LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-BOOK-409377

Sequential Bayesian Detection: A Model-Based Approach

J. V. Candy

December 12, 2008

Classical, Semi-Classical and Quantum Noise

Disclaimer

Sequential Bayesian Detection: A Model-Based Approach

James V. Candy*

Lawrence Livermore National Laboratory
P.O. 808, L-151
Livermore, CA 94551

December 8, 2008

Abstract

Sequential detection theory has been known for a long time evolving in the late 1940's by Wald and followed by Middleton's classic exposition in the 1960's coupled with the concurrent enabling technology of digital computer systems and the development of sequential processors. Its development, when coupled to modern sequential model-based processors, offers a reasonable way to attack physics-based problems. In this chapter, the fundamentals of the sequential detection are reviewed from the Neyman-Pearson theoretical perspective and formulated for both linear and nonlinear (approximate) Gauss-Markov, state-space representations. We review the development of modern sequential detectors and incorporate the sequential model-based processors as an integral part of their solution. Motivated by a wealth of physics-based detection problems, we show how both linear and nonlinear processors can seamlessly be embedded into the sequential detection framework to provid a powerful approach to solving non-stationary detection problems.

1 Introduction

Sequential detection is a methodology developed essentially by Wald [1] in the late 1940's providing an alternative to the classical batch methods evolving from the basic Neyman-Pearson theory of the 1930's [2], [3]. From the detection theoretical viewpoint, the risk (or error) associated with a

*candy1@llnl.gov; phone 1 925 422-8675

decision typically decreases as the number of measurements increase. Sequential detection enables a decision to be made more rapidly (in most cases) employing fewer measurements while maintaining the same level of risk. Thus, the aspiration is to reduce the decision time while maintaining the risk for a fixed sample size. Its significance was truly brought to the forefront with the evolution of the digital computer and the fundamental idea of acquiring and processing data in a sequential manner. The seminal work of Middleton ([2], [4]-[7]) as well as the development of sequential processing techniques ([8]-[13]) during the 1960's provided the necessary foundation for the sequential processor/detector that is applied in a routine manner today ([7], [8], [9]-[11], [13]).

This chapter investigates the idea of a sequential processor from the statistical perspective and relates its operation to that of the classical sequential Neyman-Pearson detector developed by Wald and termed the sequential probability ratio test (*SPRT*). Here the investigation of simple binary hypothesis testing evolves as a common thread that leads to Wald's *SPRT*. We start with the batch detector developing its sequential variant and then show how the *SPRT* easily follows from the results of Neyman-Pearson theory ([1], [2], [3], [9]-[11],).

In Sec. 2 we develop the sequential detector and compare it to the *SPRT* indicating the similarities and differences. In Sec. 3 we discuss how the Gauss-Markov linear processor can be applied to the model-based problem followed by the nonlinear case in Sec. 4.

2 Sequential Detection

In this section we briefly develop sequential detection theory which will be expanded to include the model-based approach in the section to follow. We start with the "batch" solution and then develop the sequential approach.

Let us assume that we have a set of measurements defined by $\mathrm{M}_{t_k} := \{m(t_k), \cdots, m(t_1), m(t_0)\}$ and we would like to decide between two hypothesis, H_0 and H_1 respectively. We choose the Neyman-Pearson criterion to develop our detector [10]. Recall that the Neyman-Pearson theorem states that a detector is optimal if it maximizes the probability of detection, P_{DET} (or minimizes the miss probability) for any false alarm rate less than a prespecified value, say P_{FA}^*. The theorem follows directly from a constrained optimization problem formulation using Lagrange multipliers (see [3], [10] for details), which yields the solution based on the ratio of likelihoods, $\Pr(M_{t_k}|H_i);\ i = 0, 1$ (for the binary case). That is, to maximize P_{DET} for a fixed value of P_{FA}^*, we have the likelihood ratio, $\mathcal{L}(t_k)$ or equivalently the sufficient statistic defined by the joint density functions

$$\mathcal{L}(t_k) := \mathcal{L}\left(\mathrm{M}_{t_k}\right) = \frac{\Pr\left(\mathrm{M}_{t_k}|H_1\right)}{\Pr\left(\mathrm{M}_{t_k}|H_0\right)} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \mathrm{T} \tag{1}$$

for $\Pr(\cdot)$ the respective joint probabilities under each hypothesis. Here the threshold T is found from integrating the false alarm density for the pre-specified value

$$P_{FA} = \int_{\mathrm{T}}^{\infty} \Pr\left(\mathrm{M}_{t_k}|H_0\right) dP = P_{FA}^{*} \tag{2}$$

Thus, based on the Neyman-Pearson criterion, the optimal detector can be found for batch mode as

$$\mathcal{L}(t_k) = \frac{\Pr\left(\mathrm{M}_{t_k}|H_1\right)}{\Pr\left(\mathrm{M}_{t_k}|H_0\right)} = \frac{\Pr\left(\mathbf{m}(t_k), \cdots, \mathbf{m}(t_0)|H_1\right)}{\Pr\left(\mathbf{m}(t_k), \cdots, \mathbf{m}(t_0)|H_0\right)} \tag{3}$$

It follows from the chain rule of probability theory [10] that

$$\begin{aligned} \Pr\left(\mathrm{M}_{t_k}|H_i\right) &= \prod_{\ell=0}^{t_k} \Pr\left(\mathbf{m}(t_{k-\ell})|\mathrm{M}_{t_{k-\ell-1}}; H_i\right) \\ &= \Pr\left(\mathbf{m}(t_k)|\mathrm{M}_{t_{k-1}}; H_i\right) \times \Pr\left(\mathbf{m}(t_{k-1})|\mathrm{M}_{t_{k-2}}; H_i\right) \\ &\times \cdots \times \Pr\left(\mathbf{m}(1)|\mathbf{m}(0); H_i\right) \times \Pr\left(\mathbf{m}(0); H_i\right) \end{aligned} \tag{4}$$

can be expressed succinctly using Bayes rule as

$$\Pr\left(\mathrm{M}_{t_k}|H_i\right) = \Pr\left(\mathbf{m}(t_k), \mathrm{M}_{t_{k-1}}|H_i\right) = \Pr\left(\mathbf{m}(t_k)|\mathrm{M}_{t_{k-1}}; H_i\right) \Pr\left(\mathrm{M}_{t_{k-1}}|H_i\right) \tag{5}$$

Substituting this expression into Eq. (1), we obtain

$$\begin{aligned} \mathcal{L}(t_k) &= \left[\frac{\Pr\left(\mathrm{M}_{t_{k-1}}|H_1\right)}{\Pr\left(\mathrm{M}_{t_{k-1}}|H_0\right)}\right] \frac{\Pr\left(\mathbf{m}(t_k)|\mathrm{M}_{t_{k-1}}; H_1\right)}{\Pr\left(\mathbf{m}(t_k)|\mathrm{M}_{t_{k-1}}; H_0\right)} \\ &= \mathcal{L}(t_{k-1}) \frac{\Pr\left(\mathbf{m}(t_k)|\mathrm{M}_{t_{k-1}}; H_1\right)}{\Pr\left(\mathbf{m}(t_k)|\mathrm{M}_{t_{k-1}}; H_0\right)} \end{aligned} \tag{6}$$

which is precisely the sequential form of the likelihood ratio. It is also clear that not just the likelihood function can be used but any monotonic function of the likelihood can also be used as well [10]. Taking natural logarithms of both sides of the equation, and defining $\Lambda(t_k) := \ln \mathcal{L}(t_k)$, we obtain the sequential log-likelihood ratio as

$$\Lambda(t_k) = \Lambda(t_{k-1}) + \ln \Pr\left(\mathbf{m}(t_k)|\mathrm{M}_{t_{k-1}}; H_1\right) - \ln \Pr\left(\mathbf{m}(t_k)|\mathrm{M}_{t_{k-1}}; H_0\right) \quad (7)$$

To construct the sequential probability ratio test (*SPRT*) or equivalently the sequential likelihood ratio test (*SLRT*) based on the Neyman-Pearson criterion, we must define two time-varying thresholds $\mathrm{T}_0(t_k)$ and $\mathrm{T}_1(t_k)$. If the likelihood ratio at time t_k is greater than threshold $\mathrm{T}_1(t_k)$, we accept hypothesis H_1. If it is less than $\mathrm{T}_0(t_k)$, we accept H_0, but if its value lies between the thresholds, we *continue* to take another sample. The sequential test differs from the fixed sample size batch test and is capable of handling non-stationary statistics.

To implement the *SLRT* at the t_k^{th}-stage, we must know the prior probabilities or equivalently the prior likelihood ratio, $\mathcal{L}(t_{k-1})$ (or $\Lambda(t_{k-1})$) containing all of the past information and the corresponding stage conditional probabilities $\Pr(\mathbf{m}(t_k)|M_{t_{k-1}}; H_i)$ to sequentially update the likelihood. The t_k^{th} stage thresholds are constructed following Neyman-Pearson theory in terms of the detection and false alarm probabilities as

$$\mathrm{T}_0(t_k) = \frac{P_{\mathrm{DET}}}{P_{\mathrm{FA}}} = \frac{1 - P_{\mathrm{MISS}}}{P_{\mathrm{FA}}}, \quad \text{and} \quad \mathrm{T}_1(t_k) = \frac{1 - P_{\mathrm{DET}}}{1 - P_{\mathrm{FA}}} = \frac{P_{\mathrm{MISS}}}{1 - P_{\mathrm{FA}}} \quad (8)$$

In the case of the log-likelihood, the thresholds are:

$$\begin{aligned} &\Lambda(t_k) \geq \ln \mathrm{T}_1(t_k) && [\mathrm{Accept}\, H_1] \\ &\ln \mathrm{T}_0(t_k) < \Lambda(t_k) < \ln \mathrm{T}_1(t_k) && [\mathrm{Continue}] \\ &\Lambda(t_k) \leq \ln \mathrm{T}_0(t_k) && [\mathrm{Accept}\, H_0] \end{aligned} \quad (9)$$

This completes the fundamental concepts for the construction of the sequential detection approach, next we investigate the development of the model-based sequential detector.

3 Model-Based Sequential Detection: Linear Case

In this section we develop the model-based approach to sequential detection employing the *SLRT* as the mechanism to implement the model-based designs starting with the linear time-varying (non-stationary) problem (see [13], [15] for more details). We begin with the development of the generic Gauss-Markov signal model defined by its *state* (signal) *vector*, $\mathbf{s}(t_k)$, where the linear state-space process model is given by

$$\mathbf{s}(t_k) = A(t_{k-1})\mathbf{s}(t_{k-1}) + B(t_{k-1})\mathbf{u}(t_{k-1}) + \mathbf{w}(t_{k-1}) \tag{10}$$

with corresponding measurement model as

$$\mathbf{m}(t_k) = C(t_k)\mathbf{s}(t_k) + \mathbf{v}(t_k) \tag{11}$$

where $\mathbf{s}$, $\mathbf{w}$ are the N_s-dimensional signal and process noise vectors, with $\mathbf{m}$, $\mathbf{v}$ the N_m-dimensional measurement and noise vectors along with $\mathbf{u}$ the N_u-dimensional known input vector. Both $\mathbf{w}$ and $\mathbf{v}$ are zero-mean Gaussian with respective covariances, $R_{\mathbf{ww}}(t_{k-1})$ and $R_{\mathbf{vv}}(t_k)$. The corresponding system, input and measurement matrices are appropriately dimensioned and given by $A(t_{k-1})$, $B(t_{k-1})$ and $C(t_k)$, respectively. The initial state vector is Gaussian with $\mathbf{s}(0) \sim \mathcal{N}(\overline{\mathbf{s}}(0), P_{\mathbf{ss}}(0))$.

Because the underlying distributions are Gaussian, we know that the optimal solution to the signal estimation problem is given by the model-based processor (Kalman filter) providing the predicted conditional mean estimate, $\hat{\mathbf{s}}(t_k|t_{k-1}) := E\{\mathbf{s}(t_k)|M_{t_{k-1}}\}$[1] with corresponding predicted conditional (error) covariance, $\tilde{P}(t_k|t_{k-1})$ [12]-[15].

With this signal model in hand we can now define the binary problem to decide whether the measurement contains the signal or just noise alone, that is, we are testing the hypotheses that:

$$\begin{array}{lrcll} H_0: & \mathbf{m}(t_k) & = & \mathbf{v}(t_k) & \text{[NOISE]} \\ H_1: & \mathbf{m}(t_k) & = & C(t_k)\mathbf{s}(t_k) + \mathbf{v}(t_k) & \text{[SIGNAL+NOISE]} \end{array}$$

From the underlying Gauss-Markov assumption, the sequential likelihood-ratio solution is specified by the ratio of conditional Gaussian posterior distributions, $\Pr(\mathbf{m}(t_k)|M_{t_{k-1}}; H_i);\ i = 0, 1$. That is,

$$\begin{array}{rcl} \Pr(\mathbf{m}(t_k)|M_{t_{k-1}}; H_0) & \sim & \mathcal{N}\left(\mathbf{m}(t_k) : \mathbf{0}, R_{\mathbf{vv}}(t_k)\right) \\ \Pr(\mathbf{m}(t_k)|M_{t_{k-1}}; H_1) & \sim & \mathcal{N}\left(\mathbf{m}(t_k) : C(t_k)\hat{\mathbf{s}}(t_k|t_{k-1}), R_{\mathbf{ee}}(t_k)\right) \end{array}$$

where the *innovations*, $\mathbf{e}(t_k)$ and its corresponding covariance, $R_{\mathbf{ee}}(t_k)$ are obtained as outputs of the model-based processor (Kalman filter) specified by:

$$\begin{array}{rcl} \mathbf{e}(t_k) & = & \mathbf{m}(t_k) - \hat{\mathbf{m}}(t_k|t_{k-1}) = \mathbf{m}(t_k) - C(t_k)\hat{\mathbf{s}}(t_k|t_{k-1}) \\ R_{\mathbf{ee}}(t_k) & = & C(t_k)\tilde{P}(t_k|t_{k-1})C'(t_k) + R_{\mathbf{vv}}(t_k) \end{array} \tag{12}$$

Thus, under the null hypothesis we have that

[1] This notation is defined in terms of predicted *conditional means* and *covariances* by: $\hat{\mathbf{s}}(t_k|t_{k-1})$ and $\tilde{P}(t_k|t_{k-1}) := \text{cov}(\tilde{\mathbf{s}}(t_k|t_{k-1}))$ for the *predicted state estimation error*, $\tilde{\mathbf{s}}(t_k|t_{k-1}) := \mathbf{s}(t_k) - \hat{\mathbf{s}}(t_k|t_{k-1})$.

$$\Pr(\mathbf{m}(t_k)|M_{t_{k-1}};H_0) = \frac{1}{(2\pi)^{N_m/2}|R_{\mathbf{vv}}(t_k)|^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{m}'(t_k)R_{\mathbf{vv}}^{-1}(t_k)\mathbf{m}(t_k)\right\} \tag{13}$$

while the conditional probability under alternate hypothesis is given by

$$\begin{aligned}\Pr(\mathbf{m}(t_k) \quad | \quad M_{t_{k-1}};H_1) = & \frac{1}{(2\pi)^{N_m/2}|R_{\mathbf{ee}}(t_k)|^{1/2}} \times \\ & \exp\left\{-\frac{1}{2}(\mathbf{m}(t_k)-\hat{\mathbf{m}}(t_k|t_{k-1}))'R_{ee}^{-1}(t_k)(\mathbf{m}(t_k)-\hat{\mathbf{m}}(t_k|t_{k-1}))\right\}\end{aligned} \tag{14}$$

or simply

$$\Pr(\mathbf{m}(t_k)|M_{t_{k-1}};H_1) = \frac{1}{(2\pi)^{N_m/2}|R_{\mathbf{ee}}(t_k)|^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{e}'(t_k)R_{\mathbf{ee}}^{-1}(t_k)\mathbf{e}(t_k)\right\} \tag{15}$$

Moving all *known* terms to the threshold, the required sequential log-likelihood of Eq. 7 for the linear Gauss-Markov signal model becomes

$$\begin{aligned}\Lambda(t_k) &= \Lambda(t_{k-1}) - \frac{1}{2}\mathbf{e}'(t_k)R_{\mathbf{ee}}^{-1}(t_k)\mathbf{e}(t_k) + \frac{1}{2}\mathbf{m}'(t_k)R_{\mathbf{vv}}^{-1}(t_k)\mathbf{m}(t_k) \underset{H_0}{\overset{H_1}{\gtrless}} \mathcal{T}(t_k) \\ \mathcal{T}(t_k) &= \ln \mathrm{T}(t_k) - \ln\left(\frac{1}{(2\pi)^{N_m/2}|R_{\mathbf{ee}}(t_k)|^{1/2}}\right) + \ln\left(\frac{1}{(2\pi)^{N_m/2}|R_{\mathbf{vv}}(t_k)|^{1/2}}\right)\end{aligned} \tag{16}$$

All that remains is to specify the predicted and corrected means and covariances to implement the sequential detector. These are available as part of the model-based algorithm given by:

$$\begin{aligned}\hat{\mathbf{s}}(t_k|t_{k-1}) &= A(t_{k-1})\hat{\mathbf{s}}(t_{k-1}|t_{k-1}) + B(t_{k-1})\mathbf{u}(t_{k-1}) && \text{[Prediction]} \\ \tilde{P}(t_k|t_{k-1}) &= A(t_{k-1})\tilde{P}(t_{k-1}|t_{k-1})A'(t_{k-1}) + R_{\mathbf{ww}}(t_{k-1}) && \text{[Prediction Cov.]}\end{aligned}$$

$$\begin{aligned}\hat{\mathbf{s}}(t_k|t_k) &= \hat{\mathbf{s}}(t_k|t_{k-1}) + K(t_k)\mathbf{e}(t_k) && \text{[Correction]} \\ \tilde{P}(t_k|t_k) &= [\mathbf{I} - K(t_k)C(t_k)]\,\tilde{P}(t_k|t_{k-1}) && \text{[Correction Cov.]}\end{aligned}$$

for $K(t_k)$ the corresponding gain (see [15] for more details), next we consider the nonlinear case.

4 Model-Based Sequential Detection: Nonlinear Case

In this section we use the development of the linear case as a road map and develop the sequential model-based detector for the nonlinear case. We start with the nonlinear (approximate) Gauss-Markov model given by:

$$\begin{aligned} \mathbf{s}(t_k) &= \mathbf{a}\left[\mathbf{s}(t_{k-1})\right] + \mathbf{b}\left[\mathbf{u}(t_{k-1})\right] + \mathbf{w}(t_{k-1}) \\ \mathbf{m}(t_k) &= \mathbf{c}\left[\mathbf{s}(t_k)\right] + \mathbf{v}(t_k) \end{aligned} \tag{17}$$

where (as before) $\mathbf{s}$, $\mathbf{w}$ are the N_s-dimensional signal and process noise vectors with $\mathbf{m}$, $\mathbf{v}$ the N_m-dimensional measurement and noise vectors along with $\mathbf{u}$ the N_u-dimensional known input vector. Both $\mathbf{w}$ and $\mathbf{v}$ are zero-mean Gaussian with respective covariances, $R_{\mathbf{ww}}(t_{k-1})$ and $R_{\mathbf{vv}}(t_k)$. The corresponding *nonlinear* system, input and measurement vector functions are appropriately dimensioned and given by $\mathbf{a}[\cdot]$, $\mathbf{b}[\cdot]$ and $\mathbf{c}[\cdot]$, respectively. The initial state vector is Gaussian with $\mathbf{s}(0) \sim \mathcal{N}(\overline{\mathbf{s}}(0), P_{\mathbf{ss}}(0))$.

Because the underlying distributions are assumed approximately Gaussian, we know that one possible solution to the signal estimation problem is given by the model-based processor (extended Kalman filter[2]) providing the predicted conditional mean estimate, $\hat{\mathbf{s}}(t_k|t_{k-1})$, with corresponding predicted conditional (error) covariance, $\tilde{P}(t_k|t_{k-1})$ [15].

With this nonlinear signal model in hand we can now define the binary problem to decide whether the measurement contains the signal or just noise alone, that is, we are testing the hypotheses that:

$$\begin{aligned} H_0: \quad \mathbf{m}(t_k) &= \mathbf{v}(t_k) && \text{[NOISE]} \\ H_1: \quad \mathbf{m}(t_k) &= \mathbf{c}\left[\mathbf{s}(t_k)\right] + \mathbf{v}(t_k) && \text{[SIGNAL+NOISE]} \end{aligned}$$

Before we proceed with the sequential detection development for nonlinear, non-Gaussian problems, we have chosen to employ the first-order Taylor series representation to approximate the nonlinear vector functions. The development evolves quite naturally from a linearized model-based processor (linearized Kalman filter) [15]. These approximations when expanded about a reference $\mathbf{s} = \mathbf{s}^*$ take the general form

$$\mathbf{c}\left[\mathbf{s}(t_k)\right] \approx \mathbf{c}\left[\mathbf{s}^*(t_k)\right] + \mathbf{C}\left[\mathbf{s}^*(t_k)\right] \times (\mathbf{s}(t_k) - \mathbf{s}^*(t_k)) + \text{H.O.T.} \tag{18}$$

[2]It is well-known that some of the modern variants currently available offer alternatives like the unscented Kalman filter or particle filter ([15]-[26]) that are better than the *EKF* but we choose this formulation since its easily tracks the linear case developed previously.

with $\mathbf{C}[\mathbf{s}^*(t_k)] := \frac{\partial}{\partial \mathbf{s}}\mathbf{c}[\mathbf{s}]\big|_{\mathbf{s}=\mathbf{s}^*(t_k)}$. Each of the nonlinear system, input and measurement functions are approximated in this manner using the Taylor series. For the ad-hoc nonlinear *EKF* processor, the most current estimate available, $\hat{\mathbf{s}}(t_k|*)$, is used as a reference during the prediction and correction steps for the linearization, that is, the state estimates are

$$\begin{aligned} \hat{\mathbf{s}}(t_k|t_{k-1}) &= \mathbf{a}\left[\hat{\mathbf{s}}(t_{k-1}|t_{k-1})\right] + \mathbf{b}\left[\mathbf{u}(t_{k-1})\right] \\ \hat{\mathbf{s}}(t_k|t_k) &= \hat{\mathbf{s}}(t_k|t_{k-1}) + \mathbf{K}\left[\hat{\mathbf{s}}(t_k|t_{k-1})\right]\mathbf{e}(t_k) \end{aligned} \tag{19}$$

where the Jacobians appear in the covariance equations.

With this in mind and from the underlying *approximate* Gauss-Markov assumption, the sequential likelihood-ratio solution is again specified by the ratio of conditional Gaussian posterior distributions, $\Pr(\mathbf{m}(t_k)|M_{t_{k-1}}; H_i);\ i = 0, 1$ as before. That is,

$$\begin{aligned} \Pr(\mathbf{m}(t_k)|M_{t_{k-1}}; H_0) &\sim \mathcal{N}\left(\mathbf{m}(t_k) : \mathbf{0}, R_{\mathbf{vv}}(t_k)\right) \\ \Pr(\mathbf{m}(t_k)|M_{t_{k-1}}; H_1) &\sim \mathcal{N}\left(\mathbf{m}(t_k) : \mathbf{c}\left[\hat{\mathbf{s}}(t_k|t_{k-1})\right], R_{\mathbf{ee}}(t_k)\right) \end{aligned}$$

where the *innovations*, $\mathbf{e}(t_k)$ and its corresponding covariance, $R_{\mathbf{ee}}(t_k)$ are obtained as outputs of the model-based processor (extended Kalman filter) specified by:

$$\begin{aligned} \mathbf{e}(t_k) &= \mathbf{m}(t_k) - \hat{\mathbf{m}}(t_k|t_{k-1}) = \mathbf{m}(t_k) - \mathbf{c}\left[\hat{\mathbf{s}}(t_k|t_{k-1})\right] \\ R_{\mathbf{ee}}(t_k) &= \mathbf{C}\left[\hat{\mathbf{s}}(t_k|t_{k-1})\right]\tilde{P}(t_k|t_{k-1})\mathbf{C}'\left[\hat{\mathbf{s}}(t_k|t_{k-1})\right] + R_{\mathbf{vv}}(t_k) \end{aligned} \tag{20}$$

Thus, under the null hypothesis we have that

$$\Pr(\mathbf{m}(t_k)|M_{t_{k-1}}; H_0) = \frac{1}{(2\pi)^{N_m/2}|R_{\mathbf{vv}}(t_k)|^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{m}'(t_k)R_{\mathbf{vv}}^{-1}(t_k)\mathbf{m}(t_k)\right\} \tag{21}$$

while the conditional probability under alternate hypothesis is given by

$$\Pr(\mathbf{m}(t_k)|M_{t_{k-1}}; H_1) = \frac{1}{(2\pi)^{N_m/2}|R_{\mathbf{ee}}(t_k)|^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{e}'(t_k)R_{\mathbf{ee}}^{-1}(t_k)\mathbf{e}(t_k)\right\} \tag{22}$$

Again moving all known terms to the threshold, the required sequential log-likelihood of Eq. 7 for the nonlinear Gauss-Markov signal model becomes

$$\Lambda(t_k) = \Lambda(t_{k-1}) - \frac{1}{2}\mathbf{e}'(t_k)R_{\mathbf{ee}}^{-1}(t_k)\mathbf{e}(t_k) + \frac{1}{2}\mathbf{m}'(t_k)R_{\mathbf{vv}}^{-1}(t_k)\mathbf{m}(t_k) \underset{H_0}{\overset{H_1}{\gtrless}} \mathcal{T}(t_k)$$

$$\mathcal{T}(t_k) \quad = \quad \ln \mathrm{T}(t_k) - \ln\left(\frac{1}{(2\pi)^{N_m/2}|R_{\mathbf{ee}}(t_k)|^{1/2}}\right) + \ln\left(\frac{1}{(2\pi)^{N_m/2}|R_{\mathbf{vv}}(t_k)|^{1/2}}\right) \tag{23}$$

All that remains is to specify the predicted and corrected covariances to implement the sequential detector, since the state estimates we given in Eq. 19 above. These are available as part of the nonlinear model-based algorithm (*EKF*) given by:

$$\begin{aligned}
\tilde{P}(t_k|t_{k-1}) &= \mathbf{A}\left[\hat{\mathbf{s}}(t_k|t_{k-1})\right]\tilde{P}(t_{k-1}|t_{k-1})\mathbf{A}'\left[\hat{\mathbf{s}}(t_k|t_{k-1})\right] + R_{\mathbf{ww}}(t_{k-1}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{[Prediction Cov.]} \\
\tilde{P}(t_k|t_k) &= \left[\mathbf{I} - \mathbf{K}\left[\hat{\mathbf{s}}(t_k|t_{k-1})\right]\mathbf{C}\left[\hat{\mathbf{s}}(t_k|t_{k-1})\right]\right]\tilde{P}(t_k|t_{k-1}) \quad \text{[Correction Cov.]}
\end{aligned}$$

for $\mathbf{K}\left[\hat{\mathbf{s}}(t_k|t_{k-1})\right]$ the corresponding gain (see [15] for more details), This completes the development.

5 Summary

In this chapter we have motivated and summarized the development of sequential detection theory from the original Neyman-Pearson theory following the Wald approach [1]. We then showed how a model-based approach can be incorporated into the sequential paradigm. We developed the Gauss-Markov (state-space) representation of both linear and nonlinear systems capable of capturing the dynamics of many physics-based problems [15] and showed how they can be embedded into the sequential theoretical framework. The solution lead to the implementation of a linear model-based processor (linear Kalman filter) as well as the nonlinear model-based scheme (extended Kalman filter).

References

[1] A. Wald, *Sequential Analysis*, New York: Dover Publications, 1973.

[2] D. Middleton, *Introduction to Statistical Communication Theory*, New York: McGraw-Hill, 1960.

[3] H. Van Trees, *Detection, Estimation and Modulation Theory*, Pt. 1, New York: John Wiley, 1968.

[4] D. Middleton and R. Esposito, "Simultaneous optimum detection and estimation of signals in noise," *IEEE Trans. Inform. Theory*, IT-14, (3), 434-444, 1968.

[5] D. Middleton and R. Esposito, "New results in the theory of simultaneous optimum detection and estimation of signals in noise," *Problemy Peredachi Informatsii*, 6, (3), 3-20, 1970.

[6] E. Sullivan and D. Middleton, "Estimation and detection issues in matched-field processing," *IEEE Trans. Oceanic Engr.*, **18**, (3), 156-167, 1993.

[7] J. Candy and D. Middleton, *Joint Detection-Estimation: A Model-Based Solution*, **LLNL Report**, UCRL-JC-120219, 1995.

[8] M. Basseville and I. Nikiforov, *Detection by Abrupt Changes: Theory and Application*, New Jersey: Prentice-Hall, 1993.

[9] L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison-Wesley, Massachusetts, 1991.

[10] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, New Jersey: Prentice-Hall, 1998.

[11] C. Therrian, *Random Signal Processing: Detection Theory*, New Jersey: Prentice-Hall, 1991.

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

[13] A. Sage and J. Melsa, *Estimation Theory with Applications to Communcations and Control*, New York: McGraw-Hill, 1971.

[14] B. Anderson and J. Moore, *Optimum Filtering*, New Jersey: Prentice-Hall, 1979.

[15] J. V. Candy, *Model-Based Signal Processing*, John Wiley/IEEE Press: New Jersey, 2006.

[16] S. J. Julier, J. K. Uhlmann and H. F. Durrant-Whyte, A new method for the nonlinear transformation of means and covariances in filters and estimators, *IEEE Trans. Autom. Contr.*, Vol. 45, No. 3, 2000.

[17] S. Haykin and N. de Freitas, "Special Issue: Sequential state estimation: from Kalman filters to particle filters." *Proc. IEEE*, vol. **92**, No. 3, 399-574, 2004.

[18] S. Godsill and P. Djuric, "Special Issue: Monte Carlo methods for statistical signal processing." *IEEE Trans. Signal Proc.*, vol. 50, 2002.

[19] B. Ristic, S. Arulampalam and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, Boston: Artech House, 2004.

[20] J. V. Candy, *Bayesian Signal Processing: Classical, Modern and Particle Filtering*, John Wiley/IEEE Press: New Jersey, 2009.

[21] N. Gordon, D. Salmond and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proc.-F.*, vol. 140, pp. 107-113,1993.

[22] M. Arulampalam, S. Maskell, N. Gordon and T. Clapp "A tutorial on particle filters for online nonlinear/non-gaussian Bayesian tracking." *IEEE Trans. Signal Proc.*, vol. **50**, no. 2, 174-188, 2002.

[23] P. Djuric, J. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. Bugallo and J. Miguez, "Particle Filtering." *IEEE Signal Proc. Mag.* vol. **20**, No. 5, 19-38, 2003.

[24] A. Doucet and X. Wang, "Monte Carlo methods for signal processing," *IEEE Signal Proc. Mag.* vol. **24**, No. 5, 152-170, 2005.

[25] O. Cappe, S. Godsill and E. Moulines, "An overview of existing methods and recent advances in sequential Monte Carlo," *Proc. IEEE*, Vol. 95, No. 5, pp. 899-924, 2007.

[26] J. Candy, "Bootstrap Particle Filtering," *IEEE Signal Proc. Magz.*, Vol. 24, No. 4, pp. 73-85, 2007.

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.