

Term Paper Topics

A. Proportionate and Affine Adaptation Algorithms

Several variants of the LMS algorithm have been proposed recently to deal with sparse echo cancellation paths or sparse communication channels. Such channels include zero as well as nonzero taps. To deal with the wide range of values in channel taps, Duttweiler [1] proposed the proportional normalized LMS (PNLMS) algorithm which adapts the step size of each individual channel tap. Variations combining this algorithm with affine adaptation algorithms [4] were proposed in Chapter 2 of [2]. A term paper could focus on the properties of PNLMS or affine adaptation algorithms or their application to echo cancellation and channel equalization.

1. D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancellers," *IEEE Trans. Speech Audio Proc.*, vol. 8, pp. 508–517, Sep. 2000.
2. J. Benesty and T. Gänslér and D. R. Morgan and M. M. Sondhi and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*, Springer Verlag, New York, 2001.
3. Z. Chen and S. L. Gay and S. Haykin, "Proportionate adaptation: new paradigms in adaptive filters," in *Least-Mean-Square Adaptive Filters*, Chap. 8, pp. 293–334, S. Haykin and B. Widrow, eds., Wiley-Interscience, New York, 2003.
4. K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electronics Commun. Japan*, vol. 67-A, pp. 19–27, 1984.
5. S. L. Gay, "Affine projection algorithms" in *Least-Mean-Square Adaptive Filters*, Chap. 7, pp. 241–291, S. Haykin and B. Widrow, eds., Wiley-Interscience, New York, 2003.
6. H.-C. Shin and A. H. Sayed, "Mean-square performance of a family of affine projection algorithms," *IEEE Trans. Signal Proc.*, vol. 52, pp. 90–102, Jan. 2004.

B. Adaptive Maximum Likelihood/Infomax Algorithm

The stochastic gradient algorithm described in class leads to a simple adaptive maximum likelihood (ML) estimation procedure, which is often called the Infomax algorithm. Let Y_t with $t \geq 0$ denote a stationary sequence of random vectors with probability density $f(y_t|\theta)$, where θ denotes an unknown parameter vector which needs to be estimated. Let θ_0 denote

the true parameter vector. The Infomax algorithm minimizes adaptively the Kullback-Leibler divergence

$$\begin{aligned} K(\theta_0|\theta) &= D(f(y_t|\theta_0), f(y_t|\theta)) \\ &= \int \ln \left(\frac{f(y_t|\theta_0)}{f(y_t|\theta)} \right) f(y_t|\theta_0) dy_t \end{aligned} \quad (1)$$

of the true density compared to the density $f(y_t|\theta)$. Since $K(\theta_0|\theta) \geq 0$ with equality whenever $\theta = \theta_0$, as long as the initial estimate is close enough to the correct value θ_0 , the adaptive minimization will converge to θ_0 . Note that since the only term of (1) that depends on θ is

$$E_{\theta_0}[-\ln f(y_t|\theta)] = - \int \ln(f(y_t|\theta)) f(y_t|\theta_0) dy_t$$

The minimization of $K(\theta_0|\theta)$ is equivalent to the minimization of the expected value of $-\ln f(y_t|\theta)$ or equivalently the maximization of the expected value of the log-likelihood function $\ln f(y_t|\theta)$ where the expectation is taken with respect to the true density. This explains why the method represents an adaptive ML estimation algorithm.

By replacing the gradient of $K(\theta_0|\hat{\theta}_t)$ by its instantaneous value, like for all stochastic gradient algorithms, the adaptive ML estimation algorithm takes the form

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \mu \nabla_{\hat{\theta}_t} \ln f(y_t|\hat{\theta}_t) \quad (2)$$

Applications of the Infomax algorithm include the blind source separation problem described in [1,2]. The LMS algorithm and constant modulus blind equalization algorithm [3-6] can also be viewed as special cases of this algorithm (in the LMS case the density is Gaussian, whereas for the constant modulus algorithm it is a sub-Gaussian density). See also [7] for a recent application to blind nonlinear noise cancellation. A term paper could focus on a discussion and analysis of the Infomax algorithm and its applications.

1. A. J. Bell and T.J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution, *Neural Computation*, vol. 7, pp. 1129–1159, 1995.
2. J.-F. Cardoso, "Infomax and maximum likelihood for blind source separation," *Signal Processing Letters*, vol. 4, pp. 112–114, April. 1997.
3. D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Communications*, vol. 28, pp. 1867–1875, Nov. 1980.
4. A. Benveniste and G. Goursat, "Blind equalizers," *IEEE Trans. Communications*, vol. 32, pp. 871–883, Aug. 1984.
5. J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans Acoust., Speech, Signal Proc.*, vol. 31, pp. 459–472, Apr. 1983.
6. Z. Ding and Y. G. Li, *Blind Equalization and Identification*, Marcel Dekker, NY, 2001.
7. N. C.-J. Chang, B. C. Levy, S. H. Lewis, and P. J. Hurst, "Blind substrate noise cancellation for single step analog-to-digital converters," submitted for publication.

C. Convergence and Performance Analysis of Adaptive Filters

Two general approaches have been developed for the analysis of linear and nonlinear adaptive algorithms. The first approach described in [1,2] is based on an averaging viewpoint and consists of describing the evolution of a given adaptive algorithm in terms of the solution of a corresponding ordinary differential equation (ODE). This method is therefore also called the ODE method. The second technique was proposed more recently and is described in detail in chapters 6-9 of [6]. It relies on an energy conservation relation which holds for a large class of adaptive algorithms. Both methods of analysis have advantages and disadvantages. The averaging method requires that the adaptation step size should be very small, but the conditions imposed on the observation data and the process to be estimated are very weak, so that this method is extremely general and applies to a wide variety of situations. In contrast, the energy conservation method does not require that the adaptation step size should be small, but it requires an independence assumption for the observed data which is unrealistic in many cases. Nevertheless, both methods yield results which are usually consistent with each other. Typically, both methods establish the convergence of the estimated parameters to the correct model parameters for vanishingly small noise. They also describe the error variance, i.e., the performance of adaptive algorithms, as a function of noise and of the step size in steady state. Finally, they characterize the transient behavior, i.e., the speed of convergence, of adaptive algorithms to their steady state. A term paper could focus on either convergence analysis technique or a comparison of the two methods. The papers [2-5] represent a case study of the averaging approach for the analysis of blind equalization algorithms, whereas [7-12] characterize the properties of the energy conservation technique for either LMS or blind equalization algorithms. A nice project would be to perform an ODE analysis of the generalization of the constant modulus algorithm for continuous-phase modulated signals described in [13].

1. A. Benveniste, M. Metivier, and P. Priouret, *Adaptive Algorithms and Stochastic Approximations*, Springer-Verlag, Berlin, 1990.
2. V. Solo and X. Kong, *Adaptive Signal Processing Algorithms— Stability and performance*, Prentice-Hall, Upper Saddle River, NJ, 1985.
3. A. Benveniste and M. Goursat, “Blind equalizers,” *IEEE Trans. Commun.*, vol. 32, pp. 871–883, Aug. 1984.
4. A. Benveniste, M. Goursat, and G. Ruget, “Blind identification of a nonminimum phase system: Blind adjustment of a linear equalizer in data communications,” *IEEE Trans. Automat. Control*, vol. 25, pp. 385–399, June 1980.
5. I Fijalkow, C. E. Manlove, and C. R. Johnson, Jr., “Adaptive fractionally spaced blind CMA equalization: Excess MSE,” *IEEE Trans. Signal Proc.*, vol. 46, pp. 227–231, Jan 1998.
6. A. H. Sayed, *Fundamentals of Adaptive Filtering*, Wiley, New York, 2003.
7. M. Rupp and A. H. Sayed, “A time-domain feedback analysis of filtered-error adaptive gradient algorithms,” *IEEE Trans. Signal Proc.*, vol. 44, pp. 1428–1439, June 1996.

8. N. R. Yousef and A. H. Sayed, "A unified approach to the steady-state and tracking analyses of adaptive filters," *IEEE Trans. Signal Proc.*, vol. 49, pp. 314–324, Feb. 2001.
9. J. Mai and A. H. Sayed, "A feedback approach to the steady-state performance of fractionally-spaced blind adaptive equalizers," *IEEE Trans. Signal Proc.*, vol. 48, pp. 80–91, Jan. 2000.
10. N. R. Yousef and A. H. Sayed, "A unified approach to the steady-state and tracking analyses of adaptive filters," *IEEE Trans. Signal Proc.*, vol. 49, pp. 314–324, Feb. 2001.
11. V. H. Nascimento and A. H. Sayed, "on the learning mechanism of adaptive filters," *IEEE Trans. Signal Proc.*, vol. 48, pp. 1609–1625, June 2000.
12. A. H. Sayed and V. Nascimento, "Energy conservation and the learning ability of LMS adaptive filters," in *Least-Mean-Square Adaptive Filters*, Chap. 3, pp. 79–104, S. Haykin and B. Widrow, eds., Wiley-Interscience, New York, 2003.
13. P. Bianchi and P. Loubaton, "On the blind equalization of continuous phase modulated signals using the constant modulus criterion," *IEEE Trans. Signal Processing*, vol. 55, pp. 1047–1061, March 2007.

D. Natural Gradient Adaptive Algorithms

It was shown in class that quasi-Newton algorithms converge much faster than gradient iterations. It is therefore desirable to develop adaptive algorithms which automatically find the best scaling or coordinate system for optimizing gradient searches. A class of algorithms which achieves this objective are the "natural gradient" algorithms proposed by Amari [1,2] and further examined in [3]. In this approach, the adaptation rule uses an appropriately selected Riemannian metric to scale the gradient locally so as to maximize the rate of decay of the objective function. This class of algorithms is closely related to relative gradient adaptation rules introduced in [4,5] for source separation. See also [6]. A term paper could focus on either natural gradient adaptation rules, or on their application to blind source separation problems.

1. S.-I. Amari, "Natural gradient works efficiently in learning," *Neural Computation*, vol. 10, pp. 251–276, 1998.
2. S. Amari and S. C. Douglas, "Why natural gradient?," in *Proc. 1998 Int. Conf. Acoustics, Speech, Signal Proc. (ICASSP 98)*, vol. 2, pp. 1213–1216, May 1998.
3. T. K. Moon and J. H. Gunther, "Contravariant adaptation on structured matrix spaces," *Signal Processing*, vol. 82, pp. 1389–1410, 2002.
4. J.-F. Cardoso and B. H. Laheld, "Equivariant source separation," *IEEE Trans. Signal Processing*, vol. 44, pp. 3017–3030, Dec. 1996.

5. J.-F. Cardoso, "Blind signal separation: statistical principles," *Proc. IEEE*, vol. 86, pp. 2009-2025, Oct. 1998.
6. S. Amari, "Superefficiency in blind source separation", *IEEE Trans. Signal Processing*, vol. 47, pp. 936-944, Apr. 1999.

E. Robust State-Space Filtering based on Model Perturbation Bounds

In deriving the Kalman filter, it is assumed that the state dynamics and observations are assumed to be known precisely. Unfortunately, in most situations the state-space model is known only approximately. In such situations, the estimates produced by ordinary Kalman filters based on the nominal model may be less accurate than those generated by more conservative filters that take the possible existence of modelling errors into account. This concern has prompted the development of state-space filtering methods which expressly take into account the possibility that the models of dynamics and observations may be imprecise. The first class of robust filtering methods was proposed in the early 1970s by Bertsekas et al. [1], Schweppe [2], as well as by Kurzhansky and other researchers [3] in Russia, for estimating the state variables of dynamic models corrupted by unknown disturbances and noises. The key feature of this approach is the assumption that the noises or disturbances describing the unmodelled state dynamics are completely unknown, but *bounded*. Under this assumption, the state is confined inside an ellipsoidal set of minimum size, where the center of the ellipsoid can be viewed as the state estimate, and its orientation and principal axes provide information equivalent to the state covariance matrix. The propagation of the ellipsoid of confidence can then be accomplished through Kalman filtering-like recursions, so that even though the state estimation formulation is purely deterministic, its solution is similar to the usual Kalman filter. While the original research on robust state-space filtering assumed that the nominal model was perturbed only by unknown additive signals and noises, in recent years, Petersen and Savkin [4,5], Sayed [6], and El Ghaoui et al. [7] have proposed various methods for handling the case when the true model includes both perturbations to the model dynamics and additive unknown signals.

A term paper on robust filtering based on model perturbation bounds could focus either on a description of the algorithms presented in [1]-[6], or on simulations for a well selected application.

1. D. P. Bertsekas and I. Rhodes, "Recursive state estimation for a set membership description of uncertainty," *IEEE Trans. Automat. Control*, *IEEE Trans. Automat. Control*, vol. 16, pp. 117-128, 1971.
2. F. C. Schweppe, *Uncertain Dynamic Systems*, Prentice Hall, Englewood Cliffs, NJ, 1973.
3. A. Kurzhanski and I Valyi, *Ellipsoidal Calculus for Estimation and Control*, Birkhauser, Boston, MA, 1997
4. A. V. Savkin and I. R. Petersen, "Robust state estimation and model validation for discrete-time uncertain systems with a deterministic description of noise and uncertainty," *Automatica*, vol. 34, pp. 271-274, 1998.

5. I. R. Petersen and A. V. Savkin, *Robust Kalman Filtering for Signals and Systems with Large Uncertainties*, Birkhauser, Boston, MA, 1999.
6. A. H. Sayed, "A framework for state-space estimation with uncertain models," *IEEE Trans. Automat. Control*, vol. 46, pp. 998-1013, July 2001.
7. L. El Ghaoui and G. Calafiore, "Robust filtering for discrete-time systems with bounded noise and parametric uncertainty," *IEEE Trans. Automat. Control*, vol. 46, pp. 1084-1089, July 2001.

F. Risk-sensitive and Robust Filtering

Unlike ellipsoid of confidence filters which were designed explicitly to deal with signal and dynamic perturbations, in their original incarnation [1-3], risk sensitive filters tackled the robustness problem only indirectly. Specifically, since these filters minimize an exponential of quadratic loss function for the filtering error, they penalize very severely large estimation errors for the nominal state-space model. As such, there is nothing robust in these filters, since they do not consider modelling errors. However, it was realized by Boel et al. [4] and later by Levy and Nikoukhah [9] that risk-sensitive filters arise as the solution of a minimax problem consisting of finding the best least-squares filter for the least favorable model located in a neighborhood of the nominal model. In this approach, the neighborhood is specified by a Kullback-Leibler divergence tolerance. Note that although it is not a conventional distance, the Kullback-Leibler (KL) divergence is a standard criterion used in statistics for fitting statistical models. It turns out that the solution of the minimax filtering problem with a KL tolerance is a risk sensitive filter. The KL tolerance can be imposed in several ways, either for the overall statistical model for the entire filtering interval [5-9], or incrementally [10]. In addition to deriving the robust risk-sensitive filter, this minimax approach also allows the explicit construction of a corresponding least-favorable model [8,10] which can be used to compare the robust filter to other filters, such as the conventional Kalman filter.

A term paper could focus on a comparison of the martingale derivation of the robust filter described in [6]-[8] to the game theoretic approach of [10]. Other issues such as filter convergence, or a comparison of robust filters with or without commitment [6,7,10] could be examined as part of a project.

1. J. L. Speyer, J. Deyst, and D. H. Jacobson, "Optimization of stochastic linear systems with additive measurement and process noise using exponential performance criteria," *IEEE Trans. Automat. Control*, vol. 19, pp. 358-366, 1974.
2. P. Whittle, *Risk-sensitive Optimal Control*, J. Wiley, New-York, 1990.
3. J. L. Speyer and W. H. Chung, *Stochastic Processes, Estimation and Control*, "Soc. Indust. Applied Math., Philadelphia, PA, 2008.
4. R. K. Boel, M. James, and I. R. Petersen, "Robustness and risk-sensitive filtering," *IEEE Trans. Automat. Control*, vol. 47, pp. 451-461, Mar. 2002.

5. M.-G. Yoon, V. A. Ugrinovskii, and I. R. Petersen, "Robust finite horizon minimax filtering for discrete-time stochastic uncertain systems," *System & Control Letters*, vol. 52, pp. 99–112, 2004.
6. L. P. Hansen and T. J. Sargent, "Robust estimation and control under commitment," *J. Economic Theory*, vol. 124, pp. 258–301, 2005.
7. L. P. Hansen and T. J. Sargent, "Recursive robust estimation and control without commitment," *J. Economic Theory*, vol. 136, pp. 1–27, 2007.
8. L. P. Hansen and T. J. Sargent, *Robustness*, Princeton University Press, Princeton, NJ, 2008.
9. B. C. Levy and R. Nikoukhah, "Robust least-squares estimation with a relative entropy constraint," *IEEE Trans. Informat. Theory*, vol. 50, pp. 89–104, Jan. 2004.
10. B. C. Levy and R. Nikoukhah, "Robust state-space filtering with an incremental relative entropy tolerance," preprint, available at <http://arxiv.org/pdf/10004.2519.pdf>, April 2010.

G. Nonlinear Kalman Filtering

Consider a discrete-time Markov process X_t defined for $t \geq 0$ with transition density $\phi_t(x_{t+1}|x_t)$ and initial density $f_0(x_0)$. This process is partially observed through some observations Y_t , $t \geq 0$, where conditioned on $X_t = x_t$, the density of Y_t is given by $\theta_t(y_t|x_t)$. Let \mathbf{Y}_0^t denote the vector formed by regrouping all observations up to time t . Then the conditional filtered and predicted densities $f_t(x_t|\mathbf{Y}_0^t)$ and $f_t(x_t|\mathbf{Y}_0^{t-1})$ of X_t based on all observations up to time t and $t-1$, respectively, can be propagated by employing in alternance the following two recursions. The measurement update recursion is given by

$$f_t(x_t|\mathbf{Y}_0^t) = \theta_t(y_t|x_t)f_t(x_t|\mathbf{Y}_0^{t-1})/\gamma_t(y_t|\mathbf{Y}_0^{t-1}) \quad (3)$$

where

$$\gamma_t(y_t|\mathbf{Y}_0^{t-1}) = \int \theta_t(y_t|x_t)f_t(x_t|\mathbf{Y}_0^{t-1})dx_t$$

denotes the conditional density of observation Y_t given the past observations \mathbf{Y}_0^{t-1} up to time $t-1$. The time update recursion is given by

$$f_{t+1}(x_{t+1}|\mathbf{Y}_0^t) = \int \phi_t(x_{t+1}|x_t)f_t(x_t|\mathbf{Y}_0^t)dx_t. \quad (4)$$

The Bayesian recursions (3) and (4) depend exclusively on the Markov structure of the process X_t and on the fact that the observations Y_t depend only on the state X_t at time t . Thus they remain valid even if the system dynamics are nonlinear and the noises affecting the dynamics are non-Gaussian. The standard Kalman filter relies on the fact that when the dynamics are linear and process and measurement noises are Gaussian, then all the

densities appearing in recursions (3) and (4) remain Gaussian (recall that the class of Gaussian densities is stable under conditioning and marginalization). Accordingly

$$\begin{aligned} f_t(x_t|\mathbf{Y}_0^t) &\sim \mathcal{N}(\hat{X}_{t|t}, P_{t|t}) \\ f_t(x_t|\mathbf{Y}_0^{t-1}) &\sim \mathcal{N}(\hat{X}_{t|t-1}, P_{t|t-1}) \end{aligned}$$

are Gaussian distributed, so that instead of having to propagate the functions $f_t(x_t|\mathbf{Y}_0^t)$ and $f_t(x_t|\mathbf{Y}_0^{t-1})$ through recursions (3) and (4), we only need to propagate the mean vectors $\hat{X}_{t|t}$, $\hat{X}_{t|t-1}$ and covariance matrices $P_{t|t}$ and $P_{t|t-1}$. In other words the Kalman filter exploits *both* the Markov structure of the model, and its linear Gaussian property, which allows all conditional densities to be *finitely parameterized* by their mean vector and covariance matrix. Another case for which the conditional distributions of X_t given the past observations up to time t or up to time $t - 1$ are finitely parametrized is when X_t is described by a finite-state Markov chain. Let n denote the number of states. If the states are denoted as $1 \leq i \leq n$, the conditional distributions $f_t(x_t|\mathbf{Y}_0^t)$ and $f_t(x_t|\mathbf{Y}_0^{t-1})$ can be represented as n -dimensional vectors, and the integrations appearing in recursions (3) and (4) can be replaced by summations over the n possible values of x_t . So in the finite state Markov chain case, the recursions (3) and (4) admit also a simple implementation.

In the general nonlinear or non-Gaussian case, it is necessary to propagate the full conditional densities $f_t(x_t|\mathbf{Y}_0^t)$ and $f_t(x_t|\mathbf{Y}_0^{t-1})$. The general nonlinear filtering problem is discussed in detail in [1]. For the continuous-time case, a partial differential equation analogous to (3) and (4) was derived by the Russian mathematician Stratonovitch and by Kushner in [2,3] for the conditional density. It was subsequently shown by Zakai [4] that this equation can be simplified if we elect to propagate an unnormalized density, which is analogous to skipping the rescaling operation by $\gamma_t(y_t|\mathbf{Y}_0^{t-1})$ in (3). Since propagating the conditional density of the state given the past observations could not be realistically considered in the 1960s, researchers examined approximations, such as the extended Kalman filter (EKF). This filter, which was first proposed by Stanley Schmidt, linearizes the system dynamics about the current estimate and applies the Kalman filter to the linearized system. It is described in detail in [1] and [5] (see also [13]). This procedure works properly only when deviations of the actual trajectory compared to the nominal trajectory are small, and detailed conditions for the EKF convergence are presented in [6]. More recently another approximation technique called the unscented Kalman filter (UKF) was proposed by Julier, Uhlmann and other researchers in [7],[8]. Another approach to the nonlinear filtering problem which has become popular over the last 15 years relies on Monte Carlo methods. The corresponding nonlinear filters are called particle Kalman filters. In this approach, the integrals needed to evaluate the conditional mean and error variances for the densities obeying (3) and (4) are evaluated by Monte Carlo simulations involving the simultaneous generation of multiple system trajectories. Several methods of this type are described in [9]-[12]. Note that to limit the number of samples required, importance sampling techniques play a major role in particle filters [12].

Term paper topics related to nonlinear filtering would include: (i) a comparison of EKF and UKF filters, their properties and limitations; or (ii) a discussion of particle filters, their implementation, and application, to say, target tracking.

1. A. H. Jaswinski, *Stochastic Processes and Filtering Theory*, Academic Press, San

- Diego, CA, 1970; reprinted by Dover Publ., 2007.
2. R. L. Stratonovich, "Conditional Markov processes," *Theor. Probability Appl.*, vol. 5, pp. 156–178, 1960.
 3. H. J. Kushner, "Dynamical equations for optimal nonlinear filtering," *J. Diff. Equations*, col. 3, pp. 179–190, 1967.
 4. M. Zakai, "On the optimal filtering of diffusion processes," *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, vol. 11, pp. 230–243, 1969.
 5. A. Gelb, editor, *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1974.
 6. A. J. Krener, "The convergence of the extended Kalman filter," in *Directions in Mathematical Systems Theory and Optimization*, A. Rantzer and C. I. Byrnes, editors, Lecture Notes in Computer and Information Sciences no. 286, pp. 173–182, Springer Verlag, Berlin, 2003.
 7. S. Julier, J. Uhlmann, and H. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. Automat. Control*, vol. 45, pp. 477–482, March 2000.
 8. S. L. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, pp. 401–422, March 2004.
 9. G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian nonlinear state-space models," *J. Computational Graphical Statistics*, vol. 5, pp. 1–25, Mar. 1996.
 10. M. S. Arulampalam, S. Maskell, N. Gordon, T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Proc.*, vol. 50, pp. 174–188, Feb. 2002.
 11. B. Ristic, S. Arulampalam and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, Artech House, Norwell, MA, 2004.
 12. O. Cappè, S. J. Godsill, and E. Moulines, "An overview of existing methods and recent advances in sequential Monte Carlo," *Proc. IEEE*, vol. 95, pp. 899–923, May 2007.
 13. D. Simon, *Optimal State Estimation: Kalman, H^∞ , and Nonlinear Approaches*, Wiley-Interscience, Hoboken, NJ, 2006.