The 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'07)

AN MSE LOWER BOUND FOR PARAMETRIC AND NONPARAMETRIC CHANNEL ESTIMATION

D. Van Welden, F. Simoens, H. Steendam and M. Moeneclaey
TELIN Department, Ghent University
Sint-Pietersnieuwstraat 41, 9000 Gent, Belgium
E-mail: {dmvwelde,fsimoens,hs,mm}@telin.ugent.be

ABSTRACT

This paper compares parametric and nonparametric channel estimation methods in a multipath environment. For both methods, the MSE on the received symbol pulse is shown to consist of a modeling error term and an estimation error term. We lower bound the latter term by computing the associated Cramer-Rao bounds. We point out that the MSE can be minimized by a proper selection of the number of parameters to be estimated. In addition, we present simulation results from practical estimators that confirm our theoretical analysis. Our numerical results indicate that the parametric method outperforms the nonparametric approach.

I. INTRODUCTION

Multipath channels give rise to inter-symbol interference (ISI) at the receiver. The traditional approach to counter the effects of ISI is to equalize the received symbol pulse. In order to successfully mitigate the ISI, an accurate knowledge of the received signal pulse is paramount. Hence, channel estimation is an important task of the receiver in digital communication systems.

A *parametric* channel estimation method exploits the multipath structure by estimating the path gains and delays, and computing from these estimates the corresponding received symbol pulse. In a *nonparametric* channel estimation method the samples of the received symbol pulse are estimated without taking the multipath structure into account. Empirical results have indicated that the estimation accuracy improves by exploiting the underlying channel structure [1, 2, 3].

In this contribution, we investigate the mean squared error (MSE) on the received symbol pulse, resulting from the two estimation methods. We present a lower bound on the MSE, which is based on the Cramer-Rao lower bound (CRB), and provide simulation results that confirm our theoretical derivations. Our numerical results illustrate that the parametric estimation method yields the smaller MSE.

II. SYSTEM MODEL AND ESTIMATION STRATEGIES

A sequence of data symbols $\{a(k)\}$ is modulated on a band limited transmit pulse $p(t)$ with bandwidth B and transmitted over a channel with impulse response $h_{channel}(t)$. The received symbol pulse $h(t)$ is defined as the convolution of the transmit pulse and the channel impulse response

$$h(t) = \int_{-\infty}^{+\infty} p(t-\tau)\, h_{channel}(\tau)\, d\tau$$

In order to estimate $h(t)$, a sequence of K pilot symbols $\{a(k),\ k = 0,\ \ldots,\ K-1\}$ is transmitted over the channel. We define $r(t)$ as the received signal that corresponds to this data sequence

$$\begin{aligned} r(t) &= s(t) + w(t) \\ &= \sum_{k=0}^{K-1} a(k)\, h(t-kT) + w(t) \end{aligned} \qquad (1)$$

where $s(t)$ represents the useful part of the received signal and $w(t)$ is zero-mean complex-valued white Gaussian noise with spectral power density N_0. The next subsections are devoted to the estimation of the received pulse $h(t)$.

A. Parametric channel estimation

In the parametric approach, the receiver assumes that $h(t)$ corresponds to some parametric model $h_0(t;\ \mathbf{x})$ characterized by a set of real-valued parameters $\mathbf{x} = [x(1),\ \ldots, x(N_{par})]^T$, where $[.]^T$ denotes transposition. Here we consider the following multipath model :

$$h_0(t;\ \mathbf{x}) = \sum_{i=0}^{L-1} \alpha_i p(t-\tau_i) \qquad (2)$$

where L denotes the number of paths, $\{\alpha_i\}$ and $\{\tau_i\}$ denote the gains and the delays, and $\mathbf{x}$ contains the delays and the real and imaginary parts of the gains (a total of $N_{par} = 3L$ real parameters).

We define $s_0(t;\ \mathbf{x})$ as the received signal without noise in the case that the received pulse $h(t)$ is equal to the model $h_0(t;\ \mathbf{x})$

$$s_0(t;\ \mathbf{x}) = \sum_{k=0}^{K-1} a(k)\, h_0(t-kT;\ \mathbf{x})$$

Note that the structure of $s_0(t;\ \mathbf{x})$ is known to the receiver, since we consider estimation based on a pilot sequence.

In practice, a modelling error usually occurs : there is no value of $\mathbf{x}$ such that $h(t) = h_0(t;\ \mathbf{x})$. We define $\mathbf{x}_0$ as the value of the parameter vector $\mathbf{x}$ that results in the best fit for $h_0(t;\ \mathbf{x})$ according to the following criterion

$$\mathbf{x}_0 = \arg\min_{\mathbf{x}} \int |h(t) - h_0(t;\ \mathbf{x})|^2\, dt \qquad (3)$$

As $\mathbf{x}_0$ is unknown to the receiver, it has to be estimated. The maximum-likelihood (ML) estimator according to the observation model

$$r(t) = s_0(t;\ \mathbf{x}) + w(t) \qquad (4)$$

is given by

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \int |r(t) - s_0(t;\ \mathbf{x})|^2\, dt \quad (5)$$

We will consider the estimate from (5), but with $r(t)$ given by (1). This estimate is used to compute samples of the corresponding symbol pulse according to

$$\hat{h}(mT_s) = h_0(mT_s;\ \hat{\mathbf{x}}) \quad (6)$$

In (6), the sampling rate $1/T_s$ should satisfy $2BT_s \leq 1$ in order to avoid aliasing.

B. *Nonparametric channel estimation*

The nonparametric estimation method involves the estimation of the samples $h(mT_s)$, with $2BT_s \leq 1$, ignoring the underlying model. The received pulse has an infinite duration so in theory this would result in an infinite number of samples to be estimated. In practice this is not possible, so only the samples of $h(t)$ that belong to a predefined range $[-L_1T_s,\ L_2T_s]$ are estimated. The number of samples that have to be estimated is equal to $N = L_1 + L_2 + 1$. The samples of the received signal $r(t)$ and of $h(t)$ can be arranged in $N_s = T/T_s$ independent vectors $\mathbf{r}_i$ and $\mathbf{h}_i$ respectively $(i = 1,\ \ldots,\ N_s)$

$$\mathbf{r}_i = [r(-L_1T_s + (i-1)T_s),\ r(-L_1T_s + T + (i-1)T_s)\ \ldots]^T$$

$$\mathbf{h}_i = [h(-L_1T_s + (i-1)T_s),\ h(-L_1T_s + T + (i-1)T_s)\ \ldots]^T$$

where $\mathbf{h}_i$ has N_i entries so that $N = \sum_{i=1}^{N_s} N_i$. Assuming $h(t)$ is limited to N samples, each vector $\mathbf{r}_i$ can be written as

$$\mathbf{r}_i = \mathbf{A}_i\mathbf{h}_i + \mathbf{w}_i \quad (7)$$

where $\mathbf{A}_i$ is a $(K + N_i - 1) \times N_i$ Toeplitz matrix with first column $[a_0 \ldots a_{K-1}\ \mathbf{0}_{1\times N_i - 1}]^T$ and $\mathbf{w}_i$ is the noise vector. The ML estimate $\hat{\mathbf{h}}_i$ corresponding to the observation model (7) is given by [4]

$$\hat{\mathbf{h}}_i = \left(\mathbf{A}_i^H\mathbf{A}_i\right)^{-1}\mathbf{A}_i^H\mathbf{r}_i \quad (8)$$

where $[.]^H$ denotes the Hermitian operator. We will consider the estimate (8), but with $\mathbf{r}_i$ corresponding to the observation model (1), with $h(t)$ not necessarily limited to N samples.

III. Lower bounds on the MSE on the received pulse $h(t)$

The performance of both estimation methods is compared in terms of the MSE on the received pulse. This MSE is defined as

$$MSE = \mathrm{E}\left[\int_{-\infty}^{+\infty}\left|h(t) - \hat{h}(t)\right|^2 dt\right] \quad (9)$$

$$= T_s\mathrm{E}\left[\sum_{l=-\infty}^{+\infty}\left|h(lT_s) - \hat{h}(lT_s)\right|^2 dt\right] \quad (10)$$

with E [.] denoting expectation over the noise and over all possible pilot sequences[1]. In this section we are going to derive some theoretic lower bounds for the MSE.

A. *Parametric channel estimation*

The MSE (9) for the parametric channel estimation method is defined as

$$MSE = \mathrm{E}\left[\int_{-\infty}^{+\infty}|h(t) - h_0(t;\ \hat{\mathbf{x}})|^2\, dt\right] \quad (11)$$

It can be shown that $h_0(t;\ \hat{\mathbf{x}})$ is an asymptotically unbiased estimate of $h_0(t;\ \mathbf{x}_0)$, so this MSE can be written as

$$MSE = \int_{-\infty}^{+\infty}|h(t) - h_0(t;\ \mathbf{x}_0)|^2\, dt + \mathrm{E}\left[\int_{-\infty}^{+\infty}|h_0(t;\ \hat{\mathbf{x}}) - h_0(t;\ \mathbf{x}_0)|^2\, dt\right] \quad (12)$$

We see that the MSE (12) consists of a term caused by the modelling error and a term caused by the estimation error $\hat{\mathbf{x}} - \mathbf{x}_0$.

We introduce the CRB corresponding to the observation model (4) to lower bound the second term in (12). This yields (see appendix A)

$$MSE \geq \int_{-\infty}^{+\infty}|h(t) - h_0(t;\ \mathbf{x}_0)|^2\, dt + \frac{N_0N_{par}}{2KE_s} \quad (13)$$

We see that this MSE bound consists of two parts: a part caused by the modeling error and a part caused by additive noise. Apparently, the second term is proportional to the number N_{par} of estimated parameters.

B. *Nonparametric channel estimation*

For the nonparametric channel estimation method, the MSE (9) on the received pulse can be expressed as

$$MSE = T_s\mathrm{E}\left[\sum_{m=-\infty}^{+\infty}\left|h(mT_s) - \hat{h}(mT_s)\right|^2\right] \quad (14)$$

Taking into account that $\hat{h}(mT_s) = 0$ for $m < -L_1$ and $m > L_2$, (14) reduces to

$$MSE = T_s\sum_{m\notin[-L_1,\ L_2]}|h(mT_s)|^2 + T_s\mathrm{E}\left[\sum_{m=-L_1}^{+L_2}\left|h(mT_s) - \hat{h}(mT_s)\right|^2\right] \quad (15)$$

The first part of this expression can be considered as a modeling error (when the duration of the actual pulse $h(t)$ exceeds NT_s). The second term of the MSE can again be lower bounded by

[1]This MSE also holds for a fixed pseudo-random pilot sequence of sufficient length

the CRB, corresponding to the observation model (7) (see appendix B). This yields for the MSE (14)

$$MSE \geq T_s \sum_{m \notin [-L_1,\, L_2]} |h\,(mT_s)|^2 + \frac{N_0 2N}{2KE_s} \qquad (16)$$

The second term of (16) is proportional to the number of estimated real-valued parameters $2N$. Compared to (13), we observe that both expressions (13) and (16) have the same structure: they both consist of a term caused by a modeling error and a part caused by the additive noise. For the parametric method the modeling error is caused by the difference between $h\,(t)$ and $h_0\,(t;\, \mathbf{x}_0)$ while for the nonparametric method the modeling error is caused by the duration of $h\,(t)$ exceeding NT_s.

IV. Numerical Results

In this section, we provide numerical results to illustrate the obtained MSE lower bounds and to compare the two estimation methods.

The transmit pulse is a square root raised cosine pulse with 25% roll-off. The pilot sequence consists of 20 BPSK symbols, that are randomly changed from one block to the next. We consider a 3-path channel with delays $[0T,\ 0.55T,\ 1.65T]$ and complex path gains $\left[0.8085e^{j\frac{\pi}{8}},\ 0.5659e^{j\frac{2\pi}{8}},\ 0.1617e^{-j\frac{3\pi}{8}}\right]$. The sample period T_s is set to $T/2$ so that no aliasing occurs.

Figs. 1 and 2 are related to parametric and nonparametric estimation, respectively. They show the theoretical lower bound on the MSE (solid lines), along with simulation results (dashed lines) obtained from an actual channel estimator. The simulation results confirm the trend observed in the MSE lower bound.

For parametric estimation method, results in Fig. 1 are shown for different values of L in the channel model (2), with the true number of paths equal to 3; the corresponding values of $\mathbf{x}_0$ from (3) are given in Table 1. The simulation results hold for the estimator introduced in [3] based on the iterative SAGE algorithm [5].

For nonparametric estimation, results in Fig. 2 are shown for different values of the number N of estimated coefficients $\{h\,(mT_s)\}$. The simulation results correspond to the estimator (8).

For both methods, we observe from Figs. 1 and 2 that the modelling error gives rise to an MSE floor at large E_s/N_0. This floor can be reduced by increasing the number of real-valued parameters ($3L$ or $2N$) to be estimated, at the expense of an increase of the noise-dependent term of the MSE. Hence, for each value of E_s/N_0 an optimum value of $3L$ or $2N$ exists, that minimizes the MSE.

Fig. 3 shows the MSE lower bound for nonparametric estimation at $E_s/N_0 = 10$ dB as a function of the number of estimated coefficients $\{h\,(mT_s)\}$, along with the contributions caused by the modeling error and the estimation error. Also shown is the MSE lower bound for parametric estimation with $L = 3$. We observe that for the given setup the MSE lower bound for nonparametric estimation is minimum for $N = 8$, but still exceeds the MSE lower bound for parametric estimation with $L = 3$.

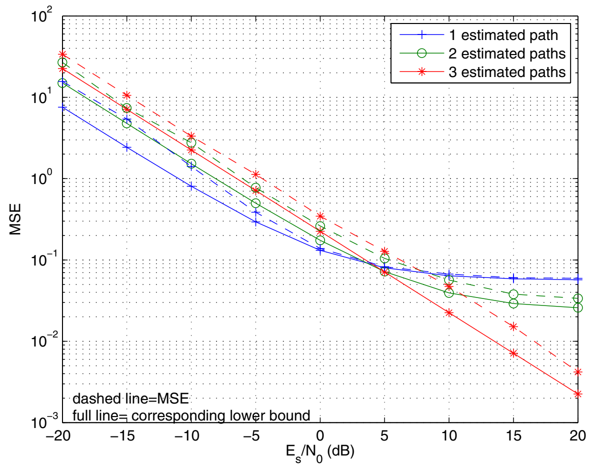

Figure 1: Influence of the number of estimated paths on the MSE for parametric channel estimation

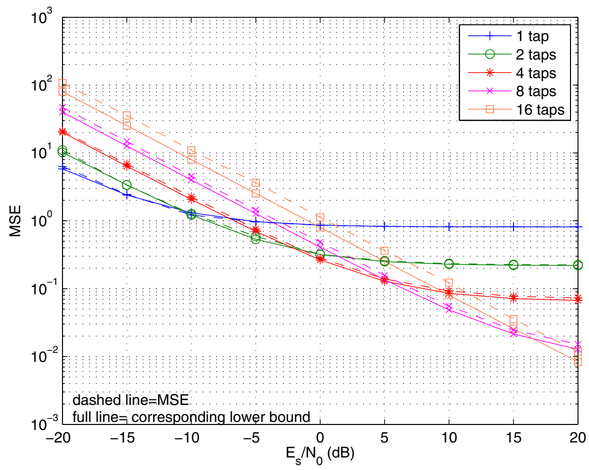

Figure 2: Influence of the number of estimated taps on the MSE for nonparametric channel estimation

Table 1: Values for the parameters $\mathbf{x}_0$ for the different numbers of estimated paths

Number of estimated paths	delays	amplitudes
3	$[0T,\ 0.55T,\ 1.65T]$	$\left[0.8085e^{j\frac{\pi}{8}},\ 0.5659e^{j\frac{2\pi}{8}},\ 0.1617e^{-j\frac{3\pi}{8}}\right]$
2	$[0T,\ 0.481T]$	$[0.8425+j0.33,\ 0.5343+j0.5488]$
1	$[0.216T]$	$[1.2439+j0.7949]$

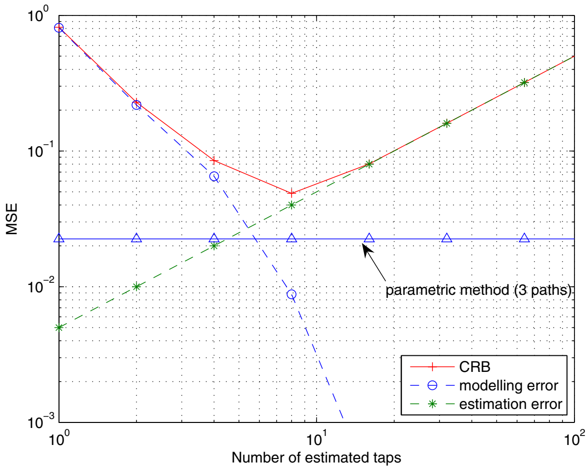

Figure 3: Influence of the number of estimated taps on the CRB for $E_s/N_0 = 10$ dB.

V. Conclusion

We have compared a parametric and a nonparametric channel estimation method for multipath channels in terms of the MSE on the received symbol pulse. We have shown that the MSE consist of a modeling error term and an estimation error term. We derived the respective Cramer-Rao lower bound on the estimation error term for both methods. We investigated the influence of number of estimated paths for the parametric estimation method and the number of estimated channel taps for the nonparametric estimation method. We noticed that for every E_s/N_0 there exists an optimal number of parameters to be estimated, which minimizes the MSE. Finally, our analysis showed that the parametric method yields the lower MSE.

Appendix A

The second term of (12) can be lower bounded by the corresponding Cramer Rao Lower bound (CRB) [6]

$$\mathrm{E}\left[\int_{-\infty}^{+\infty}\left|h_0\left(t;\ \hat{\mathbf{x}}\right)-h_0\left(t;\ \mathbf{x}_0\right)\right|^2 dt\right] \geq \mathrm{E}\left[\int_{-\infty}^{+\infty}\Re\left\{\mathbf{v}\left(t\right)^H CRB\left(\mathbf{x}_0\right)\mathbf{v}\left(t\right)\right\}dt\right] \quad (17)$$

where $CRB\left(\mathbf{x}_0\right)$ is the inverse of the Fischer information matrix $\mathbf{J}_p$ corresponding to $\mathbf{x}_0$ and where

$$\mathbf{v}\left(t\right) = \left[\frac{\partial h_0\left(t;\ \hat{\mathbf{x}}\right)}{\partial \hat{\mathbf{x}}}\right]_{\hat{\mathbf{x}}=\mathbf{x}_0}$$

$\mathbf{J}_p$ is defined as [7]

$$\mathbf{J}_p = \frac{2}{N_0}\int\Re\left\{\left[\frac{\partial s_0\left(t;\ \mathbf{x}\right)}{\partial \mathbf{x}}\right]_{\mathbf{x}=\mathbf{x}_0}\left[\frac{\partial s_0\left(t;\ \mathbf{x}\right)}{\partial \mathbf{x}}\right]_{\mathbf{x}=\mathbf{x}_0}^H\right\}dt$$

By averaging over the pilot symbols, (17) can be further simplified

$$\mathrm{E}_{a(k)}\left[\int_{-\infty}^{+\infty}\Re\left\{\mathbf{v}\left(t\right)^H\mathbf{J}_p^{-1}\mathbf{v}\left(t\right)\right\}dt\right] = \mathrm{tr}\left(\int_{-\infty}^{+\infty}\Re\left\{\mathbf{v}\left(t\right)\mathbf{v}\left(t\right)^H\right\}dt\ \mathrm{E}_{a(k)}\left[\mathbf{J}_p^{-1}\right]\right) \quad (18)$$

where tr $(.)$ denotes the trace

Since the inverse of a matrix is a matrix convex function, Jensen's inequality for matrices [8] can be applied

$$\mathrm{E}_{a(k)}\left[\mathbf{J}_p^{-1}\right] \geq \left(\mathrm{E}_{a(k)}\left[\mathbf{J}_p\right]\right)^{-1}$$

The averaging of $\mathbf{J}_p$ over the pilot symbols yields

$$\mathrm{E}_{a(k)}\left[\mathbf{J}_p\right] = \frac{2KE_s}{N_0}\int\Re\left\{\mathbf{v}\left(t\right)\mathbf{v}\left(t\right)^H\right\}dt \quad (19)$$

so expression (18) is lower bounded by

$$\frac{N_0}{2KE_s}\mathrm{tr}\left(\int_{-\infty}^{+\infty}\Re\left\{\mathbf{v}\left(t\right)\mathbf{v}\left(t\right)^H\right\}dt\left[\int\Re\left\{\mathbf{v}\left(t\right)\mathbf{v}\left(t\right)^H\right\}dt\right]^{-1}\right) \quad (20)$$

This yields for (20)

$$\frac{N_0}{2KE_s}\mathrm{tr}\left(\mathbf{I}_{N_{par}}\right) = \frac{N_0 N_{par}}{2KE_s}$$

where $\mathbf{I}_m$ is the $m\times m$ identity matrix.

Appendix B

The second term of (15) can be rewritten as

$$T_s\mathrm{E}\left[\sum_{m=-L_1}^{+L_2}\left|h\left(mT_s\right)-\hat{h}\left(mT_s\right)\right|^2\right] = T_s\sum_{i=1}^{N_s}\mathrm{E}\left[\left|\mathbf{h}_i-\hat{\mathbf{h}}_i\right|^2\right]$$

Every term of the summation can be lower bounded by the corresponding CRB

$$T_s \sum_{i=1}^{N_s} \mathrm{E}\left[\left|\mathbf{h}_i - \hat{\mathbf{h}}_i\right|^2\right] \geq T_s \mathrm{tr}\left(\mathbf{J}_{npi}^{-1}\right)$$

where $\mathbf{J}_{npi}$ is the corresponding Fischer information matrix. For a model like (7) the expression for the Fischer information matrix is a well known result [7] and given by

$$\mathbf{J}_{npi} = \frac{2T_s}{N_0} \begin{bmatrix} \Re\left(\mathbf{A}_i^H \mathbf{A}_i\right) & -\Im\left(\mathbf{A}_i^H \mathbf{A}_i\right) \\ \Im\left(\mathbf{A}_i^H \mathbf{A}_i\right) & \Re\left(\mathbf{A}_i^H \mathbf{A}_i\right) \end{bmatrix}$$

This results in

$$T_s \sum_{i=1}^{N_s} \mathrm{E}\left[\left|\mathbf{h}_i - \hat{\mathbf{h}}_i\right|^2\right] \geq N_0 \sum_{i=1}^{N_s} \mathrm{tr}\left(\left(\Re\left[\mathbf{A}_i^H \mathbf{A}_i\right]\right)^{-1}\right)$$

Averaging this result over the pilot symbols yields

$$T_s \sum_{i=1}^{N_s} \mathrm{E}\left[\left|\mathbf{h}_i - \hat{\mathbf{h}}_i\right|^2\right] \geq N_0 \sum_{i=1}^{N_s} \mathrm{tr}\left(\mathrm{E}_{a(k)}\left[\left(\Re\left[\mathbf{A}_i^H \mathbf{A}_i\right]\right)^{-1}\right]\right)$$

We can apply Jensen's inequality for matrices [8] which results in

$$T_s \sum_{i=1}^{N_s} \mathrm{E}\left[\left|\mathbf{h}_i - \hat{\mathbf{h}}_i\right|^2\right] \geq N_0 \sum_{i=1}^{N_s} \frac{N_i}{KE_s}$$

which can be further simplified using the fact that $N = \sum_{i=1}^{N_s} N_i$

$$T_s \sum_{i=1}^{N_s} \mathrm{E}\left[\left|\mathbf{h}_i - \hat{\mathbf{h}}_i\right|^2\right] \geq \frac{N_0 2N}{2KE_s}$$

References

[1] J.-T. Chen, J. Kim and J.-W. Liang. Multichannel MLSE Equalizer with Parametric FIR Channel Identification. *IEEE Trans. Veh. Technol.*, 48(6):1923–1935, November 1999.

[2] J.-T. Chen and Y.-C. Wang. Performance Analysis of the Parametric Channel Estimators for MLSE Equalization in Multipath Channels with AWGN. *IEEE Trans. Comm.*, 49(3):393–396, March 2001.

[3] D. Van Welden, M. Moeneclaey, H. Steendam. Parametric versus nonparametric data-aided channel estimation in a multipath fading environment. In *Proc. IEEE Symposium on Communications and Vehicular Technology*, Liège, Belgium, November 2006.

[4] H. Meyr, M. Moeneclaey, S.A. Fechtel. *Synchronization, Channel Estimation, and Signal Processing*, volume 2 of *Digital Communication Receivers*. John Wiley & Sons, 1997.

[5] B. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. Pedersen. Channel Parameter Estimation in Mobile Radio Environments Using the SAGE Algorithm. *IEEE Journal on Selected Areas in Communications*, 17(3):434–450, March 1999.

[6] B. Friedlander and J.M. Francos. Estimation of Amplitude and Phase Parameters of Multicomponent Signals. *IEEE Trans. Signal Processing*, 43(4):917–926, April 1995.

[7] H.L. Van Trees. *Detection, Estimation, and Modulation Theory, Part I.* Wiley and Sons, October 2001.

[8] R. Ravikanth and S.P. Meyn. Bounds on Achievable Performance in the Identification and Adaptive Control of Time-Varying Systems. *IEEE Trans. Automatic Control*, 44(4):670–682, April 1999.