Minimal Required Excitation for Closed-Loop Identification: Implications for PID Control Loops*

Yuri A.W. Shardt, *Member IEEE*, Biao Huang, *Member IEEE*

Abstract**—***In industry, developing efficient and effective methods for using all the available data is important. Previous work has examined the ability to identify process models using closed-loop data where the reference signal is not being excited. However, it was shown that only with sufficiently large time delays or fast sampling rates could the data be used. Therefore, this paper will examine, in the context of proportional, integral, and derivative controllers (PID), the minimal required excitation conditions for a reference signal in order to identify the process. Similarly to the previous case, it is shown that the complexity of the required reference signal depends strongly on the sampling rate and time delay. However, since many fast processes without time delay can be modelled as first-order systems, they can indeed be identified when the excitation in the reference signal is a simple function or sequence of functions. Using a heated tank simulation, the effect on the continuous time parameters is investigated for different sampling rates and excitations signals. It is shown that as expected an external reference signal can identify previously unidentifiable cases.*

1. INTRODUCTION

In industry applications, the development of process models from closed-loop data is an important exercise with many different applications, including fault-detection and process control. Although it is preferable to use data obtained without perturbing the system, recently, it has been shown that such an approach cannot be used in all cases, for instance, in processes with small time delays or slow sampling rates. In such cases, there is a need to perturb the system. Although it is well known that white noise or other sufficiently complex perturbations can excite the system, such random or large excitations can cause unnecessary process variability or angst amongst the operators. Therefore, there is a need to determine the relevant minimal excitations or perturbation to identify a model.

Determining the conditions for identifying a discrete model from closed-loop data has a long history. The earliest work in discrete closed-loop identification, which includes work by Box, McGregor, Söderström, and Stoica, sought to determine the theoretical limits on the delay to guarantee consistency of parameter estimates in the absence of a reference signal. [1, 2, 3]. This research led to the development of general conditions for closed-loop identification. In order to obtain a solution various assumptions were made including dealing with an autoregressive moving average model with exogenous input (ARMAX) with at least a single sample delay [3] or various degrees of *a priori* knowledge of pole-zero cancellations in the closed-loop transfer function [4, 5, 6]. Removing these conditions would provide general results that would then have broad applicability. The general case for closed-loop identification with no perturbations in the reference signal has been recently completed [7]. However, there has not been any work done to extend these results to the case where there are perturbations in the reference signal.

Therefore, the objectives of this paper are 1) to extend the previously developed closed-loop identification results to the case where there are changes in the references signal; 2) using the conditions to obtain relationships between required reference signal excitation, model orders, and PID controller; and 3) finally, to provide simulation experiments to verify the theoretical results.

2. IDENTIFIABLITY IN CLOSED-LOOP SYSTEMS

A. Theoretical Results

Assume that the process of interest can be described as a closed-loop prediction error (PE) system, similar to that shown in Figure 1, that is,

$$G_c = \frac{X\left(z^{-1}\right)}{Y\left(z^{-1}\right)} y_t, G_p = \frac{z^{-n_k} B\left(z^{-1}\right)}{A\left(z^{-1}\right) F\left(z^{-1}\right)}, G_l = \frac{C\left(z^{-1}\right)}{A\left(z^{-1}\right) D\left(z^{-1}\right)} \quad (1)$$

where the X-polynomial is given as

$$X\left(z^{-1}\right) = 1 + \sum_{i=1}^{n_X} x_i z^{-i} \quad (2)$$

n_X is the order of the polynomial; the Y-, A-, C-, D-, and F-polynomials are defined similarly to the X-polynomial; the B-polynomial is defined as

$$B\left(z^{-1}\right) = \sum_{i=1}^{n_B} \beta_i z^{-i} \quad (3)$$

n_B is the order of the B-polynomial; and n_k is the time delay in the process, which excludes the one sample time delay introduced by the sampler. For simplicity of presentation, the backshift operator, z^{-1}, will be dropped in the following sections.

For such a process, the one-step ahead predictor, $y(t \mid t-1, \theta)$ can be written as

*Research supported by the National Science and Research Council (NSERC) of Canada.

Y. A. W. Shardt was with the University of Alberta when this research was performed. He is currently an Alexander von Humboldt Fellow at the University of Duisburg-Essen, Duisburg, North Rhine-Westphalia, Germany (e-mail: yuri.shardt@ualberta.ca).

B. Huang is with the University of Alberta, Edmonton, Alberta, Canada (e-mail: biao.huang@ualberta.ca).

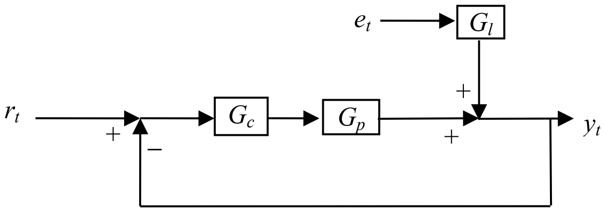

Figure 1: Generic Closed-loop Process

$$y\left(t \mid t-1,\theta\right) = \underbrace{G_l^{-1} G_p}_{W_u} u_t + \underbrace{\left(1 - G_l^{-1}\right)}_{W_y} y_t \tag{4}$$

A quasistationary vector signal, ψ_t, is persistently exciting if $\bar{E}(\psi_t \psi_t^T) > 0$ [8]. Furthermore, a quasistationary vector signal, r_t, is sufficiently rich of order n_r if the following regressor is persistently exciting,

$$\phi_{1,n} \equiv \begin{bmatrix} r(t-1) \\ r(t-2) \\ \vdots \\ r(t-n) \end{bmatrix} = \begin{bmatrix} z^{-1} \\ z^{-2} \\ \vdots \\ z^{-n} \end{bmatrix} r_t \tag{5}$$

In order to distinguish between any two candidate models for a given closed-loop data set, it is necessary that the following two conditions hold [8]:

$$\begin{cases} E\left(\left(\Delta W_y r_t\right)^2\right) = 0 \\ \quad \Delta W_y \equiv G_c \Delta W_u \end{cases} \tag{6}$$

where

$$W_u = G_l^{-1} G_p = \frac{z^{-n_k} DB}{CF}, W_y = 1 - G_l^{-1} = \frac{C - AD}{C} \tag{7}$$

and Δ represents the difference between the two candidate models 1 and 2, that is,

$$\Delta W = W_1 - W_2 \tag{8}$$

Substituting the results from Equation (7) into Equation (6) gives

$$\begin{cases} E\left(\left[\left(\frac{A_2 D_2}{C_2} - \frac{A_1 D_1}{C_1}\right) r_t\right]^2\right) = 0 \\ \left(\frac{X}{Y}\right)\left(\frac{z^{-n_k} D_1 B_1}{C_1 F_1} - \frac{z^{-n_k} D_2 B_2}{C_2 F_2}\right) = \left(\frac{A_2 D_2}{C_2} - \frac{A_1 D_1}{C_1}\right) \end{cases} \tag{9}$$

In order to solve for the closed-loop conditions, consider that

1) There are possible cancellations between D_1 and F_1, so that $D_1 = H\bar{D}_1$ and $F_1 = H\bar{F}_1$, where H is a polynomial with order n_H and $\bar{D}_1$ and $\bar{F}_1$ are coprime.
2) There are possible cancellations between $C_1\bar{F}_1 Y$ and $\left(\bar{F}_1 A_1 HY + z^{-n_k} B_1 X\right)$, so that

$$\begin{gathered} C_1 \bar{F}_1 Y = T\overline{C_1 \bar{F}_1 Y} = T\bar{C}_1 \bar{\bar{F}}_1 \bar{Y} \\ \left(\bar{F}_1 A_1 HY + z^{-n_k} B_1 X\right) = T\overline{\left(\bar{F}_1 A_1 HY + z^{-n_k} B_1 X\right)} \end{gathered} \tag{10}$$

where T is a polynomial with order $n_T = \min(n_C + n_F + n_Y,\ n_F + n_A + n_Y,\ n_B + n_X)$ and $T\bar{C}_1\bar{\bar{F}}_1\bar{Y}$ and $\overline{\left(\bar{F}_1 A_1 HY + z^{-n_k} B_1 X\right)}$ are coprime. This takes into consideration any potential pole-zero cancellations in the closed-loop system transfer function. It should be noted that since the common terms in the denominator may only appear after the terms in the denominator have been combined, it can be shown that [7]

$$\bar{F}_1 A_1 HY + z^{-n_k} B_1 X = M_1 N + P_1 O = T\bar{M}_1 \bar{N} + T\bar{P}_1 \bar{O} \tag{11}$$

where M_1, N, P_1, and O are polynomials constructed so that the common terms between numerator and denominator of the closed-loop transfer function given by the T-polynomial, appear in both terms of the sum, the orders of the M_1 and P_1 are, respectively, equal to that of the sum of the A_1- and F_1-polynomials and the B_1-polynomials, and, finally, $\bar{N}$ and $\bar{O}$ are coprime. The number of overbars placed over the polynomials represents the number of potential reductions in the order of the polynomial due to non-coprimedness of the selected polynomials.

Theorem 1: (*Routine-Operating Case*) Assume that there are no excitations in the reference signal and the assumptions described above hold, then the system can be identified if the following relationship holds among the orders of the polynomials and the discrete time delay:

$$\max\begin{pmatrix} n_X + n_k - n_F - n_A, \\ n_Y - n_B \end{pmatrix} \geq n_D + \min\begin{pmatrix} n_C + n_F + n_Y, \\ n_A + n_F + n_Y, \\ n_B + n_X \end{pmatrix} \tag{12}$$

Proof: Since the proof follows the same approach as used in [9] and is fully proven in [10], it is omitted.

Theorem 2: *(Excited Reference Signal Case*) Assume that the assumptions described above hold, then the minimal excitation required for the reference signal is

$$\begin{aligned} n_r \geq\ & n_D + \min\left(n_C + n_F + n_Y, n_A + n_F + n_Y, n_B + n_X\right) \\ & + \min\left(n_F + n_A - n_X - n_k, n_B - n_Y\right) \end{aligned} \tag{13}$$

with the caveat that the points of support for r_t do not coincide with any possible zeroes of X on the unit circle. The richness order of the reference signal is denoted by n_r and is defined by Equation (5)

Proof: The proof of this theorem is given in Appendix I.

Corollary 1: (*General PI controller*) Solving Equation (13) for a proportional, integral (PI) controller shows that the reference signal must have a persistent excitation order of

$$n_r \geq n_D + \min\left(n_C + n_F, n_A + n_F, n_B\right) + \min\left(n_F + n_A - n_k, n_B\right) \quad (14)$$

Proof: This corollary results by letting $n_Y = n_X = 1$ in Equation (13) and simplifying the resulting equation.

Corollary 2: (*First-Order Box-Jenkins Process with PI controller*) Solving Equation (13) for a proportional, integral (PI) controller shows that the reference signal must have a persistent excitation order of

$$n_r \geq 2 + n_D - n_k \quad (15)$$

Proof: This corollary results by letting $n_Y = n_X = n_F = n_B = 1$ and $n_A = 0$ in Equation (13) and simplifying the resulting equation.

Corollary 3: (*General PID controller*) Solving Equation (13) for a proportional, integral, derivative (PID) controller shows that the reference signal must have a persistent excitation order of

$$n_r \geq n_D + \min\left(n_C + n_F, n_A + n_F, n_B + 1\right) + \min\left(n_F + n_A - 1 - n_k, n_B\right) \quad (16)$$

Proof: This corollary results by letting $n_Y = 1$ and $n_X = 2$ in Equation (13) and simplifying the resulting equation.

Corollary 4: (*First-Order Box-Jenkins Process with PID controller*) Solving Equation (13) for a proportional, integral, derivative (PID) controller shows that the reference signal must have a persistent excitation order of

$$n_r \geq 1 + n_D - n_k \quad (17)$$

Proof: This corollary results by letting $n_Y = n_F = n_B = 1$, $n_A = 0$, and $n_X = 2$ in Equation (13) and simplifying the resulting equation.

B. Practical Implications for a PID Controller

In previous work, it has been noted that a PID controller is not complex enough to provide sufficient excitation to identify a process from routine-operating data [7]. These results, especially Corollary 2 show the additional excitation that needs to be provided in order to identify the process. Furthermore, it can be seen that the discrete time delay impacts on the ability to identify the process, where the relationship between the continuous and discrete time delays is

$$n_k = \frac{\tau_d}{\tau_s} \quad (18)$$

where τ_d is the continuous time delay and τ_s is the sampling time. This implies that decreasing the sampling time will decrease the required signal complexity. However, in practice, it may not always be feasible to reduce the sampling time so that the reference signal's excitation is minimised or not required. Firstly, the process may have an extremely small continuous time delay which would in turn require a correspondingly small sampling rate (say every nanosecond or faster). Secondly, a process could have no observable time delay which would imply that it would be impossible to identify it without any excitation. However, it is often the case that such processes can be modelled with simple dynamics such as a first-order Box-Jenkins (BJ) model. In such cases, it can be seen from Corollary 4 that, as long as the disturbance model is low order, then for PID controllers the required excitation are either sine waves (with persistent order of 2) or step changes (with persistent order of 1). Even more complex disturbance models would require only a sine wave that contains multiple frequencies. On the other hand, it can be seen that from Corollary 2, for PI controllers, a more complex result is required from the start. This agrees well with the observation that a PI controller is less complex than a PID controller and therefore cannot provide as much excitation. In such cases, it can be seen by comparing Corollaries 2 and 4 that the difference in complexity between the two controllers is 1 persistent excitation order. This also holds between the general results given by Corollaries 1 and 3, except that there exist regions where the performance is equivalent.

3. SIMULATION EXAMPLE

In order to test the above results, a simulation of a large heated tank will be run, whose characteristics can be described as follows:

$$G_p = \frac{1.54}{200s+1}e^{-30s}, G_l = \frac{1}{200s+1} \quad (19)$$

A discrete proportional, integral, and derivative (PID) controller will be used with parameters given as

$$K_c = \frac{1}{1.54}, \tau_I = 200, \tau_D = 4.5 \quad (20)$$

and sampling time, $\tau_s \in \{1, 2, 10, 15, 50, 100, 200\}$. The discrete PID controller can be written as

$$G_c = K_c\left(1 + \frac{\tau_s}{\tau_I\left(1 - z^{-1}\right)} + \tau_D\tau_s\left(1 - z^{-1}\right)\right) \quad (21)$$

It can be noted that a PI controller was also be designed by setting $\tau_D = 0$. A first-order Box-Jenkins model with time delay given as

$$y_t = \frac{\beta z^{-1-n_k}}{1-\alpha z^{-1}}u_t + \frac{1+c_1 z^{-1}}{1+d_1 z^{-1}}e_t \quad (22)$$

will be used, where the time delay, n_k, is computed as

$$n_k = \left\lfloor \frac{\theta}{\tau_s} \right\rfloor + 1 \quad (23)$$

where $\lfloor\cdot\rfloor$ is the floor function. This process was simulated for 5,000 seconds for both the case where there was no change in the reference and the case where the reference signal was changed from 0 to 1 at 2,500 seconds under identical noise conditions. For each sampling time, the above simulation was repeated 100 times and the gain and process time constants were obtained for both cases.

The discrete-time parameters were converted into the continuous time parameters using the following exact discretisation formulae [11]:

$$\hat{G}_p\left(z^{-1}\right) = \frac{\hat{\beta} z^{-1-n_k}}{1-\hat{\alpha} z^{-1}} \Rightarrow K = \hat{\beta} / \left(1-\hat{\alpha}\right); \tau_p = \frac{-\tau_s}{\ln\left(\hat{\alpha}\right)} \tag{24}$$

where τ_s is the sampling time, $\hat{\beta}$ is the estimated value of β, and $\hat{\alpha}$ is the estimated value of α.

It is expected that as the sampling time increases, the accuracy of the estimated parameters should decrease. The exact point at which is occurs depends on the conditions of the given experiment. Since there may not be pole-zero cancellations at the given points, then the system may be identifiable even if the condition is not satisfied. At fast sampling rates, it is expected that all methods should give better parameter estimates than at slow sampling rates.

The results are shown in Figure 2 and Figure 3. Figure 2 shows the mean deviation between the estimated process time constant and the true value (200 s) as a function of the sampling time for all four cases: PI and PID controllers with or without a change in the reference signal. All time constants greater than 400 s were removed. This was only the case for the no reference signal cases, where for each signal about 20 such cases were found. The largest number occurred at higher sampling rates and for the PI controller. Figure 3 shows the same results except that it is now for the deviation in the estimated process gain (K). Gains greater than 5.0 in absolute value were removed. Again, the only place here there were any such removals was in the case without any reference signal.

From Figure 2 and Figure 3, it can be seen that the cases with a reference signal provides a much better estimate that those without. This is expected as a reference signal will provide more excitations in the process and so that the signal-to-noise ratio will be greater and hence it will be easier to estimate the parameters. As expected, the parameter estimates for the case with a reference signal are well estimated up to about 50 samples. This is a slightly larger limit than expected ($n_r = 1 \geq 2 + 1 - \tau_d / \tau_s = 2$). This implies that the given system does not have the full number of pole-zero cancellations that are expected. Furthermore, it can be noted that in this particular example, the performance of the PI and PID controllers are similar. This is to be expected given the rather small value selected for the derivative term. Secondly, the behaviour of the case without a reference signal shows that if the sampling time is appropriately selected, it is possible to identify the process parameters. However, given the minimal excitation present, there will be a larger spread in the process values. In fact when comparing the standard deviations for the time constant estimates, the values for the case with a reference signal is about 5.5 s, while for the case without a reference signal is 90 s. Finally, it can be noted that cases with no reference signal only produces accurate estimates to a sampling time of 30 s. As expected, this is a smaller value than for the reference signal case.

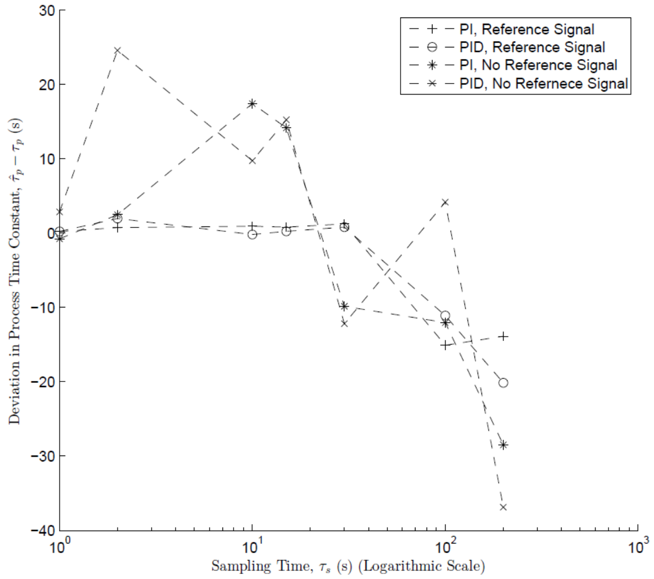

Figure 2: Mean Deviation between Estimated and True Time Constant (τ_p) for all 4 Cases

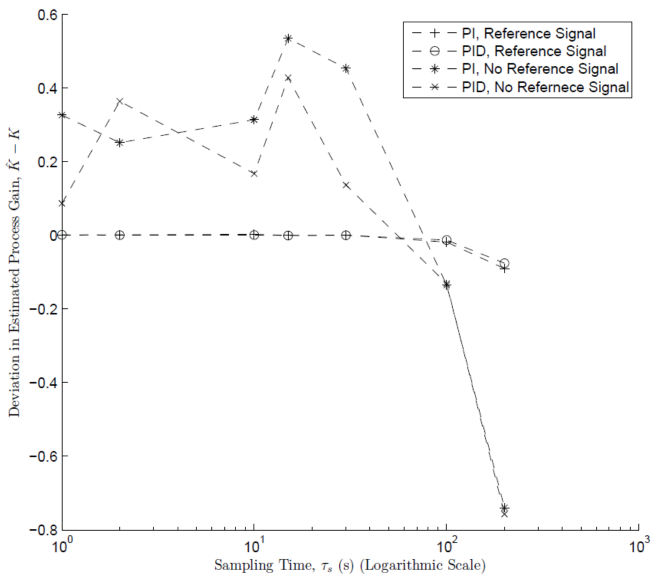

Figure 3: Mean Deviation between Estimated and True Process Gain (K) for all 4 Cases

In order to investigate further the effect of sampling time and reference signal on the distribution of the parameter estimates, the histograms of the process time constants will be examined for both controller types and cases. The results for the gain are similar. The same bounds for rejecting values will be used. The results are shown in Figure 4 and Figure 5. As expected, the distribution of the parameter estimates is much greater for the no reference signal case, where the excitation is weaker than for the reference signal case.

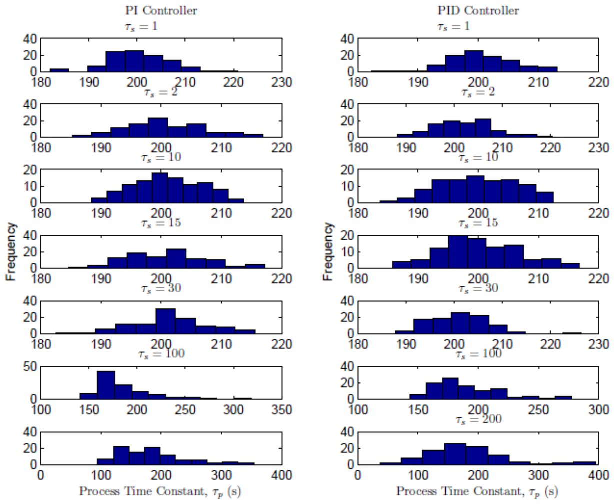

Figure 4: Histograms of $\hat{\tau_p}$ for the PI and PID Controllers for the Reference Signal Case

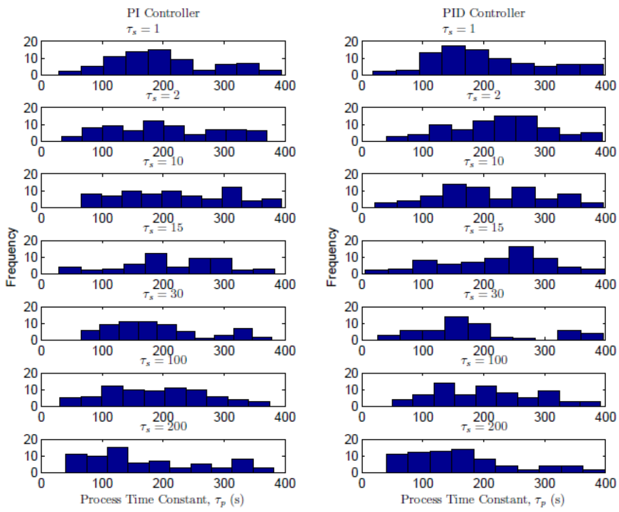

Figure 5: Histograms of $\hat{\tau_p}$ for the PI and PID Controllers for the No Reference Signal Case

4. CONCLUSIONS

In this paper, the conditions for closed-loop identification of a process so as to minimise the required excitation of the reference signal were developed. Special cases were developed for the proportional, integral (PI) and proportional, integral, and derivative (PID) controllers. These results were compared with the previously developed closed-loop conditions for identifiability in the absence of a reference signal. They are both similar in that the sampling time can affect the region over which the system can be identified. Similar to the no reference signal case, a fast sampling rate can improve the identifiability of the system. With external excitation provided by the reference signal, the region of identifiability can be expanded to include a much larger region.

Practically, these results provide a complete picture that will allow arbitrary systems to be identified irrespective of the actual time delay.

REFERENCES

[1] G. E. Box and J. F. MacGregor, “The Analysis of Closed-Loop Dynamic-Stochastic Systems,” *Technometrics,* vol. 16, no. 3, pp. 391-398, 1974.

[2] G. E. Box and J. F. MacGregor, “Parameter Estimation with Closed-Loop Operating Data,” *Technometrics,* vol. 18, no. 4, pp. 371-380, 1976.

[3] T. Söderström and P. Stoica, System Identification, New York, New York: Prentice Hall, 1989.

[4] T. Söderström, I. Gustavsson and L. Ljung, “Identifiability conditions for linear systems operating in closed loop,” *International Journal of Control,* vol. 21, no. 2, pp. 243-255, 1975.

[5] M. Gevers, A. Bazanella and M. Ljubiša, "Informative data: how to get just sufficiently rich?," in *Proceedings of the 47th IEEE Conference on Decision and Control*, Cancun, 2008.

[6] M. Gevers, A. S. Bazanella, X. Bombois and L. Mišković, “Identification and the information matrix: how to get just sufficiently rich?,” *IEEE Transactions on Automatic Control,* vol. 54, no. 12, pp. 2828-2840, 2009.

[7] Y. A. W. Shardt and B. Huang, “Closed-loop identification with routine operating data: Effect of time delay and sampling time,” *Journal of Process Control,* vol. 21, pp. 997-1010, 2011b.

[8] L. Ljung, System Identification Theory for the User, Upper Saddle River, New Jersey: Prentice Hall PTR, 1999.

[9] Y. A. W. Shardt and B. Huang, “Closed-Loop Identification Condition for ARMAX Models Using Routine Operating Data,” *Automatica,* pp. 1534-1537, 2011a.

[10] Y. A. W. Shardt, “Data Quality Assessment for Closed-Loop System Identification and Forecasting with Application to Soft Sensors,” University of Alberta, Edmonton, 2012.

[11] B. Huang and R. Kadali, Dynamic Modeling, Predictive Control and Performance Monitoring: A Data-driven Subspace Approach, London, England: Springer-Verlag London Limited, 2008.

APPENDIX I: PROOF OF THEOREM 2

Assume that the second condition of Equation (9) does not hold, that is, the model cannot be identified using routine operating data. As well, assume that the same cancellations given as 1 to 2 in Section 2.A apply.

From the first component of Equation (9), the term of interest is ΔW_y, which can be written using the simplifications as

$$\Delta W_y = \frac{A_2 D_2}{C_2} - \frac{A_1 D_1}{C_1} \qquad (25)$$

Re-arranging Equation (25) gives

$$\frac{1}{C_1 C_2}\left(C_1 A_2 D_2 - C_2 A_1 D_1\right) \qquad (26)$$

Before we can proceed with the final result, it will be useful to establish some results that fall out from both the given assumptions and Theorem 1, which deals with the routine-operating case.

Lemma 1: Based on the cancellations previously mentioned

$$C_2F_2Y = \lambda \bar{C}_1\bar{\bar{F}}_1\bar{Y} \tag{27}$$

where λ is the polynomial of order $n_\lambda = n_H + n_T$.

Proof: Taking into consideration the cancellations previously mentioned and substituting into the second component of Equation (9) gives

$$\frac{\left(\bar{M}_1\bar{N} + \bar{P}_1\bar{O}\right)\bar{D}_1}{\bar{C}_1\bar{\bar{F}}_1\bar{Y}} = \frac{\left(M_2N + P_2O\right)D_2}{C_2F_2Y} \tag{28}$$

Therefore, since the terms on the left are coprime by the 2 assumptions, based on the theory of Diophantine equations, the general solution can be written as

$$\begin{aligned} C_2F_2Y &= \lambda\bar{C}_1\bar{\bar{F}}_1\bar{Y} \\ \left(M_1N + P_1O\right)D_2 &= \lambda\left(\bar{M}_1\bar{N} + \bar{P}_1\bar{O}\right)\bar{D}_1 \end{aligned} \tag{29}$$

where λ is the polynomial of order $n_\lambda = n_H + n_T$. *Q.E.D.*

Lemma 2: It can be shown that

$$\left(T\bar{M}_2D_2 - \lambda\bar{M}_1\bar{D}_1\right) = \gamma\bar{O} \tag{30}$$

where γ is a polynomial whose leading term is zero with order $\min(n_M - n_O, n_P - n_N) + n_D + n_T$.

Proof: Rewriting the second component of Equation (29) obtained in solving Lemma 1 to isolate N, O, $\bar{N}$, and $\bar{O}$, we have that

$$\left(T\bar{M}_2D_2 - \lambda\bar{M}_1\bar{D}_1\right)\bar{N} + \left(T\bar{P}_2D_2 - \lambda\bar{P}_1\bar{D}_1\right)\bar{O} = 0 \tag{31}$$

Since $\bar{O}$ and $\bar{N}$ are assumed coprime, Equation (31) can be written as

$$\begin{cases} \left(T\bar{M}_2D_2 - \lambda\bar{M}_1\bar{D}_1\right) = \gamma\bar{O} \\ \left(T\bar{P}_2D_2 - \lambda\bar{P}_1\bar{D}_1\right) = -\gamma\bar{N} \end{cases} \tag{32}$$

where γ is a polynomial whose leading term is zero, since all the polynomials on the left hand side have a leading term of 1, which upon subtraction will be zero. Its order is given as $\min(n_M - n_O, n_P - n_N) + n_D + n_T$. *Q.E.D.*

Using Lemma 1, Equation (26) can be rewritten as

$$\frac{1}{F_2YC_1C_2}\left(C_1F_2YA_2D_2 - \lambda\bar{C}_1\bar{\bar{F}}_1\bar{Y}A_1D_1\right) \tag{33}$$

Since $C_1\bar{F}_1Y = T\bar{C}_1\bar{\bar{F}}_1\bar{Y}$, Equation (33) can be rewritten as

$$\frac{1}{TF_2YC_1C_2}\left(C_1YTF_2A_2D_2 - \lambda C_1Y\bar{F}_1A_1D_1\right) \tag{34}$$

Simplifying Equation (34) by cancelling the common terms C_1 and Y gives

$$\frac{1}{TF_2C_2}\left(TF_2A_2D_2 - \lambda\bar{F}_1A_1D_1\right) \tag{35}$$

Finally, since $D_1 = H\bar{D}_1$ and $F_1 = H\bar{F}_1$, Equation (35) can be rewritten as

$$\frac{1}{TF_2C_2}\left(TF_2A_2D_2 - \lambda F_1A_1\bar{D}_1\right) \tag{36}$$

Now, since we are interested only in the orders of terms and their effect on identifiability, it can be noted that by construction the order of M and FA are the same. Therefore, it is possible to exchange these terms in Equation (36) to obtain

$$\frac{1}{TF_2C_2}\left(TM_2D_2 - \lambda M_1\bar{D}_1\right) \tag{37}$$

Let $\bar{T}$ be the defined as those common terms of T that are found in M to give $\bar{M}$. Thus, Equation (37) can be written as

$$\frac{\bar{T}}{TF_2C_2}\left(T\bar{M}_2D_2 - \lambda\bar{M}_1\bar{D}_1\right) \tag{38}$$

Using Lemma 2, Equation (38) can be written as

$$\frac{\bar{T}\gamma\bar{O}}{TF_2C_2} \tag{39}$$

Therefore, using Equation (39) the first component of Equation (9) can be written as

$$E\left(\left(\frac{\bar{T}\gamma\bar{O}}{TF_2C_2}r_t\right)^2\right) = 0 \tag{40}$$

The only way to guarantee that Equation (40) will be satisfied is to have that the reference signal have a persistent excitation greater than or equal to the order of γ, that is,

$$n_r \geq \min(n_M - n_O, n_P - n_N) + n_D + n_T \tag{41}$$

provided that the points of the support of r_t do not coincide with the zeroes of X on the unit circle.

Based on the definitions of the original polynomials in terms of the new polynomials, Equation (41) can be rewritten in terms of the original polynomial orders to give

$$\begin{aligned} n_r \geq\ & n_D + \min\left(n_C + n_F + n_Y, n_A + n_F + n_Y, n_B + n_X\right) \\ & + \min\left(n_F + n_A - n_X - n_k, n_B - n_Y\right) \end{aligned} \tag{42}$$

which is identical with Equation (12). *Q.E.D.*