

Gaussian Filtering for Polynomial Systems Based on Moment Homotopy

Marco F. Huber
AGT International
Darmstadt, Germany
Email: marco.huber@ieee.org

Uwe D. Hanebeck
Intelligent Sensor-Actuator-Systems Laboratory (ISAS)
Institute for Anthropomatics
Karlsruhe Institute of Technology (KIT), Germany
Email: uwe.hanebeck@ieee.org

Abstract—This paper proposes Gaussian filters for polynomial systems with efficient solutions for both the prediction and the filter step. For the prediction step, computationally efficient closed-form solutions are derived for calculating the exact moments. In order to achieve a higher estimation quality, the filter step is solved without the usual additional assumption that state and measurement are jointly Gaussian distributed. As this significantly complicates the required moment calculation, a homotopy continuation method is employed that yields almost optimal results.

I. INTRODUCTION

Closed-form recursive Bayesian state estimation can only be performed for a few special systems, such as linear continuous systems or systems with finite state and measurement spaces. The famous Kalman filter is the best linear estimator, being optimal for the linear Gaussian case. For finite state and measurement spaces, grid-based filters are optimal [1]. For arbitrary nonlinearities, however, that are typical in real-world applications such as target tracking, financial forecasting, medical surveillance, or robotics, recursive Bayesian state estimation requires approximate solutions.

A practical approximation known as Gaussian assumed density filtering restricts the state estimate to be Gaussian distributed [2]. Preserving the mean and variance of the state via moment matching requires the solution of expectation integrals. Since closed-form expectation calculation is not possible in general and numerical integration is computationally infeasible, many fast approximate Gaussian assumed density filters utilizing the Kalman filter equations have been developed in the past. Employing first-order Taylor-series expansion to linearize the given system and measurement models leads to the extended Kalman filter [3]. This fast filter is only applicable to mild nonlinearities and requires differentiability. To cover also stronger nonlinearities, so-called linear regression Kalman filters have become popular in the recent years, where the approximation relies on deterministic sampling. Part of this group of Gaussian filters are the unscented Kalman filter [4], the cubature Kalman filter [5], or the Gaussian estimator [6].

Compared to these generic Gaussian filters, improved estimation performance can be achieved by focusing on a particular type of nonlinearity. In this paper, the focus is on polynomial nonlinearities. First results on Bayesian estimation for polynomial dynamics but linear measurement models can be found in [7]. The more general case is treated in [8], where exact moment calculation for the prediction step is derived based on Taylor-series expansion.

In Section IV of this paper, special properties of exponential densities with polynomial exponents are exploited to efficiently calculate the moments after a polynomial transformation. Compared to [8], this leads to a simplified and computationally cheaper calculation of moments after a prediction step. A straightforward application of these insights to the measurement step requires the assumption of a jointly Gaussian distributed state and measurement, which is a typical assumption in Gaussian filtering. For polynomial nonlinearities, however, the posterior density is an exponential density and thus, it is a conjugate density to the prior Gaussian density. Unfortunately, exponential densities allow no closed-form calculation of the moments in general, which is necessary for Gaussian filtering. To overcome this limitation, a homotopy continuation approach for calculating the posterior moments is proposed in Section V. The continuation starts with the known moments of the Gaussian prior, while the likelihood, which depends on the polynomial nonlinearity, is gradually introduced into the measurement update step. This causes a continuous transformation of the prior moments towards the posterior moments. The transformation can be expressed via a system of first-order ordinary differential equations, for which a plethora of efficient numerical solvers exists.

The proposed Gaussian filters for polynomial nonlinearities are compared to the state-of-the-art by means of numerical simulations in Section VI. The paper closes with a conclusion and an outlook to future work.

II. PROBLEM FORMULATION

In this paper, nonlinear discrete-time system and measurement equations

$$\mathbf{x}_{k+1} = a_k(\mathbf{x}_k) + \mathbf{w}_k, \quad (1)$$

$$\mathbf{z}_k = h_k(\mathbf{x}_k) + \mathbf{v}_k, \quad (2)$$

are considered, where \mathbf{x}_k is the scalar system state at time step $k = 0, 1, \dots$ and \mathbf{z}_k is the scalar measurement. An actual measurement value \hat{z}_k is a realization of the random variable \mathbf{z}_k . Both \mathbf{w}_k and \mathbf{v}_k are white zero-mean Gaussian noise processes with variance $(\sigma_k^w)^2$ and $(\sigma_k^v)^2$, respectively.

In Bayesian estimation, two alternating steps, i.e., prediction and measurement update, are performed for estimating the system state \mathbf{x}_k . The latest estimate of \mathbf{x}_{k-1} is propagated to time step k by means of the system equation (1) in the prediction step. In the measurement update, a given measurement

value \hat{z}_k is exploited for updating \mathbf{x}_k under consideration of the measurement equation (2).

Exact closed-form solutions for the prediction and the measurement update are not available for arbitrary $a_k(\cdot)$, $h_k(\cdot)$ and arbitrarily distributed random variables. This paper is restricted to polynomial system and measurement functions $a_k(\cdot)$ and $h_k(\cdot)$, respectively. It is further assumed, that the system state \mathbf{x}_k can be represented by means of Gaussian distributions for all time steps k . Thus, it is sufficient to investigate polynomial transformations of the form

$$\mathbf{y} = g(\mathbf{x}) + \mathbf{w} = \sum_{i=0}^n c_i \cdot \mathbf{x}^i + \mathbf{w}, \quad (4)$$

where a Gaussian $\mathbf{x} \sim \mathcal{N}(x; \mu_x, \sigma_x^2)$ with mean μ_x and variance σ_x^2 is mapped to a random variable \mathbf{y} . In case of a prediction, \mathbf{y} corresponds to \mathbf{x}_{k+1} , while for a measurement update, \mathbf{y} is the measurement \mathbf{z}_k . The transformation is affected by zero-mean Gaussian noise $\mathbf{w} \sim \mathcal{N}(w; 0, \sigma_w^2)$, which is assumed to be uncorrelated with \mathbf{x} .

The goal now is two-fold: (i) calculating the mean and variance of \mathbf{y} for prediction purposes and (ii) incorporation of a realization $\hat{\mathbf{y}}$ of \mathbf{y} to perform a measurement update.

III. EXPONENTIAL DENSITIES

At first, a brief introduction to the so-called family of *exponential densities with polynomial exponents* is provided. Their properties regarding recursive moment calculation play a significant role for solving the problem at hand.

A. Definition

An unnormalized one-dimensional exponential density is defined as

$$f(x) = \exp\left(\sum_{i=0}^n \eta_i \cdot x^i\right) = \exp(\underline{\eta}^T \cdot \underline{x}),$$

with parameter vector $\underline{\eta}^T \triangleq [\eta_0, \eta_1, \dots, \eta_n]$ and vector of monomials $\underline{x}^T \triangleq [1, x, x^2, \dots, x^n]$. To ensure that the exponential density is non-negative for all $x \in \mathbb{R}$ and has finite moments, the maximum degree $n \in \mathbb{N}$ must be even and the highest-order coefficient η_n must be negative, i.e., $\eta_n < 0$. If desired, the exponential density can be normalized by adding a term $\log(c(\underline{\eta}))$ to the first coefficient η_0 , where $c(\underline{\eta})$ is a normalization constant.

An important special case of the family of exponential densities is the (unnormalized) Gaussian density

$$f(x) = \exp\left(-\frac{1}{2} \left(\frac{x - \mu_x}{\sigma_x}\right)^2\right) = \exp(\eta_0 + \eta_1 \cdot x + \eta_2 \cdot x^2),$$

with $\eta_0 = -\mu_x^2/(2\sigma_x^2)$, $\eta_1 = \mu_x/\sigma_x^2$, and $\eta_2 = -1/(2\sigma_x^2)$. To obtain a normalized Gaussian density, the first coefficient is modified according to $\eta_0 = -\mu_x^2/(2\sigma_x^2) + \log(c(\underline{\eta}))$ with $c(\underline{\eta}) = 1/\sqrt{2\pi}\sigma_x$, while η_1 and η_2 remain unchanged¹.

B. Recursive Moment Calculation

In general, no analytic expressions for the (non-central) moments

$$E_i \triangleq E\{x^i\} = \int x^i \cdot f(x) dx \quad (5)$$

of an exponential density for $i \in \mathbb{N}_0$ exist, not even for the zeroth-order moment E_0 , which is required for determining the normalization constant $c(\underline{\eta})$. Only for some special cases like the Gaussian density, it is possible to derive analytic expressions. However, if at least the first n moments E_0, E_1, \dots, E_{n-1} are given, all higher-order moments can be determined recursively. As shown in [9], [10], integrating (5) by parts with respect to x yields

$$\begin{aligned} E_i &= \underbrace{\left(\frac{x^{i+1}}{i+1} f(x)\right)}_{=0} \Big|_{-\infty}^{\infty} - \int \frac{x^{i+1}}{i+1} \frac{\partial f(x)}{\partial x} dx \\ &= - \int \frac{x^{i+1}}{i+1} \left(\sum_{j=1}^n j \cdot \eta_j \cdot x^{j-1}\right) \cdot f(x) dx, \end{aligned}$$

which finally gives

$$E_i = - \sum_{j=1}^n \frac{j}{i+1} \eta_j E_{i+j}. \quad (6)$$

Thus, if the n lower-order moments $\underline{E}_{0:n-1} \triangleq [E_0, \dots, E_{n-1}]$ are given and the moments up to E_m , $m \geq n$ are of interest, solving the linear system of equations

$$\mathbf{Q}(\underline{\eta}) \cdot \underline{E}_{0:n-1} = \mathbf{R}(\underline{\eta}) \cdot \underline{E}_{n:m} \quad (7)$$

gives the desired higher-order moments $\underline{E}_{n:m} \triangleq [E_n, \dots, E_m]$. The linear system of equations in (7) follows from rearranging the result in (6), where the matrices $\mathbf{Q}(\underline{\eta}) \triangleq [\mathbf{A}(\underline{\eta})]_{0:n-1}$ and $\mathbf{R}(\underline{\eta}) \triangleq -[\mathbf{A}(\underline{\eta})]_{n:m}$ are based on the $(m-n+1) \times (m+1)$ matrix $\mathbf{A}(\underline{\eta})$ in (8). Here, $[\mathbf{A}]_{n:m}$ indicates the columns n to m , $n \leq m$, of matrix \mathbf{A} . The matrix $\mathbf{R}(\underline{\eta})$ is triangular with zeros everywhere except of the main diagonal and the n diagonals below the main diagonal. Thus, (7) can be efficiently solved by means of forward substitution.

It is worth mentioning that it is not possible in general to deduce the parameter vector $\underline{\eta}$ from given moments $\underline{E}_{0:n-1}$. Again, the Gaussian density is an exception from this general statement.

¹The term $\mathcal{N}(x; \mu_x, \sigma_x^2)$ always refers to a *normalized* Gaussian density in this paper.

$$\mathbf{A}(\underline{\eta}) = \underbrace{\begin{bmatrix} 1 & \frac{1}{1}\eta_1 & \frac{2}{2}\eta_2 & \dots & \frac{n}{1}\eta_n & 0 & 0 & \dots & 0 \\ 0 & 1 & \frac{1}{2}\eta_1 & \frac{2}{2}\eta_2 & \dots & \frac{n}{2}\eta_n & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & \frac{1}{m-n}\eta_1 & \frac{2}{m-n}\eta_2 & \dots & \frac{n}{m-n}\eta_n & 0 \\ 0 & \dots & 0 & 0 & 1 & \frac{1}{m-n+1}\eta_1 & \frac{2}{m-n+1}\eta_2 & \dots & \frac{n}{m-n+1}\eta_n \end{bmatrix}}_{=\mathbf{Q}(\underline{\eta})} = \begin{cases} 1 & i = j \\ \frac{j-i}{i} \eta_{j-i} & i < j \leq i+n \\ 0 & \text{otherwise} \end{cases} = \underbrace{\begin{bmatrix} 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{bmatrix}}_{=-\mathbf{R}(\underline{\eta})} \quad (8)$$

IV. GAUSSIAN FILTERING

Based on the properties of exponential densities, it is now possible to derive closed-form and computationally efficient expressions for the mean and variance of the transformed random variable \mathbf{y} .

A. Mean Propagation

When propagating the Gaussian random variable \mathbf{x} through the polynomial transformation $g(\cdot)$ in (4), the mean μ_y of \mathbf{y} can be expressed as

$$\begin{aligned}\mu_y &= \mathbb{E}\{g(\mathbf{x}) + \mathbf{w}\} = \mathbb{E}\{g(\mathbf{x})\} \\ &= \sum_{i=0}^n c_i \cdot \int x^i \cdot \mathcal{N}(x; \mu_x, \sigma_x^2) dx \\ &= \sum_{i=0}^n c_i \cdot \underbrace{\mathbb{E}\{x^i\}}_{=E_i}.\end{aligned}\quad (9)$$

Thus, the mean μ_y results in a weighted sum of non-central moments of a Gaussian random variable. Given the first two moments $E_0 = 1$ and $E_1 = \mu_x$ of \mathbf{x} , all remaining moments up to order n can be calculated by means of solving (7). In doing so, (9) can be expressed as

$$\mu_y = \underline{c}_n^T \cdot \underline{E}_{0:n} = \underline{c}_n^T \cdot \begin{bmatrix} \mathbf{I}_2 \\ \mathbf{L} \end{bmatrix} \cdot \underline{E}_{0:1}, \quad (10)$$

with $\mathbf{L} \triangleq (\mathbf{R}(\eta))^{-1} \mathbf{Q}(\eta)$, where the parameter vector η comprises the parameters of a (normalized) Gaussian density as defined in Section III-A. Furthermore, $\underline{c}_n^T \triangleq [c_0, c_1, \dots, c_n]$ is the vector of polynomial coefficients and \mathbf{I}_n is the $n \times n$ identity matrix.

It is important to note that the second equation in (10) is merely of formal use. From a computational and numerical point of view, it is recommended to first determine the missing higher-order moments $\underline{E}_{2:n}$ as described in Section III-B by solving the linear system of equations (7) via forward substitution. In a second step, the solution for $\underline{E}_{2:n}$ is applied to the first equation in (10).

B. Variance Propagation

In a similar fashion as before, the variance σ_y^2 of \mathbf{y} can be determined. For this purpose, the relation

$$\begin{aligned}\sigma_y^2 &= \mathbb{E}\{(\mathbf{y} - \mu_y)^2\} = \mathbb{E}\{(g(\mathbf{x}) + \mathbf{w} - \mu_y)^2\} \\ &= \mathbb{E}\{g(\mathbf{x})^2\} - \mu_y^2 + \sigma_w^2\end{aligned}\quad (11)$$

is exploited, where both the noise variance σ_w^2 and the propagated mean μ_y are known. Merely the first term (11) has to be determined, which yields

$$\begin{aligned}\mathbb{E}\{g(\mathbf{x})^2\} &= \int g(x) \cdot g(x) \cdot \mathcal{N}(x; \mu_x, \sigma_x^2) dx \\ &= \sum_{i=0}^n \sum_{j=0}^n c_i \cdot c_j \cdot \underbrace{\int x^{i+j} \cdot \mathcal{N}(x; \mu_x, \sigma_x^2) dx}_{=E_{i+j}}.\end{aligned}$$

Thus, in order to calculate the variance σ_y^2 , it is necessary to consider all moments up to order $2n$. Given these moments and exploiting the fact that the product of two polynomials

corresponds to a discrete convolution of the polynomials' coefficients, the variance calculation can be compactly written as

$$\begin{aligned}\sigma_y^2 &= (\underline{c}_n * \underline{c}_n)^T \cdot \underline{E}_{0:2n} - \mu_y^2 + \sigma_w^2 \\ &= (\mathbf{T} \cdot \underline{c}_n)^T \cdot \underline{E}_{0:2n} - \mu_y^2 + \sigma_w^2,\end{aligned}\quad (12)$$

where $*$ is the discrete convolution operator. The second equality indicates an efficient matrix-vector realization of the convolution by means of the matrix \mathbf{T} with entries $t_{i,j} = t_{i+1,j+1} = c_{i-j}$ if $i \in [j, j+n]$ and $t_{i,j} = 0$ otherwise, where $i = 1, 2, \dots, 2n+1$ and $j = 1, 2, \dots, n+1$. Hence, \mathbf{T} is special type of matrix, namely a triangular Toeplitz matrix with only the main diagonal and n diagonals below the main diagonal being non-zero and all elements on individual diagonals being equal.

C. Covariance Calculation

For exploiting the ability of calculating moments of non-linear mappings, a common assumption for performing the measurement update in Gaussian filtering is to assume that the state and the measurement are jointly Gaussian distributed. This only requires the calculation of the covariance between state and measurement, which coincides with the covariance σ_{xy} between \mathbf{x} and \mathbf{y} for the considered generic transformation (4). Similar to (11), the covariance can be formulated as

$$\begin{aligned}\sigma_{xy} &= \mathbb{E}\{(\mathbf{x} - \mu_x) \cdot (\mathbf{y} - \mu_y)\} \\ &= \mathbb{E}\{\mathbf{x} \cdot g(\mathbf{x})\} - \mu_x \cdot \mu_y,\end{aligned}\quad (13)$$

where the expected value corresponds to

$$\begin{aligned}\mathbb{E}\{\mathbf{x} \cdot g(\mathbf{x})\} &= \int x \cdot g(x) \cdot \mathcal{N}(x; \mu_x, \sigma_x^2) dx \\ &= \sum_{i=0}^n c_i \cdot \underbrace{\int x^{i+1} \cdot \mathcal{N}(x; \mu_x, \sigma_x^2) dx}_{=E_{i+1}}.\end{aligned}$$

This is almost identical to the mean calculation in (9) except for the shift by one in the order of the involved moments. Thus, the covariance is given by

$$\sigma_{xy} = \underline{c}_n^T \cdot \underline{E}_{1:n+1} - \mu_x \cdot \mu_y, \quad (14)$$

where μ_y is already known from (10).

D. Polynomial Kalman Filter

The results derived in the previous sections allow the formulation of a Gaussian state estimator for polynomial nonlinearities. For this purpose, the well-known structure of the Kalman filter is exploited. The resulting *polynomial Kalman filter* (PKF) is listed in Algorithm 1 and described in detail in the following paragraphs.

1) *Prediction*: Given the posterior state estimate $\mathbf{x}_{k-1}^e \sim f_{k-1}^e(x_{k-1}) \triangleq \mathcal{N}(x_{k-1}; \mu_{k-1}^e, (\sigma_{k-1}^e)^2)$ of the previous measurement update, the prediction from the previous time step $k-1$ to the current time step k requires the calculation of the predicted mean μ_k^p and variance $(\sigma_k^p)^2$. As the system function $a_k(\cdot)$ is assumed to be a polynomial of degree $n_p \in \mathbb{N}$ with coefficient vector $\underline{c}_{n_p}^p$, equation (10) and (12) can be directly applied in order to determine the desired predicted moments.

Algorithm 1 Polynomial Kalman Filter (PKF)

▷ *Prediction*

- 1: Determine moment vector $\underline{E}_{0:2n_p}$ of posterior state \mathbf{x}_{k-1}^e by solving (7)
- 2: Predicted mean: $\mu_k^p = \underline{c}_{n_p}^p \cdot \underline{E}_{0:n_p}$
- 3: Predicted variance:
 $(\sigma_k^p)^2 = (\mathbf{T} \cdot \underline{c}_{n_p}^p)^\top \cdot \underline{E}_{0:2n_p} - (\mu_k^p)^2 + (\sigma_k^w)^2$

▷ *Measurement Update*

- 4: Determine moment vector $\underline{E}_{0:2n_e}$ of predicted state \mathbf{x}_k^p by solving (7)
 - 5: Measurement mean: $\mu_k^z = \underline{c}_{n_e}^e \cdot \underline{E}_{0:n_e}$
 - 6: Measurement variance:
 $(\sigma_k^z)^2 = (\mathbf{T} \cdot \underline{c}_{n_e}^e)^\top \cdot \underline{E}_{0:2n_e} - (\mu_k^z)^2 + (\sigma_k^v)^2$
 - 7: Covariance: $\sigma_k^{xz} = (\underline{c}_{n_e}^e)^\top \cdot \underline{E}_{1:n_e+1} - \mu_k^p \cdot \mu_k^z$
 - 8: Kalman gain: $K_k = \sigma_k^{xz} / (\sigma_k^z)^2$
 - 9: Calculate posterior mean μ_k^e according to (15)
 - 10: Calculate posterior variance $(\sigma_k^e)^2$ according to (16)
-

2) *Measurement Update*: The measurement update aims at updating the prediction $\mathbf{x}_k^p \sim f_k^p(x_k) \triangleq \mathcal{N}(x_k; \mu_k^p, (\sigma_k^p)^2)$ with the latest measurement value \hat{z}_k . To allow for a closed-form and computationally efficient update, a common assumption in nonlinear Kalman filtering—as in the extended Kalman filter or the unscented Kalman filter—is that the state \mathbf{x}_k^p and the measurement z_k are jointly Gaussian. The implication of this so-called *joint Gaussian assumption* is discussed in Section IV-E. It requires to compute the joint mean vector and joint covariance matrix

$$\underline{\mu}_k^{xz} = \begin{bmatrix} \mu_k^p \\ \mu_k^z \end{bmatrix}, \quad \mathbf{C}_k^{xz} = \begin{bmatrix} (\sigma_k^p)^2 & \sigma_k^{xz} \\ \sigma_k^{xz} & (\sigma_k^z)^2 \end{bmatrix},$$

respectively. The posterior mean and variance are then calculated according to

$$\mu_k^e = \mu_k^p + K_k \cdot (\hat{z}_k - \mu_k^z), \quad (15)$$

$$(\sigma_k^e)^2 = (\sigma_k^p)^2 - K_k \cdot \sigma_k^{xz}, \quad (16)$$

which coincides with the well-known Kalman filter update step, where $K_k \triangleq \sigma_k^{xz} / (\sigma_k^z)^2$ is the Kalman gain. This measurement update requires determining the measurement mean μ_k^z and variance $(\sigma_k^z)^2$ as well as the covariance σ_k^{xz} of state and measurement. Given that the measurement function $h_k(\cdot)$ is a polynomial of degree $n_e \in \mathbb{N}$ with coefficient vector $\underline{c}_{n_e}^e$, all three values can be calculated by means of (10), (12), and (14), respectively. Thanks to these closed-form expressions and the simple Kalman filter equations, the measurement update is straightforward to realize and computationally undemanding.

E. Discussion

When comparing the PKF with the approach proposed in [8], it becomes apparent that both approaches are equivalent regarding the calculated mean and variance values. However, the PKF has the following benefits. First, the involved matrices for calculating the desired moments are straightforward to determine. For instance, the matrices $\mathbf{Q}(\eta)$ and $\mathbf{R}(\eta)$ merely depend linearly on the parameters of the Gaussian density. In [8], however, the involved matrices depend on binomials

coefficients, powers of the mean value, and weighted scalar products of the coefficient vector. This leads to a high computational load for determining the matrices and may cause numerical instability. Second, also the worst-case complexity is higher. While calculating the variance (14) can be performed in $\mathcal{O}(n \cdot \log n)$ if the convolution is realized via fast Fourier transform, the variance calculation in [8] scales with $\mathcal{O}(n^2)$, where $\mathcal{O}(\cdot)$ is the big O in Landau notation. This difference is especially of importance in case of polynomials with a high degree.

The PKF makes two different Gaussian assumptions. First, it assumes the predicted or posterior density to be Gaussian. Second, in order to perform the measurement update, it assumes that the joint density of state and measurement is Gaussian as well. If only the first Gaussian assumption would be in place, the PKF would be an exact Gaussian assumed density filter as it performs moment matching, i.e., the mean and variance calculated by PKF coincide with the true mean and variance. The additional joint Gaussian assumption, however, can result in a poor approximation of the true mean and variance, which may cause a significant loss in estimation performance or even a divergence of the estimator.

To demonstrate the effect of the joint Gaussian assumption on the estimation performance, the polynomial model

$$\mathbf{z} = \mathbf{x}^i + \mathbf{v} \quad (17)$$

is considered in the following, where $i > 0$ is even and the state is $\mathbf{x} \sim \mathcal{N}(x; 0, \sigma_x^2)$. According to (9), (11), and (13), the mean μ_z , variance σ_z^2 , and covariance σ_{xz} are given by

$$\mu_z = E_i, \quad \sigma_z^2 = E_{2i} - E_i + \sigma_v^2, \quad \sigma_{xz} = E_{i+1}, \quad (18)$$

respectively. Since \mathbf{x} has zero mean, it follows that $\eta_1 = 0$. Thus, the matrix $\mathbf{Q}(\eta)$ has only two non-zero elements $q_{11} = q_{22} = 1$, where q_{ij} is the element at row i and column j of matrix \mathbf{Q} . Furthermore, the matrix $\mathbf{R}(\eta)$ is zero everywhere except on the main diagonal and the second diagonal below the main diagonal. This special structure of $\mathbf{Q}(\eta)$ and $\mathbf{R}(\eta)$ leads to the conclusion that all even moments of \mathbf{x} are non-zero and all odd moments are zero, i.e., $E_i \neq 0$ and $E_{i+1} = 0$ for all i being even. Hence, the covariance σ_{xz} in (18) is zero. As a result, state \mathbf{x} and measurement z are uncorrelated and the joint Gaussian of state and measurement is axis-aligned. In Fig. 1(a), the joint Gaussian for $i = 2$, $\sigma_x^2 = 1$, and $\sigma_v^2 = 0.1$ is depicted.

As the covariance σ_{xz} is zero, the Kalman gain K in (15) and (16) is zero as well and no update of the predicted state occurs, i.e., the posterior state \mathbf{x}^e is identical to the predicted state \mathbf{x}^p . In this case, a given measurement value has no impact on the estimation. This, however, is not the case, if the joint Gaussian assumption is not made. In order to demonstrate this, the measurement update is now treated from a strict Bayesian perspective. Here, the posterior state \mathbf{x}^e is represented by the conditional density $f^e(x) \triangleq f(x|z)$ resulting from Bayes' rule

$$f^e(x) = \frac{f(z|x) \cdot f(x)}{f(z)} = \frac{f(x, z)}{f(z)}, \quad (19)$$

where $f(z|x)$ is the likelihood and $f(x)$ is the prior density of \mathbf{x} , which corresponds to the predicted Gaussian density $f^p(x) = \mathcal{N}(x; \mu^p, (\sigma^p)^2)$ in case of the considered recursive state estimation.

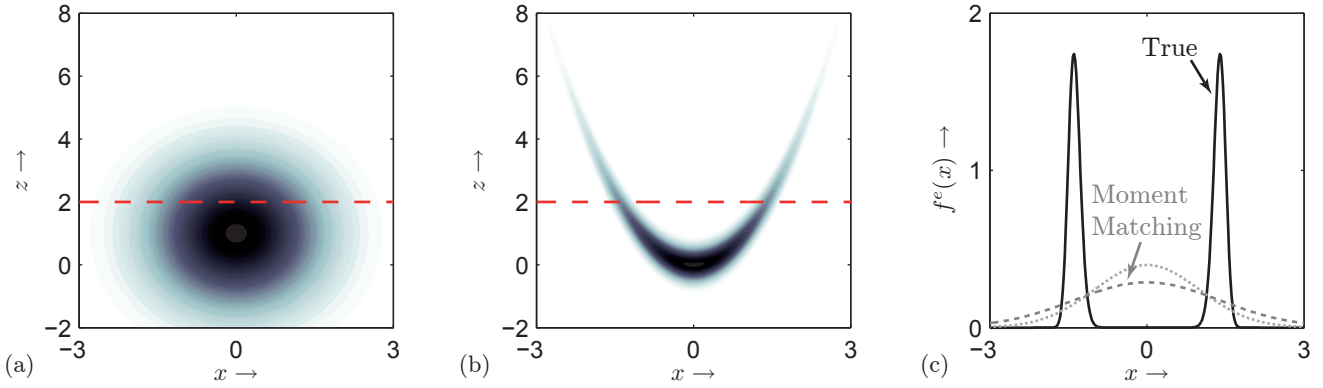


Fig. 1. Joint density $f(x, z)$ and posterior density $f^e(x)$ for $i = 2$, i.e., for a quadratic polynomial. (a) Gaussian approximation of the joint density. The red line indicates the measurement value $\hat{z} = 2$. (b) True joint density. (c) True posterior density (black) and Gaussian approximations, where one is obtained based on the joint Gaussian assumption (dotted) and the other via moment matching (dashed), i.e., its mean and variance coincide with the true posterior moments.

For the considered model (17) with a state x having zero mean, the joint Gaussian assumption leads to a factorization of the joint density $f(x, z) = f(x) \cdot f(z)$ as x and z are uncorrelated, which is equivalent to independence for Gaussian random variables. Hence, the Bayesian update in (19) degenerates to $f^e(x) = f(x) = f^p(x)$. Actually, the joint density $f(x, z)$ is an exponential density for polynomial nonlinearities. This follows from the fact that the likelihood $f(z|x)$ is defined as

$$f(z|x) \triangleq \int \delta(z - x^i - v) \cdot f(v) dv = \mathcal{N}(z; x^i, \sigma_v^2), \quad (20)$$

where $\delta(\cdot)$ is the Dirac delta distribution and the second equality results from exploiting the sifting property of the Dirac delta distribution. The product of likelihood and prior density leads to the exponential density

$$\begin{aligned} f(x, z) &= f(z|x) \cdot f(x) = \mathcal{N}(z; x^i, \sigma_v^2) \cdot \mathcal{N}(x; 0, \sigma_x^2) \\ &= \exp\left(-\log(2\pi\sigma_x\sigma_v) \right. \\ &\quad \left. - \frac{1}{2\sigma_v^2} \cdot (z^2 + \frac{\sigma_v^2}{\sigma_x^2}x^2 - 2zx^i + x^{2i})\right). \end{aligned} \quad (21)$$

This exponential joint density is depicted in Fig. 1(b) for $i = 2$. By comparing Fig. 1(a) with Fig. 1(b) the difference between the true joint density and its Gaussian approximation becomes apparent. Given a measurement value $\hat{z} = 2$, Fig. 1(c) depicts the posterior densities obtained for the Gaussian joint density and the true exponential joint density. It can be seen that the true posterior is bimodal, which only can be coarsely approximated by a Gaussian density. Furthermore, due to the joint Gaussian assumption, the Gaussian posterior does not even match the true posterior mean and variance.

The true posterior is an exponential density, since the joint density $f(x, z = \hat{z})$ is exponential and $f(\hat{z})$ is merely a normalization constant for a given measurement value \hat{z} . Thus, the posterior is a conjugate density of the prior density in the case of polynomial nonlinearities. Unfortunately, a general exponential density is not well suited for recursive processing for mainly two reasons: First, the prediction step as described above requires the availability of the moments $E_{0:2i-1}$, but for exponential densities the calculation of these moments cannot be performed in closed form. Second, even if the moments were available, the prediction step itself merely provides the

predicted moments and no analytic density representation. Determining an exponential density that matches given moments is also not possible in closed form. To overcome these limitations, a novel approach for accurately determining the true posterior mean and variance is proposed in the next section. In doing so, a computationally efficient, recursive Gaussian filter without the joint Gaussian assumption is obtained.

V. HOMOTOPIC BAYESIAN MEASUREMENT UPDATE

In this section, a new method for directly calculating the moments of the posterior density will be introduced that does not require the joint Gaussian assumption and that provides a much higher estimation quality.

The key idea is to transform the known moments of the prior Gaussian density continuously into the desired posterior moments. For this purpose, *homotopy continuation* for calculating the moments of exponential densities as proposed in [9] is exploited. By means of a so-called progression parameter $\gamma \in [0; 1]$ the posterior density $f^e(x)$ is parameterized in such a way that for $\gamma = 0$ the posterior density corresponds to prior Gaussian density $f^p(x)$ and for $\gamma = 1$ the posterior density corresponds to the true exponential density. For the initial value $\gamma = 0$, the moments are known as they coincide with the moments of the Gaussian prior. Incrementing the progression parameter causes moment variations described by means of a *system of ordinary differential equations* (ODEs). Solving this system of ODEs for $\gamma \in [0; 1]$ gives the desired posterior moments.

A. Parameterization

To allow for homotopy continuation, the Bayesian measurement update in (19) is parameterized according to

$$f^e(x; \gamma) = \left(\frac{1}{f(\hat{z})} \cdot f(\hat{z}|x)\right)^\gamma \cdot f^p(x) \quad (22)$$

for a given measurement value \hat{z} , with likelihood $f(\hat{z}|x) = \mathcal{N}(\hat{z}; h(x), \sigma_v^2)$ according to (20) for a polynomial measurement model (2). Further, $f^e(x; \gamma)$ is a parameterized version of the posterior density. For $\gamma = 1$, this parameterized measurement update corresponds to the standard Bayes' rule, while for $\gamma = 0$, the prior density $f^p(x)$ is directly assigned to the posterior density, i.e., no measurement update is performed.

In order to simplify the following calculations, the normalization constant $1/f(\hat{z})$ in (22) is ignored, which is without any disadvantages. Since the zeroth-order moment E_0 , which is reciprocal to the normalization constant, will be calculated as well, ex post division of all higher-order moments by E_0 leads to the correct results (see Section V-C).

Due to ignoring the normalization constant, merely the proportional relation

$$f^e(x; \gamma) \propto f(x, \hat{z}; \underline{\eta}(\gamma)) \triangleq f(\hat{z}|x)^\gamma \cdot f^p(x) \quad (23)$$

is considered instead of (22), where the *parameterized joint density* $f(x, \hat{z}; \underline{\eta}(\gamma)) = \exp(\underline{\eta}(\gamma)^T \cdot \underline{x})$ is an exponential density similar to (21) with parameter vector

$$\underline{\eta}(\gamma) \triangleq \underline{\eta}^p + \gamma \cdot \underline{\eta}^l \in \mathbb{R}^{2n_e+1}$$

depending on γ . Here, $\underline{\eta}^p$ is the parameter vector of the Gaussian prior $f^p(x)$ and $\underline{\eta}^l$ is the parameter vector of the likelihood $f(\hat{z}|x)$ according to

$$\underline{\eta}^l = \begin{bmatrix} -\log(\sqrt{2\pi}\sigma_v) - \frac{\hat{z}^2}{2\sigma_v^2} \\ \frac{\hat{z}}{\sigma_v^2} \cdot \underline{c}_{n_e}^e \\ \underline{0} \end{bmatrix} - \frac{1}{2\sigma_v^2} \cdot (\underline{c}_{n_e}^e * \underline{c}_{n_e}^e),$$

with $\underline{c}_{n_e}^e$ being the coefficient vector of the measurement function $h(\cdot)$ and $\underline{0}$ being a vector of zeros of appropriate dimension. The parameter vector $\underline{\eta}(\gamma)$ directly reflects the continuation in (23).

It is worth mentioning that the parameterized joint density always is a valid exponential density for each $\gamma \in [0; 1]$. As it directly depends on the Gaussian measurement noise v and the Gaussian prior $f^p(x)$, the highest-order monomial in \underline{x} is even and the last element in $\underline{\eta}^l$ is negative.

B. System of Ordinary Differential Equations

By a continuous modification of the progression parameter γ , a continuous variation of the parameter vector $\underline{\eta}(\gamma)$ is achieved. This in turn results in a variation of the moments $E_i(\underline{\eta}(\gamma))$, $i = 0, \dots, 2n_e - 1$, of the parameterized joint density $f(x, \hat{z}; \underline{\eta}(\gamma))$. These moment variations depending on γ can be described by means of a system of ODEs by calculating the partial derivatives $\dot{E}_i \triangleq \frac{\partial E_i(\underline{\eta}(\gamma))}{\partial \gamma}$ for $i = 0, \dots, 2n_e - 1$. The partial derivative of the i th-order moment is given by

$$\begin{aligned} \dot{E}_i &= \frac{\partial E_i(\underline{\eta}(\gamma))}{\partial \gamma} = \left[\frac{\partial E_i}{\partial \underline{\eta}} \Big|_{\underline{\eta}=\underline{\eta}(\gamma)} \right]^T \cdot \frac{\partial \underline{\eta}(\gamma)}{\partial \gamma} \\ &= \left(\frac{\partial \underline{\eta}(\gamma)}{\partial \gamma} \right)^T \cdot \int x^i \cdot \frac{\partial f(x, \hat{z}; \underline{\eta})}{\partial \underline{\eta}} \Big|_{\underline{\eta}=\underline{\eta}(\gamma)} dx \\ &= \left(\frac{\partial \underline{\eta}(\gamma)}{\partial \gamma} \right)^T \cdot \int x^i \begin{bmatrix} 1 \\ x \\ \vdots \\ x^{2n_e} \end{bmatrix} \exp(\underline{\eta}(\gamma)^T \cdot \underline{x}) dx \\ &= [E_i(\underline{\eta}(\gamma)) \quad E_{i+1}(\underline{\eta}(\gamma)) \quad \cdots \quad E_{i+2n_e}(\underline{\eta}(\gamma))] \cdot \underline{\eta}^l, \quad (24) \end{aligned}$$

which relates the variation of the i th-order moment to moments of order up to $i + 2n_e$. In the following, $E_i^{(\gamma)} \triangleq E_i(\underline{\eta}(\gamma))$ is used as shorthand term.

With the result in (24), the system of ODEs comprising the moment variations of all moments up to order $2n_e - 1$ is

$$\begin{aligned} \dot{\underline{E}}_{0:2n_e-1} &= (\mathbf{T}(\underline{\eta}^l))^T \cdot \underline{E}_{0:4n_e-1}^{(\gamma)} \\ &= \mathbf{T}^l \cdot \underline{E}_{0:2n_e-1}^{(\gamma)} + \mathbf{T}^h \cdot \underline{E}_{2n_e:4n_e-1}^{(\gamma)}, \end{aligned}$$

where $\mathbf{T}(\underline{\eta}^l)$ is a Toeplitz matrix with entries $t_{i,j} = t_{i+1,j+1} = \eta_{i-j}^l$ if $i \in [j, j + 2n_e]$ and $t_{i,j} = 0$ otherwise, where $i = 1, 2, \dots, 4n_e$ and $j = 1, 2, \dots, 2n_e$. The $2n_e \times 2n_e$ matrices \mathbf{T}^l and \mathbf{T}^h are sub-matrices of $\mathbf{T}(\underline{\eta}^l)$ according to $\mathbf{T}(\underline{\eta}^l) = [\mathbf{T}^l \quad \mathbf{T}^h]^T$. Besides the lower-order moments $\underline{E}_{0:2n_e-1}^{(\gamma)}$, the system of ODEs also depends on the higher-order moments $\underline{E}_{2n_e:4n_e-1}^{(\gamma)}$. Fortunately, with the result of (7), the dependence on the higher-order moments can be resolved. In doing so, the system of ODEs can be reformulated into

$$\dot{\underline{E}}_{0:2n_e-1} = \left(\mathbf{T}^l + \mathbf{T}^h (\mathbf{R}(\underline{\eta}(\gamma)))^{-1} \mathbf{Q}(\underline{\eta}(\gamma)) \right) \cdot \underline{E}_{0:2n_e-1}^{(\gamma)} \quad (25)$$

with matrices $\mathbf{R}(\underline{\eta}(\gamma))$ and $\mathbf{Q}(\underline{\eta}(\gamma))$ according to (8), which vary with γ as they depend on the parameters of the parameterized joint density $f(x, \hat{z}; \underline{\eta}(\gamma))$.

C. Initialization and Solution

The system of ODEs in (25) describes the moment variations caused by homotopy continuation of the Bayesian measurement update (23) in a very elegant manner. For solving this system of ODEs, standard numerical solvers based on the Runge-Kutta method [11] can be employed. The solution describes a trajectory of the moments $\underline{E}_{0:2n_e-1}^{(\gamma)}$ depending on different values of the progression parameter γ . The desired moments of the posterior density $f^e(x)$ are obtained for $\gamma = 1$, i.e., $\underline{E}_{0:2n_e-1}^{(1)}$ comprises the result. As mentioned above, the moments in $\underline{E}_{0:2n_e-1}^{(1)}$ are unnormalized as merely the proportional relation (23) was considered. Multiplying $\underline{E}_{0:2n_e-1}^{(1)}$ with the normalization constant

$$\alpha \triangleq 1/f(\hat{z}) = 1/E_0^{(1)} \quad (26)$$

yields the actual posterior moments.

Please note that the matrix $\mathbf{R}(\underline{\eta}(\gamma))$ in (25) is singular for $\gamma = 0$. To avoid an inversion of this matrix for $\gamma = 0$, an initialization procedure is proposed that determines an initial solution $\underline{E}_{0:2n_e-1}^{(\Delta\gamma)}$ for the first solution step, with $\Delta\gamma$ being a small positive step value². Based on the initial solution $\underline{E}_{0:2n_e-1}^{(\Delta\gamma)}$, the system of ODEs in (25) is then solved in a standard fashion for $\gamma \in [\Delta\gamma; 1]$.

To determine the initial solution, the moment calculation (5) is expanded around $\gamma = 0$ via a first-order Taylor-series according to

$$\begin{aligned} E_i^{(\Delta\gamma)} &\approx E_i^{(0)} + \Delta\gamma \cdot \frac{\partial E_i^{(\gamma)}}{\partial \underline{\eta}^T} \cdot \frac{\partial \underline{\eta}}{\partial \gamma} \Big|_{\underline{\eta}=\underline{\eta}(\gamma), \gamma=0} \\ &= E_i(\underline{\eta}^p) + \Delta\gamma \cdot \left[\frac{\partial E_i(\underline{\eta})}{\partial \underline{\eta}} \Big|_{\underline{\eta}=\underline{\eta}^p} \right]^T \cdot \underline{\eta}^l \quad (27) \end{aligned}$$

for each moment i , where $E_i(\underline{\eta}^p)$ are the moments of the predicted Gaussian state $\mathbf{x}^p \sim f^p(x) = \mathcal{N}(x; \mu^p, (\sigma^p)^2)$. The

²In the simulations in Section VI, $\Delta\gamma = 10^{-7}$ is used.

Algorithm 2 Homotopic Polynomial Gaussian Filter (HPGF)

▷ *Prediction*

- 1: Determine moment vector $\underline{E}_{0:2n_p}$ of posterior state \mathbf{x}_{k-1}^e by solving (7)
- 2: Predicted mean: $\mu_k^p = \underline{c}_{n_p}^p \cdot \underline{E}_{0:2n_p}$
- 3: Predicted variance:

$$(\sigma_k^p)^2 = (\mathbf{T} \cdot \underline{c}_{n_p}^p)^T \cdot \underline{E}_{0:2n_p} - (\mu_k^p)^2 + (\sigma_k^w)^2$$

▷ *Measurement Update*

- 4: Determine initial solution according to (27)
 - 5: Solve system of ODEs (25) for $\gamma \in [\Delta\gamma; 1]$
 - 6: Calculate posterior mean μ_k^e according to (28)
 - 7: Calculate posterior variance $(\sigma_k^e)^2$ according to (29)
-

second summand in (27) is given by (24) for $\eta(\gamma) = \eta^p$. This derivative merely depends on the Gaussian prior density $f^p(x)$. Thus, all higher-order moments in (24) can be determined via solving the moment recursion in (7).

D. Homotopic Polynomial Gaussian Filter

The novel homotopic polynomial Gaussian filter (HPGF) for polynomial nonlinearities is summarized in Algorithm 2. The prediction step coincides with the prediction of the PKF as proposed in Section IV. The measurement update utilizing homotopy continuation for calculating the posterior moments consists of three operations for each time step k : First, initialization as proposed in the previous section. Second, solving the system of ODEs (25). Finally, calculating the posterior mean and variance by correcting the ODE solution $\underline{E}_{0:2n_e}^{(1)}$ with the normalization constant (26) according to

$$\mu_k^e = \alpha \cdot E_1^{(1)}, \quad (28)$$

$$(\sigma_k^e)^2 = \alpha \cdot E_2^{(1)} - (\mu_k^e)^2, \quad (29)$$

which yields the desired posterior state estimate $\mathbf{x}_k^e \sim f_k^e(x_k) = \mathcal{N}(x_k; \mu_k^e, (\sigma_k^e)^2)$.

VI. RESULTS

In the following, both Gaussian filters the PKF and the HPGF proposed in this paper are compared to state-of-the-art filters by means of numerical simulations.

A. Moment Homotopy Examples

For the first simulation, the polynomial measurement model (17) is revisited, where now merely the quadratic (order $i = 2$) and the cubic (order $i = 3$) case are considered. Furthermore, the state estimate $\mathbf{x} \sim \mathcal{N}(x; 0, 1)$ is standard Gaussian distributed, the measurement value is $\hat{z} = 1$, and $(\sigma_v)^2 = 0.1$ is the variance of the measurement noise.

In Table I, the posterior moments calculated by the proposed HPGF and PKF are compared with the true moments and the results obtained by means of the extended Kalman filter (EKF). The true moments have been calculated via numerical integration. It can be seen that the results of the HPGF coincide with the true moments. Compared with numerical integration, the HPGF has the benefits of a significantly lower computational burden and that no integration interval needs

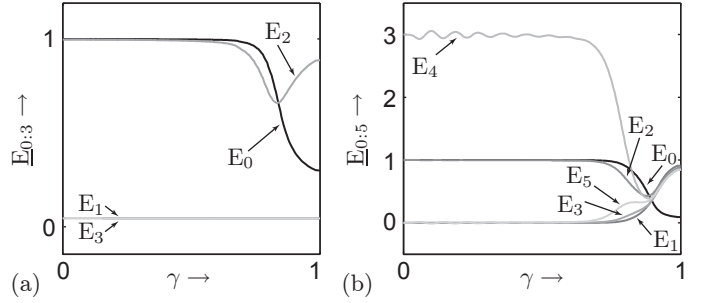


Fig. 2. Trajectories of posterior moments. (a) Quadratic model $z = x^2 + v$. (b) Cubic model $z = x^3 + v$.

to be determined. In Fig. 2, the trajectories of the posterior moments resulting from the homotopy continuation are shown. It can be seen how the moments of the prior Gaussian are transformed into the true posterior moments.

EKF and PKF both rely on the joint Gaussian assumption. Thus, the moments of both filters differ significantly from the true moments (see Table I). As discussed in Section IV-E, for $i = 2$ no update of the state estimate can be performed. For the cubic case, however, an update should be available, which is true for the PKF. The EKF, however, suffers from the linearization, which results in a zero Kalman gain. Thus, no update is performed at all. This example clearly shows the benefits of avoiding the joint Gaussian assumption.

B. Chaotic Synchronization

In this example, the polynomial system model

$$\mathbf{x}_{k+1} = T_4(\mathbf{x}_k) + \mathbf{w}_k \quad (30)$$

as used in [8] is considered, where $T_i(x) = 2x \cdot T_{i-1}(x) - T_{i-2}(x)$ for $i = 2, 3, \dots$ is the i th Chebyshev polynomial, with $T_0(x) = 1$ and $T_1(x) = x$. It is known that models as in (30) generate chaotic sequences [12], which is of practical use in securing communication systems. The true initial state \mathbf{x}_0 a time step $k = 0$ is assumed to be Gaussian with mean $\mu_0^x = 0.3$ and variance $(\sigma_0^x)^2 = 0.25$.

At first, a linear measurement model

$$\mathbf{z}_k = \mathbf{x}_k + \mathbf{v}_k \quad (31)$$

is employed, with measurement noise variance $(\sigma_v)^2 = 10^{-2}$. $(\sigma_w)^2$ and system noise variance being $(\sigma_w)^2 = 10^{-2}$ (high noise) or $(\sigma_w)^2 = 10^{-3}$ (low noise). As the measurement model is linear, the joint Gaussian assumption is correct. Thus, the HPGF does not need to be considered here. PKF is compared against EKF, unscented Kalman filter (UKF), and

TABLE I. COMPARISON OF THE POSTERIOR MOMENTS CALCULATED VIA DIFFERENT APPROACHES FOR THE QUADRATIC AND CUBIC MODEL.

	Quadratic				Cubic			
	True	HPGF	EKF	PKF	True	HPGF	EKF	PKF
E_0	0.2664	0.2664	-	-	0.0932	0.0932	-	-
E_1	0.0	0.0	0.0	0.0	0.9113	0.9113	0.0	0.1987
E_2	0.8820	0.8820	1.0	1.0	0.8760	0.8760	1.0	0.4434
E_3	0.0	0.0	-	-	0.8526	0.8525	-	-
E_4	-	-	-	-	0.8442	0.8442	-	-
E_5	-	-	-	-	0.8457	0.8456	-	-

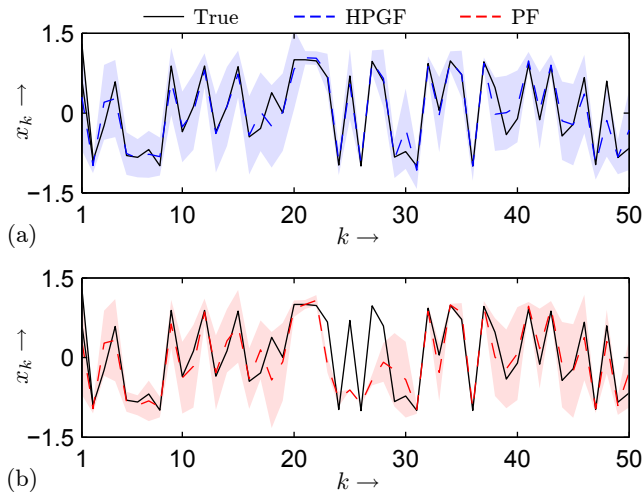


Fig. 3. State trajectory (black, solid line) and the estimates of (a) HPGF and (b) PF together with the corresponding 2-sigma confidence regions.

a particle filter (PF) with systematic resampling [13] and 500 samples. The latter is the only non-Gaussian filter. For all filters, 50 Monte Carlo simulation runs with identical noise sequences are performed, where the estimates are calculated for 50 time steps. As performance indicators, the root mean square error (rmse), the normalized estimation error squared (nees), and the runtime for 50 time steps are employed.

In Table II, the average rmse, nees, and runtime over all Monte Carlo runs are listed for all filters and for both noise cases. For high noise, the proposed PKF outperforms all Gaussian filters in terms of rmse and nees, i.e., its estimates are closest to the true system state (low rmse) and at the same time the estimates are not overly confident (low nees). Furthermore, the matrix-vector terms proposed for the PKF allow for a runtime being close to the EKF, which is known to be the fastest Gaussian filter.

For the low noise case, UKF performs best in terms of estimation error, but PKF is very close to it. PF occasionally suffers from particle depletion, i.e., most of the particles converge towards the same state, which coincides with an overconfident estimate and thus an exceedingly high nees value. Even significantly increasing the number of particles or using different resampling techniques yields no improvement.

If a cubic measurement model $z_k = \frac{x_k^3}{20} + v_k$ with measurement noise $v_k \sim \mathcal{N}(v_k; 0, 10^{-5})$ instead of the linear model (31) is utilized, it turns out that *all* Gaussian filters relying on the joint Gaussian assumption diverge. The HPGF, however, is able to provide valid estimates. In Fig. 3(a), an exemplary state trajectory is depicted. The estimates of HPGF accurately follow the true state. Furthermore, the true state is always within the 2-sigma confidence region of the estimates.

TABLE II. AVERAGE RMSE, NEES, AND RUNTIME FOR THE CHAOTIC SYSTEM MODEL (30) AND THE LINEAR MEASUREMENT MODEL (31).

	$\sigma_w^2 = 10^{-2}$				$\sigma_w^2 = 10^{-3}$			
	EKF	UKF	PF	PKF	EKF	UKF	PF	PKF
rmse	0.410	0.336	0.292	0.316	0.148	0.118	0.268	0.118
nees	4.737	1.550	1.168	1.041	7.279	1.110	–	1.129
time	0.016	0.038	0.109	0.017	0.017	0.037	0.102	0.018

The result of the PF depicted in Fig. 3(b) is less accurate and shows sample depletion from time step $k = 20$ to $k = 27$.

VII. CONCLUSION AND FUTURE WORK

Two methods for the efficient calculation of moments have been introduced in this paper. The first method named polynomial Kalman filter (PKF) efficiently calculates the moments of a polynomial mapping of a Gaussian random variable. When applied to the prediction step, the moment calculation is exact. For the filter step, this method leads to a superior estimation performance compared to existing Gaussian filters. However, the typical additional Gaussian assumption for the joint density of state and measurement is required that can cause highly inaccurate or even diverging estimates. Hence, in order to avoid this assumption, a second method named homotopic polynomial Gaussian filter (HPGF) for the almost exact calculation of the posterior moments in the filter step is introduced. This method is based on a homotopy continuation for polynomial nonlinearities. Combining both methods results in a Gaussian assumed density filter for polynomial nonlinearities that can compete even with non-Gaussian filters.

Future work is devoted to extend the proposed Gaussian filters to the multi-dimensional case. Furthermore, resolving the second limitation mentioned at the end of Section IV-E allows an extension towards a full exponential filter.

REFERENCES

- [1] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb. 2002.
- [2] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. The MIT Press, 2005.
- [3] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. Dover Publications, Inc., 2007.
- [4] E. A. Wan and R. van der Merwe, "The Unscented Kalman Filter," in *Kalman Filtering and Neural Networks*, S. Haykin, Ed. John Wiley & Sons, Inc., 2001, ch. The Unscented Kalman Filter, pp. 221–280.
- [5] I. Arasaratnam and S. Haykin, "Cubature Kalman Filters," *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1254–1269, Jun. 2009.
- [6] M. F. Huber and U. D. Hanebeck, "Gaussian Filter based on Deterministic Sampling for High Quality Nonlinear Estimation," in *Proceedings of the 17th IFAC World Congress*, Seoul, Republic of Korea, Jul. 2008.
- [7] M. Basin, "Optimal Filtering for Partially Measured Polynomial System States," in *Proceedings of the 2005 American Control Conference (ACC)*, Portland, OR, USA, Jun. 2005, pp. 4022–4027.
- [8] M. B. Luca, S. Azou, G. Burel, and A. Serbanescu, "On Exact Kalman Filtering of Polynomial Systems," *IEEE Transactions on Circuits and Systems—I: Regular Papers*, vol. 53, no. 6, pp. 1329–1340, Jun. 2006.
- [9] A. Rauh and U. D. Hanebeck, "Calculating Moments of Exponential Densities Using Differential Algebraic Equations," *IEEE Signal Processing Letters*, vol. 10, no. 5, pp. 144–147, May 2003.
- [10] D. Brigo, B. Hanzon, and F. L. Gland, "A Differential Geometric Approach to Nonlinear Filtering: the Projection Filter," *Institut National De Recherche en Informatique et en Automatique*, Tech. Rep. 2598, Jun. 1995.
- [11] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. Cambridge University Press, 2007, ch. 17: Integration of Ordinary Differential Equations.
- [12] T. Rivlin, *Chebyshev Polynomials*. New York: Wiley, 1990.
- [13] J. Carpenter, P. Clifford, and P. Fearnhead, "Improved particle filter for nonlinear problems," in *IEE Proceedings Radar, Sonar and Navigation*, vol. 146, no. 1, Feb. 1999, pp. 2–7.