Approximate Nonlinear Filtering and Its Applications for GPS

B. Azimi-Sadjadi and P.S. Krishnaprasad
Institute for Systems Research
University of Maryland at College Park
College Park, MD 20742

Abstract

In this paper we address the problem of nonlinear filtering in the presence of integer uncertainty. This setup is specially important for the case of differential GPS with carrier phase measurements. In simulation results we show that Particle Filtering is capable of resolving integer ambiguity in the given nonlinear setup. Motivated by these results we introduce a new Particle Filtering algorithm that can reduce the computational complexity for a certain class of problems. In this class, it is assumed that the conditional density of the state of the system given the observations is close to a known exponential family of densities.

1 Introduction

GPS provides world wide positioning with acceptable accuracy, if four or more satellites are in view.Using Differential GPS (DGPS) in conjunction with the carrier measurement, allows users to reach centimeter, or in the static case, millimeter level accuracy. This can happen only if one is able to estimate the number of full cycles of the carrier phase. This estimation problem is called integer ambiguity resolution [1][2] [3].

Although Carrier Phase DGPS (CPDGPS) allows for very accurate positioning, it is very sensitive to obstacles that can block satellite signals and cycle slips. A good estimation algorithm should be able to quickly estimate the integer ambiguity on the fly. Most of the algorithms use integer least square methods for this [3][1]. In [1] a Kalman filter setup is used to estimate the integer ambiguity.

In most of the applications, integrated INS/GPS [4], dead reckoning/GPS [5], or vehicle dynamics/GPS [6], linearization of the dynamics and the GPS observation is the main tool for estimation [3][1]. It can be shown [7] that when the number of satellites is below a certain level linearization degrades the position estimation significantly. In this case, it is important to use a nonlinear setup for the estimation problem. In [7] this case is studied by using an approximation for nonlinear filtering [8][9].

Except for very special cases in nonlinear settings, estimating the state given the observations results in an infinite dimensional filter [10]. The most widely used approximation filtering method is the Extended Kalman Filter (EKF)[10]. EKF is computationally simple but, often fails to track the conditional distribution. Projection Filtering (PrF) is another approximation method [11][12][13]. In PrF it is assumed that the conditional density of the state of the system can be approximated by a member of a parametric family of densities. In this case, estimating the conditional density is equivalent to estimating the parameter of the family. In [11][12] the exponential family of densities is chosen as the parametric family. In [13] the approach is different; there a Galerkin approximation is used for solving the Fokker-Planck equation [10]. An entirely different approach to approximate the conditional density was proposed in [14][8][9]. This method is based on the Monte Carlo method and is called Particle Filtering (PaF). In [8] it is proved that with sufficiently many particles, one can get an approximate conditional distribution that is arbitrarily close to the true conditional distribution.

In cases where some prior information about the distribution is available, reduction in the computational cost and increase in the convergence rate might be possible. Here, we assume that the conditional density lies in an exponential family of densities, or at least stays close to it in a sense that will be defined. Using this assumption, we replace the empirical distribution in [8] with the Maximum Likelihood Estimate (MLE) of the distribution.

To use nonlinear filtering methods for CPDGPS, one

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

should be able to include the integer ambiguity resolution in these methods. In this paper we present some simulation results which show that PaF, with minor modifications, is capable of resolving integer uncertainties present in a problem similar to CPDGPS. One problem of PaF is the need for large number of particles. This problem is even more important for the cases where integer uncertainty is present. The writers consider PaF for exponential families of distributions to be more suitable for nonlinear filtering with integer ambiguity.

In this paper, in Section 2 we explain particle filtering and we state the results in [8][9]. In Sections 3 and 4 we introduce a new PaF algorithm and we state the main results of this paper. In Section 5 we apply the PaF method to a nonlinear system with integer uncertainty and we present the simulation results. In Section 6 we discuss future directions for research.

2 Particle Filtering

Consider the continuous time state dynamics and discrete time observation

$$\begin{aligned} d\mathbf{x}_t &= \mathbf{f}_t(\mathbf{x}_t)dt + G_t(\mathbf{x}_t)d\mathbf{w}_t, \quad \mathbf{x}_0 \\ \mathbf{y}_{n\tau} &= \mathbf{h}_n(\mathbf{x}_{n\tau}) + \mathbf{v}_n, \end{aligned} \tag{1}$$

where $\mathbf{x}_t \in \mathcal{R}^n$, $\mathbf{y}_{n\tau} \in \mathcal{R}^d$, $\mathbf{w}_t \in \mathcal{R}^q$ is a vector from an independent Brownian motion process, and $\mathbf{v}_n \in \mathcal{R}^d$ is a white Gaussian noise process. $\mathbf{f}_t(\cdot)$, $\mathbf{h}_n(\mathbf{x}_{n\tau})$, and the matrix $G_t(\cdot)$ are assumed to have the proper sizes. The noise processes $\{\mathbf{w}_t,\ t \geq 0\}$, and $\{\mathbf{v}_n,\ n = 0, 1, \cdots\}$, and the initial condition $\mathbf{x}_0$ are assumed independent. We use Q_t and R_n for the covariance matrices of the processes $\mathbf{w}_t$ and $\mathbf{v}_n$, respectively. We assume that R_n is invertible for all n's. We have the following additional assumptions [12]:

(A1) LOCAL LIPSCHITZ CONTINUITY: $\forall\ \mathbf{x},\ \mathbf{x}' \in B_r$ and $t \in [0, T]$, where B_r is a ball of radius r, we have $\|\mathbf{f}_t(\mathbf{x}) - \mathbf{f}_t(\mathbf{x}')\| \leq k_r\|\mathbf{x} - \mathbf{x}'\|$, and $\|G_t(\mathbf{x})Q_tG_t^T(\mathbf{x}) - G_t(\mathbf{x}')Q_tG_t^T(\mathbf{x}')\| \leq k_r\|\mathbf{x} - \mathbf{x}'\|$.

(A2) NON-EXPLOSION: $\forall\ t \in [0, T]$ and $\forall\ \mathbf{x} \in \mathcal{R}^n$ there exists $k > 0$ such that, $\mathbf{x}^T\mathbf{f}_t(\mathbf{x}) \leq k(1 + \|\mathbf{x}\|^2)$, and $trace(G_t(\mathbf{x})Q_tG_t^T(\mathbf{x})) \leq k(1 + \|\mathbf{x}\|^2)$.

Under assumptions (A1) and (A2), there exists a unique solution $\{\mathbf{x}_t,\ t \in [0, T]\}$ to the state equation, and $\mathbf{x}_t$ has finite moment of any order [12].

One can also consider discrete dynamics and discrete observation

$$\begin{aligned} \mathbf{x}_{n+1} &= \mathbf{f}_n(\mathbf{x}_n) + G_n(\mathbf{x}_n)\mathbf{w}_n, \quad \mathbf{x}_0 \\ \mathbf{y}_n &= \mathbf{h}_n(\mathbf{x}_n) + \mathbf{v}_n. \end{aligned} \tag{2}$$

where $\mathbf{w}_n \in \mathcal{R}^q$ is a white Gaussian noise process, and other vectors and functions are similar to the ones in(1).

We assume that in both cases, the initial density for $\mathbf{x}_0$ is given. Hereafter, the index n for continuous time processes should be read as the time instant $n\tau$. The propagation of the conditional density, at least conceptually, can be expressed as follows [10]:

- Step 1 . Initialization:
  $$p_{0|0}(\mathbf{x}_0|\mathbf{y}_0) = p(\mathbf{x}_0).$$
- Step 2 . Diffusion:
  $$p_{n+1|n}(\mathbf{x}_{n+1}|\mathcal{Y}_n) = \int p(\mathbf{x}_{n+1}|\mathbf{x}_n)p_{n|n}(\mathbf{x}_n|\mathcal{Y}_n)d\mathbf{x}_n,$$
  where $\mathcal{Y}_n = \{\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n\}$.
- Step 3 . Bayes' rule update:
  $$p_{n+1|n+1}(\mathbf{x}_{n+1}|\mathcal{Y}_{n+1}) = \frac{p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1})p_{n+1|n}(\mathbf{x}_{n+1}|\mathcal{Y}_n)}{\int p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1})p_{n+1|n}(\mathbf{x}_{n+1}|\mathcal{Y}_n)d\mathbf{x}_{n+1}}.$$
- Step 4 . $n \leftarrow n + 1$; go to Step 2.

The density given by the above steps is exact, but in general it is an infinite dimensional filter, thus, not implementable. PaF, in brief, is an approximation method that mimics the above calculations with a finite number of operations using the Monte Carlo method. The procedure for PaF is as follows [14][8]:

Algorithm 1 *Particle Filtering*

- *Step 1 . Initialization*
  Sample $\mathbf{x}_0^1, \cdots,\ \mathbf{x}_0^N$, *$N$ i.i.d. random vectors with the distribution* $P_0(\mathbf{x})$.
- *Step 2 . Diffusion*
  Find $\hat{\mathbf{x}}_{n+1}^1, \cdots,\ \hat{\mathbf{x}}_{n+1}^N$ *from the given* $\mathbf{x}_n^1, \cdots,\ \mathbf{x}_n^N$, *using the dynamics in (1) or (2).*
- *Step 3 . Find the empirical distribution*
  $$P_{n+1|n}^N(\mathbf{x}) = \frac{1}{N}\sum_{j=1}^N \delta_{\hat{\mathbf{x}}_{n+1}^j}(\mathbf{x})$$
- *Step 4 . Use Bayes' Rule*
  $$P_{n+1|n+1}^N(\mathbf{x}) = \frac{\frac{1}{N}\sum_{j=1}^N \delta_{\hat{\mathbf{x}}_{n+1}^j}(\mathbf{x})\cdot\Psi_{n+1}(\mathbf{x})}{\frac{1}{N}\sum_{j=1}^N \delta_{\hat{\mathbf{x}}_{n+1}^j}(\hat{\mathbf{x}}_{n+1}^j)\cdot\Psi_{n+1}(\hat{\mathbf{x}}_{n+1}^j)}$$
- *Step 5 . Resample*
  Sample $\mathbf{x}_{n+1}^1, \cdots,\ \mathbf{x}_{n+1}^N$ *according to* $P_{n+1|n+1}^N(\mathbf{x})$
- *Step 6 .* $n \leftarrow n + 1$*; go to Step (2).*

where $\delta_{\mathbf{v}}(\mathbf{w}) = 1$ if $\mathbf{w} = \mathbf{v}$ and 0 otherwise, and $\Psi_{n+1}(\mathbf{x})$ is the conditional density of the observation given the state at time $n + 1$.

It is customary to call $\mathbf{x}_n^1, \cdots,\ \mathbf{x}_n^N$ particles. In the next few lines, we try to explain in words the evolution of these particles using the above algorithm.

Let $\hat{\mathbf{x}}_n^1, \cdots, \hat{\mathbf{x}}_n^N$ be the distinct particles at time n before incorporating the observations at time n. The probability of each particle is $\frac{1}{N}$. After incorporating the observations, the conditional probability of each particle changes. Some will have small, and some large probabilities. Therefore, in the process of resampling, it is very likely that some particles will never be used and instead some other particles (with high probabilities) will be sampled more than once. Therefore, after resampling, some particles have repeated versions, but in the diffusion phase they go through different paths and at the end of the diffusion phase, it is very likely, we would have N distinct particles. This automatically makes the approximation one of better resolution in the areas where the probability is higher.

In [8] for every bounded Borel test function $f(\cdot)$ it is proved under some conditions that $E\|\frac{1}{N}\sum_{i=1}^N f(\hat{\mathbf{x}}_n^i) - E_{P_n}(f(\mathbf{x}))\| \longrightarrow 0$ as N goes to 0.

One problem in using the PaF method is the computational cost. Specially for high dimensional systems, getting reasonable accuracy means using a large N, which generates heavy computational cost. In the next section, we propose a method that can reduce the number of particles for a certain class of problems.

3 PaF for Exponential Families of Densities

In the PaF method, we saw that the conditional distribution is approximated by the empirical distribution. In most cases, the actual conditional distribution is smooth, unlike the empirical distribution. It seems that if, before hand, we know some properties of the distribution, we can do better in performance than just using the empirical distribution. In the following, we assume that the conditional density lies in a parametric family of densities. We will see that with this assumption, we can show the convergence of the approximated density to the actual one. Forcing the density to lie in a parametric family induces some error in the estimate of the density, but we hope to find the proper family that results in an acceptable error.

Definition 1 *Let $\{c_1, \cdots, c_p\}$ be affinely independent* [1] *scalar functions defined on $\mathcal{R}^n$. Let $\mathbf{c}(\mathbf{x}) = (c_1(\mathbf{x}), \cdots, c_p(\mathbf{x}))^T$, and assume that the convex set $\Theta_0 = \{\theta \in \mathcal{R}^p : \Upsilon(\theta) = log \int exp(\theta^T \mathbf{c}(\mathbf{x}))d\mathbf{x} < \infty\}$, has nonempty interior. Then, $\mathcal{S} = \{p(\cdot, \theta),\ \theta \in \Theta\}$, $p(\mathbf{x}, \theta) := exp\left[\theta^T \mathbf{c}(\mathbf{x}) - \Upsilon(\theta)\right]$, where $\Theta \subseteq \Theta_0$ is open, is called an exponential family of probability densities.*

For System (1), we assume that the probability density of $\mathbf{x}_t$, given the observation, is in a family of exponential densities $\mathcal{S}$. This assumption is rather strong. We will drop this assumption later, and we will only assume that there exists a known family of densities that approximates the real density well, i.e., with acceptable accuracy. With this assumption, the proposed algorithm is as follows:

Algorithm 2 *PaF for exponential family of densities.*

- *Step 1 . Initialization*
  Sample $\mathbf{x}_0^1, \cdots, \mathbf{x}_0^N$, N i.i.d. random vectors with the density, $p_0(\mathbf{x})$.
- *Step 2 . Diffusion*
  Find $\hat{\mathbf{x}}_{n+1}^1, \cdots, \hat{\mathbf{x}}_{n+1}^N$ from the given $\mathbf{x}_n^1, \cdots, \mathbf{x}_n^N$, using the dynamics in (1).
- *Step 3 . Find the MLE of $\hat{\theta}_{n+1}$ given $\hat{\mathbf{x}}_{n+1}^1, \cdots, \hat{\mathbf{x}}_{n+1}^N$ [15]*
  $$\hat{\theta}_{n+1} = \arg\max_\theta \prod_{i=1}^N \exp(\theta^T \mathbf{c}(\hat{\mathbf{x}}_{n+1}^i) - \Upsilon(\theta))$$
- *Step 4 . Use Bayes' Rule*
  $$p_{n+1|n+1}(\mathbf{x}, \hat{\theta}_{n+1}) = \frac{\exp(\hat{\theta}_{n+1}^T \mathbf{c}(\mathbf{x}) - \Upsilon(\hat{\theta}_{n+1}))\Psi_{n+1}(\mathbf{x})}{\int \exp(\hat{\theta}_{n+1}^T \mathbf{c}(\mathbf{x}) - \Upsilon(\hat{\theta}_{n+1}))\Psi_{n+1}(\mathbf{x})d\mathbf{x}}$$
- *Step 5 . Resample*
  Sample $\mathbf{x}_{n+1}^1, \cdots, \mathbf{x}_{n+1}^N$ according to $p_{n+1|n+1}(\mathbf{x}, \hat{\theta}_{n+1})$.
- *Step 6 . $n \leftarrow n+1$; go to Step (2).*

To generate $\mathbf{x}_{n+1}^1, \cdots, \mathbf{x}_{n+1}^N$, a Gibbs sampler can be used [16]. This brings an extra computational cost, which should be taken into account when choosing Algorithm 2 over Algorithm 1.

Let $\hat{\mathbf{x}}_n^1, \cdots, \hat{\mathbf{x}}_n^N$ be the values of the particles, just before the measurement at time n. The MLE of θ_n, $\hat{\theta}_n$, satisfies the first order necessary condition

$$\frac{1}{N}\sum_{i=1}^N c_j(\hat{\mathbf{x}}_n^i) = E_{\hat{\theta}_n}(c_j(\mathbf{x})), \quad \text{for} \quad j = 1, \cdots, p \qquad (3)$$

where from now on, by $E_\theta(u(\mathbf{x}))$ we mean the expectation of $u(\mathbf{x})$ under the probability density $p(\mathbf{x}, \theta)$. Equation (3) suggests that the sample average of $c_j(\mathbf{x})$ and its probabilistic average, evaluated at $\hat{\theta}_n$, should be equal.

In each iteration, Algorithm 2 starts from the density $p_{\hat{\theta}_t}\left(\mathbf{x}_t|\mathbf{y}^t\right)$, $t = \tau n$, where $\hat{\theta}_t$ is the best estimate θ_t according to the algorithm. After a full iteration the algorithm yields $\hat{\theta}_{t+1}$ which is the best estimate of θ_{t+1}. The error in $\hat{\theta}_{t+1}$ is a combination of a series of possible errors for which we want to find an upper bound. The first source of error is the error in $\hat{\theta}_t$ which will propagate even if no other error is considered. The other source comes from the fact that in each iteration new particles are resampled based on

[1] $\{c_1, \cdots, c_p\}$ are affinely independent if for distinct points $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{p+1}$, $\sum_{i=1}^{p+1} \lambda_i \mathbf{c}(\mathbf{x}_i) = 0$ and $\sum_{i=1}^{p+1} \lambda_i = 0$ implies $\lambda_1 = \lambda_2 = \cdots = \lambda_{p+1} = 0$

the estimated density which is different from the actual density. Finally, the last source of error comes from the discretization of the stochastic dynamics of the system. We want to emphasize that here, we assume that the density of $\mathbf{v}_n$ is stationary and Gaussian, and $\exp(-\frac{1}{2}(\mathbf{y}_{n\tau}-\mathbf{h}(\mathbf{x}_{n\tau}))^T R^{-1}(\mathbf{y}_{n\tau}-\mathbf{h}(\mathbf{x}_{n\tau})))$ lies in the family of the densities. Therefore, no other error is added to the estimate because of the Bayes' correction. We recall the following fact [15]:

Fact 1 *For the family of densities $\mathcal{S}$ with probability density $p(\mathbf{x},\theta)=\exp(\theta^T\mathbf{c}(\mathbf{x})-\Upsilon(\theta))$, assume that the Fisher information matrix, $g(\theta)=(E(c_i(\mathbf{x})c_j(\mathbf{x}))-E(c_i(\mathbf{x}))E(c_j(\mathbf{x})))_{i,j}$, is positive definite. This implies that the likelihood function $l(\theta)=\theta^T\mathbf{C}(\mathbf{x})-\Upsilon(\theta)$, is strictly concave. Also if $c_1(\mathbf{x}),\cdots,c_p(\mathbf{x})$, the components of $\mathbf{c}(\mathbf{x})$, are affinely independent almost everywhere, then for the system of equations*

$$c_j(\mathbf{x})=E_\theta[c_j(\mathbf{x})], \qquad j=1,\cdots,p,$$

if a solution exists[2]*, it is unique. In addition, if $\mathbf{x}_1,\cdots,\ \mathbf{x}_N$ are N i.i.d. random variables distributed according to $f_\theta(\mathbf{x})=\exp(\theta^T\mathbf{c}(\mathbf{x})-\Upsilon(\theta))$, then the MLE of θ, $\hat{\theta}_N$, has the following property:*

$$\sqrt{N}(\hat{\theta}_N-\theta)\sim\mathcal{N}(0,g^{-1}(\theta)).$$

Using this fact, it is easy to see that $E(\|\hat{\theta}_N-\theta\|^2)\simeq\frac{1}{N}\ trace(g^{-1}(\theta))$, therefore, when $N\longrightarrow\infty$, $\hat{\theta}_N\longrightarrow\theta$ in the m.s. sense. On the other hand, $\hat{\theta}_N$ is the solution to (3). Using the strong law of large numbers, when $N\to\infty$ the LHS in (3) goes to $E_\theta(c_j(\mathbf{x}))$, $j=1,\cdots,p$, with probability one. In other words, the solution to (3) when the LHS is the exact $E_\theta(c_j(\mathbf{x}))$, $j=1,\cdots,p$, gives the exact solution for θ. Using this argument, one can expect that by finding a good estimate of the left hand side of (3), a good estimate of θ is accessible. In each iteration of the algorithm presented in this section the estimate of the LHS of (3) is found by using the Monte Carlo method and the approximate solution for the stochastic differential equation (1)proposed in [18].

Definition 2 *We say that a function $u(\cdot)$ belongs to the class $\mathcal{F}$, written as $u\in\mathcal{F}$, if we can find constants, $k>0$, and κ, such that for all $\mathbf{x}\in\mathcal{R}^n$, the following inequality holds:*

$$\|u(\mathbf{x})\|\le k\left(1+\|\mathbf{x}\|^\kappa\right).$$

The next lemma relates the approximate solution to the stochastic differential equation and the estimate of the parameter θ. This lemma is the main building block for our result in this section.

[2] In [17] it is shown that if $N>p$, the solution exists almost surely.

Lemma 1 *For the stochastic differential equation*

$$d\mathbf{x}_t=\mathbf{f}_t(\mathbf{x}_t)dt+G_t(\mathbf{x}_t)d\mathbf{w}_t,\ \mathbf{x}_0,\ t\in[0,t_f],$$

assume that $\mathbf{f}_t(\cdot)$, $G_t(\cdot)$ are such that for the Brownian motion, $\mathbf{w}_t$, the probability density of the state $\mathbf{x}_t$ lies in the family $\mathcal{S}$ for Θ bounded, with $g(\theta)$ positive definite and bounded away from zero. We also assume the conditions in `Fact 1` *and in* `Theorem 2` *of* `[19]` *with $\mathbf{c}(\mathbf{x})$ replacing $u(\mathbf{x})$. Then, there exist k_1 and k_2 such that*

$$E[\|\theta_t-\hat{\theta}_t\|]\le k_1h^2+\frac{k_2}{N^{1/2}},\quad t\in[0,t_f]$$

where $\hat{\theta}_t$ is the estimate of θ_t, and N and h are the number of particles and the time step, respectively.
Proof: `[19]`.

Now we are ready to present the main result of this section.

Theorem 1 *For System (1), assume that $\mathbf{f}_t(\cdot)$, $G_t(\cdot)$, and $\mathbf{h}(\cdot)$ are such that for Brownian motion $\mathbf{w}_t$, and the Gaussian noise $\mathbf{v}_n$, the conditional probability density of the state $\mathbf{x}_t$, conditioned on the observations, lies in the family $\mathcal{S}$ for Θ bounded and for $t\in[0,T]$. We also assume the conditions in* `Fact 1` *and in* `Theorem 2` *of* `[19]` *with $\mathbf{c}(\mathbf{x})$ replacing $u(\mathbf{x})$. Then, if $g^{-1}(\theta)\,E_{\theta_t}(\mathcal{L}_t\mathbf{c}(\mathbf{x}))$ is Lipschitz with the Lipschitz constant L and $g(\theta)$ is positive definite and bounded away from zero, there exist l_1 and l_2 such that*

$$E\|\theta_{n\tau}-\hat{\theta}_{n\tau}\|\le\sum_{i=0}^{n-1}\exp(Li\tau)(l_1h^2+\frac{l_2}{N^{1/2}}),\ n\tau\in[0,T],$$

where $\hat{\theta}_{n\tau}$ is the estimate of $\theta_{n\tau}$, and N and h are the number of particles and the time step, respectively.
Proof: `[19]`.

4 Projection Particle Filtering for Exponential Families of Densities

In this section, we drop the assumption that the conditional density of the state given the observation lies in an exponential family of densities. Instead we have the following assumptions:

(A3) The conditional density stays close to the given exponential family $\mathcal{S}$ in a weak sense, i.e. $\forall t\in[0,T]$, $\forall u\in\mathcal{F}\ \ \exists\theta_t^*\in\Theta^*$ s.t. $\|E_{p_t}(u(\mathbf{x}))-E_{\theta_t^*}(u(\mathbf{x}))\|\le\epsilon$, where Θ^* is closed and bounded.

(A4) $\forall\theta_1,\theta_2\in\Theta^*$ and $\forall u\in\mathcal{F}\ \exists K_1,K_2$ such that $\|E_{\theta_1}u(\mathbf{x})-E_{\theta_2}u(\mathbf{x})\|\le K_1\|\theta_1-\theta_2\|$ and $\|\theta_1-\theta_2\|\le K_2\|E_{\theta_1}u(\mathbf{x})-E_{\theta_2}u(\mathbf{x})\|$.

Assume $\hat{\theta}_n$ is calculated according to Algorithm 2 and assume $p_{n|n}(\mathbf{x},\hat{\theta}_n)$ is such that $\forall u\in\mathcal{F}$

$$E\|E_{\hat{\theta}_{n|n}}u(\mathbf{x})-E_{\theta^*_{n|n}}u(\mathbf{x})\|\quad\le\quad\delta. \qquad (4)$$

where $\theta^*_{n|n}$ satisfies

$$\|E_{p_{n|n}}u(\mathbf{x}) - E_{\theta^*_{n|n}}u(\mathbf{x})\| \quad \leq \quad \epsilon. \tag{5}$$

We define the function $\mathbf{r}(\mathbf{x})$ as follows:

$$\mathbf{r}(\mathbf{x}) = E\mathbf{c}(\hat{\mathbf{x}}_{n,\mathbf{x}}((n+1)\tau))$$

where $\hat{\mathbf{x}}_{n,\mathbf{x}}((n+1)\tau)$ is the approximate solution of the stochastic differential equation at time $(n+1)\tau$ with the given initial condition $\mathbf{x}$ at time $n\tau$. Since according to our assumption $\mathbf{c} \in \mathcal{F}$, then by using lemma 9.1 in [18], we have

$$\|\mathbf{r}(\mathbf{x})\| \leq K_3(1 + \|\mathbf{x}\|^{\mu})$$

where K_3 and μ only depend on the function $\mathbf{c}(\cdot)$ and the dimension of $\mathbf{x}$.

Theorem 2 *For the system (1) assume (A1), (A2), and (A3). We also assume (A4) and the conditions in* `Fact 1` *and in* `Theorem 2` *of* `[19]` *with* $\mathbf{c}(\mathbf{x})$ *replacing* $u(\mathbf{x})$*, and we assume* $\mathbf{r} \in \mathcal{F}$*. Then in* `Algorithm 2` *if*

$$E\|E_{\hat{\theta}_{n|n}}u(\mathbf{x}) - E_{\theta^*_{n|n}}u(\mathbf{x})\| \leq \delta$$

we get

$$E\|E_{\theta^*_{n+1|n+1}}u(\mathbf{x}((n+1)\tau)) - E_{\hat{\theta}_{n+1|n+1}}u(\mathbf{x}((n+1)\tau))\| \\ \leq K_6K_1K_2(\delta + 2\epsilon + K_4h^2 + \tfrac{K_5}{N^{1/2}}),$$

for some $K_1, \cdots, K_6$*, and* N *and* h *are the number of particles and the time step, respectively.*
Proof:`[19]`.

Using the estimate in Theorem 2 repeatedly for the time interval $[0, T]$, with $T = M\tau$, and $\|E_{\hat{\theta}_{0|0}}u(\mathbf{x}) - E_{\theta^*_{0|0}}u(\mathbf{x})\| \leq \delta_0$, $\exists\alpha_1, \alpha_2, \alpha_3$, and α_4 we have

$$E\|E_{\theta^*_{n|n}}u(\mathbf{x}((n)\tau)) - E_{\hat{\theta}_{n1|n}}u(\mathbf{x}((n)\tau))\| \leq \\ \alpha_1^n\delta_0 + \sum_{i=0}^{n-1}\alpha_1^i(\alpha_2\epsilon + \alpha_3h^2 + \alpha_4N^{-1/2}),$$

for $0 \leq n \leq M$.

5 Particle Filtering for Nonlinear Systems with Constant Integer Uncertainty

Consider the following nonlinear dynamics and observation

$$\begin{aligned} \mathbf{x}_{n+1} &= \mathbf{f}_n(\mathbf{x}_n) + G_n(\mathbf{x}_n)\mathbf{w}_n \\ \mathbf{y}_n &= \mathbf{h}_n(\mathbf{x}_n) + J_n\mathbf{z} + \mathbf{v}_n \end{aligned}$$

where the assumptions and the dimensions for $\mathbf{x}_n$, $\mathbf{y}_{n\tau}$, $\mathbf{w}_n$, and $\mathbf{v}_n$ are the same as in the previous sections. We assume that $\mathbf{z}$ is a random integer vector, i.e. $\mathbf{z} \in \mathcal{Z}^m$ and J_n has the proper dimension. Vector $\mathbf{z}$ is assumed to be constant in time. This problem can be set up in continuous time as well. In both setups we assume that the integer uncertainty affects only some components of the observation, and other components are unaffected by $\mathbf{z}$. The affected components have associated noise components in $\mathbf{v}_n$ that have considerably lower energy. This suggests that an accurate estimation of $\mathbf{z}$ can increase the accuracy of the estimate of the state of the system significantly. The INS/CPDGPS, dead reckoning/CPDGPS, or vehicle dynamics/CPDGPS problem can be formulated using this treatment.

To incorporate the integer uncertainty into the dynamics, we can augment the state $\mathbf{x}_t$ with the integer ambiguity $\mathbf{z}$. We assume that the initial distribution of the augmented state, $(\mathbf{x}_0^T, \mathbf{z}_0^T)^T$, is known. With this assumption, the state dynamics and the observation have the same form that was studied in Section 2. Therefore, we can apply particle filtering to find the conditional probability distribution of the augmented state. Since $\mathbf{z}_t$ does not change, due to resampling, the part of the particles associated to $\mathbf{z}_t$ tends to cover smaller and smaller portions of the state space. This problem should be addressed properly otherwise the particle filtering fails to find a proper estimate of the conditional distribution [19].

In a two dimensional space, three transmitters are mounted at three known points. The moving object can measure its distance from these transmitters. For each pseudo satellite, two types of measurement are possible, one with high measurement noise and the other with low measurement noise. For the low measurement noise, though, there is an integer ambiguity. The dynamics of the moving object for this example is considered to be in discrete time and linear time invariant (this problem can be thought of as vehicle dynamics/CPDGPS in two dimensions). In the simulation, it is assumed that the initial condition for the position is distributed in a square of size 200×200 units, symmetric with respect to the origin. The results of the simulation are shown in Figure 1 and 2. To display the estimated integers, we simply used the mean value. In this simulation we forced one of the integers to have a jump. Although our algorithm is not designed for these kinds of changes, we see that it can estimate the new integer values. In future, we use special treatment for the times when these kinds of jumps happen. As can be seen, the estimates for the integers are reasonably good. The reliability of the estimates for the integers depends on the energy of the noise.

6 Conclusions and Future Work

The simulation results show that our method is capable of estimating the integer ambiguity and the position. There are certain issues that need further investigation. In the following, we itemize these issues:

- What are the proper criteria to stop the integer ambiguity estimation? What happens when a cycle slip occurs, i.e. one or more of the integers have a jump?

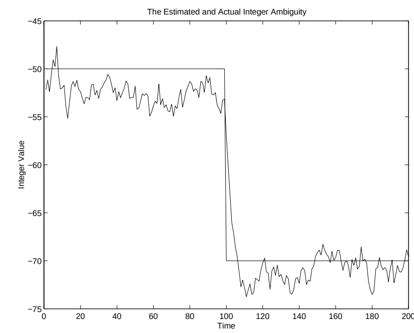

Fig 1: Estimated integer ambiguity versus the actual integer ambiguity of pseudo satellite (1).

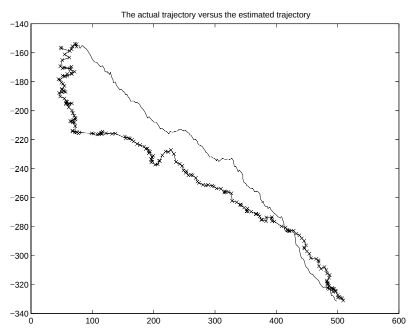

Fig 2: Estimated trajectory versus the actual trajectory of the moving object.

- How much improvement does the method of Section 3 for integer ambiguity and position estimation have over PaF?

These questions are to be answered in future work. In addition to these, we shall be more specific in our simulations, and use real GPS data for our results.

References

[1] A. Hassibi and S. Boyd. Integer Parametric Estimation in Linear Models with Applications to GPS. *IEEE Trans. on Signal Processing*, 46(11):2938–52, Nov. 1998.

[2] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins. *Global Positioning System: Theory and Practice*. Springer-Verlag, Second edition, 1993.

[3] P.J.G. Teunissen. The Least Square Ambiguity Decorrelation Adjustment: a Method for Fast GPS Integer Ambiguity Estimation. *Journal of Geodesy*, 70(1-2):65–83, Nov. 1995.

[4] N. El-Sheimy and K.P. Schwarz. Navigating Urban Areas by VISAT-A Mobile Mapping System Integrating GPS/INS/Digital Camera for GIS Applications. *Navigation*, 45(4):275–285, Winter 1999.

[5] R. Zickel and N. Nehemia. GPS Aided Dead Reckoning Navigation. *Proc. ION Nat. Tech. Meeting*, pages 577–586, 1994.

[6] G. Elkaim, M. O'Connor, T. Bell, and B. Parkinson. System Identification of a Farm Vehicle Using Carrier-Phase Differential GPS. *ION Conference*, Sept. 1996.

[7] H. Carvalho, P. Del Moral, A. Monin, and G. Salut. Optimal Nonlinear Filtering in GPS/INS Integration. *IEEE Trans. on Aerospace and Electronic Systems*, 33(3):835–850, July 1997.

[8] P. Del Moral. Non Linear Filtering: Interacting Particle Solution . *Markov Processes and Related Fields*, 2(4):555–580, 1996.

[9] F. LeGland, C. Musso, and N. Oudjane. An Analysis of Regularized Interacting Particle Methods in Nonlinear Filtering,. In J. Rojicek, M. Valeckova, M. Karry, and K. Warwick, editors, *Proceedings of the 3rd IEEE European Workshop on Computer-Intensive Methods in Control and Signal Processing*, volume 1, pages 167–174. Prague, September 7–9 1998.

[10] P.M. Maybeck. *Stochastic Models, Estimation, and Control*, volume 2. Academic Press, 1982.

[11] D. Brigo, B. Hanzon, and F. LeGland. A Differential Geometric Approach to Nonlinear Filtering : the Projection Filter. *Proceedings of the 34th IEEE Conference on Decision and Control*, pages 4006–4011, December 1995.

[12] D. Brigo. *Filtering by Projection on the Manifold of Exponential Densities*. Ph.D. Thesis Department of Economics and Econometrics, Vrije Universiteit, Amsterdam, October 29 1996.

[13] R. Beard, J. Gunther, J. Lawton, and W. Stirling. Nonlinear Projection Filter Based on Galerkin Approximation. *Journal of Guidance Control, and Dynamics*, 22(2):258–266, March–April 1999.

[14] N.J. Gordon, D.J. Salmond, and A.F.M. Smith. Novel Approach to Nonlinear/NonGaussian Bayesian State Estimation. *IEE Proceedings-F (Radar and Signal Processing)*, 140(2):107–113, 1993.

[15] E.L. Lehmann and G. Casella. *Theory of Point Estimation*. Springer-Verlag, New York, second edition, 1998.

[16] A.E. Gelfand and A.F.M. Smith. Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, 85(410):398–409, 1990.

[17] B.R. Crain. Exponential Models, Maximum Likelihood Estimation, and the Haar Condition. *Journal of the American Statistical Association*, 71(355):737–740, 1976.

[18] G.N. Milstein. *Numerical Integration of Stochastic Differential Equations*. Kluwer Academic Publishers, New York, 1995.

[19] B. Azimi-Sadjadi and P.S. Krishnaprasad. Approximate Nonlinear Filtering and Its Application for GPS. *Intitute for Systems Research Technical Report TR 2000-37*, September 2000.