Cluster Parameter Extraction Using the Space-Alternating Generalized Expectation-Maximization Algorithm

Dan J. Dechene, *Student Member, IEEE*
Department of Electrical and Computer Engineering
University of Western Ontario
London, Ontario, N6A 5B9, Canada
Email: ddechene@uwo.ca

***Abstract*—Cluster parameter extraction is an important step in successful channel modeling. In this paper, we provide an overview of the Space-Alternating Generalized Expectation-Maximization (SAGE) algorithm. This algorithm is used to extract joint cluster parameters with relatively low computational complexity. We discuss the details of SAGE in the application of cluster parameter estimation in multipath propagation environments. Results are provided for the application of the SAGE algorithm to synthetic data.**

***Index Terms*—Channel estimation, ECE9931, channel modelling, multipath propogation, optimization**

I. Introduction

In multipath propagation environments, signals often travel through a variety of paths to reach the destination. Individual paths taken by a signal can be characterized through a number of parameters. A useful and necessary step to understanding the multiple paths taken by a signal relies on estimation of these characterization parameters. Consequently four major parameters that characterize a signal are the incidence azimuth, doppler frequency, relative delay and complex gain.

Resolution of all possible paths taken by a signal is non-trivial. As such, the cluster model of wireless channels has became a widely acceptable method of characterizing the wireless channel. This involves grouping of closely related paths into a single path as resolving of closely related paths is very difficult. Using this paradigm, the characterization of a wireless channel can be done by resolving a finite set of propagation path parameters.

Moreover, research into methods of parameter extraction has become increasingly important in order to effective model realistic wireless channels. Several recent methods include the multiple signal classification (MUSIC) algorithm [1], [2] and the estimation of signal parameter via rotational invariance techniques (ESPRIT) algorithm [3], [4], [5]. MUSIC has been used for relative delay and azimuth estimation while ESPRIT has been used to estimate the jointly estimate delay and azimuth. More modern techniques for joint parameter estimation include the expectation-maximization (EM) algorithm [6] and its extension the space-alternating generalized expectation-maximization (SAGE) algorithm [7].

In this work we examine the application of SAGE in estimation of joint parameters including delay, incidence azimuth, doppler frequency and complex amplitude as in [8].

The remainder of this paper is organized as follows. In Section II we will discuss the details of the SAGE algorithm and its application to cluster parameter extraction problem. In Section III we will provide some simulation results of the SAGE algorithm and in Section IV we will draw some conclusions on this work.

II. SAGE Algorithm

The Space-Alternating Generalized Expectation-Maximization (SAGE) algorithm [7] is a method of solving maximum-likelihood estimation (MLE) problems to estimate hidden datasets, where observations of the data may not be complete. This is similar to the Expectation-Maximization (EM) algorithm [6]. SAGE differs from the EM algorithm in that estimation of the hidden dataset can be reduced to estimation of multiple subsets of hidden data. This can result in a complexity reduction from $O(N^x)$ to $O(xN)$ where x is the original problem complexity.

A. EM Algorithm

As in [8], we will describe the operation of the EM [6] algorithm and extend these concepts to describe the operation of SAGE. The operation of the EM algorithm is divided into two phases: the expectation step and the maximization step. The maximization step estimates the hidden dataset by performing MLE. This estimation however requires knowledge of the complete data which is not always known. The expectation step provides the maximization step with an estimate of the complete data using the incomplete data in conjunction with the previous estimate of the hidden dataset.

To explain this in further detail, we examine the application of the EM algorithm to a set of L superimposed signals received by an linear antenna array with M elements. Without loss of generality we define

$$\mathbf{s}(t;\theta_\ell) \triangleq \mathbf{c}(\phi_\ell)\alpha_\ell \exp\{j2\pi v_{D\ell} t\} u(t-\tau_\ell) \tag{1}$$

where $u(t-\tau_\ell)$ can be any periodically repeated signal with period T_a, $\mathbf{s}(t;\theta_\ell)$ is an individual signal and can also be referred to as the complete data of signal ℓ, θ_ℓ is the hidden dataset of this individual signal where $\theta_\ell \triangleq [\tau_\ell, \phi_\ell, v_{D\ell}, \alpha_\ell]$. The characterization of the signal is given by the doppler frequency $v_{D\ell}$, the incidence azimuth ϕ_ℓ, the relative delay τ_ℓ and the complex amplitude α_ℓ. The steering vector is given as

$$\mathbf{c}(\phi_\ell) \triangleq exp\{-j2\pi[0,\ldots,M-1]^T d\lambda^{-1}\sin(\phi_\ell)\} \quad (2)$$

where d is the spacing between two antennas and λ is the wavelength of the signal. The observed superimposed signal with additive noise will then be given by

$$\mathbf{Y}(t) = \sum_{\ell=1}^{L} \mathbf{s}(t;\theta_\ell) + \sqrt{\frac{N_0}{2}}\mathbf{N}(t) \quad (3)$$

where N_0 is the noise power spectral density, $\mathbf{N}(t)$ is the noise process and $\mathbf{Y}(t)$ is the observed signal which we will denote as the incomplete observable data as measurements of $\mathbf{Y}(t)$ do not directly yield $\mathbf{S}(t)$, where $\mathbf{S}(t) \triangleq [\mathbf{s}(t;\theta_1), \mathbf{s}(t;\theta_2), \ldots, \mathbf{s}(t;\theta_L)]$.

The problem statement for the EM (and SAGE) algorithm is given as follows. Estimate the full hidden dataset (θ, where $\theta \triangleq [\theta_1,\ldots,\theta_L]$) given the observable data $\mathbf{Y}(t)$ and knowledge of the input signal shaping ($s(t;\theta_i)$) for a known set θ_i.

1) Expectation Step: In order to perform MLE of the hidden parameter set θ, we must first find an estimate the complete data. We denote the complete data as

$$\mathbf{X}_\ell(t) \triangleq \mathbf{s}(t;\theta_\ell) + \sqrt{\frac{\beta_\ell N_0}{2}}\mathbf{N}_\ell(t), \quad \ell = 1,\ldots,L \quad (4)$$

where $\sum_{\ell=1}^{L}\beta_\ell = 1$ and each $\mathbf{X}_\ell$ is independent. From this it is clear that the incomplete, observable data $\mathbf{Y}(t)$ is related to the complete data by

$$\mathbf{Y}(t) = \sum_{\ell=1}^{L}\mathbf{X}_\ell(t) \quad (5)$$

We also denote the estimate of the complete data as $\widehat{\mathbf{x}}_\ell(t;\widehat{\theta}')$ where $\widehat{\theta}'$ is the previously gathered estimate of the hidden dataset. The natural estimate of the complete data is given as in [8]

$$\begin{aligned}\widehat{\mathbf{x}}_\ell(t;\widehat{\theta}') &\triangleq \mathbf{E}_{\widehat{\theta}'}[\mathbf{X}_\ell|\mathbf{Y}(t)], \quad \ell=1,\ldots,L \quad (6)\\ &= \mathbf{s}(t;\widehat{\theta}_\ell{}') + \beta_\ell \underbrace{\left[\mathbf{Y}(t) - \sum_{\ell'=1}^{L}\mathbf{s}(t;\widehat{\theta}_\ell{}')\right]}_{\text{Noise Estimate} \quad \sqrt{\frac{N_0}{2}}\mathbf{N}(t)} \quad (7)\end{aligned}$$

Using the estimate of the complete data, estimation can be done on the hidden dataset during the maximization step.

2) Maximization Step: In the maximization step, an estimate of the hidden dataset $\widehat{\theta}_\ell$ for $\ell = 1,\ldots,L$ is computed by maximizing the log-likelihood function given in [9]

$$\begin{aligned}\Lambda(\theta_\ell;\mathbf{x}_\ell) \triangleq \frac{1}{\beta_\ell N_0}\Bigg[&2\int_D \Re\{\mathbf{s}^\dagger(t';\theta_\ell)\mathbf{x}_\ell(t')\}dt' \\ &- \int_D \|\mathbf{s}(t';\theta_\ell)\|^2 dt'\Bigg] \quad (8)\end{aligned}$$

where D interval of observation and the superscript † denotes the complex conjugate transpose operation.

Denoting the MLE estimate of the parameter set as $(\widehat{\theta}_\ell)_{ML}(\mathbf{x}_\ell)$, we can solve iteratively for the parameter set using our estimation of the complete data $\widehat{\mathbf{x}}_\ell(t;\widehat{\theta}')$.

Using the example of superimposed waves, the MLE of the individual parameters of $\widehat{\theta}_\ell$ become

$$\left(\widehat{\tau}_\ell, \widehat{\phi}_\ell, \widehat{v}_{D\ell}\right)_{ML}(\mathbf{x}_\ell) = \arg\max_{[\tau,\phi,v]}\{|z(\tau,\phi,v;\mathbf{x}_\ell)|\} \quad (9)$$

$$\begin{aligned}(\widehat{\alpha}_\ell)_{ML}(\mathbf{x}_\ell) &= \frac{1}{I\|\mathbf{c}((\widehat{\phi}_\ell)_{ML}(\mathbf{x}_\ell))\|^2 T_a P_u} \\ &\cdot z\left(\left(\widehat{\tau}_\ell, \widehat{\phi}_\ell, \widehat{v}_{D\ell}\right)_{ML}(\mathbf{x}_\ell);\mathbf{x}_\ell\right) \quad (10)\end{aligned}$$

where I is the length of the observation interval D, P_u is the power in $u(t)$ and the function $z(\tau,\phi,v;\mathbf{x}_\ell)$ denotes the cost function given by

$$\begin{aligned}z(\tau,\phi,v;\mathbf{x}_\ell) \triangleq &\int_D u^*(t'-\tau)\exp\{-j2\pi v t'\} \\ &\cdot \mathbf{c}^\dagger(\phi)\mathbf{x}_\ell(t')dt' \quad (11)\end{aligned}$$

where the superscript $*$ denotes the complex conjugate operation. From the above equations we can observe two phenomenon. The first is that given an estimate of the hidden subset $(\widehat{\tau}_\ell, \widehat{\phi}_\ell, \widehat{v}_{D\ell})$, we can obtain a closed form expression of the estimate of $\widehat{\alpha_\ell}$. The second thing to note is that the MLE of the parameter set must be computed numerically. This problem becomes L 3D optimization problems. Achieving fine granularity of accuracy on the set of parameters invokes a large search space for the EM algorithm. To estimate the entire hidden dataset $\widehat{\theta}$, we vary the index ℓ from 1 to L for each iteration of the optimization process successively. This process continues until convergence is reached or a maximum number of iterations has occurred.

B. SAGE Extension of EM Algorithm

The benefits of SAGE involve solving for a subset of the hidden dataset during each iteration of the optimization process. In this way, the L 3D optimization problems are converted into $3L$ sets of 1D optimization problems. This is accomplished by updating each parameter in sequence for each iteration step ℓ. The estimation of each parameter becomes:

$$\begin{aligned}
\{1\} \quad \widehat{\tau}''_\ell &= \arg\max_{\tau} \left\{ \left| z\left(\tau, \widehat{\phi}'_\ell, \widehat{v}'_{D\ell}; \widehat{\mathbf{x}}_\ell\left(t; \widehat{\theta}'\right)\right) \right| \right\} \\
\{2\} \quad \widehat{\phi}''_\ell &= \arg\max_{\phi} \left\{ \left| z\left(\widehat{\tau}''_\ell, \phi, \widehat{v}'_{D\ell}; \widehat{\mathbf{x}}_\ell\left(t; \widehat{\theta}'\right)\right) \right| \right\} \\
\{3\} \quad \widehat{v}''_{D\ell} &= \arg\max_{v} \left\{ \left| z\left(\widehat{\tau}''_\ell, \widehat{\phi}''_\ell, v; \widehat{\mathbf{x}}_\ell\left(t; \widehat{\theta}'\right)\right) \right| \right\} \\
\{4\} \quad \widehat{\alpha}''_\ell &= \frac{1}{I\|\mathbf{c}(\widehat{\phi}''_\ell)\|T_a P_u} z\left(\widehat{\tau}''_\ell, \widehat{\phi}''_\ell, \widehat{v}''_{D\ell}; \widehat{\mathbf{x}}_\ell\left(t; \widehat{\theta}'\right)\right)
\end{aligned}$$

where quantities with the superscript $'$ denotes values from the previous estimation iteration. Although there exists no close expression for steps $\{1\} - \{3\}$, these values can be solved by varying the parameters over the range of possible values.

III. Simulation Results

In order to observe the performance of the SAGE algorithm for synthetic data, it is implemented using MATLAB. The parameter set is shown in Table I unless otherwise listed. For an input function, $u(t) = \sum_{i-\infty}^{\infty} \delta(t - iT_a)$ is used. In this case, SAGE is used to jointly estimate incidence azimuth, doppler frequency and complex gain as the relative delay is assumed known. In addition, we quantize all parameters. Resolution and ranges for each parameter is given as follows

- $\widetilde{\tau} = 1/F_s, \tau \in [1/F_s, 30/F_s]$
- $\widetilde{\phi} = \pi/100, \phi \in [-\pi/2, \pi/2)$
- $\widetilde{v_D} = 0.1Hz, v_D \in [10, 15]$

There is no minimum resolution for α as it is derived in closed-form using the estimation of the above parameters. In order to ensure some separability of the waves, the minimum parameter spacings are also listed in Table I (under $\Delta\tau, \Delta\phi$ and Δv_D). Noise is a zero-mean complex-valued gaussian random variable with variance N_0 and the complex gain α is a zero-mean complex-valued variable with variance $1/L$.

In Figures 1-3, the estimation error versus the received SNR is shown. These results are ensemble averaged over 20 realizations and averaged over all L waves. We observe a

TABLE I: Simulation Parameters

Parameter	Value
$\text{SNR}_{(dB)}$	10dB
N_0	1/SNR
D	0s to 0.2s
M	16
L	8
Antenna Spacing	λ/2
Resolution (F_s)	5kHz
T_a	0.01s
$\Delta\tau$	2/F_s
$\Delta\phi$	$\pi/16$
Δv_D	$0.25Hz$

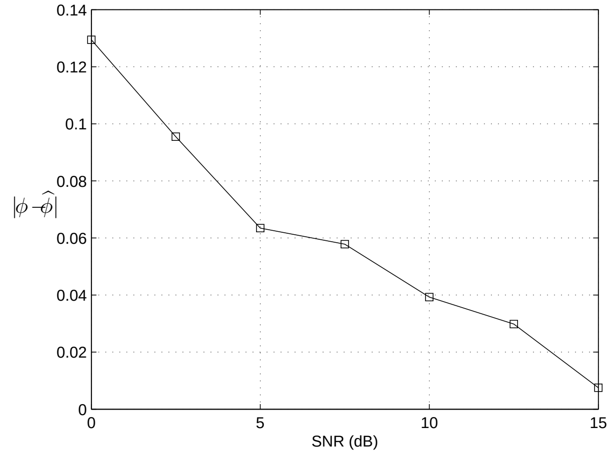

Fig. 1: Estimation Error of ϕ versus SNR, $L = 8, M = 16$

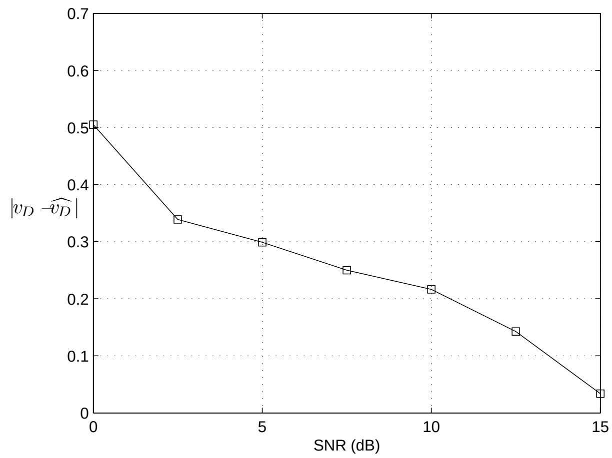

Fig. 2: Estimation Error of v_D versus SNR, $L = 8, M = 16$

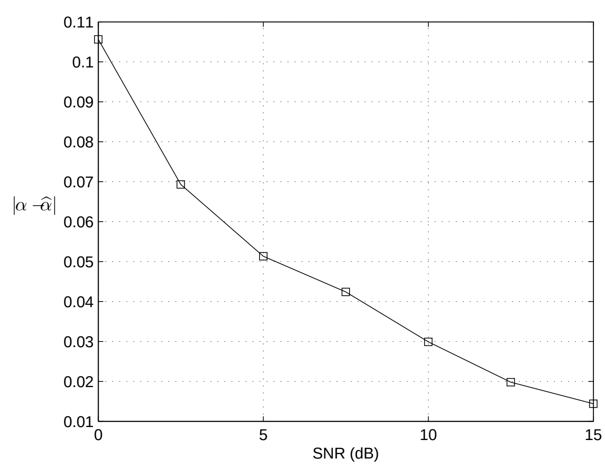

Fig. 3: Estimation Error of α versus SNR, $L = 8, M = 16$

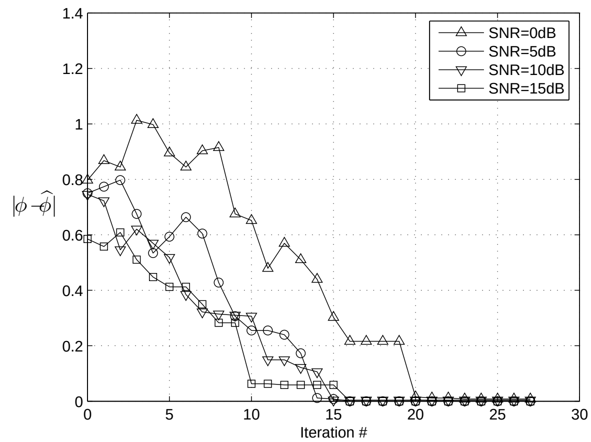

Fig. 4: Convergence of ϕ for varying SNR, $L = 8, M = 16$

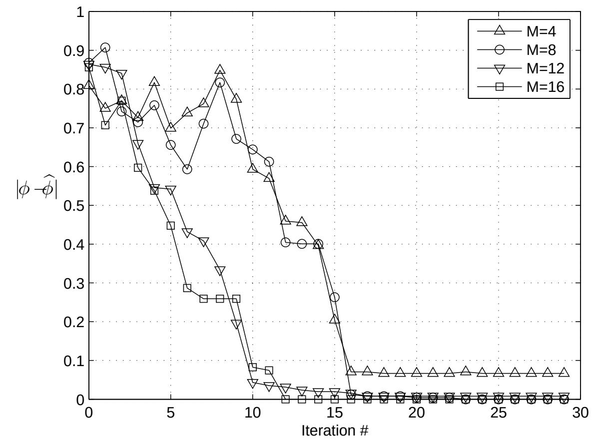

Fig. 7: Convergence of ϕ for varying M, $L = 8, SNR = 10dB$

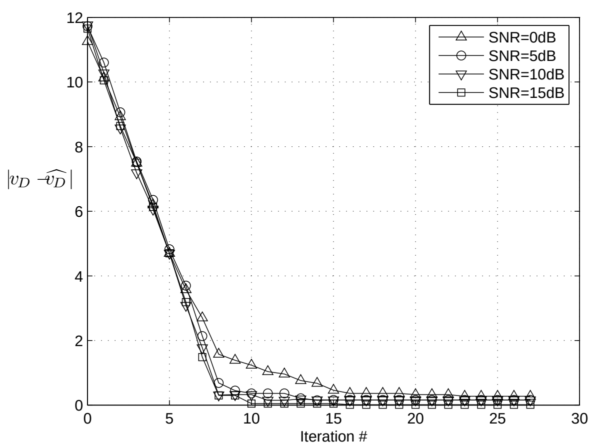

Fig. 5: Convergence of v_D for varying SNR, $L = 8, M = 16$

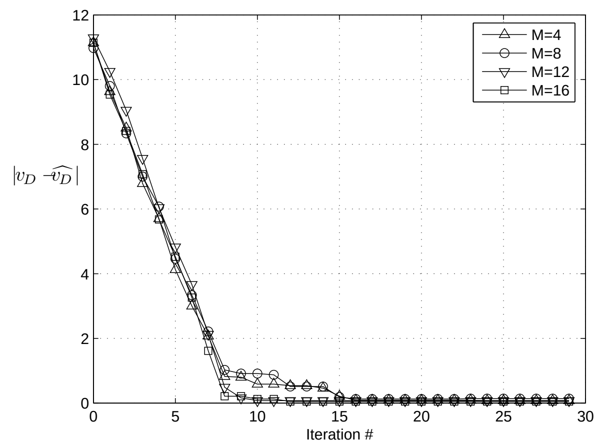

Fig. 8: Convergence of v_D for varying M, $L = 8, SNR = 10dB$

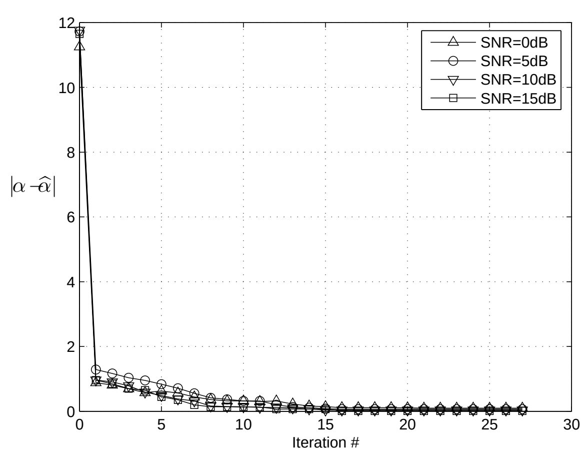

Fig. 6: Convergence of α for varying SNR, $L = 8, M = 16$

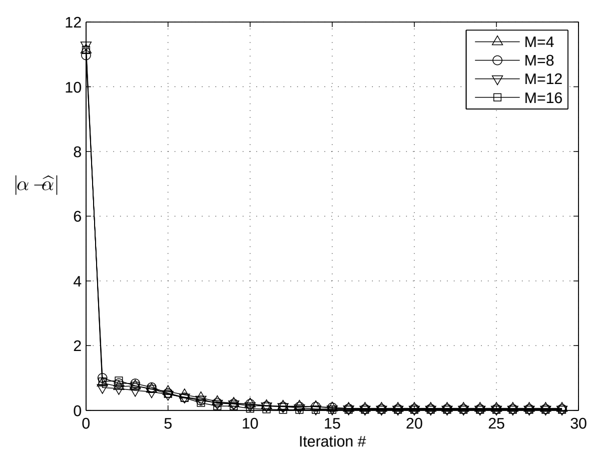

Fig. 9: Convergence of α for varying M, $L = 8, SNR = 10dB$

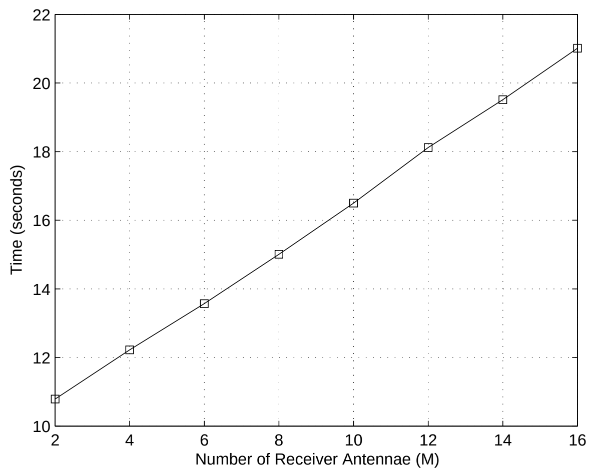

Fig. 10: Computational Time vs Number of Receiver Antennae

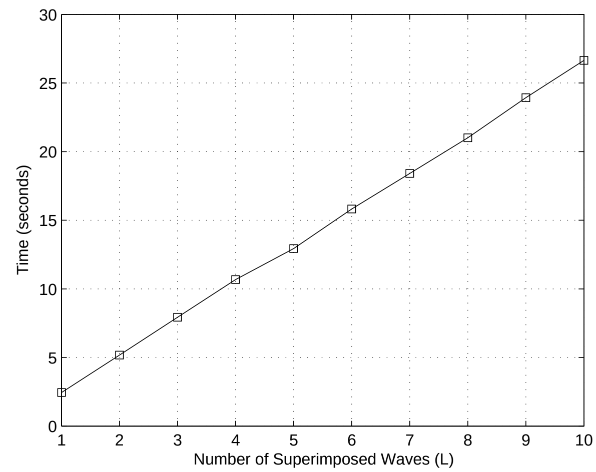

Fig. 11: Computational Time vs Number of Estimated Paths

consistent trend that an increase in SNR results in a decrease the estimation error. The relationship is closely linear with some exponential characteristics for the complex gain (α). The reason for the non-smooth relationship is due in part to a) the number of realizations required and b) the quantization of possible estimation values. Overall, these Figures show that larger SNR improves estimation accuracy.

Figures 4-6 and 7-9 show the convergence for a single set estimation realization (averaged over L waves) for both a varying SNR and a varying number of receive antennae (M) respectively. We observe that for lower number of antennas and for a lower SNR, the estimation of both ϕ and v_D take more iterations (and therefore more time) than those of higher SNR and more antennae. The convergence of the gain magnitude varies very little in time. Overall, we observe that higher SNR and larger number of antenna cannot impact the system negatively in terms of number of iterations for convergence and estimation error.

Figure 10 shows the effect of an increasing number of antennas on the time taken for each iteration. This time measured is the time taken to update the parameters for all waves once.

The Figure shows that the time taken is linear, however only scales with a fraction of the linear increase in M. For example, an increase from 2 to 16 antenna only doubles the estimation time. These increase is important to note as this must be doubly considered when choosing the number of antenna. An increase may improve accuracy, however increases convergence time.

Finally we show the time dependance on the number of estimated waves in Figure 11. This is an important metric as we originally discussed the ability of SAGE to reduce the computational complexity. This Figure confirms this result as the dependance on computational time versus number of estimated waves is linear.

IV. Conclusion

Cluster parameter extraction is a vital component of channel modelling. In this paper we have shown that use of the SAGE algorithm can not only accurately estimate cluster parameters, but also perform this estimation in a time which scales $O(L)$ with L superimposed cluster paths.

In future work we will utilize the SAGE algorithm to estimate parameters using real-world data and study its performance in this environment.

References

[1] R. O. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propagation*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[2] T. Lo, J. Litva, and H. Leung, “A new approach for estimating indoor radio propagation characteristics,” *IEEE Trans. Antennas Propagation*, vol. 42, no. 10, pp. 1369–1376, Oct. 1994.

[3] R. Roy and T. Kailaith, “Esprit - estimation of signal parameters via rotational invariance techniques,” *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 37, no. 7, pp. 984–994, July 1989.

[4] M. Haardt and J. Nossek, “Unitary esprit: How to obtain increased estimation accuracy with a reduced computational burden,” *IEEE. Trans. Signal Processing*, vol. 43, no. 5, pp. 1232–1242, May 1995.

[5] A. J. van der Veen, M. C. Vanderveen, and A. Paulraj, “Joint angle and delay estimation using shift-invariance techniques,” *IEEE Trans. Signal Processing*, vol. 46, no. 2, pp. 405–418, Feb. 1998.

[6] M. Feder and E. Weinstein, “Parameter estimation of superimposed signals using the em algorithm,” *IEEE Trans. Acoustic, Speech and Signal Processing*, vol. 36, no. 4, pp. 477–489, Apr. 1988.

[7] J. Fessler and A. Hero, “Space-alternating generalized expectation-maximization algorithm,” *IEEE Trans. Signal Process.*, vol. 42, no. 10, pp. 2664–2677, October 1994.

[8] B. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. Pedersen, “Channel parameter estimation in mobile radio environments using the SAGE algorithm,” *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 434–450, March 1999.

[9] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1988.