

# Classifying Sums of Exponentially Damped Sinusoids Using an Associated Numerical Range

Raymundo Albert<sup>b,\*</sup>, Cecilia Galarza<sup>a,b,\*</sup>

<sup>a</sup>*Centro de simulación computacional, CONICET, Godoy Cruz 2390, C1425FQD Buenos Aires, Argentina*

<sup>b</sup>*Facultad de Ingeniería, Universidad de Buenos Aires, Av. Paseo Colón 850, C1063ACV Buenos Aires, Argentina*

---

## Abstract

The matrix pencil method (MPM) is a well-known technique for estimating the parameters of exponentially damped sinusoids in noise by solving a generalized eigenvalue problem. However, in several cases, this is an ill-conditioned problem whose solution is highly biased under small perturbations. When the estimation is performed to classify the observed signal into two categories, the estimation errors induce several misclassifications. In this work we propose a novel signal classification criteria by exploiting the relationship between the generalized eigenvalue problem posed in the MPM and the numerical range of a pair of rectangular matrices. In particular, the classification test is formulated as a set inclusion problem, and no spectrum estimation is required. The technique is applied to a problem of electromagnetic scattering to classify dielectric materials using the scattering signal observed when a target is illuminated by an ultra-wideband signal. The performance of the classification scheme is assessed in terms of error rate and it is compared to another classification technique, the generalized

---

\*Corresponding author

*Email addresses:* `araymundo@fi.uba.ar` (Raymundo Albert),  
`cgalarza@csc.conicet.gov.ar` (Cecilia Galarza)

likelihood rate test (GLRT).

*Keywords:* Target classification, matrix pencil method, numerical range, singularity expansion method

---

## 1. Introduction

Problems dealing with a mixture of closely spaced sinusoidal signals in a noisy environment are regularly observed in engineering. This is the case of the scattering collected from a dielectric target illuminated by an Ultra-WideBand (UWB) signal, where the particular mixture is a feature that characterizes the target. Identification of the complex frequencies embedded in the scattered signal becomes relevant for target identification and classification. Motivated by this problem, we have designed a novel strategy for classifying mixtures of sinusoidal signals. In particular, this procedure exploits the underlying structure among the complex frequencies of the observed signal without actually identifying their values.

### 1.1. Related Work

Traditional approaches to this problem include the matrix pencil method (MPM)[1]. This is a spectrum estimation technique that constructs a generalized eigenvalue problem from a partition of the Hankel matrix of the observed signal. In this setup, the matrix pencil is non-square and the eigenvalue computation can become unstable or the solutions may fail to exist. This problem is particularly acute when the number of sinusoids in the original mixture is very large and the complex frequencies are close to each other in the complex plane. These facts make very difficult to accurately identify the parameters of the exponentially damped sinusoidal signals.

An alternative was studied in [2, 3], where a generalized likelihood test was posed to classify perfectly conducting objects using scattered signals. Nevertheless, GLRT is sensitive to unmodelled dynamics, which could be the case when a large number of sinusoids form the observed mixture.

On the other hand, research on eigenvalue computation for a matrix leads to the analysis of its numerical range, which is a set in the complex plane that contains the spectrum of the matrix [4, 5]. Analysis of the numerical range of matrices has been an active field of research in the past. Some of those results have set the ground for the classification criteria that we present in this paper. In particular, we return to the definition for the numerical range of rectangular matrices introduced in [6]. The authors of this paper defined the numerical range of a pair of rectangular matrices that resulted in a compact and convex set of the complex plane.

### *1.2. Contributions*

Making a hinge on the MPM, in particular on the associated generalized eigenvalue problem, we formulate a signal classification criteria using the numerical range for the matrices of the matrix pencil. In particular, the classification is formulated as a set inclusion problem, hence avoiding the need to compute eigenvalues of probably large ill-conditioned matrices.

We start our paper with a review of the MPM in section 2. In particular, we discuss how the MPM poses a generalized eigenvalue problem that may be complicated to solve accurately. In section 3, we reproduce the definition for the numerical range of a pair of rectangular matrices and we show the properties that are relevant to our development. Along these lines, we show that the rectangular numerical range of a pair of matrices contains the corresponding generalized eigenvalues. Making ground on this observation,

in section 4 we propose a new classification technique that is solved as a set inclusion problem. To check the validity of our approach, we analyze in section 5 the behavior of the rectangular numerical range related to noisy signals for different values of the signal to noise ratio (SNR). Finally, in section 6 we show numerical results when classifying sums of exponentially damped sinusoids. Also, we apply the new procedure to scattering signals from different dielectric materials and compare its performance to an already known classification technique, the Generalized Likelihood Ratio Test (GLRT).

## 2. Matrix Pencil Method

The Matrix Pencil Method (MPM) has been used as a high-resolution spectrum estimation technique on different applications for several years now [7, 8]. In this section, we review briefly the main aspects of this technique.

Suppose that we want to analyze a signal that is composed by a sum of damped complex exponentials perturbed by noise. We assume that different looks of a single experiment are acquired. For example, in the case of the electromagnetic scattering, an array of antennas spatially distributed may be collecting the scattered signals from different observation angles. On a general framework, we consider the signal vector  $\mathbf{y}_t$  with  $K$  components

$$\mathbf{y}_t = \sum_{i=1}^M \mathbf{c}_i z_i^t + \mathbf{w}_t, \quad t = 0, 1, \dots \quad (1)$$

where  $K$  is the number of looks,  $M$  is the number of exponentials,  $z_i \in \mathbb{C}$  is a complex resonant frequency, and  $\mathbf{c}_i \in \mathbb{C}^K$  is the vector of residues associated to  $z_i$ .

Now, for two integers  $s$  and  $n$ ,  $s > n$ , consider  $s + n - 1$  samples of  $\mathbf{y}_t$ . Let  $m = sK$  and define  $H_{\mathbf{y}}$  as the following  $(m \times (n + 1))$ -block Hankel matrix

$$H_{\mathbf{y}} = \begin{bmatrix} \mathbf{y}_0 & \mathbf{y}_1 & \cdots & \mathbf{y}_n \\ \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{y}_{s-1} & \mathbf{y}_s & \cdots & \mathbf{y}_{s+n-1} \end{bmatrix} \in \mathbb{C}^{m \times (n+1)}. \quad (2)$$

The integer  $n$  is known as the pencil parameter and it is a resource to mitigate the effect of  $\mathbf{w}_t$  on  $\mathbf{y}_t$ <sup>1</sup>, provided that  $M \leq n \leq s - M$ .

A well-known result states that in the noiseless case, the rank of  $H_{\mathbf{y}}$  is  $M$  [9]. To show that, recall that the entries of  $H_{\mathbf{y}}$  satisfy the following recurrence relation of order  $M$

$$\mathbf{y}_t = \sum_{k=1}^M a_k \mathbf{y}_{t-k}, \quad t = M, M + 1, \dots, \quad (3)$$

and  $M$  is the least integer number satisfying (3) for all  $t$ . Hence, every column of  $H$  is a linear combination of the first  $M$  columns, and  $H$  has indeed rank equal to  $M$ .

Define now two  $(m \times n)$ -block matrices,  $A_{\mathbf{y}}$  and  $B_{\mathbf{y}}$ , where the first one is obtained by deleting the first column of  $H_{\mathbf{y}}$ , and the second one is obtained from  $H_{\mathbf{y}}$  by deleting its last column. A matrix pencil is then defined as

$$A_{\mathbf{y}} - \lambda B_{\mathbf{y}}. \quad (4)$$

When the signal  $\mathbf{y}_t$  is noiseless, the rank of the pencil (4) is  $M$  as long as  $\lambda \neq z_i$ ,  $i = 1, \dots, M$ . When  $\lambda = z_i$ , the rank of the pencil is reduced

---

<sup>1</sup>It has been shown that when using the MPM for estimating  $z_i$ , its variance is minimum for values of  $n$  between  $s/2$  and  $2s$  [1].

by one. Furthermore,  $z_i, i = 1, \dots, M$  can be computed as the generalized eigenvalues of  $(A_{\mathbf{y}}, B_{\mathbf{y}})$ . Alternatively, when  $B_{\mathbf{y}}$  is full rank,  $z_i$  solves the eigenvalue problem

$$B_{\mathbf{y}}^{\dagger} A_{\mathbf{y}} - \lambda I_n, \quad (5)$$

where  $B_{\mathbf{y}}^{\dagger}$  is the pseudoinverse and  $I_n$  is the identity matrix of order  $n$ . This is the principle for using the MPM for estimating the values of  $z_i$ .

However, when considering noisy signals, the pencil (4) has usually full rank for any complex scalar  $\lambda$ . To overcome this problem, several procedures have been proposed in the past. All of them obtain an approximation for the original Hankel matrix before computing the pencil. A first simple approach was followed in [10], where the singular value decomposition of the original Hankel matrix was computed and only the directions corresponding to the dominant singular values were kept. The authors in [7, 8] noticed that this procedure did not guarantee a Hankel structure for the resulting approximation of  $H_{\mathbf{y}}$ . Following [11], they proposed an iterative procedure that kept the Hankel structure in the approximated matrix. Alternatively, in [12, 13], the problem was tackled by solving a total least-squares problem. In both procedures, knowledge of  $M$  was key, and it had to be estimated when not known a priori.

Nevertheless, even when the value of  $M$  is known, the computation of generalized eigenvalues may be a delicate problem. For a better understanding of this issue, and following [14], we define the function  $g : \mathbb{C} \rightarrow \mathbb{R}_{\geq 0}$

$$g(\lambda) = \sigma_{\min}\{A_{\mathbf{y}} - \lambda B_{\mathbf{y}}\} \quad (6)$$

where  $\sigma_{\min}$  is the smallest singular value. For a pencil  $A_{\mathbf{y}} - \lambda B_{\mathbf{y}}$  that has a rank-reducing solution at  $\lambda = z$ ,  $g(z) = 0$ . Therefore, the generalized

eigenvalues of  $(A_{\mathbf{y}}, B_{\mathbf{y}})$  are the local minima of this function.

Consider now a signal as in (1) with  $K = 1$  and  $M = 4$ , under two scenarios: one without noise ( $\mathbf{w}_t = 0$ ), and one with noise ( $\mathbf{w}_t \neq 0$ ). Using  $n = M$ , we build  $g^{noiseless}(\lambda)$  from the noiseless signal and  $g^{noisy}(\lambda)$  from the noisy one. Fig. 1 shows the level curves of function  $g^{noiseless}(\lambda)$ . Observe that in this case, its minima coincide with the location of  $z_i, i = 1, \dots, 4$ .

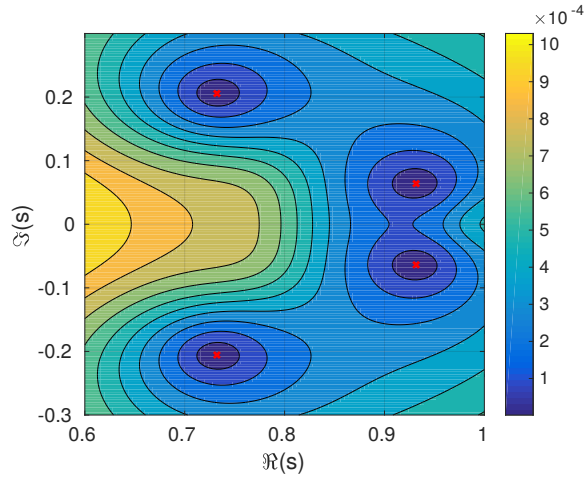


Figure 1: Function  $g^{noiseless}(\lambda)$ , and  $z_1, \dots, z_4$  ('x').

Now in Fig. 2, we have plotted the level curves for  $g^{noisy}(\lambda)$  for a signal to noise ratio 30 dB. Notice that in this case, the local minima move away from the actual values of  $z_1$  through  $z_4$ . It is clear from this very simple example that there are major difficulties that we may face when identifying the oscillation modes from noisy signals. These difficulties are even more serious when the value of  $M$  is unknown and the order of the model needs to be identified beforehand [15, 16].

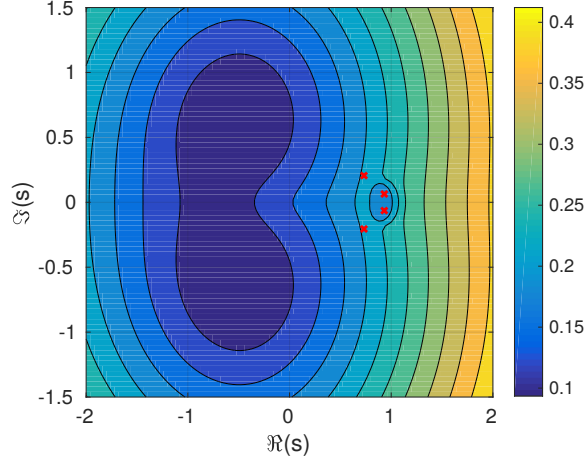


Figure 2: Function  $g^{noisy}(\lambda)$ , and  $z_1, \dots, z_4$  ('x').

### 3. The numerical range of a matrix pencil

As shown previously, solving (1) may be difficult when dealing with perturbed matrices. Nevertheless, it is interesting to define a region on the complex plane where the solutions would be. A candidate for that is the numerical range associated to  $(A_{\mathbf{y}}, B_{\mathbf{y}})$ . In this section, we present background material necessary to present the main result of the paper. It is not our intention to provide a thorough analysis of the properties and problems related to the numerical range of a matrix. For an in-depth tutorial on this matter, we refer the reader to [4].

Let  $A$  be a complex  $n \times n$ -matrix. Its numerical range or field of values is defined as

$$W(A) := \{\mathbf{x}^H A \mathbf{x} : \mathbf{x} \in \mathbb{C}^n, \mathbf{x}^H \mathbf{x} = 1\} \subset \mathbb{C}. \quad (7)$$

#### Properties of $W(A)$

1. All the eigenvalues of  $A$  lie in  $W(A)$ .



2.  $W(A)$  is a closed convex set.
3. It has the form [5]

$$W(A) = \{\theta \in \mathbb{C} : \|A - \lambda I_n\|_2 \geq |\theta - \lambda|, \forall \lambda \in \mathbb{C}\}. \quad (8)$$

where  $\|\cdot\|_2$  is the spectral norm. The set  $W(A)$  appears naturally when recursively computing the eigenvalues of  $A$ . For instance, the Arnoldi recurrence, which is the universally used algorithm for this matter<sup>2</sup>, approximates the sought eigenvalues with elements of  $W(A)$ .

The concept of eigenvalues is generalized when dealing with two rectangular matrices. Let  $A, B \in \mathbb{C}^{m \times n}$ . A scalar  $\beta$  is a generalized eigenvalue of  $(A, B)$  if there exists  $\mathbf{x} \in \mathbb{C}^n$  such that

$$A\mathbf{x} = \beta B\mathbf{x}, \quad \mathbf{x} \neq \mathbf{0}. \quad (9)$$

Following (8), we introduce the definition for the numerical range of the pencil  $(A; B)$  as

$$\begin{aligned} W_{\|\cdot\|_2}(A; B) &= \{\theta \in \mathbb{C} : \|A - \lambda B\| \geq |\theta - \lambda|, \forall \lambda \in \mathbb{C}\} \\ &= \bigcap_{\lambda \in \mathbb{C}} \mathcal{D}(\lambda, \|A - \lambda B\|). \end{aligned} \quad (10)$$

This definition was first proposed in [6]. We state, without proofs, some properties of  $W_{\|\cdot\|_2}(A; B)$ .

**Proposition 1.**  $W_{\|\cdot\|_2}(A; B)$  is non-empty if and only if  $\|B\| \geq 1$ .

Notice that from (10), we have that  $\forall \theta \in W_{\|\cdot\|_2}(A; B)$ ,  $|\theta| \leq \|A\|_2$ . Moreover, for  $\lambda$  such that  $|\theta| < |\lambda|$ , the following is true

$$|\lambda| - |\theta| \leq \|A - \lambda B\|_2 \leq \|A\|_2 + |\lambda| \|B\|_2.$$

---

<sup>2</sup>A variant of the Arnoldi algorithm has been implemented in ARPACK.

Hence  $|\lambda|(1 - \|B\|_2) \leq 2\|A\|_2$ . If  $\|B\|_2 < 1$ , this is satisfied by  $\lambda$  on a bounded set only, and not on  $\mathbb{C}$ . Therefore,  $\theta \notin W_{\|\cdot\|_2}(A; B)$ . Then if  $\|B\|_2 < 1$ , for any  $\theta \in W_{\|\cdot\|_2}(A; B)$  it is possible to find  $\lambda$  that does not satisfy (10). We conclude that  $W_{\|\cdot\|_2}(A; B) = \emptyset$ . It can be proved that  $\|B\|_2 > 1$  is also a necessary condition for non-emptiness.

**Proposition 2.** *Any generalized eigenvalue  $\beta$  of  $(A, B)$  with associated eigenvector  $\mathbf{x} \in \mathbb{C}^n$  such that  $\|B\mathbf{x}\| \geq 1$ , lies in  $W_{\|\cdot\|_2}(A; B)$ .*

Here, the condition  $\|B\mathbf{x}\| \geq 1$  is only sufficient but not necessary. Experimentally, it has been verified that the generalized eigenvalues lied in  $W_{\|\cdot\|_2}(A; B)$  although  $\|B\mathbf{x}\| < 1$ .

**Proposition 3.** *Let  $\alpha \in \mathbb{C}$  such that  $|\alpha| \geq 1$ . Then for any  $A, B \in \mathbb{C}^{m \times n}$ ,*

$$W_{\|\cdot\|_2}(A, B) \subset W_{\|\cdot\|_2}(\alpha A, \alpha B).$$

**Proposition 4.** *When the Frobenious norm is used in (10), and  $\|B\|_F \geq 1$ ,*

$$W_{\|\cdot\|_F}(A; B) = \mathcal{D}\left(\frac{\langle A, B \rangle}{\|B\|_F^2}, \left\|A - \frac{\langle A, B \rangle}{\|B\|_F^2} B\right\|_F \frac{\sqrt{\|B\|_F^2 - 1}}{\|B\|_F}\right). \quad (11)$$

In this case, the numerical range is characterized by a closed disc centered at  $\frac{\langle A, B \rangle}{\|B\|^2}$ , which is a superset that contains  $W_{\|\cdot\|_2}(A, B)$ , because of the equivalence between  $\|\cdot\|_2$  and  $\|\cdot\|_F$ .

**Lemma 1.** *For any  $A, B \in \mathbb{C}^{m \times n}$  ( $m \geq n$ ),  $\|B\|_2 = 1$ , it holds that  $W_{\|\cdot\|_2}(A; B) \subseteq W(B^\dagger A)$ .*

*Proof.* When  $m > n$  and using the definition of the Moore-Penrose pseudoinverse  $B^\dagger$ , we have that

$$\begin{aligned} W_{\|\cdot\|_2}(A; B) &= \{\theta \in \mathbb{C} : \|(B^\dagger A - \lambda I_n)B\|_2 \geq |\theta - \lambda|, \forall \lambda \in \mathbb{C}\} \\ &\subseteq \{\theta \in \mathbb{C} : \|B^\dagger A - \lambda I_n\|_2 \geq |\theta - \lambda|, \forall \lambda \in \mathbb{C}\} \\ &= W(B^\dagger A), \end{aligned}$$

where the last equality follows from (8).  $\square$

Recall that Eq. (5) formulates the generalized eigenvalue problem as a regular one using  $B^\dagger A$ . As a consequence, if  $\beta$  in (9) exists, it lies in  $W(B^\dagger A)$ . In the particular case of Lemma 1, we have that  $W_{\|\cdot\|_2}(A; B) \subseteq W(B^\dagger A)$ .

## 4. Main Result

### 4.1. A Classification Problem

Our quest is for a criterion to classify pairs of rectangular matrices  $(A, B)$  according to their generalized eigenvalues. A direct approach to this problem would compute the generalized eigenvalues. However, this would require lengthy computations when considering large size matrices, and the overall approach might fail to obtain a valid solution when using noisy experimental data. To construct a more robust strategy, we propose to use the set  $W_{\|\cdot\|_2}(A; B)$  to characterize the behavior of  $(A; B)$ , rather than the actual values of the eigenvalues. Resorting to a minor abuse of notation, we will say that a matrix class  $\Theta$  is defined by a finite set of complex numbers that would also be called  $\Theta$ .

**Definition 1.** Let  $\Theta = \{\theta_1, \theta_2, \dots, \theta_p\} \subset \mathbb{C}$ . Consider two matrices  $A, B \in C^{m \times n}$ , with  $\|B\|_2 \geq 1$ . We say that  $(A, B)$  is of class  $\Theta$  if and only if  $\theta_i \in W_{\|\cdot\|_2}(A; B)$  for all  $i = 1, \dots, p$ .

The set  $\{\theta_1, \theta_2, \dots, \theta_p\}$  is the candidate set,  $\Theta$  is the candidate class, and  $(A; B)$  is the observed pair.

The classification procedure requires verifying whether  $\theta_i$  is a member of  $W_{\|\cdot\|_2}(A; B)$  or not. Next theorem shows how to answer this question by solving a minimization problem.

**Theorem 1.** *Given  $\theta \in \mathbb{C}$ , define the function  $f_\theta : \mathbb{C} \rightarrow \mathbb{R}$*

$$f_\theta(\lambda) = \|A - \lambda B\|_2 - |\theta - \lambda|.$$

*Then,  $\theta \in W_{\|\cdot\|_2}(A; B)$  if and only if*

$$\inf_{\lambda \in \mathbb{C}} f_\theta(\lambda) \geq 0$$

*Proof.* Using (10), we have that the following statements are equivalent:

$$\begin{aligned} \theta \in W_{\|\cdot\|_2}(A; B) &\Leftrightarrow \\ \|A - \lambda B\|_2 &\geq |\theta - \lambda|, \forall \lambda \in \mathbb{C} \Leftrightarrow \\ \|A - \lambda B\|_2 - |\theta - \lambda| &\geq 0, \forall \lambda \in \mathbb{C} \Leftrightarrow \\ f_\theta(\lambda) &\geq 0, \forall \lambda \in \mathbb{C}. \end{aligned}$$

The result follows from here. □

#### 4.2. Classification of exponentially damped sinusoids

Suppose now that the measured outcome of an experiment is modeled as in (1). In particular, consider that the set of complex resonant frequencies,  $\{z_i, i = 1, \dots, M\}$ , is a feature that characterizes the experimental observation. As such, we associate a class of observed signals to a set of complex frequencies. For instance, when observing a scattering phenomenon, the class of signals represents the material from which the illuminated target is made of. Then, given a finite sample collection of  $\mathbf{y}_t$ , the question at hand is whether this observation lies within the class defined a priori.

To proceed along this path, we assume that we know a priori all or a subset of the complex resonant frequencies present in the signals of interest. That will be, we know  $Z = \{z_i, i = 1, \dots, p\}$ , where  $p \leq M$ . Consider now that we collect  $s + n - 1$  samples of the observed signal  $\mathbf{y}_t$  and we use them

to build the Hankel matrix  $H_{\mathbf{y}}$  and the associated matrices  $A_{\mathbf{y}} \in C^{m \times n}$  and  $B_{\mathbf{y}} \in C^{m \times n}$  as in section 2. A straightforward extension of Definition 1 leads to the following definition

**Definition 2.** Let  $Z = \{z_1, z_2, \dots, z_p\} \subset \mathbb{C}$ . Consider the matrices  $A_{\mathbf{y}}, B_{\mathbf{y}}$  built from  $\mathbf{y}_t, t = 0, \dots, s + n - 1$ . We say that  $\mathbf{y}_t$  is of class  $Z$  if and only if  $z_i \in W_{\|\cdot\|_2}(A_{\mathbf{y}}; B_{\mathbf{y}})$  for all  $i = 1, \dots, p$ .

*The set  $Z$  is the candidate set that characterizes the candidate class,  $\mathbf{y}_t$  is the observed signal, and  $(A_{\mathbf{y}}, B_{\mathbf{y}})$  is the observed pair.*

Definition 2 allows us to classify the observed signal without explicitly solving the generalized eigenvalue problem formulated in the MPM. We call this procedure the Classification by Rectangular Numerical Range (CRNR). Algorithm 1 shows an implementation of this classification criteria.

When considering a single candidate class, we commit a classification error if  $Z \subset W_{\|\cdot\|_2}(A_{\mathbf{y}}; B_{\mathbf{y}})$  given that  $\mathbf{y}_t$  is not from class  $Z$ . When  $c$  candidate classes are considered, we deal with  $c$  sets  $Z_k, k = 1, \dots, c$ . The construction of each one depends on the signals being classified. On the other hand, when we observe  $\mathbf{y}_t$  of class  $Z_j$ , and consider the candidate class  $Z_k$ , with  $k \neq j$ , we commit a classification error when  $Z_k \subset W_{\|\cdot\|_2}(A_{\mathbf{y}}; B_{\mathbf{y}})$ . This observation will be used to characterize the performance of the classification criteria.

In the general case, the condition  $\|B_{\mathbf{y}}\|_2 \geq 1$  for having a non-empty rectangular numerical range for the pair  $(A_{\mathbf{y}}, B_{\mathbf{y}})$  cannot be guaranteed. However, an adequate scaling of  $\mathbf{y}_t$  may solve this problem. On the other hand, since any complex frequency  $z_i$  of  $\mathbf{y}_t$  is also a complex frequency of  $\alpha \mathbf{y}_t$  for any complex scalar  $\alpha$ , scaling is a suitable resource for applying the results of subsection 4.1 to the classification of a sum of complex exponen-

---

**Algorithm 1**

---

```
1: Given  $\mathbf{y}_t$ ,  $Z = \{z_1, z_2, \dots, z_p\}$ , and a scalar  $D > 1$ 
2: Obtain a reduced-rank approximation for  $H_{\mathbf{y}}$ 
3: Compute  $(A_{\mathbf{y}}, B_{\mathbf{y}})$  from  $H_{\mathbf{y}}$  .
4: if  $\|B\|_2 < 1$  then
5:    $A_{\mathbf{y}} = \frac{D}{\|B_{\mathbf{y}}\|_2} A_{\mathbf{y}}$ 
6:    $B_{\mathbf{y}} = \frac{D}{\|B_{\mathbf{y}}\|_2} B_{\mathbf{y}}$ 
7: end if
8: for  $k = 1, \dots, p$  do
9:   if  $r(A_{\mathbf{y}}, B_{\mathbf{y}}) < |z_k - c(A_{\mathbf{y}}, B_{\mathbf{y}})|$  then
10:     $\mathbf{y}_t$  is not of class  $Z$ . return
11:   else
12:     $\delta_k = \inf_{\lambda \in \mathbb{C}} \|A_{\mathbf{y}} - \lambda B_{\mathbf{y}}\|_2 - |z_k - \lambda|$ 
13:    if  $\delta_k < 0$  then
14:       $\mathbf{y}_t$  is not of class  $Z$ . return
15:    end if
16:   end if
17: end for
18:  $\mathbf{y}_t$  is of class  $Z$ .
```

---

tials. The appropriate scaling factor depends on the set  $Z$ . One may think that  $\alpha$  is selected during a setup process and kept fixed afterward.

A sensible step when dealing with noisy data is to obtain a reduced-rank approximation for  $H_{\mathbf{y}}$ . In this paper, we follow the ideas from [11]. For completeness, we have included the procedure in Appendix A. We will analyze the relevance of this step in the following section.

As mentioned in Proposition 4,  $W_{\|\cdot\|_2}$  is a subset of  $W_{\|\cdot\|_F}$ , which is a disk as in (11). Let  $c(A_{\mathbf{y}}, B_{\mathbf{y}})$  and  $r(A_{\mathbf{y}}, B_{\mathbf{y}})$  be the center and radius of that disk. Exploiting the structure of  $A_{\mathbf{y}}$  and  $B_{\mathbf{y}}$  we obtain the following closed-form expressions:

$$c(A_{\mathbf{y}}, B_{\mathbf{y}}) = \frac{\sum_{i=0}^{m+n-2} \eta_i \mathbf{y}_{i+1}^* \mathbf{y}_i}{\sum_{i=0}^{m+n-2} \eta_i |\mathbf{y}_i|^2} \quad (12)$$

$$r(A_{\mathbf{y}}, B_{\mathbf{y}}) = \left[ \sum_{i=0}^{m+n-2} \eta_i |\mathbf{y}_{i+1}|^2 - \frac{|\sum_{i=0}^{m+n-2} \eta_i \mathbf{y}_{i+1}^* \mathbf{y}_i|^2}{\sum_{i=0}^{m+n-2} \eta_i |\mathbf{y}_i|^2} \right]^{\frac{1}{2}} \\ \times \sqrt{\frac{\sum_{i=0}^{m+n-2} \eta_i |\mathbf{y}_i|^2 - 1}{\sum_{i=0}^{m+n-2} \eta_i |\mathbf{y}_i|^2}}, \quad (13)$$

where  $\eta_i$  is the number of elements in the  $i$ -th anti-diagonal of  $A_{\mathbf{y}}$  (or  $B_{\mathbf{y}}$ ). Any  $z$  such that  $|z - c(A_{\mathbf{y}}, B_{\mathbf{y}})| > r(A_{\mathbf{y}}, B_{\mathbf{y}})$ , does not lie within  $W_{\|\cdot\|_F}(A_{\mathbf{y}}, B_{\mathbf{y}})$ . Hence  $z \notin W_{\|\cdot\|_2}(A_{\mathbf{y}}, B_{\mathbf{y}})$ . In Algorithm 1 we include this simple check to quickly reject the candidate frequencies outside  $W_{\|\cdot\|_F}(A_{\mathbf{y}}, B_{\mathbf{y}})$ . Following Th. 1, for any other candidate frequency, we compute

$$\delta = \inf_{\lambda \in \mathbb{C}} \|A_{\mathbf{y}} - \lambda B_{\mathbf{y}}\|_2 - |z - \lambda|.$$

If  $\delta < 0$ , then  $z \notin W_{\|\cdot\|_2}(A_{\mathbf{y}}, B_{\mathbf{y}})$ .

## 5. Numerical range for noisy signals

The observed signal  $\mathbf{y}_t$  has a noise component that will alter the boundaries of  $W_{\|\cdot\|_2}(A_{\mathbf{y}}, B_{\mathbf{y}})$ . In general, large noise terms tend to expand the

boundaries of the numerical range. To mitigate this problem, we perform a reduced-rank approximation of  $H_{\mathbf{y}}$  as in step 2 in Algorithm 1. In this section, we analyze a particular example to show the relevance of this procedure by studying the behavior of  $W_{\|\cdot\|_F}(A_{\mathbf{y}}, B_{\mathbf{y}})$  for different noise levels.

We build a signal as in (1), using  $K = 1$ ,  $M = 10$ , and  $z_i \in Z_1$  defined as

$$Z_1 = \{0.4474 \pm 0.5822j; 0.4447 \pm 0.5782j; 0.4236 \pm 0.5874j; \\ 0.4166 \pm 0.5959j; 0.3871 \pm 0.5858j\}. \quad (14)$$

The noise term  $\mathbf{w}_t$  is a white Gaussian process with known variance. The signal to noise ratio (SNR) is defined as usual as the ratio between the signal power to the noise power. For each SNR, we perform a Montecarlo experiment with 10000 realizations. For each record of  $\{\mathbf{y}_t, t = 0, \dots, s + n - 1\}$ , we compute (12) and (13).

Figures 3 and 4 show the ensemble averages of  $c(A_{\mathbf{y}}, B_{\mathbf{y}})$  and  $r(A_{\mathbf{y}}, B_{\mathbf{y}})$ , for each SNR, before and after performing the reduced-rank approximation.

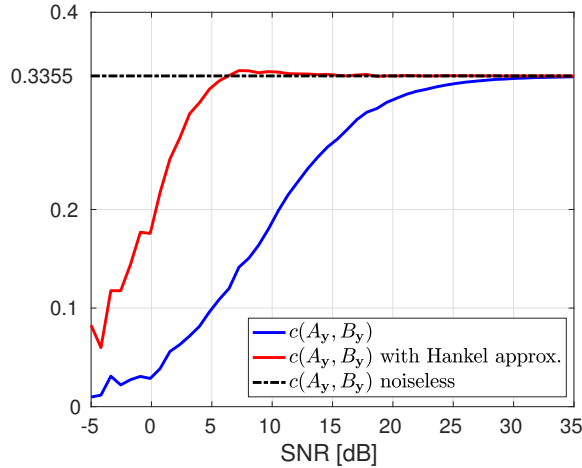


Figure 3:  $c(A_{\mathbf{y}}, B_{\mathbf{y}})$  vs. SNR with and without reduced-rank Hankel approximation.

We observe from Fig. 3 and 4 that  $W_{\|\cdot\|_F}(A_{\mathbf{y}}, B_{\mathbf{y}})$  gets enlarged as



the SNR deteriorates. For instance, the expected value of  $r(A_{\mathbf{y}}, B_{\mathbf{y}})$  grows as the noise power increases. However, the effect is dramatically reduced when we perform the reduced-rank approximation of  $H_{\mathbf{y}}$ . Also, the center  $c(A_{\mathbf{y}}, B_{\mathbf{y}})$  remains close to its noiseless value when the Hankel approximation is performed. In general, the rank reduction procedure contributes to mitigate the effect of noise on the boundaries of the numerical range. Since  $W_{\|\cdot\|_2}(A_{\mathbf{y}}, B_{\mathbf{y}})$  is a subset of  $W_{\|\cdot\|_F}(A_{\mathbf{y}}, B_{\mathbf{y}})$ , similar conclusions hold for the numerical range defined with the 2-norm. This is an important feature, since the classification process is defined by the boundaries of  $W_{\|\cdot\|_2}(A_{\mathbf{y}}, B_{\mathbf{y}})$  and its ability to enclose the complex frequencies from one class and to leave aside the complex frequencies from other classes.

Although we are showing these results on a particular example, the same conclusions were observed elsewhere. An important note to make is that this was the case when the approximated  $H_{\mathbf{y}}$  kept the Hankel structure. Simple truncation of the singular value decomposition of  $H_{\mathbf{y}}$  did not perform adequately.

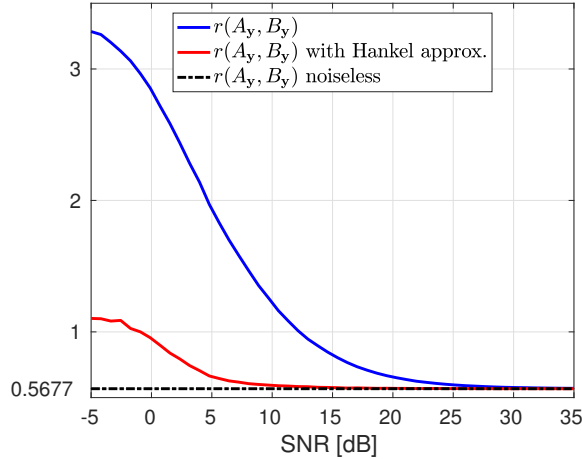


Figure 4:  $r(A_{\mathbf{y}}, B_{\mathbf{y}})$  vs. SNR with and without reduced-rank Hankel approximation.

## 6. Numerical results

### 6.1. Generalized likelihood test

Using the statistical description of  $\mathbf{y}_t$ , we can formulate a hypothesis testing problem and solve it by using a likelihood ratio test. We will use the GLRT procedure to assess the performance of the classification methodology proposed above. In this section, we succinctly describe the application of the GLRT in the context described so far.

Let the observed signal  $\mathbf{y}_t$  be from class  $Z_1$  as before, and consider a second class of signals,  $Z_2$  as the candidate class. Each signal in each class has its own vector of residues in (1). As such,  $\mathbf{c}_{i,1}$  is related to signals of class  $Z_1$ , and  $\mathbf{c}_{i,2}$  to signals of class  $Z_2$ . In general,  $\mathbf{c}_{i,1}$  and  $\mathbf{c}_{i,2}$  are unknown vectors. To robustify the classification, a generalized likelihood test (GLRT) is formulated as in [2]. Given  $Z_1 = \{z_{i,1}, i = 1, \dots, p\}$  and  $Z_2 = \{z_{i,2}, i = 1, \dots, p\}$ , and the observed signal  $\mathbf{y}_0, \dots, \mathbf{y}_l, l = s + n - 1$ , we define the hypothesis  $\mathbb{H}_1$  and  $\mathbb{H}_2$  as follows

$$\mathbb{H}_i : \begin{bmatrix} \mathbf{y}_0^t \\ \vdots \\ \mathbf{y}_l^t \end{bmatrix} = F_i \begin{bmatrix} \mathbf{c}_{1,i}^t \\ \vdots \\ \mathbf{c}_{p,i}^t \end{bmatrix}, \text{ where } F_i = \begin{bmatrix} z_{1,i}^0 & \cdots & z_{p,i}^0 \\ & \ddots & \\ z_{1,i}^l & \cdots & z_{p,i}^l \end{bmatrix},$$

is a matrix formed from the complex frequencies associated with  $Z_1$  and  $Z_2$  respectively. Notice that  $z_{k,i}, k = 1, \dots, p, i = 1, 2$ , are known frequencies. Then, the GLRT is posed as

$$\frac{\max_{\mathbf{c}_{i,1}} p(\mathbf{y}_0, \dots, \mathbf{y}_l | H_1)}{\max_{\mathbf{c}_{i,2}} p(\mathbf{y}_0, \dots, \mathbf{y}_l | H_2)} \underset{H_2}{\overset{H_1}{\gtrless}} 1.$$

## 6.2. Known model order

In this section, we test the classification procedure when the model order is known a priori. For that, we build the observed signal  $\mathbf{y}_t$  from class  $Z_1$  as before for different SNRs. For each SNR, the algorithm selects an appropriate scaling factor as stated in Algorithm 1. We consider the candidate class  $Z_2$

$$Z_2 = \{0.0429 \pm 0.0825j; -0.4130 \pm 0.1176j; -0.3118 \pm 0.2127j; \\ -0.1951 \pm 0.3642j; -0.3385 \pm 0.1249j\}. \quad (15)$$

Fig. 5 shows a particular realization of  $W_{\|\cdot\|_2}(A_{\mathbf{y}}, B_{\mathbf{y}})$  for different SNRs. We have also incorporated on this figure the sets  $Z_1$  and  $Z_2$ . Notice that  $Z_1$  is always within the boundaries of  $W_{\|\cdot\|_2}(A_{\mathbf{y}}, B_{\mathbf{y}})$ . In this case, the set  $Z_2$  is close to the boundaries of  $W_{\|\cdot\|_2}(A_{\mathbf{y}}, B_{\mathbf{y}})$  but it is not included on it. This may not be the case on all the realizations. Erroneous classifications are expected on those realizations.

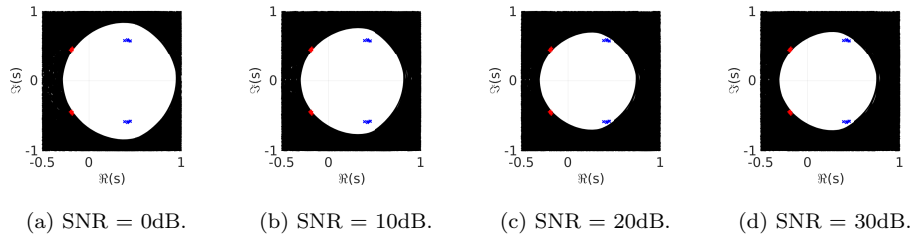


Figure 5: Numerical Range  $W_{\|\cdot\|_2}(A_{\mathbf{y}}, B_{\mathbf{y}})$  for different SNR. Sets  $Z_1$  (blue  $\times$ ),  $Z_2$  (red  $+$ ).

We compare the performances of CRNR and GLRT by computing the error rates. Fig. 6 shows the results. For low SNRs, from -5dB to 0dB, GLRT and CRNR have similar performances. However, when the SNR increases, there is approximately 2 dB difference between both methods.

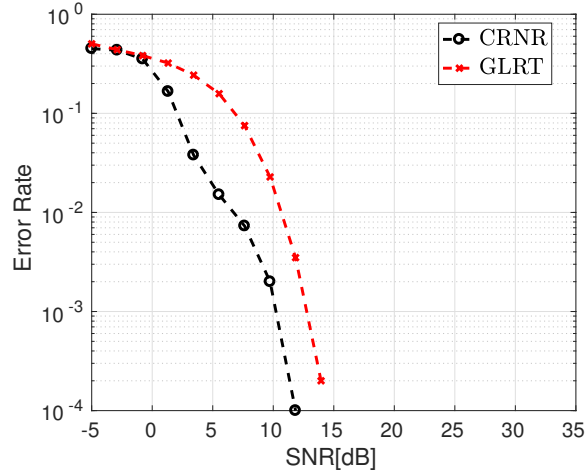


Figure 6: Error rates for observed signal  $\mathbf{y}_t \in Z_1$  and classification set  $Z_2$ .

### 6.3. An unknown number of components

In this section, we compare the performance of the CRNR with the GLRT when  $p \neq M$ . The observed signal is as before built from  $Z_1$ . However, in this case, we estimate the order of the model using [16] and with this order we consider the number of complex frequencies for the classification test. As before, we run a Montecarlo experiment with 10000 realizations. Both methods are applied to each realization. The results are shown in Fig. 7. We observe that the CRNR outperforms the GLRT, which suffers when a wrong model order is considered.

### 6.4. Scattering problem

A scattering phenomenon occurs when an electromagnetic wave strikes a small object and the propagation continues over new directions, different from the incident one. Accurate mathematical models are very complex as they require the solution of Maxwell's equations for appropriate boundary conditions. The author in [17] introduced an approximation known as the

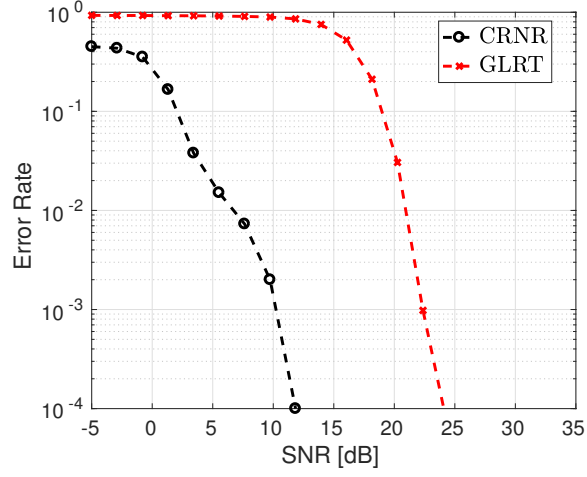


Figure 7: Error rates for observed signal  $\mathbf{y}_t$  and candidate set  $Z_2$ .

singularity expansion method (SEM) to describe the scattered field arriving at a given point in space as the combination of two components: a first one dominated by direct reflexions from the object, called the early-time component; a second one, which is due to induced surface currents on the object or cavity waves, called the late-time component. The SEM models the signal scattered from a target and observed from different angles as follows:

$$\mathbf{y}_t = \sum_{i=1}^M \mathbf{c}_i z_i^{(t-\tau_i)} + \hat{\mathbf{y}}_t, \quad t = 0, 1, \dots \quad (16)$$

where  $z_i$  and  $\mathbf{c}_i$  are the complex natural resonances of the scattering object and the associated vectors of residues,  $\tau_i$  is a constant delay, usually unknown, and  $\hat{\mathbf{y}}_t$  is an entire function that represents the early-time component.

Notice that the residues associated with each resonant modes are aspect dependent. This implies that some natural mode may not be significantly observed when its associated residue has very low energy or even zero along

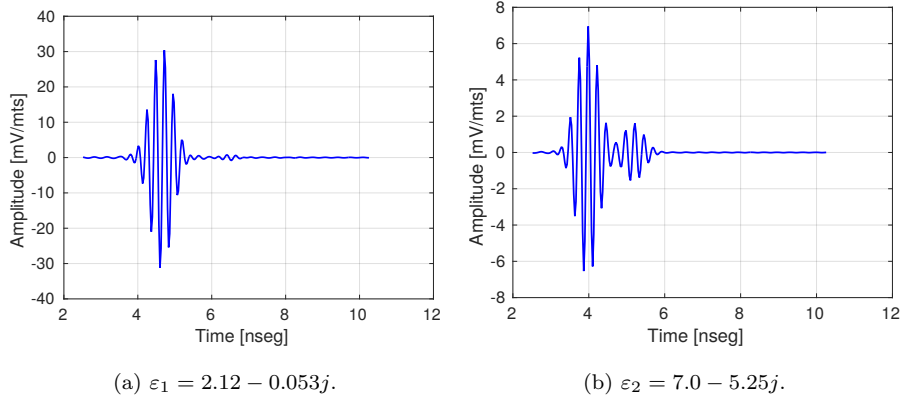


Figure 8: Scattered electric field obtained from spheres made of different materials.

certain observation angle. Having multiple observation angles will be beneficial to overcome this difficulty.

Internal resonances are associated with the composition of the particular object being illuminated. For example, Fig. 8 shows the backscattered signals for two different spheres of the same radius and different material. In the past, it has been proposed to use internal resonances to classify objects using scattering fields [18, 19, 20]. All of these methods work in the resonance region, namely the late-time component of the signal. However, a clear separation between the early-time and the late-time components is not possible [21]. Usually, involved frequency-time techniques are implemented to achieve the signal separation in these two portions [3].

To perform our study, we will consider spherical objects made of a homogeneous material and illuminated by a plane wave. In this case, the solution to Maxwell's equations is known as the Mie series [22]. We use this result to simulate the scattered field from a sphere with radius,  $a = 0.07$  meters made of dielectric material with dielectric constant  $\epsilon$ . We observe the response to a Gaussian pulse in the frequency range  $1GHz$  to  $5.9GHz$  for  $K = 3$  differ-

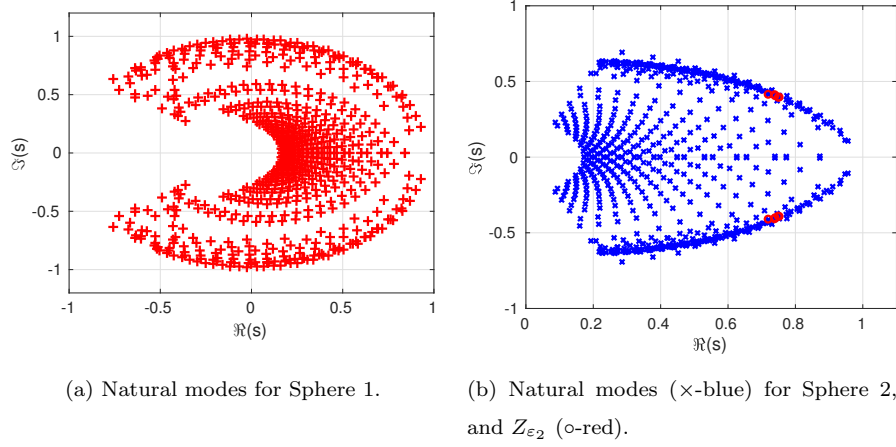


Figure 9: Natural modes for Sphere 1 and 2

ent observation angles,  $\phi = [0^\circ, 45^\circ, 180^\circ]$ . To set the classification problem, we consider two different materials, Class 1 with  $\varepsilon_1 = 2.12 - 0.053j$ , and Class 2 with  $\varepsilon_2 = 7.0 - 5.25j$ .

Fig. 9 shows some of the natural frequencies of Class 1 and Class 2. Both classes have a very large number of natural frequencies. However, it is possible to select disjoint subsets to characterize each class. The appropriate selection of the candidate modes is a delicate problem that determines the performance of the selection technique. In this paper, we choose a subset of  $p = 10$  modes to make the candidate class  $Z_{\epsilon_2}$  as shown in Fig. 9b.

Using the observed signal from Sphere 1 and the classification set  $Z_{\epsilon_2}$ , we follow Algorithm 1. We compare the performance of CRNR with the GLRT approach. In this case, we have used  $p = 10$  known resonances to build the test. Fig. 10 shows the classification error rates obtained for these methods. The CRNR outperforms the GLRT, which shows a poor performance when a reduced order model is considered.

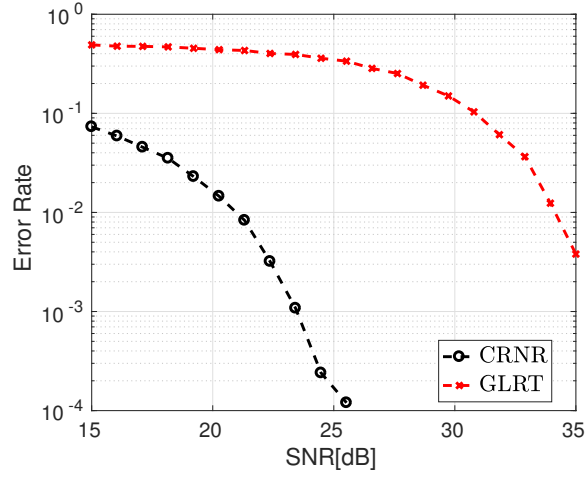


Figure 10: Error rate for observation pair  $(A_{\varepsilon_1}, B_{\varepsilon_1})$  and candidate set  $Z_{\varepsilon_2}$

## 7. Conclusion

In this work, we have proposed a novel classification algorithm for signals composed of sums of exponentially damped sinusoids. Instead of estimating the frequencies and compared them to theoretical frequencies in a database, we have formulated an inverse problem. That is, given a set of frequencies, we evaluate if they correspond to the signal that we are observing. For that, we test the boundaries of the rectangular numerical range of the matrix pencil associated with the observed signal. With this proposition, we overcome some drawbacks observed in the matrix pencil method, just like the instability of the computation of the eigenvalues and the estimation of the model order.

We have applied this new scheme to the classification of dielectric material using the backscattered response of spherical objects when illuminated with ultra-wideband pulses. We have shown that the CRNR shows a good performance. As we mentioned, the resonant frequencies are independent



of the aspect angle and polarization but their associated residues are not. So, theoretically we will have a unique set of infinite resonant frequencies associated with an object but when we measure the scattering signals we will observe only a subset of resonant frequencies. This makes the choice of candidate frequencies a critical step for the classification scheme. Nevertheless, the proposed scheme allows reducing by at least an order of magnitude the classification error compare to other existing techniques.

## Appendix A. Reduced rank Hankel matrix approximation

In this section, we summarize the procedure presented in [10, 7]. Consider  $H$  as in section 2, and define its singular value decomposition

$$H = \sum_{i=1}^{n+1} \sigma_i \mathbf{u}_i \mathbf{v}_i^H, \quad (\text{A.1})$$

where  $\mathbf{u}_i \in \mathbb{C}^m$ ,  $\mathbf{v}_i \in \mathbb{C}^{n+1}$ , and  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{n+1} \geq 0$ .

Under the presence of noise,  $H$  is full rank and has Hankel structure. The idea is to obtain a reduced-rank approximation that preserves the Hankel structure. This is important since it has been noticed that there is a one-to-one correspondence between data sequences consisting of superimposed damped sinusoids and rank-deficient Hankel matrices [7].

Let us define two matrix operators. The first one, performs an  $M$ -order rank reduction, i.e.,  $\mathcal{L} : \mathbb{C}^{m \times (n+1)} \rightarrow \mathbb{C}^{m \times (n+1)}$ ,

$$\mathcal{L}\{H\} = \sum_{i=1}^M \sigma_i \mathbf{u}_i \mathbf{v}_i^H. \quad (\text{A.2})$$

The second operator, obtains a Hankel approximation, i.e.,  $\mathcal{T} : \mathbb{C}^{m \times (n+1)} \rightarrow \mathbb{C}^{m \times (n+1)}$ ,  $Y = \mathcal{T}\{X\}$ , where the  $(k, l)$ -th element of  $Y$  ( $k = 1, \dots, m$ ,

$l = 1, \dots, n + 1$ ) is obtained by averaging the elements of the anti-diagonal of  $X$  as

$$y_{k,l} = \frac{1}{|\Lambda_{k+l}|} \sum_{(k',l') \in \Lambda_{k+l}} x_{k'l'}. \quad (\text{A.3})$$

Here,  $\Lambda_{k+l}$  is the set of indices corresponding to the  $(k+l)$ -th anti-diagonal of  $X$ , and  $|\Lambda_{k+l}|$  is the cardinality of this set.

To obtain a rank-deficient Hankel matrix we use an iterative approach by applying alternately these two operators. First, we make a low-rank approximation and then we apply the Hankel approximation procedure. The procedure is repeated until the following stopping criterion is met

$$\|(\mathcal{TL})^r\{X\} - \mathcal{L}(\mathcal{TL})^{r-1}\{X\}\|_F < \epsilon, \quad (\text{A.4})$$

where  $r$  is the number of iterations performed and  $\epsilon$  is a constant previously defined.

## Acknowledgment

Raymundo Albert is recipient of a doctoral scholarship from CONICET - Argentina. This work was part of the projects UBACyT 20020130100751BA, funded by the University of Buenos Aires, and PIP-112 201101 00997, funded by CONICET .

## References

## References

- [1] Y. Hua, T. Sarkar, Matrix pencil method for estimating parameter of exponentially damped/undamped sinusoids in noise, *IEEE Transactions on Acoustics Speech and Signal Processing* 38 (5) (1990) 814–824.

- [2] J. Mooney, Z. Ding, L. S. Riggs, Robust target identification in white gaussian noise for ultra wide-band radar systems, *IEEE Transactions on Antennas and Propagation* 46 (1998) 12.
- [3] W. C. Chen, N. Shuley, Robust target identification using a modified generalized likelihood ratio test, *IEEE Transactions on Antennas and Propagation* 62 (1) (2014) 264–273.
- [4] R. A. Horn, C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
- [5] F. F. Bonsall, J. Duncan, *Numerical Ranges of Operators on Normed Space and of elements of Normed Algebras*, no. v. 1 in *London Mathematical Society Lecture Note Series*, Cambridge University Press, 1971. doi:10.1017/CBO9781107359895.
- [6] C. Chorianopoulos, S. Karanasios, P. Psarrakos, A definition of numerical range of rectangular matrices, *Linear and Multilinear Algebra* 57 (5) (2009) 459–475. doi:10.1080/03081080802466365.  
URL <https://doi.org/10.1080/03081080802466365>
- [7] Y. Li, K. J. R. Liu, J. Razavilar, A parameter estimation scheme for damped sinusoidal signals based on low-rank hankel approximation, *IEEE Transactions on Signal Processing* 45 (2) (1997) 481–486. doi:10.1109/78.554314.
- [8] B. Lu, D. Wei, B. L. Evans, A. C. Bovik, Improved matrix pencil methods, in: *Conference record of Thirty-Second Asilomar Conference on Signals, Systems and Computers* (Cat. No.98CH36284), Vol. 2, 1998, pp. 1433–1437. doi:10.1109/ACSSC.1998.751563.

- [9] F. R. Gantmacher, The Theory of matrices, no. v. 2 in The Theory of Matrices, Chelsea Pub. Co., 1960.
- [10] Y. Hua, T. K. Sarkar, On svd for estimating generalized eigenvalue of singular matrix pencil in noise, *IEEE Transactions on Signal Processing* 39 (4) (1991) 892–900. doi:10.1109/78.80911.
- [11] J. A. Cadzow, Signal enhancement a composite property mapping algorithm, *IEEE Transactions on Acoustics Speech and Signal Processing* 36 (1) (1988) 49–62. doi:10.1109/29.1488.
- [12] H. Park, L. Zhang, J. B. Rosen, Low rank approximation of a hankel matrix by structured total least norm, *BIT Numerical Mathematics* 39 (4) (1999) 757–779. doi:10.1023/A:1022347425533.  
URL <https://doi.org/10.1023/A:1022347425533>
- [13] C. Dicle, O. I. Camps, M. Sznajder, The way they move: Tracking multiple targets with similar appearance, in: 2013 IEEE International Conference on Computer Vision, 2013, pp. 2301–2311. doi:10.1109/ICCV.2013.286.
- [14] G. Boutry, M. Elad, G. H. Golub, P. Milanfar, The generalized eigenvalue problem for nonsquare pencils using a minimal perturbation approach, *SIAM Journal on Matrix Analysis and Applications* 27 (2) (2005) 582–601. doi:10.1137/S0895479803428795.  
URL <https://doi.org/10.1137/S0895479803428795>
- [15] P. Stoica, Y. Selen, Model-order selection: a review of information criterion rules, *IEEE Signal Processing Magazine* 21 (4) (2004) 36–47. doi:10.1109/MSP.2004.1311138.

- [16] M. Gavish, D. L. Donoho, The optimal hard threshold for singular values is  $4/\sqrt{3}$ , *IEEE Transactions on Information Theory* 60 (8) (2014) 5040–5053. doi:10.1109/TIT.2014.2323359.
- [17] C. E. Baum, On the singularity expansion method for the solution of electromagnetic interaction problems, Defense Technical Information Center, 1971.
- [18] M. H. Bannis, F. M. El-Hefnawi, H. M. A. E. Kader, K. ElMahgoub, A. Z. Elsherbeni, Breast cancer detection and identification using prony’s method, in: 2014 IEEE Antennas and Propagation Society International Symposium (APSURSI), 2014, pp. 1926–1927. doi:10.1109/APS.2014.6905290.
- [19] W. Lee, T. K. Sarkar, H. Moon, M. Salazar-Palma, Computation of the natural poles of an object in the frequency domain using the cauchy method, *IEEE Antennas and Wireless Propagation Letters* 11 (2012) 1137–1140. doi:10.1109/LAWP.2012.2219846.
- [20] G. Turhan-Sayan, Natural resonance-based feature extraction with reduced aspect sensitivity for electromagnetic target classification, *Pattern Recognition* 36 (7) (2003) 1449 – 1466. doi:http://dx.doi.org/10.1016/S0031-3203(02)00290-X.  
URL <http://www.sciencedirect.com/science/article/pii/S003132030200290X>
- [21] L. W. Pearson, D. R. Wilton, R. Mittra, Some implications of the laplace transform inversion on sem coupling coefficients in the time domain, *electromagnetics* 2 (3) (1982) 181–200.

- [22] C. F. Bohren, D. Huffman, Absorption and Scattering of light by small particle, Wiley science paperback series, Wiley, 1983.