amcs

# FUSION OF MULTIPLE ESTIMATES BY COVARIANCE INTERSECTION: WHY AND HOW IT IS SUBOPTIMAL

JIŘÍ AJGL [a,*], ONDŘEJ STRAKA [a]

[a]European Centre of Excellence—New Technologies for Information Society & Department of Cybernetics
Faculty of Applied Sciences, University of West Bohemia
Univerzitní 8, 306 14, Pilsen, Czech Republic
e-mail: {jiriajgl,straka30}@kky.zcu.cz

The fusion under unknown correlations tunes a combination of local estimates in such a way that upper bounds of the admissible mean square error matrices are optimised. Based on the recently discovered relation between the admissible matrices and Minkowski sums of ellipsoids, the optimality of existing algorithms is analysed. Simple examples are used to indicate the reasons for the suboptimality of the covariance intersection fusion of multiple estimates. Further, an extension of the existing family of upper bounds is proposed, which makes it possible to get closer to the optimum, and a general case is discussed. All results are obtained analytically and illustrated graphically.

**Keywords:** decentralised estimation, fusion under unknown correlations, covariance intersection, Minkowski sum.

## 1. Introduction

The optimal processing of data is a subject of the classical estimation theory (Lehmann and Casella, 1998). Supposing that a probabilistic description of the observed data is available, unknown parameters of the description are to be inferred. A standard application is the estimation of the state of a dynamical system (Bar-Shalom *et al.*, 2001; Simon, 2006), where an assessment of the (in)accuracy of the processing is needed for recursive calculations.

A direct combination of data originating from multiple sensors may be impossible due to various reasons such as bandwidth limitation. The data, a probabilistic description of which is available only locally, can be transformed into local estimates of the unknown parameters or their functions. If the functions are globally known, the estimates can be communicated to other sensors and combined with the other local estimates.

Besides the description of the local estimates, the optimal linear fusion (Li *et al.*, 2003) also needs the knowledge of the cross-correlations of the local estimation errors. However, a local evaluation of cross-correlations requires additional information about the other estimates, which may be unavailable. The decentralised estimation fusion problems therefore make a shift in the sense of

optimality. They focus on worst cases and typically, tune the combination of estimates according to upper bounds of mean square error matrices.

Such a linear fusion rule has been proposed by Julier and Uhlmann (1997) under the name *covariance intersection* (CI). Various generalisations of the fusion have been presented later by, e.g., Julier and Uhlmann (2001), Arambel *et al.* (2001), Uhlmann (2003) or Wu *et al.* (2018). Unfortunately, the rule is often presented in a compressed form that suggests that an upper bound is computed first and the combination is defined by the upper bound next. This mental shortcut conceals that the upper bound is not unique. A discussion of general bounds can be found in the works of Chen *et al.* (2002) or Ajgl and Šimandl (2014), for example. The overall optimality of the CI fusion of two estimates has been justified by Reinhardt *et al.* (2015).

The virtue of CI is that the mean square error matrices of the local estimates need not be known exactly, but their upper bounds can be used for the fusion. Although the optimality of the fused bound is lost in such a case, the fusion rule still provides an upper bound for the fused estimate. Therefore, a sequential fusion of multiple estimates is possible (Deng *et al.*, 2012) and various fusion strategies can be adopted (Hu *et al.*, 2012; Kowalczuk and Domżalski, 2013; Ajgl and Straka, 2016a;

---

[*]Corresponding author

2016b). A weak point of the sequential fusion is that in case pairwise optimisations are used, the batch-optimal parameters of the fusion need not be found. Ajgl *et al.* (2014) show that this quantitative weakness is more profound.

A qualitative deficiency of the fusion of multiple estimates has been observed by Ajgl and Straka (2017). The matrices provided by the CI fusion can be strictly larger than it is needed to guarantee that they are valid upper bounds. Nevertheless, the reasons have not been explored in depth.

Thus, the goal of this paper is to provide a deeper insight into the problem of fusion under unknown correlations. To achieve it, mean square error matrices and their upper bounds are visualised by geometrical equivalents.

The paper is organised as follows. Section 2 formulates the estimation fusion problem and rises optimality questions. Section 3 discusses the optimisation process and demonstrates the suboptimality of the CI fusion on the simplest examples. Section 4 inspects the reason of the suboptimality and points to a refinement. Section 5 summarises the contribution.

## 2. Problem formulation

### 2.1. Fusion under unknown correlations.
Let $X$ be an $n$-dimensional column vector to be estimated and let $N$ local estimates $Z_i$ be available, $i = 1, \ldots, N$. According to Li *et al.* (2003, Section III.A), the local estimates can be viewed as measurements of $X$. A general case is given by considering $n_i$-dimensional $Z_i$. If the estimation error $E_i$ of $Z_i$ is defined as $E_i = H_i X - Z_i$, where $H_i$ is a measurement matrix, a generic measurement equation $Z_i = H_i X + (-E_i)$ can be used to describe $Z_i$. Typical examples of $H_i$ are the identity matrix $I$, which appears in classical estimation fusion problems, or a selection of its rows, which appears when only some components of $X$ are estimated locally. The global measurement matrix $H$, $H = [H_1^\top, \ldots, H_N^\top]^\top$, is assumed to be of full column rank in the sequel.

The fused estimate $\hat{X}$ is considered to be given by a linear fusion of $Z_i$,

$$\hat{X} = \sum_{i=1}^{N} W_i Z_i, \quad \sum_{i=1}^{N} W_i H_i = I, \tag{1}$$

which can be expressed in a compact form as $\hat{X} = WZ$, $WH = I$, where the global weight $W$ is introduced as $W = [W_1, \ldots, W_N]$ and the vector of stacked local estimates is denoted as $Z$, $Z = [Z_1^\top, \ldots, Z_N^\top]^\top$. The equality condition $WH = I$ preserves unbiasedness of the local estimates[1] and is needed for evaluation of the quality

of the fused estimate $\hat{X}$.

The mean square error matrices $R_i$ of the local estimates, $R_i = \mathrm{E}\{E_i E_i^\top\}$, are assumed to be known and invertible. The cross-correlations $R_{i,j}$ of the local estimation errors are defined as $R_{i,j} = \mathrm{E}\{E_i E_j^\top\}$ for $i = 1, \ldots, N$, $j = 1, \ldots, N$, $i \neq j$. Let $R$ denote a matrix with diagonal blocks $R_i$ and off-diagonal blocks $R_{i,j}$. The mean square error matrix $P$ of the fused estimate $\hat{X}$ is given by $P = \mathrm{E}\{(X - \hat{X})(X - \hat{X})^\top\}$, where the dependence of $P$ on $W$ is not denoted explicitly, and $P = WRW^\top$. The mean square error matrix $P$ is optimised in classical fusion problems, where some strictly monotonically increasing cost functions $J(P)$, e.g., the trace or the determinant, are used. The optimal fusion weight $W^*(R)$ is given by

$$W^*(R) = \underset{W:WH=I}{\arg\min} J(P) \tag{2}$$

in such a case. However, the cross-correlations $R_{i,j}$ have to be known for the optimisation.

If the cross-correlations $R_{i,j}$ are assumed to be unknown, the minimisation with respect to the worst case is considered. Now, the optimal weight $W^*$ is given by

$$W^* = \underset{W:\ WH=I}{\arg\min} \left( \underset{B:\ B \geq WAW^\top}{\min} J(B) \right), \tag{3}$$

with $A$ being a family of admissible matrices $R$, i.e., of positive semi-definite matrices with the known blocks $R_i$ on the diagonal and arbitrary off-diagonal blocks $R_{i,j}$. The inequality $B \geq WAW^\top$ means that for a selected weight $W$, the difference $B - WAW^\top$ is a positive semi-definite matrix for all matrices in the family $A$. That is, $B$ is an upper bound of all admissible $P$. The optimal weight $W^*$ optimises the best upper bound.[2]

Besides providing the bound-optimal estimate $\hat{X}^*$, which is given by (1) for the optimal choice $W = W^*$, fusion rules evaluate the quality of the provided estimate $\hat{X}$. If the cross-correlations $R_{i,j}$ are assumed to be unknown, the mean square error matrix $P$ cannot be computed. Hence, an upper bound of $P$ is provided instead. The best upper bound corresponding to the selected fusion weight $W$ is a natural choice, see the inner minimum term in the definition (3) of $W^*$, and will be denoted as $B^*(W)$,

$$B^*(W) = \underset{B:\ B \geq WAW^\top}{\arg\min} J(B). \tag{4}$$

The optimal fusion rule computes the optimal weight $W^*$ exactly and provides the best achievable upper bound, which will be denoted shortly as $B^*$, $B^* = B^*(W^*)$.

---

[1]If the local estimates are unbiased, $\mathrm{E}\{Z_i - H_i X\} = 0$ holds for each $i$. A weighted sum of the equalities gives $\mathrm{E}\{WZ - WHX\} = 0$, i.e. $\mathrm{E}\{\hat{X} - WHX\} = 0$. Then $\hat{X}$ is unbiased if $WH = I$.

[2]If an upper bound $B$ of all admissible $P$ is not required for a further processing, the worst case can be defined in a weaker sense. Gao *et al.* (2016) use the constraint $J(B) \geq J(WAW^\top)$ for tuning the weights instead of the constraint $B \geq WAW^\top$ in (3).

## 2.2. State of the art.

**Lemma 1.** (Chen *et al.*, 2002) *A family of upper bounds of the fused mean square error matrix $P$ for a given weight $W$, i.e., a subset of $\{B : B \geq WAW^\top\}$, is given by*

$$B(W, \omega) = \sum_{i=1}^{N} \frac{1}{\omega_i} W_i R_i W_i^\top, \qquad (5)$$

*where $\omega_i$ are free parameters given by a standard simplex, $0 < \omega_i < 1$, $\sum_{i=1}^{N} \omega_i = 1$, and $\omega = [\omega_1, \ldots, \omega_N]$.*

*Proof.* Uhlmann (2003, Appendix A) provides an alternative derivation. An upper bound $Q$ of $A$ (i.e., of all matrices in the family $A$) is constructed first, $Q \geq A$. Here, $Q$ is a block-diagonal matrix with the blocks $\frac{1}{\omega_i} R_i$. Since $B$ is designed as $B = WQW^\top$, it is an upper bound of $WAW^\top$. Further comments can be found in the work of Ajgl and Straka (2017, Theorem 6). ∎

**Remark 1.** If the local mean square error matrices $R_i$ are replaced by their upper bounds, a family of upper bounds $B(W, \omega)$ of $P$ is still obtained.

**Lemma 2.** (CI fusion) *Consider a family of weights $W(w)$ which satisfies*

$$W_i(w) = (\sum_{j=1}^{N} w_j H_j^\top R_j^{-1} H_j)^{-1} w_i H_i^\top R_i^{-1}, \quad (6)$$

*where $w_i$ are free parameters given by a probability simplex, $0 < w_i < 1$, $\sum_{i=1}^{N} w_i = 1$, and $w = [w_1, \ldots, w_N]$. For a selected $w$, an upper bound of $P$ is given by*

$$B(w) = (\sum_{j=1}^{N} w_j H_j^\top R_j^{-1} H_j)^{-1}. \qquad (7)$$

*Proof.* Choose $\omega = w$ in the family (5). Denoting the upper bound $Q$ for the choice $\omega = w$ as $Q(w)$, we have $W(w) = (H^\top (Q(w))^{-1} H)^{-1} H^\top (Q(w))^{-1}$ and $B(w) = W(w)Q(w)W(w)^\top = (H^\top (Q(w))^{-1} H)^{-1}$. In fact, $B(w)$ is a short notation of $B(W(w), w)$. Various derivations of the lemma can be found in the works of Julier and Uhlmann (1997), Chen *et al.* (2002), Ajgl and Šimandl (2014) or Reinhardt *et al.* (2015), among others. ∎

The optimal CI weight $W^{\text{CI}}$, $W^{\text{CI}} = W(w^*)$ uses $w^*$ given by

$$w^* = \underset{w:\, w1_N = 1,\, 0 \leq w_i}{\arg\min} \; J(B(w)), \qquad (8)$$

where $1_N$ is a column vector of ones and length $N$, $1_N = [1, \ldots, 1]^\top$. The corresponding upper bound $B^{\text{CI}}$, $B^{\text{CI}} = B(w^*)$, is the best matrix from the family $B(w)$ of particular bounds corresponding to different weights $W(w)$.
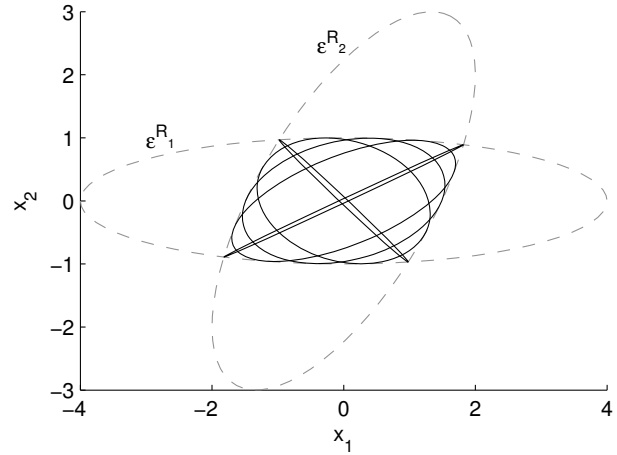


Fig. 1. Illustration of a necessary condition on $B$: Local ellipsoids $\varepsilon^{R_i}$ (dashed ellipses) and selected admissible optimal ellipsoids for different weights $W^*(A)$ (solid ellipses).

## 2.3. Visualisation of the matrices.

A symmetric positive semi-definite matrix $R$ defines a $1$-$\sigma$ ellipsoid $\varepsilon^R$ (Kurzhanski and Vályi, 1997, Section 2.1) as

$$\varepsilon^R = \{x | \, \forall v : \, v^\top x \leq (v^\top Rv)^{\frac{1}{2}}\}, \qquad (9)$$

where $x$ and $v$ have compatible dimensions. For positive definite matrices $R$, an equivalent definition can be stated as $\varepsilon^R = \{x | x^\top R^{-1} x \leq 1\}$.

CI was proposed by Julier and Uhlmann (1997) as a rule based on a geometrical concept, that can be paraphrased as "the intersection of $\varepsilon^{R_1}$ and $\varepsilon^{R_2}$ has to be a subset of $\varepsilon^B$". In fact, the argument is only a necessary condition for $B$ being an upper bound of $WAW^\top$ for a selected weight $W$. The reason is that the intersection describes the set of all admissible optimal (under known correlations) fused mean square error matrices, each of which is based on a different weight $W^*(A)$. Figure 1 illustrates that the union of the admissible optimal ellipsoids is equal to the intersection of the ellipsoids $\varepsilon^{R_1}$, $\varepsilon^{R_2}$. The parametrisation of the used matrices can be found in Appendix A.

The rule involving the family of weights (6) and the selected matrices (7) was proved to fulfil the upper bounding property of $B(w)$ already by Julier and Uhlmann (1997). Later on, it has been shown that the optimal CI is family-optimal for $J(B)$ being trace (Chen *et al.*, 2002), family-optimal for general $J(B)$ (Ajgl and Šimandl, 2014) and, for the fusion of two estimates (Reinhardt *et al.*, 2015), even optimal in the general sense of (3), (4), i.e., $W^* = W^{\text{CI}}$, $B^* = B^{\text{CI}}$. Figure 2 illustrates the upper bounding ellipsoids given by (5) for a fixed weight, namely $W = W([1/2, 1/2])$, and the ellipsoids corresponding to the admissible mean square error matrices $WAW^\top$. The necessary condition
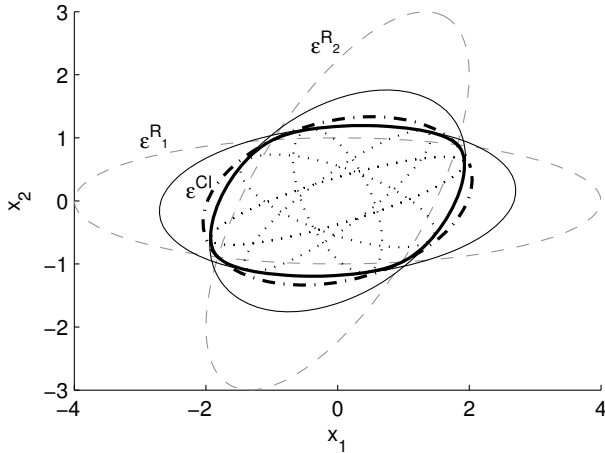
Fig. 2. Examples of $B$. Local ellipsoids $\varepsilon^{R_i}$ (dashed ellipses), four selected admissible ellipsoids (dotted ellipses) for a fixed weight $W$, bounding loop $\sigma$ (bold solid) and three important upper bounding ellipsoids $\varepsilon^B$ (solid ellipses, dash-dotted ellipse $\varepsilon^{\mathrm{CI}}$ for $B = B([1/2, 1/2])$).

from Fig. 1 is fulfilled, i.e., the ellipsoids $\varepsilon^B$ cover the intersection of $\varepsilon^{R_i}$, nevertheless, it is evident that the intersection itself is not sufficient to describe the union of the admissible ellipsoids. The parametrisation of the used matrices can be found in Appendix A.

Besides showing the optimality in the sense of (3), (4) of the fusion of two estimates by the rule (6), (7), Reinhardt *et al.* (2015) question the optimality of the fusion of multiple estimates. Ajgl and Straka (2017, Section IV.B) show that the CI rule is indeed not optimal for the fusion of multiple estimates, at least in the sense of the bound (4); it may happen that no parameter $\omega$ satisfies the equality $B^*(W(w)) = B(W(w), \omega)$. However, the weight optimality (3) has not been treated and no guideline on improving the fusion rule has been provided.

**2.4. Problem delimitation.** A necessary condition for the weight optimality (3) is that the bound $B$ inspected in (3) and (4) is minimal in some sense. A weak and a strong senses are defined as follows.

**Definition 1.** (*Tangent bound*) A bound $B$ is tangent if no matrix $\beta B$, $0 < \beta < 1$, is a bound.

**Definition 2.** (*Tight bound*) A tangent bound $B$ is tight if no matrix $S$, $S \neq B$, $B \geq S$, is a bound.

**Problem 1.** (*CI optimality*) Does $W^* = W^{\mathrm{CI}}$ hold also for $N > 2$?

**Problem 2.** (*Optimal fusion of multiple estimates*) Are there simple families of *tight* upper bounds $B$ of the fused mean square error matrix $P$ for $N > 2$?

## 3. Discussion and examples

The linear fusion of estimates has been formulated in Section 2.1 in the logical order. That is, the fusion weight $W$ is optimised first and an upper bound $B$ of the mean square error of the fused estimate is constructed after the fused estimate has been defined.

The fusion problem can be approached jointly (see, e.g., Reinhardt *et al.*, 2015),

$$(W^*, B^*) = \underset{(W, B):\, WH=I,\, B \geq WAW^\top}{\arg\min} J(B). \quad (10)$$

Instead of dealing with bounds $B$ of fused matrices, the optimal solution $(W^*, B^*)$ can be found via proposing upper bounds $Q$ of the family $A$ of all admissible global mean square error matrices $R$,

$$(W^*, Q^*) = \underset{(W, Q):\, WH=I,\, Q \geq A}{\arg\min} J(WQW^\top), \quad (11)$$

$$B^* = W^*Q^*(W^*)^\top \quad (12)$$

(see, e.g., Hanebeck *et al.*, 2001; Uhlmann, 2003). The solution can be found via minimising a minimum, $Q^* = \arg\min_{Q:\, Q \geq A} \min_{W:\, WH=I} J(WQW^\top)$. That is, the optimal fusion weight for each given bound $W^*(Q)$ is found in the inner minimum and the bounds $Q$ are optimised in the outer minimum. Since $W^*(Q)$ is given by a standard result as

$$W^*(Q) = (H^\top Q^{-1} H)^{-1} H^\top Q^{-1}, \quad (13)$$

the optimal bound $Q^*$ is given by

$$Q^* = \underset{Q:\, Q \geq A}{\arg\min}\, J((H^\top Q^{-1} H)^{-1}). \quad (14)$$

That is, the search for the optimal fusion weight $W^*$ can be replaced by the search for the optimal global bound $Q^*$. The optimal weight $W^*$ is then given by $W^* = W^*(Q^*)$ and the optimal fused bound $B^*$ by $B^* = (H^\top (Q^*)^{-1} H)^{-1}$.

The CI fusion uses a specific block-diagonal family of upper bounds $Q$ instead of considering all upper bounds of $A$. Although the bounds from the family are tight, this does not imply that the matrices $WQW^\top$ are tight bounds of $WAW^\top$; see Fig. 5 and the comments in Section 4.1. Also, the family of $Q$ does not contain all tight upper bounds of $A$. Based on the explicit formulation of the fusion under unknown correlations, it is thus easy to see that

$$J(B^{\mathrm{CI}}) \geq J(B^*(W^{\mathrm{CI}})) \geq J(B^*). \quad (15)$$

Reinhardt *et al.* (2015) show that the family contains the optimal bound $Q^*$ in the fusion of two estimates, i.e., equalities replace both inequalities in (15) for $N = 2$. On the other hand, the family may fail to contain the optimal

bound $Q^*$ in the fusion of multiple estimates, $N > 2$. This paper confirms the possibility of the failure and evidences that both inequalities in (15) can be strict.

The possible suboptimality of the weight $W^{CI}$ (and bound $B^{CI}$) has not been evidenced before. The routinely referred suboptimality of CI (see, e.g., Bar-Shalom *et al.*, 2011, Section 9.3.7), points to the comparison with the mean square error of the fused estimate, $B^{CI} \geq W^{CI}R(W^{CI})^\top$, or even to the comparison with the optimal value for known correlations, $B^{CI} \geq W^*(R)R(W^*(R))^\top$. This paper does not point to the price of unavailability of correlations given by $J(B^{CI}) - J(W^*(R)R(W^*(R))^\top)$, but it does point to the price $J(B^*(W^{CI})) - J(B^*)$ of the naivety of fusion weight that follows from using a family of bounds that is not rich enough to contain the optimal bound.

Before showing that both inequalities in (15) can be strict, a lemma enabling a visual comparison will be given.

**Lemma 3.** (Ajgl and Straka, 2017) *The union of all admissible fused ellipsoids $\varepsilon^{WRW^\top}$ is equal to the Minkowski sum $\sigma$ of the ellipsoids $\varepsilon^{W_i R_i W_i^\top}$.*

**3.1. Fusion of three partial estimates.** The following example is purely academic. Its purpose is to show that there exists a case such that the first inequality in (15) can be strict, while the equality holds in the second one. The example also offers a simple visualisation and the possibility to verify the results by hand calculation.

Let the local mean square error matrices be equal and unit, $R_i = 1$, and the local matrices $H_i$ be given by

$$H_1 = [0, 1], \quad H_2 = H_1 T^{-1}, \quad H_3 = H_1 T, \quad (16)$$

where $T$ is a rotation matrix for the angle $\pi$ divided by the number of estimates,

$$T = \begin{bmatrix} \cos\frac{\pi}{3} & -\sin\frac{\pi}{3} \\ \sin\frac{\pi}{3} & \cos\frac{\pi}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}. \quad (17)$$

Using the rotational symmetry of the setting (16) (see $\varepsilon^{R_i}$ in Fig. 3), it can be shown that the determinant cost function, $J(B) = \det(B)$, leads to the weight $W^*$ given by $W_2^* = TW_1^*$ and $W_3^* = T^{-1}W_1^*$, i.e., the optimal weights $W_i^*$ are related by the rotation (17). The condition $W^*H = I$ then leads to $W_1^* = [0, 2/3]^\top$. It also holds that $W^{CI}$ is given by $w_1^* = w_2^* = 1/3$ and that it is equal to $W^*$. Since the weights are equal, the second inequality in (15) reduces to equality.

Figure 3 illustrates the fusion that applies the optimal weight $W^*$. The local estimates $Z_i$ use general $H_i$, which does not allow a direct comparison with the fused estimate $\hat{X}$. Ellipsoids corresponding to local mean square error matrices $R_i$ will therefore be defined alternatively as $\varepsilon^{R_i} = \{x | x^\top H_i^\top R_i^{-1} H_i x \leq 1\}$. Here, the ellipsoids
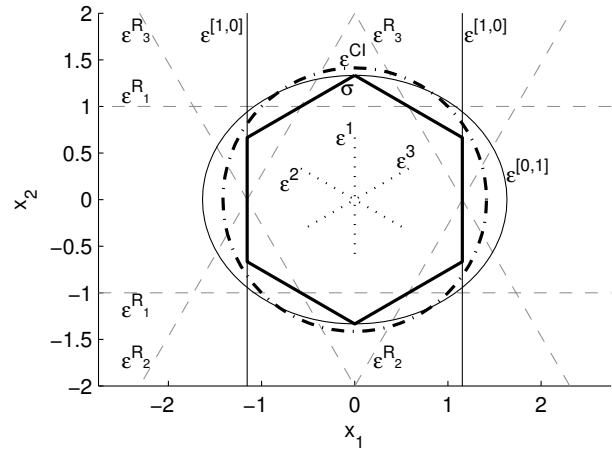


Fig. 3. Suboptimality of the family of fused bounds $B(W, \omega)$: local ellipsoids $\varepsilon^{R_i}$ (dashed parallel lines), ellipsoids $\varepsilon^{W_i R_i W_i^\top}$ (dotted line segments; denoted shortly as $\varepsilon^i$) and their Minkowski sum $\sigma$ (thick hexagon), CI ellipsoid $\varepsilon^{CI}$ (dash-dotted circle) and tangent ellipsoids $\varepsilon^{[0,1]}$, $\varepsilon^{[1,0]}$ (solid ellipses).

$\varepsilon^{R_i}$ degenerate to parallel lines. The matrices $W_i R_i W_i^\top$ dealt with in Lemma 3 are positive semi-definite and the corresponding ellipsoids degenerate to line segments in this example. The Minkowski sum $\sigma$ circumscribes the intersection of the ribbons delimited by the parallel lines $\varepsilon^{R_i}$. Thus, the necessary condition of a fusion rule that $\varepsilon^B$ has to circumscribe the intersection is clearly not sufficient in the fusion of multiple estimates.

Further, Fig. 3 visualises three fused bounds $B$. Two of them are given by choosing the normal vector $u$ from Lemma B1 in Appendix B as $[1, 0]^\top$ and $[0, 1]^\top$, respectively. The ellipsoid denoted as $\varepsilon^{[1,0]}$ degenerates to parallel lines, since $\omega_1 = 0$ holds and the division by zero (in the limit sense) produces an infinite element, $B(W, [1, 0]^\top) = \mathrm{diag}(4/3, \infty)$. The bound is tight only in the limit sense and leads to ignoring the second component. The ellipsoid denoted as $\varepsilon^{[0,1]}$ is evidently not tight, $B(W, [0, 1]^\top) = \mathrm{diag}(8/3, 16/9)$, and touches the hexagon $\sigma$ in two opposite vertices.

Since the points of tangency lie on the lines given by the choices $[1, 0]^\top$ and $[0, 1]^\top$ of the vector $u$, it is easy to compute the matrices corresponding to the inscribing and circumscribing circles of the Minkowski sum $\sigma$. They are given by $([1, 0]B(W, [1, 0]^\top)[1, 0]^\top)I$, i.e., $(4/3)I$, and $([0, 1]B(W, [0, 1]^\top)[0, 1]^\top)I$, i.e., $(16/9)I$.

The third fused bound visualised in Fig. 3 is given by the CI fusion, $\omega = w$, and it is even not tangent, $B^{CI} = 2I$, since the radius of the $\varepsilon^{CI}$ circle is $\sqrt{2}$, whereas the hexagon $\sigma$ can be circumscribed by a circle with the radius $4/3$, $B^* = (16/9)I$. By showing the suboptimality of the CI bound, it has been verified that the first inequality in (15) can be strict.

**3.2. Suboptimality of CI weight.** The following example evolves from the preceding one. It shows that both inequalities in (15) can be strict. Again, the results can be verified by hand calculations.

Let the local mean square error matrices $R_i$ be equal for $i = 1, 2, 3$, and

$$R_i = \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix},$$
$$H_1 = I, \quad H_2 = T^{-1}, \quad H_3 = T, \tag{18}$$

where $T$ is given by (17). Consider a family of fusion weights $W$ parametrised by $\theta$, $0 \le \theta \le 2/3$,

$$W_1 = \begin{bmatrix} \theta & 0 \\ 0 & \frac{2}{3} - \theta \end{bmatrix},$$
$$W_2 = TW_1, \quad W_3 = T^\top W_1. \tag{19}$$

For the determinant cost function $J(B)$, the optimal bound within the family (5) is given by $\omega_1 = \omega_2 = 1/3$ here and we have

$$B(W, \omega) = \frac{9}{2} \left( 5\theta^2 + \left( \frac{2}{3} - \theta \right)^2 \right) I$$

for such $\omega$ and $W$ (19). The optimal bound $B^*(W)$ for the given weight corresponds (in this example) to the circle that circumscribes the Minkowski sum $\sigma$. Lemma B1 can again be used for $u = [0, 1]^\top$, $u = [1, 0]^\top$, although care must be taken regarding which matrix $(u^\top B(W, u)u)I$ corresponds to the circumscribing circle. For

$$\theta \le \frac{\sqrt{5} - 1}{6},$$

the circle $\varepsilon^{B^*(W)}$ circumscribing $\sigma$ is given by

$$B^*(W) = \left( \left( \left( \frac{2}{3} - \theta \right) + \left( 3 \cdot 5\theta^2 + \left( \frac{2}{3} - \theta \right)^2 \right)^{\frac{1}{2}} \right)^2 \right) I,$$

for

$$\theta \ge \frac{\sqrt{5} - 1}{6},$$

we have

$$B^*(W) = \left( \sqrt{5}\theta + \left( 3 \left( \frac{2}{3} - \theta \right)^2 + 5\theta^2 \right)^{\frac{1}{2}} \right)^2 I.$$

The optimal CI weight $W^{\mathrm{CI}}$ is given by $w_1^* = w_2^* = 1/3$ and it can be found within the family (19) for $\theta = 1/9$. The CI bound is given by $B^{\mathrm{CI}} = (5/3)I$, which is also the optimal bound within $B(W, \omega)$, and is greater than $B^*(W^{\mathrm{CI}})$,

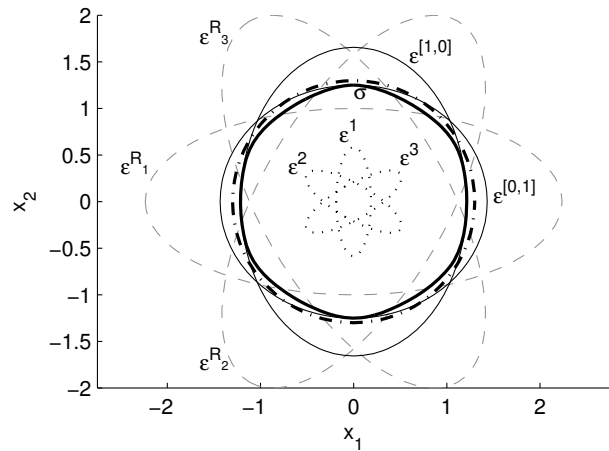$$B^*(W^{\mathrm{CI}}) = \frac{65 + 20\sqrt{10}}{81} I \approx 1.5833I.$$



Fig. 4. Comparison with Fig. 3: the dash-dotted circle now corresponds to the optimal bound within the family (5), the other lines have the same meaning.

Further, the optimal bound within $B^*(W)$ is achieved for the weight (19) given by $\theta = 1/12$ and we have

$$B^* = \frac{25}{16} I = 1.5625I.$$

That is, a weight that is not a member of the CI family (6) has been found such that a better fused upper bound $B$ can be found. The strictness of the inequality $B^*(W^{\mathrm{CI}}) > B^*$ ultimately proves that both inequalities in (15) can be strict.

*The question posed in Problem 1 has a negative answer; $W^* \ne W^{\mathrm{CI}}$ holds in this example.*

Figure 4 illustrates the optimal fusion given by $\theta = 1/12$. A graphical comparison with the example from the preceding section can be made, since the fusion given by $\theta = 0$ is illustrated by Fig. 3 in fact (with the exception of different local ellipsoids $\varepsilon^{R_i}$). Note also that no ellipsoid degenerates for $0 < \theta < 2/3$.

## 4. Extension of the family of upper bounds

The examples in Sections 3.1 and 3.2 have shown that the block-diagonal family of global bounds $Q(w)$ used by the CI fusion can fail to contain the optimal bound $Q^*$ in the fusion of multiple estimates. That means that the fusion can be improved by enlarging the family. Note that an enlarging neither guarantees that it contain $Q^*$ nor it guarantees any improvement. The following lemma will be used for an analysis.

**Lemma 4.** (Ajgl and Straka, 2017) *The union of all admissible ellipsoids $\varepsilon^R$ is equal to the Minkowski sum of the ellipsoids $\varepsilon^{M_i R_i M_i^\top}$, where the matrices $M_i$ have dimensions $(\sum_{j=1}^N n_j) \times n_i$ and all their $n_j \times n_i$ dimensional blocks except for the $i$-th one are zero, while the $i$-th block is given by the identity matrix $I$ of the size $n_i$.*

## 4.1. Analysis of the existing family of upper bounds.

The simplest case of the fusion of multiple estimates is the fusion of three scalar estimates.[3] Without a loss of generality, it is possible to assume that the local mean square error matrices are unit, $R_i = 1$, $i = 1, 2, 3$. Otherwise, the local estimates $Z_i$ and the measurement matrices $H_i$ can be formally pre-multiplied by $R_i^{\frac{1}{2}}$ to satisfy the assumption. That is, the family $A$ of the admissible matrices is the same as in the example in Section 3.1. With respect to Lemma 4, the union of all admissible $\varepsilon^R$ is the cube with the vertices $[\pm 1, \pm 1, \pm 1]$.

Figure 5 draws selected upper bounds $Q$ of $A$. The ellipsoid for the choices $Q = Q(w)$ with $w = [1/3, 1/3, 1/3]$, $w = [1/2, 1/4, 1/4]$ and $w = [0, 1/2, 1/2]$ are denoted as $\varepsilon^{\mathrm{CI}}$, $\varepsilon^{[0,1]}$ and $\varepsilon^{[1,0]}$, respectively. The chosen global bounds $Q(w)$ are tight, since the ellipsoids touch the cube at all its vertices. The correspondence with the bounds $B(W, \omega)$ of the fused mean square error matrix for the given weight $W^*$ is visible from the comparison with Fig. 3 after applying the following two steps. An orthogonal projection of the cube and ellipsoids onto a plane containing the origin of the coordinates and having a normal vector $[1, -1, -1]$ is made first. New coordinates are introduced then in order to arrive at Fig. 3. Note that whereas all global bounds $Q$ considered in Fig. 5 are tight, some fused bounds in Fig. 3 are not; see $\varepsilon^{\mathrm{CI}}$ and $\varepsilon^{[0,1]}$. The reason is that the existence of $R$ in $A$ such that $Q - R$ is rank deficient (this is a necessary condition on tangent $Q$; see Definition 1) does not imply rank deficiency of the difference $WQW^\top - WRW^\top$ (i.e., the null space of $Q - R$ need not lie in the row space of $W$).

Further, Fig. 5 shows an ellipsoid that does not have its axis aligned with the coordinate system. The ellipsoid $\varepsilon^*$ is a cylinder with the axis given by the vector $[1, -1, -1]$, which is orthogonal to the weight $W^*$ used in the example in Section 3.1. The cylinder circumscribes the cube tightly, although it does not touch the cube at all its vertices. The cylinder touches only the vertices that are projected to the hexagon in Fig. 3 and corresponds to the optimal bound $Q^*$, which is given as the limit case of $Q_\gamma$,

$$
Q_\gamma = \frac{8}{9\gamma} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}
$$
$$
+ \frac{1}{9(1-\gamma)} \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}. \tag{20}
$$

with the parameter $\gamma$, $8/9 < \gamma < 1$, approaching 1.

[3]The regularity condition that $H$ is of full column rank implies that the dimension $n$ of the estimated vector $X$ can be 1, 2 or 3. The global bounds are not related to $n$, but usefulness of an extensions of the existing family depends on $n$ significantly in the fusion problem. Namely, the extension proposed in Section 4.2 can be useful for $n = 2$, but not for $n = 1$.
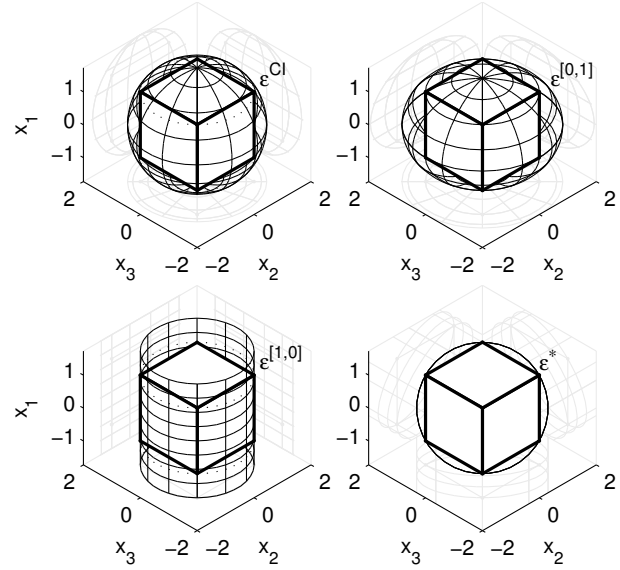


Fig. 5. Examples of global bounds $Q$: union of all admissible $\varepsilon^R$ (white cube), global ellipsoids $\varepsilon^{\mathrm{CI}}$ (sphere), $\varepsilon^{[0,1]}$ (spheroid), $\varepsilon^{[1,0]}$ (cylinder with axis $[1, 0, 0]$) and $\varepsilon^*$ (cylinder with axis $[1, -1, -1]$). The corresponding fused bounds can be found in Fig. 3.

Alternatively, the optimal bound $Q^*$ can be described by the inverse, which is directly needed in (13), (14) anyway. The inverse $(Q_\gamma)^{-1}$,

$$
(Q_\gamma)^{-1} = \frac{\gamma}{8} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}
$$
$$
+ (1 - \gamma) \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}, \tag{21}
$$

is a positive semi-definite matrix for $\gamma = 1$. That is, the optimal global bound $Q^*$ is improper (infinite). Nevertheless, the optimal fused bound $B^*$ is finite.

Finally, it has to be noted that the way to meet the assumption of unit $R_i$ is not unique. If the local estimates $Z_i$ and the measurement matrices $H_i$ are pre-multiplied by $-R_i^{\frac{1}{2}}$ instead of $R_i^{\frac{1}{2}}$, the family $A$ of admissible matrices $R$ remains the same. However, the parametrisation of the family changes signs. As a result, the elements of the corresponding row and column of the resulting bound $Q_\gamma$ change their signs as well (the corresponding diagonal element changes its sign twice, i.e., remains positive). Therefore, it is proposed to extend the family $Q(w)$ of block-diagonal upper bounds of $A$ by non-diagonal matrices in the way that is described in the following section.

**4.2. Extension of the family for a special case.** For the case $N = 3$, $n_i = 1$, it is proposed to extend the block-diagonal family $Q(w)$ by four improper upper bounds $Q_j$, $j = 1, \ldots, 4$, the inverses of which are non-diagonal and given by

$$Q_1^{-1} = \frac{1}{8} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix},$$

$$Q_2^{-1} = \frac{1}{8} \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix},$$

$$Q_3^{-1} = \frac{1}{8} \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix},$$

$$Q_4^{-1} = \frac{1}{8} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}. \tag{22}$$

As a weighted harmonic mean of upper bounds is another upper bound, the family can be further extended by any such a combination, e.g., by

$$Q = (\delta Q_j^{-1} + (1 - \delta)(Q(w))^{-1})^{-1}, \tag{23}$$

where $\delta$, $0 \leq \delta \leq 1$, is a parameter. Note also that $Q_\gamma$ is actually given by

$$\left( \delta Q_1^{-1} + (1 - \delta) \left( Q \left( \left[ \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right] \right) \right)^{-1} \right)^{-1},$$

where $\delta$ is given by $\delta = 9(\gamma - 8/9)$.

The larger the family of $Q$, the bigger the chance to contain the optimal bound $Q^*$. Nevertheless, it is useless to inspect all combinations of all $Q_j^{-1}$, $j = 1, \ldots, 4$, with all $Q(w)$. For example, the combination $(\sum_{j=1}^{4} (1/4) Q_j^{-1})^{-1}$ gives $Q = 4I$, i.e., a bound larger than $Q([1/3, 1/3, 1/3])$, which is equal to $3I$. Unfortunately, there is no guideline for constructing a useful family that is guaranteed to contain $Q^*$ and smaller than the set of all matrices $Q$ fulfilling $Q \geq A$.

**4.3. Comments to a general case.** It is also not clear how to propose improper upper bounds $Q$ for $n_i \geq 1$ with at least one $i$ such that $n_i > 1$. Some remarks are given below.

Since the dimension of the global mean square error matrix $R$ is at least four, visualisation of the global upper bounds $Q$ is not tractable. If the case $n_1 = 2$, $n_2 = n_3 = 1$ is considered, different families of bounds are needed for $n = 4$, $n = 3$ and $n = 2$. A general guideline is that the rank of the inverses $Q^{-1}$ of the improper upper bounds $Q$ should be equal to $n$, as well as the rank of $H^\top Q^{-1} H$. However, the design of the range of such $Q^{-1}$ is equivalent to the design of weight $W$ and the design of

a particular $Q^{-1}$ with the given range is equivalent to the design of upper bounds of $WAW^\top$. That is, the bounds $Q$ to be inspected in (13), (14) are designed by the complete search for (3), (4). Although the family is smaller than the set of all matrices $Q$ fulfilling $Q \geq A$, it is not simple, since the complete search has to be done.

As the extension of the $Q(w)$ family for a general case is either not guaranteed to contain the optimal $Q^*$ or is equivalent to the complete search over all weights $W$ and bounds $B$, it is not dealt with here. Since the global bounds $Q$ generate the fused bounds $B$, the following conclusion is made.

*The question posed in Problem 2 has a negative answer*; the families are either not simple (brute force design has to be adopted), or not tight.

## 5. Summary

Fusion of estimates has been discussed under the assumption that the cross-correlations of the estimation errors are unknown. An explicit expression for the union of all admissible fused ellipsoids has been used to demonstrate that the CI fusion of multiple estimates is suboptimal in the sense of the provided upper bounds as well as in the quality of the weights of the linear combination. An analysis has shown the reason for the suboptimality: the existing family of tight upper bounds of the global mean square error matrices does not contain all tight upper bounds. A remedy has been indicated next. Four improper non-diagonal bounds have been presented for a special case and the extension by combining the bounds has been proposed. Since the candidate bounds for the optimal fusion are based on the direct solution to the fusion problem, no simple family of bounds can be given for a general case. The search for suboptimal solutions has been left for the future.

## References

Ajgl, J. and Šimandl, M. (2014). On linear estimation fusion under unknown correlations of estimator errors, *Proceedings of the 19th IFAC World Congress, Cape Town, South Africa*, pp. 2364-2369.

Ajgl, J., Šimandl, M., Reinhardt, M., Noack, B. and Hanebeck, U.D. (2014). Covariance Intersection in state estimation of dynamical systems, *Proceedings of the 17th International Conference on Information Fusion, Salamanca, Spain*.

Ajgl, J. and Straka, O. (2016a). Covariance Intersection in track-to-track fusion with memory, *Proceedings of the*

*2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, Baden-Baden, Germany*, pp. 359–364.

Ajgl, J. and Straka, O. (2016b). Covariance Intersection in track-to-track fusion without memory, *Proceedings of the 19th International Conference on Information Fusion, Heidelberg, Germany*.

Ajgl, J. and Straka, O. (2017). A geometrical perspective on fusion under unknown correlations based on Minkowski sums, *Proceedings of the 20th International Conference on Information Fusion, Xi'an, China*, pp. 725–732.

Arambel, P.O., Rago, C. and Mehra, R.K. (2001). Covariance intersection algorithm for distributed spacecraft state estimation, *Proceedings of the American Control Conference, Arlington, VA, USA*, pp. 4398–4403.

Bar-Shalom, Y., Li, X.-R. and Kirubarajan, T. (2001). *Estimation with Applications to Tracking and Navigation*, Wiley, Hoboken, NJ.

Bar-Shalom, Y., Willet, P.K. and Tian, X. (2011). *Tracking and Data Fusion: A Handbook of Algorithms*, YBS Publishing, Storrs, CT.

Chen, L., Arambel, P.O. and Mehra, R.K. (2002). Fusion under unknown correlation—covariance intersection as a special case, *Proceedings of the 5th International Conference on Information Fusion, Annapolis, MD, USA*, pp. 905–912.

Deng, Z., Zhang, P., Qi, W., Liu, J. and Gao, Y. (2012). Sequential covariance intersection fusion Kalman filter, *Information Sciences* **189**: 293–309.

Gao, Y., Li, X.R. and Song, E. (2016). Robust linear estimation fusion with allowable unknown cross-covariance, *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **46**(9): 1314–1325.

Hanebeck, U.D., Briechle, K. and Horn, J. (2001). A tight bound for the joint covariance of two random vectors with unknown but constrained cross-correlation, *Proceedings of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, Baden-Baden, Germany*, pp. 147–152.

Hu, J., Xie, L. and Zhang, C. (2012). Diffusion Kalman filtering based on covariance intersection, *IEEE Transactions on Signal Processing* **60**(2): 891–902.

Julier, S.J. and Uhlmann, J.K. (1997). A non-divergent estimation algorithm in the presence of unknown correlations, *Proceedings of the American Control Conference, Albuquerque, NM, USA*, pp. 2369–2373.

Julier, S.J. and Uhlmann, J.K. (2001). General decentralized data fusion with covariance intersection, *in* D.L. Hall and J. Llinas (Eds.), *Handbook of Multisensor Data Fusion*, CRC Press, Boca Raton, FL.

Kowalczuk, Z. and Domżalski, M. (2013). Asynchronous distributed state estimation for continuous-time stochastic processes, *International Journal of Applied Mathematics and Computer Science* **23**(2): 327–339, DOI: 10.2478/amcs-2013-0025.

Kurzhanski, A.B. and Vályi, I. (1997). *Ellipsoidal Calculus for Estimation and Control*, Birkhäuser, Boston, MA.

Lehmann, E.L. and Casella, G. (1998). *Theory of Point Estimation*, 2nd Edn., Springer, Berlin/Heidelberg.

Li, X.-R., Zhu, Y., Wang, J. and Han, C. (2003). Optimal linear estimation fusion—Part I: Unified fusion rules, *IEEE Transactions on Information Theory* **49**(9): 2192–2208.

Reinhardt, M., Noack, B., Arambel, P.O. and Hanebeck, U.D. (2015). Minimum covariance bounds for the fusion under unknown correlations, *IEEE Signal Processing Letters* **22**(9): 1210–1214.

Seeger, A. (1993). Calculus rules for combinations of ellipsoids and applications, *Bulletin of the Australian Mathematical Society* **47**(1): 1–12.

Simon, D. (2006). *Optimal State Estimation: Kalman, H-Infinity, and Nonlinear Approaches*, Wiley, Hoboken, NJ.

Uhlmann, J.K. (2003). Covariance consistency methods for fault-tolerant distributed data fusion, *Information Fusion* **4**(3): 201–215.

Wu, Z., Cai, Q. and Fu, M. (2018). Covariance intersection for partially correlated random vectors, *IEEE Transactions on Automatic Control* **63**(3): 619–629.

**Jiří Ajgl** received the MSc degree in cybernetics and control engineering and the PhD degree in cybernetics from the University of West Bohemia, Pilsen, Czech Republic, in 2008 and 2013, respectively. He is with the European Centre of Excellence—New Technologies for Information Society at the University of West Bohemia. He is the author of twenty papers. His research interest include state estimation methods with the focus on data and information fusion. Dr. Ajgl received the Best Paper Award at the 2016 IEEE International Conference on *Multisensor Fusion and Integration for Intelligent Systems*.

**Ondřej Straka** received the MSc degree in cybernetics and control engineering and the PhD degree in cybernetics from the University of West Bohemia, Pilsen, Czech Republic, in 1998 and 2004, respectively. Since 2015, he has been an associate professor with the Department of Cybernetics there. He is the head of the Identification and Decision Making Research Group. He has participated in several research projects (e.g., GNSS-based safe train localization and an attitude and heading reference system) and has published over 70 papers. He focuses on state estimation, system identification and fault detection. Dr. Straka was a recipient of the Werner von Siemens Excellence Award in 2014 for the most important result in basic research.

## Appendix A

The matrices used to generate Figs. 1 and 2 are described below.

The classic full-state estimation problem is considered, $H_1 = H_2 = I$. The local mean square error matrices $R_i$ are given as

$$R_1 = \begin{bmatrix} 16 & 0 \\ 0 & 1 \end{bmatrix}, \quad R_2 = \begin{bmatrix} 4 & 4 \\ 4 & 9 \end{bmatrix}. \qquad (A1)$$

The generalised eigenvalue decomposition of the matrix pair $(R_1, R_2)$ is used, i.e., a diagonal matrix $D$ and an invertible matrix $V$ are computed such that, $R_1 V = R_2 V D$. The diagonal elements $d_j$, $j = 1, \ldots, n$, are sorted in the descending order and the columns of $V$ are permuted correspondingly. The local matrices $R_i$ can then be expressed as

$$R_1 = (V^{-1})^\top D V^{-1}, \quad R_2 = (V^{-1})^\top I V^{-1} \quad \text{(A2)}$$

and the optimal processing of hypothetical variables can be considered as follows.

The mean square error matrix $R_c$ of a hypothetical common prior estimate (measurement) is given by

$$R_c = (V^{-1})^\top \max\{D, I\} V^{-1}, \quad \text{(A3)}$$

where the maximum is meant element-wise. Hypothetical individual measurement matrices $C_a$, $C_b$ are given by the $j$-th rows of $V^\top$, where $j$ are given by $d_j < 1$ and by $d_j > 1$, respectively. Hypothetical individual mean square error matrices $R_a$, $R_a$ are introduced as diagonal matrices with elements $d_j(1-d_j)^{-1}$ selected for $j$ given by $d_j < 1$ and elements $d_j(d_j-1)^{-1}$ selected for $j$ given by $d_j > 1$, respectively. Hypothetical cross-correlations $R_{c,a}$, $R_{c,b}$ are assumed to be zero. Then, we have

$$\begin{aligned} R_1^{-1} &= R_c^{-1} + C_a^\top R_a^{-1} C_a, \\ R_2^{-1} &= R_c^{-1} + C_b^\top R_b^{-1} C_b. \end{aligned} \quad \text{(A4)}$$

If a hypothetical cross-correlation $R_{a,b}$ is assumed to be arbitrary, it can be parametrised as

$$R_{a,b} = R_a^{\frac{1}{2}} \Omega R_b^{\frac{1}{2}}, \quad \text{(A5)}$$

where the square roots (of the diagonal matrices $R_a$, $R_b$) are meant component-wise and $\Omega$ is any matrix of appropriate dimensions satisfying $I \geq \Omega\Omega^\top$.

Thus, the selected admissible cross-correlations $R_{1,2}$ are parametrised as

$$R_{1,2} = R_1(R_c^{-1} + C_a^\top R_1^{-\frac{1}{2}} \Omega R_2^{-\frac{1}{2}} C_b)R_2 \quad \text{(A6)}$$

and in Fig. 1, the values $\Omega = -0.999, -0.5, 0, 0.5, 0.999$ are used.

In Fig. 2, the weights $W(w)$ (6) are used with $w_1 = 1/2$ and the upper bounds $B(W(w), \omega)$ (5) with $\omega_1 = 0.271, 0.5, 0.751$; see the work of Ajgl and Straka (2017, Theorem 6). The illustrated admissible fused mean square error matrices are given by cross-correlations $R_{1,2}$ parametrised as

$$\begin{aligned} R_{1,2} &= R_1^{\frac{1}{2}} \Omega (R_2^{\frac{1}{2}})^\top, \\ \Omega &= \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{bmatrix}, \end{aligned} \quad \text{(A7)}$$

where the square roots of $R_i$ are the lower triangular matrices obtained by the Cholesky decomposition and the angles $\alpha$ are selected as $\alpha = 0, \pi/2, \pi, 3\pi/2$; see also the work of Ajgl and Straka (2017, Section IV.A).

The choice of $\Omega$ in (A7) is based on experiments wherein only several structures were considered. Note, however, that other choices can also be used. For example, an $n$-dimensional column vector $v$ can be chosen first (without a loss of generality as a unit vector) and an admissible fused mean square error matrix can be searched next, such that the boundary of the Minkowski sum $\sigma$ is tangent to the corresponding admissible fused ellipsoid and $v$ is a normal vector at a point of tangency. That is, elements of $\arg\max_{\Omega:I \geq \Omega\Omega^\top}(v^\top W A(\Omega) W^\top v)$ are sought. It can be shown that

$$\Omega = \frac{(R_1^{\frac{1}{2}})^\top W_1^\top v v^\top W_2 R_2^{\frac{1}{2}}}{(v^\top W_1 R_1 W_1^\top v)^{\frac{1}{2}} (v^\top W_2 R_2 W_2^\top v)^{\frac{1}{2}}} \quad \text{(A8)}$$

is such an element. The parametrisation (A8) of $\Omega$ can, therefore, be chosen for an illustration instead of (A7).

## Appendix B

**Lemma B1.** (*External ellipsoidal approximation*) *Let $u$ be a unit column vector of length $n$. If the parameter $\omega$ of the family* (5) *is chosen according to*

$$\omega_i \propto (u^\top W_i R_i W_i^\top u)^{\frac{1}{2}}, \quad \text{(B1)}$$

*where $\propto$ means equality up to a multiplicative constant guaranteeing $\sum_{i=1}^N \omega_i = 1$, the corresponding upper bound $B(W, u)$ of $WAW^\top$ is tangent. Further, a point on the boundary of the ellipsoid $\varepsilon^{B(W,u)}$ exists such that it is a point of the boundary of the Minkowski sum $\sigma$ from Lemma 3 and $u$ is a normal vector of the ellipsoid at that point.*

*Proof.* Lemma 3 links matrices with geometrical objects. Seeger (1993, Theorem 2.3) provides a relation of the family (5) with the Minkowski sum $\sigma$, while Kurzhanski and Vályi (1997, Section 2.2) give details about the parametrisation. The lemma describes the geometrical interpretation of tangency; see also the work of Ajgl and Straka (2017, Lemma 5). ∎