Calculating Moments of Exponential Densities Using Differential Algebraic Equations

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

***Abstract*—This letter introduces an efficient approach for calculating moments of exponential densities. Usually, the desired moments are obtained by means of numerical integration, which is impractical due to its computational complexity and the underlying infinite integration intervals. The new approach relies on an exact conversion of these integrals into a system of ordinary differential equations with algebraic constraints. The desired moments are then obtained by solving this system of differential algebraic equations over a finite "time" interval. The resulting procedure is simple to implement and typically reduces the computational burden by one order of magnitude.**

***Index Terms*—Differential algebraic equations (DAEs), exponential densities, moment calculation, Runge–Kutta.**

I. Introduction

Calculating moments of exponential densities is a central problem in nonlinear filtering and state estimation [1]. This letter is concerned with the efficient calculation of moments of exponential densities with *polynomial exponents* [2], [3].

Analytic expressions for the required moments are only available for a few special cases, e.g., Gaussian densities [4]–[6]. Hence, standard approaches to calculating moments of exponential densities rely on numerical integration techniques. Numerical integration is of high computational complexity and is impractical for higher dimensional problems. Furthermore, moments are defined by infinite integration intervals. This introduces additional difficulties, as the support of an exponential density in general cannot be simply deduced from its parameters.

The new approach converts the set of infinite integrals into a system of ordinary differential equations with algebraic constraints. The desired moments are obtained by solving the system of differential algebraic equations (DAEs) over a *finite* interval with a suitable vector of initial values. The implementation of the new approach is simple and typically reduces the computational burden by one order of magnitude. It is especially efficient for the simultaneous calculation of several moments of different orders for the same set of density parameters. In contrast to numerical integration, the computational complexity of the new approach is mainly influenced by the number of density parameters.

In Section II, the moment calculation problem is formulated. The new solution is derived in Section III, where attention is restricted to the case of one-dimensional (1-D) densities. Extension to the vector case is straightforward. Moment variations caused by small parameter modifications are discussed in Section III-A. For large parameter modifications, a DAE-based approach is derived in Section III-B. An efficient solution of the DAE system is given in Section IV. Its performance is demonstrated by means of an example. Finally, conclusions are given in Section V.

II. Problem Formulation

We consider unnormalized 1-D exponential densities of the form

$$f(x) = \exp\left(\eta_0 + \eta_1 x^1 + \cdots + \eta_{2n} x^{2n}\right) = \exp\left(\underline{\eta}^T \underline{x}\right)$$

characterized by the parameter vector $\underline{\eta} = [\eta_0, \eta_1, \ldots, \eta_{2n}]^T$ with $\underline{x} = [1, x, x^2, \ldots, x^{2n}]^T$ and $n \in \mathbb{N}$. A negative parameter η_{2n} guarantees finite moments. Moments of order i of the density are defined by

$$M_i\left(\underline{\eta}\right) = \int_{-\infty}^{\infty} x^i f(x) dx. \tag{1}$$

For a given parameter vector denoted by $\underline{\eta}_1$ we desire to calculate arbitrary moments $M_i(\underline{\eta}_1)$ for $i \in \mathbb{N} \cup \{0\}$.

Analytic expressions for the moments of exponential densities are only available for some special cases such as Gaussian densities

$$f(x) = \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) = \exp\left(\eta_0 + \eta_1 x + \eta_2 x^2\right)$$

with $\eta_0 = -\mu^2/(2\sigma^2)$, $\eta_1 = \mu/\sigma^2$, and $\eta_2 = -1/(2\sigma^2)$. Hence, the moments corresponding to a general parameter vector $\underline{\eta}_1$ must be calculated numerically.

III. Efficient Moment Calculation

The new approach is based on calculating moment variations caused by modification of the parameter vector. For that purpose, we assume another parameter vector $\underline{\eta}_0$ with known corresponding moments to be available. When $\underline{\eta}_1$ is close to $\underline{\eta}_0$, the moments corresponding to $\underline{\eta}_1$ can be expressed in terms of moments corresponding to $\underline{\eta}_0$ by means of a linear perturbation analysis.

Manuscript received February 13, 2002; revised August 29, 2002. This work was performed while the authors were with the Institute of Automatic Control Engineering, Technische Universität München, München, Germany. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Dimitris A. Pados.

A. Rauh is with the Department of Measurement, Control, and Microtechnology, University of Ulm, 89069 Ulm, Germany (e-mail: andreas.rauh@e-technik.uni-ulm.de).

U. D. Hanebeck is with the Institute of Computer Design and Fault Tolerance, University of Karlsruhe, 76128 Karlsruhe, Germany (e-mail: Uwe.Hanebeck@ieee.org).

Digital Object Identifier 10.1109/LSP.2003.810022

In the case of large parameter modifications, the key idea is to construct a system of ordinary differential equations describing the moment variation. Solving this system over a finite interval with the moments $M_i(\underline{\eta}_0)$, $i = 0, \ldots, 2n-1$, as initial values gives the desired moments $M_i(\underline{\eta}_1)$, $i = 0, \ldots, 2n-1$. However, the system of differential equations for calculating the moments $M_i(\underline{\eta}_1)$ requires higher order moments. These are obtained from additional algebraic constraints that relate the higher order moments to the set of desired moments.

A. Small Parameter Modifications

When the parameter modification $\Delta\underline{\eta} = \underline{\eta}_1 - \underline{\eta}_0$ is small, the moments corresponding to $\underline{\eta}_1$ can be approximated by

$$M_i\left(\underline{\eta}_1\right) = M_i\left(\underline{\eta}_0 + \Delta\underline{\eta}\right) = \int_{-\infty}^{\infty} x^i f\left(x, \underline{\eta}_0 + \Delta\underline{\eta}\right) dx$$
$$= M_i\left(\underline{\eta}_0\right) + \Delta M_i \approx M_i\left(\underline{\eta}_0\right) + \left[\left.\frac{\partial M_i}{\partial \underline{\eta}}\right|_{\underline{\eta}=\underline{\eta}_0}\right]^T \Delta\underline{\eta}.$$

The partial derivative of the ith moment (1) with respect to the density parameter vector $\underline{\eta}$ is expressed in terms of moments of up to order $i + 2n$ according to

$$\frac{\partial M_i}{\partial \underline{\eta}} = \int_{-\infty}^{\infty} x^i \frac{\partial \exp\left(\sum_{j=0}^{2n} \eta_j x^j\right)}{\partial \underline{\eta}} dx$$
$$= \int_{-\infty}^{\infty} x^i \begin{bmatrix} 1 \\ x \\ \vdots \\ x^{2n} \end{bmatrix} \exp\left(\sum_{j=0}^{2n} \eta_j x^j\right) dx$$
$$= \left[M_i \quad M_{i+1} \quad \cdots \quad M_{i+2n}\right]^T. \tag{2}$$

B. Moments of Arbitrary Exponential Densities

Now large scale modifications of the parameter vector are considered. Continuous variations of the corresponding moments $M_i(\underline{\eta}(\gamma))$ are achieved by a continuous modification of the parameter vector according to $\underline{\eta}(\gamma) = \underline{\eta}_0 + \gamma(\underline{\eta}_1 - \underline{\eta}_0)$. For that purpose, a scalar "time" variable $\gamma \in [0; 1]$ is defined.

1) System of Ordinary Differential Equations: Ordinary differential equations for the moment variation are derived by calculating the partial derivative $\dot{M}_i = \partial/\partial\gamma M_i(\underline{\eta}(\gamma))$ of the ith-order moment with respect to γ. With (2) we obtain

$$\frac{\partial M_i(\underline{\eta}(\gamma))}{\partial \gamma} = \left[\left.\frac{\partial M_i}{\partial \underline{\eta}}\right|_{\underline{\eta}=\underline{\eta}(\gamma)}\right]^T \frac{\partial \underline{\eta}(\gamma)}{\partial \gamma}$$
$$= \left[M_i\left(\underline{\eta}(\gamma)\right) \quad M_{i+1}\left(\underline{\eta}(\gamma)\right) \quad \cdots \quad M_{i+2n}\left(\underline{\eta}(\gamma)\right)\right] \cdot \left(\underline{\eta}_1 - \underline{\eta}_0\right) \tag{3}$$

which relates moment variations to moments of up to order $i + 2n$.

Appropriate initial parameter vectors $\underline{\eta}_0$ are characterized by the fact that the moments of the corresponding density are given. A convenient parameter vector $\underline{\eta}_0$ corresponding to a Gaussian density yields a time-variant probability density

$$f(x, \gamma) = \exp\left(-(1-\gamma)\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 + \gamma\, \underline{\eta}_1^T \underline{x}\right).$$

Without loss of generality, a standard normal density, i.e., $\mu = 0$ and $\sigma = 1$, corresponding to

$$\underline{\eta}_0 = \left[0 \quad 0 \quad -\tfrac{1}{2} \quad 0 \quad \ldots\right]^T \tag{4}$$

is used to simplify the following expressions. With (3) we obtain

$$\frac{\partial M_i\left(\underline{\eta}(\gamma)\right)}{\partial \gamma} = \frac{1}{2} M_{i+2}\left(\underline{\eta}(\gamma)\right) + \sum_{k=0}^{2n} \eta_k M_{i+k}\left(\underline{\eta}(\gamma)\right)$$

where η_k, $k = 0, \ldots, 2n$ from now on denote the components of the parameter vector $\underline{\eta}_1$.

Example III.1: For $n = 2$, the following system of differential equations is obtained:

$$\underline{\dot{M}}^{(\gamma)}_{(0:2n-1)} = \overbrace{\begin{bmatrix} \eta_0 & \eta_1 & \eta_2 + \frac{1}{2} & \eta_3 \\ 0 & \eta_0 & \eta_1 & \eta_2 + \frac{1}{2} \\ 0 & 0 & \eta_0 & \eta_1 \\ 0 & 0 & 0 & \eta_0 \end{bmatrix}}^{\mathbf{D}_{(0:2n-1,0:2n-1)}} \underline{M}^{(\gamma)}_{(0:2n-1)} + \underbrace{\begin{bmatrix} \eta_4 & 0 & 0 & 0 \\ \eta_3 & \eta_4 & 0 & 0 \\ \eta_2 + \frac{1}{2} & \eta_3 & \eta_4 & 0 \\ \eta_1 & \eta_2 + \frac{1}{2} & \eta_3 & \eta_4 \end{bmatrix}}_{\mathbf{D}_{(0:2n-1,2n:4n-1)}} \underline{M}^{(\gamma)}_{(2n:4n-1)} \tag{5}$$

where $\underline{M}^{(\gamma)}_{(i:j)}$ is the vector of moments of orders i to j depending on γ. The variations of the desired $2n$ moments of orders zero to $2n-1$ depend on moments of up to order $4n-1$ that are split up into lower order moments $\underline{M}_L = \underline{M}^{(\gamma)}_{(0:2n-1)} = [M_0(\underline{\eta}(\gamma)), \ldots, M_{2n-1}(\underline{\eta}(\gamma))]^T$ and higher order moments $\underline{M}_H = \underline{M}^{(\gamma)}_{(2n:4n-1)} = [M_{2n}(\underline{\eta}(\gamma)), \ldots, M_{4n-1}(\underline{\eta}(\gamma))]^T$. The coefficient matrices $\mathbf{D}_{()}$ corresponding to the respective moment vectors do not depend on γ. ■

2) System of Algebraic Constraints: The system of differential equations (5) for calculating the lower order moments $\underline{M}_L$ also involves higher order moments $\underline{M}_H$. Hence, additional algebraic constraints are required that are obtained by integration by parts of (1) with respect to x

$$M_i\left(\underline{\eta}(\gamma)\right) = \underbrace{\left.\left(\frac{x^{i+1}}{i+1} f(x,\gamma)\right)\right|_{-\infty}^{\infty}}_{=0} - \int_{-\infty}^{\infty} \frac{x^{i+1}}{i+1} \frac{\partial f(x,\gamma)}{\partial x} dx$$
$$= -\int_{-\infty}^{\infty} \frac{x^{i+1}}{i+1}\left(-(1-\gamma)x + \gamma \sum_{k=1}^{2n} k\, \eta_k x^{k-1}\right) \cdot f(x,\gamma) dx$$
$$= \frac{1-\gamma}{i+1} M_{i+2}(\underline{\eta}(\gamma)) - \frac{\gamma}{i+1} \sum_{k=1}^{2n} k\, \eta_k M_{i+k}\left(\underline{\eta}(\gamma)\right).$$

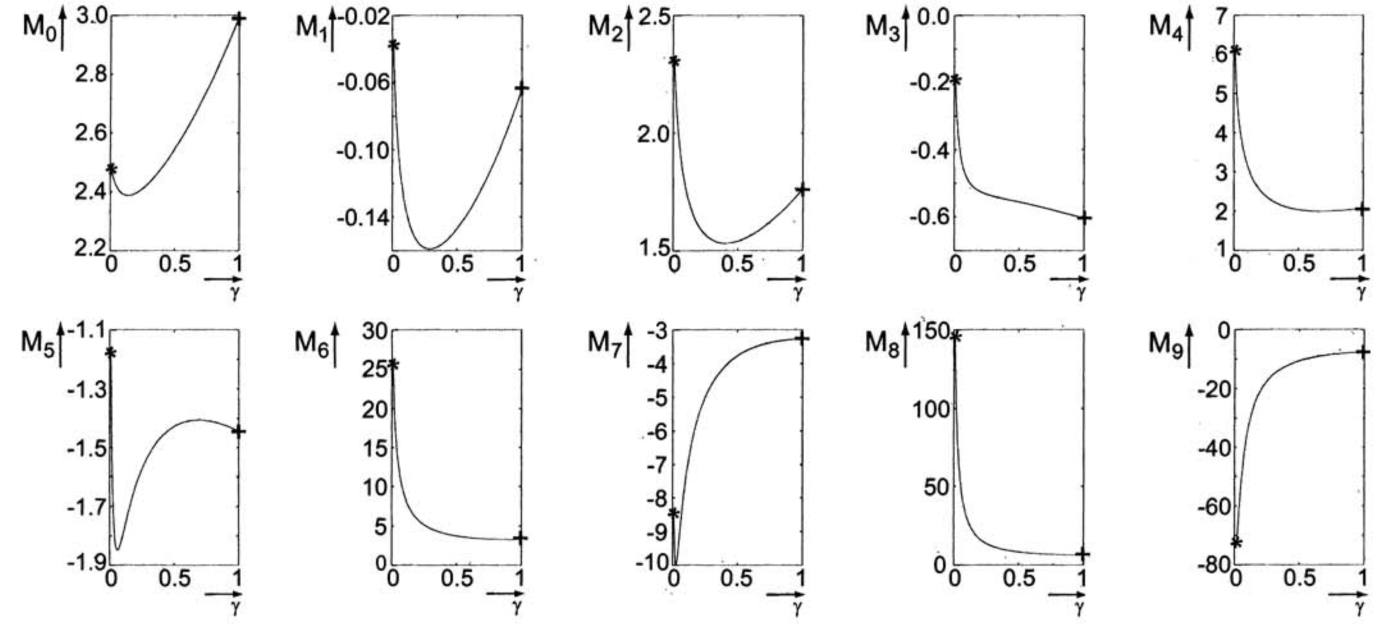

Fig. 1. Trajectories of the moments $M_0, \ldots, M_9$ calculated by the proposed DAE approach with initial moments marked by "*" and final moments marked by "+".

Example III.2: For $n = 2$, the following system

$$-\underbrace{\begin{bmatrix} 1 & \frac{\gamma\eta_1}{1} & \frac{2\gamma\eta_2-(1-\gamma)}{1} & \frac{3\gamma\eta_3}{1} \\ 0 & 1 & \frac{\gamma\eta_1}{2} & \frac{2\gamma\eta_2-(1-\gamma)}{2} \\ 0 & 0 & 1 & \frac{\gamma\eta_1}{3} \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{A}^{(\gamma)}_{(0:2n-1,0:2n-1)}} \underline{M}^{(\gamma)}_{(0:2n-1)} = \underbrace{\begin{bmatrix} \frac{4\gamma\eta_4}{1} & 0 & 0 & 0 \\ \frac{3\gamma\eta_3}{2} & \frac{4\gamma\eta_4}{2} & 0 & 0 \\ \frac{2\gamma\eta_2-(1-\gamma)}{3} & \frac{3\gamma\eta_3}{3} & \frac{4\gamma\eta_4}{3} & 0 \\ \frac{\gamma\eta_1}{4} & \frac{2\gamma\eta_2-(1-\gamma)}{4} & \frac{3\gamma\eta_3}{4} & \frac{4\gamma\eta_4}{4} \end{bmatrix}}_{\mathbf{A}^{(\gamma)}_{(0:2n-1,2n:4n-1)}} \underline{M}^{(\gamma)}_{(2n:4n-1)} \quad (6)$$

of algebraic constraints is obtained. ■

The algebraic constraints relate higher order moments $\underline{M}_H$ to lower order moments $\underline{M}_L$. In contrast to the matrices $\mathbf{D}_{()}$ in the system of differential equations, the matrices $\mathbf{A}^{(\gamma)}_{()}$ of the algebraic constraints explicitly depend on γ. For an exponential density of order $2n$, the algebraic constraints can be used to recursively calculate all required higher order moments on the basis of the parameter vector $\underline{\eta}_1$ and the given lower order moments of up to order $2n-1$ [3].

IV. Solution of the DAE

The DAE system comprising the system of ordinary differential (5) and the system of algebraic constraints (6) is solved by taking into account that the matrix $A^{(\gamma)}_{(0:2n-1,2n:4n-1)}$ is ill-conditioned for "small" values of γ and singular for $\gamma = 0$. Hence, inversion of this matrix must be avoided for $\gamma \approx 0$.

A modified fourth-order Runge–Kutta method [7]

$$\underline{M}^{(\gamma+s_\gamma)}_{(0:2n-1)} = \underline{M}^{(\gamma)}_{(0:2n-1)} + s_\gamma \underline{m} \quad (7)$$

with step size $s_\gamma = -\Delta\gamma$ is proposed for calculating the moments $\underline{M}^{(\gamma)}_{(0:2n-1)}$ for $\gamma \in [\Delta\gamma; 1]$. The vector $\underline{m}$ is expressed as a linear combination of moments of order zero to $10n-1$

$$\underline{m} = \mathbf{N}\underline{M}^{(\gamma)}_{(0:10n-1)}$$

where the matrix $\mathbf{N}$ only depends on the density parameters $\underline{\eta}$ and not on γ. $\mathbf{N}$ is obtained by isolating $\underline{M}^{(\gamma)}_{(0:10n-1)}$ in

$$\underline{m} = \frac{1}{6}\Big[\mathbf{E}_{(2n\times 8n)}\underline{m}_1 + 2\mathbf{E}_{(2n\times 6n)}\underline{m}_2 + 2\mathbf{E}_{(2n\times 4n)}\underline{m}_3 + \mathbf{E}_{(2n\times 2n)}\underline{m}_4\Big]$$

where $\mathbf{E}_{()}$ are identity matrices of appropriate dimensions. The vectors $\underline{m}_i$, $i = 1, \ldots, 4$, are defined as

$$\begin{aligned} \underline{m}_1 &= \mathbf{D}_{(0:8n-1,0:10n-1)}\underline{M}^{(\gamma)}_{(0:10n-1)} \\ \underline{z}_1 &= \underline{M}^{(\gamma)}_{(0:8n-1)} - 0.5\Delta\gamma\mathbf{D}_{(0:8n-1,0:10n-1)}\underline{M}^{(\gamma)}_{(0:10n-1)} \\ \underline{m}_2 &= \mathbf{D}_{(0:6n-1,0:8n-1)}\underline{z}_1 \\ \underline{z}_2 &= \underline{M}^{(\gamma)}_{(0:6n-1)} - 0.5\Delta\gamma\mathbf{D}_{(0:6n-1,0:8n-1)}\underline{z}_1 \\ \underline{m}_3 &= \mathbf{D}_{(0:4n-1,0:6n-1)}\underline{z}_2 \\ \underline{z}_3 &= \underline{M}^{(\gamma)}_{(0:4n-1)} - \Delta\gamma\mathbf{D}_{(0:4n-1,0:6n-1)}\underline{z}_2 \\ \underline{m}_4 &= \mathbf{D}_{(0:2n-1,0:4n-1)}\underline{z}_3. \end{aligned}$$

Substitution of $\underline{m}$ in (7) gives

$$\underline{M}^{(\gamma-\Delta\gamma)}_{(0:2n-1)} = \underline{M}^{(\gamma)}_{(0:2n-1)} - \Delta\gamma\mathbf{N}\underline{M}^{(\gamma)}_{(0:10n-1)}. \quad (8)$$

The moments $\underline{M}^{(\gamma)}_{(0:10n-1)}$ are expressed in terms of $\underline{M}^{(\gamma)}_{(0:2n-1)}$ by means of the algebraic constraints according to

$$\underline{M}^{(\gamma)}_{(0:10n-1)} = \begin{bmatrix} \mathbf{E}_{(2n\times 2n)} \\ -\left(\mathcal{A}^{(\gamma)}\right)^{-1}\mathbf{A}^{(\gamma)}_{(0:8n-1,0:2n-1)} \end{bmatrix} \underline{M}^{(\gamma)}_{(0:2n-1)} \quad (9)$$

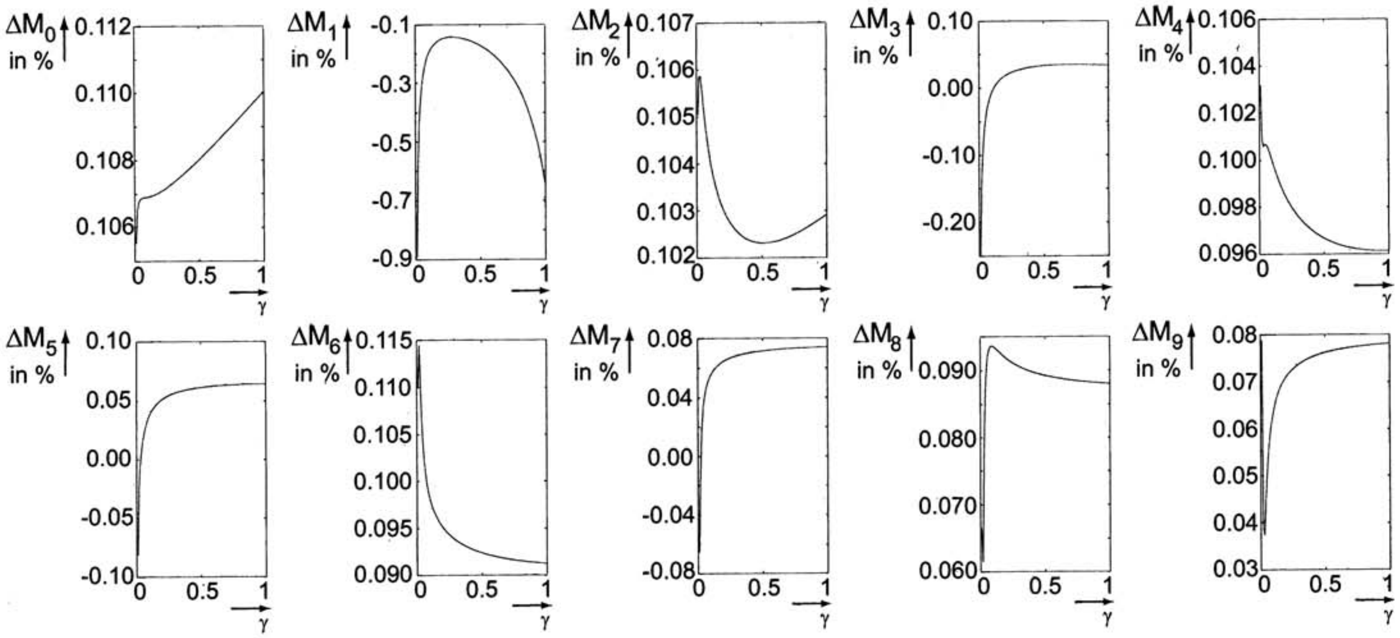

Fig. 2. Relative error (in percent) between the true moments and the moments calculated by the proposed approach.

with the abbreviation

$$\mathcal{A}^{(\gamma)} = \mathbf{A}^{(\gamma)}_{(0:8n-1,2n:10n-1)}.$$

Finally, (8) can be solved for the moments $\underline{M}^{(\gamma)}_{(0:2n-1)}$ to obtain the update rule

$$\underline{M}^{(\gamma)}_{(0:2n-1)} = \left(\mathbf{P}^{(\gamma)}\right)^{-1} \underline{M}^{(\gamma-\Delta\gamma)}_{(0:2n-1)} \qquad (10)$$

where $\mathbf{P}^{(\gamma)}$ is given by

$$\mathbf{P}^{(\gamma)} = \mathbf{E}_{(2n\times 2n)} - \Delta\gamma\mathbf{N}\begin{bmatrix} \mathbf{E}_{(2n\times 2n)} \\ -\left(\mathcal{A}^{(\gamma)}\right)^{-1}\mathbf{A}^{(\gamma)}_{(0:8n-1,0:2n-1)} \end{bmatrix}.$$

The initial moment vector is given by $\underline{M}^{(0)}_{(0:2n-1)} = \underline{M}_{(0:2n-1)}(\underline{\eta}_0)$.

For an example parameter vector $\underline{\eta}_1 = [0, 1, 1, -1, -1]$ with $n = 2$ and $\underline{\eta}_0$ in (4), the moments of up to order $2n - 1 = 3$ are calculated by (10). Some higher order moments, in this case M_4 to M_9, are calculated recursively by (9).[1] The resulting trajectories of the moments M_0 to M_9 are shown in Fig. 1. The desired moments are obtained for $\gamma = 1$, i.e., $\underline{M}_{(0:9)}(\underline{\eta}_1) = \underline{M}^{(1)}_{(0:9)}$. In this example, the relative error between the true moments and the moments calculated by the proposed approach is less than 1% (see Fig. 2).

V. Conclusion

An efficient procedure for calculating moments of exponential densities with polynomial exponents has been developed. Rather than performing numerical integration, the new approach relies on converting the original problem into a DAE system, which is solved over a finite "time" interval to obtain the desired moments.

The new procedure is simple to implement and, compared to numerical integration, typically reduces the computational burden significantly. Efficiency increases with the number of moments calculated.

When moment vectors corresponding to several parameter vectors $\underline{\eta}_1, \ldots, \underline{\eta}_L$ are required, the proposed approach can easily be generalized. For that purpose, a DAE system is constructed that modifies a convenient initial parameter vector $\underline{\eta}_0$ along a trajectory of intermediate parameter vectors $\underline{\eta}_i$ for $i = 1, \ldots, L-1$ resulting in the final parameter vector $\underline{\eta}_L$. Hence, in addition to $\underline{M}(\underline{\eta}_L)$, the solution of the modified DAE system also yields the moment vectors $\underline{M}(\underline{\eta}_i)$ for $i = 1, \ldots, L-1$.

For real-time applications, it is important that the run time of the proposed procedure does mainly depend on the problem order and not upon specific values of the parameters or the actual shape of the considered density.

[1] Example code can be found at http://www.nonlinear-filters.com.

References

[1] A. Budhiraja and H. Kushner, "A nonlinear filtering algorithm based on an approximation of the conditional distribution," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 580–585, Mar. 2000.

[2] D. Brigo, B. Hanzon, and F. L. Gland, "A differential geometric approach to nonlinear filtering: The projection filter," *IEEE Trans. Automat. Contr.*, vol. 43, pp. 247–252, Feb. 1998.

[3] ——, "Approximate nonlinear filtering by projection on exponential manifolds of densities," *Bernoulli*, vol. 5, no. 3, pp. 495–534, 1999.

[4] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Stat. Comput.*, vol. 10, pp. 197–208, 2000.

[5] H. Tanizaki, "Nonlinear and non-Gaussian state space modeling using sampling techniques," *Ann. Inst. Stat. Math.*, vol. 53, no. 1, pp. 63–81, 2001.

[6] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*. London, U.K.: Academic, 1984.

[7] W. H. Press, A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C, The Art of Scientific Computing*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992, ch. 16.