Interactive Multiple Model Estimation of Doubly-Selective Channels for OFDM systems

Mahmoud Ashour† and Amr El-Keyi‡
† Electrical Engineering Dept., Pennsylvania State University, PA, USA.
‡ Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada.

Abstract—In this paper, we propose an algorithm for channel estimation, acquisition and tracking, for orthogonal frequency division multiplexing (OFDM) systems. The proposed algorithm is suitable for vehicular communications that encounter very high mobility. A preambule sequence is used to derive an initial estimate of the channel using least squares (LS). The temporal variation of the channel within one OFDM symbol is approximated by two complex exponential basis expansion models (CE-BEM). One of the Fourier-based BEMs is intended to capture the low frequencies in the channel (slow variations corresponding to low Doppler), while the other is destined to capture high frequencies (fast variations corresponding to high Doppler). Kalman filtering is employed to track the BEM coefficients iteratively on an OFDM symbol-by-symbol basis. An interactive multiple model (IMM) estimator is implemented to dynamically mix the estimates obtained by the two Kalman filters, each of which matched to one of the BEMs. Extensive numerical simulations are conducted to signify the gain obtained by the proposed combining technique.

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is famous for its robustness against frequency-selective wireless channels. Spectrally shaped channels are transformed into a set of parallel flat-fading subchannels to enable high data rate transmission. OFDM is widely used in new emerging standards for mobile communications, e.g., long-term evolution (LTE) and WiMax. However, the design of OFDM systems in the presence of high mobility is challenging. The relative motion between the transmitter and receiver induces temporal variations in the channel which destroys orthogonality among the subcarriers. Thus, OFDM receivers require accurate channel information for reliable detection of the transmitted signals.

Accurate channel estimation in the presence of high mobility conditions has received a lot of attention [1], [2]. A pilot-aided technique is developed in [1], where the received signal samples containing pilot tones in the frequency domain are used to estimate and track the channel impulse response (CIR) coefficients via a recursive least-squares (RLS) algorithm. In [2], the authors propose a pilot-aided technique that exploits the channel correlations in time and frequency domains. Most of the computations are performed offline to reduce the complexity of the algorithm.

Another widely investigated trend is the use of parametric channel models. The temporal variation of the channel coefficients within one OFDM symbol is approximated by a basis expansion model (BEM). The time-varying channel taps are represented as a weighted sum of time-varying basis functions. Among the candidate basis functions are the complex exponential (Fourier) functions [3], [4], polynomials [5], and discrete Karhuen-Loeve functions that correspond to the most significant eigenvectors of the channel correlation matrix [6].

In this paper, we propose an algorithm for channel estimation, acquisition and tracking, in the presence of high mobility. The algorithm uses a preambule sequence to derive an initial estimate of the channel coefficients using least squares (LS). Then, the temporal variation of the channel coefficients is modelled by two complex exponential BEMs. One of the Fourier-based BEMs is intended to capture the low frequencies in the channel (slow variations corresponding to low Doppler), while the other aims at capturing high frequencies (fast variations corresponding to high Doppler). Kalman filtering proposed in [7] is then used to track the BEM coefficients under each individual BEM. An interactive multiple model (IMM) estimator is implemented to dynamically combine the estimates obtained by the two Kalman filters. It is shown through numerical simulations that the proposed combining technique outperforms the use of only one Kalman filter matched to a BEM containing a concatenation of the basis vectors available in both of the used BEMs.

Notation: We denote scalars by lower-case letters (e.g. $x$), vectors by lower-case boldface letters (e.g. $x$) and matrices by upper-case boldface letters (e.g. $X$). A hat over a variable refers to its estimate (e.g. $\hat{x}$ is an estimate of $x$). Superscripts $T$ and $H$ denote transpose and hermitian transpose, respectively. We reserve $E$ for the statistical expectation operator. The $N \times N$ identity matrix is denoted by $I_N$. The $(k,m)$th entry of the matrix $X$ is denoted by $X_{k,m}$. A diagonal matrix with $x$ on its main diagonal is denoted by $\text{diag}(x)$, $\text{blkdiag}(X, Y)$ is a block diagonal matrix with $X$ and $Y$ on its main diagonal.

II. SYSTEM MODEL

In this section, we present a matrix-vector model for a discrete-time baseband-equivalent OFDM system. We consider a system with $N$ subcarriers. Let $x_n$ denote the frequency-domain transmitted symbols of the $n$th OFDM symbol. The corresponding time-domain samples, $\tilde{x}_n$, are obtained through an $N$-point inverse discrete Fourier transformation (IDFT) of $x_n$. Thus, $\tilde{x}_n = F^H x_n$. $F$ denotes the $N$-point DFT matrix. A cyclic prefix (CP) of length $N_g$ is augmented to the time-domain transmitted samples to avoid inter-symbol interference (ISI) between consecutive OFDM symbols. Therefore, the OFDM symbol duration is given by $T = N_s T_s$, where $T_s$ denotes the sampling interval of the system and $N_s = N + N_g$.
After transmission over a multipath fading channel with $L < N_q$ taps, the received time-domain signal (after CP removal) corresponding to the $n$th OFDM symbol is written following the model presented in [2] as

$$y_n = H_n x_n + w_n \quad (1)$$

where $H_n$ is the $N \times N$ time-domain channel matrix with the following structure

$$H_n = \begin{bmatrix}
h_{0,n}(0) & 0 & \ldots & h_{1,n}(0) \\
h_{1,n}(1) & h_{0,n}(1) & \ldots & h_{2,n}(1) \\
\vdots & \ddots & \ddots & \vdots \\
h_{L-1,n}(L-1) & h_{L-2,n}(L-1) & \ldots & 0 \\
0 & h_{L-1,n}(L) & \ldots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & h_{0,n}(N-1) \\
\end{bmatrix} \quad (2)$$

and $h_{l,n}(q)$ is the complex channel gain of the $l$th tap in the $n$th OFDM symbol at time instant $q$, and $w_n$ is the time-domain additive zero-mean complex Gaussian noise with covariance matrix $\sigma_n^2 I_N$ added to the $n$th received OFDM symbol. Applying DFT to (1), we obtain the frequency domain received symbols

$$Y_n = G_n X_n + WN \quad (3)$$

where $G_n = FH_n F^H$ is the frequency domain channel matrix at the $n$th OFDM symbol, and $WN = FW_n$ is the frequency domain noise vector added to the $n$th OFDM symbol.

If the channel is quasi-static, i.e., the channel coefficients are constant within the same OFDM symbol, then the channel matrix $H_n$ becomes circulant. Hence, it can be diagonalized through applying DFT and IDFT operations yielding a diagonal matrix $G_n$. Thus, the subcarriers keep their orthogonality and no inter-carrier interference (ICI) is introduced. However, in very high mobility conditions, the relative motion between the transmitter and the receiver induces temporal variations in the channel. Therefore, the matrix $H_n$ can no longer be considered circulant. This introduces off-diagonal elements to the corresponding matrix $G_n$ causing ICI. When the channel is fast time varying, ICI becomes significant and severely degrades the system performance. This stimulates the necessity of estimating the channel coefficients at every sampling time instant. Next, we introduce the BEM that helps us do that job.

### III. Basis Expansion Model

In this section, we introduce the approach of using a BEM to model the temporal variation of the channel coefficients within one OFDM symbol. The channel taps are considered wide sense stationary (WSS) zero-mean complex Gaussian processes of variances $\{\sigma_i^2\}_{i=0}$. The $N$-dimensional vector of the $l$th channel tap at the $n$th OFDM symbol is defined as

$$h_{l,n} = [h_{l,n}(0), h_{l,n}(1), \ldots, h_{l,n}(N-1)]^T \quad (4)$$

Following Jakes’ power spectral model of maximum Doppler frequency $f_d$ [8], the temporal correlation matrix for a time-lag $p$, $R_{h_l}^{(p)} = E[h_{l,n} h_{l,n+p}^H]$, is given by

$$R_{h_l}^{(p)} = \sigma_h^2 \delta_0(2\pi f_d T_s (k - m + pN_s)) \quad (5)$$

There exist $N$ samples for each channel tap in every OFDM symbol. This yields a total number of $LN$ samples for the whole channel per OFDM symbol. We use a BEM to reduce the dimensions of the space of parameters required to be estimated. The main goal of the BEM is to accurately model the temporal variation of the channel coefficients within the same OFDM symbol. This variation is approximated by a linear combination of a few bases vectors, $b_j$, as follows

$$h_{l,n} = B c_{l,n} + \nu_{l,n} \quad (6)$$

where $B = [b_0, b_1, \ldots, b_{N_r-1}]$ is the $N \times N_c$ basis matrix containing $N_c$ bases vectors, $c_{l,n}$ is the $N_c \times 1$ vector of the BEM coefficients corresponding to the $l$th channel tap at the $n$th OFDM symbol and $\nu_{l,n}$ is the modelling error. Thus, the optimal BEM coefficients [9] are given by

$$c_{l,n} = (B^H B)^{-1} B^H h_{l,n} \quad (7)$$

From (7), the correlation matrix of the BEM coefficients for a time-lag $p$, $R^{(p)}_{c_l} = E[c_{l,n} c_{l,n+p}^H]$ is given by

$$R^{(p)}_{c_l} = (B^H B)^{-1} B^H R_{h_l}^{(p)} B (B^H B)^{-1} \quad (8)$$

In [5] and [7], (1) is derived in terms of the BEM neglecting the modelling error yielding

$$y_n = S_n c_n + w_n \quad (9)$$

where the $LN_c \times 1$ vector $c_n$ and the $N \times LN_c$ matrix $S_n$ are given by

$$c_n = [c_{0,n} c_{1,n}^T, \ldots, c_{N_{r-1},n}^T]^T \quad (10)$$

$$S_n = \frac{1}{\sqrt{N}} [V_{0,n}, V_{1,n}, \ldots, V_{L-1,n}] \quad (11)$$

$$V_{l,n} = \left[ D_0 \text{diag} \{X_l f_1, \ldots, D_{N_c-1} \text{diag} \{X_l f_1} \} \right] \quad (12)$$

$l_f$ is the $l$th column of the DFT matrix $F$, and the $N \times N$ matrix $D_d$ is given by

$$D_d = \text{diag} \{b_d\} F^H \quad (13)$$

### IV. Kalman Filtering Matched to a BEM

In this section, we introduce the acquisition phase of the proposed algorithm. Then, we briefly present the Kalman filtering approach proposed by [7], where the Kalman filter is matched to a given BEM.

#### A. Acquisition

A preamble sequence is transmitted in the beginning of each frame to obtain an initial estimate of the BEM coefficients.
Consider $K$ training OFDM symbols transmitted as a preamble. Then, the received preamble signal is given by

$$ y = Sc + w $$

where

$$ y = [y_T^1, \ldots, y_T^{K-1}]^T $$
$$ S = \text{blkdiag}(S_0, \ldots, S_{K-1}) $$
$$ c = [c_T^0, \ldots, c_T^{K-1}]^T $$
$$ w = [w_0^T, \ldots, w_{K-1}^T]. $$

Thus, the initial LS estimate of the BEM coefficients is

$$ \hat{c} = (S^H S)^{-1} S^H y. $$

B. Tracking

Simon et al. proposed a Kalman filter-based tracking for the BEM coefficients in [7]. The dynamics of $c_{l,n}$ is modelled via a first order auto-regressive (AR) model as follows

$$ c_{l,n} = A_l c_{l,n-1} + u_{l,n} $$

where $A_l$ is the transition $N_c \times N_c$ matrix of BEM coefficients of the $l$th channel taps across two consecutive OFDM symbols and $u_{l,n}$ is the complex Gaussian noise vector of covariance matrix $U_l$. The AR model parameters can be computed through Yule-Walker equations as

$$ A_l = R_l^{(0)} (R_l^{(0)})^{-1} $$
$$ U_l = R_l^{(0)} + A_l R_l^{(-1)}. $$

Using (20), the AR model for $c_n$ is given by

$$ c_n = A c_{n-1} + u_n $$

where $A = \text{blkdiag} \{A_0, \ldots, A_{L-1}\}$ and the $LN_c \times LN_c$ matrix $u_n = [u_{0,n}^T, \ldots, u_{L-1,n}^T]^T$ is the complex Gaussian noise vector of covariance matrix $U = \text{blkdiag} \{U_0, \ldots, U_{L-1}\}$. Using the state equation (23) and the measurement equation (9), the Kalman filter equations (10) presented in [7] are time update equations (TUE):

$$ \hat{c}_{n|n-1} = A \hat{c}_{n|n-1} $$
$$ M_{n|n-1} = AM_{n-1|n-1}A^H + U. $$

Measurement update equations (MUE):

$$ \hat{c}_{n|n} = \hat{c}_{n|n-1} + K_n (y_n - S_n c_{n|n-1}) $$
$$ M_{n|n} = [I_{LN_c} - K_n S_n] M_{n|n-1}. $$

where the filter gain $K_n$ and the innovation covariance matrix $Q_n$ are given respectively by

$$ K_n = M_{n|n-1} S_n^H Q_n^{-1} $$
$$ Q_n = \sigma^2_w I_{N_c} + S_n M_{n|n-1} S_n^H. $$

The Kalman filter is implemented in a decision-directed algorithm, where the channel estimate at the $(n-1)$th iteration is used to construct the channel matrix $H_n$ in (7) at the $n$th iteration for equalization. Thus, we get the transmitted symbols $x_n$ which we use to construct the matrix $S_n$.

V. INTERACTIVE MULTIPLE MODEL ESTIMATOR

In multiple model environments, the optimal state estimate is a function of the elemental state estimates obtained by the estimators tuned to all possible parameter histories. Thus, with time, an exponentially increasing number of filters is required to keep track of all possible model parameter histories. Many suboptimal techniques are proposed to overcome this complexity issue among which the IMM is the most cost efficient.

The structure of the IMM estimator is shown in Fig. 1. At time $n$, the state estimate is computed under each possible current model using two filters, with each filter using a different combination of the previous model-conditioned state estimates (mixed initial condition). The model switching process is assumed to be a Markov chain conditioned transition probability matrix

$$ P = \begin{bmatrix} p^{(11)} & p^{(12)} \\ p^{(21)} & p^{(22)} \end{bmatrix}, $$

where $p^{(ij)}$ denotes the transition probability from model $i$ to model $j$. A cycle of the algorithm consists of the following steps mentioned in [11]:

1) Calculation of the mixing probabilities: The probability that model $i$ is in effect at time $(n-1)$ given that model $j$ is in effect at time $n$ and given the measured data up to step $(n-1)$ is

$$ \mu_{n-1|n-1}^{(i|j)} = \frac{1}{\bar{\mu}_{j}} \mu_{n-1|n}^{(i)} \mu_{n-1|n}^{(j)}, \quad i, j = 1, 2 $$

where $\mu_{n-1|n}^{(j)}$ is the probability that model $i$ is in effect at time $(n-1)$ and the normalization constants are

$$ \bar{\mu}_{j} = \sum_{i=1}^{2} p^{(ij)} \mu_{n-1|n}^{(i)}. $$

2) Mixing: Starting with the state estimates of both filters and their associated covariances at time $(n-1)$, one computes the mixed initial condition $\hat{c}_{n-1|n-1}^{(0|j)}$ and $M_{n-1|n-1}^{(0|j)}$ matched to filter $j$ as

$$ \hat{c}_{n-1|n-1}^{(0|j)} = \sum_{i=1}^{2} \hat{c}_{n-1|n-1}^{(i)} \mu_{n-1|n-1}^{(i)} $$
$$ M_{n-1|n-1}^{(0|j)} = \sum_{i=1}^{2} \mu_{n-1|n-1}^{(i)} \left\{ M_{n-1|n-1}^{(i)} + \hat{c}_{n-1|n-1}^{(i)} \hat{c}_{n-1|n-1}^{(i)^H} \right\}. $$

3) Mode matched filtering: The state estimate (32) and its associated covariance (33) are used as input to the filter matched to model $j$ to yield $\hat{c}_{n|n}^{(0|j)}$ and $M_{n|n}^{(0|j)}$. The likelihood function corresponding to both filters which is the probability of the measurement $y_n$, given that model $j$ is in effect at time $n$ and all the data history up to
time \((n-1)\) are computed as
\[
\Lambda_n^{(j)} = \frac{1}{|[2\pi Q_n^{(j)}]|} \exp\{-\frac{1}{2} [y_n - \hat{y}_{n-1}^{(j)}]^H \times (Q_n^{(j)})^{-1} [y_n - \hat{y}_{n-1}^{(j)}]\}.
\]
(34)
where \(\hat{y}_{n-1}^{(j)}\) is the predicted measurement at time \(n\) given data history up to time \((n-1)\) and \([y_n - \hat{y}_{n-1}^{(j)}]\) is the innovation at step \(n\) with covariance \(Q_n^{(j)}\).

4) Mode probability update: Mode probability is updated using the likelihood function as follows
\[
\mu_n^{(j)} = \frac{1}{c} \Lambda_n^{(j)} \hat{c}^{(j)}.
\]
(35)
where \(\hat{c}^{(j)}\) is given by \([31]\) and \(c\) is a normalization constant given by
\[
c = \sum_{j=1}^{2} \Lambda_n^{(j)} \hat{c}^{(j)}.
\]
(36)

5) Estimate and covariance combination: Combination of the model conditioned estimates and their associated covariances is done according to the mixing equations
\[
\hat{c}_{n|n} = \sum_{j=1}^{2} \hat{c}_{n|n}^{(j)} \mu_{n}^{(j)}.
\]
(37)
\[
M_{n|n} = \sum_{j=1}^{2} \left[ M_{n|n}^{(j)} \left[ \hat{c}_{n|n}^{(j)} - \hat{c}_{n|n} \right] [\hat{c}_{n|n}^{(j)} - \hat{c}_{n|n}]^H \right].
\]
(38)

VI. SIMULATION RESULTS

In this section, the performance of the proposed algorithm is evaluated in terms of the mean square error (MSE) in the estimate of the most significant channel tap. We consider a SISO-OFDM system employing quadrature phase shift keying (QPSK) with \(N = 64\), \(N_g = N/4\), \(N_d = N/16\) and \(1/T_s = 20\)MHz. The transmitted frame consists of 200 OFDM symbols. The first \(K = 2\) OFDM symbols of each transmitted frame are used as a preamble sequence utilized to derive the initial channel estimate. Binary Phase shift keying (BPSK) is the modulation technique employed for the preamble. Simulation results are averaged over \(10^3\) Monte Carlo runs.

We consider a multipath channel with 4 taps (equispaced with respect to delay) whose power delay profile is given by \((0, -1, -3, -9)\) dB. Therefore, the maximum delay spread of the channel is given by \(4T_s\). The proposed combining technique (IMM) utilizes two complex exponential BEMs with \(N_c = 3\) coefficients each. The first BEM intended for capturing the slow variations in the channel coefficients, within one OFDM symbol, uses a basis matrix
\[
[B_L])_{k,m} = e^{j\frac{2\pi}{N} k (m - \frac{N_c - 1}{2})}
\]
(39)
while the second BEM intended for capturing the fast variations has a basis matrix
\[
[B_H])_{k,m} = e^{j\frac{2\pi}{N} k (2m - (N_c - 1))}
\]
(40)
where the indices \(k\) and \(m\) start at zero.

From the definition of the two bases matrices, \(B_L\) and \(B_H\), we note that both are sub-matrices obtained from the N-point DFT matrix \(F\). The \(N_c = 3\) bases vectors of \(B_L\) correspond to the DC vector and the first two discrete frequencies \(\pm \left( \frac{2\pi}{N} \right)\), while the bases vectors of the matrix \(B_H\) correspond to the DC vector and the second two discrete frequencies \(\pm \left( \frac{4\pi}{N} \right)\). We conjecture that the temporal variation of the channel is better modelled by \(B_L\) in low Doppler scenarios, while \(B_H\) performs better in high Doppler scenarios. Thus, the proposed IMM-based combining technique is expected to perform well in a wide span of possible communication scenarios. The performance of the IMM is compared to algorithm [7] utilizing

Fig. 1: Structure of the IMM estimator

where $h$ with $N_a$ single Kalman filter matched to a basis matrix with
\[ \text{with larger basis given the same } L \text{ and } B \],

\[ \text{IMM outperforms algorithm [7] with } \mu_n(1) \text{ and } \mu_n(2) \text{ to indicate no prior knowledge of the channel.} \]

For illustration purposes, the random multipath channel is generated according to the BEMs. The channel follows $B_L$ in the first half of the transmitted frame (first 100 OFDM symbols), while it follows $B_H$ in the second half of the frame (second 100 OFDM symbols). The MSE of the estimate of the $l$th channel tap is defined as

\[ \text{MSE}_l = \frac{\| h_l - \hat{h}_l \|^2}{200N} \tag{41} \]

where $h_l$ denotes the $200N \times 1$ vector of the $l$th channel tap.

The MSE in the estimate of the most significant tap is plotted in Fig. 2 versus $E_b/N_0$. It is shown that the proposed IMM outperforms algorithm [7] with $N_c = 5$ (with bases vectors corresponding to all frequencies available in $B_L$ and $B_H$). This can be explained as follows: The number of parameters to be estimated using the IMM is $LN_c = 12$, which is less than $LN_c = 20$ in the case of algorithm [7] with larger basis given the same $N$-dimensional measurement set $y_n$. Thus, the Cramer Rao bound is in favour of the IMM.

The dynamics of the proposed IMM is shown for one transmitted frame in Fig. 3. The two mode probabilities $\{\mu_n(1), \mu_n(2)\}^2$ are plotted versus the OFDM symbol index. Fig. 3 shows the capability of the IMM to lock to the true BEM.

The sharp transition in the middle of the frame illustrates the enhanced sensitivity of the IMM to variations of the channel dynamics within the same frame.

VII. CONCLUSION

In this paper, we propose an algorithm for channel estimation, acquisition and tracking, using parametric channel models. The temporal variation of the channel coefficients within one OFDM symbol is modelled by two BEMs. Kalman filtering is employed to track the coefficients of both BEMs. Then, an IMM estimator is used to mix the estimates obtained via each Kalman filter. The message behind the paper is to show that we can achieve higher estimation accuracy through dynamically mixing the estimates obtained via different BEMs. These BEMs are not limited to $B_L$ and $B_H$. The same approach generalizes to any two (or more) basis matrices. Moreover, the algorithm adapts itself to dramatic variations in the underlying environment even within the same OFDM symbol. Numerical simulations signify the gain obtained through using the proposed combining technique.

REFERENCES

[1] H. Nguyen-Le, T. Le-Ngoc, and C. Ko, “RLS-based joint estimation and tracking of channel response, sampling, and carrier frequency offsets for OFDM,” IEEE Transactions on Broadcasting, vol. 55, no. 1, pp. 84–94, December 2009.
[2] N. A.-D. Tareq Y Al-Naffouri, KM Zahidul Islam and S. Lu, “A model reduction approach for OFDM channel estimation under high mobility conditions,” IEEE Transactions on Signal Processing, vol. 58, no. 4, pp. 2181–2193, 2010.
[3] X. Ma, G. B. Giannakis, and S. Ohno, “Optimal training for block transmissions over doubly selective wireless fading channels,” IEEE Transactions on Signal Processing, vol. 51, no. 5, pp. 1351–1366, 2003.
[4] S. He and J. K. Tugnait, “Doubly-selective channel estimation using exponential basis models and subblock tracking,” in Global Telecommunications Conference, 2007. GLOBECOM’07. IEEE, 2007, pp. 2847–2851.
[5] H. Hijazi and L. Ros, “Joint data QR-detection and Kalman estimation for OFDM time-varying Rayleigh channel complex gains,” IEEE Transactions on Communications, vol. 58, no. 1, pp. 170–178, 2010.
[6] M. Visintin, “Karlhunen-Loeve expansion of a fast Rayleigh fading process,” Electronics Letters, vol. 32, no. 18, p. 1712, 1996.
[7] E. P. Simon, M. Berbineau, and M. Liénard, “Joint CFO and channel acquisition and tracking based on parametric channel modeling for OFDM systems in the presence of high mobility,” in 2011 11th International Conference on Telecommunications (ICT). IEEE, 2011, pp. 565–570.
[8] W. C. Jakes, Microwave Mobile Communications. IEEE press, 1983.
[9] G. Leus, “On the estimation of rapidly time-varying channels,” in Proc. EUSIPCO, 2004, pp. 2227–2230.
[10] S. Kay, Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory. Englewood Cliffs, NJ: Prentice Hall, 1993.
[11] Y. Bar-Shalom, X. Li, and T. Kirubarajan, Estimation with applications to tracking and navigation: theory algorithms and software. Wiley-Interscience, 2001.