PLS-Based Memory Control Scheme for Enhanced Process Monitoring

Fouzi Harrou, *Member, IEEE*, Ying Sun

King Abdullah University of Science and Technology (KAUST),
Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division,
Thuwal, 23955-6900, Saudi Arabia, Email: fouzi.harrou@kaust.edu.sa

Abstract—**Fault detection is important for safe operation of various modern engineering systems. Partial least square (PLS) has been widely used in monitoring highly correlated process variables. Conventional PLS-based methods, nevertheless, often fail to detect incipient faults. In this paper, we develop new PLS-based monitoring chart, combining PLS with multivariate memory control chart, the multivariate exponentially weighted moving average (MEWMA) monitoring chart. The MEWMA are sensitive to incipient faults in the process mean, which significantly improves the performance of PLS methods and widen their applicability in practice. Using simulated distillation column data, we demonstrate that the proposed PLS-based MEWMA control chart is more effective in detecting incipient fault in the mean of the multivariate process variables, and outperform the conventional PLS-based monitoring charts.**

I. INTRODUCTION

Fault detection and diagnosis are central to build safe and reliable modern industrial processes [1]. If anomalies in industrial systems are not detected promptly, they can affect plant productivity, profitability, and safety [2]. Process monitoring is employed by various process industries [1], [3], [4], [5]. Partial least square (PLS) is among the most widely used multivariate statistical process monitoring method for monitoring multivariate processes [6], [7]. Roughly speaking, PLS aims to extract from the predictor (input) variables latent variables which are the most correlated and relevant to predicted (output) variables [8]. By extracting the useful data information from the original dataset, and then using monitoring indices such as T^2 and Q statistics, PLS is a major statistical tool in multivariate techniques and is widely used with highly cross-correlated data [9]. In this framework, several PLS variants have been proposed to overcome some shortcomings of the classical PLS, such as multiway PLS [10], a modified PLS [11], [12], multi-block PLS [13], a multi-scale PLS [14], dynamic PLS [15], recursive PLS [16] as well as multi-phase PLS [17]. Therefore, PLS-based process monitoring method as well its variants has been largely exploited and used to different engineering applications [18], [7], [19].

Handling incipient anomalies is a key challenges in building safe and reliable processes. The detection of incipient anomalies is central in maintaining the normal operation of a system by providing an early warning which helps in avoiding serious damage and subsequent economic loss. Indeed, incipient anomalies is characterized by a weak signature that requires detection indicators that have high sensitivity to small changes. However, PLS-based monitoring indices such as T^2 and Q statistics utilize information solely from the actual observation and they are relatively insensitive for detecting incipient faults in process mean. Therefore, this potentially makes conventional PLS-based monitoring less efficient in this case. To handle this problem alternative chart, such as the MEWMA monitoring chart, which is rested on a decision statistic that takes into account information from past observations with that of current observations, can be used [20]. Thus, the purpose of this work was to exploit the advantage of the PLS modeling and those of memory control chart MEWMA by developing PLS-based MEWMA monitoring scheme to achieve enhanced detection performance compared with the traditional PLS monitoring techniques.

The next section briefly introduces PLS-based fault detection and Section III review the MEMA monitoring chart. Next, the proposed PLS-based MEWMA chart is presented in section IV. The proposed technique is validated on simulated distillation column data in Section V. Subsequently this paper is concluded with recommendation for future work in Section VI.

II. PARTIAL LEAST SQUARES (PLS)

PLS consists to find the latent variables (unmeasured variables), LVs, from the process data by capturing the largest variability in the data and achieves the maximum cross-correlation among the input and the output variables [8]. Given an input data matrix $\mathbf{X} \in R^{n \times m}$ having n observations and m variables, and an output data matrix $\mathbf{Y} \in R^{n \times p}$ consisting of p response variables, a PLS model is formally determined by the inner model and the outer model [8]. The outer model, which links LVs, and the response and predictor matrices, can be expressed as [21], [9]:

$$\begin{cases} \mathbf{X} = \widehat{\mathbf{X}} + \mathbf{E} = \sum_{i=1}^{l} \mathbf{t}_i \mathbf{P}_i^T + \mathbf{E} = \mathbf{TP}^T + \mathbf{E} \\ \mathbf{Y} = \widehat{\mathbf{Y}} + \mathbf{F} = \sum_{i=1}^{l} \mathbf{u}_i \mathbf{q}_i^T + \mathbf{F} = \mathbf{UQ}^T + \mathbf{F} \end{cases} \quad (1)$$

where $\widehat{\mathbf{X}}$ and $\widehat{\mathbf{Y}}$ represent modeling matrices of $\mathbf{X}$ and $\mathbf{Y}$, respectively, the matrices $\mathbf{T} \in R^{n \times l}$ and $\mathbf{U} \in R^{n \times q}$ consist of l kept LVs of the predictor and response data,

respectively. Blocks $\mathbf{E} \in R^{n \times m}$ and $\mathbf{F} \in R^{n \times p}$ are the residuals matrices that were the unexplained variance of the predictor and response data, respectively, Blocks $\mathbf{P} \in R^{m \times l}$ and $\mathbf{Q} \in R^{p \times q}$ are the loading of matrices $\mathbf{X}$ and $\mathbf{Y}$, respectively. The number of LVs, l, can be determined via cross-validation for example [22]. The inner relationship can then be described by:

$$\mathbf{U} = \mathbf{TB} + \mathbf{H}, \tag{2}$$

where $\mathbf{B}$ represents a regression matrix and $\mathbf{H}$ represents a residual matrix. The response $\mathbf{Y}$ can be computed as:

$$\mathbf{Y} = \mathbf{TBQ}^T + \mathbf{F}^*. \tag{3}$$

In the conventional PLS-based monitoring, two statistics, the T^2 and Q statistics, are usually utilized for fault detection. The T^2 statistic describes the variability caught by the PLS model while the Q statistic computes the residuals that are missed by the model. The T^2 statistic is used in detecting anomalies corresponded at atypical variations within a LVs subspace. T^2 can be computed as [23]:

$$T^2 = \sum_{i=1}^{l} \frac{t_i^2}{\sigma_i^2}, \tag{4}$$

where, σ_i^2 represents the estimated variance of the corresponding latent variable t_i. When the monitored process is conform with the desired performances at the $i - th$ time point, it is obvious that $T^2 < T_{l,n,a}^2$, where $T_{l,n,a}^2$ is a control limit given in [24]. The T^2 chart gives a signal of the presence of an anomaly when $T^2 > T_{l,n,a}^2$.

The Q statistic, on the other hand, which is defined as [24]:

$$Q = \mathbf{e}^T \mathbf{e} \tag{5}$$

captures the changes in the residual subspace. Where $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$ represents the residuals vector, which is the difference between the new observation, $\mathbf{x}$, and its prediction, $\hat{\mathbf{x}}$, via PLS model. Once Q statistic of the new observation exceed the control limits Q_α at significance level α given in [24]. Figure 1 shows an example of a data with one observation has a large Q value while an other with large T^2.
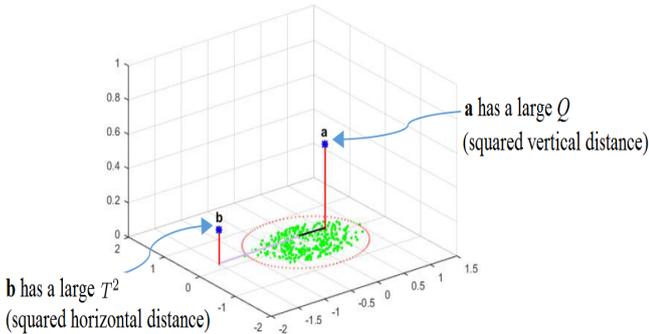

Fig. 1. A set of projected observations (green) and two original three-dimensional observations a and b.

Although the two methods have their advantages and disadvantages, both tend to fail to detect small faults [20]. Motivated by the sensitivity of MEWMA for detecting incipient shifts in the multivariate process mean [20], we propose an innovative PLS-based MEWMA FD method for multivariate processes. Towards this end, a PLS is used as model and MEWMA scheme is used to monitor the LVs of the input variable obtained from PLS model. In the next section, we briefly describe the MEWMA chart.

III. MULTIVARIATE EWMA MONITORING CHART

Statistical process control has been successfully used in applications across diverse industries as an efficient tool for monitoring the status of a process and for helping to identify abnormalities [20]. Multivariate statistical monitoring charts such as MEWMA and multivariate cumulative sum (MCUSUM) were developed to inspect several process variables simultaneously [25], [20].

The MEWMA monitoring chart was initially proposed by Lowry et al. [25], [20]. This chart is constructed based on exponential weighting of available observations, a design that provides improved sensitivity to small changes in the mean of a multivariate process. The main advantage of this chart is that it can be easily implemented in real time because of their low computational cost. Assume that there is $\mathbf{X}_t = (X_1, X_2, \ldots, X_m)^T$, an m-dimensional set of observations at time t. The MEWMA charting statistic is defined as follows [25]:

$$\begin{cases} \mathbf{Z}_t = \Lambda \mathbf{X}_t + (\mathbf{I}_{m \times m} - \Lambda)\mathbf{Z}_{t-1} & \text{if } t > 0 \\ \mathbf{Z}_0 = \mu_0, & \text{if } t = 0 \end{cases} \tag{6}$$

where $\mathbf{Z}_0 = \mathbf{0}$, $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_m)$ and $\lambda_j \in (0, 1]$ is a weighting parameter for the j-th component of $\mathbf{X}$, for $j = 1, 2, \ldots, m$, and $\mathbf{I}_{m \times m}$ is the identity matrix. TThe MEWMA chart is generally suitable for detecting relatively small changes when λ is chosen small, and it is suitable for detecting relatively large changes when λ is chosen large [20]. If there is no practical reason to give different weight for each component, then we can choose $\lambda_1 = \lambda_2 = \cdots = \lambda_m = \lambda$. In such cases,

$$\mathbf{Z}_t = \lambda \mathbf{X}_t + (1 - \lambda)\mathbf{Z}_{t-1}. \tag{7}$$

The MEWMA monitoring statistic is:

$$\mathbf{V}_t^2 = \mathbf{Z}_t^T \Sigma_{\mathbf{Z}_t}^{-1} \mathbf{Z}_t, \tag{8}$$

where $\Sigma_{\mathbf{Z}_t}$ is the covariance matrix of $\mathbf{Z}_t$. The chart declares the presence of anomaly when $\mathbf{V}_t^2$ exceeds a predefined threshold h. The value of h can be computed by simulation to obtain a specific false-alarm rate [20].

In the next section, the MEWMA monitoring chart will be integrated with PLS to extend its FD abilities for the case where small or incipient faults are of interest.

IV. Proposed Fault Detection Algorithm

The implementation of the developed monitoring method comprises two stages: off-line modelling and on-line monitoring. In the off-line modelling phase, a reference PLS model is constructed in the normal operating condition. Then, the fault detection procedure is executed by using the reference PLS model with MEWMA chart in the on-line monitoring phase. The PLS-based MEWMA monitoring strategy is outlined as follows.

1) **Construction of the reference PLS model (off-line)**: This phase is composed of four steps presented as follows:
   - Collect the training data set representative of a healthy condition.
   - Scale the data to zero mean and unit variance.
   - Construct the reference PLS model.
     - Select the number of LVs,
     - Decompose the input/output data as given in equation (1),
     - Calculate the regression matrix $\mathbf{B}$
   - Compute the MEWMA control limits using the ignored LVs of the input variables obtained from the reference PLS model.

2) **Detection of fault (on-line)**: This phase is composed of four steps,
   - For a new sample data, apply the same scaling used in modeling stage.
   - Compute the ignored LVs of the input variables.
   - Compute the MEWMA monitoring statistic, $\mathbf{V}^2$.
   - Verify the presence of faults. If the monitoring statistic ($\mathbf{V}^2$) exceeds the control limits a fault is declared.

V. Results

The method is tested using a distillation column process simulated by Aspen (see [7] for details) with added zero-mean Gaussian noise, where the predictor variables consist of ten temperatures (Tc) in different stages of the monitored column, feed flow rates and reflux stream, and the composition of the light component in the distillate stream represents the response variable. The Aspen simulator is used to generate 1024 data samples to be used in constructing the reference PLS model. Figure 2 shows the training and testing datasets, each of size 512, from the dynamic input-output model for the case where the signal-to-noise ratio is 10. A PLS model is fitted to the scaled training dataset. Based on the cross-validation technique, three LVs were required for the PLS model. Now, the fault detection abilities of the proposed PLS-based MEWMA method will be assessed.

Bias sensor fault:
In this case study, the detection of an abrupt anomaly in temperature sensor is investigated. An abrupt anomaly in the temperature measurement T_{c3} with a magnitude of 2% is introduced at the 100th sample and removed at the
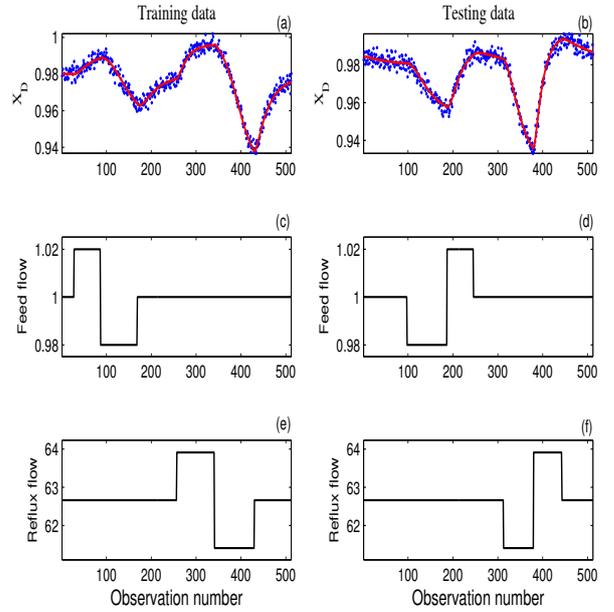

Fig. 2. Dynamic input-output data used for training and testing the model with noise SNR=10 (solid red line: noise-free data; blue dots: noisy data).

150th sample of the testing data. Results of the Q and T^2 statistics are demonstrated in Figure 3, top and middle panels, respectively. Monitoring results of the three control charts are shown in Figure 3. The dashed horizontal lines represent a 95% confidence limit used to identify the possible anomalies. It can be observed that the T^2 and Q statistics are completely insensitive to this anomaly. From this case study, it can be seen that the traditional PLS-based charts are indeed ineffective in detecting small and persistent mean shifts. This is mainly due to the fact that the PLS-based T^2 and Q charts use only the observed data at that instant to take decision about the process performance and they ignore the past data. On the other hand, the result of PLS-based MEWMA chart for the considered bias fault, which are displayed in the bottom panel of Figure 3, clearly indicate that the proposed strategy can successfully detect this fault. The weighting parameter of MEWMA is chosen to be $\lambda = 0.3$. It can also be concluded that the detection ability and sensitivity to small anomalies can be increased by taking into account the information of the actual and past data in the decision rule.

VI. Conclusion

In this paper, a fault detection approach that is based on PLS is proposed. PLS has been used in this work as a modeling framework for fault detection using MEWMA control chart. The greater ability of the MEWMA chart to detect small faults makes it very attractive compared to the conventional PLS monitoring statistics. The main contribution of this work is to integrate PLS modeling with the MEWMA control scheme to improve fault especially in the presence of small or incipient faults. To achieve this objective, the MEWMA control scheme is applied on the ignored PCs obtained from
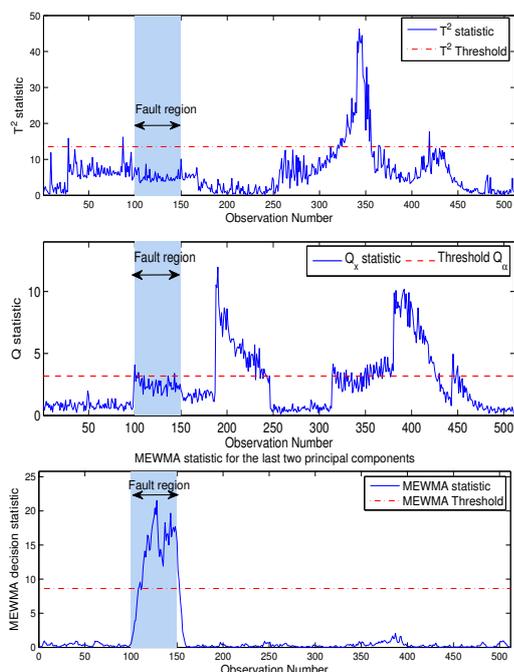
Fig. 3. The evolution of the T^2 (top), Q (middle), and MEWMA (bottom) statistics in the presence of bias fault in $\mathbf{x}_5$. The dashed horizontal lines in plots (top), (middle) and (bottom) are control limits of the related control charts.

the PLS outer model. The simulation results obtained using simulated distillation column data demonstrate the advantage of the proposed PLS-based MEWMA FD methods over the conventional PLS-based methods (Q and T^2) in detecting small faults in highly correlated data.

REFERENCES

[1] R. Isermann, "Model-based fault-detection and diagnosis: status and applications," *Annual Reviews in Control*, vol. 29, pp. 71–85, 2005.

[2] S. Yin, S. Ding, X. Xie, and H. Luo, "A review on basic data-driven approaches for industrial process monitoring," *IEEE Transactions on Industrial Electronics,*, vol. 61, no. 11, pp. 6418–6428, 2014.

[3] S. Ding, S. Yin, K. Peng, H. Hao, and B. Shen, "A novel scheme for key performance indicator prediction and diagnosis with application to an industrial hot strip mill," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 4, pp. 2239–2247, 2013.

[4] F. Harrou, F. Kadri, S. Chaabane, C. Tahon, and Y. Sun, "Improved principal component analysis for anomaly detection: Application to an emergency department," *Computers & Industrial Engineering*, vol. 88, pp. 63–77, 2015.

[5] F. Harrou, M. Nounou, and H. Nounou, "A statistical fault detection strategy using PCA based EWMA control schemes," in *9th Asian Control Conference (ASCC)*. IEEE, 2013, pp. 1–4.

[6] S. Wold, H. Ruhe, H. Wold, and W. D. III, "The collinearity problem in linear regression. the partial least squares (PLS) approach to generalized inverses," *SIAM Journal on Scientific and Statistical Computing*, vol. 5, no. 3, pp. 735–743, 1984.

[7] F. Harrou, M. Nounou, H. Nounou, and M. Madakyaru, "PLS-based EWMA fault detection strategy for process monitoring," *Journal of Loss Prevention in the Process Industries*, vol. 36.

[8] P. Geladi and B. Kowalski, "Partial least-squares regression: a tutorial," *Analytica chimica acta*, vol. 185, pp. 1–17, 1986.

[9] F. Harrou and Y. Sun, "Enhanced anomaly detection via PLS regression models and information entropy theory," in *2015 IEEE Symposium Series on Computational Intelligence*. IEEE, 2015, pp. 383–388.

[10] P. Nomikos and J. MacGregor, "Multi-way partial least squares in monitoring batch processes," *Chemometrics and intelligent laboratory systems*, vol. 30, no. 1, pp. 97–108, 1995.

[11] Y. Zhang and L. Zhang, "A modified pls regression model for quality prediction," in *Intelligent Control and Automation (WCICA), 2014 11th World Congress on*. IEEE, 2014, pp. 1383–1387.

[12] G. Wang and S. Yin, "Quality-related fault detection approach based on orthogonal signal correction and modified PLS," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 2, pp. 398–405, 2015.

[13] J. MacGregor, C. Jaeckle, C. Kiparissides, and M. Koutoudi, "Process monitoring and diagnosis by multiblock PLS methods," *AIChE Journal*, vol. 40, no. 5, pp. 826–838, 1994.

[14] P. Teppola and P. Minkkinen, "Wavelet-PLS regression models for both exploratory data analysis and process monitoring," *Journal of chemometrics*, vol. 14, no. 5-6, pp. 383–399, 2000.

[15] G. Lee, C. Han, and E. Yoon, "Multiple-fault diagnosis of the tennessee eastman process based on system decomposition and dynamic PLS," *Industrial & engineering chemistry research*, vol. 43, no. 25, pp. 8037–8048, 2004.

[16] S. Qin, "Recursive PLS algorithms for adaptive data modeling," *Computers & Chemical Engineering*, vol. 22, no. 4, pp. 503–514, 1998.

[17] N. Lu, F. Gao, and F. Wang, "Sub-pca modeling and on-line monitoring strategy for batch processes," *AIChE journal*, vol. 50, no. 1, pp. 255–259, 2004.

[18] A. Kembhavi, D. Harwood, and L. Davis, "Vehicle detection using partial least squares," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 6, pp. 1250–1265, 2011.

[19] R. Muradore and P. Fiorini, "A PLS-based statistical approach for fault detection and isolation of robotic manipulators," *Industrial Electronics, IEEE Transactions on*, vol. 59, no. 8, pp. 3167–3175, 2012.

[20] D. C. Montgomery, "Introduction to statistical quality control," *John Wiley& Sons, New York*, 2005.

[21] T. Kourti and J. MacGregor, "Process analysis, monitoring and diagnosis using multivariate projection methods: A tutorial," *Chemometrics and Intelligent Laboratory Systems*, vol. 28, no. 3, pp. 3–21, 1995.

[22] B. Li, J. Morris, and E. Martin, "Model selection for partial least squares regression," *Chemometrics and Intelligent Laboratory Systems*, vol. 64, no. 1, pp. 79–89, 2002.

[23] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *Journal of Educational Psychology*, vol. 24, pp. 417–441, 1933.

[24] S. Qin, "Statistical process monitoring: Basics and beyond," *Journal of Chemometrics*, vol. 17, no. 8/9, pp. 480–502, 2003.

[25] C. Lowry, W. Woodall, C. Champ, and S. Rigdon, "A multivariate exponentially weighted moving average control chart," *Technometrics*, vol. 34, no. 1, pp. 46–53, 1992.