

Parameterized Joint Densities with Gaussian and Gaussian Mixture Marginals

Felix Sawo, Dietrich Brunn, and Uwe D. Hanebeck

Intelligent Sensor-Actuator-Systems Laboratory
 Institute of Computer Science and Engineering
 Universität Karlsruhe (TH)
 Karlsruhe, Germany

{sawo,brunn}@ira.uka.de, uwe.hanebeck@ieee.org

Abstract - *In this paper we attempt to lay the foundation for a novel filtering technique for the fusion of two random vectors with imprecisely known stochastic dependency. This problem mainly occurs in decentralized estimation, e.g., of a distributed phenomenon, where the stochastic dependencies between the individual states are not stored. Thus, we derive parameterized joint densities with both Gaussian marginals and Gaussian mixture marginals. These parameterized joint densities contain all information about the stochastic dependencies between their marginal densities in terms of a parameter vector ξ , which can be regarded as a generalized correlation parameter. Unlike the classical correlation coefficient, this parameter is a sufficient measure for the stochastic dependency even characterized by more complex density functions such as Gaussian mixtures. Once this structure and the bounds of these parameters are known, bounding densities containing all possible density functions could be found.*

Keywords: Stochastic systems, data fusion, decentralized estimation, parameterized joint density, generalized correlation parameter

1 Introduction

In a wide variety of applications estimating the state of a dynamic system by fusing uncertain information is a topic of extraordinary importance, e.g., robot localization [1], multiple target tracking [2], and decentralized observation of a distributed phenomenon, to name just a few. In most cases, an appropriate system model together with a stochastic noise model is given and then the state is estimated by means of a Kalman filter or one of its variations [3].

Problems mainly occur when the stochastic dependency between the states and/or the sensor noises are not precisely known. In that case, classical concepts like the Kalman filter conveniently assume stochastic independence or precisely known correlation, which automatically leads to unjustified improvement of estimation results. For understanding the source of unknown stochastic dependency a typical application problem is considered: observation of a distributed phenomenon

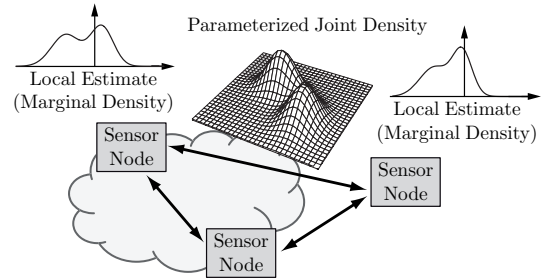


Figure 1: Simple example scenario for a decentralized sensor-actuator-network used for the observation of a distributed phenomenon.

by means of a sensor-actuator-network, see Fig. 1.

This scenario suffers from two types of *imprecisely known* stochastic dependencies. The first type is inherent to the system and caused by partially stochastic dependent noise sources for different sensor nodes. In other words, there are usually additional external disturbances affecting more than one sensor, e.g. sunshine, wind, or the same origin of a pollutant cloud. Even for ideal sensor properties, this already would lead to a partially stochastic dependency between the measured states.

In general, using a centralized approach, the distributed phenomenon can be easily observed by applying a standard estimator to the augmented state vector containing all the physical quantities estimated by individual sensor nodes; in the linear case this can be achieved by a Kalman filter, see [4]. In that case, the estimator stores the associated stochastic dependencies and use them for the next estimation step. However, for practical applications it is often desirable to reduce the heavy computational burden and to reduce communication activities between the individual nodes to a minimum. This leads to a decentralized estimation approach, which implies that just parts of the state vector are manipulated at each update step.

Hence, the decentralized estimation process itself causes a second source of imprecisely known stochastic dependency. Let us assume that a physical quantity is measured and processed only locally. In order to estimate the physical quantity at non-measurement points, i.e., by means of a distributed parameterized system (PDE), the individual sensor nodes exchange

and fuse their local estimates. In this case the resulting estimates become automatically stochastic dependent after the first fusion of the individual estimates. Unfortunately, applying the Kalman filter for decentralized problems while ignoring the existing dependencies between the individual states leads to overoptimistic estimation results. Coping with such problems is one of the main justifications for robust filter design [1].

Based on the correlation coefficient r , which is a measure for the stochastic dependency of random variables, two robust estimators have been introduced to solve the previous mentioned problems, namely Covariance Intersection [5, 6] and Covariance Bounds [7]. Basically, these filters do not neglect the unknown stochastic dependency, but consider them by producing conservative estimates for prediction and measurement step, which are compatible with all correlations within an assumed correlation structure. Besides, the generalization of the Covariance Intersection which is based on a minimization of the Chernoff information is worthwhile mentioning [8].

Although robust filters like Covariance Intersection and Covariance Bounds are efficient for linear state-space models and linear measurement models, they cannot be directly applied to nonlinear models, e.g. of a distributed phenomenon. In addition, these filters are not able to work with more complex density functions such as Gaussian mixtures, which are known as universal approximators and thus, well-suited for nonlinear estimation [9].

In this paper, we attempt to lay the foundation for a novel filtering technique, which we believe could possibly cope with all previously mentioned drawbacks. The main problem of a robust decentralized estimation based on more complex density functions, such as Gaussian mixtures, is that for these functions the *classical* correlation coefficient r is not a sufficient measure for the stochastic dependency. This means that for such density functions no obvious correlation coefficient exists that can be bounded and thus, no bounding density can be found in the classical sense.

The main contribution of this paper is to derive *parameterized joint densities* with both Gaussian marginals and Gaussian mixture marginals. These parameterized joint densities contain all information about the stochastic dependency between their marginal density functions in terms of a parameter vector ξ . This parameter vector can be regarded as a kind of *generalized correlation parameter* for the assumed structure of stochastic dependency. Once this structure and the bound of the correlation parameter vector ξ are known, bounding densities which are compatible with all stochastic dependency structures can be found. In [10], a definition for lower and upper bounding densities is given, and their potential use in robust estimation is shown.

The remainder of this paper is structured as follows. In Section 2, a rigorous formulation of the problem of state estimation with unknown correlations is given. Section 3 then derives several types of parameterized joint densities with Gaussian marginals. These joint densities are useful for robust linear estimation.

Furthermore, they are helpful to gain some insight about the *generalized correlation parameter vector* ξ . Whereas, in section 4 we generalize these ideas to parameterize joint densities with *Gaussian mixture marginals*. Bounding the parameter vector of these joint densities, it is possible to find a conservative estimation even for nonlinear problems and for more complex density functions.

Throughout this paper, we use the notation $\mathcal{N}(z, \mathbf{C}(r))$ to represent a joint Gaussian density, which is defined as

$$\mathcal{N}(z, \mathbf{C}(r)) = \frac{1}{2\pi |\mathbf{C}(r)|} \exp \left\{ -\frac{1}{2} z^T \mathbf{C}(r)^{-1} z \right\} ,$$

where

$$z = \begin{bmatrix} x - \hat{x} \\ y - \hat{y} \end{bmatrix} , \quad \mathbf{C}(r) = \begin{bmatrix} C_x & r \sqrt{C_x C_y} \\ r \sqrt{C_y C_x} & C_y \end{bmatrix} ,$$

are state vector and covariance matrix, respectively. The expected value of the states are denoted by \hat{x} and \hat{y} , and $r \in [-1, 1]$ denotes the *classical* correlation coefficient.

2 Problem Formulation

As mentioned in the introduction, there are many sources of stochastic dependencies. In this section, we take up the previously mentioned example, the observation of a distributed phenomenon with a sensor-actuator-network. By this means we are able to clarify the main problem common to all sources of imprecisely known stochastic dependency: *imprecisely* known joint densities with given marginal densities.

For the sake of simplicity, consider a simple discrete-time dynamic model with the system state $x_k \in \mathbb{R}$, and the system input $u_k \in \mathbb{R}$ according to

$$x_{k+1} = a_k(x_k) + b_k(u_k) . \quad (1)$$

The corresponding density functions of the estimates are given by $f_x^e(x_k)$ and $f_u^e(u_k)$.

In the case of a precisely known joint density $f_k^e(x_k, u_k)$, the predicted density $f_{k+1}^p(x_{k+1})$ is given by

$$\begin{aligned} & f_{k+1}^p(x_{k+1}) \\ &= \int_{\mathbb{R}^2} \delta(x_{k+1} - a_k(x_k) - b_k(u_k)) f_k^e(x_k, u_k) dx_k du_k , \end{aligned} \quad (2)$$

where δ denotes the Dirac delta distribution.

With the justification of the considered example scenario we assume that the state estimate x_k and the system input u_k are stochastically dependent with an *imprecisely* known structure. Hence, although the marginal density functions $f_x^e(x_k)$ and $f_u^e(u_k)$ are known, the joint density $f_k^e(x_k, u_k)$, which contains all information about the stochastic dependency, is unknown. However, as it can be seen in (2), the knowledge of the joint density or at least its parameterization in terms of a correlation parameter is essential for the estimation process.

If the joint density, i.e., the correlation structure, would be known precisely, this correlation could easily be tracked and be considered in the next processing step. However, when the correlation structure is not precisely known, the joint densities need somehow to be reconstructed.

In the next section, the reconstruction of joint densities $f(x, y)$ based on known marginal densities $f_x(x)$ and $f_y(y)$ is discussed in more detail. For all introduced types this leads to a joint density parameterized by a generalized correlation parameter vector ξ containing the information about the stochastic dependency between the considered random variables.

3 Gaussian Marginals for Linear Estimation

In this section, we describe possible parameterizations of joint densities for given Gaussian marginal densities. We have identified three types of parameterized joint densities. The first two types mainly serve for didactic purposes for understanding the problem of imprecisely known stochastic dependency. The third type, however, is useful for robust decentralized linear estimation and thus, is of practical relevance.

3.1 Piecewise Gaussian Densities with Different Weighting Factors

This section consists of a description for the simplest non-Gaussian joint density with Gaussian marginals; this type frequently appears in informal discussions on the internet. Basically, a jointly Gaussian density is “sliced” into different pieces, which are raised or lowered, respectively. The visualization of two examples for this simple type of parameterized joint density is shown in Fig. 2.

THEOREM 3.1 *Given two Gaussian marginal densities $f_x(x) = \mathcal{N}(\hat{x}, C_x)$ and $f_y(y) = \mathcal{N}(\hat{y}, C_y)$, a family of possible joint densities $f(x, y)$ can be parameterized by*

$$f(x, y) = \begin{cases} \xi_1 \tilde{f}(x, y) & x \cdot y \geq 0 \\ \xi_2 \tilde{f}(x, y) & x \cdot y < 0 \end{cases} \quad (3)$$

where $\xi_1 = 2(1 - \xi)$ and $\xi_2 = 2\xi$.

The free parameter $\xi \in [0, 1]$ can be regarded as a generalized correlation parameter, which specifies the member of the family. To assure that the parameterized joint density $f(x, y)$ is a valid joint density for the given marginals $f_x(x)$ and $f_y(y)$, the density $\tilde{f}(x, y)$ must be defined according to

$$\tilde{f}(x, y) = \mathcal{N}(x - \hat{x}, C_x) \mathcal{N}(y - \hat{y}, C_y) .$$

PROOF. To prove this, it must be shown that the marginal densities of $f(x, y)$ in (3) are represented by $f_x(x) = \mathcal{N}(x - \hat{x}, C_x)$ and $f_y(y) = \mathcal{N}(y - \hat{y}, C_y)$, respectively. The marginal density $f_x(x)$ can be derived

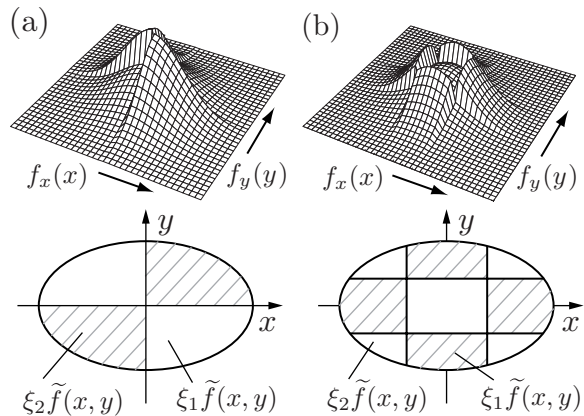


Figure 2: Gaussian density with piecewise different weighting factors; “sliced” in (a) four pieces and (b) nine pieces with different weighting factors.

by direct integration over y . Assuming $x > 0$, it follows

$$\begin{aligned} f_x(x) &= \int_0^\infty \xi_1 \tilde{f}(x, y) dy + \int_{-\infty}^0 \xi_2 \tilde{f}(x, y) dy \\ &= 2(1 - \xi) \int_0^\infty \tilde{f}(x, y) dy + 2\xi \int_{-\infty}^0 \tilde{f}(x, y) dy \\ &= 2 \int_0^\infty \tilde{f}(x, y) dy - 2\xi \int_0^\infty \tilde{f}(x, y) dy + 2\xi \int_{-\infty}^0 \tilde{f}(x, y) dy \end{aligned}$$

Due to the symmetry property of the Gaussian density this integral can be simplified, which yields

$$f_x(x) = \int_{-\infty}^\infty \tilde{f}(x, y) dy = \mathcal{N}(x - \hat{x}, C_x) .$$

Similar calculations for $x \leq 0$ justify the choices for ξ_1 and ξ_2 . This proves Theorem 3.1. \square

This simple parameterized joint density is depicted in Fig. 2 (a). It is possible to extend this type of parameterized joint density to more complex ones, i.e., the joint density could be “sliced” into more pieces, as shown in Fig. 2 (b).

3.2 Sum of Positive and Negative Gaussians

A second type of non-Gaussian density with Gaussian marginals could be constructed by the sum of *positive* and *negative* jointly Gaussian densities. The mean values, variances, and weighting factors of the individual joint densities need to be chosen appropriately, i.e., in such way that the joint density is both a valid density function and the marginals are Gaussian densities. For the sake of simplicity we consider only three components in the following theorem.

THEOREM 3.2 *Given two Gaussian densities $f_x(x) = \mathcal{N}(x - \hat{x}, C_x)$ and $f_y(y) = \mathcal{N}(y - \hat{y}, C_y)$, the unknown joint density $f(x, y)$ can be defined by the sum of positive and negative Gaussians according to*

$$f(x, y) = f_1(x, y) + f_2(x, y) - f_3(x, y) , \quad (4)$$

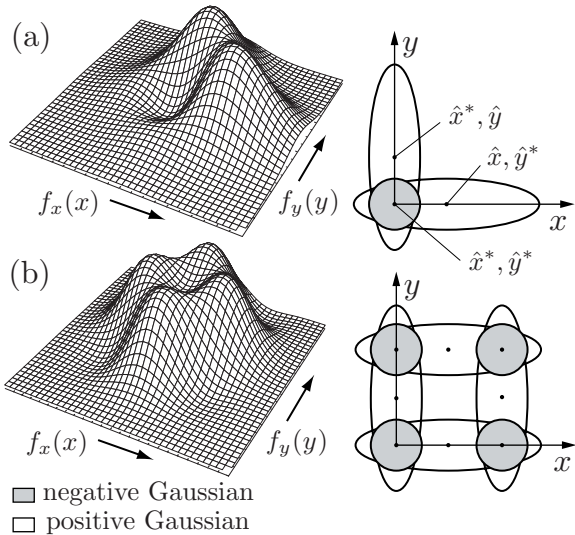


Figure 3: Joint density consisting of the sum of positive and negative Gaussian densities, (a) three components, (b) eight components.

where the individual densities $f_1(x, y)$, $f_2(x, y)$, and $f_3(x, y)$ are given by

$$\begin{aligned} f_1(x, y) &= \mathcal{N}(x - \hat{x}^*, C_y) \mathcal{N}(y - \hat{y}, C_x) , \\ f_2(x, y) &= \mathcal{N}(x - \hat{x}, C_x) \mathcal{N}(y - \hat{y}^*, C_y) , \\ f_3(x, y) &= \mathcal{N}(x - \hat{x}^*, C_y) \mathcal{N}(y - \hat{y}^*, C_y) . \end{aligned}$$

To assure that the parameterized joint density $f(x, y)$ is a valid joint density, the mean values and variances of the individual joint densities need to be chosen so that $f(x, y) \geq 0$ holds.

PROOF. The marginal density $f_x(x)$ can be derived by direct integration over y according to

$$\begin{aligned} f_x(x) &= \int_{\mathbb{R}} f(x, y) dy \\ &= \mathcal{N}(x - \hat{x}^*, C_y) + \mathcal{N}(x - \hat{x}, C_x) - \mathcal{N}(x - \hat{x}^*, C_y) \\ &= \mathcal{N}(x - \hat{x}, C_x) . \end{aligned}$$

Similar calculations for the marginal density $f_y(y)$ justify the choices for individual mean values and variances of the parameterized joint density, and finally conclude the proof. \square

Fig. 3 (a) visualizes the simplest case for this type of parameterized joint density consisting of three components. It is possible to extend this type to joint densities with more components, for example with eight components as depicted in Fig. 3 (b).

3.3 Mixtures of Correlated Jointly Gaussian Densities

In this section, we derive a parameterized joint density with practical relevance, which is based on the integral of jointly Gaussian densities with different classical correlation coefficients. The weighting factors of the individual joint densities need to be chosen in such a way that the marginals are represented by the given Gaussian marginal densities.

THEOREM 3.3 Given two Gaussian marginal densities $f_x(x) = \mathcal{N}(\hat{x}, C_x)$ and $f_y(y) = \mathcal{N}(\hat{y}, C_y)$, a family of possible joint densities, depending on the generalized correlation function $\xi(r)$, can be parameterized by

$$f(x, y) = \int_{-1}^1 \xi(r) \mathcal{N}\left(\begin{bmatrix} x - \hat{x} \\ y - \hat{y} \end{bmatrix}, \mathbf{C}(r)\right) dr , \quad (5)$$

where $\xi(r)$ is defined on $r \in [-1, 1]$.

The parameterized continuous Gaussian mixture $f(x, y)$ is a valid normalized density function for

$$\xi(r) \geq 0 , \quad \int_{-1}^1 \xi(r) dr = 1 .$$

PROOF. The results directly follow from the integration of the joint density $f(x, y)$ over y and x , respectively. Hence, the marginal density $f_x(x)$ can be derived by direct integration of the joint density $f(x, y)$ over y according to

$$\begin{aligned} f_x(x) &= \int_{\mathbb{R}} \int_{-1}^1 \xi(r) \mathcal{N}\left(\begin{bmatrix} x - \hat{x} \\ y - \hat{y} \end{bmatrix}, \mathbf{C}(r)\right) dr dy \\ &= \int_{-1}^1 \xi(r) \int_{\mathbb{R}} \mathcal{N}\left(\begin{bmatrix} x - \hat{x} \\ y - \hat{y} \end{bmatrix}, \mathbf{C}(r)\right) dy dr . \end{aligned}$$

With reference to [3] it can be shown that the solution of the integral does not depend on the correlation coefficient r at all. Thus, it can easily be obtained

$$f_x(x) = \int_{-1}^1 \xi(r) \mathcal{N}(x - \hat{x}, C_x) dr = \mathcal{N}(x - \hat{x}, C_x) ,$$

which justifies the condition $\int_{-1}^1 \xi(r) dr = 1$. Similar calculations for the marginal density $f_y(y)$ leads to the same condition and finally concludes the proof. \square

Two examples for this type of parameterized joint density are depicted in Fig. 4.

In the following, a rough idea is shown on how this type of parameterized joint density can be used for the development of a novel estimator for an imprecisely known correlation coefficient. Consider two marginal densities $f_x(x)$ and $f_y(y)$. Let us assume that just the mean value \hat{r} and variance C_r of their classical correlation coefficient r is known. In this case, a density function $\xi(r)$ for the correlation coefficient can be defined according to

$$\xi(r) = c_n \mathcal{N}(r - \hat{r}, C_r) \quad r \in [-1, 1] ,$$

with normalization constant c_n

$$(c_n)^{-1} = \int_{-1}^1 \mathcal{N}(r - \hat{r}, C_r) dr .$$

The density function $\xi(r)$, which is depicted in the right column of Fig. 4 (b), can be regarded as a generalized correlation function. That means, this function contains all information about the correlation structure. Using this information, the joint density $f(x, y)$ can be parameterized by

$$f(x, y) = \int_{-1}^1 c_n \mathcal{N}(r - \hat{r}, C_r) \mathcal{N}\left(\begin{bmatrix} x - \hat{x} \\ y - \hat{y} \end{bmatrix}, \mathbf{C}(r)\right) dr . \quad (6)$$

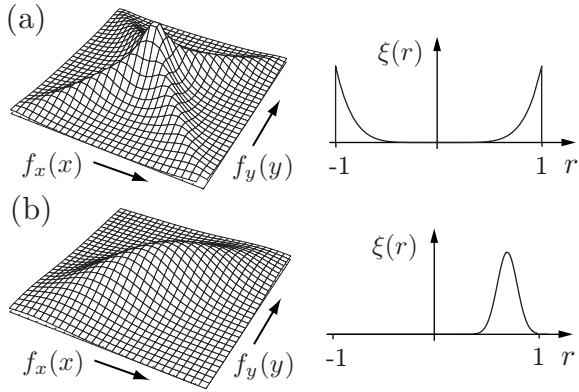


Figure 4: Joint density consisting of sum of correlated Gaussian density for different generalized correlation function $\xi(r)$, (a) $\xi(r) = \frac{7}{2} r^6$, and (b) $\xi(r) = \mathcal{N}(r - \hat{r}, C_r)$.

It is possible to use this parameterized joint density as the estimated joint density $f_k^e(x_k, u_k)$ in (2). The predicted density $f_{k+1}^p(x_{k+1})$ depending on the parameters of the generalized correlation function $\xi(r)$ can then be derived. This idea could lay the foundation for a novel filtering technique taking into account the imprecisely known classical correlation coefficient r and gives a bounding density for the prediction result. The actual calculation of a bounding density is left for future research work.

4 Gaussian Mixture Marginals for Nonlinear Estimation

So far, we have only considered parameterized joint densities with Gaussian marginals. In this section, we generalize these ideas to the parameterization of joint densities with Gaussian mixture marginals. Gaussian mixtures consist of the convex sum of weighted Gaussian densities. Due to the fact that Gaussian mixtures are universal approximators, they are well-suited for nonlinear estimation problems [11].

Thus, finding a parameterization for the imprecisely known joint density with Gaussian mixture marginals, it is possible to develop a novel filtering technique, which could possibly cope with both nonlinear system models and nonlinear measurement models in a robust manner. As it was mentioned in the introduction, the challenge of a robust decentralized estimation based on Gaussian mixtures is that the classical correlation coefficient r is not a sufficient measure for the stochastic dependency of Gaussian mixtures. Therefore, we define in this section a *generalized correlation parameter vector* ξ for Gaussian mixtures. Finding bounding densities, which are compatible with all stochastic dependency structures in terms of ξ , it is possible to derive a robust filtering technique for distributed nonlinear problems.

For the sake of simplicity consider two scalar

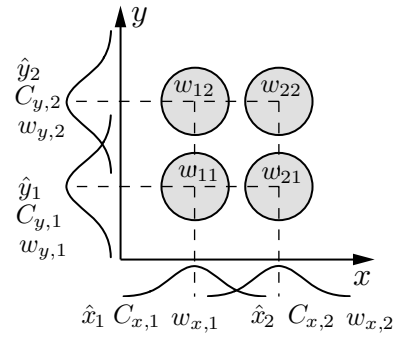


Figure 5: Notation for marginal densities and joint density for Gaussian mixtures.

Gaussian mixtures marginals according to

$$f_x(x) = \sum_{i=1}^m w_{x,i} \mathcal{N}(x - \hat{x}_i, C_{x,i}) , \quad (7)$$

$$f_y(y) = \sum_{i=1}^n w_{y,i} \mathcal{N}(y - \hat{y}_i, C_{y,i}) , \quad (8)$$

where \hat{x}_i and \hat{y}_i are the individual means, $C_{x,i}$ and $C_{y,i}$ are the individual variances, and $w_{x,i}$ and $w_{y,i}$ are the individual weighting factors, which must be positive and sum to one. The notation for the marginals and their joint density are visualized in Fig. 5.

THEOREM 4.1 *Given two Gaussian mixture marginal densities $f_x(x)$ and $f_y(y)$, a family of possible joint densities $f(x, y)$ depending on the weighting factors w_{ij} is defined by*

$$f(x, y) = \sum_{i=1}^m \sum_{j=1}^n w_{ij} \mathcal{N} \left(\begin{bmatrix} x - \hat{x}_i \\ y - \hat{y}_j \end{bmatrix}, \mathbf{C}_{ij}(r_{ij}) \right) . \quad (9)$$

To assure that the parameterized Gaussian mixture $f(x, y)$ is a valid normalized density function, the weighting factors w_{ij} must be positive and sum to one

$$w_{ij} \geq 0 , \quad \sum_{i=1}^m \sum_{j=1}^n w_{ij} = 1 . \quad (10)$$

In addition, the weighting factors must satisfy

$$\sum_{j=1}^n w_{ij} = w_{x,i} , \quad \sum_{i=1}^m w_{ij} = w_{y,j} , \quad (11)$$

to ensure that the parameterized joint density $f(x, y)$ is a valid joint density for the marginal density functions $f_x(x)$ and $f_y(y)$.

PROOF. These properties and conditions can be proven by showing that the marginal densities of $f(x, y)$ are represented by $f_x(x)$ and $f_y(y)$, given by (7) and (8), respectively. The marginal density $f_x(x)$ can be derived by direct integration over y according to

$$\begin{aligned} f_x(x) &= \int_{\mathbb{R}} \sum_{i=1}^m \sum_{j=1}^n w_{ij} \mathcal{N} \left(\begin{bmatrix} x - \hat{x}_i \\ y - \hat{y}_j \end{bmatrix}, \mathbf{C}_{ij}(r_{ij}) \right) dy \\ &= \sum_{i=1}^m \sum_{j=1}^n w_{ij} \int_{\mathbb{R}} \mathcal{N} \left(\begin{bmatrix} x - \hat{x}_i \\ y - \hat{y}_j \end{bmatrix}, \mathbf{C}_{ij}(r_{ij}) \right) dy \end{aligned}$$

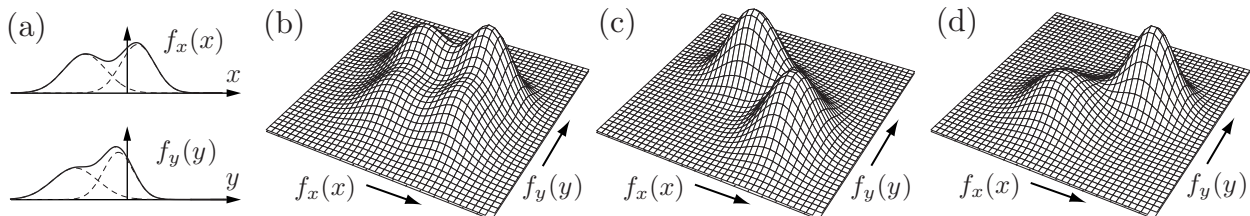


Figure 6: **(a)** Gaussian mixture marginal densities with two individual components ($w_{x,i} = w_{y,i} = 0.5$), $f_x(x)$ with $\hat{x}_1 = -3$, $\hat{x}_2 = 1$, $C_{x,1} = 2$, $C_{x,2} = 1.6$, and $f_y(y)$ with $\hat{y}_1 = -4$, $\hat{y}_2 = -1$, $C_y = 2.1$, $C_y = 1.4$. Parameterized joint densities for various generalized correlation parameter ξ , **(b)** $\xi = 0$, **(c)** $\xi = -1$, **(d)** $\xi = 1$.

Referring to [3] it can be shown that the integral solution does *not* depend on the correlation coefficient r_{ij} at all. Thus,

$$f_x(x) = \sum_{i=1}^m \sum_{j=1}^n w_{ij} \mathcal{N}(x - \hat{x}_i, C_{x,i})$$

can easily be obtained. Now, it is obvious that for the condition $\sum_{j=1}^n w_{ij} = w_{x,i}$ this leads to the desired Gaussian mixture marginals

$$f_x(x) = \sum_{i=1}^m w_{x,i} \mathcal{N}(x - \hat{x}_i, C_{x,i}) .$$

Similar calculations for marginal density function $f_y(y)$ leads to the desired condition $\sum_{i=1}^m w_{ij} = w_{y,j}$. This concludes the proof. \square

For the following calculations it is more convenient to rearrange the weighting factors of the joint Gaussian mixture density w_{ij} from matrix form to vector form according to

$$\underline{w} = [w_{11} \ \cdots \ w_{1n}, \ w_{21} \ \cdots \ w_{2n}, \ \cdots \ w_{mn}]^T$$

The weighting factors of marginals are given by

$$\begin{aligned} \underline{w}_x &= [w_{x,1} \ \cdots \ w_{x,m}]^T , \\ \underline{w}_y &= [w_{y,1} \ \cdots \ w_{y,n}]^T . \end{aligned}$$

The conditions (11) for a valid joint density $f(x, y)$ with marginals $f_x(x)$ and $f_y(y)$ can be expressed by a linear equation, which transforms weighting factors of joint density to weighting factors of marginals. This linear transformation is given by

$$\mathbf{T} \underline{w} = \begin{bmatrix} \underline{w}_x \\ \underline{w}_y \end{bmatrix} .$$

Due to the fact that the matrix $\mathbf{T} \in \mathbb{R}^{(n+m) \times nm}$ in general is not a square matrix and does *not* have full rank, there exists a null space (kernel) $\ker \mathbf{T}$, which is given by

$$(\ker \mathbf{T}) \underline{w} = \underline{\xi} .$$

The null space $\ker \mathbf{T}$ is spanned by the free parameter vector $\underline{\xi}$. Thus, for the calculation of valid weighting factors \underline{w} we propose following Lemma.

LEMMA 4.2 *The weighting factors \underline{w} of the joint density $f(x, y)$ can be derived according to*

$$\underline{w} = \mathbf{T}_e^\dagger \begin{bmatrix} \underline{w}_x \\ \underline{w}_y \\ \underline{\xi} \end{bmatrix} , \quad \mathbf{T}_e = \begin{bmatrix} \mathbf{T} \\ \ker \mathbf{T} \end{bmatrix} ,$$

where \mathbf{T}_e describes a unique transformation of weighting factors for valid joint densities to weighting factors of given marginal densities. The pseudo-inverse is denoted by \mathbf{T}_e^\dagger .

Similar to the other types of parameterized joint densities the free parameter vector $\underline{\xi}$ can be regarded as a kind of generalized correlation parameter for Gaussian mixtures. This parameter vector needs to be specified in order to define the joint density $f(x, y)$ uniquely.

Here, the individual components of the Gaussian mixture are assumed to be uncorrelated, i.e., $r_{ij} = 0$. That's why the resulting generalized correlation vector $\underline{\xi}$ is derived only by the weighting factors \underline{w} ; and not the correlation coefficients r_{ij} .

EXAMPLE 4.1 Now, we will consider possible joint densities $f(x, y)$ for two given Gaussian mixture marginals $f_x(x)$ and $f_y(y)$, depicted in Fig. 6 (a). The weighting factors for components of the joint density are obtained by

$$\begin{bmatrix} w_{11} \\ w_{12} \\ w_{21} \\ w_{22} \end{bmatrix} = \frac{1}{8} \begin{bmatrix} 3 & -1 & 3 & -1 & -2 \\ -1 & 3 & 3 & -1 & 2 \\ -1 & 3 & -1 & 3 & 2 \\ 3 & -1 & -1 & 3 & -2 \end{bmatrix} \cdot \begin{bmatrix} w_{x,1} \\ w_{x,2} \\ w_{y,1} \\ w_{y,2} \\ \xi \end{bmatrix} ,$$

where ξ is the free parameter. Possible joint densities for various parameters ξ are depicted in Fig. 6(b)–(d).

Furthermore, it is possible to combine all the types of parameterized joint densities introduced in this paper. The most relevant combination is applying the mixtures of correlated jointly Gaussian (Sec. 3.3) to the individual components of a Gaussian mixture. For simplicity we consider the previous example and assume the *generalized correlation function* $\xi_{ij}(r)$ for the ij -th component to be given by

$$\xi_{ij}(r) = \delta(r - \alpha) ,$$

where δ is the Dirac delta distribution and α denotes the *classical correlation coefficient*. In Fig. 7 the parameterized joint density is depicted for different α and ξ .

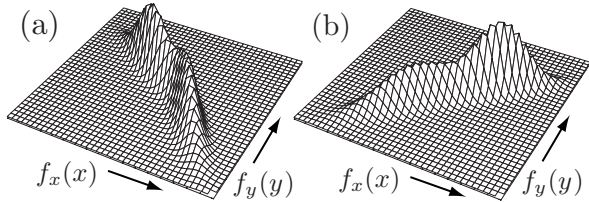


Figure 7: Combination of two types of parameterized joint densities. (a) $\xi = -1$, $\alpha = -1$, (b) $\xi = 1$, $\alpha = 1$.

It turns out that in the case of $\alpha = -1$ and $\xi = -1$ the two considered random variables, which are described by Gaussian mixtures, are fully correlated, i.e., similar to the Gaussian case $r \Rightarrow -1$.

Similar to the previous example and with reference to (2), let us consider possible joint densities $f_k^e(x_k, u_k)$ for two given Gaussian mixture marginals, $f_x^e(x_k)$ and $f_u^e(u_k)$. The Gaussian mixtures consisting of three components is visualized in Fig. 8 (a). It turns out that the free parameter vector is given by $\underline{\xi} = [\xi^{(1)}, \dots, \xi^{(4)}]^T$, which completely describes the stochastic dependency between the random variables x_k and u_k .

In Fig. 9 (a)–(d) the parameterized joint density is depicted for variation of the individual free parameters $\xi^{(i)}$ from their minimum value to their maximum value, respectively.

In order to visualize that *all* possible joint densities can be described just by the generalized correlation parameter vector $\underline{\xi}$, various joint densities are depicted in Fig. 10 (a)–(d).

Furthermore, the predicted density $f_{k+1}^p(x_{k+1})$ depending on the generalized correlation parameter vector $\underline{\xi}$ can be derived, which is shown in the lower row in Fig. 9 (a)–(d) and Fig. 10 (a)–(d).

For a next step toward a nonlinear decentralized filtering technique robust against unknown correlation it would be necessary to find a bounding density which somehow contains the density functions of all possible, regarding $\underline{\xi}$ and/or $\xi(r)$.

5 Conclusion and Future Work

This paper focuses on the parameterization of different types of joint densities both with Gaussian marginals and Gaussian mixture marginals. It is shown that assuming a specific stochastic dependency structure these joint densities contain all information about this dependency in terms of a *generalized correlation parameter vector* $\underline{\xi}$ and/or a *generalized correlation function* $\xi(r)$. Unlike the *classical* correlation coefficient r the generalized correlation parameter vector $\underline{\xi}$ is a sufficient measure for the stochastic dependency between two random variables represented by Gaussian mixtures.

More detailed prediction results and measurement results depending on the generalized correlation parameter will be covered in a forthcoming paper. Finding a bounding density fulfilling the stochastic dependency

constraints, and thus containing all possible values of the generalized correlation parameter $\underline{\xi}$, is left for future research work; a possible direction for solving this problem can be found in [10]. It is obvious that once such a bounding density is found, a filtering technique can be derived, which can cope with *nonlinear models* and is robust against imprecisely known stochastic dependencies.

6 Acknowledgement

This work was partially supported by the German Research Foundation (DFG) within the Research Training Group GRK 1194 “Self-organizing Sensor-Actuator-Networks”.

References

- [1] J. A. Castellanos, J. D. Tardos, G. Schmidt, “Building a Global Map of the Environment of a Mobile Robot: The Importance of Correlations,” *IEEE International Conference on Robotics and Automation (ICRA’97)*, Albuquerque, New Mexico, USA, 1997.
- [2] S. Oh, S. Sastry, L. Schenato, “A Hierarchical Multiple-Target Tracking Algorithm for Sensor Networks,” *Proc. of the 2005 IEEE International Conference on Robotics and Automation (ICRA’05)*, Barcelona, Spain, 2005.
- [3] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw Hill Book Company, 1984.
- [4] F. Sawo, K. Roberts, U. D. Hanebeck, “Bayesian Estimation of Distributed Phenomena Using Discretized Representations of Partial Differential Equations,” *Proc. of the 3rd International Conference on Informatics in Control, Automation and Robotics (ICINCO’06)*, Setúbal, Portugal, 2006.
- [5] J. Uhlmann, S. Julier, M. Csorba, “Nondivergent Simultaneous Map Building and Localization Using Covariance Intersection,” *Proc. of the 1997 SPIE Aerosense Conference*, Orlando, Florida, 1997.
- [6] C. Y. Chong, S. Mori, “Convex Combination and Covariance Intersection Algorithms in Distributed Fusion,” *Proc. of the 4th International Conference on Information Fusion 2001*, Montreal, Canada, 2001.
- [7] U. D. Hanebeck, K. Briechle, J. Horn, “A Tight Bound for the Joint Covariance of Two Random Vectors with Unknown but Constrained Cross-Correlation,” *IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI’01)*, Baden-Baden, Germany, 2001.
- [8] M. B. Hurley, “An Information Theoretic Justification for Covariance Intersection and Its Generalization,” *Proc. of the 5th International Conference on Information Fusion 2002*, Annapolis, Maryland, USA, 2002.
- [9] D. L. Alspach, H. W. Sorenson, “Nonlinear Bayesian Estimation Using Gaussian Sum Approximation,” *IEEE Transactions on Automatic Control*, Vol. 17, No. 4, pp. 439–448, 1972.
- [10] V. Klumpp, D. Brunn, U. D. Hanebeck, “Approximate Nonlinear Bayesian Estimation Based on Lower and Upper Densities,” *Proc. of the 9th International Conference on Information Fusion 2006*, Florence, Italy, 2006.
- [11] U. D. Hanebeck, K. Briechle, A. Rauh, “Progressive Bayes: A New Framework for Nonlinear State Estimation,” *Proceedings of SPIE, AeroSense Symposium*, 2003.

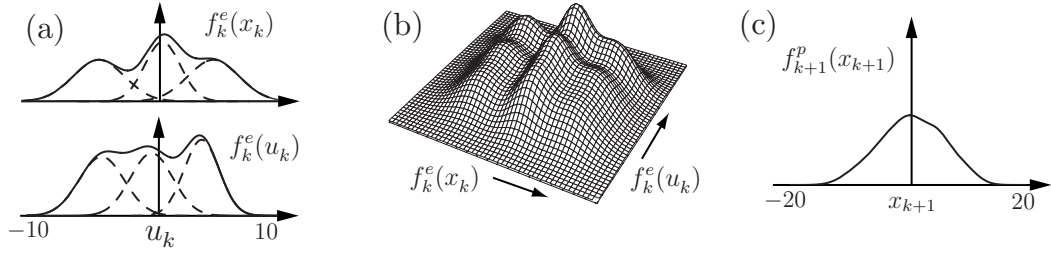


Figure 8: (a) Gaussian mixture marginal densities $f_k^e(x_k)$ and $f_k^e(u_k)$, and (b) their joint density $f(x_k, u_k)$ for the uncorrelated case, i.e. $\xi^{(i)} = 0$, (c) predicted density $f_{k+1}^p(x_{k+1})$.

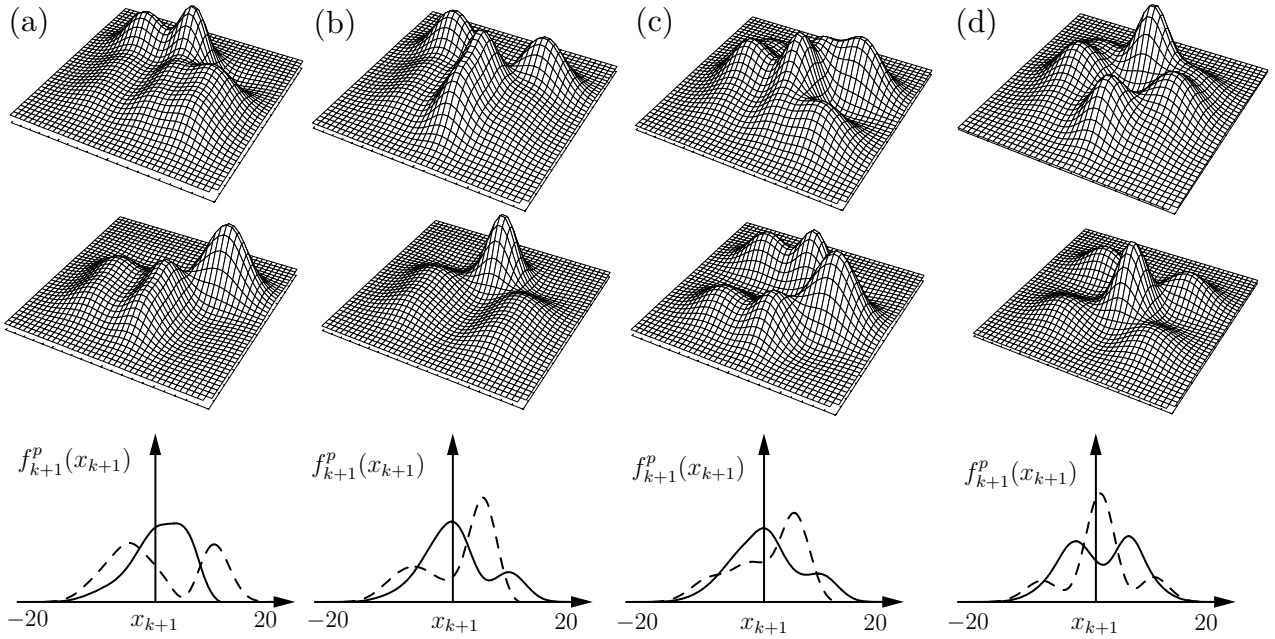


Figure 9: Parameterized joint density $f(x_k, u_k)$ and predicted density $f_{k+1}^p(x_{k+1})$ for variation of the individual free parameter $\xi^{(i)}$ from their minimum values to their maximum values, $\xi^{(i)} = 0$ except (a) $\xi^{(1)} = -1/3 \dots 2/3$, (b) $\xi^{(2)} = -1/3 \dots 2/3$, (c) $\xi^{(3)} = -1/3 \dots 2/3$, and (d) $\xi^{(4)} = -1/3 \dots 2/3$.

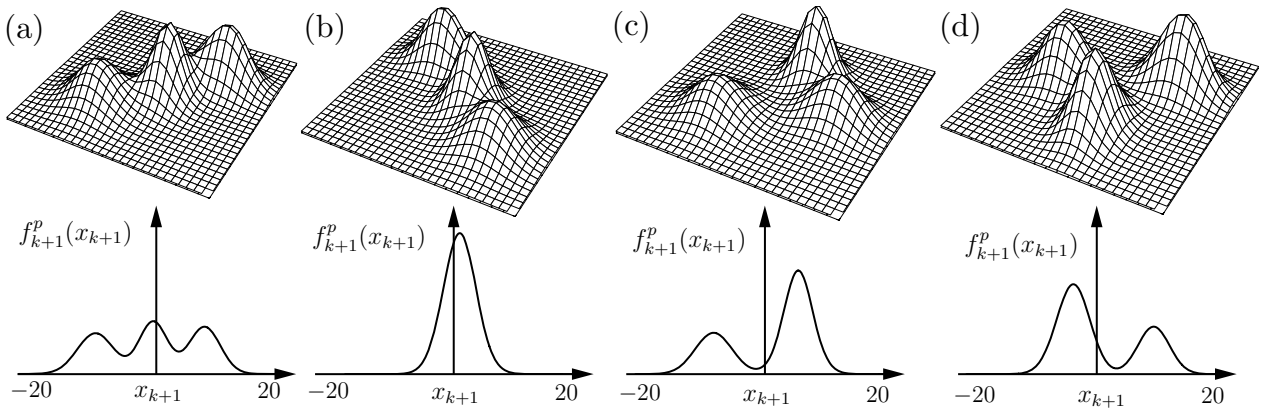


Figure 10: Parameterized joint density $f(x_k, u_k)$ and predicted density $f_{k+1}^p(x_{k+1})$ for various free parameter vectors, (a) $\underline{\xi} = [\frac{2}{3}, \frac{1}{3}, \frac{1}{3}, \frac{2}{3}]^T$, (b) $\underline{\xi} = [-\frac{2}{3}, -\frac{1}{3}, -\frac{1}{3}, \frac{1}{3}]^T$, (c) $\underline{\xi} = [\frac{1}{3}, \frac{2}{3}, \frac{2}{3}, \frac{1}{3}]^T$, and (d) $\underline{\xi} = [\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}, -\frac{2}{3}]^T$.