Decentralized Data Fusion with Probabilistically Conservative Ellipsoidal Intersection

Aryan Naveen and Clark Taylor
Autonomy and Navigation Technology (ANT) Center
Air Force Institute of Technology
{aryanpn, clark.n.taylor}@gmail.com

***Abstract*— When performing decentralized data fusion, one of the major challenges is the well-known "track-to-track" correlation problem. One approach to handling this correlation is to use correlation-agnostic fusion techniques that generate *conservative* estimates of the fused probability distribution, irrespective of the amount of common information present. When working with Gaussian distributions, past approaches have only considered the covariance of the input distributions to compute the covariance of the fused distribution. In this paper, we introduce another constraint that considers the means of the input distributions. While this constraint cannot 100% guarantee the output is conservative, it can "probabilistically" – to a user-specified certainty – ensure the output is conservative. This introduction of a probabilistic bound enables fused probability distributions with less excess covariance. Specifically, we modify the ellipsoidal intersection technique to include the probabilistically conservative constraint. We present results that show using the probabilistic bound generates a fused distribution that is still conservative but with less excess covariance than Ellipsoidal Intersection without the probabilistic constraint.**

I. Introduction

With the significant reduction in size, weight, and power requirements for both sensors and computational devices, coupled with increased battery capacity and the increased ability to wirelessly communicate data between sensors, there has been significant interest in the concept of distributed data networks collecting data and performing fusion over these networks [1] [2] [3].

The basic method for "combining" probability distribution functions (PDFs) from multiple sensors is to use Bayes' rule, which (roughly) involves multiplying the input PDFs together and then normalizing. This approach, however, leads to one of the central problems in enabling distributed data fusion networks: the "track-to-track correlation" problem. Using Bayes' rule assumes that the input distributions to the fusion center are independent, or that the correlation between the inputs is known and can be removed. In a distributed data fusion technique, correlations are introduced both by "feedback" (when a node's own estimates returns back to itself after being used by other nodes), or by different "tracks" of data coming from a single node to the node performing fusion. Unless these correlations are explicitly considered, simply multiplying PDFs together will result in fusion estimates that are over-confident.

Addressing this correlation problem has been the focus of significant efforts in the past, including enforcing specific network structures that eliminate double-counting[3], maintaining bookkeeping on each piece of data to enable the proper removal of the correlation through "channel filters"[4], and proposing data fusion techniques that yield conservative probability distributions regardless of the correlation on the inputs[5]. In this paper, we focus on this last technique as it is applicable to multiple, time-varying network structures and does not require any additional bookkeeping information be transmitted between the nodes, helping minimize the communication required for data fusion.

Starting with the introduction of Covariance Intersection (CI)[5] in 1997, several different techniques have been introduced to enable correlation-agnostic fusion. These techniques have ranged from considering *only* Gaussian distributions (and providing Gaussian distribution outputs)[5], [6], [7], [8], [9] and more general methods that allow any type of probability distribution functions (PDFs) to be considered[10], [11], [12], [13], [14]. In this paper, we are focused on the problem where Gaussian distributions are received and a Gaussian distribution output is desired.

In previous works considering correlation-agnostic fusion of Gaussian PDFs, the covariance of the fusion technique's output was determined solely by the covariances of the input distributions. Specifically, output covariances were computed such that any (admissible) quantity of common information could be present in the input distributions and the output covariance would still be conservative. The primary contribution of this paper is a technique for considering both the covariance and mean information of the input distributions when performing data fusion. The mean information is used to (probabilistically) bound the amount of common information that is admissible in the input distributions. Utilizing this extra constraint, we are able to obtain fusion results that, while still conservative compared with the "optimal" (known correlation) fusion result, are less conservative than prior correlation-agnostic fusion techniques.

This paper is organized as follows. In Section II, we define the data fusion problem and introduce notation utilized for the remainder of this paper. We also review the Ellipsoidal Intersection approach to correlation-agnostic fusion. In Section III, we introduce our definition of *probabilistically admissible* using the means of the input distributions to bound the amount of common information between the input distributions. We use this definition in Section IV to modify Ellipsoidal Intersection, deriving a new correlation-agnostic fusion algorithm for Gaussian distributions. Section V presents results demonstrating the efficacy of our proposed fusion approach. Section VI concludes the paper.

II. Review of Fusion Problem

To perform fusion, assume two sensors or sources of information about a quantity are available, each providing a probability distribution function (PDF) representing its knowledge (P_a and P_b). When the two PDFs are (conditionally) independent, the optimal fused PDF (P_ϕ) is:

$$P_\phi(x) = \frac{P_a(x)P_b(x)}{\int_S P_a(x)P_b(x)dx} \; \forall x \in S \tag{1}$$

where S is the region of support for both input PDFs (i.e. $P(x) = 0$ for all $x \notin S$).

When the PDFs P_a and P_b contain some common information, we can write the inputs in terms of the common and independent information as:

$$P_a(x) \simeq P_{a\backslash c}(x)P_c(x) \tag{2}$$

$$P_b(x) \simeq P_{b\backslash c}(x)P_c(x), \tag{3}$$

where $\simeq$ denotes equality up to a normalization factor. The *optimal* data fusion result in this case is:

$$P_\phi(x) \simeq P_{a\backslash c}(x)P_{b\backslash c}(x)P_c(x). \tag{4}$$

Because we are focusing on Gaussian distributions in this paper, the input, common, and fused distributions can be completely represented by the mean (μ) and covariance matrix (C) associated with that distribution. Therefore, we have (μ_a, C_a) and (μ_b, C_b) representing the information to be fused and the common information represented by (μ_c, C_c). The input distribution (μ_a, C_a) can be composed from independent distributions $(\mu_{a\backslash c}, C_{a\backslash c})$ and (μ_c, C_c) as:

$$\mu_a = C_a(C_{a\backslash c}^{-1}\mu_{a\backslash c} + C_c^{-1}\mu_c) \tag{5}$$

$$C_a^{-1} = C_{a\backslash c}^{-1} + C_c^{-1} \tag{6}$$

with P_b composed from $(\mu_{b\backslash c}, C_{b\backslash c})$ and (μ_c, C_c) similarly.

Using this specialization of (2) and (3), we can obtain the optimal fusion result (denoted by ϕ) for Gaussians:

$$\mu_\phi = C_\phi(C_{a\backslash c}^{-1}\mu_{a\backslash c} + C_{b\backslash c}^{-1}\mu_{b\backslash c} + C_c^{-1}\mu_c) \tag{7}$$

$$C_\phi^{-1} = C_{a\backslash c}^{-1} + C_{b\backslash c}^{-1} + C_C^{-1} \tag{8}$$

Note that all covariance matrices are required to be positive semi-definite (p.s.d.). A matrix (M) is p.s.d. if

$$x^\top M x \geq 0 \; \forall x. \tag{9}$$

Ensuring all matrices are p.s.d. leads to the following definition:

Definition 1. *A common covariance matrix C_c is* admissible *if the matrices $C_c - C_a$ and $C_c - C_b$ are p.s.d.*

A. Ellipsoidal Intersection Fusion Algorithm

To create a fusion algorithm that considers probabilistic constraints on the common information, we will be modifying the Ellipsoidal Intersection (EI) algorithm[6]. To review, the EI algorithm attempts to solve the optimization problem

$$\begin{aligned} C_c = &\operatorname*{argmin}_{\Upsilon} |\Upsilon| \\ s.t.\ &x^\top(\Upsilon - C_a)x \geq 0 \; \forall x. \\ &x^\top(\Upsilon - C_b)x \geq 0 \; \forall x. \end{aligned} \tag{10}$$

where $|\Upsilon|$ represents the determinant of the matrix and the constraints express the admissible constraint of Definition 1.

To find the matrix that has the minimum determinant, given the constraints, ellipsoidal geometry is used to derive the closed form expression

$$\begin{aligned} D_{C_c} &= \begin{cases} \max(D_2,\, 1) & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \\ C_c &= E_a D_a^{0.5} Q_b D_{C_c} Q_b^T D_a^{0.5} E_a^T + \epsilon I \end{aligned} \tag{11}$$

where ϵ is a small numerical constant used for numerical stability when computing the mean and

$$\begin{aligned} C_a &= E_a D_a E_a^T \\ Q_b D_b Q_b^T &= D_a^{-0.5} E_a^T C_b E_a D_a^{-0.5} \end{aligned} \tag{12}$$

From the calculated C_c, EI finds the common mean as

$$\mu_c = (C_a^{-1} + C_b^{-1} - 2C_c^{-1})^{-1}(C_{a\backslash c}^{-1}\mu_a + C_{b\backslash c}^{-1}\mu_b) \tag{13}$$

Using the estimated common mean and covariance, the fused estimate (μ_f, C_f) is found similarly to Equations (7) and (8).

III. The probabilistically Admissible Constraint

Note that EI and other correlation agnostic fusion techniques have, in essence, used the covariance matrices of the inputs to impose constraints on what the fused result can be. By finding the maximum amount of common information that could be present, *within the constraints imposed by the input covariance matrices*, a correlation agnostic fusion technique is achieved. We believe that the input distribution means can also be used to indicate how much common information is present in the inputs. Conceptually, if the means of two input distributions are different, even if their covariance is the same, that is an indicator of independent information being present in the inputs.

We assume that the means of each input distribution can be constructed as

$$\mu = x + \nu \tag{14}$$

where μ is the mean of the input distribution, x is the true value of the quantity being estimated, and ν is a random variable sampled from a zero mean Gaussian variable with covariance C. When an input distribution is received by the fusion algorithm, it receives both μ and C. Because x from Equation (14) is assumed constant across all inputs to a fusion algorithm, this leads to the following key proposition:

Proposition 2. *If multiple Gaussian PDFs express uncertainty about a single estimate, the differences in their means are due solely to the presence of noise.*

Corollary 3. *A difference of means between two inputs implies the presence of independent information in the inputs.*

Based on Proposition 2 and assuming $\mu_{a\backslash c}$, $\mu_{b\backslash c}$, and μ_c are estimates of the same quantity, then the differences in mean value should be due solely to noise, i.e. different estimates of the quantity being estimated. This proposition can be expressed as a null hypothesis test

$$H_0 : \mu_{a\backslash c} - \mu_{b\backslash c} = 0 \tag{15}$$

$$H_A : \mu_{a\backslash c} - \mu_{b\backslash c} \neq 0 \tag{16}$$

where the alternative hypothesis is that $\mu_{a\backslash c}$ and $\mu_{b\backslash c}$ are estimates of different quantities.[1] If the alternative hypothesis were to be accepted, that would indicate that the two input data sources were actually measurements of different quantities and should not be fused together. This leads to the following definition:

Definition 4. *An estimate of common information is* probablistically admissible *if there is not sufficient evidence to reject the null hypothesis that they are estimates of the same quantity.*

By constraining the common information estimates to probabilistically admissible estimates, the estimated quantity of common information will be reduced, leading to fusion results that are less conservative than previous techniques.

For Gaussian distributions, the null hypothesis test can be expressed as a threshold on the Mahalanobis distance

$$\Psi(C_c) = (\mu_{a\backslash c} - \mu_{b\backslash c})^\top (C_{a\backslash c} + C_{b\backslash c})^{-1} (\mu_{a\backslash c} - \mu_{b\backslash c}) \tag{17}$$

where $C_{a\backslash c}$, $C_{b\backslash c}$, $\mu_{a\backslash c}$, and $\mu_{b\backslash c}$ are all functions of C_c using Equation (6). The threshold value γ is derived from a user selected level of significance $\alpha = p(\Psi \geq \gamma | H_0)$, or the probability of rejecting the null hypothesis when the null hypothesis was actually true. Given α, the threshold for the Mahalanobis distance γ can be computed from a chi-squared distribution of degree equal to the number of dimensions in the state space. If the Mahalanobis distance (Ψ) is greater than γ, the null hypothesis is rejected.

A. Impact of Probabilistically Admissible Constraint

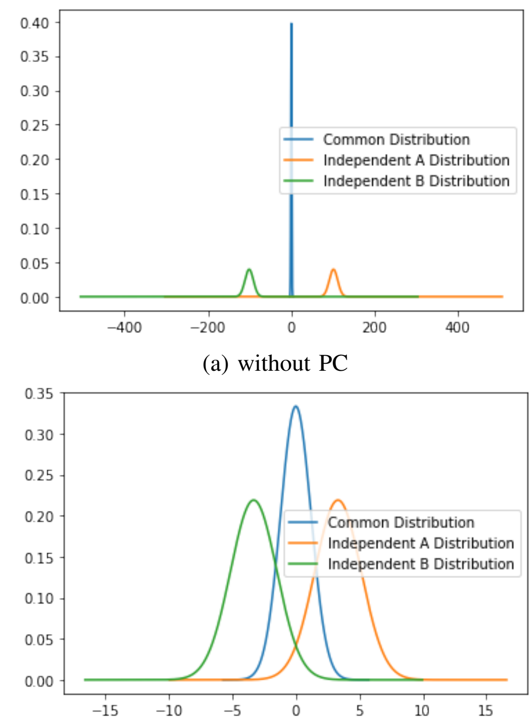
This subsection analyzes various situations illustrating the impact of the the probabilistically admissible constraint. For pedagogical purposes, we focus on one-dimensional distributions. Note that in 1D, traditional Ellipsoidal Intersection always sets C_c to the larger input covariance as that is the smallest common covariance that meets the admissible constraints of Definition 1. For each case study, we consider a certain characteristic in the initial distributions p_a and p_b, and analyze the impact of adding a probabilistic constraint to the estimate of C_c.

*1) **Case 1. Input Distribution Covariances are the same**:* The impact of adding the probabilistically admissible constraint is most obvious in the case when the two input distributions have the same covariance. In this scenario, Equation (13) devolves into a 0 over 0 equation, except for the addition of the ϵI term in Equation (11). Therefore, the mean values $\mu_{a\backslash c}$ and $\mu_{b\backslash c}$ will be "pushed" very far away from the input means μ_a and μ_b, with the "pushed" distance determined by the ratio of C_a (or C_b) to ϵ.

[1]We have selected $\mu_{a\backslash c}$ and $\mu_{b\backslash c}$ because the most likely μ_c will always be "in-between" these two μ values. Therefore, using $\mu_{a\backslash c}$ and $\mu_{b\backslash c}$ will give us the most restrictive constraints on common covariance.

In Figure 1a, the effect of this ϵI term is illustrated. To create this figure, we assumed two input distributions of $p_a = (1, 1)$ and $p_b = (-1, 1)$, creating an estimated common covariance of 1 (the blue plot in Figure 1a. Using an ϵ value of 0.01 results in $p_{a\backslash c} = (101, 101)$ and $p_{b\backslash c} = (-101, 101)$ (both plotted). Note that these values for μ are (1) very large and do not lead to probabilistically consistent values ($\Psi \approx 100.5$) and (2) should actually be much larger for more reasonable (smaller) values of ϵ.

If instead, we constrain C_c to be probabilistically admissible with an α of .05, the independent distributions will be much closer together as shown in Figure 1b. Specifically, setting $C_c \approx 1.43$ results in a p-value of 5% and the distributions $p_{a\backslash c} \approx (3.32, 3.32)$ and $p_{b\backslash c} \approx (-3.32, 3.32)$. Note that when the probabilistic constraint is added: (1) the estimated common covariance is larger, (2) the estimated independent covariances are smaller, (3) the independent means are much closer to the input distributions means, and (4) the fused result has a smaller covariance (.77 vs 1).

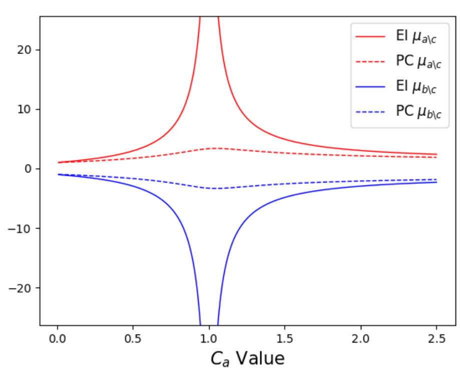


(b) with PC

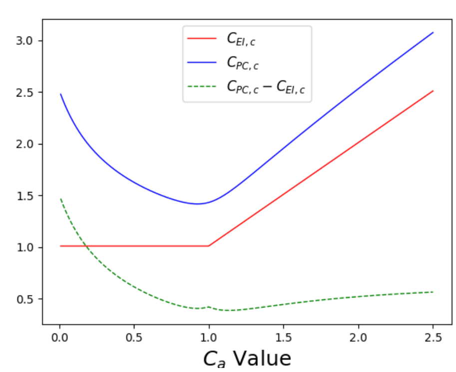
Fig. 1: Visualization of the estimated 1-dimensional independent distributions both without (a) and with (b) the probabilistic constraints. Input distributions have covariances of 1 and means of 1 and -1. Note the significantly different ranges on the x axis between the subfigures.

*2) **Case 2: Input Distribution Covariances are different**:* To demonstrate the impact the probabilistically admissible constraint has when the input covariances are different, we set $p_b = (1, 1)$. The other distribution is set to have a mean value of -1, but the covariance (C_a) varies. In Figure 2a, we plot, as a function of C_a (the x-axis), what the estimated independent mean values are without (the solid, outside lines) and with (the interior, dashed lines) the probabilistically admissible constraint enforced. Note that as C_a approaches 1 (i.e. the same value as C_b), the mean values asymptotically go to $\pm\infty$ without the probabilistically admissible constraint. (The graph is truncated in the y direction for readability.) This is the "same covariance value" case discussed above. To be probabilistically admissible, however, the mean values will stay closer to the average of the two values. This is illustrated by the interior mean value lines in 2a.

In addition to revised mean values, Figure 2b shows the effect of the probabilistic constraint (PC) on the estimated common covariance. In this figure, the red line indicates what EI would have chosen as the common covariance, while the blue line shows the minimum common variance that is allowed by the probabilistic constraint. The green line shows the difference between the two estimates of common covariance. Note that the traditional covariance estimate will always be smaller than what the probabilistically admissible constraint allows. The significant increase in the estimated common covariance on the left side (as C_a decreases below 1) is because as C_a decreases, the probability of having two means spaced 2 units apart is very small if there were significant common information.



(a) Computed *Independent* Means of EI and PC fusion



(b) Common Covariances of EI and PC with varying C_a

Fig. 2: Effect of probabilistic constraint (PC) on the independent mean estimates (subfigure a) and the common covariance estimate (subfigure b) as C_a varies. Other values are $\mu_a = 1$ and $p_b = (1, 1)$.

*3) **Case 3: Input Distribution Means are equal:*** Note that when the initial distribution means are equal (i.e. $\mu_a = \mu_b$), the addition of the probabilistic constraint does not influence estimated common covariance.

IV. Fusion using the Probabilistically Admissible Constraint

The central idea behind Ellipsoidal Intersection is a constrained optimization problem (Equation (10)) wherein the determinant of the common covariance is the optimization variable to be minimized, while constraining the common covariance to *admissible* values. Similarly, we can create a probabilistically constrained (PC) fusion technique described using the following optimization problem:

$$C_c = C_{c,EI} + SS^T \tag{18}$$

$$\begin{aligned} S := &\operatorname*{argmin}_{\Upsilon} |\Upsilon| \\ &s.t.\ \Psi(C_{c,EI} + \Upsilon\Upsilon^T) \leq \gamma \end{aligned} \tag{19}$$

where Υ is a lower triangular matrix, γ is the chi-square threshold selected by the user for a problem of dimension N, $C_{c,EI}$ is the common covariance computed by ellipsoidal intersection, and $\Psi(\ldots)$ is described by (17). This optimization problem finds the smallest matrix SS^T (where small is defined by its determinant) required to make C_c meet the probabilistic constraint. We use the SS^T parameterization of symmetric matrices (where S is a lower triangular matrix) because (1) this ensures that we are generating a p.s.d. matrix and (2) the determinant of a triangular matrix is just the product of its diagonals, making it easy to compute. Note that because SS^T is added to the common covariance from ellipsoidal intersection, the computed common covariance will be both admissible *and* probabilistically admissible.

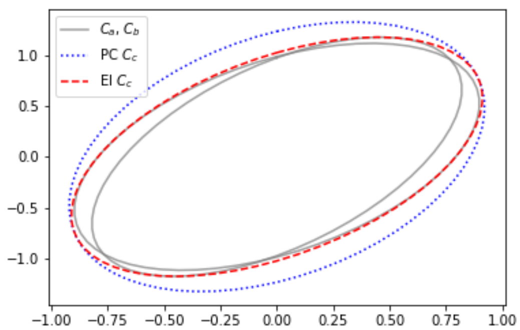
V. Fusion Results

To evaluate the effectiveness of the probabilistically constrained fusion algorithm, we evaluate two different scenarios in this section. First, we consider the scenario where two distributions are being fused together and evaluate both the output covariance and mean squared error when using the probabilistically constrained fusion method (PC) and the traditional ellipsoidal intersection (EI) method. Second, we look at a distributed estimation scenario where either EI or PC is applied repeatedly across the links of the network until the entire network converges. Each of these are described in more detail in the following subsections.

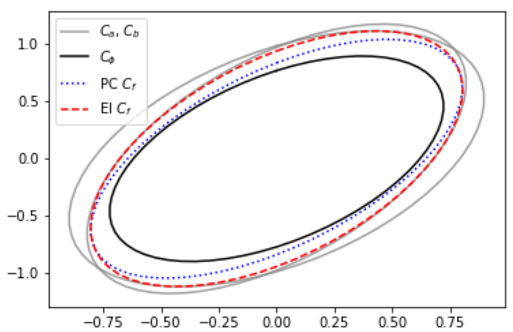
A. Two input distributions

In this section, we analyze the results of performing fusion on two input Gaussian distributions. We analyze a single instance to show the potential results of PC fusion and some Monte Carlo simulations showing more general performance.

In Figure 3, we show, for a single instance, the covariance output from both PC and EI fusion. In subfigure (a), we plot the level set of the common covariance matrix derived by both EI and the PC techniques. Note that the estimated common covariance of the PC method is larger, meaning it estimates less common information between the two inputs. This decrease in common information leads to the fused covariance outputs illustrated in subfigure (b). Note that the fused covariance when probabilistically constrained is both smaller and closer to the truth than when using EI.



(a) Common covariances



(b) Fused covariances

Fig. 3: Comparison of PC (blue dotted) and EI (red dashed) fusion approaches. Light gray ellipses denote the input covariances and dark black the optimal fusion result (subfigure (b)). Note that PC has a larger estimated common covariance (a) leading to a smaller fused covariance output (b).

While the results illustrated in Figure 3 are promising, we wanted to know how often PC leads to a significant improvement in fusion compared with EI. To perform this analysis, we ran Monte Carlo simulations over 35 different sets of 2 input distributions. To generate input distributions, we used an Inverse Wishart distribution to generate three covariance matrices corresponding with $C_{a\backslash c}$, $C_{b\backslash c}$, and C_c, where each matrix is a 2×2 matrix. A random sample from a Gaussian with the corresponding covariance is then drawn to generate $\mu_{a\backslash c}$, $\mu_{b\backslash c}$, and μ_c. These values are then used to generate the inputs (μ_a,C_a) and (μ_b,C_b) according to Equation (6). To evaluate the effect of how similar the input distributions are, we varied the degrees of freedom of the Inverse Wishart distribution used to generate the covariances over the values 2, 10, and 100 (leading to a total of 105 sets of input distributions.) As the degrees of freedom increased, the sampled covariances will approach the two-dimensional identity matrix. Note that with higher degrees of freedom, all three covariances will be more similar, making C_a and C_b more similar as well. Each set of input distributions was fused with the proposed technique (PC fusion) at α values of .05 and .01 and with EI fusion.

The effectiveness of the different fusion techniques were evaluated using two metrics: (1) the determinant of the output (fused) covariance and (2) the mean squared error (MSE) of the fused means. The goal of using the PC is to compute fused covariances that are less conservative than EI while still being conservative compared to the optimal fusion result. Therefore, we would expect the determinants of the fused covariances from the proposed technique to be smaller than the EI-derived covariances. The error used to compute MSE is defined as $\mu_{f,PC} - \mu_\phi$, where $\mu_{f,PC}$ is the mean of the fused distribution using PC and μ_ϕ is the optimal mean as defined in Equation (8). The MSE due to EI was computed similarly. The results of the Monte Carlo simulations are summarized in Table I. Note that:

- Using the PC leads to a significant reduction in the size of the fused covariance determinant, an indicator that it is less conservative.
- The MSE is also significantly smaller with PC fusion
- As the α value gets smaller, the PC approaches the EI technique, as would be expected.
- As the input covariances become more similar (the degrees of freedom increase), the advantage of using PC fusion increases.

TABLE I: Monte-carlo inverse wishart simulation result comparison between EI and PC for different degrees of freedom.

	PC, α =0.05	**PC, α =0.01**	**EI**
	df=2		
Avg. Determinant	0.7229	0.7418	0.7946
% smaller than EI	9.02%	6.64%	
MSE	1.927	1.9442	2.019
% smaller than EI	4.56%	3.70%	
	df=10		
Avg. Determinant	0.415	0.435	0.485
% smaller than EI	14.47%	10.28%	
MSE	0.574	0.588	0.640
% smaller than EI	10.32%	8.09%	
	df=100		
Avg. Determinant	0.537	0.565	0.651
% smaller than EI	17.62%	13.24%	
MSE	0.672	0.684	0.867
% smaller than EI	22.46%	21.15%	

Figure 4 also shows a box and whisker plot showing the percentage decrease in determinate using PC fusion at $\alpha = .01$ compared with EI fusion for the same monte-carlo simulations described previously. This chart demonstrates that (1) PC fusion is always as good as and often better than EI fusion (no negative percentages) and (2) as the input covariances become more similar, the advantage of PC fusion over EI fusion increases.

B. Decentralized Network Convergence

In addition to showing results for the 2-input fusion problem, we would also like to demonstrate the utility of PC fusion in a distributed network fusion scenario. Note that in a distributed network, the estimated covariances may

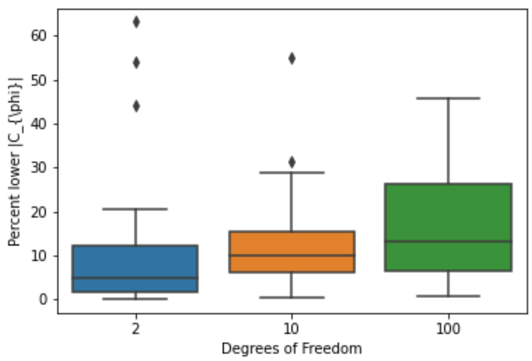


Fig. 4: Decrease in fused covariance determinate (y-axis) using PC ($\alpha = 0.01$) as opposed to EI fusion at different inverse wishart degrees of freedom (x-axis). Each degree of freedom had 35 runs of data.

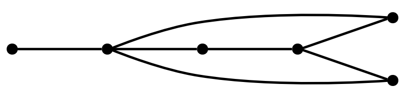


Fig. 5: Network used to demonstrate PC fusion in a distributed setting

very quickly converge, often without the means converging, leading to a possibly significant gain for PC fusion over previous techniques.

To demonstrate the performance of PC fusion in a distributed network, we simulated a network of six nodes with the topology shown in Figure 5. Each node started with an independent Gaussian distribution with a covariance sampled from an Inverse Wishart distribution with 30 degrees of freedom and the mean sampled from that covariance. For a single iteration, the EI or PC fusion was applied to each edge in the graph, replacing the estimate in that node with the fused (between both ends of the edge) distribution. We ran this procedure 25 times to ensure convergence of the network. In Figure 6, we show the resulting covariances from a centralized (true) fusion technique, EI, and PC. As shown, adding the probabilistic constraints leads to a fused covariance that is much less conservative than the traditional EI technique.

VI. Conclusion

This paper has described a new correlation-agnostic technique for data fusion that attempts to address the "double counting" phenomenon. This technique proposes a novel concept of leveraging not only the independent distribution's covariances to define "admissible" common covariances C_c, but also the means. We utilize a chi-squared test to evaluate whether a particular C_c also "probabilistically" admissible. This additional constraint to the optimization problem for C_c results in tighter covariance bounds, especially when the input covariances are similar. Fusion using this constraint is achieved using a constrained optimization framework.

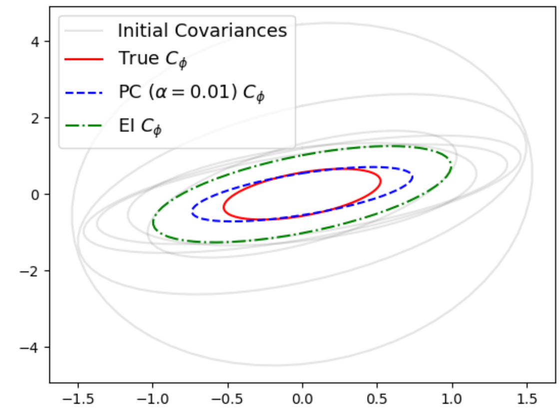


Fig. 6: Covariance results of decentralized data fusion using probabilistically constrained (PC) and elliposidal intersection (EI) approaches, with the input covariances and centralized fusion covariance also shown for comparison

References

[1] M. A. Bakr and S. Lee, "Distributed multisensor data fusion under unknown correlation and data inconsistency," *Sensors*, vol. 17, no. 11, p. 2472, 2017.

[2] R. Brooks, "Data fusion for a distributed ground-based sensing system," *Handbook of Multisensor Data Fusion, CRC Press, Boca Raton, FL*, 2001.

[3] M. E. Liggins, C.-Y. Chong, I. Kadar, M. G. Alford, V. Vannicola, and S. Thomopoulos, "Distributed fusion architectures and algorithms for target tracking," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 95–107, 1997.

[4] S. Grime and H. F. Durrant-Whyte, "Data fusion in decentralized sensor networks," *Control engineering practice*, vol. 2, no. 5, pp. 849–863, 1994.

[5] S. Julier and J. Uhlmann, "A non-divergent estimation algorithm in the presence of unknown correlations," *Proceedings of the 1997 American Control Conference (Cat. No.97CH36041)*, 1997.

[6] J. Sijs and M. Lazar, "State fusion with unknown correlation: Ellipsoidal intersection," *Automatica*, vol. 48, no. 8, p. 1874–1878, 2012.

[7] B. Noack, J. Sijs, and U. D. Hanebeck, "Inverse covariance intersection: New insights and properties," *2017 20th International Conference on Information Fusion (Fusion)*, 2017.

[8] N. R. Ahmed, W. Whitacre, S. Moon, and E. W. Frew, "Scalable decentralized partial state estimation with sensor uncertainties using factorized data fusion," *AIAA Infotech @ Aerospace*, 2016.

[9] H. Li, F. Nashashibi, and M. Yang, "Split covariance intersection filter: Theory and its application to vehicle localization," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, p. 1860–1871, 2013.

[10] N. Ahmed, J. Schoenberg, and M. Campbell, "Fast weighted exponential product rules for robust general multi-robot data fusion," in *Robotics: Science and Systems*, vol. 8, 2013, pp. 9–16.

[11] I. L. Manuel and A. N. Bishop, "Distributed monte carlo information fusion and distributed particle filtering," *IFAC Proceedings Volumes*, vol. 47, no. 3, p. 8681–8688, 2014.

[12] C. N. Taylor and A. N. Bishop, "Homogeneous functionals and bayesian data fusion with unknown correlation," *Information Fusion*, vol. 45, p. 179–189, 2019.

[13] N. R. Ahmed and M. Campbell, "Fast consistent chernoff fusion of gaussian mixtures for ad hoc sensor networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 12, p. 6739–6745, 2012.

[14] S. J. Julier, T. Bailey, and J. K. Uhlmann, "Using exponential mixture models for suboptimal distributed data fusion," in *2006 IEEE Nonlinear Statistical Signal Processing Workshop*. IEEE, 2006, pp. 160–163.