Multivariable System Identification for Integral Controllability – Computational Issues

Mark L. Darby, Michael Nikolaou

Chemical & Biomolecular Engineering Department, University of Houston, Houston, TX 77204-4004 USA (darbymark@sbcglobal.net, nikolaou@uh.edu)

Abstract: A process model satisfies the integral controllability (IC) condition if the model can be used in a model-based controller that can be arbitrarily detuned without jeopardizing closed-loop stability. For decoupling multivariable control this requirement is equivalent to the inequality $\mathrm{Re}\left[\lambda\left(\mathbf{G}\hat{\mathbf{G}}^{-1}\right)\right] > 0$ for the actual and estimated process steady-state gain matrices $\mathbf{G}$ and $\hat{\mathbf{G}}$. This necessitates experiments for identification of $\hat{\mathbf{G}}$ that satisfies the IC inequality. In this work we explore, via computer simulations, computational issues related to the design of such experiments for an FCC process. The proposed approach is based on a general mathematical optimization framework we presented in prior work.

Keywords: Identification, Integral controllability, multivariable systems

1. INTRODUCTION

A process model satisfies the integral controllability (IC) condition if the model can be used in a model-based controller that can be arbitrarily detuned without jeopardizing closed-loop stability. For decoupling multivariable control this requirement is equivalent to the inequality

$$\mathrm{Re}\left[\lambda\left(\mathbf{G}\hat{\mathbf{G}}^{-1}\right)\right] > 0 \tag{1}$$

for the actual and estimated process steady-state gain matrices $\mathbf{G}$ and $\hat{\mathbf{G}}$ (Garcia and Morari 1985). The problem is acute for ill-conditioned processes. This necessitates experiments for identification of a model $\hat{\mathbf{G}}$ that satisfies the IC inequality. "The main weakness of the eigenvalue conditions [eqn. (1)] is that they consist of a coupling between the plant model and the true plant which is highly cumbersome for use in robust control analysis and design." (Featherstone and Braatz 1998b). A number of attempts have been made to address this weakness. Featherstone and Braatz (1998a) showed that for processes with constant rotation matrices in the singular-value decomposition (svd) of their transfer matrix the problem reduces to D-optimal design of experiments. Using insightful geometric reasoning to ensure IC for general linear 2×2 systems, Koung and MacGregor (1993) introduced experiment designs in terms of rotated PRBS input vectors, with power of each component of the rotated input vector reciprocally proportional to the corresponding singular value of $\hat{\mathbf{G}}$. Koung & MacGregor (1994) heuristically extended these design rules to $n\times n$ multivariable systems. The same rules were also used by Bruwer & MacGregor (2006) for the design of identification experiments subject to input and output bounds in the time domain. Darby and Nikolaou (2008) showed that the design rules proposed by Koung and MacGregor (1993; 1994) accept the same deep theoretical justification for both 2×2 and $n\times n$ systems in a number of cases. However, Darby and Nikolaou (2008) also showed that these design rules are not optimal for a number of typical cases, such as when outputs and/or inputs are constrained or when input rather than output variance alone must be maintained at a minimum. Furthermore, the same authors provided rigorous design rules for optimal inputs in a number of such cases. These design rules from solution of corresponding optimization problems. The purpose of this article is to explore the nature of the optimal input designs produced by the mathematical framework introduced by Darby and Nikolaou (2008) when applied to a realistic system, such as a 5×5 fluid catalytic cracking (FCC) unit.

2. BACKGROUND: EXPERIMENT DESIGN FOR IC

Consider a stable, linear, time-invariant, multivariable system with steady-state input-output relationship

$$\mathbf{y} = \mathbf{G}\mathbf{m} \tag{2}$$

where $\mathbf{y},\mathbf{m} \in \mathbb{R}^n$, $\mathbf{G}$ and $\hat{\mathbf{G}} \triangleq [\hat{\mathbf{g}}_1,\ldots,\hat{\mathbf{g}}_n]^T \in \mathbb{R}^{n\times n}$. Because the IC condition, eqn. (1), involves the real process $\mathbf{G}$ and identified model $\hat{\mathbf{G}}$, it cannot directly guide input design for an $n\times n$ system. The following results (Darby and Nikolaou 2008) avoid that difficulty and can be used directly to design experiments pursuing IC.

Theorem 1. Experiment design for IC. Let the model uncertainty matrix $\mathbf{D}=\mathbf{G}-\hat{\mathbf{G}}\in\mathbb{R}^{n\times n}$ *belong to the ellipsoidal uncertainty set*

$$D \triangleq \left\{[\mathbf{d}_1 \ldots \mathbf{d}_n]^T \in \mathbb{R}^{n\times n} \middle| \mathbf{d}_k^T \mathbf{M}^T \mathbf{M} \mathbf{d}_k \le c^2,\ 1\le k\le n\right\}. \tag{3}$$

Then, an experiment design guarantees IC if the resulting information matrix $\mathbf{M}^T\mathbf{M}$ *and identified model* $\hat{\mathbf{G}}$ *satisfy the inequality*

$$\sum_{k=1}^{n} a_k \sqrt{\hat{\mathbf{v}}_k^T \left(\mathbf{M}^T\mathbf{M}\right)^{-1} \hat{\mathbf{v}}_k} < 1. \tag{4}$$

where

$$a_k \triangleq c\|\hat{\mathbf{u}}_k\|_1 / \hat{\sigma}_k, \quad k=1,\ldots,n \tag{5}$$

and $\hat{\sigma}_1 \ge \cdots \ge \hat{\sigma}_n$, $\hat{\mathbf{u}}_k$, $\hat{\mathbf{v}}_k$ *are defined through the svd*

$$\hat{\mathbf{G}} = \hat{\mathbf{U}}\hat{\mathbf{\Sigma}}\hat{\mathbf{V}}^T = \sum_{k=1}^{n} \hat{\sigma}_k \hat{\mathbf{u}}_k \hat{\mathbf{v}}_k^T. \tag{6}$$

Eqn. (4) clearly suggests that IC can be satisfied if the information matrix $\mathbf{M}^T\mathbf{M}$ is "large enough". Given bounds on the input vector $\mathbf{m}$, a large enough $\mathbf{M}^T\mathbf{M}$ can be achieved if (a) the identification experiment is run long enough, or (b) $\mathbf{m}$ is shaped appropriately. While the first alternative is straightforward, it is far less desirable than the second one. Therefore, the essence of experiment design for IC is how to *shape process inputs that satisfy eqn. (4) subject to relevant constraints*. Darby and Nikolaou (2008) showed that numerical or analytical solutions can be developed for a number of cases. While for some of these cases the resulting designs are similar to designs that have appeared in literature, for others the resulting designs are entirely different.

2.1. Analytical solutions

Suppose that the quantity $\sum_{k=1}^{n} a_k [\hat{\mathbf{v}}_k^T (\mathbf{M}^T\mathbf{M})^{-1} \hat{\mathbf{v}}_k]^{1/2}$ in eqn. (4) is to be minimized with respect to the zero-mean random input $\mathbf{m}$, subject to the total weighted variance inequality

$$x\,\mathrm{var}(\mathbf{y}) + (1-x)\,\mathrm{var}(\mathbf{m}) \le W^2 \tag{7}$$

where $0\le x\le 1$ and $0<a_1<\ldots<a_n$. Then it can be shown that the optimal input vector $\mathbf{m}$ is

$$\mathbf{m} = \hat{\mathbf{V}}\boldsymbol{\xi}, \tag{8}$$

where $\boldsymbol{\xi}$ is a zero-mean multivariable PRBS with

$$E\left[\xi_k^2\right] = \left(\frac{a_k}{b_k^2}\right)^{2/3} \frac{W^2}{\left[\sum_{j=1}^{n}\left(a_j b_j\right)^{2/3}\right]^2}, \quad k=1,\ldots,n \tag{9}$$

$$b_k^2 \triangleq x\sigma_k^2 + 1 - x, \quad k=1,\ldots,n. \tag{10}$$

and $b_1 \ge \ldots \ge b_n > 0$. Reversing the role of the above objectives and constraints, the minimum of the cost function $x\,\mathrm{var}(\mathbf{y}) + (1-x)\,\mathrm{var}(\mathbf{m})$ subject to eqn. (4) can be shown to be attained at an optimal $\mathbf{m}$ satisfying eqn. (8) with

$$E\left[\xi_k^2\right] > \left(\frac{a_k}{b_k^2}\right)^{2/3} \left[\sum_{j=1}^{n}\left(a_j b_j\right)^{2/3}\right]^2, \quad k=1,\ldots,n. \tag{11}$$

Note that for $x=1$ (all cost on output variance, as is desirable in early stages of an identification experiment) both eqns. (9) and (11) result in the well known design rule

$$r_{kj} \triangleq \sqrt{\frac{E\left[\xi_k^2\right]}{E\left[\xi_j^2\right]}} = s_{kj}\frac{\hat{\sigma}_j}{\hat{\sigma}_k} \approx \frac{\hat{\sigma}_j}{\hat{\sigma}_k}. \tag{12}$$

where $s_{kj} \triangleq (\|\hat{\mathbf{u}}_k\|_1 / \|\hat{\mathbf{u}}_j\|_1)^{1/3} \approx 1$ for most cases of practical interest, and $1/n^{1/6} \le s_{kj} \le n^{1/6}$ when $\hat{\mathbf{u}}_k$, $\hat{\mathbf{u}}_j$ are any orthonormal vectors in $\mathbb{R}^n$. However, for $x=0$ (all cost on input variance) we get the new input design

$$r_{kj} \triangleq \sqrt{\frac{E\left[\xi_k^2\right]}{E\left[\xi_j^2\right]}} = s_{kj}\left(\frac{\hat{\sigma}_j}{\hat{\sigma}_k}\right)^{1/3} \approx \left(\frac{\hat{\sigma}_j}{\hat{\sigma}_k}\right)^{1/3}. \tag{13}$$

The above design would keep inputs small to avoid inadvertent loss of IC by failure to excite the process by inputs along directions corresponding to small singular values.

Finally, a D-optimal design subject to eqns. (4) and (7) can be shown to be attained, if feasible, at an optimal $\mathbf{m}$ as in eqn. (8) with

$$E\left[\xi_k^2\right] = W^2 / \left(n b_k^2\right), \quad k=1,\ldots,n, \tag{14}$$

if eqn. (4) is satisfied by the above ξ_k, or each $E\left[\xi_k^2\right]$ equal to the unique positive solution of the equation

$$\frac{\rho}{t-1}\left(\frac{a_k}{\sqrt{(t-1)E\left[\xi_k^2\right]}} - \frac{1}{n}\right) = b_k^2 E\left[\xi_k^2\right] - \frac{W^2}{n} \tag{15}$$

for a value of $\rho>0$ such that eqn. (4) is satisfied. It has been shown that eqn. (15) guarantees that $E\left[\xi_{k+1}^2\right] > E\left[\xi_k^2\right]$ and that D-optimality is compatible with IC by Cauchy's inequality (Darby and Nikolaou 2008). Note that as $t\to\infty$ eqns. (14) and (15) coincide asymptotically.

2.2. Numerical solutions

The preceding section 2.1 summarized analytical solutions for simple cases, offering insight into the nature of corresponding solutions. However, in many practical situations individual constraints on m_i and y_i may be present, such as

$$\sum_{\tau=1}^{t} {m_{i,\tau}}^2 = [\mathbf{M}^T\mathbf{M}]_{ii} \le (t-1)M_i^2 \tag{16}$$

$$\sum_{\tau=1}^{t} y_{i,\tau}{}^2 = [\mathbf{Y}^T\mathbf{Y}]_{ii} = [\hat{\mathbf{G}}\mathbf{M}^T\mathbf{M}\hat{\mathbf{G}}^T]_{ii} \le (t-1)Y_i^2 \tag{17}$$

corresponding to bounds on the variance of individual inputs m_k or outputs y_k. In such cases a numerical solution is required. To obtain a numerical solution, assume a zero-mean input vector $\mathbf{m}$, approximate the information matrix as $\mathbf{M}^T\mathbf{M} \approx (t-1)\mathbf{C}_m$, parametrize the input covariance matrix $\mathbf{C}_m$ in terms of the triangular matrix $\mathbf{Q}$ through the Cholesky factorization $\mathbf{C}_m = \mathbf{Q}\mathbf{Q}^T$, and substitute $\mathbf{M}^T\mathbf{M}$ into eqn. (4), to get

$$\frac{c}{\sqrt{t-1}} \underbrace{\sum_{i=1}^{n} \frac{\|\hat{\mathbf{u}}_k\|_1}{\hat{\sigma}_k} \sqrt{\hat{\mathbf{v}}_k^T \left(\mathbf{Q}\mathbf{Q}^T\right)^{-1} \hat{\mathbf{v}}_k}}_{\beta} \triangleq \frac{c}{\sqrt{t-1}}\beta < 1. \tag{18}$$

(Other parametrizations of a symmetric matrix in terms of corresponding basis matrices could be used. This is subject of ongoing investigation.) We can then design experiments for IC using eqn. (18) as a constraint or by minimizing β with respect to $\mathbf{Q}$ subject to input and output constraints such as in eqns. (16) and (17). Then, the corresponding optimal input $\mathbf{m}$ is

$$\mathbf{m} = \mathbf{Q}_{\text{opt}}\mathbf{z} \tag{19}$$

where $\mathbf{z}$ is a zero-mean PRBS with $\text{cov}(\mathbf{z}) = \mathbf{I}$. Even though β is not convex, resulting design problems are not prohibitively large for realistic systems, as demonstrated in section 4. It should be stressed that minimizing β may result to neither uncorrelated rotated inputs ξ, nor magnitudes of rotated input components reciprocally proportional to corresponding singular values of the steady-state gain matrix, eqn. (12). In fact, the advantage of the above numerical formulation is that no such underlying assumptions on the nature of optimal inputs are necessary. Rather, the numerical optimization determines the nature of optimal input designs.

3. SUMMARY OF PROPOSED APPROACH

a. Establish constraints commensurate with time t available for identification experiments.

b. Obtain preliminary estimates of $\hat{\mathbf{G}}$ and c^2.

c. Compute the svd of $\hat{\mathbf{G}}$ to get $\hat{\mathbf{U}}$, $\hat{\mathbf{\Sigma}}$, $\hat{\mathbf{V}}^T$, eqn. (6).

Case I

d. Compute $\lambda_{k,\text{opt}}$ via eqns. (9) or similar

e. Design $\mathbf{m} = \hat{\mathbf{V}}\xi$ with ξ zero-mean PRBS and $\text{cov}(\xi) = \text{diag}(\lambda^2_{\ell_1,\text{opt}},\ldots,\lambda^2_{\ell_n,\text{opt}})/(t-1)$

Case II

d. Compute $\mathbf{Q}_{\text{opt}} = \arg\min \beta$ subject to constraints.

e. Design $\mathbf{m} = \mathbf{Q}_{\text{opt}}\mathbf{z}$ with $\mathbf{z}$ zero-mean PRBS and $\text{cov}(\mathbf{z}) = \mathbf{I}$ (eqn. (19)).

f. Implement $\mathbf{m}$ and collect data, to update $\hat{\mathbf{G}}$ and c^2.

g. If $\hat{\mathbf{G}}$ is adequate, stop. Otherwise go to step c.

4. CASE STUDY

A steady-state gain matrix is obtained from a linear empirical dynamic model of an industrial reactor-regenerator from a FCC unit, identified from plant testing (Harmse 2007). Note that the specific inputs and outputs are not indicated. Scaling is performed according to the inverse of the typical operating ranges of the inputs and outputs. The resulting gain matrix is

$$\mathbf{G} = \begin{bmatrix} 0.3866 & 0.0 & 0.1192 & 0.0 & 0.0630 \\ 0.0 & -0.6935 & 1.5463 & -0.1311 & -0.2462 \\ 0.0 & 0.0 & 0.5225 & -0.1298 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.1058 & 0.0 \\ 0.0 & -0.5803 & -0.3669 & -0.2057 & -0.4435 \end{bmatrix} \tag{20}$$

We tested the designs shown in Table 1. The constraints shown in Table 2 were considered.

Table 1. Summary of experimental designs tested

Design	Objective	Constraints
ICmin	$\min_{\mathbf{Q}} \beta$	$\mathbf{Q}$ triangular
KM	(Koung and MacGregor 1994)	
PRBSmax	$\min_{\mathbf{C}_m} -\log\left(\det \mathbf{C}_m\right)$	$\mathbf{C}_m \triangleq \text{diag}(v_i) \succ \mathbf{0}$

Table 2. Constraints considered for experiment design

Design case	Bounds	
A	$\text{var}(m_i) \le \infty$	$\text{var}(y_1) \le 0.50$
	$i = 1,\ldots,5$	$\text{var}(y_2) \le 0.47$
		$\text{var}(y_3) \le 0.44$
		$\text{var}(y_4) \le 0.41$
		$\text{var}(y_5) \le 0.38$
B	$\text{var}(m_i) \le 0.5$	$\text{var}(y_i) \le 0.5$
	$i = 1,\ldots,5$	

In all simulations, parameter estimation is initiated at time step 5 and is performed at each subsequent time step. The IC condition, eqn. (1), is calculated based on the true gain matrix and the inverse of the gain estimate at each time step. For simulations of case A, independent Gaussian noise of zero-mean and unit variance is added to all outputs.

Realizations of the inputs and outputs for each design are shown in Figure 1. We see that due to the high noise levels, the actual output variances are significantly higher than the optimal output variances (which are based on the model without noise). Note that the relatively low signal-to-noise ratio allows us to observe the evolution of gains over a longer period of time.

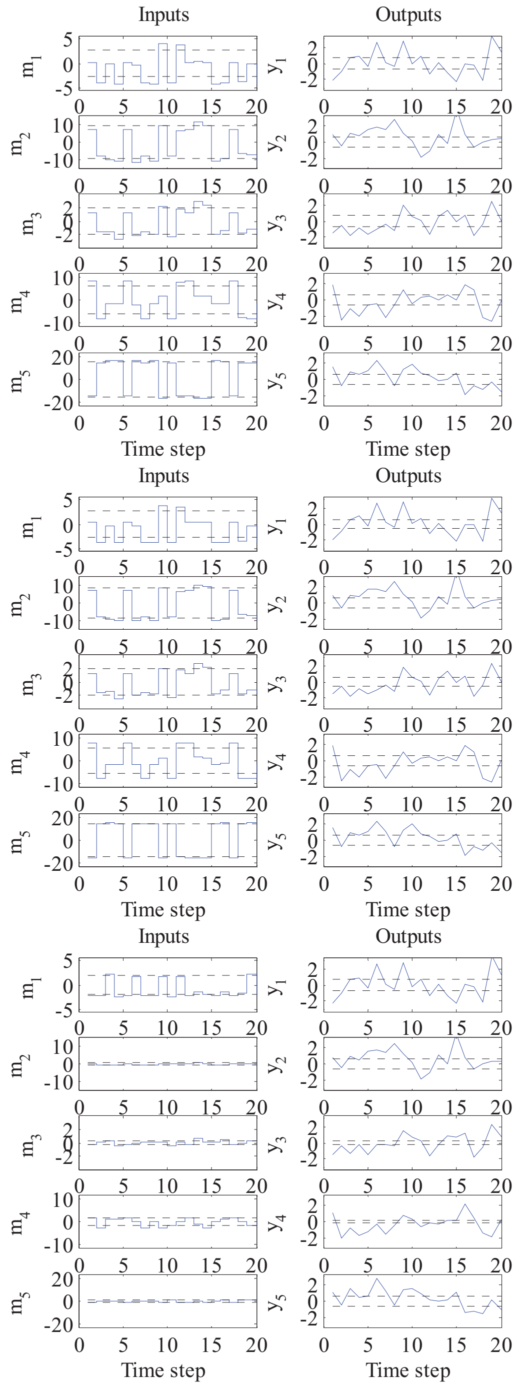

Figure 1 – Example realization of reactor regenerator for the three designs of case A (Table 1). The dotted lines represent $\pm(\mathrm{var}(m_i)^{\mathrm{opt}})^{1/2}$ and $\pm(\mathrm{var}(y_i)^{\mathrm{opt}})^{1/2}$ values.

Trends of the gain errors and an indicator of the IC condition are shown in Figure 2 For this realization, we see that IC is achieved first by ICmin at time step 7, followed by KM at time step 8, and finally PRBSopt at time step 13. The evolution of the gain errors is similar for ICmin and KM, whereas the PRBSopt results show higher gain errors and slower error reduction over time, consistent with the lower value of $\det(\mathrm{cov}(\mathbf{m}))$ for PRBSopt.

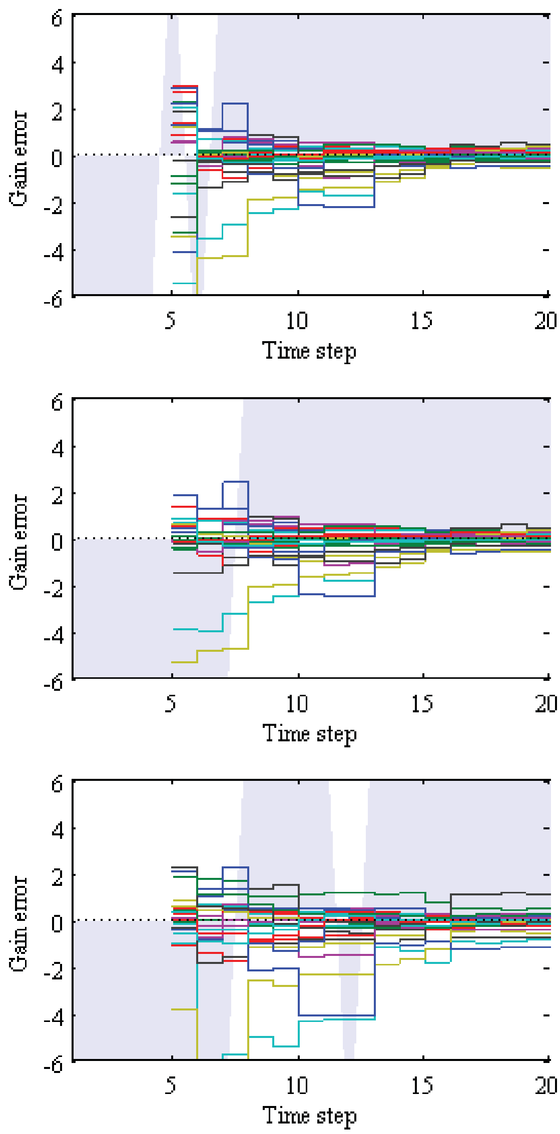

Figure 2 – Gain errors and IC condition for example realization of design case A (Table 1). Satisfaction or violation of the IC condition corresponds to shading above or below 0, respectively.

For the simulations of case B, independent, Gaussian noise of zero mean and variance (0.15^2) is added to all outputs. Realizations of the inputs and outputs for each design are shown in Figure 3. Note that the smaller signals for the KM design, compared to the other two designs, are due to the fact that inputs must satisfy all inequality constraints, as well as equality constraints on rotated input ratios. The latter are clearly not optimal for case B.

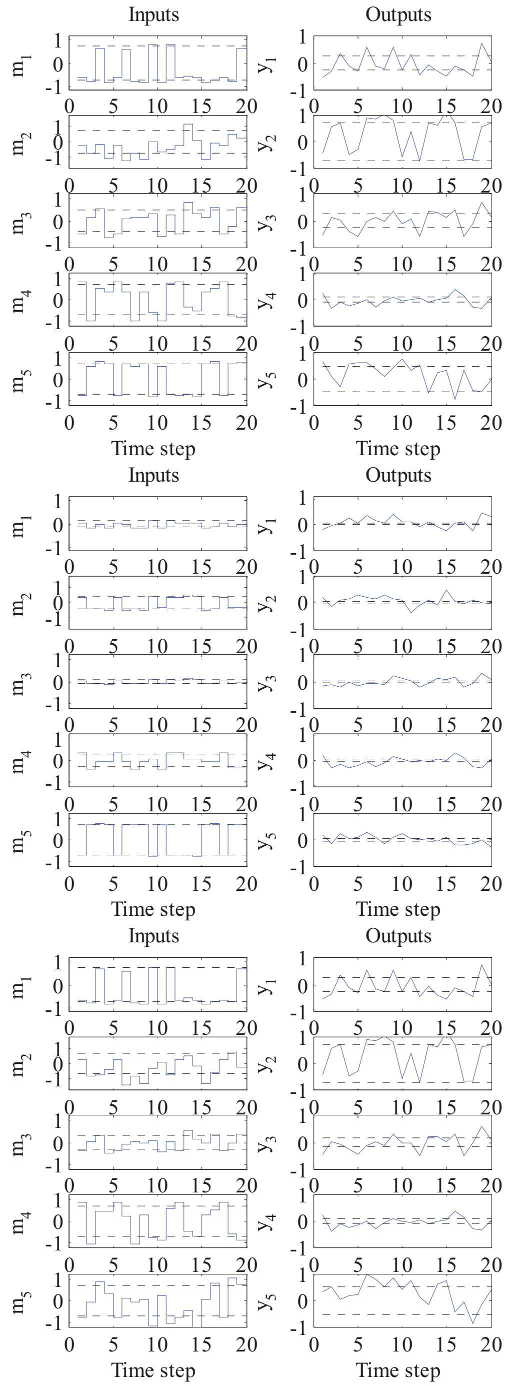

Figure 3 – Example realization of reactor regenerator case B (Table 1). The dotted lines represent $\pm\sqrt{\mathrm{var}(m_i)}^{\mathrm{opt}}$ and $\pm\sqrt{\mathrm{var}(y_i)}^{\mathrm{opt}}$ values.

Trends of the gain errors and an indicator of the IC condition are shown in Figure 4. We see that the parameters estimates for the KM design are significantly inferior to the ICmin and PRBSopt designs, reflecting the much smaller value of $\det(\mathrm{cov}(\mathbf{m}))$ for KM. Further, while IC is achieved by ICmin at time step 6 and by PRBSopt at time step 5, the KM design does not satisfy the IC condition by time step 20 (it is actually achieved at time step 22 – not shown – more then 3 times longer than required for either ICmin or PRBSopt).

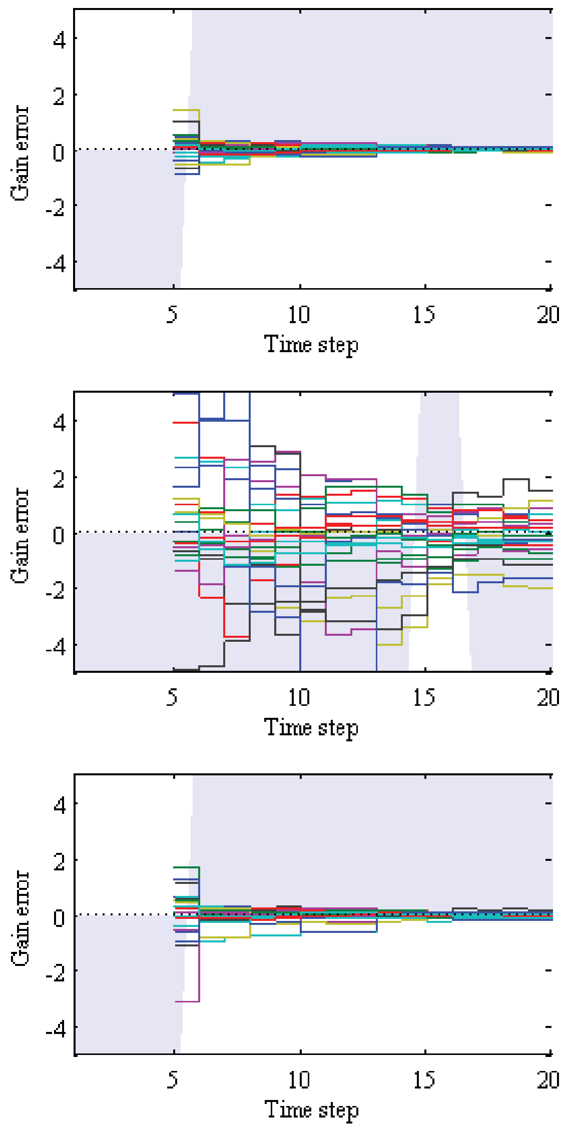

Figure 4 – Gain errors and IC condition for example realization of design case B (Table 1). Satisfaction or violation of the IC condition corresponds to shading above or below 0, respectively.

To develop these designs we used the Matlab routine fmincon with multiple starting points, to reduce the possibility of missing a global optimum for non-convex optimization problems. Run time for all simulations was, on the average, of the order of 0.25 seconds when using the Yalmip interface, and of the order of 0.1 seconds without it.

5. CONCLUSIONS

The purpose of this paper was to explore numerical aspects of a numerical optimization approach proposed in prior work for the design of experiments targeting IC. Given that analytical solutions for this approach are available only for some cases, it is natural to ask how well numerical optimization can work, given that the proposed problems are non-convex. In this work, we develop what appear to be useful designs for a 5×5 multivariable FCC unit. Convergence time appeared not to be an issue. These results suggest that the proposed approach can work reasonably well for problems of that size. Clearly, other optimization methods (either deterministic or probabilistic) can be explored.

Acknowledgments

Partial support from the National Science Foundation (Grant No. CTS-0227232) is gratefully acknowledged.

6. REFERENCES

Bruwer, M. J. and J. F. MacGregor (2006). "Robust multi-variable identification: Optimal experimental design with constraints." Journal Of Process Control **16**(6): 581-600.

Darby, M. L. and M. Nikolaou (2008). "Multivariable system identification for integral controllability." Automatica: submitted.

Featherstone, A. P. and R. D. Braatz (1998a). "Input Design for Large-Scale Sheet and Film Processes." Industrial & Engineering Chemistry Research **37**: 449-454.

Featherstone, A. P. and R. D. Braatz (1998b). "Integrated robust identification and control of large-scale processes." Industrial & Engineering Chemistry Research **37**(1): 97-106.

Garcia, C. E. and M. Morari (1985). "Internal model control: 2. design procedure for multivariate systems." Ind. Eng. Chem. Process Des. Dev. **24**: 472-484.

Harmse, M. (2007). "Personal communication."

Koung, C. W. and J. F. MacGregor (1993). "Design Of Identification Experiments For Robust-Control - A Geometric Approach For Bivariate Processes." Industrial & Engineering Chemistry Research **32**(8): 1658-1666.

Koung, C. W. and J. F. MacGregor (1994). "Identification For Robust Multivariable Control - The Design Of Experiments." Automatica **30**(10): 1541-1554.