Proceedings Paper:
Kim, J orcid.org/0000-0002-3456-6614 and Richardson, R (2016) Negative-free approximation of probability density function for nonlinear projection filter. In: 2016 IEEE 55th Conference on Decision and Control (CDC). 55th IEEE Conference on Decision and Control, 12-14 Dec 2016, Las Vegas, Nevada, United States. IEEE , pp. 3738-3743. ISBN 978-1-5090-1837-6

https://doi.org/10.1109/CDC.2016.7798832

Negative-free approximation of probability density function for nonlinear projection filter*

Jongrae Kim[1], and Robert Richardson[1]

Abstract— **Several approaches have been developed to estimate probability density function (pdf). The pdf has two important properties: the integration of pdf over whole sampling space is equal to 1 and the value of pdf in the sampling space is greater than or equal to zero. The first constraint can be easily achieved by the normalisation. On the other hand, it is very hard to impose the non-negativeness in the sampling space. In the pdf estimation, some areas in the sampling space might have negative pdf values. It produces unreasonable moment values such as negative probability or variance. A transformation to guarantee the negative-free pdf over a chosen sampling space is presented and it is applied to the nonlinear projection filter. The filter approximates the pdf to solve nonlinear estimation problems. For simplicity, one-dimensional nonlinear system is used as an example to show the derivations and it can be readily generalised for higher dimensional systems. The efficiency of the proposed method is demonstrated by numerical simulations. The simulations also show that to achieve the same level of approximation error in the filter the required number of basis functions with the transformation is a lot smaller compared to the ones without transformation. This will be hugely benefited when the filter is used for high dimensional systems, which requires significantly less computational cost.**

I. INTRODUCTION

Nonlinear estimation has been one of the most studied topics in control theory in the last half century since the success of Kalman filter [1] in the Apollo mission [2]. The orbit estimation problem in the Apollo mission is nonlinear and the initial success of the Kalman filter relied on the assumption of the reasonable error bounds in the perturbed state, which can be approximated by a linear dynamics, and it is called the extended Kalman filter. The extended Kalman filter has been applied to many dynamical systems including the attitude estimation problem for satellite [3], the target tracking problem such as α-β filter [4] and α-β-γ filter [5], a traffic management [6], and plethora of many other examples.

It was, however, quickly realised that the limitation of the extended Kalman filter or linearised approach to nonlinear estimation, in general. If the initial guess is not close enough to the region, where the linearisation is valid, then the estimated state could diverge. In order to resolve the issues, unscented Kalman filter is proposed [7]. Using a nonlinear transformation, the unscented Kalman filter improves the linearisation accuracy up to the 3rd-order in Taylor series expansion [8]. Several examples of the unscented Kalman filter applications can be found in [9]–[11].

True nonlinear estimation can be obtained by solving the Fokker-Plank equation [12], which governs the propagation of the underlying joint probability density function (pdf) for a given nonlinear system, while the propagation is updated using the Bayes' rule whenever a measurement is available. The particle filter is based on these mechanisms [13], where the required multi-dimensional integration is performed by the Monte-Carlo sampling method. One of the main issues in the particle filter is that the sampling points are quickly concentrated in a few sampling spaces and the re-sampling methods are developed to rectify the problem, which does not solve another problem with highly variable cases [14].

The nonlinear projection filter presented in [15] solves the Fokker-Planck equation in a different way compared to the particle filter. The solution of the Fokker-Planck equation is assumed to be a linear combination of a set of orthogonal basis functions. The dynamics of the time-varying variables to combine the basis function is obtained and used for propagating the time-varying variables; and they are updated based on the Bayes' rule when a measurement is available. It has been applied to a target tracking problem [16]–[18] and there are a few variations of the filter via combining with particle samplings [19], [20] or performing the multi-dimensional integration combining with multiple sensor measurements [21].

As it is emphasised, however, in [21], the approximated pdf might not satisfy all requirements for any pdf to content. The integration of pdf over whole sampling space must be equal to 1. This condition can be easily met for the filter by normalising the estimated pdf every propagation and/or update step. On the other hand, the non-negativeness condition, which requires the pdf greater than equal to zero in every point in the sampling space, is the one that hard to comply with. It is not easy to impose the non-negativeness during the propagation and the update steps. The negativeness issue will not be resolved via simply increasing the number of basis functions. This is a similar problem in the Fourier transform called the Gibbs phenomena [22].

The negativeness in the approximated pdf causes inaccurate and/or physically impossible values in the moment calculation. The first moment, i.e. the average, would not suffer directly from the discrepancy in the pdf. Higher moments including the variance would be directly affected: for example, a calculated variance would be negative, which is impossible by the definition. In order to impose the non-negative condition, a novel transformation approach is

*This research is supported by EPSRC Research Grant, EP/N010523/1, Balancing the impact of city infrastructure engineering on natural systems using robots.

[1]J.Kim and R.Richardson are with Institute of Design, Robotics & Optimisation, School of Mechanical Engineering University of Leeds, Leeds LS2 9JT, UK. {menjkim, r.c.richardson}@leeds.ac.uk

proposed using a log function. This is partially inspired by the method presented in [23].

This paper organises as follows: firstly, the summary of the nonlinear projection filter is presented; secondly, a transformation for prohibiting the negativeness of pdf is presented along with the detail derivations of the filter for a one-dimensional nonlinear system, where we highlight that for many cases analytic expression can be obtained for a general type of multi-dimensional nonlinear systems. Thirdly, numerical simulations are performed to show the effectiveness of the proposed transformation and to compare the results with the ones without the transformation. Finally, the conclusions are presented.

II. NONLINEAR PROJECTION FILTER

One of the standard forms of nonlinear stochastic system is given by

$$d\mathbf{x} = \mathbf{f}(\mathbf{x})dt + G(\mathbf{x})d\beta \tag{1}$$

for $t \geq t_0$, where $\mathbf{x}$ is an n-dimensional state, $\beta(t)$ is a b-dimensional Brownian motion, $\mathrm{E}(\beta\beta^T) = Q(t)dt$, $E(\cdot)$ is the expectation, $(\cdot)^T$ is the transpose, $\mathbf{f}(\cdot)$ is an n-dimensional nonlinear function, and $G(\cdot)$ is an $n \times b$ matrix. The measurement is given by

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \tag{2}$$

for $k \geq 1$, where $\mathbf{y}_k$ is an m-dimensional real vector, $\mathbf{v}_k$ is a white noise whose covariance is R_k, and $\mathbf{h}(\cdot)$ is the m-dimensional measurement function. It is assumed that $\beta(t)$ and $\mathbf{v}_k$ are not correlated.

The conditional probability density function can be written as follows:

$$p(t,\mathbf{x}|\ Y_k) = \frac{p(t,\mathbf{x},Y_k)}{p(Y_k)},$$

where Y_k is the collection of all measurement up to $t_k \leq t$. The time evolution of the conditional pdf is governed by the Fokker-Planck equation:

$$\frac{\partial p}{\partial t} = -\sum_{i=1}^{n} \frac{\partial (p f_i)}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 \left[p \left(GQG^T\right)_{i,j}\right]}{\partial x_i \partial x_j} \tag{3}$$

where x_i and f_i are the i-th element of $\mathbf{x}$ and $\mathbf{f}(\mathbf{x})$, respectively, $(GQG^T)_{i,j}$ is the i-th row and j-th column element of the matrix, GQG^T, and the initial condition is given by $p(t_0,\mathbf{x})$. Once the conditional pdf is obtained, any moment can be calculated as follows:

$$E\left(x_1^{\ell_1} x_2^{\ell_2} \ldots x_n^{\ell_n}\right) = \int_\Omega \left(x_1^{\ell_1} x_2^{\ell_2} \ldots x_n^{\ell_n}\right)\ p(\mathbf{x},t|\ Y_k)d\mathbf{x},$$

where Ω is a closed bounded subset of $\mathbb{R}^n$ and ℓ_i for $i = 1,2,\ldots,n$ is a real number.

The nonlinear projection filter is derived by assuming the pdf as a linear combination of basis functions as follows:

$$p(t,\mathbf{x}|\ Y_k) \approx p_N(t,\mathbf{x}|\ Y_k) = \sum_{\ell=1}^{N} c_\ell(t)\phi_\ell(\mathbf{x}), \tag{4}$$

where the basis functions are orthogonal

$$\int_\Omega \phi_i(\mathbf{x})\ \phi_j(\mathbf{x})d\mathbf{x} = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j \end{cases} \tag{5}$$

for $i, j = 1,2,3,\ldots,N-1,N$.

*P*ropagation: Substituting (4) into (3), projecting onto ϕ_q and integrating over Ω provide the following differential equation for $\mathbf{c}(t)$:

$$\dot{\mathbf{c}}(t) = (\mathbf{A}_1 + \mathbf{A}_2)\,\mathbf{c}(t), \tag{6}$$

where $(\dot{})$ is the derivative by time, the initial condition is equal to

$$\mathbf{c}(t_0) = \int_\Omega p(t_0,\mathbf{x})\phi(\mathbf{x})d\mathbf{x},$$

$\phi(\mathbf{x}) = [\phi_1(\mathbf{x}),\ \phi_2(\mathbf{x}),\ \ldots,\ \phi_N(\mathbf{x})]^T$, and the ith-row and jth-column element of $\mathbf{A}_1$ or $\mathbf{A}_2$ is given by

$$[\mathbf{A}_1]_{i,j} = -\sum_{k=1}^{n} \int_\Omega \frac{\partial\left[\phi_j f_k\right]}{\partial x_k}\phi_i d\mathbf{x},$$

$$[\mathbf{A}_2]_{i,j} = \frac{1}{2}\sum_{k=1}^{n}\sum_{\ell=1}^{n} \int_\Omega \frac{\partial^2 \left[\phi_j \left(GQG^T\right)_{k,\ell}\right]}{\partial x_k \partial x_\ell}\phi_i d\mathbf{x}.$$

Update: The conditional probability density function is updated by the Bayes' rule at the measurement available from a sensor [13]. Substituting (4) into the Bayes' rule, the update equation is obtained as follows:

$$\mathbf{c}(t_k^+) = \eta\left[Y(\mathbf{y}_k)\ \mathbf{c}(t_k^-)\right] \tag{8}$$

where t_k^+ and t_k^- indicate after and before the k-th measurement are considered,

$$Y(\mathbf{y}_k) = \int_\Omega p(\mathbf{y}_k|\mathbf{x})\phi\phi^T d\mathbf{x}$$

and η is the normalising constant. The full details on the nonlinear projection filter derivation can be found in [15].

III. NEGATIVE-FREE PDF APPROXIMATION

In the nonlinear projection filter with the basis function approach, the main issue is the negativeness of the probability in some places of the sampling space. To reduce the size of the negative probability area, the number of basis functions must increase quite significantly. The computational cost increases exponentially as the number of basis functions increases. This would be the one of main obstacles in the applications of the filter to some reasonable size systems

For simplicity, all derivations from now on are based on the following nonlinear system [15]:

$$dx = \sin(x)dt + d\beta \tag{9}$$

with a measurement equation equal to

$$y_k = x_k + v_k \tag{10}$$

where the variance of the process noise, $\beta(t)$, is equal to q, and the variance of the measurement noise, v_k, is equal to r. Note that the similar procedures can be applied to various types of nonlinear function with higher-dimensional states.

A. Transformation

Define $\rho(t,x|Y_k)$ using the joint pdf, $p(t,x|Y_k)$

$$\rho(t,x|Y_k) = \ln[p(t,x|Y_k) + \varepsilon] \tag{11}$$

where ε is a small positive real number in order to prevent the log-function becoming the negative infinity or a large negative value for $p(t,x|Y_k)$ to approach to zero arbitrary close. As the domain of x, i.e., Ω, less likely includes the regions where the joint pdf is exactly equal to zero, the following assumption is introduced without imposing any strong restriction:

Assumption 3.1: Ω is chosen such that δ in the following inequality is close to zero but not exactly zero.

$$0 < \delta \leq \min_{\substack{t\geq 0\\ x\in\Omega}}[p(t,x|Y_k)] \tag{12}$$

Hence, ε can be set to very close to zero and a lower bound of $\rho(t,x|Y_k)$ is given by

$$-\infty < \ln(\delta+\varepsilon) \leq \rho(t,x|Y_k) \tag{13}$$

and $\rho(t,x)$ can be negative unlike $p(t,x)\geq 0$ for all $t\in[0,\infty)$ and $x\in\Omega$. The inverse transformation is simply

$$p(t,x|Y_k) = e^{\rho(t,x|Y_k)} - \varepsilon \tag{14}$$

In practice, ε can be chosen to be very close to zero allowed by the given numerical precision.

B. Propagation

Take the time derivative of the inverse transformation

$$\frac{dp(t,x|Y_k)}{dt} = e^{\rho(t,x|Y_k)}\frac{d\rho(t,x|Y_k)}{dt} \tag{15}$$

Substitute (14) and (15) into the Fokker-Planck equation, (3),

$$e^{\rho}\frac{d\rho}{dt} = -e^{\rho}\sin(x)\rho' + (e^{\rho}-\varepsilon)\cos(x) + \frac{q}{2}\left[e^{\rho}(\rho')^2 + e^{\rho}\rho''\right]$$

where $(\cdot)' = d(\cdot)/dx$ and $(\cdot)'' = d^2(\cdot)/dx^2$. Divide both sides by e^{ρ}, which is greater than ε,

$$\dot\rho = -\sin(x)\rho' + \frac{e^{\rho}-\varepsilon}{e^{\rho}}\cos(x) + \frac{q}{2}\left[(\rho')^2 + \rho''\right] \tag{16}$$

In the nonlinear projection filter [15], it is proposed that the joint pdf is approximated by (4). The approximation uses a set of basis functions and it does not force the approximated values at all x in Ω be non-negative. On the other hand, as $\rho(t,x|Y_k)$ can be positive or negative, it is natural to approximate $\rho(t,x|Y_k)$ instead of $p(t,x|Y_k)$ by a set of basis functions without concerning to impose non-negativeness. Let

$$\rho(t,x) = \mathbf{c}^T(t)\phi(x) \tag{17}$$

and substitute the approximation into (16)

$$\begin{aligned}\dot{\mathbf{c}}^T\phi = &-\sin(x)\left(\mathbf{c}^T\phi'\right) + \frac{e^{\mathbf{c}^T\phi}-\varepsilon}{e^{\mathbf{c}^T\phi}}\cos(x)\\ &+\frac{q}{2}\left[(\mathbf{c}^T\phi')^2 + \mathbf{c}^T\phi''\right]\end{aligned} \tag{18}$$

where the N-number of cosine-basis functions are chosen

$$\phi_i(x) = \begin{cases} 1/\sqrt{\Delta_x} & \text{for } i=1,\\ \sqrt{\frac{2}{\Delta_x}}\cos\left[\kappa_i(s-a_x)\right] & \text{for } 2\leq i\leq N\end{cases} \tag{19}$$

Δ_x is equal to $b_x - a_x$, a_x is the lower bound of x, b_x is the upper bound of x, i.e., $x\in[a_x,b_x]$ and $\kappa_i = [(i-1)\pi]/\Delta_x$ for $i = 1,2,\ldots,N$. The basis functions satisfy the orthogonality condition.

Proposition 3.2: The basis functions are satisfied

$$\phi'(x) = -K\psi(x) \tag{20a}$$

$$\phi''(x) = -K^2\phi(x) \tag{20b}$$

where

$$K = \operatorname{diag}\begin{bmatrix}\kappa_1 & \kappa_2 & \ldots & \kappa_N\end{bmatrix}, \tag{21a}$$

$$\psi(x) = \sqrt{\frac{2}{\Delta_x}}\begin{bmatrix}0\\ \sin[\kappa_2(s-a_x)]\\ \sin[\kappa_3(s-a_x)]\\ \vdots\\ \sin[\kappa_N(s-a_x)]\end{bmatrix}, \tag{21b}$$

Proof: The proof is straightforward and omitted. The full derivations can be found in [18]. ■

Using the proposition 3.2, (18) becomes

$$\begin{aligned}\dot{\mathbf{c}}^T\phi = &-\sin(x)\left(\mathbf{c}^TK\psi\right) + \frac{e^{\mathbf{c}^T\phi}-\varepsilon}{e^{\mathbf{c}^T\phi}}\cos(x)\\ &+\frac{q}{2}\left[(\mathbf{c}^TK\psi)^2 - \mathbf{c}^TK^2\phi\right]\end{aligned} \tag{22}$$

Multiply ϕ both sides and integrate over Ω

$$\begin{aligned}\dot{\mathbf{c}} = &-\int_\Omega \sin(x)\left(\psi^TK\phi\right)dx\ \mathbf{c} + \int_\Omega \frac{e^{\mathbf{c}^T\phi}-\varepsilon}{e^{\mathbf{c}^T\phi}}\cos(x)\phi dx\\ &+\frac{q}{2}\int_\Omega(\mathbf{c}^TK\psi)^2\phi dx - \frac{q}{2}K^2\mathbf{c}\end{aligned} \tag{23}$$

Re-arrange it

$$\begin{aligned}\dot{\mathbf{c}} = &-\left[\int_\Omega \sin(x)\left(\psi^TK\phi\right)dx + \frac{q}{2}K^2\right]\mathbf{c}\\ &+\int_\Omega \frac{e^{\mathbf{c}^T\phi}-\varepsilon}{e^{\mathbf{c}^T\phi}}\cos(x)\phi dx + \frac{q}{2}\int_\Omega(\mathbf{c}^TK\psi)^2\phi dx\end{aligned} \tag{24}$$

Some of the integrals in the right hand side of (24) have analytic expression. The first row of the first integral is zero,

$$\left[\int_\Omega \sin(x)\left(\psi^TK\phi\right)dx\right]_{1j} = 0 \tag{25}$$

for $j = 1,2,\ldots,N$. The diagonal terms are given by

$$\left[\int_\Omega \sin(x)\left(\psi^TK\phi\right)dx\right]_{ii} = \frac{2(i-1)^2\pi^2[\sin(b)-\sin(a)]}{\Delta_x[\Delta_x^2 - 4\pi^2(i-1)^2]} \tag{26}$$

for $i = 2,3,\ldots,N$. Note that Δ_x must not be equal to $2\pi(i-1)$ and this is always possible as the size of the sampling space can be always chosen to be different from an integer multiple of π. Hence, the following assumption is introduced.

Assumption 3.3: Δ_x is not equal to $\ell\pi$ for all ℓ in $\mathbb{N}$, where $\mathbb{N}$ is the set of natural number.

The assumption is not strong as Ω can be always adjusted so that the boundary length is not equal to $\ell\pi$. The non-diagonal terms have an analytic expression but the expression is too long and omitted.

Consider the magnitude difference between the second term in (24) and the one with ε equal to zero.

$$\begin{aligned}\Delta &= \left\| \int_\Omega \frac{e^{\mathbf{c}^T\phi}-\varepsilon}{e^{\mathbf{c}^T\phi}} \cos(x)\phi dx - \int_\Omega \cos(x)\phi dx \right\|_\infty \\ &= \frac{\varepsilon}{e^{\mathbf{c}^T\phi}} \left\| \int_\Omega \cos(x)\phi dx \right\|_\infty \le \frac{\varepsilon}{e^{\mathbf{c}^T\phi}} \le \frac{\varepsilon}{\delta+\varepsilon} \end{aligned} \tag{27}$$

where $\|\cdot\|_\infty$ is the infinity norm. Set ε equal to γ-times smaller than δ, where $0 \le \gamma \ll 1$, the error is bounded by the following inequality:

$$\Delta \le \frac{\varepsilon}{\delta+\varepsilon} = \frac{\gamma\delta}{\delta+\gamma\delta} = \frac{\gamma}{1+\gamma} \ll 1 \tag{28}$$

where ε is very small by Assumption 3.1. Given the choice of $\varepsilon = \gamma\delta$, the second integration term is approximated by

$$\int_\Omega \frac{e^{\mathbf{c}^T\phi}-\varepsilon}{e^{\mathbf{c}^T\phi}} \cos(x)\phi dx \approx \int_\Omega \cos(x)\phi dx \tag{29}$$

and the integral with the approximation has an analytic expression. Define $\mathbf{b} = \int_\Omega \cos(x)\phi dx$ and it becomes

$$\mathbf{b} = \begin{bmatrix} \frac{\sin(b_x)-\sin(a_x)}{\sqrt{\Delta_x}} \\ \frac{\sqrt{2}\ \Delta_x^{3/2}\left[\sin(b_x)(-1)^{2+1}-\sin(a_x)\right]}{\Delta_x^2-(2-1)^2\pi^2} \\ \frac{\sqrt{2}\ \Delta_x^{3/2}\left[\sin(b_x)(-1)^{3+1}-\sin(a_x)\right]}{\Delta_x^2-(3-1)^2\pi^2} \\ \vdots \\ \frac{\sqrt{2}\ \Delta_x^{3/2}\left[\sin(b_x)(-1)^{N+1}-\sin(a_x)\right]}{\Delta_x^2-(N-1)^2\pi^2} \end{bmatrix} \tag{30}$$

The k-th element of the last integration term in the right hand side of (24) can be reformulated as follows:

$$\begin{aligned}\left[\int_\Omega (\mathbf{c}^T K\psi)^2\phi dx\right]_k &= \int_\Omega (\mathbf{c}^T K\psi)^2\phi_k dx \\ &= \int_\Omega \sum_{i=1}^N \sum_{j=1}^N c_i c_j \kappa_i \kappa_j \psi_i \psi_j \phi_k dx \\ &= \left[\int_\Omega (\psi^T \otimes \psi^T)\phi_k dx\right](K\otimes K)(\mathbf{c}\otimes\mathbf{c}) \\ &= \mathbf{s}_k^T (K\otimes K)(\mathbf{c}\otimes\mathbf{c}) \end{aligned} \tag{31}$$

where $\mathbf{s}_k^T$ is equal to the integration of $(\psi^T\otimes\psi^T)\phi_k$ over Ω, and its size is $N^2\times 1$. Define

$$S = \begin{bmatrix}\mathbf{s}_1 & \mathbf{s}_2 & \dots & \mathbf{s}_N\end{bmatrix}^T \tag{32}$$

whose size is $N\times N^2$. Note that S has most of the terms equal to zero. The last integration term of (24) is given by

$$\int_\Omega (\mathbf{c}^T K\psi)^2\phi dx = S(K\otimes K)(\mathbf{c}\otimes\mathbf{c}) \tag{33}$$

Consider ℓ-th element of $\mathbf{s}_k$

$$(\mathbf{s}_k)_\ell = \int_\Omega \psi_i\psi_j\phi_k dx \tag{34}$$

where $\ell = (i-1)N+j$. It is trivial that all integrations with ψ_1 is equal to zero as $\psi_1 = 0$. Also, it can be easily shown that all integrations in the following index set are equal to zero:

$$\{(i,j,k) | i \ne j+k-1 \text{ or } j \ne i+k-1 \text{ or } k \ne i+j-1\} \tag{35}$$

for i, j, k are in $[1,N]$. For the case of either $i = j+k-1$ or $j = i+k-1$, $(\mathbf{s}_k)_\ell$ is equal to $1/\sqrt{2\Delta_x}$. For the case of $k = i+j-1$, $(\mathbf{s}_k)_\ell$ is equal to $-1/\sqrt{2\Delta_x}$.

Finally, the differential equation of $\mathbf{c}(t)$ can be written in a compact form as follows:

$$\dot{\mathbf{c}} = \mathbf{f}_c(\mathbf{c}) = -A\mathbf{c} + \mathbf{b} + \frac{q}{2}\left[S(K\otimes K)\right](\mathbf{c}\otimes\mathbf{c}) \tag{36}$$

where

$$A = \int_\Omega \sin(x)\left(\psi^T K\phi\right)dx + \frac{q}{2}K^2 \tag{37}$$

Note that S and $(K\otimes K)$ are sparse and can be obtained for large size N without requiring long computation and large memory size. For each sampling time, the above differential equation is used to propagate $\mathbf{c}(t)$ from $t = t_k$ to $t = t_{k+1}$ as follows:

$$\mathbf{c}(t_{k+1}) = \Phi[t_{k+1}, t_k, \mathbf{c}(t_k)] \tag{38}$$

where $t_{k+1} = t_k + \Delta t_k$, $\Phi(\cdot)$ is the transition function. In the original propagation with $p(t,x|Y_k) = \mathbf{c}^T(t)\phi(x)$ in [15], the differential equation for $\mathbf{c}(t)$ is linear and the transition function can be obtained off-line. The propagation equation with $\rho(t,x|Y_k)$, (36) is nonlinear. Hence, the transition function has to be calculated on-line. The following algorithm is proposed to propagate $\mathbf{c}(t_k)$:

Algorithm 3.4: $\mathbf{c}(t_k)$ propagation to $\mathbf{c}(t_{k+1})$

1) Select n_p, a positive integer, and set $\Delta t = \Delta t_k/n_p$, $\ell = 0$, and $\mathbf{c}_\ell = \mathbf{c}(t_k)$
2) Calculate

$$\mathbf{c}_{new} = \mathbf{c}_\ell + \Delta t\ \mathbf{f}_c[\mathbf{c}_\ell] \tag{39}$$

3) Increase ℓ by one, i.e., $\ell \leftarrow \ell+1$
4) Set $\mathbf{c}_\ell = \mathbf{c}_{new}$
5) If $\ell = n_p$, set $\mathbf{c}(t_{k+1}) = \mathbf{c}_\ell$ and stop. Otherwise, go to step 2).

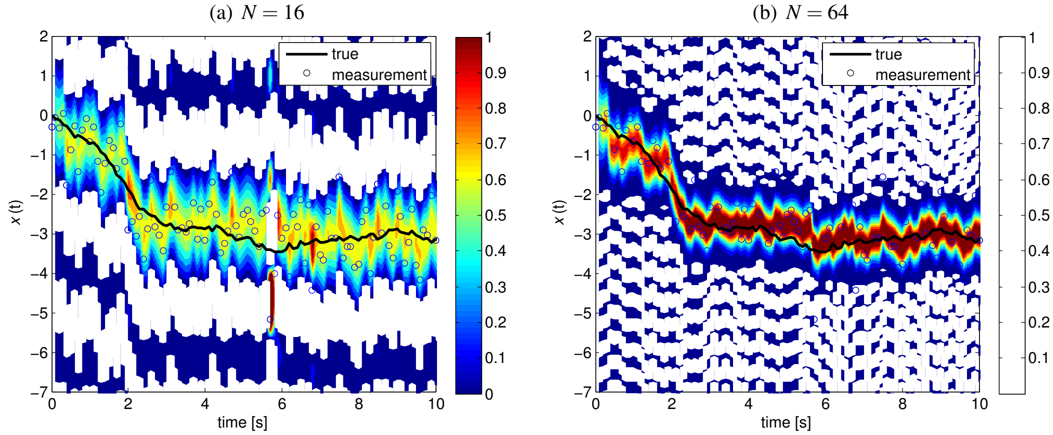

Fig. 1. Without the transformation: the regions in white colour are the space where the pdf is negative. The heat-map indicates the time-history of the pdf over the sampling space. The true state is shown by the black solid line and the measurements are depicted by the hollow circles.

C. Update

The conditional probability density function is updated using the current measurement by Bayes' rule [13]:

$$p(t_k^+, x|Y_k) = \eta\, p(y_k|x)\; p(t_k^-, x|Y_{k-1}) \tag{40}$$

where η is the normalising constant to be determined, and $p(y_k|x)$ is the sensor model. To determine the normalising constant, η, integrate (40) over the sampling space, Ω,

$$\int_\Omega p(t_k^+, x|Y_k)dx = \eta \int_\Omega p(y_k|x)\; p(t_k^-, x|Y_{k-1})dx \tag{41}$$

where the left hand side is equal to 1 as it is the total probability. Hence,

$$\eta = \left[\int_\Omega p(y_k|x)\; p(t_k^-, x|Y_{k-1})dx\right]^{-1} \tag{42}$$

Substituting the inverse transformation, (14), into, (42), and η is calculated as

$$\eta = \left\{\int_\Omega p(y_k|x)\; \left[e^{\mathbf{c}^T(t_k)\phi(x)} - \varepsilon\right] dx\right\}^{-1} \tag{43}$$

This integral must be performed on real-time and the Monte-Carlo integration method was proposed in [21]. Sample the random points based on the sensor model provides an efficient integration performance, which reduces the required sampling points significantly. See [21] for details.

Substituting the pdf approximation with the transformation into (40)

$$e^{\rho(t_{k+1}, x|Y_k)} - \varepsilon = \eta\, p(y_k|x)p(t_k^-, x|Y_{k-1}) \tag{44}$$

Take log both sides

$$\rho(t_{k+1}, x|Y_k) = \log\left[\eta\, p(y_k|x)p(t_k^-, x|Y_{k-1}) + \varepsilon\right] \tag{45}$$

and replace $\rho(t_{k+1})$ by the approximation

$$\mathbf{c}^T(t_{k+1})\phi(x) = \log\left[\eta\, p(y_k|x)p(t_k^-, x|Y_{k-1}) + \varepsilon\right] \tag{46}$$

Finally, multiply $\phi^T(x)$ both sides and integrate over Ω

$$\mathbf{c}(t_{k+1}) = \int_\Omega \log\left[\eta\, p(y_k|x)p(t_k^-, x|Y_{k-1}) + \varepsilon\right] \phi(x)\; dx, \tag{47}$$

which provides the update equation. Again, this integration must be performed on real-time and the efficient sampling could be implemented based on the a priori joint pdf, $p(t_k^-, x|Y_{k-1})$.

IV. Simulation

The variances of the noises in the system and the sensor are set to $q = 0.5$ and $r = 0.5$, respectively. The measurement is taken every 0.1s. The state variable, x, is in the range of $[-2.2\pi, 2.2\pi]$ and the initial state is equal to zero. The sensor is assumed to have the following noise model:

$$p(y_k|x) = \frac{1}{\sqrt{2\pi r}}\; e^{-(y_k - x)^2/(2r)} \tag{48}$$

To compare the performance, the same measurement set is used for the filter without the transformation and the one with the transformation. ε for the transformation is set to 10^{-300}.

Figure 1 shows the results without the transformation for the number of basis functions equal to 16 or 64. Both cases have large areas with the negative pdf, which is depicted by the white colour. For $N = 16$, around the time equal to 5.8s, the negative region even appears across the true value. For $N = 64$ the negativeness still exist in the wide areas. In order to reduce the negative region significantly, N must be very large. This causes serious issues in applying the filter for higher dimensional systems as the number of basis function increases with N^n, where n is the dimension of state. Hence, it is important to keep N small. The filtering results for the same data set with the transformation are shown in Figure 2. As shown in the figure, they do not have any negative region by the definition and the quality of estimation is not distinguishable in the results for $N = 16$ and $N = 64$. Hence, the number of basis functions could be kept lower, for example, as small as $N = 5$ in this case (simulation result is not shown).

V. Conclusions

A novel transformation is presented to obtain the negative-free approximation of the probability density function. The

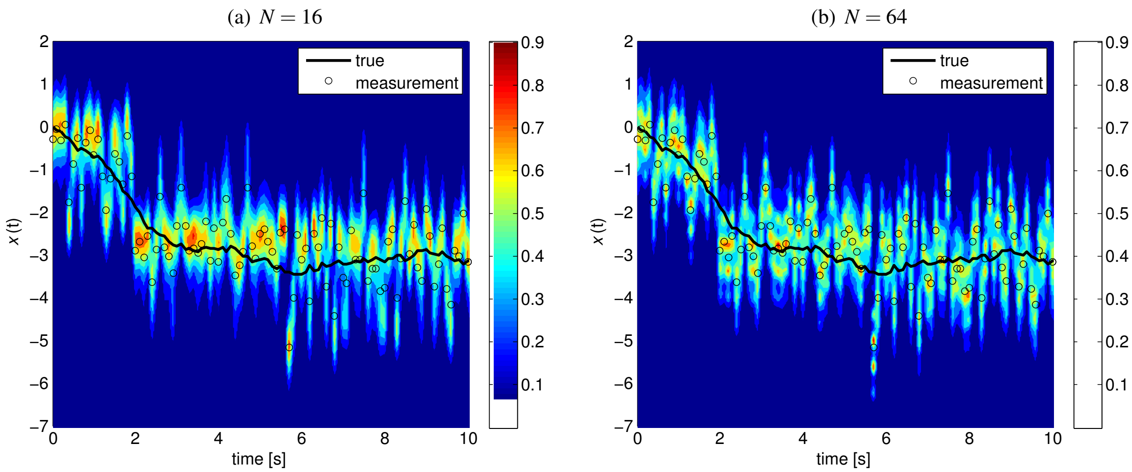

Fig. 2. With the transformation: the pdf does not have negative values by the definition. The heat-map indicates the time-history of the pdf over the sampling space. The true state is shown by the black solid line and the measurements are depicted by the hollow circles.

approximation is used to derive the nonlinear projection filter, where the joint probability density function is estimated. It is shown that some analytic expressions can be derived with the set of cosine basis functions and this could be applicable for various types of nonlinear functions. In addition, the required number of basis functions could be reduced significantly and it would enable to apply the filter to higher-dimensional systems. The effectiveness of the proposed method is demonstrated by some numerical simulations.

ACKNOWLEDGEMENT

This research is supported by EPSRC Research Grant, EP/N010523/1, Balancing the impact of city infrastructure engineering on natural systems using robots.

REFERENCES

[1] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME Journal of Basic Engineering*, no. 82 (Series D), pp. 35–45, 1960.
[2] M. S. Grewal and A. P. Andrews, "Applications of Kalman Filtering in Aerospace 1960 to the Present [Historical Perspectives]," *Control Systems, IEEE*, vol. 30, pp. 69–78, June 2010.
[3] E. J. Lefferts, F. L. Markley, and M. D. Shuster, "Kalman filtering for spacecraft attitude estimation," *Journal of Guidance, Control, and Dynamics*, vol. 5, pp. 417–429, September-October 1982.
[4] S. Rogers, "Alpha-beta filter with correlated measurement noise," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. AES-23, pp. 592–594, July 1987.
[5] J. Gray and W. Murray, "A derivation of an analytic expression for the tracking index for the alpha-beta-gamma filter," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 29, pp. 1064–1065, July 1993.
[6] C. Antoniou, M. Ben-Akiva, and H. N. Koutsopoulos, "Nonlinear kalman filtering algorithms for on-line calibration of dynamic traffic assignment models," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, pp. 661–670, Dec 2007.
[7] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, pp. 401–422, Mar. 2004.
[8] E. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," in *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, pp. 153–158, IEEE, Aug. 2000.
[9] J. L. Crassidis and F. L. Markley, "Unscented filtering for spacecraft attitude estimation," *Journal of Guidance, Control, and Dynamics*, vol. 26, pp. 536–542, Jul 2003.
[10] C. Liu, P. Shui, G. Wei, and S. Li, "Modified unscented kalman filter using modified filter gain and variance scale factor for highly maneuvering target tracking," *Journal of Systems Engineering and Electronics*, vol. 25, pp. 380–385, June 2014.
[11] M. Partovibakhsh and G. Liu, "An adaptive unscented kalman filtering approach for online estimation of model parameters and state-of-charge of lithium-ion batteries for autonomous mobile robots," *IEEE Transactions on Control Systems Technology*, vol. 23, pp. 357–363, Jan 2015.
[12] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry, Third Edition (North-Holland Personal Library)*. North Holland, 3 ed., May 2007.
[13] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *Signal Processing, IEEE Transactions on*, vol. 50, pp. 174–188, Feb. 2002.
[14] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," vol. 10, no. 3, pp. 197–208, 2000.
[15] R. Beard, J. Kenney, J. Gunther, J. Lawton, and W. Stirling, "Nonlinear Projection Filter Based on Galerkin Approximation," *Journal of Guidance, Control, and Dynamics*, vol. 22, pp. 258–266, Mar. 1999.
[16] Y. Zhai, M. Yeary, and D. Zhou, "Target tracking using a particle filter based on the projection method," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 3, pp. III–1189–III–1192, April 2007.
[17] G. Qian, K. Shafique, and P. Wang, "Fusion of nonlinear motion dynamics using fokker-planck equation and projection filter," in *Information Fusion (FUSION), 2014 17th International Conference on*, pp. 1–7, July 2014.
[18] J. Kim, "Nonlinear projection filter for target tracking using range sensor & optical tracker," in *International Federation of Automatic Control, 20th IFAC Symposium on Automatic Control in Aerospace - ACA 2016, submitted in February 2016*.
[19] B. Azimi-Sadjadi and P. Krishnaprasad, "Approximate nonlinear filtering and its application in navigation," *Automatica*, vol. 41, no. 6, pp. 945 – 956, 2005.
[20] M. Kumar and S. Chakravorty, "A nonlinear filter based on fokker planck equation and mcmc measurement updates," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, pp. 7357–7362, Dec 2010.
[21] T. Single-Liertz, J. Kim, and R. Richardson, "Nonlinear projection filter with parallel algorithm and parallel sensors," in *Decision and Control (CDC), 2015 IEEE 54th Annual Conference on*, pp. 2432–2437, Dec 2015.
[22] C. Pan, "Gibbs phenomenon removal and digital filtering directly through the fast fourier transform," *IEEE Transactions on Signal Processing*, vol. 49, pp. 444–448, Feb 2001.
[23] A. R. Barron and C.-H. Sheu, "Approximation of density functions by sequences of exponential families," *Ann. Statist.*, vol. 19, pp. 1347–1369, 09 1991.