Tan, C. M., Nix, A. R., & Beach, M. A. (2002). Problems with direction finding using linear array with element spacing more than half wavelength. (pp. 6 p). (COST 273), (WP 02 008).

Problems with direction finding using linear array with element spacing more than half wavelength

C. M. Tan†, M. A. Beach and A. R. Nix

Centre for Communications Research, University of Bristol,
Merchant Venturers Building, Woodland Road, Bristol, BS8 1UB, UK.
Tel: +44 (0)117 954 5202, Fax: +44 (0)117 954 5206
†Email: Chor.Min.Tan@bristol.ac.uk

ABSTRACT

This paper investigates the performance of different parameter estimation algorithms in direction finding using a linear array with element spacing of more than half a wavelength. The study shows that all the common algorithms namely the beamformer methods, MUSIC, Min-Norm, JoDeG, ESPRIT, and SAGE, generate an ambiguous error in the estimated direction-of-arrival results, when the antenna element spacing is more than half the carrier wavelength. Here, the source of this ambiguity is identified, and our proposition confirmed by applying the algorithms on the real measured data in a controlled environment. Furthermore, the relative sensitivity of each candidate algorithm is appraised.

I. INTRODUCTION

The performance of future wireless communication systems can be significantly improved by applying antenna arrays. Channel information like the direction-of-arrival (DoA) of the signals is crucial to those techniques that exploit the channel spatial information, e.g. beamforming methods. Following the undesirable consequences of poor resolution of the classical beamforming techniques (e.g. Bartlett's method, Min-Variance) in estimating the DoA, a number of efficient high resolution algorithms have been developed to enhance the resolution of the estimated parameters beyond the Fourier limit.

Some of the most commonly used high resolution algorithms are MUltiple SIgnal Classification (MUSIC) [1], Min-Norm [2], JoDeG[1] [3], Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) [4], and Space-Alternating Generalised Maximisation-expectation (SAGE) [5]. Disregarding the resolution capability of these algorithms, they generate an ambiguity error when applied to estimate the DoA of radio signals impinging on a uniform linear array (ULA), when the element spacing is more than half wavelength. This ambiguity error leads to unexpected wrong estimation of the DoA and has direct impact on the antenna array application that relies on the DoA information.
In this paper, the source of this ambiguity in all the aforementioned algorithms is identified and our proposition confirmed by applying these algorithms to estimate the DoA of the signals from the real measurement data in a controlled environment (an anechoic chamber). The ULA used in the test had an element spacing (d) of 0.563*λ, where λ is the wavelength of the radiated signal at 2.12 GHz. The results show that the ambiguous error will only affect the estimated DoA values, as these are evaluated in the spatial domain. Other signal parameters, such as Time-Delay-of-Arrival (TDoA) and path weights, are unaffected since they are not a spatial domain estimation problem and thus not influenced by the antenna element separation.

Throughout the paper, it is assumed that the readers have a fundamental knowledge of the herein algorithms under investigation. The paper is organised as follows. Section 2 explains the ambiguity in the classical beamformer methods, Section 3 explains the ambiguity in MUSIC, Min-Norm and the JoDeG algorithms. Section 4 and 5 explain the ambiguity in ESPRIT and SAGE respectively. Section 6 describes the setup procedure for the test measurement and presents the estimation results. Finally, Section 7 concludes the paper.

II. AMBIGUITY IN THE CLASSICAL BEAMFORMER METHODS

The classical beamformer methods (e.g. Bartlett's method, Min-Variance) used in the past are among the simplest in the direction finding algorithms. They are based on the beamforming, where a beam is formed at different directions across the azimuth and the DoA is located by the peak in the spatial power spectrum. Since this is a beamforming application, its performance is degraded by the grating lobes that occur when $d > \lambda/2$ [6].

For the ULA, grating lobes will occur when equation (1) is satisfied

$$\pi \frac{d}{\lambda}\left[\sin\theta_g - \sin\theta_o\right] = \pm n\pi \qquad (1)$$

where θ_g is the azimuth angular position of the grating lobe, θ_o is the azimuth angular position of the main lobe, and n is a non-zero integer. For d = 0.563λ, grating lobes will occur when $|\theta_o|$>50º (eq. 1). Figure 1 illustrates the occurrence of the

[1] The name 'JoDeG' was derived from the names of the authors [3] in order to identify the method proposed by them.

undesirable grating lobe at -73º when a beam is formed at 55º. However, note that the array does not suffer from grating lobe when a beam is formed at 20º.

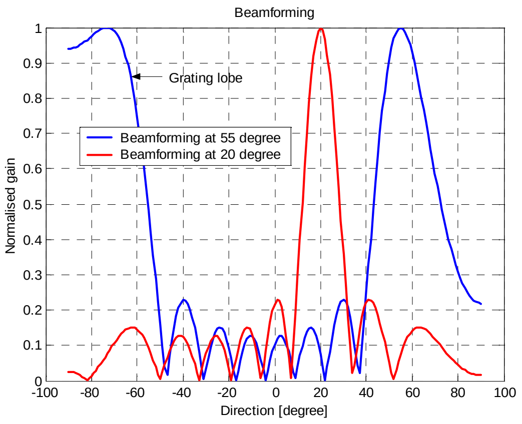

Figure 1 Beamforming with 8-element ULA, d=0.563λ

Hence, the false peak (corresponding to the grating lobe) as well as the true peak (corresponding to the true DoA) will both occur in the power spectrum, implying that there are two different sources in the far field. The wrong DoA estimate will be obtained if the grating lobe has got strong and distinct peak as the main lobe. This ambiguous problem is unavoidable since in practice the number of impinging signals and their true DoAs are unknown.

III. AMBIGUITY IN MUSIC, MIN-NORM AND JODEG

The development and implementation of the MUSIC, Min-Norm and JoDeG algorithms on an ULA are well documented in [1-3]. Although the implementation of each of these algorithms is different, their fundamental philosophy is the same. They identify the signal parameters by exploiting the property of the noise subspace of the covariance matrix through eigen-decomposition. Generally, these algorithms search the array manifold for the steering vector (corresponding to the true DoA) that is closest to the signal subspace, or orthogonal to the noise subspace.

Assuming only one signal is impinging on a ULA with an element spacing greater than half wavelength ($d > \lambda/2$), ambiguity is said to arise when there are two steering vectors that are (almost) orthogonal to the noise subspace. This is similar to the occurrence of grating lobes in a beamforming application discussed in Section 2 when $d > \lambda/2$.

This has a direct impact on the performance of MUSIC, Min-Norm, and JoDeG algorithms. For example, if the true DoA of the signal impinged on a ULA with 0.563λ element spacing is 55º, there will be two steering vectors that are very close to the signal subspace. One of them corresponds to the true DoA at 55º, while the other one corresponds to a false DoA at -73º (from eq. 1), where the peak of the grating lobe occurs. In this case, there will be two peaks in the spectrums of MUSIC, Min-Norm and JoDeG algorithms (at 55º and -73º respectively), and the search function[2] of these algorithms will provide two different DoA estimates for the same impinging signal. This leads to misleading results, since there is only one valid DoA and this is unknown to the user.

IV. AMBIGUITY IN ESPRIT

The main objective of the ESPRIT-type[3] algorithms is to compute the so-called '*invariance phase shift*' (Φ) corresponding to the DoA [4], given by

$$\Phi = 2\pi \frac{d}{\lambda} \sin\theta \qquad (2)$$

where θ is the DoA measured from the ULA broadside. Assuming that only one signal impinges on the ULA and the effective angular range of the ULA is restricted to ±90º (measured from the array broadside), in order to avoid mirror image ambiguity[4]. For any incident DoA (θ) and element spacing (d), Φ will have a value $-180^\circ \le \Phi \le 180^\circ$. To avoid any cyclically ambiguity in Φ, Φ should be restricted to $0^\circ \le \Phi \le 180^\circ$ for $0^\circ \le \theta \le 90^\circ$, and $-180^\circ \le \Phi \le 0^\circ$ for $-90^\circ \le \theta \le 0^\circ$. This ambiguity-free condition sets an upper limit of d at $\lambda/2$ (spatial Nyquist sampling theorem).

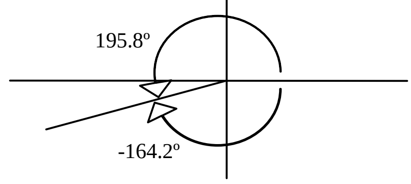

Figure 2 Illustration of cyclically wrapped around of Φ 195.8º = -164.2º

Now suppose that d=0.563λ (>λ/2) and θ=75º. Substituting these values into (2) yields Φ=195.8º. In this case, since $-180^\circ \le \Phi \le 180^\circ$, a phase shift of 195.8º is cyclically equivalent to -164.2º (Figure 2). ESPRIT will misinterpret this to be -164.2º and compute the estimated DoA at -54º (from eq. 2), which is very different from the true DoA at 75º. In short, for d=0.563λ, ambiguous error due to misinterpretation of cyclically ambiguous phase shift in ESPRIT will occur if |θ|>62.6º (eq. 2). Also, the DoA ambiguity reported is different to that of the previous methods.

2 To search for the peaks in the spectrum.

3 This includes the whole family of algorithms based on the '*invariance phase shift*' property of ESPRIT, eg. LS ESPRIT, TLS ESPRIT, SLS ESPRIT, Multiple Invariance ESPRIT, Unitary ESPRIT, etc.

4 Angular range is usually restricted to less than ±90° depending on the beamwidth of the individual elements.

V. AMBIGUITY IN SAGE

The underlying concept of the SAGE algorithm [5] is the same as the classical Expectation-Maximisation (EM) algorithm, except that the SAGE algorithm updates its multi-dimensional (m-D) parameters sequentially in different 1-D parameter space domains. This parameter updating procedure (M-step) relies on the maximum likelihood of occurrence of the estimated parameter value that is calculated from a cost function.

In terms of the DoA estimation using an *N*-element ULA, the M-step in SAGE is given by [5]

$$\theta' = \arg\max_{\theta}\left\{\left|z\left(\theta;\hat{X}_l\right)\right|^2\right\} \qquad (3)$$

where

$$z\left(\theta;\hat{X}_l\right)=\sum_{n=1}^{N} e^{j2\pi(n-1)d/\lambda\sin(\theta)}\cdot\hat{X}_l \qquad (4)$$

is the spatial correlation function, and $\hat{X}_l$ is the *complete data* of the *l*-path obtained in the *Expectation* step (E-step).

Note that effectively (4) represents a 1-D beamforming process where the estimated DoA (θ') of the *l*-path is given by the value of θ (within the θ space of eq. 3) which maximises the energy of the correlation function. This in turn is sensitive to the grating lobes problem described in Section 2 when $d > \lambda/2$, which creates a high possibility that the false θ_g (corresponding to the grating lobe) will maximise the energy of the correlation function (eq. 4), rather than a true DoA. Indeed, the ambiguous error in SAGE due to $d > \lambda/2$ is said to arise when the correlation in (4) with θ_g in the M-step is higher than that of the true θ_o.

VI. MEASUREMENT AND ESTIMATED RESULTS

In order to confirm the above observations, a controlled measurement has been conducted in the University of Bristol's anechoic chamber with a single source impinging on the 8-element ULA. The measurement was performed in a 20 MHz bandwidth centred at 2.12 GHz, with d=0.563λ (ULA element separation of 80 mm). The channel complex frequency response data was recorded using the Medav RUSK BRI channel sounder [7]. Both the transmitter and the ULA were fixed at one position in such a way that the ULA broadside directly faced the transmitter at the same height. Measurements were taken for each of the ULA's orientation from -90º to 90º in 1º steps (angle measured from ULA's broadside). An electrical positioner was used to rotate the ULA in every 1º step to guarantee absolute accuracy. The additional undesired responses of the ULA caused by its manufacturing imperfections and antenna mutual coupling had been calibrated out from the measured data before applying the post-processing described below.

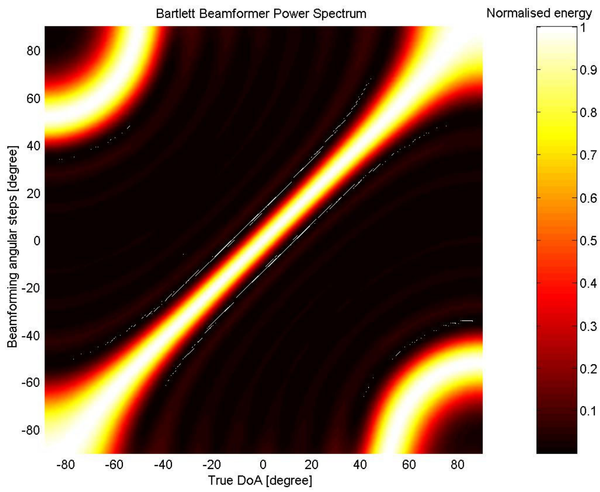

Figure 3 Power spectrum of the Bartlett's beamforming method

Figure 3 shows the power spectrum of the Bartlett's beamforming method when applied to the measured data. As discussed in Section 2, since d=0.563λ, grating lobes start to occur when $|\theta_o|$>50º. It can be clearly seen that two peaks of (almost) equal energy appear in the region of $|\theta_o|$>50º.

The channel data was recorded as a frequency response over 20 MHz of bandwidth. Thus the 2-D high resolution algorithms were applied to jointly estimate the DoA and TDoA of the signals together. In each of the following 2-D high resolution algorithms that based on eigen-decomposition (except 2-D SAGE), a 2-D forward-backward spatial smoothing [8] process had been applied prior to implementing the algorithms[5].

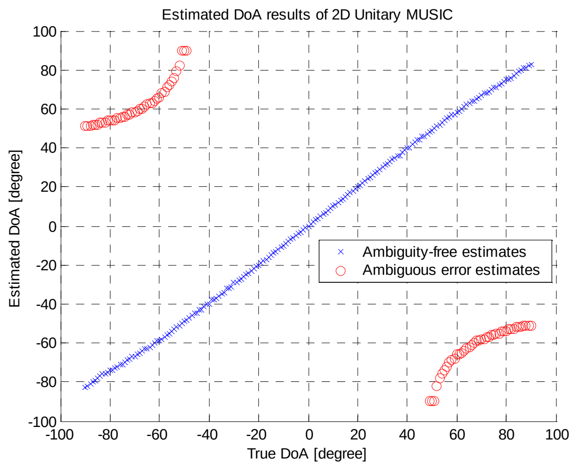

Figure 4 Estimated DoA results of 2-D Unitary MUSIC algorithm

[5] The reason of doing this is to follow the standard procedure in the implementation of eigen-decomposition based algorithms so that a fair judgement can be obtained, although there is no need for applying it in this case.

Figures 4-6 show the estimated DoA results using the 2-D Unitary MUSIC, 2-D Unitary JoDeG, and 2-D Min-Norm algorithms respectively. As discussed in Section 3, since these algorithms produce two distinct peaks in their respective spectrums[6] that correspond to the true DoA and the false DoA when $|\theta_o|$>50º, two different estimates are obtained in $|\theta_o|$>50º. The false estimates are indicated in the circles since the true DoA is known in an anechoic chamber. In practice, since the true DoA is unknown, one might think that there are two different sources since there are two distinct peaks in these spectrums when $d > \lambda/2$. However, the algorithms are able to produce (almost) correct estimates when $|\theta_o|\leq 50º$ since this is the error-free region. Note that the ambiguity-free estimates (indicated in blue crosses) in the region of $|\theta_o|$>80º have an error within ±5º. This is caused by low signal-to-noise ratio (SNR) in the measured data since the antennas of the ULA have directional beam pattern.

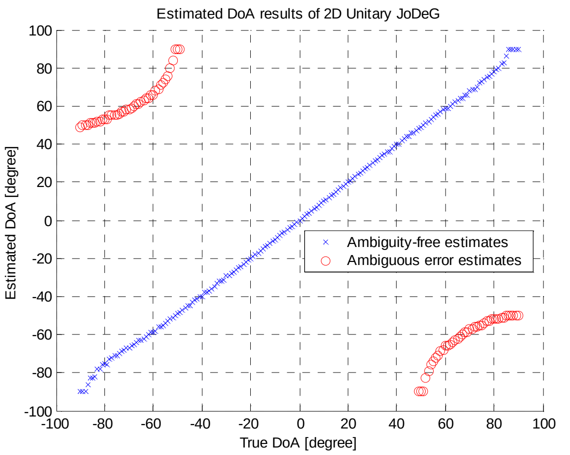

Figure 5 Estimated DoA results of 2-D Unitary JoDeG algorithm

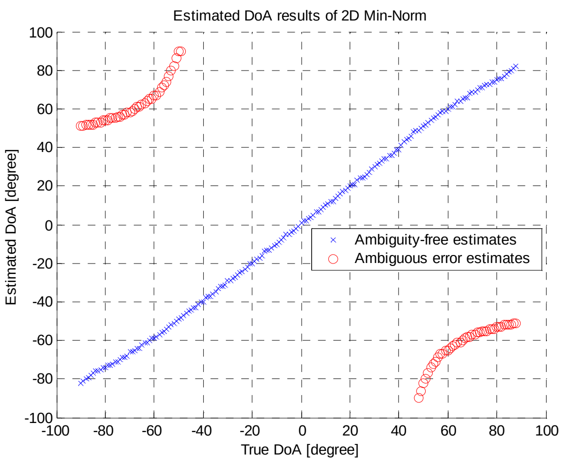

Figure 6 Estimated DoA results of 2-D Min-Norm algorithm

[6] The spectrums are not displayed in this paper for brevity. Here we estimated the DoAs and TDoAs based on the spatial and temporal location of the peaks in the spectrums.

On the other hand, 2-D Unitary ESPRIT [9] algorithm provides consistent wrong estimates when $|\theta_o|$>62.6º, since the estimated result given by ESPRIT is strictly determined by (2). Results show that ESPRIT has a larger error-free angular range compared to the rest of the algorithms. This is because ESPRIT is not affected by the effect of the grating lobes (strictly speaking), but it is due to the misinterpretation of the cyclically ambiguous phase shift when $d > \lambda/2$ (violation of spatial Nyquist sampling theorem).

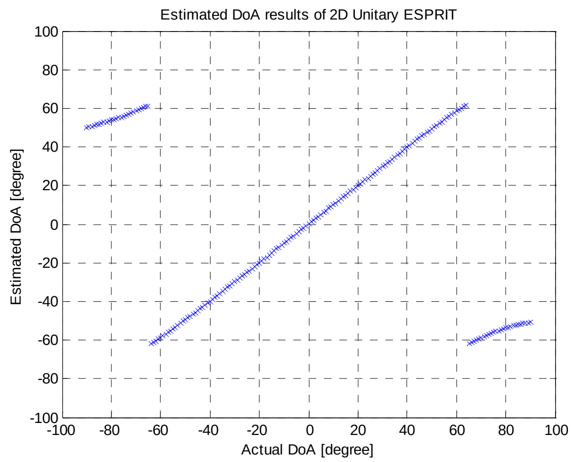

Figure 7 Estimated DoA results of 2-D Unitary ESPRIT algorithm

The estimated results of the 2-D SAGE algorithm are shown in Figure 8. SAGE is able to estimate correct DoA results when $|\theta_o|\leq 50º$, but the probability of estimating the correct results and the wrong results when $|\theta_o|$>50º is equal. As discussed in Section 5, the M-step of the SAGE algorithm is based on the spatial correlation given in (4). If the steering vector corresponding to the false DoA (due to the grating lobe) has a higher correlation with the *complete signal* of the *l*-path than the steering vector corresponding to the true DoA, then SAGE will produce a wrong estimate, and vice-versa. Since this is a grating lobes related problem, the probabilities of both events are equal.

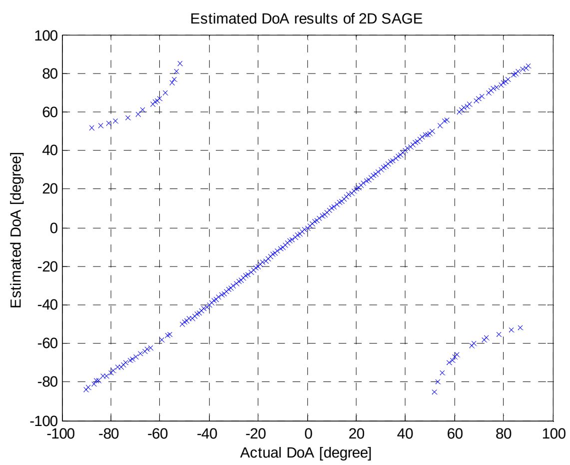

Figure 8 Estimated DoA results of 2-D SAGE algorithm

Since the M-step of the SAGE algorithm is based on the correlation property, if we could increase the correlation corresponding to the true DoA, then the chances of obtaining the correct estimate will be higher. Note that the result in Figure 8 is obtained by correlating the data with the ideal array manifold of the ULA. When $|\theta_o|>50°$, correlation with the ideal manifold provides equal probability of obtaining the correct and the false estimates. The probability of obtaining the correct estimate can be increased by correlating the data with the real measured array manifold in the M-step. This is demonstrated in Figure 9 in which the result is obtained by using the proposed correlation method in the 2-D SAGE algorithm. By doing this, the SAGE algorithm is able to produce accurate estimate up to $|\theta_o|>75°$ even in the occurrence of the grating lobes when $|\theta_o|>50°$. Furthermore, this also implies that the spatial calibration procedure [10] is not needed in implementing the SAGE algorithm, provided that an exact knowledge of the array manifold is available. However, in doing this, more memory would be required to store the array manifold data and the angular resolution of the SAGE algorithm in this case depends on the angular resolution of the array manifold (i.e. the step size of the spatial sampling grid of the measured array manifold).

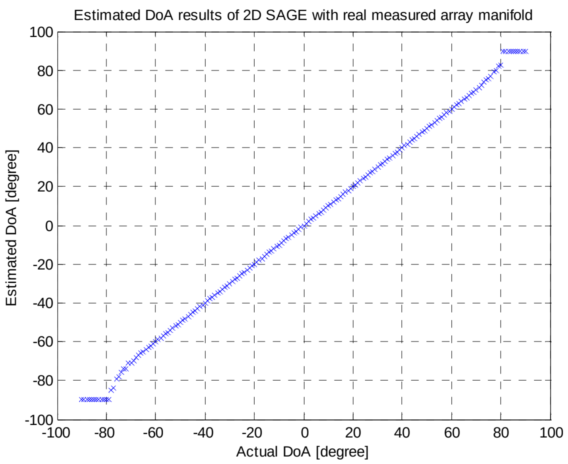

Figure 9 Estimated DoA results of 2-D SAGE algorithm with real measured array manifold

Note that using the reference data of the array manifold (measured in an anechoic chamber) in the correlation process in the M-step is merely a suggested alternative method of implementation of the SAGE algorithm. This does not constitute a solution to the ambiguity problem in the SAGE algorithm when the element spacing is more than half wavelength. The result shown in Figure 9 demonstrates a huge improvement since the measurement in the anechoic chamber has a high SNR and only a single source is present. In the real environment, we might still obtain the wrong estimate since more than one path is present. But the probability of estimating the wrong results could be reduced by using the reference data with a small spatial sampling grid in the correlation process.

Although all the aforementioned algorithms produce the ambiguous error when $d > \lambda/2$, they are still able to produce accurate TDoA estimates and path gain values[7] since these parameters are not affected by the antenna element separation (assuming both the transmitter and receiver are properly synchronised [7]). All the TDoA estimates produced by these 2-D algorithms are very similar to each other. The variance of the TDoA estimates is less than 5 ns (the resolution of the channel sounder is 50 ns in 20 MHz bandwidth) and the path gain values are very consistent (almost equal to the element beam pattern) throughout the azimuth range. Hence, as expected, it can be verified that the antenna element spacing in the ULA will only affect the applications in the spatial domain (e.g. DoA estimation, beamforming), but has virtually no effect on other domains (e.g. temporal and frequency domains).

VII. CONCLUSION

It is shown that the DoA estimation algorithms (beamformer methods, MUSIC, Min-Norm, JoDeG, ESPRIT, and SAGE) rely on certain array geometry to generate the best performance. A ULA with element spacing of more than half a wavelength will produce ambiguity problems to the direction finding algorithms. As a short-term solution, a ULA with highly directive elements to restrict the effective visible range of the array into the error-free zone can solve this problem. However, this effectively reduces the visible range of the ULA into a smaller region, which in fact is not practically feasible since the ULA will reject all useful information from all other directions. A long-term solution, on the other hand, would be to use a ULA of no more than half wavelength spacing, with the expense of increased correlation between elements and increased difficulties in putting the elements and feedlines this close together. This work provides useful information to array designers if DoA estimation based on the aforementioned algorithms is to be implemented.

ACKNOWLEDGEMENT

The authors thank Allgon Systems AB for the loan of their 2GHz UMTS-FDD panel array as well as the Mobile VCE (www.mobilevce.com) for the financial support of Chor Min Tan.

[7] The estimated TDoA values and the path gain values are not displayed here for brevity.

REFERENCES

[1] R. O. Schmidt, '*Multiple emitter location and signal parameters estimation*', IEEE Trans. Antenna and Propagation, Mar 86, pp. 276-280.

[2] R. Kumaresan, D. W. Tufts, '*Estimating the angles of arrival of multiple plane waves*', IEEE Trans. Aero. Elect. Sys., Vol. 19 No. 1, Jan 83, pp. 134-139.

[3] D. H. Johnson, S. R. DeGraaf, '*Improving the resolution of bearing in passive sonar arrays by eigenvalue analysis*', IEEE Trans. ASSP, Vol. 30 No. 4, Aug 82, pp. 638-647.

[4] R. Roy, T. Kailath, '*ESPRIT – Estimation of Signal Parameters via Rotational Invariance Techniques*', IEEE Trans. ASSP, Vol. 37 No. 7, July 89, pp. 984-995.

[5] C. C. Chong, D. I. Laurenson, C. M. Tan, S. McLaughlin, M. A. Beach, A. R. Nix, '*Joint detection estimation of directional channel parameters using the 2-D frequency domain SAGE algorithm with Serial Interference Cancellation*', conference proceeding in ICC 2002, New York City, Apr 28- May 2 2002.

[6] M. I. Skolnik, '*Introduction to radar systems*', 2nd edition, McGraw Hill 1980, Chapter 8 pp. 283.

[7] R. S. Thomä, D. Hampicke, A. Richter, G. Sommerkorn, A. Schneider, U. Trautwein, and W. Wirnitzer, '*Identification of time-variant directional mobile radio channels*', IEEE Trans. Instrum. and Meas., Vol. 49, Apr 2000, pp. 357–364.

[8] S. U. Pillai, B. H. Kwon, '*Forward/backward spatial smoothing techniques for coherent signal identification*', IEEE Trans. ASSP, Vol. 37 No. 1, Jan 89, pp8-15.

[9] M. Haardt, '*Efficient one-, two-, and multidimensional high-resolution array signal processing*', PhD. Thesis, ISBN 3-8265-2220-6, 1996.

[10] G. Sommerkorn, D. Hampicke, R. Klukas, A. Richter, A. Schneider, R. Thomä, '*Reduction of DoA estimation errors caused by antenna array imperfections*', 29th Euro. Microwave Conf., Munich, Oct 5-7, 1999.