

Density Approximation Based on Dirac Mixtures with Regard to Nonlinear Estimation and Filtering

Oliver C. Schrempf, Dietrich Brunn, Uwe D. Hanebeck
Intelligent Sensor-Actuator-Systems Laboratory
Institute of Computer Science and Engineering
Universität Karlsruhe (TH), Germany
schrempf@ieee.org, brunn@ira.uka.de, uwe.hanebeck@ieee.org

Abstract—A deterministic procedure for optimal approximation of arbitrary probability density functions by means of Dirac mixtures with equal weights is proposed. The optimality of this approximation is guaranteed by minimizing the distance of the approximation from the true density. For this purpose a distance measure is required, which is in general not well defined for Dirac mixtures. Hence, a key contribution is to compare the corresponding cumulative distribution functions.

This paper concentrates on the simple and intuitive integral quadratic distance measure. For the special case of a Dirac mixture with equally weighted components, closed-form solutions for special types of densities like uniform and Gaussian densities are obtained. Closed-form solution of the given optimization problem is not possible in general. Hence, another key contribution is an efficient solution procedure for arbitrary true densities based on a homotopy continuation approach.

In contrast to standard Monte Carlo techniques like particle filters that are based on random sampling, the proposed approach is deterministic and ensures an optimal approximation with respect to a given distance measure. In addition, the number of required components (particles) can easily be deduced by application of the proposed distance measure. The resulting approximations can be used as basis for recursive nonlinear filtering mechanism alternative to Monte Carlo methods.

I. INTRODUCTION

Bayesian methods are very popular for dealing with systems suffering from uncertainties and are used in a wide range of applications. For nonlinear systems, unfortunately the processing of the probability density functions involved in the estimation procedure, typically cannot be performed exactly. This effects especially the type of density in recursive processing, which changes and increases the complexity. Hence, nonlinear estimation in general requires the approximation of the underlying true densities by means of generic density types.

In literature different types of parametric continuous densities have been proposed for approximation, including Gaussian mixtures [1], Edgeworth series expansions [2], and exponential densities [3]. Furthermore, discrete approximations are very popular. A well known approach is to represent the true density by means of a set of samples [4]. This is used by the class of particle filters [5]. Typically, the locations and weights of the particles are determined by means of Monte Carlo techniques [6], [7].

In this paper we provide a different view on such a discrete representation. The given data points are interpreted as a mixture of Dirac delta components in order to systematically

approximate an arbitrary density function. The proposed method differs from the deterministic type of particle filters in [8] as a distance measure is employed to find an optimal approximation of the true density. However, typical distance measures quantifying the distance between two densities are not well defined for the case of Dirac mixtures. Examples are the Kullback–Leibler distance [9] and integral quadratic distances between the densities. Hence, in this paper the corresponding cumulative distribution functions of the true density and its approximation are compared in order to define an optimal Dirac Mixture approximation. This can be viewed as a reversal of the procedure introduced in [10], where a distribution distance, in that case the Kolmogorov–Smirnov test statistic, is used to calculate optimal parameters of a density given observed samples.

Here, we apply the integral quadratic distance between the cumulative distributions, which is simple and intuitive. Other possible distribution distances include the Kolmogorov–Smirnov distance [11], or the Cramér–von Mises distance [12], [13]. For the special case of a Dirac mixture with equally weighted components, closed-form solutions for special types of densities like uniform and Gaussian densities are obtained. The more general case for non-equally weighted components is discussed in [14]. Since a closed-form solution of the given optimization problem is not possible in general, an efficient solution procedure for arbitrary true densities based on a homotopy continuation approach similar to the approach introduced in [15] is applied.

The approximations yielded by the approach presented in this paper can immediately be used for implementing a recursive nonlinear filter that could serve as an alternative to the popular particle filters. In contrast to a standard particle representation, the proposed approach provides an optimal approximation with respect to a given distance measure. Furthermore, the approach is deterministic, since no random numbers are involved. In addition, the number of required components (particles) can easily be deduced by taking the distance measure presented into account.

The paper is organized as follows. After the problem formulation in Section II, the conversion of the approximation problem into an equivalent optimization problem is explained in Section III. Closed-Form solutions of this optimization problem for special types of densities are given in Section IV. A general solution approach for the case of arbitrary densities is then given in Section V. Conclusions

and a few remarks about possible extensions and future work are given in Section VI.

It is important to note that this paper is restricted to the case of scalar random variables. Furthermore, the focus is on the special case of Dirac mixtures with equally weighted components. This dramatically simplifies the derivations and allows for closed-form approximations in some important cases.

II. PROBLEM FORMULATION

We consider a given density function $\tilde{f}(x)$. The goal is to approximate this density by means of a Dirac mixture given by

$$f(x, \underline{\eta}) = \sum_{i=1}^L w_i \delta(x - x_i) , \quad (1)$$

where in the context of this paper, the weighting factors w_i are assumed to be equal and given by $w_i = 1/L$. The parameter vector $\underline{\eta}$ contains the positions of the individual Dirac functions according to

$$\underline{\eta} = [x_1, x_2, \dots, x_L]^T .$$

For the remainder of this paper, it is assumed that the positions are ordered according to

$$x_1 < x_2 < \dots < x_{L-1} < x_L .$$

Our goal is to minimize a certain distance measure G between the given density $\tilde{f}(x)$ and its approximation $f(x, \underline{\eta})$. However, standard measures of deviation are not well defined for Dirac mixture densities.

III. THE OPTIMIZATION APPROACH

The first key idea is to reformulate the above approximation problem as an optimization problem by minimizing a certain distance between the true density $\tilde{f}(x)$ and its approximation $f(x)$. Instead of comparing the densities directly, which does not make sense for Dirac Delta functions, the corresponding (cumulative) distribution functions are employed for that purpose.

The distribution function corresponding to the true density $\tilde{f}(x)$ is given by

$$\tilde{F}(x) = \int_{-\infty}^x \tilde{f}(t) dt ,$$

the distribution function corresponding to the Dirac mixture approximation can be written as

$$F(x, \underline{\eta}) = \int_{-\infty}^x f(t, \underline{\eta}) dt = \sum_{i=1}^L w_i H(x - x_i) , \quad (2)$$

where $H(\cdot)$ denotes the Heaviside function defined as

$$H(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & x = 0 \\ 1, & x > 0 \end{cases} .$$

A distance measure can then given by

$$G(\underline{\eta}) = \int_{-\infty}^{\infty} \left(\tilde{F}(x) - F(x, \underline{\eta}) \right)^2 dx . \quad (3)$$

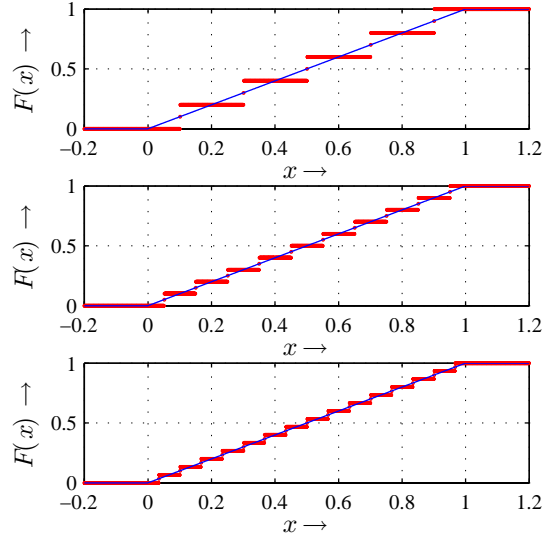


Fig. 1. Approximation of the uniform distribution for a different number of components $L = 5$, $L = 10$, and $L = 15$.

Theorem III.1 The optimal parameters x_i , $i = 1, \dots, L$ of the Dirac mixture approximation $f(x, \underline{\eta})$ according to (1) of a given density $\tilde{f}(x)$ with respect to the distance measure (3) are obtained by solving

$$\tilde{F}(x_i) = \frac{2i-1}{2L} , \quad (4)$$

for $i = 1, \dots, L$.

PROOF. The necessary condition for a minimum of the distance measure $G(\underline{\eta})$ is satisfied by the roots of the derivative of $G(\underline{\eta})$ with respect to the parameter vector $\underline{\eta}$ according to

$$\frac{\partial G(\underline{\eta})}{\partial \eta} = \underline{0} .$$

For the individual parameters x_i we obtain

$$\frac{\partial G(\underline{\eta})}{\partial x_i} = -2 \int_{-\infty}^{\infty} \left(\tilde{F}(x) - \sum_{j=1}^L w_j H(x - x_j) \right) \delta(x - x_i) dx$$

for $i = 1, \dots, L$. Using the fundamental property of the Dirac delta function gives

$$\frac{\partial G(\underline{\eta})}{\partial x_i} = -2 \left(\tilde{F}(x_i) - \sum_{j=1}^L w_j H(x_i - x_j) \right) .$$

Evaluating the Heaviside function and setting the result to zero finally gives the desired result

$$\tilde{F}(x_i) - \frac{2i-1}{2L} \stackrel{!}{=} 0$$

for $i = 1, \dots, L$. \square

IV. SOLUTION FOR SPECIAL CASES

For illustrating the usefulness of the result given in Theorem III.1, we consider two special types of densities that admit a closed-form solution for the desired parameters of the Dirac mixture approximation.

A. Special Case: Uniform Density

Without loss of generality, we consider the uniform density

$$\tilde{f}(x) = \begin{cases} 0, & x < 0 \\ 1, & 0 \leq x < 1 \\ 0, & x \geq 1 \end{cases} . \quad (5)$$

The corresponding distribution function is given by

$$\tilde{F}(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases} .$$

Hence, in (4) we have $\tilde{F}(x_i) = x_i$, and the parameters of the Dirac mixture approximation are immediately given by

$$x_i = \frac{2i-1}{2L} \quad (6)$$

for $i = 1, \dots, L$.

The true distribution $\tilde{F}(x)$ and its Dirac mixture approximation are shown for a different number of components L in Figure 1. Obviously, the Dirac mixture approximation converges to the true density for $L \rightarrow \infty$, which will be shown more formally in the next theorem.

Theorem IV.1 *The distance measure $G(\underline{\eta})$ between the uniform density $\tilde{f}(x)$ and its Dirac mixture approximation with L components is given by*

$$G(\underline{\eta}) = \frac{1}{12} \frac{1}{L^2}$$

and decreases quadratically towards zero as the number of components L increases.

PROOF. The distance measure is now given by

$$G(\underline{\eta}) = \int_0^1 \left(x - \sum_{i=1}^L \frac{1}{L} H \left(x - \frac{2i-1}{2L} \right) \right)^2 dx ,$$

which is equivalent to

$$G(\underline{\eta}) = \sum_{i=1}^{2L} \left(\int_0^{\frac{1}{2L}} x^2 dx \right) .$$

Evaluating the integral gives

$$G(\underline{\eta}) = \sum_{i=1}^{2L} \frac{1}{3} \frac{1}{(2L)^3} = \frac{1}{3} \frac{1}{(2L)^2} ,$$

which concludes the proof. \square

The expected value and the variance of the given true density $\tilde{f}(x)$ are given by

$$E_{\tilde{f}}\{x\} = \frac{1}{2}$$

and

$$E_{\tilde{f}} \left\{ \left(x - E_{\tilde{f}}\{x\} \right)^2 \right\} = \frac{1}{12} .$$

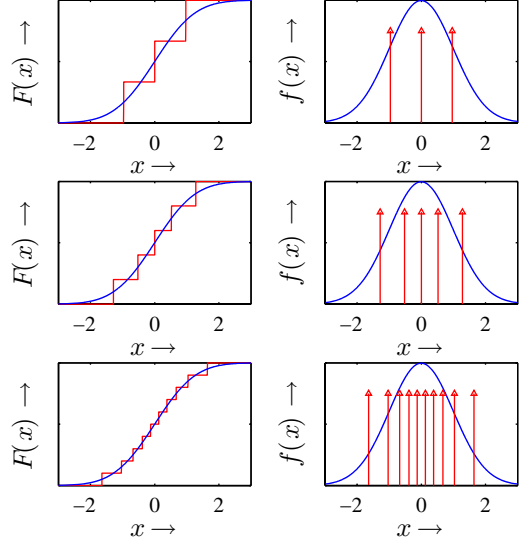


Fig. 2. Approximation of the standard normal distribution and density for a different number of components $L = 3, L = 5$ and $L = 10$.

Lemma IV.1 *The expected value of the Dirac mixture approximation of the uniform distribution is (independent of the number of components L) given by*

$$E_f\{x\} = \sum_{i=1}^L \frac{1}{L} \frac{2i-1}{2L} = \frac{1}{2} .$$

PROOF.

$$E_f\{x\} = \int_{-\infty}^{\infty} x \left(\sum_{i=1}^L w_i \delta \left(x - \frac{2i-1}{2L} \right) \right) dx ,$$

which upon using the fundamental property of the Dirac delta function gives

$$E_f\{x\} = \sum_{i=1}^L \frac{1}{L} \frac{2i-1}{2L} .$$

Simplification yields

$$E_f\{x\} = \frac{L(L+1) - L}{2L^2} = \frac{1}{2} . \quad \square$$

Hence, the expected value of the approximation density is equal to the expected value of the true density independent of the number of components.

Lemma IV.2 *The variance of the Dirac mixture approximation is given by*

$$E_f \left\{ \left(x - E_f\{x\} \right)^2 \right\} = \frac{L^2 - 1}{12L^2} .$$

PROOF. The variance is given by

$$\begin{aligned} E_f \left\{ \left(x - E_f\{x\} \right)^2 \right\} &= \left(\sum_{i=1}^L \frac{1}{L} \frac{(2i-1)^2}{(2L)^2} \right) - \frac{1}{4} \\ &= \frac{(2L+1)(2L-1)}{12L^2} - \frac{1}{4} , \end{aligned}$$

which gives the above result. \square

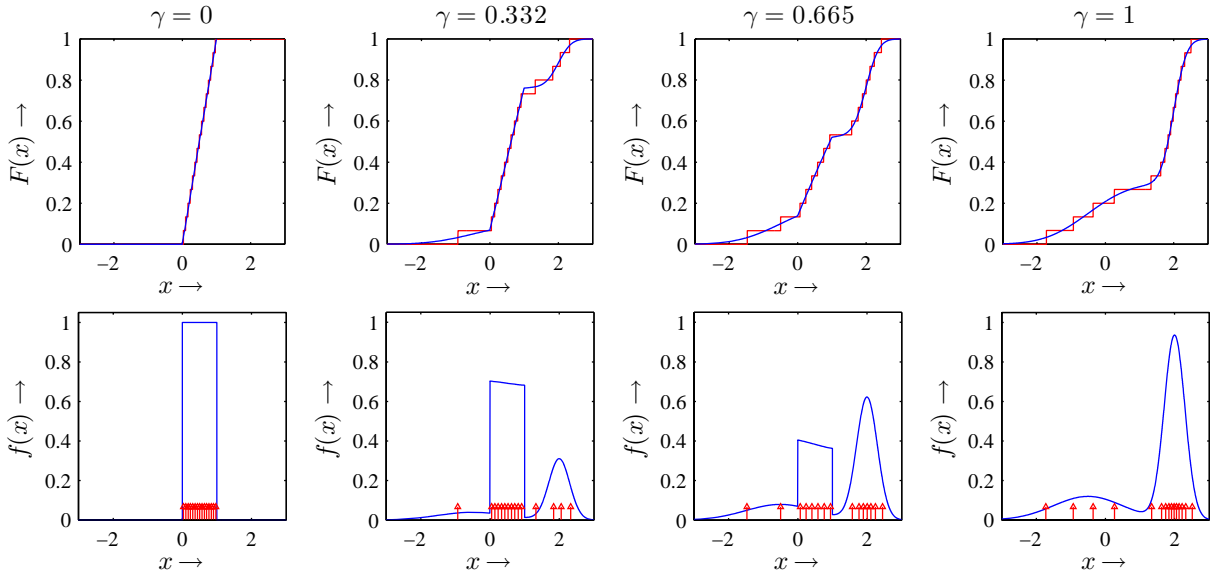


Fig. 3. Progressive approximation of the Gaussian mixture with two components Example V.2 for $\gamma = 0 \dots 1$. The approximation density consists of $L = 15$ components. The height of each Dirac component in the density approximation corresponds to its weight value.

The variance of the approximation density converges to the true variance when the number of components goes to infinity.

B. Special Case: Gaussian Density

Without loss of generality, we consider a standard normal density with distribution function

$$\tilde{F}(x) = \frac{1}{2} \operatorname{erf} \left(\frac{x}{\sqrt{2}} \right) + \frac{1}{2},$$

where $\operatorname{erf}(\cdot)$ denotes the error function.

Lemma IV.3 *The parameters of a Dirac mixture approximation of the standard normal density are given by*

$$x_i = \sqrt{2} \operatorname{erf}^{-1} \left(\frac{2i - 1 - L}{L} \right)$$

for $i = 1, \dots, L$.

PROOF. With (4) we have

$$\frac{1}{2} \operatorname{erf} \left(\frac{x_i}{\sqrt{2}} \right) + \frac{1}{2} = \frac{2i - 1}{2L},$$

which immediately gives the desired result. \square

The Dirac mixture approximation of the standard normal density and the corresponding distribution for a different number of components L is shown in Figure 2.

Remark IV.1 *We assume, that a suitable implementation of the inverse error function $\operatorname{erf}^{-1}(\cdot)$ ¹ is available.*

¹The corresponding MATLAB function is denoted by `erfinv()`

V. SOLUTION FOR THE GENERAL CASE

For general true densities $\tilde{f}(\cdot)$, a closed-form solution for the parameters of the approximating Dirac mixture density is not possible. Hence, we have to resort to a numerical solution of (4) in Theorem III.1.

Of course, in the scalar case considered in this paper, a wealth of numerical procedures for solving (4) are readily available. However, the multidimensional case calls for more advanced approaches. This is even more important, when Dirac mixture approximations with non-equally weighted components are considered [14]. Hence, we provide an efficient solution procedure, that will be derived and explained in simple scalar cases, but is also very well suited for the more advanced cases.

The approach pursued here is based on the intuition that in typical density approximation scenarios, a certain prior density is given, which is transformed by the considered type of processing step. This includes transformation of random variables, the Bayesian filter step (measurement update), and the prediction step (time update) for propagating a given prior density through a nonlinear dynamic system.

Instead of performing the considered processing step at once, the effect on the resulting density is introduced gradually. For that purpose, a continuous transformation of the given density towards the desired density is employed. This typically allows us to start with a “simple” density $s(x)$, for which the approximation is either already known or can easily be constructed.

A progression parameter γ is introduced, which is used to parameterize the true density $\tilde{f}(\cdot)$. Without loss of generality, the progression parameter γ is assumed to range in the interval $\gamma \in [0, 1]$, such that $\tilde{f}(x, \gamma = 0)$ corresponds to the simple density $s(x)$ and $\tilde{f}(x, \gamma = 1)$ corresponds to the original true density, i.e., $\tilde{f}(x, \gamma = 0) = s(x)$ and $\tilde{f}(x, \gamma = 1) = \tilde{f}(x)$.

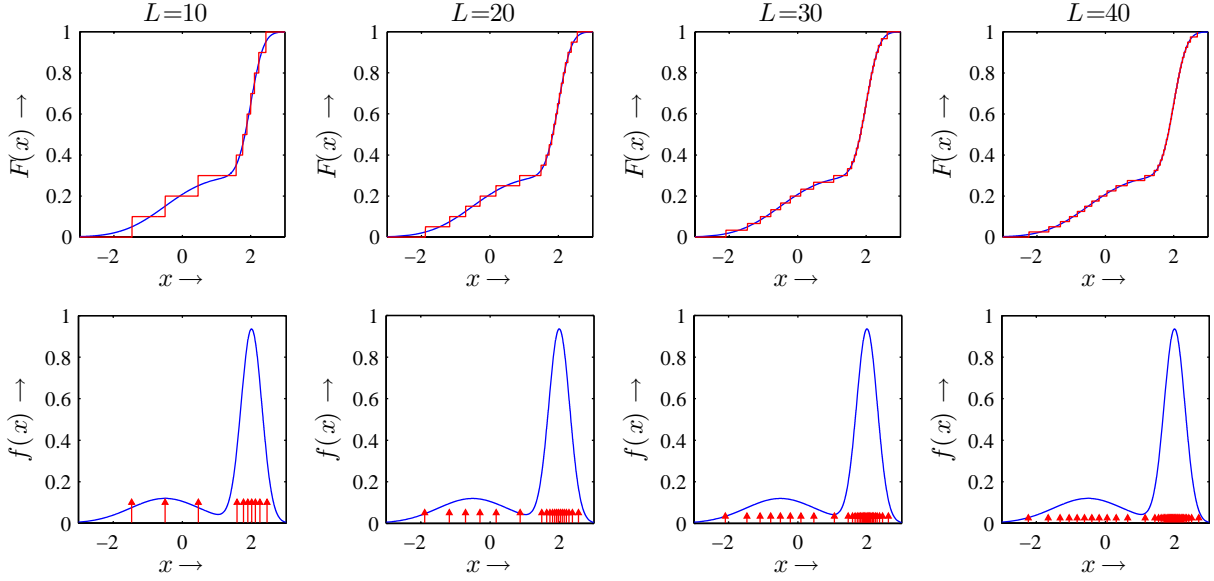


Fig. 4. Approximation of the Gaussian mixture from Example V.2 with a Dirac mixture with $L = 10$, $L = 20$, $L = 30$, and $L = 40$ components.

A straightforward progression from the simple density $s(x)$ to the original true density $\tilde{f}(x)$ is given by

$$\tilde{f}(x, \gamma) = (1 - \gamma) s(x) + \gamma \tilde{f}(x) , \quad (7)$$

which is demonstrated in the next example. Different progression schedules are possible as shown in [14].

Example V.1 We consider a Gaussian Mixture density with two components given by

$$\tilde{f}(x) = w_1 N(x, m_1, \sigma_1) + w_2 N(x, m_2, \sigma_2) \quad (8)$$

with

$$\begin{aligned} w_1 &= 0.3, \quad w_2 = 0.7 , \\ m_1 &= -0.5, \quad m_2 = 2.0 , \\ \sigma_1 &= 1, \quad \sigma_2 = 0.3 . \end{aligned}$$

As simple density, we select the uniform density given in (5).

The resulting homotopy is visualized in Figure 3 for various values of $\gamma \in [0, 1]$. In addition, the optimal Dirac mixture approximations with $L = 15$ components are shown for the corresponding values of γ . This approximation will be discussed in what follows.

The approximation now tracks the true density that is progressively modified by increasing γ . In order for the parameter vector $\underline{\eta}(\gamma)$ to track the optimum, we require a differential relation between the progression parameter γ and the parameter vector $\underline{\eta}$. For that purpose, we consider the cumulative version of (7) given by

$$\tilde{F}(x, \gamma) = (1 - \gamma) S(x) + \gamma \tilde{F}(x) ,$$

where $S(x)$ is the cumulative distribution corresponding to the simple density $s(x)$. The resulting progressive version $\tilde{F}(x, \gamma)$ of $\tilde{F}(x)$ is then plugged into (4) and we take the derivative with respect to γ . Since $\tilde{F}(x_i, \gamma)$ is both an explicit and due to $x_i = x_i(\gamma)$ an implicit function of γ , we obtain

$$\frac{\partial \tilde{F}(x_i, \gamma)}{\partial \gamma} + \frac{\partial \tilde{F}(x_i, \gamma)}{\partial x_i} \frac{\partial x_i(\gamma)}{\partial \gamma} = 0 ,$$

for $i = 1, \dots, L$. With

$$\frac{\partial \tilde{F}(x_i, \gamma)}{\partial x_i} = \tilde{f}(x_i, \gamma) ,$$

we obtain

$$-\frac{\partial \tilde{F}(x_i, \gamma)}{\partial \gamma} = \tilde{f}(x_i, \gamma) \dot{x}_i(\gamma) ,$$

where $\dot{x}_i(\gamma)$ denotes the derivative of x_i with respect to the progression parameter γ .

In vector–matrix–notation, we obtain the following system of explicit first–order ordinary differential equations (ODE)

$$\underline{b}(\underline{\eta}(\gamma), \gamma) = \mathbf{P}(\underline{\eta}(\gamma), \gamma) \cdot \dot{\underline{\eta}}(\gamma) , \quad (9)$$

with

$$\underline{b}(\underline{\eta}(\gamma), \gamma) = \begin{bmatrix} -\frac{\partial \tilde{F}(x_1, \gamma)}{\partial \gamma} \\ -\frac{\partial \tilde{F}(x_2, \gamma)}{\partial \gamma} \\ \vdots \\ -\frac{\partial \tilde{F}(x_L, \gamma)}{\partial \gamma} \end{bmatrix}$$

and

$$\mathbf{P}(\underline{\eta}(\gamma), \gamma) = \text{diag} \left(\tilde{f}(x_1, \gamma), \tilde{f}(x_2, \gamma), \dots, \tilde{f}(x_L, \gamma) \right) .$$

For the specific progression given in (7), we obtain

$$\frac{\partial \tilde{F}(x_i, \gamma)}{\partial \gamma} = \tilde{F}(x_i(\gamma)) - \tilde{S}(x_i(\gamma)) .$$

As a result, we obtain $\underline{b}(\underline{\eta}(\gamma), \gamma) = \underline{b}(\underline{\eta}(\gamma))$ with

$$\underline{b}(\underline{\eta}(\gamma)) = \begin{bmatrix} \tilde{S}(x_1(\gamma)) - \tilde{F}(x_1(\gamma)) \\ \tilde{S}(x_2(\gamma)) - \tilde{F}(x_2(\gamma)) \\ \vdots \\ \tilde{S}(x_L(\gamma)) - \tilde{F}(x_L(\gamma)) \end{bmatrix} . \quad (10)$$

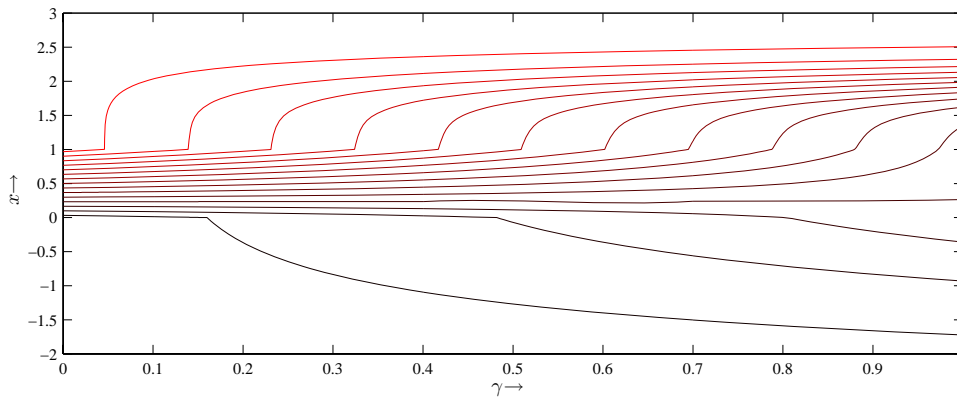


Fig. 5. Trace of the parameter vector $\underline{\eta}(\gamma)$ with $L = 15$ samples. Each line in the plot shows the evolution of a particular parameter.

Example V.2 This example demonstrates the approximation of the Gaussian mixture density from Example V.1 by means of a Dirac mixture. For that purpose, we use the simple progression scheme given in (7) also used in Example V.1.

For tracking the parameter vector, the system of ODE (9) is solved for $\gamma \in [0, 1]$ with $\underline{b}(\underline{\eta}(\gamma), \gamma)$ given in (10). The progression is started with a parameter vector $\underline{\eta}(\gamma = 0)$ corresponding to the optimal approximation of the uniform density given in (6). For $\gamma = 1$, the parameter vector $\underline{\eta}(\gamma = 1)$ corresponding to the desired optimal Dirac mixture approximation of the true density (8) is obtained. This is demonstrated in Figure 3 for a fixed number of $L = 15$ components and a few selected values of $\gamma \in [0, 1]$. Please note that γ continuously covers the interval $[0, 1]$.

The resulting Dirac mixture approximations are shown in Figure 4 for a different number of mixture components.

Figure 5 shows the evolution of the parameters of the Dirac mixture, i.e., the positions of the individual components, as the progression parameter γ varies from $\gamma = 0$ to $\gamma = 1$.

VI. DISCUSSION AND FUTURE WORK

In this paper a systematic procedure for approximating an arbitrary probability density function by means of a Dirac mixture with equally weighted components as a special case of the method proposed in [14] has been introduced. The resulting approximation can be used immediately as a basis for recursive nonlinear filtering mechanisms. This can be seen as an alternative to Monte Carlo based particle filters. A special benefit of the proposed method lies in the fact, that the approximation is deterministic and hence, successive application yields identical results, which is not the case for random number sampling based Monte Carlo methods.

The proposed procedure has been introduced in the context of scalar random variables for the sake of simplicity. It can, however, be generalized to random vectors in a straightforward manner.

Of course, the procedure is not limited to integral quadratic distance measures. Similar derivations as the ones given in this paper can be performed for different distance measures.

REFERENCES

- [1] D. L. Alspach and H. W. Sorenson, "Nonlinear Bayesian Estimation Using Gaussian Sum Approximation," *IEEE Transactions on Automatic Control*, vol. AC-17, no. 4, pp. 439–448, 1972.
- [2] S. Challa, Y. Bar-Shalom, and V. Krishnamurthy, "Nonlinear Filtering via Generalized Edgeworth Series and Gauss–Hermite Quadrature," *IEEE Transactions on Signal Processing*, vol. 48, no. 6, pp. 1816–1820, 2000.
- [3] D. Brigo, B. Hanzon, and F. LeGland, "A Differential Geometric Approach to Nonlinear Filtering: The Projection Filter," *IEEE Transactions on Automatic Control*, vol. 42, no. 2, pp. 247–252, 1998.
- [4] N. Oudjane and S. Rubenthaler, "Stability and Uniform Particle Approximation of Nonlinear Filters in Case of non–Ergodic Signal," Prépublication, Laboratoire de Probabilités et Modèles Aléatoires, Université de Paris VI, Tech. Rep. PMA786, January 2003. [Online]. Available: <http://www.proba.jussieu.fr/mathdoc/textes/PMA-786.pdf>
- [5] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A Tutorial on Particle Filters for On–line Non–linear/Non–Gaussian Bayesian Tracking," *IEEE Transactions of Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.
- [6] E. Bølviken, P. J. Acklam, N. Christophersen, and J.-M. Størdal, "Monte Carlo Filters for Non–Linear State Estimation," *Automatica*, vol. 37, pp. 177–183, 2001.
- [7] A. Doucet, S. Godsill, and C. Andrieu, "On Sequential Monte Carlo Sampling Methods for Bayesian Filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.
- [8] E. Bølviken and G. Storvik, "Deterministic and Stochastic Particle Filters in State Space Models," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. de Freitas, and N. Gordon, Eds. Springer Series of "Statistics for Engineering and Information Science", 2001.
- [9] S. Kullback and R. A. Leibler, "On Information and Sufficiency," *Annals of Mathematical Statistics*, vol. 22, no. 2, pp. 79–86, 1951.
- [10] M. D. Weber, L. M. Leemis, and R. K. Kincaid, "Minimum Kolmogorov–Smirnov Test Statistic Parameter Estimates," *Journal of Statistical Computation and Simulation*, vol. 76, no. 3, pp. 196–206, 2006.
- [11] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, "Kolmogorov–Smirnov Test," in *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 1992, pp. 623–628.
- [12] D. A. Darling, "The Kolmogorov–Smirnov, Cramér–von Mises Tests," *The Annals of Mathematical Statistics*, vol. 28, no. 4, pp. 823–83, 1957.
- [13] D. D. Boos, "Minimum Distance Estimators for Location and Goodness of Fit," *Journal of the American Statistical Association*, vol. 76, no. 375, pp. 663–670, 1981.
- [14] O. C. Schrempf, D. Brunn, and U. D. Hanebeck, "Dirac Mixture Density Approximation Based on Minimization of the Weighted Cramér–von Mises Distance," in *Proceedings of the International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2006)*, Heidelberg, Germany, September 2006, pp. 512–517.
- [15] U. D. Hanebeck and O. Feiermann, "Progressive Bayesian Estimation for Nonlinear Discrete–Time Systems: The Filter Step for Scalar Measurements and Multidimensional States," in *Proceedings of the 2003 IEEE Conference on Decision and Control (CDC'03)*, Maui, Hawaii, 2003, pp. 5366–5371.