PARTICLE FILTERING WITH OBSERVATIONS IN A MANIFOLD

Salem Said

CNRS, Université de Bordeaux
Laboratoire IMS, (UMR 5218)
salem.said@u-bordeaux.fr

Jonathan H. Manton, IEEE Senior Member

The University of Melbourne
Department of Electrical and Electronic Engineering
jmanton@unimelb.edu.au

ABSTRACT

This paper describes the application of particle filtering to the solution of the problem of filtering with observations in a manifold. Mathematically, this is based on an original use of so-called connector maps. It is shown that well-chosen connector maps can be used to transform successive samples from a continuous time observation process, evolving on a manifold, into a discrete sequence of random vectors, which are asymptotically independent and normally distributed, in the limit where the sampling interval goes to zero. Roughly speaking, this "innovation sequence" can be used as the input of a sequential Monte Carlo algorithm. As a concrete application, numerical simulation results are presented, for the problem of estimating the angular velocity of a rigid body from noisy observations of its attitude.

Index Terms— Stochastic filtering, Particle filtering, Differentiable manifold, Lie group, Angular velocity

1. PROBLEM STATEMENT

The problem considered in this paper belongs to the general class of continuous time stochastic filtering problems. Such problems are stated in terms of two stochastic processes, the signal process $\{X_t; t \geq 0\}$, and the observation process $\{Y_t; t \geq 0\}$. Given some dynamical models for the processes X and Y, the aim is to compute the posterior distribution π_t, distribution of X_t given observations $\mathcal{Y}_t = \{Y_s; s \leq t\}$.

The case which has received most attention is the "additive white noise case", where X is a Markov process with state space S, and Y is a diffusion process with values in a Euclidean space $\mathbb{R}^d$ [1, 2]. Precisely, Y is given by a dynamical model, (stochastic differential equation),

$$dY_t = H(X_t)dt + dB_t \tag{1}$$

where $H : S \to \mathbb{R}^d$ is called the sensor function, and B is a standard Brownian motion with values in $\mathbb{R}^d$. This setting will be said to define a *classical filtering problem*.

Starting in the 1980s, there has been interest in the problem of filtering with observations in a manifold, which is a generalisation of the classical filtering problem [3, 4, 5, 6]. With X just as before, this new problem assumes Y is a diffusion process with values in a smooth manifold M. A dynamical model for Y is defined by a mapping $H : S \times M \to TM$, such that $H(s, y) \in T_yM$, (this means H is a vector field on M, parameterised by $s \in S$), some vector fields $V_1, \ldots, V_q$ on M, and a standard Brownian motion B in $\mathbb{R}^q$. Then Y is given by the stochastic differential equation [5, 6]

$$dY_t = H(X_t, Y_t)dt + \sum_{r=1}^{q} V_r(Y_t) \circ dB_t^r \tag{2}$$

where $\circ dB_t^r$ denotes the Stratonovich differential.

In the following, Section 2 recalls the closed form solution of the problem of filtering with observations in a manifold, in the form of a generalised Kallianpur-Striebel, (KS), formula. Section 3 is concerned with the use of connector maps. Section 4, describes the application of particle filtering to the evaluation of the KS formula of Section 2, and Section 5 applies the results of Section 4 to the problem of estimating the angular velocity of a rigid body from noisy observations of its attitude. Before going on, the reader should consult Section 6, for background and references to previous work.

2. CLOSED FORM SOLUTION : THE KS FORMULA

The closed form solution of the classical filtering problem is given by the KS formula, which may be thought of as an abstract Bayes formula for the posterior distribution π_t [1]. In existing literature [3, 4, 5], various generalisations of the KS formula, to the problem of filtering with observations in a manifold, have been proposed. Recently [6], the authors presented a new generalisation, which is now described. This is formula (6) below.

Assume first the process Y, as given by (2), is elliptic. This means that for $y \in M$, the vectors $V_1(y), \ldots, V_q(y)$ span the tangent space T_yM. Then, there exist on M a Riemannian metric $\langle \cdot, \cdot \rangle$ and a connection ∇, uniquely determined by the following properties [6],

$$\langle K, E \rangle = \sum_{r=1}^{q} \langle K, V_r \rangle \langle V_r, E \rangle \tag{3}$$

$$\sum_{r=1}^{q} \nabla_{V_r} V_r = 0 \tag{4}$$

While the connection ∇ is compatible with the metric $\langle \cdot, \cdot \rangle$, it is different from the corresponding Levi-Civita connection, in general.

The generalised KS formula (6) involves the Itô integral of a vector field along the observation process Y. If K is a vector field along Y, then its Itô integral along Y is defined by

$$\int_0^t \langle K, dY_s \rangle = \int_0^t \langle K, H_s \rangle dt + \sum_{r=1}^{q} \int_0^t \langle K, V_r(Y_s) \rangle dB_s^r \tag{5}$$

where $H_t = H(X_t, Y_t)$. While (5) is found from (2) by formally replacing the Stratonovich differentials $\circ dB_t^r$ by the Itô differentials dB_t^r, it takes some work to show this formula indeed has all the required properties of a stochastic integral, (for instance, that it only depends on the vector field K and the observation process Y). For this, the reader is referred to [7, 8].

It is now possible to state the closed form solution of the problem of filtering with observations in a manifold, in the form of a generalised KS formula [6]. Assume the state space S of the signal process X is a complete separable metric space.

Proposition 1 *The posterior distribution π_t is given by $\int \varphi(s)\pi_t(ds) \propto \int \varphi(s)\rho_t(ds)$, where the unnormalised posterior ρ_t is given by*

$$\int \varphi(s)\rho_t(ds) = \mathbb{E}\left[\varphi(\tilde{X}_t)L_t(\tilde{X})\Big|\,\mathcal{Y}_\infty\right] \tag{6}$$

for any bounded Lipschitz continuous function φ on S. Here, "$\propto$" means "proportional to", where the constant of proportionality ensures that $\int \pi_t(ds) = 1$. Moreover, $\tilde{X}$ is a process having the same distribution as X, but independent of the observations Y. The likelihood process $L(\tilde{X})$ is given by

$$L_t(\tilde{X}) = \exp\left(\int_0^t \langle \tilde{H}, dY_s\rangle - \frac{1}{2}\int_0^t \|H_s\|^2\, ds\right) \tag{7}$$

where $\tilde{H}_t = H(\tilde{X}_t, Y_t)$, and $\|H_t\|^2 = \langle H_t, H_t\rangle$.

For any practical implementation of the KS formula (6), it is necessary to discretise it in a suitable way. This will be done using connector maps, which are considered in the following section.

3. NUMERICAL SOLUTION : CONNECTOR MAPS

This section considers the use of connector maps, in the solution of the problem of filtering with observations in a manifold. Recall that a connector map is a smooth function $I : M \times M \to TM$, such that $I(p,q) \in T_pM$, for all $p,q \in M$. The vector $I(p,q)$ is intended to be a tangent vector at p, which "connects p and q" [9].

To consider connector maps in the context of the problem of filtering with observations in a manifold, assume samples $\{Y_{k\delta}; k \in \mathbb{N}\}$, of the observation process Y, are given, and let $\Delta Y_k = I(Y_{k\delta}, Y_{(k+1)\delta})$. Below, it is stated that connector maps, when chosen adequately, can be used in two ways.

First, they allow for Itô integrals, as defined in (5), to be approximated by Riemann sums, with limit in the square mean,

$$\int_0^t \langle K, dY_s\rangle = \lim_{\delta\to 0} \sum_{(k+1)\delta < t} \langle K_{k\delta}, \Delta Y_k\rangle \tag{8}$$

Second, the vectors $\{\Delta Y_k; k \in \mathbb{N}\}$ are asymptotically independent and normally distributed, in the limit $\delta \to 0$. This justifies thinking of the vectors ΔY_k as an "innovation sequence" for the samples Y_k, once δ has been taken small enough.

To state this property precisely, let $E^1, \ldots, E^d$ be a parallel orthonormal frame along Y. That is, $E^1, \ldots, E^d$ are vector fields along Y, with $\langle E^i_t, E^j_t\rangle = \delta_{ij}$ for all $t \geq 0$, ($\delta_{ij} = 1$ if $i = j$ and $= 0$ if $i \neq j$), and such that each E^i verifies the equation of stochastic parallel transport along Y [8]. Each ΔY_k is determined by its components $\Delta Y^1_k, \ldots, \Delta Y^d_k$ in the orthonormal basis $E^1_{k\delta}, \ldots, E^d_{k\delta}$. The joint distribution of these components verifies

$$\mathbb{P}\left(\delta^{-\frac{1}{2}}\left(\Delta Y^i_k - \delta \times H^i_{k\delta}\right)\Big|\, \Delta Y_0, \ldots \Delta Y_{k-1}, X_{k\delta}\right) \to \mathcal{N}_d \tag{9}$$

where the limit is taken as $\delta \to 0$. Here, the left hand side denotes conditional distribution, $H^i_t = \langle E^i, H_t\rangle$, and $\mathcal{N}_d$ denotes the standard normal distribution on $\mathbb{R}^d$. Asymptotic independence follows since $\mathcal{N}_d$ does not depend on $\Delta Y_0, \ldots \Delta Y_{k-1}$.

It is clear that both properties (8) and (9) are desirable. However, they will not hold for an arbitrary choice of the connector map I. Rather, it is necessary to choose I in a way which is attuned to the dynamics of the observation process Y. Proposition 2, below, states that (8) and (9) hold if I verifies the following conditions, (where I_p is the mapping $I_p(q) = I(p,q)$, for $p,q \in M$),

$$dI_p(p)(V) = V \qquad \nabla^2 I_p(p)(V,V) = 0 \tag{10}$$

for all $V \in T_pM$, where $dI_p, \nabla^2 I_p$ denote the derivative and the Hessian of I_p, the latter being with respect to the connection ∇ defined by (4) — See definition in [10].

Proposition 2 *Assume the manifold M is compact. If I is a connector map which verifies (10), then (8) and (9) are verified.*

To appreciate the usefulness of this proposition, consider its conditions. Compactness of M is imposed to ensure the limit (8) holds in the square mean. It could be dropped if a weaker notion of convergence is used, (convergence in probability). Condition (10) defines a choice of connector maps which guarantee (8) and (9).

The proposition tells us that, in principle, *any* connector map can be used, as long as it verifies (10). Thus, in practice, if a mapping I is being used, which is difficult to compute, (*e.g.* highly nonlinear), it is possible to replace it by another mapping I', which is much easier to deal with. This will incur no noticeable loss in performance if the sampling interval δ is sufficiently small.

The class of connector maps I which verify (10) is never empty. For any manifold M with the metric and connection (3) and (4), the mappings I known as *geodesic connectors* will verify (10). Roughly [9], a connector map I is a geodesic connector if for any $p,q \in M$ which are close enough to each other, $I(p,q)$ is the initial velocity of a geodesic curve $\gamma : [0,1] \to M$ such that $\gamma(0) = p, \gamma(1) = q$. In this case, one writes $I_p(q) = \log_p(q)$, where $\log_p$ is the Riemannian logarithm of the metric (3).

An elementary example of a geodesic connector is when $M = \mathbb{R}^d$, with (3) and (4) defined based on the classical filtering problem (1). Then, a geodesic connector is naturally given by $I(p,q) = q - p$. For most cases beyond this elementary example, geodesic connectors are highly nonlinear mappings and one should prefer to avoid them in numerical implementations. Here is an example of how Proposition 2 allows this to be done, (a further, related, example can be found in Section 5).

Let M be the group of rotation matrices, $SO(3)$. Let the vector fields V_1, V_2, V_3 be left invariant vector fields, corresponding to rotations around the three axes of a positive orthonormal frame [11]. The metric (3) and connection (4), are then defined by $\langle V_i, V_j\rangle = \delta_{ij}$ and $\nabla_{V_i} V_j = 0$, for $i,j = 1,2,3$. A geodesic connector can be defined by

$$I(p,q) = \log_p(q) = p\log(p^\dagger q) \tag{11}$$

with † denoting transposition, and $\log$ the antisymmetric matrix logarithm. Proposition 2 allows for the highly nonlinear formula (11) to be replaced by its first order approximation,

$$I'(p,q) = \frac{1}{2}\left(q - p\,q^\dagger p\right) \tag{12}$$

This construction may be used in filtering problems arising in navigation, localisation and computer vision [12, 13].

For lack of space, a detailed proof of Proposition 2 cannot be given, but will be included in an upcoming journal submission. The main idea of the proof is to consider the process y with values in $\mathbb{R}^d$ and whose components are $y^i_t = \int_0^t \langle E^i, dY_s\rangle$, where $E^1, \ldots, E^d$ is a parallel orthonormal frame along Y. Let $\Delta y^i_k = y^i_{(k+1)\delta} - y^i_{k\delta}$. The proposition follows by showing ΔY^i_k and Δy^i_k are asymptotically equal, as $\delta \to 0$. For example, (9) follows from the fact that y satisfies a stochastic differential equation of the form (1), so the Δy_k are independent normally distributed, (this was shown in [6]).

4. NUMERICAL SOLUTION : PARTICLE FILTERING

This section describes a sequential Monte Carlo algorithm for sampling from the posterior distribution π_t. Starting from the closed form representation of π_t, given by the KS formula (6), the design of the proposed algorithm will involve two steps, i) obtaining a suitable discretisation of formula (6), and ii) using sequential importance-resampling, (SIS), to evaluate this discretisation. These are now presented in detail.

i) Discretisation of the KS formula : Consider the issue of discretising the KS formula (6). Thinking of this formula as an abstract Bayes formula, it is natural to attempt a discretisation which replaces it with a "concrete" Bayes formula, *i.e.* one of the usual form

$$\text{posterior} \propto \text{prior} \times \text{likelihood} \tag{13}$$

Such a discretisation can also be viewed as replacing the original continuous time filtering problem, with a discrete time filtering problem. The key to achieving this is property (9) of Section 3.

Indeed, recall that a connector map I can be used to construct vectors $\Delta Y_k = I(Y_{k\delta}, Y_{(k+1)\delta})$ from successive samples $Y_{k\delta}$ of the observation process Y. Property (9) states that, when the sampling interval δ becomes small, these vectors are asymptotically independent and normally distributed. Precisely, the asymptotic conditional distribution of the components ΔY_k^i does not depend on $\Delta Y_0, \ldots, \Delta Y_{k-1}$, but only on the signal X. This distribution is normal, with covariance $\delta \times I_d$, (here, I_d is the $d \times d$ identity matrix), and with mean $\delta \times H^i_{k\delta}$.

Taking this to hold exactly, (mathematically, this means replacing exact expressions with asymptotic ones), it is possible to write

$$\mathbb{P}(\Delta Y_k | X_{k\delta}) \propto \exp\left(\langle H_{k\delta}, \Delta Y_k\rangle - \frac{\delta}{2} \times \|H_{k\delta}\|^2\right) \tag{14}$$

for the conditional distribution of ΔY_k given $X_{k\delta}$. Indeed, this follows immediately by writing down the expression of the normal distribution of the ΔY_k^i, with covariance and mean described above.

Let π_k^δ denote the distribution of $X_{k\delta}$ given $\Delta Y_0, \ldots, \Delta Y_k$. If the transition kernel of the Markov sequence $\{X_{k\delta}; k \in \mathbb{N}\}$ is known, then it is straightforward to write π_k^δ in the form (13). Indeed, the sequences $\{X_{k\delta}\}$ and $\{\Delta Y_k\}$ satisfy the usual assumptions of a discrete time filtering problem.

Moreover, see [14], expression (13) for π_k^δ can be reformulated as follows

$$\int \varphi(s)\pi_k^\delta(ds) \propto \mathbb{E}\left[\varphi(\tilde{X}_{k\delta})L_k^\delta \,\middle|\, \Delta Y_k; k \in \mathbb{N}\right] \tag{15}$$

Here, the constant of proportionality ensures $\int \pi_k^\delta(ds) = 1$, and the likelihood $L_k^\delta = L_k^\delta(\tilde{X}_0, \ldots, \tilde{X}_{k\delta})$ is equal to the product $l(\tilde{X}_0, \Delta Y_0) \times \ldots \times l(\tilde{X}_{k\delta}, \Delta Y_k)$, where

$$l(s, \Delta Y_k) = \exp\left(\langle H_k(s), \Delta Y_k\rangle - \frac{\delta}{2} \times \|H_k(s)\|^2\right) \tag{16}$$

with, $H_k(s) = H(s, Y_{k\delta})$ for any s in the state space S, (this is the same as the right hand side of (14)).

ii) SIS implementation : Consider the task of sampling from the distribution π_k^δ. In view of the representation (15) of this distribution, this task can be realised using the SIS strategy [15, 16]. This implies simulating N particles $\{\tilde{x}^i; i = 1, \ldots, N\}$, which approximately follow the trajectories of the sequence $\{X_{k\delta}\}$, and associating importance weights $\{w^i; i = 1, \ldots, N\}$ to these particles, which are updated according to the "marginal likelihoods" (16).

As X is a continuous time process, it is often not possible to simulate its trajectories exactly. Assume however that a transition kernel $q(s, ds')$ is given on S, such that if $\{x_k\}$ is a Markov sequence with the same initial distribution μ of X_0 and with transition kernel q, it holds that,

$$\mathbb{E}\left[d_S^2(X_t, x_k)\right] = O(\delta) \qquad k\delta \le t < (k+1)\delta \tag{17}$$

where $d_S(\cdot, \cdot)$ is the metric in the state space S. One says that $\{x_k\}$ converges to X with strong order of convergence 0.5 [17].

The following algorithm can now be stated. At its kth iteration, it processes the vector ΔY_k to produce particles $\hat{x}_k^1, \ldots, \hat{x}_k^N$ which approximately sample from π_k^δ.

When ΔY_j becomes available

(1) if $j = 0$: generate N particles $\tilde{x}_j^i \sim \mu$
set $w_{-1}^i = 1/N$
if $j > 0$: generate N particles $\tilde{x}_j^i \sim q(\hat{x}_{j-1}^i, ds)$

(2) compute normalised weights, $w_j^i \propto w_{j-1}^i l(\tilde{x}_j^i, \Delta Y_k)$

(3) generate $(n_j^1, \ldots, n_j^N) \sim \text{multinomial}(w_j^1, \ldots, w_j^N)$ and replace $\tilde{x}_j^i$ by n_j^i particles with same value

(4) relabel the new particles $\hat{x}_j^1, \ldots, \hat{x}_j^N$; set $w_j^i = 1/N$

The above algorithm is one of several possible variants of an SIS strategy for sampling from π_k^δ. However [15, 16], other variants differ only by modifications to instruction (3), (resampling), and are essentially the same for the present purpose.

Proposition 3, below, gives the convergence of the algorithm. To state this proposition, let $\hat{\pi}_k^\delta$ be the empirical distribution of the particles $\hat{x}_k^1, \ldots, \hat{x}_k^N$. That is,

$$\int \varphi(s)\hat{\pi}_k^\delta(ds) = \frac{1}{N}\sum_{i=1}^{N} \varphi(\hat{x}_k^i) \tag{18}$$

For $t \ge 0$, let $k(t)$ be the largest k such that $k\delta \le t$ and put $\hat{\pi}_{k(t)}^\delta \equiv \hat{\pi}_t$.

The following condition will be required for the statement. Assume there exist constants A, B such that for all $s, s' \in S$ and $y \in M$,

$$\|H(s, y)\| \le A \;\text{ and }\; \|H(s, y) - H(s', y)\| \le B \tag{19}$$

Proposition 3 *Assume the manifold M is compact. Assume also conditions (10), (17) and (19) are verified. For any bounded Lipschitz continuous function φ on S, and any $t \ge 0$*

$$\lim_{\delta \downarrow 0} \lim_{N \uparrow \infty} \mathbb{E}\left|\int \varphi(s)\hat{\pi}_t(ds) - \int \varphi(s)\pi_t(ds)\right|^2 = 0 \tag{20}$$

A detailed proof of proposition 3 will be included in an upcoming journal submission. The main argument in this proof is the following. By a result in [18], $\hat{\pi}_k^\delta$ converges to π_k^δ, (in the sense of (20)), as $N \uparrow \infty$. On the other hand, property (8) can be used to show that $\pi_{k(t)}^\delta$ converges to π_t, as $\delta \downarrow 0$. Then, (20), follows easily by combining these two limits.

The convergence result (20) may seem weak, as it involves a fixed "test function" φ. However, this result implies almost sure weak convergence of $\hat{\pi}_t$ to π_t, by a usual argument using the fact that S is separable and complete [1].

5. NUMERICAL EXAMPLE

This section presents numerical simulation results, for the application of the sequential Monte Carlo algorithm of the previous section, to the problem of estimating the angular velocity of a rigid body, from noisy observations of its attitude. This problem arises in spacecraft navigation, under the name of *gyroless estimation* [19, 20].

Assume a rigid body, (*e.g.* robot, vhicle, or satellite), is performing rotational motion. Determination of its attitude would require knowing its orientation with respect to three, reference, orthonormal directions. However, here, it is assumed this orientation is only known with respect to one reference direction. Mathematically, at time $t \geq 0$, this is represented by a unit vector Y_t.

Consider the resulting process Y with values on the unit sphere $S^2 \subset \mathbb{R}^3$. In the absence of any noise or uncertainty, Y exactly represents the orientation of the body, so its evolution is given by [21]

$$\frac{dY_t}{dt} = -Y_t \times X_t \tag{21}$$

where X_t, a vector in $\mathbb{R}^3$, is the angular velocity and $\times$ denotes cross product. Here, the prsence of noise or observation uncertainty is modeled by the stochastic differential equation

$$dY_t = -Y_t \times \{X_t dt + \circ dB_t\} \tag{22}$$

where B is a Brownian motion in $\mathbb{R}^3$.

A continuous time filtering problem is obtained if one wishes to estimate the angular velocity X, from observations of the attitude Y. Recall [1, 2], indeed, that the optimal least squares estimator of X_t, based on observations $\mathcal{Y}_t$, is the first order moment of the posterior distribution π_t, distribution of X_t given $\mathcal{Y}_t$.

In the following, it will be shown that the dynamical model (22) can be recast in the form (2). Then, the algorithm of Section 4 will be specified and applied to the case where X is a constant, $X_t \equiv x$.

To rewrite equation (22), consider three vector fields V_1, V_2, V_3, defined on S^2 as follows, for $y \in S^2$ with $y = (y_1, y_2, y_3)$,

$$V_r^i(y) = \sum_{j=1}^{3} \epsilon_{r,ij} y_j \tag{23}$$

where $V_r^i(y)$ are the coordinates of the vector $V_r(y)$, and $\epsilon_{r,ij}$ is alternating, (changes sign when two indices are exchanged), with $\epsilon_{1,23} = -1$. Also, let $H : \mathbb{R}^3 \times S^2 \to TS^3$ be given by

$$H^i(s, y) = \sum_{j=1}^{3} \sum_{r=1}^{3} \epsilon_{r,ij} s_r y_j \tag{24}$$

Replacing (23) and (24) in (2), (with the Brownian motion B being written $B = (B^1, B^2, B^3)$), a direct calculation leads to (22). This shows the problem under consideration is of the general form introduced in Section 1.

In order to apply the algorithm of Section 4 to this problem, it is necessary to find a suitable connector map I, which satisfies condition (10). Recall this expression involves the metric and connection (3) and (4).

By verification of (3) and (4), it is possible to show the required metric $\langle \cdot, \cdot \rangle$ and connection ∇ are none other than the usual Riemannien metric and Levi-Civita connection on S^2. Therefore, as remarked in the discussion after Proposition 2, Section 3, it is always possible to take for I a geodesic connector. In the present case, this can be written

$$I(p, q) = \frac{\arcsin \|\Pi_p(q)\|}{\|\Pi_p(q)\|} \, \Pi_p(q) \tag{25}$$

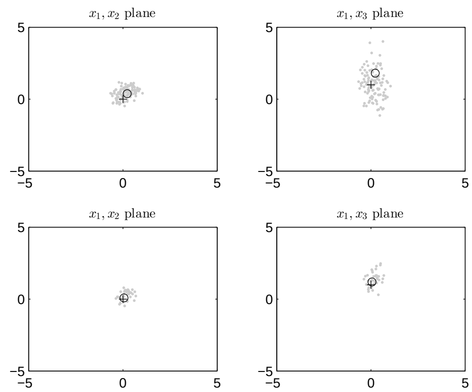

Fig. 1. Particles distribution (grey); estimate ($\circ$); true value ($+$)

where $\Pi_p(q) = p \times q \times p$, and $\| \cdot \|$ denotes Euclidea length in $\mathbb{R}^3$. Another connector map, which also verifies condition (10), but which is much less costly, in terms of computation, is

$$I(p, q) = \Pi_p(q) \tag{26}$$

which results from (25), in the limit where $\|\Pi_p(q)\|$ is small.

This connector map was used in implementing the algorithm of Section 4, for the case where $X_t \equiv x$ is constant, $x = (0, 0, 1)$. Figure 1 shows the distribution of $N = 1000$ particles in the (x_1, x_2) and (x_1, x_3) planes at times $T = 1.5$ (top row) and $T = 3$ (bottom row), with the sampling interval δ being equal to 0.02.

In the figure, the $+$ designates the position of x, while the $\circ$ designates the arithmetic mean of the particles. The particles were initially generated from a normal distribution μ with mean $(0.5, 0.5, 1)$ and variance 1. A large value of N was chosen for visualisation. It is possible to use $N = 100$ with a similar performance.

Figure 1 shows the algorithm of Section 4 is able to recover x using 150 samples $Y_{k\delta}$, from the observation process. It is interesting to note the larger variability of the particle distribution in the x_3 direction, apparent in the right column of the figure. This is because $Y_0 = (0, 0, 1)$, so that, initially, the component of x along this direction has no effect on the position of Y.

6. RELATION TO PRIOR WORK

A special case of the filtering problem studied in the present paper was treated in [12], using a similar sequential Monte Carlo approach. In [12], the authors only considered the case where the manifold M is a Stiefel manifold. The present paper provides a fully general framework.

Theoretical foundation for both papers comes from [6], whose main idea is to extend the approach of [13], from Lie groups to general manifolds.

Readers who are unfamiliar with stochastic calculus in manifolds should consult the highly readable introduction [7].

In the mathematics literature, connector maps were defined in [9]. The present paper introduces their application to the problem of filtering with observations in a manifold. In the field of optimisation on manifolds, a closely related general construction for iterative algorithms can be found in [22].

7. REFERENCES

[1] A. Bain and D. Crisan, *Fundamentals of stochastic filtering*, Springer, 2009.

[2] A. H. Jazwinski, *Stochastic processes and filtering theory*, Academic Press, 1970.

[3] T. E. Duncan, "Some filtering results in Riemannian manifolds," *Information and Control*, vol. 35, pp. 182–195, 1977.

[4] S. K. Ng and P. E. Caines, "Nonlinear filtering in Riemannian manifolds," *IMA J. Math. Control I.*, vol. 2, pp. 25–36, 1985.

[5] M. Pontier and J. Szpirglas, "Filtrage non linéaire avec observations sur une variété," *Stochastics*, vol. 15, no. 2, pp. 121–148, 1985.

[6] S. Said and J. H. Manton, "On filtering with observation in a manifold: Reduction to a classical filtering problem," *SIAM J. Control. Optim.*, vol. 51, no. 1, pp. 767–783, 2013.

[7] J. H. Manton, "A primer on stochastic differential geometry for signal processing," *IEEE J. Sel. Topics Signal Process*, vol. 7, no. 4, pp. 681–699, 2013.

[8] M. Emery, *Stochastic calculus in manifolds*, Springer-Verlag (Universitext), 1989.

[9] J. Picard, "Calcul stochastic avec sauts sur une variété," *Séminaire de Probabilités (Strasbourg)*, vol. 25, pp. 196–219, 1991.

[10] I. Chavel, *Riemannian geometry: A modern introduction*, Cambridge University Press, 2006.

[11] J. Faraut, *Analysis on Lie groups: An introduction*, Cambridge University Press, 2008.

[12] J. Boulanger, N. Le Bihan, S. Said, and J.H. Manton, "Monte-carlo estimation from observation on stiefel manifold," in *ICASSP*, 2014, pp. 4195–4199.

[13] J. T. Lo, "Signal detection for bilinear systems," *Information Sciences*, vol. 9, pp. 249–378, 1975.

[14] A. Budhiraja, L. Chen, and C. Lee, "A survey of numerical methods for nonlinear filtering problems," *Physica D: Nonlinear Phenomena in Data Assimilation*, vol. 230, no. 1-2, pp. 27–36, 2007.

[15] A. Doucet and A. M. Johansen, "A tutorial on particle filtering and smoothing: Fifteen years later," in *The oxford handbook of nonlinear filtering*, D. Crisan and B. Rozovskii, Eds. Oxford University Press, 2011.

[16] P. Del Moral, *Feynam-Kac formulae: Geealogical and interacting particle systems with applications*, Springer (Probability and Applications), 2004.

[17] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, Springer (Stochastic modelling and applied probability), 1992.

[18] D. Crisan and A. Doucet, "A survey of convergence results on particle filtering methods for practioners," *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 736–746, 2002.

[19] Z. Hongmei and D. Zhenglong, "UKF-based attitude determination method for gyroless satellite," *Journal of Systems Engineering and Electronics*, vol. 15, no. 2, pp. 105–109, 2004.

[20] M.S. Grewal and M. Shiva, "Application of kalman filtering to gyroless attitude determination and control system for environmental satellites," in *CDC*, 1995, pp. 1544–1552.

[21] P. D. Groves, *Principles of GNSS, Inertial, and multisensor integrated navigation systems*, Artech House, 2008.

[22] J.H. Manton, "A framework for generalising the newton method and other iterative methods from euclidean space to manifolds," *Numerische Mathematik*, pp. 1–35, 2014.