Input Design for Subspace-Based Fault Detection

Alireza Esna Ashari and Laurent Mevel

Abstract—This paper considers the problem of auxiliary input design for subspace-based fault detection methods. In several real applications, particularly in the damage detection of mechanical structures and vibrating systems, environment noise is the only input to the system. In some applications, white noise produces low quality output data for the subspace-based fault detection method. In those methods, a residual is calculated to detect the fault based on the output information. However, some modes of the system may not influence the outputs and the residual appropriately if the input is not exciting enough for those modes. In this paper, “rotated inputs” method is implemented to excite the system modes. In addition to produce a residual more sensitive to the weak modes, it is possible to detect system order changes due to the fault using the rotated inputs. Simulation results demonstrate the efficiency of injecting these auxiliary inputs to improve the subspace-based fault detection methodology.

I. INTRODUCTION

Over the last decades subspace-based identification have been an active domain of research. These methods are based on geometric concepts including the calculation of certain matrices, geometric manipulation of the row spaces and computation of projections of data on certain subspaces. A comprehensive survey of subspace-based identification approaches can be found in [1], [2]. The identification problem consists of obtaining the state-space representation of the system from input-output data using linear algebra tools, up to a similarity transformation.

Fault detection problem is one of the related topics to system identification and has been studied using several methods [3], [4]. In some fault detection and health monitoring problems, detecting the changes of system eigenstructure is the main subject of interest. On the other hand, subspace identification is well suited to estimate the system eigenstructure. Based on the subspace identification methods, a fault detection methodology is developed in [5] to detect the eigenstructure changes.

Subspace-based fault detection method has already been used in several practical applications [6], [7], [8], [9], [10], [11], [12], [13]. In these applications, the natural unknown and unmeasured environment noise, considered as white noise, is the only input to the system. This input excites the system modes and produces output data for identification. A residual signal is calculated based on the observed outputs of the system. The residual is around zero for the nominal system and a fault alarm is raised if the mean value of the residual reaches a threshold. However, in some applications, this input cannot stimulate some of the system modes enough, and consequently the corresponding singular values would be very small and would be considered as noise effects. Therefore these weak modes will not have considerable effects on the residual and their changes will be overlooked. In some dynamic multi-input multi-output systems the direction of the input is important to produce output data for detection and identification purposes and for some input directions the outputs are much larger than the others [14].

As the subspace-based fault detection method is derived from the subspace-based eigenstructure identification, it inherits the merits and also difficulties of those approaches. One important issue in application of the subspace identification method is to know or estimate the system order, in order to obtain precise results. Every subspace-base method includes a common step of performing singular value decomposition (SVD) on a data matrix. The order of the system is usually determined by the number of “large” singular values. Small singular values are considered as the effect of noise on data. This may affect the system order estimation if the real order is not given. In subspace-based fault detection, it prevent us from the detection of system order changes.

The problem of ill-conditioning and the error analysis of subspace identification algorithms in such a case is discussed in details in [15]. In these systems, some of the singular values are very small when white noise is the only input to the system. Hence, it is very hard to select the real nonzero singular values corresponding to the system modes in this case. In order to overcome this issue, “rotated input” approach is proposed in literature. In this approach, rotated inputs are considered, the best angles between the inputs are calculated and applied to the system [16], [17], [14]. Application of this pre-designed test input helps to increase the ratio between real singular values and the rest of singular values due to the noise. Consequently, the weak modes may not be negligible anymore and the order of the system should be better estimated. While the application of the rotated input in system identification may lead to a better system order recovery, it will not increase the accuracy of the identification. On the other hand, an important problem with using the rotated inputs for identification purposes is that a preliminary model of the system is required in order to calculate the rotation angle, while it is not available in most applications. Therefore, the rotation angle is calculated by trial and error in [14]. Also, the effectiveness of using rotated inputs is strongly dependent on the subspace algorithm implemented.
Traditionally most fault detection approaches, including the subspace-based method in [5], are passive in that system inputs and outputs are monitored and decisions made. Recently, there has been more interest in designing inputs to improve fault detection methods [18], [19], [20], [21], [22], [23], [24].

While the application of the rotated input method in subspace identification is limited to system order recovery of ill-conditioned systems, the idea is implemented to improve the subspace-based fault detection technique in this paper. The sensitivity of the residual to the changes of the system modes is amplified using the input. Injecting the rotated input to the system, the impact of the weak modes on the residual will be increased. In addition, the application of this input provides useful information to detect the system order change due to a fault. Unlike the identification case, there exist a nominal model for fault detection tests and it is possible to calculate the test input precisely. Simulation results demonstrate the surprising advantages of this method to improve the quality of subspace-based fault detection.

The paper is organized as follows: in section 2, problem formulation and preliminary material are provided. Section 3 presents subspace-based fault detection. In sections 4, input design for subspace-based fault detection is discussed. In sections 5 simulation results are presented to show the efficiency of the method and concluding remarks follow in section 6.

II. Problem Formulation and Preliminaries

Consider the discrete-time model in state space form:

\[ x(k+1) = Ax(k) + w(k), \]
\[ y(k) = Cx(k), \]

where \( x(k) \in \mathbb{R}^n \), \( y(k) \in \mathbb{R}^r \), \( A \in \mathbb{R}^{n \times n} \) and \( C \in \mathbb{R}^{r \times n} \) are the state vector, the output vector, the state transient matrix, and the output matrix, respectively. The state noise \( w(k) \in \mathbb{R}^p \) is unmeasured and Gaussian, zero mean with covariance \( \Sigma_w \). It is assumed that the measurement noise is zero to simplify calculations. However, the results can be extended for systems with measurement noise. The fault detection problem considered in this paper consists of monitoring the eigenstructure of system (1)–(2), and detect its changes. The eigenstructure of a system is the set of eigenvalues of the system (3), and the corresponding eigenvectors. The eigenvalues and eigenvectors are represented by \( \lambda_i \) and \( v_i \) for \( i = 1, \ldots, n \), respectively. A pair of an eigenvalue and the corresponding eigenvector is called a mode. The set of all \( n \) modes of the system makes system parameter \( \theta \)

\[ \theta = \left( \begin{array}{c} \Lambda \\ v_{\Lambda \Phi} \end{array} \right), \]

where \( \Lambda \) is the vector whose elements are \( \lambda_i \)'s and \( \Phi \) is the matrix whose columns are \( v_i \)'s. Let \( p \) and \( q \) be chosen parameters with \( n \leq (p+1)r \leq qr \). From the output data \( y(k), k = 1, \ldots, N + p + q \) a matrix \( \mathcal{H}_{p+1,q} \in \mathbb{R}^{(p+1)r \times qr} \) is built according to a chosen subspace identification (detection) algorithm (see [2] for an overview of several algorithms). The matrix \( \mathcal{H}_{p+1,q} \) is called “subspace matrix”. There exist different subspace-based algorithms in literature and each algorithm uses a particular subspace matrix (more information can be found in [1]). In this paper the following Hankel matrix is implemented for simulations (see [25])

\[ \mathcal{H}_{p+1,q} = Y^+ Y^{-T} (Y^{-} Y^{-T})^{-1} Y^{-}, \]

where

\[ Y^+ = \begin{pmatrix} y(k) \\ \vdots \\ y(k+p) \end{pmatrix}, \quad Y^{-} = \begin{pmatrix} y(k-l-1) \\ \vdots \\ y(k-l-q) \end{pmatrix}. \]

Apart from the selection of the matrix \( \mathcal{H}_{p+1,q} \), it always has the following factorization property asymptotically

\[ \mathcal{H}_{p+1,q} = \mathcal{O}_{p+1} \mathcal{Z}_q, \]

where \( \mathcal{O}_{p+1} \) is the observability matrix,

\[ \mathcal{O}_{p+1} = (C^T \ A^T \ \cdots \ \ (CA^p)^T)^T, \]

and \( \mathcal{Z}_q \) depends on the selected subspace identification algorithm. The observability matrix \( \mathcal{O}_{p+1} \) is obtained from the Singular Value Decomposition (SVD) of the matrix \( \mathcal{H}_{p+1,q} \). Considering the SVD of \( \mathcal{H}_{p+1,q} \)

\[ \mathcal{H}_{p+1,q} = U \Delta V^T, \]

\[ = (U_1 \ U_2) \begin{pmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix}, \]

where \( \Delta_2 \) contains the singular values corresponding to the weak modes which are almost zero. Assume that \( \Delta_2 \approx 0 \). Using this assumption

\[ \mathcal{H}_{p+1,q} = U_1 \Delta_1 V_1^T, \]

and one can obtain

\[ \mathcal{O}_{p+1} = U_1 \Delta_1^{\frac{1}{2}}. \]

Reducing the dimension of the matrix \( \mathcal{H}_{p+1,q} \) and going from (8) to (9) needs some attention that we discuss later in Section IV-D. For the sake of simplicity, let \( p \) and \( q \) be given. Hence, the subscripts \( p \) and \( q \) of \( \mathcal{H}_{p+1,q} \), \( \mathcal{O}_{p+1} \) and \( \mathcal{Z}_p \) are dropped and \( \mathcal{H} \), \( \mathcal{O} \) and \( \mathcal{Z} \) are substituted respectively in the following.

III. Subspace-Based Fault Detection

In [5] a statistical fault detection method is proposed based on subspace algorithms satisfying factorization property (6). This method consists of comparing characteristics of the nominal system with a subspace matrix \( \mathcal{H} \) computed on a new data sample \( y(k), k = 1, \ldots, N + p + q \), corresponding to an unknown, possibly damaged state, assuming that \( \mathcal{H} \) is a consistent estimate of \( \mathcal{H} \).

To compare the new data with nominal characteristics, the left null space matrix \( S \) of the observability matrix corresponding to the nominal system is computed, which is also the left null space of the subspace matrix for the
nominal system because of factorization property (6). The characteristic property of a system in the nominal state is $S^T \mathcal{H} = 0$ and the residual vector
\[
\zeta \overset{\text{def}}{=} \sqrt{N} \text{vec}(S^T \mathcal{H}),
\] (11)
demonstrates the difference between the nominal system and the current situation of the system.

Let $\theta$ be a vector containing a canonical parametrization of the system under monitoring (see [5] for details) and $\theta_0$ the parametrization of the nominal system. The fault detection problem is that whether or not the subspace matrix $\mathcal{H}$ from the monitored system (corresponding to $\theta$) is still well described by the characteristics of the nominal system (corresponding to $\theta_0$). This is done by hypothesis test
\[
H_0 : \theta = \theta_0 \quad \text{(reference system)},
\]
\[
H_1 : \theta = \theta_0 + \delta \theta / \sqrt{N} \quad \text{(faulty system)},
\] (12)
where $\delta \theta$ is unknown but fixed. This is called the local approach, and the following theorem is used to decide a hypothesis.

**Theorem 3.1 ([5]):** The residual $\zeta$ is asymptotically Gaussian for large $N$, and the hypotheses test between $H_0$ and $H_1$ is performed using the $\chi^2$-test
\[
\chi^2 = \zeta^T \Sigma^{-1} \mathcal{J}(\mathcal{J}^T \Sigma^{-1} \mathcal{J})^{-1} \mathcal{J}^T \Sigma^{-1} \zeta,
\] (13)
which is compared with a threshold, where $\mathcal{J}$ and $\Sigma$ are consistent estimates of the mean residual sensitivity and residual covariance, respectively. $\mathcal{J}$ and $\Sigma$ are defined as follows
\[
\mathcal{J} = \lim_{n \to \infty} -\frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta} \mathbf{E}_{\theta_0} \zeta(\theta)|_{\theta = \theta_0},
\]
\[
\Sigma = \lim_{n \to \infty} \mathbf{E}_{\theta_0} (\zeta \zeta^T).
\] (14)
Here, $\mathbf{E}_{\theta}(.)$ represents the expected value of $(.)$ when the current system parameter is $\theta$.

The computation of the Jacobian $\mathcal{J}$ needs a parametrization of the system, where the eigenvalues and mode shapes of the nominal system must be known, and is explained in detail in [5].

**IV. INPUT DESIGN FOR SUBSPACE-BASED FAULT DETECTION**

In this section, the necessity of designing auxiliary inputs to better detect the faults in some dynamic systems is explained. The method of input design is explained thereafter.

**A. Fault Detection using Natural Excitations**

In this section, the difficulties with the traditional subspace-based fault detection method are explained briefly. The method in [5] gives excellent results when it is used in practice [6], [7], [9], [10], [12], [13]. This method uses white noise as input to excite the system. In practice, the environment noise is the only input that excites the system and produces the data for fault monitoring. The environment noise usually can be considered as white noise. However, practical experiences show that some particular damages are not well reflected in the residual. Hence, their changes cannot be easily monitored. To illustrate the reason, $Z$-transformation is used and the frequency representation of the system is obtained. Consider a system that has big condition number, particularly in steady state case. Suppose that the transfer function of the system is $G(z)$, and $u(z)$ and $y(z)$ are the input and output signals in frequency domain
\[
y(z) = G(z)u(z),
\] (15)
and the SVD of $G(z)$ is
\[
G(z) = \Upsilon(z)\Sigma(z)(\Omega^T(z)).
\] (16)
Consequently,
\[
y(z) = \left(\begin{array}{c}
u_{11} \\
\vdots \\
u_{rt}
\end{array}\right) (\begin{array}{c}
s_1 \\
0 \\
s_2 \\
0 \\
\vdots \\
t_m \\
\end{array}) \times \left(\begin{array}{c}
\omega_{11} \\
\omega_{1t} \\
\vdots \\
\omega_{rt}
\end{array}\right) \left(\begin{array}{c}
u_1 \\
\vdots \\
u_m \\
\end{array}\right),
\] (17)
where $t = \min(m, r)$. In (17), the dependence on $z$ is dropped for simplicity. Assuming
\[
\xi_i = s_i \sum_{j=1}^m \omega_i \nu_j,
\] (18)
one obtains
\[
y(z) = \xi_1 \left(\begin{array}{c}
u_{11} \\
\vdots \\
u_{1t}
\end{array}\right) + \cdots + \xi_r \left(\begin{array}{c}
u_{rt} \\
\vdots \\
u_{rt}
\end{array}\right).
\] (19)
The system $G(z)$ is ill-conditioned if the biggest singular value $\sigma(z)$ is much larger than the smallest singular value $\sigma(z)$. In this case, some of the singular values are negligible considering that $\Upsilon(z)$ is a unitary matrix. The important consequence of (19) is that there is a linear dependency between the outputs, and they are in the range of the first columns of $\Upsilon(z)$, corresponding to the biggest singular values. Particularly, in the case that one singular value is much bigger than the others $y_i \approx \frac{\omega_{i1}}{\sigma_{11}} y_1$, for $i = 2, \cdots, r$. This output is used to compute the Hankel matrix $\mathcal{H}$. This justifies the failure of white inputs to reflect the effects of all the system modes in the residual. A second difficulty, as it is shown in [14], is to distinguish the smallest system singular value of $\mathcal{H}$ from the other small nonzero noise singular values. It leads to the incorrect estimation of the system order, if it is not given.

**B. Subspace-Based Fault Detection using Rotated Input**

It is desired to design an input such that the residual is more sensitive to the fault. In order to improve the quality of fault detection and increase the sensitivity of the residual to the changes of the weak modes, the effect of all the singular values in (17) on the residual should be equal. Rotated inputs
will be designed and injected into the system to satisfy this condition.

The detection method in Section III is based on calculating $S$, the left kernel of the observability matrix. The selection of $S$ depends on the SVD of $\mathbf{H}$. Only the $n$ first columns are taken to make the residual and $n$ is the number of large singular values of $\mathbf{H}$. If the number of large singular values, the values which are considerably larger than the others, was exactly equal to the order of the system, it would be possible to detect system order changes when the number of large singular values changes.

An input is desired to be injected into the system which can improve the fault detection method in two directions:

1) to strengthen the effect of weak modes on the residual.
2) to detect the change of the system order due to a fault.

This paper focuses more on the first point, which is crucial to detect some minor damages.

### C. Rotated Input Design

The method of rotated input design is implemented in this section to achieve the detection objectives explained in Section IV-B. The method is already used for identification experiment design [16] and later for order determination together with subspace identification method [14], [17]. In this section, another representation of the method is provided for the systems with two inputs and two outputs to demonstrate the idea and explain why it is called rotated input method. However, it can be easily extended to the general case using any method that equalizes the effect of the modes. The $2 \times 2$ orthogonal matrices $\mathbf{Y}(z)$ and $\Omega(z)$ in (16) can be written as

$$
\mathbf{Y}(z) = \begin{pmatrix}
\cos\varphi(z) & -\sin\varphi(z) \\
\sin\varphi(z) & \cos\varphi(z)
\end{pmatrix},
\Omega^T(z) = \begin{pmatrix}
\cos\theta(z) & -\sin\theta(z) \\
\sin\theta(z) & \cos\theta(z)
\end{pmatrix}.
$$

(20)

(21)

Considering the singular value decomposition of the system $G(z)$ introduced in (16), one may obtain

$$
y(z) = G(z)u(z) = \mathbf{Y}(z) \begin{pmatrix}
\sigma_1(z)\cos\theta(z)u_1(z) - \sigma_1(z)\sin\theta(z)u_2(z) \\
\sigma_2(z)\sin\theta(z)u_1(z) + \sigma_2(z)\cos\theta(z)u_2(z)
\end{pmatrix}.
$$

(22)

To avoid collinearity problems, the outputs should be as independent as possible. To satisfy this condition, the inputs will be designed. The key point is that no terms in (22) should be negligible. Hence, it is necessary to make each term of the summation contribute in the value of $y$ equally. From (19), considering the energy of each term contributes in $y_i$, and the fact that $\Psi_i(z)$ is a unitary matrix, the criterion becomes

$$
\int_0^{2\pi} |\xi_i(e^{j\omega})|^2 d\omega = constant,
$$

(23)

for $i = 1, 2$. This occurs if

$$
\xi_i(e^{j\omega}) = constant.
$$

(24)

To satisfy (24) it is enough to have

$$
\sigma_1(z)\cos\theta(z)u_1(z) - \sigma_1(z)\sin\theta(z)u_2(z) \\
= \sigma_2(z)\sin\theta(z)u_1(z) + \sigma_2(z)\cos\theta(z)u_2(z),
$$

(25)

or equivalently

$$
u_2(z) = \sigma_1(z)\cos\theta(z) - \sigma_2(z)\sin\theta(z) \\
\sigma_2(z)\cos\theta(z) + \sigma_1(z)\sin\theta(z)u_1(z).
$$

(26)

A drawback of using the method for identification purposes is that the angle between the inputs cannot be obtained directly as the model is unknown and it should be calculated experimentally by trial and error. In fault detection, the nominal model of the system is given, and the rotated input can be calculated without approximation using (26) and the inverse Laplace transform. However, an approximate approach is enough for the most of the practical applications. Now assume that $\frac{\sigma_1}{\sigma_2} = \kappa$. It follows that

$$
u_2(z)|_{\kappa \to \infty} = cot\theta(z)u_1(z).
$$

(27)

From (27), it can be understood why the method is called rotated input. To approximate the solution one can calculate the solution in the steady state (frequency zero) assuming that $\Omega(z) \approx \Omega$ where $\Omega = \Omega(0)$. Hence,

$$
u_2(k) \approx cot\theta u_1(k),
$$

(28)

where $\theta = \theta(0)$. To calculate $u_2(k)$, $\theta$ is computed first from the nominal model. The input $u_1(k)$ is chosen as a zero mean white noise and $u_2(k)$ is obtained from (28).

### D. Detection of Model Order Change

One of the advantages of the new subspace-based fault detection method excited by the rotated input is to detect system order changes, while it is not possible in most of the fault detection methods in literature. In some practical applications, system may lose some part of its dynamic and therefore the system order may change. Consider again (8) and (9) and the fact that the selection of $\Delta_2$ is not easy using the traditional subspace-based method when the system is subjected to noise. It is difficult to distinguish between the singular values due to the system and noise. Small singular values should be related to noise, however in some systems the some singular values of the system are very small if white noise is the only input to the system. Using the rotated input, these modes are pushed up and consequently, they will be distinguishable from the effects of noise. Therefore the system order change can be detected if the number of large singular values of $\mathbf{H}$ changes. Note that this is the second advantage of using rotated inputs and the method still improves the detection as we described in Section IV-C even if the system order is not changed.

### V. Numerical Examples

In order to verify the efficiency of the proposed method, several simulations are performed. In this section, the results are summarized. To show the efficiency of the proposed method, the $\chi^2$ values of Theorem 3.1, generated by the traditional method and the new approach are plotted. Note
that each point of the residual plot represents the $\chi^2$ value calculated from the output data in 200 sample-times. It is called an “experience” in this paper. In each simulation, 1000 experiences are performed and the system changes after 500 experiences. The aim is to detect this change.

To perform the simulations, an ill-conditioned system is selected which has been studied in several papers e.g. [14] and [17]. This system is a high-purity distillation column in $LV$ control configuration. The transfer matrix of the continuous-time system is

$$
\begin{pmatrix}
y_1(s) \\
y_2(s)
\end{pmatrix} = \begin{pmatrix}
\frac{87.8}{1+108.2s} & \frac{-87.8}{1+108.2s} \\
\frac{0.14}{1+194s} & \frac{0.14}{1+194s}
\end{pmatrix}
\begin{pmatrix}
u_1(s) \\
u_2(s)
\end{pmatrix}.
$$

In order to use the method proposed in this paper, the continuous-time model should be transformed to discrete-time. Two different simulations are performed to study the effect of using the rotated input on residual and system order estimation. To produce the rotated input in each test, the first input channel $u_1$ is excited by white noise with covariance one and $u_2$ is the same as $u_1$ but rotated by angle $\theta$ and distorted slightly using a weak white noise. The angle is calculated using the SVD according to Section IV-C.

In the first experiment, the advantage of using the new method is studied to improve the detection of faults that are hard to be detected or remain hidden using the traditional method. Assume that the fault occurs at 500-th experiences. The faulty system model is

$$
\begin{pmatrix}
y_1(s) \\
y_2(s)
\end{pmatrix} = \begin{pmatrix}
\frac{87.8}{1+108.2s} & \frac{-87.8}{1+108.2s} \\
\frac{0.14}{1+194s} & \frac{0.14}{1+194s}
\end{pmatrix}
\begin{pmatrix}
u_1(s) \\
u_2(s)
\end{pmatrix}.
$$

The traditional subspace-base detection method and the new approach are both used to detect this change. The $\chi^2$ values of both methods are plotted in Figure 1. The green line is the $\chi^2$ value when the system is excited by white noise input and the blue line shows the $\chi^2$ value when the rotated input is implemented. Figure 1 shows that white noise can hardly reveal this fault, but the residual changes significantly using the rotated input when the fault happens.

The probability density function (PDF) of the $\chi^2$ value for the first 500 experience steps (blue bars) vs the second 500 steps (white bars), in the case of using the rotated input and white noise are demonstrated in Figure 2 and Figure 3, respectively. It can be concluded that the rotated input can reveal this fault, but the residual changes significantly using the rotated input.

The new method can even be more helpful to detect minor changes to which the traditional approach is not sensitive and does not detect anything. An example is the fault which leads to the following system

$$
\begin{pmatrix}
y_1(s) \\
y_2(s)
\end{pmatrix} = \begin{pmatrix}
\frac{87.8}{1+108.2s} & \frac{-87.8}{1+108.2s} \\
\frac{0.14}{1+194s} & \frac{0.14}{1+194s}
\end{pmatrix}
\begin{pmatrix}
u_1(s) \\
u_2(s)
\end{pmatrix}.
$$

Figure 4 depicts the $\chi^2$ value corresponding to this experiment. It can be clearly seen that the mean of the $\chi^2$ value changes in the faulty case using the new method.

The corresponding PDF plots are shown in Figure 5 and Figure 6 where the rotated input and white noise are implemented, respectively.
VI. CONCLUSIONS

Improving the subspace-based fault detection method, detecting the system order changes and strengthening the effect of weak modes on the residual are considered in this paper. It has been shown that for a group of systems, some fault cannot be detected using random inputs. The rotated inputs are implemented to better excite weak modes of the system and finally to get better detection results. Using the new method, the sensitivity of the residual to the fault is increased. Unlike the identification case, where usually there is no model of the system and the rotation angle should be found by trial and error, a nominal model exists for fault detection tests and the test input can be calculated precisely. Simulation results show the advantages of using the rotated input approach. In addition, the singular values corresponding to the weak modes are pushed up using the rotated inputs, the modes that are considered as the effect of noise and are neglected if random input is used. Therefore, it is possible to detect the system order changes, when a sever damage happens to the system and it loses part of dynamics.

REFERENCES