

# A Particle Filtering Approach to Change Detection for Nonlinear Systems

**Babak Azimi-Sadjadi**

*Electrical, Computer, and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, NY 12180-3590, USA  
Email: babak@ecse.rpi.edu*

*The Institute for Systems Research, University of Maryland, College Park, MD 20742, USA*

**P. S. Krishnaprasad**

*The Institute for Systems Research, University of Maryland, College Park, MD 20742, USA  
Email: krishna@isr.umd.edu*

*Received 13 September 2003; Revised 22 March 2004*

We present a change detection method for nonlinear stochastic systems based on particle filtering. We assume that the parameters of the system before and after change are known. The statistic for this method is chosen in such a way that it can be calculated recursively while the computational complexity of the method remains constant with respect to time. We present simulation results that show the advantages of this method compared to linearization techniques.

**Keywords and phrases:** nonlinear filtering, generalized likelihood ratio test, CUSUM algorithm, online change detection.

## 1. INTRODUCTION

Page states the change detection problem as follows [1]: “Whenever observations are taken in order it can happen that the whole set of observations can be divided into subsets, each of which can be regarded as a random sample from a common distribution, each subset corresponding to a different parameter value of the distribution. The problems to be considered in this paper are concerned with the identification of the subsamples and the detection of changes in the parameter value.”

We refer to a change or an abrupt change as any change in the parameters of the system that happens either instantaneously or much faster than any change that the nominal bandwidth of the system allows.

The key difficulty of all change detection methods is that of detecting intrinsic changes that are not necessarily directly observed but are measured together with other types of perturbations [2].

The change detection could be offline or online. In online change detection, we are only interested in detecting the change as quickly as possible (e.g., to minimize the detection delay with fixed mean time between false alarms), and the estimate of the time when the change occurs is not of importance. In offline change detection, we assume that the whole observation sequence is available at once and finding the estimate of the time of change could be one of the goals of the detection method. In this paper, we limit our concern to online detection of abrupt changes.

The change detection methods that we consider here can

be classified under the general name of likelihood ratio (LR) methods. Cumulative sum (CUSUM) and generalized LR (GLR) tests are among these methods. CUSUM was first proposed by Page [1]. The most basic CUSUM algorithm assumes that the observation signal is a sequence of stochastic variables which are independent and identically distributed (i.i.d.) with known common probability density function before the change time, and i.i.d. with another known probability density function after the change time. In the CUSUM algorithm, the log-likelihood ratio for the observation from time  $i$  to time  $k$  is calculated and its difference with its current minimum is compared with a certain threshold. If this difference exceeds the threshold, an alarm is issued.

Properties of the CUSUM algorithm have been studied extensively. Its most important property is the asymptotic optimality, which was first proven in [3]. More precisely, CUSUM is optimal, with respect to the worst mean delay, when the mean time between false alarms goes to infinity. This asymptotic point of view is convenient in practice because a low rate of false alarms is always desirable.

In the case of unknown system parameters after change, the GLR algorithm can be used as a generalization of the CUSUM algorithm. Since, in this algorithm, the exact information of the change pattern is not known, the LR is maximized over all possible change patterns.<sup>1</sup>

---

<sup>1</sup>If the maximum does not exist, the supremum of the LR should be calculated.

For stochastic systems with linear dynamics and linear observations, the observation sequence is not i.i.d. Therefore, the regular CUSUM algorithm cannot be applied for detection of changes in such systems. However, if such systems are driven by Gaussian noise, the innovation process associated with the system is known to be a sequence of independent random variables. The regular CUSUM algorithm or its more general counterpart, GLR, can be applied to this innovation process [2, 4].

In this paper, we are interested in the change detection problem for stochastic systems with nonlinear dynamics and observations. We show that for such systems, the complexity of the CUSUM algorithm grows with respect to time. This growth in complexity cannot be tolerated in practical problems. Therefore, instead of the statistic used in the CUSUM algorithm, we introduce an alternative statistic. We show that with this statistic, the calculation of the LR can be done recursively and the computational complexity of the method stays constant with respect to time.

Unlike the linear case, change detection for nonlinear stochastic systems has not been investigated in any depth. In the cases where a nonlinear system experiences a sudden change, linearization and change detection methods for linear systems are the main tools for solving the change detection problem (see, e.g., [5]). The reason this subject has not been pursued is clear; even when there is no change, the estimation of the state of the system, given the observations, results in an infinite-dimensional nonlinear filter [6], and the change in the system can only make the estimation harder.

In the last decade, there has been an increasing interest in simulation-based nonlinear filtering methods. These filtering methods are based on a gridless approximation of the conditional density of the state, given the observations. Gridless simulation-based filtering, now known by many different names such as particle filtering (PAF) [7, 8], the condensation algorithm [9], the sequential Monte Carlo (SMC) method [10], and Bayesian bootstrap filtering [11], was first introduced by Gordon et al. [11] and then it was rediscovered independently by Isard and Blake [9] and Kitagawa [12].

The theoretical results regarding the convergence of the approximate conditional density given by PAF to the true conditional density (in some proper sense) suggest that this method is a strong alternative for nonlinear filtering [7]. The advantage of this method over the nonlinear filter is that PAF is a finite-dimensional filter. The authors believe that PAF and its modifications are a starting point to study change detection for nonlinear stochastic systems. In this paper, we use the results in [13] and we develop a new change detection method for nonlinear stochastic systems.

In [13], we showed that when the number of satellites is below a critical number, linearization methods such as extended Kalman filtering (EKF) result in an unacceptable position error for an integrated inertial navigation system/global positioning system (INS/GPS). We also showed that the approximate nonlinear filtering methods, the projection particle filter [13], in particular, are capable of providing an acceptable estimate of the position in the same situation.

If the carrier phase is used for position information in an integrated INS/GPS, one sudden change that happens rather often is the cycle slip. A cycle slip happens when the phase of the received signal estimated by the phase lock loop in the receiver has a sudden jump. An integrated INS/GPS with carrier phase receiver is used as an application for the method introduced in this paper for detection of a cycle slip with known strength.

In Section 2, we state the approximate nonlinear filtering method used in this paper. In Section 3, we briefly define the change detection problem. In Section 4, we review the CUSUM algorithm for linear systems with additive changes. Then, in Sections 5 and 6, we present a new change detection method for nonlinear stochastic systems. In Section 7, we lay out the formulation for an integrated INS/GPS. In Section 8, we present some simulation results. In Section 9, we summarize the results and lay out the future work.

## 2. APPROXIMATE NONLINEAR FILTERING

Consider the dynamical system

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathbf{f}_k(\mathbf{x}_k) + G_k(\mathbf{x}_k)\mathbf{w}_k, \\ \mathbf{y}_k &= \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k,\end{aligned}\quad (1)$$

where the distribution of  $\mathbf{x}_0$  is given,  $\mathbf{x}_k \in \mathbb{R}^n$ ,  $\mathbf{y}_k \in \mathbb{R}^d$ , and  $\mathbf{w}_k \in \mathbb{R}^q$ , and  $\mathbf{v}_k \in \mathbb{R}^d$  are white noise processes with known statistics, and the functions  $\mathbf{f}_k(\cdot)$  and  $\mathbf{h}_k(\cdot)$  and the matrix  $G_k(\cdot)$  have the proper dimensions. The noise processes  $\mathbf{w}_k$ ,  $\mathbf{v}_k$ ,  $k = 0, 1, \dots$ , and the initial condition  $\mathbf{x}_0$  are assumed independent.

We assume the initial distribution for  $\mathbf{x}_0$  is given. The goal is to find the conditional distribution of the state, given the observation, that is,  $P_k(d\mathbf{x}_k | \mathcal{Y}_1^k)$ , where  $\mathcal{Y}_1^k = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k\}$  is the observation up to and including time  $k$ . The propagation of the conditional distribution, at least conceptually, can be expressed as follows [6].

Step (1). Initialization:

$$P_0(d\mathbf{x}_0 | \mathbf{y}_0) = P(d\mathbf{x}_0). \quad (2)$$

Step (2). Diffusion:

$$\begin{aligned}P_{(k+1)^-}(d\mathbf{x}_{k+1} | \mathcal{Y}_1^k) \\ = \int P(d\mathbf{x}_{k+1} | \mathbf{x}_k) P_k(d\mathbf{x}_k | \mathcal{Y}_1^k).\end{aligned}\quad (3)$$

Step (3). Bayes' rule update:

$$\begin{aligned}P_{(k+1)}(d\mathbf{x}_{k+1} | \mathcal{Y}_1^{k+1}) \\ = \frac{p(\mathbf{y}_{k+1} | \mathbf{x}_{k+1}) P_{(k+1)^-}(d\mathbf{x}_{k+1} | \mathcal{Y}_1^k)}{\int p(\mathbf{y}_{k+1} | \mathbf{x}_{k+1}) P_{(k+1)^-}(d\mathbf{x}_{k+1} | \mathcal{Y}_1^k)}.\end{aligned}\quad (4)$$

Step (4).  $k \leftarrow k + 1$ ; go to Step (2).

We have assumed that  $P(d\mathbf{y}_{k+1} | \mathbf{x}_{k+1}) = p(\mathbf{y}_{k+1} | \mathbf{x}_{k+1}) d\mathbf{y}_{k+1}$  and  $p(\mathbf{y}_{k+1} | \mathbf{x}_{k+1})$  is the conditional density of the observation, given the state at time  $k + 1$ .

The conditional distribution given by the above steps is exact, but in general, it can be viewed as an infinite-dimensional filter, thus not implementable. PAF, in brief,

Step (1). Initialization.  
Sample  $N$  i.i.d. random vectors  $\mathbf{x}_0^1, \dots, \mathbf{x}_0^N$  with the initial distribution  $P_0(d\mathbf{x})$ .

Step (2). Diffusion.  
Find  $\hat{\mathbf{x}}_{k+1}^1, \dots, \hat{\mathbf{x}}_{k+1}^N$  from the given  $\mathbf{x}_k^1, \dots, \mathbf{x}_k^N$ , using the dynamic rule:

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k) + G_k(\mathbf{x}_k)\mathbf{v}_k.$$

Step (3). Use Bayes' rule and find the empirical distribution

$$P_{k+1}^N(d\mathbf{x}) = \sum_{j=1}^N \frac{p(\mathbf{y}_{k+1} | \hat{\mathbf{x}}_{k+1}^j)}{\sum_{i=1}^N p(\mathbf{y}_{k+1} | \hat{\mathbf{x}}_{k+1}^i)} \delta_{\hat{\mathbf{x}}_{k+1}^j}(d\mathbf{x}).$$

Step (4). Resampling.  
Sample  $\mathbf{x}_{k+1}^1, \dots, \mathbf{x}_{k+1}^N$  according to  $P_{k+1}^N(d\mathbf{x})$ .

Step (5).  $k \leftarrow k + 1$ ; go to Step (2).

ALGORITHM 1: Particle filtering.

is an approximation method that mimics the above calculations with a finite number of operations using the Monte Carlo method. Algorithm 1 shows one manifestation of PAF [7, 11].

It is customary to call  $\mathbf{x}_k^1, \dots, \mathbf{x}_k^N$  particles. The key idea in PAF is to eliminate the particles that have low importance weights  $p(\mathbf{y}_k | \mathbf{x}_k)$  and to multiply particles having high importance weights [11, 14]. The surviving particles are thus approximately distributed according to  $P_k^N(d\mathbf{x})$ . This automatically makes the approximation one of better resolution in the areas where the probability is higher.

In the simulations in this paper, we use a modified version of the classical PAF method called projection PAF. For completeness sake, we repeat the algorithm that was given in [13]. In projection PAF, we assume that the conditional density of the state of the system, given the observation, is close to an exponential family of densities  $\mathcal{F}$  defined as follows.<sup>2</sup>

*Definition 1* (Brigo [15]). Let  $\{c_1(\cdot), \dots, c_p(\cdot)\}$  be affinely independent<sup>3</sup> scalar functions defined on  $k$  dimensional Euclidean space  $\mathbb{R}^k$ . Assume that

$$\Theta_0 = \left\{ \theta \in \mathbb{R}^p : \Upsilon(\theta) = \log \int \exp(\theta^T \mathbf{c}(\mathbf{x})) d\mathbf{x} < \infty \right\} \quad (5)$$

is a convex set with a nonempty interior, where  $\mathbf{c}(\mathbf{x}) = [c_1(\mathbf{x}), \dots, c_p(\mathbf{x})]^T$ . Then  $\mathcal{F}$ , defined as

$$\begin{aligned} \mathcal{F} &= \{p(\cdot, \theta), \theta \in \Theta\}, \\ p(\mathbf{x}, \theta) &:= \exp[\theta^T \mathbf{c}(\mathbf{x}) - \Upsilon(\theta)], \end{aligned} \quad (6)$$

where  $\Theta \subseteq \Theta_0$  is open, is called an exponential family of probability densities.

<sup>2</sup>For details of the assumptions and the convergence results for the projection PAF, see [13].

<sup>3</sup> $\{c_1, \dots, c_p\}$  are affinely independent if for distinct points  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{p+1}$ ,  $\sum_{i=1}^{p+1} \lambda_i c(\mathbf{x}_i) = 0$  and  $\sum_{i=1}^{p+1} \lambda_i = 0$  imply  $\lambda_1 = \lambda_2 = \dots = \lambda_{p+1} = 0$  [16].

Step (1). Initialization.  
Sample  $N$  i.i.d. random vectors  $\mathbf{x}_0^1, \dots, \mathbf{x}_0^N$  with the density  $p_0(\mathbf{x})$ .

Step (2). Diffusion.  
Find  $\hat{\mathbf{x}}_{k+1}^1, \dots, \hat{\mathbf{x}}_{k+1}^N$  from the given  $\mathbf{x}_k^1, \dots, \mathbf{x}_k^N$ , using the dynamic rule:

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k) + G_k(\mathbf{x}_k)\mathbf{v}_k.$$

Step (3). Find the MLE of  $\hat{\theta}_{(k+1)^-}$ , given  $\hat{\mathbf{x}}_{k+1}^1, \dots, \hat{\mathbf{x}}_{k+1}^N$ :

$$\hat{\theta}_{(k+1)^-} = \arg \max_{\theta} \prod_{i=1}^N \exp(\theta^T \mathbf{c}(\hat{\mathbf{x}}_{k+1}^i) - \Upsilon(\theta)).$$

Step (4). Use Bayes' rule

$$\begin{aligned} p(\mathbf{x}, \hat{\theta}_{(k+1)}) &= \frac{\exp(\hat{\theta}_{(k+1)}^T \mathbf{c}(\mathbf{x}) - \Upsilon(\hat{\theta}_{(k+1)})) p(\mathbf{y}_{k+1} | \mathbf{x})}{\int \exp(\hat{\theta}_{(k+1)}^T \mathbf{c}(\mathbf{x}) - \Upsilon(\hat{\theta}_{(k+1)})) p(\mathbf{y}_{k+1} | \mathbf{x}) d\mathbf{x}}. \end{aligned}$$

Step (5). Resampling.  
Sample  $\mathbf{x}_{k+1}^1, \dots, \mathbf{x}_{k+1}^N$  according to  $p(\mathbf{x}, \hat{\theta}_{(k+1)})$ .

Step (6).  $k \leftarrow k + 1$ ; go to Step (2).

ALGORITHM 2: Projection particle filtering for an exponential family of densities.

With this definition for the exponential family of densities, the projection PAF algorithm is stated as in Algorithm 2.

### 3. CHANGE DETECTION: PROBLEM DEFINITION

Online detection of a change can be formulated as follows [2]. Let  $\mathcal{Y}_1^k$  be a sequence of observed random variables with conditional densities  $p_{\theta}(\mathbf{y}_k | \mathbf{y}_{k-1}, \dots, \mathbf{y}_1)$ . Before the unknown change time  $t_0$ , the parameter of the conditional density  $\theta$  is constant and equal to  $\theta_0$ . After the change, this parameter is equal to  $\theta_1$ . In online change detection, one is interested in detecting the occurrence of such a change. The exact time and the estimation of the parameters before and after the change are not required. In the case of multiple changes, we assume that the changes are detected fast enough so that in each time instance, only one change can be considered. Online change detection is performed by a stopping rule [2]

$$t_a = \inf \{k : g_k(\mathcal{Y}_1^k) \geq \lambda\}, \quad (7)$$

where  $\lambda$  is a threshold,  $\{g_k\}_{k \geq 1}$  is a family of functions, and  $t_a$  is the alarm time, that is, the time when change is detected.

If  $t_a < t_0$ , then a false alarm has occurred. The criterion for choosing the parameter  $\lambda$  and the family of functions  $\{g_k\}_{k \geq 1}$  is to minimize the detection delay for the fixed mean time between false alarms.

### 4. ADDITIVE CHANGES IN LINEAR DYNAMICAL SYSTEMS

Consider the system

$$\begin{aligned} \mathbf{x}_{k+1} &= F_k \mathbf{x}_k + G_k \mathbf{w}_k + \Gamma_k \Upsilon_{\mathbf{x}}(k, t_0), \\ \mathbf{y}_k &= H_k \mathbf{x}_k + \mathbf{v}_k + \Xi_k \Upsilon_{\mathbf{y}}(k, t_0), \end{aligned} \quad (8)$$

where  $\mathbf{x}_k \in \mathbb{R}^n$ ,  $\mathbf{y}_k \in \mathbb{R}^d$ , and  $\mathbf{w}_k \in \mathbb{R}^q$  and  $\mathbf{v}_k \in \mathbb{R}^d$  are white noise processes with known statistics.  $F_k$ ,  $G_k$ ,  $H_k$ ,  $\Gamma_k$ , and  $\Xi_k$  are matrices of proper dimensions, and  $Y_x(k, t_0)$  and  $Y_y(k, t_0)$  are the dynamic profiles of the assumed changes, of dimensions  $\tilde{n} \leq n$  and  $\tilde{d} \leq d$ , respectively.  $\mathbf{w}_k$  and  $\mathbf{v}_k$  are white Gaussian noise processes, independent of the initial condition  $\mathbf{x}_0$ . It is assumed that  $Y_x(k, t_0) = 0$  and  $Y_y(k, t_0) = 0$  for  $k < t_0$ , but we do not necessarily have the exact knowledge of the dynamic profile and the gain matrices  $\Gamma_k$  and  $\Xi_k$ .

For the case of known parameters before and after change, the CUSUM [2] algorithm can be used, and it is well known that the change detection method has the following form:

$$\begin{aligned} t_a &= \min \{k \geq 1 \mid g_k \geq \lambda\}, \\ g_k &= \max_{1 \leq j \leq k} S_j^k, \\ S_j^k &= \ln \frac{\prod_{i=j}^k P_{\rho(i,j)}(\epsilon_i)}{\prod_{i=j}^k p_0(\epsilon_i)}, \end{aligned} \quad (9)$$

where  $\epsilon_i$  is the innovation process calculated using Kalman filtering assuming that no change occurred, and  $\rho(i, j)$  is the mean of the innovation process at time  $j$  conditioned on the change occurring at the time  $i$ .  $p_0$  and  $p_{\rho(\cdot, \cdot)}$  are Gaussian densities with means 0 and  $\rho(\cdot, \cdot)$ , respectively. The covariance matrix for these two densities is the same and is calculated using Kalman filtering.  $S_j^k$  is the LR between two hypotheses: change occurrence at  $j$  and no change occurrence.

When the parameter after change is not known, GLR can be used to calculate  $g_k$  [4]:

$$g_k = \max_{1 \leq j \leq k} \sup_{Y_x, Y_y} S_j^k. \quad (10)$$

The solution for (10) is well known and can be found in many references (see [2]).

Similar to nonlinear filtering, change detection for nonlinear stochastic systems results in an algorithm that is infinite dimensional. Linearization techniques, whenever applicable, are the main approximation tool for studying the change detection problem for nonlinear systems. Although linearization techniques are computationally efficient, they are not always applicable. In the sections to come, we propose a new method based on nonlinear PAF that can be used for change detection for nonlinear stochastic systems.

## 5. NONLINEAR CHANGE DETECTION: PROBLEM SETUP

Consider the nonlinear system

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{f}_k^{i_k}(\mathbf{x}_k) + \mathbf{G}_k^{i_k}(\mathbf{x}_k)\mathbf{w}_k, \\ \mathbf{y}_k &= \mathbf{h}_k^{i_k}(\mathbf{x}_k) + \mathbf{v}_k, \end{aligned} \quad (11)$$

where

$$i_k = \begin{cases} 0, & k < t_0, \\ 1, & k \geq t_0, \end{cases} \quad (12)$$

and the functions  $\mathbf{f}_k^0(\cdot)$ ,  $\mathbf{f}_k^1(\cdot)$ ,  $\mathbf{h}_k^0(\cdot)$ ,  $\mathbf{h}_k^1(\cdot)$  and the matrices  $\mathbf{G}_k^0(\cdot)$ ,  $\mathbf{G}_k^1(\cdot)$  have the proper dimensions. The sudden change occurs when  $i_k$  changes from 0 to 1.

In this setup,  $S_j^k$  can be written as follows:

$$S_j^k = \ln \frac{p(\mathcal{Y}_j^k | \mathcal{Y}_1^{j-1}, t_0 = j)}{p(\mathcal{Y}_j^k | \mathcal{Y}_1^{j-1}, t_0 > k)}. \quad (13)$$

Writing (13) in a recursive form, we get

$$p(\mathcal{Y}_j^k | \mathcal{Y}_1^{j-1}, t_0 = j) = \prod_{i=j}^k p(\mathbf{y}_i | \mathcal{Y}_1^{i-1}, t_0 = j), \quad (14)$$

where  $p(\mathbf{y}_i | \mathcal{Y}_1^{i-1}, t_0 = j)$  can be written as follows:

$$p(\mathbf{y}_i | \mathcal{Y}_1^{i-1}, t_0 = j) = \int p(\mathbf{y}_i | \mathbf{x}_i) P(d\mathbf{x}_i | \mathcal{Y}_1^{i-1}, t_0 = j). \quad (15)$$

To find  $P(d\mathbf{x}_i | \mathcal{Y}_1^{i-1}, t_0 = j)$  or equivalently to find the density  $p(\mathbf{x}_i | \mathcal{Y}_1^{i-1}, t_0 = j)$ <sup>4</sup> in (15), one needs to find an approximation for the corresponding nonlinear filter. We assume that this approximation is done using either PAF or projection PAF [13].

To calculate the LR in (13), we must calculate the conditional densities of the state, given the observation for two hypotheses (change occurrence at  $j$  and change occurrence after  $k$ ). This means that two nonlinear filters should be implemented just to compare these two hypotheses. Therefore, it is clear that to use an algorithm similar to (9),  $k$  parallel nonlinear filters should be implemented. In Figure 1, we see that the computational complexity of the CUSUM algorithm grows linearly with respect to time. In most applications, this growth is not desirable. One possible way to approximate the CUSUM algorithm is to truncate the branches that fork from the main branch in Figure 1. We will explain this truncation procedure and its technical difficulties in the next few lines.

Recall that the main branch (horizontal) and the branches forked from it in Figure 1 are representing a series of nonlinear filters with specific assumptions on the change time. The dynamic and observation equations for all forked branches are the same and the only difference is the initial density. If the conditional density of the state, given the observation for a nonlinear system, with the wrong initial density converges (in some meaningful way) to the true conditional density (initialized by the true initial density), we say that the corresponding nonlinear filter is *asymptotically stable* [17].

For asymptotically stable nonlinear filters, the forked branches in Figure 1 converge to a single branch, therefore, there is no need to implement several parallel nonlinear filters. In other words, after each branching, the independent nonlinear filter is used for a period of time and then this branch converges to the branches that have forked earlier,

<sup>4</sup>If the density exists.

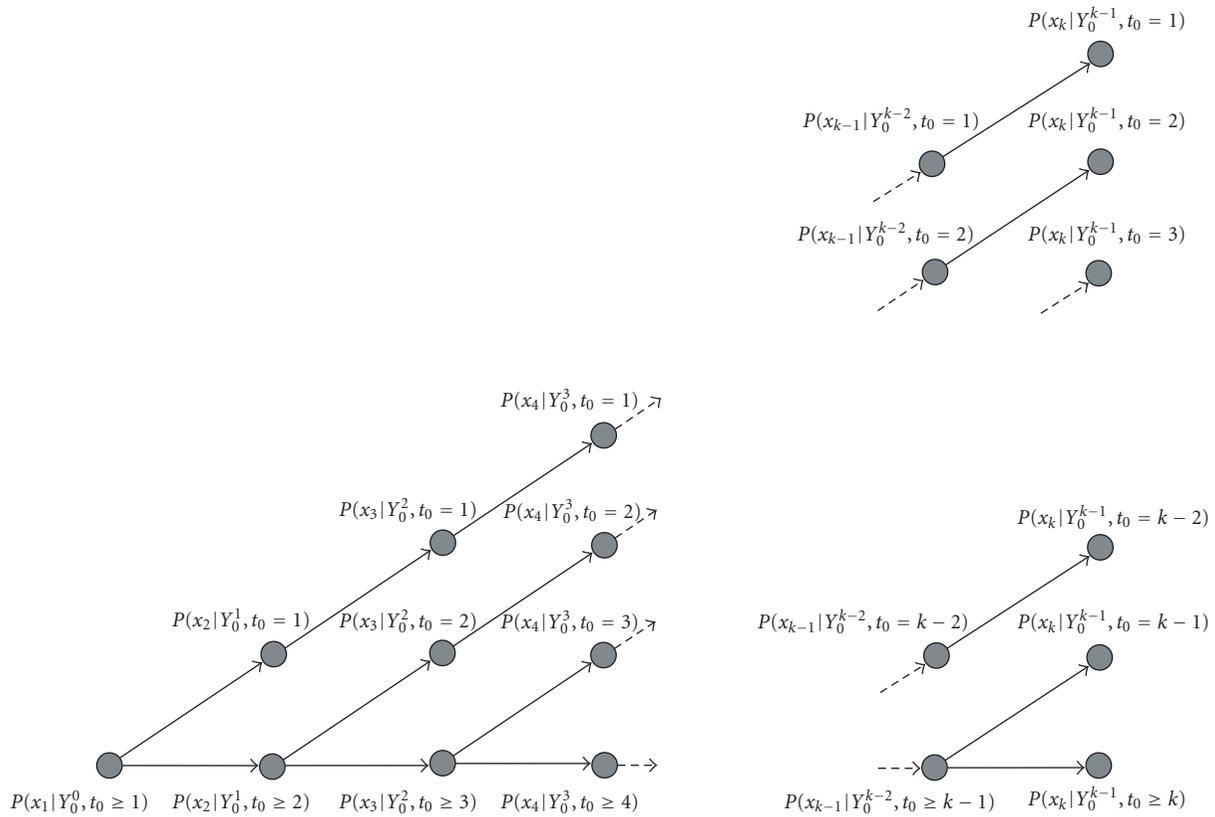


FIGURE 1: Combination of nonlinear filters used in the CUSUM change detection algorithm.

that is, joins them. The time needed for the branch of the independent nonlinear filter to join the other forked branches depends on the convergence rate and the target accuracy of the approximation.

Although the procedure mentioned above can be used for asymptotically stable nonlinear filters, there are several problems associated with this method. The known theoretical results for identifying asymptotically stable filters are limited to either requiring ergodicity and the compactness of the state space [18, 19, 20], or very special cases of the observation equation [17]. The rate of convergence of the filters in different branches is another potential shortcoming of the mentioned procedure. If the convergence rate is low in comparison to the rate of parameter change in the system, then the algorithm cannot take advantage of this convergence.

### 6. NONLINEAR CHANGE DETECTION: NONGROWING COMPUTATIONAL COMPLEXITY

In this section, we introduce a new statistic to overcome the problem of growing computational complexity for the change detection method. We emphasize that the parameters of the system before and after change are assumed to be known. Therefore, the conditional density of the state of the system, given the observation, can be calculated using a nonlinear filter. We show that this statistic can be calculated recursively.

Consider the following statistic:

$$T_j^k = \ln \frac{p(\mathcal{Y}_j^k | \mathcal{Y}_1^{j-1}, t_0 \in \{j, \dots, k\})}{p(\mathcal{Y}_j^k | \mathcal{Y}_1^{j-1}, t_0 > k)}. \quad (16)$$

The change detection algorithm based on statistic  $T_j^k$  can be presented as

$$t_a = \min \{k \geq j \mid T_j^k \geq \lambda \text{ or } T_j^k \leq -\alpha\}, \quad (17)$$

where  $j$  is the last time when  $T_j^k \geq \lambda$  or  $T_j^k \leq -\alpha$ , and  $\lambda > 0$  and  $\alpha > 0$  are chosen such that the detection delay is minimized for a fixed mean time between two false alarms.

For the rest of this paper, we assume that the probability of change at time  $i$  condition on no change before  $i$  is  $q$ , that is,

$$P(t_0 = i \mid t_0 \geq i) = q. \quad (18)$$

Without loss of generality and for simplifying the notation, we assume that  $j = 1$ . To calculate the statistic  $T_1^k$ , it is sufficient to find  $P(d\mathbf{x}_k, t_0 \leq k | \mathcal{Y}_1^k)$  and  $P(d\mathbf{x}_k, t_0 > k | \mathcal{Y}_1^k)$ . Then  $T_1^k$  is given by

$$\begin{aligned} T_1^k &= \ln \left( \frac{(1-q)^k \int P(d\mathbf{x}_k, t_0 \leq k | \mathcal{Y}_1^k)}{(1-(1-q)^k) \int P(d\mathbf{x}_k, t_0 > k | \mathcal{Y}_1^k)} \right) \\ &= \ln \left( \frac{(1-q)^k}{(1-(1-q)^k)} \frac{W_k^1}{W_k^0} \right), \end{aligned} \quad (19)$$

where  $W_k^0 = \int P(d\mathbf{x}_k, t_0 > k | \mathcal{Y}_1^k)$  and  $W_k^1 = \int P(d\mathbf{x}_k, t_0 \leq k | \mathcal{Y}_1^k)$ . Therefore, to calculate  $T_1^k$  recursively, it is sufficient to

calculate  $P(d\mathbf{x}_k, t_0 \leq k | \mathcal{Y}_1^k)$  and  $P(d\mathbf{x}_k, t_0 > k | \mathcal{Y}_1^k)$  recursively.  $P(d\mathbf{x}_{k+1}, t_0 \leq k+1 | \mathcal{Y}_1^{k+1})$  can be written as follows:

$$\begin{aligned}
 & P(d\mathbf{x}_{k+1}, t_0 \leq k+1 | \mathcal{Y}_1^{k+1}) \\
 &= \frac{P(d\mathbf{x}_{k+1}, t_0 \leq k+1, \mathbf{y}_{k+1} | \mathcal{Y}_1^k)}{\int P(d\mathbf{x}_{k+1}, t_0 \leq k+1, \mathbf{y}_{k+1} | \mathcal{Y}_1^k) + \int P(d\mathbf{x}_{k+1}, t_0 > k+1, \mathbf{y}_{k+1} | \mathcal{Y}_1^k)} \\
 &= \frac{p(\mathbf{y}_{k+1} | \mathbf{x}_{k+1}, t_0 \leq k+1)P(d\mathbf{x}_{k+1}, t_0 \leq k+1 | \mathcal{Y}_1^k)}{\int p(\mathbf{y}_{k+1} | \mathbf{x}_{k+1}, t_0 \leq k+1)P(d\mathbf{x}_{k+1}, t_0 \leq k+1 | \mathcal{Y}_1^k) + \int p(\mathbf{y}_{k+1} | \mathbf{x}_{k+1}, t_0 > k+1)P(d\mathbf{x}_{k+1}, t_0 > k+1 | \mathcal{Y}_1^k)} \\
 &= \frac{p_1(\mathbf{y}_{k+1} | \mathbf{x}_{k+1})P(d\mathbf{x}_{k+1}, t_0 \leq k+1 | \mathcal{Y}_1^k)}{\int p_1(\mathbf{y}_{k+1} | \mathbf{x}_{k+1})P(d\mathbf{x}_{k+1}, t_0 \leq k+1 | \mathcal{Y}_1^k) + \int p_0(\mathbf{y}_{k+1} | \mathbf{x}_{k+1})P(d\mathbf{x}_{k+1}, t_0 > k+1 | \mathcal{Y}_1^k)},
 \end{aligned} \tag{20}$$

where  $p_1(\mathbf{y}_{k+1} | \mathbf{x}_{k+1})$  is the conditional density of the state  $\mathbf{x}_{k+1}$ , given the observation  $\mathbf{y}_{k+1}$ , under the hypothesis that change has occurred and  $p_0(\mathbf{y}_{k+1} | \mathbf{x}_{k+1})$  is the same condi-

tional density under the hypothesis that no change has occurred. Similarly, for  $P(d\mathbf{x}_{k+1}, t_0 > k+1 | \mathcal{Y}_1^{k+1})$ , we have the following:

$$\begin{aligned}
 & P(d\mathbf{x}_{k+1}, t_0 > k+1 | \mathcal{Y}_1^{k+1}) \\
 &= \frac{p_0(\mathbf{y}_{k+1} | \mathbf{x}_{k+1})P(d\mathbf{x}_{k+1}, t_0 > k+1 | \mathcal{Y}_1^k)}{\int p_1(\mathbf{y}_{k+1} | \mathbf{x}_{k+1})P(d\mathbf{x}_{k+1}, t_0 \leq k+1 | \mathcal{Y}_1^k) + \int p_0(\mathbf{y}_{k+1} | \mathbf{x}_{k+1})P(d\mathbf{x}_{k+1}, t_0 > k+1 | \mathcal{Y}_1^k)}.
 \end{aligned} \tag{21}$$

Also, we have

$$\begin{aligned}
 & P(d\mathbf{x}_{k+1}, t_0 \leq k+1 | \mathcal{Y}_1^k) \\
 &= \int P(d\mathbf{x}_{k+1}, t_0 \leq k+1 | \mathbf{x}_k, t_0 \leq k, \mathcal{Y}_1^k)P(d\mathbf{x}_k, t_0 \leq k | \mathcal{Y}_1^k) \\
 &\quad + \int P(d\mathbf{x}_{k+1}, t_0 \leq k+1 | \mathbf{x}_k, t_0 > k, \mathcal{Y}_1^k)P(d\mathbf{x}_k, t_0 > k | \mathcal{Y}_1^k) \\
 &= W_k^1 \int P_1^{k+1}(d\mathbf{x}_{k+1} | \mathbf{x}_k)P(d\mathbf{x}_k | t_0 \leq k, \mathcal{Y}_1^k) \\
 &\quad + qW_k^0 \int P(d\mathbf{x}_{k+1} | \mathbf{x}_k, t_0 = k+1)P(d\mathbf{x}_k | t_0 > k, \mathcal{Y}_1^k),
 \end{aligned} \tag{22}$$

$$\begin{aligned}
 & P(d\mathbf{x}_{k+1}, t_0 > k+1 | \mathcal{Y}_1^k) \\
 &= (1-q)W_k^0 \int P_0^{k+1}(d\mathbf{x}_{k+1} | \mathbf{x}_k)P(d\mathbf{x}_k | t_0 > k, \mathcal{Y}_1^k).
 \end{aligned} \tag{23}$$

In (22) and (23),  $P_0^{k+1}(d\mathbf{x}_{k+1} | \mathbf{x}_k)$  and  $P_1^{k+1}(d\mathbf{x}_{k+1} | \mathbf{x}_k)$  are the Markov transition kernel under the hypothesis that no change has occurred before  $k+1$  and change has occurred before  $k$ , respectively.  $P(d\mathbf{x}_{k+1} | \mathbf{x}_k, t_0 = k+1)$  is the Markov transition kernel under the hypothesis that the change has occurred at  $k+1$ . For the dynamics in (11), we have  $P(d\mathbf{x}_{k+1} | \mathbf{x}_k, t_0 = k+1) = P_0^{k+1}(d\mathbf{x}_{k+1} | \mathbf{x}_k)$ .

Equations (19)–(23) show that the statistic  $T_1^k$  can be calculated recursively. They also show that in the prediction step of the nonlinear filter at each time instance, only two conditional distributions should be calculated,  $P(d\mathbf{x}_{k+1}, t_0 \leq k+1 | \mathcal{Y}_1^k)$  and  $P(d\mathbf{x}_{k+1}, t_0 > k+1 | \mathcal{Y}_1^k)$ . Therefore, if a PAF method is used to approximate (22) and (23), we only need two sets of particles to approximate these two conditional distributions. In the Bayes' rule update step and the resampling step of the PAF, (20) and (21) are used. One possible way of implementing (19) in (23) using a PAF method is as follows.

In Algorithm 3, the same number of particles is used to find the conditional distribution before and after change. This guarantees that always enough numbers of particles are available for approximating the conditional densities before and after change.

In the remaining sections, we use the introduced statistic  $T_j^k$  for detecting a sudden change in the phase measurement of an integrated INS/GPS.

## 7. INTEGRATED INS/GPS

GPS provides worldwide accurate positioning if four or more satellites are in view of the receiver. Although the satellite

Step (1). Initialization.

Sample  $2N$  i.i.d. random vectors  $\mathbf{x}_0^1, \dots, \mathbf{x}_0^N$  with the weight  $W_0^0/N$ , and  $\mathbf{x}_0^{N+1}, \dots, \mathbf{x}_0^{2N}$  with the weight  $W_0^1/N$  and  $W_0^1 = 1 - W_0^0 = \epsilon$ , with the initial distribution  $P_0(d\mathbf{x}, t_0 > 0)$ . (Since we start with the assumption that no change has happened,  $\epsilon$  is either 0 or a very small number.)

Step (2). Diffusion.

Find  $\hat{\mathbf{x}}_{k+1}^1, \dots, \hat{\mathbf{x}}_{k+1}^N$  from the given  $\mathbf{x}_k^1, \dots, \mathbf{x}_k^N$  using the dynamic rule:

$$\mathbf{x}_{k+1} = \mathbf{f}_k^0(\mathbf{x}_k) + \mathbf{G}_k^0(\mathbf{x}_k)\mathbf{v}_k,$$

and find  $\hat{\mathbf{x}}_{k+1}^{N+1}, \dots, \hat{\mathbf{x}}_{k+1}^{2N}$  from the given  $\mathbf{x}_k^{N+1}, \dots, \mathbf{x}_k^{2N}$  using the dynamic rule:

$$\mathbf{x}_{k+1} = \mathbf{f}_k^1(\mathbf{x}_k) + \mathbf{G}_k^1(\mathbf{x}_k)\mathbf{v}_k.$$

Step (3). Find an estimate for  $P(d\mathbf{x}_{k+1} | t_0 > k, \mathcal{Y}_1^k)$  and  $P(d\mathbf{x}_{k+1} | t_0 \leq k, \mathcal{Y}_1^k)$  either by using empirical distributions

$$P_{(k+1)^-}^N(d\mathbf{x}_{k+1} | t_0 > k, \mathcal{Y}_1^k) = \frac{1}{N} \sum_{j=1}^N \delta_{\hat{\mathbf{x}}_{k+1}^j}(d\mathbf{x}_{k+1}),$$

$$P_{(k+1)^-}^N(d\mathbf{x}_{k+1} | t_0 \leq k, \mathcal{Y}_1^k) = \frac{1}{N} \sum_{j=N+1}^{2N} \delta_{\hat{\mathbf{x}}_{k+1}^j}(d\mathbf{x}_{k+1}),$$

or by using Step (3) of Algorithm 2. Then

$$\begin{aligned} P_{(k+1)^-}^N(d\mathbf{x}_{k+1}, t_0 > k + 1 | \mathcal{Y}_1^k) \\ &= (1 - q) W_k^0 P_{(k+1)^-}^N(d\mathbf{x}_{k+1} | t_0 > k, \mathcal{Y}_1^k), \\ P_{(k+1)^-}^N(d\mathbf{x}_{k+1}, t_0 \leq k + 1 | \mathcal{Y}_1^k) \\ &= q W_k^0 P_{(k+1)^-}^N(d\mathbf{x}_{k+1} | t_0 > k, \mathcal{Y}_1^k) \\ &\quad + W_k^1 P_{(k+1)^-}^N(d\mathbf{x}_{k+1} | t_0 \leq k, \mathcal{Y}_1^k). \end{aligned}$$

Step (4). Use Bayes' rule in (20) and (21) to calculate  $P_{k+1}^N(d\mathbf{x}_{k+1}, t_0 > k + 1 | \mathcal{Y}_1^{k+1})$  and  $P_{k+1}^N(d\mathbf{x}_{k+1}, t_0 \leq k + 1 | \mathcal{Y}_1^{k+1})$ . Then set  $W_{k+1}^0 = \int P_{k+1}^N(d\mathbf{x}_{k+1}, t_0 > k + 1 | \mathcal{Y}_1^{k+1})$  and  $W_{k+1}^1 = \int P_{k+1}^N(d\mathbf{x}_{k+1}, t_0 \leq k + 1 | \mathcal{Y}_1^{k+1})$ .

Step (5). Resampling.

Sample  $\mathbf{x}_{k+1}^1, \dots, \mathbf{x}_{k+1}^N$  according to  $P_{k+1}^N(d\mathbf{x}_{k+1}, t_0 > k + 1 | \mathcal{Y}_1^{k+1})/W_{k+1}^0$ .

Sample  $\mathbf{x}_{k+1}^{N+1}, \dots, \mathbf{x}_{k+1}^{2N}$  according to  $P_{k+1}^N(d\mathbf{x}_{k+1}, t_0 \leq k + 1 | \mathcal{Y}_1^{k+1})/W_{k+1}^1$ .

Step (6).  $k \leftarrow k + 1$ ; go to Step (2).

ALGORITHM 3: Change detection using particle filtering.

constellation guarantees availability of four or more satellites worldwide, natural or man-made obstacles can block the satellite signals easily. Integrating dead reckoning or INS with GPS [21, 22, 23, 24] is a method to overcome this vulnerability. Here, INS or the dead reckoning provides positioning that is calibrated by the GPS. In this section, we consider the case of an integrated INS/GPS. In [25], we showed that using nonlinear filtering for positioning is essential. We compared the proposed PAF with regular PAF and EKF.

TABLE 1: Definition of the parameters for WGS84 reference frame.

a	6378137.0 m	Semimajor axis
b	6356752.3142 m	Semiminor axis
$\omega_{ie}$	$7.292115 \times 10^{-5}$	Earth angular velocity
e	$(\sqrt{b(a-b)})/a$	Ellipsoid eccentricity

Using carrier phase measurements enables the differential GPS to reach centimeter-level accuracy. A phase lock loop cannot measure the full-cycle part of the carrier phase. This unmeasured part is called integer ambiguity that requires to be resolved through other means. In this paper, we assume that the integer ambiguity is resolved (see, e.g., [26]). However, the measured phase can experience a sudden change undetected by the phase lock loop. This sudden change is called the cycle slip and if it is undetected by the integrated INS/GPS, it results in an error in the estimated position. We will use the method introduced in this paper to detect such changes.

We consider the observation equation provided by the  $i$ th GPS satellite at time  $k$  to have the form

$$y_k^i = \rho^i(p_x, p_y, p_z) - \rho^i(b_x, b_y, b_z) + n_i(t, t_0) + c\delta + v_k^i, \quad (24)$$

where  $[p_x, p_y, p_z]^T$  and  $[b_x, b_y, b_z]^T$  are the rover and (known) base coordinates at time  $k$ , respectively,  $\rho^i(x_1, x_2, x_3)$  is the distance from point  $[x_1, x_2, x_3]^T$  to satellite  $i$ ,  $\delta$  is a combination of the receiver clock bias in the base and the rover,  $c$  is the speed of light,  $v_k^i$  is the measurement noise for the  $i$ th satellite signal,  $t_0$  is the unknown moment of the cycle slip, and  $n_i(t, t_0) = 0$  for  $t < t_0$  and  $n_i(t, t_0) = n_i$  for  $t \geq t_0$ , where  $n_i$  is the change in phase measurement due to the cycle slip.

The main goal in the simulations is to detect the change in the phase measurement as soon as it happens.

Here we point out that the nonlinearity in measurement is not solely due to the function  $\rho$ . Integrated INS/GPS requires coordinate transformations between INS parameters and GPS parameters, which contributes to the nonlinearity of the measurement.

Parameters of an integrated INS/GPS are expressed in different coordinate systems. The GPS measurements are given in an earth-centered earth-fixed (ECEF) frame [27, 28]. The GPS position is given either in the rectangular coordinate system (see (24)) or in the geodetic coordinate system with the familiar latitude, longitude, and height coordinate vector  $[p_\lambda, p_\phi, p_h]^T$ . For the latter, the earth's geoid is approximated by an ellipsoid. Table 1 shows the defining parameters for the geoid according to the WGS84 reference frame. The parameters and measurements of INS are given in the local geographical frame or in the body frame system, where the velocity is given by the north-east-down velocity vector  $[v_N, v_E, v_D]^T$ . The transformation matrix from the body frame to the local geographical frame is given by the matrix  $R_{b2g}$ . In this paper, we assume the estimation problem for the gyro measurements is solved, hence  $R_{b2g}$  is known.

The GPS clock drift and the INS equations constitute key dynamics in integrated INS/GPS. The INS dynamic equation can be expressed as follows (see [25] for details):

$$\begin{aligned}
 d \begin{pmatrix} p_\lambda \\ p_\phi \\ p_h \end{pmatrix} &= \begin{pmatrix} \frac{1}{R_\lambda + p_h} & 0 & 0 \\ 0 & \frac{1}{(R_\phi + p_h) \cos(p_\lambda)} & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} v_N \\ v_E \\ v_D \end{pmatrix} dt, \\
 d \begin{pmatrix} v_N \\ v_E \\ v_D \end{pmatrix} &= \begin{pmatrix} -\frac{v_E^2}{R_\phi + p_h} \tan(p_\lambda) - 2\omega_{ie} \sin(p_\lambda) v_E \\ \frac{v_E v_N}{R_\lambda + p_h} \tan(\lambda) + \omega_{ie} \sin(p_\lambda) v_N \\ \frac{v_E v_D}{R_\phi + p_h} + 2\omega_{ie} \cos(p_\lambda) v_D \\ -\frac{v_N^2}{R_\lambda + p_h} - \frac{v_E^2}{R_\phi + p_h} - 2\omega_{ie} \cos(p_\lambda) v_E \end{pmatrix} \\
 &\quad + R_{b2g} \left( \begin{pmatrix} \tilde{a}_u \\ \tilde{a}_v \\ \tilde{a}_w \end{pmatrix} + \begin{pmatrix} b_u \\ b_v \\ b_w \end{pmatrix} \right) + \begin{pmatrix} 0 \\ 0 \\ g \end{pmatrix} dt + d\mathbf{w}_t^v, \quad (25)
 \end{aligned}$$

where  $g = 9.780327 \text{ m/s}^2$  is the gravitational acceleration,  $R_\lambda = a(1 - e^2)/(1 - e^2 \sin^2(p_\lambda))^{3/2}$ ,  $R_\phi = a/(1 - e^2 \sin^2(p_\lambda))^{1/2}$ ,  $[\tilde{a}_u, \tilde{a}_v, \tilde{a}_w]^T$  and  $[b_u, b_v, b_w]^T$  are the accelerometer measurement and the accelerometer measurement bias, respectively, both expressed in the body frame, and  $\mathbf{w}^v$  is a vector-valued Brownian motion process with zero mean and known covariance matrix. The measurement bias  $\mathbf{b} = [b_u, b_v, b_w]^T$  has the dynamics

$$d\mathbf{b} = -a_b \mathbf{b} dt + d\mathbf{w}_t^b, \quad (26)$$

where  $\mathbf{w}_t^b$  is a vector-valued Brownian motion with zero mean and known covariance matrix, and  $a_b$  is a small positive constant. The receiver clock drift  $\delta_t$  is represented by the integration of an exponentially correlated random process  $\rho_t$  [24]:

$$\begin{aligned}
 d\rho_t &= -\frac{1}{500} \rho_t dt + d\mathbf{w}_t^\rho, \\
 d\delta_t &= \rho_t dt, \quad (27)
 \end{aligned}$$

where  $\mathbf{w}_t^\rho$  is a zero-mean Brownian motion process with variance  $\sigma_\rho^2 = 10^{-24}$ . This dynamic model is typical for a quartz TCXO with frequency drift rate  $10^{-9} \text{ s/s}$  [24].

## 8. SIMULATIONS AND RESULTS

The dimension of the dynamical system in the simulation is eleven. The state of the dynamical system  $\mathbf{x}$  is given as follows:

$$\mathbf{x} = [p_\lambda, p_\phi, p_h, v_N, v_E, v_D, b_u, b_v, b_w, \rho, \delta]^T. \quad (28)$$

The differential equation describing the dynamics of the system is the combination of the differential equations in (25), (26), and (27). Here we assume that  $a_b = 0.001$  and that the covariance matrices for the Brownian motions in the INS dynamic equations  $\Sigma_b$  and  $\Sigma_v$  are diagonal. To be more specific,  $\Sigma_b = 10^{-6} I$  and  $\Sigma_v = 10^{-4} I$ , where  $I$  is the identity matrix of the proper size. The observation equation is given in (24), where  $y_i$  is one component of the observation vector. The dimension of the observation vector is the same as the number of available satellites. In (24), the observation is given as a function of the position in the ECEF rectangular coordinate system that is then transformed to the geodetic coordinate system [25].

For the simulation, we simply chose an eleven-dimensional Gaussian density for the projection PAF. This choice of density makes the random vector generation easy and computationally affordable. To be able to use the projection PAF, we used the maximum likelihood criteria to estimate the parameters of the Gaussian density before and after Bayes' correction.

We used two Novatel GPS receivers to collect navigation data (April 2, 2000). From this data, we extracted position information for the satellites, the pseudorange, and the carrier phase measurement noise powers for the L1 frequency. From the collected information we generated the pseudorange and the carrier phase data for one static and one moving receiver (base and rover, respectively). We assume that for the carrier phase measurement, the integer ambiguity problem is already solved (see [26] for details). The movement of the INS/GPS platform was simulation based and it was the basis for the measurement data measured by the accelerometers, the gyros, the GPS pseudorange, and the GPS carrier phase data.

As a precursor, we note that in the simulation in [13] we showed that for an integrated INS/GPS when the number of satellites is less than a critical number, projection PAF provides a very accurate estimate of the position, while the position solution given by EKF is unacceptable. In Figures 2 and 3, a comparison of the position estimation error in the rectangular coordinate system for *one typical run* of each method is shown. For that simulation, we assumed that the GPS receiver starts with six satellites. At time  $t = 100$ , the receiver loses the signal from three satellites, and it gains one satellite signal back at  $t = 400$ . From these two figures, it is clear that when the number of satellites in view is below a certain number (here four satellites), the EKF is unable to provide a reasonable estimate of the position for the integrated INS/GPS. Since the error of the position estimate of linearization methods is unacceptable even when no change in the phase measurement occurs, using these methods in the presence of an abrupt change is fruitless as well.

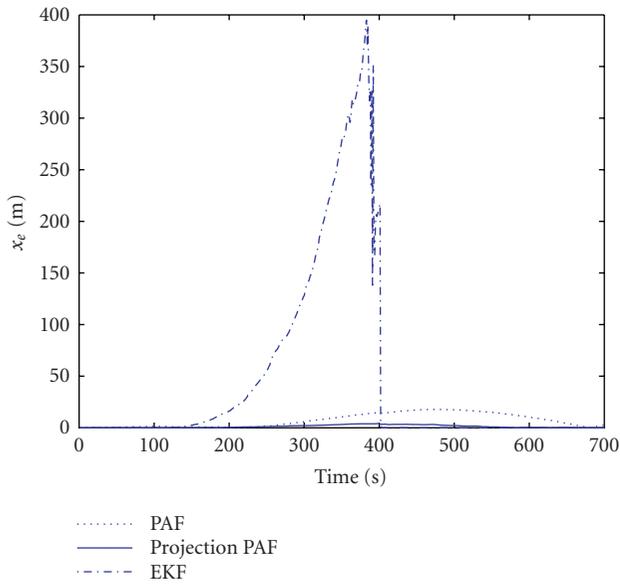


FIGURE 2: Position estimation error (= Euclidean distance between Cartesian position vector and its estimate) for three methods: EKF, PAF, and projection PAF. The system starts with six satellites in view. At  $t = 100$  seconds, the signals from three satellites are lost. At  $t = 400$  seconds, the system regains the signal from one satellite.

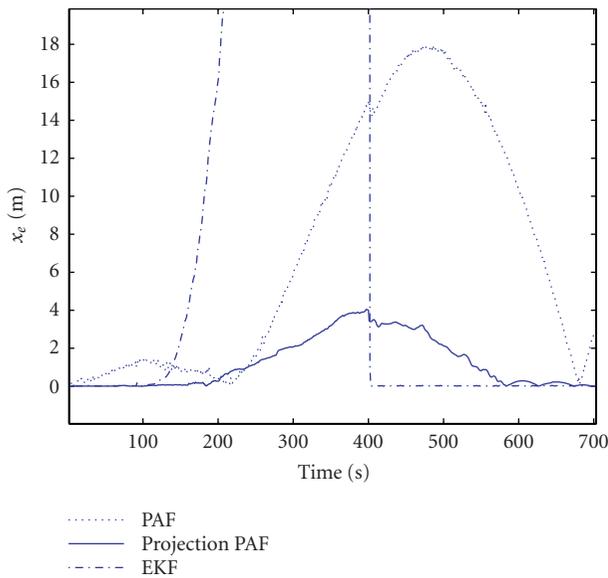


FIGURE 3: Details of Figure 2, where the difference between the projection PAF method and the PAF method is clear.

To apply our method described in (17) for sudden phase change detection in an integrated INS/GPS, we use projection PAF as our nonlinear filtering method. We use the CUSUM algorithm to evaluate the proposed changed detection scheme. We compare the statistic in (17) with that of the CUSUM algorithm. We wish to emphasize that in the example given in this section, we assume that the parameter of change, before and after the change, is known and the only

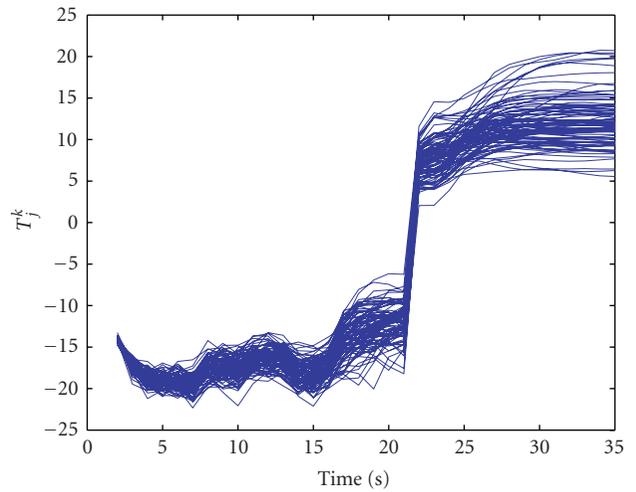


FIGURE 4: The plot of  $T_j^k$  with respect to time for 100 independent runs of an INS/GSP system. At time  $t = 15$ , the receiver loses three satellites. The cycle slip in channel one occurred at  $t = 20$  seconds.

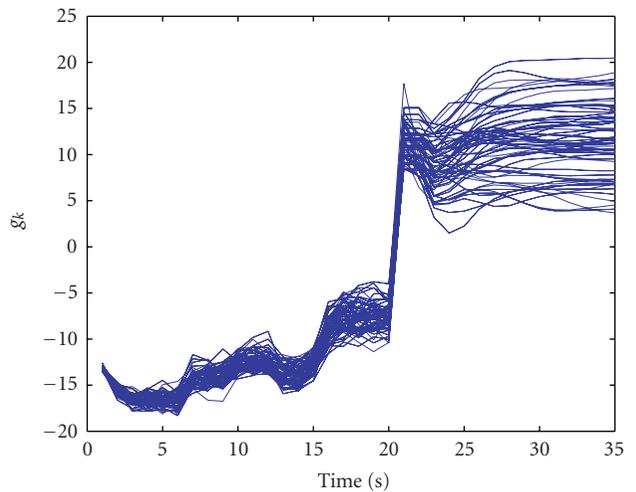


FIGURE 5: The plot of  $g_k$  with respect to time for 100 independent runs of an INS/GSP system.  $g_k$  is the statistics used in the CUSUM algorithm. At time  $t = 15$ , the receiver loses three satellites. The cycle slip in channel one occurred at  $t = 20$  seconds.

unknown parameter is the change time. In the future, we will address the more general problem of unknown change parameters.

For the simulation in this paper, we assumed that the phase lock loop associated to satellite one experiences a cycle slip at time  $t = 20$  and the phase changes suddenly. The size of the change is one cycle. We assumed that the GPS receiver starts with six satellites. At time  $t = 15$ , the receiver loses three satellites (we eliminate them from the data). We used Algorithm 2 to calculate the statistic in the CUSUM algorithm  $g_k$  and Algorithm 3 and projection PAF to calculate  $T_j^k$ . In Figure 4, we have plotted  $T_j^k$  with respect to time for

100 independent runs. In Figure 5, we have plotted the statistic  $g_k$  for the same 100 independent runs. The figures show that there are sudden changes both in  $T_j^k$  and  $g_k$  when a cycle slip occurs, and this is true for all 100 runs. These figures also show that for this simulation, the performance of the algorithm given in this paper is comparable to the performance of the CUSUM algorithm. This simulation shows that  $T_j^k$  can be used successfully to detect the cycle slip with known strength.

## 9. CONCLUSION AND FUTURE WORK

In this paper, we developed a new method for the detection of abrupt changes for known parameters after change. We showed that unlike the CUSUM algorithm, the statistic in this method can be calculated recursively for nonlinear stochastic systems. In the future, we intend to extend our results to the case where the parameters after change are unknown. The major obstacle in this extension is the complexity of the change detection method.

## ACKNOWLEDGMENT

This research was supported in part by Army Research Office under ODDR&E MURI97 Program Grant no. DAAG55-97-1-0114 to the Center for Dynamics and Control of Smart Structures (through Harvard University), under ODDR&E MURI01 Program Grant no. DAAD19-01-1-0465 to the Center for Communicating Networked Control Systems (through Boston University), and by the National Science Foundation Learning and Intelligent Systems Initiative Grant no. CMS9720334.

## REFERENCES

- [1] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, pp. 100–115, 1954.
- [2] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.
- [3] G. Lorden, "Procedures for reacting to a change in distribution," *Annals of Mathematical Statistics*, vol. 42, no. 6, pp. 1897–1908, 1971.
- [4] A. S. Willsky and H. L. Jones, "A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems," *IEEE Trans. Automatic Control*, vol. 21, no. 1, pp. 108–112, 1976.
- [5] I. V. Nikiforov, "New optimal approach to global positioning system/differential global positioning system integrity monitoring," *Journal of Guidance, Control and Dynamics*, vol. 19, no. 5, pp. 1023–1033, 1996.
- [6] P. S. Maybeck, *Stochastic Models, Estimation, and Control*, vol. 2, Academic Press, New York, NY, USA, 1982.
- [7] P. Del Moral, "Nonlinear filtering: interacting particle solution," *Markov Processes and Related Fields*, vol. 2, no. 4, pp. 555–580, 1996.
- [8] F. Le Gland, C. Musso, and N. Oudjane, "An analysis of regularized interacting particle methods in nonlinear filtering," in *Proc. IEEE 3rd European Workshop on Computer-Intensive Methods in Control and Signal Processing*, J. Rojicek, M. Valeckova, M. Kárný, and K. Warwick, Eds., vol. 1, pp. 167–174, Prague, Czech Republic, September 1998.
- [9] M. Isard and A. Blake, "Contour tracking by stochastic propagation of conditional density," in *Proc. European Conference on Computer Vision (ECCV '96)*, vol. 1, pp. 343–356, Cambridge, UK, April 1996.
- [10] P. Fearnhead, *Sequential Monte Carlo methods in filter theory*, Ph.D. thesis, Merton College, University of Oxford, Oxford, UK, 1998.
- [11] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proceedings Part F: Radar and Signal Processing*, vol. 140, no. 2, pp. 107–113, 1993.
- [12] G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian nonlinear state space models," *Journal of Computational and Graphical Statistics*, vol. 5, no. 1, pp. 1–25, 1996.
- [13] B. Azimi-Sadjadi and P. S. Krishnaprasad, "Approximate nonlinear filtering and its applications for an integrated INS/GPS," to appear in *Automatica*.
- [14] A. Doucet, N. de Freitas, and N. Gordon, "An introduction to sequential Monte Carlo methods," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. de Freitas, and N. Gordon, Eds., pp. 1–14, Springer, New York, NY, USA, 2001.
- [15] D. Brigo, *Filtering by projection on the manifold of exponential densities*, Ph.D. thesis, Department of Economics and Econometrics, Vrije Universiteit, Amsterdam, The Netherlands, October 1996.
- [16] B. R. Crain, "Exponential models, maximum likelihood estimation, and the Haar condition," *Journal of the American Statistical Association*, vol. 71, no. 355, pp. 737–740, 1976.
- [17] A. Budhiraja and D. Ocone, "Exponential stability in discrete-time filtering for non-ergodic signals," *Stochastic Processes and Their Applications*, vol. 82, no. 2, pp. 245–257, 1999.
- [18] R. Atar and O. Zeitouni, "Exponential stability for nonlinear filtering," *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques*, vol. 33, no. 6, pp. 697–725, 1997.
- [19] R. Atar, "Exponential stability for nonlinear filtering of diffusion processes in a noncompact domain," *The Annals of Probability*, vol. 26, no. 4, pp. 1552–1574, 1998.
- [20] F. LeGland and N. Oudjane, "Stability and approximation of nonlinear filters in the Hilbert metric, and application to particle filters," in *Proc. IEEE 39th Conference on Decision and Control*, vol. 2, pp. 1585–1590, Sydney, Australia, December 2000.
- [21] E. Abbott and D. Powell, "Land-vehicle navigation using GPS," *Proceedings of the IEEE*, vol. 87, no. 1, pp. 145–162, 1999.
- [22] G. Elkaim, M. O'Connor, T. Bell, and B. Parkinson, "System identification of a farm vehicle using carrier-phase differential GPS," in *Proc. of 9th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-96)*, pp. 485–494, Kansas City, Mo, USA, September 1996.
- [23] R. Zickel and N. Nehemia, "GPS aided dead reckoning navigation," in *Proc. 1994 National Technical Meeting*, pp. 577–586, San Diego, Calif, USA, January 1994.
- [24] H. Carvalho, P. Del Moral, A. Monin, and G. Salut, "Optimal nonlinear filtering in GPS/INS integration," *IEEE Trans. on Aerospace and Electronics Systems*, vol. 33, no. 3, pp. 835–850, 1997.
- [25] B. Azimi-Sadjadi, *Approximate nonlinear filtering with applications to navigation*, Ph.D. thesis, University of Maryland at College Park, College Park, Md, USA, August 2001.
- [26] B. Azimi-Sadjadi and P. S. Krishnaprasad, "Integer ambiguity resolution in GPS using particle filtering," in *Proc. American Control Conference*, vol. 5, pp. 3761–3766, Arlington, Va, USA, June 2001.

- [27] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins, *Global Positioning System: Theory and Practice*, Springer, New York, NY, USA, 2nd edition, 1993.
- [28] J. A. Farrell and M. Barth, *The Global Positioning System and Inertial Navigation*, McGraw-Hill, New York, NY, USA, 1998.

**Babak Azimi-Sadjadi** received his B.S. degree from Sharif University of Technology in 1989, his M.S. degree from Tehran University in 1992, and his Ph.D. degree from the University of Maryland at College Park in 2001, all in electrical engineering. He is currently with the Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, where he is a Research Assistant Professor. He also has a joint appointment with the Institute for Systems Research, University of Maryland, as a Research Associate. His research interests include nonlinear filtering, networked control systems, and wireless networks.



**P. S. Krishnaprasad** received the Ph.D. degree from Harvard University in 1977. He was in the faculty of the Systems Engineering Department at Case Western Reserve University from 1977 to 1980. He has been with the University of Maryland since August 1980, where he has held the position of Professor of electrical engineering since 1987, and has a joint appointment with the Institute for Systems Research since 1988. He is also a member of the Applied Mathematics faculty. He has held visiting positions with Erasmus University (Rotterdam); the Department of Mathematics at University of California, Berkeley; the Department of Mathematics at University of Groningen (The Netherlands); the Mathematical Sciences Institute at Cornell University; and the Mechanical and Aerospace Engineering Department at Princeton University. Krishnaprasad's interests include geometric control theory, filtering and signal processing theory, robotics, acoustics, and biologically-inspired approaches to control, sensing, and computation. He is currently actively engaged in the study of collectives, that is, communicating networked control systems. P. S. Krishnaprasad was elected a Fellow of the IEEE in 1990.

