arXiv:1610.01045v1 [cs.SY] 4 Oct 2016

# A Game-Theoretic Approach to Robust Fusion and Kalman Filtering Under Unknown Correlations

Spyridon Leonardos[1] and Kostas Daniilidis[2]

*Abstract*— **This work addresses the problem of fusing two random vectors with unknown cross-correlations. We present a formulation and a numerical method for computing the optimal estimate in the minimax sense. We extend our formulation to linear measurement models that depend on two random vectors with unknown cross-correlations. As an application we consider the problem of decentralized state estimation for a group of agents. The proposed estimator takes cross-correlations into account while being less conservative than the widely used Covariance Intersection. We demonstrate the superiority of the proposed method compared to Covariance Intersection with numerical examples and simulations within the specific application of decentralized state estimation using relative position measurements.**

## I. INTRODUCTION

State estimation is one of the fundamentals problem in control theory and robotics. The most common state estimators are undoubtedly the Kalman filter [9], which is optimal in the minimum mean squared error for the case of linear systems, and its generalizations for nonlinear systems: the Extended Kalman Filter (EKF) [16] and the Uscented Kalman Filter (UKF) [6].

In multi-agent systems, the task of state estimation takes a collaborative form in the sense that it involves inter-agent measurements and constraints. Examples are cooperative localization in robotics [15] using relative pose measurements, camera network localization using epipolar constraints [17] and many more. On the one hand, a decentralized solution that scales with the number of agents is necessary. On the other hand, the state estimates become highly correlated as information flows through the network. Ignoring these correlations has grave consequences: estimates become optimistic and result in divergence of the estimator. This phenomenon is analogous to the rumor spreading in social networks.

Unknown correlations may be present in other scenarios as well. A popular simplification, that significantly reduces computations and enables the use of EKF-based estimators, is that noise sources are independent. For instance, a common assumption in vision-aided inertial navigation is independence of the image projection noises for each landmark [5], [12] although in reality, they are coupled with the motion of the camera sensor. In other cases, it might be impractical to store the entire covariance matrix due to storage limitations. For instance, in Simultaneous Localization and Mapping (SLAM) problems, there have been several approaches that decouple the sensor state estimate from the estimates of the landmark positions [8], [12] to increase the efficiency of the estimator and reduce the storage requirements.

The most popular algorithm for fusion under the presence of unknown correlations is the Covariance Intersection (CI) method which was introduced by Julier and Uhlmann [7]. In its simplest form, the Covariance Intersection algorithm is designed to fuse two random vectors whose correlation is not known by forming a convex combination of the two estimates in the information space. Covariance Intersection produces estimates that are provably consistent, in the sense that estimated error covariance is an upper bound of the true error covariance. However, it has been observed [1], [13], [18] that Covariance Intersection produces estimates that are too conservative which may decrease the accuracy and convergence speed of the overall estimator when used as a component of an online estimator.

One of the most prominent applications of the proposed fusion algorithm is distributed state estimation in an EKF-based framework. However, the problem of distributed state estimation is far from new. There have been numerous approaches for EKF-based distributed state estimation and EKF-based cooperative localization. Yet, some of them require that each agent maintains the state of the entire network [1], [15], which is impractical and does not scale with the number of agents, while others ignore correlations [11], [14] in order to simplify the estimation process or use Covariance Intersection and variations of it [3], [10] despite its slow convergence.

The contributions of this work are summarized as follows. First of all, we propose a method for fusion of two random vectors with unknown cross-correlations. The proposed approach is less conservative than the widely used Covariance Intersection (CI) while taking cross-correlations into account. Second of all, we extend our formulation for the case of a linear measurement model. Finally, we present numerical examples and simulations in a distributed state estimation scenario which demonstrate the validity and comparative performance of the proposed approach compared with the Covariance Intersection.

The paper is structured as follows: in Section II we include definitions of consistency and related notions and we introduce the problem at hand. The game-theoretic approach to fusing two random variables with unknown correlations is the topic of Section III which is generalized for arbitrary linear measurement models in Section III. In Section V we include details on the implemented numerical algorithm. Numerical examples and simulation results are presented in Sections VI and VII respectively.

[1,2]The authors are with the Department of Computer and Information Science, University of Pennsylvania Philadelphia, PA 19104, USA {spyridon,kostas}@seas.upenn.edu

## II. PROBLEM FORMALIZATION

In this section, we formalize the problem at hand. First, we need a precise definition of *consistency*.

*Definition 2.1 (Consistency [7]):* Let $z$ be a random vector with expectation $E[z] = \bar{z}$. An estimate $\widetilde{z}$ of $\bar{z}$ is another random vector. The associated error covariance is denoted $\widetilde{\Sigma}_{zz} \doteq \mathrm{Cov}\,(\widetilde{z} - \bar{z})$. The pair $(\widetilde{z}, \Sigma_{zz})$ is *consistent* if $E[\widetilde{z}] = \bar{z}$ and

$$\Sigma_{zz} \succeq \widetilde{\Sigma}_{zz} \tag{1}$$

***Problem Statement** 1 (Consistent fusion):* Given two consistent estimates $(\widetilde{x}, \Sigma_{xx})$, $(\widetilde{y}, \Sigma_{yy})$ of $\bar{z}$, where $\Sigma_{xx}, \Sigma_{yy}$ are known upper bounds on the true error covariances. The problem at hand consists of fusing the two consistent estimates $(\widetilde{x}, \Sigma_{xx})$, $(\widetilde{y}, \Sigma_{yy})$ in a single consistent estimate $(\widetilde{z}, \Sigma_{zz})$, where $\widetilde{z}$ is of the form

$$\widetilde{z} = W_x \widetilde{x} + W_y \widetilde{y} \tag{2}$$

with $W_x + W_y = I$ in order to preserve the mean.

The most widely used solution of the above problem is the Covariance Intersection algorithm [7]. Given upper bounds $\Sigma_{xx} \succeq \widetilde{\Sigma}_{xx}$, $\Sigma_{yy} \succeq \widetilde{\Sigma}_{yy}$ the Covariance Intersection equations read

$$\begin{aligned} \widetilde{z} &= \Sigma_{zz} \left\{ \omega \Sigma_{xx}^{-1} \widetilde{x} + (1-\omega) \Sigma_{yy}^{-1} \widetilde{y} \right\} \\ \Sigma_{zz}^{-1} &= \omega \Sigma_{xx}^{-1} + (1-\omega) \Sigma_{yy}^{-1} \end{aligned} \tag{3}$$

where $\omega \in [0, 1]$. It can be immediately seen that $\Sigma_{zz} \left\{ \omega \Sigma_{xx}^{-1} + (1-\omega) \Sigma_{yy}^{-1} \right\} = I$ which implies $E[\widetilde{z}] = \bar{z}$. Moreover, it is easy to check that $(\widetilde{z}, \Sigma_{zz})$ is consistent. The above can be easily generalized for the case of more than 2 random variables, for partial measurements and for the linear measurement model we consider in Section IV. Usually, $\omega$ is chosen such that either $\mathrm{tr}(\Sigma_{zz})$ or $\log\det(\Sigma_{zz}^{-1})$ is minimized.

Next, we introduce a notion related to consistency but with relaxed requirements. Let $S_+^n$ denote the positive semidefinite cone, that is the set of $n \times n$ positive semidefinite matrices. First, recall that a function $f : S_+^n \to \mathbb{R}$ is called $S_+^n$-*nondecreasing* [2] if

$$X \succeq Y \Rightarrow f(X) \geq f(Y) \tag{4}$$

for any $X, Y \in S_+^n$. An example of such a function is $f(X) = \mathrm{tr}(X)$. Now, we are ready to introduce the notion of *consistency with respect to a $S_+^n$-nondecreasing function*.

*Definition 2.2 ($f$-Consistency):* Let $f : S_+^n \to \mathbb{R}$ be a nondecreasing function (with respect to $S_+^n$) satisfying $f(\mathbf{0}) = 0$. Let $z$ be a random vector with expectation $E[z] = \bar{z}$ and $\widetilde{z}$ be an estimate of $\bar{z}$ with associated error covariance $\widetilde{\Sigma}_{zz}$. The pair $(\widetilde{z}, \Sigma_{zz})$ is *$f$-consistent* if $E[\widetilde{z}] = \bar{z}$ and

$$f(\Sigma_{zz}) \geq f(\widetilde{\Sigma}_{zz}) \tag{5}$$

*Remark 1:* Observe that consistency implies $f$-consistency. However, the converse in not necessarily true.

***Problem Statement** 2 (Trace-consistent fusion):* Given two consistent estimates $(\widetilde{x}, \Sigma_{xx})$, $(\widetilde{y}, \Sigma_{yy})$ of $\bar{z}$, where $\Sigma_{xx}, \Sigma_{yy}$ are known upper bounds on the true error variances. The problem at hand consists of fusing the two consistent estimates $(\widetilde{x}, \Sigma_{xx})$, $(\widetilde{y}, \Sigma_{yy})$ in a single trace-consistent estimate $(\widetilde{z}, \Sigma_{zz})$, where $\widetilde{z}$ is a linear combination of $x$ and $y$ and

$$\mathrm{tr}(\Sigma_{zz}) \geq \mathrm{tr}(\widetilde{\Sigma}_{zz}) \tag{6}$$

Next, we introduce a game-theoretic formulation for the problem of trace-consistent fusion. Relaxing the consistency constraint to the trace-consistency constraint enables us to estimate the weighting matrices $W_x, W_y$ according to some optimality criterion, which is none other than the minimax of the trace of the covariance matrix.

*Remark 2:* No assumptions on the distribution of the estimates $\tilde{x}$ and $\tilde{y}$ have been made so far.

## III. ROBUST FUSION

The goal of this section is the derivation of our minimax approach. First, we need some basic notions from game theory. A zero-sum, two-player game on $\mathbb{R}^m \times \mathbb{R}^n$ is defined by a pay-off function $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$. Intuitively, the first player makes a move $u \in \mathbb{R}^m$ then, the second player makes a move $v \in \mathbb{R}^m$ and receives payment from the first player equal to $f(u, v)$. The goal of the first player is to minimize its payment and the goal of the second player is to maximize the received payment. The game is *convex-concave* if the pay-off function $f(u, v)$ is convex in $u$ for fixed $v$ and concave in $v$ for fixed $u$. For a review minimax and convex-concave games in the context of convex optimization, we refer the reader to [4].

Let $z$ be a random vector with expectation $E[z] = \bar{z}$. Assume we have two estimates $(\widetilde{x}, \Sigma_{xx})$, $(\widetilde{y}, \Sigma_{yy})$ of $\bar{z}$ where $\Sigma_{xx}, \Sigma_{yy}$ are approximations to the true error covariances $\widetilde{\Sigma}_{xx}$, $\widetilde{\Sigma}_{yy}$. Based on the discussion of Section II, the fused estimate is of the form

$$\widetilde{z} = (I - K)\widetilde{x} + K\widetilde{y} \tag{7}$$

and the associated error covariance $\widetilde{\Sigma}_{zz} \doteq \mathrm{Cov}\,(\widetilde{z} - \bar{z})$ is given by

$$\widetilde{\Sigma}_{zz} = \begin{bmatrix} I - K & K \end{bmatrix} \begin{bmatrix} \widetilde{\Sigma}_{xx} & \widetilde{\Sigma}_{xy} \\ \widetilde{\Sigma}_{xy}^T & \widetilde{\Sigma}_{yy} \end{bmatrix} \begin{bmatrix} I - K^T \\ K^T \end{bmatrix} \tag{8}$$

However, $\widetilde{\Sigma}_{xx}$, $\widetilde{\Sigma}_{yy}$ are not known. Therefore, we define

$$\Sigma_{zz} \doteq \begin{bmatrix} I - K & K \end{bmatrix} \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy}^T & \Sigma_{yy} \end{bmatrix} \begin{bmatrix} I - K^T \\ K^T \end{bmatrix} \tag{9}$$

where we have the following Linear Matrix Inequality (LMI) constraint on $\Sigma_{xy}$

$$\begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy}^T & \Sigma_{yy} \end{bmatrix} \succeq 0 \tag{10}$$

*Remark 3:* It can be seen that $\mathrm{tr}(\Sigma_{zz})$ is convex in $K$ for a fixed $\Sigma_{xy}$ satisfying (10). Therefore, the supremum of $\mathrm{tr}(\Sigma_{zz})$ over all $\Sigma_{xy}$ satisfying (10) is a convex function of $K$. Moreover, for a fixed $K$, $\operatorname{tr}(\Sigma_{zz})$ is linear, and thus concave as well, in $\Sigma_{xy}$ with a convex domain defined by (10). It follows that $\operatorname{tr}(\Sigma_{zz})$ is a convex-concave function in $(K, \Sigma_{xy})$.

As anticipated, we formulate the problem of finding the weighting matrix $K$ as a zero-sum, two-player convex-concave game: the first player chooses $K$ to minimize $\operatorname{tr}(\Sigma_{zz})$ whereas the second player chooses $\Sigma_{xy}$ to maximize $\operatorname{tr}(\Sigma_{zz})$. More specifically, let $(K^\star, \Sigma_{xy}^\star)$ be the solution to the following minimax optimization problem

$$\begin{aligned} &\underset{K}{\text{minimize}} \quad \sup_{\Sigma_{xy}} \quad \operatorname{tr}(\Sigma_{zz}) \\ &\text{subject to} \quad \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy}^T & \Sigma_{yy} \end{bmatrix} \succeq 0 \end{aligned} \tag{11}$$

Then, the fused estimated and the associated error covariance are given by

$$\begin{aligned} \widetilde{z} &= (I - K^\star)\widetilde{x} + K^\star \widetilde{y} \\ \Sigma_{zz}^\star &= \begin{bmatrix} I - K^\star & K^\star \end{bmatrix} \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy}^\star \\ \Sigma_{xy}^{\star T} & \Sigma_{yy} \end{bmatrix} \begin{bmatrix} I - K^{\star T} \\ K^{\star T} \end{bmatrix} \end{aligned} \tag{12}$$

Naturally, we have the following lemma.

*Lemma 3.1:* If $(\widetilde{x}, \Sigma_{xx})$ and $(\widetilde{y}, \Sigma_{yy})$ are consistent, then the pair $(\widetilde{z}, \Sigma_{zz}^\star)$ given by (12) is trace-consistent.

A proof of lemma 3.1 is presented in AppendixVIII-A.

The problem of numerically solving problem (11) is the topic of subsequent sections. The case under consideration in this section can be viewed as a special case of the next section.

## IV. ROBUST LINEAR UPDATE

In this section, we explore a more general setting. We assume we have two random vectors $x$, $y$ with expectations $E[x] = \overline{x}$ and $E[y] = \overline{y}$. We have some estimates $\widetilde{x}$ and $\widetilde{y}$ of $\overline{x}$ and $\overline{y}$ respectively with associated error covariances $\widetilde{\Sigma}_{xx}$ and $\widetilde{\Sigma}_{yy}$. As before, we assume that the true error covariances are only approximately known. Let $\Sigma_{xx}$ and $\Sigma_{yy}$ denote these approximate values. We assume we have a linear measurement model of the form

$$z = C\overline{x} + D\overline{y} + \eta \tag{13}$$

where $\eta$ is a zero-mean noise process with covariance $\Sigma_\eta$. We assume that the measurement noise process $\eta$ is independent to the estimates $\widetilde{x}$ and $\widetilde{y}$. As in the classic Kalman filter derivation, we are looking for an update step of the form

$$\widetilde{x}^+ = \widetilde{x} + K(z - \widetilde{z}) \tag{14}$$

where $\widetilde{z} \doteq C\widetilde{x} + D\widetilde{y}$. The error of the update is given by

$$\widetilde{x}^+ - \overline{x} = (I - KC)(\widetilde{x} - \overline{x}) - KD(\widetilde{y} - \overline{y}) + K\eta \tag{15}$$

and the associated error covariance is defined as $\widetilde{\Sigma}_{xx}^+ \doteq \operatorname{Cov}(\widetilde{x}^+ - \overline{x})$ and is given by

$$\begin{aligned} \widetilde{\Sigma}_{xx}^+ = &\begin{bmatrix} I - KC & -KD \end{bmatrix} \begin{bmatrix} \widetilde{\Sigma}_{xx} & \widetilde{\Sigma}_{xy} \\ \widetilde{\Sigma}_{xy}^T & \widetilde{\Sigma}_{yy} \end{bmatrix} \begin{bmatrix} I - C^T K^T \\ -D^T K^T \end{bmatrix} \\ &+ K\Sigma_\eta K^T \end{aligned} \tag{16}$$

However, the true error covariances $\widetilde{\Sigma}_{xx}$ and $\widetilde{\Sigma}_{yy}$ are not known. Therefore, we define

$$\begin{aligned} \Sigma_{xx}^+ \doteq &\begin{bmatrix} I - KC & -KD \end{bmatrix} \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy}^T & \Sigma_{yy} \end{bmatrix} \begin{bmatrix} I - C^T K^T \\ -D^T K^T \end{bmatrix} \\ &+ K\Sigma_\eta K^T \end{aligned} \tag{17}$$

where $\Sigma_{xy}$ should satisfy (10) in order to be a valid cross-correlation. To alleviate notation, let $X = K^T$ and define

$$f(X, \Sigma_{xy}) \doteq \operatorname{tr}(\Sigma_{xx}^+) \tag{18}$$

By rewriting (10) using Schur complement, the minimax formulation is written as follows

$$\begin{aligned} &\underset{X}{\text{minimize}} \quad \sup_{Q} \quad f(X, Q) \\ &\text{subject to} \quad \Sigma_{yy}^{-1/2} Q^T \Sigma_{xx}^{-1} Q \Sigma_{yy}^{-1/2} - I \preceq 0 \end{aligned} \tag{19}$$

Let $(X^\star, Q^\star)$ be the optimal solution of problem (19) and let $(K^\star, \Sigma_{xy}^\star) = (X^{\star T}, Q^\star)$. Then,

$$\begin{aligned} \widetilde{x}^+ &= (I - K^\star C)\widetilde{x} - K^\star D\widetilde{y} \\ \Sigma_{xx}^{+\star} &= \begin{bmatrix} I - K^\star C & -K^\star D \end{bmatrix} \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy}^\star \\ \Sigma_{xy}^{\star T} & \Sigma_{yy} \end{bmatrix} \begin{bmatrix} I - C^T K^{\star T} \\ -D^T K^{\star T} \end{bmatrix} \\ &\quad + K\Sigma_\eta K^T \end{aligned} \tag{20}$$

Naturally, we have the following lemma.

*Lemma 4.1:* If $(\widetilde{x}, \Sigma_{xx})$ and $(\widetilde{y}, \Sigma_{yy})$ are consistent, then the pair $(\widetilde{x}^+, \Sigma_{xx}^{+\star})$ given by (20) is trace-consistent

The proof of lemma 4.1 is exactly analogous to the proof of lemma 3.1 presented in AppendixVIII-A.

*Remark 4:* When $C = I$, $D = -I$, $\Sigma_\eta = 0$, we recover the case of the simple fusion of two random vectors.

## V. INTERIOR POINT METHODS FOR CONVEX-CONCAVE GAMES

In this section, we describe the numerical method we use to solve Problem (19). First, we will look at the simpler case of an unconstrained convex-concave game with pay-off function $f(u, v)$. A point $(u^\star, v^\star)$ is a *saddle* point for an unconstrained convex-concave game with pay-off function $f(u, v)$ if

$$f(u^\star, v) \leq f(u^\star, v^\star) \leq f(u, v^\star) \tag{21}$$

and the optimality conditions for differentiable convex-concave pay-off function are

$$\nabla_u f(u^\star, v^\star) = 0, \qquad \nabla_v f(u^\star, v^\star) = 0 \tag{22}$$

We use the infeasible start Newton method [2], outlined in Algorithm 1, to find the optimal solution of the unconstrained problem:

$$\underset{u}{\text{minimize}}\ \underset{u}{\text{maximize}}\ f(u, v) \tag{23}$$

Intuitively, at each step the directions $\Delta u_{nt}, \Delta v_{nt}$ are the solutions of the first order approximation

$$0 = r(u + \Delta u_{nt}, v + \Delta v_{nt}) \approx r(u, v) + Dr(u, v)[\Delta u_{nt}, \Delta v_{nt}] \tag{24}$$

where $r(u,v) = [\nabla_u f(u,v)^T, \nabla_v f(u,v)^T]^T$. Then, a backtracking line search is performed on the norm of the residual along the previously computed directions.

**Algorithm 1** Infeasible start Newton method.

**given:** starting points $u, v \in \mathbf{dom} f$,
tolerance $\epsilon > 0$, $\alpha \in (0, 1/2)$, $\beta \in (0, 1)$.
**Repeat**
1. $r(u,v) = [\nabla_u f(u,v)^T, \nabla_v f(u,v)^T]^T$
2. Compute Newton steps by solving

$$Dr(u,v)[\Delta u_{nt}, \Delta v_{nt}] = -r(u,v)$$

3. Backtracking line search on $\|r\|_2$.
$t = 1$.
$u_t = u + t\Delta u_{nt}$, $v_t = v + t\Delta v_{nt}$.
**While** $\|r(u_t, v_t)\|_2 > (1 - \alpha t)\|r(u,v)\|_2$
$t = \beta t$.
$u_t = u + t\Delta u_{nt}$, $v_t = v + t\Delta v_{nt}$.
**EndWhile**
4. Update: $u = u + t\Delta u_{nt}$, $v = v + t\Delta v_{nt}$.

**until** $\|r(u,v)\|_2 \leq \epsilon$

However, the problem at hand is slightly more complicated since it involves a linear matrix inequality. Therefore, we use the barrier method [2]. Intuitively, a sequence of unconstrained minimization problems is solved, using the last point iteration is the starting point for the next iteration. Define for $t > 0$, the cost function $f_t(X, Q)$ by

$$f_t(X,Q) = tf(X,Q) + \log\det(-f_1(Q)) \tag{25}$$

where $f(X,Q)$ as defined in (18) and

$$f_1(Q) = \Sigma_{yy}^{-1/2} Q^T \Sigma_{xx}^{-1} Q \Sigma_{yy}^{-1/2} - I \tag{26}$$

Intuitively, $\frac{1}{t} f_t$ approaches $f$ as $t \to \infty$. Note that $f_t(X,Q)$ is still convex-concave for $t > 0$. The optimality conditions for a fixed $t > 0$ are given by

$$\nabla_X f_t(X^\star, Q^\star) = 0, \quad \nabla_Q f_t(X^\star, Q^\star) = 0 \tag{27}$$

where explicit expressions for $\nabla_X f_t$ and $\nabla_Q f_t$ are presented in AppendixVIII-B along with the linear equations for computing $\Delta X_{nt}, \Delta Q_{nt}$ .

Finally, the structure of the problem allows us to easily identify a strictly feasible initial point $(X_0, Q_0)$ where $Q_0 = 0$ and $X_0$ is given by

$$(C\Sigma_{xx}C^T + D\Sigma_{yy}D^T + \Sigma_\eta)X_0 = C\Sigma_{xx} \tag{28}$$

For details on the convergence of the infeasible start Newton method and the barrier method for convex-concave games, we refer the reader to [2], [4].

*Remark 5:* Notation: $Df(x)[h]$ denotes the *(Fréchet) derivative* or *differential* of $f$ at $x$ along $h$. Similarly, $Df(x,y)[h_x, h_y]$ denotes the *differential* of $f$ at $(x,y)$ along $(h_x, h_y)$.

## VI. NUMERICAL EXAMPLES

In this section, we present two numerical examples which shed light on the differences between the Covariance Intersection (CI) and the proposed Robust Fusion (RF) approaches. First, consider the example of fusing two random variables with means $\tilde{x} = \tilde{y} = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$ and covariances

$$\Sigma_{xx} = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}, \ \Sigma_{yy} = \begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix} \tag{29}$$

Let $(\tilde{z}_{CI}, \Sigma_{CI})$ and $(\tilde{z}_{RF}, \Sigma_{RF})$ be the fused estimates and the corresponding error covariances obtained from Covariance Intersection and Robust Fusion. We have that $\tilde{z}_{CI} = \tilde{z}_{RF} = [0\ 0]^T$ and

$$\Sigma_{CI} = \begin{bmatrix} 3.79 & 0 \\ 0 & 5.79 \end{bmatrix}, \ \Sigma_{RF} = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}, \tag{30}$$

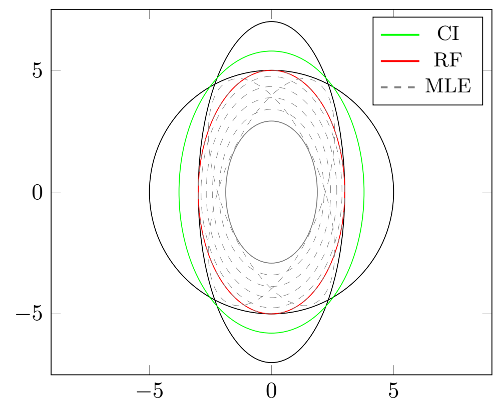


Fig. 1. Confidence ellipses: given a covariance matrix $\Sigma$ we draw the set $\{x : x^T \Sigma^{-1} x = 1\}$. Initial confidence ellipse (black), Maximum likelihood Estimate (MLE) confidence ellipse (gray dashed) for various values of correlation, CI confidence ellipse (green) and RF confidence ellipse (red). The confidence ellipses obtained from MLE lie in the intersection of the two ellipsoids $\{x : x^T \Sigma_{xx}^{-1} x \leq 1\}$ and $\{x : x^T \Sigma_{yy}^{-1} x \leq 1\}$. Since the proposed approach is equivalent to MLE for some worst-case correlation, the RF confidence ellipse lies in the intersection of the two ellipsoids as well. When correlation increases, the trace of the covariance of MLE approaches the trace of $\Sigma_{RF}$. CI is not maximum likelihood for any value of correlation but produces a guaranteed upper bound on the true error covariance.

In the second example, we consider the case of partial measurements. More specifically, using notation of Section IV, let

$$\tilde{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \ \Sigma_{xx} = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} \tag{31}$$

and $C = \begin{bmatrix} 1 & 0 \end{bmatrix}$, $z = \tilde{z} = 0$, $\Sigma_{yy} = 1$, $D = 1$ and $\Sigma_\eta = 0$. Both Covariance Intersection and Robust Fusion yield $\tilde{z}^+ = 0$ but

$$\Sigma_{CI} = \begin{bmatrix} 3 & 0 \\ 0 & 6 \end{bmatrix}, \ \Sigma_{RF} = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}, \tag{32}$$

Observe that despite we have a measurement of only the first coordinate, the error variance of the second coordinate increased! The reason for this phenomenon is that the CI updates the current estimate and the associated error covariance along a predefined direction only. Although $\mathrm{tr}(\Sigma_{CI}) < \mathrm{tr}(\Sigma_{xx})$, the bound on the true error covariance estimated by Covariance Intersection is very conservative.

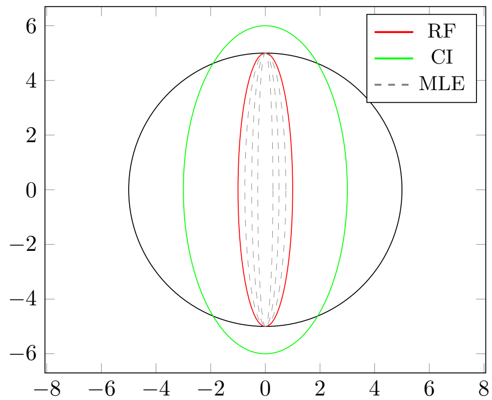


Fig. 2. Illustration of the second numerical example. Initial confidence ellipse (black), Maximum likelihood Estimate (MLE) confidence ellipse (gray dashed) for various values of correlation, CI confidence ellipse (green) and RF confidence ellipse (red).

## VII. SIMULATIONS

Finally, we consider an application in distributed state estimation using relative position measurements. We experiment with a group of $n = 4$ agents on the plane with a communication network topology as depicted in Fig. 3. If there is an edge from $i$ to $j$, then agent $i$ transmits its current state estimate and the corresponding error covariance estimate to agent $j$ which upon receipt, takes a measurement of the relative position and updates its own state estimate and associated error covariance estimate. All agents have identical dynamics described by

$$\begin{bmatrix} x_i(t+1) \\ v_i(t+1) \end{bmatrix} = \begin{bmatrix} I & I \\ 0 & I \end{bmatrix} \begin{bmatrix} x_i(t) \\ v_i(t) \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix} w_i(t) \tag{33}$$

where $x_i(t) \in \mathbb{R}^2$ and $v_i(t) \in \mathbb{R}^2$ denote respectively the position and velocity of agent $i$ at time instance $t$ and $w_i(t) \sim \mathcal{N}(0, Q_i(t))$ is the noise process. If $\mathbf{x}_i(t) \doteq \begin{bmatrix} x_i(t)^T & v_i(t)^T \end{bmatrix}^T$, then let $A, B$ such that

$$\mathbf{x}_i(t+1) = A\mathbf{x}_i(t) + Bw_i(t) \tag{34}$$

Only agent 1 is equipped with global position system (GPS), that is we have a measurement of the form

$$y_1(t) = x_1(t) + \eta_1(t) \tag{35}$$

where $\eta_1(t) \sim \mathcal{N}(0, R_1)$. Agent 1 performs a standard Kalman Filter update step after a GPS measurement. For each edge $(i, j)$ we have a pairwise measurement of the form

$$y_{ij}(t) = x_j(t) - x_i(t) + \eta_{ij}(t) \tag{36}$$

where $\eta_{ij}(t) \sim \mathcal{N}(0, R_{ij})$. Updates of the state estimates can be performed by either ignoring cross-correlations (Naive Fusion) or by one of Covariance Intersection or the proposed Robust Fusion.

Each agent maintains only its one state and communicates its to each neighbors at each time instance. The individual prediction step is the same as the Kalman Filter (KF) prediction step, that is

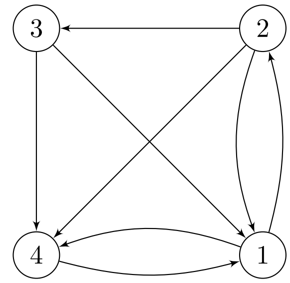


Fig. 3. Network topology.

$$\widetilde{\mathbf{x}}_i(t+1|t) = A\widetilde{\mathbf{x}}_i(t|t) \tag{37}$$

$$\Sigma_i(t+1|t) = A\Sigma_i(t+1|t)A^T + BQ_i(t)B^T \tag{38}$$

where $\widetilde{\mathbf{x}}_i(t+1|t)$ denotes the estimate of agent $i$ for its state at time $t+1$ having received measurements up to time $t$ and $\Sigma_i(t+1|t)$ is the associated error covariance.

We evaluate four estimator, three decentralized and one centralized: Naive Fusion (NF) which ignores correlations, Robust Fusion (RF), Covariance Intersection (CI) and Centralized Kalman Filter (CKF). The Centralized Kalman Filter (CKF) is simply a standard Kalman Filter containing all agent states. It serves as a measure of how close the decentralized estimators are to the optimal centralized estimator. Results can be seen in Figure 4 and Table I. We used the following values for the noise parameters: $Q_i = 10^{-6}I_2$ for all agents, $R_1 = I_2$ and $R_{ij} = 10^{-2}I_2$ for all pairwise measurements. The Robust Fusion based estimator significantly outperforms the Covariance Intersection based estimator which produces particularly noisy velocity estimates.

TABLE I

POSITION ERRORS

| Agent # | CKF | RF | CI |
|---|---|---|---|
| 1 | $0.174 \pm 0.107$ m | $0.222 \pm 0.091$ m | $0.230 \pm 0.093$ m |
| 2 | $0.166 \pm 0.098$ m | $0.248 \pm 0.108$ m | $0.286 \pm 0.145$ m |
| 3 | $0.170 \pm 0.099$ m | $0.248 \pm 0.114$ m | $0.343 \pm 0.187$ m |
| 4 | $0.161 \pm 0.085$ m | $0.238 \pm 0.090$ m | $0.285 \pm 0.128$ m |

## VIII. APPENDIX

### A. *Proof of lemma 3.1*

First, it is easy to see that if $E[\widetilde{x}] = E[\widetilde{y}] = \bar{z}$ then $E[\widetilde{z}] = \bar{z}$. Now, one has to show that if $\Sigma_{xx} \succeq \widetilde{\Sigma}_{xx}$ and $\Sigma_{yy} \succeq \widetilde{\Sigma}_{yy}$ then $\mathrm{tr}(\Sigma_{zz}^\star) \geq \mathrm{tr}(\widetilde{\Sigma}_{zz})$. We have that $\Sigma_{zz}^\star - \widetilde{\Sigma}_{zz}$ is equal to

$$\begin{aligned} &(I - K^\star)(\Sigma_{xx} - \widetilde{\Sigma}_{xx})(I - K^{\star T}) + K^\star(\Sigma_{yy} - \widetilde{\Sigma}_{yy})K^{\star T} \\ &+ K^\star(\Sigma_{xy}^{\star T} - \widetilde{\Sigma}_{xy}^T)(I - K^{\star T}) + (I - K^\star)(\Sigma_{xy}^\star - \widetilde{\Sigma}_{xy})K^{\star T} \\ &\succeq K^\star(\Sigma_{xy}^{\star T} - \widetilde{\Sigma}_{xy}^T)(I - K^{\star T}) + (I - K^\star)(\Sigma_{xy}^\star - \widetilde{\Sigma}_{xy})K^{\star T} \end{aligned}$$

since $\Sigma_{xx} \succeq \widetilde{\Sigma}_{xx}$ and $\Sigma_{yy} \succeq \widetilde{\Sigma}_{yy}$. Since trace is $S_+^n$-nondecreasing, we get

$$\mathrm{tr}(\Sigma_{zz}^\star - \widetilde{\Sigma}_{zz}) \geq 2\,\mathrm{tr}\left(K^T(I - K)(\Sigma_{xy}^\star - \widetilde{\Sigma}_{xy})\right) \geq 0$$

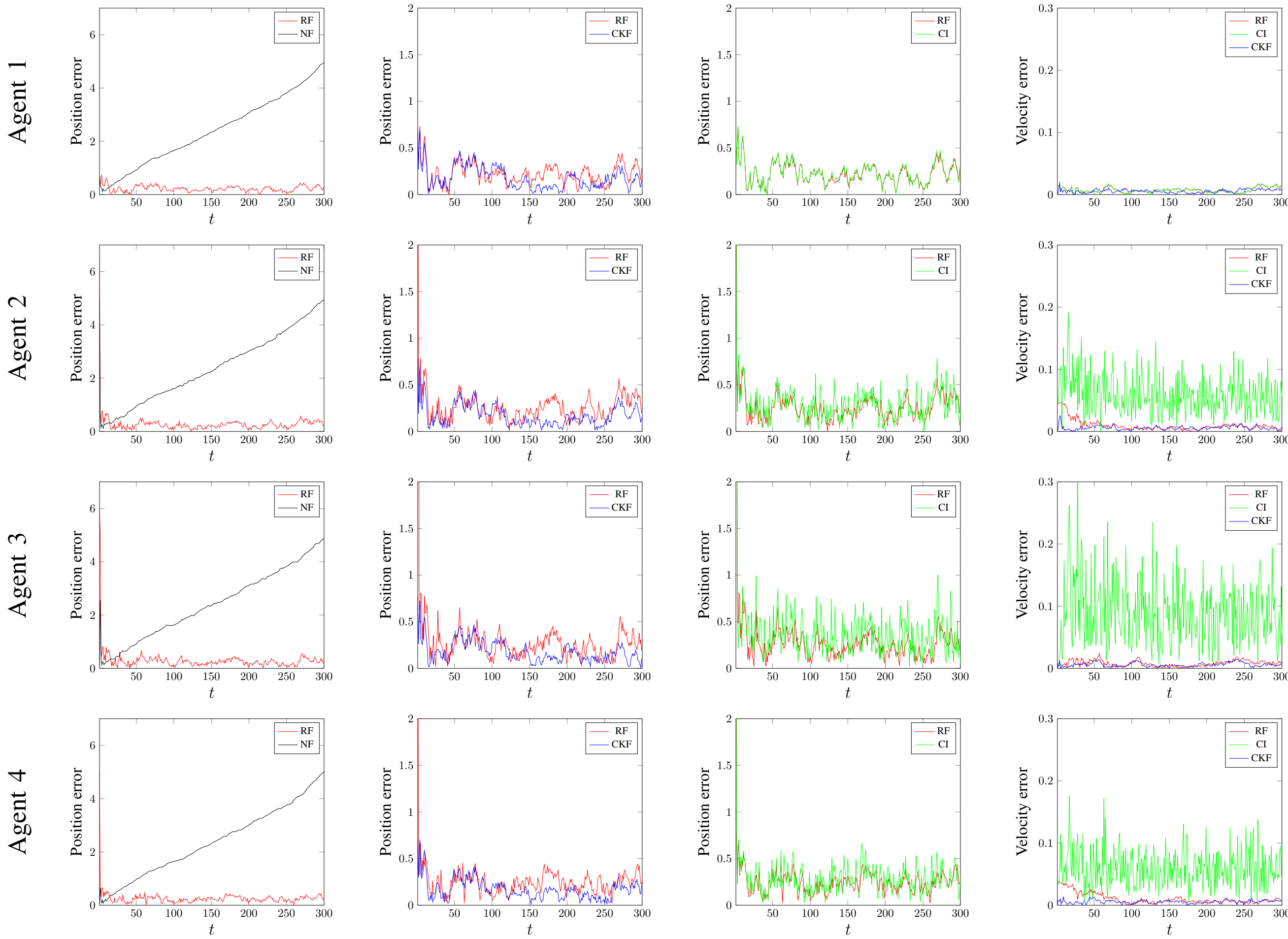


Fig. 4. Comparison of the four methods. The Naive Fusion estimator quickly diverges, whereas the Robust Fusion and Covariance Intersection do not. Clearly the Robust Fusion estimator is more accurate than the Covariance Intersection and produces less noisy estimates due to its less conservative nature. The steady state error covariance estimate of the Covariance Intersection is much larger than the actual error covariance making the estimator more susceptible to measurement noise.

since $\operatorname{tr}\left(K^{\star T}(I-K^{\star})\Sigma_{xy}^{\star}\right) \geq \operatorname{tr}\left(K^{\star T}(I-K^{\star})\Sigma_{xy}\right)$ for every $\Sigma_{xy}$ satisfying (10) due to optimality of $\Sigma_{xy}^{\star}$. Verifying that $\widetilde{\Sigma}_{xy}$ satisfies (10) is straightforward. ■

### *B. Formulas for computing the Newton steps*

First of all, the differential of $f_1(Q)$ at the direction of $\Delta Q$ is given by

$$Df_1(Q)[\Delta Q] = \Sigma_{yy}^{-1/2}\left(\Delta Q^T \Sigma_{xx}^{-1} Q + Q^T \Sigma_{xx}^{-1}\Delta Q\right)\Sigma_{yy}^{-1/2} \tag{39}$$

For small $\Delta X$, we have the first order approximation [2]:

$$\log\det(X+\Delta X) \approx \log\det(X) + \operatorname{tr}(X^{-1}\Delta X) \tag{40}$$

and thus, using the chain rule, we obtain

$$\nabla_Q \log\det(-f_1(Q)) = 2\Sigma_{xx}^{-1} Q \Sigma_{yy}^{-1/2} f_1(Q)^{-1}\Sigma_{yy}^{-1/2} \tag{41}$$

Moreover, we have

$$\begin{aligned}\nabla_X f(X,Q) =& 2\left(\begin{bmatrix} C & D\end{bmatrix}\begin{bmatrix}\Sigma_{xx} & Q \\ Q^T & \Sigma_{yy}\end{bmatrix}\begin{bmatrix}C^T \\ D^T\end{bmatrix} + \Sigma_\eta\right)X \\ & -2(C\Sigma_{xx} + DQ^T)\end{aligned} \tag{42}$$

and

$$\nabla_Q f(X,Q) = 2(C^T X X^T D - X^T D) \tag{43}$$

Let $g_1(Q) \doteq \Sigma_{yy}^{-1/2} f_1(Q)^{-1}\Sigma_{yy}^{-1/2}$. Using

$$(X+\Delta X)^{-1} \approx X^{-1} - X^{-1}\Delta X X^{-1} \tag{44}$$

for small $\Delta X$ and the chain rule, we obtain the following system of linear equations for $(\Delta X_{nt}, \Delta Q_{nt})$:

$$\begin{aligned}&2t\left(\begin{bmatrix} C & D\end{bmatrix}\begin{bmatrix}\Sigma_{xx} & Q \\ Q^T & \Sigma_{yy}\end{bmatrix}\begin{bmatrix}C^T \\ D^T\end{bmatrix} + \Sigma_\eta\right)\Delta X_{nt} \\ &+2t(C\Delta Q_{nt} D^T X - D\Delta Q_{nt}^T(I - C^T X)) = -\nabla_X f_t(X,Q)\end{aligned} \tag{45}$$

and

$$\begin{aligned}&2t(C^T \Delta X_{nt} X^T D - (I - C^T X)\Delta X_{nt}^T D) \\ &-2\Sigma_{xx}^{-1} Q g_1(Q)\left(\Delta Q_{nt}^T \Sigma_{xx}^{-1} Q + Q^T \Sigma_{xx}^{-1}\Delta Q_{nt}\right) g_1(Q) \\ &+2\Sigma_{xx}^{-1}\Delta Q_{nt} g_1(Q) = -\nabla_Q f_t(X,Q)\end{aligned} \tag{46}$$

# References


[1] Pablo O Arambel, Constantino Rago, and Raman K Mehra. Covariance intersection algorithm for distributed spacecraft state estimation. In *American Control Conference, 2001. Proceedings of the 2001*, volume 6, pages 4398–4403. IEEE, 2001.

[2] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[3] Luis C Carrillo-Arce, Esha D Nerurkar, José L Gordillo, and Stergios I Roumeliotis. Decentralized multi-robot cooperative localization using covariance intersection. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1412–1417. IEEE, 2013.

[4] A Ghosh and S Boyd. Minimax and convex-concave game. *lecture notes for course EE392:'Optimization Projects" Stanford Univ., Stanford, CA*, 2003.

[5] J. Hernandez, K. Tsotsos, and S. Soatto. Observability, identifiability and sensitivity of vision-aided inertial navigation. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 2319–2325, May 2015.

[6] Simon J Julier and Jeffrey K Uhlmann. New extension of the kalman filter to nonlinear systems. In *AeroSense'97*, pages 182–193. International Society for Optics and Photonics, 1997.

[7] Simon J Julier and Jeffrey K Uhlmann. A non-divergent estimation algorithm in the presence of unknown correlations. In *In Proceedings of the American Control Conference*. Citeseer, 1997.

[8] Simon J Julier and Jeffrey K Uhlmann. Using covariance intersection for slam. *Robotics and Autonomous Systems*, 55(1):3–20, 2007.

[9] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45, 1960.

[10] Hao Li and Fawzi Nashashibi. Cooperative multi-vehicle localization using split covariance intersection filter. *IEEE Intelligent transportation systems magazine*, 5(2):33–44, 2013.

[11] Agostino Martinelli. Improving the precision on multi robot localization by using a series of filters hierarchically distributed. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1053–1058. IEEE, 2007.

[12] Anastasios I Mourikis and Stergios I Roumeliotis. A multi-state constraint kalman filter for vision-aided inertial navigation. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pages 3565–3572. IEEE, 2007.

[13] Esha D Nerurkar and Stergios I Roumeliotis. Power-slam: A linear-complexity, consistent algorithm for slam. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 636–643. IEEE, 2007.

[14] Stefano Panzieri, Federica Pascucci, and Roberto Setola. Multirobot localisation using interlaced extended kalman filter. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2816–2821. IEEE, 2006.

[15] Stergios I Roumeliotis and George A Bekey. Distributed multi-robot localization. *IEEE Transactions on Robotics and Automation*, 18(5):781–795, 2002.

[16] Harold Wayne Sorenson. *Kalman filtering: theory and application*. IEEE, 1985.

[17] Roberto Tron and Rene Vidal. Distributed 3-d localization of camera sensor networks from 2-d image measurements. *Automatic Control, IEEE Transactions on*, 59(12):3325–3340, 2014.

[18] Xun Xu and Shahriar Negahdaripour. Application of extended covariance intersection principle for mosaic-based optical positioning and navigation of underwater vehicles. In *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, volume 3, pages 2759–2766. IEEE, 2001.