Data Compression for Subspace-Based Identification Using Periodic and Nonperiodic Inputs[1]

Islam I. Hussein
University of Michigan
ihussein@umich.edu

Seth L. Lacy
University of Michigan
sethlacy@engin.umich.edu

Dennis S. Bernstein
University of Michigan
dsbaero@umich.edu

1 Introduction

In time-domain identification there is a distinction between batch and recursive identification methods. In batch identification, all of the data is assumed to be available before identification begins, while in recursive identification, parameter estimates are updated as data become available. The classical SISO least squares technique can be implemented in either batch or recursive form; these forms are mathematically equivalent, that is, the final estimates obtained from the recursive form are exactly equal (ignoring numerical effects) to the batch estimates. However, the recursive form allows the inclusion of a forgetting factor, which does not have a counterpart in the batch form.

One advantage of recursive identification is the fact that there is no constraint on the amount of data that can be processed since additional data can be included ad infinitum. Unfortunately, not all identification algorithms can be recast in recursive form. For example, the subspace identification algorithms [1, 2, 3, 4] are batch identification algorithms, and recursive forms have not been developed. Consequently, the amount of data that can be processed by a subspace algorithm is limited by the memory requirements of the singular value decompositions used by the algorithm.

As in batch time-domain identification, frequency domain identification is also limited to small amounts of data. In practice, however, frequency domain identification is implemented repetitively, that is, the identification signal is generated by the spectrum analyzer multiple times, and the generated frequency spectra are successively averaged to produce successively higher quality estimates of the transfer function [5, 6, 7, 8]. In this way, the amount of data utilized by frequency domain identification methods is effectively quite large.

The objective of this paper is to develop techniques for implementing subspace identification algorithms that utilize larger amounts of data than can be handled by a single batch run of the algorithm. Our approach is to compress the available data, that is, to reduce a large data set to a smaller data set prior to implementing the algorithm. The motivation for this problem is the fact that the main calculation in subspace algorithms is a singular value decomposition. If the algorithm is to be implemented in, for instance, Matlab®, there is a memory requirement of approximately $3\times8\times lp\times N = 24lpN$ bytes of memory, where l is the number of outputs, p is the "window size" and N is the length of the data record. The factor 3 appears because the SVD matrices U, S and V may need to be stored simultaneously in the memory. For example, consider a machine with 1GB of RAM. If the plant has 6 outputs and we choose $p = 20$, we can process only $N = 2777$ data points. Of course, we must also consider the memory occupied by the operating system.

The idea of compressing identification data prior to batch processing is not new. In [9, 10, 11], for example, it is suggested that when periodic identification signals are used for identification, the data can be averaged prior to processing. This approach is roughly equivalent to repetitive frequency domain identification except that the averaging is performed before the algorithm is implemented rather than after.

In Section 2 we review a subspace identification algorithm. In Section 3 an averaging scheme is discussed. In Section 4 we analyze the effect of averaging input-output data when the input signal is periodic. In particular, we account for the transient and periodic steady responses, and analyze the signal-to-noise ratio that results from averaging. In Section 5 we discuss the effect of "slicing up" the data and we develop guidelines for data slicing. In Section 6, we consider data compression for non-periodic inputs.

2 Review of a Subspace Algorithm

Several approaches for identifying a discrete-time state space model of a system using subspace methods exist. Below, we describe one such approach [1].

Consider the state space model

$$x_{k+1} = Ax_k + Bu_k + w_k, \tag{1}$$

$$y_k = Cx_k + Du_k + v_k, \tag{2}$$

where $x_k \in \Re^n$, $y_k \in \Re^l$, $u_k \in \Re^m$, $A \in \Re^{n\times n}$, $B \in \Re^{n\times m}$, $C \in \Re^{l\times n}$, $D \in \Re^{l\times m}$, and with $w_k \in \Re^m$ and $v_k \in \Re^l$ zero mean, stationary, white noise vector

sequences with positive semidefinite covariance matrix

$$E\left[\begin{pmatrix} w_p \\ v_p \end{pmatrix} \cdot \begin{pmatrix} w_q & v_q \end{pmatrix}\right] = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \delta_{pq}, \quad (3)$$

$Q \in \Re^{n\times n}$, $S \in \Re^{n\times l}$, and $R \in \Re^{l\times l}$. The identification problem is: Given N input and output measurements $u_1, \ldots, u_N$ and $y_1, \ldots, y_N$, respectively, estimate the system order n and the system matrices A, B, C, D, Q, R and S.

Omitting noise terms, (2) can be written in the form

$$Y_{p_1,p_2} = \Gamma_p \tilde{X}_{p_1} + H_p U_{p_1,p_2}. \quad (4)$$

The output Hankel matrix $Y_{p_1,p_2} \in \Re^{l(p_2-p_1+1)\times N}$ is defined by

$$Y_{p_1,p_2} \triangleq \begin{bmatrix} y_{p_1} & y_{p_1+1} & \cdots & y_{p_1+N-1} \\ y_{p_1+1} & y_{p_1+2} & \cdots & y_{p_1+N} \\ \vdots & \vdots & \ddots & \vdots \\ y_{p_2} & y_{p_2+1} & \cdots & y_{p_2+N-1} \end{bmatrix} \quad (5)$$

and the input Hankel matrix $U_{p_1,p_2} \in \Re^{m(p_2-p_1+1)\times N}$ is defined analogously. The past input, past output, future input and future output Hankel matrices are defined by $U_- \triangleq U_{0,p-1}$, $Y_- \triangleq Y_{0,p-1}$, $U_+ \triangleq U_{p,2p-1}$ and $Y_+ \triangleq Y_{p,2p-1}$, respectively, where p is a user defined window length such that $1 < p < N$.

The extended observability matrix $\Gamma_p \in \Re^{lp\times n}$ is given by

$$\Gamma_p \triangleq \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{p-1} \end{bmatrix}. \quad (6)$$

The estimated state sequence $\tilde{X}_{p_1} \in \Re^{n\times N}$ is defined as

$$\tilde{X}_{p_1} \triangleq \begin{bmatrix} \tilde{x}_{p_1} & \cdots & \tilde{x}_{p_1+N-1} \end{bmatrix}, \quad (7)$$

where $\tilde{x}_k$ is an estimate of the state of the system at time step k. For $p_1 = 0$, we have $\tilde{X}_0 = \begin{bmatrix} \tilde{x}_0 & \tilde{x}_1 & \cdots & \tilde{x}_{N-1} \end{bmatrix}$. The lower triangular Toeplitz matrix $H_p \in \Re^{lp\times mp}$ is given by

$$H_p \triangleq \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{p-2}B & CA^{p-3}B & \cdots & D \end{bmatrix}. \quad (8)$$

For $G \in \Re^{s\times N}$, define

$$\Pi_G \triangleq G^T(GG^T)^\dagger G, \quad (9)$$

the orthogonal projection onto the row space of G, and

$$\Pi_{G^\perp} \triangleq I_N - \Pi_G, \quad (10)$$

the orthogonal projection onto the orthogonal complement of the row space of G, where I_N is the $N\times N$ identity matrix. Thus, the orthogonal projection of the row space of $F \in \Re^{r\times N}$ onto the row space of $G \in \Re^{s\times N}$ is given by

$$F/G \triangleq F\Pi_G \in \Re^{r\times N}. \quad (11)$$

The oblique projection of the row space of $F \in \Re^{r\times N}$ along the row space of $G \in \Re^{s\times N}$ onto the row space of $J \in \Re^{q\times N}$ is given by

$$F/_G J \triangleq F\begin{bmatrix} J^T & G^T \end{bmatrix}\begin{bmatrix} JJ^T & JG^T \\ GJ^T & GG^T \end{bmatrix}^\dagger J \in \Re^{r\times N}, \quad (12)$$

where only the first q columns of the pseudoinverse are used. Define the matrices W_-, Z_-, $\bar{Z}_-$ and $\bar{W}_-$ by

$$W_- \triangleq \begin{bmatrix} U_- \\ Y_- \end{bmatrix}, \; Z_- \triangleq Y_+/\begin{bmatrix} W_- \\ U_+ \end{bmatrix},$$
$$\bar{Z}_- \triangleq Y_{p+1,2p-1}/\begin{bmatrix} \bar{W}_- \\ U_{p+1,2p-1} \end{bmatrix}, \; \bar{W}_- \triangleq \begin{bmatrix} U_{0,p} \\ Y_{0,p} \end{bmatrix}.$$

The correlation between two signals a_k and e_k is given by

$$\mathbf{E}[a_k e_k^T] \triangleq \lim_{N\to\infty}\left[\frac{1}{N}\sum_{i=0}^{N} a_i e_i^T\right]. \quad (13)$$

Finally, the covariance $\Phi_{[F,G]} \in \Re^{r\times s}$ of $F \in \Re^{r\times N}$ and $G \in \Re^{s\times N}$ is defined as

$$\Phi_{[F,G]} \triangleq \mathbf{E}[FG^T]. \quad (14)$$

Since N is finite, $\Phi_{[F,G]}$ is approximated by

$$\Phi_{[F,G]} \simeq \frac{1}{N}FG^T. \quad (15)$$

The following assumptions will also be made:
(1) u_k is uncorrelated with the process noise w_k and measurement noise v_k.
(2) u_k is persistently exciting; $\text{rank}\,(U_{0,2p-1}) = 2mp$.
(3) w_k and v_k are not both identically zero.
(4) Since only observable modes can be identified, the matrix pair (A, C) is observable. Also, assume that the matrix pair $\left(A, \begin{bmatrix} B & Q^{1/2} \end{bmatrix}\right)$ is controllable.

N4SID, MOESP and CVA are three approaches to subspace identification. For each, there is defined weighting matrices $W_1 \in \Re^{lp\times lp}$ and $W_2 \in \Re^{N\times N}$, where W_1 is full rank and W_2 satisfies: $\text{rank}(W_-) = \text{rank}(W_- \cdot W_2)$ (see Table 1). These weighting matrices determine the state space basis in which the system model will be identified.

The following algorithm provides estimates of the system matrices:

1. Define the oblique projection

$$\mathcal{O}_+ \triangleq Y_+/_{U_+}W_- \quad (16)$$

Table 1: Subspace weighting matrices

Method	W_1	W_2
N4SID	$I_{lp\times lp}$	$I_{N\times N}$
MOESP	$I_{lp\times lp}$	$\Pi_{U_+^\perp}$
CVA	$\Phi^{-1/2}_{[Y_+/U_+^\perp, Y_+/U_+^\perp]}$	$\Pi_{U_+^\perp}$

and compute the singular value decomposition

$$\begin{aligned} W_1\mathcal{O}_+W_2 &= [U_1\ U_2]\begin{bmatrix} S_1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \\ &= U_1S_1V_1^T. \end{aligned} \tag{17}$$

2. Then $\mathcal{O}_+$ satisfies (see [1])

$$\mathcal{O}_+ = \Gamma_p\tilde{X}_p. \tag{18}$$

From assumption 4, we have $\text{rank}(\Gamma_p) = n$ and $\text{rank}(\tilde{X}_p) = n$. Since $\text{rank}(\mathcal{O}_+) = \text{rank}(\Gamma_p\tilde{X}_p)$, then $\text{rank}(\mathcal{O}_+) = n$. W_1 is nonsingular. Thus, $\text{rank}(W_1\mathcal{O}_+W_2) = \text{rank}(\mathcal{O}_+W_2)$. And since $\text{rank}(W_+W_2) = \text{rank}(W_+)$ then W_2 is nonsingular and $\text{rank}(W_2) = N$. Moreover, $\text{rank}(\mathcal{O}_+W_2)$ implies $\text{rank}(W_1\mathcal{O}_+W_2) = \text{rank}(\mathcal{O}_+W_2) = n$. Thus, the system order can be determined from the number of the nonzero singular values in (17).

3. Note that (17) and (18) imply

$$W_1\Gamma_p\tilde{X}_pW_p = \left(U_1S^{1/2}T\right)\left(T^{-1}S^{1/2}V_1^T\right),$$

where T is an arbitrary transformation matrix whose value affects the choice of basis for the identified system model. Next, estimate Γ_p by

$$\hat{\Gamma}_p = W_1^{-1}U_1S_1^{1/2}T. \tag{19}$$

From this, $\tilde{X}_p$ can be estimated by

$$\hat{\tilde{X}}_p = \hat{\Gamma}_p^\dagger\mathcal{O}_+. \tag{20}$$

4. Define

$$K(B,D) \triangleq \begin{bmatrix} \begin{bmatrix} B & \hat{\Gamma}^\dagger_{p-1}H_{p-1}\end{bmatrix} - A\hat{\Gamma}_p^\dagger H_p \\ [D\ \ 0] - C\hat{\Gamma}_p^\dagger H_p \end{bmatrix} \tag{21}$$

and solve the least squares problem

$$\begin{aligned} \hat{A},\hat{C},\hat{K} &= \arg\min \left\| \begin{bmatrix} \hat{\Gamma}^\dagger_{p-1}\bar{Z}_- \\ Y_{p,p} \end{bmatrix} \right. \\ &\quad - \left. \begin{bmatrix} A \\ C \end{bmatrix}\hat{\Gamma}_p^\dagger Z_- - K(B,D)U_+ \right\|_{\text{F}}^2 \end{aligned} \tag{22}$$

5. Solve the least squares problem

$$\begin{aligned} \hat{B},\hat{D} &= \arg\min \left\| \begin{bmatrix} \hat{\Gamma}^\dagger_{p-1}\bar{Z}_- \\ Y_{p,p} \end{bmatrix} \right. \\ &\quad - \left. \begin{bmatrix} \hat{A} \\ \hat{C} \end{bmatrix}\hat{\Gamma}_p^\dagger Z_- - \hat{K}(B,D)U_+ \right\|_{\text{F}}^2 . \end{aligned} \tag{23}$$

6. Use the residues of the minimization (22) as estimates for $\begin{bmatrix} \rho_w^{\text{T}} & \rho_v^{\text{T}} \end{bmatrix}^{\text{T}}$ to compute Q, S and R

$$\begin{bmatrix} \hat{Q} & \hat{S} \\ \hat{S}^{\text{T}} & \hat{R} \end{bmatrix} = \mathbf{E}\left(\begin{bmatrix} \rho_w \\ \rho_v \end{bmatrix}\begin{bmatrix} \rho_w^T & \rho_v^T \end{bmatrix}\right) \tag{24}$$

The above algorithm does not guarantee the stability of the realization. However, in [1, 12, 13] techniques are proposed that guarantee the stability of the identified model. First, define

$$\Gamma_p^0 \triangleq \begin{bmatrix} \bar{\Gamma}_p \\ 0 \end{bmatrix}, \tag{25}$$

where $\bar{\Gamma}_p$ is Γ_p without the first l rows. From (19), determine $\hat{C}$ as the first l rows of $\hat{\Gamma}_p$ and a stable $\hat{A}$ as $\hat{A} = \hat{\Gamma}_p^\dagger\hat{\Gamma}_p^0$. From these values of $\hat{A}$ and $\hat{C}$, recompute $\hat{\Gamma}_p$ and $\hat{\Gamma}_{p-1}$ and then proceed to step 5.

3 An Averaging Method

Here we study the LTI discrete-time system described by (1) and (2). To simplify the analysis, we assume that the noise sequence w_k is not present. The averaging method described here requires that the data be sliced up into batches of data points, each of a user-defined slice size q. Also, let M be the number of slices for a total of $N = Mq$ data points. Note that if u_k is a periodic sequence, with a period of q^* data-points, q does not necessarily equal q^*. Thus the i^{th} ($i = 1, 2, 3, \ldots, M$) output slice $Y^{(i)} \in \Re^{lq}$ is given by

$$\begin{aligned} Y^{(i)} &\triangleq \begin{bmatrix} y_{(i-1)q} \\ y_{(i-1)q+1} \\ \vdots \\ y_{iq-1} \end{bmatrix} = \Gamma_q x_0^{(i)} + H_qU^{(i)} + V^{(i)} \\ &= \Upsilon^{(i)} + V^{(i)}, \end{aligned} \tag{26}$$

where $x_0^{(i)} = x_{(i-1)q}$ is the initial condition of the i^{th} slice and $\Upsilon^{(i)}$, $U^{(i)}$ and $V^{(i)}$ are defined analogous to $Y^{(i)}$ in (26).

It follows from (1) and (2) that

$$x_0^{(i)} = x_{(i-1)q} = A^{(i-1)q}x_0 + \sum_{j=0}^{(i-1)q-1} A^{(i-1)q-j-1}Bu_j,$$

where x_0 is the process initial condition. This expression can be re-written in the compact form

$$\begin{aligned} x_0^{(1)} &= x_0, \\ x_0^{(i)} &= A^{(i-1)q}x_0 + \Lambda_i\hat{U}_i,\ i \geq 2, \end{aligned} \tag{27}$$

where

$$\Lambda_i \triangleq \begin{bmatrix} A^{(i-1)q-1}B & A^{(i-1)q-2}B & \cdots & AB & B \end{bmatrix} \tag{28}$$

and

$$\hat{U}_i \triangleq \begin{bmatrix} U^{(1)} \\ U^{(2)} \\ \vdots \\ U^{(i-1)} \end{bmatrix} \in \Re^{(i-1)mq}. \quad (29)$$

Thus, (26) can be further simplified into

$$\begin{aligned} Y^{(1)} &= \Gamma_q A^{(i-1)q} x_0 + H_q U^{(1)} + V^{(1)}, \\ Y^{(i)} &= \Gamma_q \left[A^{(i-1)q} x_0 + \Lambda_i \hat{U}_i \right] + H_q U^{(i)} + V^{(i)}, \end{aligned}$$

$i \geq 2$. Summing the output over all slices, we have

$$\begin{aligned} Y_M &\triangleq \frac{1}{M} \sum_{i=1}^{M} Y^{(i)} \\ &= \Gamma_q \left[\frac{1}{M} \left(x_0 + \sum_{i=2}^{M} A^{(i-1)q} x_0 + \Lambda_i \hat{U}_i \right) \right] \\ &\quad + H_q \left[\frac{1}{M} \sum_{i=1}^{M} U^{(i)} \right] + \frac{1}{M} \sum_{i=1}^{M} V^{(i)} \\ &= \Gamma_q \bar{x}_0 + H_q U_M + V_M \\ &= \Upsilon_M + V_M, \end{aligned} \quad (30)$$

where

$$\bar{x}_0 \triangleq \left[\frac{1}{M} \left(x_0 + \sum_{i=2}^{M} A^{(i-1)q} x_0 + \Lambda_i \hat{U}_i \right) \right] \quad (31)$$

is a weighted average initial condition,

$$U_M \triangleq \frac{1}{M} \sum_{i=1}^{M} U^{(i)} \quad (32)$$

is the averaged input sequence,

$$V_M \triangleq \frac{1}{M} \sum_{i=1}^{M} V^{(i)} \quad (33)$$

is the averaged output noise sequence, and

$$\Upsilon_M \triangleq \frac{1}{M} \sum_{i=1}^{M} \Upsilon^{(i)} = \Gamma_q \bar{x}_0 + H_q U_M \quad (34)$$

is the averaged system output (uncorrupted with the averaged noise V_M).

Though the signals themselves are now different, (34) has the same input/output relation defined through Γ_q and H_q as in (4). Due to system linearity, this shows that the averaged sequence Y_M and the averaged input sequence U_M can be used for the subspace-based identification of a system that gives the output sequence y_k when excited by an input signal u_k.

Now define the signal-to-noise ratio

$$\text{SNR} \triangleq \sqrt{\frac{\|\Phi_{[\Upsilon_M, \Upsilon_M]}\|}{\|\Phi_{[V_M, V_M]}\|}}, \quad (35)$$

where $\|\cdot\|$ denotes the induced 2-norm. Here we treat the input sequence as a deterministic signal or as a realization of a measured random signal. Referring to the approximation made in (15), it follows that $\|\Phi_{[\Upsilon_M, \Upsilon_M]}\| \cong \|\Upsilon_M \Upsilon_M^{\mathrm{T}}\|$, which is independent of M. For the denominator of (35) we have

$$\begin{aligned} \|\Phi_{[V_M, V_M^{\mathrm{T}}]}\| &\cong \|V_M V_M^{\mathrm{T}}\| \\ &= \left\| \left(\frac{1}{M} \sum_{i=1}^{M} V^{(i)} \right) \left(\frac{1}{M} \sum_{i=1}^{M} V^{(i)} \right)^{\mathrm{T}} \right\| \\ &= \left\| \frac{\sigma_v^2}{M} I_{lq} \right\| = \frac{\sigma_v^2}{M}, \end{aligned} \quad (36)$$

where σ_v is the variance of the noise sequence. Equations (35) and (36) suggest that SNR $\propto \sqrt{M}$ whether the input is periodic or not and regardless of how we choose the slice size q. In the following sections we numerically investigate this statement. It was found that this result is only true when two conditions are satisfied: (1) the input u_k is a periodic signal that repeats every q^* points and (2) when we choose $q = q^*$.

4 Averaging Using Periodic Signals

In this and the following sections we consider the system

$$A = \begin{bmatrix} 0.67 & 0.67 & 0 & 0 \\ -0.67 & 0.67 & 0 & 0 \\ 0 & 0 & -0.67 & -0.67 \\ 0 & 0 & 0.67 & -0.67 \end{bmatrix} \quad B = \begin{bmatrix} 0.6598 \\ 1.9698 \\ 4.3171 \\ -2.6436 \end{bmatrix}$$

$$C = [-0.5749 \ 1.0751 \ -0.5225 \ 0.1830] \quad D = -0.7139$$

This system has poles at $0.67 \pm 0.67\imath$ and $-0.67 \pm 0.67\imath$, which correspond to the frequencies 1.87×10^5 rad/sec and 2.36×10^6 rad/sec, respectively. System zeros are at $1.03 \pm 1.08\imath$, -1.6 and -1.86, which correspond to frequencies 9.03×10^5 rad/sec, 3.18×10^6 rad/sec and 3.2×10^6 rad/sec, respectively. The sampling frequency is $T_s = 10^{-6}$ sec. This system has the frequency response plot shown in Figure 1. We designed the simulations such that the SNR of the entire N-point data record is 0.1.

To evaluate the averaging procedure, we use the discrete time H_2-norm of the difference between the true and the identified systems. Letting $\hat{A}$, $\hat{B}$, $\hat{C}$ and $\hat{D}$ denote the identified system matrices, the performance measure is

$$J = \tilde{C} Q \tilde{C}^{\mathrm{T}} + \tilde{D}^{\mathrm{T}} \tilde{D}, \quad (37)$$

where

$$\tilde{A} \triangleq \begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & \hat{A} \end{bmatrix},$$

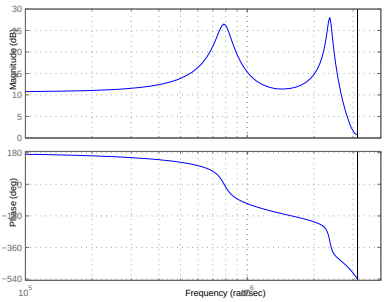

Figure 1: System frequency response plot

$\tilde{B} \triangleq \begin{bmatrix} B^{\mathrm{T}} & \hat{B}^{\mathrm{T}} \end{bmatrix}$, $\tilde{C} \triangleq \begin{bmatrix} C & -\hat{C} \end{bmatrix}$, $\tilde{D} = D - \hat{D}$ and Q is the solution to the following discrete-time Lyapunov equation

$$Q = \tilde{A} Q \tilde{A}^{\mathrm{T}} + \tilde{B} \tilde{B}^{\mathrm{T}}. \tag{38}$$

As mentioned above, the two conditions for obtaining SNR $\propto \sqrt{M}$ are the periodicity of the input signal (with period q^*) and having $q = q^*$. Here we use three types of signals (with $q^* = 931$): a repeated realization of a zero mean, stationary random sequence, a repeated chirp signal and a repeated multitonal signal.

Figure 2 shows SNR as a function of M. By making a data fit, it was found that SNR $= \alpha\sqrt{M}$ with $\alpha = 0.1, 0.073$ and 0.17 for noise-, chirp- and multitonal-based identification, respectively. Note that the periodic multitonal signal results in the highest value of averaged SNR whereas the chirp results in the lowest. This may lead us to conclude that the multitonal signal should result in better identification results. However, Figure 3 shows that this is not necessarily true. The reason is that the multitonal input sequence is not as sufficiently exciting as the chirp or random signals and, therefore, its identification performance is worse than that of the other two signal types. The reason for the multitonal input signal resulting in the highest averaged SNR is because the denominator of (35) is the same regardless of the input signal type. However, these signals would result in an output having different values of $\|\Phi_{[\Upsilon_M, \Upsilon_M]}\|$. The periodic multitonal input sequence results in the highest value for $\|\Phi_{[\Upsilon_M, \Upsilon_M]}\|$ and, consequently, it has the highest SNR values. Figure 4 shows the frequency response as a function of M.

5 Averaging Using Sinusoidal Signals

Care should be taken when a multitonal input signal is used when frequency components of this signal coincide with the zeros of the system. When this happens, these signal components will not contribute to the output. When the averaging method described above is performed, the averaged input signal will retain these

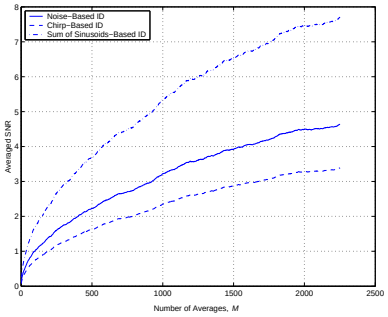

Figure 2: SNR as a function of M

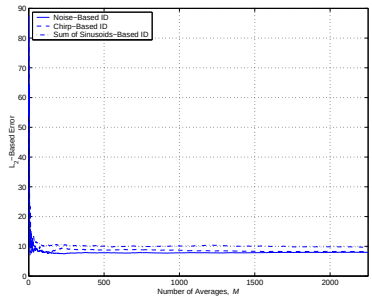

Figure 3: Error as a function of M

components but the averaged output will not. This "missing information" may have some negative effects.

For instance, Figure 5 shows a multitonal signal that has two frequency components at 3.18×10^6 rad/sec and 3.2×10^6 rad/sec, which do not show up in the output because they coincide with two system zeros. The corresponding identified frequency response is compared to the true one in Figure 7, where the result is poor. When the signal shown in Figure 6 is used instead, the model obtained from the identification algorithm has the frequency response plot shown in Figure 8, and the result is satisfactory.

In reality, of course, information such as the location of the system zeros will not be available beforehand. The aim here is to show that multitonal signals may in some cases produce very poor results.

6 Effect of Slice Size q

If the input signal is periodic and its periodicity is known, then choosing $q = q^*$ guarantees constructive averaging of the input signal and, thus, avoids the loss of spectral content. Otherwise, the averaging process may result in spectral degradation of the signal.

This, however, does not imply that choosing $q = q^*$ results in the least performance error. Using the input signals mentioned in the previous section, we obtained the H_2-based error vs. slice size as shown in Figures 9, 10 and 11. Only when the multitonal signal was used, $q = q^* = 931$ resulted in the least error. For the

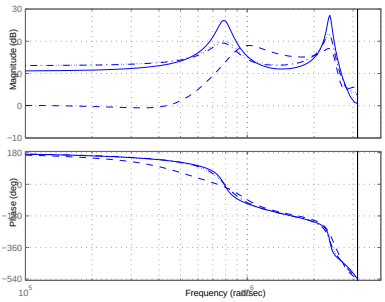

Figure 4: System frequency response (solid), $M = 1$ (dashed), $M = 100$ (dash-dotted) and $M = 800$ (dotted) with periodic multitonal signal

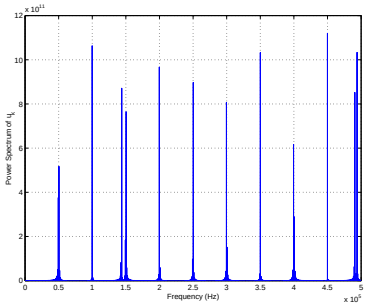

Figure 5: Spectrum of u_k with harmonics at system zeros

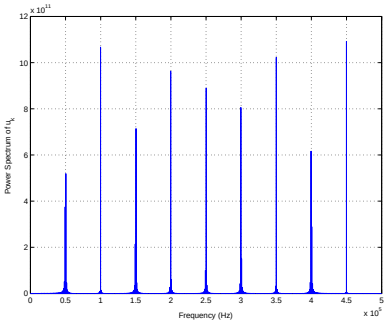

Figure 6: Spectrum of u_k with no harmonics at system zeros

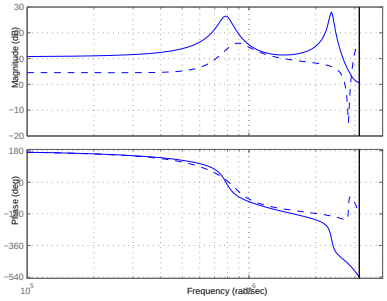

Figure 7: System frequency response (solid) and identified model (dashed) using signal in Figure 5

noise signal, $q = 1064$ resulted in the least error. For the chirp signal, $q = 399$ resulted in the least error. Note that $931 = 133 \times 7$, $1064 = 133 \times 8$ and $399 = 133 \times 3$. This result may be further appreciated when we investigate plots of the averaged SNR as a function of slice size as in Figures 12, 13 and 14. Note again the harmonic peaks that take place at multiples of 133.

If q^* is known beforehand, choosing $q = q^*$ guarantees satisfactory performance. Still, there may exist slice sizes that are equal to some rational multiple of q^* and that result in even better performance. It is not clear what determines this slice size. If, on the other hand, knowledge of q^* was not available, one may look to the power spectral density of the input signal and from it estimate the value of $q^* = \frac{2\pi}{T_s f^*}$, where f^* is the frequency at which there is a sharp peak in the spectral content of u_k.

7 Averaging Using Non-Periodic Signals

If u_k is known to be non-periodic (for instance if its power spectrum has no sharp peaks), the question then arises: when we average, will we get better or worse results than without averaging? To answer this question, we used non-periodic signals of lengths that enable us to run the identification code without running out of memory. Here we chose u_k to be a realization of a zero mean, stationary gaussian random sequence with $N = 4104$ and $q = 513$ (chosen arbitrarily). The result was quite poor and is shown in Figure 15.

This result may be due to two effects. First, since u_k is a random sequence, arbitrarily slicing up the data and averaging may have resulted in the destructive averaging of u_k. The second reason may be that not enough averages were sufficiently taken. To investigate this, another run was carried out but with $N = 1,862,000$ data points and $q = 931$ (i.e. $M = 2000$) with u_k being in one trial a realization of a zero mean, stationary gaussian random sequence and in the second chirp signal. The result is shown in Figure 16. Even with more averages, the results are poor.

8 Discussion

In this paper we explored the possibility of applying data compression techniques to subspace methods. We have shown that when the input sequence is periodic, then satisfactory results are obtained when a subspace algorithm is used to identify the system by slicing and averaging the input and output signals. A consequence of this result is that subspace methods can be implemented in a fashion that utilizes large input/output measurements which are otherwise discarded due to memory constraints. We have also discussed the impact of the choice of the slice size q. Applying the averaging method with non-periodic input sequences

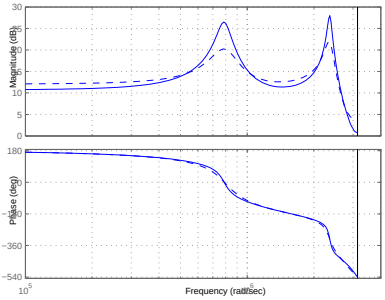

Figure 8: System frequency response (solid) and identified model (dashed) using signal in Figure 6

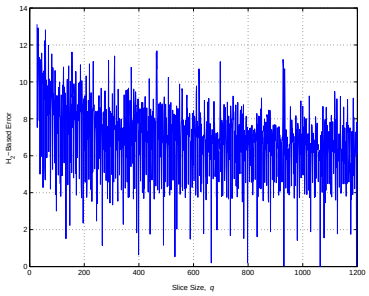

Figure 9: Error vs. q with periodic noise input, $q^* = 931$. Minimum error at $q = 1064$

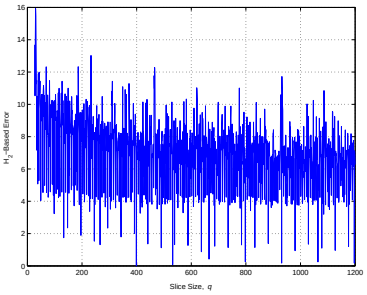

Figure 10: Error vs. q with periodic chirp input, $q^* = 931$. Minimum error at $q = 399$

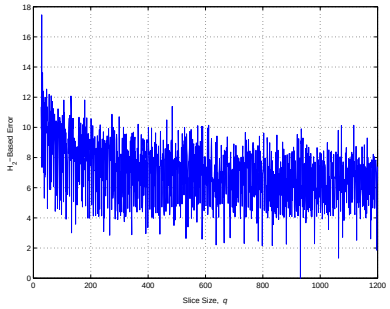

Figure 11: Error vs. q with periodic multitonal input, $q^* = 931$. Minimum error at $q = 931$

was shown not to be beneficial as with the case of periodic signals. Future work will focus on the methods for utilizing large data sets with nonperiodic inputs.

References

[1] P. Van Overschee and B. De Moor. *Subspace Identification For Linear Systems*. Kluwer Academic Publishers, 1996.

[2] P. Van Overschee and B. De Moor. N4SID: Subspace Algorithms for the Identification of combined Deterministic-Stochastic Systems. *Automatica*, Vol. 30, No. 1, pp. 75-93, 1994.

[3] M. Viberg. Subspace-based Methods for the Identification of Linear Time-invariant Systems. *Automatica*, Vol. 31, No. 12, pp. 1835-1851, 1995.

[4] P. Van Overschee and B. De Moor. A Unifying Theorem for Three Subspace System identification Algorithms. *Automatica*, Vol. 31, No. 12, pp. 1853-1864, 1995.

[5] Coherent Phase Line Enhancer: a Method of Spectral Analysis. *NASA Tech Briefs*, Vol. 25, No. 7, pp. 36, July 2001.

[6] D. K. De Vries and P. M. Van Den Hof. Quantification of Uncertainty in Transfer Function Estimation: a Mixed Probabilistic-Worst-case Approach. *Automatica*, Vol. 31, No. 4, pp. 543-557, 1995.

[7] M. Nakamura, S. Nishida and H. Shibasaki. Evaluation of the Signal-to-Noise Ratio for Average Evoked Potentials Contaminated by Colored Noise. *IFAC Identification and System Parameter Estimation*, pp. 891-896, 1988.

[8] I. Kollár. Signal Enhancement Using Non-Synchronized Measurements. *IEEE Instrumentation and Measurement Technology Conference*, pp. 365-367, 1991.

[9] R. G. Hakvoort and P. M. J. Van Den Hof. Consistent Parameter Bounding Identification for Linearly Parametrized Model Sets. *Automatica*, Vol. 31, No. 7, pp. 957-969, 1995.

[10] S. J. Orfanidis. *Optimum Signal Processing, Second Edition.*MacMillan Publishing Company, 1988.

[11] P. M. Mäkilä. Robust Identification and Galois Sequences. Int. J. Control, Vol. 54, No. 5, pp. 1189-1200, 1991.

[12] J. M. Maciejowski. Guaranteed stability with subspace methods. *Systems and Control Letters*, Vol. 26, pp. 153-156, 1995.

[13] N. L. C. Chui and J. M. Maciejowski. Realization of stable models with subspace methods. *Automatica*, Vol. 32, No. 11, pp. 1587-1595, 1996.

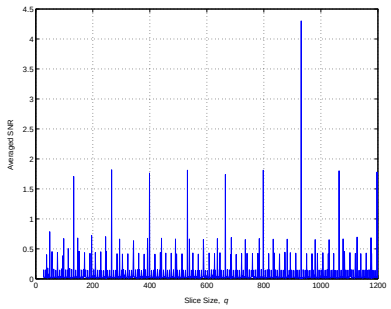

Figure 12: Averaged SNR vs. q (periodic noise input)

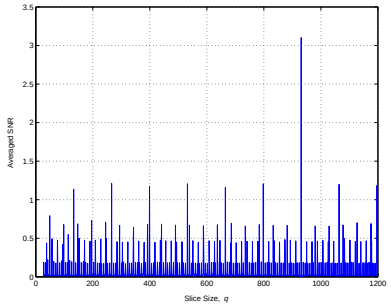

Figure 13: Averaged SNR vs. q (periodic chirp input)

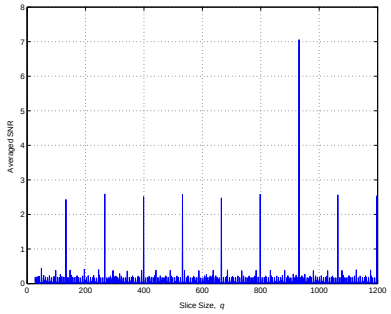

Figure 14: Averaged SNR vs. q (periodic multitonal input)

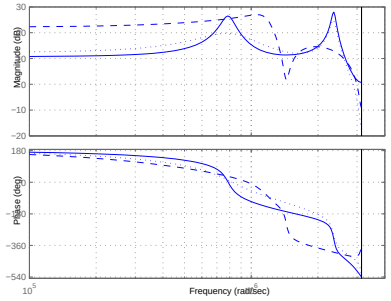

Figure 15: System frequency response (solid), non-averaged identified model (dotted) and averaged identified model (dashed)

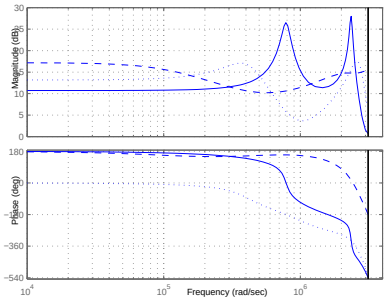

Figure 16: System frequency response (solid), chirp input identified model (dotted) and random input identified model (dashed)