A Novel Approach to Plant-Friendly Multivariable Identification of Highly Interactive Systems

Hyunjin Lee, Daniel E. Rivera*
Control Systems Engineering Laboratory
Department of Chemical and Materials Engineering
Ira A. Fulton School of Engineering

Hans D. Mittelmann
Department of Mathematics and Statistics
College of Liberal Arts and Science
Arizona State University, Tempe, AZ 85287

AICHE Annual Meeting, November 16-21, 2003
San Francisco

* daniel.rivera@asu.edu  (480)-965-9476

Presentation Outline

• Multivariable System Identification using Multisine Signals
  ➢ Extension to highly interactive systems using modified “zippered” spectra
  ➢ Optimization-based formulations for minimum crest factor signals, conducive to “plant-friendliness”

• Case Study: High-Purity Distillation Column (Weischedel-McAvoy)
  ➢ Optimization-based design using an a priori ARX model
  ➢ Closed-loop evaluation of data effectiveness with MPC
  ➢ Extension to input signal design for nonlinear identification using NARX models

• Latest Efforts:
  ➢ Input signal design for data-centric estimation (such as MoD)
System Identification Challenges Associated with Highly Interactive Processes:

- Need to capture both low and high gain directions under noisy conditions
- Plant-friendliness must be achieved during identification testing

Some Solutions to the Highly Interactive Identification Problem

- Chien and Ogunnaike (1992 AIChE Mtg.) and Ogunnaike et al. (1993 ECC) use “high frequency” linear models and nonlinear empirical models, respectively.
- Li and Lee (1996 Comp. Chem. Eng) and Varga and Jorgensen (1994 AIChE) examine the problem using both open and closed-loop identification tests.
- Koung and MacGregor (1993 I&EC Res.) use correlated input signals based on a priori knowledge of high/low gain directions.
- Stec and Zhu (2001 ACC) and Butoyi and Zhu (2002 CEP) apply a sequential combination of correlated and uncorrelated signals of varying magnitudes to enhance the low gain information in the data.
Identifying Highly Interactive Systems –
(Stec and Zhu, 2001 ACC)

The sequential cycles of **correlated** and **uncorrelated** signals provide a mechanism for generating a data set with good information content in both high and low gain directions (e.g., tested with a simple model).

Multisine Input Signals

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

\[
u_j(k) = \sum_{i=1}^{m\delta} \delta_{ji} \cos(\omega_i kT + \phi_{ji}^a) + \sum_{i=m\delta+1}^{m(\delta+n_a)} \alpha_{ji} \cos(\omega_i kT + \phi_{ji}^a) + \sum_{i=m(\delta+n_a)+1}^{m(\delta+n_a+n_\alpha)} \tilde{a}_{ji} \cos(\omega_i kT + \phi_{ji}^a), \quad j = 1, \ldots, m
\]

where \(T\) is sampling time, \(N_a\) is the sequence length, \(m\) is the number of channels, \(\delta, n_a, n_\alpha\) are the numbers of sinusoids per channel \(m(\delta+n_a+n_\alpha) = N_a/2\), \(\phi_{ji}^a, \phi_{ji}^a, \phi_{ji}^a\) are the phase angles, \(\alpha_{ji}\) represents the Fourier coefficients defined by the user, \(\delta_{ji}\), \(\tilde{a}_{ji}\) are the "snow effect" Fourier coefficients.
Multisine Signal Design Guideline
(H. Lee, D. Rivera, H. Mittelmann, SYSID 2003)

For signal bandwidth denoted by \((\omega_0, \omega^*)\),  
\(n_s, N_s,\) and \(T\) must satisfy the inequalities:

\[
(1 + \delta) \frac{\omega^*}{\omega_0} \leq n_s \leq \frac{N_s}{2m}
\]

\[
T \leq \min\left\{ \frac{\pi}{\omega^*}, \frac{\pi}{\omega^* - \omega_0}\left(1 - \frac{1 + \delta}{n_s}\right) \right\}
\]

\[
\max\left\{ 2mn_s, \frac{2\pi mn(1 + \delta)}{\omega_0 T} \right\} \leq N_s \leq \frac{2\pi mn_s}{\omega^* T}
\]

Finally, design values (denoted by the superscript "d") should satisfy:

\[
\frac{(\omega^* - \omega_0)}{2\pi m} N^d_s T^d + (1 + \delta) \leq n^d_s \leq \frac{N^d_s}{2m}
\]

Modified Zippered Spectrum

Primary excitation frequency band

\(\hat{a}_u\)

\(2\pi mn(1 + \delta)\)

\(\omega_0\)

\(\omega^*\)

\(\frac{2\pi mn_s}{N_s T}\)

\(\pi\)

\(T\)

\(\gamma\)

Coefficients selected by optimizer

Correlated harmonic

Channel 1

Channel 2
Modified Zippered Spectrum Design

Utilize the steady-state gain matrix from a priori model:

\[
K = \begin{bmatrix}
    k_{11} & k_{12} \\
    k_{21} & k_{22}
\end{bmatrix}
\]

\[
\frac{\min(\frac{k_{11}^2, k_{12}^2, k_{21}^2, k_{22}^2}{(k_{11} + k_{12})^2, (k_{21} + k_{22})^2})}{\max(\frac{k_{11}^2, k_{12}^2, k_{21}^2, k_{22}^2}{(k_{11} + k_{12})^2, (k_{21} + k_{22})^2})} \leq \gamma^2 \leq \frac{\max(\frac{k_{11}^2, k_{12}^2, k_{21}^2, k_{22}^2}{(k_{11} + k_{12})^2, (k_{21} + k_{22})^2})}{\min(\frac{k_{11}^2, k_{12}^2, k_{21}^2, k_{22}^2}{(k_{11} + k_{12})^2, (k_{21} + k_{22})^2})}
\]

For Weischedel-McAvoy Case Study: \(\{\gamma\} = \{10.32 \leq \gamma \leq 15.67\}\)

Plant-Friendly Identification Testing

- A plant-friendly input signal should:
  - be as short as possible
  - not take actuators to limits, or exceed move size restrictions
  - cause minimum disruption to the controlled variables (i.e., low variance, small deviations from setpoints)
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 input sequence are located near the min. and max. values of the sequence.

**Problem Statement #1**

$$\min_{\{\phi^{a}_{ji}, \{\phi^{a}_{ji}\}, \{\phi^{d}_{ji}\}, \{\hat{a}_{ji}\}, \{\hat{d}_{ji}\}} \max_j \text{CF}(u_j) \quad j = 1, \ldots, m$$

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

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

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

$$u_j^{\text{min}} \leq u_j(k) \leq u_j^{\text{max}} \quad \forall \, k, j$$
Problem Statement #2

\[
\begin{align*}
\min_{\{\phi_j^a\}, \{\phi_j^b\}, \{\phi_j^c\}, \{\phi_j^d\}, \{\phi_j^e\}} & \quad \max_{z} \quad CF(y_z) \\
\end{align*}
\]

subject to constraints in input

\[
\begin{align*}
|\Delta u_j(k)| & \leq \Delta u_{j}^{\max} \quad \forall \ k, j \\
u_j^{\min} & \leq u_j(k) \leq u_j^{\max} \quad \forall \ k, j \\
\end{align*}
\]

and output

\[
\begin{align*}
|\Delta y_z(k)| & \leq \Delta y_z^{\max} \quad \forall \ k, z \\
y_z^{\min} & \leq y_z(k) \leq y_z^{\max} \quad \forall \ k, z \\
\end{align*}
\]

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

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.
Case Study: High-Purity Distillation

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.

Fig. 2. Two-product distillation column.

Standard & Modified Zippered Spectrum Design

Standard Zippered Spectrum

Modified Zippered Spectrum

For $\tau_r^{L}$ = 5, $\tau_r^{H}$ = 20 min, $\delta$ = 0, $\alpha_N$ = 2, and $\beta_N$ = 3, feasible design choices are $T$ = 2 min, $n_s$ = 25, $N_s$ = 378, and $\gamma$ = 15.
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

Standard & Modified Zippered Spectrum Design

Input State-Space

Output State-Space

Output Power Density
min CF signal design: time-domain

- min CF(\(u\)) signal with Standard Zippered Spectrum
- min CF(\(y\)) signal with ARX model and Modified Zippered Spectrum

Noise SNR [-0.04, -1.12]dB
Noise SNR [-5.0, -5.0]dB

Case Study: High-Purity Distillation

min CF Signal Design: Test signals statistics

<table>
<thead>
<tr>
<th>Type</th>
<th>Signal (a)</th>
<th>CF(a)</th>
<th>PIPS(%)</th>
<th>max((\Delta x))</th>
<th>max(x)</th>
<th>min(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>min CF((u)) design: standard zippered spectrum</td>
<td>(u_1)</td>
<td>1.227516</td>
<td>81.485337</td>
<td>0.002737</td>
<td>0.002000</td>
<td>-0.002000</td>
</tr>
<tr>
<td></td>
<td>(u_2)</td>
<td>1.227516</td>
<td>81.485337</td>
<td>0.002277</td>
<td>0.002000</td>
<td>-0.002000</td>
</tr>
<tr>
<td></td>
<td>(y_1)</td>
<td>2.524868</td>
<td>44.969514</td>
<td>0.003417</td>
<td>0.003565</td>
<td>-0.004796</td>
</tr>
<tr>
<td></td>
<td>(y_2)</td>
<td>2.531872</td>
<td>43.891194</td>
<td>0.004876</td>
<td>0.020027</td>
<td>-0.025283</td>
</tr>
</tbody>
</table>

- min CF(\(y\)) design: modified zippered spectrum by ARX model \([\Delta t] \leq 0.0075\) and \(|\Delta a| \leq 0.01\)

| | \(u_1\) | 2.900540 | 43.027738 | 0.009979 | 0.019783 | -0.024907 |
| | \(u_2\) | 2.683625 | 40.314130 | 0.009999 | 0.021987 | -0.025803 |
| | \(y_1\) | 1.607535 | 61.615468 | 0.004484 | 0.011356 | -0.012208 |
| | \(y_2\) | 1.945124 | 58.748109 | 0.007131 | 0.009280 | -0.012428 |

- min CF(\(y\)) design: modified zippered spectrum by NARX model \([\Delta t] \leq 0.0075\) and \(|\Delta a| \leq 0.01\)

| | \(u_1\) | 2.676489 | 37.697322 | 0.009999 | 0.025999 | -0.025734 |
| | \(u_2\) | 2.834289 | 35.288221 | 0.010000 | 0.025969 | -0.027428 |
| | \(y_1\) | 1.348449 | 74.850385 | 0.005174 | 0.008878 | -0.008709 |
| | \(y_2\) | 1.341305 | 75.176406 | 0.007500 | 0.008769 | -0.008606 |

- min CF(\(y\)) design: modified zippered spectrum with even harmonic suppression by ARX model \([\Delta t] \leq 0.007, |\Delta a| \leq 0.0075\) and \(|\Delta a| \leq 0.01\)

| | \(u_1\) | 2.902927 | 34.447985 | 0.009600 | 0.019552 | -0.019552 |
| | \(u_2\) | 2.537524 | 39.100325 | 0.010000 | 0.013227 | -0.013227 |
| | \(y_1\) | 1.607000 | 64.257126 | 0.003982 | 0.009401 | -0.008804 |
| | \(y_2\) | 1.674822 | 62.556892 | 0.005353 | 0.007887 | -0.008455 |
Closed-loop Performance Comparison using MPC Setpoint Tracking: models obtained from noise-free data

MPC Tuning Parameters:
- Prediction Horizon PHOR: 100
- Move Horizon: 25
- Output Weighting: [1 1]
- Input Weighting: [0.2 0.2]

Closed-loop Performance Comparison using MPC Setpoint Tracking: models obtained from noisy data conditions

MPC Tuning Parameters:
- Prediction Horizon PHOR: 100
- Move Horizon: 25
- Output Weighting: [1 1]
- Input Weighting: [0.2 0.2]
Rely on a NARX model equation to predict the system outputs during optimization:

\[
y(k) = \theta_0^{(0)} + \sum_{i=1}^{n_y} \theta_1^{(i)} y(k-i) + \sum_{i=1}^{n_y} \theta_2^{(i)} u(k-i) + \sum_{i=1}^{n_y} \sum_{j=1}^{d} \theta_{i,j}^{(2)} y(k-i)y(k-j) + \sum_{i=1}^{n_y} \sum_{j=1}^{d} \theta_{i,j}^{(4)} u(k-i)u(k-j) + \sum_{i=1}^{n_y} \sum_{j=1}^{d} \theta_{i,j}^{(5)} y(k-i)u(k-j) + \ldots
\]

Evaluation criterion (Srinivas et al., 1995):

\[
I = \frac{\sum_{k=1}^{N}[y(k) - \hat{y}(k)]^2}{\sum_{k=1}^{N}[y(k) - \bar{y}(k)]^2} \times 100\%
\]
ARX vs. NARX Model Predictions

ARX Model

NARX Model

+ (blue) : Model Prediction
* (red) : Weischedel-McAvoy Distillation Simulation

Model-on-Demand Estimation
(Stenman, 1999)

- A modern data-centric approach developed at Linkoping University
- Identification signals geared for MoD estimation should consider the geometrical distribution of data over the state-space.
Data-Centric Output Distribution Approach

Modified Zippered, min CF (\( y \)) Signal

Modified Zippered, Data Centric Signal

Optimize the magnitudes of the correlated harmonics to obtain evenly distributed outputs signals in the state-space.
Summary and Conclusions

- A comprehensive multisine signal design applicable to the identification of highly interactive systems has been presented.
- A modified zippered spectrum design is proposed for highly interactive systems, which combined with constrained optimization leading to informative data under “plant-friendly” operation.
- Models estimated from modified zippered signals are more effective under noisy conditions compared to standard designs.
- NARX model estimation leads to less distortion in the model output predictions for the nonlinear Weischedel-McAvoy column.
- The effective use of a priori knowledge is critical in the solution of this (or any other) control-relevant, plant-friendly identification problem.

Acknowledgements

This research has been supported by the American Chemical Society
- Petroleum Research Fund,
  Grant No. ACS PRF#37610-AC9.