Joint symbol detection and timing estimation with Stochastic M-algorithm

Yufang Yin, Yufei Huang
Department of Electrical Engineering
The University of Texas at San Antonio
San Antonio, TX 78249-0669
E-mail: yyin@lonestar.utsa.edu, yhuang@utsa.edu

Jianqiu (Michelle) Zhang
Department of Electrical and Computer Engineering
University of New Hampshire
Durham, NH 03824.
Email: jianqiu.zhang@unh.edu

***Abstract*— In digital communications, the symbol timing estimation is an very important element for high quality data detection. This paper considers the problem of joint symbol detection and timing estimation, whose optimal solution is analytically intractable. In this paper, a stochastic M-algorithm is proposed for the solution. The stochastic M-algorithm is a novel efficient particle filtering algorithm designed for the discrete unknowns. To accommodate in the stochastic M-algorithm the continuous unknown symbol timing, the unscented Kalman filter is introduced, which leads to very efficient implementation. The simulation results illustrated that the stochastic M-algorithm achieves similar performance as particle filtering with only less than 1/25 of the complexity.**

I. Introduction

In digital communications, data transmission experiences distortion due to the transmission channel. The data detection performance at the receiver strongly depends on the quality of the estimation of channel parameters, such as the symbol timing, the carrier phase, and the channel gain. We focus our attention in this paper on joint data symbol detection and timing estimation.

Timing recovery is a nonlinear problem whose optimal solution is analytically intractable. The symbol timing is also varying constant with time, and the requirement to track its variation adds an additional difficulty to timing recovery. In the past, suboptimal adaptive estimation techniques have been reported in the literature. The extended Kalman filter (EKF) was proposed in [1], which approximates the nonlinear estimation problem with a linear solution. The unscented Kalman filter (UKF) was proposed in [2], which approximates the nonlinearity with an unscented transform and was shown to achieve improved results. However, there, for timing recovery, data symbols are assumed to be known, but in practice, joint symbol detection is required. The optimal joint solution requires calculation on all the combination of symbols over time, a formidable task with complexity exponentially increase with time. To reduce the complexity, decision-directed schemes are often adopted, whose performance is, however, limited by error propagation, an inherent shortfall of the decision-directed detection.

Recently, particle filtering, or sequential Monte Carlo sampling was applied to joint symbol detection and timing recovery [3]. In a particle filtering solution, a set of properly weighted samples, which approximate the desired posterior distribution, is generated sequentially over time. A prominent feature of particle filtering is that the optimal solution to problem can be approximated easily with the weighted samples with high accuracy. In addition, particle filtering provides soft information about the unknowns, which can be used for iterative processing. Nevertheless, the high computation complexity of particle hinders its practical implementation.

In [4], a new particle filtering algorithm, called the stochastic M-algorithm (SMA) is proposed for problems with the discrete unknowns. For detection in BLAST systems, the SMA is shown to provide better performance than the generic particle filtering solution with much less complexity. Motivated by the appealing feature of the SMA, we propose in this paper a novel solution under the SMA framework for joint symbol detection and timing estimation. To accommodate in the SMA the continuous unknown symbol timing, the UKF is introduced, which leads to very efficient implementation. The simulation results illustrated that the SMA achieves similar performance as particle filtering with only less than 1/25 of the complexity.

The remaining of the paper is organized as follows. Section II describes the system model. The proposed algorithm is developed in Section III. In Section VI, simulation results are presented. Finally, conclusions are drawn in Section IV.

II. Problem Formulation

Consider digital signals transmitted over a communication channel. The received signal envelope has the following form [3]:

$$y(t) = \sum_{n=0}^{\infty} s_n g(t - nT + \tau(t)) e^{j(\theta+\omega t)} + v(t) \qquad (1)$$

where $\{s_m\}$ represents the BPSK modulated symbols transmitted during the nth symbol period, $g(t)$ is the modulation pulse waveform, T is the symbol period, $\tau(t)$ is the time varying delay, θ and ω are the carrier phase and carrier frequency offsets, respectively, and $v(t)$ is the complex additive white Gaussian noise with 0 mean and the power spectrum density $N_0/2$.

Let us assume that $g(t)$ is a causal pulse with finite duration (i.e., the raised cosines with finite duration), which happens in practice due to the use of truncated Nyquist pulse. Then we can write the equivalent discrete-time signal model after sampling as [3]

$$y_k = \sum_{n=k-L}^{k} s_n g(kT_s - nT + \tau_k) e^{j\theta} e^{j\frac{2\pi}{N_s}k\nu} + v_k \tag{2}$$

where $L+1$ is the intersymbol interference (ISI) span ($L < M$), T_s is the sampling period, (for convenience, set $T_s = T$), $y_k = y(kT_s)$, $v_k = v(kT_s)$, $\tau_k = \tau(kT_s)$, and ν is the normalized frequency offset.

The variation of the timing delay can be modeled using a first order AR process [5]

$$\tau_k = a\tau_{k-1} + u_k \tag{3}$$

where a is the known model coefficient and u_k is the Gaussian noise with zero mean and the variance σ_u^2.

In this paper, we only focus on the situation that the symbols and the delays are unknown and we thus assume that the carrier phase and the frequency offsets are given. Under the above assumption, we can express the system in a more compact format through a dynamic state space model (DSSM)

$$\begin{cases} \tau_k &= a\tau_{k-1} + u_k & \textit{state equation} \\ \mathbf{s}_k &= \mathbf{S}\mathbf{s}_{k-1} + \mathbf{d}_k & \textit{state equation} \\ y_k &= \mathbf{s}_k^\top \mathbf{g}(\tau_k) + v_k & \textit{observation equation} \end{cases} \tag{4}$$

where $\mathbf{s}_k = [s_{k-L}, s_{k-L+1}, ..., s_k]^\top$, $\mathbf{g}(\tau_k) = [g(LT + \tau_k), g((L-1)T + \tau_k), ..., g(\tau_k)]^\top$, $\mathbf{d} = [\mathbf{0}^\top, s_k]^\top$,

$$\mathbf{S} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_L \\ 0 & \mathbf{0}^\top \end{bmatrix}$$

$\mathbf{0}$ is an $L \times 1$ vector of zeros, and $\mathbf{I}_L$ is an $L \times L$ identity matrix.

Based on this dynamic state space model, our objective is to sequentially estimate symbols s_k and timing delays τ_k from the observations $y_{0:k}$, a collection of the observations from y_0 to y_k. Note from (2) that τ is nonlinear.

To form the solution, we adopt the maximum *a posteriori* (MAP) criterion, i.e.,

$$\{\hat{s}_k, \hat{\tau}_k\} = \arg\max_{s_k,\tau_k} \; p(s_k, \tau_k | y_{0:k}) \tag{5}$$

III. The Stochastic M-algorithm

In this paper, we consider a stochastic M-algorithm for obtaining the MAP solution.

A. The optimal solution

To calculate the MAP solution, the posterior distribution $p(\mathbf{s}_k, \tau_k | y_{0:k})$ is required, which can be obtained by marginalizing of the joint posterior distribution $p(\mathbf{s}_{0:k}, \tau_{0:k}|y_{0:k})$ over $\mathbf{s}_{0:k-1}$ and $\tau_{0:k-1}$, i.e.,

$$p(\mathbf{s}_k, \tau_k|y_{0:k}) = \sum_{\mathbf{s}_{0:k-1}} \int_{\tau_{0:k-1}} p(\mathbf{s}_{0:k}, \tau_{0:k}|y_{0:k}) d\tau_{0:k-1} \tag{6}$$

where $\mathbf{s}_{0:k}$ and $\tau_{0:k}$ are the respective collections of the symbols and the timing delays. However, no analytical expression of $p(\mathbf{s}_k, \tau_k|y_{0:k})$ is available. It is because that, firstly, the joint posterior distribution $p(\mathbf{s}_{0:k}, \tau_{0:k}|y_{0:k})$ is obtained by

$$\begin{aligned} & p(\mathbf{s}_{0:k}, \tau_{0:k}|y_{0:k}) \\ = \; & \frac{p(y_k|\mathbf{s}_k, \tau_k) p(\mathbf{s}_k, \tau_k|\mathbf{s}_{k-1}, \tau_{k-1})}{p(y_k|y_{0:k-1})} \\ & p(\mathbf{s}_{0:k-1}, \tau_{0:k-1}|y_{0:k-1}) \\ = \; & \frac{p(y_k|\mathbf{s}_k, \tau_k) p(\tau_k|\tau_{k-1}) p(\mathbf{s}_k|\mathbf{s}_{k-1})}{p(y_k|y_{0:k-1})} \cdot \\ & p(\mathbf{s}_{0:k-1}, \tau_{0:k-1}|y_{0:k-1}) \\ \propto \; & p(y_k|\mathbf{s}_k, \tau_k) p(\tau_k|\tau_{k-1}) p(\mathbf{s}_{0:k-1}, \tau_{0:k-1}|y_{0:k-1}) \end{aligned} \tag{7}$$

where the second equality is arrived since s_k is independent from τ_k. When k increases, the computation for (7) increases exponentially. Also, due to the nonlinearity of τ, the analytical expression of (7) cannot be obtained. Second, notice from (6) that the marginalization requires multiple summation and high dimensional integration. The multiple summation has a complexity increasing exponentially with k, and the high dimensional integration is analytically unattainable due to the nonlinearity in τ. As a result, the exact MAP solution is prohibited. We then resort to a random numerical method known as the stochastic M-algorithm.

B. The Stochastic M-algorithm

The stochastic M-algorithm is a very efficient particle filtering algorithm designed for problems with discrete variables. Like all particle filtering algorithms, the SMA approximates the joint posterior distribution $p(\mathbf{s}_{0:k}, \tau_{0:k}|y_{0:k})$ with a discrete random measure represented by a set of $2M$ samples, or trajectories, and weights, $\{\mathbf{s}_{0:k}^{(m)}, w_{0:k}^{(m)}\}_{m=1}^{2M}$, and the approximation can be expressed by

$$p(\mathbf{s}_{0:k}, \tau_{0:k}|y_{0:k}) \approx \sum_{m=1}^{2M} w_k^{(m)} \delta(\mathbf{s}_{0:k} - s_{0:k}^{(m)}) \delta(\tau_{0:k} - \tau_{0:k}^{(m)}) \tag{8}$$

where $\delta(\cdot)$ is the Dirac delta function. The MAP solution (5) can be then obtained by

$$\{\hat{s}_k, \hat{\tau}_k\} = \{s_k^{(J)}, \tau_k^{(J)}\} \tag{9}$$

where $J = \arg\max_{m\in\{1,\cdots,2M\}} w_k^{(m)}$.

To calculate the MAP solution, the weights and the sample trajectories are needed. To this end, the SMA follows the particle filtering framework using the idea of sequential importance sampling. To put the notion on firmer ground, let us assume that, at time $k-1$, we have obtained M trajectories $\{\mathbf{s}_{0:k-1}^{(m)}, \tau_{0:k-1}^{(m)}\}_{m=1}^{M}$ and the corresponding weights $\{w_{k-1}^{(m)}\}_{m=1}^{M}$ that approximate the joint posterior distribution $p(\mathbf{s}_{0:k-1}, \tau_{0:k-1}|y_{0:k-1})$. At time k, we produce totally $2M$ samples, two samples from each trajectory, and particularly for the two from the mth trajectory, we assign $s_k^{(2m-1)} = +1$ and $s_k^{(2m)} = -1$. In addition, we obtain $\tau_k^{(2m-1)}$ and $\tau_k^{(2m)}$ deterministically from the UKF, the detail of which will be discussed in section III-C. In the context of particle filtering

[6], the two samples can be considered from the following importance functions

$$\pi(\mathbf{s}_k, \tau_k|\mathbf{s}_{0:k-1}^{(m)}, \tau_{0:k-1}^{(m)}, y_{0:k}^{(m)}) = \delta_{s_k,1}\delta_{s_{k-1},s_{k-1}^{(m)}} \cdots \delta_{s_{k-L},s_{k-L}^{(m)}} \delta(\tau_k - \tau_k^{(2m-1)}) \quad (10)$$

and

$$\pi(\mathbf{s}_k, \tau_k|\mathbf{s}_{0:k-1}^{(m)}, \tau_{0:k-1}^{(m)}, y_{0:k}^{(m)}) = \delta_{s_k,-1}\delta_{s_{k-1},s_{k-1}^{(m)}} \cdots \delta_{s_{k-L},s_{k-L}^{(m)}} \delta(\tau_k - \tau_k^{(2m)}) \quad (11)$$

where $\delta_{i,j}$ is the Kronecker delta function. Note that the 'sampling' procedure is completely deterministic, and no random sampling is actually performed. By far, we have obtained $2M$ updated trajectories $\{\mathbf{s}_{0:k}^{(\bar{m})}, \tau_{0:k}^{(\bar{m})}\}_{\bar{m}=1}^{2M}$, where $\mathbf{s}_{0:k-1}^{(\bar{m})} = \mathbf{s}_{0:k-1}^{(m)}$ and $\tau_{0:k-1}^{(\bar{m})} = \tau_{0:k-1}^{(m)}$ for $\bar{m} = 2m-1$ and $2m$. The corresponding weights can be then calculated according to particle filtering as

$$w_k^{(\bar{m})} \propto w_{k-1}^{(\bar{m})} p(y_k|\mathbf{s}_k^{(\bar{m})}, \tau_k^{(\bar{m})}) p(\mathbf{s}_k^{(\bar{m})}|\mathbf{s}_{k-1}^{(\bar{m})}) p(\tau_k^{(\bar{m})}|\tau_{k-1}^{(\bar{m})}) \quad (12)$$

where $w_{k-1}^{(\bar{m})} = w_{k-1}^{(m)}$ for $\bar{m} = 2m-1$ and $2m$, $p(y_k|\mathbf{s}_k^{(\bar{m})}, \tau_k^{(\bar{m})}) \sim \mathcal{N}((\mathbf{s}_k^{(\bar{m})})^\top \cdot \mathbf{g}(\tau_k^{(\bar{m})}), N_0/2)$, $p(\mathbf{s}_k^{(\bar{m})}|\mathbf{s}_{k-1}^{(\bar{m})}) = 0.5$, and $p(\tau_k^{(\bar{m})}|\tau_{k-1}^{(\bar{m})}) \sim \mathcal{N}(a \cdot \tau_{k-1}^{(\bar{m})}, \sigma_u^2)$. It is necessary to note that the importance function does not appear in the weight calculation. It is because that it carries the same value for all the trajectories and therefore can be removed after normalization. The normalized weights are attained from

$$w_k^{(\bar{m})} = \frac{w_k^{(\bar{m})}}{\sum_{\bar{m}=1}^{2M} w_k^{(\bar{m})}} \quad (13)$$

With these $2M$ trajectories and the weights, we can obtain the MAP solution from (9).

Finally, a random selection scheme is performed to choose M from a total $2M$ trajectories. The selection is necessary since otherwise the number of trajectories will increase exponentially. We want to emphasize that the selection is a sampling-without-replacement process, i.e., no replicate of trajectories will be produced after selection. Thus, the popular resampling schemes such as the residual resampling are not suited. Here, we use the optimal resampling algorithm [7]. The optimal resampling is a sampling-without-replacement algorithm, and it is optimal in the sense that the mean square error between the original weights and the sampled weights is minimized.

The proposed SMA can be now summarized in the following chart

---

At the kth symbol time,

- **Trajectory expansion**
  For $m = 1$ to M
  - Set $s_k^{(2m-1)} = 1$ and calculate $\tau_k^{(2m-1)}$ by UKF based on $\mathbf{s}_k^{(2m-1)}$ and $\mathbf{s}_{0:k-1}^{(m)}$.
  - Set $s_k^{(2m)} = -1$ and calculate $\tau_k^{(2m)}$ by UKF based on $\mathbf{s}_k^{(2m)}$ and $\mathbf{s}_{0:k-1}^{(m)}$.
- Calculate the weights $w_k^{(\bar{m})}$ by (12) and normalize them according to (13).
- Obtain the MAP solution according to (9).
- **Trajectory selection**
  Select M trajectories from $2M$ trajectories using the optimal resampling algorithm.

We want to also point out that the SMA algorithm starts at $k = \lfloor log_2 M \rfloor$, where $\lfloor x \rfloor$ denotes the rounding-off operation on x. We have initially 2^k trajectories, each of which contains one possible combination of the first k symbols and an estimate of τ based on the symbol combination. The corresponding initial weights are the joint posterior distribution evalued at the initial samples. This measure coupled with the SMA algorithm guarantees that no two trajectories will be the same at any time k, which is however hardly true with particle filtering. This distinct feature implies that the SMA carries more diversity for a given M than the generic particle filtering. As a result, the SMA is more efficient. Further, since no random sampling is involved, the SMA has less complexity.

C. *Timing estimation using the unscented Kalman filter*

In this section, we briefly describe the steps for obtaining samples for τ_k. At time k, based on each trajectory, say $\{\mathbf{s}_{0:k-1}^{(m)}, \tau_{0:k-1}^{(m)}\}$ and $\mathbf{s}_k^{(2m)}$, our objective is to obtain an estimate and use it as the sample for τ_k. The problem is equivalent to the parameter estimation in a nonlinear DSSM, and we obtain the estimate using the UKF [8]. For convenience of composition, we drop the superscript $^{(\bar{m})}$ in the proceeding discussion.

In the UKF, the analytical intractable posterior distribution of τ_k is approximated by a set of appropriately weighted sigma points obtained through the unscented transformation. The advantage is that the use of sigma points permits the accurate propagation of the mean and the variance of τ_k through the nonlinear equation (4) in a computationally efficient manner. Refer to [9], [10] for more detailed discussion on the topic.

The estimation of τ_k by the UKF is also proceeded in an iterative fashion. At time k, base on the estimate $\bar{\tau}_{k-1}$ and its variance P_{k-1} obtained at time $k-1$, we first define a vector $\mathcal{X}_{k-1} = [\ \bar{\tau}_{k-1} \ \ \bar{\tau}_{k-1} + \gamma\sqrt{P_{k-1}} \ \ \bar{\tau}_{k-1} - \gamma\sqrt{P_{k-1}}\ ]$ where γ is a pre-defined constant by the UKF. We then perform a prediction step on τ_k by

$$\bar{\tau}_{k|k-1} = \sum_{i=1}^{3} W_i^{(mean)} \mathcal{X}_{i,k|k-1}$$

and

$$P_{k|k-1} = \sum_{i=1}^{3} W_i^{(mean)} \left(\mathcal{X}_{i,k|k-1} - \bar{\tau}_{k|k-1}\right)^2 + \sigma_u^2$$

where $\mathcal{X}_{i,k|k-1} = a\mathcal{X}_{i,k-1}$ with $\mathcal{X}_{i,k-1}$ being the ith element of $\mathcal{X}_{k-1}$, and $W_i^{(mean)}$s are the parameters defined in the UKF. The estimate of τ can thus be obtained from the measurement steps as

$$\bar{\tau}_k = \bar{\tau}_{k|k-1} + \mathcal{K}_k(y_k - \bar{y}_{k|k-1}) \quad (14)$$

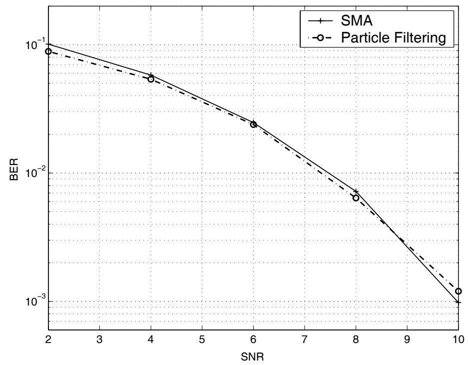

Fig. 1. BERs as functions of SNR

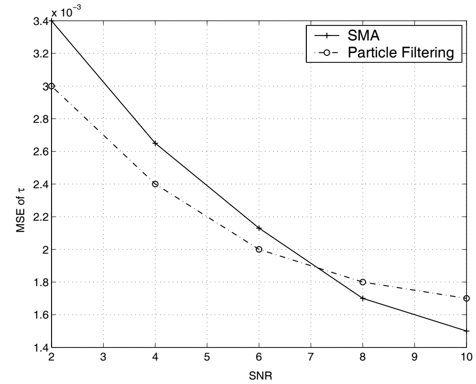

Fig. 2. MSEs as functions of SNR

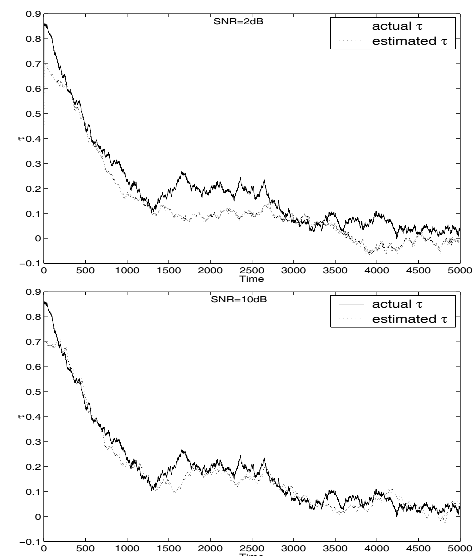

Fig. 3. Actual τ and estimated τ for two different SNRs

where $\mathcal{K}_k = P_{\tau_k y_k} P^{-1}_{y_k y_k}$, $P_{y_k y_k} = \sum_{i=1}^{3} W_i^{(cov)} \big(\mathcal{Y}_{i,k|k-1} - \bar{y}_{k|k-1}\big)^2 + N_0/2$, $P_{\tau_k y_k} = \sum_{i=1}^{3} W_i^{(cov)} \big(\mathcal{X}_{i,k|k-1} - \bar{\tau}_{k|k-1}\big)^2$, $\bar{y}_{k|k-1} = \sum_{i=1}^{3} W_i^{(mean)} \mathbf{s}_k^\top \mathbf{g}(\mathcal{X}_{i,k|k-1})$, and $W_i^{(cov)}$s are again parameters defined in the UKF.

IV. Simulation Results

We demonstrate in this section the performance of the proposed SMA through simulations. In the simulation, a system with BPSK modulation, ISI span of $L+1=3$, and the raised cosine pulse with roll-off factor of 0.7 was used. In addition, in the time delay AR model, the coefficient a was set to be 0.999, and σ_u^2 is 10^{-5}.

We compared the performance of the SMA with $M = 2$ with particle filtering with 50 samples [3]. In particle filtering, τ_ks and s_ks are attained according to the minimum mean square error criterion. Figure 1 demonstrates the symbol detection performance of the two algorithms using the bit error rates (BER) at signal-to-noise (SNR) from 2 dB to 10 dB. It is obvious that the SMA performs closely with particle filtering and even a little better at high SNR. Note that the complexity of the SMA is less 1/25 of particle filtering.

We also show in Figure 2 the mean squared error (MSE) of the estimates obtained by the two algorithms. Similar relationship as in the symbol detection performance is observed. Again, SMA can achieve smaller MSE in high SNR region.

Finally, we depict in Figure 3 the tracking ability of SMA on τ at two different SNRs. We see that the tracking performance improves with the increase of SNR, and at SNR=10 dB, the SMA can track symbol timing accurately.

V. Conclusions

In this paper, we proposed to solve joint symbol detection and timing estimation with the stochastic M-algorithm. Our simulation results showed that the stochastic M-algorithm attains the performance of particle filtering, but with significantly reduced complexity.

References

[1] T. J. Lim and L. K. Rasmussen, "Adaptive symbol and parameter estimation in asynchronous multiuser cdma detectors," *IEEE Trans. Commun.*, vol. 45, pp. 213–220, Jan. 1997.

[2] J. Caffery and G. L. Stüber, "Nonlinear multiuser parameter estimation and tracking in CDMA systems," *IEEE Trans on Commun.*, vol. 48, no. 12, pp. 2053–2063, Dec 2000.

[3] T. Ghirmai, M. F. Bugallo, J. Migues, and P. M. Djurić, "Joint symbol detection and timing estimation using particle filtering," *in proceedings of ICASSP*, vol. IV, pp. 596–599, 2003.

[4] Y. Huang, J. Zhang, and P. M. Djurić, "Bayesian detection for BLAST," *accepted for publication in IEEE Transactions on Signal Processing*, 2004.

[5] R. Iltis, "Joint estimation of PN code delay and multipath using the extended Kalman filter," *IEEE Transactions on Communications*, vol. 38, pp. 1677–1685, Oct. 1990.

[6] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10(3), pp. 197–208, 2000.

[7] P. Fearnhead, "Sequential Monte Carlo method in filter theory," *Ph.D. thesis, Oxford University*, 1998.

[8] S. J. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems," *Proc. American Control Conf.*, pp. 1628–32, 1995.

[9] S. Haykin, *Adaptive Filter Theory*, 4th ed. Prentice Hall, 2002.

[10] E. Wan and R. V. D. Merwe, *The unscented Kalman filter, in Kalman Filtering and Neural Networks, Ed. S. Haykin*. Wiley: New York, 2001, ch. 7.