

SELECTING THE MODELING ORDER FOR THE ESPRIT HIGH RESOLUTION METHOD: AN ALTERNATIVE APPROACH

Roland Badeau, Bertrand David and Gaël Richard

École Nationale Supérieure des Télécommunications
46 rue Barrault, 75634 Paris Cedex 13 France
roland.badeau, bertrand.david, gael.richard@enst.fr

ABSTRACT

High Resolution methods, such as the ESPRIT algorithm, perform an accurate representation of an harmonic signal as a sum of exponentially damped sinusoids. However, in coding applications, the signal must be represented with a minimum number of parameters. Unfortunately, it is well known that applying the ESPRIT algorithm with an under-estimated model order generates biased frequency estimates. In this paper, we propose a new method for selecting an appropriate modeling order, which minimizes this bias. This approach was applied to both synthetic and musical signals and outperformed the classical information theoretic criteria.

1. INTRODUCTION

Estimating a line spectrum is an important task for many applications, such as speech signals analysis and synthesis [1] and musical signals modification [2]. Although the Fourier transform remains a prominent tool for frequency estimation, the ESPRIT algorithm [3] overcomes the resolution limit of the Fourier analysis and provides straight estimates of the model parameters. This method relies on the rotational invariance property of the signal subspace spanned by the sinusoids. Its drawback is that the model order is supposed to be known, which is not the case in many applications.

Many methods have been proposed in the literature for estimating the number of sinusoids. Most of them rely on an additive white noise hypothesis. This is the case of the maximum likelihood method [4], the information theoretic criteria [5] and the Wishart matrices approach [6]. In the presence of a correlated noise, these methods tend to overestimate the model order. Consequently, specific methods have been designed to address the colored noise case, such as new information theoretic criteria [7] and a technique based on a band noise covariance matrix model [8].

In this paper, we show that applying the ESPRIT high resolution method with an erroneous model order perturbs the estimation of the sinusoids, and we derive error bounds which can be easily computed (a similar bias phenomenon was analyzed in [9] in the case of the MUSIC algorithm). Based on this result, our new model order selection method consists in minimizing this perturbation. Contrary to the above mentioned methods, which achieve the model order selection by analyzing the spectral properties of the additive noise, our approach focuses on the signal itself. Although it relies on a noiseless model, we observed that it outperforms the classical information theoretic criteria, even in low SNR scenarios.

The paper is organized as follows. Section 2 summarizes the principles of the ESPRIT method. In section 3, the perturbation of the poles induced by an under-estimated model order is analyzed.

Then our new model order selection method, referred to as the ESTER method, is introduced in section 4, where a fast implementation is proposed. In section 5, the performance of this method is compared to that of the AIC and MDL criteria [5]. Finally, the main conclusions of this paper are summarized in section 6.

2. THE ESPRIT METHOD FOR ESTIMATING MULTIPLE FREQUENCIES

The noiseless Exponentially Damped Sinusoids (EDS) model [10] defines the discrete signal as $x(t) = \sum_{k=1}^r \alpha_k z_k^t$, where r is the order of the model, $\alpha_k \in \mathbb{C}^*$ are the complex amplitudes, and $z_k \in \mathbb{C}^*$ denote the complex poles. The vector $\mathbf{x}(t) = [x(t - n + 1), \dots, x(t)]^T$ (with $n > r$) belongs to the *signal subspace* spanned by the $n \times r$ Vandermonde matrix

$$\mathbf{V} = \begin{bmatrix} 1 & \dots & 1 \\ z_1 & \dots & z_r \\ \vdots & \vdots & \vdots \\ z_1^{n-1} & \dots & z_r^{n-1} \end{bmatrix}.$$

If the r poles $\{z_1, \dots, z_r\}$ are distinct, \mathbf{V} is full rank, so that the signal subspace has dimension r . Let \mathbf{V}_\perp be the matrix extracted from \mathbf{V} by deleting the last row. Similarly, let \mathbf{V}_\uparrow be the matrix extracted from \mathbf{V} by deleting the first row. Then \mathbf{V} satisfies the so-called *rotational invariance property*:

$$\mathbf{V}_\uparrow = \mathbf{V}_\perp \mathbf{D} \quad (1)$$

where \mathbf{D} is the diagonal matrix $\mathbf{D} = \text{diag}(z_1, \dots, z_r)$.

In practice, \mathbf{V} is unknown, but the signal subspace can be obtained by computing the singular value decomposition (SVD) of a Hankel or Toeplitz data matrix with $n > r$ rows and $l > r$ columns, or the eigenvalue decomposition (EVD) of a square data matrix [10]. Indeed, if $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ are the left singular vectors associated to the singular values $\sigma_1 \geq \dots \geq \sigma_n \geq 0$ sorted in decreasing order, then the signal subspace is spanned by the $n \times r$ orthonormal matrix $\mathbf{W}(r) = [\mathbf{w}_1, \dots, \mathbf{w}_r]$ (the $n - r$ last singular values being equal to 0). Since $\mathbf{W}(r)$ and \mathbf{V} span the same subspace, there is a non-singular matrix \mathbf{G} of dimension $r \times r$ such that $\mathbf{V} = \mathbf{W}(r) \mathbf{G}$. Then $\mathbf{W}(r)$ satisfies an equation similar to equation (1): $\mathbf{W}(r)_\uparrow = \mathbf{W}(r)_\perp \mathbf{\Phi}(r)$, where $\mathbf{\Phi}(r)$ is defined by its eigenvalue decomposition: $\mathbf{\Phi}(r) \triangleq \mathbf{G} \mathbf{D} \mathbf{G}^{-1}$. Finally, the ESPRIT algorithm [3] consists of the following steps:

- computing $\mathbf{W}(r)$ via a singular value decomposition

- computing $\Phi(r) = \mathbf{W}(r)_\downarrow^\dagger \mathbf{W}(r)_\uparrow$ (where the symbol \dagger denotes the Moore-Penrose pseudo-inverse)
- extracting the poles z_k as the eigenvalues of $\Phi(r)$

In practice, the model order r is unknown. We assume below that the ESPRIT algorithm is applied with an erroneous model order p and we focus on how the estimation of the poles is affected. For all $p \in \{1, \dots, n\}$, let $\mathbf{W}(p) \triangleq [\mathbf{w}_1, \dots, \mathbf{w}_p]$ and

$$\Phi(p) \triangleq \mathbf{W}(p)_\downarrow^\dagger \mathbf{W}(p)_\uparrow. \quad (2)$$

The estimated poles are defined as the eigenvalues of $\Phi(p)$. If $p > r$, the following proposition shows that the r poles do no longer form the whole set of eigenvalues of $\Phi(p)$, but they are part of it in the general case.

Proposition 1 *Suppose that $p \geq r$ and $\mathbf{W}(p)_\downarrow$ is full rank. Then $\forall k \in \{1, \dots, r\}$, z_k is an eigenvalue of $\Phi(p)$.*

Proof Let \mathbf{v}_k be the right eigenvector of $\Phi(r)$ associated to the eigenvalue z_k and consider the p -dimensional vector $\bar{\mathbf{v}}_k \triangleq \begin{bmatrix} \mathbf{v}_k \\ \mathbf{0} \end{bmatrix}$. Note that $\mathbf{W}(p)_\downarrow \bar{\mathbf{v}}_k = \mathbf{W}(r)_\downarrow \mathbf{v}_k$. Consequently, $\mathbf{W}(p)_\uparrow \bar{\mathbf{v}}_k = \mathbf{W}(r)_\uparrow \mathbf{v}_k = \mathbf{W}(r)_\downarrow \Phi(r) \mathbf{v}_k = z_k \mathbf{W}(r)_\downarrow \mathbf{v}_k = z_k \mathbf{W}(p)_\downarrow \bar{\mathbf{v}}_k$. Since $\mathbf{W}(p)_\downarrow$ is full rank, left multiplying the previous equality by $\mathbf{W}(p)_\downarrow^\dagger$ yields $\Phi(p) \bar{\mathbf{v}}_k = z_k \bar{\mathbf{v}}_k$. \square

3. UNDER-ESTIMATION OF THE MODEL ORDER

If $p < r$, it is well known that the eigenvalues of $\Phi(p)$ do not match the poles in the general case. Let $\hat{\mathbf{v}}$ be a unitary eigenvector of $\Phi(p)$, associated to the eigenvalue \hat{z} . In this section, it will be shown that \hat{z} approximates one of the eigenvalues of $\Phi(r)$, and that an error bound can be easily computed. First, we need to define the *upper condition number* of the Vandermonde matrix \mathbf{V} : $\kappa_2 = \frac{\sigma_{\max}(\mathbf{V})}{\sigma_{\min}(\mathbf{V}_\downarrow)}$, where $\sigma_{\max}(\cdot)$ denotes the largest singular value of a matrix, and $\sigma_{\min}(\cdot)$ denotes the smallest one. This condition number characterizes the noiseless signal itself, and does not depend on p . It is an unknown constant for our problem, which does not need to be calculated. It is involved in the following theorem, whose proof can be found in the appendix.

Theorem 2 (a priori error bound) *Let $\hat{\mathbf{v}}$ be a given unitary vector, let \hat{z} be a given complex number. Define the residual vector*

$$\mathbf{e}(p) \triangleq \mathbf{W}(p)_\uparrow \hat{\mathbf{v}} - \hat{z} \mathbf{W}(p)_\downarrow \hat{\mathbf{v}}. \quad (3)$$

Then there is an eigenvalue z_k of $\Phi(r)$ for which

$$|\hat{z} - z_k| \leq \kappa_2 \|\mathbf{e}(p)\|_2. \quad (4)$$

In theorem 2, $\|\cdot\|_2$ denotes the Euclidian vector norm (or 2-norm). Note that the bound $\|\mathbf{e}(p)\|_2$ can be computed without knowing the exact value of r . Corollary 3 follows from theorem 2. It gives an error bound valid for all the eigenvalues of $\Phi(p)$. Again, this bound can be computed without knowing the exact value of r . It involves the spectral norm of a matrix (or 2-norm), also denoted $\|\cdot\|_2$, defined as $\|\mathbf{M}\|_2 \triangleq \max_{\|\mathbf{u}\|_2=1} \|\mathbf{M}\mathbf{u}\|_2$.

Corollary 3 (a posteriori error bound) *For each eigenvalue \hat{z} of $\Phi(p)$, there is an eigenvalue z_k of $\Phi(r)$ for which*

$$|\hat{z} - z_k| \leq \kappa_2 \|\mathbf{E}(p)\|_2 \quad (5)$$

where

$$\mathbf{E}(p) = \mathbf{W}(p)_\uparrow - \mathbf{W}(p)_\downarrow \Phi(p). \quad (6)$$

Proof Let \hat{z} be an eigenvalue of $\Phi(p)$ and $\hat{\mathbf{v}}$ a unitary eigenvector associated with \hat{z} . Then equation (3) yields $\mathbf{e}(p) = (\mathbf{W}(p)_\uparrow - \mathbf{W}(p)_\downarrow \Phi(p)) \hat{\mathbf{v}}$. Since $\hat{\mathbf{v}}$ is unitary, applying the 2-norm yields

$$\|\mathbf{e}(p)\|_2 \leq \|\mathbf{W}(p)_\uparrow - \mathbf{W}(p)_\downarrow \Phi(p)\|_2 \quad (7)$$

Finally substituting equation (7) into equation (4) yields equation (5). \square

4. SELECTION OF AN APPROPRIATE MODELING ORDER BASED ON THE ESTIMATION ERROR

The practical interest of corollary 3 is that $\|\mathbf{E}(p)\|_2$ (which will be referred to as the *a posteriori error bound*) can be computed for all $p \in \{1, \dots, r\}$, giving a quantitative criterion for selecting an appropriate modeling order lower than r , such that the estimation error bound is minimum. In practice, r is unknown, and the bound will be computed for all $p \in \{1, \dots, p_{\max}\}$, where $1 \leq p_{\max} < n - 1$. If p_{\max} happens to be greater than r , then for all $p > r$ the a posteriori error bound is positive¹, whereas it is zero for $p = r$. Consequently, the global minimum is zero and it is reached for $p = r$, which is the exact model order. In any case, detecting p such that the a posteriori error bound is minimum in the range $[1, p_{\max}]$ is a relevant approach for selecting the modeling order. Following this remark, the ESTimation ERror (ESTER) method consists in maximizing the inverse error function $J : p \mapsto \frac{1}{\|\mathbf{E}(p)\|_2^2}$. Such a function is plotted in section 5. Proposition 4 shows that its values are in the interval $[1, +\infty[$ (the proof is given in the appendix).

Proposition 4 *For all $p \in \{1, \dots, n\}$, $\|\mathbf{E}(p)\|_2 \leq 1$.*

The drawback of the ESTER method is that a direct implementation would lead to a very computationally demanding algorithm. Indeed, once the SVD of the data matrix has been computed, the matrix $\mathbf{E}(p)$ must then be calculated for all $p \in \{1, \dots, p_{\max}\}$. Such a computation would involve $3np^2$ flops for each p , so that the overall complexity would be np_{\max}^3 flops. This complexity is to be compared to that of a direct calculation of the AIC and MDL criteria [5] from the singular values of the data matrix; both of them involve only $2np_{\max}$ flops. Therefore, we developed a recursive implementation of the ESTER method, presented in table 1, which involves only $6np$ flops for each p , so that its overall complexity is $3np_{\max}^2$. In particular, computing the matrices $\mathbf{E}(p)$ for all $p \in \{1 \dots p_{\max}\}$ in this way is not more computationally demanding than computing $\mathbf{E}(p_{\max})$ directly. Although the mathematical derivations leading to the equations in table 1 are not detailed here (they will be presented in a forthcoming paper), note that this algorithm recursively updates the auxiliary matrices $\Psi(p) \triangleq \mathbf{W}(p)_\downarrow^H \mathbf{W}(p)_\uparrow$ and $\Xi(p) \triangleq \mathbf{W}(p)_\uparrow - \mathbf{W}(p)_\downarrow \Psi(p)$.

5. SIMULATION RESULTS

5.1. Synthetic signal

The test signal is a sum of $r = 5$ complex exponentials plus a complex colored noise, defined in the time window $t \in [0, 254]$. For each $k \in \{1, \dots, r\}$, the k^{th} exponential is characterized by

¹Let $\mathbb{P}(p) = \mathbf{I}_{n-1} - \mathbf{W}_\downarrow(p) \mathbf{W}_\downarrow(p)^\dagger$ be the projector onto the orthogonal complement of $\mathcal{Im}(\mathbf{W}_\downarrow(p))$. If $p > r$, $\mathcal{Im}(\mathbf{W}_\uparrow(r)) = \mathcal{Im}(\mathbf{W}_\downarrow(r)) \subset \mathcal{Im}(\mathbf{W}_\downarrow(p))$, therefore $\mathbb{P}(p) \mathbf{W}_\uparrow(r) = \mathbf{0}$. Since $\mathbf{E}(p) = \mathbb{P}(p) \mathbf{W}_\uparrow(p)$, the r left columns of $\mathbf{E}(p)$ are zero. However, the $p - r$ right columns of $\mathbf{W}_\uparrow(p)$ do not belong to $\mathcal{Im}(\mathbf{W}_\downarrow(p))$ in the general case, therefore the $p - r$ right columns of $\mathbf{E}(p)$ are non-zero, and the a posteriori error bound is positive.

Table 1. Recursive computation of $\mathbf{E}(p)$

$$\begin{aligned}
 & \overline{\mathbf{W}}(0) = \mathbf{I}, \quad \overline{\Psi}(0) = \mathbf{I}, \quad \overline{\Xi}(0) = \mathbf{0} \\
 & \text{For } p = 1 \text{ to } p_{\max} \\
 & \left[\begin{array}{l}
 \psi_r(p) = \mathbf{W}(p-1)_{\downarrow}^H \mathbf{w}(p)_{\uparrow} \\
 \psi_l(p) = \mathbf{W}(p-1)_{\uparrow}^H \mathbf{w}(p)_{\downarrow} \\
 \psi_{lr}(p) = \mathbf{w}(p)_{\downarrow}^H \mathbf{w}(p)_{\uparrow} \\
 \Psi(p) = \begin{bmatrix} \overline{\Psi}(p-1) & \psi_r(p) \\ \psi_l(p)^H & \psi_{lr}(p) \end{bmatrix} \\
 \xi(p) = \mathbf{w}(p)_{\uparrow} - \mathbf{W}(p-1)_{\downarrow} \psi_r(p) - \mathbf{w}(p)_{\downarrow} \psi_{lr}(p) \\
 \Xi(p) = \begin{bmatrix} \overline{\Xi}(p-1) - \mathbf{w}(p)_{\downarrow} \psi_l(p)^H & \xi(p) \end{bmatrix} \\
 \mathbf{W}(p) = \begin{bmatrix} \mathbf{W}(p-1) & \mathbf{w}(p) \end{bmatrix} \\
 \mathbf{u}(p) = \mathbf{W}(p)_{(n,:)}^H \\
 \varphi(p) = \Psi(p)^H \mathbf{u}(p) \\
 \mathbf{E}(p) = \overline{\Xi}(p) - \frac{1}{1 - \|\mathbf{u}(p)\|^2} (\mathbf{W}(p)_{\downarrow} \mathbf{u}(p)) \varphi(p)^H
 \end{array} \right.
 \end{aligned}$$

its amplitude α_k and its pole $z_k = e^{-\delta_k + i2\pi f_k}$, where $f_k \in \mathbb{R}$ is its frequency and $\delta_k \in \mathbb{R}$ is its damping factor. The values of the parameters f_k are 0, 0.1, 0.3, 0.6 and 0.8 Hz, and for all $k \in \{1, \dots, r\}$, $\delta_k = 0$ and $\alpha_k = 1$. The additive noise has been obtained by applying the filter $H(z) = \frac{1}{1 - 0.99z^{-1}}$ to a complex white gaussian noise, whose variance has been chosen so that the resulting Signal Noise Ratio (SNR) is 10 dB. The ESTER method is compared to three Information Theoretic Criteria (ITC), known as the *Akaike Information Criterion* (AIC) [5], the *Minimum Description Length* (MDL) [5], and the *Efficient Detection Criterion* (EDC) [11] which is known to be a robust generalization of AIC and MDL. These methods consist in minimizing a cost function which involves the singular values $\{\sigma_1, \dots, \sigma_n\}$:

$$\text{ITC}(p) = -(n-p) l \ln \left(\frac{\left(\prod_{q=p+1}^n \sigma_q^2 \right)^{\frac{1}{n-p}}}{\frac{1}{n-p} \sum_{q=p+1}^n \sigma_q^2} \right) + p(2n-p) C(l) \quad (8)$$

where $C(l)$ is a function of l . The AIC criterion is defined by choosing $C(l) = 1$ and the MDL criterion is defined by choosing $C(l) = \frac{1}{2} \ln(l)$. The EDC criteria are obtained for all functions of l such that $\lim_{l \rightarrow +\infty} \frac{C(l)}{l} = 0$ and $\lim_{l \rightarrow +\infty} \frac{C(l)}{\ln(\ln(l))} = +\infty$. We chose $C(l) = \sqrt{l \ln(\ln(l))}$, for which we obtained the best results. The singular values have been obtained by computing the SVD of a Hankel data matrix containing $n = 128$ rows and $l = 128$ columns, involving the $n + l - 1 = 255$ samples of the whole signal. Figure 1-a displays the AIC (solid line), MDL (dashed line) and EDC (dotted line) criteria, such as formulated in the above equation. None of them reaches a minimum at $p = r = 5$. This failure might be explained by the presence of the surrounding noise, whose power spectral density is not uniform, contrary to the additive white noise hypothesis on which these estimators basically rely. As expected, the EDC criterion is more robust than AIC and MDL, but its minimum is obtained for $p = 7$.

Figure 1-b displays new ITC criteria proposed in [7] to address the colored noise case². The best results were obtained with the

²These new criteria are referred to as C_1 (solid line), C_2 (dashed line), C_{m1} (dotted line) and C_{m2} (circles line). The common value of the parameters M_1 and M_2 defined in [7] was set to $\frac{n}{2} - 1$.

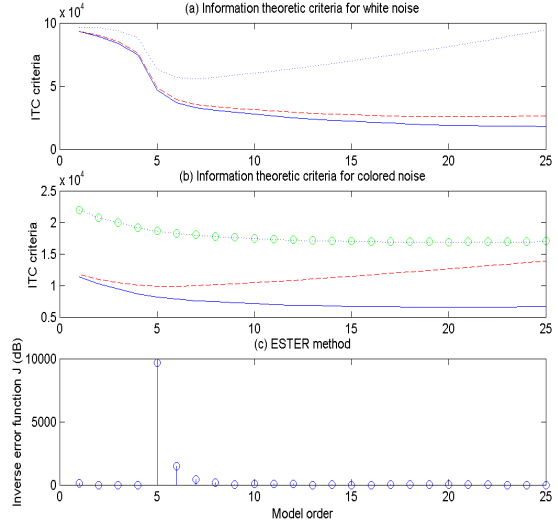


Fig. 1. Model order selection for a synthetic signal

dashed line criterion, which reaches a minimum at $p = 6 \simeq r$.

Concurrently, figure 1-c displays the inverse error function J for $p \in \{1, \dots, p_{\max}\}$. It can be noticed that the global maximum is reached at $p = r = 5$, despite the surrounding noise, which was not included in the model.

5.2. Musical signal

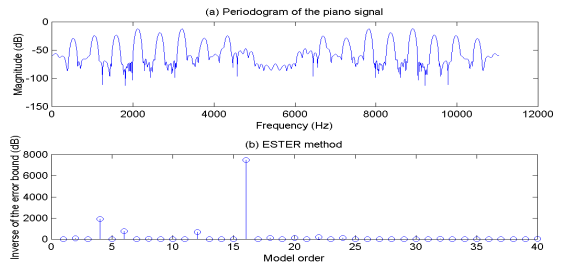


Fig. 2. Periodogram and ESTER criterion for a piano signal

This section illustrates the application of the ESTER method to a musical signal. The study deals with a piano tone, C5, sampled at 11025 Hz, from which a segment of 255 samples (23 ms) has been extracted. Since the energy of audio signals is not evenly distributed over frequencies, we used a pre-emphasis filter obtained by linear prediction at order 7 to compensate the energy decrease. The periodogram of the filtered piano signal is displayed in figure 2-a. We used a Blackman window, because of its high leakage rejection (-57 dB), and despite its lobe width (three times larger than that of the rectangular window). In this figure, sixteen sharp spectral peaks clearly upraise from the surrounding noise level. We observed that the ITC criteria mentioned in section 5.1 present

a high convexity at $p = 16$, but their penalty term is not sufficient to make this point be the global minimum. Figure 2-b displays the inverse error function J for all $p \in \{1, \dots, p_{\max}\}$. The global maximum is reached at $p = 16$. It can also be noticed that the error bounds obtained for lower values of p are relevant. Indeed, high values are reached at $p = 4, 6, 12$, which in fact correspond to small jumps in the singular values decrease (not represented here). Therefore, the ESTER method gives the expected model order, and moreover the error bounds can be used to quantify the adequacy of a possible lower modeling order. In particular, it can be noticed that odd model orders do not fit the signal. Indeed, since this signal is real valued and centered, its spectrum is hermitian symmetric with no constant component, which underlies an even model order.

6. CONCLUSIONS

In this paper, we presented some results on how the estimation of a noiseless EDS model is affected by applying the ESPRIT algorithm with an erroneous model order. If the model order is overestimated, the true poles are among the estimated ones. On the contrary, if the order is underestimated, the estimated poles can be seen as approximations of some of the true ones. In this last case, an *a posteriori error bound* was given, which can be computed without knowing the exact model order. Following from this observation, we introduced the *ESTER* method for selecting an appropriate model order. Since the basic method was computationally expensive, we proposed a fast algorithm for recursively computing the a posteriori error bounds. Then, we illustrated the performance of the ESTER method on a synthetic signal and on a piano signal. Moreover, we noted that the error bounds could be used to quantify the adequacy of a possible lower modeling order. Finally, although it was designed for the ESPRIT algorithm, the ESTER criterion can be used with any High Resolution method.

Appendix

Proof of theorem 2 If $\hat{z} = z_k$ for some $k \in \{1 \dots r\}$, the assertion is trivial, so we may assume that $\forall k \in \{1 \dots r\}, \hat{z} \neq z_k$. Note that $\mathbf{W}(p) \hat{\mathbf{v}} = \mathbf{W}(r) \begin{bmatrix} \hat{\mathbf{v}} \\ \mathbf{0} \end{bmatrix}$, so that equation (3) yields

$$\mathbf{e}(p) = (\mathbf{W}(r)_\uparrow - \hat{z} \mathbf{W}(r)_\downarrow) \begin{bmatrix} \hat{\mathbf{v}} \\ \mathbf{0} \end{bmatrix}. \quad (9)$$

Since $\mathbf{V} = \mathbf{W}(r) \mathbf{G}$, we have $\mathbf{W}(r)_\downarrow = \mathbf{V}_\downarrow \mathbf{G}^{-1}$, and equation (1) yields $\mathbf{W}(r)_\uparrow = \mathbf{V}_\uparrow \mathbf{D} \mathbf{G}^{-1}$. Therefore, equation (9) yields

$$\mathbf{e}(p) = \mathbf{V}_\downarrow (\mathbf{D} - \hat{z} \mathbf{I}_r) \mathbf{G}^{-1} \begin{bmatrix} \hat{\mathbf{v}} \\ \mathbf{0} \end{bmatrix}. \quad (10)$$

Since $\forall k \in \{1 \dots r\}, \hat{z} \neq z_k$, $\mathbf{D} - \hat{z} \mathbf{I}_r$ is non-singular. Therefore, equation (10) yields

$$\begin{bmatrix} \hat{\mathbf{v}} \\ \mathbf{0} \end{bmatrix} = \mathbf{G} (\mathbf{D} - \hat{z} \mathbf{I}_r)^{-1} \mathbf{V}_\downarrow^\dagger \mathbf{e}(p). \quad (11)$$

Then applying the 2-norm into equation (11) yields

$$\|\hat{\mathbf{v}}\|_2 \leq \|\mathbf{G}\|_2 \left\| (\mathbf{D} - \hat{z} \mathbf{I}_r)^{-1} \right\|_2 \left\| \mathbf{V}_\downarrow^\dagger \right\|_2 \|\mathbf{e}(p)\|_2. \quad (12)$$

Since $\mathbf{W}(r)$ is orthonormal and $\mathbf{V} = \mathbf{W}(r) \mathbf{G}$, $\|\mathbf{G}\|_2 = \|\mathbf{V}\|_2 = \sigma_{\max}(\mathbf{V})$. Moreover, $(\mathbf{D} - \hat{z} \mathbf{I}_r)^{-1}$ is diagonal with diagonal entries $\frac{1}{z_k - \hat{z}}$, thus $\left\| (\mathbf{D} - \hat{z} \mathbf{I}_r)^{-1} \right\|_2 = \frac{1}{\min_{k \in \{1 \dots r\}} |z_k - \hat{z}|}$. The singular values

of $\mathbf{V}_\downarrow^\dagger$ are the inverses of those of \mathbf{V}_\downarrow , so that $\left\| \mathbf{V}_\downarrow^\dagger \right\|_2 = \frac{1}{\sigma_{\min}(\mathbf{V}_\downarrow)}$. Finally, since $\hat{\mathbf{v}}$ is unitary, equation (12) yields

$$\min_{k \in \{1 \dots r\}} |\hat{z} - z_k| \leq \frac{\sigma_{\max}(\mathbf{V})}{\sigma_{\min}(\mathbf{V}_\downarrow)} \|\mathbf{e}(p)\|_2. \quad (13)$$

Finally, the assertion follows from equation (13). \square

Proof of proposition 4 It can be noticed that

$$\mathbf{E}(p) = (\mathbf{W}(p)_\uparrow \mathbf{W}(p)_\uparrow^\dagger - \mathbf{W}(p)_\downarrow \mathbf{W}(p)_\downarrow^\dagger) \mathbf{W}(p)_\uparrow$$

Applying the 2-norm yields

$$\|\mathbf{E}(p)\|_2 \leq \text{dist}(\mathbf{W}(p)_\uparrow, \mathbf{W}(p)_\downarrow) \|\mathbf{W}(p)_\uparrow\|_2 \quad (14)$$

where $\text{dist}(\mathbf{W}(p)_\uparrow, \mathbf{W}(p)_\downarrow) \triangleq \left\| \mathbf{W}(p)_\uparrow \mathbf{W}(p)_\uparrow^\dagger - \mathbf{W}(p)_\downarrow \mathbf{W}(p)_\downarrow^\dagger \right\|_2$ is the distance between subspaces $\text{Span}(\mathbf{W}(p)_\uparrow)$ and $\text{Span}(\mathbf{W}(p)_\downarrow)$, which satisfies $\text{dist}(\mathbf{W}(p)_\uparrow, \mathbf{W}(p)_\downarrow) \leq 1$, as shown in [12, pp. 76-77]. Since $\|\mathbf{W}(p)_\uparrow\|_2 \leq 1$, the result follows from equation (14). \square

7. REFERENCES

- [1] R. J. McAulay and T. F. Quatieri, "Speech analysis and synthesis based on a sinusoidal representation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, no. 4, 1986.
- [2] B. David, G. Richard, and R. Badeau, "An EDS modeling tool for tracking and modifying musical signals," in *Proc. of SMAC 03*, Stockholm, Sweden, 2003, vol. 2, pp. 715-718.
- [3] R. Roy and T. Kailath, "ESPRIT-Estimation of Signal Parameters via Rotational Invariance Techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 7, 1989.
- [4] G. Bienvenu and L. Kopp, "Optimality of high-resolution array processing using the eigensystem method," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 31, no. 5, 1983.
- [5] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, no. 2, pp. 387-392, Apr. 1985.
- [6] J. Grouffaud, P. Larzabal, and H. Clergeot, "Some properties of ordered eigenvalues of a Wishart matrix: application in detection test and model order selection," in *Proc. of IEEE ICASSP*, 1996, vol. 5, pp. 2465-2468.
- [7] Q. T. Zhang and K. M. Wong, "Information theoretic criteria for the determination of the number of signals in spatially correlated noise," *IEEE Trans. Signal Processing*, vol. 41, no. 4, pp. 1652-1663, Apr. 1993.
- [8] J. J. Fuchs, "Estimation of the number of signals in the presence of unknown correlated sensor noise," *IEEE Trans. Signal Processing*, vol. 40, no. 5, pp. 1053-1061, May 1992.
- [9] H. Saarnisaari, "Robustness of the MUSIC algorithm to Errors in Estimation the Dimensions of the Subspaces: Delay Estimation in DS/SS in the Presence of Interference," in *Proc. of MILCOM'99*, 1999.
- [10] R. Badeau, R. Boyer, and B. David, "EDS parametric modeling and tracking of audio signals," in *Proc. Int. Conf. on Digital Audio Effects DAFX-02*, Sept. 2002, pp. 139-144.
- [11] L. C. Zhao, P. R. Krishnaiah, and Z. D. Bai, "On detection of the number of signals in presence of white noise," *Journal of Multivariate Analysis*, vol. 20, no. 1, pp. 1-25, 1986.
- [12] G. H. Golub and C. F. Van Loan, *Matrix computations*, The Johns Hopkins University Press, Baltimore and London, third edition, 1996.