To be presented at the 13th IFAC Symposium on System Identification (SYSID 2003)
Rotterdam, The Netherlands, August 27-29, 2003

Session: Identification for Process Control - Input Design

"PLANT-FRIENDLY" SYSTEM IDENTIFICATION: A CHALLENGE FOR THE PROCESS INDUSTRIES

Daniel E. Rivera, [*,1] **Hyunjin Lee,** [*] **Martin W. Braun** [*,2] **and Hans D. Mittelmann** [**]

[*] *Control Systems Engineering Laboratory*
Department of Chemical and Materials Engineering
[**] *Department of Mathematics and Statistics*
Arizona State University, Tempe, Arizona 85287

Abstract: The term "plant-friendly" system identification has been used within the chemical process control research community in reference to the broad-based goal of accomplishing informative identification testing while meeting the demands of industrial practice. While many different identification topics (such as control-relevant identification, closed-loop identification and optimal input design) can be said to contribute to plant-friendliness in identification, the problem has some unique character of its own. This paper describes some of the issues that motivate plant-friendly identification and presents an overview of some approaches that have been proposed in this topic. The problem of *identification test monitoring* is presented as a novel means for accomplishing plant-friendly identification.

Keywords: plant-friendly system identification, chemical process control, identification test monitoring

1. INTRODUCTION

The term "plant-friendly" identification is one that seems to be exclusive to the chemical process control community. The exact origin of the term is not clear, but it first appears mentioned in print in the paper by Pearson *et al.* (1993). However, articulation of issues related to plant-friendly considerations can be found in literature based on the experience of industrial practitioners (Rivera *et al.*, 1992) or influenced by the demands of practice (Godfrey, 1993).

The concept of plant-friendliness in system identification for the process industries stems from the fundamental need for informative experiments despite practical requirements to the contrary. Broadly speaking, a plant-friendly identification test will produce data leading to a suitable model within an acceptable time period, while keeping the variation in both input and output signals within user-defined constraints. Examples of plant-friendly constraints (and their impact on process operations) include:

- keeping output deviations low to minimize variability in product quality,
- implementing a signal of sufficiently short duration to minimize the amount of off-spec product and reduce engineering time associated with an identification test,
- keeping move sizes small to satisfy actuator constraints and minimize "wear and tear" on process equipment.

These practical considerations are often in conflict with theoretical requirements (e.g., asymptotic operation, persistence of excitation, etc.) that demand long identification tests under high signal-to-noise ratios. As a result, plant-friendliness often involves a compromise between the demands of theory (which are for the most part "plant-hostile") and the demands of

---

[1] to whom all correspondence should be addressed; phone:(480) 965-9476, email:daniel.rivera@asu.edu
[2] Currently with Texas Instruments Inc., 13570 N. Central Expressway, MS 3701, Dallas, TX, 75265

plant engineers (who would prefer no changes in the process as a result of identification testing).

The main objective of this paper is to create increased awareness of the plant-friendly identification problem within the larger system identification community. The paper is organized as follows: Section 2 discusses the motivation for plant-friendliness, using industrially-relevant examples to justify why the problem goes beyond classical input signal design approaches. Section 3 presents two plant-friendliness measures and surveys some existing approaches to the problem. In Section 4 the problem of *identification test monitoring* is proposed as a novel means for accomplishing plant-friendly identification; this discussion is supported with some analysis and an example. The paper concludes with a summary section. Ultimately, it is our desire that this paper will stimulate multidisciplinary research efforts in this important problem to the practicing process control community.

2. MOTIVATION FOR PLANT-FRIENDLY IDENTIFICATION

System identification in practice is an iterative procedure. The lack of *a priori* information regarding the plant model will require that initially each step be examined in a superficial manner. After each stage, the user must discern if the previous stages were properly accomplished; if this is not the case, the stage(s) must be redone until a "satisfactory" model is obtained. A satisfactory model is one that meets the requirements of the intended application (e.g., simulation, prediction, or control).

The quality of the data generated from the experimental design stage is critical to the success of the comprehensive system identification and subsequent control design procedures. In the chemical process industries, identification testing is by necessity conducted while the plant is in normal operation, and as such represents one of the most expensive and time consuming steps in the application of advanced control in the process industries. There are many industrial examples of the significant disruption and cost that identification testing represents to plant operations; here we just discuss a few. As a research engineer at Shell Development Company in the late 1980's, the first author was involved in implementing individual identification tests lasting five days or more. More recently, while consulting for a major chemical company, the first author was presented with the identification testing guidelines of a major control software vendor. To identify an ethylene furnace characterized by 17 independent variables and a six hour settling time, the vendor's guideline suggested nearly a month (25.5 days) of identification tests. Mathur and Conroy (2002) cite an example where the total costs of step testing (taking into account reductions in throughput, off-spec product, and engineering time) were estimated at $270,000. It comes as no surprise that model development has been reported to account for 75% of the costs associated with an advanced control project (Hussain, 1999).

In addition, "fanout" issues are a problem - there are a large percentage of loops in industrial practice for which significant performance improvements would be possible as a result of system identification (Ogawa and Douke, 2002). This task, however, has to be done in an acceptable manner to operations. This represents yet another significant incentive for accomplishing plant-friendly, informative identification testing.

One could expect that for a mature area such as system identification, existing literature would address these problems. For example, the identification literature abounds with optimal input design theorems which, one would expect, should be well suited to address the plant-friendly identification problem. However, the majority of these "optimal" solutions are not formulated in a manner meaningful to the chemical process engineer. For starters, most optimal input design results depend on knowledge of the true plant and noise models (Ljung, 1999), which is information that is largely unknown to the user (at least at the start of identification testing). Furthermore, process control engineers tend to think more in terms of maintaining high/low limits, move size constraints, and test duration during identification testing and less in terms of the norm criteria that are typically used in the classical optimal experimental design formulations.

Control-relevant system identification has been an important research topic in the community since the late 1980's. However, control-relevance in and of itself does not automatically imply plant-friendliness. Taking control performance requirements into consideration during identification testing will often lead to more focused input signal with a narrower bandwidth and emphasizing the important regimes of time and frequency needed for the intended application (Rivera *et al.*, 1992; Cooley *et al.*, 1998). However, control-relevant input designs require additional scrutiny (principally in the form of constraint enforcement) to insure that they will promote plant-friendliness. Similar statements can be made for closed-loop identification and its role in plant-friendly identification.

Ultimately, the main question in plant-friendly identification (and which distinguishes it from other problems in the field) is this: how does one build process knowledge relevant to system identification in a systematic and nearly automatic way, with little user intervention and without demanding significant computational time and effort? If adequate process knowledge is absent at the start of experimental testing, how can this knowledge be systematically acquired in the course of identification testing, for the purposes of improving the identification test?

3. SURVEY OF PLANT-FRIENDLY IDENTIFICATION APPROACHES

Current approaches to plant-friendly identification range from the use of simple measures to sophisticated constrained formulations. Doyle *et al.* (1999) first proposed a *friendliness index* f for an arbitrary input sequence $u(k)$, $k = 1 \cdots N$. The measure is defined as a percentage according to

$$f = 100 \times \left(1 - \frac{n_T}{N-1}\right) \qquad (1)$$

where N is the sequence length and n_T constitutes the number of transitions (i.e., situations where $u(k) \neq u(k-1)$) in the input signal. This measure is also examined in Rengasamy *et al.* (2000) and Parker *et al.* (2001). According to this measure, a constant sequence is "100% plant-friendly", while any sequence that changes value at every instant is "0% plant friendly." A stochastic interpretation and its use in the design of input signals for identifying Volterra series models are presented in the aforementioned papers.

Another measure that has been proposed to determine plant-friendliness is the Crest Factor (CF) (Guillaume *et al.*, 1991). The crest factor is defined as the ratio of the ℓ_∞ norm and the ℓ_2-norm of a signal

$$CF(u) = \frac{\ell_\infty(u)}{\ell_2(u)} \qquad (2)$$

and provides a measure of how well distributed the signal values are over the input span. A low crest factor indicates that most of the elements in the input sequence are distributed near the minimum and maximum values of the sequence. Reducing the crest factor of the input or output signals (or both) can significantly contribute to plant-friendliness during experimental testing. For example, if two signals with equivalent power spectral densities are to be evaluated for identification purposes, the one with lower crest factor is preferred because it will deliver the same power over a lower overall span. An alternative measure to crest factor is the Performance Index for Perturbation Signals (PIPS) (Godfrey *et al.*, 1999).

$$PIPS(\%) = 200 \frac{\sqrt{u_{rms}^2 - u_{mean}^2}}{u_{max} - u_{min}} \qquad (3)$$

The PIPS measure ranges between 0 and 100% (compared to 1 versus ∞ for crest factor), which gives it an intuitive, practical appeal.

In Braun *et al.* (2002), crest factor minimization is used as the basis for designing multisine inputs meaningful for plant-friendly identification of both linear and nonlinear systems. Constrained extensions have been presented for both the single input (Rivera *et al.*, 2002) and multivariable cases (Lee *et al.*, 2003); these formulations allow specifying both frequency domain signal requirements (meaningful from a system theoretic perspective) and time-domain constraints (meaningful from the user's standpoint).

We previously noted that control-relevant approaches to input signal design can be formulated to promote plant-friendliness; examples of this approach include the work of J.H. Lee and co-workers (Chikkula and Lee, 1997; Cooley *et al.*, 1998; Cooley and Lee, 2001). Actuator constraints are recognized in these problem formulations. Other control-relevant approaches taking into account constraints include the work of Li and Georgakis (2002) who focus on highly interactive systems. The recent work Narasimhan *et al.* (2003) considers a multi-objective framework in which the tradeoffs associated with maximizing input-output friendliness, constraints, and other criteria are explored in the context of a mixed-integer nonlinear programming problem.

4. IDENTIFICATION TEST MONITORING

Identification test monitoring is proposed as a novel approach to the plant-friendly identification problem. Factors that influence this approach include the following:

(1) *a priori* knowledge of the system dynamics available to the engineer is often sketchy at the start of identification testing. To maintain flexibility in the design the user is forced to make some generous assumptions. A framework is needed that allows the user to systematically acquire process knowledge which in turn is used to refine the experimental test, without resorting to a completely new experimental testing procedure.

(2) The use of periodic, deterministic inputs such as multisine or pseudo-random signals provides distinct advantages. The timespan defined by one period of a periodic signal provides a natural window or examination point for analysis of the data and signals. As a result, monitoring procedures can be established that work in "quasi-real" time, where the data resulting from prior cycles of a multisine signal are analyzed while the current cycle is being implemented.

The iterative evaluation and refinement of identification signals implied by the identification test monitoring problem contrasts much of current industrial practice. The tendency in practice is to collect and analyze data in one batch, and determine (after pursuing the complete identification cycle) if the data is adequate. This results in costly re-testing and requires significant user intervention and effort. In contrast, a systematic monitoring approach based on periodic inputs allows the opportunity for users to improve tests without having to perform a full comprehensive analysis.

4.1 *Towards Identification Test Monitoring: Integrating Identification and Robust Loopshaping*

In addition to system identification, it is possible to draw from the fields of signal processing, robust control, and optimization to synergistically create novel frameworks for identification test monitoring. Data generated from periodic inputs can be used to calculate an empirical transfer function estimate (ETFE) (Ljung, 1999). According to Bayard (1993), a confidence region in the Nyquist plane for the ETFE $p^*(\omega_i)$ is a perfect circle centered at $p^*(\omega_i)$ of radius ϵ_i where

$$\epsilon_i^2 = \frac{\hat{\sigma}^2|\overline{W}(e^{-j\omega_i T})|^2 F_{1-\kappa}(2,\nu)}{\Phi_u(\omega_i)m} \quad (4)$$

$\overline{W}(z)$ is the estimated noise model while $\hat{\sigma}^2$ is the estimated variance from the residual output spectrum. m corresponds to the number of cycles of the periodic input, while $F_{1-\kappa}(2,\nu)$ is the 2-way Fisher statistic computed for a specified statistical confidence of $(1-\kappa)\times 100\%$. Noting that $F_{1-\kappa}(2,\nu)$ is bounded as ν becomes large (e.g., $F_{1-\kappa}(2,\nu) \leq 3$ for $1-\kappa = .95$ and $\nu > 120$), it becomes clear that the uncertainty region increases with the noise-to-signal ratio $\hat{\sigma}^2|\overline{W}|/\Phi_u$ and decreases as the number of input signal cycles m increases. Thus (4) provides significant insight regarding the important practical issues of signal magnitude and test length in system identification. Noise in the data set can be overcome by either increasing signal power or lengthening the test duration. The decision to follow one approach over the other is dependent upon the circumstances being faced during identification testing, such as operational restrictions, and so forth. These are the types of tradeoffs that the identification test monitoring problem seeks to address.

The confidence regions defined by (4) can be expressed as norm-bounded multiplicative uncertainty

$$\begin{aligned} |(p(e^{j\omega_i T}) - p^*(\omega_i))p^{*-1}(\omega_i)| &\leq \bar{\ell}_m(\omega_i) \quad (5) \\ &= \epsilon_i/|p^*(\omega_i)| \end{aligned}$$

which in turn can be used to assess robust performance, such as the μ analysis measure

$$\mu^* = \sup_{\omega} |\eta^*(e^{j\omega T})\bar{\ell}_m(\omega)| + |\epsilon^*(e^{j\omega T})w_P(j\omega)| \quad (6)$$

w_P weights the sensitivity function $\epsilon = (1+pc)^{-1}$; $\eta = pc(1+pc)^{-1}$ is the complementary sensitivity function. η^* and ϵ^* are the frequency responses of the closed-loop transfer functions based on the estimated frequency response p^*. Whenever $\mu^* < 1$, the following condition is satisfied for the closed-loop system

$$|\epsilon| \leq 1/|w_P| \quad \forall p \in \bar{\ell}_m \quad 0 \leq \omega \leq \pi/T$$

The paper by Braatz *et al.* (1991) provides a procedure based on robust loopshaping for determining necessary and sufficient bounds on the nominal closed-loop transfer functions from knowledge of the process uncertainty and performance specifications on the closed-loop system. For a SISO system, sufficient bounds on η^* and ϵ^* are

$$|\eta^*| < \frac{1-|w_P|}{|w_P|+\bar{\ell}_m} \quad |\epsilon^*| < \frac{1-\bar{\ell}_m}{|w_P|+\bar{\ell}_m} \quad (7)$$

Additional bounds (namely necessary upper and lower bounds on $|\eta^*|$ and $|\epsilon^*|$) can be found in Braatz *et al.* (1991). Using this analysis, one realizes that the robust loopshaping bounds on η^* and ϵ^* can be computed in real-time, *during identification testing*. The decision to halt or modify an identification test can be determined on the basis of how these bounds evolve with increasing number of cycles of the input m.

4.2 *An Identification Test Monitoring Example*

The "thought process" that may be involved in the case of identification test monitoring of a SISO system is presented in this section using a simulated representative scenario. This is depicted in Figure 1 with relevant statistics summarized in Table 1. The simulation considers a first-order system per the transfer function

$$p(s) = \frac{1}{10s+1}$$

the parameters for this system are assumed to be vaguely known *a priori*. We further assume that the plant is facing significant disturbances, but plant operating restrictions dictate that the test be carried out under low Signal-to-Noise Ratios (SNRs). Multisine input signals relying on the constrained minimum crest factor design procedure per Rivera *et al.* (2002) and Lee *et al.* (2003) are used to construct this identification test monitoring scheme. The initial signal (for Stage 1) is designed per the guidelines in (Rivera *et al.*, 2002; Lee *et al.*, 2003) with $\alpha_s = 2, \beta_s = 3, \tau_{dom}^L = 5$ min, $\tau_{dom}^H = 20$ min. The result is a signal of very wide bandwidth, implemented using a low amplitude span (± 0.75) to avoid undesirable levels of off-spec product. This signal is represented by Stage 1 in Table 1 and Figure 1. Analysis of the Stage 1 data results in a preliminary model that can then serve as a basis for the design of the signal in Stage 2. This preliminary model will most likely display high variance in the parameters, as a result of the low signal-to-noise ratio of the data. Nonetheless the information contained in this model is useful for determining the design parameters of Stage 2. In this simulated scenario the model time constant range is narrowed to between 8 and 12 minutes; the resulting use of the guidelines produces a signal of much shorter duration. Furthermore, the initial estimate of system gain grants the user confidence to increase the input span to ± 1.5, improving the signal to noise ratio. The model estimated from the the two cycles of data can be used to generate the third and final stage; in this simulated scenario an input design with a control-relevant power spectrum as defined per Rivera *et al.* (1992) and with higher span (± 2.5) is used. Given the improved model knowledge from analysis of the first two cycles, the increased input span does not result in unacceptable swings in the output. This final design takes advantage of improved process knowledge to incorporate control requirements which ultimately results in a high performance control system.

It was previously stated that robust loopshaping measures can be used to determine in real-time, during identification testing, limitations to achievable control performance from the data. Consider the signal design per Stage 1 in the previous example. Here a Type-I weight function per Skogestad and Postlethwaite (1996) is chosen:

$$w_P = \frac{s/M_p + \omega_B^*}{s + \omega_B^* A} \tag{8}$$

where M_p and ω_B^* represent the upper maximum peak and lower bandwidth bounds in ϵ, respectively. A (set to zero in this analysis) is a zero-frequency upper bound. Figure 2 presents the results of the various bounds in Braatz *et al.* (1991) for two cycles of the signal per Stage 1 with SNR = 1 and performance specs defined by $M_p = 2$ and $\omega_B^* = 0.1$. Figure 2 shows that a nominal control design with sensitivity and complementary sensitivity operators per

$$\epsilon(s) = \frac{10s}{10s+1}, \qquad \eta(s) = \frac{1}{10s+1}$$

satisfies robust performance conditions. Having confirmed via this analysis that the data is sufficiently informative to yield a model leading to acceptable performance requirements, the experimental test can be halted.

5. SUMMARY AND CONCLUSIONS

We have presented the plant-friendly identification problem as an important issue meaningful to the process industries. In spite of the significant advances in the system identification field, this problem remains open to further avenues of research. Identification test monitoring was presented as one approach to address this gap, and a basic identification test monitoring scheme relying on the integration of identification concepts with robust loopshaping was presented. These measures, while useful, are conservative, which is a consideration that should be addressed in future work. Furthermore, non-expert usage is critical in the implementation of this concept in practice.

6. ACKNOWLEDGEMENT

Support for this work from the American Chemical Society - Petroleum Research Fund, Grant No. ACS PRF# 37610-AC9 is gratefully acknowledged.

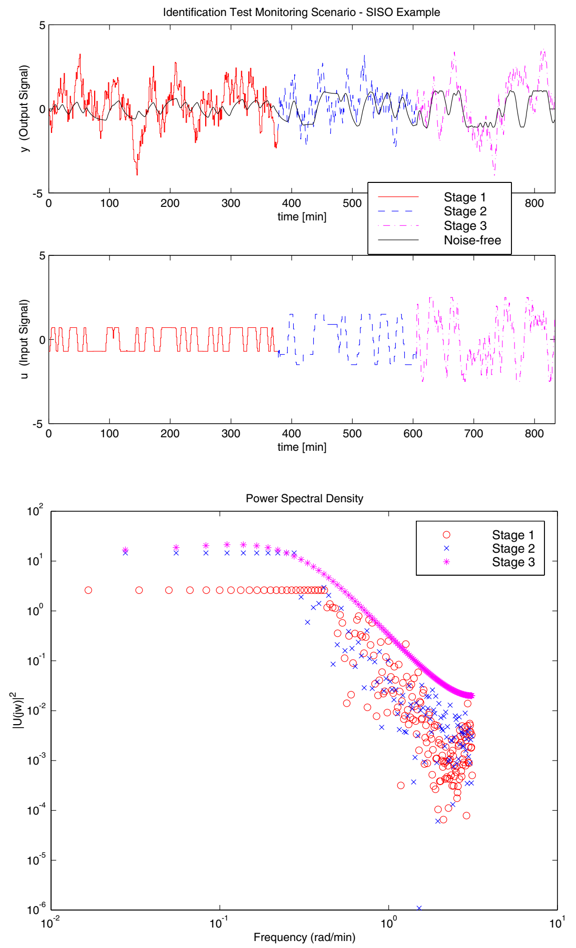

Fig. 1. Simulated identification test monitoring scenario for a single-input, single-output plant. a) time series b) power spectra. The "snow effect" is enabled only in the high frequency range of Stages 1 and 2.

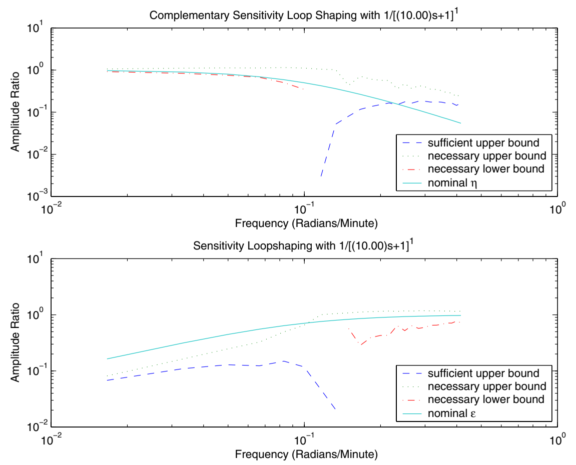

Fig. 2. Robust loopshaping example results.

7. REFERENCES

Bayard, D.S. (1993). Statistical plant set estimation using schroeder-phased multisinusoidal input design. *J. Applied Mathematics and Computation* **58**, 169.

Braatz, R.D., M. Morari and J.H. Lee (1991). Necessary/sufficient loopshaping bounds for robust

Type	Signal (x)	CF(x)	PIPS(%)	$\max|\Delta x|$	$\max x$	$\min x$	SNR (db)
Stage 1: flat PSD (with snow);	u	1.0898	91.7598	0.4999	0.7042	-0.7042	
$|\Delta u| \leq 0.5$	y	2.0623	50.4701	0.2437	0.6174	-0.6705	-18.99
Stage 2: flat PSD (with snow);	u	1.2665	78.9565	0.4999	1.4999	-1.4999	
$|\Delta u| \leq 0.5$, $|u| \leq 1.5$	y	1.4816	67.4283	0.6497	1.0248	-1.0247	-13.67
Stage 3: control-relevant PSD;	u	1.6620	60.1671	0.6999	2.4999	-2.4999	
$|\Delta u| \leq 0.7$, $|u| \leq 2.5$	y	1.3619	75.3906	0.3079	1.0850	-1.1457	-13.46

Table 1. Summary of results for the simulated SISO identification test monitoring problem scenario. All statistics for y (except SNR) are calculated on the noise-free portion of the signal.

performance. In: *1991 AIChE Annual Meeting*. Los Angeles, CA.

Braun, M.W., R. Ortiz-Mojica and D.E. Rivera (2002). Application of minimum crest factor multisinusoidal signals for "plant-friendly" identification of nonlinear process systems. *Control Engineering Practice* **10**, 301.

Chikkula, Y. and J.H. Lee (1997). Input sequence design for parametric identification of nonlinear systems. In: *American Control Conference*. Albuquerque, New Mexico. pp. 3037–3041.

Cooley, B.L. and J.H. Lee (2001). Control-relevant experiment design for multivariable systems described by expansions in orthonormal bases. *Automatica* **37**, 273–281.

Cooley, B.L., J.H. Lee and S.P. Boyd (1998). Control-relevant experiment design: a plant-friendly, LMI-based approach. In: *American Control Conference*. Vol. 2. Philadelphia , PA. pp. 1240–1244.

Doyle, F.J., R.S. Parker, R.K. Pearson and B.A. Ogunnaike (1999). Plant-friendly identification of second-order volterra models. In: *European Control Conference*. Karlsruhe, Germany.

Godfrey, K., Ed.) (1993). *Perturbation Signals For System Identification*. Prentice Hall International (UK) Limited. Hertfordshire, UK.

Godfrey, K.R., H.A Barker and A.J. Tucker (1999). Comparison of perturbation signal for linear system identification in the frequency domain. *IEE. Proc. Control Theory Appl.* **146**, 535.

Guillaume, P., J. Schoukens, R. Pintelon and I. Kollár (1991). Crest-factor minimization using nonlinear chebyshev approximation methods. *IEEE Trans. on Inst. and Meas.* **40**(6), 982–989.

Hussain, M.A. (1999). Review of the applications of neural networks in chemical process control-simulation and on-line implementation. *Artificial Intelligence in Engineering* **13**(1), 55–68.

Lee, H., D.E. Rivera and H. Mittelmann (2003). Constrained minimum crest factor multisine signals for plant-friendly identification of highly interactive systems. In: *SYSID 2003*. Rotterdam, The Netherlands.

Li, T. and C. Georgakis (2002). Design of multivariable identification signals for constrained systems. In: *Annual AIChE 2002 Meeting*. Indianapolis, IN. paper 255g.

Ljung, L. (1999). *System Identification: Theory for the User*. 2nd ed.. Prentice-Hall. New Jersey.

Mathur, U. and R.J. Conroy (2002). Multivariable control without plant tests. In: *Annual AIChE 2002 Meeting*. Indianapolis, IN. paper 254g.

Narasimhan, S., S. Rengaswamy and R. Rengasamy (2003). Multiobjective input signal design for plant-friendly identification. In: *SYSID 2003*. Rotterdam, The Netherlands.

Ogawa, M. and H. Douke (2002). Process control at Mitsubishi. `http://www.che.utexas.edu / twmcc / Presentation0203/ Douke03spring.pdf` .

Parker, R.S., D. Heemstra, J.D. Doyle III, R.K. Pearson and B.A. Ogunnaike (2001). The identification of nonlinear models for process control using tailored "plant-friendly" input sequences. *J. of Process Control* **11**(2), 237–250.

Pearson, R.K., B.A. Ogunnaike and F.J. Doyle III (1993). Identification of nonlinear input/output models using non-gaussian input sequences. In: *American Control Conference*. San Francisco, CA. pp. 1465–1469.

Rengasamy, R., R.S. Parker and F.J. Doyle III (2000). Issues in design of input signals for the identification of nonlinear models of process systems. In: *ADCHEM 2000*. Pisa, Italy.

Rivera, D.E., J.F. Pollard and C.E. García (1992). Control-relevant prefiltering: A systematic design approach and case study: Special issue on system identification for robust control design. *IEEE Trans. Autom. Cntrl.* **37**, 964–974.

Rivera, D.E., M.W. Braun and H.D. Mittelmann (2002). Constrained multisine inputs for plant-friendly identification of chemical processes. *15th IFAC World Congress, Barcelona, Spain.*

Skogestad, S. and I. Postlethwaite (1996). *Multivariable feedback control: analysis and design*. Wiley. New York, NY.