

# Robust Nonlinear Generalised Minimum Variance Control and Fault Monitoring

Sung-ho Hur and Michael John Grimble

**Abstract:** The first part of this paper extends the Nonlinear Generalised Minimum Variance (NGMV) controller to improve the robustness of its control or set-point tracking performance. This is achieved by replacing the Kalman filter included in the original NGMV controller with an observer to minimise the effect of uncertainty, which includes unknown disturbance, modelling error, and faults. The observer design also allows the NGMV controller to be utilised in fault monitoring. More specifically, the second part of this paper obtains the observer gain by solving a multi-objective optimisation problem through the application of a genetic algorithm so that the residual signal becomes sensitive to faults and insensitive to any other uncertainty. The control and fault monitoring performance of the extended NGMV controllers is tested by application to a nonlinear tank model.

**Keywords:** Nonlinear generalised minimum variance, observer, robust control, fault monitoring.

---

## 1 INTRODUCTION

In order to deal with the effect of nonlinearities, the Nonlinear Generalised Minimum Variance (NGMV) controller was proposed in [1] based on an extension of the well known Minimum Variance (MV) [2] and Generalised Minimum Variance (GMV) [3] controllers, which are suitable only for linear systems. Since then, many papers have been published with improvement and for application to various types of the plant model structure.

This paper is concerned with one of the latest versions introduced in [4]. The structure of the model to which the NGMV controller can be applied includes nonlinear, linear, reference, disturbance, and error weighting subsystems. The Kalman filter is utilised for estimating the states of the combined model of these subsystems, enabling the predicted values of the signal to be obtained. The optimal control law is chosen such that the predicted value approaches zero by minimising the variance of the signal.

The Kalman filter estimation assumes that the disturbance model can be identified online or given in advance. When the disturbance model can be neither identified online nor given in advance, or when any other uncertainty, which includes unknown disturbance, modelling error, and faults, emerges the performance of the Kalman filter, and hence the NGMV controller, could degrade significantly. As a solution, the Kalman filter is replaced with an observer whose gain is obtained in the frequency domain such that the NGMV controller becomes less sensitive to

uncertainty allowing it to be more robust and easier to tune. The observer design is based on the well-known Luenberger observer [5, 6] and could be applied to any other control algorithms that utilise an estimator or an internal model such as Linear Quadratic Gaussian (LQG), Model Predictive Control (MPC), and Inverse Model Control (IMC). NGMV control method that is truly nonlinear has not previously been improved from a robustness view point. This is the first attempt to gain good performance of NGMV on nonlinear systems and, at the same time, have reasonable robustness.

Not only does replacing the Kalman filter with the observer improve the control (or set-point tracking) performance, it also enables fault monitoring. The second observer design allows the residual that is the difference between the measurement and the estimate of the output signal to become sensitive to faults but insensitive to any other uncertainty including disturbance and modelling error. Sensitivity to faults and insensitivity to any other uncertainty are typically conflicting problems. Therefore, obtaining an observer gain that strikes a good balance is essential and can be achieved by solving a multi-objective optimisation problem through the application of a genetic algorithm.

The extra computation for the extended NGMV controllers is performed offline and only takes several seconds. The addition of the fault monitoring feature provides the NGMV controller with additional capabilities at only modest cost.

The original NGMV controller is introduced in Section 2, and the proposed improvement to the controller is reported in Section 3. Further extension to the NGMV controller for fault monitoring purposes is described in Section 4. The control and fault monitoring performance of

---

This work was supported by the UK Engineering and Physical Sciences Research Council on the Platform Grant No: EP/F026781/1.

Sung-ho Hur and Michael Grimble are with the department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, UK (e-mail: sung.h.hur@strath.ac.uk).

the extended NGMV controllers is tested by application to a nonlinear tank model in Sections 3 and 4, respectively. Conclusions are drawn in Section 5.

## 2 NONLINEAR GENERALISED MINIMUM VARIANCE CONTROL

This section briefly introduces one of the latest versions of the NGMV family – more details can be found in [4]. The structure of the model that the controller can be applied to is described in Section 2.1, and the NGMV controller is reported in Section 2.2.

### 2.1 System Models

The structure of the plant model to which the NGMV controller can be applied to is of limited generality and carefully chosen so that simple results can be obtained. The plant model can include nonlinear dynamics, but the reference and disturbance signals are assumed to have linear model representations as depicted in Fig. 1. The zero mean white measurement noise  $v(t)$  has a covariance matrix  $R_f$ , and there is no loss of generality in assuming that the zero-mean white noise signals  $\omega(t)$ ,  $\xi_d(t)$ , and  $\xi_0(t)$ , which feed the reference, disturbance, and linear dynamics submodels, have identity covariance matrices. The distribution of the noise sources does not need to be determined because the special structure shown below leads to a prediction equation that is rather dependent on the linear disturbance and reference submodels.

Error, plant, reference, output, observation, and noise error signals shown in Fig. 1 can respectively be summarised as follows:

$$e(t) = r(t) - y(t) \quad (1)$$

$$y(t) = d(t) + (\mathcal{W}u)(t) \quad (2)$$

$$r(t) = W_r \omega(t) \quad (3)$$

$$d(t) = W_d \xi_d(t) \quad (4)$$

$$z(t) = y(t) + v(t) \quad (5)$$

$$e_o(t) = r(t) - z(t) \quad (6)$$

The submodels for a  $r$  by  $m$  Multi-Input and Multi-Output (MIMO) system shown in Fig. 1 can now be mathematically presented as follows.

#### 2.1.1 Reference

$$x_r(t+1) = A_r x_r(t) + B_r \omega(t) \quad (7)$$

$$r(t) = C_r x_r(t) \quad (8)$$

$$W_r(q^{-1}) = C_r(qI - A_r)^{-1} B_r \quad (9)$$

where  $x_r(t) \in \mathbb{R}^{n_r}$ , and  $W_r(q^{-1})$  represents the reference model.  $q^{-1}$  is the delay (shift) operator such that  $x(t) = q^{-1}x(t+1)$ .

#### 2.1.2 Disturbance

$$x_d(t+1) = A_d x_d(t) + B_d \xi_d(t) \quad (10)$$

$$d(t) = C_d x_d(t) \quad (11)$$

$$W_d(q^{-1}) = C_d(qI - A_d)^{-1} B_d \quad (12)$$

where  $x_d(t) \in \mathbb{R}^{n_d}$ , and  $W_d(q^{-1})$  represents the disturbance model.

#### 2.1.3 Linear Dynamics

If the plant includes a linear dynamics submodel that could be either stable or unstable, it could be included as follows:

$$x_0(t+1) = A_0 x_0(t) + q^{-k} B_0 u_0(t) + D_0 \xi_0(t) \quad (13)$$

$$y_0(t) = C_0 x_0(t) + q^{-k} E_0 u_0(t) \quad (14)$$

where  $x_0(t) \in \mathbb{R}^{n_0}$ .

Equations (13) and (14) are equivalent to

$$y_0(t) = q^{-k} W_{0k}(q^{-1}) u_0(t) + W_{0d}(q^{-1}) \xi_0(t) \quad (15)$$

where

$$W_{0k}(q^{-1}) = C_0(qI - A_0)^{-1} B_0 + E_0 \quad (16)$$

$$W_{0d}(q^{-1}) = C_0(qI - A_0)^{-1} D_0 \quad (17)$$

#### 2.1.4 Error Weighting

$$x_p(t+1) = A_p x_p(t) + B_p(r(t) - d(t) - y_0(t)) \quad (18)$$

$$y_p(t) = C_p x_p(t) + E_p(r(t) - d(t) - y_0(t)) \quad (19)$$

where  $x_p(t) \in \mathbb{R}^{n_p}$ . The last terms of these equations, for the cost-function error weighting term, could be written as follows:

$$x_p(t+1) = A_p x_p(t) + B_p(C_r(x_r(t) - C_d x_d(t)) \quad (20)$$

$$- C_0 x_0(t) - q^{-k} E_0 u_0(t))$$

$$y_p(t) = C_p x_p(t) + E_p(C_r(x_r(t) - C_d x_d(t)) \quad (21)$$

$$- C_0 x_0(t) - q^{-k} E_0 u_0(t))$$

#### 2.1.5 Nonlinear Dynamics

The nonlinear dynamics submodel is represented by  $\mathcal{W}_{1k}(t)$ . Any unstable modes are assumed to be included in a stable or unstable linear block of state-space equation form,  $q^{-k} W_{0k}(q^{-1})$  from Equation (15). Hence, the total plant model can be described as follows:

$$\mathcal{W}(t) = q^{-k} W_{0k}(q^{-1}) \mathcal{W}_{1k}(t) \quad (22)$$

where  $k$  denotes the magnitude of the common delay elements in the output signal paths.

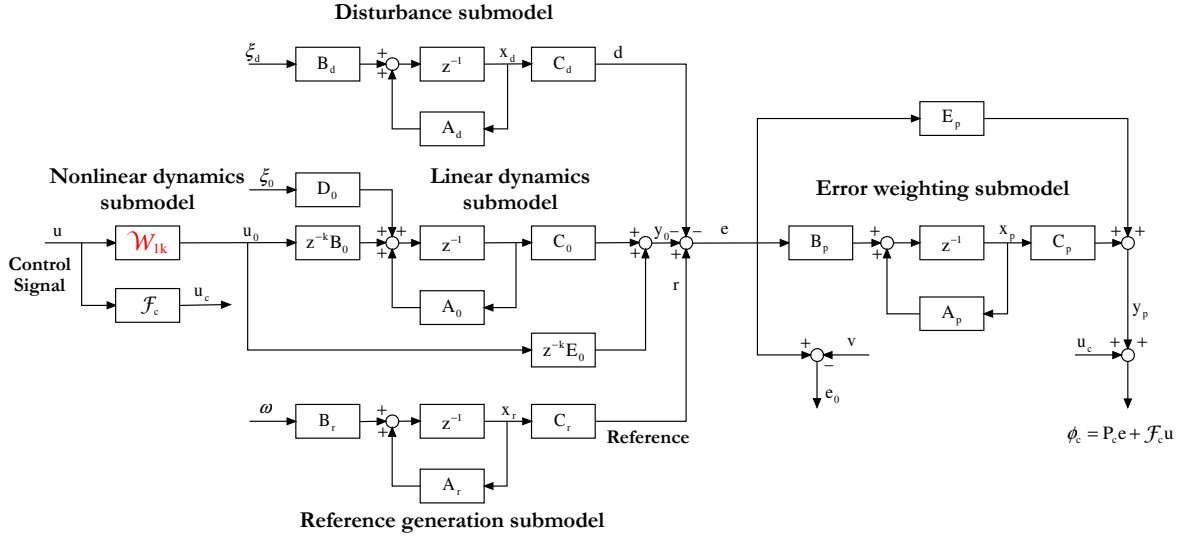


Fig 1: Nonlinear, linear, disturbance, reference, and error weighting submodels

### 2.1.6 Combined State-space Model

The equations in Section 2.1 except Equation (22) can be combined to yield the following augmented state-space equations:

$$x(t+1) = Ax(t) + Bu_0(t-k) + D\xi(t) \quad (23)$$

$$y(t) = Cx(t) + Eu_0(t-k) \quad (24)$$

$$z(t) = y(t) + v(t) \quad (25)$$

$$y_p(t) = C_\phi x(t) + E_\phi u_0(t-k) \quad (26)$$

$$e_0(t) = C_e x(t) - q^{-k} E_0 u_0(t) - v(t) \quad (27)$$

where  $x(t) \in \mathbb{R}^n$ ,  $E_\phi = -E_p E_0$ , and  $z(t)$  and  $y_p(t)$  respectively denote measurement and output to be costed.

By introducing the following resolvent operator

$$\Phi(q^{-1}) = (qI - A)^{-1} \quad (28)$$

Equation (23) can be rewritten as follows:

$$x(t) = \Phi(q^{-1})Bu_0(t-k) + \Phi(q^{-1})D\xi(t) \quad (29)$$

### 2.1.7 Stability

The nonlinear dynamics submodel  $\mathcal{W}_{1k}(t)$  is assumed to be finite gain stable for simplicity, but the linear dynamics submodel  $W_{0k}(q^{-1})$  could contain unstable modes. A system is finite-gain stable system if there exists a positive gain  $\gamma$  such that  $\|y_\tau\| \leq \gamma \|u_\tau\|$  for all  $t$ , where  $y_\tau$  and  $u_\tau$  are the output and input signals of the system, respectively, and  $\|\cdot\|$  is a norm such as 2-norm or  $\infty$ -norm.  $u_\tau$  represents a truncated signal equal to  $u(t)$  for  $t \in [0, \tau]$  and zero for  $t > \tau$  (see [7]).

## 2.2 Control

The NGMV controller requires the states  $\hat{x}(t|t)$  of the combined model. The state estimation is briefly presented in Section 2.2.1 and the NGMV control law is reported in Section 2.2.2.

### 2.2.1 State Estimation via Kalman Filter

The states of the combined model are obtained using the Kalman filter [8] as follows:

$$\hat{x}(t|t) = T_{f1}(q^{-1})e_0(t) + T_{f2}(q^{-1})u_0(t) \quad (30)$$

where

$$T_{f1}(q^{-1}) = (I - q^{-1}(A - K_f C_e A))^{-1} K_f \quad (31)$$

$$T_{f2}(q^{-1}) = (I - q^{-1}(A - K_f C_e A))^{-1} q^{-k} (K_f E_0 + q^{-1}(I - K_f C_e)B) \quad (32)$$

$K_f$  denotes the Kalman filter gain.

The state estimates computed for  $k$  steps ahead can be derived as follows:

$$\begin{aligned} \hat{x}(t+k|t) &= A^k \hat{x}(t|t) + T_0(k, q^{-1})Bu_0(t) \\ &= A^k \hat{x}(t|t) + A^{k-1}Bu_0(t-k) + A^{k-2}Bu_0(t-k+1) + \dots \\ &\quad + ABu_0(t-2) + Bu_0(t-1) \end{aligned} \quad (33)$$

where the transfer operator with finite pulse response  $T_0(k, q^{-1})$  can be written as

$$\begin{aligned} T_0(k, q^{-1}) &= (I - A^{-k}q^{-k})\Phi(q^{-1}) \\ &= q^{-1}(I + q^{-1}A + q^{-2}A^2 + \dots + q^{-k+1}A^{k-1}) \end{aligned} \quad (34)$$

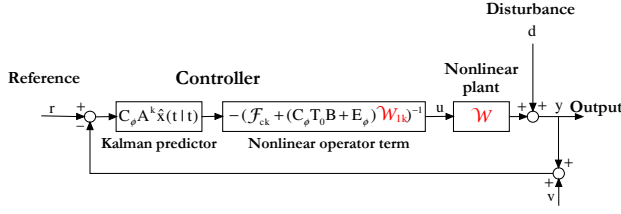


Fig 2: NGMV controller structure

### 2.2.2 Control Law

The optimal NGMV control problem minimises the variance of the signal  $\phi_c(t)$  shown in Fig. 1. This signal involves an error signal dynamic cost weighting matrix  $P_c(q^{-1})$  represented by a linear state-space equation and a nonlinear dynamics control signal costing operator term  $(\mathcal{F}_c u)(t)$ . The choice of dynamic weightings [9] is critical to the design and typically  $P_c$  and  $\mathcal{F}_c$  are low and high-pass filters, respectively.

The cost function  $J$  to be minimised is the variance of  $\phi_c(t)$  as follows:

$$J = E [\phi_c^T(t) \phi_c(t)] \quad (35)$$

If the smallest delay in each output channel of the plant is of magnitude  $k$  steps, the control at time  $t$  affects the output at least  $k$  steps later. Therefore, the control signal costing can be defined to have the following form:

$$(\mathcal{F}_c u)(t) = q^{-1} (\mathcal{F}_{ck} u)(t) \quad (36)$$

This will normally be a linear operator, but it could also be nonlinear in order to cancel the plant input nonlinearities.

Before presenting the main control theorem, the assumptions to be made are summarised as follows:

1. *The nonlinear dynamics submodel  $\mathcal{W}_{1k}$  needs to be finite gain stable.*

The input subsystem to the plant model does not need to be specified in usual equation form, so no structure is assumed known. This is one of the most powerful features of the NGMV control technique, that is, only a black box type of model is assumed. When it comes to proving overall stability on the system it is necessary to know that at least this subsystem is stable. The type of stability assumed for the total nonlinear system depends upon the requirements of the application. This black box subsystem is assumed to have the same type of stability as that to be proven for the closed-loop. This then enables the stability of the total system to be demonstrated.

2. *Reference and disturbance models are assumed to be linear subsystems.*

It is this assumption which leads to relatively simple control structures and solutions. Without this assumption one would require one of the more sophisticated and difficult nonlinear control methods such as the Lyapunov methods. A further assumption made on the system model is that the transport delays are the same in each channel which makes the solution a little easier, but it is not essential.

3. *The cost weightings are chosen such that the operator  $P_c \mathcal{W}_k - \mathcal{F}_{ck}$  is finite gain stable.*

It is equivalent to the polynomial solution of this type of optimal control problem for linear systems. In this case the requirement is that the equivalent operator is minimum phase so that its inverse is stable. This can normally be achieved even for this nonlinear system since a free choice of weightings is available.

4. *The control signal weighting operator  $\mathcal{F}_{ck}$  is assumed to be full rank and thus invertible.*

It is necessary since it becomes part of the controller structure as can be seen in the figure of the controller structure (Fig. 2). In traditional optimal control this is related to the requirement that the control costing matrix be positive definite [10]. It plays a similar role in that it is the penalty on the control action.

**Theorem 1:** As shown in Fig. 2, the NGMV optimal controller to minimise the variance of the weighted error and control signals in Equation (35) is computed as

$$u(t) = -(\mathcal{F}_{ck} + (C_\phi T_0(k, q^{-1})B + E_\phi)W_{1k})^{-1} C_\phi A^k \hat{x}(t|t) \quad (37)$$

**Proof:** The proof of this equation can be found in [4].  $\square$

For more details of the NGMV controller including tuning of the controller and simulation results, readers are referred to [4].

### 3 EXTENSION TO NGMV CONTROL

Although the Kalman filter in Section 2.2 estimates the states only, it could estimate the output signal, too. By minimising the residual which is the difference between the measurement and the estimate of the output signal, the performance of the NGMV controller could improve significantly.

The Kalman filter assumes that the disturbance model can be identified online or in advance (hence “known” disturbance) and when this identification is not possible (in other words, in the presence of “unknown” disturbance) or in the presence of any other uncertainty (which include modelling error and faults besides unknown disturbance), the performance of the controller could degrade significantly because the estimate can no longer follow the measurement resulting in a large residual signal.

On the other hand, the observer introduced here, to replace the Kalman filter, is designed such that the residual

is small even in the presence of not only unknown disturbance but also any other uncertainty. Consequently, the performance of the NGMV controller could improve significantly in the presence of uncertainty, as demonstrated here, thereby allowing the extended NGMV controller to be more robust and more straightforward to tune.

So called extended NGMV controller that includes the observer could be suboptimal in the sense that in the absence of any uncertainty, the original NGMV controller may provide superior results than the extended version since the Kalman filter is supposed to be optimal. However, in most situations, uncertainty is unavoidable and therefore the extended version should give improved results in most situations.

### 3.1 Observer Design

The mathematical description of the observer is as follows:

$$\tilde{x}(t+1) = A\tilde{x}(t) + Bu_0(t-k) + K(y(t) - \tilde{y}(t)) \quad (38)$$

$$\tilde{y}(t) = C\tilde{x}(t) + Eu_0(t-k) \quad (39)$$

where  $y(t) \in \mathbb{R}^n$  and  $\tilde{y}(t) \in \mathbb{R}^n$  denote the plant measurements and model estimates, respectively, and  $u_0(t) \in \mathbb{R}^m$  represents the control action. The term involving the observer gain  $K$  should correct the observer estimate continuously such that  $\tilde{y}(t)$  follows  $y(t)$  more closely. This implies that the effects of any uncertainty can be reduced by optimising  $K$ . Derivation of an optimal gain  $K$  is summarised here.

#### 3.1.1 Observer Gain in Frequency Domain

Any uncertainty may be described by incorporating additional terms  $d_1(t) \in \mathbb{R}^r$  and  $d_2(t) \in \mathbb{R}^n$  into the combined model in Equations (23) and (24) as follows:

$$x(t+1) = Ax(t) + Bu_0(t-k) + d_1(t) \quad (40)$$

$$y(t) = Cx(t) + Eu_0(t-k) + d_2(t) \quad (41)$$

where  $d_i(t)$  ( $i = 1, 2$ ) represents the signal of uncertainty, which in this paper is defined as any unknown input including modelling error, unknown disturbance, and faults. It could be re-defined as follows in order to improve the accuracy of the optimisation problem discussed in this section.

$$d_1(t) = \Delta Ax(t) + \Delta Bu(t) + E_1 \tilde{d}_1(t) \quad (42)$$

$$d_2(t) = \Delta Cx(t) + \Delta Eu(t) + E_2 \tilde{d}_2(t) \quad (43)$$

where  $\Delta A$ ,  $\Delta B$ ,  $\Delta C$  and  $\Delta E$  represent modelling errors, and  $\tilde{d}_1(t)$  and  $\tilde{d}_2(t)$  the disturbance distribution matrices. These matrices, however, are often not available in practice, and it is also assumed in this paper that they are unknown.

Subtracting  $\tilde{x}(t+1)$  in Equation (38) from  $x(t+1)$  in Equation (40), the equation for the residual  $\bar{r}(t)$  can be

derived as follows:

$$e(t+1) = (A - KC)e(t) + d_1(t) - Kd_2(t) \quad (44)$$

$$\bar{r}(t) = Ce(t) + d_2(t) \quad (45)$$

where

$$e(t+1) = x(t+1) - \tilde{x}(t+1) \quad (46)$$

The z-transform of Equation (45) is thus

$$\bar{r}(z) = C(zI - A + KC)^{-1}d_1(z) + (I - C(zI - A + KC)^{-1})Kd_2(z) \quad (47)$$

Subsequently, the effects of uncertainty can be reduced by minimising the following performance indices, expressed in terms of the z-transform complex number:

$$J_1(K) = \|C(zI - A + KC)^{-1}\|_\infty \quad (48)$$

$$J_2(K) = \|(I - C(zI - A + KC)^{-1}K)\|_\infty \quad (49)$$

where  $\|\cdot\|_\infty$  denotes  $H_\infty$  norm.

By minimising  $J_1(K)$  and  $J_2(K)$ , the maximums of  $C(zI - A + KC)^{-1}$  and  $I - C(zI - A + KC)^{-1}K$  in the case of Single-Input and Single-Output (SISO) and the maximums of the largest singular values in the case of MIMO, which correspond to the peak gains of the frequency response, are minimised. Hence, the effects of uncertainty can be minimised.

Therefore, the problem now is to find  $K$  such that  $J_1(K)$  and  $J_2(K)$  are minimised. However, it is likely that  $K$  causes instability; that is, an optimised  $K$  could cause the observer to become unstable. This can be prevented by parameterising  $K$  via the Eigenstructure Assignment Method (EAM) summarised here.

#### 3.1.2 Parameterisation via Eigenstructure Assignment Method

When conducting an optimisation to minimise  $J_1(K)$  and  $J_2(K)$  in Equations (48) and (49), it is important to ensure that the stability of the observer is always guaranteed under normal operating conditions, and this leads to more complex constrained optimisation problem.

To guarantee the stability condition, Chen and Patton [11] suggest the use of the EAM which parameterises  $K$ . The method has an advantage of allowing the eigenvalues in predefined regions and is summarised below.

First, it is assumed that the eigenvalues are always real for the sake of brevity. Since the observer design problem is the ‘‘dual problem’’ [12] of the controller design,  $v_i$  is the  $i_{th}$  eigenvector of  $A^T - C^T K^T$  corresponding to the  $i_{th}$  eigenvalue  $\lambda_i$  as follows:

$$(A^T - C^T K^T)v_i = \lambda_i v_i \quad (50)$$

$$v_i = -(\lambda_i I - A^T)^{-1} C^T w_i \quad (51)$$

where  $w_i = K^T v_i$ . There are now two design parameters  $w_i$  and  $\lambda_i$  instead of one design parameter  $K$ . These design parameters still do not guarantee the stability of the observer.

The eigenvalue  $\lambda_i$ , one of the design parameters, is generally not required to be placed at a specific point in the  $s$  or  $z$ -planes but rather in a predefined region to satisfy the stability condition. This in turn provides more relaxed design freedom as follows:

$$\lambda_i \in [L_i, U_i] \quad (52)$$

where  $L_i$  and  $U_i$  ( $i = 1, \dots, n$ ) respectively denote the upper and lower bounds. By defining an equation for the eigenvalue as [13]

$$\lambda_i = L_i + (U_i - L_i) \sin^2(\bar{z}_i) \quad (53)$$

Equation (53) ensures that  $\lambda_i$  remains within the upper  $U_i$  and lower bounds  $L_i$  for any  $\bar{z}_i \in \mathbb{R}$  ( $i = 1, \dots, n$ ).  $\bar{z}_i$  becomes a design parameter instead of  $\lambda_i$ , and any  $\bar{z}_i$  subsequently guarantees the stability condition.

Finally, the two design parameter vectors  $W$  and  $Z$  have been defined and the performance indices in Equations (48) and (49) can be rewritten as follows:

$$J_1(W, Z) = \|C(zI - A + KC)^{-1}\|_\infty \quad (54)$$

$$J_2(W, Z) = \|I - C(zI - A + KC)^{-1}K\|_\infty \quad (55)$$

where

$$K = [WV^{-1}]^T \quad (56)$$

Each element of  $V$  is given in Equation (51).

Having redefined the optimisation problem as finding  $Z$  and  $W$  from finding  $K$  only, the stability condition is always guaranteed.

Since two performance indices Equations (54) and (55) need to be minimised simultaneously, a multi-objective optimisation technique could be exploited. However, unlike the multi-objective problem addressed in Section 4, the minimisation of  $J_1$  does not apt to contradict with the minimisation of  $J_2$ , and vice versa. This implies that the minimisation of either  $J_1$  and  $J_2$  only could suffice.

### 3.2 Simulation Results

To assess the extended NGMV controller, both the extended and original NGMV controllers are applied to a nonlinear tank model.

The nonlinear equations for variation in the tank variable are

$$Q_i - Q_o = A \frac{dH}{dt} \quad (57)$$

$$Q_o = C_d a_d \sqrt{2gH} \quad (58)$$

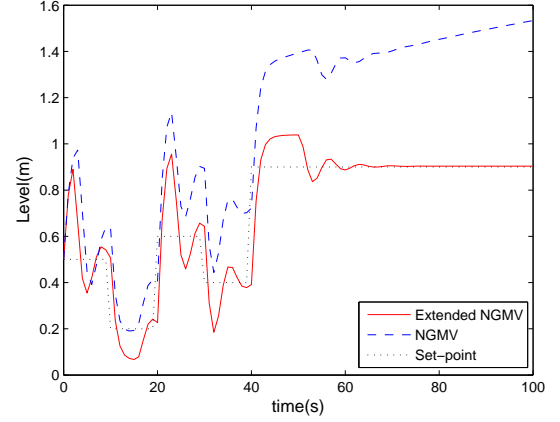


Fig 3: Set-point (black/dotted) tracking performance of NGMV (blue/dashed) and extended NGMV (red/solid) in the presence of uncertainty

where  $Q_i$  (in  $m^3/s$ ) denotes flow-in,  $Q_o$  (in  $m^3/s$ ) flow-out,  $A$  (in  $m^2$ ) area of tank,  $H$  (in  $m$ ) height of tank,  $C_d$  discharge coefficient,  $a_d$  (in  $m^2$ ) area of discharge orifice, and  $g$  (in  $ms^{-2}$ ) acceleration due to gravity. This differential equation can be rearranged to give expression for the change in output level (i.e., height of tank) as follows:

$$\frac{dH}{dt} = \frac{Q_i}{A} - \frac{C_d a_d \sqrt{2g}}{A} \sqrt{H} \quad (59)$$

Simulation results in [4] demonstrate that the original NGMV controller could be applied to a nonlinear process successfully in the absence of uncertainty. In order to test the original and extended NGMV controllers in the presence of uncertainty, random disturbance is incorporated into the nonlinear tank model at all times. More specifically, a disturbance signal, which is the unit step response of a first order system with the gain of 0.3 (almost 30 % of the steady state output) and the time constant of 10 s, is added to the change in output level at all times. The simulations are carried out in Matlab/SIMULINK®.

In the presence of the random disturbance (or uncertainty), the set-point (black/dotted) tracking performance of the original NGMV controller (blue/dashed) degrades as illustrated in Fig. 3. The extended NGMV controller, however, demonstrates improved set-point tracking performance as depicted in the same figure (red/solid). The improvement is mainly due to the observer design that causes the NGMV controller to be less sensitive to uncertainty than the original NGMV controller.

## 4 FAULT MONITORING

The observer design introduced in Section 3 can be extended further in order for the controller to be utilised not only as a controller but also as a fault monitoring system

(or fault detector) at the same time. A fault can be defined as an unpermitted deviation of at least one characteristic property or variable of the process [14]. The residual that is the difference between the measurement and the estimate of the output signal can be exploited to detect faults. It should become large when a fault emerges violating the threshold but needs to be as small as possible under normal operating conditions and even when any other uncertainty exists. Therefore, when an observer is designed for fault monitoring purposes, it is essential to distinguish faults from any other uncertainty. The use of an observer for residual generation is one of the most popular methods alongside parity relations [15–17] and parameter estimation [5, 18].

In contrast, for control purposes, the residual should be small at all times, hence there is no need to distinguish faults from any other uncertainty in Section 3. By allowing the residual to become sensitive to faults for fault monitoring purposes in this section, the performance of the controller degrades when a fault appears. However this should not cause any problem in most situations as the plant should normally be stopped as soon as a fault emerges for the plant operators to remove the fault before restarting the plant.

A fault detector is typically designed for a specific fault. Therefore, the number of fault detectors to be designed should depend on the number of faults to be detected. For simplicity, this section is only concerned with a sensor fault.

Because a sensor fault affects the measurement  $y(t)$  without affecting the states, Equation (40) can be modified to

$$x(t+1) = Ax(t) + Bu_0(t) + d_1(t) \quad (60)$$

$$y(t) = Cx(t) + Eu_0(t) + f_1(t) \quad (61)$$

where  $f_1(t)$  denotes a fault signal.

Following the steps shown in Section 3, the performance indices can be obtained as follows:

$$J_3(W, Z) = \|C(zI - A + KC)^{-1}\|_\infty \quad (62)$$

$$J_4(W, Z) = \|I - C(zI - A + KC)^{-1}K\|_\infty \quad (63)$$

where  $J_3(W, Z)$  and  $J_4(W, Z)$  are identical to  $J_1(W, Z)$  and  $J_2(W, Z)$ .

In contrast to the optimisation of the observer reported in Section 3,  $J_4(W, Z)$  needs to be maximised, and  $J_3(W, Z)$  still minimised, leading to a multi-objective optimisation problem. The minimisation of  $J_3(W, Z)$  and the maximisation of  $J_4(W, Z)$  are conflicting problems.

Several analytical methods for solving multi-objective optimisation problems have been proposed, including the optimisation via Singular Value Decomposition (SVD) and the optimisation via optimal projection which can be found in [16] and [11]. These methods can often solve optimisation problems accurately but, in some situations, can lead

to local minima due to the calculation of cost function gradients, especially when much noise is present. By contrast, we break this tradition and utilise an evolutionary algorithm [19], or more specifically, a genetic algorithm, for solving this multi-objective optimisation problem. By employing a genetic algorithm, the possibility of finding the global optimisation solution (global minimum) could increase by performing a parallel search of the solution space or the objective function as opposed to the calculation of cost function gradients [20]. Another advantage is that it is relatively straightforward to implement.

A combination of the method of inequalities, the moving-boundaries algorithm, and a genetic algorithm is exploited for solving this multi-objective optimisation problem and is explained below. This combination has been exploited in [11] for designing a fault detector for a flight simulation model but has never been incorporated into a controller in any literature.

## 4.1 Multi-objective Optimisation

### 4.1.1 Method of Inequalities

The method of inequalities transforms the problem of the minimisation or maximisation of the performance indices to the problem of the satisfaction of a set of inequalities; that is, the problem becomes searching for a parameter set that satisfies the following inequalities:

$$J_i(p) \leq \varepsilon_i \quad (64)$$

where  $p = \{W, Z\}$  and  $\varepsilon_i$  ( $i = 1, 2$ ) is the bound on the performance index  $J_i(p)$  chosen by the designer. By restricting or relaxing the bound  $\varepsilon_i$ , the designer can place a different emphasis. If  $J_1^*(p)$  and  $J_2^*(p)$  are the minimum values that can be achieved, the designer should define  $\varepsilon_i$  ( $i = 1, 2$ ) as

$$J_i^*(p) \leq \varepsilon_i \quad (65)$$

### 4.1.2 Moving-boundaries Algorithm

To help solving the design problem presented above, Zakian and Al-Naib [21] suggest an algorithm which they call the moving boundaries algorithm. The performance indices are first normalised as follows:

$$\phi_i(p) = J_i(p)/\varepsilon_i \quad (66)$$

In turn, the problem becomes satisfying

$$\phi_i(p) \leq 1 \quad (67)$$

To solve Equation (67), let  $P_i$  be the parameter that satisfies the  $i^{\text{th}}$  performance index

$$P_i = \{p : \phi_i(p) \leq 1\} \quad (68)$$

and  $P$  be the parameter that satisfies both performance indices

$$P = \left\{ p : \max_{i=1}^2 \{\phi_i(p) \leq 1\} \right\} \quad (69)$$

The search for an optimal  $P$  can be achieved by solving the following optimisation problem:

$$\min \left\{ \max_{i=1}^2 \{ \phi_i(p) \} \right\} \leq 1 \quad (70)$$

In order to solve Equation (70), let  $P^k$  be the parameter at step  $k$ , and define

$$P_i^k = \left\{ p : \phi_i(p) \leq \Delta^k \right\} \quad (71)$$

where

$$\Delta^k = \max_{i=1}^2 \left\{ \phi_i(P^k) \right\} \quad (72)$$

Now, let the problem become finding a new parameter  $p$  that reduces the largest performance index  $\Delta^k$  such that

$$\Delta^{k+1} \leq \Delta^k \quad (73)$$

The optimisation process terminates either when  $\Delta^k$  is less than 1 or when  $\Delta^k$  cannot be reduced further. If  $\Delta^k$  cannot be reduced further and persists being larger than 1, the appropriate bound should be relaxed. The most difficult part of this process can be the provision of a trial parameter  $P^{k+1}$  given  $P^k$ . As a solution, a genetic algorithm can be utilised as follows.

#### 4.1.3 Multi-objective optimisation via Genetic Algorithm

This paper assumes that the readers are familiar with genetic algorithms – detailed introduction to genetic algorithms can be found in books and papers such as [22] and [23].

The multi-objective optimisation procedures that utilise the combination of the method of inequalities, the moving boundaries algorithm, and a genetic algorithm for satisfying the performance indices are briefly summarised as follows.

##### Step 1: Determination of Bounds

The bounds  $\varepsilon_i$  ( $i = 1, 2$ ) in Equation (64) are chosen.  $\varepsilon_i$  ( $i = 1, 2$ ) are the design parameters for the optimisation algorithm. By restricting  $\varepsilon_1$ , for instance, stronger emphasis is placed on the optimisation of  $J_1(p)$ , and by relaxing  $\varepsilon_1$ , weaker emphasis is placed on the optimisation of  $J_1(p)$ .

##### Step 2: Generation of Initial Population

A random matrix, whose numbers of columns and rows respectively represent the number of variables in the fitness function (i.e., the length of  $p$ ) and the size of population, is generated.

##### Step 3: Evaluation

$J_3$  and  $J_4$  in Equations (62) and (63) are calculated using each row (i.e., individual) of the matrix from Step 1 as the variables. Using the bounds from Step 1 and Equations (64) to (70),  $\Delta^k$  in Equation (72) is calculated. This value known as a “score” is used in the following steps.

##### Step 4: Reproduction

Using “ranking”, which is one of the available options, new children are created.

##### Step 5: Elitism

A certain percentage of individuals in the current generation with the lowest fitness values are selected as “elites” and are passed on to form the population for the next generation. This step ensures that the best individuals are not lost, but it can also increase the chance of being dominated by the elite individuals prematurely.

##### Step 6: Recombination

Using “scattering”, which is one of the available options, cross-over children are created. This step is motivated by the assumption that the best solution is contained in the population as a whole rather than in each individual and thus can be found by combining individuals.

##### Step 7: Mutation

Mutation children are created by randomly changing the genes of parents’ individuals. We employ the Gaussian distribution to add a random vector to the parents. This step is motivated by the probability that the initial random population did not contain all the information necessary to solve the problem and also that the individuals that were not allowed to reproduce offspring during the previous steps may have had some information necessary to solve the problem.

##### Step 8: Termination Checking

Step 3 to 7 are repeated until either the following stopping criteria,  $\Delta^k \leq 1$ , has been met or it cannot be minimised further, in which case, either  $\varepsilon_1$  or  $\varepsilon_2$  should be relaxed.

#### 4.1.4 Tuning of the Genetic Algorithm

In order to execute the genetic algorithm, the Genetic Algorithm and Direct Search Toolbox<sup>TM</sup>2 in Matlab<sup>®</sup> is utilised. For the simulation illustrated in Fig. 4, the tuning parameters are set as follows: population size: 20, number of generations: 100, reproduction method: ranking, elite count: 2 out of 20, cross-over fraction: 14 out of 20, cross-over function: scattering, mutation function: Gaussian, mutation fraction: 4 out of 20.

## 4.2 Simulation Results

The observer in the controller is optimised using the multi-objective optimisation technique reported in Section 4.1. It is named “extended NGMV controller B” and the version developed in Section 3 “extended NGMV controller A”. Both controllers are applied to the nonlinear tank model introduced in Section 3.2 for comparison purposes. A sensor fault is added at 100s and the residual (i.e.,



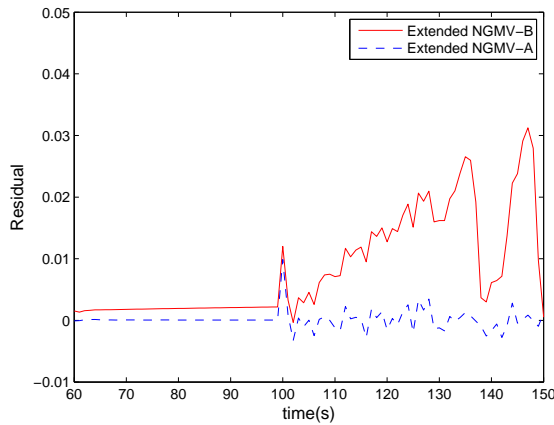


Fig 4: Red/solid residual by extended NGMV controller B and blue/dashed residual by extended NGMV controller A; a fault appears at 100s

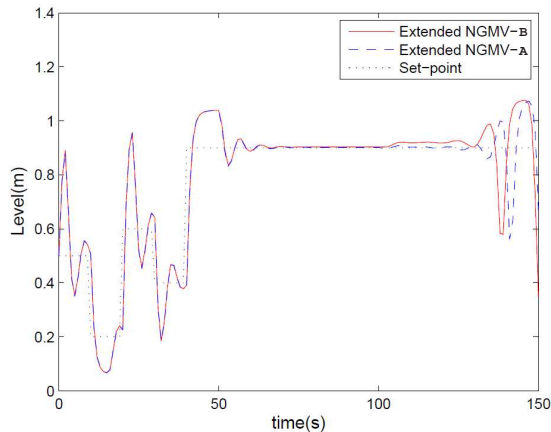


Fig 5: Set-point (black/dotted) tracking performance of extended NGMV-B (red/solid) and extended NGMV-A (blue/dashed) in the presence of fault

the difference between the measurement and estimate of  $y(t)$  from the extended NGMV controllers A and B are plotted in Fig. 4. Moreover, the unknown disturbance that is present in Section 3.2 still exists at all times.

For good fault detection, the residuals are expected to be insensitive to the disturbance but sensitive to the fault. Note that the original NGMV controller reported in Section 2 is not shown here because it fails the simulation due to the disturbance as shown in Fig. 3. The results in Fig. 4 demonstrate that the residual from the extended NGMV controller B becomes significantly larger than the residual from the NGMV controller A when the fault appears. By increasing the sensitivity of the residual to the fault, its sensitivity to the disturbance is slightly compromised as depicted in the figure before 100s, i.e. the residual by the extended NGMV controller B is larger than the

residual by the extended NGMV controller A even before the fault appears. However, as depicted in Fig. 5, it does not have a significant impact on the control performance; that is, the set-point tracking performance of the extended NGMV controller B is still almost identical to the extended NGMV controller A in the absence of the fault (before the fault appears at 100s).

This indicates that the control performance is not significantly sacrificed for improving the fault monitoring performance. When the fault appears, the control performance of the controller degrades, but this should not cause any problem in most situations because the plant should typically be stopped as soon as a fault emerges. The fault would then be removed by the plant operators before restarting the process.

In more detail, with the extended NGMV controller A in place (which is introduced in Section 3 for improved control performance without considering fault detection), as the fault appears at 100 s in Fig 5 (blue/dashed), the controller causes the faulty sensor measurement to track the reference signal until approximately 130 s. Thereafter, the fault becomes too large for the controller to cope with, and the controller eventually becomes unstable. This is also evident in Fig 4: when the fault appears at 100 s, the residual signal (blue/dashed) suddenly ramps up, and because the controller subsequently causes the faulty sensor measurement to track the reference signal, the residual becomes smaller, again.

In comparison, with the extended NGMV controller B in place (which is introduced in Section 4 considering fault detection), the controller is designed to produce a large residual signal as soon as a sensor fault emerges. Therefore, the residual becomes larger as soon as the fault appears at 100 s as depicted in Fig 4, allowing the sensor fault to be more readily detected.

## 5 CONCLUSIONS

The original NGMV controller revised in this paper is suitable for controlling nonlinear systems. However, as with the most of other controllers, the performance of the controller can degrade in the presence of uncertainty although it typically copes well with disturbances that can be identified online or in advance. The first part of the paper extends this original NGMV controller to allow it to be less sensitive to uncertainty. The application of the extended controller to a nonlinear tank model demonstrates that it can cope with the situation that the original NGMV controller can not. The improvement is mainly due to the observer whose gain is obtained in the frequency domain to minimise the effects of uncertainty. Another benefit of this approach is that the extra computation is performed offline and only takes a few seconds on an AMD™ Phenom X4 955 3.2 GHz machine. For future work, it would be interesting to find out how other types of observers, such

as specialised Kalman filter [24] and Sliding Mode Observer [25], would perform.

The second part of this paper extends the NGMV controller further for utilisation in fault monitoring. The performance of this version is also assessed by application to the nonlinear tank model, and the results demonstrate improvement in fault monitoring performance in comparison to the previous version without significantly sacrificing the control performance. The improvement is mainly due to the observer design whose gain is obtained through the application of the multi-objective optimisation technique that utilises a genetic algorithm to allow the residual to become more sensitive to faults but insensitive to any other uncertainty at the same time. The extra computation is also performed offline. It takes slightly longer than the first version but still takes a number of seconds on the same machine. The addition of the fault monitoring feature provides the NGMV controller with additional capabilities at only modest cost. Similarly to the control section of this paper, it would be interesting to discover how other types of observers would perform, such as Unknown Input Observer [11] that would allow more degree of freedom for decoupling control and fault monitoring.

## 6 ACKNOWLEDGMENTS

The authors would like to acknowledge financial support from the Engineering and Physical Sciences Research Council on the Platform Grant N<sup>o</sup> EP/F026781/1. The authors are grateful for the previous contributions of Dr Pawel Majecki (previously University of Strathclyde and now University of Grenoble) and the useful discussions with Dr Hossein Javaherian (General Motors Research & Development), Dr Gerrit M. van der Molen (Industrial Systems and Control Ltd), Dr Reza Katebi (University of Strathclyde), Prof Ron Patton (University of Hull), and Dr Jie Chen (Brunel University).

## REFERENCES

- [1] M. Grumble, "GMV control of nonlinear multivariable systems," in *Proceedings of the International Conference Control 2004 (UKACC)*, 2004.
- [2] K. J. Åström, *Introduction to Stochastic Control Theory*. Academic Press, 1970.
- [3] M. J. Grumble, *Robust industrial control*. Prentice Hall, Hemel Hempstead, 1994.
- [4] M. Grumble, "GMV control of non-linear continuous-time systems including common delays and state-space models," *International Journal of Control*, vol. 80, pp. 150–165, 2007.
- [5] L. Chiang, E. Russell, and R. Braatz, *Fault Detection and Diagnosis in Industrial Systems*, ser. Advanced Textbooks in Control and Signal Processing. Springer, 2001.
- [6] R. Isermann, *Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault*. Springer, 2006.
- [7] R. Hoseinnezhad and P. Harding, "A novel hybrid angle tracking observer for resolver to digital conversion," in *Proceedings of the 44th IEEE Conference on Decision and Control, and European Control Conference, Seville, Spain December 12-15, 2005*.
- [8] M. J. Grumble and M. A. Johnson, *Optimal control and stochastic estimation: theory and applications*. Wiley, 1988.
- [9] M. J. Grumble, *Robust industrial control: Optimal Design Approach For Polynomial Systems*. John Wiley, 2006.
- [10] M. J. Grumble and P. Majecki, "Nonlinear generalized minimum variance control under actuator saturation," in *Proceedings of the IFAC World Congress, Prague, Friday 8 July, 2005*.
- [11] J. Chen and R. J. Patton, *Robust model-based fault diagnosis for dynamic systems*. Kluwer Academic Publishers, 1999.
- [12] K. Ogata, *Modern Control Engineering*, 4th ed. Pearson Education International, 2002.
- [13] S. P. Burrows and R. J. Patton, "Design of a low sensitivity, minimum norm and structurally constrained control law using eigenstructure assignment," *Optim. Contr. Appl. & Methods*, vol. 12(3), pp. 131–140, 1991.
- [14] R. Isermann and P. Ball, "Trends in the application of model based fault detection and diagnosis of technical processes," in *Proceedings of the 13th IFAC World Congress*, vol. N. IEEE Press, 1996, pp. 1–12.
- [15] P. M. Frank, "Fault diagnosis in dynamic systems using analytical and knowledge based redundancy - a survey and some new results," *Automatica*, vol. 26(3), pp. 459–474, 1990.
- [16] X. Lou, A. S. Willsky, and G. C. Verghese, "Optimally robust redundancy relations for failure detection in uncertain systems," *Automatica*, vol. 22(3), pp. 333–344, 1986.
- [17] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. Kavuri, "A review of process fault detection and diagnosis, part i: Quantitative model-based methods," *Computers and Chemical Engineering*, vol. 27, pp. 293–311, 2003.
- [18] D. M. Himmelblau, *Fault Detection and Diagnosis in Chemical and Petrochemical Processes*. Elsevier Scientific Publishing Company, 1978.
- [19] A. Konak, D. W. Coit, and A. E. Smith, "Multi-objective optimization using genetic algorithms: A tutorial," *Reliability Engineering and System Safety*, vol. 91, pp. 992–1007, 2006.
- [20] R. J. Patton, J. Chen, and G. P. Liu, "Robust fault detection of dynamic systems via genetic algorithms," *IMechE Part I-J of Systems and Control Engineering*, vol. 211(5), pp. 357–364, 1997.

- [21] V. Zakian and U. Al-Naib, "Design of dynamical and control systems by the method of inequalities," *IEE Transactions on Control Theory and Application*, vol. 120(11), pp. 1421–1427, 1973.
- [22] J. F. Frenzel, "Genetic algorithms," *IEEE Potentials*, vol. 12(3), pp. 21–24, 1993.
- [23] L. Davis, *Handbook of Genetic Algorithms*. Van Nostrand Reinhold Company, 1991.
- [24] J. Y. Keller and M. Darouach, "Reduced-order kalman filter with unknown inputs," *Automatica*, vol. 34, no. 11, pp. 1463–1468, 1997.
- [25] S. V. Drakunov, "Sliding-mode observers based on equivalent control method," in *Proceedings of the 31st IEEE Conference on Decision and Control*, 1992, pp. 2368–2370.



**Sung-ho Hur** received the B.Eng. degree in Electronics and Electrical Engineering from the University of Glasgow in 2004 and the M.Sc. degree with Distinction in Electronic and Electrical Engineering from University of Strathclyde in 2005. From 2005 to 2006, he was a Research Assistant in the Industrial Control Centre within the Department of Electronic and Electrical Engineering at University of Strathclyde.

He received the Ph.D. from University of Strathclyde in 2010. Throughout his Ph.D. he worked very closely with DuPont Teijin Films UK Ltd, a sponsor of the project. Since then, he has been working as a Research Associate in the Industrial Control Centre within the Department of Electronic and Electrical Engineering at University of Strathclyde. His research interests are in control, in particular with applications to wind turbines and farms and cross-directional processes such as plastic film manufacturing processes; condition monitoring; and modelling.



**Mike J. Grimble** was appointed to the Professorship of Industrial Systems by the University of Strathclyde, Glasgow in 1981, and established the Industrial Control Centre. He is the Managing Editor of the International Journal of Adaptive Control and Signal Processing and of the International Journal of Robust and Non-linear Control. He was more recently appointed the editor of *Optimal Control Applications and Methods*, all published by John Wiley. He is a joint editor of the Springer Monograph Series on Advances in Industrial Control and also the joint editor of the Springer text book series on Control and Signal Processing. He established a consultancy company Industrial Systems and Control Ltd. for which he acts as the Technical Director and Deputy Chairman that has been in business for almost 3 decades.

He became a Fellow and Chartered Engineer for the Institution of Electrical Engineers in 1974 and Fellow and Chartered Mathematician of the Institute of Mathematics and its Applications in 1982. He was awarded an IEEE Fellowship in 1992 and became a founder member, and the first Chairman of the IEEE Control Systems Society UK and Ireland Control Systems Chapter, in 1975. He became a Fellow of the Institute of Measurement and Control in 1990 and a member of the Institute of Directors in 1993. He was appointed a Fellow of the Royal Society of Edinburgh in 1999. The Institution of Electrical Engineers presented him with the Heaviside Premium in 1978 for his papers on control engineering. The following year, 1979, he was awarded jointly the Coopers Hill War Memorial Prize and Medal by the Institutions of Electrical, Mechanical and Civil Engineering. The Institute of Measurement and Control awarded him the 1991 Honeywell International Medal. He was recognised at the 1993 Edinburgh International Science Festival as one of Scotland's four most cited Scientists. He was awarded the Institution of Engineering and Technology Heaviside Medal for achievements in Control, 2008.