State Space Modelling of Electromagnetic Responses — A Practical Approach to Extract Parameters from Simulated or Measured Data

Krishna Naishadham*

Abstract—As computing power and algorithmic advances have evolved rapidly in the recent past, it is now feasible to solve complex electromagnetic (EM) problems involving scattering, radar cross section, antenna design, microwave circuit design, artificial EM materials, etc., using full-wave numerical methods. Several general-purpose commercial software packages are routinely used in industry in all these domains for EM analysis or design. However, the task of processing large sets of data output from these design studies and analyses is generally beyond the realm of commercial software packages, and the designer spends many hours writing problem-specific computer programs to extract the desired performance parameters. Some examples where auxiliary processing is needed for the extraction of EM parameters of interest include determination of coupling coefficients or the unloaded quality factor of a dielectric resonator, de-embedding feed lines from antenna currents, removal of discontinuity effects, and the extraction of equivalent circuit models. The same considerations as simulated data apply to the parametric analysis of measured data in the presence of noise. This paper presents a versatile data-driven spectral model derived from a state-space system representation of the computed or measured EM fields, from which all the parameters of interest can be extracted. An attractive feature of the state space method is its ability to identify a small number of system transfer function poles uniquely associated with a specific scattering mechanism or modal response, thereby enabling its isolation from the total response for detailed study. For example, using SSM, specular reflection and creeping waves on a smooth convex surface can be analyzed, and the diffraction at the edges can be isolated from the composite RCS of a large body. The desired field parameter is extracted or estimated from synthetic or measured data using a linear system of a relatively small model order that characterizes the specific modal response of interest. Illustrative examples will be presented to demonstrate the usefulness of the proposed approach for parametric extraction.

1. INTRODUCTION

With prolific algorithmic advances and significant improvement of computational resources in recent years, full-wave electromagnetic (EM) simulation techniques, such as the method of moments (MoM) [1], finite element method (FEM) [2], finite-difference time-domain (FDTD) method [3], and transmission line matrix method [4], have been increasingly used to solve complex EM problems involving scattering, radar cross section (RCS), antenna design, microwave circuit design, artificial EM materials, etc. These rigorous techniques account for physical phenomena such as surface-wave coupling and radiation, dispersion, frequency-dependent metallization and dielectric losses, proximity effects, and near-field coupling. Although EM simulation methods are computationally intensive, algorithmic advances and the large-scale GPU and distributed computing platforms enable commercially available full-wave EM analysis software packages to be applied to challenging problems such as scattering by electrically large objects (e.g., computing the RCS of an aircraft) and modeling the EMI between interconnects.
on a printed circuit board containing dense RF and/or high-speed digital circuits. The outputs of these software packages are generally voltages, currents, scattering (S) parameters or some derived EM field variable such as the RCS. In design studies, large datasets of these output parameters need to be processed, for example, to remove the effects of transmission line discontinuities on the S-parameters or to extract the modal response of a particular wave species and isolate the scattering centers causing undesirable RCS. Therefore, it is very desirable to investigate independent signal models (in time domain) and spectral models (in frequency domain), which accurately represent the simulated or measured output field responses. One can then utilize this model for removing the effects of discontinuities or parasitic scattering mechanisms so that the desired EM performance can be characterized accurately.

Several researchers have extracted parameters of interest from scattered fields or circuit response using signal processing techniques such as Prony’s method [5, 6], pencil of functions [7–11], autoregressive moving average (ARMA) [12, 13], estimation of signal parameters via rotational invariance techniques (ESPRIT) [14–16], multiple signal classification (MUSIC) [17], and the state space method (SSM) [18–21]. Applications include computation of complex natural resonances and eigenmodes [22–28], impulse response characterization of time-domain signatures [29–34], broadband equivalent circuit parameter extraction [35, 36], identification of radar target’s features [37–40], extraction of biomedical vital signs from UWB radar measurements [41–43], and location of buried targets using ground penetrating radar [44]. The basis behind such signal processing applications is that the EM field scattered by an object can be adequately represented as a sum of damped sinusoids, whose amplitude and phase are closely related to the physical parameters of interest. For example, poles of the transfer function for the exponential signal model locate discrete scattering centers useful in object typing and feature identification in radar target identification.

This paper presents a versatile data-driven spectral model derived from a state-space system representation of the computed or measured scattering parameters and EM fields, from which all the parameters of interest can be extracted. Parametric extraction from measured data is especially challenging because of noise and random measurement errors. The efficacy of spectral estimation methods to extract parameters of damped sinusoids embedded in noise has been studied by many researchers. For example, performance analysis of MUSIC is treated in [45, 46]; estimation of the direction of arrival of radar returns using ESPRIT is presented in [47]; a performance study of matrix pencil method in the presence of noise is described in [48, 49], and sensitivity analysis of the state space method is treated in [50, 51]. SSM [18–21] has been extended to ultra-wideband coherent processing of range-Doppler data for radar target identification and validated with static range measurements [52, 53]. In EM problems, the matrix pencil method has been widely used (cf. [25, 26, 32]). A quantitative comparison between state space and matrix pencil methods shows similar performance with comparable accuracy when being implemented on canonical harmonic retrieval problems with noisy data [49, 54, 55]. It is emphasized that the current paper focuses only on SSM, and a review of other signal processing methods applied to EM problems is not undertaken. The interested reader refers to the relevant references cited above.

The state space method can be applied in either time domain or frequency domain. In this paper, we focus on the frequency domain problems. The reader refers to [34] for the application of SSM to time-domain EM problems. The state space system identification approach in [34] is derived from the subspace identification method for linear systems introduced by Overschee and De Moor in [58]. In this paper, we apply the subspace system identification method to frequency domain. We describe the problem formulation and the modeling process in detail so that an uninitiated reader can follow the steps to program and execute the algorithm and generate results of interest. An attractive feature of SSM is its ability to identify and associate a small number of poles of the system transfer function with a specific scattering mechanism or modal response, such as scattering parameters at the transition (discontinuity) between a coplanar connector and a microstrip transmission line in measured data [36], and isolated scattering by edges and seams from the composite RCS of a large body [40, 52]. Thus, the desired field parameter can be extracted or estimated from synthetic or measured data using a linear system with relatively small model order. Illustrative examples will be presented to demonstrate the usefulness of the SSM for parametric extraction in EM problems involving simulated as well as measured

\[ \text{The author is thankful to a reviewer for pointing this out.} \]
data.

The first two problems pertain to simulated data. First, we consider a planar dielectric slab illuminated by a plane wave at normal incidence and isolate the reflection off the front face using range-classified poles pertinent to the specular reflection. This process is akin to de-embedding transmission lines at the ports to evaluate the circuit behavior of an antenna or a discontinuity. In the second example, appealing to the canonical problem of Mie scattering by a sphere, creeping waves are extracted using SSM to model the RCS response and validated against high-frequency asymptotic approximations. Next, random white Gaussian noise is added to the Mie series solution, and Monte Carlo simulation is performed to examine robustness of the SSM estimates to noise. Numerical considerations such as dynamic range, signal-to-noise ratio (SNR), and model order determination are addressed in detail for both examples.

The application of SSM to the measured data on relatively complex EM problems, including canonical radar target scattering, biomedical radar and microwave circuits, is reviewed. A state-space spectral model is presented for the parametric characterization of low-loss wideband transmission lines enclosed in a MEMS package. SSM isolates and removes discontinuity effects at the CPW-microstrip transitions leading to the line, and filters out measurement noise, enabling accurate extraction of distributed circuit parameters, such as propagation constant and characteristic impedance.

The paper is organized as follows. In Section 2 the state space algorithm is reviewed in detail following [53, 56]. Section 3 presents basic numerical considerations on the state space approach, such as estimation of the SNR and the model order. Section 4 presents illustrative examples on parametric extraction for simulated EM problems, as discussed above. Factors affecting accuracy and numerical efficiency of the proposed technique are discussed. Section 5 reviews the application of SSM to the measured data alluded to earlier. Finally, the paper is summarized and concluding remarks are presented in Section 6.

2. STATE SPACE METHOD

In recent years, there has been a great deal of attention devoted to model-based eigen-decomposition methods derived from a state-space realization of the system identification problem, which is defined as the determination of the internal states of a linear time-invariant (LTI) system given a set of inputs and outputs [18, 21, 52–62]. The idea of using state-space methods to estimate frequencies and amplitudes of damped sinusoids was first suggested in [18], where Kung et al. developed a system identification approach based on singular value decomposition (SVD) for the harmonic retrieval (or spectral estimation) problem. The foundation of state-space signal modeling is based on representation of a linear rational system, popular in linear systems and control theory [63–65], in which the difference equations for the discrete-time signal are converted into state equations and the model parameters are estimated in terms of the state matrices characterizing the system. State-space parameterization enables reduction of parameter sensitivity, as demonstrated by several examples relevant to the sinusoid retrieval problem in the tutorial article by Rao and Arun [21]. Unlike polynomial-based signal processing methods such as Prony’s, MUSIC and maximum likelihood estimation, in which the frequencies of damped sinusoids are obtained by cumbersome root-search of polynomials, state-space methods simultaneously yield complex amplitudes (with initial phases) and frequencies directly from three state matrices. The frequencies are calculated from the eigenvalues of the state transition matrix, and the amplitudes are derived by eigen-decomposition of the state equations using two auxiliary matrices, the control and observation matrices.

As shown in the sequel, the decomposition of the data into a sum of damped sinusoids representing an LTI system enables direct computation of the state matrices from a low-rank truncation of the Hankel (or forward-prediction) matrix representing the data. We show that each entry of the Hankel matrix can be expressed in terms of the impulse response derived from the state space matrices. The SVD of the Hankel matrix is a product of the observability and controllability matrices which yield all the model parameters in the harmonic retrieval problem.
2.1. ARMA Signal Model

The scattered field output data sequence \( y(k) \) comprises \( N \) uniformly spaced frequency samples (see Eq. (1) below), each represented as a sum of \( M \) complex sinusoids (or scattering centers) corrupted by measurement noise \( w(k) \), assumed to be white Gaussian with zero mean. In deterministic data modeling, we assume that \( w(k) = 0, k \neq 0 \) and characterize the impulse response, thus estimating the system parameters using only the output. When the input is known and needs to be considered (e.g., short pulses or modulated waveforms), the system can be identified using both input \( u(k) \) and output \( y(k) \) [34]. For completeness, we retain \( w(k) \) in the formulation even if we may not consider it in a given case. Thus, over a given bandwidth, the signal measurements at \( N \) frequencies are modeled as

\[
y(k) = \sum_{i=1}^{M} a_i p_i^k + w(k); \quad k = 1, \ldots, N, \tag{1}
\]

The difference between the “true” signal, \( y(k) \), and the state space model, \( \hat{y}(k) \), is the random noise, measurement error or modeling error, \( w(k) \). The signal \( y(k) \) represents the response of a linear time invariant system where \( \hat{y}(k) \) and \( w(k) \) correspond to the homogeneous solution and the particular solution, respectively. The poles \( p_i \) are given by

\[
p_i = \exp \left[ -\Delta f (\alpha_i + j 2\pi \tau_i) \right] \tag{2}
\]

where \( a_i \) (in Eq. (1)) and \( \alpha_i \) represent the amplitude and its spectral rate of decay (for negative \( \alpha_i \)) or growth (for positive \( \alpha_i \)), respectively, for the \( i \)-th scattering center, and \( \Delta f \) is the sampling frequency. Both positive and negative \( \alpha_i \) are needed to model the peaks and dips in the frequency response (for example at resonance). It is emphasized that positive \( \alpha_i \) does not imply instability or violation of any physical constraint such as the radiation condition. The parameter \( \tau_i \) denotes time delay of the \( i \)-th scatterer at the observation point and is related to range \( R_i \) by \( \tau_i = 2R_i/c \), where \( c \) is the speed of light. Equation (1) is set up in terms of the monostatic RCS that an observer would measure for a given target at range \( R_i \). Therefore, \( 2R_i \) is simply the round-trip distance between the transmitter and the target. If bistatic radar is used, then \( 2R_i \) should be replaced by \( R_t + R_r \), where \( R_t \) and \( R_r \) denote the distance to the target from the transmitter and the receiver, respectively. It follows that \( 2\pi \tau_i f_k = (\omega_k/c) (2R_i) = \beta_k (2R_i) \) represents the phase of the \( i \)-th sinusoid, with \( \beta_k \) being the phase constant at the frequency \( f_k \). The frequency vector is specified in terms of the carrier frequency \( f_1 \) as

\[
f_k = f_1 + (k-1)\Delta f, \quad k = 1, \ldots, N, \tag{3}
\]

where \( \Delta f \) is the sampling frequency. The primary interest in state space system identification is to estimate the parameters \( a_i \), \( \alpha_i \), and \( R_i \) (or \( \tau_i \)), which are embedded in the data sequence \( y(k) \). The latter two parameters are computed from the eigenvalues of an open-loop state matrix, to be defined shortly. Once these parameters are estimated, the amplitudes \( a_i \) can be readily derived from the state matrices using a modal decomposition method based on least squares [53].

2.2. ARMA Transfer Function

The model in Eq. (1) is interpreted as an LTI system with input given by the sequence \( u(k) \) and the output by the data sequence \( y(k) \). The goal in system identification is to compute the coefficients of the ARMA transfer function (TF) characterizing the discrete signal model in Eq. (1). Since the transfer function is defined for impulse response, we first set \( w(k) = \delta(k) \) in Eq. (1) and take the \( z \)-transform after substituting \( f_k \) from Eq. (3), to obtain

\[
Y(z) = \sum_{i=1}^{M} \frac{B_i}{1 - p_i z^{-1}}, \quad \tag{4}
\]

\[
B_i = a_i e^{\alpha_i f_1} e^{j 2\pi \tau_i f_1} = A_i e^{j \theta_i} \tag{5}
\]
In the above, \( B_i \) denote the complex amplitudes; \( \theta_i \) is the initial (static) phase; and \( p_i \) are the poles in the complex \( z \)-plane given in Eq. (2). The finite summation in Eq. (4) leads to a TF comprising rational polynomials in the numerator and denominator [10], given by

\[
T(z) = \frac{Y(z)}{W(z)} = \frac{c_0 + \sum_{i=1}^{M-1} c_i z^{-i}}{1 - \sum_{i=1}^{M} d_i z^{-i}} \triangleq \frac{N(z)}{D(z)} \tag{6}
\]

The TF \( T(z) \) has \( M \) poles and \( M - 1 \) zeros, located at the roots of \( D(z) \) and \( N(z) \), respectively. The TF in Eq. (6) represents a special case of a more general ARMA TF given by

\[
T(z) = \frac{c_0 + \sum_{i=1}^{Q} c_i z^{-i}}{1 - \sum_{i=1}^{P} d_i z^{-i}} \triangleq \frac{N(z)}{D(z)} \tag{7}
\]

which has \( P \) poles and \( Q \) zeros (\( P \) and \( Q \) can assume arbitrary integer values). The special case \( Q = P - 1 \) listed in Eq. (6) is also known as “strictly proper” TF and the Prony’s model belongs to this category. Another important special case is the purely Auto Regressive (AR) TF with \( Q = 0 \).

The input-output relationship for the general ARMA model with the TF in Eq. (7) is characterized by the difference equation [10]

\[
y(k) = \sum_{i=1}^{P} d_i y(k - i) + \sum_{j=1}^{Q} c_j w(k - j) + c_0 w(k) \tag{8}
\]

Eq. (8) is valid for the TF in Eq. (6) as well with \( P = M \) and \( Q = (M - 1) \). From linear systems and control theory [63, 64], one can show that the difference equation in Eq. (8) may be written alternatively in terms of the state-space description characterized by the difference equations

\[
\begin{align*}
x(k + 1) &= Ax(k) + Bw(k) \tag{9} \\
y(k) &= Cx(k) + w(k), \tag{10}
\end{align*}
\]

where \( x(k) \in C^{M \times 1} \) is the state vector; \( A \in C^{M \times M} \) is the state transition matrix; \( B \in C^{M \times 1} \) and \( C \in C^{1 \times M} \) are constant matrices known as control matrix and observation matrix, respectively [64]. Our goal is to identify the matrices \( A, B, \) and \( C \) given the data sequence \( y(k) \) and the input \( w(k) \). The transfer function, \( T(z) \), is obtained by taking the \( z \)-transform of Eqs. (9) and (10) with \( w(k) = \delta(k) \) and evaluating the ratio \( Y(z)/W(z) \):

\[
T(z) = C(zI - A)^{-1}B + 1. \tag{11}
\]

It follows from Eq. (11) that the poles of the model (i.e., the roots of \( D(z) \) in Eq. (7)) are the eigenvalues of the open-loop matrix \( A \) and the zeros (i.e., the roots of \( N(z) \) in Eq. (7)) are the eigenvalues of the matrix \( (A - BC) \) [52, 53].

### 2.3. State Space System Identification

The transfer function in Eq. (7) can be written equivalently in terms of its impulse response sequence as

\[
T(z) = h(0) + h(1)z^{-1} + \ldots + h(n)z^{-n} + \ldots \tag{12}
\]

which is aptly referred to as the infinite impulse response (IIR) transfer function [64]. The expansion of the inverse matrix in Eq. (11) into an infinite series yields

\[
(zI - A)^{-1} = I z^{-1} + A z^{-2} + A^2 z^{-3} + \ldots \tag{13}
\]
where $I$ is $M \times M$ identity matrix. By inserting Eq. (13) into Eq. (11), we obtain

$$T(z) = 1 + CBz^{-1} + CABz^{-2} + CA^2Bz^{-3} + \ldots$$

Equating coefficients of like-powers of $z$ in Eqs. (12) and (14) and realizing that $h(k) = y(k)$ for the impulse input, we obtain

$$y(0) = 1$$
$$y(1) = CB$$
$$y(3) = CAB$$
$$\vdots$$
$$y(k) = CA^{k-1}B$$
$$\vdots$$

Therefore, the relationship between the impulse response of the model and the state-space parameters for any positive value of $k$ is defined by

$$y(k) = CA^{k-1}B, \quad k > 0.$$  \hspace{1cm} (16)

Eq. (16) indicates that a Hankel (or forward-prediction) matrix, $H$, formed from the IIR sequence of a system as

$$H = \begin{bmatrix} y(1) & y(3) & y(4) & \ldots \\ y(3) & y(4) & y(5) & \ldots \\ y(4) & y(5) & y(??) & \ldots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \triangleq \Omega \Gamma,$$  \hspace{1cm} (17)

can be decomposed into a product of two matrices given by

$$H = \begin{bmatrix} C \\ CA \\ CA^2 \end{bmatrix} \begin{bmatrix} B & AB & A^2B & \ldots \end{bmatrix} \triangleq \Omega \Gamma,$$  \hspace{1cm} (18)

where $\Omega$ and $\Gamma$ are known as observability and controllability matrices, respectively [65]. It is important to note that despite the infinite dimensions of $H$ in Eq. (17), in practice, the impulse response is always finite. Thus, for a given set of measurements the rank of the Hankel matrix $H$, and by inference the rank of $\Omega$ and $\Gamma$, will always be finite. As described later in this Section, $\Omega$ and $\Gamma$, consequently $H$, can be truncated to a low rank $r \leq M$, where $M$ is the number of complex sinusoids in the model (see Eq. (1)). In summary, the Hankel matrix $H$ may be interpreted as an operator constructed from a set of measurements $y(k)$ that maps the past input vector $w^-$ to the future output $y^+$. Causality of SSM is explicit in Eqs. (8)–(10), which emphasize that the current output is dependent on the past output and the current as well as the past input signals. Next, we present a method to derive the state-space matrices from the Hankel matrix constructed using a finite set of output data samples.

The first step in computing the triplet $(A, B, C)$ is to form the Hankel matrix using available data samples $y(k), k = 1, 2, \ldots, N$.

$$H = \begin{bmatrix} y(1) & y(3) & \ldots & y(L) \\ y(3) & y(4) & \ldots & y(L+1) \\ \vdots & \vdots & \vdots & \vdots \\ y(N - L + 1) & y(N - L + 2) & \ldots & y(N) \end{bmatrix}$$  \hspace{1cm} (19)

where the parameter $L$ denotes length of the correlation window, heuristically chosen to be $L = \lfloor N/2 \rfloor$, and the brackets denote the smallest integer less than or equal to the inserted quantity. Note that the Hankel matrix in Eq. (19) is a truncated version of the IIR in Eq. (17). Subspace decomposition
methods exploit the eigenstructure of Hankel matrices to compute the state matrices of the LTI system and estimate the signal models [18, 21, 25, 26, 52–56, 60–62]. Accordingly, we partition the Hankel matrix into signal and noise subspaces using the singular value decomposition of \( H \). Organizing the SVD in terms of the singular values \( \{ \sigma_i \} \) in descending order of magnitude, the Hankel matrix may be written as

\[
H = \begin{bmatrix} U_s & U_n \end{bmatrix} \begin{bmatrix} \Sigma_s & 0 \\ 0 & \Sigma_n \end{bmatrix} \begin{bmatrix} V_s^* \\ V_n^* \end{bmatrix} = U \Sigma V^*
\]  

(20)

where the subscripts \( s \) and \( n \) denote the signal and noise subspaces, respectively, and the asterisk refers to matrix conjugate transpose. The matrices \( U_s \) and \( U_n \) are the signal and noise components, respectively, of the left-unitary matrix \( [U_s \ U_n] \). Likewise, \( V_s \) and \( V_n \) denote the signal and noise components, respectively, of the right-unitary matrix \( [V_s \ V_n] \). Furthermore, \( \Sigma_s \) and \( \Sigma_n \) are diagonal matrices comprising signal and noise singular values, respectively. It is understood that the signal components in the SVD are entirely characterized by the dominant singular values \( \Sigma_s \). The classification between signal and noise subspaces is achieved by parsing the singular value spectrum in descending order of magnitude and removing the noise singular values \( \Sigma_n \) [18, 50–53]. To increase the accuracy of the state-space estimation, the Hankel matrix \( \hat{H} \) may then be truncated by suppressing the noise singular values and their associated unitary matrix components. This results in a reduced-rank approximation to the Hankel matrix in Eq. (20), obtained by retaining only the dominant singular values (cf. [51]), i.e.,

\[
\hat{H} = U_s \Sigma_s V_s^*.
\]  

(21)

Let the computed singular values in \( \Sigma \) be arranged in descending order of magnitude

\[
\sigma_1 > \sigma_2 > \ldots > \sigma_r > 0
\]  

(22)

where \( r \) is chosen subject to the threshold [55] \( \sigma_r / \sigma_1 \approx 10^{-p} \), and \( p \) is the number of significant decimal digits in the truth data. For example, if the data are accurate to three significant digits, then the singular values with upper bound \( r \) for which the ratio is less than 0.001 are considered as noise singular values and excluded from the model. The largest index \( r \) of the singular values in the signal subspace is the rank of the Hankel matrix. It can be shown that the largest retained singular value minimizes the error between the Hankel matrices \( H \) and \( \hat{H} \) in the spectral norm sense [49], i.e.,

\[
\sigma_r \approx \left\| H - \hat{H} \right\|_s,
\]  

(23)

where the subscript \( s \) denotes the spectral or \( L_2 \) norm [64]. As demonstrated in Section 3, magnitude of the dominant singular values in the signal subspace may be conveniently used to estimate the “optimal” model order \( M \) in Eq. (1). Next, we address how one can compute the state matrices from the SVD.

Akin to Eq. (18), \( \hat{H} \) is obtained in factored form as

\[
\hat{H} = U_s \Sigma_s V_s^* = \hat{\Omega} \hat{\Gamma}
\]  

(24)

By using the balanced coordinate transformation method proposed in [18], one can compute the finite-rank observability matrix \( \hat{\Omega} \) and controllability matrix \( \hat{\Gamma} \) from the SVD in Eq. (24). These matrices are given by

\[
\hat{\Omega} = U_s \Sigma_s^{1/2} \quad \text{and} \quad \hat{\Gamma} = \Sigma_s^{1/2} V_s^*.
\]  

(25)

Then, the open-loop matrix \( A \in \mathbb{C}^{M \times M} \) can be derived from either \( \hat{\Omega} \) or \( \hat{\Gamma} \) as shown in Appendix A. If the derivation of \( A \) is based on the observability matrix \( \hat{\Omega} \), then [40]

\[
A = \left( \hat{\Omega}_{-r \ell}^* \hat{\Omega}_{-r \ell} \right)^{-1} \hat{\Omega}_{-r \ell}^* \hat{\Omega}_{-r 1}.
\]  

(26)

The matrices \( \hat{\Omega}_{-r 1} \) and \( \hat{\Omega}_{-r \ell} \) in Eq. (26) are obtained by deleting the first and last rows, respectively, of \( \hat{\Omega} \). Alternatively, in terms of the controllability matrix, \( A \) is given by [40]

\[
A = \hat{\Gamma}_{-c 1} \hat{\Gamma}_{-c \ell}^* \left( \hat{\Gamma}_{-c \ell} \hat{\Gamma}_{-c \ell}^* \right)^{-1}
\]  

(27)

where the matrices \( \hat{\Gamma}_{-r 1} \) and \( \hat{\Gamma}_{-r \ell} \) are obtained by deleting the first and last rows, respectively, of \( \hat{\Gamma} \). The reader is referred to Appendix A for the least-squares computation of \( B \) and \( C \), also using either
As noted earlier, the matrices $B$ and $C$ contribute to zeros of the transfer function. In the AR SSM that we have employed in [39], system zeros are not used, because the system is entirely characterized by the poles determined from $A$. In general, if the EM frequency response is monotonic, an AR model would suffice. Practical systems such as microwave resonant structures and antennas have non-monotonic responses, and it is desirable to use the complete model including the zeros.

2.4. Modal Decomposition

Once the state matrices ($A, B, C$) are known, the model parameters in Eq. (1) may be computed using a modal (or eigen-) decomposition method [52, 53]. If the complex eigenvalues of $A$ are assumed to be distinct, one has

$$\lambda \{A\} = \{ \lambda_1, \lambda_2, \ldots, \lambda_M \}.$$  

The magnitude of the eigenvalues $\{\lambda_i\}$ determines amplitude decay/growth rate (with respect to frequency) of the system output, and their phase provides the time delay $\tau_i$ or the range $R_i$ in Eq. (1). As noted in Eq. (11), the eigenvalues of $A$ represent the poles of the transfer function, $T(z)$. The eigenvalues $\{\lambda_i\}$ and eigenvectors $[\psi_i]$ of the state matrix $A$ satisfy

$$A\psi_i = \lambda_i \psi_i, \quad i = 1, \ldots, M \tag{29}$$

Now let us form a modal matrix $\Psi_i$ from these eigenvectors by stacking them row-wise:

$$\Psi = [\psi_1 \psi_2 \ldots \psi_M] \tag{30}$$

It follows from Eqs. (28)–(30) that

$$A\Psi = \Psi \Lambda, \quad \Lambda = \text{diag} \{ \lambda_1, \lambda_2, \ldots, \lambda_M \} \tag{31}$$

Therefore, we may calculate $\Lambda$ from Eq. (31) as

$$\Lambda = \Psi^{-1} A \Psi \tag{33}$$

The entries on its main diagonal are exactly the eigenvalues of the state transition matrix $A$. Therefore, the modal decomposition approach provides valid information to identify and characterize the discrete identify point scatterers (sinusoids) embedded in the data set. The impulse response approximation, $\hat{y}(k)$, of the data sequence $y(k)$ can be computed in terms of the state matrices using Eq. (16) as

$$\hat{y}(k) = C A^{k-1} B, \quad k = 1, 2, \ldots, N. \tag{34}$$

After substituting for $A$ from Eq. (31) and realizing that $A^n = \Psi \Lambda^n \Psi^{-1}$, we obtain

$$\hat{y}(k) = C \Psi \Lambda^{k-1} \Psi^{-1} B. \tag{35}$$

The amplitudes of the scatterers in Eq. (1) may be computed from Eq. (34). Let

$$\Psi^{-1} = [\nu_1 \nu_2 \ldots \nu_M]^T \tag{36}$$

where superscript $T$ denotes the matrix transpose. After inserting Eqs. (30), (32), and (36) into Eq. (34), we obtain

$$\hat{y}(k) = \sum_{i=1}^{M} (C \psi_i) (\nu_i B) \lambda_i^{k-1}, \quad k = 1, \ldots, N \tag{37}$$

Upon comparison with Eq. (1), we deduce that the magnitude and phase of the eigenvalues (poles) are related to the model parameters $\alpha_i$, $\tau_i$, and $R_i$, respectively, by

$$\alpha_i = \frac{\log |\lambda_i|}{\Delta f}, \quad \tau_i = \frac{\phi_i}{2\pi \Delta f}, \quad \text{and} \quad R_i = \frac{c \phi_i}{4\pi \Delta f}, \quad i = 1, \ldots, M \tag{38}$$

In Eq. (38), $\phi_i$ refers to phase of the eigenvalue $\lambda_i$. The sinusoid amplitudes $a_i$ defined in Eq. (1) are then obtained from Eq. (37) using these eigenvalues, the corresponding eigenvectors, and the state matrices $B$ and $C$:

$$a_i = \frac{(C \psi_i) (\nu_i B)}{(\lambda_i)^{\frac{1}{\Delta f}}}; \quad i = 1, \ldots, M. \tag{39}$$
The frequency dependence of the amplitudes and their decay/growth rate is emphasized in Eqs. (38) and (39).

Next, we discuss the physical interpretation of the state space parameters. The correspondence between the scattered field response and its state space representation in terms of complex sinusoids becomes apparent when one compares Eqs. (1) and (37). The poles $\lambda_i$ of the transfer function yield the localized scattering centers on the scattering object. As derived in Appendix A, the matrices $B$ and $C$ follow from least squares solution applied to linear systems represented by the observability and controllability matrix, respectively (see Eqs. (A15) and (A19)). Therefore, the matrices $B$ and $C$ are directly related to the observability and controllability of the linear system in Eq. (34). The observability matrix $\tilde{\Omega}$ affects the shape of the natural response of the scattered field (directly proportional to the natural modes or eigenvectors of the state transition matrix $A$) within a given frequency band. The controllability matrix $\tilde{\Gamma}$ limits the effect of an input such as an impulse on the natural mode. In other words, the controllability matrix determines how much of the excitation is coupled to the natural modes of a particular scattering center within the defined bandwidth.

3. NUMERICAL CONSIDERATIONS

3.1. Estimation of the Model Order

In any spectral estimation problem, the selection of model order of the underlying system is a critical issue. The difference between the desired and modeled signal, as defined in Eq. (1), is a combination of model mismatch error, measurement error, and noise. If this error is Gaussian distributed, then the minimization of the error criterion in Eq. (23) represents the maximum likelihood estimate (MLE) for the ARMA problem [18, 66]. In the case of MLE, the Akaike information criterion (AIC) or minimum description length (MDL) criterion are often used for model order determination [67–69]. An estimator that yields the true number of signals with probability one, as the sample size increases to infinity, is said to be “consistent.” It has been shown in [68] that MDL is a consistent estimator of order, whereas AIC is inconsistent and often overestimates the model order.

SSM has been used with simulated as well as measured data. In deterministic data modeling such as simulated data where the SNR is quite high, we employ the singular value matrix $\Sigma$ to estimate model order in the SSM. The large singular values in $\Sigma$ correspond to strong signal components, while the smaller values are generally attributed to noise. For low noise levels, there is a sharp transition between the large and small singular values. This transition point can be used as an estimate of the model order. At higher noise levels, especially for measured data, the transition from large to small singular values may not be well defined, making model order estimation more difficult. In this case, probabilistic methods such as AIC and MDL criteria have been used for model order estimation (cf. [33]). No general criterion exists to determine which of these two methods gives the “optimal” model order, and none of them provide necessary and sufficient conditions to guarantee that all the required signals are included in the model. It is expedient to consider the underlying physics of the scattering mechanism to guide the determination of the number of scatterers using a combination of these methods and to examine the goodness of fit for each model order by computing the mean square error in the model [40]. For convenience, the equations to compute AIC and MDL from the singular values, $\sigma_i$, of the Hankel matrix approximation, $\tilde{H}$, are summarized below [68].

$$AIC(r) = -2(M - r)N \ln \left( \prod_{i=r+1}^{M} \sigma_i^{-(M-r)} \right) + 2r(2M - r)$$

$$= -2(M - r)N \ln \left( \frac{1}{M - r} \sum_{i=r+1}^{M} \sigma_i \right) + 2r(2M - r)$$

(40)
\[
MDL(r) = -(M - r)N \ln \left( \prod_{i=r+1}^{M} \sigma_i^{-1} \right) + \frac{1}{2} r(2M - r) \ln N \tag{41}
\]

In the above, \( M \) is the maximum model order in Eq. (1), and \( N \) is the number of data samples. For a given number \( r \) of signals, the term in the parentheses simply denotes the ratio of geometric mean to arithmetic mean of the smallest \( M - r \) singular values, and it is a measure of the error between the truth and the model. The number of signals, i.e., the model order, is determined as the value of \( r \in \{1, 2, \ldots (M - 1)\} \) for which either the AIC or the MDL is minimized [68].

As an example, Fig. 1 shows the singular value spectrum (amplitudes in descending order of magnitude) plotted against the model order for experimental RCS data on a right circular conical stationary metallic target with and without a lossy dielectric coating [40]. It is seen that the transition from strong signal components to small singular values is not well defined because of measurement noise. AIC yields orders of 15 and 13 for bare (metallic) and dielectric-coated cones, while MDL yields model orders of 4 and 7, respectively. Due to measurement noise, AIC usually gives a larger estimate than MDL, and modelers usually choose an order no smaller than the AIC prediction. Therefore, an order of 15 has been chosen in [40] to extract the wave features of interest for both metallic and coated cones.

![Figure 1. Singular value spectrum vs. model order for measured RCS on a cone. Circles denote MDL and diamonds denote AIC [40].](image)

Figure 1. Singular value spectrum vs. model order for measured RCS on a cone. Circles denote MDL and diamonds denote AIC [40].

### 3.2. Estimation of Signal-to-Noise Ratio

The SNR of the radar system is influenced by noise contributed by several sub-systems, such as oscillator phase noise, mixer 1/f noise, thermal noise, antenna leakage, and mixer leakage. It is cumbersome and often inaccurate to estimate these random noise levels in the radar sub-systems using circuit models [70]. SSM considers this noise through decomposition of the Hankel matrix into signal and noise subspaces using the singular value decomposition.

The threshold for such decomposition is based on the dynamic range and SNR, which can be estimated from power spectrum of the original measured or simulated data, as discussed next. First, we consider measured data on the chest wall displacement of a human subject to extract the heart and respiration rates [41]. UWB radar with a center frequency of 2.4 GHz, bandwidth of 2 GHz, and pulse repetition frequency (PRF) of 75 Hz is used to collect the data. The complex signal \( I + jQ \) of the radar returning from a human subject located 1 m from the radar is plotted in Fig. 2 [41]. Random fluctuations
caused by measurement noise are evident. The complex signal in Fig. 2 is transformed with FFT using a Hamming window to suppress the sidelobes. This does not affect the peak signal but clearly defines the noise floor relative to the peak. Therefore, the dynamic range and SNR can be estimated from the compressed (or transformed) signal plotted as a function of Doppler frequency proportional to the chest wall displacement. Fig. 3 shows that the SNR for Channel 3 data is moderately high (around 24 dB), as one may expect from indoor laboratory RCS measurements, which are not subject to multipath and clutter encountered by fielded radar systems. Next, we consider SNR evaluation for simulated data.

In simulated data, the difference between the desired and modeled signals, as defined in Eq. (1), is a combination of model mismatch error and any noise intentionally added to the signal to determine statistical effects of random variation in parameters (e.g., Monte Carlo trials). In the absence of any added noise, the model mismatch error can be treated as noise. The SNR can be defined as the ratio of variance of the signal data sequence to the variance of the noise sequence,

\[
SNR = 10 \log \left( \frac{V \{ \hat{y}(k) \}}{V \{ y(k) - \hat{y}(k) \}} \right), \quad k = 1, 2, \ldots, N
\]

(42)

where \( V \) stands for the variance defined by

\[
V \{ f(k) \} = \frac{1}{N-1} \sum_{k=1}^{N} |f(k) - \mu|^2,
\]

(43)

and \( \mu \) is the mean of \( f(k) \), given by

\[
\mu = \frac{1}{N} \sum_{k=1}^{N} f(k).
\]

(44)

In a noise-free situation, if the SSM estimates defined by Eq. (34) closely match the data samples, the SNR defines the dynamic range of the model. Thus, the number of scatterers or signals giving rise to the estimates corresponds exactly to the number of sinusoids embedded in the data. In a noisy situation, it is important to use caution so that the estimates are not direct replicas of the noisy samples. Therefore, an “optimal” model order should provide appropriate SNR or dynamic range for a sinusoid being modelled. This will be illustrated by examples next.

4. RESULTS AND DISCUSSION

In this section, two examples based on simulated data are presented to illustrate the state space method. In the first example, we consider a planar dielectric slab illuminated by a plane wave at normal incidence.
and isolate the reflection off the front face using range-classified poles pertinent to the specular reflection. The subsequent radar-return signals also enable the classification of multiple internal reflections. In the second example, we consider Mie scattering from a sphere in the absence of noise, extract the creeping wave that circumnavigates the shadow zone and returns to the radar, and examine accuracy of the SSM estimates by comparison with an analytically derived expression of the creeping wave. Next, random white Gaussian noise is added to the Mie series solution, and Monte Carlo simulation is performed to examine robustness of the estimates to noise. Numerical considerations such as SNR, singular value spectrum, and model order estimation are addressed in detail for both of these examples.

4.1. Reflection by a Dielectric Slab

Plane wave reflection from a 15-mm thick lossy dielectric slab with dielectric constant \( \varepsilon_r = (5, -0.01) \) is considered, and it is shown that the response of the front face of the slab as well as multiple internal reflections can all be isolated from the composite response using SSM. Each of these contributions is specifically mapped to a range-gated pole. The slab of thickness \( d \) is located in free space. Its Fresnel reflection coefficient is calculated as

\[
y(\omega) = \Gamma = \frac{(\eta_1^2 - \eta_0^2) \tanh (\gamma_1 d)}{2 \eta_1 \eta_0 + (\eta_1^2 + \eta_0^2) \tanh (\gamma_1 d)},
\]

\[
\gamma_1 = j k_1, \quad k_1 = k_0 \sqrt{\varepsilon_r}, \quad k_0 = \omega \sqrt{\mu_0 \varepsilon_0},
\]

\[
\eta_1 = \eta_0 / \sqrt{\varepsilon_r}, \quad \eta_0 = 120 \pij.
\]

Without loss of generality, normal incidence is considered for illustration. The complex propagation constant and intrinsic impedance in the lossy dielectric are given by \( \gamma_1 \) and \( \eta_1 \), respectively. The phase constant and the intrinsic impedance of free space are denoted as \( k_0 \) and \( \eta_0 \), respectively, while \( \omega \) is the angular frequency. The reflection coefficient is calculated at normal incidence over 2–20 GHz bandwidth and Fourier transformed, using a Hamming window to suppress the sidelobes, to obtain the composite response shown in Fig. 4. The transform employs distance to the scatterer (or range) instead of time, an operation known as pulse compression.

Because the phase reference is at the front face, the pulse with peak at zero range is the reflection off the slab-front; the second pulse is the first reflection off the back face arriving coherently at the “receiver”; the third pulse is the second re-reflection off the back, and so on, as illustrated geometrically in Fig. 5 using transmission line analogy [71].

Next, a 10th order state-space model (see Eq. (1)) is estimated from the simulated data. The model parameters are listed in Table 1, where the poles and signal amplitudes are classified in terms
Table 1. SSM parameters for model order $M = 10$.

<table>
<thead>
<tr>
<th>Index</th>
<th>Pole Location</th>
<th>Complex Amplitude</th>
<th>Abs (Amp)</th>
<th>Range (m)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1.0000 - j0.0000$</td>
<td>$-0.3820 + j0.0004$</td>
<td>$0.381967$</td>
<td>$0.0000$</td>
<td>Front</td>
</tr>
<tr>
<td>2</td>
<td>$1.0000 - j0.0070$</td>
<td>$-0.3077 - j0.1057$</td>
<td>$0.325323$</td>
<td>$0.0150$</td>
<td>Refl. 1</td>
</tr>
<tr>
<td>3</td>
<td>$0.9999 - j0.0140$</td>
<td>$0.0374 + j0.0290$</td>
<td>$0.047331$</td>
<td>$0.0300$</td>
<td>Refl. 2</td>
</tr>
<tr>
<td>4</td>
<td>$0.9998 - j0.0211$</td>
<td>$-0.0038 - j0.0058$</td>
<td>$0.006886$</td>
<td>$0.0450$</td>
<td>Refl. 3</td>
</tr>
<tr>
<td>5</td>
<td>$0.9996 - j0.0281$</td>
<td>$0.0002 + j0.0010$</td>
<td>$0.001002$</td>
<td>$0.0600$</td>
<td>Refl. 4</td>
</tr>
<tr>
<td>6</td>
<td>$0.9993 - j0.0351$</td>
<td>$0.0000 - j0.0001$</td>
<td>$0.000146$</td>
<td>$0.0750$</td>
<td>Refl. 5</td>
</tr>
<tr>
<td>7</td>
<td>$0.9991 - j0.0421$</td>
<td>$0.0000 + j0.0000$</td>
<td>$0.000021$</td>
<td>$0.0900$</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>$0.9987 - j0.0492$</td>
<td>$0.0000 - j0.0000$</td>
<td>$0.000003$</td>
<td>$0.1050$</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>$0.9984 - j0.0562$</td>
<td>$0.0000 + j0.0000$</td>
<td>$0.000000$</td>
<td>$0.1200$</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>$0.9979 - j0.0632$</td>
<td>$0.0000 - j0.0000$</td>
<td>$0.000000$</td>
<td>$0.1350$</td>
<td></td>
</tr>
</tbody>
</table>

Figure 6. Corroboration between SSM model and the truth (Eq. (45)) using only the first five signals in Table 1. (a) Magnitude, (b) Phase.

of scatterers of interest highlighted in the last column. The rows are sorted in descending order of the peak amplitude. All the poles lie within or on the unit circle, validating the stability of the system.

It is interesting that the signal peaks are separated in SSM-computed range (see Eq. (38)) by exactly 15 mm, the thickness of the slab. The dominant pole, located at $(1, 0)$ with amplitude of 0.382 and phase $\pi$, is found to have zero range. This signal corresponds to the isolated front face reflection. The subsequent poles, consecutively spaced at 15 mm, represent the isolated secondary reflections off the back face, as illustrated in Fig. 4. The first five signals of interest in Table 1 appear to have the maximum spectral content, as the sixth term is three orders of magnitude smaller than the dominant signal. Using the coherent summation of only these first five terms in the state space model of Eq. (1), we have computed the frequency response of the composite signal. Fig. 6 compares the “truth” with the state space model in both magnitude and phase, and the two sets of curves overlap each other, signifying excellent model fidelity.

As alluded, one can isolate the front face response by using only the dominant pole at zero range. The reflection coefficient of this isolated pulse is the same as that of the half-space problem because the delayed reflections off the back face are not included. Therefore, using the constant amplitude of $\Gamma_f = -0.382 + j0.004$ from Table 1, the slab’s dielectric constant can be determined as

$$
\varepsilon_r = \left( \frac{1 - \Gamma_f}{1 + \Gamma_f} \right)^2.
$$

(47)
Eq. (47) yields $\varepsilon_r = 5 - j0.01$ exactly, verifying the accuracy of the signal extraction. It is imperative that isolation of individual signals of interest requires adequate bandwidth to provide the spectral resolution needed to separate the peaks.

When the thickness of the slab is reduced to 5 mm, one obtains the compressed range response shown in Fig. 7, which clearly shows the impending merging of the first two peaks. Using a bandwidth of 18 GHz, these two peaks are very close in amplitude and may not be adequately resolved in range, but when the bandwidth is increased to 28 GHz, the peaks can be isolated. Furthermore, in the latter case, the markers also indicate secondary fluctuations appearing at approximately 10 and 15 mm, corresponding to re-reflections off the back face. However, in measured data, such small fluctuations of the main lobe may be masked by noise, and a larger bandwidth may not necessarily improve the SNR.

![Figure 7. Range response of the 5-mm thick slab reflection coefficient.](image)

![Figure 8. Model error in magnitude of the SSM frequency response.](image)

Figure 8 plots the model error magnitude of the frequency-domain response for SSM with order $M = 5$ (the first five terms in Table 1) and $d = 15$ mm. The mean error is $-77.3$ dB, proving that all the significant signals have been modeled. A similar agreement between raw data and SSM is observed in the phase too, with mean error of $0.2^\circ$.

The model error for $M = 10$ and $M = 15$ does not decrease beyond the error depicted in Fig. 8 for $M = 5$. A performance measure for evaluation of the model error is the SNR in Eq. (42). Table 2 lists the SNR for a few model orders ranging from 5 to 15. It is evident that SSM with $M = 10$ gives the highest SNR, and any further increase in the order does not impact the error.

**Table 2. SNR as a function of model order for the slab problem.**

<table>
<thead>
<tr>
<th>Model Order</th>
<th>5</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNR (dB)</td>
<td>67.4</td>
<td>101.1</td>
<td>101.1</td>
</tr>
</tbody>
</table>

A deeper understanding of the sources of modeling error in SSM can be obtained by examining the error for various model orders relative to the compressed SSM signal. Along with the singular value spectrum, to be discussed shortly, this will help us in evaluating impact of the signals discarded in the model as “noise.” Fig. 9 displays the error responses for model orders $M = 5$ and $M = 10$, compressed using a 4,096-point FFT on sequences of length 3,601, with a Hamming window to suppress the noise sidelobes. For comparison, the compressed SSM signal estimate, identical for the two model orders, is also shown. The range coordinate for each peak (in mm) is annotated in the graph. We observe that the SSM signal, $\hat{y}(k)$, models the five dominant reflections signified by the first five poles (see Table 1), and the remaining signals are embedded in noise. Because these discarded signals are there in the truth data of Eq. (45), it is not surprising that the range spectra of the error signal, $y(k) - \hat{y}(k)$, depict these discarded noise peaks precisely. The number of such noise peaks embedded in the error sequence...
depends on the model order. For example, with $M = 5$, we calculate a noise floor of $-110$ dB (the mean square of the compressed error pulse) and identify the three noise peaks at 75, 90 and 105 mm. With $M = 10$, we can identify the noise peaks at 105, 120 and 135 mm.

The noise spectrum did not change with any further increase in model order beyond 10. Thus, the last noise peak that one can identify is at 135 mm. As the highest noise peak (at 75 mm) is 90 dB down from the main reflection at zero range, none of these discarded noise peaks has any influence on the SNR.

Lastly, we estimate the model order by examining the singular value spectrum plotted in Fig. 10 as a function of its index, $r$. The SVs are normalized to the largest value ($r = 1$). The model order can be determined subject to the threshold $56 \sigma_r/\sigma_1 \approx 10^{-p}$, where $p$ is the number of significant decimal places in the truth data. Thus, if the data are accurate to 7 places, the SV spectrum predicts a model order of $r = 10$. In practice, the data may be accurate to the third or fourth decimal place, but not till the seventh as the simulated reflection coefficient in Eq. (45) for this simple illustrative problem. For three-digit accuracy, Table 1 indicates that a model order of 10 yields 5 signal poles and 5 noise poles (weak signals embedded in noise). Indeed, the error analysis in Figs. 8 and 9 validates this observation.

### 4.2. Scattering by a Sphere

Let us consider the extraction of creeping waves from the canonical problem of scattering by a perfect electrically conducting (PEC) sphere using the SSM. The scattered field for plane wave incidence is computed analytically using the Mie series [72], first in the absence of noise, and then with the influence of noise.
of additive Gaussian noise. The model order is estimated in each case by examining the singular value spectrum, and the accuracy of the extracted creeping wave is established by comparison with an analytically derived asymptotic expression of the creeping wave [73].

4.2.1. Baseline without Noise

The monostatic RCS is calculated for a PEC sphere of radius \( a = 17.7 \) cm using the Mie series and is plotted against frequency over a range of 2 to 20 GHz in Fig. 11. As the radius becomes much larger than the wavelength, the normalized RCS asymptotically approaches \( \pi a^2 \). The contribution to back-scattered field comes from a specularly reflected ray emanating from point A, and a creeping wave, which attaches to the sphere tangentially at point B, navigates half the circumference and detaches at the symmetrical tangential location C, as shown in the inset in Fig. 11. The specular contribution is governed by geometrical optics (GO) approximation and can be analytically calculated [74]. The creeping wave diffracted field can be approximated by asymptotic evaluation of the Fock integrals [73, 74].

In order to isolate the scattering centers pertinent to the specular and the creeping wave, we first compute the range profile of the back-scattered field, displayed in Fig. 12, by FFT of the Mie series frequency response using a Hamming window to suppress the sidelobes. The phase reference is chosen such that the point of specular reflection is at zero range. The second peak which corresponds to the creeping wave is at 0.454 m range, in agreement with the range predicted from ray path geometry (see the inset of Fig. 11) as 0.455 m, i.e., \( R = (\pi/2 + 1)a \).

Next, SSM is applied to extract range-isolated poles specific to the creeping wave, and their cumulative frequency response is calculated. A 10th order state space model is computed from the Mie series frequency response shown in Fig. 11. We have observed excellent corroboration between the estimated SSM model and the truth data in Fig. 11 over the entire 18 GHz band. For brevity, we focus only on representation of the creeping wave using the corresponding range-isolated poles from the SSM. As described in Section 2, an advantage of using range processing with the spectral estimation method is the direct relation between range and pole phase (see Eq. (38)). Thus, one may isolate a given scattering mode by adding contributions from only the poles associated with the range window of that mode. Of the 10 poles used in the SSM to represent the entire signal in Fig. 11, only two poles are identified with model-computed range of 0.454 m, relevant to the creeping wave peak. By coherently summing the contributions of only these two poles, we obtain the extracted frequency response shown in Fig. 13.

In order to demonstrate baseline validation with no noise present in the data, we also plot in Fig. 13 an analytical solution for the spherical creeping wave from [73], which essentially overlaps with the SSM-estimated data. The signal amplitudes for the two creeping wave poles are observed to be around \(-36 \) dB, signifying high accuracy in the model even for small signals. Later, we will evaluate model robustness using Monte Carlo analysis with Gaussian noise added to the Mie series. The absolute model prediction error for the creeping wave estimation is plotted in Fig. 14. A similar agreement with the analytical solution is also observed for the specular peak extracted from the Mie series in Fig. 11.

![Figure 13. Comparison of SSM-extracted creeping wave with the analytical solution from [73].](image)

![Figure 14. Model prediction error for the extracted creeping wave relative to the analytical solution from [73].](image)
4.2.2. Monte Carlo Analysis

The SSM extraction of signals of interest from Mie series in the presence of noise is considered next. Monte Carlo (MC) analysis is performed by adding random noise of a given SNR to the Mie series solution for the PEC sphere and evaluating the accuracy of the SSM-extracted components, namely, the creeping wave and the specular. The measurement noise \( w(k) \) (see Eq. (1)) is assumed to be complex white Gaussian with variance \( \sigma_n^2 \), defined by peak signal-to-noise ratio

\[
SNR = 20 \log \left( \frac{\sigma_{sn}}{\sigma_n} \right),
\]

(48)

where \( \sigma_{sn}^2 \) denotes the signal variance. For a given SNR, 1000 independent trials are executed on the Mie series solution, and the state space method is applied to process each noise-corrupted data sequence and extract the wave constituents of interest. In order to assess the quality of the data, the noise-corrupted Mie series (truth) data is plotted in Fig. 15 against SSM model (with order \( M = 10 \)) of the total response for the mean of 1000 MC trials with \( SNR = 20 \) dB. It is observed that coherent averaging of the MC trials reduces the influence of noise considerably. As the electrical size of the sphere increases, the asymptotic limit of SSM correctly reaches \( \pi a^2 \).

Next, we consider the extraction of the specular wave. Fig. 16 depicts the frequency response corresponding to SSM processing of the average of 1000 MC trials, with SNR ranging from 5 dB to 20 dB. Only two poles are used in the SSM for the extraction. The response for each SNR is observed to track the analytical (noise-free) solution quite well. The worst-case error relative to the reference solution is about 0.05 dB and occurs for \( SNR = 5 \) dB. More importantly, because of the relatively large amplitude of the specular (about \(-10 \) dB), we have observed good correlation between corresponding pole locations for each trial. For brevity, the pole plots are not included.

![Figure 15](image1.png)

**Figure 15.** Average of the SSM-extracted Mie series composite signal over 1000 Monte Carlo trials.

![Figure 16](image2.png)

**Figure 16.** Average of the SSM-extracted specular wave signal over 1000 Monte Carlo trials.

We address Monte Carlo analysis of the creeping wave next. As seen in Fig. 13, the noise-free creeping wave signal has an amplitude of \(-20 \) to \(-70 \) dB, and therefore, adding noise before the creeping wave extraction would considerably stress the state space algorithm. Fig. 17 displays the performance of SSM in extracting the creeping wave for various SNRs, relative to the reference analytical solution from [73].

The extracted signal is in reasonable agreement with the analytical solution for \( SNR \geq 10 \) dB. The SSM estimate deviates significantly from the reference solution at frequencies greater than 10 GHz for \( SNR = 5 \) dB. For SNRs between 5 and 15 dB, the error increases with frequency for \( f > 16 \) GHz, but the RCS response is around \(-70 \) to \(-65 \) dB, which is approaching the noise floor. It has been observed that the pole drift from trial to trial becomes significant for \( SNR = 5 \) dB, and in fact, for signals smaller than \(-60 \) dB, it becomes difficult to discriminate noise poles from the signal poles. Therefore, for weak signals such as creeping waves, caution should be exercised in extracting the signal under low SNR conditions.
Nevertheless, the worst-case performance of SSM under the stressing conditions depicted in Fig. 17 is gratifying, given that the frequency response of the creeping wave signal for each SNR, extracted from the average of 1000 MC trials of the Mie series, follows the overall analytical trend. One can improve the noise performance by employing two-dimensional data, e.g., aspect- and frequency-dependent RCS, which improves the SNR by coherent integration of frequency samples over many pulses [61, 62].

The absolute model prediction error for the creeping wave estimation from the average of 1000 MC trials on the Mie series is plotted in Fig. 18 in terms of the SNR. It is evident that the worst-case error for $\text{SNR} = 5$ dB is between two and four orders of magnitude larger than the model error for the noise-free case shown in Fig. 14. For $\text{SNR} \geq 10$ dB, the error is only slightly larger than that in Fig. 14.

Lastly, we investigate the model order by computing the singular value spectrum as well as AIC and MDL estimates. It is seen in Fig. 19 that the transition from strong signal components to small singular values is not well defined because of noise. AIC yields model order of 20 while MDL yields 7. Due to measurement noise, AIC usually gives a larger estimate than MDL, and modelers usually choose an order no smaller than the AIC prediction. Therefore, an order of 20 has been chosen to extract the wave features of interest for the sphere. Compared to estimates in the dielectric slab case, which have a very high SNR (see Table 2), the accuracy in the SSM estimate for the MC trials is much lower, of the order 0.001.
5. EXTRACTION OF PARAMETERS FROM MEASURED DATA

Section 4 presented a detailed description of the process to extract parameters of interest from simulated data on two canonical objects, namely, dielectric slab and sphere. In this Section, we briefly summarize efforts on the application of SSM to the measured data on relatively complex EM problems, including canonical radar target scattering, biomedical radar and microwave circuits.

5.1. Canonical Target Scattering

The reader refers to [40] for application of the SSM to wave-oriented feature identification using monostatic RCS measured data on a large conical metallic target with and without a lossy dielectric coating. The results therein demonstrate the isolation of electromagnetic wave species of interest, such as creeping waves, multiply diffracted waves and specular scattering, using a small number of range-classified poles specific to each species. Detailed study of the isolated wave species yields a better understanding of the physics behind wave propagation around curved dielectric and coated structures, thereby improving the accuracy of feature extraction or target identification. For example, it has been shown that a thin dielectric coating enhances the creeping wave contribution significantly in certain directions compared to the metallic cone, suggesting efficacy of the coating to absorb EM waves at certain incident angles and frequencies [40], which may be of significant interest in military applications.

5.2. Biomedical Radar

SSM has been successfully applied to estimate cardiopulmonary parameters of a stationary human subject in [41–43]. The biomedical UWB radar [41–43] uses narrow pulses to probe the human body and track tiny cardiopulmonary chest movements by spectral analysis of the backscattered EM field. With the help of super-resolution spectral algorithms [52], the UWB radar is capable of increased accuracy in estimating vital signs such as heart and respiration rates in adverse signal-to-noise conditions. A major challenge for the biomedical radar system is detection of the heartbeat with high accuracy, because of minute thorax motion (less than 0.5 mm) caused by heartbeat. The problem becomes compounded by EM clutter and noise in the ambient environment. We have shown that SSM processing of the UWB radar data on an indoor stationary human subject consistently produces accurate estimates of the vital signs for several independent channels of UWB data without producing harmonics and inter-modulation products that plague signal resolution in the widely used FFT spectrograms [41]. Our extensive work on state-space methods to process biomedical radar data [41–43, 75, 76] suggests that the SNR of vital sign detection can be substantially improved, and additional features such as subject localization, motion compensation, gait analysis, etc., can be estimated, using block-processing of one-dimensional data (in the frequency domain) to attain considerably improved accuracy over FFT-based methods.

5.3. Isolation of Discontinuity Effects in a MEMS Interconnect

A variety of MEMS series and shunt switches, varactors and other circuits have been developed at MIT Lincoln Laboratory which operate over 100 GHz bandwidth with low insertion loss (cf. [77]). Muldavin et al. reported a low-loss MEMS packaging technology which facilitates these devices to be integrated with other RF circuits on the same wafer using CMOS-compatible fabrication methods [78]. This package features interconnects (transmission lines) based on inverted suspended microstrip configuration, with insertion loss typically less than 0.1 dB for a 100 µm long line over 50 GHz bandwidth and absolute phase error less than 1° for 10.8 mm long line.

The physical layout of the packaged interconnect is shown in Fig. 20 [78]. The circuit consists of an inverted microstrip transmission line, 10.8 mm long, on a 25 µm-thick silicon-on-insulator (SOI) layer, with low-inductance vias through the wafer providing the ground connection. Electrically isolated posts in the cavity provide support for bump bonds on the center conductor. The S-parameters of this packaged transmission line were measured from 100 MHz to 40 GHz using an Agilent Vector Network Analyzer (VNA) and a Cascade Summit Probe Station [78]. An SOLT calibration substrate was used to calibrate the measurements to the probe tips. The circuit between the probe tips includes two CPW-to-inverted microstrip transitions through the package vias (see Fig. 20), and their influence has
to be “gated out” in order to extract the desired transmission line parameters of the interconnect from measurements.

SSM has been used to derive an AR (all pole) system model from the measured $S$-parameters, which extracts the transmission line parameters for the packaged MEMS circuit in Fig. 20 by isolating and removing the CPW discontinuity effects [36]. Fig. 21 displays the range profile of the wideband measured data for $S_{11}$ and $S_{21}$ on a 10.8 mm long MEMS line. This plot is similar to time-gated waveforms seen on a VNA, except that post-processing allows better noise filtering and improves the resolution. The peak of $S_{21}$ corresponds to the leading edge of the line, and the small $S_{11}$ peak directly below it gives the reflection off the line. The two larger peaks in $S_{11}$ (on either side of the center) correspond to the leading and trailing returns from the input and output transitions, respectively.

Good agreement is observed between the measured $S$-parameters and the 4-pole SSM model derived in [36]. In general, the magnitude discrepancy for the model fit is less than 0.5 dB, and the phase discrepancy is within $0.4^\circ$ over 40 GHz bandwidth. Of these four poles, only one has zero range that corresponds to the line input reference. The other poles are associated with the transitions and are not considered in evaluation of the line parameters.
The attenuation and phase constants are computed using the single pole corresponding to the peak of $S_{21}$ and are plotted in Fig. 22. The line loss is determined to be 0.75 dB/cm at 20 GHz and 1.04 dB/cm at 40 GHz, which compare very well with the values estimated in [78] using differential line measurements. The phase changes linearly with frequency and yields an effective dielectric constant of 2.85 to 2.9 between 1 GHz and 40 GHz, again in excellent corroboration with independent calculations in [78]. Using the isolated peak of $S_{11}$ at zero range (see Fig. 21), we have also computed the characteristic impedance of the line, assuming a reference termination of 50 Ω. The resulting plot (Fig. 23) shows that the impedance (real part) is within 0.1 Ω of the nominal value over the wide 40 GHz bandwidth, attesting to the accuracy of the state space model. The imaginary part is less than 65 m Ω over the entire 40 GHz band. It is emphasized that network synthesis using lumped equivalent circuit models over such a large bandwidth is excessively cumbersome and entails a large model order even if a model can be found [79, 80].

6. CONCLUSIONS

A spectral model based on state-space ARMA representation has been presented to isolate and extract modal EM responses, such as creeping waves and multiply reflected or diffracted waves, which are of interest in radar target identification, and feature extraction. The eigenvalues of the open-loop system matrix, i.e., poles of the ARMA transfer function, provide the range and frequency-dependent amplitude decay/growth rate of the field data. The amplitudes in the complex exponential model are obtained by least squares modal decomposition involving the state matrices. Range classification of the poles allows for the isolation of desired responses in the EM signature via pulse compression and spectral decomposition. For purposes of illustration, in this paper the SSM is reviewed by application to two simple EM problems, namely, Fresnel reflection by a dielectric slab and the extraction of creeping waves using Mie scattering by a PEC sphere. The method has been applied to isolate the leading-edge reflection and multiple internal reflections of the slab. It is shown that range-classification of the singularities in the response enables characterization of the isolated reflected waves using a low model-order SSM representation. In the second example, we have analyzed the extraction of creeping waves from Mie scattering by a PEC sphere in the absence of noise. The SSM estimates of the extracted creeping wave have been validated by comparison with an analytically derived expression of the creeping wave. Next, random white Gaussian noise is added to the Mie series solution, and Monte Carlo simulation is performed to examine robustness of the SSM estimates to noise. Numerical considerations such as SNR, singular value spectrum, and order determination are addressed in detail for both of these examples.

The extraction of EM parameters using SSM on the measured data for relatively complex EM problems, including canonical radar target scattering, biomedical radar, and microwave circuits has been summarized. A state-space spectral model has been presented for the parametric characterization of wideband transmission lines in a MEMS package, including a detailed description on the removal
of package-specific discontinuity effects. SSM enables the representation of measured S-parameters using a compact low-order all-pole model, which filters out measurement noise and facilitates extraction of distributed circuit parameters, such as propagation constant and characteristic impedance. It is anticipated that direct equivalent circuit synthesis from the system model, as opposed to the measured data, leads to more efficient optimization in circuit simulators.

**APPENDIX A.**

In terms of the finite-rank observability matrix \( \tilde{\Omega} \in \mathbb{C}^{(N-L+1) \times M} \) computed from the SVD in (25) as

\[
\tilde{\Omega} = \begin{bmatrix} C & CA & CA^2 & \ldots & CA^{N-L-1} & CA^{N-L} \end{bmatrix}^T,
\]

the state transition or the open-loop matrix \( A \in \mathbb{C}^{M \times M} \) is determined by the solution to the matrix equation

\[
\tilde{\Omega}_{-r\ell} A = \tilde{\Omega}_{-r1},
\]

with

\[
\tilde{\Omega}_{-r1} = \begin{bmatrix} CA & CA^2 & CA^3 & \ldots & CA^{N-L-1} \end{bmatrix}^T,
\]

\[
\tilde{\Omega}_{-r\ell} = \begin{bmatrix} C & CA & CA^2 & \ldots & CA^{N-L-1} \end{bmatrix}^T.
\]

It is observed that the matrices \( \tilde{\Omega}_{-r1} \) and \( \tilde{\Omega}_{-r\ell} \) are obtained by deleting the first and last rows, respectively, of the matrix \( \tilde{\Omega} \) in Eq. (A1). Using the pseudo-inverse least squares on Eq. (A2), we obtain the state transition matrix \( A \) given in Eq. (26) and repeated below.

\[
A = \left( \tilde{\Omega}_{-r\ell}^* \tilde{\Omega}_{-r\ell} \right)^{-1} \tilde{\Omega}_{-r\ell}^* \tilde{\Omega}_{-r1}.
\]

This matrix may also be derived from the controllability matrix \( \tilde{\Gamma} \in \mathbb{C}^{M \times L} \) derived from the SVD in Eq. (25) as

\[
\tilde{\Gamma} = \begin{bmatrix} B & AB & A^2B & A^3B & \ldots & A^{L-1}B \end{bmatrix}.
\]

It follows that the state transition matrix \( A \) satisfies the matrix equation

\[
A\tilde{\Gamma}_{-c\ell} = \tilde{\Gamma}_{-c1},
\]

with

\[
\tilde{\Gamma}_{-c1} = \begin{bmatrix} AB & A^2B & A^3B & \ldots & A^{L-1}B \end{bmatrix},
\]

\[
\tilde{\Gamma}_{-c\ell} = \begin{bmatrix} B & AB & A^2B & \ldots & A^{L-2}B \end{bmatrix}.
\]

Eqs. (A8) and (A9) are obtained by deleting the first and last columns, respectively, of the controllability matrix \( \tilde{\Gamma} \) in Eq. (A6). By solving Eq. (A7) for \( A \) using the pseudo-inverse least squares method, equivalent to Eq. (A5), we obtain (also see Eq. (27))

\[
A = \tilde{\Gamma}_{-c1} \tilde{\Gamma}_{-c\ell}^* (\tilde{\Gamma}_{-c\ell} \tilde{\Gamma}_{-c\ell}^*)^{-1}.
\]

Next, we address the computation of control and observation matrices, \( B \) and \( C \), respectively, using two alternative approaches. In the first approach, it follows from Eq. (A1) that the observation matrix \( C \) is simply given by the first row of the observability matrix,

\[
C = \tilde{\Omega}(1,:) .
\]

Alternatively, the IIR approximation, \( \hat{y}(k) \), of the data sequence \( y(k), k = 1, 2, \ldots, N \), can be computed in terms of the state matrices using Eq. (16) as

\[
\hat{y}(k) = CA^{k-1}B, \quad k = 1, 2, \ldots, N.
\]

We define the augmented observability matrix employing all the \( N \) samples:

\[
\tilde{\Omega}_N = \begin{bmatrix} C & CA & CA^2 & \ldots & CA^{N-1} \end{bmatrix}^T,
\]
which is related to the impulse response in Eq. (A12) as
\[ \hat{\Omega}_N B = \hat{y}^T. \] (A14)

Therefore, in order to minimize the error between the model \( \hat{y} \) and the measured data vector \( y \), the control matrix \( B \) is computed by the pseudo-inverse least squares method, as shown below.
\[ B = \left( \hat{\Omega}_N^* \hat{\Omega}_N \right)^{-1} \hat{\Omega}_N^* \hat{y}^T. \] (A15)

In summary, in the first approach, one may calculate the matrices \( C \) and \( B \) from the observability matrix \( \hat{\Omega} \), using Eqs. (A11) and (A15), respectively.

Alternatively, in the second approach, \( B \) is computed from the first column of the controllability matrix in Eq. (A6) as
\[ B = \hat{\Gamma}(,1), \] (A16)
and \( C \) follows from the least squares fit between the data sequence \( y(k) \), and the state-space IIR approximation \( \hat{y}(k) \) in Eq. (A12), as shown next. In terms of the augmented controllability matrix
\[ \hat{\Gamma}_N = \begin{bmatrix} B & AB & A^2B & \ldots & A^{N-1}B \end{bmatrix}, \] (A17)
the estimation problem for \( C \) may be written as
\[ C \hat{\Gamma}_N = \hat{y}, \] (A18)
and its pseudo-inverse least squares solution yields
\[ C = \frac{\hat{y}^{\Gamma}_N (\hat{\Gamma}_N^{\Gamma}_N)^{-1}}{\hat{\Gamma}_N (\hat{\Gamma}_N^{\Gamma}_N)^{-1}}. \] (A19)

To summarize, it is emphasized that there are two expressions for the state transition matrix, \( A \), and two corresponding alternative approaches to computing \( B \) and \( C \). The matrix \( A \) may be computed from either the observability matrix using Eq. (A5) or the controllability matrix using Eq. (A10). Correspondingly, the matrices \( B \) and \( C \) must be computed by using either Eqs. (A15) and (A11), or Eqs. (A16) and (A19), respectively. In order to improve the estimation accuracy in a low SNR environment, we almost always employ the least-squares computations in Eqs. (A15) and (A19) to estimate \( B \) and \( C \), respectively, and seldom use Eqs. (A16) and (A11).

ACKNOWLEDGMENT

The author is grateful to his former colleague, Dr. Jean Piou at MIT Lincoln Laboratory, for introducing him to the state space method. This paper reviews the results from several publications co-authored with Dr. Piou. The author also acknowledges Dr. Jeremy Muldavin for supplying the wideband measured data used to characterize the MEMS transmission lines (see Section 5).

REFERENCES


