# Linear Fusion under Random Correlation of Estimation Errors

1st Jiří Ajgl 2nd Ondřej Straka
*European Centre of Excellence – New Technologies for Information Society and*
*Department of Cybernetics, Faculty of Applied Sciences, University of West Bohemia*
Pilsen, Czech Republic
jiriajgl@kky.zcu.cz, straka30@kky.zcu.cz

***Abstract*—Linear fusion of estimates has been studied from the perspectives of known and unknown correlations of estimation errors. Whereas optimal linear combinations can be designed in the former case, a robust approach is usually chosen in the latter one. The loss of performance may be unacceptably high, which raises the need to find a middle ground. This paper reviews various approaches to information fusion, formulates the problem of random correlation and presents the solution. Monte Carlo verification of the results is discussed and an illustration is provided.**

***Index Terms*—stochastic systems, linear estimation, information fusion, unknown correlation, random correlation**

## I. Introduction

The probabilistic calculus can be used in many ways. Interpretation of probability [1]–[3] is not enforced by the nature, but it is a choice of the user. Nevertheless, once an interpretation is chosen, the framework has to be kept to. The authors of this paper distinguish the following elementary approaches to fusion inference under different interpretations.

The information geometry [4], [5] introduces differential geometry structures. The inference is based on choosing a parameter value from a statistical manifold. Various divergences can be proposed in order to define information fusion. The inferred point representing a probability distribution can lie between the points/distributions to be fused [6] or not far from those [7] for example. Fusion rules [8] have often the form of averages like the arithmetic or geometric means of densities [9], which are also called as linear and logarithmic opinion pools [10]. To sum up, the geometric approach is used to model the knowledge and its processing. It is therefore questionable, whether any properties can be attributed not only to the inference process, but to the results of the inference themselves. This approach may be called as subjective Bayesian, since it does not assume any objective ground truth.

Other approaches are based on the assumption of existence of a system. An abstract system, i.e. a model, is defined first for a real system. It is the user of the probability theory who is responsible for the correctness of the model; the profit is that properties can be attributed to the results of the inference. Once the measurement process is described, the observed data can be used.

This work was supported by the Czech Science Foundation, project no. P103/20-06054J.

If the quantity of the interest is considered to be stochastic, its distribution conditioned on the data can be searched. This approach may be called objective Bayesian, since it tries to find an objective information. Various approximations may be unavoidable due to computational complexity, nevertheless, it is not possible to depart from the philosophical framework, as it has been shown in [11].

The classical estimation deals with designing functions of data. The quality of these functions is evaluated after an experiment is specified, but before the data are observed. A typical criterion is the mean square error, which is to be minimised. For set-valued functions of data, the goal may be to achieve a confidence level, i.e. the probability that the random set contains the quantity of the interest [12]–[14]. Among the sets with the same confidence levels, the sets with small size are preferred.

The possibility to design the optimal solution is limited by the availability of the system description, which is not assumed in decentralised estimation. Local data can be processed locally, without knowing the whole system. Linear fusion under unknown correlation is considered in [15]–[17] for example. The estimates are constructed in a robust sense and they are accompanied with a self-assessment matrix. If the device that provides the estimate is correctly designed, the provided matrix is greater than or equal to the unknown mean square error matrix. The optimality in the robust sense has been shown in [18], observations useful for the construction of confidence sets have been presented in [19].

The fusion under unknown correlation can be improved by supplying a partial knowledge. Various sets of admissible correlation are discussed in [20]–[22] for example. There, the sets are smaller than the sets stemming from the condition of positive semi-definiteness of mean square error matrices. The idea of considering random correlation has appeared in [23] or [24] for example, nevertheless, incompatible philosophical frameworks have been tangled there.

The goal of the paper is to provide a proper formulation of the problem of fusion under random correlation and to find its solution. Section II formulates the problem. Notes on the existing approaches are given in Section III. The analysis of the problem and its solution are provided in Section IV, illustrations can be found in Section V. Section VI summarises the paper.

## II. Problem Formulation

It is assumed that a stochastic system exists and that its description is known. Let $x$ denote its state (i.e. the quantity to be estimated) and let the state lie within a known set $X$. If the state is stochastic, let its probability density function $p(x)$ be known.[1] Denote the measured data by $z$ and let the probability density $p(z|x)$ be known for each admissible value of $x$.

The mean-square-error-based estimation constructs functions $\hat{x}(z)$ of the random measurement $z$. The resulting random value $\hat{x}$ will be referred to as the estimate. The estimation error $\tilde{x}$ defined as $\tilde{x} = x - \hat{x}$ is random as well. The mean square error matrix $\Sigma$ is given by $\Sigma = \mathrm{E}\{\tilde{x}\tilde{x}^T\}$; it is a constant by definition and it does not depend on the realisation of $z$. If it holds $\mathrm{E}\{\tilde{x}\} = 0$, the estimator $\hat{x}(z)$ (or the random estimate $\hat{x}$) is said to be unbiased and the mean square error matrix $\Sigma$ is equal to the covariance matrix of the error $\tilde{x}$; note that unbiasedness is not related to a particular realisation of $z$.

The confidence sets $\mathcal{X}$ are random subsets of the set $X$ that are obtained by a mapping of $z$. The confidence level $\ell$, $0 \leq \ell \leq 1$, is a number with the property $P(x \in \mathcal{X}) \geq \ell$ and its typical choices are close to 1.

The formulation of the fusion problem follows by adding more assumptions. Section II-A proceeds with the linear fusion and the problem of random correlation is introduced in Section II-B.

### A. Linear Fusion

Let the measurement $z$ consist of two parts $z_1$ and $z_2$ that are used to produce estimates $\hat{x}_i$, $i = 1, 2$, with the errors $\tilde{x}_i$ and mean square error matrices $\Sigma_i$, respectively. The cross-correlation matrix $\Sigma_{1,2}$ of the estimation errors is given by $\Sigma_{1,2} = \mathrm{E}\{\tilde{x}_1\tilde{x}_2^T\}$. The fused estimate $\hat{x}$ is constructed by a linear combination of these estimates

$$\hat{x} = W_1\hat{x}_1 + W_2\hat{x}_2, \tag{1}$$

where $W_1$, $W_2$ are matrix weights constrained by a regularity condition $W_1 + W_2 = I$ with $I$ denoting the identity matrix. If both individual estimates $\hat{x}_i$ are unbiased, the condition guarantees the fused estimate $\hat{x}$ to be unbiased as well. Due to the condition, the mean square error matrix $\Sigma$ can be expressed as $\Sigma = W_1\Sigma_1W_1^T + W_1\Sigma_{1,2}W_2^T + W_2\Sigma_{1,2}^TW_1^T + W_2\Sigma_2W_2^T$. That is, it is parametrised neither by the unknown value of $x$ in the case of deterministic state nor by the distribution of $x$ in the case of stochastic state.

For presentation simplicity, the following joint vectors and matrices are introduced,

$$\hat{x}_J = \begin{bmatrix}\hat{x}_1\\ \hat{x}_2\end{bmatrix}, \quad \tilde{x}_J = \begin{bmatrix}\tilde{x}_1\\ \tilde{x}_2\end{bmatrix}, \quad J = \begin{bmatrix}I\\ I\end{bmatrix},$$
$$W_J = \begin{bmatrix}W_1 & W_2\end{bmatrix}, \quad \Sigma_J = \begin{bmatrix}\Sigma_1 & \Sigma_{1,2}\\ \Sigma_{1,2}^T & \Sigma_2\end{bmatrix}. \tag{2}$$

[1]The distinction of deterministic/stochastic state is used in the discussion of various approaches, but it is irrelevant in the problem formulation.

Thus, the fused estimate and its mean square error matrix can be expressed as $\hat{x} = W_J\hat{x}_J$ and $\Sigma = W_J\Sigma_JW_J^T$. The fusion optimal in the mean square error sense designs the weights as $W_J = (J^T\Sigma_J^{-1}J)^{-1}J^T\Sigma_J^{-1}$ and the corresponding matrix is then given by $\Sigma = (J^T\Sigma_J^{-1}J)^{-1}$.

For the construction of confidence sets, certain probability density functions have to be specified. Therefore, it is assumed that the estimates are unbiased and that the estimation errors are jointly Gaussian,

$$\tilde{x}_J \sim \mathcal{N}([0,0]^T, \Sigma_J). \tag{3}$$

The linear fusion (1) then provides an estimate with a Gaussian distributed error, $\tilde{x} \sim \mathcal{N}(0, \Sigma)$. The standard confidence sets $\mathcal{X}$ with confidence level $\ell$ are constructed as ellipsoids $\varepsilon(\hat{x}, \Sigma, q)$ with random centre $\hat{x}$, shape matrix $\Sigma$ and size parameter $q$ as

$$\varepsilon(\hat{x}, \Sigma, q) = \{s | (s - \hat{x})^T\Sigma^{-1}(s - \hat{x}) \leq q\}, \tag{4}$$

where the parameter $q$ is given by the $\ell$ quantile of the chi-square distribution with the degrees of freedom corresponding to the dimension of the state $x$. This construction is valid even for suboptimal fusion weights $W_i$ and it does not require the state $x$ to be stochastic.

### B. Problem of Random Correlation

Whereas the distribution of the joint estimation error $\tilde{x}_J$ is known in the standard case, this paper considers existence of a random nuisance parameter $\rho$. It is assumed that the density $p(\tilde{x}_1, \tilde{x}_2|\rho)$ is available and that the density $p(\rho)$ has been specified. Moreover, independence of the local errors $\tilde{x}_i$ of the nuisance parameter $\rho$ is assumed,

$$p(\tilde{x}_i|\rho) = p(\tilde{x}_i). \tag{5}$$

That is, the local estimates $\hat{x}_i$ and confidence ellipsoids $\varepsilon(\hat{x}_i, \Sigma_i, q)$ can be constructed locally, but the fusion is affected by the nuisance parameter $\rho$. The parameter describes the dependence structure. For example, if the density $p(\tilde{x}_1, \tilde{x}_2|\rho)$ is jointly Gaussian for each $\rho$ and fulfils the condition (5), the nuisance parameter can be given by the cross-correlation matrix $\Sigma_{1,2}$.

The elementary property of mean square error matrices $\Sigma_J$ of being positive semi-definite enables the matrix $\Sigma_{1,2}$ to be parametrised by a matrix $\Omega$ (which plays the role of $\rho$) as

$$\Sigma_{1,2}(\Omega) = \Sigma_1^{\frac{1}{2}}\Omega\Sigma_2^{\frac{T}{2}}, \quad \Sigma_i = \Sigma_i^{\frac{1}{2}}\Sigma_i^{\frac{T}{2}}, \quad I - \Omega\Omega^T \geq 0, \tag{6}$$

where the inequality means that the left hand side term is a positive semi-definite matrix. It is therefore possible to specify the density $p(\Omega)$ with a standardised support instead of the density $p(\Sigma_{1,2})$ with the support dependent on $\Sigma_1$ and $\Sigma_2$. For more details, see e.g. [25]–[27].

The fusion problem under random correlation is thus defined as the problem of designing the fusion weights $W_i$, finding the mean square error matrix $\Sigma$ corresponding to the fused estimate and constructing a confidence set $\mathcal{X}$ with a prescribed confidence level $\ell$. Availability of the density $p(\rho)$ and a Gaussian density $p(\tilde{x}_1, \tilde{x}_2|\rho)$ is assumed here, together with the validity of (5).

## III. Notes on Existing Approaches

In order to highlight the philosophical background, connection to and differences from other approaches are discussed next. Unknown correlation can be found in Section III-A, Bayesian estimation is commented in Section III-B.

### A. Unknown Correlation

If the matrix parameter $\Omega$ is limited only by the regularity condition (6) and no distribution of $\Omega$ is considered, the robust approach is used. The unknown joint mean square error matrix $\Sigma_J$ is replaced by an upper bound $B_J$, i.e. by a matrix that fulfils the inequality $B_J \geq \Sigma_J$ for all admissible values of $\Omega$. Then, the construction for fusion under known correlation is applied to obtain the weights, $W_J = (J^T B_J^{-1} J)^{-1} J^T B_J^{-1}$, $\hat{x}_J$, and subsequently the self-assessment matrix, $B = W_J B_J W_J^T$, which is an upper bound of the mean square error matrix of the fused estimate, $B \geq \Sigma$. In the Gaussian case (3), the confidence set can be constructed as follows,

$$\mathcal{X} = \bigcup_{\Omega} \varepsilon(\hat{x}, \Sigma(\Omega), q) = \tag{7a}$$

$$= \{\hat{x}\} \oplus \varepsilon(0, W_1 \Sigma_1 W_1^T, q) \oplus \varepsilon(0, W_2 \Sigma_2 W_2^T, q) \subseteq \tag{7b}$$

$$\subseteq \varepsilon(\hat{x}, B, q), \tag{7c}$$

where $\oplus$ denotes the Minkowski sum operation. The first equality (7a) is a simple suboptimal construction with the dependence of $\Sigma$ on $\Omega$ explicitly denoted; the set for the actual value $\Omega$ is covered by the union for all admissible values of this unknown parameter. The second equality (7b) follows from [19]. The inclusion (7c) is a standard relation of ellipsoids for $B \geq \Sigma$.

### B. Bayesian Estimation

The classical estimation fuses the estimates $\hat{x}_i$ by arbitrary weights $W_i$. The optimal weights can vary according to the selection of an optimality criterion. The mean square error matrix $\Sigma$ is an expectation, wherein the estimates are random and the state $x$ need not be random.

In the literature, the estimates $\hat{x}_i$ are frequently referred as "means". This incorrect notation likely stems from the fact that for a stochastic state $x$, the optimal estimate for a specific criterion (the mean square error one) is the conditional expectation. The objective Bayesian approach aims at finding the conditional densities $p(x|z_i)$, $p(x|z_1, z_2)$ of a stochastic state, while the estimate construction can be postponed. Even if all densities are Gaussian and the means are linearly related, there is no space for designing arbitrary weights. Also, the conditional covariance matrices of the state $x$ are functions of the measurements $z_i$, even if their values are constant. The conditional densities are frequently presented for particular realisations of the measurements $z_i$ and denoted simplistically as $p_i(x)$ and $p_{1,2}(x)$. This suggests that the state $x$ has multiple densities at the same time, which makes no sense in the objective framework.

In the subjective Bayesian case, the densities $p_i(x)$ and $p_{1,2}(x)$ are expert opinions and can exist simultaneously. However, no concepts like random variables, unbiasedness, mean square error matrices or confidence sets make sense there. The subjective combinations like the weighted arithmetic mean, $p_{1,2}(x) = \omega p_1(x) + (1-\omega) p_2(x)$, or the weighted geometric one, $p_{1,2}(x) \propto (p_1(x))^{\omega} + (p_2(x))^{1-\omega}$, can be designed. The latter do show some similarities for Gaussian densities with the fusion under unknown correlation, but the philosophies are utterly incompatible. The subjective case also tempts to mimic the joint distribution of errors by a joint density $p(x, x)$ of the same state $x$ and to define the fused density by some marginalisation or conditioning. Such an approach has poor theoretical foundations as well.

## IV. Analysis of the Problem

A part of the problem of linear fusion under random correlation has an analytic solution, the rest requires numerical computations or crude approximations. Namely, the computation of the mean square error matrices $\Sigma_J$ and $\Sigma$ can use the law of total expectation, which allows an unconditional expectation to be computed as an expectation of a conditional expectation, $\mathrm{E}\{\tilde{x}_1 \tilde{x}_2^T\} = \mathrm{E}\{\mathrm{E}\{\tilde{x}_1 \tilde{x}_2^T | \Omega\}\}$. The factorisation (6) and the linearity of expectation thus give $\Sigma_{1,2} = \Sigma_1^{\frac{1}{2}} \mathrm{E}\{\Omega\} \Sigma_2^{\frac{T}{2}}$. The weights $W_J$ and the mean square error matrix $\Sigma$ are designed according to the standard formulas shown in Section II-A.

The problem of confidence set design has no analytic solution for a general density $p(\Omega)$. Although the densities $p(\tilde{x}_1, \tilde{x}_2|\Omega)$ are assumed to be Gaussian for each $\Omega$, the density $p(\tilde{x}_1, \tilde{x}_2)$ is no more Gaussian. An analogy holds for the density $p(\tilde{x}|\Omega)$, which is Gaussian, and the density $p(\tilde{x})$, which is not. It is possible to consider only the support of $p(\Omega)$, i.e. those $\Omega$ with nonzero values of the density, and use the construction (7) designed for unknown correlation. The numerical solution is based on approximation of the density $p(\tilde{x})$, which is computationally expensive. These two approaches will be further commented.

The approach based on the support of $p(\Omega)$ yields large confidence sets. Moreover, a question arises whether the mean square error matrix or the size of the confidence sets are minimised. If the weights $W_J$ are designed by the standard formulas with the expected correlation matrix $\Sigma_{1,2}$, the confidence set (7) will be larger than in the case of the weights $W_J$ designed according to the confidence set size. Vice versa, the latter case produces mean square error matrix $\Sigma$ with a size larger than the former case.

The numerical approach computes integrals in order to obtain probabilities $P(x \in \mathcal{X})$, e.g. over a grid of points. Since typical confidence levels $\ell$ are close to one, high numerical precision is required. There is the curse of dimensionality; the approach becomes costly very fast with increasing dimensions of the state $x$. A crude solution to the confidence set design is to pretend that the density $p(\tilde{x})$ is Gaussian with the covariance matrix given by the expectation over the covariance matrices related to the Gaussian densities $p(\tilde{x}|\Omega)$. Unfortunately, the confidence level $\ell$ need not be achieved by designing the corresponding confidence ellipsoid.

The following section illustrates the designs.

## V. ILLUSTRATION

Let the state $x$ be deterministic and scalar. Let the measurements $z_1$, $z_2$ have an additive Gaussian error with zero mean and variances $\sigma^2$ and 1 respectively, i.e. $z_1 \sim \mathcal{N}(x, \sigma^2)$ and $z_2 \sim \mathcal{N}(x, 1)$. Let the joint density of the measurement errors be Gaussian and denote the correlation coefficient of the errors by $\rho$. Let the estimates $\hat{x}_i$ of the state $x$ be given directly by the measurements, $\hat{x}_i = z_i$. To fulfil the assumptions posed in Section II-B, it is further needed the specify the density of the nuisance parameter $\rho$. Let it be uniform. Thus, it holds

$$\Sigma_J(\rho) = \begin{bmatrix} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{bmatrix}, \qquad p(\rho) = \begin{cases} \frac{1}{2} & -1 \le \rho \le 1 \\ 0 & otherwise \end{cases}. \quad (8)$$

The local estimates $\hat{x}_i$ are unbiased. Since the expectation of $\rho$ is zero, $\mathrm{E}\{\rho\} = 0$, the mean square error optimal fusion weights are given by

$$W_1 = \frac{1}{\sigma^2 + 1}, \quad W_2 = \frac{\sigma^2}{\sigma^2 + 1} \quad (9)$$

and the fused estimate is unbiased. For these weights, the conditional variances of the Gaussian densities $p(\tilde{x}|\rho)$ and the unconditional variance of the estimation error $\tilde{x}$ are given respectively by

$$\Sigma(\rho) = \frac{\sigma^2}{\sigma^2 + 1} + \frac{2\sigma^3}{(\sigma^2 + 1)^2}\rho, \quad \Sigma = \frac{\sigma^2}{\sigma^2 + 1}. \quad (10)$$

Fig. 1 illustrates the densities of local and fusion errors. The variance $\sigma^2$ is chosen as $\sigma^2 = 4$. It the top left figure, ellipses $\varepsilon(0, \Sigma_J, q)$ are shown for the Gaussian densities $p(\tilde{x}_1, \tilde{x}_2|\rho)$ for $\rho = -0.99, -0.5, 0, 0.5, 0.99$. The quantile $q$ corresponds to the 0.95 confidence level $\ell$. Since the variance of error of the optimal estimate for known $\rho$ is given by $\Sigma = ([1,1]\Sigma_J^{-1}[1,1]^T)^{-1}$ and the ellipses are given by (4), the results of the $\rho$-optimal fusion can be read on the line passing through the origin and $[1,1]^T$ (the dashed line). If the parameter $\rho$ is unknown, the fusion cannot be better than that corresponding to the ellipse with the largest radius in the direction $[1,1]^T$, i.e. that for $\rho = 0.5$ in this example. Nevertheless, the fusion weights are selected only once, which means that the actual fusion results can be read as certain marginal variances. These variances can be found by projecting the ellipses onto the dashed line, where the direction is indicated by the dotted lines and is orthogonal to $W_J$. The direction corresponds to the direction of the tangent line to the ellipse for the zero-valued expectation of $\rho$ (8), where the tangent is taken in the point of intersection of the ellipse and the dashed line.

The bottom left figure of Fig. 1 shows the conditional densities of the fused error $p(\tilde{x}|\rho)$ and the $\rho$-optimal 0.95 confidence intervals $\varepsilon(0, \Sigma, q)$ (delimited by the dots). The intervals may also be read from the dashed line in the top left figure. The right column of Fig. 1 shows the unconditional densities $p(\tilde{x}_1, \tilde{x}_2)$ and $p(\tilde{x})$ and the moment-matched Gaussian densities. It can be seen from the bottom right figure that the interval (delimited by the dots) based on the Gaussian approximation is slightly smaller than that (delimited by the circles) based on the exact density $p(\tilde{x})$. The probabilities that the error $\tilde{x}$ lies within these intervals are 0.9428 and 0.95, respectively. That is, the approximation need not provide a sufficiently high confidence level.

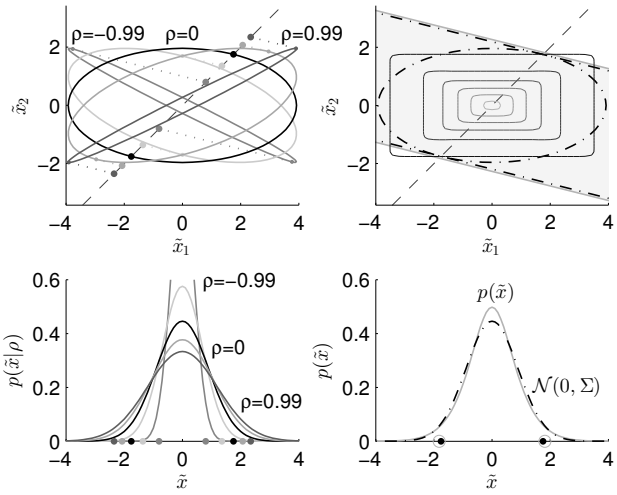


Fig. 1. Conditional and unconditional densities of local and fusion errors – $p(\tilde{x}_1, \tilde{x}_2|\rho)$ (top left – five examples), $p(\tilde{x}|\rho)$ (bottom left – five examples), $p(\tilde{x}_1, \tilde{x}_2)$ (top right – six density levels and approximation) and $p(\tilde{x})$ (bottom right – density [solid line] and its Gaussian approximation [dash-dotted line]).

The top right figure of Fig. 1 shows six levels of the density $p(\tilde{x}_1, \tilde{x}_2)$, namely $0.01, 0.03, \ldots, 0.11$, and the ellipse $\varepsilon(0, \Sigma_J, q)$. The shaded area corresponds to the local errors $\tilde{x}_1$, $\tilde{x}_2$ that lead to the fusion errors $\tilde{x}$ within the indicated 0.95 confidence interval. The values of the density $p(\tilde{x})$ can be obtained by integrating the joint density over the lines parallel with the axis of the shaded area, i.e. the direction shown by the dotted lines in the top left figure. They can also be obtained from the Gaussian densities $p(\tilde{x}|\rho)$ and the density $p(\rho)$.

The design of confidence sets by pretending unknown correlation is discussed next. If the weights $W_J$ are optimal according to the mean square error criterion, i.e. given by (9), the confidence sets $\mathcal{X}$ designed by (7b) are given by $\{\hat{x}\} \oplus [-\frac{\sigma}{\sigma^2+1}\sqrt{q}, \frac{\sigma}{\sigma^2+1}\sqrt{q}] \oplus [-\frac{\sigma^2}{\sigma^2+1}\sqrt{q}, \frac{\sigma^2}{\sigma^2+1}\sqrt{q}]$, that is $\mathcal{X} = [\hat{x} - \frac{\sigma^2+\sigma}{\sigma^2+1}\sqrt{q}, \hat{x} + \frac{\sigma^2+\sigma}{\sigma^2+1}\sqrt{q}]$. These intervals with random centres $\hat{x}$ have constant lengths $2\frac{\sigma^2+\sigma}{\sigma^2+1}\sqrt{q}$. On the other hand, the choice $\sigma^2 = 4$ made in this example implies that the estimate $\hat{x}_2$ has smaller error variance than the first estimate $\hat{x}_1$. The weights designed as $W_1 = 0$ and $W_2 = 1$ lead to $\hat{x} = \hat{x}_2$. The corresponding mean square error $\Sigma = 1$ is higher than in the preceding case (10), but the confidence intervals are given by $\mathcal{X} = [\hat{x}_2 - \sqrt{q}, \hat{x}_2 + \sqrt{q}]$, i.e. they are shorter.

The upper bound $B$ used in (7c) can be given by a standard construction as $\omega^{-1}W_1\Sigma_1W_1^T + (1-\omega)^{-1}W_2\Sigma_2W_2^T$ with the choice $\omega = \sqrt{W_1\Sigma_1W_1^T}(\sqrt{W_1\Sigma_1W_1^T} + \sqrt{W_2\Sigma_2W_2^T})^{-1}$, i.e. by $B = (\frac{\sigma^2+\sigma}{\sigma^2+1})^2$. This bound $B$ is equal to the conditional covariance matrix $\Sigma(\rho)$ for maximal positive correlation; see(10) for $\rho = 1$. This also illustrates that in the one-dimensional case, equality of the sets in (7b) and (7c) can be achieved.

The analytical results can be supported by Monte Carlo simulations; a pseudocode is provided in Algorithm 1. Parameters are set first, functions are called next. The actual Matlab code for a specific setting can be found next to the % character. The purpose of the pseudocode is to highlight the philosophy of the linear fusion.

**Algorithm 1** Pseudocode of Monte Carlo simulation

```
Set x or define p(x).      % EX=5; CX=0;
Define z1, z2.  % H1=1; H2=1; E1=0; E2=0; C1=4; C2=1;
Define p(ρ), e.g. as uniform on the interval [−1, 1].
Choose the number of Monte Carlo runs.      % N=10000;

X=MCstate();      % X=mvnrnd(EX,CX,N)';
rho=MCcorrelation();      % rho=2*rand(1,N)-1;
[Z1,Z2]=MCmeasurements(X,rho);
    % for k=1:N,
    %    C12=sqrt(C1)*rho(k)*sqrt(C2);
    %    V=mvnrnd([E1 E2],[C1 C12;C12 C2])';
    %    Z1(k)=H1*X(k)+V(1); Z2(k)=H2*X(k)+V(2);
    % end
[hatX1,S1]=estimator1(Z1);      % hatX1=Z1; S1=4;
[hatX2,S2]=estimator2(Z2);      % hatX2=Z2; S2=1;
[hatX,S]=fusion(hatX1,hatX2);
    % hatX=1/5*hatX1+4/5*hatX2; S=4/5;
[trueS,Gauss_level]=evaluation(X,hatX);
    % histogram(X-hatX,'Normalization','pdf')
    % trueS=cov(X-hatX);
    % Gauss_level=mean((X-hatX).^2/S<=chi2inv(0.95,1));
```

It has to be stressed that the random parameter $\rho$ must not be an input parameter of the estimators and the fusion. These functions can be realised in hardware devices, where the inputs $z_i$ and outputs $\hat{x}_i$ are physical signals and the self-assessment matrices correspond to the data in data-sheets. The data-sheets cannot depend on any signal and their data need not match the reality. The true performance is not evaluated over a single run and can be verified by repeating the experiments. The pseudocode generates estimates of the density $p(\tilde{x})$, covariance of $\tilde{x}$ or confidence levels resulting from the Gaussian approximation as representatives of the evaluation. The code can be easily modified for providing conditional quantities by setting a constant "rho".

## VI. Summary

Fusion under random correlation has been formulated and in order to stress the extent of the philosophical framework, various incompatible approaches have been commented. The solution to the chosen approach has been presented. Whereas the minimisation of the mean square error matrix is straightforward, the design of confidence sets is intricate. The former is based on computing the expectation of the cross-correlation matrix of the estimation errors. The latter is either computationally expensive, relies on crude approximations or is designed for unknown correlation. The theoretical background has been illustrated by a simple analytical example. Guidelines for verifying the results via a simulation have been given last.

## References


[1] D. Gillies, *Philosophical Theories of Probability*. Routledge, 2000.

[2] D. P. Rowbottom, *Probability*. Polity, 2015.

[3] A. Khrennikov, *Interpretations of Probability*, 2nd ed. de Gruyter, 2009.

[4] S. Amari and H. Nagaoka, *Methods of Information Geometry*, ser. Translation of Mathematical Monographs. AMS, 2000, vol. 191.

[5] F. Nielsen, "An elementary introduction to information geometry," *Entropy*, vol. 22, no. 10, September 2020.

[6] J. Ajgl and M. Šimandl, "Conservative merging of hypotheses given by probability densities," in *Proceedings of the 15th International Conference on Information Fusion*, Singapore, July 2012.

[7] ——, "Design of a robust fusion of probability densities," in *2015 American Control Conference*, Chicago, IL, USA, July 2015.

[8] G. Koliander, Y. El-Laham, P. M. Djurić, and F. Hlawatsch, "Fusion of probability density functions," *Proceedings of the IEEE*, vol. 110, no. 4, pp. 404–453, April 2022.

[9] R. Kulhavý and F. J. Kraus, "On duality of regularized exponential and linear forgetting," *Automatica*, vol. 32, no. 10, pp. 1403–1415, 1996.

[10] C. Genest and J. V. Zidek, "Combining probability distributions: A critique and an annotated bibliography," *Statistical Science*, vol. 1, no. 1, pp. 114–148, 1986.

[11] J. Ajgl and M. Šimandl, "On conservativeness of posterior density fusion," in *Proceedings of the 16th International Conference on Information Fusion*, Istanbul, Turkey, July 2013.

[12] W. Q. Meeker, G. J. Hahn, and L. A. Escobar, *Statistical Intervals: A Guide for Practitioners and Researchers*, 2nd ed. Wiley, 2017.

[13] J. Ajgl and O. Straka, "Design of confidence sets for estimation with approximate piecewise linear constraint," in *European Control Conference (ECC 2020)*, Saint Petersburg, Russia, May 2020.

[14] ——, "Comparison of confidence sets designs for various degrees of knowledge," in *Proceedings of the 24th International Conference on Information Fusion*, Sun City, South Africa, November 2021.

[15] U. D. Hanebeck and K. Briechle, "New results for stochastic prediction and filtering with unknown correlations," in *Proceedings of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Baden-Baden, Germany, August 2001, pp. 85–90.

[16] P. O. Arambel, C. Rago, and R. K. Mehra, "Covariance intersection algorithm for distributed spacecraft state estimation," in *Proceedings of the American Control Conference*, Arlington, Virginia, USA, June 2001, pp. 4398–4403.

[17] L. Chen, P. O. Arambel, and R. K. Mehra, "Estimation under unknown correlation: Covariance intersection revisited," *IEEE Transactions On Automatic Control*, vol. 47, no. 11, pp. 1879–1882, 2002.

[18] M. Reinhardt, B. Noack, P. O. Arambel, and U. D. Hanebeck, "Minimum covariance bounds for the fusion under unknown correlations," *IEEE Signal Processing Letters*, vol. 22, no. 9, pp. 1210–1214, 2015.

[19] J. Ajgl and O. Straka, "A geometrical perspective on fusion under unknown correlations based on Minkowski sums," in *20th International Conference on Information Fusion*, Xi'an, China, 2017, pp. 725–732.

[20] S. J. Julier and J. K. Uhlmann, *Handbook of multisensor data fusion*. CRC Press, 2001, ch. General Decentralized Data Fusion with Covariance Intersection.

[21] Z. Wu, Q. Cai, and M. Fu, "Covariance intersection for partially correlated random vectors," *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 619–629, March 2018.

[22] J. Ajgl and O. Straka, "Covariance intersection fusion with element-wise partial knowledge of correlation," *Automatica*, vol. 139, p. 110168, May 2022.

[23] M. Reinhardt, B. Noack, M. Baum, and U. D. Hanebeck, "Analysis of set-theoretic and stochastic models for fusion under unknown correlations," in *Proceedings of the 14th International Conference on Information Fusion*, Chicago, Illinois, USA, July 2011.

[24] H. Zhu, Q. Zhai, M. Yu, and C. Han, "Estimation fusion algorithms in the presence of partially known cross-correlation of local estimation errors," *Information Fusion*, vol. 18, pp. 187–196, 2014.

[25] P. J. Forrester and J. Zhang, "Parametrising correlation matrices," *Journal of Multivariate Analysis*, vol. 178, p. 104619, July 2020.

[26] M. Laurent and S. Poljak, "On a positive semidefinite relaxation of the cut polytope," *Linear Algebra and its Applications*, vol. 223/224, pp. 439–461, July 1995.

[27] H. Joe, "Generating random correlation matrices based on partial correlations," *Journal of Multivariate Analysis*, vol. 97, no. 10, pp. 2177–2189, 2006.