Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Moment-Based Prediction Step for Nonlinear Discrete-Time Dynamic Systems Using Exponential Densities

Andreas Rauh and Uwe D. Hanebeck, *Member, IEEE*

Abstract— **In this paper, an efficient approach for a moment-based Bayesian prediction step for both linear and nonlinear discrete-time dynamic systems using exponential densities with polynomial exponents is proposed. The exact solution of the prediction step is approximated by an exponential density which minimizes the Kullback-Leibler distance. Compared to other approaches, the user of this procedure can specify the approximation quality by controlling the deviation between the moments of the exact and the approximated solution. Furthermore, this algorithm can also be used for the adaptation of the order of the exponential densities either to improve the approximation quality or to reduce the computational effort.**

I. Introduction

Estimating predicted probability densities is a common problem in filtering theory for nonlinear dynamic systems [1]. Only for a few special cases this problem can be solved analytically, e.g. if prior information about the state variable is described by Gaussian densities. Even then, system state equations have to be restricted to linear mapping functions with additive Gaussian system noise to obtain exact closed form solutions for the predicted densities. These restrictions are equivalent to the prerequisites of the prediction step considered in linear filtering problems solved exactly by the Kalman filter.

Usually, the exact prediction step can only be calculated by numerical evaluation of convolution integrals over infinite regions of the state space, if the system is characterized by nonlinear state equations, non-Gaussian a priori information about the state variables, non-Gaussian system noise, and non-additive uncertainties. Therefore, techniques relying on particle filters or grid-based approximations of the predicted probability densities are often applied [2], [3].

Other common methods, using no discretization of the state space, rely on analytical density descriptions. In this case, the exact predicted density is approximated by optimization of the parameters of an assumed density such that a distance measure, e.g. the Kullback-Leibler distance [4] between the exact and the approximated density, is minimized. These parameterizable densities are often chosen as exponential densities with polynomial exponents [5], [6]. The case of Gaussian density approximations corresponds to finding optimal parameters of an exponential density with second order polynomial exponent. It has to be pointed out, that even for linear mapping functions with additive Gaussian system noise the prediction step cannot be calculated analytically, if non-Gaussian exponential densities are assumed.

In this paper, a general approach for a prediction step using exponential densities to approximate the predicted density is considered for nonlinear discrete-time state equations. This approximation of the prediction step is neither restricted to additive system noise nor to linear mapping functions.

The new approach significantly reduces the difficulties of minimization of the Kullback-Leibler distance between the exact predicted density and an assumed approximation. In general, such optimization criteria are nonlinear in the components of a multi-dimensional parameter space. In this paper, the distance between the exact density and its approximation by an exponential density with fixed or variable order polynomial exponents is minimized without explicitly evaluating the integral defining the exact predicted density.

For that purpose, a progression approach corresponding to a continuous variation of the state equation from an identity mapping to the desired nonlinear state equation defines continuous variations of the moments of the predicted density starting with an initial parameter set with known moments. By solving a set of ordinary differential equations relating continuous moment variations to continuous variations of the density parameters an approximation of the predicted density is determined. In the case of exponential densities this procedure minimizes the Kullback-Leibler distance to the exact solution. The corresponding proof is given.

The advantages of the presented algorithm are: First, instead of numerical integration over infinite regions of the state space, a differential algebraic approach is applied to the moment calculation of exponential densities [7]. Second, the order of the polynomial exponent can be adapted, if the approximation quality of the moments of the predicted density is decreasing. Third, the computational effort of the prediction algorithm can be influenced by the user by limiting the order of the exponential density. If the presented prediction algorithm is applied to state estimation of dynamical systems a filter step which incorporates measured data has to be implemented additionally. For exponential densities such techniques have e.g. been described in [8].

In Section II, a precise problem formulation is given. In Section III, the idea of approximating the predicted density

This work was performed while the authors were with the Institute of Automatic Control Engineering (LSR), Technische Universität München, München, Germany.

A. Rauh is with the Department of Measurement, Control, and Microtechnology (MRM), University of Ulm, D-89069 Ulm, Germany Andreas.Rauh@uni-ulm.de.

U. D. Hanebeck is with the Institute of Computer Science and Engineering, Universität Karlsruhe (TH), D-76128 Karlsruhe, Germany Uwe.Hanebeck@ieee.org.

A. Rauh wants to thank G. Schmidt, former head of LSR, for the opportunity to perform this work at his institute and E. P. Hofer, head of MRM, for helpful suggestions for improvement of this paper.

0-7803-9568-9/05/$20.00 ©2005 IEEE

by an exponential density with the same moments as the exact density is explained. Then, in Section IV, an overview of the key components of the moment-based prediction algorithm is given. In Section V, the implementation of the prediction algorithm is explained in more detail. In Section VI, an example demonstrating the performance of this algorithm for recursive prediction is given for a nonlinear state equation. Finally, in Section VII this paper is concluded.

II. Problem Formulation

In this paper, a prediction step for scalar nonlinear discrete-time system state equations

$$x_{k+1} = a_k\left(x_k, \underline{v}_k\right) \tag{1}$$

is proposed for uncertain state variables x_k and x_{k+1}, both described by exponential densities with polynomial exponents. The system state equation $a_k\left(x_k, \underline{v}_k\right)$ represents a nonlinear relation between the uncertain state variables x_k and x_{k+1}, as well as uncertain, stochastic parameters and system noise described by the vector $\underline{v}_k$. For additive uncertainties v_k the state equation

$$x_{k+1} = a_k\left(x_k\right) + v_k \ ,$$

leads to the exact predicted density $\tilde{f}^x_{k+1}\left(x_{k+1}\right)$ given by the convolution integral

$$\tilde{f}^x_{k+1}\left(x_{k+1}\right) = \int_{-\infty}^{\infty} f^x_k\left(x_k\right) f^v_k\left(x_{k+1} - a_k\left(x_k\right)\right) dx_k \ . \tag{2}$$

In general, analytic expressions for this integral are only available for a few special cases such as Gaussian (mixture) densities $f^x_k\left(x_k\right)$, additive Gaussian noise v_k with the density $f^v_k\left(v_k\right)$, and linear mapping functions $a_k\left(x_k\right)$.

To enable efficient recursive prediction for nonlinear state equations, the exact predicted densities $\tilde{f}^x_{k+1}\left(x_{k+1}\right)$ have to be approximated adequately. In recursive application, the number of parameters of the approximated density should only increase moderately. Users should be able to control quality and computational effort of this approximation.

In this paper, the densities $\tilde{f}^x_k\left(x_k\right)$ and $\tilde{f}^x_{k+1}\left(x_{k+1}\right)$ are approximated by exponential densities

$$f^x_k\left(\underline{\eta}^x_k, x_k\right) = c\left(\underline{\eta}^x_k\right) \exp\left(\sum_{j=1}^{2n} \eta^x_{k,j} x^j_k\right) \tag{3}$$

of order $2n$ with $n \in \mathbb{N}$ and the density parameters $\underline{\eta}^x_k = \begin{bmatrix}\eta^x_{k,1} & \cdots & \eta^x_{k,j} & \cdots & \eta^x_{k,2n}\end{bmatrix}^T$. To guarantee correct probability densities, the highest-order coefficient $\eta^x_{k,2n}$ has to be negative. Then, finite moments

$$M^x_{k,i}\left(\underline{\eta}^x_k\right) = \int_{-\infty}^{\infty} x^i_k f^x_k\left(\underline{\eta}^x_k, x_k\right) dx_k < \infty \tag{4}$$

exist for $i \in \mathbb{N}_0$. The negativity of the highest-order coefficient $\eta^x_{k,2n}$ is guaranteed by the parameter vector

$$\underline{\eta}^x_k = \left[\eta^x_{k,1} \ \cdots \ \eta^x_{k,j} \ \cdots \ -\left(\eta^x_{k,2n}\right)^2\right]^T \ . \tag{5}$$

After the prediction step, the normalization factor $c\left(\underline{\eta}^x_{k+1}\right)$ has to be chosen such that the predicted density $f^x_{k+1}\left(\underline{\eta}^x_{k+1}, x_{k+1}\right)$ is normalized again, i.e.,

$$M^x_{k+1,0}\left(\underline{\eta}^x_{k+1}\right) = \int_{-\infty}^{\infty} f^x_{k+1}\left(\underline{\eta}^x_{k+1}, x_{k+1}\right) dx_{k+1} = 1. \tag{6}$$

The presented approach relies on calculating the exact moments $\tilde{M}^x_{k+1,j}$ of the predicted state variable x_{k+1}, in terms of the moments of x_k and v_k. Then, an approximation $f^x_{k+1}\left(\underline{\eta}^x_{k+1}, x_{k+1}\right)$ of the predicted density is determined which has approximately the same moments

$$\underline{M}^x_{k+1,[1:2n]}\left(\underline{\eta}^x_{k+1}\right) \approx \underline{\tilde{M}}^x_{k+1,[1:2n]} \ , \tag{7}$$

where the index $[1:2n]$ denotes a vector of moments of orders $1, \ldots, 2n$.

III. Motivation of the Moment-Based Prediction Step for Exponential Densities

In this Section, the important property is derived that a normalized exponential density $f^x_{k+1}\left(\underline{\eta}^x_{k+1}, x_{k+1}\right)$ with the same moments of orders $1, \ldots, 2n$ as the exact probability density $\tilde{f}^x_{k+1}\left(x_{k+1}\right)$ minimizes the Kullback-Leibler distance

$$G\left(\tilde{f}^x_{k+1} \| f^x_{k+1}\right) = \int_{-\infty}^{\infty} g\left(\underline{\eta}^x_{k+1}, x_{k+1}\right) dx_{k+1} \tag{8}$$

with

$$g\left(\underline{\eta}, x\right) = \tilde{f}\left(x\right) \ln\left(\frac{\tilde{f}\left(x\right)}{f\left(\underline{\eta}, x\right)}\right) \ . \tag{9}$$

Note, that a differential geometric filtering approach for continuous-time problems which also aims at the minimization of a distance measure between the exact and approximate density functions has been published in [5], [6]. In contrast to that work, the approach presented in this paper introduces a progression approach which defines continuous moment variations and thus continuous variations of the density parameters for discrete-time systems.

Theorem: If the first $2n$ moments $M^x_{k+1,i}\left(\underline{\eta}^x_{k+1}\right)$ of the approximate, normalized predicted density $f^x_{k+1}\left(\underline{\eta}^x_{k+1}, x_{k+1}\right)$ are equal to the moments $\tilde{M}^x_{k+1,i}$ of the true density, i.e.,

$$M^x_{k+1,[1:2n]}\left(\underline{\eta}^x_{k+1}\right) = \tilde{M}^x_{k+1,[1:2n]} \ , \tag{10}$$

the Kullback-Leibler distance $G\left(\tilde{f}^x_{k+1} \| f^x_{k+1}\right)$ between $\tilde{f}^x_{k+1} = \tilde{f}^x_{k+1}\left(x_{k+1}\right)$ and its approximation $f^x_{k+1} = f^x_{k+1}\left(\underline{\eta}^x_{k+1}, x_{k+1}\right)$ is minimized. ■

Proof: The derivative of $G\left(\tilde{f}^x_{k+1} \| f^x_{k+1}\right)$ with respect to the

parameter vector $\underline{\eta}^x_{k+1}$ is set to zero according to

$$\begin{aligned}\frac{\partial G\left(\tilde{f}^x_{k+1}\|f^x_{k+1}\right)}{\partial \underline{\eta}^x_{k+1}} &= -\int_{-\infty}^{\infty}\frac{\tilde{f}^x_{k+1}}{f^x_{k+1}}\frac{\partial f^x_{k+1}}{\partial \underline{\eta}^x_{k+1}}dx_{k+1}\\ &= -\int_{-\infty}^{\infty}\frac{\tilde{f}^x_{k+1}}{f^x_{k+1}}\left(\frac{\partial c\left(\underline{\eta}^x_{k+1}\right)}{\partial \underline{\eta}^x_{k+1}}\frac{f^x_{k+1}}{c\left(\underline{\eta}^x_{k+1}\right)}\right.\\ &\qquad\left.+f^x_{k+1}\begin{bmatrix}x_{k+1}\\ x^2_{k+1}\\ \vdots\\ -2\left(\eta^x_{k+1,2n}\right)x^{2n}_{k+1}\end{bmatrix}\right)dx_{k+1}=0\end{aligned} \tag{11}$$

which leads to

$$\begin{aligned}&-\left(\frac{\partial c\left(\underline{\eta}^x_{k+1}\right)}{\partial \underline{\eta}^x_{k+1}}\frac{1}{c\left(\underline{\eta}^x_{k+1}\right)}\right)\underbrace{\int_{-\infty}^{\infty}\tilde{f}^x_{k+1}\left(x_{k+1}\right)dx_{k+1}}_{=1}\\ &-\begin{bmatrix}\tilde{M}^x_{k+1,1}\\ \tilde{M}^x_{k+1,2}\\ \vdots\\ -2\left(\eta^x_{k+1,2n}\right)\tilde{M}^x_{k+1,2n}\end{bmatrix}=0.\end{aligned} \tag{12}$$

Choosing the normalization factor $c\left(\underline{\eta}^x_{k+1}\right)$ such that

$$M^x_{k+1,0}=\int_{-\infty}^{\infty}f^x_{k+1}\left(\underline{\eta}^x_{k+1},x_{k+1}\right)dx_{k+1}=1\ ,$$

the derivative of the zero-order moment with respect to the parameter vector yields

$$\frac{\partial M^x_{k+1,0}\left(\underline{\eta}^x_{k+1}\right)}{\partial \underline{\eta}^x_{k+1}}=\frac{\partial}{\partial \underline{\eta}^x_{k+1}}\int_{-\infty}^{\infty}f^x_{k+1}\left(x_{k+1}\right)dx_{k+1}\overset{!}{=}\underline{0}\ . \tag{13}$$

Equation (13) is rewritten as

$$-\frac{\frac{\partial c\left(\underline{\eta}^x_{k+1}\right)}{\partial \underline{\eta}^x_{k+1}}}{c\left(\underline{\eta}^x_{k+1}\right)}=\begin{bmatrix}M^x_{k+1,1}\left(\underline{\eta}^x_{k+1}\right)\\ M^x_{k+1,2}\left(\underline{\eta}^x_{k+1}\right)\\ \vdots\\ -2\left(\eta^x_{k+1,2n}\right)M^x_{k+1,2n}\left(\underline{\eta}^x_{k+1}\right)\end{bmatrix}. \tag{14}$$

Substituting this expression in (12), the gradient

$$\begin{aligned}\frac{\partial G\left(\tilde{f}^x_{k+1}\|f^x_{k+1}\right)}{\partial \underline{\eta}^x_{k+1}} &= -\begin{bmatrix}\tilde{M}^x_{k+1,1}\\ \tilde{M}^x_{k+1,2}\\ \vdots\\ -2\left(\eta^x_{k+1,2n}\right)\tilde{M}^x_{k+1,2n}\end{bmatrix}\\ &\quad+\begin{bmatrix}M^x_{k+1,1}\left(\underline{\eta}^x_{k+1}\right)\\ M^x_{k+1,2}\left(\underline{\eta}^x_{k+1}\right)\\ \vdots\\ -2\left(\eta^x_{k+1,2n}\right)M^x_{k+1,2n}\left(\underline{\eta}^x_{k+1}\right)\end{bmatrix}=0\end{aligned} \tag{15}$$

is expressed in terms of the difference of the moments of the exact and the approximated density. From (15) it is obvious that the minimum of the Kullback-Leibler distance corresponds to equal moments

$$M^x_{k+1,[1:2n]}\left(\underline{\eta}^x_{k+1}\right)=\tilde{M}^x_{k+1,[1:2n]} \tag{16}$$

with $\eta^x_{k+1,2n}\neq 0$. The parameter vector $\underline{\eta}^x_{k+1}$ cannot be determined explicitly if the corresponding desired moment vectors are given. However, as shown in the following Sections, small variations of the density parameters $\underline{\eta}^x_{k+1}$ are related to small variations of the moments of the exponential density by the diffeomorphism

$$\begin{aligned}&\frac{\partial \underline{M}^x_{k+1,[1:2n]}\left(\underline{\eta}^x_{k+1}\right)}{\partial \underline{\eta}^x_{k+1}}=\mathbf{J}\quad\Longleftrightarrow\\ &\partial \underline{M}^x_{k+1,[1:2n]}\left(\underline{\eta}^x_{k+1}\right)=\mathbf{J}\,\partial \underline{\eta}^x_{k+1}\ .\end{aligned} \tag{17}$$

■

IV. Overview of the Moment-Based Prediction Algorithm

In this Section, an overview of the basic concepts of the prediction step with exponential densities are introduced. In Subsection IV-A, the calculation of the moments of the exact predicted density is described.

In Subsection IV-B, the differential relation between moment and parameter variations, mentioned in (17), is discussed.

A. Calculation of the Moments of the Exact Prediction

If the moments of the density $f^x_k\left(x_k\right)$ of the uncertain state variable x_k and the moments of $f^v_k\left(\underline{v}_k\right)$ describing the uncertainties $\underline{v}_k$ are known, the moments of the predicted random variable x_{k+1} can easily be calculated as a linear combination of the prior density's moments. This holds for polynomial, trigonometric, and exponential functions with polynomial exponents as state equations $a_k\left(x_k,\underline{v}_k\right)$.

In the following it is assumed that the state equation

$$x_{k+1}=a_k\left(x_k,\underline{v}_k\right)=a_{1,k}\left(x_k,\underline{v}_{1,k}\right)+v_{2,k}$$

can be separated into a nonlinear function $a_{1,k}\left(x_k,\underline{v}_{1,k}\right)$ of the prior state variable with non-additive uncertainties $\underline{v}_{1,k}\in\mathbb{R}^{v-1}$ and additive noise $v_{2,k}\in\mathbb{R}$ with $\underline{v}_k=\left[\underline{v}^T_{1,k}\ ;\ v_{2,k}\right]^T$.

According to [9] and generalization of equation (2), the exact prediction is given by

$$\begin{aligned}f^x_{k+1}\left(x_{k+1}\right)=&\int_{\mathbb{R}^{v-1}}\int_{\mathbb{R}}\left\{f^x_k\left(x_k\right)f^{v_1}_k\left(\underline{v}_{1,k}\right)\right.\\ &\left.\cdot f^{v_2}_k\left(x_{k+1}-a_{1,k}\left(x_k,\underline{v}_{1,k}\right)\right)\right\}dx_k d\underline{v}_{1,k}\ .\end{aligned} \tag{18}$$

Substituting (18) into the definition of the i-th moment

$$\begin{aligned}M^x_{k+1,i}&=\int_{\mathbb{R}}x^i_{k+1}f^x_{k+1}\left(x_{k+1}\right)dx_{k+1}\\ &=\int_{\mathbb{R}}\int_{\mathbb{R}^{v-1}}\int_{\mathbb{R}}\left\{\left(a_{1,k}\left(x_k,\underline{v}_{1,k}\right)+v_{2,k}\right)^j\right.\\ &\quad\left.\cdot f^x_k\left(x_k\right)\,f^{v_1}_k\left(\underline{v}_{1,k}\right)f^{v_2}_k\left(v_{2,k}\right)\right\}dx_k d\underline{v}_{1,k}dv_{2,k}\end{aligned} \tag{19}$$

with $\frac{\partial x_{k+1}}{\partial v_{2,k}} = 1$ leads to the moments of the predicted density in terms of the moments of $f_k^x(x_k)$ and $f_k^v(\underline{v}_k)$ for the above mentioned types of nonlinear state equations for which closed form solutions exist.

B. Differential Relation between Moments and Parameter Variations of Exponential Densities

As described in Section III, the Kullback-Leibler distance $G\left(\tilde{f}_{k+1}^x \| f_{k+1}^x\right)$ between the exact predicted probability density and its approximation by an exponential density is minimized by a parameter vector $\underline{\eta}_{k+1}^x$ which minimizes the distance between the moments of $\tilde{f}_{k+1}^x(x_{k+1})$ and $f_{k+1}^x\left(\underline{\eta}_{k+1}^x, x_{k+1}\right)$. In general, it is difficult to determine the parameter vector $\underline{\eta}_{k+1}^x$ corresponding to the *global* minimum of G if only the desired moments $\underline{\tilde{M}}_{k+1,[1:2n]}^x$ of the predicted density are given and no prior knowledge about the parameters is available.

This approximation problem is simplified significantly, if parameters of a density are to be determined which is described by moments that are only slightly different from an initial density with known moments and parameters. Introducing an artificial "time" variable $\gamma \in [0\ ;\ 1]$, the infinitesimal relation between the moment variation $\frac{\partial \underline{\tilde{M}}_{k+1,[1:2n]}^x}{\partial \gamma}$ and the corresponding parameter variation $\frac{\partial \underline{\eta}_{k+1}^x}{\partial \gamma}$ is given by

$$\frac{\partial \underline{M}_{k+1,[1:2n]}^x}{\partial \underline{\eta}_{k+1}^x} \frac{\partial \underline{\eta}_{k+1}^x}{\partial \gamma} = \frac{\partial \underline{\tilde{M}}_{k+1,[1:2n]}^x}{\partial \gamma} \quad . \tag{20}$$

This differential equation is solved successfully if continuous parameter variations exist which correspond to the given continuous moment variations $\frac{\partial \underline{\tilde{M}}_{k+1,[1:2n]}^x}{\partial \gamma}$. Mathematically, a unique vector of infinitesimal parameter variations $\frac{\partial \underline{\eta}_{k+1}^x}{\partial \gamma}$ can be calculated for nonsingular Jacobians $\mathbf{J} = \frac{\partial \underline{M}_{k+1,[1:2n]}^x}{\partial \underline{\eta}_{k+1}^x}$. In this paper, for non-singular Jacobians the ordinary differential equation

$$\frac{\partial \underline{\eta}_{k+1}^x}{\partial \gamma} = \mathbf{J}^{-1} \frac{\partial \underline{\tilde{M}}_{k+1,[1:2n]}^x}{\partial \gamma} \tag{21}$$

is solved for the parameter interval $\gamma \in [0; 1]$. During the variation of γ the state equation is transferred, e. g. from a unit mapping function $a_k(\gamma = 0, x_k, \underline{v}_k) = x_k$, to the desired nonlinear state equation for $\gamma = 1$ according to

$$a_k(\gamma, x_k, \underline{v}_k) = (1-\gamma)\, x_k + \gamma a_k(x_k, \underline{v}_k) \quad . \tag{22}$$

The remainder of this paper considers a proposed solution to the three problems arising from this approach. How can the parameter dependent entries of the Jacobian, i. e., the moments of the approximated density be calculated efficiently? What can be done if the matrix $\mathbf{J}$ is (nearly) singular? How can the approximation quality and the computational effort be influenced by the user of this prediction algorithm?

V. Prediction Algorithm for Exponential Densities

This Section is concerned with the implementation of the moment-based prediction algorithm and the questions at the end of Section IV. The basics of the moment calculation for exponential densities by a system of differential algebraic equations are described in Subsection V-A.

In Subsection V-B, the solution of the differential equations (21) between moment and parameter variations is discussed and methods for the control of the approximation quality of the prediction algorithm are proposed. In Fig. 1, a block diagram summarizing the proposed algorithm is shown.

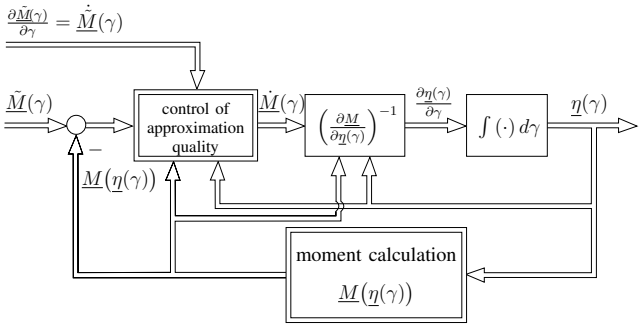

Fig. 1. Overview of the proposed moment-based prediction step.

A. Moment Calculation of Exponential Densities

The scope of this subsection is the review of the basic idea of moment calculation of exponential densities by a set of differential algebraic equations, see also [7]. At the same time, the relation between moment calculation and adaptation of density parameters (21) via the diffeomorphism (17) should become clear.

It is assumed that the moments $\underline{M}_{1,[0:2n-1]}^x\left(\underline{\eta}_1^x\right)$ of an, in general unnormalized, exponential density function $f_1^x\left(\underline{\eta}_1^x, x\right)$ of order $2n$ are given. The goal of the differential algebraic approach for moment calculation is to determine the moments $\underline{M}_{2,[0:\zeta]}^x\left(\underline{\eta}_2^x\right)$ belonging to another exponential density $f_2^x\left(\underline{\eta}_2^x, x\right)$ up to an *arbitrary* order ζ.

1) Definition of Continuous Parameter Variations: In a first step, a continuous variable $\tau \in [0; 1]$ is introduced, which transfers the initial parameter vector $\underline{\eta}_1^x = \left[\eta_{1,0}^x \quad \eta_{1,1}^x \quad \ldots \quad -\left(\eta_{1,2n}^x\right)^2\right]$ into the final parameter vector $\underline{\eta}_2^x = \left[\eta_{2,0}^x \quad \eta_{2,1}^x \quad \ldots \quad -\left(\eta_{2,2n}^x\right)^2\right]$ according to

$$\underline{\eta}_{12}^x(\tau) = (1-\tau)\,\underline{\eta}_1^x + \tau \underline{\eta}_2^x = \underline{\eta}_1^x + \tau\left(\underline{\eta}_2^x - \underline{\eta}_1^x\right) \quad . \tag{23}$$

For normalized densities $f_1^x\left(\underline{\eta}_1^x, x\right)$, the parameter $\eta_{1,0}^x = \ln(c)$, where c is the density's normalization factor.

2) Algebraic Constraints between Lower-Order Moments and Higher-Order Moments of Exponential Densities with Polynomial Exponents: The i-th order moment of the density $f_{12}^x\left(\underline{\eta}_{12}^x(\tau), x\right)$ is defined by the integral

$$M_i^x\left(\underline{\eta}_{12}^x(\tau)\right) = \int_{-\infty}^{\infty} x^i f_{12}^x\left(\underline{\eta}_{12}^x(\tau), x\right) dx \quad . \tag{24}$$

Integration by parts of (24) with respect to x yields

$$M_i^x(\underline{\eta}_{12}^x(\tau)) = -\frac{1}{i+1}\sum_{j=1}^{2n}\alpha_j M_{i+j}^x(\underline{\eta}_{12}^x(\tau)) \tag{25}$$

with

$$\alpha_j = \begin{cases} j\left((1-\tau)\,\eta_{1,j}^x + \tau\eta_{2,j}^x\right) & \text{for } 1 \le j \le 2n-1 \\ -2n\left((1-\tau)\left(\eta_{1,2n}^x\right)^2 + \tau\left(\eta_{2,2n}^x\right)^2\right) & \text{for } j = 2n. \end{cases}$$

relating lower-order moments up to order $2n-1$ to a linear combination of higher-order moments up to order $4n-1$ [6].

Example V.1 For $2n = 4$ and $i = 0, \ldots, 3$, the algebraic constraints (25) are re-written in matrix-vector-notation:

$$\begin{aligned} &-\begin{bmatrix} 1 & \frac{\alpha_1}{1} & \frac{\alpha_2}{1} & \frac{\alpha_3}{1} \\ 0 & 1 & \frac{\alpha_1}{2} & \frac{\alpha_2}{2} \\ 0 & 0 & 1 & \frac{\alpha_1}{3} \\ 0 & 0 & 0 & 1 \end{bmatrix} \underline{M}_{[0:2n-1]}^x\left(\underline{\eta}_{12}^x(\tau)\right) \\ &= \begin{bmatrix} \frac{\alpha_4}{1} & 0 & 0 & 0 \\ \frac{\alpha_3}{2} & \frac{\alpha_4}{2} & 0 & 0 \\ \frac{\alpha_2}{3} & \frac{\alpha_3}{3} & \frac{\alpha_4}{3} & 0 \\ \frac{\alpha_1}{4} & \frac{\alpha_2}{4} & \frac{\alpha_3}{4} & \frac{\alpha_4}{4} \end{bmatrix} \underline{M}_{[2n:4n-1]}^x\left(\underline{\eta}_{12}^x(\tau)\right) \ . \end{aligned} \tag{26}$$

Arbitrary higher-order moments $\underline{M}_{[2n:4n-1]}^x\left(\underline{\eta}_{12}^x(\tau)\right)$ can obviously be calculated analytically, for known lower-order moments $\underline{M}_{[0:2n-1]}^x\left(\underline{\eta}_{12}^x(\tau)\right)$ up to order $2n-1$. ■

3) Set of Differential Equations: Now, variations of the moments $M_i^x\left(\underline{\eta}_{12}^x(\tau)\right)$, $i = 0, \ldots, 2n-1$, are expressed in terms of moments up to order $i+2n$ for $\tau \in [0;1]$. This set of under-determined ordinary differential equations

$$\frac{\partial M_i^x\left(\underline{\eta}_{12}^x(\tau)\right)}{\partial\tau} = \sum_{j=0}^{2n}\beta_j M_{i+j}^x\left(\underline{\eta}_{12}^x(\tau)\right) \tag{27}$$

with

$$\beta_j = \begin{cases} \eta_{2,j}^x - \eta_{1,j}^x & \text{for } 0 \le j \le 2n-1 \\ -\left(\eta_{2,2n}^x\right)^2 + \left(\eta_{1,2n}^x\right)^2 & \text{for } j = 2n \end{cases}$$

can be solved under consideration of the additional algebraic constraints (25). Especially, if only small variations of the density parameters are of interest, this method is much more efficient than moment calculation by numerical integration of the density function.

B. Solution of the Differential Equations of the Parameter Variations and Control of the Approximation Quality

After summarizing a procedure for the calculation of the moments of an exponential density up to an arbitrary order, the implementation of the prediction step is described in this Subsection. The Jacobian $\mathbf{J} = \frac{\partial \underline{M}_{k+1,[1:2n]}^x}{\partial \underline{\eta}_{k+1}^x}$ in the differential equation (21) has to be calculated explicitly.

For that purpose, the derivative of $M_{k+1,[1:2n]}^x$ with respect to the parameter vector $\underline{\eta}_{k+1}^x$ is calculated. Using the results of equations (13) and (14), the components

$$\frac{\partial \underline{M}_{k+1,[1:2n]}^x}{\partial \underline{\eta}_{k+1}^x} = \frac{\partial}{\partial \underline{\eta}_{k+1}^x}\int_{-\infty}^{\infty} x_{k+1}^i f_{k+1}^x\left(\underline{\eta}_{k+1}^x, x_{k+1}\right) dx_{k+1} \tag{28}$$

of the Jacobian

$$\begin{aligned} \mathbf{J} = &-M_{k+1,i}^x\begin{bmatrix} M_{k+1,1} \\ \vdots \\ -2\eta_{k+1,2n}^x M_{k+1,2n} \end{bmatrix} \\ &+\begin{bmatrix} M_{k+1,i+1} \\ \vdots \\ -2\eta_{k+1,i+2n}^x M_{k+1,2n} \end{bmatrix} \end{aligned} \tag{29}$$

are given in terms of the moments of $f_{k+1}^x\left(\underline{\eta}_{k+1}^x, x_{k+1}\right)$ which are determined by the differential algebraic approach in Subsection V-A.

If the differential algebraic approach for moment calculation is applied to the unnormalized density

$$f_{k+1}^{*,x} = \exp\left(\eta_{k+1,1}^x x_{k+1} + \ldots - \left(\eta_{k+1,2n}^x\right)^2 x_{k+1}^{2n}\right)$$

the normalization factor $c\left(\underline{\eta}_{k+1}^x\right)$ is given by $\frac{1}{M_{k+1,0}^{*,x}}$, i.e.,

$$\underline{M}_{k+1}^x = \frac{1}{M_{k+1,0}^{*,x}}\underline{M}_{k+1}^{*,x} \ .$$

During the solution of the differential equation (21), inversion of the Jacobian may be prevented by singularities of this matrix. In practical applications, this problem can be avoided by adding a regularization matrix $\lambda\mathbf{I}$ to the Jacobian matrix, where $\mathbf{I}$ is an identity matrix of appropriate dimension and λ is a small real number.

If the deviation between the true and the approximated moments is exceeding a user-defined bound, several strategies are possible to control the approximation quality. For a fixed value of γ, the differential equation (21) can be used to reduce the difference between the true moments and the approximation. If the approximation error cannot be reduced sufficiently, not all given moments can be approximated simultaneously by the assumed number of density parameters. Then, not all $2n$ moment conditions can be minimized. Alternatively, an under-determined set of differential equations, described by a rectangular Jacobian matrix, can be used to determine the minimum-norm parameter variations relying on the pseudo-inverse of the Jacobian matrix. The differential equation (21) is also used to reduce or increase the number of density parameters, with only small changes of the moments. E.g. the reduction from order $2n$ to $2n-2$ can be calculated by

$$\frac{\partial \underline{M}_{[1:2n]}^x}{\partial \underline{\eta}_{[1:2n-2]}^x}\frac{\partial \underline{\eta}_{[1:2n-2]}^x}{\partial\gamma} = -\frac{\partial \underline{M}_{[1:2n]}^x}{\partial \underline{\eta}_{[2n-1:2n]}^x}\frac{\partial \underline{\eta}_{[2n-1:2n]}^x}{\partial\gamma} \tag{30}$$

with the given parameter variations $\frac{\partial \underline{\eta}_{[2n-1:2n]}^x}{\partial\gamma}$, which are chosen such that $\eta_{2n-1}^x(\gamma=1) = \eta_{2n}^x(\gamma=1) = 0$. Similarly, this idea can be applied to increase the absolute value of the parameter η_{2n}^x to increase the distance to the singularity at $\eta_{2n}^x = 0$ and $\eta_{2n-1}^x \neq 0$ or to increase the order of the density from $2n$ to $2n+2$.

VI. Simulation Results for Recursive Prediction

In this Section, simulation results for a nonlinear, recursive prediction step with the discrete-time state equation

$$x_{k+1} = 2x_k - 0.5x_k^3 + v_k \tag{31}$$

are given. Fig. 2 shows the state equation $a_k(x_k)$ as well as the unit mapping $x_{k+1} = x_k$. The linear function and $a_k(x_k)$ have three points in common. Two stable equilibria exist for $x_k = \pm\sqrt{2}$. For $x_k = 0$, an instable equilibrium exists. In recursive prediction, the stable equilibria correspond to the maxima of the predicted density functions.

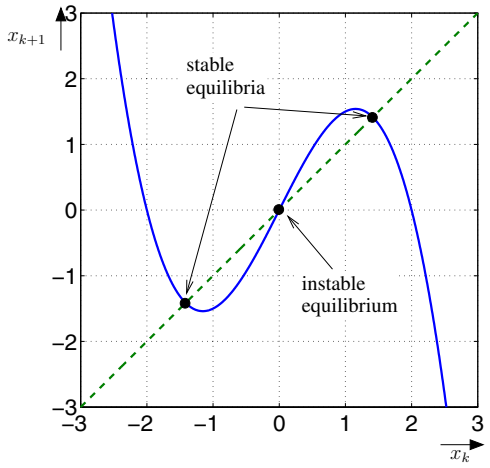

Fig. 2. Stable and instable equilibria of the nonlinear system state equation.

The initial exponential density function of the uncertain state variable x_0 is described by the parameter vector

$$\underline{\eta}_0^x = \begin{bmatrix} 0.2 & 0.05 & 0 & 0 & 0 & 0 & 0 & \sqrt{1.5} \end{bmatrix}^T$$

with $2n = 8$. The time-invariant system noise v_k is assumed to be Gaussian with $\mu_k^v = 0.0$ and $\sigma_k^v = 0.15$. In Fig. 3, the resulting densities are depicted for 15 subsequent time steps. These are compared to a numerical integration of (2). The predicted parameter vector after the first prediction step is

$$\begin{aligned} \underline{\eta}_1^x \approx \; & [0.112 \quad 0.245 \quad 0.010 \quad -0.095 \\ & -0.011 \quad 0.160 \quad 0.0028 \quad 0.286]^T \end{aligned}$$

The deviations of the moments in the same time-step are shown in Fig. 4.

VII. Conclusions

In this paper, an efficient moment-based approach for a Bayesian prediction step for linear and nonlinear discrete-time systems has been proposed in order to minimize the deviation between the exact solution and a parameterizable exponential density. This novel approach uses a set of ordinary differential equations to calculate the variation of the density parameters due to given moment variations. As demonstrated in simulations the approach can be applied successfully for prediction using exponential densities. Furthermore, the computational effort for moment calculation, on which this algorithm relies, is significantly reduced by the use of a set of differential algebraic equations compared to numerical integration of the density function.

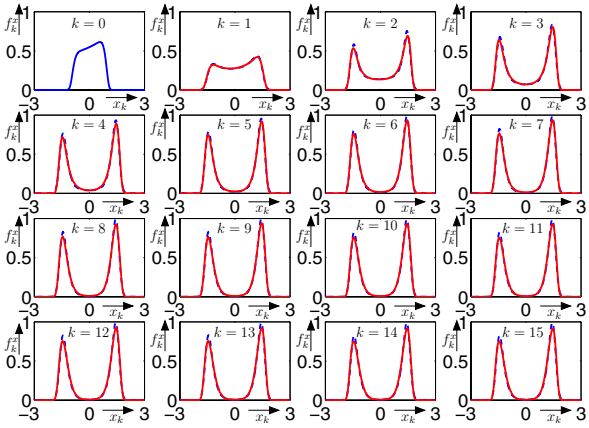

Fig. 3. Approximated (solid lines) and exact predicted probability density functions (dashed lines) for a recursive, nonlinear prediction step with additive Gaussian noise.

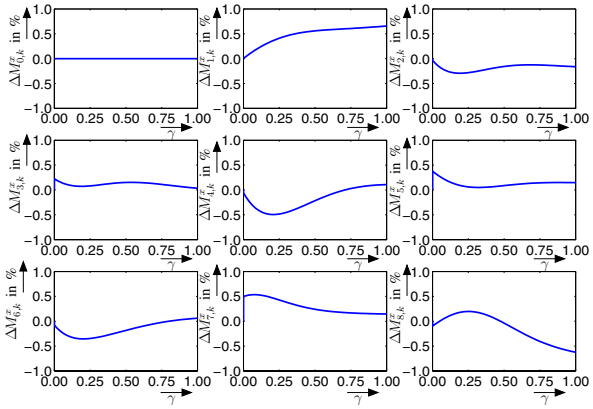

Fig. 4. Relative deviation (in %) between the moments of the exact predicted density and the approximation.

References

[1] A. Budhiraja and H. Kushner, "A Nonlinear Filtering Algorithm Based on an Approximation of the Conditional Distribution," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 580–585, March 2000.

[2] A. Doucet, S. Godsill, and C. Andrieu, "On Sequential Monte Carlo Sampling Methods for Bayesian Filtering," *Statistics and Computing*, vol. 10, pp. 197–208, 2000.

[3] H. Tanizaki, "Nonlinear and Non-Gaussian State Space Modeling Using Sampling Techniques," *Annals of the Institute of Statistical Mathematics*, vol. 53, no. 1, pp. 63–81, 2001.

[4] S. Kullback and R. A. Leibler, "On Information and Sufficiency," *Annals of Mathematical Statistics*, vol. 22, pp. 79–86, 1951.

[5] D. Brigo, B. Hanzon, and F. L. Gland, "A Differential Geometric Approach to Nonlinear Filtering: the Projection Filter," *IEEE Transactions on Automatic Control*, vol. 43, no. 2, pp. 247–252, February 1998.

[6] ——, "Approximate Nonlinear Filtering by Projection on Exponential Manifolds of Densities," *Bernoulli*, vol. 5, no. 3, pp. 495–534, 1999.

[7] 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.

[8] R. Kulhavý, "Recursive Nonlinear Estimation: Geometry of a Space of Posterior Densities," *Automatica*, vol. 28, no. 2, pp. 313–323, 1992.

[9] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. Tokyo: McGraw-Hill, 1965.