Phase Drift Estimation and Symbol Detection in Digital Communications: A Stochastic Recursive Filtering Approach

P. Pedrosa[(1,2)], R. Dinis[(1,3)], F. Nunes[(1,2)], and J. Bioucas-Dias[(1,2)]
[(1)]Instituto de Telecomunicações, 1049-001, Lisboa, Portugal
[(2)]Instituto Superior Técnico, Universidade Técnica de Lisboa, 1049-001 Lisboa, Portugal
[(3)]Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal
e-mail:{ppedrosa, rdinis, nunes, bioucas}@lx.it.pt

***Abstract*—This paper proposes a novel Bayesian stochastic filtering approach for the simultaneous phase drift estimation and symbol detection in digital communications. The posterior density of the phase drift is propagated in a recursive fashion by implementing a prediction and a filtering step in each iteration. The prediction step is supported on a random walk model playing the role of prior for the phase drift process; the filtering step is supported on a Gaussian sum approximation for the probability density of the current observation, *i.e.*, the so-called *sensor factor*. The Gaussian sum approximation turns out to be the key element allowing to derive a fast and efficient stochastic filter, which otherwise would be very hard to compute. The detection of the digital symbols is then carried out based on the inferred statistics of the phase drift. The effectiveness of the proposed method is illustrated for BPSK signals in the presence of strong phase drift.**

***Index Terms*—Stochastic recursive filtering; Gaussian sum filter; phase drift; state estimation; burst communications.**

I. Introduction

In digital communications, the lack of phase synchronization between the receiver oscillator and the received signal often renders communication systems useless. Estimating the phase drift, which includes phase noise due to oscillators instabilities and frequency offset caused by Doppler effect and/or poor frequency alignment between oscillators, is of utmost relevance.

Under a Bayesian perspective, the phase drift estimation from the observation of the channel output amounts to determining the posterior probability density function (pdf) of the state (*i.e.*, the current phase drift) conditioned on all measurement data, thus providing the means to compute an optimal estimate with respect to any criterion, *e.g.*, minimum mean-squared error (MMSE). The determination of the posterior pdf is usually extremely difficult. A remarkable exception occurs when the state and the measurement equations are linear and the noise is additive and Gaussian distributed. In this case, the posterior pdf is efficiently computed by the well known Kalman filter.

In this paper, we introduce a stochastic recursive filter to propagate the posterior pdf based on two central features:

1) the phase drift is modeled by a random walk model playing the role of prior.
2) the sensor factor is approximated by a sum of Gaussian functions.

These two features allow to exploit well known properties of Gaussian functions to derive an efficient and effective stochastic filter, in a vein similar to that of the Gaussian sum filter introduced in [1]. The detection of the digital symbols is then carried out based on the inferred statistics of the phase drift.

To illustrate the effectiveness of the proposed filter, we compare the performance results obtained when modeling the pdf by a weighted sum of Gaussian functions with that achieved using a single Gaussian. We assume perfect channel estimation and M-phase shift-keying (M-PSK) modulation with perfect symbol synchronization.

The paper is organized in the following manner. Section II establishes the models for the channel output and for the phase drift. Section III defines the Gaussian sum filter, with subsections dedicated to the Gaussian sum approximation, the Gaussian fitting of the sensor factor, the evaluation of the filtering and prediction densities, the filter initialization, and the implementation aspects of the algorithm. Section IV is devoted to the symbol detection, while Section V presents the performance results and Sec. VI concludes the paper.

II. System Model

Consider the transmission of the BPSK symbols $\{s_n = e^{j\phi_n}; n = 0,\ldots,N-1\}$, where $\phi_n \in \{0,\pi\}$, over an additive white Gaussian noise (AWGN) in the presence of phase drift. The received signal at the detector output is given by

$$y_n = e^{j(\phi_n+\theta_n)} + v_n; \qquad n = 0,\ldots,N-1. \tag{1}$$

In (1), $\{v_n; n = 0,\ldots,N-1\}$ is a sequence of complex zero-mean white Gaussian noise random variables with uncorrelated real and imaginary components each one with variance $\sigma_v^2 = N_0/(2E_b)$, where E_b stands for the bit energy.

The phase drift θ_n is modeled by a random walk process described by the linear stochastic difference equation

$$\theta_n = \theta_{n-1} + w_n; \qquad n = 0,\ldots,N-1, \tag{2}$$

where $\{w_n; n = 0,\ldots,N-1\}$ is a sequence of real zero-mean independent Gaussian random variables with variance σ_w^2.

The noise sequences $\{v_n; n = 0, \ldots, N-1\}$ and $\{w_n; n = 0, \ldots, N-1\}$ are assumed to be mutually independent and independent of the initial state θ_{-1}. The pdfs of v_n and w_n are denoted by p_v and p_w, respectively.

A. Optimal Symbol Detection

In our setup, we aim at detecting the nth symbol based on the set of observations $Y_n = [y_0, y_1, \ldots, y_n]$. According to the Bayesian paradigm, the posterior density $p(\phi_n|Y_n)$ summarizes all the statistical information about ϕ_n, given the observed data Y_n. In this work, we compute the *maximum a posteriori probability* (MAP) estimate, which minimizes the probability of error.

The MAP estimate of ϕ_n is given by

$$\begin{aligned} \widehat{\phi}_n &\in \arg \max_{\phi_n \in \{0,\pi\}} p(\phi_n|Y_n) \\ &= \arg \max_{\phi_n \in \{0,\pi\}} p(\phi_n, Y_n). \end{aligned} \tag{3}$$

The joint density $p(\phi_n, Y_n)$ can be obtained as

$$\begin{aligned} p(\phi_n, Y_n) &= \int_{-\infty}^{+\infty} p(y_n, \phi_n, \theta_n, Y_{n-1}) \, \mathrm{d}\theta_n \\ &= p(\phi_n) \int_{-\infty}^{+\infty} p(y_n|\phi_n, \theta_n) p(\theta_n|Y_{n-1}) p(Y_{n-1}) \, \mathrm{d}\theta_n, \end{aligned} \tag{4}$$

where we have used repeatedly the Bayes law and the facts that $p(y_n|\phi_n, \theta_n, Y_{n-1}) = p(y_n|\phi_n, \theta_n)$ and ϕ_n is independent of (θ_n, Y_{n-1}). Therefore, the main ingredients to compute the MAP estimate (3) are the sensor factor $p(y_n|\phi_n, \theta_n)$ and the posterior $p(\theta_n|Y_{n-1})$. The former is easily obtained from the observation model (1). The computation of the latter is a very hard problem owing to the Markovian statistical dependencies introduced by the random walk model (2) on the phase drift random sequence θ_n. In the next section, we introduce a stochastic recursive filtering approach allowing an efficient computation of $p(\theta_n|Y_{n-1})$ from the computational point of view.

B. Bayesian Filtering

The conditional density function $p(\theta_n|Y_n)$ summarizes all required information to derive any estimate of the phase drift $\hat{\theta}_n$ and can be determined recursively by the following equations [2]:

Prediction step

$$p(\theta_n|Y_{n-1}) = \int_{-\infty}^{+\infty} p(\theta_n|\theta_{n-1}) p(\theta_{n-1}|Y_{n-1}) \, \mathrm{d}\theta_{n-1}, \tag{5}$$

where $p(\theta_n|\theta_{n-1})$ is the so-called *transition* pdf.

Filtering step

$$p(\theta_n|Y_n) = \frac{p(y_n|\theta_n) p(\theta_n|Y_{n-1})}{\int_{-\infty}^{+\infty} p(y_n|\theta_n) p(\theta_n|Y_{n-1}) \, \mathrm{d}\theta_n} \tag{6}$$

where the pdf $p(y_n|\theta_n)$ is the so-called *sensor factor*.

From the random walk model (2), we have

$$p(\theta_n|\theta_{n-1}) = p_w(\theta_n - \theta_{n-1}), \tag{7}$$

implying that the operation in (5) is a convolution between $p_w(\cdot)$ and $p(\cdot|Y_{n-1})$ evaluated at θ_n.

The integral in the denominator of (6) is just a normalizing factor and, in certain cases, it need not to be evaluated.

The initial conditional density function $p(\theta_0|Y_{-1})$ is assumed to be

$$p(\theta_0|Y_{-1}) = p(\theta_0) \tag{8}$$

with $p(\theta_0)$ known.

Our approach for the non-linear problem is to approximate the non-Gaussian function $p(y_n|\theta_n)$ by a weighted sum of Gaussian functions such that (5) and (6) can be computed analytically. The Gaussian sum filter approach was considered, for instance, in [1], [3]–[5].

III. Gaussian Sum Filter

A. Gaussian Fitting of the Sensor Factor

From (1) and the statistics of v_n, the joint phase drift and symbol likelihood function is given by [6]

$$p(y_n|\theta_n, \phi_n) = p_v\left(y_n - e^{j(\phi_n+\theta_n)}\right) = \frac{1}{2\pi\sigma_v^2} e^{-\frac{\left|y_n - e^{j(\theta_n+\phi_n)}\right|^2}{2\sigma_v^2}}, \tag{9}$$

which is a periodic (non-Gaussian) function of θ_n and ϕ_n. Expanding the squared modulus in (9) leads to

$$\left|y_n - e^{j(\theta_n+\phi_n)}\right|^2 = 1 + |y_n|^2 - 2|y_n| \cos(\theta_n + \phi_n - \eta_n), \tag{10}$$

where $\eta_n = \arg\{y_n\}$ denotes the argument of y_n. By using (10) in (9) we have

$$p(y_n|\theta_n, \phi_n) = \frac{1}{2\pi\sigma_v^2} e^{-\gamma_n} e^{\beta_n \cos(\theta_n+\phi_n-\eta_n)}, \tag{11}$$

where

$$\gamma_n = \frac{1 + |y_n|^2}{2\sigma_v^2}, \tag{12}$$

and

$$\beta_n = \frac{|y_n|}{\sigma_v^2}. \tag{13}$$

Assuming equiprobable transmitted symbols, *i.e.*, $\mathrm{Prob}(\phi_n = 0) = \mathrm{Prob}(\phi_n = \pi) = 1/2$, the marginal pdf $p(y_n|\theta_n)$ is given by

$$\begin{aligned} p(y_n|\theta_n) &= \frac{1}{2} \int_{-\infty}^{+\infty} p(y_n|\theta_n, \phi_n) [\delta(\phi_n) + \delta(\phi_n - \pi)] \, \mathrm{d}\phi_n = \\ &= \frac{1}{4\pi\sigma_v^2} e^{-\gamma_n} \left(e^{\beta_n \cos(\eta_n - \theta_n)} + e^{-\beta_n \cos(\eta_n - \theta_n)} \right). \end{aligned} \tag{14}$$

The solid curve of Fig. 1 plots the sensor factor for $E_b/N_0 =$ 6 dB and $\theta_n \in [\eta - \frac{3}{2}\pi, \eta + \frac{3}{2}\pi]$. We stress that the sensor factor, quite often termed the *likelihood function*, is not a density; *i.e.*, $p(y_n|\theta_n)$ as a function of θ_n does not have to integrate to one, as θ_n is the conditioning variable. In our case, the sensor factor is not even integrable, as it is periodic.

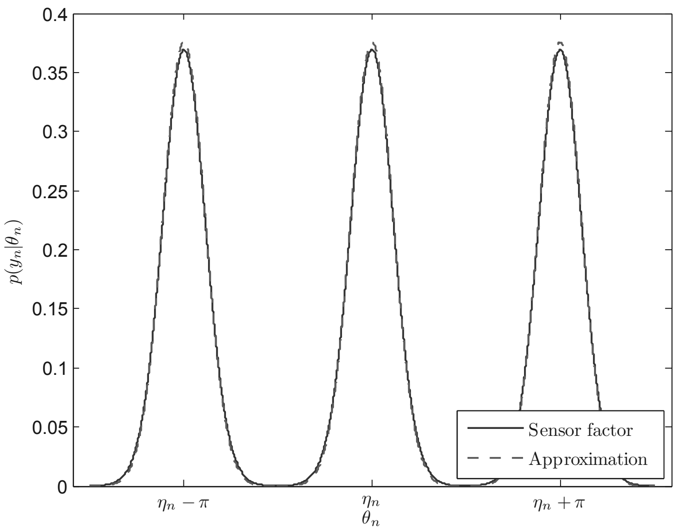

Fig. 1. Sensor factor $p(y_n|\theta_n)$ for $E_b/N_0 = 1/(2\sigma_v^2) = 6\,\text{dB}$, and the respective approximation.

It was shown in [7] that periodic replicas with the formal structure (14) are well approximated by Gaussian shapes. That is

$$p(y_n|\theta_n) \simeq \sum_{j\in\mathbb{Z}} \kappa\mathcal{N}\left(\theta_n - (\eta_n + j\pi), R_n\right) \quad (15)$$

where $\mathbb{Z}$ denotes the integers, κ is a constant, and R_n is chosen to minimize a measure of the approximation error (more on this ahead). The goodness of the approximation (15) is illustrated in Fig. 1 by the dashed line.

B. Computation of the Filtering Density

The Gaussian sum filter approach considers that both the prediction density function (5) and the filtering density function (6) are weighted sums of Gaussian functions. Therefore, the prediction density is given by

$$p(\theta_n|Y_{n-1}) = \sum_{i\in\mathbb{Z}} \alpha_n^{(i)} \mathcal{N}\left(\theta_n - \hat{\theta}_{n|n-1}^{(i)}, P_{n|n-1}\right), \quad (16)$$

with means $\hat{\theta}_{n|n-1}^{(i)}$ and common (justification given ahead) variance $P_{n|n-1}$. The filtering density is given by

$$p(\theta_n|Y_n) = \sum_{i\in\mathbb{Z}}\sum_{j\in\mathbb{Z}} \gamma_n^{(i,j)} \mathcal{N}(\theta_n - \hat{\theta}_{n|n}^{(i,j)}, P_{n|n}) \quad (17)$$

where, invoking the operation of multiplication in (6), results (see [3] for details)

$$\gamma_n^{(i,j)} = \alpha_n^{(i)} \kappa \mathcal{N}(\eta_n + j\pi - \hat{\theta}_{n|n-1}^{(i)}, R_n + P_{n|n-1}), \quad (18)$$

$$\hat{\theta}_{n|n}^{(i,j)} = \frac{1}{R_n + P_{n|n-1}}[(\eta_n + j\pi)P_{n|n-1} + \hat{\theta}_{n|n-1}^{(i)} R_n], \quad (19)$$

$$P_{n|n} = (R_n^{-1} + P_{n|n-1}^{-1})^{-1}. \quad (20)$$

The sum on the right hand side of (17) contains an infinite number of modes. In order to maintain the computational efficiency of the Gaussian sum filter, one must keep the number of modes as small as possible without losing significant information. A number of techniques have been developed to restrain the number of terms in a Gaussian mixture below a maximum value (see, *e.g.*, [3], [8], [9]). For its simplicity, our preferred component reduction procedure is to preserve the p modes in the Gaussian mixture (17) with the largest weights and discard the remaining (pruning procedure). The resulting filtering density is

$$p(\theta_n|Y_n) = \sum_{i=1}^{p} \lambda_n^{(i)} \mathcal{N}(\theta_n - \hat{\theta}_{n|n}^{(i)}, P_{n|n}), \quad (21)$$

where the weights $\lambda_n^{(i)}$ are obtained from the p largest weights $\gamma_n^{(i,j)}$, given in (18), by a re-normalization operation such that $\sum_{i=1}^{p} \lambda_n^{(i)} = 1$. Let the p largest weights be denoted as $\gamma_n^{(1)}, \ldots, \gamma_n^{(p)}$. Then $\lambda_n^{(i)} = \gamma_n^{(i)} / \sum_{i=1}^{p} \gamma_n^{(i)}$.

The Gaussian sum method is used to approximate $p(\theta_n|Y_n)$ such that the convolution in (5) and the products in (6) are determined in a simple manner, also allowing the MMSE estimate of θ_n to be computed analytically and efficiently in the filtering step as $\hat{\theta}_n = \sum_{i=1}^{p} \lambda_n^{(i)} \hat{\theta}_{n|n}^{(i)}$.

C. Computation of the Prediction Density

As the filtering density $p(\theta_n|Y_n)$ in (21) is a Gaussian sum, the prediction density $p(\theta_{n+1}|Y_n)$, resulting from the convolution of Gaussian functions, is also a Gaussian sum with weights (see (16)) $\alpha_{n+1}^{(i)} = \lambda_n^{(i)}$, means $\hat{\theta}_{n+1|n}^{(i)} = \hat{\theta}_{n|n}^{(i)}$, and variance $P_{n+1|n} = P_{n|n} + \sigma_w^2$.

The presence of a carrier frequency offset Δf can be accommodated by assuming a non-zero mean value for the noise w_n present in the random walk model (2) given by $E[w_n] = 2\pi\Delta f T/N$, where T is the bit duration. However, because we often do not have a precise estimate for the carrier frequency offset, we prefer modeling carrier frequency offsets by keeping the mean value of w_n set to zero and increasing its variance by $(2\pi\Delta f T/N)^2$. Although this is clearly a model mismatch, we will give evidence that the proposed stochastic filter is robust to it.

D. Implementation Aspects

An initial phase acquisition step is assumed to occur, *e.g.*, employing one of the techniques suggested in [10] or the references therein, and therefore, for synchronization purposes, the phase drift for the first symbol is assumed to be $\theta_0 = 0$. As a consequence, the initialization of the recursive algorithm is done by defining a filtering density with p modes and means

$$\hat{\theta}_{0|0}^{(i)} = 0, \quad i = 1, \ldots, p, \quad (22)$$

weights

$$\lambda_0^{(i)} = \begin{cases} 1, & i = 1 \\ 0, & \text{otherwise} \end{cases}, \quad i = 1, \ldots, p, \quad (23)$$

and variance $P_{0|0} = R_0$. In the following iterations ($n \geq 1$) the general propagation algorithm is applied.

To determine the variance R_n of the Gaussian modes in (15) we adopt the solution proposed in [7] which consists of minimizing the Kullback-Leibler distance between the initial pdf and the pdf used for replacement. Alternatively, we may use the approximation $R_n \approx \frac{1}{\beta_n} = \sigma_v^2/|y_n|$ which is especially tight for $\beta_n \geq 10$.

IV. MAP Symbol Detection

Let us consider that $p(\theta_n|Y_{n-1})$ has most of its probability mass concentrated in a small interval centered at $\hat{\theta}_{n|n-1}$. We have then $p(\theta_n|Y_{n-1}) \simeq \delta(\theta_n - \hat{\theta}_{n|n-1})$. Using this density in (4), we get

$$p(\phi_n|Y_n) \propto p(\phi_n)p(y_n|\phi_n, \hat{\theta}_{n|n-1}). \quad (24)$$

Assuming that $\text{Prob}(\phi_n = 0) = \text{Prob}(\phi_n = \pi) = 1/2$, *i.e.,* the symbols are equiprobable, the MAP estimate is given by

$$\hat{\phi}_n = \begin{cases} 0, & \cos(\eta_n - \hat{\theta}_{n|n-1}) > 0 \\ \pi, & \cos(\eta_n - \hat{\theta}_{n|n-1}) \leq 0 \end{cases} ; \quad n = 0, \dots, N-1. \quad (25)$$

V. Performance Results

Considering blocks with $N = 512$ BPSK data symbols, we depict in Fig. 2 the BER performance of the proposed receiver for several values of E_b/N_0. Two scenarios are considered. One in which the phase drift is a random walk process and another where, besides the random walk, there is carrier frequency offset. The parameters are $\sigma_w = \{0.05, 0.15\}$ rad and $\Delta fT = \{0, 8\}$. The case $\Delta fT = 8$ illustrates a possible scenario where the relative transmitter-receiver motion experiences Doppler frequency shift and/or the oscillators exhibit poor frequency alignment. Notice that the carrier frequency offset is not known and that the only modification in our algorithm is to include a constant term in the evaluation of the variance of the prediction density. In fact, although derived for phase trajectories defined by (2), our algorithm is able to estimate the phase errors given as

$$\theta_n = 2\pi\Delta f nT/N + \sum_{k=1}^{n} w_k = \theta_{n-1} + w_n + 2\pi\Delta fT/N. \quad (26)$$

This proves the robustness of the algorithm in the presence of carrier frequency offset.

Regarding the operation of the algorithm, two implementations are considered; one with $p = 1$ and another with $p = 3$ Gaussian modes. For comparison we also depict the curve of the theoretical BER for a BPSK constellation [11]. For $\sigma_w = 0.05$ rad and $\Delta fT = 0$ the curves for $p = 1$ and $p = 3$ display the same results. The estimator performs equally well and the curves are very close to the theoretical bit error probability. If we add carrier frequency offset the performance of the estimator degrades but, nevertheless, it still performs well. Our simulations have shown that, with $\Delta fT = 8$, if a compensation procedure is not undertaken, the BER will be 0.5. Also notable is the fact that, in the presence of carrier frequency offset the performance of the estimator, even for small values of σ_w, is better with $p = 3$ than with $p = 1$. For larger values of σ_w we must always consider $p = 3$.

VI. Conclusions

We proposed a Bayesian solution for the problem of simultaneous bit detection and phase drift estimation by approximating the posterior non-Gaussian probability density function of the phase by a weighted sum of Gaussian functions. Additionally, we compare the performance results obtained when modeling the density function by a weighted sum

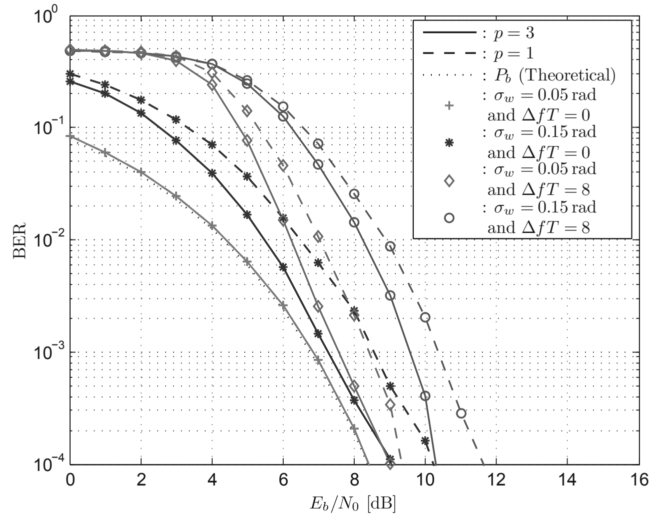

Fig. 2. BER versus E_b/N_0 for $p = 1$ and $p = 3$ Gaussian modes in the presence of phase drift with carrier frequency offset.

of Gaussian functions with those obtained using a single Gaussian. Moreover, through simulation results with BPSK signaling we observed that the proposed algorithm presents outstanding results, even in the presence of carrier frequency offset, for perfect channel estimation and symbol synchronization. The generalization of the algorithm to M-PSK signals, with $M > 2$, is straightforward, with changes affecting mainly the symbol detection scheme.

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable comments. This work was partially supported by the FCT – Fundação para a Ciência e Tecnologia (pluriannual funding, ADCOD project PTDC/EEA – TEL/099973/2008, PEst-OE / EEI / LA0008 / 2011 and Ph.D. grant SFRH/BD/40265/2007).

References

[1] D. Alspach and H. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Trans. Autom. Control*, vol. 17, no. 4, pp. 339–448, Aug. 1972.

[2] A. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic Press, 1970.

[3] J. M. N. Leitão and J. M. F. Moura, "Acquisition in phase demodulation: application to ranging in radar/sonar systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 2, pp. 581–599, Apr. 1995.

[4] W. I. Tam, K. N. Plataniotis, and D. Hatzinakos, "An adaptive Gaussian sum algorithm for radar tracking," *Signal Processing*, vol. 77, no. 1, pp. 85–104, Aug. 1999.

[5] J. M. F. Moura and A. B. Baggeroer, "Phase unwrapping of signals propagated under the Arctic ice crust: a statistical approach," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 5, pp. 617–630, May 1988.

[6] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, 1st ed. New Jersey, Englewood Cliffs: Prentice-Hall, 1993.

[7] J. M. N. Leitão and J. M. F. Moura, "Nonlinear phase estimators based on the Kullback distance," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 4, Apr. 1994, pp. 521–524.

[8] D. G. Lainiotis, "Partitioning: A unifying framework for adaptative systems I: Estimation," *Proc. IEEE*, vol. 64, pp. 1126–1143, 1976.

[9] Y. Bar-Shalom and X. R. Li, *Estimation and Tracking: Principles, Techniques and Software*. New York: Artech House, 1986.

[10] M. Chiani, "Noncoherent frame synchronization," *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1536–1545, May 2010.

[11] J. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.