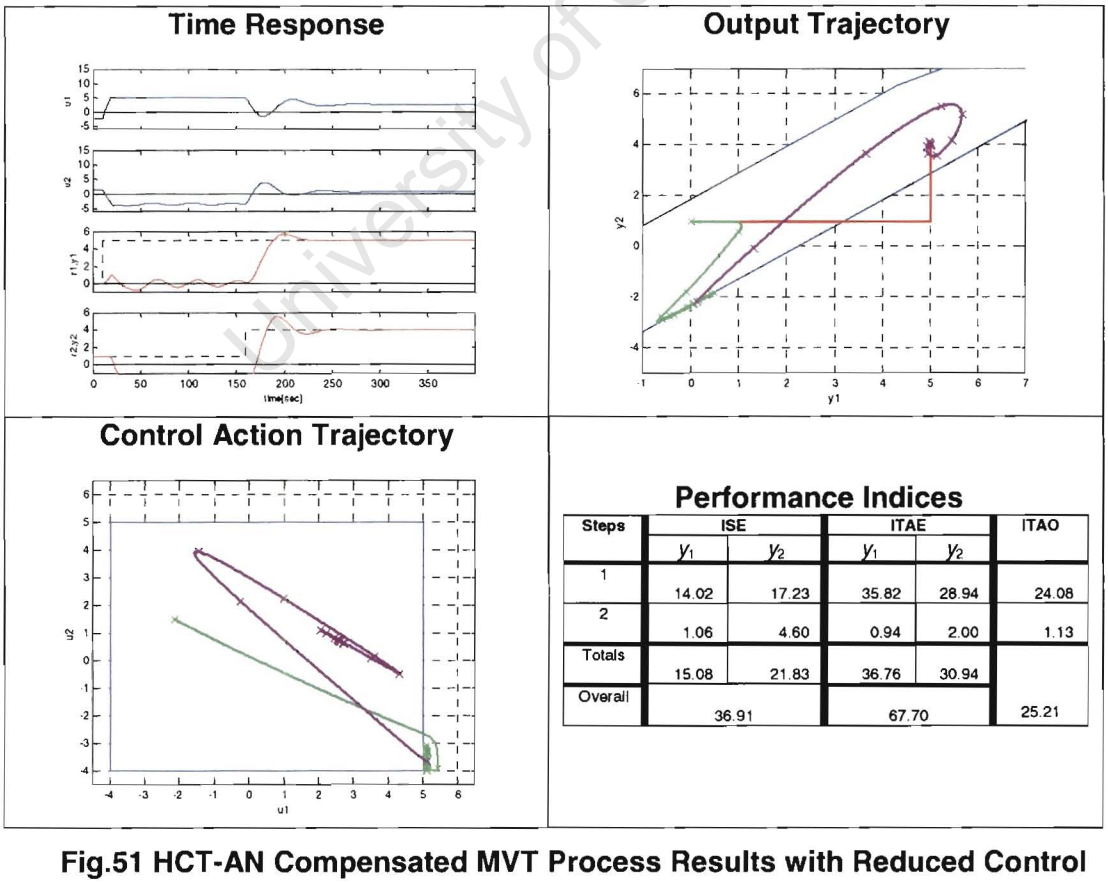
Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	2.762	0.004	13.23	0.108	7.426
2	0.105	0.938	0.28	0.911	0.457
Totals	2.867	0.942	13.51	1.019	
Overall	3.809		14.53		7.883

Fig.49 CAW Compensated MVT Process Results with Reduced Control Region



Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	1.840	0.944	8.559	6.736	5.775
2	0.054	0.591	0.214	0.700	0.351
Totals	1.894	1.535	8.773	7.436	
Overall	3.429		16.21		6.126

Fig.50 HCT-ER Compensated MVT Process Results with Reduced Control Region



Steps	ISE		ITAE		ITAO
	y_1	y_2	y_1	y_2	
1	14.02	17.23	35.82	28.94	24.08
2	1.06	4.60	0.94	2.00	1.13
Totals	15.08	21.83	36.76	30.94	
Overall	36.91		67.70		25.21

Fig.51 HCT-AN Compensated MVT Process Results with Reduced Control Region

D. Milling Circuit Results

This milling circuit example is presented in Section 5.3. In particular, this example is a 3x3 full matrix with significant dead-times on the third output, Particle Size. Because of the dead-times, the pairing of inputs and outputs have non-dominant diagonal elements. This is indicated by the high portion of the performance index value for this output, in the ideal case, coming from the first and second step.

As discussed in the main text, the analysis of this system suggests that it is not ideally suited to ER compensation. The analysis predicts that the ER will be effective in compensation of the second output, *PCD*; but due to the miss-pairing of dominant terms, little benefit is indicated for the third output, Particle Size. A number of the operating point changes presented here were chosen to demonstrate these limitations.

There are seven operating points (numbers 2 to 8) presented here, each set of results includes:

- The resultant graph of the closed loop step response for each type of system: Ideal, Saturating, AWBT and ER compensated.
- The three normalised comparative performance index tables for ISE, ITAE and ITAO indices.

Each response set is accompanied by a brief commentary highlighting the significant aspects of the operating point changes and the effectiveness of the compensation techniques.

Overall, as the discussion in Section 5.3 concluded, the use of ER compensation, where y_1 is degraded to optimise y_2; and where y_1 and y_2 are being degraded in an attempt to optimise y_3 when the effect on y_3 is expected to be so small, is not recommended. The possibility of a more complex ER function that would through it's own states deal better with the large dead-times in this system has not been excluded.

Response Set for Operating Point 2

$$r_1^2(t)=\begin{cases}0.0 & t<100\\ 10 & t\geq 100\end{cases},\ r_2^2(t)=\begin{cases}0.0 & t<5000\\ 10 & t\geq 5000\end{cases}\ \text{and}\ r_3^2(t)=\begin{cases}0.0 & t<10000\\ 2 & t\geq 10000\end{cases} \qquad (-)$$

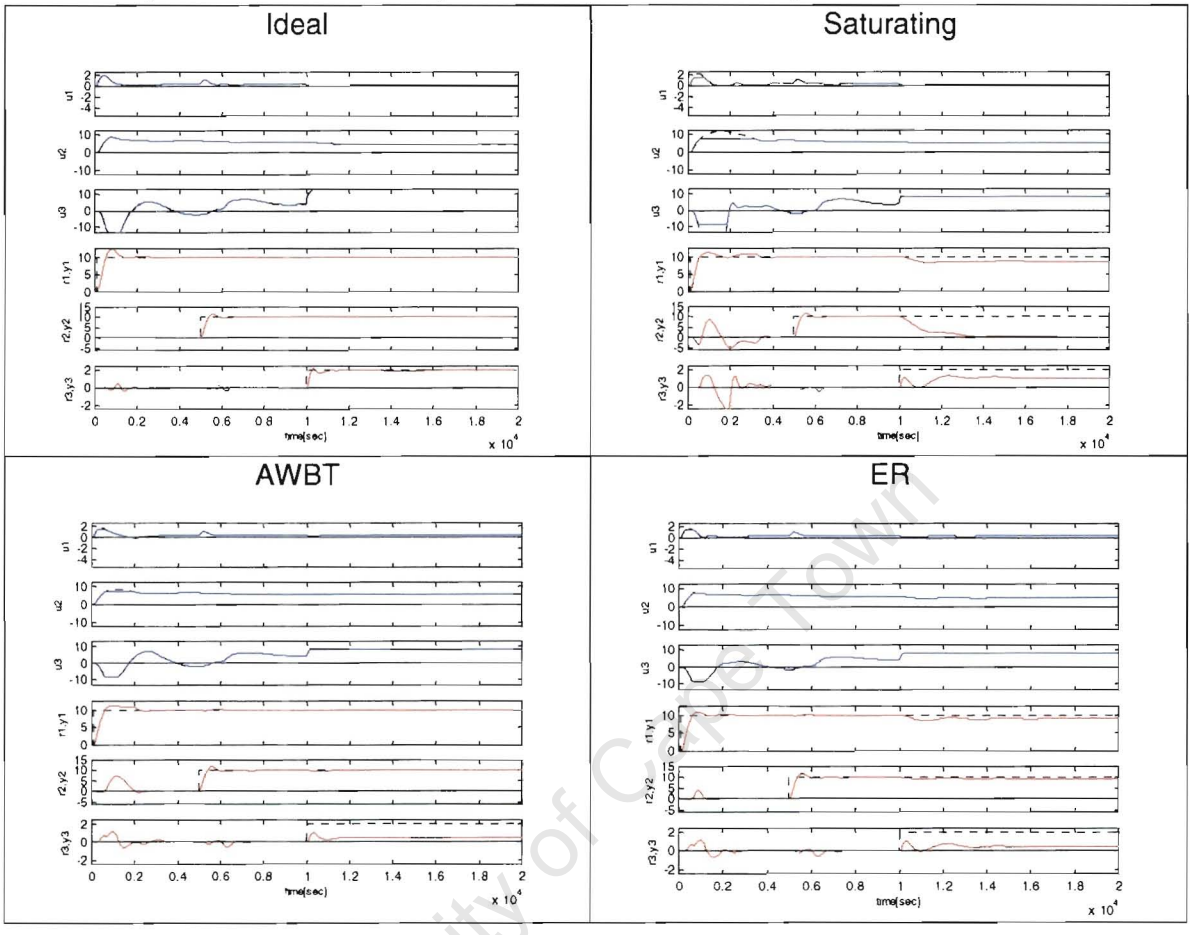


Fig.52 Milling Circuit Time Responses: Operating Point Set 2

Table 37 Milling Circuit Performance Index ISE (OP2)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.333	0.323	0.346	0.332	0.000	2.26	1.49	0.180	0.131	5.5	1.1	1.1
2	0.000	0.000	0.000	0.000	0.667	0.66	0.69	0.695	0.063	0.1	0.2	0.2
3	0.000	0.316	0.000	0.178	0.000	31.2	0.00	0.110	0.806	15.9	34.0	33.6
Total	0.333	0.639	0.347	0.510	0.667	34.1	2.18	0.985	1.00	21.5	35.3	34.9

Table 38 Milling Circuit Performance Index ITAE (OP2)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.326	0.8	0.467	0.24	0.063	10	4.58	0.9	0.298	3.5	1.6	1.3
2	0.002	0.0	0.035	0.04	0.523	0	0.55	0.6	0.246	0.3	0.6	0.5
3	0.005	12.7	0.006	9.39	0.081	261	0.03	14.4	0.456	34.9	57.5	56.6
Total	0.333	13.5	0.507	9.67	0.667	272	5.16	15.9	1.00	38.8	59.6	58.4

Table 39 Milling Circuit Performance Index ITAO (OP2)

Steps				
	IDEAL	SAT	AWBT	ER
1	0.388	3.2	1.5	1.1
2	0.221	0.3	0.5	0.5
3	0.391	33.7	49.2	48.9
Total	1.00	37.3	51.2	50.5

The ideal response clearly shows some disturbance to the third output with the steps in y_1 and y_2. For small changes in u_1, large changes are induced in u_2 and u_3. The saturating response shows the saturation of these signals greatly reduces the compensation for changes in u_1; and both y_2 and y_3 show significant disturbance to the step in y_1. During step 3, even though only u_3 is saturating, the output of all three signals is significantly degraded.

The performance indices for y_3 show that the saturating case is better than either AWBT or ER compensation. However, since the system is not operating in linear mode and will be likely to take considerable time to return to a linear mode on the changing of the operating point, this would not be a desirable mode of operation. Also it is interesting to note the distribution of the ideal performance index, indicating that the system has not been totally decoupled.

AWBT compensation improves the performance over all three steps. The ER compensated system obtains marginal benefits for y_3 in all performance indices but at considerable cost to the performance of y_2. This poor performance in y_3 was suggested by the analysis in Section 5.3.2.

Response Set for Operating Point 3

Again in the saturating case significant disturbance to the stepping y_1 is shown in y_2 and y_3. During the final state, all outputs lose their setpoints.

AWBT compensation maintains overall stability and outputs y_1 and y_2. ER compensation is effective in reducing the disturbances of the step in y_1 to y_2. In the attempt to reduce the steady-state error in y_3 after step 3, y_1 and y_2 are shifted hence ITAE for y_1 and y_2 are increased. This also introduces transients in the outputs increasing the ISE numbers for step 3 over those for AWBT compensation.

ER compensation is acting as intended, however its effectiveness during step 3 is marginal.

$$r_1^3(t)=\begin{cases}0.0 & t<100\\ 10 & t\geq 100\end{cases},\ r_2^3(t)=\begin{cases}0.0 & t<5000\\ 10 & t\geq 5000\end{cases}\ \text{and}\ r_3^3(t)=\begin{cases}0.0 & t<10000\\ -2 & t\geq 10000\end{cases} \quad (-)$$

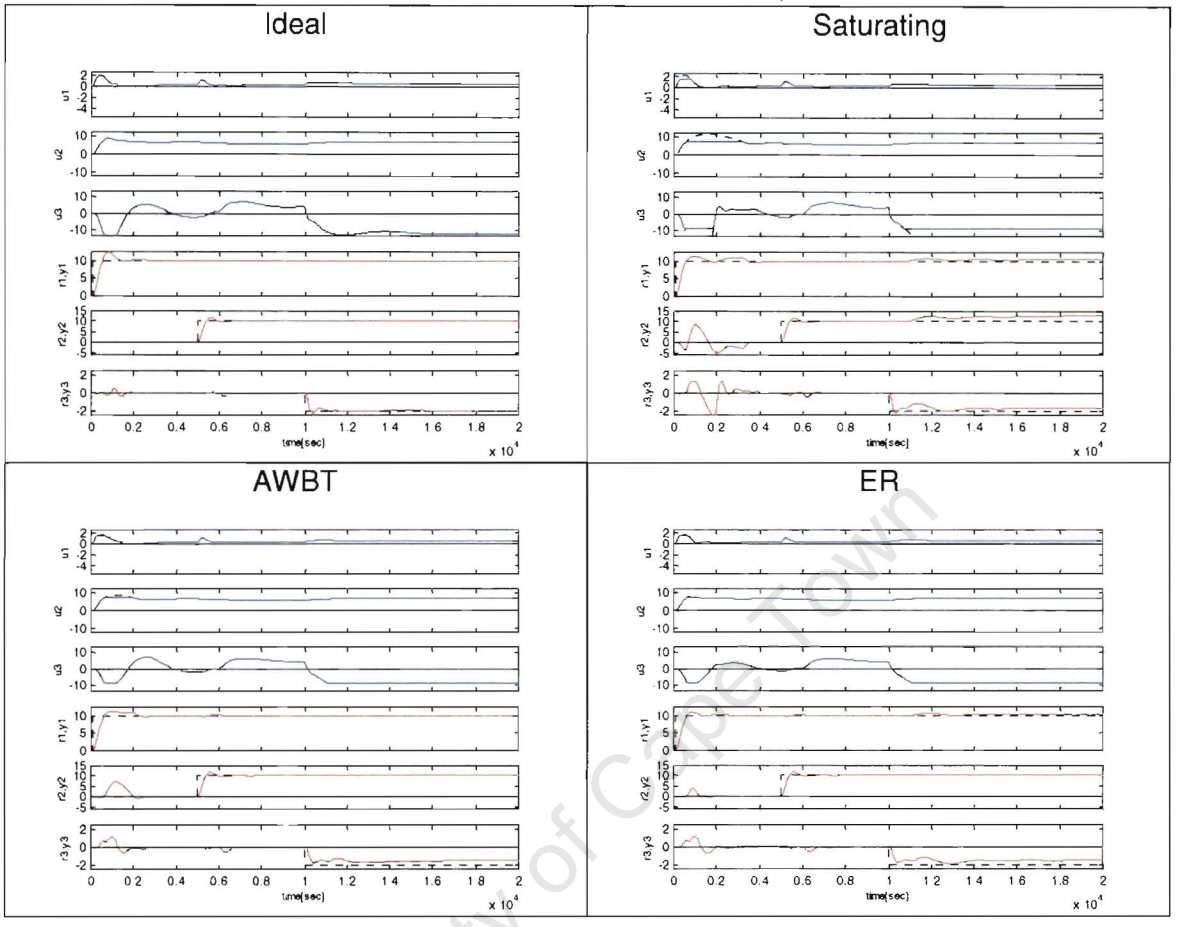


Fig.53 Milling Circuit Time Responses: Operating Point Set 3

Table 40 Milling Circuit Performance Index ISE (OP3)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.333	0.323	0.346	0.332	0.000	2.26	1.49	0.180	0.129	5.39	1.10	1.11
2	0.000	0.000	0.000	0.000	0.666	0.66	0.69	0.695	0.062	0.08	0.16	0.15
3	0.000	0.025	0.000	0.015	0.000	2.35	0.00	0.009	0.808	2.21	3.71	3.78
Total	0.333	0.348	0.346	0.347	0.666	5.27	2.19	0.884	1.00	7.67	4.97	5.05

Table 41 Milling Circuit Performance Index ITAE (OP3)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.327	0.85	0.468	0.24	0.063	10.4	4.63	0.94	0.288	3.4	1.5	1.3
2	0.002	0.01	0.035	0.04	0.529	0.6	0.55	0.56	0.238	0.3	0.5	0.5
3	0.004	3.66	0.011	2.69	0.074	74.1	0.08	4.23	0.474	10.1	16.0	15.8
Total	0.333	4.52	0.514	2.97	0.666	85.1	5.26	5.73	1.00	13.8	18.0	17.6

Table 42 Milling Circuit Performance Index ITAO (OP3)

Steps				
	IDEAL	SAT	AWBT	ER
1	0.377	3.14	1.5	1.1
2	0.245	0.29	0.5	0.5
3	0.408	9.74	13.8	13.7
Total	1.00	13.2	15.8	15.3

Response Set for Operating Point 4

$$r_1^4(t)=\begin{cases}0.0 & t<100\\ 10 & t\geq 100\end{cases},\ r_2^4(t)=\begin{cases}0.0 & t<5000\\ -10 & t\geq 5000\end{cases}\ \text{and}\ r_3^4(t)=\begin{cases}0.0 & t<10000\\ -2 & t\geq 10000\end{cases} \qquad (-)$$

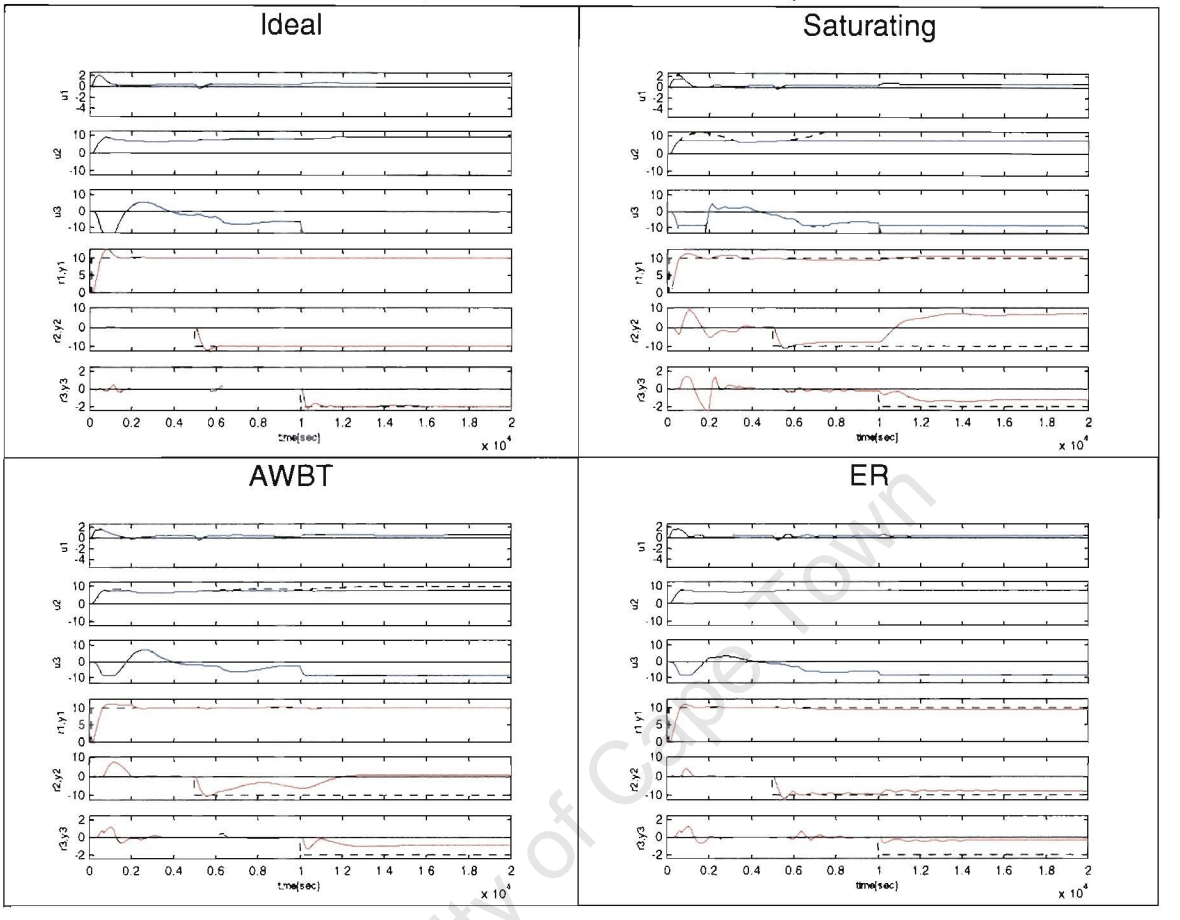


Fig.54 Milling Circuit Time Responses: Operating Point Set 4

Table 43 Milling Circuit Performance Index ISE (OP4)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.333	0.323	0.346	0.332	0.000	2	1.5	0.18	0.131	5.5	1.1	1.1
2	0.000	0.021	0.001	0.019	0.667	1	5.0	0.79	0.063	0.3	0.1	0.3
3	0.000	0.035	0.001	0.080	0.000	100	41.8	2.07	0.806	10.3	18.9	37.4
Total	0.333	0.379	0.348	0.431	0.667	103	48.3	3.04	1.00	16.1	20.1	38.8

Table 44 Milling Circuit Performance Index ITAE (OP4)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.326	0.85	0.466	0.24	0.061	10	5	0.9	0.304	3.6	1.6	1.3
2	0.002	1.22	0.207	1.22	0.524	14	34	5.1	0.227	1.7	0.7	0.9
3	0.005	4.39	0.084	6.42	0.082	449	290	61.7	0.468	26.4	40.8	61.5
Total	0.333	6.46	0.758	7.88	0.667	472	329	67.7	1.00	31.7	43.1	63.7

Table 45 Milling Circuit Performance Index ITAO (OP4)

Steps				
	IDEAL	SAT	AWBT	ER
1	0.395	3.3	1.5	1.2
2	0.205	1.8	2.0	1.3
3	0.400	32.5	37.9	52.7
Total	1.00	37.6	41.4	55.3

The saturating response is consistent with the previous results, where all outputs are degraded during step 1 and 3. In this case, u_2 also saturates during step 2, degrading the outputs during this time too.

With AWBT compensation the system bounds the controller outputs, but a significant discrepancy between the request control action and process input for u_2 suggests that the CAW value for u_2 should be increased.

The inclusion of ER shows y_1 being modified to greatly improve the response for y_2 to step 2, and u_2 is held closer to the desired value. During step 3, y_2 is degraded in an attempt to drive y_3 closer to it's setpoint. In this case compensation for y_3 is very ineffectual.

Overall, ER worked well over step 1 and 2, (sum ITAO = 2.5 AWBT = 3.5) but was out-performed by AWBT on step 3.

Response Set for Operating Point 5

Similar to Operating Point 4's results, significant improvements are seen as additional compensation is included, with the exception of y_3. The state of the controller closer responds to reality and return to linear mode is swifter with the included ER.

$$r_1^5(t)=\begin{cases}0.0 & t<100\\ -10 & t\geq 100\end{cases},\ r_2^5(t)=\begin{cases}0.0 & t<5000\\ 10 & t\geq 5000\end{cases}\ \text{and}\ r_3^5(t)=\begin{cases}0.0 & t<10000\\ 2 & t\geq 10000\end{cases} \quad (-)$$

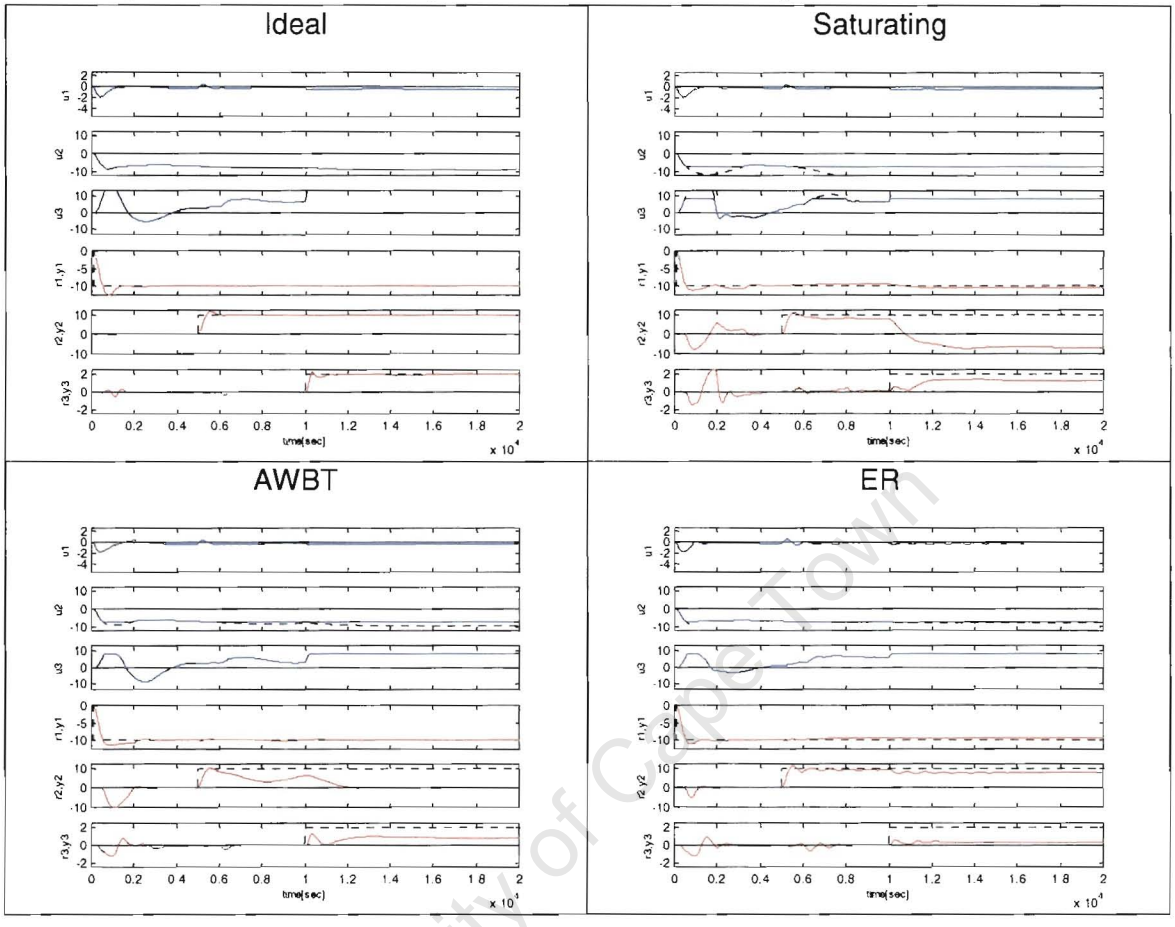


Fig.55 Milling Circuit Time Responses: Operating Point Set 5

Table 46 Milling Circuit Performance Index ISE (OP5)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.333	0.319	0.348	0.324	0.000	2	3.0	0.29	0.131	6.1	1.4	1.4
2	0.000	0.021	0.001	0.019	0.667	1	4.9	0.79	0.063	0.3	0.1	0.3
3	0.000	0.042	0.001	0.072	0.000	102	38.8	2.06	0.806	11.0	20.7	39.2
Total	0.333	0.382	0.350	0.415	0.667	105	46.7	3.13	1.00	17.8	22.2	40.9

Table 47 Milling Circuit Performance Index ITAE (OP5)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.326	0.80	0.538	0.24	0.061	10	6	1.0	0.304	4.0	1.7	1.5
2	0.002	1.24	0.195	1.22	0.524	13	34	5.1	0.227	1.7	0.7	0.9
3	0.005	4.82	0.078	6.11	0.082	453	279	61.4	0.468	27.4	43.2	62.9
Total	0.333	6.86	0.811	7.57	0.667	476	319	67.5	1.00	33.2	45.6	65.3

Table 48 Milling Circuit Performance Index ITAO (OP5)

Steps				
	IDEAL	SAT	AWBT	ER
1	0.395	3.6	1.7	1.3
2	0.205	1.9	1.9	1.3
3	0.400	33.3	39.6	53.9
Total	1.00	38.8	43.2	56.6

Response Set for Operating Point 6

$$r_1^6(t)=\begin{cases}0.0 & t<100\\ -10 & t\geq 100\end{cases},\quad r_2^6(t)=\begin{cases}0.0 & t<5000\\ 10 & t\geq 5000\end{cases}\quad \text{and}\quad r_3^6(t)=\begin{cases}0.0 & t<10000\\ -2 & t\geq 10000\end{cases} \qquad (-)$$

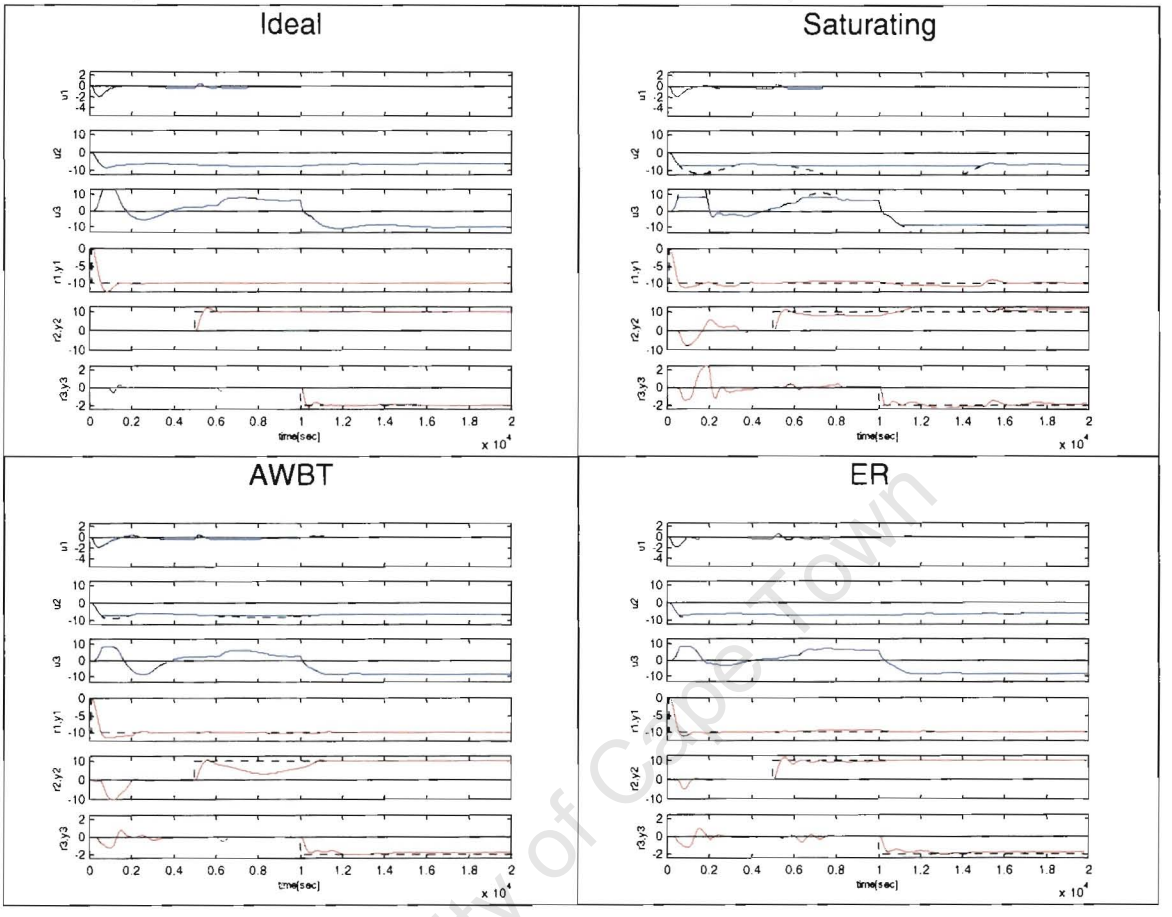


Fig.56 Milling Circuit Time Responses: Operating Point Set 6

Table 49 Milling Circuit Performance Index ISE (OP6)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.333	0.319	0.348	0.324	0.000	2.16	3.00	0.286	0.129	6.06	1.38	1.47
2	0.000	0.021	0.001	0.019	0.667	1.27	4.91	0.788	0.062	0.29	0.13	0.28
3	0.000	0.039	0.001	0.004	0.000	2.27	0.22	0.006	0.808	1.52	1.50	1.65
Total	0.333	0.379	0.350	0.347	0.667	5.70	8.13	1.08	1.00	7.87	3.01	3.40

Table 50 Milling Circuit Performance Index ITAE (OP6)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.327	0.796	0.539	0.236	0.062	10.3	6.0	1.00	0.295	3.89	1.68	1.44
2	0.002	1.25	0.195	1.22	0.532	13.6	34.7	5.16	0.220	1.66	0.64	0.90
3	0.005	3.33	0.064	1.11	0.072	46.6	0.4	1.76	0.485	5.99	6.58	6.72
Total	0.333	5.37	0.798	2.58	0.667	70.5	41.1	7.92	1.00	11.5	8.90	9.06

Table 51 Milling Circuit Performance Index ITAO (OP6)

Steps				
	IDEAL	SAT	AWBT	ER
1	0.384	3.54	1.64	1.29
2	0.200	1.83	1.89	1.31
3	0.416	6.33	5.65	5.82
Total	1.00	11.7	9.17	8.42

The system is operating in a wider plain of the realisable output region and for a nominal disturbance of y_3 during step 2 and 3, the response of y_2 is greatly improved and the steady state response of y_3 similar to the AWBT case.

Response Set for Operating Point 7

This operating point is similar to that requested in Operating Point 2. The difference is that the requested r_3 is 2.5 in this case as compared to 2.0 in OP1. The response for the first two steps is consistent with OP1. However, for step 3 the saturating case experiences a greater degradation of performance.

Note that since the performance indices are normalised the error distribution the numbers for step 1 and 2 are not directly comparable with those given for OP2. For example, in the ideal case in step 1 for OP2, the error during this step for output y3 is ISE= 0.131, which means that 13.1% of the ISE value was produced during this step. In step 1 for OP7, an ISE value of 0.065 indicates that 6.5% of the total ISE error values comes from this step. Since it is known that the systems are identical for this step, the absolute numbers represented by these percentages are equal. Hence, the conclusion is that in OP7 the ideal system experienced a greater total ISE index. This is reasonable since step 3 is greater in OP7 than in OP2 and the transient error is therefore more.

With ER compensation, a significant improvement of y_2's performance is achieved through degradation of y_1. During step 3, degradation of y_1 and y_2 allow a marginal improvement of y_3 .

The further outside of the realisable region the requested operating point is, the greater the degradation in an uncompensated system.

$$r_1^7(t)=\begin{cases}0.0 & t<100\\ 10 & t\geq 100\end{cases},\ r_2^7(t)=\begin{cases}0.0 & t<5000\\ 10 & t\geq 5000\end{cases}\ \text{and}\ r_3^7(t)=\begin{cases}0.0 & t<10000\\ 2.5 & t\geq 10000\end{cases} \quad (-)$$

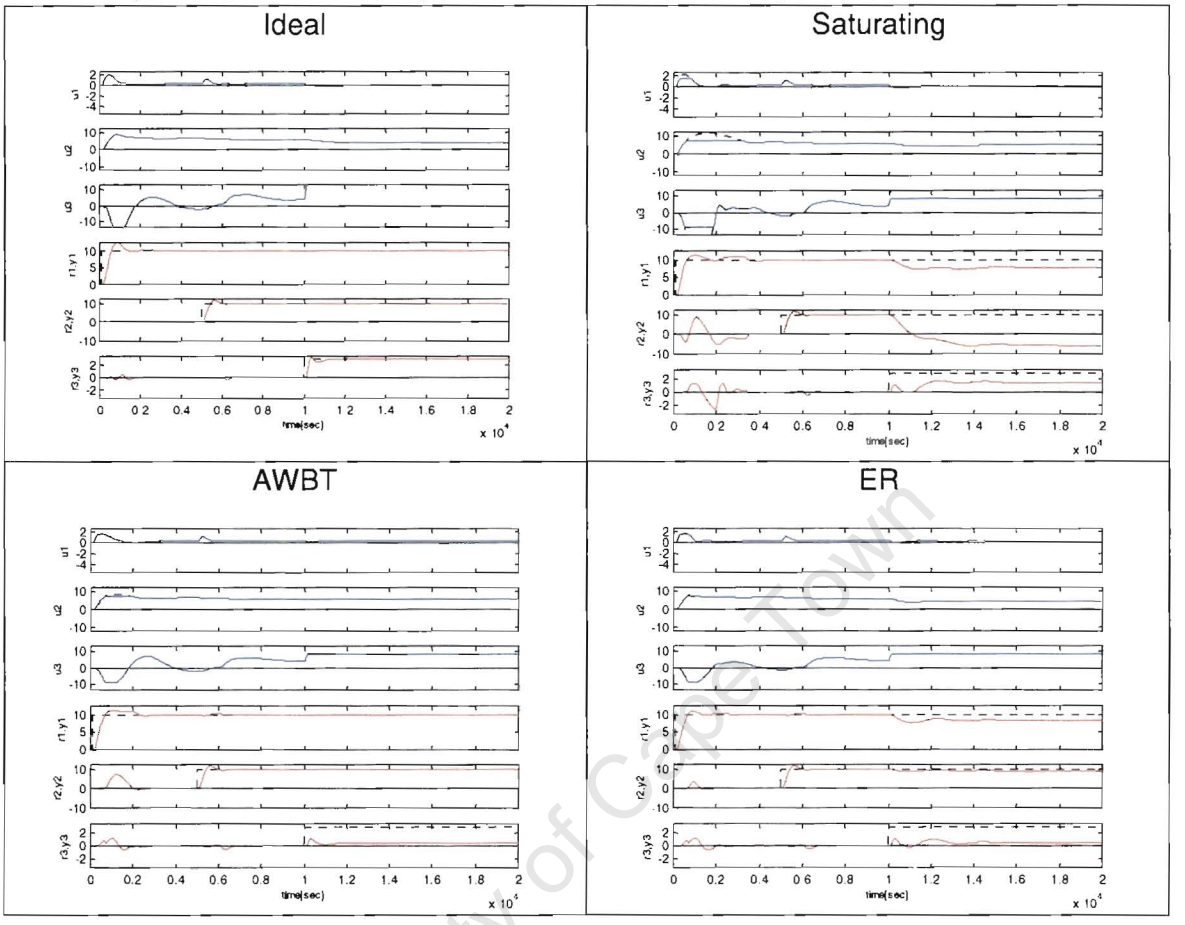


Fig.57 Milling Circuit Time Responses: Operating Point Set 7

Table 52 Milling Circuit Performance Index ISE (OP7)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.333	0.323	0.346	0.332	0.000	2.3	1.49	0.180	0.065	2.7	0.6	0.6
2	0.000	0.000	0.000	0.000	0.667	0.7	0.69	0.695	0.031	0.0	0.1	0.1
3	0.000	0.851	0.000	0.480	0.000	84.5	0.00	0.298	0.904	21.0	45.5	44.9
Total	0.333	1.17	0.347	0.813	0.667	87.4	2.18	1.17	1.00	23.7	46.2	45.6

Table 53 Milling Circuit Performance Index ITAE (OP7)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.324	0.8	0.463	0.2	0.059	10	4.33	0.88	0.241	2.8	1.3	1.1
2	0.002	0.0	0.034	0.0	0.495	0	0.52	0.52	0.199	0.3	0.5	0.4
3	0.007	20.6	0.005	15.3	0.113	406	0.03	22.4	0.560	46.3	76.2	75.0
Total	0.333	21.4	0.503	15.5	0.667	416	4.88	23.8	1.00	49.4	78.0	76.5

Table 54 Milling Circuit Performance Index ITAO (OP7)

Steps				
	IDEAL	SAT	AWBT	ER
1	0.323	2.7	1.2	1.0
2	0.184	0.3	0.4	0.4
3	0.493	46.0	67.1	66.6
Total	1.00	49.0	68.7	68.0

Response Set for Operating Point 8

$$r_1^8(t)=\begin{cases}0.0 & t<100\\ 10 & t\geq 100\end{cases},\quad r_2^8(t)=\begin{cases}0.0 & t<5000\\ -10 & t\geq 5000\end{cases}\quad \text{and}\quad r_3^8(t)=\begin{cases}0.0 & t<10000\\ 2.5 & t\geq 10000\end{cases} \qquad (-)$$

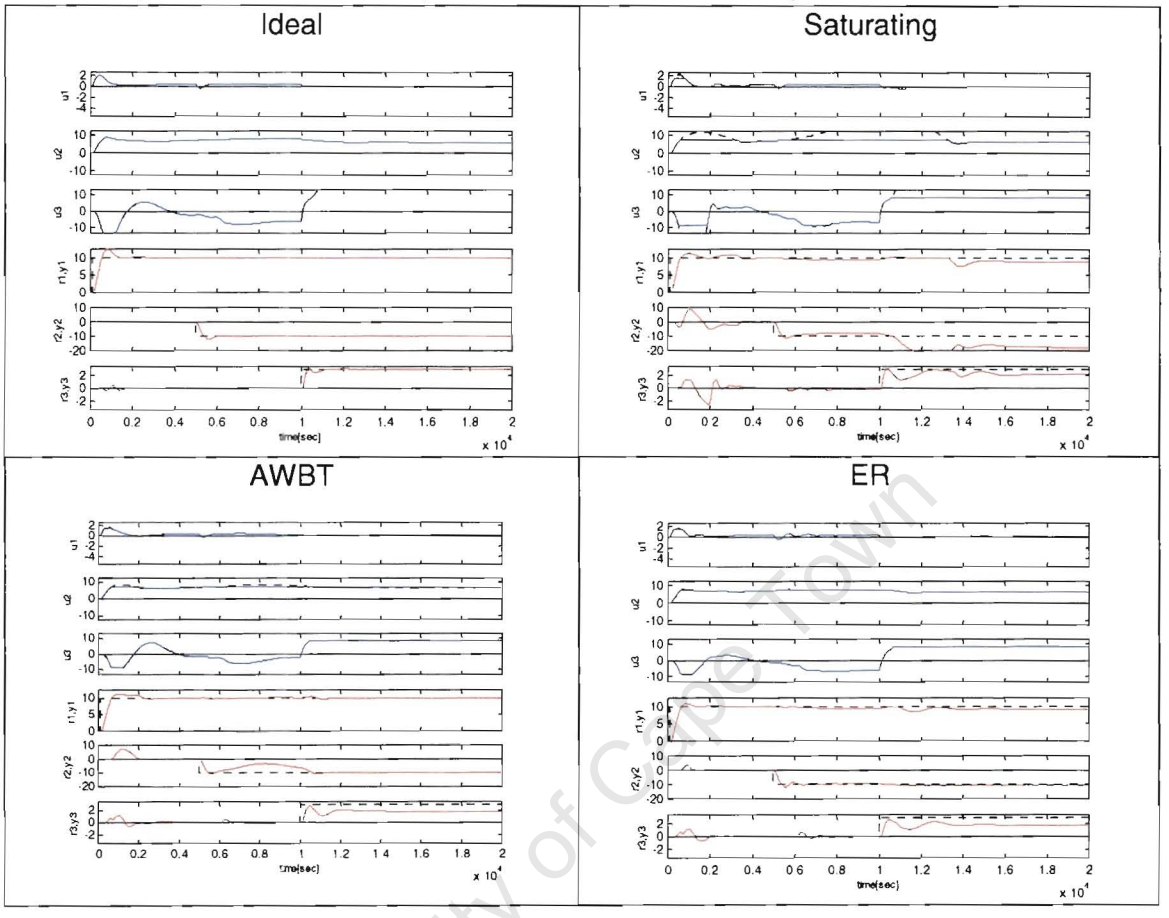


Fig.58 Milling Circuit Time Responses: Operating Point Set 8

Table 55 Milling Circuit Performance Index ISE (OP8)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.333	0.323	0.346	0.332	0.000	2.3	1.49	0.180	0.064	2.69	0.6	0.6
2	0.000	0.021	0.001	0.019	0.667	1.3	5.00	0.789	0.031	0.12	0.1	0.1
3	0.000	0.175	0.002	0.110	0.000	24.4	0.164	0.071	0.905	5.22	10.7	10.8
Total	0.333	0.519	0.349	0.461	0.667	28.0	6.65	1.04	1.00	8.03	11.4	11.5

Table 56 Milling Circuit Performance Index ITAE (OP8)

Steps	y_1				y_2				y_3			
	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER	IDEAL	SAT	AWBT	ER
1	0.324	0.84	0.464	0.24	0.059	10	4.3	0.9	0.239	2.8	1.2	1.0
2	0.002	1.21	0.206	1.21	0.504	13	33.1	4.9	0.179	1.3	0.5	0.7
3	0.007	9.39	0.068	7.26	0.103	201	0.3	10.6	0.528	21.1	35.8	35.4
Total	0.333	11.4	0.738	8.71	0.667	224	37.7	16.4	1.00	25.3	37.6	37.1

Table 57 Milling Circuit Performance Index ITAO (OP8)

Steps				
	IDEAL	SAT	AWBT	ER
1	0.320	2.7	1.2	0.9
2	0.166	1.5	1.6	1.1
3	0.513	21.7	31.6	31.5
Total	1.00	25.9	34.4	33.5

This operating point is similar to that requested in Operating Point 1 in the main text. The difference is that the requested r_3 is 2.5 in this case as compared to 2.0 in OP1. The response for the first two steps is consistent with OP1. However, for step 3 the saturating case experiences a greater degradation of performance.

The implemented AWBT compensation is not able to contain u_2, but during step3, y_2 is realisable and y3 is degraded due to the saturation of u_3.

With ER compensation, a significant improvement of y_2's performance is achieved through degradation of y_1. During step 3, degradation of y_1 and y_2 allow a marginal improvement of y_3.

E. Multivariable Thermal Process

Based on an undergraduate thesis in 1994, Professor Martin Braae suggested the Multivariable Thermal Process (MVTP) as an ideal laboratory process for the development of multivariable control strategies. As part of this thesis, the system has been designed, 10 units built and installed into the Control Laboratory, Electrical Engineering, University of Cape Town, for use in advanced control courses. This appendix gives an overview of the MVTP and touches on some of the underlying issues of operating on a real process.

Overview

The MVTP unit is a modular design, able to have up to 4 thermal sources and sensors on the matrix experiment top, with placing for the sources and sensors for experiment repeatability.

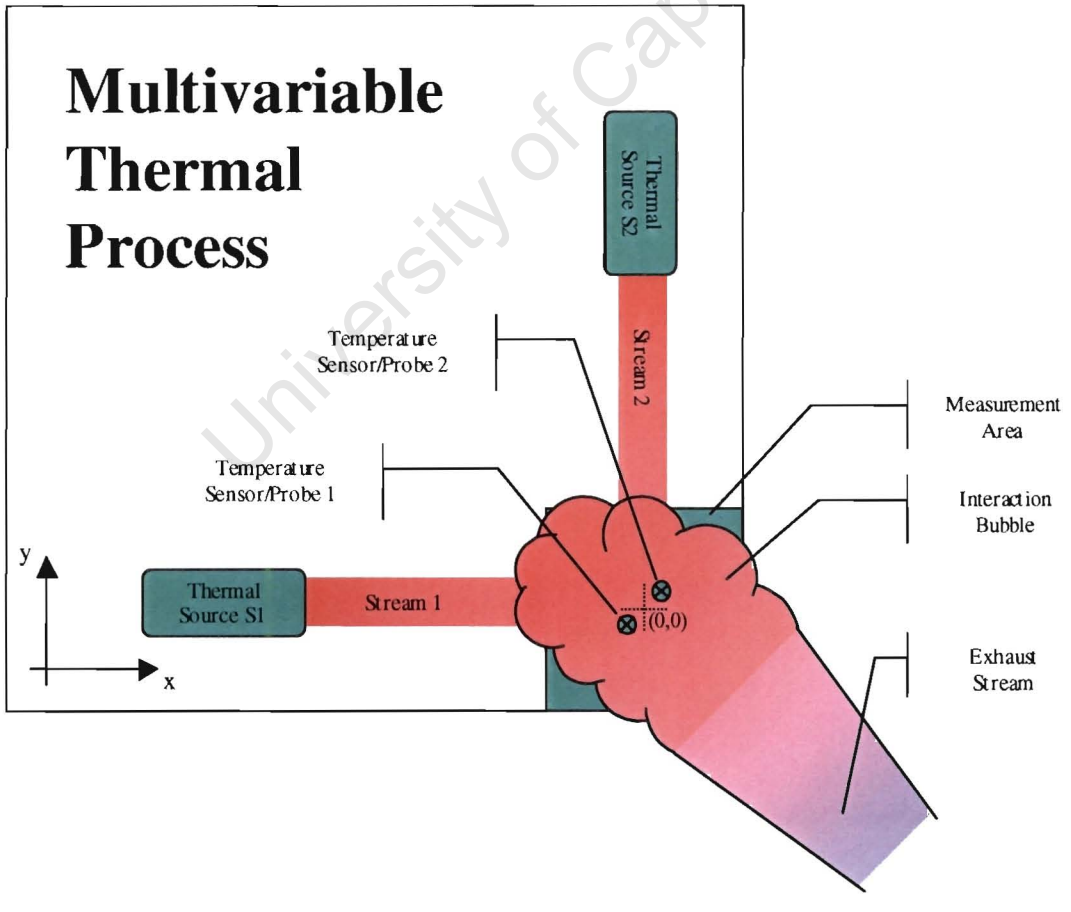


Fig. 59 The Physical Thermal Plant

The system has some key properties for research use:

- It is a clean process;
- Time constants vary from 5 to 30 seconds;
- The system's complexity is a function of the positioning of the sources and sensors relative to one another;
- The system has well defined real world non-linearities, such as dead-time, saturation and hysteresis;
- The system has at least one systematic non-linearity's that can be linearised, if desired;
- The system experiences drift and noise disturbances on a scale comparable to real-world systems.

A systematic view of the system, Fig. 59, shows the typical experimental configuration. As the discussion below will show, this is a complicated thermodynamic system and it's exact theoretical analysis is beyond the scope of this work. The discussion is present as a starting point in understanding the system about which the control examples are based.

Within the thermal source, a fan accelerates the air to the source velocity of $\underline{\mathbf{v}}_s$ and the heating element will, as a function of the control signal, be transferring Q watts of heat into the passing air, increasing the temperature of the air by ΔT_s.

The stream of air leaves the nozzle and travels towards the measurement area. As the streams travel, they lose energy through friction with the surrounding air that is peeling off the outer layer and radiation of heat energy, into the surroundings.

These effects act to reduce the velocity and temperature of the air reaching the measurement area. This also means that the density of the stream at the measurement area will have reduced from what it was at the source; but it will be greater than the surrounding air, which is cooler.

The collision of the two source streams result in a turbulent bubble over the measurement area. An exhaust stream flows away in roughly the direction of the resultant of the sum of the velocity vectors. During the turbulent bubble flow, energy is lost though convection and radiation with the temperature probe, the measurement

area and the environment. The remaining energy propels the air along the exhaust path. Presenting an obstruction in the exhaust path will change the dynamics of the system; increasing the backpressure and the pressure in the bubble, increasing the number of molecules and the time they are in contact with the sensors, and therefore increasing the temperature. A draft in the vicinity of the system can pull air away, reducing the air in the bubble and thereby the temperature.

Modeling the MVTP

A schematic representation of the system components, in Fig. 60, breaks the system into significant components. It is not the objective to give a complete model, but to discuss issues relating to controlling the system.

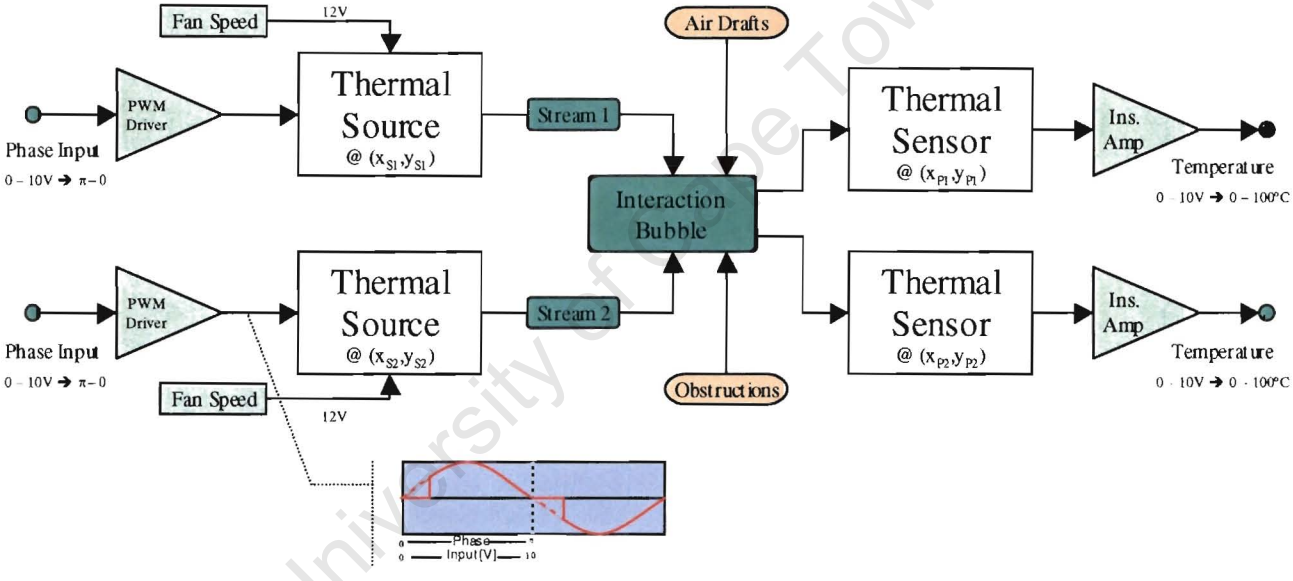


Fig. 60 Schematic of the Multivariable Thermal Process

Typically the Fan speed of the system is held constant, but this is not necessarily so. The drive to the thermal source is a 0-10V signal corresponding to a $\pi - 0$ phase input, which pulse-width-modulates the mains supply as indicated. The effects of the thermal sources are dependent on the ambient temperature, their relative positions and the stream paths. The interaction bubble is affected by air drafts, obstructions, and ambient temperature. The thermal sensors are affected by their position relative to the each other within the interaction bubble, variation in the thermal coupling due to the bubble turbulence, and measurement noise in the electronics. The sensors output a voltage 0-10V for a range of 0 to 100 degrees Celsius.

Fig.61 gives the block diagram summary of the process.

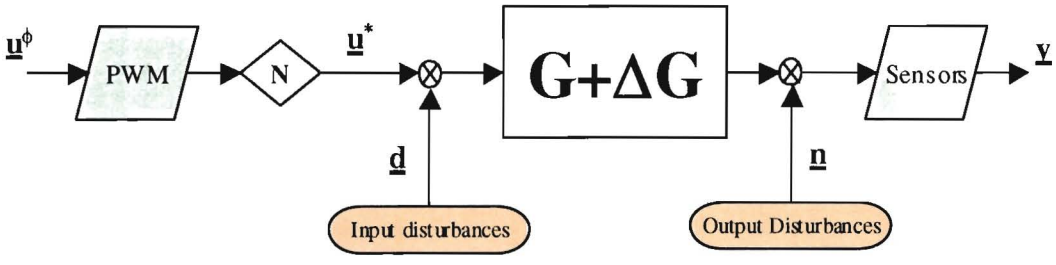


Fig.61 Multivariable Thermal Process Block Diagram

Thermodynamic Approximations

These are first approximations of the systems involved. A more detailed analysis would require a complicated thermodynamic and gas dynamics analysis which is unnecessary for the implementation of the control solution. The subscript "s" indicates parameters as measured at the source and "p" the parameters at the temperature probe.

At the Thermal source:

The heating element is resistive nichrome wire able to dissipate P_s watts of power. Therefore the available heat energy is:

$$Q_s = P_s t \tag{E-1}$$

The time with which the air is in contact with the heating element regulates the heat that can be transferred. The maximum temperature change at the source is envisioned to be ΔT_s. The specific heat capacity of air, the amount of energy required to change the temperature of air, is:

$$C_{air} = 1010\, {}^{J}\!/_{kg.K} \tag{E-2}$$

And the relationship defining the heat transfer is:

$$Q_s = \Delta T_s \cdot M_{air} \cdot C_{air} \tag{E-3}$$

The mass of air being heated depends upon two factors, the volume of the air, V_s, and its density, ρ_s, as it passes the heating element:

$$\begin{aligned} M_{air} &= \rho_s \cdot V_s \\ &= \rho_s \cdot F_s \cdot t \end{aligned} \tag{E-4}$$

Where, F_s is the flow rate at the source in $m^3 \cdot s^{-1}$.

It is clear that at some power, P_s, the increase or decrease of the volume of air passing the element would result in an corresponding increase or decrease in the amount of heat energy transferred, therefore decreasing or increasing the change in temperature. For the experiments conducted in this work, the speed was held constant at a value where at maximum power the temperature change was limited to ΔT_s.

As a first approximation, the ideal gas law allows us to calculate the density of air within a particular volume, at a particular pressure and temperature:

$$pV = nRT \tag{E-5}$$

Where, p is the pressure, V the volume, n the number of moles of gas, $R = 8.31 J \cdot mol^{-1} \cdot K^{-1}$ is the universal gas constant and T is the absolute temperature. With effective molecular mass, the ratio between the mass of air and the number of moles, is:

$$m_{air} = 0.0289 \, {}^{kg}/_{mol} \tag{E-6}$$

The density of air as a function of pressure and temperature can be derived as:

$$\rho_{air} = m_{air} \cdot \frac{n}{V} = m_{air} \cdot \frac{p}{R \cdot T} \tag{E-7}$$

At a given pressure, the cooler the air the more dense it will be, and the greater the mass to pass over the element. On leaving the nozzle of the source, the air will be heated compressed by the acceleration.

Hence the mass of air is:

$$M_{air} = m_{air} \cdot \frac{p}{R \cdot T} \cdot F \cdot t \tag{E-8}$$

Equating equations (E-1) and (E-2), and the substitution of equation (E-8) gives

$$\frac{P}{\Delta T \cdot C_{air}} = m_{air} \cdot \frac{p}{R \cdot T} \cdot F \tag{E-9}$$

Assuming the input air to the thermal source to be at room temperature, $T_A \approx 293K$, that the system is at sea level, $p_s \approx 101kPa$, and the maximum power which can dissipated in the heating element is $P_s = 1200W$; given that the maximum change in temperature is to be $\Delta T_s \approx 50K$, the minimum flow rate is:

$$F_s \geq \frac{P_s \cdot R \cdot T_A}{\Delta T_s \cdot C_{air} \cdot m_{air} \cdot p_s}$$
$$F_s \geq 0.0198\, m^3 \cdot s^{-1} \tag{E-10}$$

The inner diameter of the thermal source is estimated at 4cm, therefore the linear velocity, $\underline{\mathbf{v}}_s$ can be estimated as:

$$v_s = \frac{F_s}{A} \approx \frac{0.0198}{\pi \cdot 0.02^2} = 15.8\, m \cdot s^{-1} \tag{E-11}$$

Solving for ΔT shows that the change in temperature is a function of the ambient temperature and the power being dissipated in the heating element:

$$\Delta T_s = \frac{P \cdot R \cdot T_A}{F_s \cdot C_{air} \cdot m_{air} \cdot p_s} \tag{E-12}$$

Since the heating element is not a point source, but extends over approximately 4cm, a more accurate calculation would integrate this result over the length using the input temperature as a function of position.

Source beams of air:

Along the path to the measurement area, the streams lose energy in friction with the surrounding air, convection and radiation losses and the divergence of the beam reducing the density and therefore the temperature.

$$E_{loses} = E_{friction} - Q_{convection} - Q_{radiation} \tag{E-13}$$

Measurement Bubble:

At and within the bubble, the system is chaotic and subject to the losses described above. However, the main points are that: the net air flow passed the sensors is related to the exhaust velocity vector; and what is being measured by the probe is related to the air passing the probe with a particular density and temperature.

Therefore, the change in probe temperature can be related to its environmental conditions as:

$$\Delta T_p = \frac{P \cdot R \cdot T_p}{F_p \cdot C_{air} \cdot m_{air} \cdot p_p} \tag{E-14}$$

Non-linear effects

A number of additional physical limitations affect the performance of the system to varying degrees. These are discussed in terms of common control nonlinearities.

Dead-time:

The further the sources are from the measurement area, the longer the beam takes to travel the distance, and the more it has slowed down, results in increased time before the effect of any changes in the control action is detected at the probe. The signals changing the power of the thermal source are limited to change every half a cycle of the mains supply. In general, these times are small compared to the time constant of the system and can be ignored.

Hysteresis:

There is a detectable thermal inertia, causing discrepancies between increasing and decreasing step responses. This can be attributed to the heating and cooling of the thermal source housing and changes in the density. Again, this effect is generally small enough that when using a reasonably large controller gain it can be ignored.

Saturation:

The input to the heating element is limited between nothing and the full mains signal. Hence, the energy transfer has both a lower and upper limit. When no heating is taking place, the increased airflow has the effect of reducing the temperature below the ambient. When at maximum power and a fixed flow rate, the maximum temperature will be limited by the ambient temperature of the input and the amount of energy that can be transferred to the air as it passes over the element.

At any relative positioning of the sensors within the interaction bubble, there is a change of the plant dynamics and a maximum steady state thermal gradient that can exist between the two sensors. This occurs when one source is continually blowing cold and the other hot - thus limiting the realisable output space at any time based on the current system condition.

Pulse width modulated Phase to Power linearisation:

The electrical energy is delivered to the heating coil by pulse width modulating (PWM) the mains supply. The actual control input varies between 0 and 10V representing a switching phase between π and 0 radians per mains cycle.

It is necessary to note that the input to the PWM circuit is a phase signal and control signal is proportional to the power and heat dissipated. The relationship between power and phase is derived as follows:

$$P(t) = \frac{V^2(t)}{R} \qquad \text{(E-15)}$$

Where, R is the effective resistance, assuming that the inductive effects of the element are negligible at these low frequencies, and $V(t)$ is the instantaneous voltage across the heating element:

$$V(t) = V_p \sin(\omega \cdot t) \tag{E-16}$$

Expressing $V(t)$ as a function of phase, $\phi = \omega \cdot t$:

$$V(\phi) = V_p \sin(\phi) \tag{E-16}$$

The energy transferred over the interval $[\phi_0, \phi_1]$:

$$\begin{aligned} E(\phi_0, \phi_1) &= \int_{\phi_0}^{\phi_1} \frac{V^2(\phi)}{R} \delta\phi \\ &= \frac{V_p^2}{2R}\left[\phi - \frac{1}{2}\sin(2\phi)\right]_{\phi_0}^{\phi_1} \end{aligned} \tag{E-17}$$

Since the power to the heating element is a triac, which switches off at a zero crossing, the largest continuous on time is one half of a 50/60Hz cycle. So therefore, the rate of energy transferred to the element during phase $\phi_0 \in [0, \pi]$ to $\phi_1 = \pi$ is:

$$P(\phi) = \frac{E(\phi, \pi)}{\pi} = \frac{V_p^2}{2\pi R}\left(\pi - \phi + \frac{1}{2}\sin 2\phi\right) \tag{E-18}$$

It is clear that the power output of the element is not linearly related to the input phase signal. For the control of the temperature, it would be better to design the controller in terms of power and convert from power to phase at the input.

Using an approximation for $\sin\theta$:

$$\sin\theta \approx \begin{cases} \frac{4}{\pi^2}(\theta)(\pi - \theta) & 0 \le \theta \le \pi \\ \frac{4}{\pi^2}(\theta - \pi)(\theta - 2\pi) & \pi \le \theta \le 2\pi \end{cases} \tag{E-19}$$

Let $\theta = 2\phi$ and substitute equation (E-19) into and equation (E-18):

$$E(\phi,\pi)\approx\begin{cases}\dfrac{V_p^2}{2R}\left[\pi+\left(\dfrac{4}{\pi}-1\right)\phi-\dfrac{8}{\pi^2}\phi^2\right] & 0\le\phi\le\dfrac{\pi}{2}\\ \dfrac{V_p^2}{2R}\left[(\pi+4)-\dfrac{9\pi+4}{\pi}\phi+\dfrac{8}{\pi}\phi^2\right] & \dfrac{\pi}{2}\le\phi\le\pi\end{cases} \qquad \text{(E-20)}$$

These quadratic equations in (E-20) can then be solved for ϕ, allowing a mapping of the control action from Power, $\underline{\mathbf{u}}$, to Phase $\underline{\mathbf{u}}^{\phi}$.

References

Åström, K.J. and Hägglund, T. (1988). Automatic Tuning of PID Controllers. Instrument Society of America, Research Triangle Park.

Åström, K.J. and Rundqwist, L. (1989). Integrator Windup and how to avoid it. In *Proc. 1989 Am. Control Conf.*, Pittsburgh, 1693-1698.

Åström, K.J. and Wittenmark, B. (1984). Computer Controlled Systems Theory and Design. Prentice-Hall, Inc., Engelwood Cliffs, NJ.

Biran, A and Breiner, M (1995) MATLAB for Engineers. *Addison-Wesley Publishing Company*

Braae, M. (1996) Digital and State Space Control Theory. *UCT Press.*

Braae, M. (1994) Control Theory for Electrical Engineers. *UCT Press.*

Campo, P.J. Morari, M. and Nett, C.N. (1989). Multivariable anti-windup and bumpless transfer: a general theory. In *Proc. 1989 Am. Control Conf.*, Pittsburgh, 1706-1711.

Campo, P.J. and Morari, M. (1990) Robust control of a process subject to saturation nonlinearities. *Comput. Chem. Engng.*, 14, 343-358.

Carew, W.D. (1996-1997) Multivariable control of a non-linear process. Progress reports no. 1 through 6, Project 243951A, Mintek.

Carew, W.D. (1997) Output Prioritisation in Saturating MIMO Systems by Error Redistribution, Submitted to *Automatica* (97-TB-035)

Carew, W.D. (2001) Output Prioritisation by Error Redistribution in Saturating MIMO Systems. Submitted to *Control Systems Magazine.*

Cherukuri, M.R. and Nikolaou, M. (2001) The equivalence between Model Predictive control and anti-windup control schemes. Submitted to Automatica.

Cherukuri, M.R. (1998) The relationship between model predictive control and anti-windup control schemes. Department of Chemical Engineering, University of Houston.

De Prada, C. and Valentin, A. (1996) Setpoint optimization in multivariable constrained predictive control. In *Proc. IFAC 13th Triennial World Congress*, San Fransisco, 351-356

Doyle, J.C. and Packard, A. (1987). Uncertain Multivariable Systems from a State Space Perspective. In *Proc. 1987 Am. Control Conf.*, Minneapolis, 1034-1039.

Doyle, J.C., Smith, R.S. and Enns, D.F. (1987). Control of Plants with Input Saturation Nonlinearities. In *Proc. 1987 Am. Control Conf.*, Minneapolis, 1034-1039.

Edwards, C. and Postlethwaite, I. (1996) Antri-windup and bumpless transfer schemes. In *Proc. UKACC 1996 Control Conf.*, 394-399

Edwards, C. and Postlethwaite, I. (1998) Antri-windup and bumpless transfer schemes. *Automatica*, Vol. 34, No.2, 199-210

Fertik, H.A. and Ross, C.H. (1967) Direct digital control algorithm with anti-windup feature. Instrument Society of America Reprint, Number 10-1-1CQS-67

Fraleigh, J.B and Beauregard, R.A. (1990) Linear Algebra. Second Edition. *Addison-Wesley Publishing Company*

Gibson, J.E. (1963) Nonlinear Automatic Control. International Student Edition. *McGraw-Hill Book Company*

Hanus, R. (1980). A new technique for preventing control windup. *Journal A*, Vol. 21, 15-20.

Hanus, R. (1980). The conditioned control: a new technique for preventing windup nuisances. Proceedings of the IFIP-ASSOPO conference, 1980. 221-224.

Hanus, R., Kinnaert, M. and Henrotte, J.L. (1987). Conditioning Technique, a General Anti-windup and Bumpless Transfer Method. *Automatica*, Vol. 23, No. 6, 729-739.

Hanus, R. and Kinnaert, M. (1989). Control of constrained multivariable systems using the conditioning technique. Proceedings of the American Control Conference 1989m, 1711-1718

Hanus, R. and Peng, Y. (1992). Conditioning Technique for controllers with time delays. *IEEE Trans. Aut. Control*, 37, 689-692

Hulbert, D.G. and Braae M. (1981). Multivariable Control of a Milling Circuit at East Driefontein Gold Mine. Report No. 2113, National Institute for Metallurgy.

Jeffrey, A. (1990) Linear Algebra and Ordinary Differential Equations. *Blackewell Scientific Publications*

Kapoor, N., Teel, A.R. and Daoutidis, P. (1998). An Anti-Windup Design for Linear Systems with Input Saturation. *Automatica*, Vol. 34, No. 5, 559-574.

Kennedy, J.B. and Neville A.M. (1986) Basic statistical methods for engineers and scientists. Third Edition. *Harper & Row Publishers*

Kosut, R.L. (1983) Design of linear systems with saturating linear control and bounded states. *IEEE trans. on Automatic Control.* Vol. AC-28, No. 1, 121-124

Kramer, L.C. and Jenkins (1971). A new technique for preventing direct digital control windup. Joint Automatic Control Conference Preprints, Paper no 6-E4, 571-577

Kothare, M.V., Campo, P.J., Morari, M. and Nett, C.N. (1994). A unified framework for the study of Anti-windup designs. *Automatica*, Vol. 30, No. 12, 1869-1883.

Kuo, B.C. (1992) Digital Control Systems. International Edition, Second Edition. *Saunders College Publishing*

Larson, J.E. (1994). A simple anti-windup method. *Int. J. Control*, Vol. 60, No 5, 1025-1033.

Maciejowski, J.M. (1989) Multivariable feedback design. *Addison-Wesley Publishing Company*

Marcopoli, V.R. and Phillips, S.M. (1994) Anti-windup analysis and design approaches for MIMO systems. *NASA Contractor Report 195304*, April 1994.

Mattern, D. (1993) A comparison of two multi-variable integrator windup protection schemes. *NASA Contractor Report 194436*, August 1993.

Mayne, D.Q. and Schroeder, W.R. (1994) Nonlinear control of constrained linear systems. *Int. J. Control*, Vol. 60, No 5, 1035-1043

Peng, Y., Vrancic, D. and Hanus, R. (1996). Anti-windup, bumpless, and conditioned transfer techniques for PID controllers. *IEEE Control Systems*, August 1996, 48-57.

Peng, Y., Vrancic, D., Hanus, R. and Wellers, S.S.R. (1998). Anti-windup Designs for Multivariable Controllers. *Automatica*, Vol. 34, No. 12, 1559-1565.

Raven, F.H. (1987) Automatic Control Engineering. Fourth Edition. *McGraw-Hill International Editions*

Rosenbrock, H.H. (1974) Computer-Aided control system design. *Academic Press*

Singh, A. (1999) "Inverse Nyquist Array Design", CBO780 - Multivariable Design 780, University of Pretoria

Singh, M.G. Systems & Control encyclopaedia. *New York Pergamon Press*

Walgama, K.S., Rönnbäck, S. and Sternby, J.(1992). Generalization of conditioning technique for anti-windup compensators. *IEE Proc. D*, 139 109-118.

Wood and Berry, (1973) Wood and Berry Distillation Column, *Chemical Engineering Science*, Vol 28 1707

Vrancic, D. and Peng, Y. (1996) Some practical recommendations in anti-windup design. In *Proc. UKACC 1996 Control Conf.*, 108-113

Vrancic, Damir (1997) Desgin of Anti-windup and Bumpless Transfer Protection, Faculty of Electrical Engineering, University of Ljubljana.

Zheng, A, Kothare, M.V. and Morari, M (1994). Anti-windup design for internal model control. *Int. J. Control*, Vol 60, No 5, 1015-1024

Further Study References

Åström, K.J. (August 2001). Lecture 16 – Controller Structures

Grimm, G., Postlethwaite, I., Teel, A.R., Turner M.C. and Zaccarian, L (June 2001) Linear Matrix Inequalities for full and reduced order anti-windup synthesis In *Proc. 2001 Am. Control Conf.*, Arlington, VA, 4134-4139

Industrial Control, (2000), IGDS 2000, Department of Electrical Engineering, Hong Kong Polytechnic University

Kothare, M.V. and Morari, M (May 1999). Multiplier Theory for Stability Analysis of Anti-windup Control Systems. *Automatica*, Vol 35, No 5, 917-928

Morari. M. and Lee, J.H. (1997). *Model Predictive Control: Past, Present and Future*, Computers and Chemical Engineering, 1997.

Mulder, E.F., Kothare, M.V. and Morari, M (May 2000). Multivariable Anti-windup Controller Synthesis using Bilinear Matrix Inequalities. *European Journal of Control*, Vol 6, No 5, 455-464

Mulder, E.F., Kothare, M.V. and Morari, M (September 2001). Multivariable Anti-windup Controller Synthesis using Linear Matrix Inequalities. *Automatica*, Vol 37, No 9, 1407-1416

Mulder, E.F. and Kothare, M.V. (May 2002). Static Anti-windup Controller Synthesis using Simultaneous Convex Design. In *Proc. 2002 Am. Control Conf., Anchorage, Alaska*

Protuner Application Manual (2002), Techmation

Special issue on "Anti-windup" (2000). *European Journal of Control*, Vol 6, No 5