Input Design for Model Order Determination in Subspace Identification

Pratik Misra and Michael Nikolaou
Chemical Engineering, University of Houston, Houston, TX 77204

Subspace identification methods require that the inputs to the process being identified be persistently exciting. This may be inadequate for subspace identification of ill-conditioned multivariable processes, because the process model order may be underestimated, leading to subsequent identification of poor models. To remedy the problem, it is proposed that inputs must be used that excite a process to be identified in a way that produces as uncorrelated process outputs as possible. This can be accomplished either in open or in closed-loop fashion. Simulations on a high-purity distillation column and on a heat exchanger illustrate the merit of the proposed approach.

Introduction

Over the past decade, subspace-based identification (SI) methods have emerged as an important tool for identification of linear multivariable dynamic systems. These methods are numerically reliable and computationally less cumbersome than prediction-error-based methods. Overviews of these methods can be found in a number of publications, such as Viberg (1995), van der Veen et al. (1993), and Ljung (1999). The identification problem considered in subspace identification methods is that of finding, up to a similarity transformation, the matrices A, B, C, and D (as well as K in combined deterministic–stochastic identification) of the system

$$x(k+1) = Ax(k) + Bu(k) + Ke(k)$$
$$y(k) = Cx(k) + Du(k) + e(k) \qquad (1)$$

with state $x \in \Re^n$, input $u \in \Re^m$, output $y \in \Re^p$, and white noise $e \in \Re^p$, given observed inputs u and outputs y. Larimore (1990), Van Overschee and De Moor (1994), and Verhaegen (1994) have solved this problem in different ways. Van Overschee and De Moor (1995) have shown that these different methods are essentially similar, that is, they use the same subspace, weighted by different matrices. Shi and MacGregor (2000) have further elucidated these ideas, by identifying that SI methods essentially perform reduced-rank least squares. By doing so, Shi and MacGregor (2000) also indicated how different weights can enforce certain desired model properties, such as incorporation of prior knowledge.

In all SI approaches, the order n of the state space (Eq. 1) is determined by performing singular-value decomposition (SVD) on a data matrix (Eq. 7 in the sequel) and enumerating the number of "large" singular values, which is equal to the nonzero singular values of the same matrix in the absence of noise. The remaining singular values, which are "small," correspond to what would be zero singular values in the absence of noise. In practical identification problems it may be difficult to distinguish between "large" and "small" singular values of the data matrix. This problem is exacerbated in the case of ill-conditioned systems, where the smallest nonzero singular value of the noiseless data matrix is usually very small, making it hard to distinguish which "small" singular values would be nonzero and which ones would be zero in the absence of noise. This may lead to the wrong system order and to subsequent identification of a system model that may be seriously problematic in controller design, as illustrated later in the article.

In this work we propose to address the problem just described by appropriately designing inputs for a process to be identified by SI experiments. We demonstrate that appropriately designed inputs, which excite the system along preferred directions, increase the difference between nonzero singular values due to signal and those due to noise, and make the task of order determination easier, resulting in better subsequent model identification.

Correspondence concerning this article should be addressed to M. Nikolaou.

Subspace Identification

Because SI methods are described in the references given previously, we are only going to refer here to elements of SI methods that are relevant to the focus of this work.

Assume that the linear, time-invariant system to be identified is described by Eq. 1 with $K = 0$. Define the vector of stacked outputs

$$\boldsymbol{y}_\alpha(k) = \left[y^T(k), y^T(k+1), \ldots, y^T(k+\alpha-1) \right]^T, \quad (2)$$

where α is a parameter required to be greater than the system order n. In practice α is chosen to be sufficiently larger than the expected system order n. Vectors of stacked inputs $\boldsymbol{u}_\alpha(k)$ and innovations $\boldsymbol{e}_\alpha(k)$ can be similarly defined.

Suppose $N + \alpha + \beta - 1$ data samples are available and construct the following Hankel data matrices

$$\boldsymbol{Y}_\alpha = \left[\boldsymbol{y}_\alpha(1+\beta), \ldots, \boldsymbol{y}_\alpha(N+\beta) \right] \quad (3)$$

$$\boldsymbol{Y}_\beta = \left[\boldsymbol{y}_\beta(1), \ldots, \boldsymbol{y}_\beta(N) \right] \quad (4)$$

The quantity β is user-defined and can be chosen such that $\alpha - \beta > n$, as it will be the number of rows in the estimated extended observability matrix $\hat{\Gamma}_\alpha$. The matrices $\boldsymbol{Y}_\alpha$ and $\boldsymbol{Y}_\beta$ are sometimes loosely referred to as *future* and *past* matrices. Matrices $\boldsymbol{U}_\alpha$ and $\boldsymbol{U}_\beta$ can be defined in a similar way. Then, it can be shown that an estimate of the extended observability matrix

$$\boldsymbol{\Gamma}_\alpha = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{\alpha-1} \end{bmatrix} \quad (5)$$

can be obtained as

$$\hat{\boldsymbol{\Gamma}}_\alpha = \hat{\boldsymbol{Q}}_s, \quad (6)$$

where the matrix $\hat{\boldsymbol{Q}}_s$ is obtained from an SVD of the matrix $(1/N)\boldsymbol{Y}_\alpha \Pi_{\boldsymbol{U}_\alpha^T}^{\perp} \boldsymbol{U}_\beta^T$ as

$$\frac{1}{N}\boldsymbol{Y}_\alpha \Pi_{\boldsymbol{U}_\alpha^T}^{\perp} \boldsymbol{U}_\beta^T = \underbrace{\begin{bmatrix} \hat{\boldsymbol{Q}}_s & \hat{\boldsymbol{Q}}_n \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{S}}_s & 0 \\ 0 & \hat{\boldsymbol{S}}_n \end{bmatrix}}_{S} \begin{bmatrix} \hat{\boldsymbol{V}}_s^T \\ \hat{\boldsymbol{V}}_n^T \end{bmatrix}, \quad (7)$$

where

$$\Pi_{\boldsymbol{U}_\alpha^T}^{\perp} = \boldsymbol{I} - \boldsymbol{U}_\alpha^T \left(\boldsymbol{U}_\alpha \boldsymbol{U}_\alpha^T \right)^{-1} \boldsymbol{U}_\alpha \quad (8)$$

is a projection matrix, and the partition of the matrix of singular values, $\boldsymbol{S}$, in Eq. 7, is performed so that the matrix of "large" singular values, $\hat{\boldsymbol{S}}_s$, corresponding to the signal, is separated from the matrix of "small" (essentially zero) singular values, $\hat{\boldsymbol{S}}_n$, corresponding to noise.

The system matrices A and C can then be determined using the shift invariance property of the extended observability matrix $\boldsymbol{\Gamma}_\alpha$.

Finally, the matrices B and D can be determined by solving the overdetermined system of linear equations

$$\frac{1}{N}\hat{\boldsymbol{Q}}_n^T \boldsymbol{Y}_\alpha \boldsymbol{U}_\alpha^T \left(\boldsymbol{U}_\alpha \boldsymbol{U}_\alpha^T \right)^{-1} \boldsymbol{U}_\beta^T = \frac{1}{N}\hat{\boldsymbol{Q}}_n^T \Phi_\alpha \boldsymbol{U}_\beta^T \quad (9)$$

in the least-squares sense.

The SI algorithm shown earlier is one illustrative example of SI methods. The variant used in this work is the combined robust algorithm from Van Overschee and De Moor (1996), shown by them to work well with several practical data sets. In this algorithm the crucial order-revealing SVD step is performed on the matrix

$$\boldsymbol{M} \hat{=} \boldsymbol{Y}_\alpha /_{\boldsymbol{U}_\alpha} \boldsymbol{W}_\beta \boldsymbol{\Pi}_{\boldsymbol{U}_\alpha^T}^{\perp} \hat{=} \boldsymbol{Y}_\alpha \boldsymbol{\Pi}_{\boldsymbol{U}_\alpha^T}^{\perp} \left[\left[\boldsymbol{W}_\beta \boldsymbol{\Pi}_{\boldsymbol{U}_\alpha^T}^{\perp} \right]^\dagger \left[\boldsymbol{W}_\beta \boldsymbol{\Pi}_{\boldsymbol{U}_\alpha^T}^{\perp} \right] \right], \quad (10)$$

where

$$\boldsymbol{Y}_\alpha /_{\boldsymbol{U}_\alpha} \boldsymbol{W}_\beta \hat{=} \left[\boldsymbol{Y}_\alpha \boldsymbol{\Pi}_{\boldsymbol{U}_\alpha^T}^{\perp} \right] \left[\boldsymbol{W}_\beta \boldsymbol{\Pi}_{\boldsymbol{U}_\alpha^T}^{\perp} \right]^\dagger \boldsymbol{W}_\beta$$

$$\boldsymbol{W}_\beta \hat{=} \begin{pmatrix} \boldsymbol{U}_\beta \\ \boldsymbol{Y}_\beta \end{pmatrix} \quad (11)$$

Note that both $\boldsymbol{\Pi}_{\boldsymbol{U}_\alpha^T}^{\perp}$ and $[[\boldsymbol{W}_\beta \boldsymbol{\Pi}_{\boldsymbol{U}_\alpha^T}^{\perp}]^\dagger [\boldsymbol{W}_\beta \boldsymbol{\Pi}_{\boldsymbol{U}_\alpha^T}^{\perp}]]$ are projection matrices.

Favoreel et al. (2000) have compared SI to the prediction error method (PEM) for a variety of industrial problems, and have identified advantages and disadvantages of each method. In contrast to PEM (Söderström and Stoica, 1989; Ljung, 1999), where many free parameters need to be selected, SI methods require selection of only one free parameter, namely, the order of the state space n. The order of the state space is suggested by the SVD in Eq. 7 (or Eq. 10). In the absence of noise, the rank of the matrix on the righthand side of Eq. 7 is exactly n, the system order. Thus, there are exactly n nonzero singular values in the SVD in Eq. 7, and the size of $\hat{\boldsymbol{Q}}_s$ is easily determined. In the presence of noise, however, the data matrix on the lefthand side of Eq. 7 becomes a full-rank matrix. The selection of the sizes of $\hat{\boldsymbol{S}}_s$ and $\hat{\boldsymbol{S}}_n$ then requires determining which singular values can be considered "small," hence, essentially zero, and which ones "large." If the noise level is not too high, there is usually a significant difference between noise and signal singular values, and their separation is easily achieved. However, for ill-conditioned systems, even a small magnitude of noise can make it very difficult to determine the system order correctly.

Note that for the matrix $\Pi_{\boldsymbol{U}_\alpha^T}^{\perp}$ in Eq. 8 to exist, the matrix $\boldsymbol{U}_\alpha \boldsymbol{U}_\alpha^T$ must be invertible. This is a frequent note in SI literature, followed by the comment that the invertibility of $\boldsymbol{U}_\alpha \boldsymbol{U}_\alpha^T$ is guaranteed if the input u is persistently exciting. We will demonstrate in the next sections that the persistent excitation requirement is not adequate (that is, it may lead to incorrect identification of the system order n) when SI is used with ill-conditioned systems, and that additional requirements must be posed on the input u for the state-space order to be correctly identified.

Motivating Example: High-Purity Distillation

Consider a high-purity distillation column in LV control configuration, that is, the reflux and vapor flow rates L and V are the manipulated variables, and the distillate and bot-

tom concentrations x_D and x_B are the controlled variables. Assume the column behaves as (Skogestad and Morari, 1987)

$$\begin{bmatrix} y_1(s) \\ y_2(s) \end{bmatrix} = \underbrace{\begin{bmatrix} 87.8/(1+194s) & -87.8/(1+194s)+1.4/(1+15s) \\ 108.2/(1+194s) & -108.2/(1+194s)-1.4/(1+15s) \end{bmatrix}}_{\boldsymbol{G}(s)} \begin{bmatrix} u_1(s) \\ u_2(s) \end{bmatrix} \tag{12}$$

in the Laplace domain, where the time constants are in minutes. An identification experiment (sampling period 1 min) using independent PRBS inputs with values ± 1.5 and measurements corrupted by Gaussian white noise with $(\mu, \sigma) = (0, 0.02)$ (see Figure 6) produces the singular value plot of Figure 1 for the matrix $\boldsymbol{M}$ in Eq. 10, suggesting that the system order is 1. Following that, the resulting identified transfer matrix is $\hat{\boldsymbol{G}}(z) = C(zI - A)^{-1}B + D$ with

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 0.9733 & (-0.1009 \quad 0.0985) \\ \begin{pmatrix} -0.2121 \\ -0.2636 \end{pmatrix} & \begin{pmatrix} 0.9068\times10^{-3} & 0.6804\times10^{-3} \\ -0.3708\times10^{-3} & -0.8309\times10^{-3} \end{pmatrix} \end{bmatrix}, \tag{13}$$

making a model-based multivariable control scheme unrealistic, because $C(zI - A)^{-1}B$ is singular and the inverse of $\hat{\boldsymbol{G}}(z)$ depends mainly on D, which is entirely due to noise. Note also that if D is set identically equal to the zero matrix, then the resulting $\hat{\boldsymbol{G}}(z) = C(zI - A)^{-1}B$ is singular; and, therefore, cannot be used in model-based control. Therefore, even though white process inputs were used in this identification experiment, the results are clearly unacceptable. If a model-predictive controller was designed on the basis of the model in Eq. 13, the performance of such a controller would be inadequate, as shown in Figure 2. The reason for poor identification is the ill-conditioning of the system matrix $\boldsymbol{G}(s)$ in Eq. 12, as discussed in the next section.

Ill-Conditioned Systems

Ill-conditioned processes are multivariable processes whose transfer function matrices have a high condition number at zero frequency (steady-state gain) or even at higher frequencies. Ill-conditioned processes are common in chemical-process industries, and include high-purity distillation columns (Skogestad and Morari, 1988) and heat exchangers (Jacobsen and Skogestad, 1993). Identification of ill-conditioned systems is usually challenging. When using SI methods to identify ill-conditioned processes, identifying the dimension of the signal subspace becomes difficult, as demonstrated in the preceding section, because some of the singular values of the matrix $\boldsymbol{M}$, Eq. 10 corresponding to the signal (or of the matrix $\hat{\boldsymbol{S}}_s$ in Eq. 7) become very small, and, thus, they can erroneously be considered to be identically zero. To illustrate this, consider a 2×2 process

$$\boldsymbol{y}(s) = \boldsymbol{G}(s)\boldsymbol{u}(s) \tag{14}$$

and let the SVD of $\boldsymbol{G}(s)$ be

$$\boldsymbol{G}(s) = \Upsilon(s)\Sigma(s)\Omega(s)^T \tag{15}$$

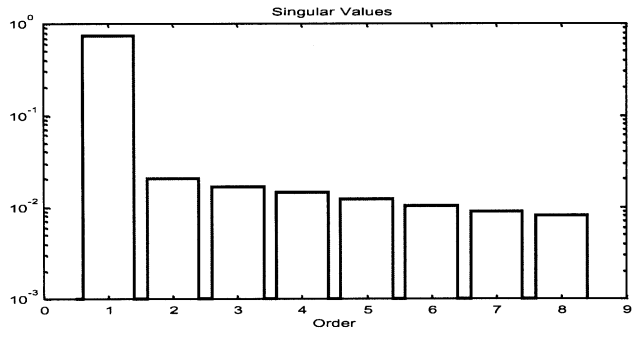


Figure 1. Singular-value plot for distillation column using random inputs.

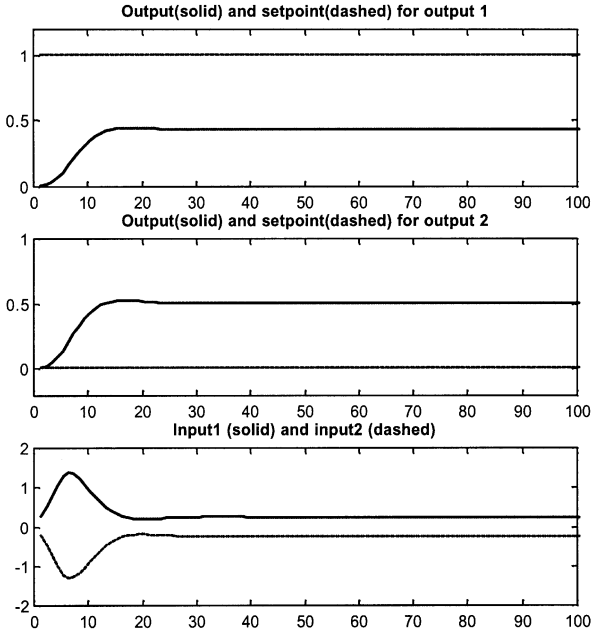


Figure 2. Outputs and inputs for distillation column using an MPC designed using first-order model.

where the 2×2 orthogonal matrices $\Upsilon(s)$ and $\Omega(s)$ can be written as

$$\Upsilon(s)=\begin{bmatrix}\cos\phi(s) & -\sin\phi(s)\\ \sin\phi(s) & \cos\phi(s)\end{bmatrix} \tag{16}$$

$$\Omega(s)^T=\begin{bmatrix}\cos\theta(s) & -\sin\theta(s)\\ \sin\theta(s) & \cos\theta(s)\end{bmatrix} \tag{17}$$

When $\boldsymbol{G}(s)$ is ill-conditioned, that is, the singular value $\sigma_1(s) \gg \sigma_2(s)$, the outputs can be written as (temporarily dropping the dependence on s for notational simplicity)

$$\begin{bmatrix}y_1\\ y_2\end{bmatrix}=\begin{bmatrix}\sigma_1\cos\phi(u_1\cos\theta-u_2\sin\theta)-\sigma_2\sin\phi(u_1\sin\theta+u_2\cos\theta)\\ \sigma_1\sin\phi(u_1\cos\theta-u_2\sin\theta)+\sigma_2\cos\phi(u_1\sin\theta+u_2\cos\theta)\end{bmatrix}\approx\begin{bmatrix}\sigma_1\cos\phi(u_1\cos\theta-u_2\sin\theta)\\ \sigma_1\sin\phi(u_1\cos\theta-u_2\sin\theta)\end{bmatrix}, \tag{18}$$

which implies that

$$y_2(s)\approx y_1(s)\tan\phi(s), \tag{19}$$

that is, the two outputs become almost collinear at all frequencies. This makes the columns of $\boldsymbol{Y}_\alpha$, Eq. 3, almost collinear, and, consequently, because $\prod_{\boldsymbol{U}_\alpha^T}^{\perp}$ and $[[\boldsymbol{W}_\beta\prod_{\boldsymbol{U}_\alpha^T}^{\perp}]^\dagger[\boldsymbol{W}_\beta\prod_{\boldsymbol{U}_\alpha^T}^{\perp}]]$ are projection matrices, the columns of $\boldsymbol{Y}_\alpha\prod_{\boldsymbol{U}_\alpha^T}^{\perp}$ in Eq. 7 or $\boldsymbol{Y}_\alpha\boldsymbol{\Pi}_{\boldsymbol{U}_\alpha^T}^{\perp}[[\boldsymbol{W}_\beta\prod_{\boldsymbol{U}_\alpha^T}^{\perp}]^\dagger[\boldsymbol{W}_\beta\prod_{\boldsymbol{U}_\alpha^T}^{\perp}]]$ in Eq. 10 are almost collinear as well. Thus, if the input matrix $\boldsymbol{U}_\beta^T$ is generated by white inputs $\boldsymbol{u}$, the matrix $(1/N)\boldsymbol{Y}_\alpha\prod_{\boldsymbol{U}_\alpha^T}^{\perp}\boldsymbol{U}_\beta^T$ or $\boldsymbol{Y}_\alpha\prod_{\boldsymbol{U}_\alpha^T}^{\perp}[[\boldsymbol{W}_\beta\prod_{\boldsymbol{U}_\alpha^T}^{\perp}]^\dagger[\boldsymbol{W}_\beta\prod_{\boldsymbol{U}_\alpha^T}^{\perp}]]$ will also have almost collinear columns, making it difficult to distinguish the smallest nonzero singular value corresponding to the signal from small nonzero singular values corresponding to noise.

It is instructive to recall that the projection matrices $\prod_{\boldsymbol{U}_\alpha^T}^{\perp}$ and $[[\boldsymbol{W}_\beta\prod_{\boldsymbol{U}_\alpha^T}^{\perp}]^\dagger[\boldsymbol{W}_\beta\prod_{\boldsymbol{U}_\alpha^T}^{\perp}]]$ have singular values equal to either 0 or 1, with the number of zero singular values being fixed as long as the matrices exist, that is, as long as $\boldsymbol{U}_\alpha\boldsymbol{U}_\alpha^T$ is invertible, even though $\boldsymbol{U}_\alpha\boldsymbol{U}_\alpha^T$ may be ill-conditioned. Thus the smallest signal singular value of $\boldsymbol{Y}_\alpha\prod_{\boldsymbol{U}_\alpha^T}^{\perp}[[\boldsymbol{W}_\beta\prod_{\boldsymbol{U}_\alpha^T}^{\perp}]^\dagger[\boldsymbol{W}_\beta\prod_{\boldsymbol{U}_\alpha^T}^{\perp}]]$ becomes very small, primarily because outputs are almost collinear. This explains the failure of white inputs to generate data that reveal the correct process order for the distillation column mentioned earlier. Indeed, the condition number of that column at steady state is about 140 and remains high over a large range of frequencies, as shown in Figure 3. Figure 3 also shows the element $\Omega_{11}(j\omega)$, Eq. 17, as a function of frequency.

For a higher-dimensional multivariable system, the same analysis can be applied. Namely, SVD of a higher-dimensional $\boldsymbol{G}(s)$ yields

$$\boldsymbol{G}(s)=\Upsilon(s)\Sigma(s)\Omega(s)^T=\sum_{i=1}^{r}\sigma_i(s)\psi_i(s)\omega_i(s)^T \approx\sum_{i=1}^{q}\sigma_i(s)\psi_i(s)\omega_i(s)^T, \tag{20}$$

where $r\triangleq rank[\boldsymbol{G}(s)]$ and $r-q$ nonzero singular values of the matrix $\boldsymbol{G}(s)$ are much smaller than the remaining nonzero singular values q, with $q<r$. Thus, Eq. 20 implies

$$\boldsymbol{y}(s)=\boldsymbol{G}(s)\boldsymbol{u}(s)=\sum_{i=1}^{r}\sigma_i(s)\psi_i(s)\underbrace{\omega_i(s)^T\boldsymbol{u}(s)}_{\xi_i(s)} \approx\sum_{i=1}^{q}\sigma_i(s)\psi_i(s)\underbrace{\omega_i(s)^T\boldsymbol{u}(s)}_{\xi_i(s)}\triangleq\sum_{i=1}^{q}\sigma_i(s)\psi_i(s)\xi_i(s) \tag{21}$$

indicating that the output would belong to a q-dimensional space, that is, it would create collinearity problems in the matrices $(1/N)\boldsymbol{Y}_\alpha\prod_{\boldsymbol{U}_\alpha^T}^{\perp}\boldsymbol{U}_\beta^T$ or $\boldsymbol{M}$, even if the input u were white.

Note that Eq. 21 defines a new input vector

$$\xi(s)\triangleq\Omega(s)^T\boldsymbol{u}(s) \tag{22}$$

corresponding to rotation of the original input vector $\boldsymbol{u}$.

For a 2×2 system, Eq. 22 combined with Eq. 17 yields

$$\xi_1(s)=u_1(s)\cos\theta(s)-u_2(s)\sin\theta(s) \tag{23}$$

$$\xi_2(s)=u_1(s)\sin\theta(s)+u_2(s)\cos\theta(s) \tag{24}$$

Assuming that $\Omega(s)\approx\Omega$ (not an unreasonable assumption as Figure 3b exemplifies), we get from the above two equations that

$$\xi_1(t)=u_1(t)\cos\theta-u_2(t)\sin\theta \tag{25}$$

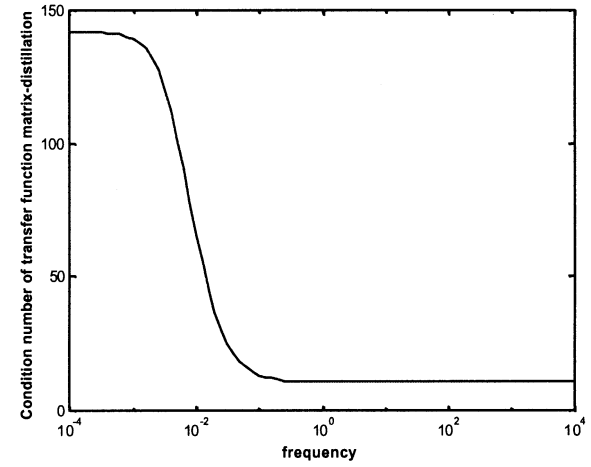
$$\xi_2(t)=u_1(t)\sin\theta+u_2(t)\cos\theta \tag{26}$$

As explained in the next section, properly designing the new inputs ξ can help alleviate the collinearity problem.

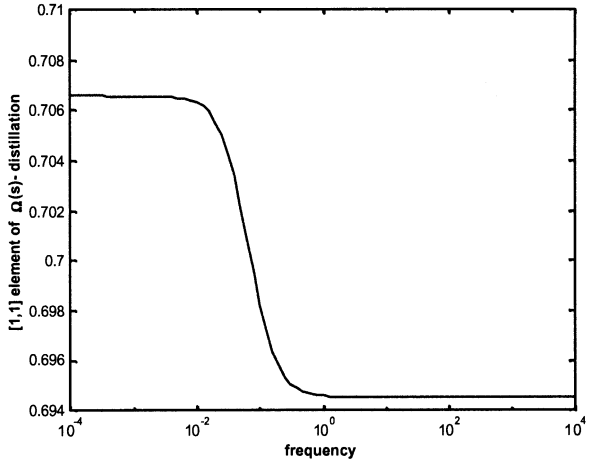
Input Design for Open-Loop SI of Ill-Conditioned Systems

The preceding discussion has established that the process outputs need to be made as uncorrelated as possible in order to determine the order of the state subspace more easily. Designing the inputs in a certain way can accomplish this. A method of input design that can make outputs uncorrelated was suggested by Koung and MacGregor (1994) for the problem of identification for robust control. The idea is directly suggested by Eq. 21: No term in the summation in Eq. 21 is negligible, and, hence, collinearity problems are avoided, if each term of the sum in Eq. 21 contributes equally to the magnitude of $\boldsymbol{y}$, that is, by Parseval's theorem, the terms $\int_0^\infty|\sigma_i(j\omega)|^2|\xi_i(j\omega)|^2\|\psi_j(j\omega)\|^2d\omega$ have the same value for all $i=1,\ldots,r$, or, because ψ_j are orthonormal

$$\int_0^\infty|\sigma_i(j\omega)|^2|\xi_i(j\omega)|^2d\omega=\text{constant} \tag{27}$$



(a)



(b)

Figure 3. (a) Condition-number $|\sigma_1(j\omega)/\sigma_2(j\omega)|$ of the distillation column transfer matrix; (b) [1,1] element of $\Omega(j\omega)$ matrix as a function of frequency.

for all $i = 1, \ldots, r$. This occurs if

$$|\sigma_i(j\omega)||\xi_i(j\omega)| = \text{constant for all } i = 1, \ldots, r \quad (28)$$

or

$$\frac{|\xi_\ell(j\omega)|}{|\xi_k(j\omega)|} = \frac{|\sigma_k(j\omega)|}{|\sigma_\ell(j\omega)|} \text{ for all } \ell, k. \quad (29)$$

The preceding ideas can be applied to the distillation column example mentioned earlier. For a 2×2 system, Eq. 29 implies that

$$\frac{|\xi_2(j\omega)|}{|\xi_1(j\omega)|} = \frac{|\sigma_1(j\omega)|}{|\sigma_2(j\omega)|} = \kappa(\omega) \quad (30)$$

Moreover, assuming that $\kappa(\omega) \approx \kappa$, we get from Eq. 30 that $\int_0^\infty |\xi_2(j\omega)|^2 d\omega \approx \kappa^2 \int_0^\infty |\xi_1(j\omega)|^2 d\omega$, which, by Parseval's theorem, implies $\int_0^\infty \xi_2(t)^2 dt \approx \kappa^2 \int_0^\infty \xi_1(t)^2 dt$ or

$$\frac{\|\xi_2\|_2}{\|\xi_1\|_2} \approx \kappa \quad (31)$$

where the preceding 2-norm is in the time domain.

Combined with Eq. 31, Eqs. 25 and 26 suggest that the inputs u_1 and u_2 can be independent signals satisfying all other properties for consistency of identification as long as the ratio of the magnitudes of the rotated signals ξ_1 and ξ_2 in Eqs. 25 and 26 is κ, as in Eq. 31.

Of course, because the process model $\boldsymbol{G}(s)$ is not known, neither Ω nor κ is known. However, if it is known that the process being identified is ill-conditioned, then knowledge of the exact value of κ is not crucial, for the following reason: Eq. 30 is satisfied if $\xi_2(t)/\xi_1(t) \approx \kappa$. Then, by Eqs. 25 and 26 it follows that

$$u_2(t) \approx \frac{\kappa\cos\theta - \sin\theta}{\cos\theta + \kappa\sin\theta} u_1(t) \quad (32)$$

For highly ill-conditioned processes $\kappa \gg 1$, therefore, Eq. 32 can be further simplified to

$$u_2(t) \approx \cot\theta\, u_1(t) \quad (33)$$

where the inputs u_1 and u_2 must be within their bounds. This leaves the rotation angle θ in Eqs. 25 and 26 as the single parameter to choose. This can be done experimentally by trial and error, as follows:

(1) Select a value for θ;

(2) Select the input u_1 as a two-level PRBS signal;

(3) Compute u_2 using Eq. 33, that is,

$$u_2(t) = \cot\theta\, u_1(t) + e(t), \quad (34)$$

where $e(t)$ is zero-mean noise with "small" amplitude. Simultaneously adjust the amplitude of u_1 and u_2 so that both are within bounds.

(4) Perform the identification and compute the singular values of the matrix $\boldsymbol{M}$, Eq. 10.

(5) Go to step 1, until maximum separation of a pair of singular values of $\boldsymbol{M}$ is observed.

The preceding kind of input design generates inputs that tend to be highly correlated, that is, almost collinear. Of course, inputs must not be *exactly* collinear, because that would make the matrix $\boldsymbol{U}_\alpha \boldsymbol{U}_\alpha^T$ in Eq. 8 noninvertible. The critical relation to be satisfied is Eq. 27. As long as the inputs are not exact linear multiples of each other, $\boldsymbol{\Pi}_{\boldsymbol{U}_\alpha^T}^{\perp}$ and $[\boldsymbol{W}_\beta \boldsymbol{\Pi}_{\boldsymbol{U}_\alpha^T}^{\perp}]^{\dagger}[\boldsymbol{W}_\beta \boldsymbol{\Pi}_{\boldsymbol{U}_\alpha^T}^{\perp}]$ can be computed. The rank properties of these matrices are independent of the degree of correlation in inputs, as they are both projection matrices.

For the processes of higher dimension, application of Eq. 28 is more involved, and inefficiencies of a combinatorial nature may be problematic when selecting input directions (Euler angles in Ω) by trial and error. To avoid such problems, the alternative of closed-loop identification can be attempted, as discussed in the sequel.

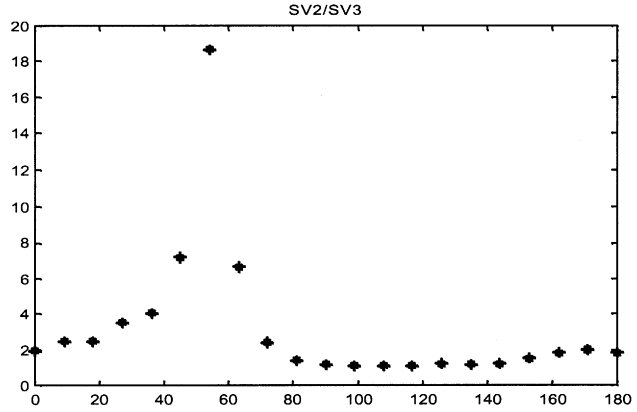


Figure 4. Ratio of second over third singular-value for rotated random inputs for distillation column.

Motivating Example Revisited

The type of input design suggested in the previous section is used on the distillation column example. The inputs u_1 and u_2 are rotated by different angles θ according to Eq. 34, where e is zero-mean Gaussian white noise with variance 0.01. The ratio of the second to third singular values of the resulting matrix $\boldsymbol{M}$, Eq. 10, is shown in Figure 4. It can be seen that, at about $\theta = 54°$, that ratio becomes high, allowing easy distinction between these singular values. Figure 5 shows the singular values of the matrix $\boldsymbol{M}$ for inputs rotated by that angle. It can easily be seen that the order of the system is 2. The difference between the second and third singular values in the case of random inputs (Figure 1) and rotated random inputs (Figure 5) can easily be seen. Figure 6 shows the input and output for random inputs, and Figure 7 shows process inputs and outputs for rotated random inputs. Note the difference between Figure 6 and Figure 7: In Figure 6 inputs are not correlated, but outputs are highly correlated. In contrast, in Figure 7 inputs are highly correlated, but outputs are much less correlated.

Heat-Exchanger Example

The same method of designing inputs is also used for a heat exchanger (Jacobsen and Skogestad, 1993). Figure 8 shows the singular-value plot for the heat exchanger using white PRBS as the input. In this case it can be seen that the system order would be determined as 1, which would be wrong, as the correct order is 2. Figure 9 shows the singular-value plot for rotated random inputs. It can easily be observed that in this case the order is identified as 2.

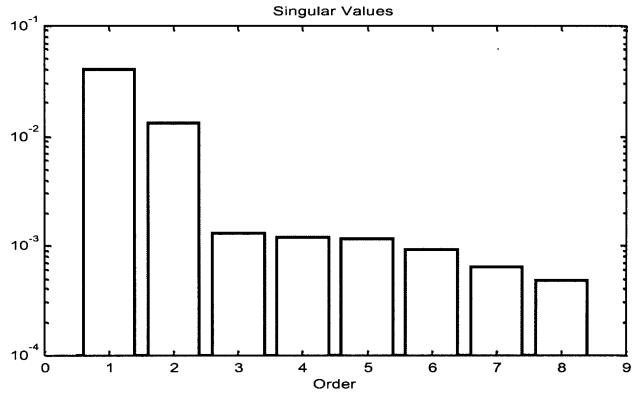
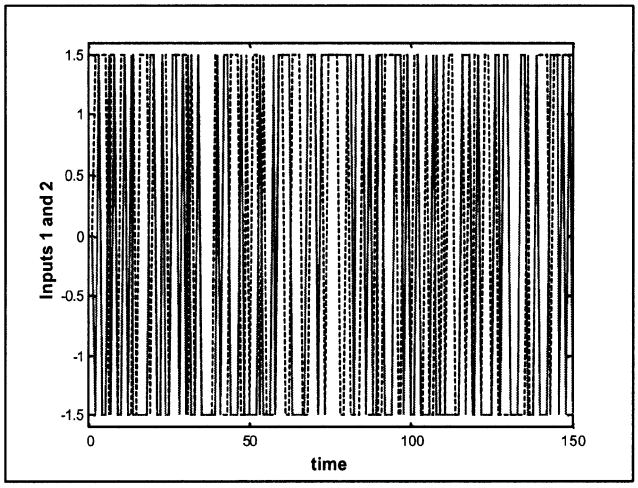
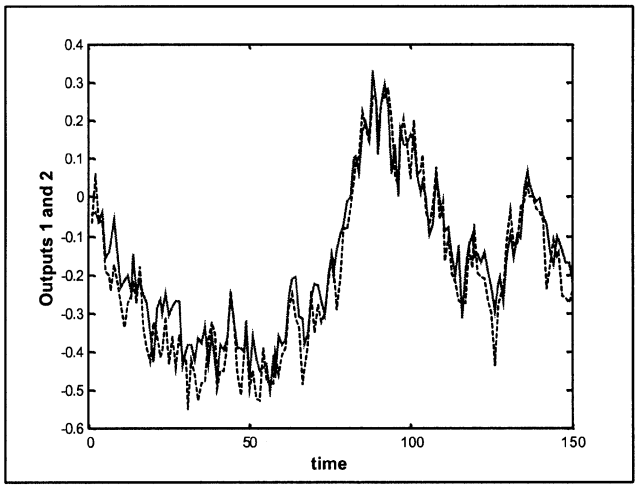


Figure 5. Singular-value plot for rotated random inputs for distillation column.



Inputs



Outputs

Figure 6. Inputs and outputs for random inputs for distillation column.

Input Design Using Closed-Loop Identification

The main objective of the open-loop input design method was to make outputs as uncorrelated as possible, to avoid making the smallest singular value of the matrix $\boldsymbol{M}$, Eq. 10, that corresponds to the signal, very small and practically indistinguishable from singular values corresponding to noise. This approach involves trial and error, because it needs good knowledge of the very system being identified, and may be overly cumbersome for high-dimensional systems, as discussed previously.

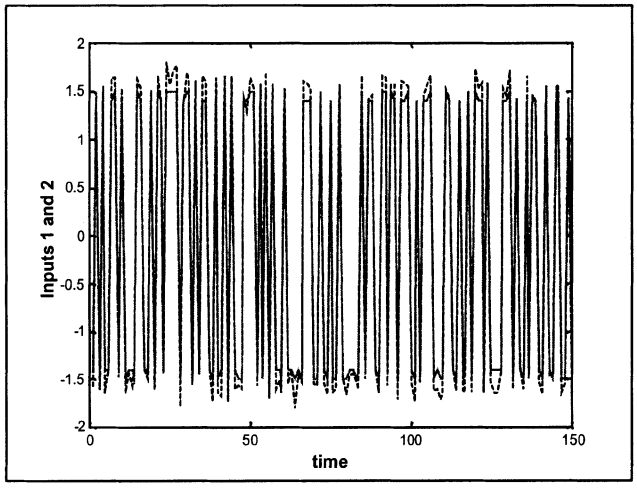


Figure 7. Inputs and outputs for rotated inputs for distillation column.

To avoid heavy reliance on the model of the process being identified or excessive experimentation for the generation of process inputs, one can let a feedback controller generate process inputs in closed-loop operation. Choi et al. (2000) have also used closed-loop identification for better identification of ill-conditioned processes. They use sequential closing of control loops and iterative identification to identify control-relevant models. In this work the entire system is being identified and all the loops are closed. The behavior of SI methods when applied to closed-loop identification has also been investigated by Verhaegen (1993). Our interest here lies mainly in understanding how closed-loop operation can help identify the correct process order. External excitation of the closed loop by orthogonal (uncorrelated) set point changes should generate approximately uncorrelated process outputs. Because feedback can tolerate model inaccuracies, the closed-loop approach can rely less on *a priori* model information than the open-loop approach. The obvious trade-off is that feedback can destabilize a stable process, a fact that can be addressed by using a conservatively tuned controller at the beginning of the experiment, and by making tuning more aggressive as modeling information becomes available. In addition, when identifying a stable process, constraints on process inputs can be used to ensure that inputs do not exceed desirable bounds.

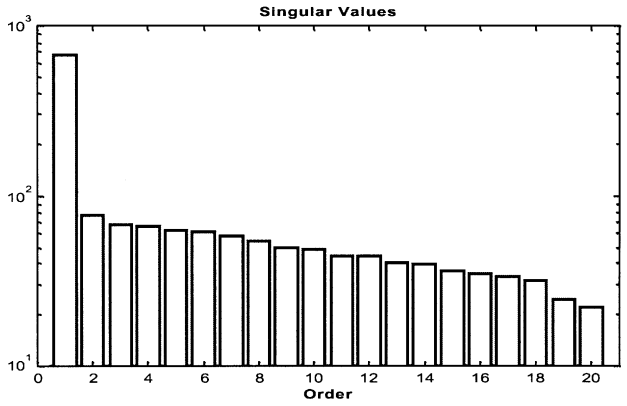


Figure 8. Singular-value plot for heat exchanger with random inputs.

Examples of Subspace Identification in Closed Loop

To conduct SI in closed loop, a preliminary controller is needed first. We use a model-predictive controller (MPC) with parameters shown in Table 1. Because MPC is based on the inverse of the process model, it cannot use the model of Eq. 13, of state-space dimension 1, because that model corresponds either to a nearly singular transfer matrix, or to a singular transfer matrix, when D is set to zero. Therefore, we develop a process model suitable for MPC by using SI with PRBS inputs in open loop, and with the assumption that the order of the system is 2. The inputs and the outputs to the distillation column are shown in Figure 10, and the resulting

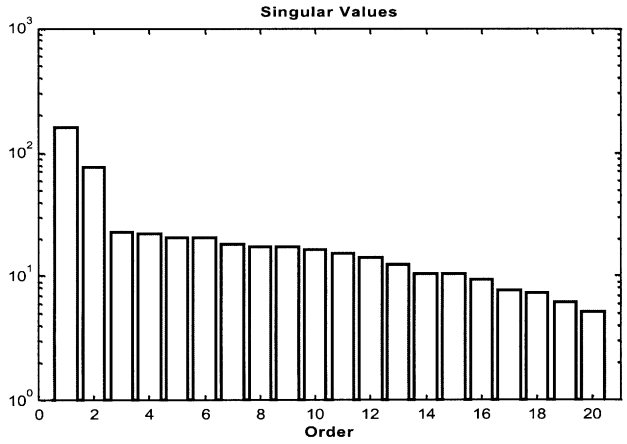


Figure 9. Singular-value plot for heat exchanger with rotated random inputs.

Table 1. MPC Parameters for Closed-Loop Experiments

Parameter	Value
Prediction horizon	20
Control horizon	7
Weights on outputs	[1, 1]
Weights on input	[0.5, 0.5]
$\boldsymbol{m}_{in}$	$[-4, -4]$
$\boldsymbol{m}_{ax}$	[4, 4]
$\Delta \boldsymbol{u}$ limits	[0.25, 0.25]

model for use with the preliminary MPC system is

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \left[\begin{array}{cc} \begin{pmatrix} 0.8043 & 0 \\ 0 & 0.8043 \end{pmatrix} & \begin{pmatrix} 0.1834 & -0.1528 \\ 0.2274 & -0.1897 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{array}\right] \tag{35}$$

The transfer matrix $\hat{\boldsymbol{G}}(z) = C(zI - A)^{-1}B + D$ corresponding to this model is not singular, but has a condition number of 3,629 at $\omega = 0$, which is quite high, that is, the distillation column model is ill-conditioned. This kind of model is used solely because we want to be able to get an initial controller, even a poorly functioning one, for closed-loop identification.

Next, orthogonal changes are made to the set points and the controller just described is used up to time 1,000. The outputs do not follow the set points well, as shown in Figure 10 for time up to 1,000.

Nevertheless, data up to time 1,000 allow identification of a good model for control purposes. The plot of singular values of $\boldsymbol{M}$ in Figure 11 shows that the process order is correctly identified as 2. The newly identified model

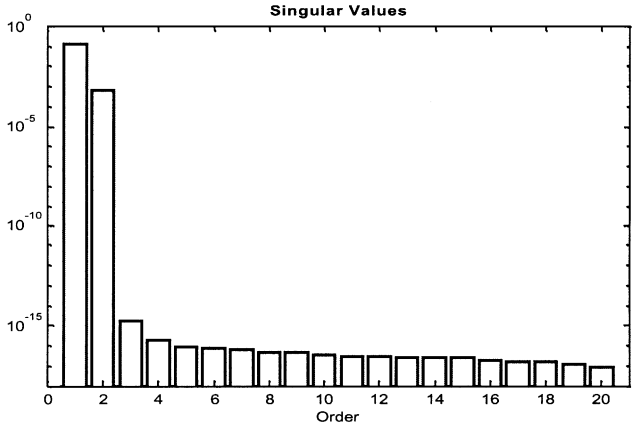


Figure 11. Singular-value plot for closed-loop identification for distillation column.

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \left[\begin{array}{cc} \begin{pmatrix} 0.9745 & 0.0008 \\ 0.0082 & 0.7166 \end{pmatrix} & \begin{pmatrix} 0.2434 & -0.2445 \\ 0.0078 & 0.3813 \end{pmatrix} \\ \begin{pmatrix} 0.0915 & 0.0105 \\ 0.1056 & -0.0099 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{array}\right] \tag{36}$$

is then used by MPC from time 1,000 on, with clear improvement in performance, as shown in Figure 10.

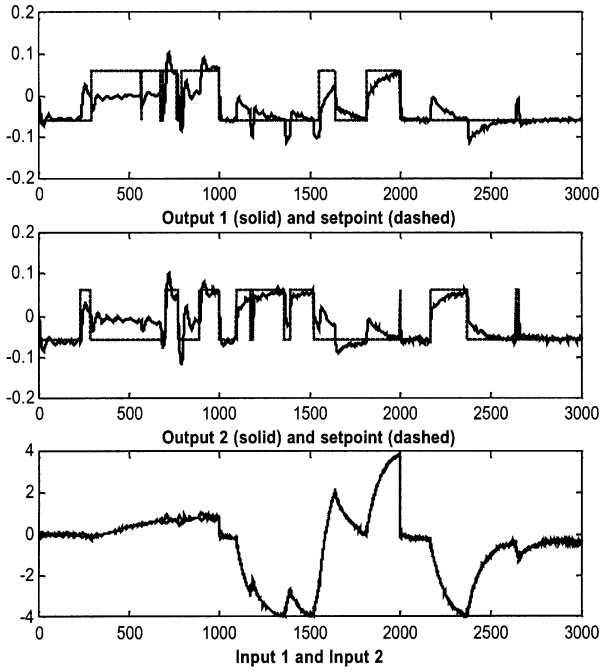


Figure 10. Inputs and outputs for closed-loop identification for distillation column.

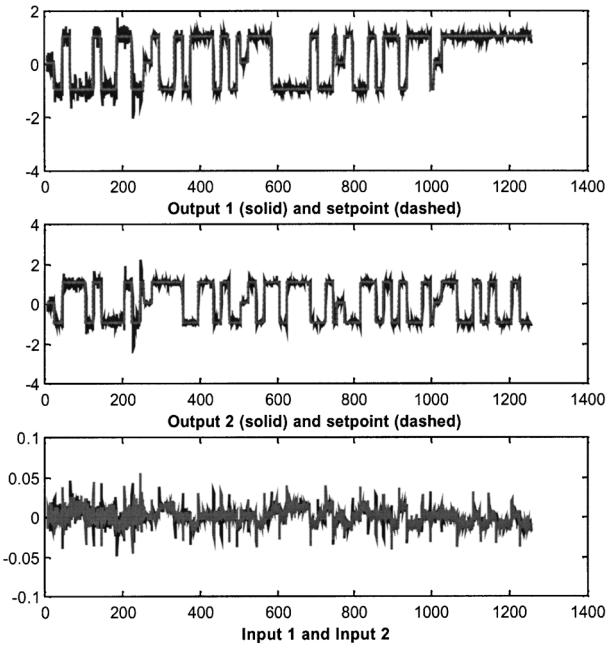


Figure 12. Inputs and outputs for closed-loop identification for heat exchanger.

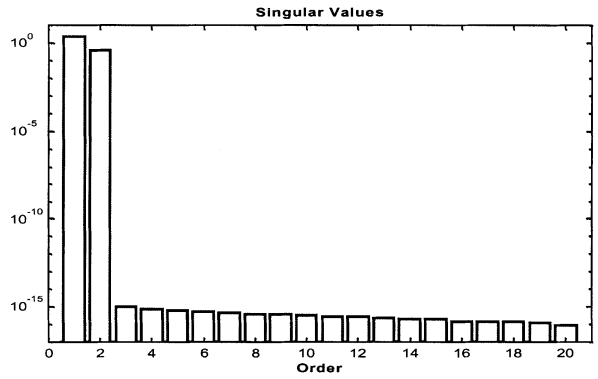


Figure 13. Singular-value plot for closed-loop identification for heat exchanger.

Similar results are obtained with the heat-exchanger example. Figure 12 shows the inputs and outputs, and Figure 13 shows the singular-value plot for the heat-exchanger example in closed loop. Note the improvement in performance after the newly identified model was employed by MPC after time 250.

Conclusion

In this work a way has been suggested to design inputs when subspace identification methods are used for ill-conditioned systems. Use of uncorrelated inputs, as usually suggested in the literature, does not produce a correct model order when the process being identified is ill-conditioned, and, thus, the resulting model may be problematic for use in model-based control. In order to overcome this problem, process inputs must be designed in such a way that process outputs are as uncorrelated as possible. This can be done in either open- or closed-loop fashion. In open-loop fashion, inputs can be selected after rotation by an angle determined by trial and error. For larger systems, SI in closed-loop fashion with orthogonal set point changes can be more efficient. Both of these approaches to input design were successfully applied to well-known examples of ill-conditioned systems from the literature.

Literature Cited

Choi, J. Y., J. Lee, and T. F. Edgar, "Use of the Sequential Loop Closing Method for Iterative Identification of Ill-Conditioned Processes," *Ind. Eng. Chem. Res.*, **39**, 2404 (2000).

Favoreel, W., B. DeMoor, and P. van Overschee, "Subspace State Space System Identification for Industrial Processes," *J. Process Control*, **10**, 149 (2000).

Jacobsen, E. W., and S. Skogestad, "Identification of Models for Ill-Conditioned Plants—A Bench Mark Problem," *Proc. European Control Conf.*, Vol. 1, Groningen, The Netherlands, p. 242 (1993).

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

Larimore, W. E., "Canonical Variate Analysis in Identification, Filtering and Adaptive Control," *Proc. IEEE Conf. on Decision and Control*, Honolulu, HI, 596 (1990).

Ljung, L., *System Identification: Theory for the User*, 2nd ed., Prentice Hall, Upper Saddle River, NJ (1999).

Shi, R., and J. F. MacGregor, "Modeling of Dynamic Systems Using Latent Variable and Subspace Methods," *J. Chemometrics*, **14**(5–6), 423 (2000).

Skogestad, S., and M. Morari, "Implications of Large RGA Elements on Control Performance," *Ind. Eng. Chem. Res.*, **26**, 2323 (1987).

Skogestad, S., and M. Morari, "LV-Control of a High Purity Distillation Column," *Chem. Eng. Sci.*, **43**, 33 (1988).

Söderström, T., and P. Stoica, *System Identification*, Prentice Hall, London (1989).

Van der Veen, A.-J., E. Deprettere, and A. Swindlehurst, "Subspace Based Signal Analysis Using Singular Value Decomposition," *Proc. IEEE*, **81**, 1277 (1993).

Van Overschee, P., and B. DeMoor, "N4SID: Subspace Algorithms for the Identification of Combined Deterministic-Stochastic Systems," *Automatica*, **30**, 75 (1994).

Van Overschee, P., and B. DeMoor, "A Unifying Theorem for Three Subspace System Identification Algorithms," *Automatica*, **31**, 1853 (1995).

Van Overschee, P., and B. DeMoor, *Subspace Identification for Linear Systems: Theory-Implementation-Applications*, Kluwer, Dordrecht, The Netherlands (1996).

Verhaegen, M., "Application of Subspace Model Identification Technique to Identify LTI Systems Operating in Closed-Loop," *Automatica*, **29**, 1027 (1993).

Verhaegen, M., "Identification of the Deterministic Part of MIMO State Space Models Given in Innovations Form From Input-Output Data," *Automatica*, **30**, 61 (1994).

Viberg, M., "Subspace-Based Methods for the Identification of Linear Time-Invariant Systems," *Automatica*, **31**, 1835 (1995).

Manuscript received Jan. 3, 2002, and revision received Feb. 4, 2003.