A Multivariate Double EWMA Process Adjustment Scheme for Drifting Processes

Enrique Del Castillo

Ramkumar Rajagopal

Dept. of Industrial & Manufacturing Engineering

The Pennsylvania State University

University Park, PA 16802

March 2001

Abstract

The "predictor-corrector" feedback controller, a process adjustment scheme proposed for semiconductor manufacturing run-to-run processes that drift, is extended to the multiple input–multiple output case. The controller is based on two coupled multivariate exponentially-weighted-moving-average (EWMA) equations, thus its performance depends on the choices of EWMA weight matrices. Stability conditions are given for a pure gain process adjusted with a MIMO double EWMA (dEWMA) controller. It is shown that the stability conditions are invariant with respect to various realistic drift disturbance models. Recommendations on how to choose the EWMA weight matrices are given. An analysis is conducted to assess the impact of errors in the estimates of the process gains. The proposed MIMO dEWMA feedback controller is compared to the common practice of using multiple single-input–single-ouput dEWMA controllers running in parallel.

1 Introduction: EWMA-based feedback methods

Recent work in the area of process control, notably in semiconductor manufacturing, has concentrated in the application of the exponentially-weighted-moving-average (EWMA) statistic for process adjustment purposes [17, 8]. Del Castillo [6] presents an analysis of an adjustment scheme based on two coupled EWMA equations, the so-called "predictor-corrector", or double EWMA, feedback controller. The double EWMA (dEWMA) control scheme was first proposed by Butler and Stefani [5], who applied it to a polysilicon gate etch process used in semiconductor manufacturing, a process known to drift considerably as the reactor ages. In

1

general, the drift is not always constant or known a priori, and a feedback controller that compensates against different types of stochastic drift is desirable. Del Castillo [6] analyzed the stability and robustness of dEWMA controllers under the assumption of a deterministic trend and later extended the analysis to random walk with drift and IMA(1,1) disturbances [7]. Adivikolanu and Zafiriou [1] analyzed a modified version of the univariate dEWMA using the internal control principle. Chen and Guo [9] modified the double EWMA equations to obtain cleaner steady-state estimates of the state variables and applied it to a polishing process. All these analyses were restricted to single input, single output (SISO) processes, although it was noticed in [6] that the analysis carried over to the multiple input, *single output* (MISO) case. In this paper, a multiple input, multiple output (MIMO) extension of the dEWMA feedback controller is presented and analyzed. The use of the multivariate EWMA statistic for *monitoring* a multivariate process, as opposed to process adjustment, was discussed by Lowry et al. [11] and Prabhu et al. [13].

In the SISO case, the process model assumed by Butler and Stefani [5] in their feedback control scheme was:

$$y_t = \alpha + \beta u_{t-1} + \delta t + \epsilon_t \tag{1}$$

where y_t is the measured quality characteristic of batch or run t, u_{t-1} is the level of the controllable factor set at the end of run $(t-1)$, δ is the average drift per run, and $\{\epsilon_t\}_{t=1}^{\infty}$ is a white noise sequence. Model (1) assumes a deterministic trend (DT) disturbance, given by the terms $\delta t + \epsilon_t$. The parameter α models any offset or bias from target and β is the input-output *gain* parameter. If an off-line estimate, b, exists for the gain, β, the dEWMA controller is given by:

$$u_t = \frac{T - a_t - D_t}{b}, \tag{2}$$

where,

$$a_t = \lambda_1(y_t - bu_{t-1}) + (1 - \lambda_1)a_{t-1}; \qquad 0 < \lambda_1 \le 1, \tag{3}$$

and

$$D_t = \lambda_2(y_t - bu_{t-1} - a_{t-1}) + (1 - \lambda_2)D_{t-1}; \qquad 0 < \lambda_2 \le 1. \tag{4}$$

2

As shown in [6], the quantity $(a_t + D_t)$ provides an asymptotically unbiased one-step-ahead estimate of where the quality characteristic would have drifted in the absence of any control action. In the case of the dEWMA, there are two weights, λ_1 and λ_2, that need to be defined in order to use the controller.

For a considerably large family of drift disturbances (including deterministic trend), the stability conditions for this SISO dEWMA feedback controller are ([6, 7]:

$$|1 - 0.5\xi(\lambda_1 + \lambda_2) + 0.5z| < 1 \tag{5}$$

$$|1 - 0.5\xi(\lambda_1 + \lambda_2) - 0.5z| < 1 \tag{6}$$

where $z = \sqrt{\xi^2(\lambda_1 + \lambda_2)^2 - 4\lambda_1\lambda_2\xi}$ and $\xi = \beta/b$ is a process-model mismatch parameter.

The dEWMA scheme is not a member of the PID family, but contains integral action that adds robustness and compensates against shifts [7]. This scheme is close, although not equal, to a minimum variance controller for an ARIMA(0,2,2) process, a process that models random changes in slope and that is useful to model some processes that drift but not in a monotonic manner [7].

In a vast majority of Quality Control applications, there are multiple quality characteristics of interest. Frequently, there are multiple controllable factors that can be manipulated to modify the quality characteristics. This underlines the practical relevance of finding multivariate process adjustment techniques. Usually, industrial practice with MIMO processes is to apply several SISO feedback controllers acting in parallel. Therefore, it is of interest to study the advantages of a *multivariate* approach to process adjustment over a "parallel SISO" strategy . As in recent work in the area of process adjustment [4], the emphasis is on simple control methods that are easy to use by process engineers yet are capable of compensating against a variety of realistic disturbances and shifts in the quality characteristics. The multivariate dEWMA approach developed in this paper has these properties.

Tseng et al. [18] have recently developed a MIMO extension of a *single* EWMA controller (i.e., an extension of the controller that results from setting $\lambda_2 = 0, D_0 = 0$). This turns

out to be equivalent to MIMO integral (I) control, a particular instance of the well-known proportional-integral-derivative (PID) controllers. As in the scalar case, single EWMA feedback schemes may result in a considerable offset if the process drifts rapidly [10, 7]. Avoiding such offset was the rationale behind Butler and Stefani's dEWMA method. Evidently, the MIMO EWMA scheme is a particular case of the MIMO dEWMA scheme developed herein.

The remainder of this paper is organized as follows. Section 2 contains the main extension of a dEWMA feedback controller to the multivariate (MIMO) case. The stability conditions for the MIMO dEWMA are studied in Section 3 for various drift disturbances. It is shown that the stability conditions are invariant for a large family of drift disturbances. The effect that the weights of the MIMO dEWMA controller have on the closed-loop performance of the process are investigated by simulation in Section 4. The effect of estimating correctly, underestimating, or overestimating the process gains is studied in Section 5. Throughout sections 4 and 5, contrast is made between using a MIMO dEWMA feedback controller or using instead multiple SISO dEWMA controllers working in parallel for individual input-output couplings. Extensions of the control methods to the non-square case, i.e., the case when there are more controllable factors than responses, is briefly addressed in Section 6. The paper concludes with a summary and directions for further research.

As in previous references that discuss EWMA control, the main interest is on minimizing the sum of squared deviations of the outputs or responses, without direct consideration for the cost of the adjustments. Readers interested in modeling the cost of the adjustments are referred to the book by Box and Luceño [4]. The controllers presented in Section 6, however, consider and limit the magnitude of the control actions.

2 Extension to MIMO processes

The process model assumed is similar in form to that of the Single Input - Single Output (SISO) case (equation 1). However, it is assumed here that there are p responses or outputs and m inputs or controllable factors. Most of the results addressed in this paper are for the

case where p and m are equal, or in other words, for square systems. The case when $m > p$ is briefly discussed in section 6 and is discussed in more detail in [15]. The assumed model is:

$$\mathbf{Y}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{U}_{t-1} + \mathbf{N}_t, \tag{7}$$

where,

$$\mathbf{N}_t = \boldsymbol{\delta}\mathbf{t} + \boldsymbol{\epsilon}_t. \tag{8}$$

In the above equations, $\mathbf{N}_t$ is assumed to be a multivariate Deterministic Trend (DT) noise model and the MIMO dEWMA controller will be derived for this type of disturbance. However, as discussed in the following sections, the stability conditions for the MIMO dEWMA scheme under other types of stochastic processes that model drift are identical as those for the DT disturbance. Here, $\boldsymbol{\delta}$ is a $p \times p$ diagonal matrix with the (i, i) entry equals to the average drift rate per time unit for response i, $\mathbf{t}$ is a $p \times 1$ vector with entries equal to t, the time index, and $\{\boldsymbol{\epsilon}_t\}$ is a multivariate white noise sequence (no assumption about its distribution is made). The $(p \times 1)$ vector $\mathbf{Y}_t$ contains the quality characteristics (outputs or responses), $\boldsymbol{\alpha}$ is a $(p \times 1)$ vector containing the offset parameter of each of the responses, $\mathbf{U}_{t-1}$ is a $(m \times 1)$ vector giving the levels of the controllable factors (or inputs), and $\boldsymbol{\beta}$ is a $(p \times m)$ process gain matrix. As assumed in the univariate case, this is a "responsive" pure-gain transfer function [4] where the dynamics come from the noise term. Such model is applicable not only in semiconductor manufacturing but in many other batch-oriented processes where there is drift in the quality characteristics owing to wear-out phenomena. From equations (7-8), it can be seen that there are three sets of model parameters to be estimated, namely, those contained in $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\delta}$, respectively. As customary in the area of run-to-run control [8], it is assumed that an estimate, $\mathbf{B}$, of the process gain, $\boldsymbol{\beta}$, is obtained off-line using methods such as Design of Experiments and linear regression techniques. Similarly than in the univariate dEWMA scheme, the intercepts, $\boldsymbol{\alpha}$, and the drift, $\boldsymbol{\delta}$, will be estimated on-line and updated after each run.

At the end of each run t, a corrective action, $\mathbf{U}_t$, is chosen, which gives a prescribed set

of inputs, or *recipe*, to the process engineer for use in the next run. This is done by choosing the control action that corrects for the one-step-ahead predicted deviation from target in a similar way as done for the univariate dEWMA. That is, $\|\hat{\mathbf{Y}}_{t+1|t} - \mathbf{T}\|^2$ is minimized, where $\mathbf{T}$ is a $(p\times1)$ vector of targets for the responses:

$$Min \; Z = \|\hat{\mathbf{Y}}_{t+1|t} - \mathbf{T}\|^2 \tag{9}$$

where,

$$\hat{\mathbf{Y}}_{t+1|t} = \hat{\boldsymbol{\alpha}} + \hat{\boldsymbol{\delta}}(\mathbf{t+1}) + \mathbf{B}\mathbf{U}_t. \tag{10}$$

In the above equation the estimate of the sum $(\boldsymbol{\alpha} + \boldsymbol{\delta}t)$ is denoted $\hat{\boldsymbol{\alpha}} + \hat{\boldsymbol{\delta}}\mathbf{t}$. If $(\mathbf{A}_t + \mathbf{D}_t)$ estimates $(\boldsymbol{\alpha} + \boldsymbol{\delta}(t+1))$, in analogy to the univariate case, then (10) may be written as follows:

$$\hat{\mathbf{Y}}_{t+1|t} = \mathbf{A}_t + \mathbf{D}_t + \mathbf{B}\mathbf{U}_t. \tag{11}$$

Hence, the objective function can be written as

$$Min \; Z = \|\mathbf{A}_t + \mathbf{D}_t + \mathbf{B}\mathbf{U}_t - \mathbf{T}\|^2, \tag{12}$$

where the minimization is with respect to $\mathbf{U}_t$. Thus,

$$\frac{\partial Z}{\partial \mathbf{U}_t} = 2\mathbf{B}'(\mathbf{A}_t + \mathbf{D}_t + \mathbf{B}\mathbf{U}_t - \mathbf{T}) = \mathbf{0}_{m\times1} \tag{13}$$

from where we obtain

$$(\mathbf{B}'\mathbf{B})\mathbf{U}_t = \mathbf{B}'(\mathbf{T} - \mathbf{A}_t - \mathbf{D}_t) \tag{14}$$

or

$$\mathbf{U}_t = (\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'(\mathbf{T} - \mathbf{A}_t - \mathbf{D}_t) \tag{15}$$

which, if $m = p$ reduces to:

$$\mathbf{U_t} = \mathbf{B}^{-1}(\mathbf{T} - \mathbf{A_t} - \mathbf{D_t})$$

which is simply the unique solution to $\hat{\mathbf{Y}}_{t+1|t} = \mathbf{T}$. This evidently solves problem (9) yielding $Z = 0$. If $m > p$, (15) should be used once the $\mathbf{B}'\mathbf{B}$ matrix is regularized, as discussed in

section 6 (see also [15]). In the case when $m = p$, $\mathbf{A}_t$ and $\mathbf{D}_t$ are obtained using two multivariate EWMA equations:

$$\mathbf{A}_t = \mathbf{\Lambda}_1(\mathbf{Y}_t - \mathbf{B}\mathbf{U}_{t-1}) + (\mathbf{I} - \mathbf{\Lambda}_1)\mathbf{A}_{t-1}, \tag{16}$$

and

$$\mathbf{D}_t = \mathbf{\Lambda}_2(\mathbf{Y}_t - \mathbf{B}\mathbf{U}_{t-1} - \mathbf{A}_{t-1}) + (\mathbf{I} - \mathbf{\Lambda}_2)\mathbf{D}_{t-1} \tag{17}$$

where $\mathbf{I}$ is a $(p \times p)$ identity matrix, and $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ are $(p \times p)$ EWMA weight matrices. It is shown in the Appendix that when $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ are diagonal matrices, $(\mathbf{A}_t + \mathbf{D}_t)$ provides an asymptotically unbiased estimate of $(\boldsymbol{\alpha} + \boldsymbol{\delta}(\boldsymbol{t} + \mathbf{1}))$. Since these are EWMA equations, the diagonal elements of $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ usually lie in the range $(0, 1]$ so that they can be interpreted as weights and the resulting computations as averages, although this is not necessary for stability. Table 1 gives a numerical illustration of the MIMO Double EWMA controller for the case of two inputs and two outputs. In this example, $\boldsymbol{\alpha} = (2.0, 2.0)'$, $\boldsymbol{\delta} = 0.1I$, and $\mathbf{T} = (0, 0)'$. The errors are assumed to be multivariate standard normal variables with variance covariance matrix $\sigma^2\mathbf{I}$. The estimated gains $\mathbf{B}$ are assumed equal to the true gains $\boldsymbol{\beta}$, where

$$\mathbf{B} = \boldsymbol{\beta} = \begin{pmatrix} 1.0 & 0.2 \\ 0.3 & 1.0 \end{pmatrix}.$$

The matrices $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ are diagonal matrices, where

$$\mathbf{\Lambda}_1 = \begin{pmatrix} 0.2 & 0.0 \\ 0.0 & 0.2 \end{pmatrix}, \mathbf{\Lambda}_2 = \begin{pmatrix} 0.3 & 0.0 \\ 0.0 & 0.3 \end{pmatrix}$$

were used. At the end of each run, the EWMA equations for $\mathbf{A}_t$ and $\mathbf{D}_t$ are calculated first, followed by $\mathbf{U}_t$ computed using equation (15), which gives the prescribed input settings to the process for the next run.

INSERT TABLE 1 ABOUT HERE

3　Stability analysis

We assume in this section that $m = p$ (same number of controllable factors than responses), and extend the stability conditions to the case when $m > p$ in Section 6.

3.1　Stability conditions for a deterministic trend disturbance.

The multivariate deterministic trend model is given by equation (8). Initially, this disturbance model will be assumed to hold, although this is generalized to other disturbances in the next section. A sample realization of the DT disturbance for the case of two responses is given in Figure 1.

INSERT FIGURE 1 ABOUT HERE

By substituting equation (15) into (7), the closed-loop expression for the response is obtained. This is given by

$$\mathbf{Y}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{B}^{-1}(\mathbf{T} - \mathbf{A}_{t-1} - \mathbf{D}_{t-1}) + \boldsymbol{\delta}\mathbf{t} + \boldsymbol{\epsilon}_t. \tag{18}$$

Following the work on univariate dEWMA schemes [6, 8], we define the system-model mismatch matrix $\boldsymbol{\xi} = \boldsymbol{\beta}\mathbf{B}^{-1}$. Then, equation (18) may be written as

$$\mathbf{Y}_t = \boldsymbol{\alpha} + \boldsymbol{\xi}\mathbf{T} - \boldsymbol{\xi}\mathbf{A}_{t-1} - \boldsymbol{\xi}\mathbf{D}_{t-1} + \boldsymbol{\delta}\mathbf{t} + \boldsymbol{\epsilon}_t \tag{19}$$

Substituting equations (7) and (15) into (16), we get

$$\mathbf{A}_t = \boldsymbol{\Lambda}_1(\boldsymbol{\alpha} + (\boldsymbol{\xi} - \mathbf{I})\mathbf{T}) + \boldsymbol{\Lambda}_1\boldsymbol{\delta}\mathbf{t} + \boldsymbol{\Lambda}_1\boldsymbol{\epsilon}_t + (\mathbf{I} - \boldsymbol{\Lambda}_1\boldsymbol{\xi})\mathbf{A}_{t-1} + \boldsymbol{\Lambda}_1(\mathbf{I} - \boldsymbol{\xi})\mathbf{D}_{t-1} \tag{20}$$

Similarly, substituting equations (7) and (15) in (17), we get

$$\mathbf{D}_t = \boldsymbol{\Lambda}_2(\boldsymbol{\alpha} + (\boldsymbol{\xi} - \mathbf{I})\mathbf{T}) + \boldsymbol{\Lambda}_2\boldsymbol{\delta}\mathbf{t} + \boldsymbol{\Lambda}_2\boldsymbol{\epsilon}_t + (-\boldsymbol{\Lambda}_2\boldsymbol{\xi})\mathbf{A}_{t-1} + (\mathbf{I} - \boldsymbol{\Lambda}_2\boldsymbol{\xi})\mathbf{D}_{t-1} \tag{21}$$

In an analogous way as in the univariate dEWMA case, state-space analysis can be performed to obtain the stability conditions from the above three coupled difference equations. Define

the state vector as $\mathbf{X}_t' = (\mathbf{A}_{t-1}', \mathbf{D}_{t-1}')$, a vector of dimension $(2p \times 1)$. The state space representation of equations (19-21) is:

$$\mathbf{X}_{t+1} = \mathbf{\Omega}\mathbf{X}_t + \mathbf{W}_t \tag{22}$$

$$\mathbf{Y}_t = \mathbf{C}'\mathbf{X}_t + \mathbf{R}_t \tag{23}$$

where, $\mathbf{\Omega}$ is a $(2p \times 2p)$ transition matrix, $\mathbf{W}$ is a $(2p \times 1)$ vector, $\mathbf{C}$ is a $(2p \times p)$ matrix and $\mathbf{R}$ is a $(p \times 1)$ vector. These are defined as

$$\mathbf{\Omega} = \begin{pmatrix} (\mathbf{I} - \mathbf{\Lambda}_1\boldsymbol{\xi}) & \mathbf{\Lambda}_1(\mathbf{I} - \boldsymbol{\xi}) \\ -\mathbf{\Lambda}_2\boldsymbol{\xi} & (\mathbf{I} - \mathbf{\Lambda}_2\boldsymbol{\xi}) \end{pmatrix}, \tag{24}$$

$$\mathbf{W}_t = \begin{pmatrix} \mathbf{\Lambda}_1(\boldsymbol{\alpha} + (\boldsymbol{\xi} - \mathbf{I})\mathbf{T} + \mathbf{N_t}) \\ \mathbf{\Lambda}_2(\boldsymbol{\alpha} + (\boldsymbol{\xi} - \mathbf{I})\mathbf{T} + \mathbf{N_t}) \end{pmatrix}, \tag{25}$$

$$\mathbf{C}' = \begin{pmatrix} -\boldsymbol{\xi} & -\boldsymbol{\xi} \end{pmatrix}, \tag{26}$$

and

$$\mathbf{R} = \begin{pmatrix} \boldsymbol{\alpha} + \boldsymbol{\xi}\mathbf{T} + \mathbf{N_t} \end{pmatrix}. \tag{27}$$

As it is well-known from state-space models [2], equation (23) is a stationary process if all the $2p$ eigenvalues of $\mathbf{\Omega}$ are less than one in magnitude. Assuming that the weight matrices are such that $\mathbf{\Lambda}_1 = \lambda_1\mathbf{I}$ and $\mathbf{\Lambda}_2 = \lambda_2\mathbf{I}$, the eigenvalues of $\mathbf{\Omega}$ can be shown to be:

- $\frac{1}{2}(2 - \xi_{jj}\lambda_1 - \xi_{jj}\lambda_2 - (\xi_{jj})^{1/2}\sqrt{\xi_{jj}\lambda_1{}^2 + 2\xi_{jj}\lambda_1\lambda_2 + \xi_{jj}\lambda_2{}^2 - 4\lambda_1\lambda_2}$,
  $j = 1, 2, ..., p$

- $\frac{1}{2}(2 - \xi_{jj}\lambda_1 - \xi_{jj}\lambda_2 + (\xi_{jj})^{1/2}\sqrt{\xi_{jj}\lambda_1{}^2 + 2\xi_{jj}\lambda_1\lambda_2 + \xi_{jj}\lambda_2{}^2 - 4\lambda_1\lambda_2}$,
  $j = 1, 2, ..., p$

These are very similar to the eigenvalues of the univariate Double EWMA scheme [6]. For this disturbance we have that $E[\mathbf{A_t} + \mathbf{D_t}] \to \boldsymbol{\alpha} + \boldsymbol{\delta}(\mathbf{t} + \mathbf{1})$ (see Appendix), so the dEWMA equations will "track" the diverging disturbance. However, since $\mathbf{C}' = (-\boldsymbol{\xi}, -\boldsymbol{\xi})$, the diverging state variable cancels at the output. The stability conditions, derived from the eigenvalues are described next.

Stability Conditions. For the process described by equation (7) and controlled by a MIMO Double EWMA controller (15–17) under the above mentioned conditions, the following inequalities provide necessary and sufficient conditions for the asymptotic stability of the controlled process:

$$|1 - 0.5\xi_{jj}(\lambda_1 + \lambda_2) + 0.5z| < 1, \quad \forall j \in \{1, 2, ..., p\} \tag{28}$$

$$|1 - 0.5\xi_{jj}(\lambda_1 + \lambda_2) - 0.5z| < 1, \quad \forall j \in \{1, 2, ...p\} \tag{29}$$

where, $z = \sqrt{\xi_{jj}^2(\lambda_1 + \lambda_2)^2 - 4\lambda_1\lambda_2\xi_{jj}}$

Remarks

- As stated, the noise terms $\{\epsilon_t\}$ are assumed to form a white noise sequence. This implies that the aforementioned conditions guarantee weakly stationary only, i.e., the first two moments of $\mathbf{Y_t}$ will be finite. If the additional assumption of multivariate normal white noise errors is made, then the conditions guarantee the strict stationarity of the output.

- The stability conditions are very similar to those obtained for the univariate Double EWMA case. It is seen here that only the main diagonal elements of the $\boldsymbol{\xi}$ matrix play a role in the stability of the process. However, the value of these main diagonal elements depend on both the main-diagonal as well as off-diagonal elements of the gain matrix, $\boldsymbol{\beta}$, and its estimate, $\mathbf{B}$.

- As it is well-known, a system of difference equations exhibits oscillatory behavior if its eigenvalues are complex or if a real eigenvalue is negative. Stability and oscillatory conditions are summarized in Table 2.

- Contrary to the stability conditions in the SISO case, these conditions in the MIMO case cannot be directly understood in terms of the accuracy in the estimation of the gain matrix. The reason for this is that the relationship between the $\xi'_{jj}s$ and each

element of the matrices, $\boldsymbol{\beta}$ and $\mathbf{B}$, is complex. However, for the special case in which the estimated matrix, $\mathbf{B}$ is diagonal, ξ_{jj} is simply $\boldsymbol{\beta}_{jj}/\mathbf{B}_{jj}$. This is the case shown in Figure 4 where a MIMO process is controlled using p univariate (SISO) controllers acting in parallel.

INSERT TABLE 2 ABOUT HERE

3.2   Stability conditions for other disturbances

The random walk with drift (RWD) noise model is defined as

$$\mathbf{N}_t = \mathbf{N}_{t-1} + \boldsymbol{\delta} + \boldsymbol{\epsilon}_t. \tag{30}$$

A sample realization of this process for the case of $p = 2$ is shown in Figure 2.

INSERT FIGURE 2 ABOUT HERE

The Integrated Moving Average process of order (1,1), or IMA(1,1) noise model [3] is described by the equation:

$$\mathbf{N}_t = \mathbf{N}_{t-1} - \boldsymbol{\theta}\boldsymbol{\epsilon}_{t-1} + \boldsymbol{\epsilon}_t \tag{31}$$

where $\boldsymbol{\theta}$ is a $(p{\times}p)$ matrix, and all the eigenvalues of $\boldsymbol{\theta}$ are less than 1 for invertibility [16]. A sample realization of a multivariate IMA(1,1) process is shown in Figure 3.

INSERT FIGURE 3 ABOUT HERE

It is relatively easy to show, by analogous development as in the DT disturbance case, that the state-space representation of the controlled process under RWD and IMA(1,1) disturbances is identical as long as we redefine the $\mathbf{N_t}$ term. In particular, the transition matrix Ω will be identical, and this implies that the stability conditions given in Table 2 also apply to the cases of RWD and IMA(1,1) disturbances.

4   Effects of weights on controller performance (DT disturbance)

It is of relevance to study what is the effect of the selection of EWMA weights on the closed-loop performance of the controller. In this section and in section 5, a MIMO dEWMA controller is contrasted against the alternative of using p univariate dEWMA controllers operating in parallel. The relevance of this comparison is that multivariate (i.e., MIMO) feedback control schemes are not implemented frequently in industry. Instead, it is common practice that each input-output loop is controlled individually with a SISO controller. Figures 4 and 5 depict the two different control strategies. Evidently, the performance of a multivariate controller will be better than that of a parallel SISO scheme if the controlled process is in fact multiple input, multiple output. It is still of interest, however, to study where the advantages lie given the common industrial practice of using SISO controllers acting in parallel.

INSERT FIGURES 4 AND 5 ABOUT HERE

The analysis was conducted under the assumption of a DT disturbance and the results are shown in Figure 6 for the case of 2-inputs and 2-outputs (for a similar analysis for $p = 3$ see [14]). In this case, a MIMO dEWMA controller is used to adjust a MIMO process (see Figure 5). In contrast, Figure 7 shows the results for two SISO dEWMA feedback controllers applied in parallel to a process with two responses. The architecture for this approach is

shown in Figure 4. MATLAB programs were written for the simulations. It is assumed in this section that the process gain is estimated perfectly, i.e., $\mathbf{B} = \boldsymbol{\beta}$ for the MIMO controller. For the parallel SISO dEWMA controller, the $\mathbf{B}$ matrix is diagonal as the SISO controllers are obviously decoupled.

To simplify the analysis, the two main diagonal elements of $\boldsymbol{\beta}$ are assumed to be equal to 1, and the two off-diagonal elements are kept equal and their value is varied from 0 to 0.9 along the x-axis of the figures. In the case where all the off-diagonal elements are zero, it simply means that the system is decoupled, i.e., each response is affected by only one of the inputs. In such case, the MIMO dEWMA and the p-parallel SISO dEWMA controllers are identical. As the value of the off-diagonal element increases, the correlation between the responses increases because both responses are affected by changes in either of the inputs. Simulations were carried out with an offset of $\alpha_i = 2$, and a target of $T_i = 0$, for $i = 1, 2$.

INSERT FIGURES 6 AND 7 ABOUT HERE

The assumed DT disturbance model has $\boldsymbol{\delta} = (0.2)I$ and a multivariate white noise sequence with errors $\epsilon_t \sim N(0, 1)$ was generated. As justified in the Appendix, the weight matrices, $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$ are diagonal. Furthermore, it is assumed that all the diagonal elements in each of the weight matrices are equal, so we can write $\boldsymbol{\Lambda}_1 = \lambda_1 I$ and $\boldsymbol{\Lambda}_2 = \lambda_2 I$.

The variability of the output was used to evaluate the performance of the different alternatives. The Mean Square Error (MSE) of each response was estimated over 500 observations:

$$\widehat{\text{MSE}}_i = \frac{1}{500} \sum_{t=1}^{500} (Y_{ti} - T_i)^2 \tag{32}$$

and the average of 1000 such estimated MSE's (each obtained from independent simulations) was computed for each response. Due to the symmetry in the assumed parameters and models, the average MSE for the two responses is practically identical. Thus, only the average MSE of one of the responses is plotted on the y-axis of Figures 6 and 7 for different combinations of the weights (indicated within parenthesis). The first number in parenthesis

gives the value of the diagonal elements in $\mathbf{\Lambda}_1$, and the second number gives those of $\mathbf{\Lambda}_2$. The following observations can be made from the figures:

- For both parallel SISO and MIMO dEWMA structures, there is an optimal weight combination λ_1, λ_2. Combinations where one weight is close to zero (e.g., when it equals 0.05) and the other weight is large (e.g., 0.35) gave lowest average MSE for the responses. This combination of weights seems to perform well irrespective of the correlation between the responses as given by the off-diagonal elements of $\boldsymbol{\beta}$.

- For both parallel SISO and MIMO structures, using high values for both weights results in large values of the average MSE. For the parallel SISO case, the average MSE can diverge when there is strong correlation between the responses. It can be seen that when both weights are chosen equal to 0.65, the process explodes when the off-diagonal element increases over 0.3 (Figure 7). On the contrary, the MIMO dEWMA controller provides stable performance across a much wider range of off-diagonal elements.

- In the parallel SISO case, when very low values of the EWMA weights are chosen, the behaviour is very strange (observe the line in Figure 7 corresponding to a value of 0.05 for both weights). This is possibly because when these weights are almost zero, the transient is longer, as known to happen in the SISO case, and this inflates the MSE. However, such behavior was not observed for the MIMO case (Figures 6, 7), indicating that the transient effects are probably less pronounced than in the parallel-SISO configuration.

- The effect of the EWMA weights matrices, $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$, appears approximately symmetrical, i.e., interchanging the values of $\mathbf{\Lambda}_1 = \lambda_1 I$ and $\mathbf{\Lambda}_2 = \lambda_2 I$ does not affect the Average MSE curve.

Choosing the EWMA weights is hence crucial for the performance of the process and it is recommended that values between 0.05 and 0.35 are used for the EWMA weights in a MIMO

Double EWMA controller. One of the weights should be close to zero and smaller than the other weight. These recommendations are similar to those given by Del Castillo [6] who proposed a "trade-off" solution to a weight determination optimization problem. In that problem, SISO DEWMA controller weights were chosen to minimize the sum of expected squared deviations from target (a measure of the severeness of the transient) and the asymptotic (long-run) variance. It was observed that the trade-off solution always results in one small weight to ensure stability, and one larger weight to minimize the aforementioned sum. Unfortunately, a similar analysis in the MIMO case is complicated algebraically. However, the simulation results in this section appear to confirm the design recommendation of the univariate dEWMA controller.

5 Effects of Errors in the Gain Estimates (DT disturbance)

In this section, the two control structures in Figures 4 and 5 are simulated under different assumptions regarding the process gain estimate $\mathbf{B}$. The performance is compared using the average Mean Square Error of the response. The following assumptions were made in the simulations of this section:

- All the diagonal elements of the two diagonal EWMA weight matrices, $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ are assumed to be equal to one another, i.e., $\mathbf{\Lambda}_1 = \lambda_1 \mathbf{I} = \mathbf{\Lambda}_2 = \lambda_2 \mathbf{I}$. Although equal weights are *not* recommended based on the results of the previous section, for the purposes of analyzing the effect of errors in the gains it is easier to vary the weights together. The conclusions of this section will not depend on this choice of weights.

- The main diagonal elements in the gain matrix are assumed to be equal to one. The off-diagonals are kept symmetrically equal and simulations are carried out by varying their value from 0.0 to 0.9.

- In the case of the SISO controllers in parallel, it is assumed that the estimated gain

matrix is a diagonal matrix. It is thus expected that in the case where the off-diagonal elements are all zero, the performance of the parallel SISO controllers and the MIMO controller is identical.

- The noise term used in these simulations is a multivariate DT disturbance, with $\boldsymbol{\delta}$ = $(0.1)\mathbf{I}$, and the errors in the multivariate white noise sequence are $N(0, 1)$ random variables.

- It is assumed that the offset is 2.0, and the target is 0 for all responses.

The simulations were carried out in MATLAB. MSE estimates based on 1000 replications of 500 runs each were conducted. In each simulation, the first 150 runs were eliminated to avoid the transient phase. The Average MSE is then calculated over 1000 such simulations. The Average MSE of one of the responses is then plotted on the y-axis and the EWMA weights are plotted on the x-axis (the MSE's for all responses is practically identical given the symmetry of model and disturbance parameters). Each curve on Figures 8–11 represents the average MSE plot for a particular value of off-diagonal element of the gain matrix which is indicated next to the curve. In what follows, the case of $\boldsymbol{\beta} = \mathbf{B}$ (known gains) is treated first. Conclusions for all cases considered are found at the end of Section 5.

5.1   Known gains

In this case, the system is assumed to have three inputs and three outputs. Similar results are obtained for other values of p ([14]). The gain estimate, $\mathbf{B}$, is assumed to be equal to the actual gain, $\boldsymbol{\beta}$, while simulating the MIMO controller. In case of the parallel SISO controller the main-diagonal elements of $\mathbf{B}$ are the same as that of $\boldsymbol{\beta}$ and the off-diagonal elements are zero. The results are given in Figures 8 and 9 for the parallel SISO and MIMO controllers respectively.

INSERT FIGURES 8 AND 9 ABOUT HERE

When the gains are known, it was observed in the simulations that the responses were on-target.

5.2 Gains underestimated

In this section the gain matrix is assumed to be under-estimated. The ratio of under-estimation is assumed to be two. That is, each element of $\mathbf{B}$ was set to be half that of $\boldsymbol{\beta}$ for the MIMO controller. For the parallel SISO controllers, the main-diagonal elements of $\mathbf{B}$ are half that of $\boldsymbol{\beta}$, and the off-diagonal elements are zero. The results are shown in Figures 10 for the parallel SISO case and Figure 9 for the MIMO case.

INSERT FIGURE 10 ABOUT HERE

5.3 Gains overestimated

In this case, the gain matrix is assumed to be over-estimated. The ratio of over-estimation is assumed to be two. That is, all the elements in $\mathbf{B}$ are twice those of $\boldsymbol{\beta}$ for the MIMO controller. For the parallel SISO controllers the main-diagonal elements of $\mathbf{B}$ are twice that of $\boldsymbol{\beta}$ and the off-diagonal elements are zero. These results are shown in Figures 11 for the parallel SISO case and in Figure 9 for the MIMO case.

INSERT FIGURE 11 ABOUT HERE

From Figures 8–11, the following observations can be made. For the case of known gains $(\mathbf{B} = \boldsymbol{\beta})$:

- The parallel SISO case produces an envelope of MSE functions that gets steeper as

17

the off-diagonal elements increase. From a stability point of view, it seems that small weights result in pronounced transients, inflating the MSE's.

- The MIMO DEWMA controller provides, for given weights, practically the same MSE performance regardless of the input-output coupling. For any value of the off-diagonal element, the performance is almost identical to that of the parallel SISO case when the system is decoupled (compare Figures 8 and 9). The MIMO controller is thus more consistent, and has more choices of weights for which it provides stable performance. From the figures, it can be seen that small weights guarantee stability. However, as in the parallel SISO case, setting both weights close to zero leads to pronounced transients.

For the case of over/under estimation of gains:

- The evidence we show for (3×3) systems indicates that it is better, from an MSE point of view, to over-estimate the gains, $\mathbf{B}$, than to under-estimate them (similar results for the case $p = 2$ can be found in [14]). Over-estimation leads to a flatter MSE Vs. (λ_1, λ_2) function than when $\mathbf{B} = \boldsymbol{\beta}$ in both parallel SISO and MIMO controllers, although the superiority of the MIMO controller is evident. A flatter MSE function implies the stability region for the controller has increased. It should be pointed out that if the over-estimation is very severe, these nice stability properties are lost.

- Under-estimation of the gain matrix results in instability. This is true for the parallel SISO controllers as well as the MIMO controller. The system diverges for all the values of the off-diagonal element, even if relatively small values of the EWMA weight are used.

- These results are quite similar to that of the scalar double EWMA case, where overestimation of the gain was found to be better than under-estimation [7].

6   Extension to non-square systems

Suppose now that the number of controllable factors (m) exceeds the number of responses (p). Then equation (15) is not applicable since the matrix $\mathbf{B'B}$ is not invertible. Two common solutions to this problem are to regularize the matrix before attempting inversion, and using a more general inverse. These give raise to two types of feedback controllers which we will refer to as the *Ridge Solution* controller and the *Right Inverse* controller. These controllers are explained in greater detail in [15].

The idea behind the Ridge Solution controller is to minimize $(E[\mathbf{Y}_t] - \mathbf{T})^2$ with the added constraint that $(\mathbf{U'U}) < \rho^2$, where ρ is the radius that bounds the values the controllable factors can take. This is analogous to the optimization methods used in Ridge Analysis of Response Surfaces [12]). If μ is the Lagrange multiplier of the constraint, then the control equation obtained is of the form:

$$\mathbf{U}_t = (\mathbf{B'B} + \mu\mathbf{I})^{-1}\mathbf{B'}(\mathbf{T} - \mathbf{A}_t - \mathbf{D}_t). \tag{33}$$

The above solution is similar to the control equation for square systems (eq. 15). The only change here is the $\mu\mathbf{I}$ term which when added to $(\mathbf{B'B})$ makes it invertible as long as $\mu \neq 0$. As this is a minimization problem, μ is chosen so that the matrix $(\mathbf{B'B} + \mu\mathbf{I})$ is positive definite. This will happen as long as $\mu > 0$. It is recommended to choose very small values of μ just enough to obtain an invertible matrix for computation purposes. All the results for stability and the conditions for oscillations in Table 2 may be applied to this controller by simply redifining $\boldsymbol{\xi} = \boldsymbol{\beta}(\mathbf{B'B} + \mu\mathbf{I})^{-1}\mathbf{B'}$.

The Right pseudoinverse controller, or Right Inverse controller for short, is based on the controller proposed by Sachs *et al.* [17] for systems with multiple outputs and a single input. It is obtained by minimizing the magnitude of the adjustments to be made, i.e., $(|\mathbf{U}_t - \mathbf{U}_{t-1}|)$, subject to the constraint that $(\mathbf{Y}_t - \boldsymbol{\alpha} - \boldsymbol{\delta}\mathbf{t} = \mathbf{T})$. More recently, this approach was used by Tseng *et al.* [18] for extending the single EWMA to the multivariate case. The form of this

dEWMA controller is:

$$\mathbf{U}_t = (\mathbf{I} - \mathbf{B}'(\mathbf{B}\mathbf{B}')^{-1}\mathbf{B})\mathbf{U}_{t-1} + \mathbf{B}'(\mathbf{B}\mathbf{B}')^{-1}(\mathbf{T} - \mathbf{A}_t). \tag{34}$$

As mentioned in [15], the adjustments prescribed by both of these controllers was observed to be almost identical, as long as the μ parameter of the Ridge Controller is numerically small.

7   Conclusions and further research

A MIMO extension to the univariate dEWMA feedback adjustment method was provided. The stability condition of the closed-loop system was shown to be invariant with respect to three practical drift disturbances, namely, a multivariate deterministic trend (DT), a multivariate random walk with drift (RWD), and a multivariate IMA(1,1) process. A simulation study was conducted to investigate the effects of the EWMA weight matrices and of the gain estimate on the mean square error of the responses. The benefits of using a MIMO dEWMA controller over the common practice of using several SISO dEWMA controllers acting in parallel loops were investigated.

Further work can be devoted to studying the performance of MIMO dEWMA feedback controllers when there exist process dynamics (as opposed to the pure gain process assumed in this paper). This might prove fruitful for the chemical industries where such dynamics are common.

References

[1] Adivikolanu, S., and Zafiriou, E., "Extensions and Performance/Robustness Tradeoffs of the EWMA Run-to-Run Controller by Suing the Internal Model Control Structure,"

IEEE Transactions on Electronics Packaging Manufacturing, 23, 1, (2000).

[2] Åström, K.J. and B. Wittenmark, *Computer Controlled Systems, Theory and Design.* 3rd ed., Prentice Hall, New Jersey, (1997).

[3] Box, G.E.P. and Jenkins, G.M., *Time Series Analysis, Forecasting and Control* (1976), Holden Day, Oakland, CA.

[4] Box, G.E.P., Luceño, A, *Statistical Control by Monitoring and Feedback Adjustment.* John Wiley & Sons, New York, NY, 1997.

[5] Butler, S.W. and Stefani, J.A., "Supervisory Run-to-Run Control of a Polysilicon Gate Etch using *in situ* Ellipsometry", *IEEE Transactions on Semiconductor Manufacturing* 7, 193-201, (1994).

[6] Del Castillo, E., "Long Run and Transient Analysis of a Double EWMA Feedback Controller", *IIE Transactions* , 31, 1-13, (1999).

[7] Del Castillo, E., "Some Properties of EWMA Feedback Quality Adjustment SChemes for Drifting Disturbances", *Journal of Quality Technology*, 33(2), 153-166, (2001).

[8] Del Castillo, E. and Hurwitz, A.M., "Run-to-Run Process Control: Literature Review and Extensions", *Journal of Quality Technology*, 29, 184-196, (1997).

[9] Chen, A., and Guo, R.S., "Age-based Double EWMA Controller and its Application to a CMP process," in *Run-to-run control in Semiconductor Manufacturing*, J. Moyne, E. Del Castillo, and A. Hurwitz, (eds.), CRC Press, Boca Raton, FL, (2001).

[10] Ingolfsson A. and E. Sachs "Stability and Sensitivity of an EWMA Controller," *Journal of Quality Technology*, 25, 4, pp. 271-287, (1993).

[11] Lowry, C.A., Woodall, W.H., Champ, C.W. and Rigdon, S.E., "A Multivariate Exponentially Weighted Moving Average Control Chart", *Technometrics* 34, 46-53, (1992).

[12] R.H. Myers, and D.C. Montgomery, *Response Surface Methodology*, John Wiley and Sons, New York, NY, (1995).

[13] Prabhu, S.S. and Runger, G.C., "Designing a Multivariate EWMA Control Chart", *Journal of Quality Technology* ,29, 8-15, (1997).

[14] Rajagopal, R., *Analysis and Multivariate Extensions of double EWMA Feedback Adjustment Scheme for Quality Control*, Master Thesis, Department of Industrial & Manufacturing Engineering, Penn State University (2000).

[15] Rajagopal, R., and Del Castillo, E., "An Analysis and MIMO Extension of a Double EWMA Run-to-Run Controller", working paper, IME department, Penn State University, (submitted for publication), (2001).

[16] Reinsel, G.C., *Elements of Multivariate Time Series Analysis* (1997), 2nd edn., Springer, New York, NY

[17] Sachs, E., Hu, A. and Ingolfsson, A., "Run by Run Process Control: Combining SPC and Feedback Control", *IEEE Transactions on Semiconductor Manufacturing*, 8, 26-43, (1995).

[18] Tseng, S.T., Chou, R.J., and Lee, S.P., "A Study of a Multivariate EWMA Controller," working paper, Institute of Statistics, National Tsing Hua University. (submitted for publication, 2000).

Appendix: Justification of diagonal weight matrices

Throughout the paper, diagonal weight matrices $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ were used. This section gives a formal justification for such choices of the weight matrices under the assumption that the noise is a deterministic trend. Writing the process model equation (Eqn. 7) for a single

response, i, where $i \in \{1, 2, ..., p\}$, the following equation is obtained:

$$Y_{ti} = \alpha_i + \sum_{j=1}^{m} \beta_{ij} U_{(t-1)j} + \delta_{ii}t + \epsilon_{ti}. \tag{35}$$

Thus,

$$\alpha_i + \delta_{ii}t = Y_{ti} - \sum_{j=1}^{m} \beta_{ij} U_{(t-1)j} + \epsilon_{ti} \tag{36}$$

Thus, an estimate of $(\alpha_i + \delta_{ii}t)$ is given by

$$\hat{\alpha}_i + \hat{\delta}_{ii}t = Y_{ti} - \sum_{j=1}^{m} B_{ij} U_{(t-1)j} \tag{37}$$

Hence,

$$\hat{\alpha}_i + \hat{\delta}_{ii}(t+1) = Y_{(t+1)i} - \sum_{j=1}^{m} B_{ij} U_{tj}. \tag{38}$$

Similarly, the EWMA equations for response i are given below.

$$A_{ti} = \Lambda_{1,ii}(Y_{ti} - \sum_{j=1}^{m} B_{ij} U_{(t-1)j}) + \sum_{k=1,k \neq i}^{p} \Lambda_{1,ik}(Y_{tk} - \sum_{j=1}^{m} B_{kj} U_{(t-1)j}) + (1 - \Lambda_{1,ii})A_{(t-1)i}$$
$$+ \sum_{k=1,k \neq i}^{p} (-\Lambda_{1,ik})A_{(t-1)k} \tag{39}$$

$$D_{ti} = \Lambda_{2,ii}(Y_{ti} - \sum_{j=1}^{m} B_{ij} U_{(t-1)j} - A_{(t-1)j}) + \sum_{k=1,k \neq i}^{p} \Lambda_{2,ik}(Y_{tk} - \sum_{j=1}^{m} B_{kj} U_{(t-1)j} - A_{(t-1)i})$$
$$+ (1 - \Lambda_{2,ii})D_{(t-1)i} + \sum_{k=1,k \neq i}^{p} (-\Lambda_{2,ik})D_{(t-1)k} \tag{40}$$

If $\Lambda_{1,ik} = 0, i \neq k$; and assuming $A_{ti} = 0$ at $t = 0$, Eqn. (39) becomes

$$A_{ti} = \Lambda_{1,ii} \sum_{l=0}^{t} (1 - \Lambda_{1,ii})^l (Y_{(t-l)i} - \sum_{j=1}^{m} B_{ij} U_{(t-l-1),j}) \tag{41}$$

So, from Eqn. 37 and Eqn. 41,

$$E[A_{ti}] = \Lambda_{1,ii} \sum_{l=0}^{t} (1 - \Lambda_{1,ii})^l (\alpha_i + \delta_{ii}(t-l)) \tag{42}$$

If $|1 - \Lambda_{1,ii}| < 1$ (which implies $0 < \Lambda_{1,ii} < 2$), we get

$$E[A_{ti}] \rightarrow \alpha_i + \delta_{ii}t + \delta_{ii}(1 - \frac{1}{\Lambda_{1,ii}}) \tag{43}$$

Similarly, if $\Lambda_{2,ik} = 0, i \neq k$; and assuming $D_{ti} = 0$ at $t = 0$, Eqn. (40) becomes

$$D_{ti} = \Lambda_{2,ii} \sum_{l=0}^{t} (1 - \Lambda_{2,ii})^l (Y_{(t-l)i} - \sum_{j=1}^{m} B_{ij} U_{(t-l-1),j} - A_{(t-l-1),i}) \tag{44}$$

From Eqn. 37, Eqn. 43 and Eqn. 44, we obtain

$$E[D_{ti}] = \Lambda_{2,ii} \sum_{l=0}^{t} (1 - \Lambda_{2,ii})^l (\alpha_i + \delta_{ii}(t - l) - (\alpha_i + \delta_{ii}(t - 1) + \delta_{ii}(1 - \frac{1}{\Lambda_{1,ii}}))) \tag{45}$$

So, if $0 < \Lambda_{2,ii} < 2$, we get

$$E[D_{ti}] \rightarrow \frac{\delta_{ii}}{\Lambda_{1,ii}} \tag{46}$$

and therefore,

$$E[A_{ti} + D_{ti}] \rightarrow \alpha_i + \delta_{ii}(t + 1) \tag{47}$$

In other words, if $\mathbf{B} = \boldsymbol{\beta}$, then $(\mathbf{A}_t + \mathbf{D}_t)$ becomes an asymptotically unbiased estimate of $(\boldsymbol{\alpha} + \boldsymbol{\delta}(\mathbf{t} + \mathbf{1}))$, if the weight matrices are diagonal. This condition is essential so that the recipe, $\mathbf{U}_t$, given at the end of each run corrects for the one-step-ahead predicted deviation from target.

run, t	$\mathbf{Y}_t$	$\mathbf{A}_t$	$\mathbf{D}_t$	$\mathbf{U}_t$
0		$\begin{pmatrix} 0.0 \\ 0.0 \end{pmatrix}$	$\begin{pmatrix} 0.0 \\ 0.0 \end{pmatrix}$	$\begin{pmatrix} 0.0 \\ 0.0 \end{pmatrix}$
1	$\begin{pmatrix} 1.813 \\ 3.216 \end{pmatrix}$	$\begin{pmatrix} 0.363 \\ 0.643 \end{pmatrix}$	$\begin{pmatrix} 0.544 \\ 0.965 \end{pmatrix}$	$\begin{pmatrix} -0.622 \\ -1.421 \end{pmatrix}$
2	$\begin{pmatrix} 1.047 \\ 1.453 \end{pmatrix}$	$\begin{pmatrix} 0.681 \\ 1.127 \end{pmatrix}$	$\begin{pmatrix} 0.858 \\ 1.401 \end{pmatrix}$	$\begin{pmatrix} -1.099 \\ -2.198 \end{pmatrix}$
3	$\begin{pmatrix} 0.714 \\ -1.381 \end{pmatrix}$	$\begin{pmatrix} 0.995 \\ 1.131 \end{pmatrix}$	$\begin{pmatrix} 1.072 \\ 0.987 \end{pmatrix}$	$\begin{pmatrix} -1.749 \\ -1.593 \end{pmatrix}$
4	$\begin{pmatrix} -0.816 \\ 1.753 \end{pmatrix}$	$\begin{pmatrix} 1.047 \\ 1.679 \end{pmatrix}$	$\begin{pmatrix} 0.828 \\ 1.513 \end{pmatrix}$	$\begin{pmatrix} -1.315 \\ -2.797 \end{pmatrix}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table 1: Numerical Example to Illustrate a MIMO Double EWMA Controller, $\mathbf{\Lambda}_1 = (0.2)I, \mathbf{\Lambda}_2 = (0.3)I$.

Range of $\xi_{jj} \forall j = 1, 2, ...p$	System Behavior
$(-\infty, 0]$	Unstable and Oscillates
$(0, \frac{4\lambda_1\lambda_2}{(\lambda_1+\lambda_2)^2})$	Stable but Oscillates
$[\frac{4\lambda_1\lambda_2}{(\lambda_1+\lambda_2)^2}, \frac{1}{\lambda_1+\lambda_2-\lambda_1\lambda_2}]$	Stable, No Oscillations
$(\frac{1}{\lambda_1+\lambda_2-\lambda_1\lambda_2}, \frac{2}{\lambda_1+\lambda_2}]$	Stable but Oscillates
$(\frac{2}{\lambda_1+\lambda_2}, +\infty)$	Unstable and Oscillates

Table 2: Stability and Oscillation conditions for MIMO DEWMA controller
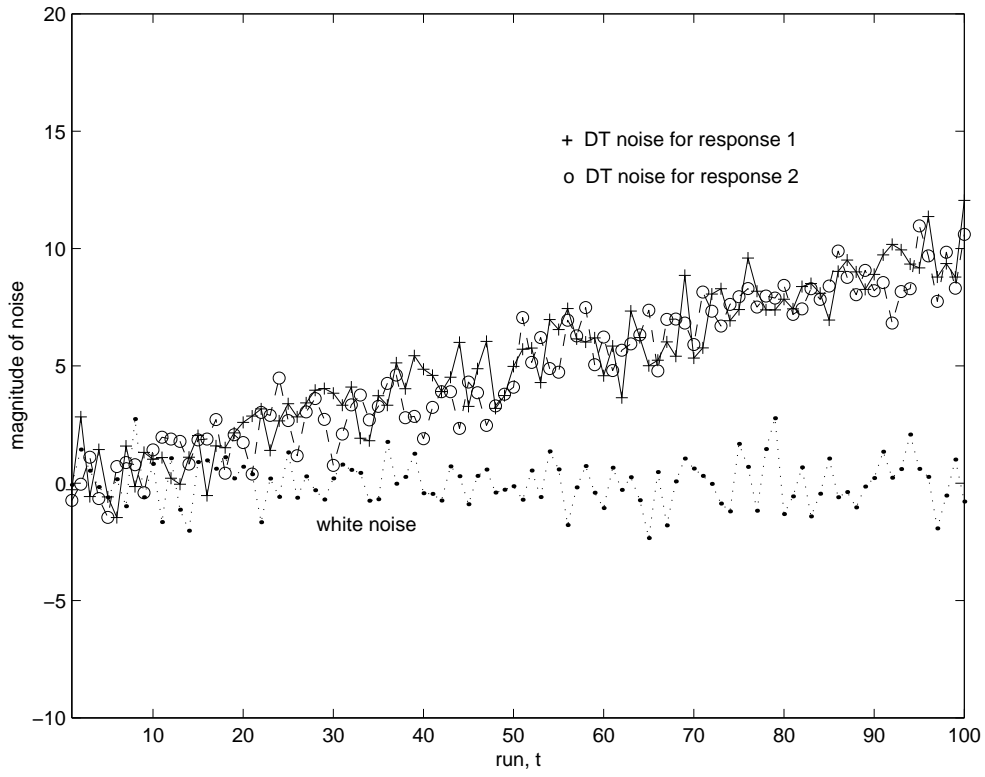
Figure 1: Time series plot of a multivariate ($p = 2$) Deterministic Trend disturbance, $\delta_{ii} = 0.1, \sigma_i = 1; i = 1, 2$.
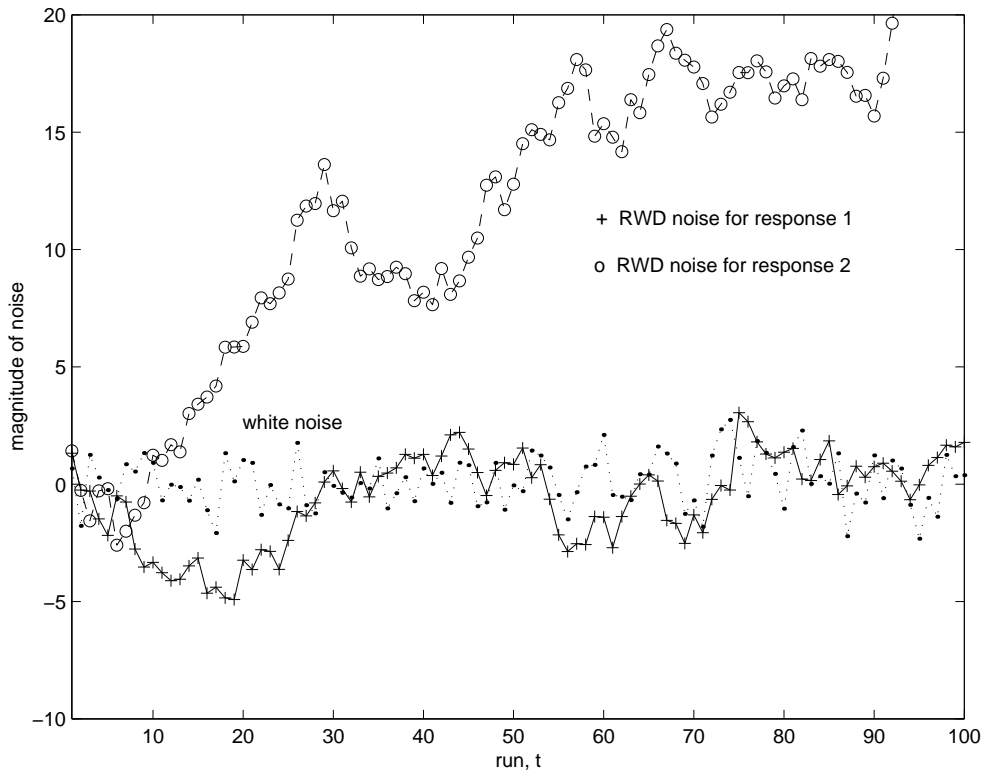
Figure 2: Time Series Plot of a multivariate $(p = 2)$ Random Walk with Drift disturbance, $\delta_{ii} = 0.1, \sigma_i = 1; i = 1, 2$.
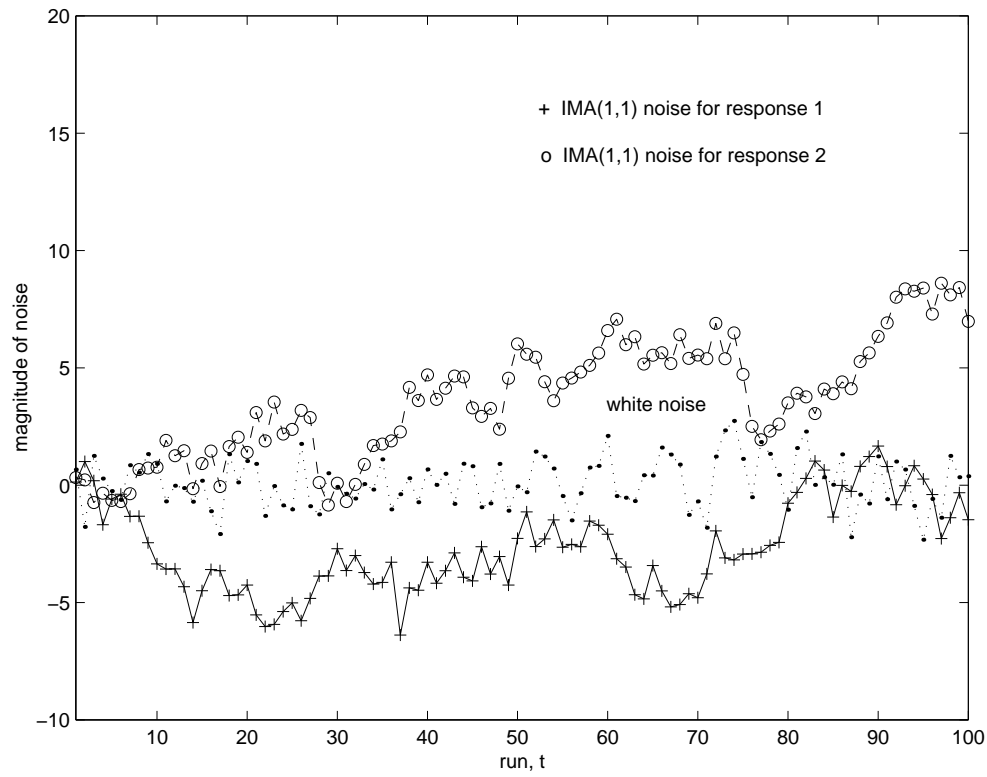
Figure 3: Time Series Plot of a multivariate ($p = 2$) IMA(1,1) disturbance, $\theta_{ii} = 0.3, \sigma_i = 1; i = 1, 2$.
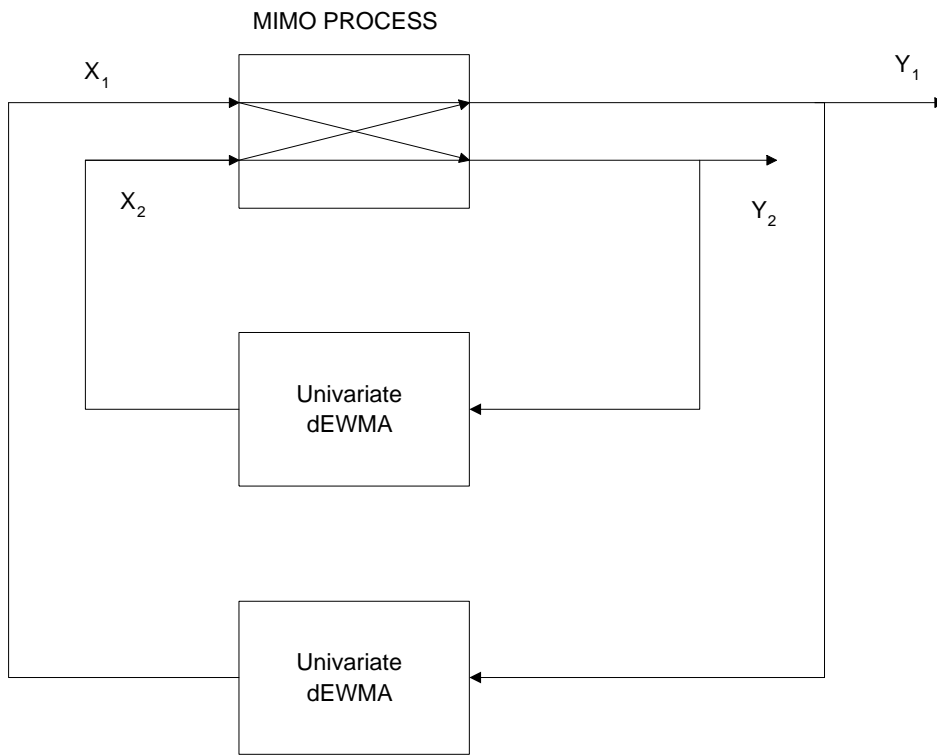
Figure 4: A Parallel SISO dEWMA Controller applied to a 2 Input, 2 Output system. Here, the matrix of estimated gains **B** is necessarily diagonal.
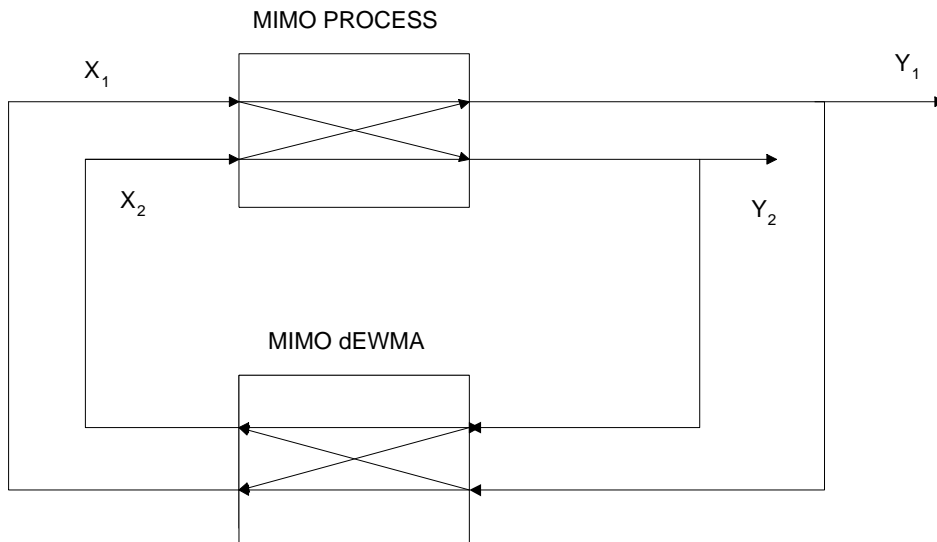
Figure 5: A MIMO dEWMA Controller applied to a 2 Input, 2 Output system. Here, the matrix of estimated gains $\mathbf{B}$ is *not* necessarily diagonal.
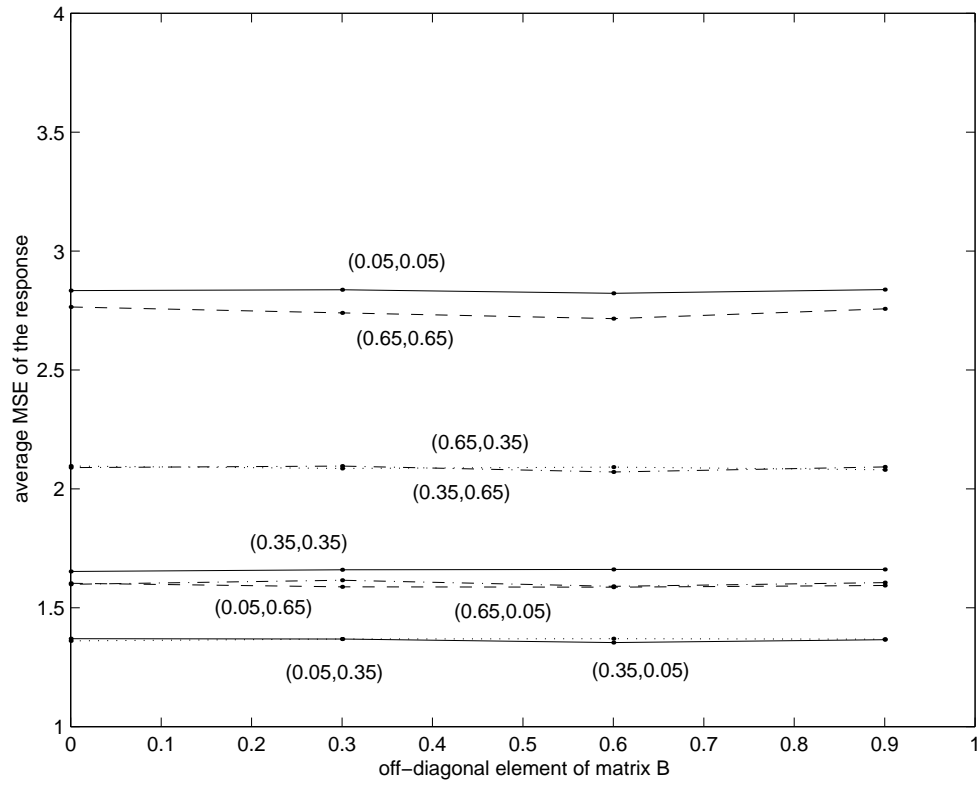
Figure 6: Effect of changing the EWMA weights on the estimated MSE of the responses ($\boldsymbol{\beta} = \mathbf{B}, \boldsymbol{\Lambda}_1 = \lambda_1\mathbf{I}, \boldsymbol{\Lambda_2} = \lambda_\mathbf{2}\mathbf{I}$, 2 Input - 2 Output MIMO case). Labels give (λ_1, λ_2).
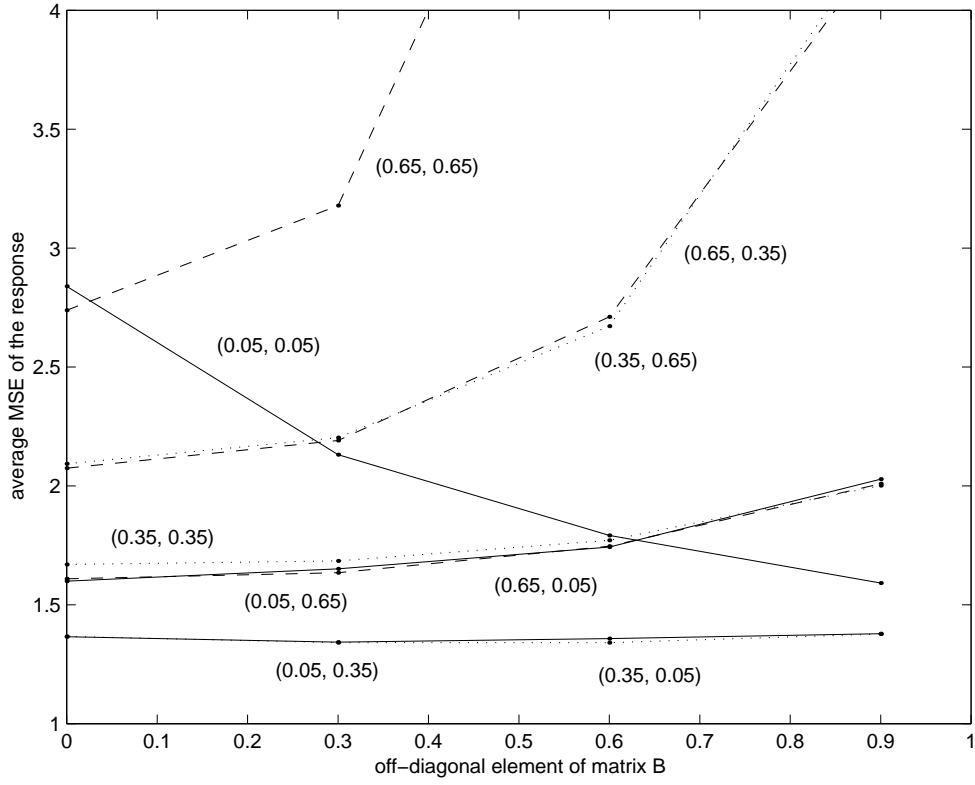
Figure 7: Effect of changing the EWMA weights on the estimated MSE of the responses ($\boldsymbol{\beta} = \mathbf{B}, \boldsymbol{\Lambda}_1 = \lambda_1\mathbf{I}, \boldsymbol{\Lambda}_2 = \lambda_2\mathbf{I}$, 2 Input - 2 Output parallel SISO case). Labels give (λ_1, λ_2).
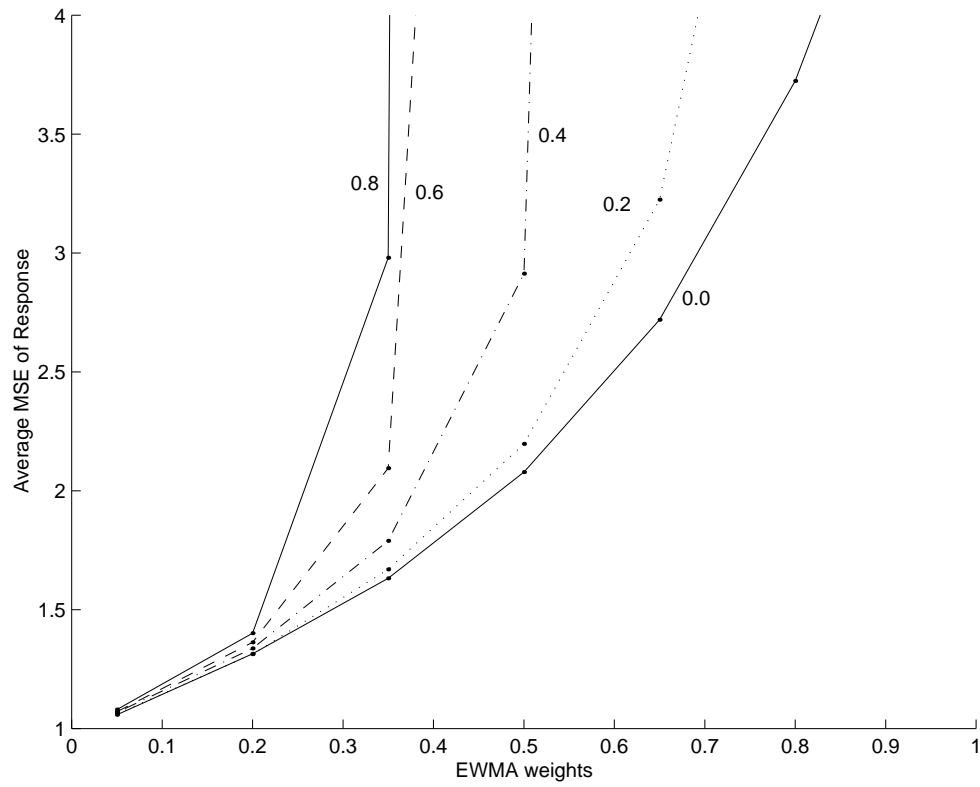
Figure 8: Estimated MSE of the responses for a 3 Input, 3 Output Parallel SISO dEWMA controllers, known gains. Labels give off-diagonal elements of $\boldsymbol{\beta}$.
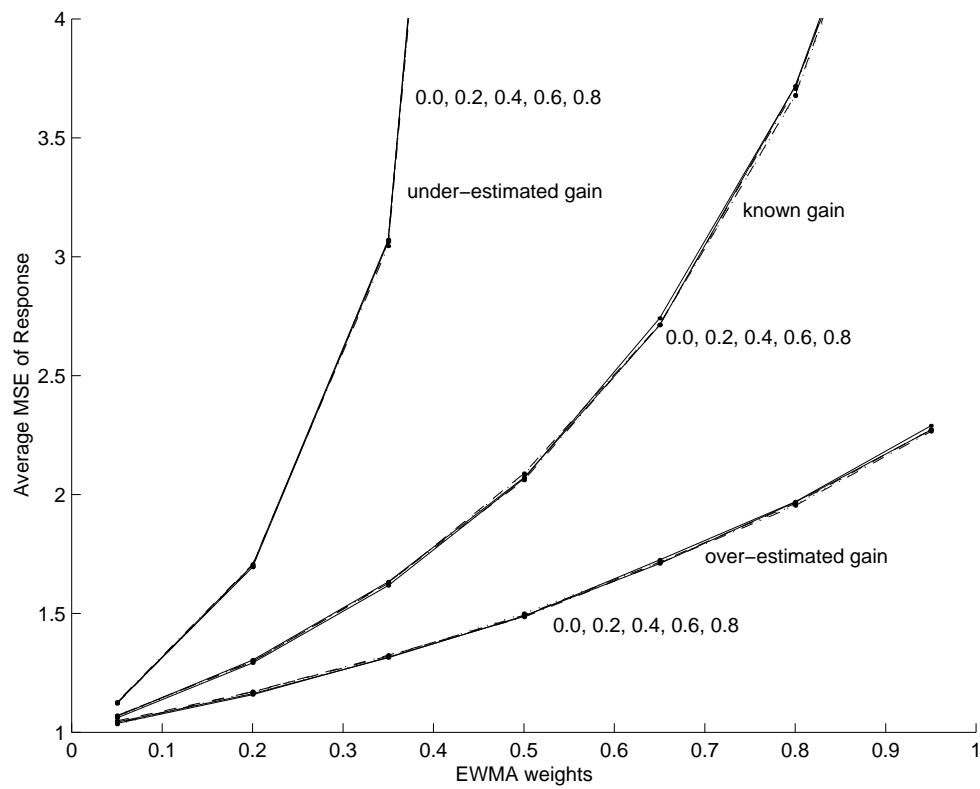
Figure 9: Estimated MSE of the responses for a 3 Input, 3 Output MIMO dEWMA controllers, when the gains are known, underestimated, and overestimated. Labels give off-diagonal elements of $\boldsymbol{\beta}$.
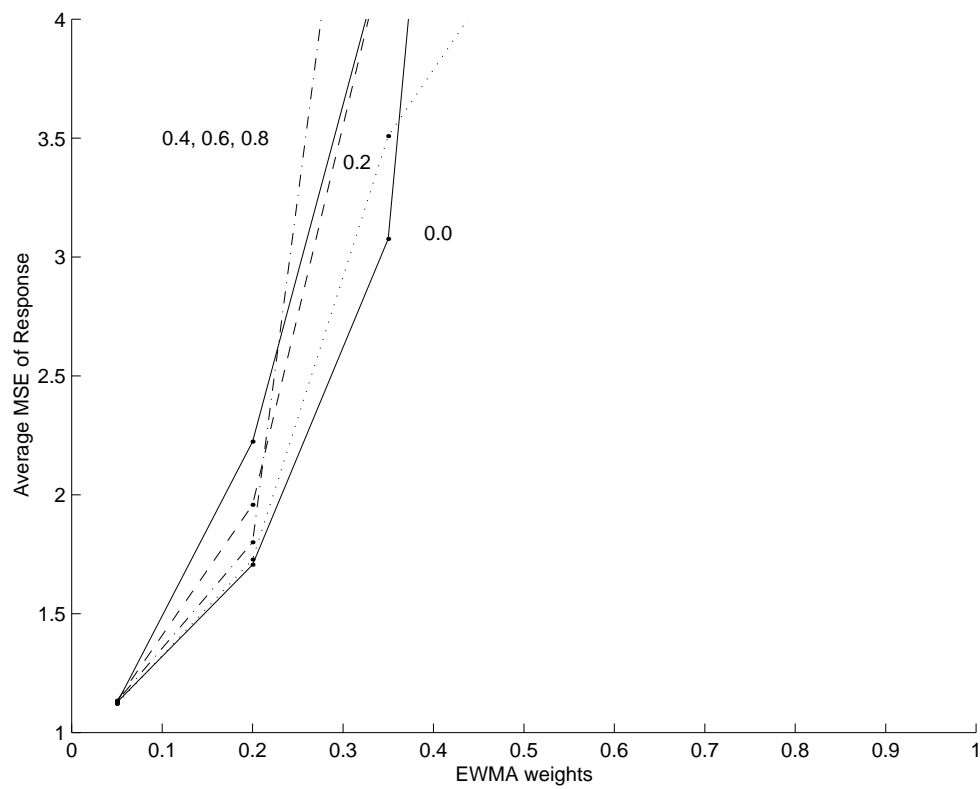
Figure 10: Estimated MSE of the responses for a 3 Input, 3 Output Parallel SISO System with Under-estimated Gain. Labels give off-diagonal elements of $\boldsymbol{\beta}$.
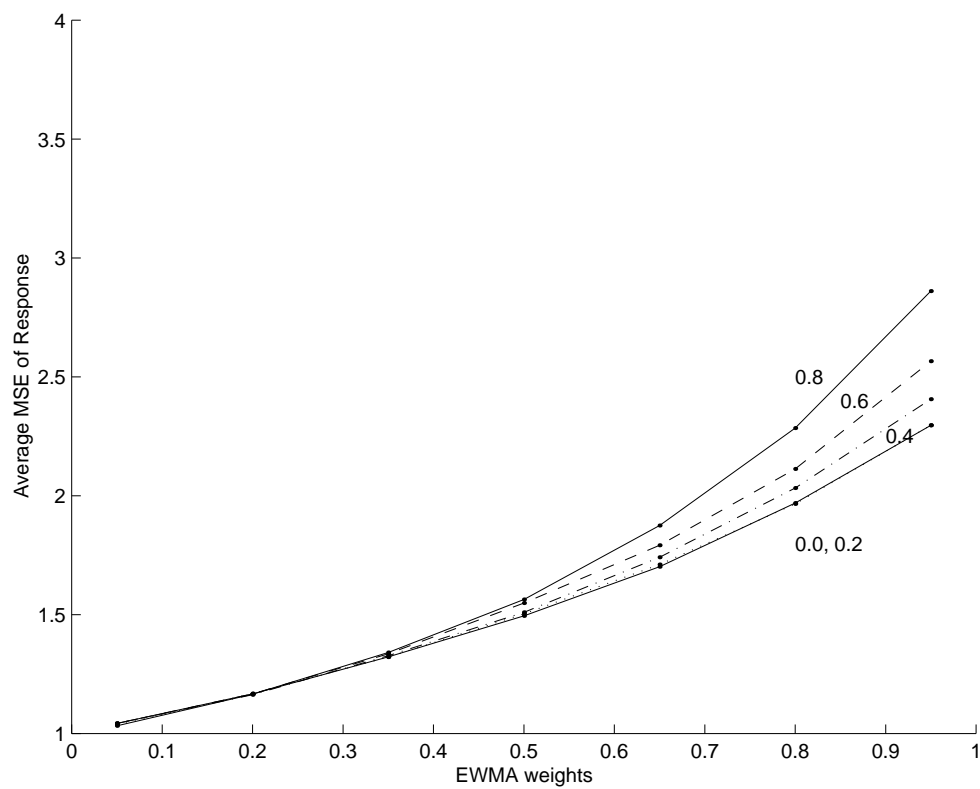
Figure 11: Estimated MSE of the responses for a 3 Input, 3 Output Parallel SISO System with Over-estimated Gain. Labels give off-diagonal elements of $\boldsymbol{\beta}$.