

Analysis of Set-theoretic and Stochastic Models for Fusion under Unknown Correlations

Marc Reinhardt, Benjamin Noack, Marcus Baum, and Uwe D. Hanebeck

Intelligent Sensor-Actuator-Systems Laboratory (ISAS),

Institute for Anthropomatics,

Karlsruhe Institute of Technology (KIT), Germany.

Email: marc.reinhardt@student.kit.edu, noack@kit.edu, marcus.baum@kit.edu, uwe.hanebeck@ieee.org

Abstract—In data fusion theory, multiple estimates are combined to yield an optimal result. In this paper, the set of all possible results is investigated, when two random variables with unknown correlations are fused. As a first step, recursive processing of the set of estimates is examined. Besides set-theoretic considerations, the lack of knowledge about the unknown correlation coefficient is modeled as a stochastic quantity. Especially, a uniform model is analyzed, which provides a new optimization criterion for the covariance intersection algorithm in scalar state spaces. This approach is also generalized to multi-dimensional state spaces in an approximative, but fast and scalable way, so that consistent estimates are obtained.

Keywords: filtering, estimation, fusion, Bayesian, correlation coefficient.

I. INTRODUCTION

In many practical applications, distributed sensor systems are utilized in order to take advantage of different angles, distances etc. By means of a Bayesian state estimator, the measurement information can be fused with the current estimate, uncertainties can be modeled and taken into account, and for further processing the obtained estimates can be predicted. From a central architecture, where all estimates and correlations between them are managed centrally, to fully distributed approaches, where data is processed and collected on different nodes and no information on cross-correlations is available, different distributed fusion architectures have been developed [1].

In this paper, we focus on linear estimation problems in distributed fusion architectures [1]–[4]. Distributed fusion algorithms have the advantage of lower infrastructure costs, such as communication or data storage expenses, and are robust to failures. The main challenge is to handle cross-correlations between the estimates, since ignoring correlations and applying standard Kalman filter equations for the fusion in general lead to inconsistent results. Suppose for example a distributed sensor network, where node B gets information from node A and the data of both nodes should be fused in node A. If we assume independence, the uncertainty is erroneously reduced due to the fusion although both nodes share the same information.

Different approaches to cope with the problem of unknown correlations have been developed. In particular, the covariance intersection algorithm (CI), which has been proposed by Julier and Uhlmann [5], [6], is often used as a baseline. Minimizing

the determinant of the fusion result is the most commonly used optimization criterion for the CI algorithm. Especially in scalar state spaces, this implies that CI does not update an estimate as long as no information with smaller variance is available. Since this is not desirable in most applications, we start our discussions with the impacts of the cross-correlation on the fusion result particularly in one-dimensional state spaces. We present closed-form equations for the interval of possible means and variances and show that the set of possible fusion results may diverge, when cross-correlations are not restricted. In a next step, we model the lack of knowledge about the correlation coefficient by a uniform distribution, i.e., as a uniform random variable on the interval $[-1, 1]$. In order to provide a practical estimator, we also derive closed-form solutions for mean and variance by marginalizing out the correlation variable. Based on these solutions, we derive a new optimization criterion for CI and generalize it to multi-dimensional state spaces.

II. PROBLEM FORMULATION

In data fusion theory, estimates are combined to yield an optimal fused estimate. The estimates characterize uncertain quantities, which are modeled by random variables.

This paper concentrates on the fusion of two estimates \underline{x} and \underline{y} to a resulting estimate \underline{z} , when the correlation between \underline{x} and \underline{y} is unknown. We denote the mean vectors by $\hat{\underline{x}}$, $\hat{\underline{y}}$, and $\hat{\underline{z}}$. The joint covariance matrix is

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}^x & \mathbf{C}^{xy} \\ \mathbf{C}^{yx} & \mathbf{C}^y \end{pmatrix}, \quad (1)$$

whereas the fused covariance matrix is \mathbf{C}^z . Let $\bar{\underline{\xi}}$ denote the true statistics, then the estimation errors are $\tilde{\underline{x}} = \bar{\underline{\xi}} - \hat{\underline{x}}$ and $\tilde{\underline{y}} = \bar{\underline{\xi}} - \hat{\underline{y}}$. Let $\mathbf{C}^{x*} = \mathbb{E}\{\tilde{\underline{x}}\tilde{\underline{x}}^T\}$ and $\mathbf{C}^{y*} = \mathbb{E}\{\tilde{\underline{y}}\tilde{\underline{y}}^T\}$ denote the unknown actual mean squared error (MSE) matrices. The input data are *consistent* estimates, if

$$\begin{aligned} \mathbf{C}^x - \mathbf{C}^{x*} &\geq \mathbf{0} \text{ and} \\ \mathbf{C}^y - \mathbf{C}^{y*} &\geq \mathbf{0}, \end{aligned}$$

i.e., if the difference between the matrices is a positive semi-definite matrix.

A central problem in distributed data fusion is to find an optimal estimate of the true statistics, if the cross-correlation matrices $\mathbf{C}^{xy} = \mathbf{C}^{yx^T}$ are unknown. Ignoring the cross-

correlations and applying a Kalman filter leads to inconsistent and possibly biased results. The standard CI algorithm, which minimizes the determinant or the trace of the fused covariance matrix, does not provide an intuitive procedure in scalar state spaces. Therefore, we investigate the optimal fusion results for different correlation coefficients to obtain approaches, that comprise all possible fusion results.

III. STATE OF THE ART

Especially in distributed fusion architectures, strong correlations between estimates can arise. For example, in distributed tracking algorithms for the same target, common process noise has to be dealt with [7], the sensors on different nodes may suffer from correlated noise, and, due to communication, network nodes may share common information. By employing hierarchical network topologies [1], [3], [4], the correlations between two estimates can be stored and exploited, when they are fused in a node. For the linear fusion of two estimates, there exists an optimal solution in the sense of a minimum mean squared error combination, which equals the best linear unbiased estimator (BLUE). This solution is given by the Bar-Shalom/Campo (BC) formulas [7]

$$\hat{\underline{z}} = (\mathbf{C}^y - \mathbf{C}^{yx}) (\mathbf{C}^x + \mathbf{C}^y - \mathbf{C}^{xy} - \mathbf{C}^{yx})^{-1} \hat{\underline{x}} + (\mathbf{C}^x - \mathbf{C}^{xy}) (\mathbf{C}^x + \mathbf{C}^y - \mathbf{C}^{xy} - \mathbf{C}^{yx})^{-1} \hat{\underline{y}} \quad (2)$$

and

$$\mathbf{C}^z = \mathbf{C}^x - (\mathbf{C}^x - \mathbf{C}^{xy}) (\mathbf{C}^x + \mathbf{C}^y - \mathbf{C}^{xy} - \mathbf{C}^{yx})^{-1} \cdot (\mathbf{C}^x - \mathbf{C}^{xy})^T. \quad (3)$$

A generalization to multiple input estimates is given by Millman's formulas, which are derived in [8].

In fully distributed networks, where the underlying network topology may remain unknown to the nodes and communication can lead to cycles, correlations between estimates can, in general, not be maintained and exploited. In this case, suboptimal fusion results can be computed by means of the CI algorithm [5], [6], [9], which yields a fused estimate with mean vector

$$\hat{\underline{z}}_\omega = \mathbf{C}_\omega^z \left(\omega \mathbf{C}^{x-1} \hat{\underline{x}} + (1 - \omega) \mathbf{C}^{y-1} \hat{\underline{y}} \right) \quad (4)$$

and covariance matrix

$$\mathbf{C}_\omega^z = \left(\omega \mathbf{C}^{x-1} + (1 - \omega) \mathbf{C}^{y-1} \right)^{-1}, \quad (5)$$

where $\omega \in [0, 1]$ is a weighting parameter. It can be shown that CI provides a conservative bound on the actual MSE matrix irrespective of the true cross-covariance matrix \mathbf{C}^{xy} and the choice of ω [5]. CI can also be derived in the joint state space of \underline{x} and \underline{y} [10], [11].

In general, ω is determined numerically in such a way that the determinant or trace of \mathbf{C}_ω^z is minimized. In [12] it has been shown, that CI with trace optimization criterion yields the covariance matrix with minimum trace in the set of all conservative covariance matrices, although only a scalar parameter is optimized. In order to avoid numerical optimization, approximate closed-form solutions have been proposed in [13], [14]. How CI works becomes apparent when multi-

dimensional estimates are fused, since CI can be considered as an outer ellipsoidal approximation of the intersection of the covariance ellipsoids that correspond to the estimates $(\hat{\underline{x}}, \mathbf{C}^x)$ and $(\hat{\underline{y}}, \mathbf{C}^y)$ centered at the origin $\underline{0}$. Minimizing the determinant of (5) then implies that the ellipsoidal approximation with minimum volume is chosen. In situations where one covariance ellipsoid is contained in the other covariance ellipsoid, i.e., $\mathbf{C}^x - \mathbf{C}^y$ or vice versa is positive definite, ω becomes 0 or 1. This means that the fusion of $(\hat{\underline{x}}, \mathbf{C}^x)$ and $(\hat{\underline{y}}, \mathbf{C}^y)$ yields the estimate with the smaller covariance matrix. Especially in one-dimensional setups, this happens in every fusion step, which is undesirable, because, for instance, a node would place greater trust in a single possible outlier with small variance than in many sources that report estimates with high variances. In this regard, information-theoretic [15], [16] and set-theoretic [17], [18] optimization criteria for ω have been proposed, where the choice of ω also depends on the means $\hat{\underline{x}}$ and $\hat{\underline{y}}$.

In this paper, we aim at modeling and analyzing the lack of knowledge about the correlation between two estimates. At first, we model the ignorance by the set of all possible correlations and set up a recursive estimator. Since we calculate an estimate for every possible correlation, we obtain a set of estimates, which corresponds to a set of Gaussian densities [19]. The second approach is to model the correlation coefficient as a uniform distribution, which is the way in which Bayesians often model ignorance. This is a special case of extending the density of the system state as it is described in [20].

IV. SET-THEORETIC APPROACH

In the following, the BC formulas for scalar valued random variables are investigated to determine the influence of the correlation coefficient, which leads to closed-form equations for the extrema of mean and variance of the fused estimates. By means of these extrema, it can be shown that further restrictions or assumptions on the correlation coefficient are necessary to constrain the solution sets.

A. Analysis of the Correlation Coefficient in Scalar State Space

The influence of the cross-correlation on the fusion result is investigated in one-dimensional state space on the basis of the equations for the fused mean (2) and variance (3).

In one-dimensional state space, the cross-correlation can be described by a single scalar correlation coefficient r , where

$$C_r^{xy} = r \sqrt{C^x C^y}. \quad (6)$$

We use (6) to obtain

$$\begin{aligned} \det(C) &= C^x C^y - r^2 C^x C^y \\ &= (1 - r^2) C^x C^y \end{aligned}$$

and finally $|r| \leq 1$ holds, since the covariance matrix is defined to be positive semi-definite. Using this restriction, it is possible to constrain the set of fusion results.

The fusion equations (2) and (3) can be simplified in one-dimensional state space to

$$\hat{z}_r = \frac{C^y \hat{x} + C^x \hat{y} - r\sqrt{C^x C^y}(\hat{x} + \hat{y})}{C^x + C^y - 2r\sqrt{C^x C^y}} \quad (7)$$

and

$$C_r^z = \frac{C^x C^y (1 - r^2)}{C^x + C^y - 2r\sqrt{C^x C^y}}. \quad (8)$$

LEMMA 1

The fused variance C^z for interval-valued $r = [-1, 1]$ and prior variances C^x and C^y is an interval with lower bound 0 and upper bound $\min\{C^x, C^y\}$.

PROOF.

Taking the derivative of (8) with respect to r leads to

$$\frac{\partial C_r^z}{\partial r} = \frac{-2C^x C^y (C^x r + C^y r - \sqrt{C^x C^y} (1 + r^2))}{(C^x + C^y - 2\sqrt{C^x C^y} \cdot r)^2}.$$

The necessary condition

$$\frac{\partial C_r^z}{\partial r} \stackrel{!}{=} 0$$

gives

$$r_1 = \sqrt{\frac{C^x}{C^y}} \text{ and } r_2 = \sqrt{\frac{C^y}{C^x}}$$

as candidates for extrema. The second derivative validates the variance with $r = \min\{r_1, r_2\}$ as true maximum and the variance with $r = \max\{r_1, r_2\}$ as true minimum. The maximum variance is obtained for the coefficient

$$r_{ext} = \sqrt{\frac{\min\{C^x, C^y\}}{\max\{C^x, C^y\}}}.$$

Inserting r_{ext} into (8) leads to $\min\{C^x, C^y\}$. \square

It follows from Lemma 1 that in scalar state space, CI provides a tight bound on the set of fused variances with trace or determinant optimization criterion. This result confirms the upper bound for the fused covariance matrix, which has been derived in [12].

LEMMA 2

The posterior mean \hat{z} for interval-valued $r = [-1, 1]$, prior means \hat{x} , \hat{y} , and variances C^x , C^y is an interval given by

$$\hat{z} = \begin{cases} [\hat{x}, \hat{x}] & \text{for } \hat{x} = \hat{y}, \\ [\hat{z}_1, \hat{z}_{-1}] & \text{for } \hat{x} > \hat{y}, C^x > C^y \text{ or } \hat{x} < \hat{y}, C^x < C^y, \\ [\hat{z}_{-1}, \hat{z}_1] & \text{elsewhere.} \end{cases}$$

PROOF.

Taking the derivative of (7) with respect to the correlation coefficient leads to

$$\frac{\partial \hat{z}_r}{\partial r} = \frac{(C^y - C^x)\sqrt{C^x C^y}(\hat{x} - \hat{y})}{(C^x + C^y - 2r\sqrt{C^x C^y})^2}. \quad (9)$$

From (9), it directly follows that for the three different cases \hat{z}_r is 1. equal, 2. strictly monotonically decreasing, or 3. strictly monotonically non-decreasing. \square

For further investigation of the influence of the correlation coefficient, the means and variances of the fused random

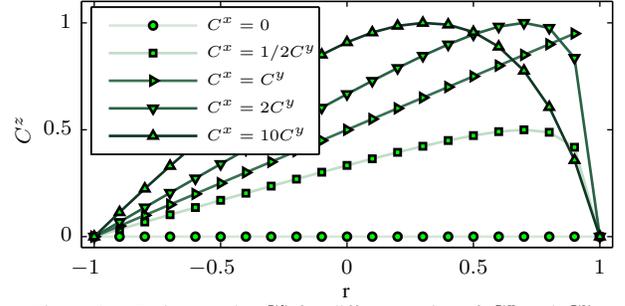


Figure 1. Fusion results C^z for different ratios of C^x and C^y .

variable are analyzed. The effects of the cross-correlation on the variance of the fused estimate can be seen in Fig. 1. For correlation coefficients $r = 1$ and $r = -1$, C_r^z equals 0. For input estimates with the same variance, C_r^z is linear with respect to the correlation coefficient on the interval $[-1, 1]$. In all other cases, C_r^z is a concave function with maximum at

$$r = \sqrt{\frac{C^x}{C^y}}.$$

In Fig. 2, the mean \hat{z}_r is plotted for a varying variance

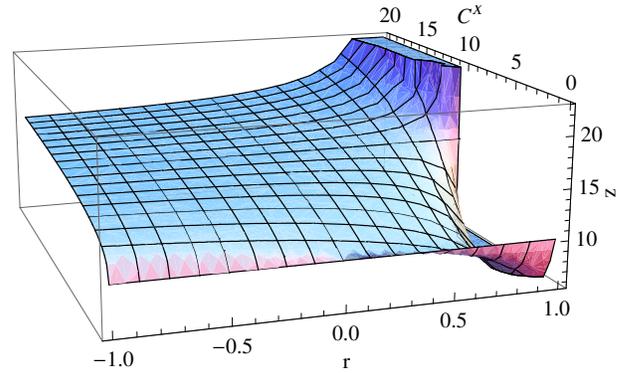


Figure 2. Fusion result \hat{z}_r depending on correlation coefficient $r \in [-1, 1]$ for different input variances C^x with $C^y = 10$, $\hat{x} = 10$ and $\hat{y} = 20$.

C^x and for all valid correlation coefficients $r \in [-1, 1]$. For $C^x = 0$, the resulting mean equals \hat{x} for all correlations, as there is no uncertainty in estimate x . The mean of the fused estimate increases with the correlation coefficient to values considerably larger than \hat{x} for

$$r > \sqrt{\frac{C^x}{C^y}}.$$

For input estimates with the same variance, \hat{z}_r is $\frac{1}{2}(\hat{x} + \hat{y})$, since both estimates have the same uncertainty and thus, they are equally weighted. If $C^x < C^y$, all possible results of \hat{z}_r are larger than $\frac{1}{2}(\hat{x} + \hat{y})$ and visa versa. For $r = 1$, the mean of the fused estimate converges for $C^x \rightarrow C^y$ to infinity or minus infinity.

B. Recursive Fusion

With the given Lemmata, it is not possible to recursively formulate the mean extrema of the fused estimates. To allow for such a description, the input parameters need to be extended to mean and variance intervals or a set of all possible estimates needs to be considered. In case of an

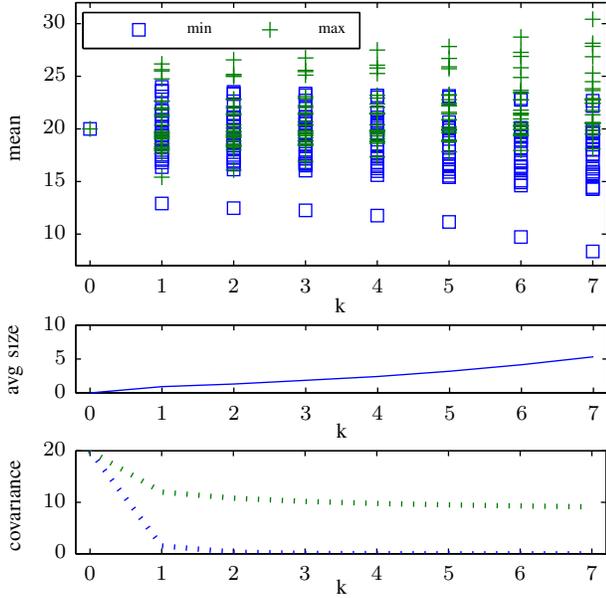


Figure 3. Minimum and maximum mean values, the average mean interval size, and the covariance extrema in 25 Monte Carlo runs for the given example.

extension to interval analysis, it is inevitable to investigate the impact of treating the means and variances independently of each other. Simulation results indicate that the interval-based approach provides similar mean and variance extrema as the set-theoretic approach, in which we consider the *combined* means and variances. However, as we do not present closed-form equations for the recursive case and the extrema of the interval-valued approach are a conservative estimate of the true extrema, we demonstrate the divergence of the minimum and maximum means by a simple static example.

Assume an estimate $\mathbf{x}_0 \sim \mathcal{N}(\hat{x}_0, C_0^x)$, which is fused with measurements $\mathbf{y}_k \sim \mathcal{N}(\hat{y}_k, C_k^y)$ at time steps $k \in \{1, \dots, T\}$. The correlation coefficient is symmetrically bounded by r_b , so that only correlation coefficients in the interval $[-r_b, r_b]$ are allowed. We discretize $[-r_b, r_b]$ at each time step at equidistant r_i and for each correlation coefficient, we simulate an estimator. With the number of discretization points H , we obtain

$$r_i = \frac{2r_b \cdot i}{H-1} - r_b \text{ for } i \in \{0, \dots, H-1\} .$$

Thus, the number of estimators at time step k is H^k . Let \hat{x}_k^{sim} denote the interval with the minimum and maximum mean of the simulation at time step k as boundaries, $\hat{x}_k^{r_b}$ the corresponding interval for continuous correlation coefficients, and \hat{x}_k^{all} the corresponding entire interval with continuous $r = [-1, 1]$. It is easy to see that

$$\hat{x}_k^{sim} \subseteq \hat{x}_k^{r_b} \subseteq \hat{x}_k^{all} .$$

We set $r_b = 0.8$, $H = 15$, $T = 7$, $\mathbf{x}_0 = \mathcal{N}(20, 20)$, $\mathbf{y}_k \sim \mathcal{N}(20, 12)$ and perform 25 Monte Carlo runs with the BC combination algorithm (7) and (8). The results are given in Fig. 3. The variances at each time instant are the same for all Monte Carlo runs since their calculation does not depend on the measurements. While the minimum variance decreases

to zero, it can be shown that for $C^y < C^x$, the maximum variance for continuous correlation coefficients converges to $r_b^2 C^y = 0.8^2 \cdot 12 = 7.68$. As can be seen, the interval of means \hat{x}_k^{sim} is increasing at each time instant.

It follows that besides the calculation and memory efforts, the naive recursive fusion with reasonably bounded correlation coefficients is not applicable over multiple time steps. Thus, it is necessary to make further assumptions or restrictions on the cross-correlation between the input estimates to provide a practical estimator.

V. STOCHASTIC CORRELATION COEFFICIENT MODEL

The idea behind the *Stochastic Correlation Coefficient Model* is to generalize the correlation coefficient from a deterministic value to a density. In the case of ignorance, it is reasonable to assume a uniform distribution for the correlation coefficient. Other densities may also be meaningful but will not be investigated here. We will assume that the correlation coefficient is uniformly distributed and give a closed-form equation for the scalar-valued fusion and an approximate solution for the vector-valued fusion. The presented procedure will be called *uniform distribution* (UD) approach in this section.

A. Optimal Fusion of Scalar-Valued Random Variables for Uniformly Distributed Correlation Coefficients

The interval of correlation coefficients is again discretized into the set $\{r_1, \dots, r_n\}$. For each r_i the BC combination is computed, and finally the average of the fused densities is calculated. The resulting density is a Gaussian Mixture with uniform weights, which is approximated with a Gaussian density by moment matching.

For $\max(|r_{i-1} - r_i|) \rightarrow 0, \forall i$, the resulting density can be written as the integral

$$f(z) = \int_{-1}^1 f(z, r) dr = \int_{-1}^1 f(z|r) f(r) dr ,$$

which can be simplified in the case of a uniformly distributed correlation coefficient to

$$f(z) = \frac{1}{2} \int_{-1}^1 f(z|r) dr = \frac{1}{2} \int_{-1}^1 \mathcal{N}(\hat{z}_r, C_r^z) dr . \quad (10)$$

Marginalization of (10) finally leads to

$$E^{GM} = \frac{\hat{x} + \hat{y}}{2} - \frac{(C^x - C^y)(\hat{x} - \hat{y}) \ln \left(\frac{C^x + C^y}{|C^x - C^y|} \right)}{4\sqrt{C^x C^y}} \quad (11)$$

and

$$C^{GM} = \frac{1}{4} (C^x + C^y + (\hat{x} - \hat{y})^2) - \frac{(C^x - C^y)^2 \ln \left(\frac{C^x + C^y}{|C^x - C^y|} \right)}{16C^x C^y} \left((\hat{x} - \hat{y})^2 \ln \left(\frac{C^x + C^y}{|C^x - C^y|} \right) + 2\sqrt{C^x C^y} \right) . \quad (12)$$

A derivation for symmetrically bounded correlation coefficients can be found in appendix A.

It can be seen in (11), that the fused mean is the average of the two input means plus a correction term, which depends on

the difference of the input means as well as the difference and the absolute values of the input covariance matrices. Furthermore, the fused mean is corrected by the difference term in direction of the input mean with the smaller covariance matrix.

The formula for the fused covariance matrix (12) describes the uncertainty of the fusion. Unlike, e.g., in CI, the means are also involved in the calculation of the covariance matrix. This allows an estimate of the uncertainty of the fusion step and gives a crude measurement of the bandwidth of the BC fusion results.

A non-recursive example proposed in [17] demonstrates the performance of the algorithm in calculating the best mean.

B. Non-recursive Example

We fuse two scalar-valued Gaussian estimates $x_i \sim \mathcal{N}(\hat{x}_i, C_i^x)$, $i \in \{1, 2\}$ of the same true underlying statistics $\bar{x} \sim \mathcal{N}(\bar{x}, \bar{C}^x)$. Assume two measurements

$$z_i = \bar{x} + v_i, i \in \{1, 2\}$$

$$\text{with } \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} C_1^v & C_{12}^v \\ C_{12}^v & C_2^v \end{pmatrix} \right).$$

The estimates x_i can be obtained by

$$\hat{x}_i = \bar{x} + \bar{C}^x (C_i^x + \bar{C}^x)^{-1} (z_i - \bar{x})$$

$$\text{and } C_i^x = \left((\bar{C}^x)^{-1} + (C_i^z)^{-1} \right)^{-1}.$$

The cross-correlation between x_1 and x_2 is

$$C_{12}^x = C_1^x \bar{C}^x C_2^x T + K_1 C_{12}^v K_2^T$$

$$\text{with } K_i = \bar{C}^x (C_i^x + \bar{C}^x)^{-1}.$$

The underlying density is simulated for different cross-correlations $C_{12}^v = r \sqrt{C_1^v C_2^v}$ in 1000 Monte Carlo runs each. If C_{12}^x is known, the best possible results are given by the BC formulas (2) and (3), which serve as a baseline. For unknown correlation coefficients, the UD approach is compared to CI with determinant optimization criterion. In

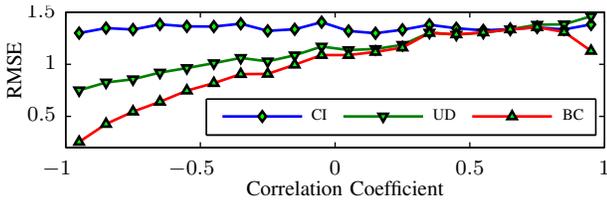


Figure 4. The results of the non-recursive example depending on a varying correlation coefficient.

Fig. 4, the Root Mean Squared Error (RMSE) for varying correlation coefficient and the input variables

$$\bar{x} \sim \mathcal{N}(1, 20) \text{ and } C_1^v = 2.22, C_2^v = 5$$

is shown. As can be seen, the algorithm performs well especially for low correlations.

C. Application as Optimization Criterion for CI

For multistep fusion and practical applications, it is essential to obtain a consistent estimate. Let \hat{z}_{ud} be the mean calculated in (11). The derived variance (12) is an approximation of the

uncertainty of the fusion step and is, in general, not consistent, because it is derived from an average over all densities. Therefore, we use CI to find an estimate that corresponds to \hat{z}_{ud} . Or the other way round: we use UD as optimization criterion for CI.

The CI equations are given by (4) and (5). We set \hat{z}_{ud} equal to the parametrized mean, which is derived by CI equations

$$\hat{z}_{ud} = \hat{z}_{\omega_{ud}} = C_{\omega_{ud}}^z \left(\frac{\omega_{ud} \hat{x}}{C^x} + \frac{(1 - \omega_{ud}) \hat{y}}{C^y} \right)$$

$$= \frac{C^x \hat{y} + \omega_{ud} (C^y \hat{x} + C^x \hat{y})}{C^x + \omega_{ud} (C^y - C^x)},$$

and obtain the closed-form solution

$$\omega_{ud} = \left(1 - \frac{C^y (\hat{z}_{ud} - \hat{x})}{C^x (\hat{z}_{ud} - \hat{y})} \right)^{-1}. \quad (13)$$

We need to show that ω_{ud} (13) is a valid CI optimization parameter for all combinations of input information. Replacing \hat{z}_{ud} by (11) and simplifying leads to

$$\frac{C^y (\hat{z}_{ud} - \hat{x})}{C^x (\hat{z}_{ud} - \hat{y})} = \frac{C^y}{C^x} \left(\frac{\ln \left(\frac{C^x + C^y}{|C^x - C^y|} \right) - 2 \frac{\sqrt{C^x C^y}}{C^x - C^y}}{\ln \left(\frac{C^x + C^y}{|C^x - C^y|} \right) + 2 \frac{\sqrt{C^x C^y}}{C^x - C^y}} \right).$$

It can be shown that this term is always negative or zero for positive C^x and C^y . If we insert this result in (13), it follows that $\omega_{ud} \in [0, 1]$. Therefore, in scalar state space, a valid CI solution can be calculated for all \hat{z}_{ud} in closed-form without the need for optimization algorithms.

D. Extension to Multi-Dimensional State Space

The generalization of UD to multi-dimensional state space is difficult to derive in closed-form because not only the number of correlation coefficients increases quadratically with the dimension of the data, but also the correlation coefficients depend on each other in more than two-dimensional joint covariance matrices. More precisely, the joint covariance matrix of dimension n can be written as in (1) with completely known C^x and C^y . The cross-correlation matrices can be described by correlation coefficients $r_{ij} \in [-1, 1]$ with

$$C^{xy} = \begin{pmatrix} r_{11} \sqrt{C_{11}^x C_{11}^y} & \cdots & r_{1n} \sqrt{C_{nn}^x C_{nn}^y} \\ \vdots & \ddots & \vdots \\ r_{n1} \sqrt{C_{nn}^x C_{11}^y} & \cdots & r_{nn} \sqrt{C_{nn}^x C_{nn}^y} \end{pmatrix},$$

where $C^{xy} = C^{yxT}$. Because covariance matrices are defined to be positive semi-definite, not all combinations of correlations coefficients are valid. There exist several algorithms to find out whether a matrix is positive semi-definite or not. Well known in literature is for example Sylvester's criterion [21]. According to Sylvester's criterion, all leading principal minors have to be non-negative. This is especially applicable when working with correlation coefficients, as the variances C_{ii}^x and C_{jj}^y can be factorized out of the determinant so that it is only necessary to check the correlation coefficient combinations.

For example in the three-dimensional case, the one- and two-dimensional leading principal minors are positive semi-

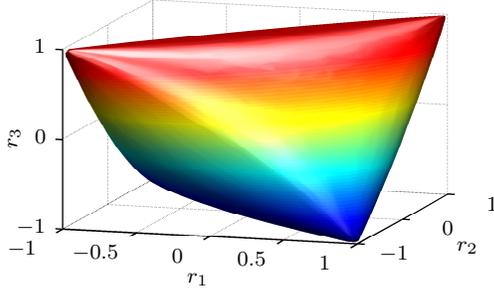


Figure 5. Structure of valid correlation coefficients in 3D space.

definite. The determinant of a covariance matrix is

$$\det(\mathbf{C}) = \begin{vmatrix} C_1^x & r_{12}\sqrt{C_1^xC_2^x} & r_{13}\sqrt{C_1^xC_3^x} \\ r_{12}\sqrt{C_1^xC_2^x} & C_2^x & r_{23}\sqrt{C_2^xC_3^x} \\ r_{13}\sqrt{C_1^xC_3^x} & r_{23}\sqrt{C_2^xC_3^x} & C_3^x \end{vmatrix} \\ = C_1^xC_2^xC_3^x(1 + 2r_{12}r_{13}r_{23} - r_{12} - r_{13} - r_{23}).$$

Thus, all correlation coefficient combinations with $|r_{12}| \leq 1$, $|r_{13}| \leq 1$, $|r_{23}| \leq 1$ that fulfill the inequality

$$1 + 2r_{12}r_{13}r_{23} - r_{12} - r_{13} - r_{23} \geq 0$$

are valid. In Fig. 5, these correlation coefficient combinations are plotted. As can be seen, the structure of valid correlation coefficients in three-dimensional space is convex, but cannot be easily described by a simple shape. To obtain a closed-form solution, it is possible to find out the multi-dimensional intervals of valid correlation coefficients and solve the multi-dimensional extension of the integral (10). Although this procedure is possible for small dimensions, the equations are complicated even in three dimensions.

Since an integral approximation by Monte Carlo methods is computationally demanding due to the curse of dimensionality, another approximate approach is presented. The idea is to fuse the multi-dimensional densities component-wise with equation (11) and find a consistent covariance matrix afterwards. Formally, we assume $r_{ij} = 0$, $i \neq j$ for correlation coefficients in the joint covariance matrix and a component-wise uniform distribution of the correlation coefficients r_{ii} .

Let n denote the dimension of the state space and $\hat{z}_{ud} = (\hat{z}_{ud1}, \dots, \hat{z}_{udn})^T$ be the vector of the means, which have been derived component-wise by scalar UD. From \hat{z}_{ud} , we obtain the gain

$$\mathbf{K} = \text{diag} \left(\frac{\hat{z}_{ud1} - \hat{y}_1}{\hat{x}_1 - \hat{y}_1}, \dots, \frac{\hat{z}_{udn} - \hat{y}_n}{\hat{x}_n - \hat{y}_n} \right),$$

which can be applied to the linear fusion equation

$$\mathbf{z} = \mathbf{K}\mathbf{x} + (\mathbf{I} - \mathbf{K})\mathbf{y}. \quad (14)$$

In [12] it has been shown, that for linear fusion in the form (14), a family of consistent estimates can be derived. This family is given by

$$\mathbf{C}^z = (1 + \gamma)\mathbf{K}\mathbf{C}^x\mathbf{K}^T + \left(1 + \frac{1}{\gamma}\right)(\mathbf{I} - \mathbf{K})\mathbf{C}^y(\mathbf{I} - \mathbf{K})^T$$

for a scalar parameter $\gamma > 0$. The covariance matrix \mathbf{C}^z with

minimum trace in this family is given for the parameter

$$\gamma = \sqrt{\frac{\text{trace} \left((\mathbf{I} - \mathbf{K})\mathbf{C}^y(\mathbf{I} - \mathbf{K})^T \right)}{\text{trace}(\mathbf{K}\mathbf{C}^x\mathbf{K}^T)}}.$$

It is also possible to optimize the determinant of \mathbf{C}^z . Since this procedure provides similar results but requires numerical optimization, we will not consider it here.

Therefore, by using the scalar UD approach to derive the mean and the procedure presented in [12] to estimate the covariance matrix, we provide a fast and scalable, suboptimal procedure with low memory requirements and low computational costs. In the following, this approach is optimized, since we do not consider all available information in naive UD.

E. Joint Diagonalization of Covariance Matrices

The presented generalization of the UD approach to multi-dimensional state spaces has its main drawback in the negligence of the cross-correlations between different parts of the state variables. In particular, we do not use the off-diagonal elements of the known input matrices \mathbf{C}^x and \mathbf{C}^y . Hence, we present a procedure to factor the known cross-correlations within the input covariances into the calculation of the mean.

Let the joint covariance matrix be as in (1) with unknown matrices \mathbf{C}^{xy} and \mathbf{C}^{yx} . Before the fusion algorithm is applied, it is possible to jointly rotate and scale the covariance matrices \mathbf{C}^x and \mathbf{C}^y to a diagonal form. This corresponds to a rotation of the covariance ellipsoids as it is described in [22] and allows an inclusion of the off-diagonal elements of the known matrices \mathbf{C}^x and \mathbf{C}^y .

The diagonal form can be obtained by the following procedure: First, an eigenvalue decomposition is performed for matrix \mathbf{C}^x . Let \mathbf{B}^x be the matrix of eigenvectors of matrix \mathbf{C}^x . With diagonal eigenvalue matrix \mathbf{D}^x and $\mathbf{B}^x\mathbf{B}^{xT} = \mathbf{I}$ it follows

$$\mathbf{C}^x = \mathbf{B}^x\mathbf{D}^x\mathbf{B}^{xT} = \mathbf{B}^x\sqrt{\mathbf{D}^x}\mathbf{B}^{xT}\mathbf{B}^x\sqrt{\mathbf{D}^x}\mathbf{B}^{xT}. \quad (15)$$

The decomposition (15) can be used to obtain

$$\mathbf{T}^x = \left(\mathbf{B}^x\sqrt{\mathbf{D}^x}\mathbf{B}^{xT}\right)^{-1} = \mathbf{B}^x\sqrt{\mathbf{D}^x}^{-1}\mathbf{B}^{xT}, \quad (16)$$

while $\mathbf{T}^x\mathbf{C}^x\mathbf{T}^{xT} = \mathbf{I}$.

In a second step, an eigenvalue decomposition for transformed \mathbf{C}^y is performed. Let $\mathbf{C}^{y'} = \mathbf{T}^x\mathbf{C}^y\mathbf{T}^{xT}$ be the rotated and scaled covariance matrix of input data y . With an eigenvalue decomposition of $\mathbf{C}^{y'}$, we obtain the matrix of eigenvectors \mathbf{B}^y and the diagonal matrix of eigenvalues

$$\mathbf{D}^y = \mathbf{B}^y\mathbf{C}^{y'}\mathbf{B}^{yT}. \quad (17)$$

Finally, we combine (16) and (17) to get a transformation matrix

$$\mathbf{T} = \mathbf{B}^y\mathbf{T}^x \quad (18)$$

and diagonal covariance matrices \mathbf{I} for x and \mathbf{D}^y for y . The back-transformation matrix is given by

$$\mathbf{T}^{-1} = \mathbf{T}^{x-1}\mathbf{B}^{y-1} = \mathbf{B}^x\sqrt{\mathbf{D}^x}\mathbf{B}^{xT}\mathbf{B}^yT.$$

We use the transformation (18) to obtain transformed input

variables

$$\bar{\underline{x}} = \mathbf{T}\hat{\underline{x}}, \bar{\underline{y}} = \mathbf{T}\hat{\underline{y}}$$

and

$$\bar{\mathbf{C}}^x = \mathbf{I}, \bar{\mathbf{C}}^y = \mathbf{D}^y, \bar{\mathbf{C}}^{xy} = \mathbf{T}\mathbf{C}^{xy}\mathbf{T}^T.$$

In order to show that the transformation does not distort the results, we proof that for given correlations, the BC equations yield the same result, when they are applied to the original data as when they are applied to the transformed data and are back-transformed afterwards:

$$\begin{aligned} \hat{\underline{z}} &= \mathbf{T}^{-1}\mathbf{T}\hat{\underline{z}} \\ &= \mathbf{T}^{-1}\mathbf{T}\hat{\underline{x}} + \mathbf{T}^{-1}\mathbf{T}(\mathbf{C}^x - \mathbf{C}^{xy})\mathbf{T}^T\mathbf{T}^{-T} \\ &\quad (\mathbf{C}^x + \mathbf{C}^y - \mathbf{C}^{xy} - \mathbf{C}^{yx})^{-1}\mathbf{T}^{-1}\mathbf{T}(\hat{\underline{y}} - \hat{\underline{x}}) \\ &= \mathbf{T}^{-1}\left(\hat{\underline{x}} + (\bar{\mathbf{C}}^x - \bar{\mathbf{C}}^{xy}) \cdot \right. \\ &\quad \left. (\bar{\mathbf{C}}^x + \bar{\mathbf{C}}^y - \bar{\mathbf{C}}^{xy} - \bar{\mathbf{C}}^{yx})^{-1}(\bar{\underline{y}} - \bar{\underline{x}})\right), \\ \mathbf{C}^z &= \mathbf{T}^{-1}\mathbf{T}\mathbf{C}^z\mathbf{T}^T\mathbf{T}^{-T} \\ &= \mathbf{T}^{-1}(\mathbf{T}\mathbf{C}^x\mathbf{T}^T - \mathbf{T}(\mathbf{C}^x - \mathbf{C}^{xy})\mathbf{T}^T\mathbf{T}^{-T} \\ &\quad (\mathbf{C}^x + \mathbf{C}^y - \mathbf{C}^{xy} - \mathbf{C}^{yx})^{-1} \\ &\quad \mathbf{T}^{-1}\mathbf{T}(\mathbf{C}^x - \mathbf{C}^{xy})\mathbf{T}^T)\mathbf{T}^{-T} \\ &= \mathbf{T}^{-1}(\bar{\mathbf{C}}^x - (\bar{\mathbf{C}}^x - \bar{\mathbf{C}}^{xy})^T \\ &\quad (\bar{\mathbf{C}}^x + \bar{\mathbf{C}}^y - \bar{\mathbf{C}}^{xy} - \bar{\mathbf{C}}^{yx})^{-1}(\bar{\mathbf{C}}^x - \bar{\mathbf{C}}^{xy}))\mathbf{T}^{-T}. \end{aligned}$$

Thus, even if we determine the mean component-wise, we include the information of the off-diagonal elements without distorting the fusion results, when the input variables are intermediately transformed. This allows an improvement of UD, which will be denoted by OPT UD in the following.

The two proposed approaches provide good estimation results, when two estimates are fused under unknown correlations, as it will be shown in the next example.

F. Dynamic Example

The performance of both approaches is demonstrated by an example that has been proposed by Wang in [18]. We assume a dynamic system model

$$\underline{x}_{k+1} = \mathbf{A}\underline{x}_k + \mathbf{B}u_k$$

characterizing a moving target with constant velocity. Let \mathbf{A} and \mathbf{B} be given as

$$\mathbf{A} = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix} \text{ and } \mathbf{B} = \begin{pmatrix} T^2/2 \\ T \end{pmatrix}.$$

The input u_k is a zero-mean white Gaussian noise with covariance W in all time steps. The measurement system consists of two sensors. One sensor is measuring the position with variance C^{v_1} , the other one observes the velocity with variance C^{v_2} . In each time step, both sensors use a Kalman filter to predict the state and fuse it with their own measurements. Subsequently, their estimates are exchanged and combined with different data fusion algorithms in each node. For the compared algorithms, the results will be the same in both nodes.

The measurements are assumed to be correlated with a varying correlation coefficient r . The quality of the algorithms is measured by the RMSE with respect to the underlying state \underline{x} after 20 time steps. Let the measurement noises be $C^{v_1} = 10$ and $C^{v_2} = 6$. The system noise is $\mathbf{u}_k \sim \mathcal{N}(0, 4)$. The initial position is a Gaussian estimate with $\underline{x}_0 \sim \mathcal{N}((10, 5)^T, \text{diag}(100, 25))$. The average results of 1000

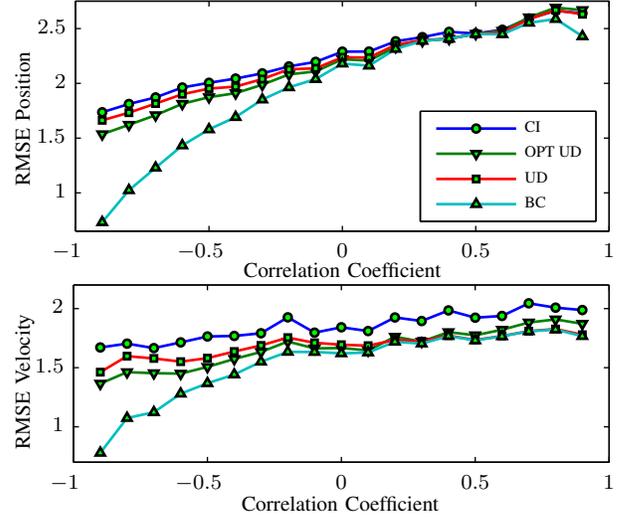


Figure 6. The results of the dynamic example for a varying correlation coefficient.

Monte Carlo runs for each correlation coefficient are plotted in Fig. 6. We compare the different optimization criteria for CI with the optimal BC combination. The optimization criteria are the minimization of the determinant (CI), UD without transformation (UD) and with transformation (OPT UD). In the given example, estimates that have been calculated by OPT UD have a lower RMSE than those calculated by CI. Especially for negative correlation, the RMSE of OPT UD decreases similarly to the RMSE of BC, while CI estimates are not improving in the same way.

For a detailed analysis of the relative performance between the different approaches, it would be necessary to run the given example with different uncertainties in system and measurement model. Tests have shown, that for low measurement uncertainties compared to the system noise, CI with determinant minimization criterion performs better than (OPT) UD for low correlation coefficients. Indeed, the computational effort especially for untransformed UD is lower since no numerical optimization is necessary.

VI. CONCLUSIONS

When cross-correlations are unknown, the fusion of two estimates can only yield suboptimal results. We have investigated the idea of calculating the entire set of possible fusion results. Considering every possibility, of course, constitutes the most careful and conservative approach to deal with an unknown parameter, which is the usual way in set-theoretic estimation. However, we have demonstrated by virtue of a simple example that the set of estimated means diverges with an increasing

number of fusion steps. As an alternative, we have modeled the uncertainty about the correlation coefficient by a probability density. In particular, we have assumed a uniform distribution for the correlation coefficient and marginalized the joint density over the correlation parameter. The resulting density has then been approximated by a Gaussian density, whose mean and variance can be computed analytically. This approach can be regarded as a weighted average over all possible fusion results. We have shown that the closed-form solution for the mean can also be employed to calculate a weight for the CI algorithm in scalar state space. As an extension to multi-dimensional state spaces, we have proposed to apply the presented concept component-wise, after rotating the covariance matrices to diagonal matrices, and to estimate the covariance afterwards. This approach guarantees conservative estimates, and yields promising fusion results.

APPENDIX A

UNIFORM DISTRIBUTED CORRELATION COEFFICIENT WITH SYMMETRIC BOUNDS

A solution for uniform distributed correlation coefficients in scalar state space with symmetric bounds r_b for mean and variance can be obtained by solving of the two integrals

$$\begin{aligned} E^{GM} &= \frac{1}{2r_b} \int_{-r_b}^{r_b} \hat{z}(r) dr \\ &= \frac{1}{2r_b} \int_{-r_b}^{r_b} \frac{C^y \hat{x} + C^x \hat{y} - r \sqrt{C^x C^y} (\hat{x} + \hat{y})}{C^x + C^y - 2r \sqrt{C^x C^y}} dr \\ &= \frac{\hat{x} + \hat{y}}{2} - (C^x - C^y) (\hat{x} - \hat{y}) \frac{\ln \left(\frac{2r_b \sqrt{C^x C^y} + C^x + C^y}{-2r_b \sqrt{C^x C^y} + C^x + C^y} \right)}{8r_b \sqrt{C^x C^y}}, \end{aligned}$$

$$\begin{aligned} C^{GM} &= \frac{1}{2r_b} \int_{-r_b}^{r_b} (C_k^z + \hat{z}(r)^2) dr - E^{GM^2} \\ &= \frac{C^{x^3} + C^x C^y (C^y (3 - 4r_b^2) - 2(\hat{x} - \hat{y})^2)}{4((C^x + C^y)^2 - 4C^x C^y r_b^2)} + \\ &\quad \frac{C^{x^2} (C^y (3 - 4r_b^2) + (\hat{x} - \hat{y})^2)}{4((C^x + C^y)^2 - 4C^x C^y r_b^2)} + \\ &\quad \frac{C^{y^2} (C^y + (\hat{x} - \hat{y})^2)}{4((C^x + C^y)^2 - 4C^x C^y r_b^2)} - \\ &\quad \frac{(C^x - C^y)^2 \ln \left(\frac{2r_b \sqrt{C^x C^y} + C^x + C^y}{-2r_b \sqrt{C^x C^y} + C^x + C^y} \right)}{64C^x C^y r_b^2} \left((\hat{x} - \hat{y})^2 \right. \\ &\quad \left. \ln \left(\frac{2r_b \sqrt{C^x C^y} + C^x + C^y}{-2r_b \sqrt{C^x C^y} + C^x + C^y} \right) + 4r_b \sqrt{C^x C^y} \right). \end{aligned}$$

REFERENCES

- [1] M. E. Liggins, D. L. Hall, and J. Llinas, Eds., *Handbook of Multisensor Data Fusion: Theory and Practice*, 2nd ed., ser. The Electrical Engineering and Applied Signal Processing Series. CRC Press, 2009.
- [2] S. Grime and H. Durrant-Whyte, "Data Fusion in Decentralized Sensor Networks," *Control Engineering Practice*, vol. 2, no. 5, pp. 849–863, 1994.
- [3] C.-Y. Chong and S. Mori, "Graphical Models for Nonlinear Distributed Estimation," in *Proceedings of the 7th International Conference on Information Fusion (Fusion 2004)*, Stockholm, Sweden, Jun. 2004.

- [4] M. E. Liggins, C.-Y. Chong, I. Kadar, M. G. Alford, V. Vannicola, and S. Thomopoulos, "Distributed Fusion Architectures and Algorithms for Target Tracking," in *Proceedings of the IEEE*, vol. 85, no. 1, 1997.
- [5] S. J. Julier and J. K. Uhlmann, "A Non-divergent Estimation Algorithm in the Presence of Unknown Correlations," in *Proceedings of the IEEE American Control Conference (ACC 1997)*, vol. 4, Albuquerque, New Mexico, USA, Jun. 1997, pp. 2369–2373.
- [6] —, *Handbook of Multisensor Data Fusion: Theory and Practice*, 2nd ed. CRC Press, 2009, ch. General Decentralized Data Fusion with Covariance Intersection, pp. 319–343.
- [7] Y. Bar-Shalom and L. Campo, "The Effect of the Common Process Noise on the Two-Sensor Fused-Track Covariance," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 22, no. 6, pp. 803–805, Nov. 1986.
- [8] Y. L. Vladimir Shin and T.-S. Choi, "Generalized Millman's Formula and its Application for Estimation Problems," vol. 86, pp. 257–266, 2006.
- [9] J. K. Uhlmann, "General Data Fusion for Estimates with Unknown Cross Covariances," in *Proceedings of SPIE Aerosense Conference*, vol. 2755, 1996, pp. 536–547.
- [10] U. D. Hanebeck and K. Briechle, "New Results for Stochastic Prediction and Filtering with Unknown Correlations," in *Proceedings of the 2001 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2001)*, Baden–Baden, Germany, Aug. 2001, pp. 147–152.
- [11] U. D. Hanebeck, K. Briechle, and J. Horn, "A Tight Bound for the Joint Covariance of Two Random Vectors with Unknown but Constrained Cross-Correlation," in *Proceedings of the 2001 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2001)*, Baden–Baden, Germany, Aug. 2001, pp. 85–90.
- [12] L. Chen, P. O. Arambel, and R. K. Mehra, "Fusion under Unknown Correlation - Covariance Intersection as a Special Case," in *Proceedings of the 5th International Conference on Information Fusion (Fusion 2002)*, Annapolis, Maryland, USA, Jul. 2002.
- [13] W. Niehsen, "Information Fusion based on Fast Covariance Intersection Filtering," in *Proceedings of the 5th International Conference on Information Fusion (Fusion 2002)*, Annapolis, Maryland, USA, Jul. 2002.
- [14] D. Fränken and A. Hüpper, "Improved Fast Covariance Intersection for Distributed Data Fusion," in *Proceedings of the 8th International Conference on Information Fusion (Fusion 2005)*, Philadelphia, Pennsylvania, USA, Jul. 2005.
- [15] M. B. Hurley, "An Information Theoretic Justification for Covariance Intersection and its Generalization," in *Proceedings of the 5th International Conference on Information Fusion (Fusion 2002)*, Annapolis, Maryland, USA, Jul. 2002.
- [16] Y. Wang and X. R. Li, "A Fast and Fault-Tolerant Convex Combination Fusion Algorithm under Unknown Cross-Correlation," in *Proceedings of 12th International Conference on Information Fusion (Fusion 2009)*, Seattle, Washington, USA, Jul. 2009.
- [17] C.-Y. Chong and S. Mori, "Convex Combination and Covariance Intersection Algorithms in Distributed Fusion," in *Proceedings of the 4th International Conference on Information Fusion (Fusion 2001)*, Montréal, Quebec, Canada, Aug. 2001.
- [18] Y. Wang and X. R. Li, "Distributed Estimation Fusion under Unknown Cross-correlation: An Analytic Center Approach," in *Proceedings of the 13th International Conference on Information Fusion (Fusion 2010)*, Edinburgh, United Kingdom, Jul. 2010.
- [19] B. Noack, V. Klumpp, D. Brunn, and U. D. Hanebeck, "Nonlinear Bayesian Estimation with Convex Sets of Probability Densities," in *Proceedings of the 11th International Conference on Information Fusion (Fusion 2008)*, Cologne, Germany, Jul. 2008, pp. 1–8.
- [20] V. Klumpp and U. D. Hanebeck, "Bayesian estimation with uncertain parameters of probability density functions," in *Proc. 12th Int. Conf. Information Fusion FUSION '09*, 2009, pp. 1759–1766.
- [21] G. T. Gilber, "Positive Definite Matrices and Sylvester's Criterion," *American Mathematical Monthly*, vol. 98, 1991.
- [22] J. Sijs, M. Lazar, and P. P. J. van den Bosch, "State-Fusion with Unknown Correlation: Ellipsoidal Intersection," in *Proceedings of the 2010 American Control Conference (ACC 2010)*, Baltimore, Maryland, USA, Jun. 2010.