ON THE MULTI-RESOLUTION ESPRIT ALGORITHM

Aweke N. Lemma¹ aweke@cas.et.tudelft.nl Alle-Jan van der Veen allejan@cas.et.tudelft.nl

Ed F. Deprettere ed@cas.et.tudelft.nl

Delft Univ. of Technology, Dept. of Electrical Eng., Mekelweg 4, 2628 CD Delft, The Netherlands Tel. +31-15-2786465 Fax. +31-15-2786190

Abstract—Multi-resolution ESPRIT is an extension of the ESPRIT direction finding algorithm to antenna arrays with multiple baselines. A short (half wavelength) baseline is necessary to avoid aliasing, a long baseline is preferred for accuracy. The MR-ESPRIT algorithm allows to combine both estimates. The ratio of the longest baseline to the shortest one is a measure of the gain in resolution. In this work, we show that because of various factors, including noise, signal bandwidth and measurement error, the achievable gain in resolution is bounded.

1 INTRODUCTION

Since its derivation, the ESPRIT [1] algorithm has been used for direction-of-arrival estimation, harmonic analysis, frequency estimation, delay estimation, and combinations thereof. In essence, the algorithm makes use of a single shift invariance structure present in the array response vector $a(\theta)$, where $\theta = e^{j\mu}$, and μ is a phase shift to be estimated. In narrowband direction-of-arrival estimation, the phase shift is due to the difference in arrival times of the wavefront at the elements of an antenna array. For a uniform linear array (ULA), it is well known that $a(\theta) = [1 \ \theta \ \theta^2 \ \cdots]^T$ and $\mu = 2\pi\Delta \sin(\alpha)$, where Δ is the distance between the elements (in wavelengths), and α is the angle of arrival measured with respect to the normal of the array axis.

In the literature [2-6], it was shown that the resolution of the estimation of $\sin(\alpha)$ is directly proportional to $\frac{1}{\Delta}$. Thus, it is preferable to have a large baseline separation Δ , so that we collect a large phase shift μ . Unfortunately, however, we cannot collect more than a single cycle, $-\pi \leq \mu < \pi$, because the inverse of the mapping $\mu \to \theta = e^{j\mu}$ is ambiguous outside this range. To prevent aliasing, we thus have to ensure that $\Delta \leq \frac{1}{2}$, which is essentially Shannon's sampling theorem in space.

The idea behind multi-resolution parameter estimation [7,8] is to obtain two or more estimates of μ : the first based on a small baseline, yielding a coarse estimate μ_1 of μ without aliasing, and the second based on a large baseline or (much) larger sampling period, providing an aliased estimate μ_2 of μ at a finer scale. These two estimates are combined to obtain a final estimate $\hat{\mu} = 2\pi n + \mu_2$, where the integer number of cycles n is estimated from μ_1 . The ratio of the longest baseline to the shortest baseline (denoted by k_s and referred to as the resolution gain factor) is a measure of the gain in resolution. Theoretically, the resolution gain factor can be made arbitrarily large. In practice, however, there are factors that bound k_s . These include, among others, the SNR and the array imperfection.

In this work, we give a theoretical analysis of the properties of the MR-ESPRIT. Particularly, we find the bounds on k_s and on the tolerated array imperfection that will allow the proper functioning of the algorithm.

¹The research of A.N. Lemma is supported by TNO-FEL, The Hague, The Netherlands.

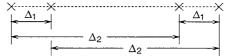


Fig. 1. Multi-resolution spatial sampling 2 THE MR-ESPRIT

The original ESPRIT algorithm is based on arrays with a doublet structure, i.e. consisting of several antenna pairs with the same baseline vectors. However, the chosen array geometries often admit other pairings with different baselines. For instance, the array structure shown in Fig. 1 combines two spatial sampling rates. The minimal number of antennas to have two baseline vector pairs is four. With more antennas, several interesting configurations are possible.

The *M*-dimensional array response vector $\mathbf{a}(\alpha)$ is defined as the response of the *M*-element antenna array to a narrowband signal from a direction α . It can be parameterized in several ways. The usual parameterization is in terms of $\theta = e^{j2\pi\Delta\sin(\alpha)}$, where Δ is a reference distance. In our case of an array with two baselines, we can (redundantly) parameterize the array by two parameters, $\theta_1 = e^{j2\pi\Delta_1\sin(\alpha)}$ and $\theta_2 = e^{j2\pi\Delta_2\sin(\alpha)}$. In the case of the array of Fig. 1, we have

$$\boldsymbol{a}(\theta_1, \theta_2) = \begin{bmatrix} 1\\ \theta_1\\ \theta_2\\ \theta_1 \theta_2 \end{bmatrix}.$$
(1)

The idea is to treat the two parameters as independent and estimate both of them from the measurement data, and only then combine them into a single estimate of $sin(\alpha)$. Estimation is done by exploiting the dual shift-invariance structure of $a(\theta_1, \theta_2)$, i.e., in the above example

where a_i is the *i*-th entry of $a(\theta_1, \theta_2)$. For more general arrays with a dual shift-invariance structure, we can define selection matrices J_{xi} and J_{yi} (i = 1, 2) such that the above relations hold for $J_{xi}a$ and $J_{yi}a$.

Let μ_i (i = 1, 2) be the argument of θ_i . Then, if the distance $\Delta_i < \frac{1}{2}$, the angle of arrival α of the wavefront can be uniquely determined from μ_i using the transformation

$$\alpha = \arcsin(\frac{\mu_i}{2\pi\Delta_i}).$$

However, when $\Delta_i > \frac{1}{2}$, because of aliasing we get a set of cyclically related candidates for α :

$$\alpha(n) = \arcsin(\frac{\mu_i + 2\pi n}{2\pi\Delta_i}).$$

0-7803-5010-3/98/\$10.00©1988 IEEE

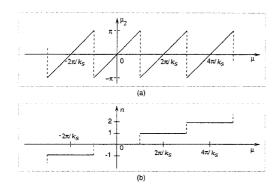


Fig. 2. (a) The aliased spatial frequency μ₂ as a function of the alias-free spatial frequency μ.
(b) The corresponding winding number n.

In MR-ESPRIT we combine non-aliased and aliased estimates of the parameters to obtain a better estimation accuracy. The resulting algorithm is very similar to the case of joint azimuth-elevation estimation [9].

Thus, to be specific, consider d narrowband sources $s_i(t)$ impinging on the antenna array. Collecting N samples of the M antenna outputs into an $M \times N$ data matrix X in the usual way, we obtain the data model

$$X = AS = a_1s_1 + \dots + a_ds_d$$

where the columns of A are the array response vectors $\{a_i\}$, and the rows of S are the sampled source signals. Assuming d < M, the first step of the algorithm is to estimate a basis U_s of the column span of X, typically using an SVD. U_s and A are related by a $d \times d$ nonsingular matrix T as

$$U_s = AT$$

The second step in the algorithm is to form submatrices of U_s using the proper selection matrices:

$$\boldsymbol{U}_{xi} = \boldsymbol{J}_{xi}\boldsymbol{U}_s, \qquad \boldsymbol{U}_{yi} = \boldsymbol{J}_{yi}\boldsymbol{U}_s. \quad (i=1,2)$$

The shift-invariance structure of the array implies that

$$\boldsymbol{U}_{xi} = \boldsymbol{A}'\boldsymbol{T}, \qquad \boldsymbol{U}_{yi} = \boldsymbol{A}'\boldsymbol{\Theta}_i\boldsymbol{T},$$

where A' is a submatrix of A and the diagonal matrix $\Theta_i = \text{diag}\{\theta_{ij}\}_{j=1}^d$ contains the d shift parameters of the d sources with reference to the *i*-th baseline. The final step is to estimate the parameters by considering

$$E_1 = U_{x1}^{\dagger} U_{y1} = T^{-1} \Theta_1 T,$$

$$E_2 = U_{x2}^{\dagger} U_{y2} = T^{-1} \Theta_2 T.$$

It is seen that the data matrices E_1 and E_2 are jointly diagonalizable by the same matrix T. There are several algorithms to compute this joint diagonalization, e.g. by means of Jacobi iterations [9] or QZ iterations [10]. For this to work, it is necessary that each submatrix U_{xi} has at least d rows. After T has been found, we also have estimates of $\{(\theta_{1k}, \theta_{2k})\}$ for each of the d sources.

It remains, for each source, to combine² θ_1 and θ_2 into an estimate of the argument μ of θ . Let us assume that $\Delta_1 \leq \frac{1}{2}$, so that μ_1 (argument of θ_1) is not aliased and is a coarse estimate of μ . Also assume that $\Delta_2 \gg \frac{1}{2}$, so that in μ_2 aliasing occurs: the estimate μ is proportional to μ_2 plus an appropriate integer multiple of 2π (see Fig. 2). It follows that we have two estimates of $2\pi \sin(\alpha)$,

$$2\pi\sin(\alpha) = \frac{1}{\Delta_1}\mu_1 = \frac{1}{\Delta_2}(2\pi n + \mu_2).$$
 (2)

The winding number n is determined as the best fitting integer to match the two right hand side expressions,

$$n = \operatorname{round}\left(\frac{1}{2\pi}(\frac{\Delta_2}{\Delta_1}\mu_1 - \mu_2)\right) =: \operatorname{round}(\hat{n}).$$
(3)

The ratio $k_s := \frac{\Delta_2}{\Delta_1}$ can be interpreted as the (spatial) gain in resolution. In particular, the estimate of $2\pi \sin(\alpha)$ based on μ_2 is a factor k_s more accurate than that based on μ_1 . Thus a more accurate estimate of the spatial frequency μ can be obtained as

$$\mu = \frac{1}{k_s} (2\pi n + \mu_2). \tag{4}$$

3 ANALYSIS

3.1 The winding number

Consider the relations given in (3) and (4), where we have tacitly assumed that $\Delta_2 = k_s \Delta_1$ holds perfectly. In practice however, due to measurement errors, this holds only approximately. Let Δk_s represent the error on k_s such that $\Delta_2 = (k_s + \Delta k_s)\Delta_1$. Also assume that μ_1 and μ_2 are determined with estimation errors $\Delta \mu_1$ and $\Delta \mu_2$, respectively. We further assume that $\Delta \mu_1$ and $\Delta \mu_2$ are independent processes, with $E\{\Delta \mu_1^2\} = E\{\Delta \mu_2^2\} = \sigma_{\mu}^2$. With these assumptions, the error Δn on \hat{n} in (3) can be approximated as

$$\Delta n \approx \frac{\partial \hat{n}}{\partial k_s} \Delta k_s + \frac{\partial \hat{n}}{\partial \mu_1} \Delta \mu_1 + \frac{\partial \hat{n}}{\partial \mu_2} \Delta \mu_2$$

Replacing the value of \hat{n} from (3) into the above equation we obtain

$$\Delta n = \frac{1}{2\pi} \mu_1 \Delta k_s + \frac{1}{2\pi} (k_s \Delta \mu_1 - \Delta \mu_2).$$
 (5)

For a given array configuration, the first term in (5) is a constant. It represents the offset on \hat{n} due to the array imperfection. On the other hand, both parameters $\Delta \mu_1$ and $\Delta \mu_2$ in the second term are zero mean Gaussian processes [4–6].³ Consequently, Δn is also a Gaussian process with a mean $\frac{1}{2\pi}\mu_1\Delta k_s$ and a variance

$$\sigma_n^2 = \mathbb{E}\{(\Delta n - \frac{1}{2\pi}\mu_1 \Delta k_s)^2\} = \frac{1}{4\pi^2}(k_s^2 + 1)\sigma_\mu^2.$$
 (6)

A typical distribution function of Δn is shown in Fig. 3. It

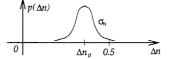


Fig. 3. A typical probability distribution function of Δn , $(\Delta n_o = \frac{1}{2\pi}\mu_1 \Delta k_s)$

is seen from (3) that n is determined correctly if $|\Delta n| < 0.5$. However, since Δn is a random process, we can satisfy this only with some uncertainty (confidence level). In particular, given a required confidence level \mathcal{L} , we find the conditions under which the probability

$$P(|\Delta n| < 0.5) > \mathcal{L} \tag{7}$$

²Here we drop the subscript k in θ_{ik} for readability purpose.

 $^{^{3}}$ More precisely, these are Gaussian processes if the input noise is Gaussian.

Assuming that $P(\cdot)$ is a Gaussian process, it can be shown that, [11]

$$P(|\Delta n| < 0.5) = \frac{1}{2} \operatorname{erf}\left(\frac{\pi + \mu_1 \Delta k_s}{\sigma_\mu \sqrt{2(k_s^2 + 1)}}\right) + \frac{1}{2} \operatorname{erf}\left(\frac{\pi - \mu_1 \Delta k_s}{\sigma_\mu \sqrt{2(k_s^2 + 1)}}\right), (8)$$

where σ_{μ} represents the root mean square measurement error in μ . A family of curves $P(|\Delta n| < 0.5)$, for $\mu_1 = \pi$ (representing worst case scenario) and an arbitrarily chosen value of $\Delta k_s = 0.75$, as functions of σ_{μ} (for different values of k_s) are shown in Fig. 4. To obtain more explicit

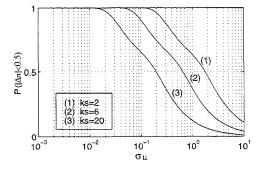


Fig. 4. A family of curves $P(|\Delta n| < 0.5)$ as functions of σ_{μ} for $\mu_1 = \pi$ and $\Delta k_s = 0.75$

expressions, let the function f(x) be defined as

$$f(x) = \frac{1}{2} \operatorname{erf} \left((\pi + \mu_1 \Delta k_s) x \right) + \frac{1}{2} \operatorname{erf} \left((\pi - \mu_1 \Delta k_s) x \right), \quad (9)$$

then $P(|\Delta n| < 0.5)$ may be expressed in terms of f(x) as

$$P(|\Delta n| < 0.5) = f\left(\frac{1}{\sigma_{\mu}^{!}\sqrt{2(k_{s}^{2}+1)}}\right).$$

Now, putting this into (7) and solving for k_s , we get

$$k_s < \sqrt{\frac{1}{2\sigma_{\mu}^2} \left(\frac{1}{f^{-1}(\mathcal{L})}\right)^2 - 1} =: k_{\max},$$
 (10)

where $f^{-1}(\cdot)$ is the inverse function of $f(\cdot)$. From this relation, it is clear that the resolution gain factor cannot be made arbitrarily large. It is bounded from above by a number which is a function of the estimation error and the array imperfection factor Δk_s . Particularly, one can clearly see that, as the estimation error increases, the maximum value of k_s decreases. This is in perfect agreement with intuitive perception. For instance, for the case $\mu_1 = \pi$, $\Delta k_s = 0$ and $\mathcal{L} = 0.998$, the bounds on k_s at $\sigma_{\mu} = 0.1$ and 0.05 are 10.1 and 20.3, respectively. Note that, when $\Delta k_s = 0$, (8) reduces to

$$P(|\Delta n| < 0.5) = \operatorname{erf}\left(rac{\pi}{\sigma_{\mu}\sqrt{2(k_{s}^{2}+1)}}
ight),$$

and the expression for k_{\max} becomes

$$k_{\max} = \sqrt{\frac{\pi^2}{2\sigma_{\mu}^2} \left(\frac{1}{\operatorname{erf}^{-1}(\mathcal{L})}\right)^2 - 1}$$
(11)

3.2 Dependence of k_{max} on SNR

To establish the relation between k_{max} and SNR, we first need to determine the dependence of σ_{μ} (the phase estimation error) on the SNR. To this end, in [2, 3], it is shown that the DOA estimation error and the SNR are related as,

$$\sigma_{\alpha i}^{2} = \frac{1}{SNR} \left(\frac{1}{M^{2}N} \left(\frac{1}{2\pi\Delta_{i}\cos(\alpha)} \right)^{2} \right), \qquad (12)$$

where $\sigma_{\alpha i}$ is the root mean square error (RMSE) obtained with reference to the *i*-th base line separation Δ_i . Recall that $\mu_i = 2\pi\Delta_i \sin(\alpha)$ and, hence

$$\mu_i + \Delta \mu_i = 2\pi \Delta_i \sin(\alpha + \Delta \alpha) \\\approx 2\pi \Delta_i (\sin(\alpha) + \Delta \alpha \cos(\alpha))$$

This implies that $\Delta \mu_i = (2\pi \Delta_i \cos(\alpha)) \Delta \alpha$ and

$$\sigma_{\mu}^{2} = (2\pi\Delta_{i}\cos(\alpha))^{2}\sigma_{\alpha i}^{2}$$
(13)

Here, the index reference to the baseline in $\sigma_{\mu}^2 = E\{(\Delta \mu_i)^2\}$ is dropped because $\Delta \mu_i$ is independent of Δ_i . Now, using (12), σ_{μ}^2 is expressed in terms of array parameters as

$$\sigma_{\mu}^{2} = \frac{1}{SNR} \left(\frac{1}{M^{2}N} \right) \tag{14}$$

Finally, putting (14) into (10), we find the following expression for k_{\max} :

$$k_{\max} = \sqrt{\frac{\mathrm{SNR}}{2}} \left(M^2 N \left(\frac{1}{f^{-1}(\mathcal{L})} \right)^2 \right) - 1 \qquad (15)$$

Note that (12) and, therefore, (15) are derived assuming that there is only one source in the channel. For more than one source (d sources, say), let $\sigma_{\mu j}^2$ represent the variance of the phase estimation error for the *j*-th source.⁴ Then, the bound on k_s is generalized as

$$k_{\max} = \min_{j=1\cdots d} \sqrt{\frac{1}{2\sigma_{\mu j}^2} \left(\frac{1}{f_j^{-1}(\mathcal{L})}\right)^2 - 1}$$

where $f_j(\cdot)$ is as defined in (9), with μ_1 replaced by μ_{1j} (the *j*-th spatial frequency measured with reference to Δ_1).

3.3 Bias on μ due to array imperfections and a self calibrating MR-ESPRIT

Once the winding number n is determined correctly, the next step is to use (4) to estimate the spatial frequency μ . If the array is imperfect, the estimate of μ will be biased. The bias (offset) $\Delta \mu$ on μ due to Δk_s can be approximated by (viz. (4))

$$\Delta \mu \approx \frac{\partial \mu}{\partial k_s} \Delta k_s = \frac{1}{k_s^2} (2\pi n + \mu_2) \Delta k_s, \qquad (16)$$

which indicates that, for a given value of k_s , angles associated with large winding numbers are more affected by Δk_s than those associated with smaller winding numbers.

To minimize the bias, a self calibrating MR-ESPRIT may be implemented. Let T be a finite positive integer, and for $j = 1, \ldots, d$, let $\mu_{1j}(t)$ and $\mu_{2j}(t)$ represent the coarse and fine spatial frequency estimates of the *j*-th wave front at a

⁴For more than one sources $\sigma_{\mu j}$ depends on the SNR in a more complicated way. Interested readers are referred to [3] and [5] for more information.

time index $t \in [t_o, t_o + T - 1]$, respectively. Let also $n_j(t)$ be the estimate of the corresponding winding number. The idea is to first estimate the resolution gain factor as

$$\hat{k}_s = \frac{1}{Td} \sum_{t=t_o+1}^{T} \sum_{j=1}^{a} \frac{2\pi n_j(t) + \mu_{2j}(t)}{\mu_{1j}(t)},$$
(17)

and then insert this estimate into (4) for the computation of the spatial frequency μ . Assuming that the mean of the estimation errors is zero, \hat{k}_s asymptotically converges to its true value. The performance of a self calibrating MR-ESPRIT is compared against a non calibrating MR-ESPRIT in the simulation results.

4 SIMULATION

In this section, we give simulation results that confirm our theory. The simulation example considers a processing band of 10 MHz and a linear antenna array with M = 4 antenna elements arranged as in Fig. 1 with $\Delta_1 = \frac{1}{2}$ and varying Δ_2 . The data is collected into a 4×64 matrix at a sampling rate of $F_1 = 20$ MHz. Two sources emitting narrowband signals (25 kHz) at center frequencies $\mathbf{f} = [6, 6.5]$ MHz, and propagating in distinct directions with DOAs $\boldsymbol{\alpha} = [40, 45]$ degrees are considered.

The results are shown in Fig. 5 through Fig. 7. From the first plot, it is seen that the accuracy of MR-ESPRIT is indeed proportional to the gain factor k_s . An upper limit for this gain is reached when the winding numbers n can no longer be estimated accurately. This is shown in Fig. 6, where the RMSE of the parameter estimator as a function of varying k_s is analyzed. To make the figure less crowded, only the behavior corresponding to DOA = 45 degrees is plotted (the same is true for Fig. 7). It is seen that, for a given SNR, there exists a limit on k_s beyond which the performance of the estimator degrades sharply. Moreover, this bound is seen to be proportional to the SNR as expected.

Finally, simulations showing the improvements in the biases due to self calibration are shown in Fig. 7. The results correspond to $k_s = 20$, and the averaging for the self calibration is taken with T = 20. It is seen that when the bias $\Delta k_s > 0.05$, the self calibrating MR-ESPRIT gives lower bias. This means that, if array placement errors are expected to be large, it is advantageous to implement a self calibrating MR-ESPRIT. Moreover, one can see that there exists a limit on Δk_s beyond which the self calibration process fails. It is further seen that the break away point gets smaller with decreasing SNR as expected.

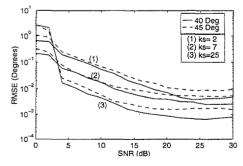


Fig. 5. The root mean square error of the frequency estimates as functions of SNR. $(k_s = 2 \text{ corresponds to ULA})$

REFERENCES

 R.H. Roy, ESPRIT-Estimation of Signal Parameters via Rotational Invariance Techniques. PhD thesis, Stanford University, Stanford, CA, 1987.

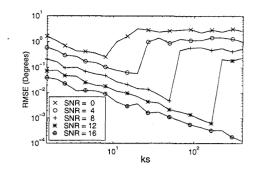


Fig. 6. Root mean square error of the DOA estimates, corresponding to the wavefront with $DOA = 45^{\circ}$, as functions of k_s .

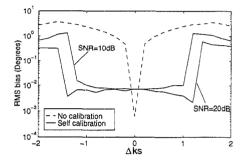


Fig. 7. The bias on DOA estimation due to imperfect array $(k_s = 20, T = 20)$. The noncalibrated behavior is shown for SNR=20dB only, its behavior at 10dB is approximately the same.

- [2] Bhaskar D. Rao and K. V. S. Hari, "Performance analysis of Root-Music," *IEEE Trans. Acous. Speech and Signal Processing*, vol. 37, pp. 1939–1949, December 1989.
- [3] Bhaskar D. Rao and K. V. S. Hari, "Performance analysis of ESPRIT and TAM in determining the direction of arrival of plane waves in noise," *IEEE Trans. Acous. Speech and Signal Processing*, vol. 37, pp. 1990–1995, December 1989.
- [4] Mats Viberg and Björn Ottersten, "Sensor array processing based on subspace fitting," *IEEE Trans. Signal Processing*, vol. 39, pp. 1110–1121, May 1991.
- [5] Björn Ottersten, Mats Viberg, and Thomas Kailath, "Performance analysis of the total least squares ESPRIT algorithm," *IEEE Trans. Signal Processing*, vol. 39, pp. 1122– 1135, May 1991.
- [6] Mats Viberg, Björn Ottersten, and Arye Nehorai, "Performance analysis of direction finding with large arrays and finite data," *IEEE Trans. Signal Processing*, vol. 43, pp. 469– 477, February 1995.
- [7] Aweke N. Lemma, A.-J. van der Veen, and Ed F. Deprettere, "Joint angle-frequency estimation using multiresolution ESPRIT," in *Proc. ICASSP'98*, (Seattle, Washington), May 1998.
- [8] Michael D. Zoltowski and Cherian P. Mathews, "Realtime frequency and 2-D angle estimation with sub-nyquist spatio-temporal sampling," *IEEE Trans. Signal Processing*, vol. SP-42, pp. 2781-2794, October 1994.
- [9] A.-J. van der Veen, P.B. Ober, and E.F. Deprettere, "Azimuth and elevation computation in high resolution DOA estimation," *IEEE Trans. Signal Processing*, vol. 40, pp. 1828-1832, July 1992.
- [10] A.-J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Processing*, vol. 44, pp. 1136–1155, May 1996.
- [11] Simon Haykin, An Introduction to Analog and Digital Communications. John Wiley & Sons, 1986.