

---

# “Plant-Friendly” System Identification: A Challenge for the Process Industries

Daniel E. Rivera

Control Systems Engineering Laboratory  
Department of Chemical and Materials Engineering  
Arizona State University  
Tempe, Arizona 85287-6006

<http://www.fulton.asu.edu/~csel>  
[daniel.rivera@asu.edu](mailto:daniel.rivera@asu.edu)

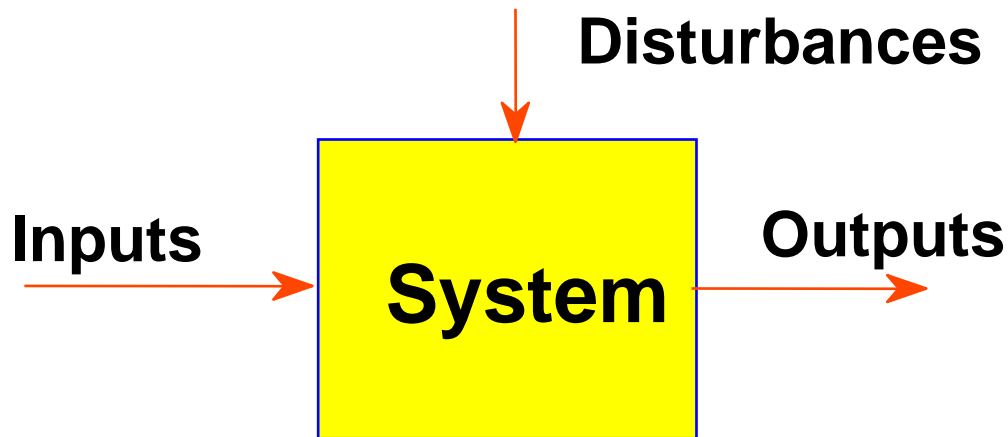
# Presentation Outline

- What is system identification?
- “Plant-friendliness” in identification testing.
- Minimum crest factor multisine signals for strongly interactive processes - *high purity distillation*.
- Identification Test Monitoring.
- Summary and Conclusions

# System Identification

“Identification is the determination, on the basis of input and output, of a system within a specified class of systems, to which the system under test is equivalent.”

- L. Zadeh, (1962)



System identification focuses on the modeling of dynamical systems from experimental data

# Some General Facts Regarding System Identification

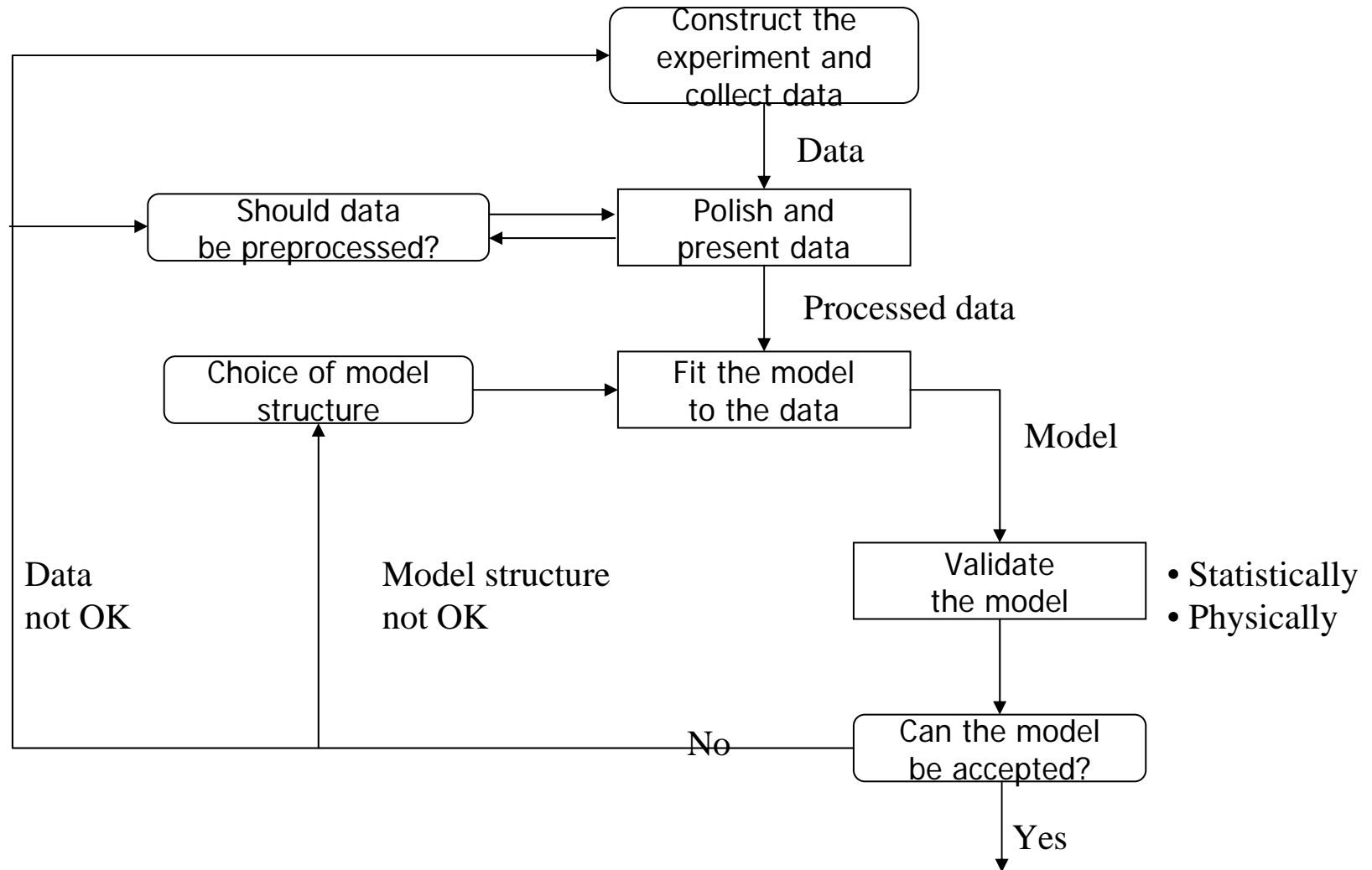
- System identification is not exclusive to control system design, although it forms a significant component of control system implementation.
- Often times, the system identification task is the most expensive and time-consuming portion of advanced control projects in industry.
- It is a broadly-applicable area with applications in many diverse fields.

---

# Steps in System Identification

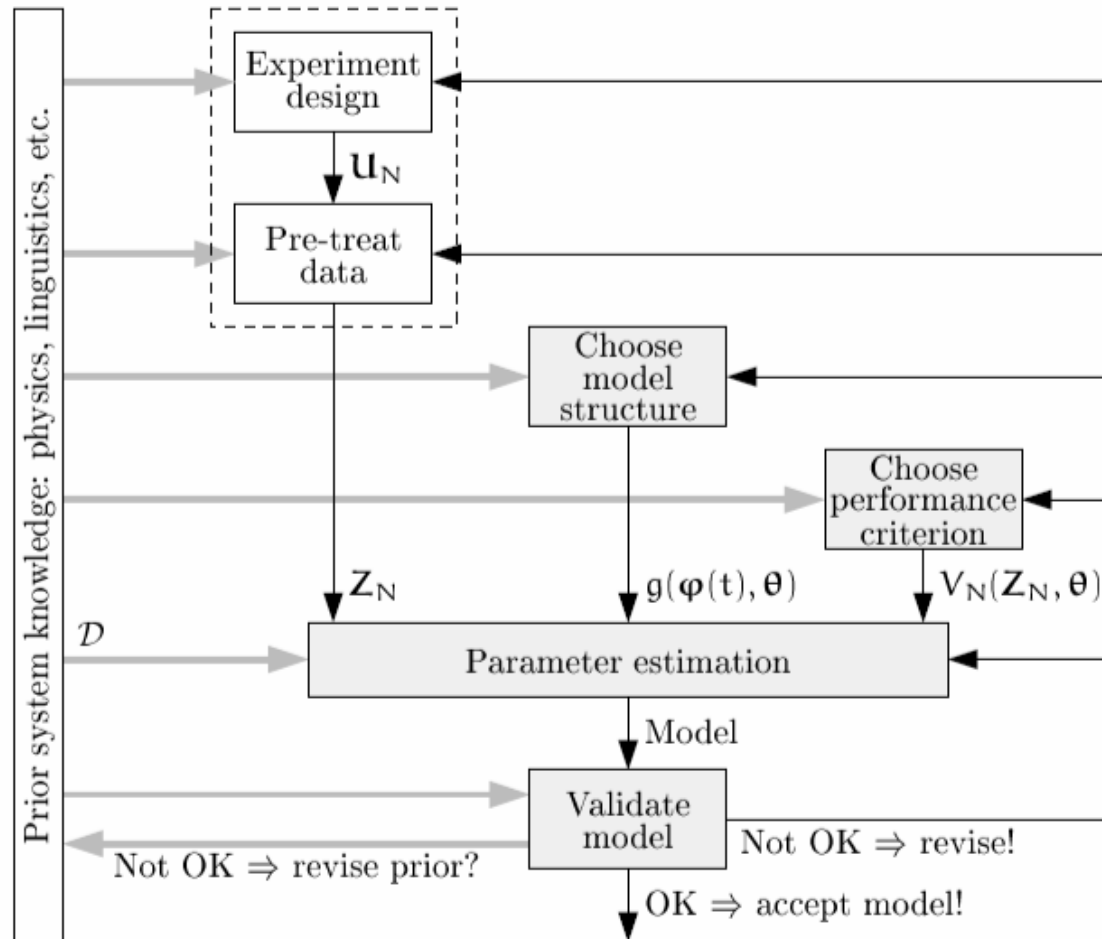
- Experiment Design and Execution
- Data preprocessing
- Model structure selection
- Parameter Estimation
- Model Validation

# System Identification Loop (Ljung and Glad, 1994)

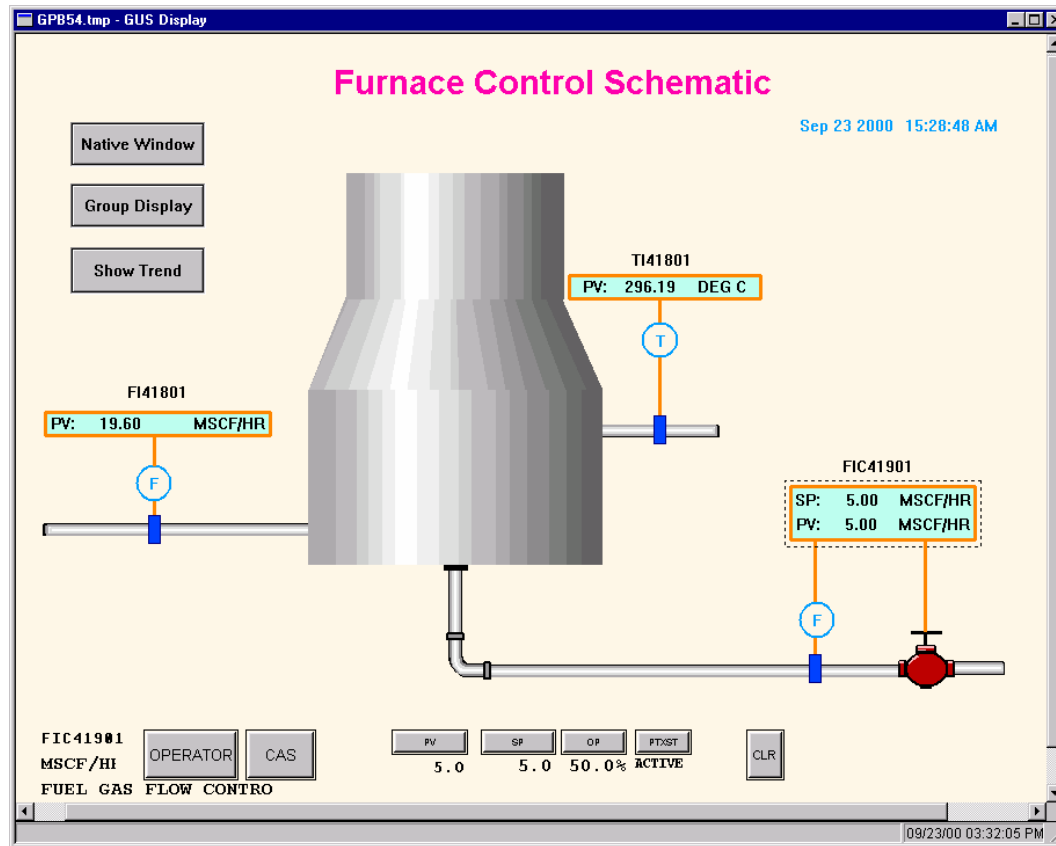


# System Identification Loop - 2

(reprinted from Lindskog (1996), with permission)



# An Industrial Process Control Problem



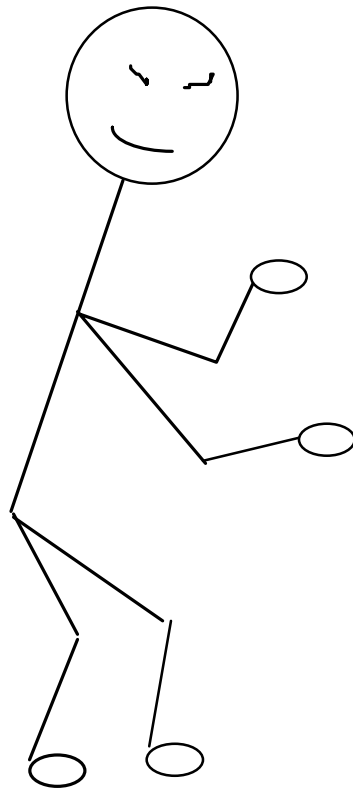
Objective: Use fuel gas flow to keep outlet temperature under control, in spite of occasional yet significant changes in the feed flowrate.



# The “Shower” Control Problem

Controlled:  
Temperature,  
Total Water Flow

Disturbances:  
Inlet Water Flows,  
Temperatures



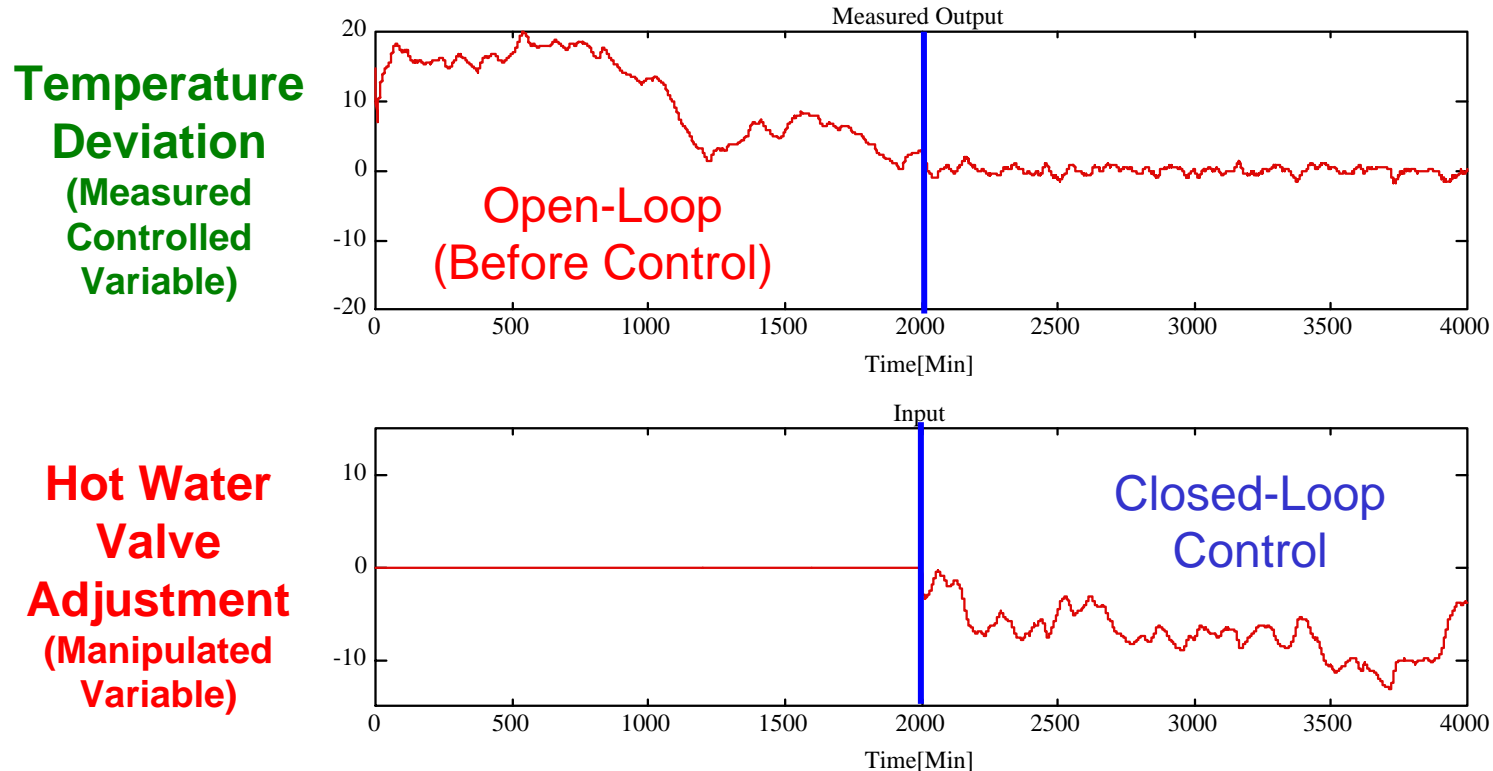
The presence of delay or  
“transportation lag”  
makes this a difficult control  
problem

Hot

Cold

Manipulated: Hot and Cold  
Water Valve Positions

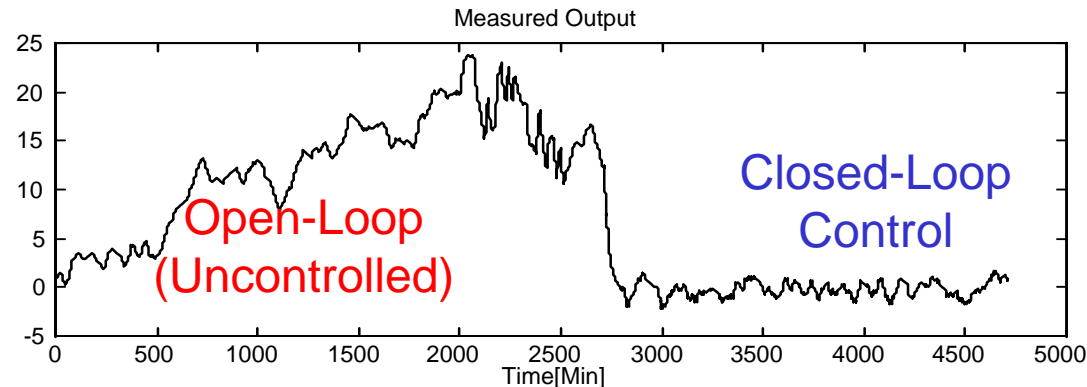
# From Open-Loop Operation to Closed-Loop Control



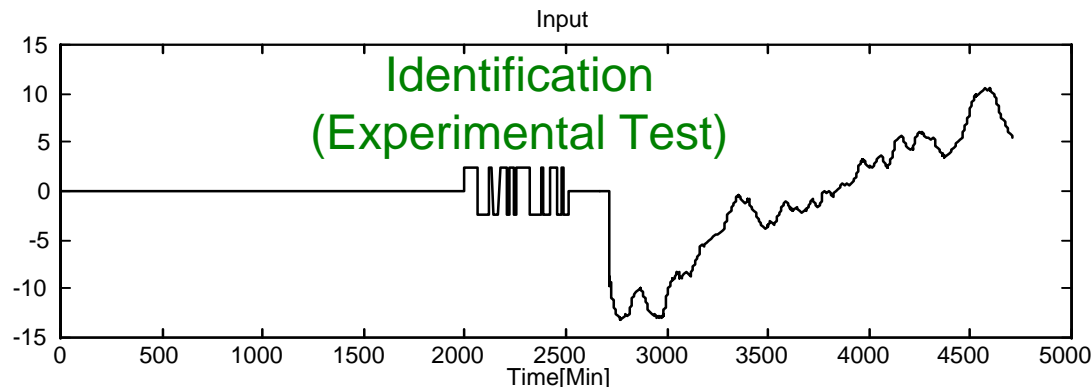
The transfer of variance from an expensive resource to a cheaper one is one of the major benefits of engineering process control

# From Open-Loop Operation to Closed-Loop Feedback Control

Temperature  
Deviation  
(Controlled  
Variable)



Fuel Flow  
(Manipulated  
Variable)



Furnace example with PRBS input, PID with filter controller

# “Plant-Friendly” Identification Testing

- The term originates from the chemical process control community; first used by Dupont control researchers and collaborators in the early 90's.
- Is principally motivated by the desire for informative identification experiments while meeting the demands of industrial practice.
- *Broadly speaking*, a plant-friendly test yields 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.

# “Plant-Friendly” Identification Testing (Continued)

The ideal plant-friendly identification test should:

- be as short as possible,
- not take actuators to limits, or exceed move size restrictions,
- cause minimal disruption to the controlled variables (i.e., low variance, small deviations from setpoint).

*Note that theoretical requirements may strongly conflict  
with "plant-friendly" operation!*

---

# Motivation for Plant-Friendly Identification

- Plant operations desires plant-friendliness, but classical identification theory is “plant-hostile”

# Reducing Variance Effects

$$\text{Model Parameter Variance} \sim \frac{n}{N} \frac{\lambda}{\mu}$$

$n \equiv$  no. of model parameters

$N \equiv$  no. of data points (length of data set)

$\lambda \equiv$  variance of the disturbance signal  $\nu$

$\mu \equiv$  variance of the input signal  $u$

$\mu/\lambda \equiv$  input signal-to-noise ratio

**Reducing the number of estimated model parameters, *increasing the length of the data set*, and *increasing the variance of the input signal* all contribute to variance reduction in system identification**

# Motivation for Plant-Friendly Identification

- Plant operations desires plant-friendliness, but classical identification theory is “plant-hostile”
- Identification testing is an expensive proposition, and improper execution can endanger a project.
- There is an absence of fundamentally based, systematic guidelines in the literature for problems of practical significance



# Process Testing Duration

(as reported by Mitsubishi Chemical engineers,  
from guidelines presented by a major process control software vendor)

Suggested Test Duration =

$(6...8) * (\text{Estimated Settling Time Process}) * (\text{Number of Independent Variables})$

*Example: Ethylene Fractionator:*

$$\begin{aligned} 6 * 6 \text{ (hrs)} * 17 &= 612 \text{ (hrs)} = 25.5 \text{ (days)} \\ 8 * 6 \text{ (hrs)} * 17 &= 816 \text{ (hrs)} = 34 \text{ (days)} \end{aligned}$$

# Incentives for “Fast” Identification Testing

Per Kothare and Mandler, Air Products & Chemicals, (presented at the 2003 AIChE Annual Mtg.)

Inputs



Outputs

**Estimate for a large Air Separation Unit:  
2 months at the plant 24 hrs/day!**

# Typical Costs of Step Testing

(from Mathur and Conroy, "Multivariable Control without Plant Tests" *2002 AIChE Annual Mtg.*)

- Cut throughput, 5-10% for 6-8 weeks \$ 50,000
  - One off-grade excursion, 100% production loss \$ 60,000
  - Engineering (testing) 6-8 weeks, 24 hours/day \$140,000
  - Engineering (commissioning), 2 weeks, 24 hours/day \$ 20,000
- Total: \$270,000**

# Motivation for Plant-Friendly Identification

- Plant operations desires plant-friendliness, but classical identification theory is “plant-hostile”
- Identification testing is an expensive proposition, and improper execution can endanger a project.
- There is an absence of fundamentally based, systematic guidelines in the literature for problems of practical significance
- Some well-established identification topics (e.g., classical optimal input design, control-relevant identification, closed-loop identification) are helpful but do not address all the issues.

# Classical Optimal Input Signal Design

- Classical formulations (summarized in Chpt. 13 of Ljung's System Identification: Theory for the User) address minimizing the constrained variance of the input and/or output signals
- The optimal experimental design depends on the (unknown) true system and noise characteristics
- In practice, process control engineers tend to think more in terms of keeping manipulated and controlled variables under constraints and minimizing overall test duration than in achieving constrained variance.

# A Benchmark Highly Interactive System: High-Purity Distillation

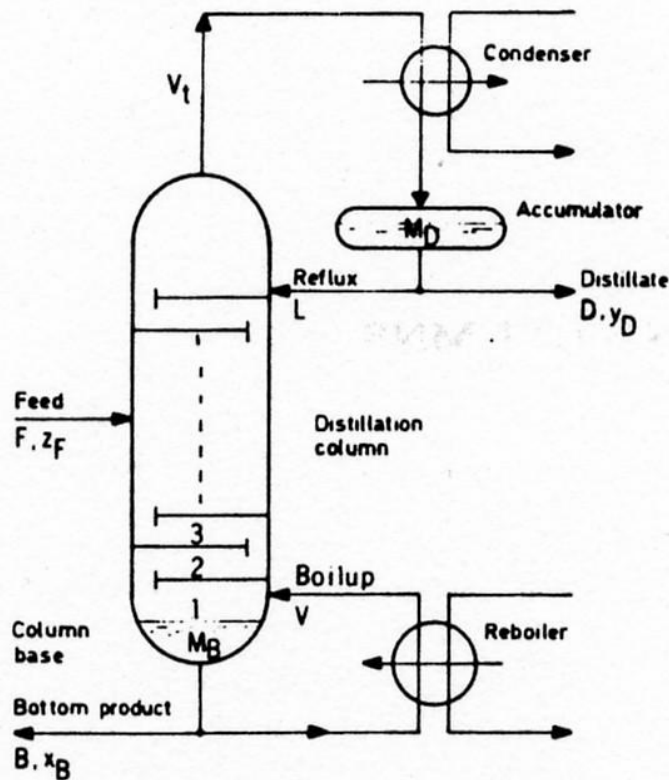
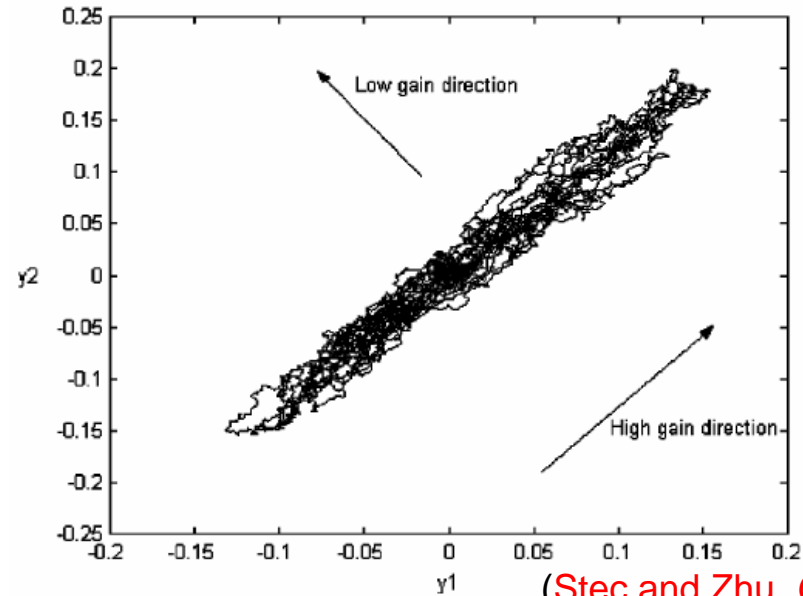


Fig. 2. Two-product distillation column.



(Stec and Zhu, CEP 2002)

High-Purity Distillation Column per Weischedel and McAvoy (1980) : a classical example of a highly interactive process system, and a challenging problem for control system design

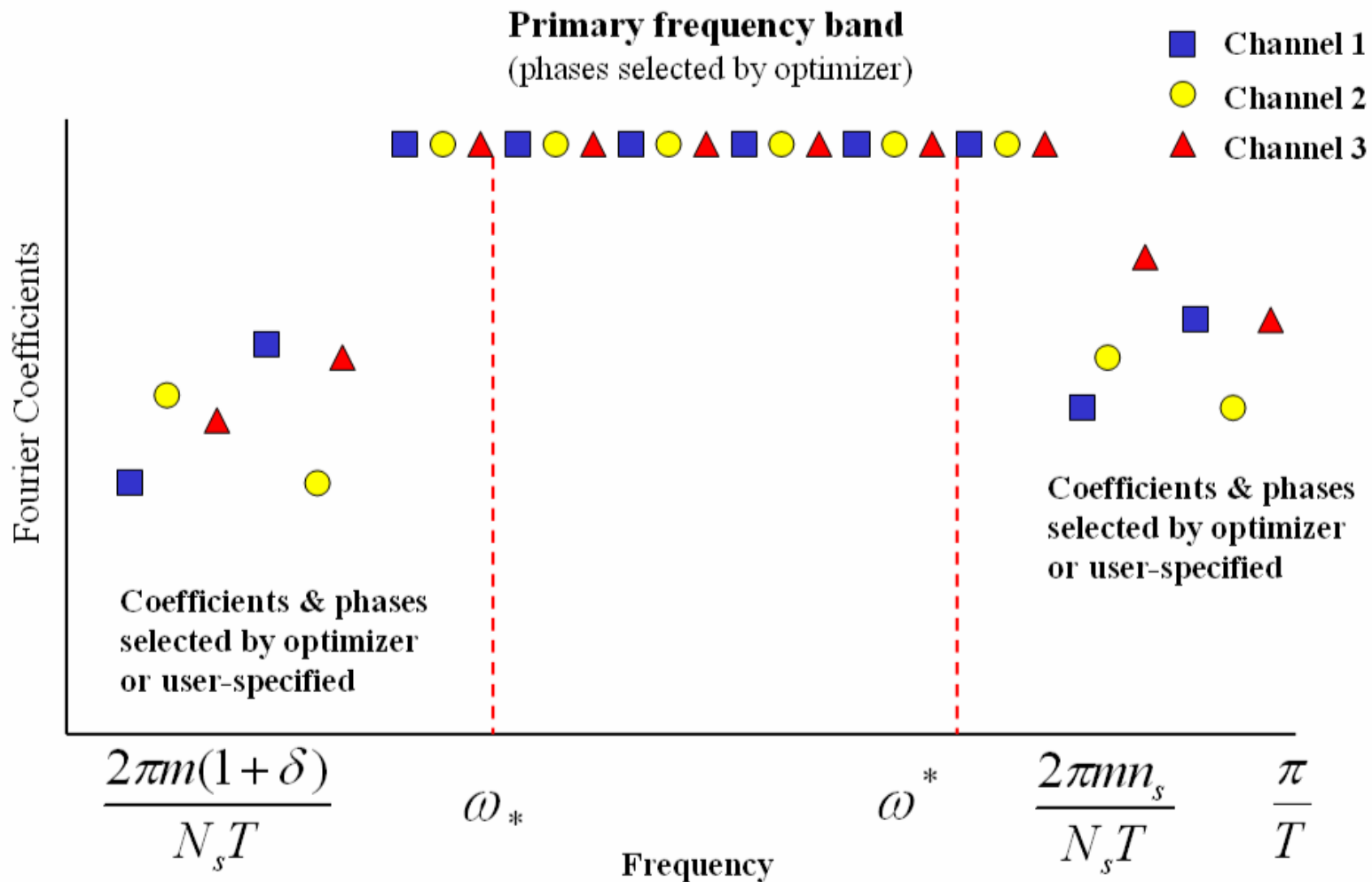
# Multisine Input Signals

A multisine input is a deterministic, periodic signal composed of a harmonically related sum of sinusoids,

$$u_j(k) = \sum_{i=1}^{m\delta} \hat{\delta}_{ji} \cos(\omega_i kT + \phi_{ji}^{\delta}) + \sum_{i=m\delta+1}^{m(\delta+n_s)} \alpha_{ji} \cos(\omega_i kT + \phi_{ji}) \\ + \sum_{i=m(\delta+n_s)+1}^{m(\delta+n_s+n_a)} \hat{a}_{ji} \cos(\omega_i kT + \phi_{ji}^a), \quad j = 1, \dots, m$$

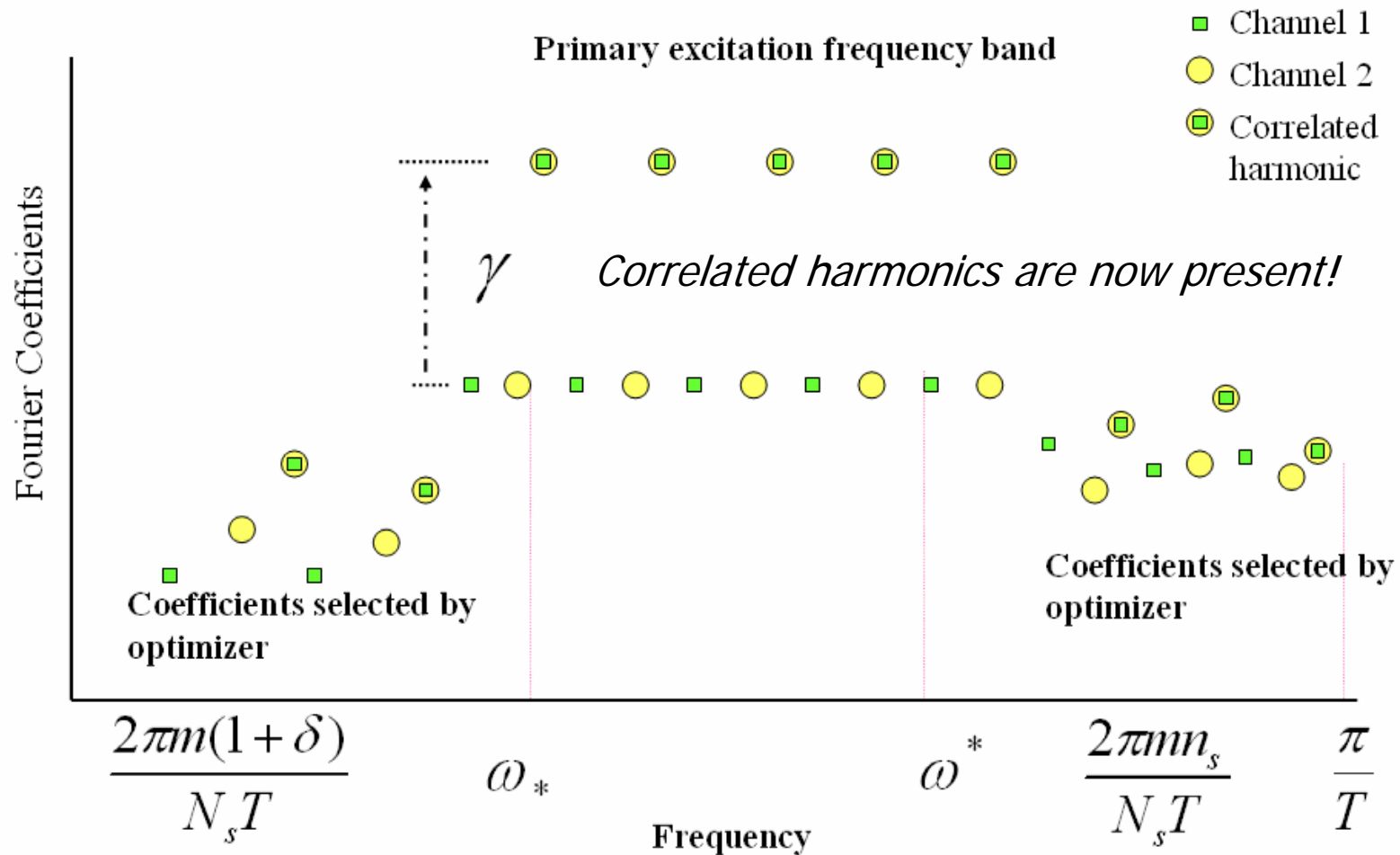
where  $T$  is sampling time,  $N_s$  is the sequence length,  $m$  is the number of channels,  $\delta, n_s, n_a$  are the number of sinusoids per channel ( $m(\delta + n_s + n_a) = N_s/2$ ),  $\phi_{ji}^{\delta}, \phi_{ji}, \phi_{ji}^a$  are the phase angles,  $\alpha_{ji}$  represents the Fourier coefficients defined by the user,  $\hat{\delta}_{ji}, \hat{a}_{ji}$  are the “snow effect” Fourier coefficients

# “Zippered” Power Spectrum





# Modified Zippered Spectrum



# Crest Factor

The Crest Factor (CF) is defined as the ratio of  $\ell_\infty$  (or Chebyshev) norm and the  $\ell_2$  norm

$$CF(x) = \frac{\ell_\infty(x)}{\ell_2(x)}$$

A low crest factor indicates that most elements in the signal are located near the minimum and maximum values of the sequence.

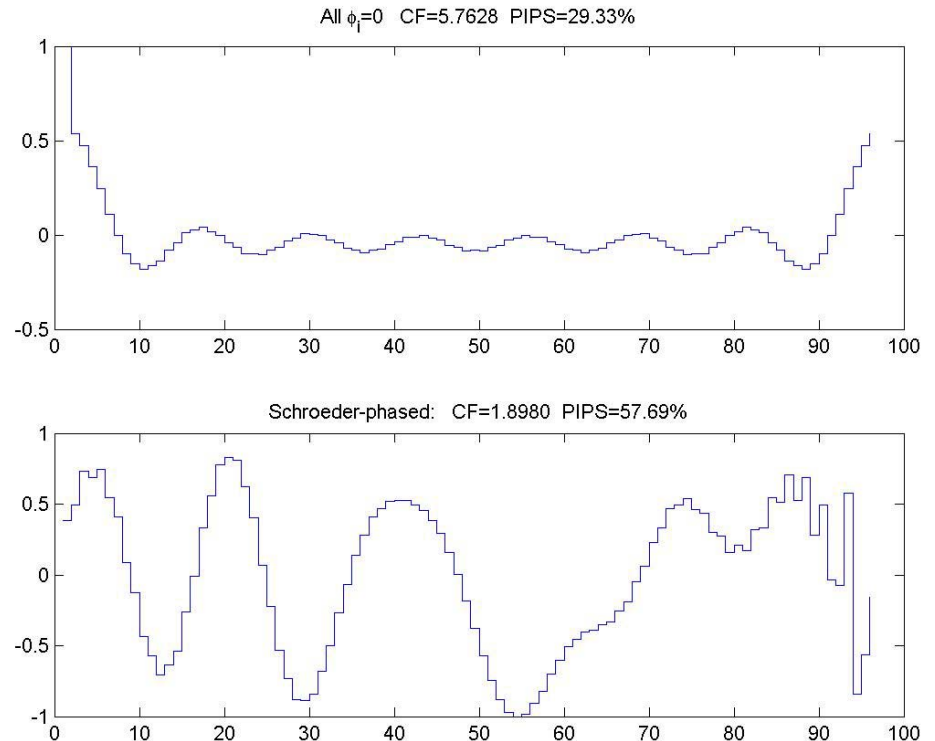
- Seminal paper by Schroeder (1970) presents an analytical formula for determining phases in multisine signals that leads to near-optimal crest factors (for wide-band signals)
- Work by Guillaume *et al.* (1991) provides a very efficient numerical technique for computing minimum crest factor multisine signals with arbitrary power spectral densities

# Crest Factor Signal Comparison

Two signals with identical spectra and different crest factors can have markedly different “plant-friendliness” properties.

The Performance Index for Perturbation Signals (PIPS) is a practical alternative (Godfrey, Barker, & Tucker, *IEE Proc. Control Theory Appl.*, 1999):

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



# Problem Statement #1

$$\min_{\{\phi_{ji}^a\}, \{\phi_{ji}^\delta\}, \{\phi_{ji}\}, \{\hat{a}_{ji}\}, \{\hat{\delta}_{ji}\}} \max_j \text{CF}(u_j) \quad j = 1, \dots, m$$

subject to maximum move size constraints on  $\{u_j(k)\}$

$$|\Delta u_j(k)| \leq \Delta u_j^{max} \quad \forall k, j$$

and high/low limits on  $\{u_j(k)\}$

$$u_j^{min} \leq u_j(k) \leq u_j^{max} \quad \forall k, j$$

## Problem Statement #2

$$\min_{\{\phi_{ji}^a\}, \{\phi_{ji}^b\}, \{\phi_{ji}\}, \{\hat{a}_{ji}\}, \{\hat{b}_{ji}\}} \max_z \text{CF}(y_z)$$
$$j = 1, \dots, m \quad z = 1, \dots, N_{outs}$$

subject to constraints in input

$$|\Delta u_j(k)| \leq \Delta u_j^{max} \quad \forall k, j$$
$$u_j^{min} \leq u_j(k) \leq u_j^{max} \quad \forall k, j$$

and output

$$|\Delta y_z(k)| \leq \Delta y_z^{max} \quad \forall k, z$$
$$y_z^{min} \leq y_z(k) \leq y_z^{max} \quad \forall k, z$$

This problem statement requires an *a priori* model to generate output predictions

# Other Problem Formulations

- Minimize worst-case of both input and output crest factors

$$\min_{\{\phi_{ji}^a\}, \{\phi_{ji}^b\}, \{\phi_{ji}\}, \{\hat{a}_{ji}\}, \{\hat{b}_{ji}\}} \max_{j, z} \{ \text{CF}(u_j), \text{CF}(y_z) \}$$
$$j = 1, \dots, m \quad z = 1, \dots, N_{outs}$$

- Incorporate controller equations in the optimization problem for signal design under closed-loop conditions
- Examine alternative criteria (e.g., geometric discrepancy via Weyl's Theorem) in lieu of crest factor.

# Constrained Solution Approach

Some aspects of our numerical solution approach:

- The problem is formulated in the modeling language AMPL, which provides exact, automatic differentiation up to second derivatives.
- A direct min-max solution is used where the nonsmoothness in the problem is transferred to the constraints.
- The trust region, interior point method developed by Nocedal and co-workers (Byrd, R., M.E. Hribar, and J. Nocedal. "An interior point method for large scale nonlinear programming." *SIAM J. Optim.*, 1999) is applied.

# Linear System Example

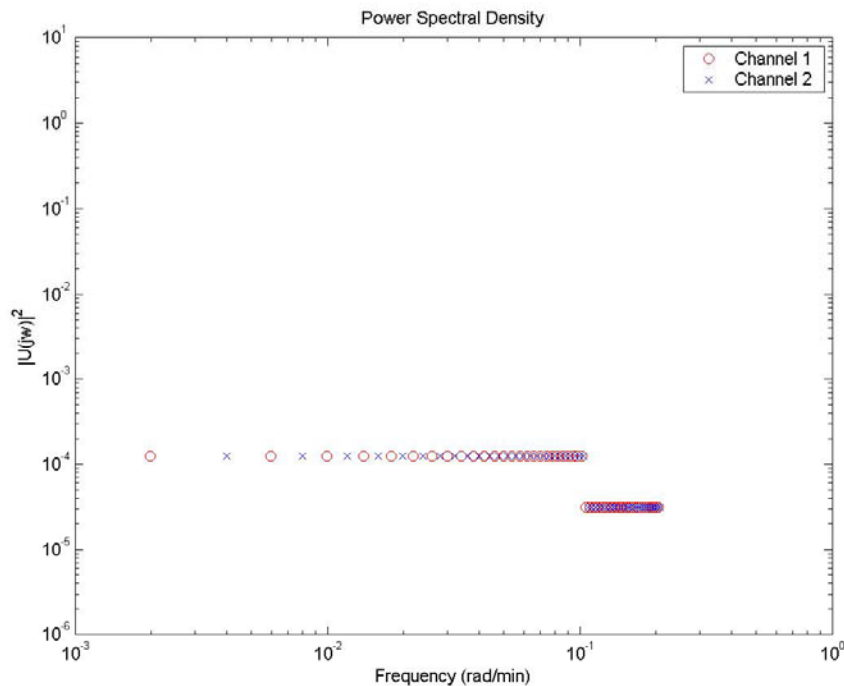
$$P(s) = \frac{1}{75s + 1} \begin{bmatrix} 87.8 & -86.4 \\ 108.2 & -109.6 \end{bmatrix}$$

- From Morari and Zafiriou, ***Robust Process Control*** (1988)
- Simplest meaningful highly interactive problem we could find...

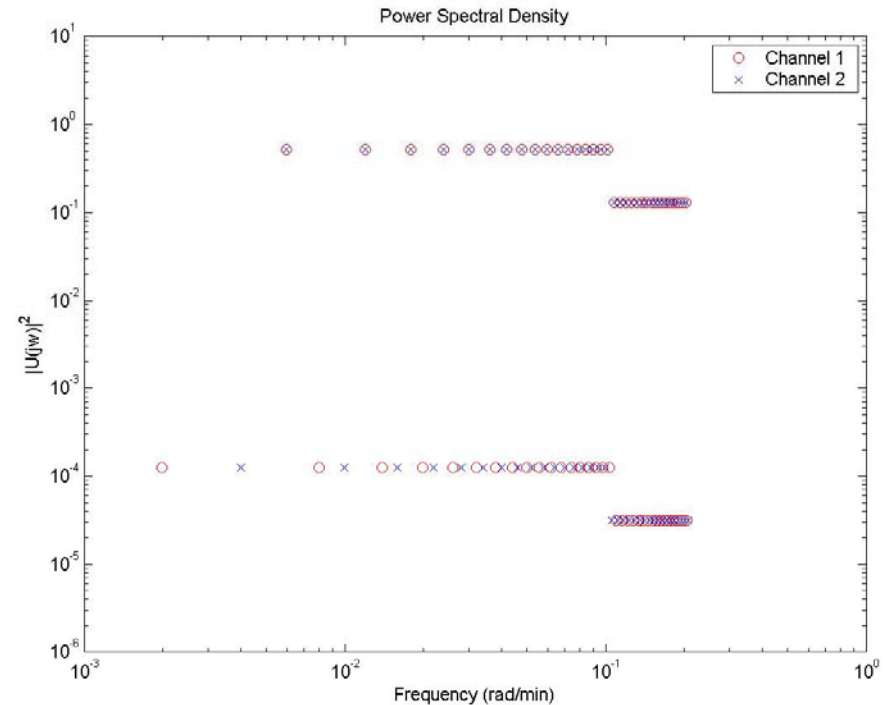


# min CF Signal Designs: power spectra

Standard Zippered Spectrum



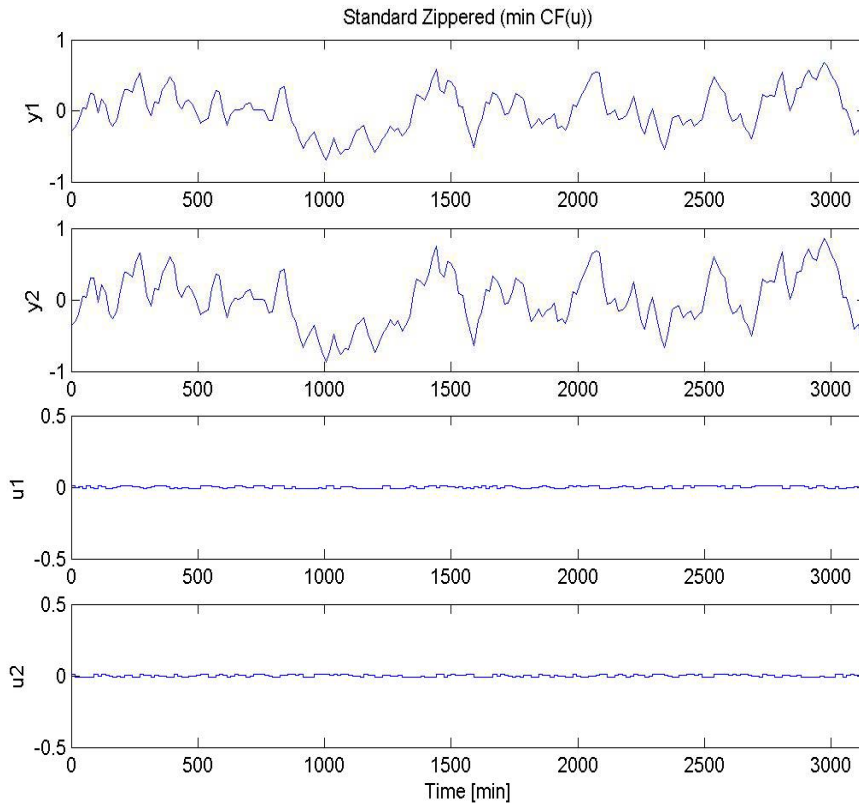
Modified Zippered Spectrum



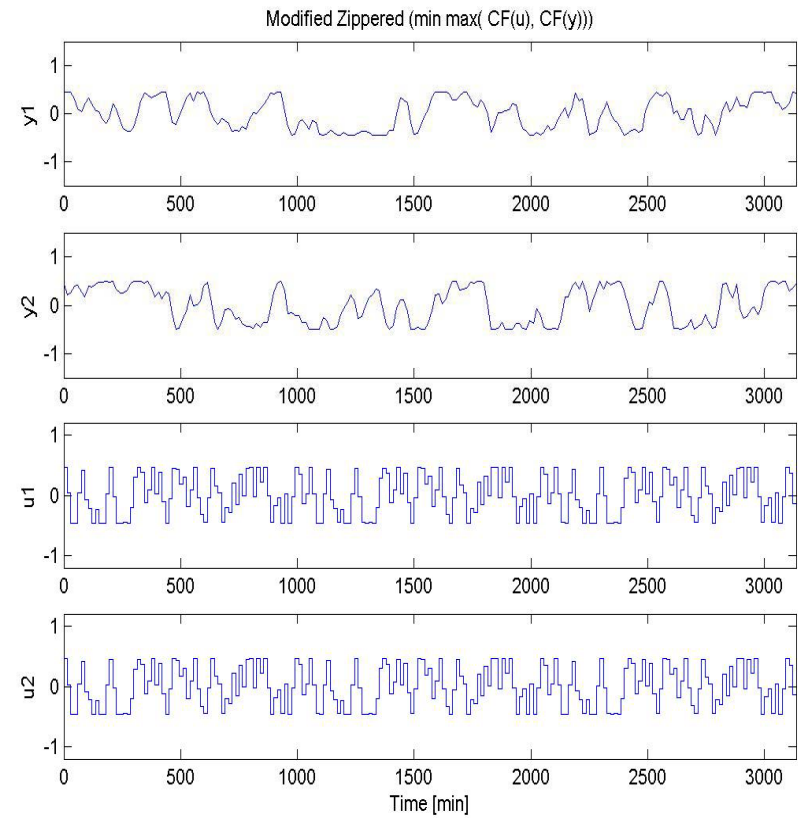
For  $\tau_{dom}^L = \tau_{dom}^H = 75$  min,  $\delta = 0$ ,  $\alpha_s = 7.5$ , and  $\beta_s = 3.33$ , feasible design choices are  $T = 15$  min,  $n_s = 26$ ,  $N_s = 210$ , and  $\gamma = 64.5$ .

# min CF Signal Designs: time-domain

$\min CF(u)$ , Standard Zippered Spectrum



$\min \max \{CF(u), CF(y)\}$   
Modified Zippered Spectrum

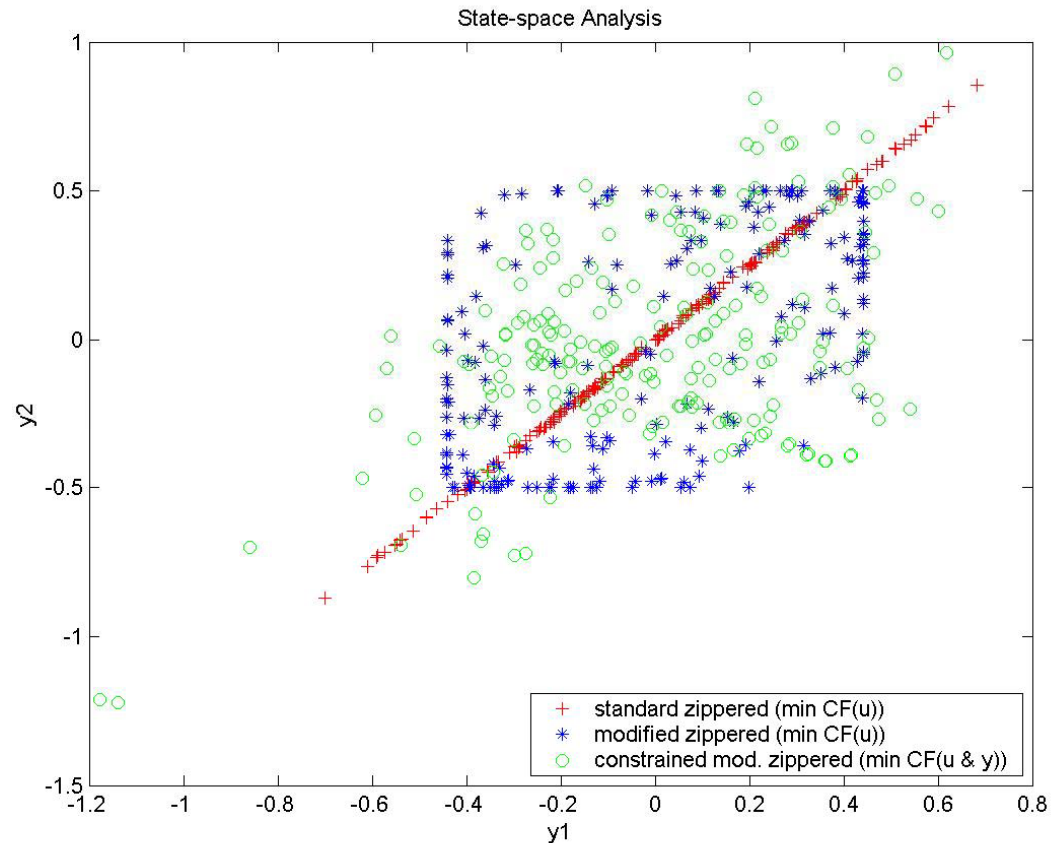


# min CF Signal Designs : State-Space Comparison

standard (+), modified  
unconstrained (o), and  
modified with  
constraints (\*)  
zippered designs

For the constrained modified  
zippered spectrum signal,

$$(\Delta u, \Delta y, |y|)^{max} \leq 0.5$$



# A Benchmark Highly Interactive System: High-Purity Distillation

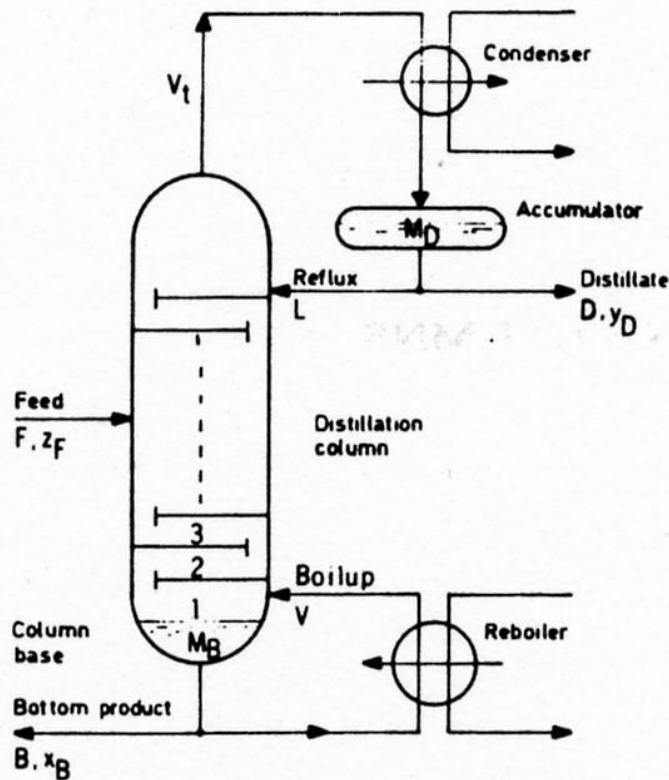
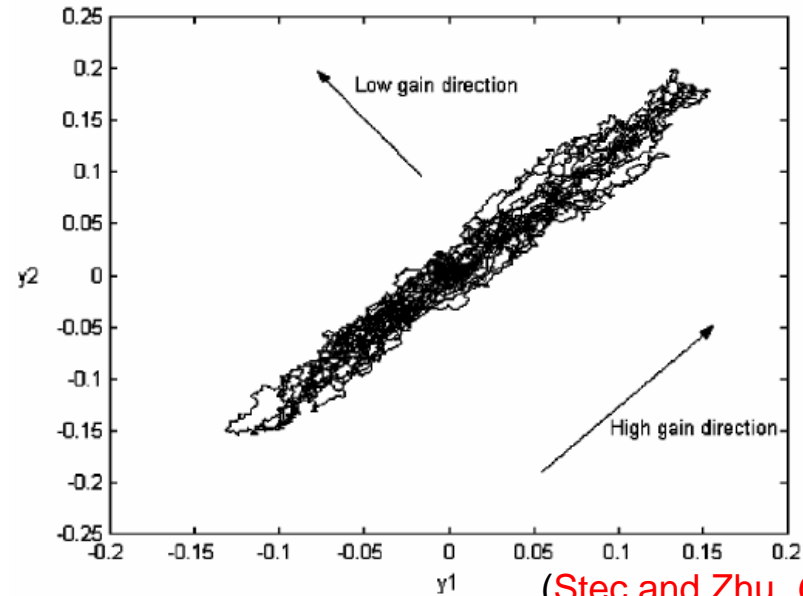


Fig. 2. Two-product distillation column.

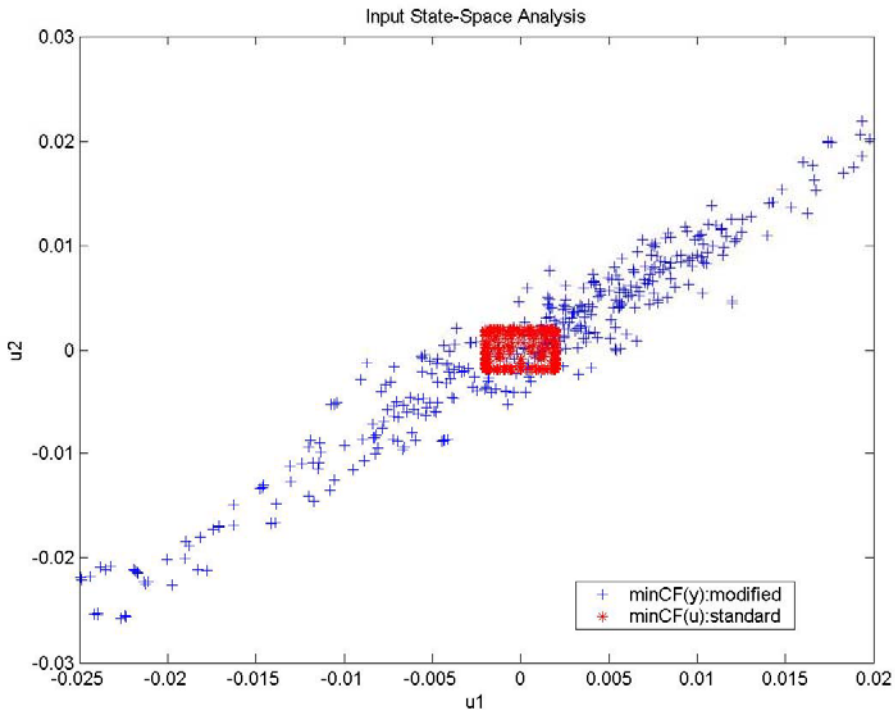


(Stec and Zhu, CEP 2002)

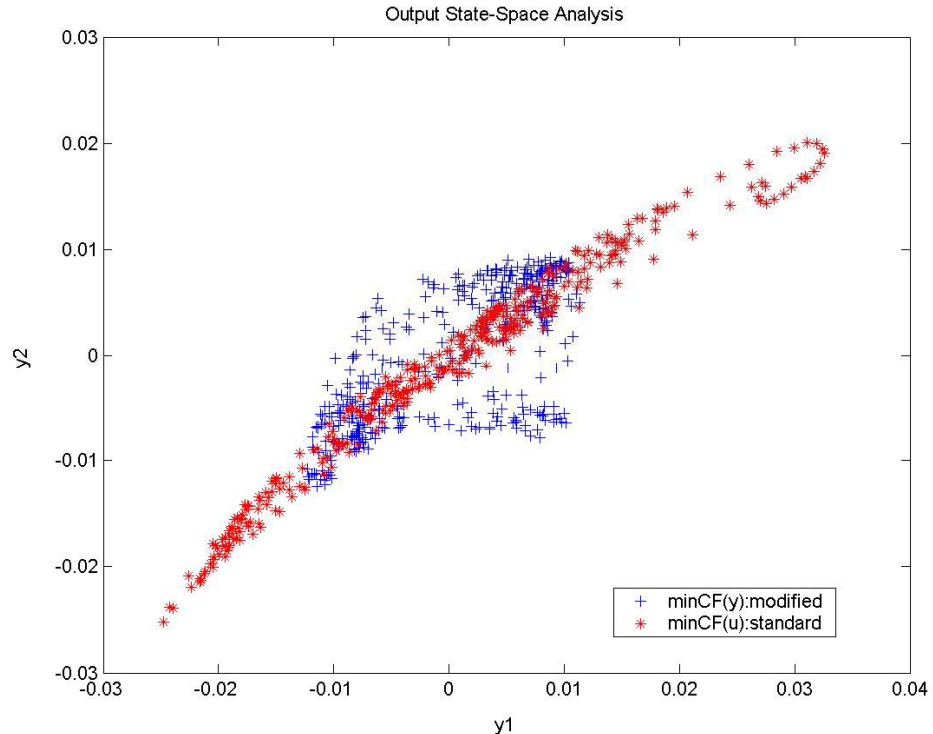
High-Purity Distillation Column per Weischedel and McAvoy (1980) : a classical example of a highly interactive process system, and a challenging problem for control system design

# State-space Analysis

## Input State-Space

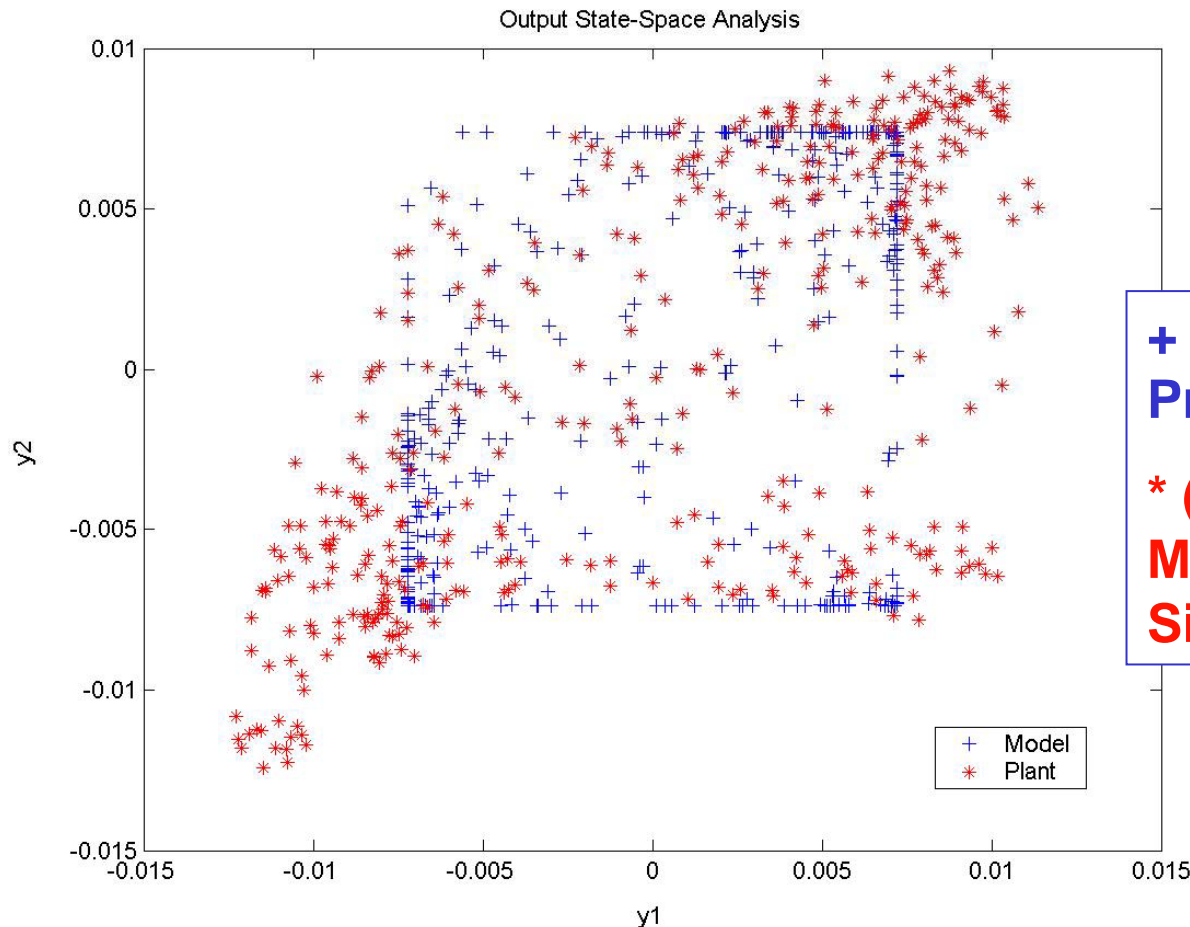


## Output State-Space



**+(blue): min CF(y) signal with a modified zippered spectrum and *a priori* ARX model**  
**\*(red) : min CF(u) signal with a standard zippered spectrum**

# Linear (ARX) Model Prediction vs. Plant Data





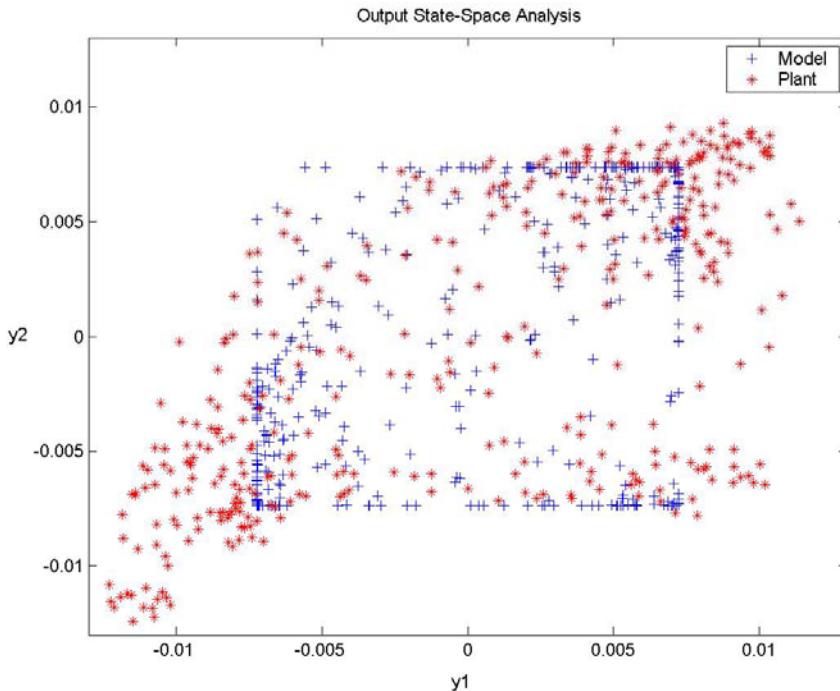
# NARX Model Estimation

**We rely on a NARX model to predict the system outputs during optimization (Srinivas *et al.*, 1995)**

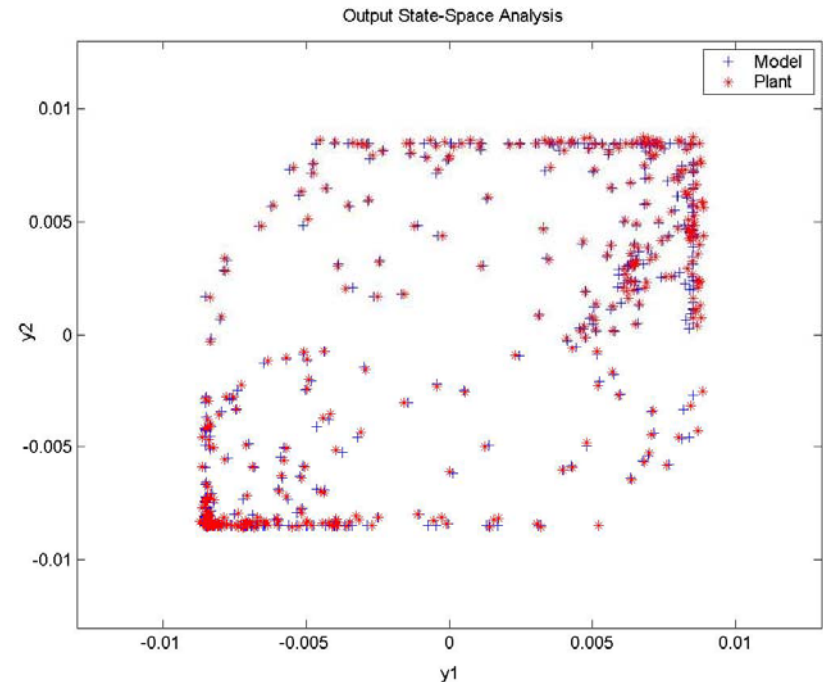
$$\begin{aligned} y(k) = & \theta^{(0)} + \sum_{i=1}^{n_y} \theta_i^{(1)} y(k-i) + \sum_{i=\rho}^{n_u} \theta_i^{(2)} u(k-i) + \sum_{i=1}^{n_y} \sum_{j=1}^i \theta_{(i,j)}^{(3)} y(k-i) y(k-j) \\ & + \sum_{i=\rho}^{n_u} \sum_{j=\rho}^i \theta_{(i,j)}^{(4)} u(k-i) u(k-j) + \sum_{i=1}^{n_y} \sum_{j=\rho}^{n_u} \theta_{(i,j)}^{(5)} y(k-i) u(k-j) + \dots \end{aligned}$$

# ARX vs. NARX Model Predictions

ARX Model



NARX Model

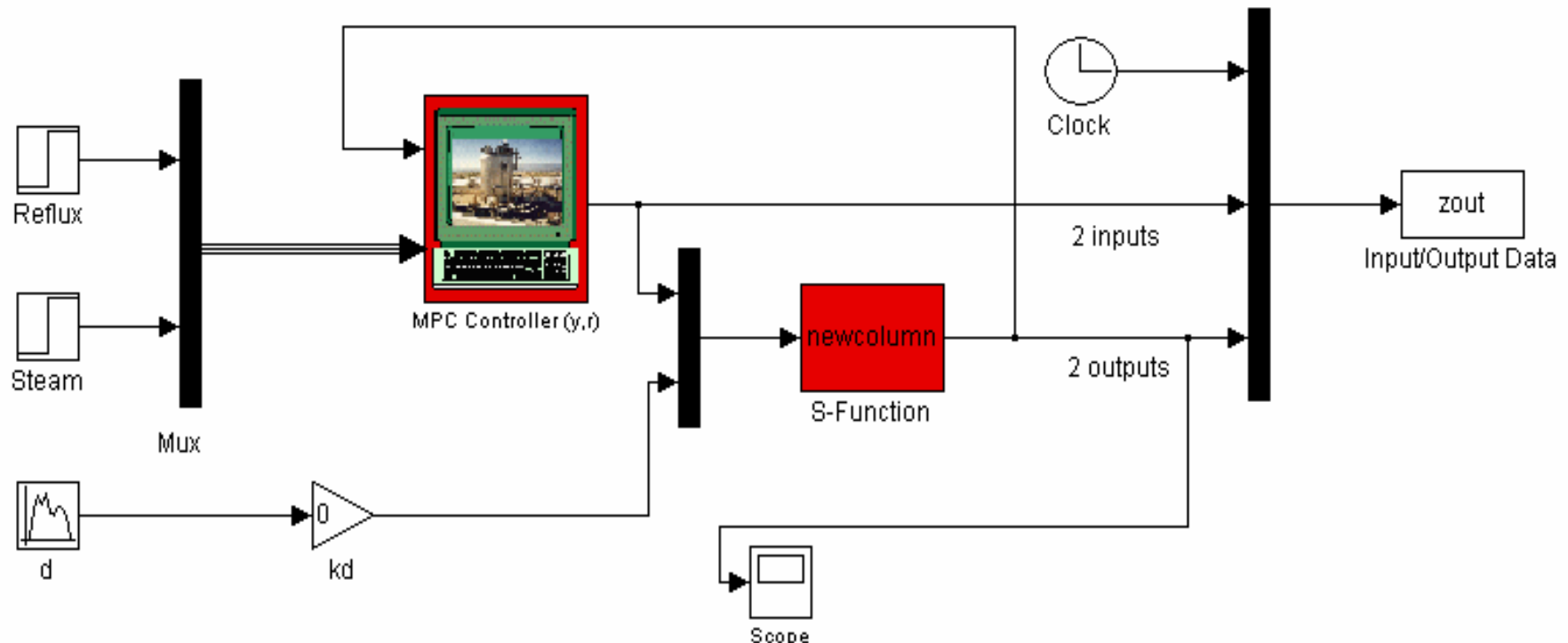


+ (blue) : Model Prediction

\* (red) : Weischedel-McAvoy Distillation Simulation



## DuPont High-Purity Distillation Column



### MPC Tuning Parameters:

Prediction Horizon PHOR : 100 , Move Horizon : 25

Output Weighting: [1 1] , Input Weighting : [0.2 0.2]

# Closed-loop Performance Comparison, MPC Setpoint Tracking: Models obtained from noisy data

## Model Predictive Control (MPC)

optimizes the predicted future values of the plant output based on previous and future information

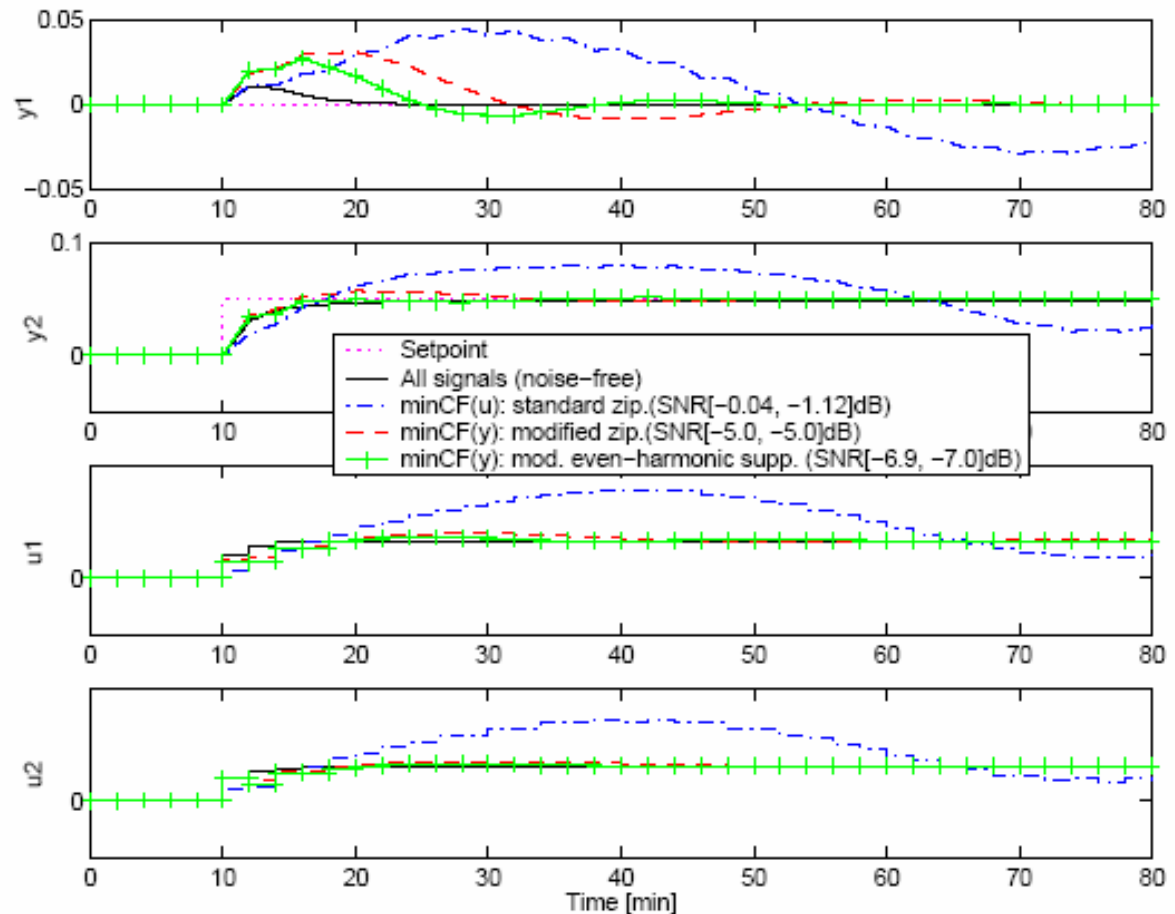
### MPC Tuning Parameters:

Prediction Horizon : 100

Move Horizon : 25

Output Weighting: [1 1]

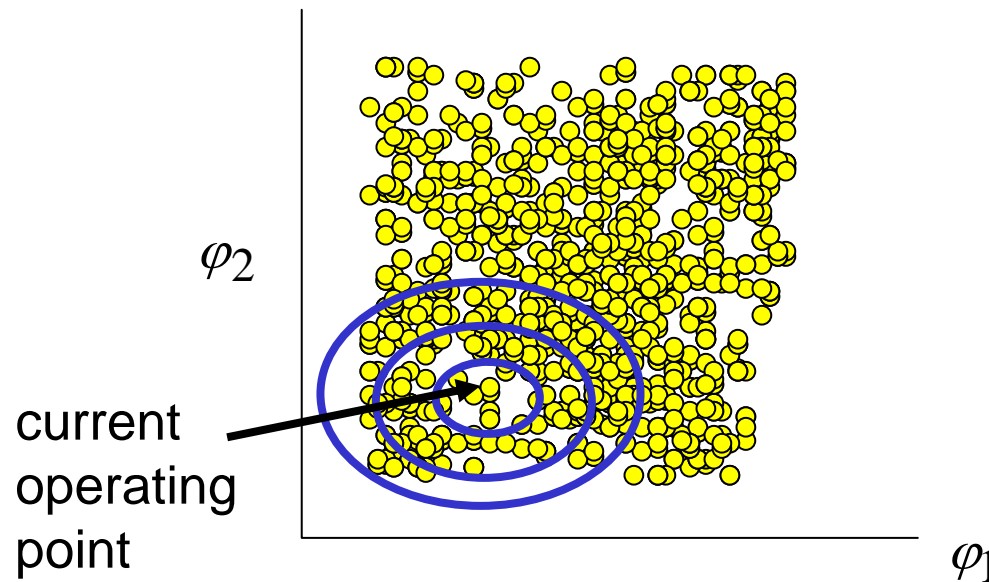
Input Weighting : [0.2 0.2]



# Model-on-Demand Estimation

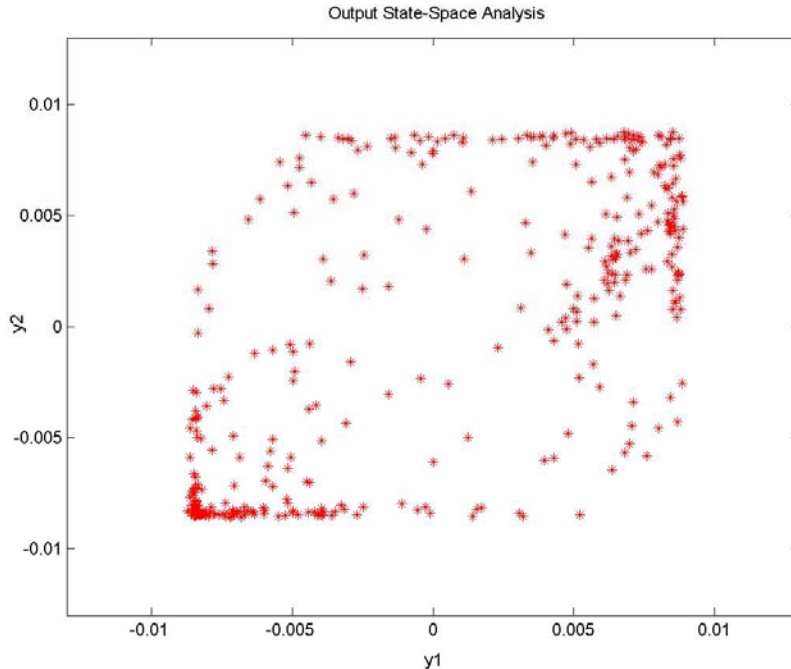
(Stenman, 1999)

- A modern data-centric approach developed at Linköping University
- Identification signals geared for MoD estimation should consider the geometrical distribution of data over the state-space.

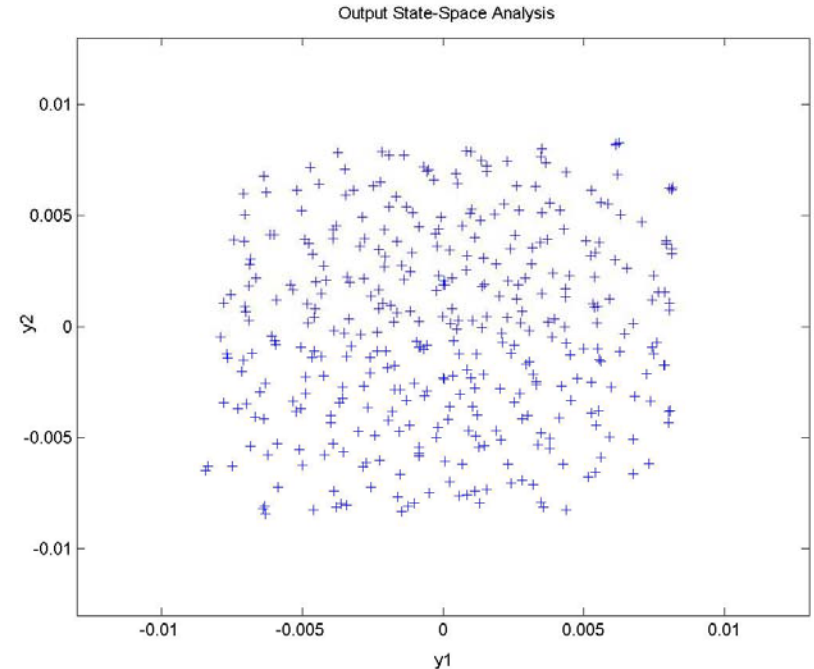


# min Crest Factor vs Weyl-based Signals: Output State-Space

Modified Zippered,  
min CF (y) Signal



Modified Zippered,  
Weyl-based signal



# Some Pertinent Questions

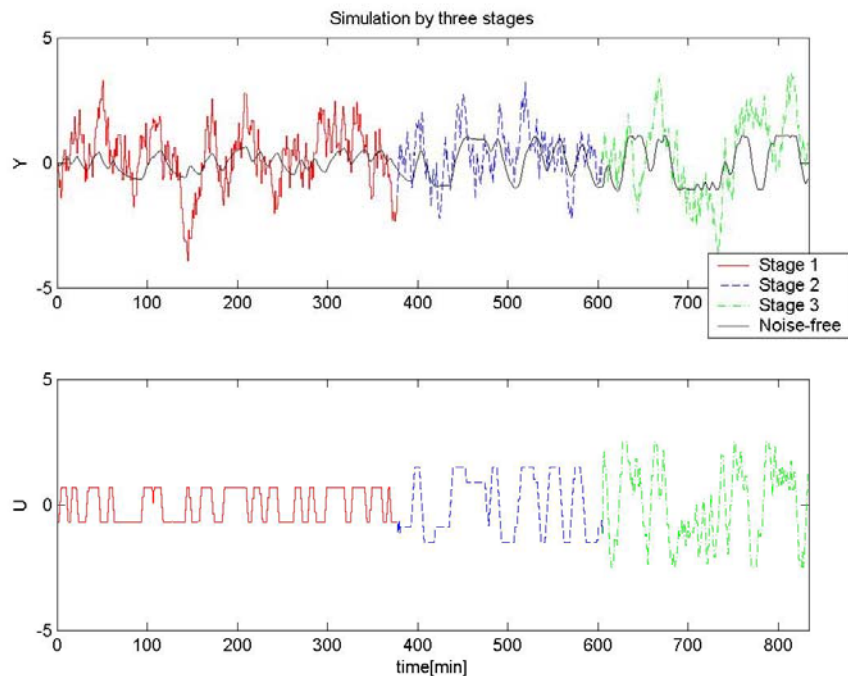
- How does one build process knowledge relevant to system identification in a systematic and (nearly) automatic way, with little user intervention and without demanding significant computational time and effort?
- How is process knowledge systematically acquired in the course of identification testing, for purposes of improving the identification test?

# Identification Test Monitoring

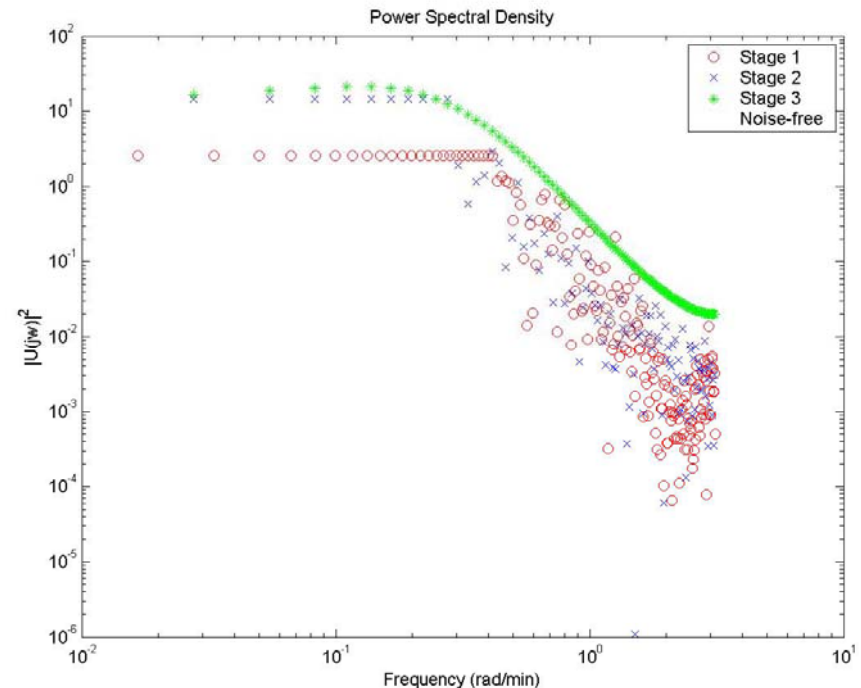
- Relies on the use of periodic, deterministic inputs (such as multisines or pseudo-random signals) to define a natural window for analysis,
- Relies on concepts from signal processing, robust control, and optimization to develop measures that systematically acquire and apply process knowledge, and use this knowledge to refine the design parameters of the identification test

# Identification Test Monitoring Scenario (from Rivera *et al.*, 2003)

## Time Series



## Input Power Spectral Density



*Input signal evolves from cautious to more informative as process knowledge increases*

# Summary and Conclusions

- “Plant-friendliness” in identification testing represents an important problem that despite advances in supporting topics (e.g, optimal input signal design, control-relevant identification, closed-loop identification) still merits focused research.
- Optimization-based design of multisine input signals can be used to achieve plant-friendliness during experimental testing for demanding process systems (such as high-purity distillation).
- *Identification Test Monitoring* has been proposed as a meaningful direction in the development of plant-friendly system id.



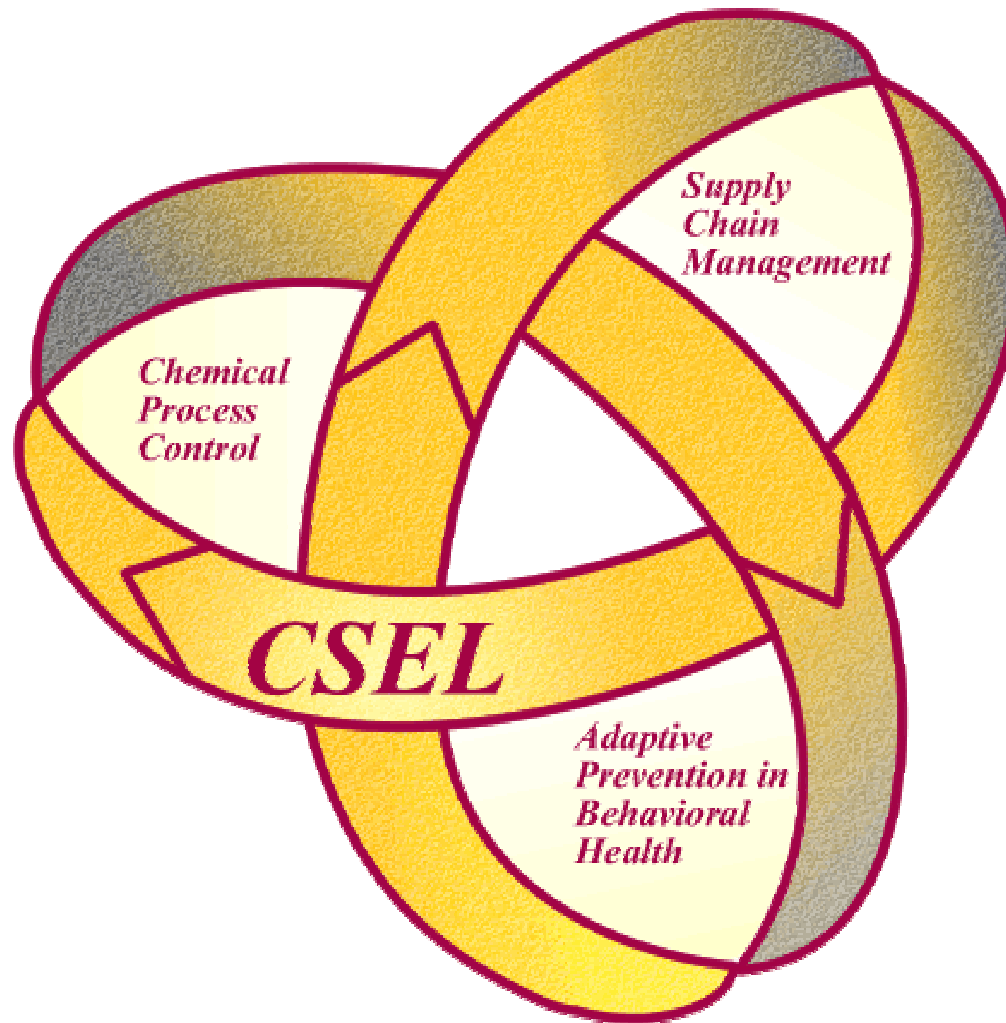
# Acknowledgements

- Hans D. Mittelmann, Gautam Pendse\*, Department of Mathematics and Statistics, Arizona State University
- Hyunjin Lee, Martin W. Braun\*\*, Department of Chemical and Materials Engineering, ASU
- Support from the American Chemical Society – Petroleum Research Fund, Grant No. ACS PRF#37610-AC9.

\* currently with McLean Hospital, Harvard Medical School, Boston.

\*\*currently with Intel Corp., Chandler, AZ.

<http://www.fulton.asu.edu/~csel>



# 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 performance. In: *Annual AIChE 1991 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.
- Byrd, R. M.E. Hribar and J. Nocedal (1999). An interior point method for large-scale nonlinear programming. *SIAM J. Optim.* **9**, 877-900.
- 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 base. *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.

# References (Continued)

- 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.R. 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. Kollar (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.
- Kothare, S. and J.A. Mandler (2003). Fast Plant Testing for MPC, In: *2003 AIChE Annual Meeting*. San Francisco, CA paper 254g.

# References (Continued)

- Lee, H., D.E. Rivera, and H. Mittelmann (2003a). A novel approach to plant-friendly multivariable identification of highly interactive systems. In: *Annual AIChE 2003 Meeting*. San Francisco, CA.
- Lee, H., D.E. Rivera, and H. Mittelmann (2003b). 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. paper255g.
- Li, T. and C. Georgakis (2003). Constrained signal design using approximate priori models with application to the Tennessee Eastman Process.
- L. Ljung (1997). Identification in closed-loop : some aspects on direct and indirect approaches. In: *SYSID 1997*, Fukuoka, Japan.
- L. Ljung (1999). *System Identification: Theory for the User*. 2<sup>nd</sup> ed.. Prentice-Hall, New Jersey.
- Mathur, U. and R.J. Conroy (2002). Multivariable control without plant tests. In: *2000 AIChE Annual Meeting*. Indianapolis, IN. paper 254g.

# References (Continued)

- Narasimhan, S. S. Rengaswamy and R. Rengasamy (2003). Multiobjective input signal design for plant-friendly identification. In: *SYSID 2003*. Rotterdam, The Netherlands.
- Narasimhan, S. and R. Rengaswamy (2004). Multi-objective input signal design for plant friendly identification of process systems. In: *American Control Conference*, Boston, MA, pp. 4891-4896.
- 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.
- Pendse, G.V. (2004). Optimization based formulations using the Weyl criterion for input signal design in system identification. Master's thesis. Arizona State University. Tempe, AZ, U.S.A.
- 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.

# References (Continued)

- Rivera, D.E., H. Lee, and H.D. Mittelmann (2003). "Plant-friendly" system identification: a challenge for the process industries," In: *SYSID 2003*. Rotterdam, The Netherlands.
- Srinivas, G.R., Y. Arkun, I-L. Chien, and B.A. Ogunnaike (1995). Nonlinear Identification and control of a high-purity distillation column: a case study. *J. Proc. Cont.* **5**, 149.
- A. Stenman (1999). Model on Demand: algorithms, analysis and applications. PhD thesis. Linköping University. Linköping, Sweden.
- H. Weyl. (1916) Über die gleichverteilung von zahlen mod eins. *Math. Ann.* **77**:313-352.