FOURIER–HERMITE RAUCH–TUNG–STRIEBEL SMOOTHER

Juha Sarmavuori

Nokia Siemens Networks
Karakaarenkuja 1
FI-02610 Espoo, Finland

Simo Särkkä

Aalto University
Rakentajanaukio 2
FI-02150 Espoo, Finland

ABSTRACT

In this article, we introduce the Fourier–Hermite Rauch–Tung–Striebel smoother which is based on expansion of nonlinear functions in a Fourier–Hermite series in same way as the traditional extended Rauch–Tung–Striebel smoother is based on the Taylor series. The first order truncation of the Fourier–Hermite series gives the previously known statistically linearized smoother.

Index Terms— nonlinear smoothing, nonlinear filtering, statistical linearization, Fourier–Hermite series

1. INTRODUCTION

In this article, we introduce a new Fourier–Hermite series based method for approximate optimal smoothing of non-linear state space models. The stochastic state space models (see, e.g., [1, 2]) are assumed to be of the form:

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{q}_{k-1} \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k) + \mathbf{r}_k, \end{aligned} \tag{1}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the state and $\mathbf{y}_k \in \mathbb{R}^d$ is the measurement of the model at the time step k. The terms $\mathbf{q}_{k-1} \sim \mathrm{N}(\mathbf{0}, \mathbf{Q}_{k-1})$ and $\mathbf{r}_k \sim \mathrm{N}(\mathbf{0}, \mathbf{R}_k)$ are the Gaussian process and measurement noises, respectively. Finally, $\mathbf{f}(\cdot)$ and $\mathbf{h}(\cdot)$ are the non-linear dynamic and measurement model functions.

Smoothing solutions are used to improve filtering solutions in a wide range of applications, such as guidance systems, integrated inertial navigation and passive sensor based target tracking [1, 2, 3, 4]. Mathematically, the smoothing solution consists of the sequence of conditional probability distributions $p(\mathbf{x}_k \,|\, \mathbf{y}_{1:T}), k = 1, \ldots, T$ of the states given all the measurements up to the step $T > k$. Computing the exact probability distributions is intractable for nonlinear functions $\mathbf{f}$ and $\mathbf{g}$. Here we follow the Gaussian smoothing approach [5], where we approximate the distributions with Gaussian distributions:

$$p(\mathbf{x}_k \,|\, \mathbf{y}_{1:T}) \approx \mathrm{N}(\mathbf{x}_k \,|\, \mathbf{m}_{k|T}, \mathbf{P}_{k|T}), \tag{2}$$

where the mean vectors $\mathbf{m}_{k|T}$ and the covariance matrices $\mathbf{P}_{k|T}$ of the approximating Gaussian distributions are computed by the smoothing algorithm.

For linear $\mathbf{f}$ and $\mathbf{g}$, the solution is exactly Gaussian and has been found already in the 1960s. The solution consists of the Kalman filter, which computes filtering distributions $p(\mathbf{x}_k \,|\, \mathbf{y}_{1:k}) = \mathrm{N}(\mathbf{x}_k \,|\, \mathbf{m}_{k|k}, \mathbf{P}_{k|k})$ of the state $\mathbf{x}_k$ given all measurements up to the step k, for $k = 1, \ldots, T$ [6, 1, 7], and the Rauch–Tung–Striebel smoother [8, 9, 10], which computes the smoothing distributions $p(\mathbf{x}_k \,|\, \mathbf{y}_{1:T}) = \mathrm{N}(\mathbf{x}_k \,|\, \mathbf{m}_{k|T}, \mathbf{P}_{k|T})$.

The filtering and smoothing solutions to non-linear state space model can be approximated with Gaussian filters and smoothers [11, 12, 5]. In that approach, we first compute Gaussian approximations to the filtering distributions with a non-linear Kalman filter (= Gaussian filter) and then Gaussian approximations to the smoothing distributions with a non-linear Rauch–Tung–Striebel smoother (= Gaussian smoother) [5].

Both the Gaussian filter and Gaussian smoother require computation of the following three kinds of expectations [11, 12, 5, 13]:

$$\hat{\mathbf{g}} = \mathrm{E}[\mathbf{g}(\mathbf{x})] \tag{3}$$

$$\mathrm{Cov}[\mathbf{g}(\mathbf{x})] = \mathrm{E}[(\mathbf{g}(\mathbf{x}) - \hat{\mathbf{g}})\,(\mathbf{g}(\mathbf{x}) - \hat{\mathbf{g}})^T] \tag{4}$$

$$\mathrm{Cov}[\mathbf{x}, \mathbf{g}(\mathbf{x})] = \mathrm{E}[(\mathbf{x} - \mathrm{E}[\mathbf{x}])\,(\mathbf{g}(\mathbf{x}) - \hat{\mathbf{g}})^T], \tag{5}$$

over Gaussian distributions. Computation of the exact values of the Gaussian expectations is intractable for most nonlinear functions. Therefore, several methods for approximating the Gaussian expectations have been developed. A Taylor series based approximation is used in the extended Kalman filter (EKF) [1]. The smoother equivalent is the extended Rauch–Tung–Striebel smoother (ERTS) [14, 15, 10]. During the last decade, various sigma-point and numerical integration based Gaussian approximations have been developed for the filtering problem: unscented Kalman filter (UKF) [16, 17], Gauss-Hermite Kalman filter (GHKF) [11], central differences Kalman filter (CDKF) [11], cubature Kalman filter (CKF) [18] and various others [19, 20, 12]. The same approximations have also been applied to non-linear smoothing [21, 22, 23, 24].

One old nonlinear filtering method, which has not gained that much attention since its discovery at 1970's, is the statistical linearization based statistically linearized filter (SLF)

[25]. The smoother equivalent has been presented in [23]. Recently, a Fourier–Hermite series based filter, Fourier–Hermite Kalman filter (FHKF) was proposed [13] and it turned out that SLF is a special case of the truncated Fourier–Hermite series based filters [13]. In Gaussian case, Taylor series can also be seen as an approximation to the Fourier–Hermite series [13]. Interestingly, several sigma-point methods can be considered as numerical approximations to the statistical linearization [20, 26] and thus it is likely that new higher order sigma-point type of methods can be developed as approximations to the Fourier–Hermite series based methods.

In this article, we describe the application of Fourier–Hermite series to non-linear smoothing and demonstrate its performance with a numerical example.

1.1. Gaussian Smoother

In Gaussian smoothing, we assume that Gaussian approximations to filtering solution $p(\mathbf{x}_k \mid \mathbf{y}_{1:k}) \approx \mathrm{N}(\mathbf{x}_k \mid \mathbf{m}_{k|k}, \mathbf{P}_{k|k})$ have already been computed. We use the following shorthand notation for Gaussian expectations:

$$\mathrm{E}[\mathbf{g}(\mathbf{x}_k) \mid \mathbf{m}_{k|k}, \mathbf{P}_{k|k}] = \int_{\mathbb{R}^n} \mathbf{g}(\mathbf{x}_k)\ \mathrm{N}(\mathbf{x}_k \mid \mathbf{m}_{k|k}, \mathbf{P}_{k|k})\, \mathrm{d}\mathbf{x}_k$$

The Gaussian smoother solution is then given by [5]:

$$\begin{aligned}
\mathbf{m}_{k+1|k} &= \mathrm{E}[\mathbf{f}(\mathbf{x}_k) \mid \mathbf{m}_{k|k}, \mathbf{P}_{k|k}] \\
\mathbf{P}_{k+1|k} &= \mathrm{E}[(\mathbf{f}(\mathbf{x}_k) - \mathbf{m}_{k+1|k}) \\
&\qquad \times (\mathbf{f}(\mathbf{x}_k) - \mathbf{m}_{k+1|k})^T \mid \mathbf{m}_{k|k}, \mathbf{P}_{k|k}] + \mathbf{Q}_k \\
\mathbf{C}_{k,k+1} &= \mathrm{E}[(\mathbf{x}_k - \mathbf{m}_{k|k})(\mathbf{f}(\mathbf{x}_k) - \mathbf{m}_{k+1|k})^T \mid \mathbf{m}_{k|k}, \mathbf{P}_{k|k}] \\
\mathbf{G}_k &= \mathbf{C}_{k,k+1}\mathbf{P}_{k+1|k}^{-1} \\
\mathbf{m}_{k|T} &= \mathbf{m}_{k|k} + \mathbf{G}_k\,(\mathbf{m}_{k+1|T} - \mathbf{m}_{k+1|k}) \\
\mathbf{P}_{k|T} &= \mathbf{P}_{k|k} + \mathbf{G}_k\,(\mathbf{P}_{k+1|T} - \mathbf{P}_{k+1|k})\,\mathbf{G}_k^T,
\end{aligned} \tag{6}$$

where we start from the filtering result on step $k = T$ and iterate the above equations back to $k = 0$. Note that all the expectations are of the form (3) - (5). The resulting approximations to the smoothing distributions are then Gaussian $p(\mathbf{x}_k|\mathbf{y}_{1:T}) \approx \mathrm{N}(\mathbf{x}_k \mid \mathbf{m}_{k|T}, \mathbf{P}_{k|T})$.

1.2. Linearization Based Smoothing

The classical approach to Gaussian filtering and smoothing is linearization. It is very natural way of extending the linear KF and RTS for nonlinear problems, because both KF and RTS are well defined even if the linear function is different at each time step. A linearized filter or smoother acts almost like a linear filter or smoother which is using different linear function at each time step. It is also easier to compute the Gaussian expectations for the linearized functions. Linearization by zeroth and first terms of the Taylor series leads to EKF [1] and ERTS [14, 15, 10].

1.3. Statistical Linearization

In statistical linearization [25] of nonlinear function $\mathbf{f}(\mathbf{x})$, we seek vector $\mathbf{b}$ and matrix $\mathbf{A}$ so that they minimize the mean square error:

$$\begin{aligned}
\mathrm{MSE}(\mathbf{b}, \mathbf{A}) = \mathrm{E}[&(\mathbf{f}(\mathbf{x}) - (\mathbf{b} + \mathbf{A}\,(\mathbf{x} - \mathbf{m})))^T \\
&\times (\mathbf{f}(\mathbf{x}) - (\mathbf{b} + \mathbf{A}\,(\mathbf{x} - \mathbf{m})))],
\end{aligned} \tag{7}$$

where the expectation is taken over the distribution of $\mathbf{x}$. The result is:

$$\begin{aligned}
\mathbf{b} &= \mathrm{E}[\mathbf{f}(\mathbf{x})] \\
\mathbf{A} &= \mathrm{E}[\mathbf{f}(\mathbf{x})(\mathbf{x} - \mathbf{m})^T]\,\mathbf{P}^{-1}.
\end{aligned} \tag{8}$$

It would be possible to use some other error criterion than the square error, but with this choice the result is simple to compute and also quite natural for our Gaussian view.

1.4. Fourier–Hermite series

The idea of linearization can be extended to higher order polynomial approximations. In principle, we could extend statistical linearization to arbitrary order statistical polynomial approximations. With arbitrary probability distribution that would be cumbersome. For example, see general second order approximation of a scalar function in [25].

For Gaussian distribution, we can simplify the computations by reformulating the problem as an orthogonal projection of the function in a Hilbert space. We define the inner product of functions $f(\mathbf{x})$ and $g(\mathbf{x})$ as [13]:

$$\langle f, g \rangle = \mathrm{E}[f(\mathbf{x})\, g(\mathbf{x}) \mid \mathbf{m}, \mathbf{P}]. \tag{9}$$

Functions for which $\langle f, f \rangle < \infty$ form a complete inner product space or a Hilbert space in other words.

The optimal polynomial approximation is easy to form with an orthogonal basis of the Hilbert space, which is given by the Hermite polynomials [27]:

$$H_{[i_1,\ldots,i_k]}(\mathbf{x}) = (-1)^n \exp\left(\frac{1}{2}\mathbf{x}^T\mathbf{x}\right) \frac{\partial^k \exp(-\frac{1}{2}\mathbf{x}^T\mathbf{x})}{\partial x_{i_1} \ldots \partial x_{i_k}}.$$

The Hermite polynomials are orthogonal with respect to the inner product (9) with $\mathbf{m} = 0$ and $\mathbf{P} = \mathbf{I}$:

$$\langle H_{[i_1,\ldots,i_k]}, H_{[j_1,\ldots,j_m]} \rangle = \begin{cases} k! & , \quad k = m,\ i_p = j_p, \\ & \quad \text{for } p = 1, \ldots, k \\ 0 & , \quad \text{otherwise} \end{cases}$$

For arbitrary $\mathbf{m}$ and $\mathbf{P}$ we have to adjust the polynomials a little to make them orthogonal:

$$H_{[i_1,\ldots,i_k]}(\mathbf{x}; \mathbf{m}, \mathbf{P}) = H_{[i_1,\ldots,i_k]}(\sqrt{\mathbf{P}}^{-1}(\mathbf{x} - \mathbf{m})), \tag{10}$$

where $\sqrt{\mathbf{P}}$ is symmetric non-negative definite square root matrix. We could use Cholesky factorization $\mathbf{L}\,\mathbf{L}^T = \mathbf{P}$ as in

[13], but due to commutativity, symmetric square root matrices are easier to use in the derivation of the equations.

With the orthogonal basis, we can expand functions in a Fourier–Hermite series as follows [13]:

$$\begin{aligned}\mathbf{f}(\mathbf{x}) &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{i_1,\ldots,i_k=1}^{n} \mathrm{E}[H_{[i_1,\ldots,i_k]}(\mathbf{x};\mathbf{m},\mathbf{P})\,\mathbf{f}(\mathbf{x}) \mid \mathbf{m},\mathbf{P}] \\ &\quad \times H_{[i_1,\ldots,i_k]}(\mathbf{x};\mathbf{m},\mathbf{P}) \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{\substack{i_1,\ldots,i_k=1 \\ j_1,\ldots,j_k=1}}^{n} \mathrm{E}\left[\frac{\partial^k \mathbf{f}(\mathbf{x})}{\partial x_{i_1} \ldots \partial x_{i_k}} \;\middle|\; \mathbf{m},\mathbf{P} \right] \\ &\quad \times \left(\prod_{r=1}^{k} \left[\sqrt{\mathbf{P}}\right]_{i_r,j_r} \right) H_{[j_1,\ldots,j_k]}(\mathbf{x};\mathbf{m},\mathbf{P}) \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{\substack{i_1,\ldots,i_k=1 \\ j_1,\ldots,j_k=1}}^{n} \frac{\partial^k \mathrm{E}[\mathbf{f}(\mathbf{x}) \mid \mathbf{m},\mathbf{P}]}{\partial m_{i_1} \ldots \partial m_{i_k}} \\ &\quad \times \left(\prod_{r=1}^{k} \left[\sqrt{\mathbf{P}}\right]_{i_r,j_r} \right) H_{[j_1,\ldots,j_k]}(\mathbf{x};\mathbf{m},\mathbf{P}). \end{aligned} \tag{11}$$

The later two forms can be derived by repeated integration by parts [13]. As $\mathbf{P} \to \mathbf{0}$, the FH-series becomes the Taylor series. The uth order truncation of this series gives the MSE-optimal uth order polynomial approximation to the function $\mathbf{f}(\mathbf{x})$.

With orthogonal basis, we can compute the covariance by Parseval relation [13]:

$$\begin{aligned}\mathrm{Cov}[\mathbf{f}(\mathbf{x})] &= \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{i_1,\ldots,i_k=1}^{n} \\ &\quad \mathrm{E}[H_{[i_1,\ldots,i_k]}(\mathbf{x};\mathbf{m},\mathbf{P})\,\mathbf{f}(\mathbf{x}) \mid \mathbf{m},\mathbf{P}] \\ &\quad \times \mathrm{E}[H_{[i_1,\ldots,i_k]}(\mathbf{x};\mathbf{m},\mathbf{P})\,\mathbf{f}(\mathbf{x}) \mid \mathbf{m},\mathbf{P}]^T \\ &= \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{\substack{i_1,\ldots,i_k=1 \\ j_1,\ldots,j_k=1}}^{n} \mathrm{E}\left[\frac{\partial^k \mathbf{f}(\mathbf{x})}{\partial x_{i_1} \ldots \partial x_{i_k}} \;\middle|\; \mathbf{m},\mathbf{P} \right] \\ &\quad \times \left(\prod_{r=1}^{k} P_{i_r,j_r} \right) \mathrm{E}\left[\frac{\partial^k \mathbf{f}(\mathbf{x})}{\partial x_{j_1} \ldots \partial x_{j_k}} \;\middle|\; \mathbf{m},\mathbf{P} \right]^T \\ &= \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{\substack{i_1,\ldots,i_k=1 \\ j_1,\ldots,j_k=1}}^{n} \frac{\partial^k \mathrm{E}[\mathbf{f}(\mathbf{x}) \mid \mathbf{m},\mathbf{P}]}{\partial m_{i_1} \ldots \partial m_{i_k}} \\ &\quad \times \left(\prod_{r=1}^{k} P_{i_r,j_r} \right) \frac{\partial^k \mathrm{E}[\mathbf{f}(\mathbf{x}) \mid \mathbf{m},\mathbf{P}]^T}{\partial m_{j_1} \ldots \partial m_{j_k}}. \end{aligned} \tag{12}$$

The covariance of the corresponding uth order approximation is given by the uth order truncation of this series.

2. FOURIER-HERMITE RTS SMOOTHER

With Fourier–Hermite series, we can compute the expectations of (6) as

$$\begin{aligned}\mathbf{m}_{k+1|k} &= \mathrm{E}[\mathbf{f}(\mathbf{x}_k) \mid \mathbf{m}_{k|k},\mathbf{P}_{k|k}] \\ \mathbf{P}_{k+1|k} &= \mathbf{Q}_k + \sum_{s=1}^{u} \frac{1}{s!} \sum_{i_1,\ldots,i_s=1}^{n} \\ &\quad \mathrm{E}[H_{[i_1,\ldots,i_s]}(\mathbf{x}_k;\mathbf{m}_{k|k},\mathbf{P}_{k|k})\,\mathbf{f}(\mathbf{x}_k) \mid \mathbf{m}_{k|k},\mathbf{P}_{k|k}] \\ &\quad \times \mathrm{E}[H_{[i_1,\ldots,i_s]}(\mathbf{x}_k;\mathbf{m}_{k|k},\mathbf{P}_{k|k})\,\mathbf{f}(\mathbf{x}_k) \mid \mathbf{m}_{k|k},\mathbf{P}_{k|k}]^T \\ &= \mathbf{Q}_k + \sum_{s=1}^{u} \frac{1}{s!} \sum_{\substack{i_1,\ldots,i_s=1 \\ j_1,\ldots,j_s=1}}^{n} \\ &\quad \mathrm{E}\left[\frac{\partial^s \mathbf{f}(\mathbf{x})}{\partial x_{i_1} \ldots \partial x_{i_s}} \;\middle|\; \mathbf{m}_{k|k},\mathbf{P}_{k|k} \right] \\ &\quad \times \left(\prod_{r=1}^{s} \left[\mathbf{P}_{k|k}\right]_{i_r,j_r} \right) \\ &\quad \times \mathrm{E}\left[\frac{\partial^s \mathbf{f}(\mathbf{x})}{\partial x_{j_1} \ldots \partial x_{j_s}} \;\middle|\; \mathbf{m}_{k|k},\mathbf{P}_{k|k} \right]^T \\ &= \mathbf{Q}_k + \sum_{s=1}^{u} \frac{1}{s!} \sum_{\substack{i_1,\ldots,i_s=1 \\ j_1,\ldots,j_s=1}}^{n} \\ &\quad \left. \frac{\partial^s \mathrm{E}[\mathbf{f}(\mathbf{x}_k) \mid \mathbf{m},\mathbf{P}_{k|k}]}{\partial m_{i_1} \ldots \partial m_{i_s}} \right|_{\mathbf{m}=\mathbf{m}_{k|k}} \\ &\quad \times \left(\prod_{r=1}^{k} \left[\mathbf{P}_{k|k}\right]_{i_r,j_r} \right) \\ &\quad \times \left. \frac{\partial^s \mathrm{E}[\mathbf{f}(\mathbf{x}_k) \mid \mathbf{m},\mathbf{P}_{k|k}]^T}{\partial m_{j_1} \ldots \partial m_{j_s}} \right|_{\mathbf{m}=\mathbf{m}_{k|k}} \end{aligned} \tag{13}$$

$$\begin{aligned}\mathbf{C}_{k,k+1} &= \mathrm{E}[(\mathbf{x}_k - \mathbf{m}_{k|k})\,(\mathbf{f}(\mathbf{x}_k) - \mathbf{m}_{k+1|k})^T \mid \mathbf{m}_{k|k},\mathbf{P}_{k|k}] \\ &= \mathbf{P}_{k|k}\, \mathrm{E}\left[\left[\frac{\partial \mathbf{f}(\mathbf{x})}{\partial x_1} \;\cdots\; \frac{\partial \mathbf{f}(\mathbf{x})}{\partial x_n} \right]^T \;\middle|\; \mathbf{m}_{k|k},\mathbf{P}_{k|k} \right] \\ &= \mathbf{P}_{k|k} \left(\begin{bmatrix} \frac{\partial}{\partial m_1} \\ \vdots \\ \frac{\partial}{\partial m_n} \end{bmatrix} \mathrm{E}[\mathbf{f}(\mathbf{x}_k) \mid \mathbf{m},\mathbf{P}_{k|k}]^T \right) \Bigg|_{\mathbf{m}=\mathbf{m}_{k|k}} \end{aligned}$$

First order truncation $u = 1$ gives the statistically linearized RTS smoother SLRTS. If we also approximate the expectations with $\mathbf{P}_{k|k} = \mathbf{0}$, we get the Taylor series based ERTS.

3. NUMERICAL EXAMPLE

To illustrate the practical applicability of the FHRTS smoother, we applied the method to the simulated pendulum model used in [13]. The dynamic and measurement models can be written

as

$$\begin{aligned} x_{k,1} &= x_{k-1,1} + x_{k-1,2}\,\Delta t \\ x_{k,2} &= x_{k-1,2} - g\,\sin(x_{k-1,1})\,\Delta t + q_{k-1} \\ y_k &= h(x_{k,1}) + r_k, \end{aligned} \tag{14}$$

where $q_{k-1} \sim \mathrm{N}(0, Q)$, $r_k \sim \mathrm{N}(0, R)$, and

$$h(x) = \begin{cases} -1 & , \text{ if } x < -a/2 + b \\ 0 & , \text{ if } -a/2 + b < x < a/2 + b \\ 1 & , \text{ if } x > a/2 + b. \end{cases} \tag{15}$$

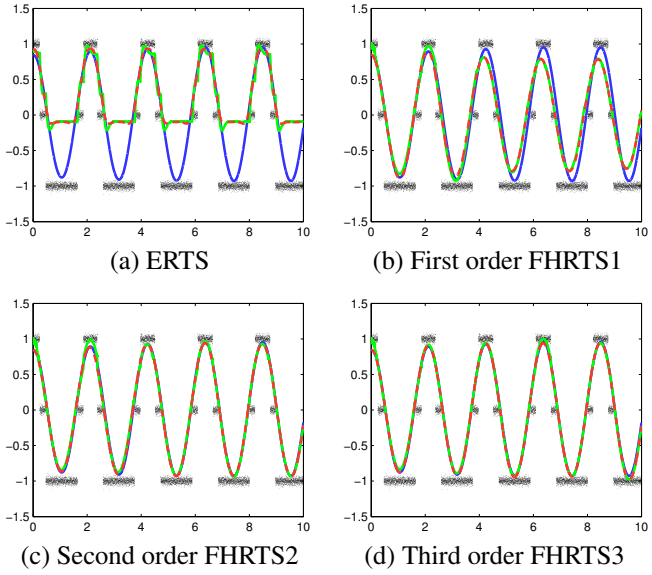
The parameters of the model were the same as in [13]. For this model the FHRTS equations can be computed in closed form.

The result of the EKF/ERTS estimate is quite far away from the correct one and smoothing does not improve the estimate much, most likely because the covariance estimate of the filtering solution is quite inaccurate. A better result is obtained with the statistical linearization (SLRTS/FHRTS1), that is, the first order FHRTS smoother, as shown in Figure 1(b). Because here the linearization takes into account the whole distribution of x, the asymmetry is better accounted and the result is much more accurate. The results of the second order, (FHRTS2) in 1(c), and the third order, (FHRTS3) in 1(d), smoothers are even more accurate. The RMSE values for each of the filters and smoothers are given in the caption of the Figure 1. The smoothing improves the estimates especially with second and third order approximations. The second order smoother is even better than the third order filter.

We implemented unscented transform, spherical cubature and Gauss-Hermite quadrature based smoothers [22, 5, 24] (URTS, CRTS and GHRTS, respectively) to the same problem. The error of CRTS and URTS with equivalent parameters ($\alpha = 1$, $\beta = 0$, $\kappa = 0$) was RMSE = 0.040, which is similar to second order FHRTS2. With URTS different set of parameters ($\alpha = 1$, $\beta = 0$, $\kappa = 1$) result in better RMSE = 0.023 that is similar to third order GHRTS (RMSE = 0.020) or to third order FHRTS3.

The theoretical advantage of the FH series based approach is that a uth order truncation of the FH series is exact for polynomials up to order u. Thus the uth order FHRTS smoother computes the Gaussian expectations exactly for functions $\mathbf{f}$ that are polynomials up to order u. In comparison, URTS and CRTS smoothers compute the Gaussian expectations (in particular, the covariance) exactly only for linear $\mathbf{f}$ [12]. Also the ERTS and SLRTS smoothers compute the Gaussian expectations exactly only for linear functions, that is, polynomials up to order $u = 1$. GHRTS is another method that can compute expectations correctly for an arbitrary order polynomial. GHRTS smoother of order p computes the Gaussian expectations exactly for functions $\mathbf{f}$ that are of order $p - 1$ [11, 12].

The disadvantage of GHRTS is that its computational complexity is exponential in the number of state dimensions.



(a) ERTS (b) First order FHRTS1

(c) Second order FHRTS2 (d) Third order FHRTS3

Fig. 1. The results of (a) EKF (RMSE = 0.413) and ERTS (RMSE = 0.406), (b) first order FHKF1 (RMSE = 0.164) and FHRTS1 (RMSE = 0.160), (c) second order FHKF2 (RMSE = 0.041) and FHRTS2 (RMSE = 0.025), and (d) third order FHKF3 (RMSE = 0.031) and FHRTS3 (RMSE = 0.018). The measurements are marked with small dots, the solid blue line is the real signal, solid green the filter estimate and dashed red the smoother estimate.

The computational complexity depends on order p and dimension n as $O(p^n)$ [11, 12]. From Equation (13) we can see that the computational complexity of uth order FHRTS is $O(n^u)$, that is, polynomial on the state dimension.

4. CONCLUSION AND DISCUSSION

In this article, we have introduced a new Fourier–Hermite series based method for approximate optimal smoothing of nonlinear state space models. We have also analyzed the connection of the new smoother to the SLRTS and ERTS smoothers. The new smoother was also compared in a numerical simulation with the ERTS and the sigma-point based URTS, CRTS and GHRTS smoothers.

5. REFERENCES

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

[2] Y. Bar-Shalom, X.R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, Wiley, New York, 2001.

[3] M.S. Grewal, L.R. Weill, and A.P. Andrews, *Global Po-*

sitioning Systems, Inertial Navigation and Integration, Wiley, New York, 2001.

[4] S. Challa, M.R. Morelande, D. Mušicki, and R.J. Evans, *Fundamentals of Object Tracking*, Cambridge University Press, 2011.

[5] S. Särkkä and J. Hartikainen, "On Gaussian optimal smoothing of non-linear state space models," *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1938–1941, August 2010.

[6] R.E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME, Journal of Basic Engineering*, vol. 82, pp. 35–45, March 1960.

[7] M.S. Grewal and A.P. Andrews, *Kalman Filtering, Theory and Practice Using MATLAB*, Wiley, New York, 2001.

[8] H.E. Rauch, F. Tung, and C.T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA Journal*, vol. 3(8), pp. 1445–1450, 1965.

[9] P. Maybeck, *Stochastic Models, Estimation and Control, Volume 2*, Academic Press, 1982.

[10] J.L. Crassidis and J.L. Junkins, *Optimal Estimation of Dynamic Systems*, Chapman & Hall/CRC, 2004.

[11] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Transactions on Automatic Control*, vol. 45(5), pp. 910–927, 2000.

[12] Y. Wu, D. Hu, M. Wu, and X. Hu, "A numerical-integration perspective on Gaussian filters," *IEEE Transactions on Signal Processing*, vol. 54(8), pp. 2910–2921, 2006.

[13] J. Sarmavuori and S. Särkkä, "Fourier-Hermite Kalman filter," *IEEE Transactions on Automatic Control*, vol. 57(6), pp. 1511–1515, 2012.

[14] H. Cox, "On the estimation of state variables and parameters for noisy dynamic systems," *IEEE Transactions on Automatic Control*, vol. 9(1), pp. 5–12, January 1964.

[15] C.T. Leondes, J.B. Peller, and E.B. Stear, "Nonlinear smoothing theory," *IEEE Transactions on Systems Science and Cybernetics*, vol. 6(1), January 1970.

[16] S.J. Julier, J.K. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *Proceedings of the 1995 American Control, Conference, Seattle, Washington*, 1995, pp. 1628–1632.

[17] S.J. Julier, J.K. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45(3), pp. 477–482, March 2000.

[18] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Transactions on Automatic Control*, vol. 54(6), pp. 1254–1269, 2009.

[19] M. Nørgaard, N.K. Poulsen, and O. Ravn, "New developments in state estimation for nonlinear systems," *Automatica*, vol. 36(11), pp. 1627 – 1638, 2000.

[20] R. Van der Merwe and E. Wan, "Sigma-point Kalman filters for probabilistic inference in dynamic state-space models," in *Proceedings of the Workshop on Advances in Machine Learning*, 2003.

[21] E. Wan and R. Van der Merwe, "The unscented Kalman filter," in *Kalman Filtering and Neural Networks*, Simon Haykin, Ed., chapter 7. Wiley, 2001.

[22] S. Särkkä, "Unscented Rauch-Tung-Striebel smoother," *IEEE Transactions on Automatic Control*, vol. 53(3), pp. 845–849, 2008.

[23] S. Särkkä and J. Hartikainen, "Sigma point methods in optimal smoothing of non-linear stochastic state space models," in *Proceedings of IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2010, pp. 184–189.

[24] I. Arasaratnam and S. Haykin, "Cubature Kalman smoothers," *Automatica*, vol. 47, no. 10, pp. 2245–2250, 2011.

[25] A. Gelb, *Applied Optimal Estimation*, The MIT Press, Cambridge, MA, 1974.

[26] T. Lefebvre, H. Bruyninckx, and J. De Schutter, "Comment on "a new method for the nonlinear transformation of means and covariances in filters and estimators" [and authors' reply]," *IEEE Transactions on Automatic Control*, vol. 47(8), pp. 1406–1409, 2002.

[27] P. I. Kuznetsov, R. L. Stratonovich, and V. I. Tikhonov, "Quasi-moment functions in the theory of random processes," *Theory of Probability and its Applications*, vol. V, no. 1, pp. 80–97, 1960.