Fourier-Hermite Kalman Filter

Juha Sarmavuori and Simo Särkkä, *Member, IEEE*

Abstract—In this note, we shall present a new class of Gaussian filters called Fourier-Hermite Kalman filters. Fourier-Hermite Kalman filters are based on expansion of nonlinear functions with the Fourier-Hermite series in same way as the traditional extended Kalman filter is based on the Taylor series. The first order truncation of the Fourier-Hermite series gives the previously known statistically linearized filter.

Index Terms—nonlinear Kalman filtering, extended Kalman filtering, statistical linearization, Fourier-Hermite series

I. INTRODUCTION

The *Kalman filter* (KF) [1] is concerned with estimation of the dynamic state from noisy measurements in the class of estimation problems where the dynamic and measurement processes can be approximated by linear Gaussian state space models. The KF is also applicable to linear state space models with a wide range of non-Gaussian noise distributions [2]. General filtering theory for non-linear and non-Gaussian models was already presented in [3], [4], but in practice, numerical solutions derived as approximations to the general theory are usually computationally more demanding than the Gaussian approximations derived as extensions to the KF. The Taylor series based *extended Kalman filter* (EKF) [4], and the Gaussian describing function based *statistically linearized filter* (SLF) [5] are the classical Gaussian approximation based extensions of the KF to nonlinear dynamic and measurement models.

Recently, numerical integration based sigma point filters [6]–[10], have been introduced as alternatives to the classical linearization based methods. Many of the sigma point methods can also be interpreted as numerical approximations to the SLF [11]–[13]. In this note, we shall take the opposite approach from the sigma point methods – instead of approximating the SLF we shall develop higher order approximations by extending SLF. In numerical comparison, the new approach is found to give similar results as the sigma point methods. The advantage of the new method compared to the sigma point methods is that it provides a closed form approximation instead of applying a numerical method directly. The implementation of the closed form solution can be more efficient. If closed form solution is not possible for some part of the problem then it is still possible to use the numerical sigma point approach for that part.

In this paper, we shall introduce a new class of filters that we call Fourier-Hermite Kalman filters (FHKF). The filters are based on a finite truncation of the Fourier-Hermite series in a similar way as the EKF is based on a truncation of the

Manuscript received October 8, 2011; revised October 8, 2011.

Juha Sarmavuori* is with Nokia Siemens Networks, Karakaarenkuja 1. FIN-02610 ESPOO, Finland. E-mail: juha.sarmavuori@nsn.com, Tel: +358 50 486 5075, Fax: +358 71 802 4210.

Simo Särkkä is with Aalto University, P.O. Box 12200. FI-00076 AALTO, Finland. E-mail: simo.sarkka@tkk.fi, Tel: +358 50 512 4393, Fax: +358 9 470 24833.

Taylor series. The first order truncation gives the previously known SLF [5] in a similar way as the first order truncation of the Taylor series gives the basic EKF. The new approach also makes it possible to use higher order truncation of the Fourier-Hermite series similar to the second order EKF. Due to the orthogonality of the Hermite polynomials, any order truncation is almost as easy to use as the first order truncation and gives the best possible polynomial approximation in the mean squared error sense. With the Taylor series the higher order truncations are much more difficult to develop than the first or the second order truncations, and a truncation of the Taylor series is not the best possible polynomial approximation under any simple criterion. Fourier-Hermite series expansions can also be derived for non-differentiable functions whereas Taylor series expansions only exist for differentiable functions. We feel that introduction of orthogonal basis functions for Gaussian filtering framework can lead to many new developments.

Fourier-Hermite expansions of non-linear functions and the related Wiener chaos expansions of stochastic functionals have previously been used for approximating the formal continuous-time filtering equations, for example, in [14]–[16]. The Edgeworth series and related approximations are based on Fourier-Hermite expansions of probability densities, and they have also been applied in optimal filtering context [17]–[19]. In [14]–[19] the Fourier-Hermite expansion is applied to probability density. In this article, the probability density is always approximated as Gaussian and the Fourier-Hermite expansion is applied to the nonlinear functions.

A. Gaussian Filter

In this article, we consider discrete-time state space models of the following form:

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

where $\mathbf{x}_k \in \mathbb{R}^n$ is the state of the model at a time step k, $\mathbf{y}_k \in \mathbb{R}^d$ is the measurement. $\mathbf{q}_k \sim \mathrm{N}(\mathbf{0}, \mathbf{Q}_k)$ and $\mathbf{r}_k \sim \mathrm{N}(\mathbf{0}, \mathbf{R}_k)$ are Gaussian process and measurement noises, respectively. The Gaussian filter [7], [20] for the above model is the following:

- *Prediction*:

$$\begin{aligned} \mathbf{m}_k^- &= \mathrm{E}[\mathbf{f}(\mathbf{x}_{k-1})] \\ \mathbf{P}_k^- &= \mathrm{E}[(\mathbf{f}(\mathbf{x}_{k-1}) - \mathbf{m}_k^-)(\mathbf{f}(\mathbf{x}_{k-1}) - \mathbf{m}_k^-)^T] + \mathbf{Q}_k, \quad (2) \end{aligned}$$

where the expectations $\mathrm{E}[\cdot]$ are with respect to $\mathbf{x}_{k-1} \sim \mathrm{N}(\mathbf{m}_{k-1}, \mathbf{P}_{k-1})$.

- *Update*:

$$\begin{aligned} \boldsymbol{\mu}_k &= \mathrm{E}[\mathbf{h}(\mathbf{x}_k)] \\ \mathbf{S}_k &= \mathrm{E}[(\mathbf{h}(\mathbf{x}_k) - \boldsymbol{\mu}_k)(\mathbf{h}(\mathbf{x}_k) - \boldsymbol{\mu}_k)^T] + \mathbf{R}_k \\ \mathbf{K}_k &= \mathrm{E}[(\mathbf{x}_k - \mathbf{m}_k^-)(\mathbf{h}(\mathbf{x}_k) - \boldsymbol{\mu}_k)^T] \mathbf{S}_k^{-1} \\ \mathbf{m}_k &= \mathbf{m}_k^- + \mathbf{K}_k (\mathbf{y}_k - \boldsymbol{\mu}_k) \\ \mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^T, \quad (3) \end{aligned}$$

where the expectations $\mathrm{E}[\cdot]$ are with respect to $\mathbf{x}_k \sim \mathrm{N}(\mathbf{m}_k^-, \mathbf{P}_k^-)$.

II. MAIN RESULTS

A. Computation of Gaussian Expectations via Fourier-Hermite Series

The Gaussian filter (see Section I-A) requires computation of three kinds of expectations:

$$\hat{\mathbf{g}} = \mathrm{E}[\mathbf{g}(\mathbf{x})] \tag{4}$$

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

$$\mathrm{Cov}[\mathbf{x}, \mathbf{g}(\mathbf{x})] = \mathrm{E}[(\mathbf{x} - \mathrm{E}[\mathbf{x}])(\mathbf{g}(\mathbf{x}) - \hat{\mathbf{g}})^T], \tag{6}$$

where $\mathbf{g}$ is $\mathbf{f}$ on the prediction step (2) or $\mathbf{h}$ on the update step (3) and $\mathbf{x} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. As discussed in the introduction, in the EKF, these expectations are approximated with Taylor series of the function and in the sigma-point filters with numerical integration. As the following theorem shows, the expectations can also be easily computed from the Fourier-Hermite series coefficients of the function $\mathbf{g}(\mathbf{x})$ (see Appendix):

Theorem 1: Assume that the Fourier-Hermite series representation (37) for the function $\mathbf{g}(\mathbf{x})$ exists. The expectation of the function is the zeroth coefficient in (38) or (39):

$$\mathrm{E}[\mathbf{g}(\mathbf{x})] = \mathbf{c}^0 \tag{7}$$

and the covariance of the function is

$$\mathrm{Cov}[\mathbf{g}(\mathbf{x})] = \sum_{k=1}^{\infty} \frac{1}{k!} \boldsymbol{\Gamma}_k, \tag{8}$$

where the matrices $\boldsymbol{\Gamma}_k \in \mathbb{R}^{n \times n}$ are given by (42) or (43). The cross-covariance between the random variable and the function can be expressed as

$$\mathrm{Cov}[\mathbf{x}, \mathbf{g}(\mathbf{x})] = \mathbf{L}\,\mathbf{C}^T = \boldsymbol{\Sigma}\,\mathbf{A}^T, \tag{9}$$

where we have arranged the vectors of (38) and (40) into matrix forms $\mathbf{C} = [\mathbf{c}_1^1 \; \mathbf{c}_2^1 \; \ldots \; \mathbf{c}_n^1]$ and $\mathbf{A} = [\mathbf{a}_1^1 \; \mathbf{a}_2^1 \; \ldots \; \mathbf{a}_n^1]$.

Proof: Equation (7) is trivial from (38) with $k = 0$. Equation (8) follows from the Parseval relation (41). Equation (9) is trivial from (38) and (39) with $k = 1$. ∎

Unfortunately, the direct use of the formulas (38) and (40) for the Fourier-Hermite series coefficient vectors would require computation of complicated expectation integrals. However, by using the idea originally presented in [5, Eq. (6.4-12)], the required number of the expectation integrals can be reduced to only one:

Theorem 2: Assume that we can compute the following integral in closed form:

$$\hat{\mathbf{g}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \int_{\mathbb{R}^n} \mathbf{g}(\mathbf{x})\, \mathrm{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})\, \mathrm{d}\mathbf{x}. \tag{10}$$

Then the derivative expectation terms in (40) can be equivalently expressed as:

$$\mathbf{a}_{j_1, \ldots, j_k}^k = \frac{\partial^k}{\partial \mu_{j_1} \ldots \partial \mu_{j_k}} \hat{\mathbf{g}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}). \tag{11}$$

Proof: $k = 0$ is trivial. For $k = 1$, we can assume that for most functions of practical interest, we can change the order

of the integration and the differentiation:

$$\mathbf{c}_i^1 = \int_{\mathbb{R}^n} \mathbf{g}(\mathbf{x})\, \left[\mathbf{H}_1(\mathbf{L}^{-1}(\mathbf{x} - \boldsymbol{\mu}))\right]_i \mathrm{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})\, \mathrm{d}\mathbf{x}$$

$$= \int_{\mathbb{R}^n} \mathbf{g}(\mathbf{x}) \sum_{j=1}^{n} L_{j,i} \frac{\partial}{\partial \mu_j}\, \mathrm{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})\, \mathrm{d}\mathbf{x}$$

$$= \sum_{j=1}^{n} L_{j,i} \frac{\partial}{\partial \mu_j} \int_{\mathbb{R}^n} \mathbf{g}(\mathbf{x})\, \mathrm{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})\, \mathrm{d}\mathbf{x}. \tag{12}$$

Further indices $k = 2, \ldots$ follow similarly. ∎

Thus with the Theorems 1 and 2 we can construct Fourier-Hermite series based approximations to the Gaussian expectations in terms of the integral (10) and its derivatives. Although, the closed form integration of (10) is not always possible, for many common functions it is possible to use the random input describing functions tabulated in [21]. Finding closed form formulas for the expectations in terms of elementary functions or special functions is also nowadays easier than before due to availability of powerful computer algebra systems (CAS). Usually, the integral is the hardest part, but the derivatives are an easy task for a CAS.

B. Fourier-Hermite Kalman Filter

We present the Fourier-Hermite Kalman filter up to third order using the method described in section II-A. It is easy to deduce how to compute fourth and higher order terms.

- *Compute mean functions*:

$$\hat{\mathbf{f}}(\mathbf{m}, \mathbf{P}) = \int_{\mathbb{R}^n} \mathbf{f}(\mathbf{x})\, \mathrm{N}(\mathbf{x}|\mathbf{m}, \mathbf{P})\, \mathrm{d}\mathbf{x}$$

$$\hat{\mathbf{h}}(\mathbf{m}, \mathbf{P}) = \int_{\mathbb{R}^n} \mathbf{h}(\mathbf{x})\, \mathrm{N}(\mathbf{x}|\mathbf{m}, \mathbf{P})\, \mathrm{d}\mathbf{x}. \tag{13}$$

- *Compute Jacobians*:

$$[\hat{\mathbf{F}}_{\mathbf{m}}(\mathbf{m}, \mathbf{P})]_{i,i'} = \frac{\partial}{\partial m_{i'}} \hat{f}_i(\mathbf{m}, \mathbf{P})$$

$$[\hat{\mathbf{H}}_{\mathbf{m}}(\mathbf{m}, \mathbf{P})]_{i,i'} = \frac{\partial}{\partial m_{i'}} \hat{h}_i(\mathbf{m}, \mathbf{P}). \tag{14}$$

- *Compute higher derivatives*:

$$\hat{\mathbf{f}}_{i,j}''(\mathbf{m}, \mathbf{P}) = \frac{\partial^2}{\partial m_i \partial m_j} \hat{\mathbf{f}}(\mathbf{m}, \mathbf{P})$$

$$\hat{\mathbf{f}}_{i,j,u}'''(\mathbf{m}, \mathbf{P}) = \frac{\partial^3}{\partial m_i \partial m_j \partial m_u} \hat{\mathbf{f}}(\mathbf{m}, \mathbf{P})$$

$$\vdots \tag{15}$$

$$\hat{\mathbf{h}}_{i,j}''(\mathbf{m}, \mathbf{P}) = \frac{\partial^2}{\partial m_i \partial m_j} \hat{\mathbf{h}}(\mathbf{m}, \mathbf{P})$$

$$\hat{\mathbf{h}}_{i,j,u}'''(\mathbf{m}, \mathbf{P}) = \frac{\partial^3}{\partial m_i \partial m_j \partial m_u} \hat{\mathbf{h}}(\mathbf{m}, \mathbf{P})$$

$$\vdots \tag{16}$$

With these approximations, the prediction and update steps of the Gaussian filter can be written as follows:

- *Prediction*:

$$\mathbf{m}_k^- = \hat{\mathbf{f}}(\mathbf{m}_{k-1}, \mathbf{P}_{k-1})$$

$$\mathbf{P}_k^- = \mathbf{Q}_k + \hat{\mathbf{F}}_{\mathbf{m}} \mathbf{P}_{k-1} \hat{\mathbf{F}}_{\mathbf{m}}^T + \frac{1}{2!} \sum_{\substack{i,j \\ u,v}} \hat{\mathbf{f}}_{i,u}'' P_{i,j} P_{u,v} \hat{\mathbf{f}}_{j,v}''^T$$

$$+ \frac{1}{3!} \sum_{\substack{i,j \\ u,v \\ a,b}} \hat{\mathbf{f}}_{i,u,a}''' P_{i,j} P_{u,v} P_{a,b} \hat{\mathbf{f}}_{j,v,b}'''^T + \dots, \qquad (17)$$

where we used shorthand notation $P_{i,j} = [\mathbf{P}_{k-1}]_{i,j}$ and the derivatives are evaluated at $\mathbf{m}_{k-1}$ and $\mathbf{P}_{k-1}$.

- *Update*:

$$\mathbf{S}_k = \mathbf{R}_k + \hat{\mathbf{H}}_{\mathbf{m}} \mathbf{P}_k^- \hat{\mathbf{H}}_{\mathbf{m}}^T + \frac{1}{2!} \sum_{\substack{i,j \\ u,v}} \hat{\mathbf{h}}_{i,u}'' P_{i,j} P_{u,v} \hat{\mathbf{h}}_{j,v}''^T$$

$$+ \frac{1}{3!} \sum_{\substack{i,j \\ u,v \\ a,b}} \hat{\mathbf{h}}_{i,u,a}''' P_{i,j} P_{u,v} P_{a,b} \hat{\mathbf{h}}_{j,v,b}'''^T + \dots \qquad (18)$$

$$\mathbf{K}_k = \mathbf{P}_k^- \hat{\mathbf{H}}_{\mathbf{m}}^T \mathbf{S}_k^{-1}$$

$$\mathbf{m}_k = \mathbf{m}_k^- + \mathbf{K}_k (\mathbf{y}_k - \hat{\mathbf{h}}(\mathbf{m}_k^-, \mathbf{P}_k^-))$$

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^T, \qquad (19)$$

where $P_{i,j} = [\mathbf{P}_k^-]_{i,j}$ and the derivatives are evaluated at $\mathbf{m}_k^-$ and $\mathbf{P}_k^-$.

III. DISCUSSION

The first order truncation of the Fourier-Hermite series gives the statistically linearized filter [5]. The first order truncation of the series (37) using the coefficients of (38) leads to the same filter form as in [5], which is also correct linearization for any non-Gaussian distribution. In the Gaussian case, the statistical linearization based approximations for (4), (5) and (6) can be written with the first order truncation of (37) using the coefficients of (40) and (39) as:

$$\hat{\mathbf{g}} = \mathrm{E}[\mathbf{g}(\mathbf{x})] \qquad (20)$$

$$\mathrm{Cov}[\mathbf{g}(\mathbf{x})] \approx \mathrm{E}[\mathbf{G}_{\mathbf{x}}(\mathbf{x})] \, \boldsymbol{\Sigma} \, \mathrm{E}[\mathbf{G}_{\mathbf{x}}(\mathbf{x})]^T \qquad (21)$$

$$\mathrm{Cov}[\mathbf{x}, \mathbf{g}(\mathbf{x})] = \boldsymbol{\Sigma} \, \mathrm{E}[\mathbf{G}_{\mathbf{x}}(\mathbf{x})]^T. \qquad (22)$$

where the expectations $\mathrm{E}[\cdot]$ are with respect to $\mathbf{x} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The corresponding approximations for EKF are very similar:

$$\hat{\mathbf{g}} \approx \mathbf{g}(\mathrm{E}[\mathbf{x}]) \qquad (23)$$

$$\mathrm{Cov}[\mathbf{g}(\mathbf{x})] \approx \mathbf{G}_{\mathbf{x}}(\mathrm{E}[\mathbf{x}]) \, \boldsymbol{\Sigma} \, \mathbf{G}_{\mathbf{x}}(\mathrm{E}[\mathbf{x}])^T \qquad (24)$$

$$\mathrm{Cov}[\mathbf{x}, \mathbf{g}(\mathbf{x})] \approx \boldsymbol{\Sigma} \, \mathbf{G}_{\mathbf{x}}(\mathrm{E}[\mathbf{x}])^T. \qquad (25)$$

The order of the expectation and the nonlinear functions are the opposite in the EKF and the SLF. Of course, with a linear function the order is exchangeable. Also, the SLF recursion becomes the EKF, if $\boldsymbol{\Sigma}$ is small and we use approximation $\boldsymbol{\Sigma} \to 0$ in the expectations.

The second order approximation for the covariance can also be written in a form closer to the approximation used in the second order EKF [22]:

$$\mathrm{Cov}[\mathbf{g}(\mathbf{x})] \approx \mathrm{E}[\mathbf{G}_{\mathbf{x}}(\mathbf{x})] \, \boldsymbol{\Sigma} \, \mathrm{E}[\mathbf{G}_{\mathbf{x}}(\mathbf{x})]^T$$

$$+ \frac{1}{2} \sum_{i,j=1}^n \mathbf{e}_i \mathbf{e}_j^T \, \mathrm{tr} \left[\mathrm{E}[\mathbf{G}_{\mathbf{xx}}^i(\mathbf{x})] \, \boldsymbol{\Sigma} \, \mathrm{E}[\mathbf{G}_{\mathbf{xx}}^j(\mathbf{x})] \, \boldsymbol{\Sigma} \right], \qquad (26)$$

where $\mathbf{e}_i$ is a coordinate axis column vector with one at index i and zeros elsewhere.

Due to the orthogonality of the Hermite polynomials, the Fourier-Hermite Kalman filter is relatively easy to compute for any given order. The Taylor series used in EKF does not have the same property. We demonstrate this with one dimensional expansion of the variance:

$$\mathrm{Var}[g(x)] = (g'(\mu))^2 \sigma^2 + \left(\frac{1}{2}(g''(\mu))^2 + g'(\mu)g'''(\mu) \right) \sigma^4$$

$$+ \dots \qquad (27)$$

The second order EKF [22] omits the $g'(\mu)g'''(\mu)$ term. It is possible to develop higher order Taylor series based EKF, but the expansion for variance will have more and more terms of the form $g^{(k)}(\mu)g^{(n)}(\mu)\sigma^{k+n}$. The Fourier-Hermite series gives only terms of the form $\mathrm{E}\left[g^{(k)}(x)\right]^2 \sigma^{2k}$.

For a monomial of order n, the nth order FHKF gives exact results for all the Gaussian expectations (4), (5) and (6). The same is true for an nth order Taylor series based EKF. The unscented transform [6] and the spherical cubature integration [9] give exact results only for first order monomials, but the unscented transform can be made exact for second order monomials with a suitable selection of parameters [8]. The Gauss-Hermite quadrature rule of order n gives exact results for $(n-1)$th order monomial [7]. More advanced inequalities for the accuracy of the FHKF could be derived from the results presented in [23], [24].

IV. NUMERICAL EXAMPLE

To illustrate the practical applicability of the proposed filters, we consider the following simulated pendulum model, whose discretized dynamic model can be written as

$$\begin{aligned} x_{k,1} &= x_{k-1,1} + x_{k-1,2} \, \Delta t \\ x_{k,2} &= x_{k-1,2} - g \, \sin(x_{k-1,1}) \, \Delta t + q_{k-1}, \end{aligned} \qquad (28)$$

where $\Delta t = 1/1000$ is the sampling period, $g = 9.81$ and $q_{k-1} \sim N(0, Q)$, with $Q = \Delta t/100$. The measurement model is the following:

$$y_k = h(x_{k,1}) + r_k, \qquad r_k \sim N(0, R), \qquad (29)$$

where $R = 1/1000$ and h is the piecewise constant function

$$h(x) = \begin{cases} -1 & , \quad \text{if } x < -a/2 + b \\ 0 & , \quad \text{if } -a/2 + b < x < a/2 + b \\ 1 & , \quad \text{if } x > a/2 + b, \end{cases} \qquad (30)$$

where the constants have the values $a = 0.5, b = 0.4$. In the simulation it was assumed that the initial conditions were known with a small error of variance 0.01 in both of the state
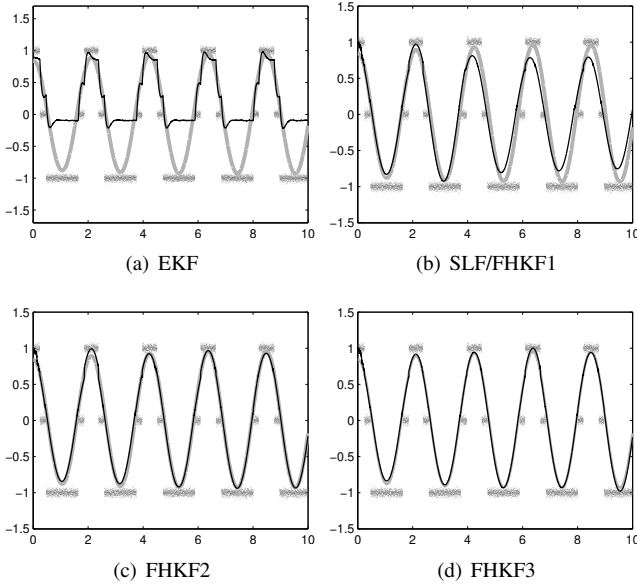
Fig. 1. The results of (a) EKF (RMSE = 0.41), (b) first order FHKF (RMSE = 0.16), (c) second order FHKF (RMSE = 0.04), and (d) third order FHKF (RMSE = 0.03). The measurements are marked with small dots, the thicker gray line is the real signal and thinner black line is the estimate.

components. Note that implementing an EKF to the model is quite hard due to the piecewise constant measurement model function. Probably, the biggest practical advantage of the SLF or FHKF is that the functions do not have to be differentiable. The expectation of the function (10) is differentiable although the function is not. A sensible way to implement an EKF (or linearized filter) is to replace the measurement model with $h(x) \approx (x - b)/a$. The result of tracking with this simple linearized model is shown in Figure 1(a). As can be seen, the result is quite far away from the correct one, because the approximation is unable to take the asymmetry in the measurements into account.

Much better result can be obtained with the statistically linearized filter (SLF/FHKF1), that is, the first order Fourier-Hermite Kalman filter, as shown in Figure 1(b). Because here the linearization takes into account the distribution of x, the asymmetry is also better accounted and the result is much more accurate. The results of the second order Fourier-Hermite Kalman filter (FHKF2) in 1(c), and the third order Fourier-Hermite Kalman filter (FHKF3) in 1(d) are even more accurate. The RMSE values for each of the filters are given in the caption of the Figure 1.

For comparison, the same simulation was repeated also with unscented Kalman filter (UKF) [6], [10], [13], cubature Kalman filter (CKF) [8], [9] and Gauss-Hermite Kalman filter (GHKF) [7]. In terms of RMSE, the GHKF gives similar results as the equal order FHKF. UKF and CKF give similar results as the second order FHKF or GHKF.

V. CONCLUSION

In this note we have introduced a novel Fourier-Hermite Kalman filter for nonlinear filtering and developed a practical method for computation. We have also analyzed, how the new filter solution relates to the older filters EKF and SLF. The

new filter was also compared in a numerical simulation with the EKF and the sigma-point methods: UKF, CKF and GHKF. The performance of the FHKF and the sigma-point methods were found equal in the simulation.

APPENDIX
FOURIER-HERMITE SERIES

We use the following definition of the Hermite polynomials [25]:

$$H_n(x) = (-1)^n \, e^{x^2/2} \, \frac{\mathrm{d}^n}{\mathrm{d}x^n} e^{-x^2/2}, \qquad n = 0, 1, \ldots, \quad (31)$$

which are orthogonal with respect to the inner product defined by the expectation with respect to the standard normal distribution $\mathrm{N}(0,1)$ as $\langle f, g \rangle = \mathrm{E}[f(x)g(x)]$. If $g(x)$ is a function such that $\mathrm{E}[g(x)^2] < \infty$, then it has the Fourier-Hermite series [25]:

$$g(x) = \sum_{k=0}^{\infty} \frac{1}{k!} \mathrm{E}[g(x)H_k(x)]H_k(x) \quad (32)$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \mathrm{E}[g^{(k)}(x)]H_k(x), \quad (33)$$

where the latter representation also requires that the kth derivatives of the function satisfy $E[g^{(k)}(x)^2] < \infty$. When the standard normal distribution $\mathrm{N}(0,1)$ is replaced with an arbitrary normal distribution $\mathrm{N}(\mu, \sigma^2)$, the representations (32) and (33) become:

$$g(x) = \sum_{k=0}^{\infty} \frac{1}{k!} \mathrm{E}\left[g(x)H_k\left(\frac{x - \mu}{\sigma} \right) \right] H_k\left(\frac{x - \mu}{\sigma} \right) \quad (34)$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \mathrm{E}[g^{(k)}(x)]H_k\left(\frac{x - \mu}{\sigma} \right) \sigma^k. \quad (35)$$

The Fourier-Hermite series (34), (35) can be generalized to higher dimensions [25]. The multidimensional Hermite polynomials are multi-index objects, whose elements are products of one dimensional Hermite polynomials. Although, the usual multi-index notation used in [16], [25], [26] is more elegant, it is difficult to manipulate non unity covariances $\boldsymbol{\Sigma} \neq \mathbf{I}$ with it. Therefore we use notation similar to [26, Section 8]. For an m order Hermite polynomial of n dimensional vector $\mathbf{x}$ we need m indices $i_1, i_2, \ldots, i_m \in \{1, 2, \ldots, m\}$. Indices are divided into n sets $(J_s = \{k | i_k = s\})_{s=1}^{n}$ of, which some may be empty. The size of the sets is $q_s = |J_s| \in \{0, \ldots, m\}$ so that $\sum_{s=1}^{n} q_s = m$. The elements of the Hermite polynomial are then:

$$[\mathbf{H}_m(\mathbf{x})]_{i_1, \ldots, i_m} = \prod_{s=1}^{n} H_{q_s}(x_s). \quad (36)$$

For example, let $n = 3$, $m = 4$, $i_1 = 3, i_2 = 2, i_3 = 2, i_4 = 1$, then $J_1 = \{4\}, J_2 = \{2, 3\}, J_3 = \{1\}, q_1 = 1, q_2 = 2, q_3 = 1$ and $[\mathbf{H}_4(\mathbf{x})]_{3,2,2,1} = H_1(x_1)H_2(x_2)H_1(x_3)$. For small m we can use other notations: $\mathbf{H}_0(\mathbf{x}) = 1$ is a scalar, $\mathbf{H}_1(\mathbf{x}) = \mathbf{x}$ is a vector and $\mathbf{H}_2(\mathbf{x}) = \mathbf{x}\mathbf{x}^T - \mathbf{I}$ is a matrix.

In a multidimensional case the standard deviation σ is replaced with the Cholesky decomposition $\mathbf{L}$ of the covariance matrix $\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^T$. One dimensional $\frac{x - \mu}{\sigma}$ is replaced with

$\mathbf{L}^{-1}(\mathbf{x} - \boldsymbol{\mu})$. The multidimensional Fourier-Hermite series can be written as:

$$\mathbf{g}(\mathbf{x}) = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{i_1,\dots,i_k=1}^{n} \mathbf{c}_{i_1,\dots,i_k}^{k} \left[\mathbf{H}_k \left(\mathbf{L}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right) \right]_{i_1,\dots,i_k}, \tag{37}$$

where the coefficients corresponding to the one-dimensional series (34) are given as:

$$\mathbf{c}_{i_1,\dots,i_k}^{k} = \mathrm{E}\left[\mathbf{g}(\mathbf{x}) \left[\mathbf{H}_k \left(\mathbf{L}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right) \right]_{i_1,\dots,i_k} \right]. \tag{38}$$

For sufficiently differentiable function, the multidimensional analog of (35) is:

$$\mathbf{c}_{i_1,\dots,i_k}^{k} = \sum_{j_1,\dots,j_k=1}^{n} \mathbf{a}_{j_1,\dots,j_k}^{k} \prod_{m=1}^{k} L_{j_m,i_m}, \tag{39}$$

where

$$\mathbf{a}_{j_1,\dots,j_k}^{k} = \mathrm{E}\left[\frac{\partial^k}{\partial x_{j_1} \dots \partial x_{j_k}} \mathbf{g}(\mathbf{x}) \right]. \tag{40}$$

Due to the orthogonality, the coefficient vectors satisfy the Parseval's relation:

$$\mathrm{E}[\mathbf{g}(\mathbf{x})\,\mathbf{g}^T(\mathbf{x})] = \sum_{k=0}^{\infty} \frac{1}{k!} \boldsymbol{\Gamma}_k, \tag{41}$$

where the matrices $\boldsymbol{\Gamma}_k$ are given by:

$$\boldsymbol{\Gamma}_k = \sum_{i_1,\dots,i_k=1}^{n} \mathbf{c}_{i_1,\dots,i_k}^{k} \left(\mathbf{c}_{i_1,\dots,i_k}^{k} \right)^T \tag{42}$$

$$= \sum_{\substack{i_1,\dots,i_k=1 \\ j_1,\dots,j_k=1}}^{n} \mathbf{a}_{i_1,\dots,i_k}^{k} \left(\mathbf{a}_{j_1,\dots,j_k}^{k} \right)^T \prod_{m=1}^{k} \Sigma_{i_m,j_m}. \tag{43}$$

REFERENCES

[1] 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.

[2] J. L. Maryak, J. C. Spall, and B. D. Heydon, "Use of the Kalman filter for inference in state-space models with unknown noise distribution," *IEEE Transactions on Automatic Control*, vol. 49(1), pp. 87–90, Jan. 2004.

[3] R. L. Stratonovich, *Conditional Markov Processes and Their Application to the Theory of Optimal Control*. American Elsevier Publishing Company, Inc., 1968.

[4] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. Academic Press, 1970.

[5] A. Gelb, Ed., *Applied Optimal Estimation*. The MIT Press, 1974.

[6] 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.

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

[8] 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.

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

[10] R. van der Merwe, "Sigma-point Kalman filters for probabilistic inference in dynamic state-space models," Ph.D. dissertation, OGI School of Science & Engineering, Oregon Health & Science University, Portland, OR, USA, April 2004.

[11] 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'," *IEEE Transactions on Automatic Control*, vol. 47(8), pp. 1406–1409, August 2002.

[12] T. Lefebvre, H. Bruyninckx, and J. De Schutter, "Kalman filters for nonlinear systems: a comparison of performance," *International Journal of Control*, vol. 77, 7, pp. 639–653, 2004.

[13] 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.

[14] S. Lototsky, R. Mikulevicius, and B. Rozovskii, "Nonlinear filtering revisited: A spectral approach," *SIAM Journal of Control and Optimization*, vol. 35, pp. 435–461, 1997.

[15] R. Mikulevicius and B. L. Rozovskii, "Fourier–Hermite expansions for nonlinear filtering," *Theory of Probability and its Applications*, vol. 44, pp. 606–612, 2000.

[16] S. V. Lototsky, "Chaos approach to nonlinear filtering," in *The Oxford Handbook of Nonlinear Filtering*, D. Crisan and B. Rozovskii, Eds. Oxford University Press, 2011, pp. 231–264.

[17] H. Singer, "Generalized Gauss-Hermite filtering," *AStA Advances in Statistical Analysis*, vol. 92, no. 2, pp. 179–195, 2008.

[18] ——, "Moment Equations and Hermite Expansion for Nonlinear Stochastic Differential Equations with Application to Stock Price Models," *Computational Statistics*, vol. 21,3, pp. 385–397, 2006.

[19] S. Challa, Y. Bar-Shalom, and V. Krishnamurthy, "Nonlinear filtering via generalized Edgeworth series and Gauss-Hermite quadrature," *IEEE Transactions on Signal Processing*, vol. 48(6), June 2000.

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

[21] A. Gelb and W. Vander Velde, *Multiple-Input Describing Functions and Nonlinear System Design*. McGraw Hill, 1968.

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

[23] W. Luo, "Wiener chaos expansion and numerical solutions of stochastic partial differential equations," Ph.D. dissertation, California Institute of Technology, Pasadena, California, May 2006.

[24] J. P. Boyd, *Chebyshev and Fourier Spectral Methods*. DOVER Publications, Inc., 2000.

[25] P. Malliavin, *Stochastic Analysis*. Springer, 1997.

[26] 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, 1, pp. 80–97, 1960.