Practical solutions to multivariate feedback control performance assessment problem: reduced a priori knowledge of interactor matrices

Biao Huang a,*,1, Steven X. Ding b, Nina Thornhill c

Abstract

The research on control loop performance monitoring and diagnostics has been and remains to be one of the most active research areas in process control community. Despite of numerous developments, it remains as a considerably challenging problem to obtain a minimum variance control benchmark from routine operating data for multivariable process since the solution relies on the interactor matrix (or inverse time delay matrix). Knowing the interactor matrix is tantamount to knowing a complete knowledge of process models that are either not available or not accurate enough for a meaningful calculation of the benchmark. However, the order of an interactor matrix (OIM) for a multivariable process, a scalar measure of multivariate time delay, is a relatively simple parameter to know or estimate a priori. This paper investigates the possibility to estimate a suboptimal multivariate control benchmark from routine operating data if the OIM is available. The relation between this suboptimal benchmark and the true multivariate minimum variance control benchmark is investigated. Analytical expressions for the lower and upper bounds of the true multivariate minimum variance are derived. Although not minimum variance control, this benchmark answers important practical questions like “at least how much potential of the improvement does the control have by tuning or redesigning?” It is further shown that the proposed suboptimal benchmark is achievable by a practical control provided that the system of interest is minimum phase. Simulation examples illustrate the feasibility of the proposed approach.

Keywords: Performance monitoring; Performance assessment; Control monitoring; Multivariate systems; Interactor matrices

1. Introduction

The research on control loop performance monitoring and diagnostics has been and remains to be one of the most active research areas in process control community. It is estimated that several hundreds of papers have published in this or related direction [5]. In practical side, Eastman Kodak recently reported regular loop monitoring on over 14,000 PID loops. Some commercial control performance assessment software including multivariate performance assessment has also been available. Despite of the success in the research and the applications of univariate control performance assessment, applications of multivariate control performance assessment remain as a challenge. In the editorial for the special issue on control loop performance monitoring [5], it is pointed out that “Various methods to handle the multivariable case have been published in recent years; however, hardly any of them has been successfully
implemented in commercial tools or in permanent customer installations”. Although this statement is unduly pessimistic, it nevertheless points out a gap between theoretical development and practical applications in the area of multivariate control performance assessment.

Among a number of approaches for control performance assessment, minimum variance control (MVC)-benchmark remains to be the most popular benchmark. One of the reasons for the suitability of MVC benchmark to assess performance of control loops in the industry is that it is non-intrusive and routine closed loop operating data is sufficient for the calculation of this benchmark. However, this convenience holds only in the univariate case where the time delay is the only a priori knowledge that needs to be available. For multivariate processes, this simplicity is lost and the time delay is no longer a simple technical concept. Termined as the interactor matrix, its calculation is beyond the knowledge of the time delay between each pair of inputs and outputs. The earlier work in this area is Harris et al. [2] and Huang et al. [11,12], both of which require an explicit knowledge of the interactor matrix.

In recent years, there are growing research interests in reducing the complexity of the a priori knowledge, such as Ko and Edgar [15], Kadali and Huang [13], and McNabb and Qin [17]. Although these attempts have reduced the complexity of the a priori knowledge to some extent, they all require certain information that is computationally simpler but fundamentally equivalent to the interactor matrices, for example, the process Markov parameter matrices, the lower triangular Toeplitz matrix $H^i_n$, or the multivariate time delay (MTD) matrix. Harris et al. [2,4] introduced an extended horizon performance monitoring approach without using the interactor matrix. Kadali and Huang [14] and Shah et al. [25] introduced curvature measures of multivariate performance without relying on the interactor matrix. But these methods stop short of quantifying the quality of such alternative performance measures.

Is it possible to calculate multivariate minimum variance without using the interactor matrix or any of its equivalences? All research conducted so far has indicated infeasibility of such a proposition. However, having to know the interactor matrix or its equivalence has been the major stumbling block to the application of the performance assessment techniques. Given such a dilemma, how to move forward?

The previous research on multivariate performance assessment [2,12,15,25,13,17] has focused on deriving an exact expression of the output error under minimum variance control (referred to as the minimum variance term in this paper). However, the minimum variance term can only be estimated from data. The uncertainty associated with the estimation of this minimum variance term is inevitable [1]. Therefore, the minimum variance estimated from data should not be a single point but an interval in any case.

Consider that we are in the process to evaluate control performance for a multivariate process. Except for the knowledge of the OIM (equivalent to the time delay in the univariate case), we do not have the complete knowledge of the interactor matrix or any of its equivalences. Under this circumstance, it is not possible to estimate the true multivariate minimum variance. Given the available knowledge (OIM only), one would like to ask what is the best we can do. Can we derive a meaningful benchmark control, possibly suboptimal control, with the knowledge of the OIM and routine operating data? For it to be meaningful, this control benchmark must at least possess three properties: (1) It must be achievable by a physically implementable control provided certain conditions are satisfied. (2) Its performance must be no poorer than the existing control performance. (3) It must yield a considerably simpler computation procedure than the calculation of the true minimum variance term. With such a suboptimal control benchmark, one could deliver conclusions such as “We have potential to improve performance of this multivariate control by at least n%”. This is therefore a minimum potential estimated from data while the well-known minimum variance control benchmark delivers a maximum potential. Although this potential may or may not be practically achievable depending on certain conditions such as invertibility of zeros and hard constraints of the actuators, it does provide one with the incentive for a further investigation of the control tuning if the potential turns out to be significant.

Motivated by the above discussions, this paper is concerned with the derivation of the practical but suboptimal performance benchmark. This suboptimal performance benchmark can be estimated from routine operating closed-loop data with the a priori knowledge of OIM. The relation between this suboptimal benchmark and the true multivariate minimum variance control benchmark is investigated. Analytical expressions for the lower and upper bounds of the true multivariate minimum variance are derived. The upper bound can be estimated from routine operating data with the a priori knowledge of the OIM while the estimation of the lower bound needs additional information known as the relative degree to be elaborated shortly. It is further shown that the proposed suboptimal benchmark is achievable by a practical control provided the system is minimum phase.

The remainder of this paper is organized as follows. In Section 2, the concept of the interactor matrix is revisited. The main theoretical derivation for the new practical benchmark and its relation with the true minimum variance control is addressed in Section 3. Simulations are provided in Section 4, followed by concluding remarks in Section 5.
2. Revisit of interactor matrix

Consider the following multivariate process:

\[ Y_t = T U_t + N a_t \]  \hspace{1cm} (1)

where \( T \) and \( N \) are proper (causal), rational transfer function matrices in the backshift operator \( q^{-1} \); \( Y_t, U_t \) and \( a_t \) are output, input and noise vectors of appropriate dimensions. \( q_t \) is further assumed to be white noise with zero mean and \( \text{Var}(a_t) = \Sigma_a \). \( N \) is rational realization of the disturbance spectrum with the standard assumptions [16] that \( N(q^{-1} - 0) = I \) and is minimum phase. Furthermore, we assume that \( T \) does not have non-minimum phase zeros in multivariable sense [7]. This last assumption is only necessary if one needs to guarantee the proposed sub-optimal control to be physically implementable. However, similar to the application of the minimum variance control benchmark if one is only interested in assessing the potential to improve control performance then this assumption is not necessary.

**Lemma 1.** For every \( n \times m \) proper, rational polynomial transfer function matrix \( T \), there is a unique, non-singular, \( n \times n \) lower left triangular polynomial matrix \( D \), such that \( |D| = q^r \) and

\[ \lim_{q^{-1} \to 0} DT = \lim_{q^{-1} \to 0} \tilde{T} = K \]

where \( K \) is a full rank constant matrix, the integer \( r \) is defined as the number of infinite zeros of \( T \), and \( \tilde{T} \) is the delay-free transfer function (factor) matrix of \( T \) which contains only finite zeros. The matrix \( D \) is defined as the interactor matrix and can be written as

\[ D = D_0 q^d + D_1 q^{d-1} + \cdots + D_r q^{d-r} \]  \hspace{1cm} (2)

where \( d \), the maximum power of \( q \) in \( D \), is denoted as the order of the interactor matrix (OIM) and is unique for a given transfer function matrix \([19,24]\). \( v \) is defined here as the relative degree of the interactor matrix (the difference between maximum power and minimum power of \( q \) in \( D \)), and \( D_i \) (for \( i = 0, \ldots, v \)) are coefficient matrices.

**Remark 2.** The OIM, the maximum power of \( q \) in \( D \), may be interpreted as the minimum time elapsed before the control actions can influence all outputs after the control actions apply to the process simultaneously. \(^{2}\) This can be done, for example, by applying impulses or steps to all inputs simultaneously. On the other hand, the minimum power of \( q \) in \( D \) may be interpreted as the minimum time elapsed before the control can influence at least one of the outputs after control actions apply to the process. Therefore, we claim that performance assessment according to OIM and/or relative degree is useful for the following reasons. (1) According to the interpretation above, the OIM \( (d) \) as well as the relative degree \( (v) \) may all be available/calculated from the engineering knowledge of the operators/engineers without actually knowing the process models. If they are not available from engineering knowledge, then simple step tests may be performed as discussed above. (2) From a different perspective, it is also possible that one knows the interactor matrix but is not interested in calculating a true minimum variance benchmark. Instead, using the known OIM, one may calculate a suboptimal but a more realistic control benchmark (to be elaborated shortly). (3) Finally, the OIM can be determined by singular values of closed-loop experiment data [7] without actually calculating the interactor matrix. These are some of the important driving forces for this work.

The interactor matrix \(^3\) \( D \) can be one of the three forms described in the sequel. If \( D \) is of the form: \( D = q^d I \), then the transfer function matrix \( T \) is regarded as having a simple interactor matrix. If \( D \) is a diagonal matrix, i.e., \( D = \text{diag}(q^{d_1}, q^{d_2}, \ldots, q^{d_m}) \), then \( T \) is regarded as having a diagonal interactor matrix. Otherwise, \( T \) is considered to have a general interactor matrix (one realization of which is a triangular interactor matrix). However, the general interactor matrix also has forms other than the lower triangular form. It can be a full matrix or an upper triangular matrix [8,24]. Rogozinski et al. [23] have introduced an algorithm for the calculation of a nilpotent interactor matrix. Peng and Kinnaert [21] have introduced the unitary interactor matrix.

**Definition 3.** Instead of taking the lower triangular form, if an interactor matrix as per Definition 1 satisfies

\[ D^r(q^{-1})D(q) = I \]  \hspace{1cm} (3)

then this interactor matrix is denoted as the unitary interactor matrix.

An expansion similar to Eq. (2) for the unitary interactor matrix can be written as

\[ D = D_0 q^d + D_1 q^{d-1} + \cdots + D_r q^{d-r} \]  \hspace{1cm} (4)

It follows from Eq. (3) that

\[ D^{-1}(q) = D^r(q^{-1}) \]  \hspace{1cm} (5)

The following are two important equations related to the interactor filtering:

\[ q^{-d}D(q) = D_0 + D_1 q^{-1} + \cdots + D_r q^{-r} \]  \hspace{1cm} (6)

---

\(^2\) This heuristic argument does not apply to a special case when the “multivariate time delay” is introduced by the perfect correlation between rows of the first non-zero Markov parameter matrix (cf. Huang and Shah (1999) for detail). In this case, the OIM should be interpreted as the minimum time elapsed before the control actions can influence all outputs in an arbitrary direction. However, this special case of perfect correlation is more of an academic interest.

\(^3\) The complete form of the interactor matrix \( D \) should be written as \( D(q) \). For simplicity, we drop \( (q) \) wherever there is no confusion.
\[ q^d(D(q))^{-1} = q^d(D_q^T + D_{q-1}^T q^{-1} + \cdots + D_0^T q^{-v}) \]  
(7)

Eq. (6) can be directly derived from Eq. (4) and Eq. (7) is the result of combining Eqs. (4) and (5).

3. Proposed performance benchmark for multivariate feedback control

3.1. The pragmatic benchmark

It has been shown in Huang and Shah [7] and Harris et al. [2] that the first \(d\) terms of the following moving average expansion of the interactor filtered multivariate closed-loop output are feedback control invariant, where \(d\) is the OIM

\[ \hat{Y}_t = q^{-d}D Y_t \]

\[ = \hat{F}_0a_t + \hat{F}_1a_{t-1} + \cdots + \hat{F}_{d-1}a_{t-(d-1)} + \hat{F}_d a_{t-d} + \cdots \]

(8)

The first \(d\) terms represent the closed-loop output of \(\hat{Y}_t\) if the minimum feedback control were implemented, where the minimum variance is in the sense of minimizing the trace of the covariance of \(\hat{Y}\). In addition, due to the property of Eq. (3), the trace of the covariance of \(\hat{Y}_t\) is the same as that of \(Y_t\). If the interactor matrix is known, then Eq. (8) can be easily obtained through time series analysis of \(Y_t\), followed by the filtering of \(q^{-d}D\) and then the moving average model expansion, and the minimum variance can be calculated from the first \(d\) terms of Eq. (8). The only problem in practical application is the interactor matrix as discussed in the last section, which is often not available, and not easily understood even if it is available due to its complexity. In particular, an experiment with a sufficient excitation and modelling effort has to be undertaken in order to calculate the interactor matrix or any of its equivalences if the process model is not available. This is the major difficulty for the application of multivariate control performance assessment algorithms.

What is the best one can do if the interactor matrix is not available? At this point, it is natural to proceed the calculation of the “minimum variance” using the counterpart of Eq. (8) with the original output \(Y_t\), if the interactor matrix is unknown, i.e. using

\[ Y_t = F_0a_t + F_1a_{t-1} + \cdots + F_{d-1}a_{t-(d-1)} + F_d a_{t-d} + \cdots \]

(9)

Then the variance of the first \(d\) terms may be used to represent the “minimum variance” as in the univariate case. These first \(d\) terms, however, represent the minimum variance \(d\)-step ahead prediction error (shown shortly) and consequently is the minimum variance control output if the interactor matrix is simply \(q^dI\), i.e. a simple interactor. Is it a meaningful benchmark? Can the use of this obviously simple solution be justified? If this “minimum variance” can be of use for performance assessment, it should at least satisfy the following three conditions:

(1) It must be achievable by a physically implementable control provided certain conditions are satisfied.
(2) Its performance must be no poorer than the existing control performance.
(3) It must yield a considerably simpler computation procedure than the calculation of the true minimum variance term.

The first condition ensures that the benchmark control is practical one and provides certain standard for the existing control to compare with. The satisfaction of this condition will be shown shortly. The second and third conditions are obvious and they ensure that the benchmark control is meaningful and that one can calculate the benchmark even if the interactor matrix is not known. The satisfaction of the second condition can be seen from Eq. (9), where the variance of the first \(d\) terms of the right hand side must be smaller than that of all terms. The third condition is satisfied since no interactor matrix is used in the derivation of Eq. (9).

We therefore formally state the pragmatic performance benchmark for multivariate control performance assessment and then analytically justify its use.

**Algorithm 1.** Let the OIM of the multivariate process under feedback control be \(d\). A set of representative multivariate closed-loop output data \(Y_t\) (error data between the output and the setpoint if setpoint is non-zero) is available. Then perform time series analysis of \(Y_t\) (e.g., using multivariate ARMA or subspace methods) to estimate a time series model in the form of \(Y_t = G_{Y,a}a_t\), where \(G_{Y,a}\) is a transfer function matrix in the backshift operator \(q^{-1}\). Expand this model into the impulse response (moving average) form shown in Eq. (9). This expansion can be easily done by calculating the impulse response coefficients of \(G_{Y,a}\). The pragmatic control performance benchmark is defined and calculated as follows:

\[ \text{MV}_d = \text{tr}(F_0 \Sigma_0 F_0^T + F_1 \Sigma_1 F_1^T + \cdots + F_{d-1} \Sigma_{d-1} F_{d-1}^T) \]

In practice, due to the limited sample size, the estimated time series model will deviate from the true one due to the noise. Thus \(F_i’s\) should be replaced by the estimated counterpart \(\hat{F}_i’s\) and \(\Sigma_a = \text{Cov}(\hat{a}_i)\) where \(\hat{a}_i\) is the residual resulted from the time series modelling of \(Y_t\).

3.2. The achievability of the pragmatic benchmark

The following theorem shows that this pragmatic benchmark control is achievable by a control that is
implementable provided that the system $T$ is minimum phase.

**Theorem 4.** If there exists a controller with $Q^* = Q_0^*$ expressed under the IMC framework in Fig. 1 such that its closed-loop output can be written as

$$Y_t = (F_0 + \cdots + F_{d-1}q^{-(d-1)} + q^{-d}R)a_t$$

where $R$ is a proper rational transfer function matrix, provided $T$ is minimum phase then there exists another physically realizable controller with $Q = Q_d$ such that its closed-loop output can be written as

$$Y_t = (F_0 + \cdots + F_{d-1}q^{-(d-1)})a_t$$

and the controller is given by

$$Q_d = Q_0 + \tilde{T}^{-1}(q^{-d}D)RN^{-1}$$

where, the relation between the IMC control, $Q^*$, and the actual feedback control, $Q$, is given by

$$Q = Q'(I - TQ')^{-1}$$

**Proof.** Under the IMC framework [18], the closed-loop output can be written as

$$Y_t = (I - TQ')Na_t$$

For controller $Q_0'$, the output should be

$$Y_t = (I - TQ_0')Na_t = (F + q^{-d}R)a_t$$

Or in terms of transfer functions, we have

$$(I - TQ_0')N = F + q^{-d}R$$

A controller $Q_d'$ that makes Eq. (11) hold should satisfy the following equation:

$$(I - TQ_d')N = F$$

Subtracting Eq. (17) from Eq. (16) yields

$$T(Q_d' - Q_0')N = q^{-d}R$$

Using the interactor decomposition of $T$ [7]

$$T = D^{-1}\tilde{T}$$

where $\tilde{T}$ is an invertible transfer function matrix, Eq. (18) can be written as

$$D^{-1}\tilde{T}(Q'_d - Q'_0)N = q^{-d}R$$

Solving Eq. (20) results in Eq. (12). Since the order or the highest power of $D$ is $q^d$, $(q^{-d}D)$ is proper (and also stable by the nature of the interactor matrix). As we have assumed that there is no non-minimum phase zeros in $T$, $\tilde{T}$ is also stable (and proper due to the factorization of the interactor). As the closed-loop response $Y_t$ is stable, $R$ must be stable too. Finally, by assumption of $\frac{N}{N'}$ is stable and proper. Therefore, $Q_d'$ as solved from Eq. (12) is proper and stable, and therefore a physically achievable control. \qed

**Remark 5.** In Theorem 4, the condition “If there exists a controller with $Q' = Q_0'$” is always satisfied since we consider a closed-loop operation in control performance assessment. The existing control in the loop is precisely the controller $Q_0'$ expressed under IMC framework.

3.3. The analytical relationship between the proposed benchmark and the true minimum variance

In this section, we shall discuss the relationship between the true minimum variance control and the proposed pragmatic benchmark control. To this end, we derive the bound on the difference between the proposed benchmark control and minimum variance control, bounds on the true minimum variance, and analytical relation between the proposed benchmark control and true minimum variance control.

**Theorem 6.** Let $d$ and $v$ be the order and relative degree of the interactor matrix $D$ for the process expressed in Eq. (1) where $D$ is a unitary interactor matrix.

(1) Then the minimum variance term of the closed-loop output $Y_t$, denoted by $MV$, is upper bounded by the following expressions:

$$MV \leq MV_d$$

where $MV_d$ is the proposed benchmark, $MV$ is the true minimum variance benchmark, and they are given by

$$MV_d = \text{tr}(F_0\Sigma_dF_0^T + F_1\Sigma_dF_1^T + \cdots + F_{d-1}\Sigma_dF_{d-1}^T)$$

$$MV = \min[\text{tr}(EY_tY_t^T)] = \min[\text{tr}(EY_tY_t^T)]$$

$$= \text{tr}(\tilde{F}_0\Sigma_d\tilde{F}_0^T + \tilde{F}_1\Sigma_d\tilde{F}_1^T + \cdots + \tilde{F}_{d-1}\Sigma_d\tilde{F}_{d-1}^T)$$

where $F_0,F_1,\ldots,F_{d-1}$ are given from the moving average expression of closed-loop output $Y_t$ as...
\[ Y_t = F_0 a_t + F_1 a_{t-1} + \cdots + F_{d-1} a_{t-(d-1)} + F_d a_{t-d} + \cdots \]

which can be obtained from time series analysis of \( Y_t \), without relying on an interactor matrix, and \( F_0, F_1, \ldots, F_{d-1} \) are given by Eq. (8), calculation of which, however, requires the knowledge of the interactor matrix. This result \( \text{MV} \leq \text{MV}_d \) also implies a less aggressive proposed benchmark control than the minimum variance control.

(2) The precise difference between \( \text{MV} \) and \( \text{MV}_d \) is given by

\[ \text{MV}_d - \text{MV} = \text{tr}(E_d \Sigma_d E_d^T + \cdots + E_{d+v-1} \Sigma_v E_{d+v-1}^T) \]

where \( E_i \)'s are given by Eq. (A.18).

(3) The following lower bound on \( \text{MV} \) holds:

\[ \text{MV} \geq \text{MV}_{d-v} \]

Furthermore, the difference between \( \text{MV} \) and \( \text{MV}_d \) is bounded by the following expression:

\[ \text{MV}_d - \text{MV} \leq \text{tr}(F_{d-v} \Sigma_d F_{d-v}^T + \cdots + F_{d-1} \Sigma_1 F_{d-1}^T) \]

where \( v \geq 1 \) otherwise \( \text{MV}_d - \text{MV} = 0 \). Therefore, if one wishes to know how far away \( \text{MV}_d \) is from \( \text{MV} \), one could use Eq. (28) to give a bounded estimate.

**Proof.** Proof is given in the appendix. \( \square \)

**Remark 7.** From Theorem 6, we can make the following observations:

(1) The precise difference between \( \text{MV} \) and \( \text{MV}_d \) given by Eq. (26) is dependent on the interactor matrix. Thus, it cannot be calculated without the knowledge of the interactor matrix.

(2) However, the upper bound on the difference between \( \text{MV} \) and \( \text{MV}_d \) given by Eq. (28) can be calculated from original output \( Y_t \) without using an interactor matrix.

(3) The number of terms \( F_i \)'s that contribute to the difference between \( \text{MV} \) and \( \text{MV}_d \) is \( v \), the relative degree of the interactor matrix. Therefore, smaller the relative degree, closer \( \text{MV}_d \) to \( \text{MV} \).

(4) Considering \( F_i \)'s are the impulse response of the closed-loop multivariable transfer function matrix of stable systems, they decay with the time. Therefore, if \( d - v \) is sufficiently large, then \( \text{MV}_d - \text{MV} \rightarrow 0 \).

**Remark 8.** The output variance under minimum variance control for a simple interactor matrix can be interpreted as the variance of the output prediction error. This fact is now elaborated using Eqs. (25) and (22). As \( a_t \) is white noise, an optimal \( j \) step ahead prediction of \( Y_t \) can be directly derived from Eq. (25) as

\[ Y_{\hat{Y}_{t-j}} = F_0 a_{t-j} + F_1 a_{t-j-1} + \cdots \]

Thus, the optimal \( j \) step prediction error can be calculated by subtracting Eq. (25) by Eq. (29)

\[ Y_t - Y_{\hat{Y}_{t-j}} = F_0 a_t + F_1 a_{t-1} + \cdots + F_{j-1} a_{t-j+1} \]

and its variance is given by

\[ \text{tr}[E(Y_t - Y_{\hat{Y}_{t-j}})(Y_t - Y_{\hat{Y}_{t-j}})^T] = \text{tr}(F_0 \Sigma_0 F_0^T + F_1 \Sigma_1 F_1^T + \cdots + F_{j-1} \Sigma_{j-1} F_{j-1}^T) \]

Therefore, \( \text{MV}_d \) in Eq. (22) can be interpreted as the variance of the optimal \( d \) step ahead prediction error.

From the perspective of prediction error discussed above, the proposed practical benchmark can be applied to process under constrained control. The rationale is that, more predictable the output data is, more potential of improvement the control system has. However, due to the constraint, the bound \( \text{MV}_d \) is not necessarily achievable. Nevertheless, it does provide us with an estimated potential to improve control performance. This is similar to the minimum variance control benchmark \( \text{MV} \) that is not necessarily achievable but does provide an estimated maximum potential of improvement.

**Remark 9.** We have so far focused on the equal-weighted quadratic performance measure \( J = EY_t^T Y_t \), i.e. all outputs are equally important in terms of their variance reduction. If one is interested in a weighted quadratic performance measure such as \( J = EY_t^T W Y_t \), then all the unitary interactor matrix should be changed to the weighted unitary interactor matrix \([7]\) and the OIM and the relative degree should now be those of the weighted interactor matrix. Then the extension of the results obtained from this paper can be done by replacing \( G_{Y_a} \) with \( W^{1/2} \) and then carrying on subsequent transformation to the moving average model as is done in Algorithm 1.

### 4. Simulation

In this section, we will perform simulations to demonstrate the calculation of the proposed benchmark \( \text{MV}_d \), bound on \( \text{MV}_d - \text{MV} \), lower and upper bounds on the true minimum variance. We will also verify the effects of the OIM and the relative degree on the upper and lower bounds. All calculations are performed using both theoretical models and estimated models from routine operating data.

Consider a 2 \( \times \) 2 minimum-phase multivariable process with the open-loop transfer function matrix \( T \) and disturbance transfer function matrix \( N \) given by
The interactor matrix can be calculated as
\[
T = \begin{bmatrix}
q^{-d-1} & 0.5q^{-d} \\
1-0.4q^{-1} & 1-0.1q^{-1} \\
0.3q^{-(d+1)} & 0.6q^{-1} \\
1-0.4q^{-1} & 1-0.8q^{-1}
\end{bmatrix}
\]
\[
N = \begin{bmatrix}
1-0.5q^{-1} & q^{-1} \\
1-0.5q^{-1} & 1-0.q^{-1} \\
q^{-1} & 1.0
\end{bmatrix}
\]

The interactor matrix can be calculated as
\[
D = \begin{bmatrix}
-0.9578q^{d-1} & -0.2873q^{d-1} \\
-0.2873q^d & 0.9578q^d
\end{bmatrix}
\]

The white noise excitation, \(a_t\), is a two-dimensional normal-distributed white noise sequence with \(\Sigma_a = I\). The output performance is measured by \(J = \mathbb{E}[Y_t^T Y_t]\).

Consider that the following multiloop controller is implemented on the process:
\[
Q = \begin{bmatrix}
0.5-0.2q^{-1} & 0 \\
1-0.5q^{-1} & 0.25-0.2q^{-1} \\
0 & (1-0.5q^{-1})(1+0.5q^{-1})
\end{bmatrix}
\]

Then the closed-loop transfer function from the white noise \(a_t\) to the output \(Y_t\) can also be calculated as
\[
Y_t = (I + TQ)^{-1}Na_t
\]

By decomposing Eq. (32) into Markov parameter (impulse response) form, one can calculate theoretically the variance of the minimum variance prediction error over any step ahead, say \(d\) step ahead, through the first \(d\) terms of the Markov parameters.

Table 1 shows the theoretical \(MV_d\) and performance index with \(MV_d\) as the benchmark. The performance index \(\eta_{MV_d}\) is defined as
\[
\eta_{MV_d} = \frac{MV_d}{\text{tr}[\text{Cov}(Y)]}
\]

where the theoretical value of \(\text{tr}[\text{Cov}(Y)]\) can be calculated according to the \(H_2\) norm of \((I + TQ)^{-1}N\Sigma_a^{1/2}\).

To estimate it from data, one can either directly calculate it from data or calculate the \(H_2\) norm of the estimated \(G_m\Sigma_a^{1/2}\) while the latter is the preferred approach.

The last row of Table 1 shows the percentage difference between the actual variance \(\text{tr}[\text{Cov}(Y)]\) and the benchmark control variance \(MV_d\). Since \(MV_d\) is guaranteed to be achievable (due to minimum phase of \(T\), there exists a practically achievable control to guarantee the percentage reduction of variance as indicated in the last row. For example, if the OIM is known as 8, then the current variance can be reduced by at least 51%, which has a sufficient incentive to re-tune the control or implement advanced multivariable control.

If we are interested in the difference between \(MV_d\) and MV, i.e. the difference between the proposed benchmark suboptimal control and the true minimum variance control, we may calculate the upper bound on the difference. The result is listed in Table 2. If we are also interested in the range of the true minimum variance MV, we can calculate a bound on the true MV. The results are plotted in Fig. 2. The calculation of bounds needs the additional a priori knowledge, the relative degree. One can see that indeed the minimum variance is bounded by the calculated lower and upper bounds. The interval between the lower and upper bounds shrinks quickly as the OIM increases. Little difference is observed between the upper and lower bound when the OIM is larger than 4. Performance indices are also plotted in Fig. 3.

So far we have simulated our results according to a known model and these results have demonstrated our conclusion. All of above results can, however, be estimated from routine operating data as discussed next with a priori knowledge of OIM and the relative degree (if the bounds are also to be estimated).

First, both \(MV_d\) and \(\eta_{MV_d}\) can be estimated from routine operating data with the a priori knowledge of OIM only. Furthermore, if the relative degree is also known, the lower and upper bounds on the true minimum variance can also be calculated from routine operating data. To see these results, with \(a_t\) being white noise of unit variance, a set of 2000 samples of closed-loop output \(Y\) are simulated. Using the time series analysis algorithm and subsequent moving average model transformation, the estimated upper and lower bounds together with the true minimum variance are shown in Fig. 4. The comparison of performance indices between the true minimum variance benchmark and estimated \(MV_d\) benchmark is shown in Fig. 5. Although slightly noisy due to the estimation error, the results generally resemble their theoretical counterparts shown in Fig. 3.

To see the effect of the relative degree on the difference between \(MV_d\) and MV, consider the following model:
\[
T = \begin{bmatrix}
q^{-3} & 0.5q^{-4} \\
1-0.4q^{-1} & 1-0.1q^{-1} \\
0.3q^{-(2+1)} & 0.6q^{-1} \\
1-0.4q^{-1} & 1-0.8q^{-1}
\end{bmatrix}
\]

where \(v\) varies from 1 to 9 in the simulation. The corresponding interactor matrix is given by
\[
D = \begin{bmatrix}
-0.9578q^3 & -0.2873q^{2+v} \\
-0.2873q^4 & 0.9578q^{3+v}
\end{bmatrix}
\]
It follows from the expression of the interactor matrix that \( m \) is indeed the relative degree of the interactor matrix. The disturbance model and control law are not changed i.e. they are the same as the previous example. The theoretical MV, its lower and upper bounds are plotted in Fig. 6. Their estimated version is shown in Fig. 8. From these two figures, one can clearly see the fact that the lower and upper bounds get wider with increasing relative degree. Performance indices with true minimum variance control as benchmark and with MV\(_d\) as benchmark respectively are shown in Fig. 7 and their estimated version is shown in Fig. 9. Once again the dif-

<table>
<thead>
<tr>
<th>( d )</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max(MV(_d) – MV)</td>
<td>2.36</td>
<td>1.13</td>
<td>0.65</td>
<td>0.44</td>
<td>0.35</td>
<td>0.31</td>
<td>0.29</td>
<td>0.28</td>
<td>0.27</td>
</tr>
<tr>
<td>MV(_d)</td>
<td>4.36</td>
<td>6.02</td>
<td>6.86</td>
<td>7.30</td>
<td>7.55</td>
<td>7.69</td>
<td>7.77</td>
<td>7.82</td>
<td>7.85</td>
</tr>
<tr>
<td>MV</td>
<td>3.77</td>
<td>5.86</td>
<td>6.75</td>
<td>7.16</td>
<td>7.37</td>
<td>7.48</td>
<td>7.54</td>
<td>7.58</td>
<td>7.60</td>
</tr>
</tbody>
</table>

Fig. 2. True minimum variance and bounds vs OIM.

Fig. 3. True performance indices (PI) vs OIM.

Fig. 5. Comparison of performance indices vs OIM.

Fig. 6. True minimum variance and bounds vs relative degree.

Fig. 7. True performance indices (PI) vs relative degree.

Fig. 8. True minimum variance and estimated bounds vs relative degree.
ference in performance indices increases as the relative degree increases.

5. Conclusion

In this paper, we have proposed a practical solution to multivariable feedback control performance assessment problem. Calculation of the proposed benchmark needs only the knowledge of the order of the interactor matrix. We have also derived the bound on the difference between the proposed benchmark variance and the true minimum variance, and the lower and upper bounds on the true multivariate minimum variance. Calculation of bounds needs an additional information, the relative degree of the interactor matrix. It is shown that there exists an implementable control to achieve the proposed benchmark performance provided the system is minimum phase. Although this new benchmark is not the same as the true minimum variance control, it provides a practical solution to multivariate control performance assessment problem and addresses some important practical questions such as at least how much potential the existing control can be improved. The simulation examples have demonstrated the feasibility of the proposed algorithms and verified the results arrived from this work.

Acknowledgment

This work is supported in part by Humboldt Research Fellowship of Germany.

Appendix A. Proof of Theorem 6

Proof. Since $d - v$ and $d$ are the minimum and maximum powers of $q$ in the interactor matrix, and $\text{MV}_d$ and $\text{MV}_{d-v}$ would be the minimum variance if the interactor matrix were $q^{d-v}I$ and $q^vI$ respectively, the two inequalities, (21) and (27), may be argued heuristically. The rigorous proof of these two inequalities together with other conclusions of the theorem will proceed in three steps following the inequalities/equalities (21), (26), (27) and (28).

Step 1:

Proof of inequality (21)

Re-write Eq. (25) as

$$\begin{align*}
Y_t &= [F_0 + F_1 q^{-1} + \cdots + F_{d-1} q^{-(d-1)} + F_d q^{-d} + \cdots]a_t \\
&= [(F_0 + F_1 q^{-1} + \cdots + F_{d-1} q^{-(d-1)}) \\
&+ q^{-d}(F_d + F_{d+1} q^{-1} + \cdots + F_{2d-1} q^{-(d-1)}) \\
&+ q^{-2d}(\cdots) + \cdots]a_t \\
&\triangleq [F_{0d-1} + q^{-d} F_{1d-1} + q^{-2d} F_{2d-1}d-1 + \cdots]a_t \quad (A.1)
\end{align*}$$

Now using the definition of the interactor filtered output $\hat{Y}_t$ and Eq. (A.1), we have

$$\hat{Y}_t = [q^{-d} DF_{0d-1} + q^{-2d} DF_{1d-1} + q^{-3d} DF_{2d-1}d-1 + \cdots]a_t \quad (A.2)$$

Using Eq. (6), the first term on right-hand side of Eq. (A.2) can be written as

$$q^{-d} DF_{0d-1} = (D_0 + D_1 q^{-1} + \cdots + D_v q^{-v}) \times (F_0 + F_1 q^{-1} + \cdots + F_{d-1} q^{-(d-1)}) = E_0 + E_1 q^{-1} + \cdots + E_{d+v-1} q^{-(d+v-1)} \quad (A.3)$$

where $E_i$’s can be calculated from

$$E_i = \sum_{m=1}^{n} D_m F_n, \quad i = 0, 1, \ldots, d + v - 1 \quad (A.4)$$

Similarly, the second term on right-hand side of Eq. (A.2) may be written as

$$q^{-2d} DF_{1d-1} = q^{-2d}[S_0 + S_1 q^{-1} + \cdots + S_{d+v-1} q^{-(d+v-1)}] \quad (A.5)$$

Similar results apply to $q^{-3d} DF_{2d-1}$ and so on.

Substituting Eqs. (A.4) and (A.5) into Eq. (A.2) yields

$$\hat{Y}_t = [E_0 + E_1 q^{-1} + \cdots + E_{d-1} q^{-(d-1)} + E_d q^{-d} + \cdots + E_{d+v-1} q^{-(d+v-1)} + S_0 q^{-2d} + S_1 q^{-(2d+1)} + \cdots]a_t$$

Comparing Eq. (A.6) with Eq. (8) immediately yields

$$\tilde{F}_0 = E_0, \quad \tilde{F}_1 = E_1, \ldots, \tilde{F}_{d-1} = E_{d-1} \quad (A.7)$$

Therefore,

$$\text{MV} = \min\{\text{tr}(E(Y_t Y_t^T))\} = \min\{\text{tr}(E(\hat{Y}_t \hat{Y}_t^T))\} \quad (A.8)$$

$$= \text{tr}(\tilde{F}_0 \Sigma \tilde{F}_0^T + \cdots + \tilde{F}_{d-1} \Sigma \tilde{F}_{d-1}^T) \quad (A.9)$$

$$= \text{tr}(E_0 \Sigma E_0^T + E_1 \Sigma E_1^T + \cdots + E_{d-1} \Sigma E_{d-1}^T) \quad (A.10)$$

$$= \text{tr}(E_0 \Sigma E_0^T + E_1 \Sigma E_1^T + \cdots + E_{d+v-1} \Sigma E_{d+v-1}^T)$$

$$- \text{tr}(E_d \Sigma E_d^T + \cdots + E_{d+v-1} \Sigma E_{d+v-1}^T) \quad (A.11)$$
\[
\text{To this end, we need similarly to define}
\]
\[
\tilde{Y}_t = [\tilde{F}_0 + \hat{F}_1 q^{-1} + \cdots + \hat{F}_{d-1} q^{-(d-1)} + \hat{F}_d q^{-d} + \cdots] a_t
\]
\[
= [(\tilde{F}_0 + \hat{F}_1 q^{-1} + \cdots + \hat{F}_{d-1} q^{-(d-1)})
+ q^{-d}(\hat{F}_d + \hat{F}_{d+1} q^{-1} + \cdots + \hat{F}_{2d-1} q^{-(d-1)})
+ q^{-2d}(\cdots) + \cdots] a_t
\]
\[
\triangleq [\tilde{F}_{0d-1} + q^{-d}\tilde{F}_{d2d-1} + q^{-2d}\tilde{F}_{2d3d-1} + \cdots] a_t
\]

Now using the definition of \(\tilde{Y}_t = q^{-d}D Y_t\) and Eq. (5), we have
\[
Y_t = q^d D^T (q^{-1}) [\tilde{F}_{0j0} + q^{-d}\tilde{F}_{d2d-1}
+ q^{-2d}\tilde{F}_{2d3d-1} + \cdots] a_t
\]

Using Eq. (7), the first term on the right hand side of Eq. (22) can be written as
\[
q^d D^T (q^{-1}) \tilde{F}_{0j0} = q^d (D^T + D_{v-1}^T q^{-1} + \cdots + D_0^T q^{-v})
\times (\tilde{F}_0 + \tilde{F}_1 q^{-1} + \cdots + \tilde{F}_{d-1} q^{-(d-1)})
= q^d (\tilde{F}_0 + \tilde{E}_1 q^{-1} + \cdots + \tilde{E}_{d+v-1} q^{-(d+v-1)})
\]

where \(\tilde{E}_i\)'s can be calculated from
\[
\tilde{E}_i = \sum_{m=n-i} D_{v-1}^T \tilde{F}_n, \quad i = 0, 1, \ldots, d + v - 1
\]

Since \(Y_t = q^d D^T (q^{-1}) \tilde{Y}_t\) must be proper, all terms in Eq. (23) that have positive power of \(q\) must be zero. As a result,
\[
q^d D^T (q^{-1}) \tilde{F}_{0j0} = \tilde{E}_0 + \tilde{E}_{v+1} q^{-1} + \cdots + \tilde{E}_{d} q^{-d-1} + \cdots
+ \tilde{E}_{d+v-1} q^{-(d+v-1)}
\]

Following the similar approach as that of deriving the upper bound,
\[
q^d D^T (q^{-1}) q^{-d} \tilde{F}_{d2d-1} = q^{-(d-1)} (D^T + D_{v-1}^T q^{-1}
+ \cdots + D_0^T q^{-v}) \tilde{F}_{d2d-1}
= \tilde{S}_0^{-(d-1)} + \tilde{S}_1 q^{-(d-2)}
+ \cdots + \tilde{S}_{d+v-1} q^{-(2d-1)}
\]

where the calculation of \(\tilde{S}_i\)'s follows the same formula as the calculation of \(\tilde{E}_i\)'s in Eq. (24). We could also continue to derive the term \(q^d D^T (q^{-1}) q^{-2d} \tilde{F}_{2d3d-1}\) using the same approach, but the leading term will then start from \(q^{-(2d-1)}\), and so on.

Substituting Eqs. (25) and (26) into (22), and then following the same argument as the proof of the upper bound, we can get the following results
\[
\tilde{E}_v = F_0, \quad \tilde{E}_{v+1} = F_1, \ldots, \tilde{E}_{d-1} = F_{d-1}
\]
One can then obtain

$$MV = \text{tr}(\tilde{F}_d \Sigma_a \tilde{F}_d^T + \cdots + \tilde{F}_{d-1} \Sigma_a \tilde{F}_{d-1}^T)$$

$$= \text{tr}(\tilde{E}_d \Sigma_a \tilde{E}_d^T + \cdots + \tilde{E}_{d+1} \Sigma_a \tilde{E}_{d+1}^T)$$

$$+ \text{tr}(\tilde{E}_d \Sigma_a \tilde{E}_d^T + \cdots + \tilde{E}_{d+v-1} \Sigma_a \tilde{E}_{d+v-1}^T)$$

$$= MV_{d-v} + \text{tr}(\tilde{E}_d \Sigma_a \tilde{E}_d^T + \cdots + \tilde{E}_{d+v-1} \Sigma_a \tilde{E}_{d+v-1}^T)$$

$$\geq MV_{d-v}.$$  \hspace{1cm} (A.28)

Thus we have proved the lower bound. Now from inequality (A.19), we obtain

$$MV_d - MV \leq MV_{d-v} - MV_{d-v}$$ \hspace{1cm} (A.29)

Substituting Eqs. (A.16) and (A.20) yields

$$MV_d - MV \leq \text{tr}(F_{d-v} \Sigma_d F_{d-v}^T + \cdots + F_{d-1} \Sigma_d F_{d-1}^T)$$ \hspace{1cm} (A.30)

where \(d - v \leq d - 1\) otherwise \(MV_d - MV = 0. \square\)

References


