Communication Efficient Decentralized Track Fusion Using Selective Information Extraction

Robin Forsling, Zoran Sjanic, Fredrik Gustafsson and Gustaf Hendeby

DOI: https://doi.org/10.23919/FUSION45008.2020.9190575

Communication Efficient Decentralized Track Fusion Using Selective Information Extraction

Robin Forsling*†, Zoran Sjanic*†, Fredrik Gustafsson*, and Gustaf Hendeby*
* Dept. of Electrical Engineering, Linköping University, Linköping, Sweden
e-mail: {firstname.lastname}@liu.se
† Saab AB, Linköping, Sweden
e-mail: {firstname.lastname}@saabgroup.com

***Abstract*—We consider a decentralized sensor network of multiple nodes with limited communication capability where the cross-correlations between local estimates are unknown. To reduce the bandwidth the individual nodes determine which subset of local information is the most valuable from a global perspective. Three *information selection methods* (ISM) are derived. The proposed ISM require no other information than the communicated estimates. The simulation evaluation shows that by using the proposed ISM it is possible to determine which subset of local information is globally most valuable such that both reduced bandwidth and high performance are achieved.**

***Index Terms*—Cooperative target tracking, information extraction, track fusion, communication efficiency, covariance intersection, consistency.**

I. Introduction

In target tracking one or multiple sensors, *e.g.*, radar, sonar, and cameras, are used to estimate the instant state of a dynamic target. When utilizing multiple sensor nodes complementary information can be extracted, resulting in synergy effects through the multi-sensor fusion. Multiple sensor nodes, however, means more measurements must be exchanged. This bandwidth issue can be partial overcome by distributed and decentralized fusion architectures since these allow for measurements to be pre-processed, within the sensor nodes, into filtered estimates which are exchanged. When speaking of estimates of a true state $\mathbf{x}$, we here mean the pair $(\hat{\mathbf{x}}, \mathbf{P})$, where $\hat{\mathbf{x}}$ is the state estimate and $\mathbf{P}$ the covariance reported by the estimator.

Benefits with decentralized solutions compared to their centralized counterparts are the distribution of data processing and the flexibility and modularity, where nodes can be added and removed on the fly. In total, these benefits provide decentralized solutions a high level of robustness.

One of the primary problems encountered in decentralized architectures, and track-to-track fusion (T2TF) in general, is the handling of the induced dependencies between the filtered estimates. Failure to handle these cross-correlations in a proper way can lead to *inconsistent* estimates as information is double counted, which, in the worst case, make the estimates start diverging. Providing consistent estimates is therefore one of the main aspects in this work. The definition of consistency used herein is formulated as [1]

$$\mathbf{P} - \mathrm{E}[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^{\mathsf{T}}] \succeq \mathbf{0}, \tag{1}$$

where $\tilde{\mathbf{x}} = \hat{\mathbf{x}} - \mathbf{x}$ is the true error of the estimate and $\mathrm{E}[\cdot]$ the expectation value. An estimator fulfilling (1) without reaching equality is called *conservative*.

The *Bar-Shalom-Campo* formulas compensate for all present cross-correlations such that optimality, in the sense of (1), can be assured [2, 3]. Unfortunately, the requirement of maintaining memory of cross-correlations makes this method inefficient for large sensor networks with many interacting nodes [4]. The recently developed *sample-based fusion* (SBF, [5, 6]) method partially comes around this tractability issue by introducing samples, which are themselves exchanged, that are used for reconstructing the cross-correlations. In the *generalized information matrix filter* (GIMF, [7]) the T2TF is solved by a decorrelation step where previously fused track information is removed prior to updating with new information.

The methods *covariance intersection* (CI, [1, 8]), *inverse covariance intersection* (ICI, [9]) and *safe fusion* (SF, [10]), the latter also called *ellipsoidal intersection* (EI, [11, 12]), all belong to the category of memory-less methods. None of these methods requires the actual cross-correlations to work, but come with the drawback that they also become conservative. Various of the mentioned fusion schemes have been compared in different contexts, see, *e.g.*, [10, 13]. In [14, 15] covariance matrix bounds for the decentralized fusion problem were derived.

In the common fusion setup the problem is to fuse two, possibly cross-correlated, estimates where all quantities are complete, *i.e.*, no partial information is missing. In [16] the problem of approximating the exchanged covariance with its diagonal was addressed. In this paper a bandwidth limited configuration is likewise assumed, but here the approach is rather to selectively choose and exchange only the most useful information.

The paper is organized as follows. Sec. II states and motivates the problem. General aspects of the considered estimation problem are provided in Sec. III. The proposed methods

This work has been supported by the Industry Excellence Center LINK-SIC funded by The Swedish Governmental Agency for Innovation Systems (VINNOVA) and Saab AB, and by the project Scalable Kalman filters funded by the Swedish Research Council (VR). G. Hendeby has received funding from the Center for Industrial Information Technology at Linköping University (CENIIT) grant no. 17.12.

are presented in Sec. IV and are evaluated experimentally in Sec. V. Concluding remarks are given in Sec. VI.

II. Problem Statement

The problem is to develop methods, denoted *information selection methods* (ISM), for determination of which information is the most useful to exchange such that high performance can be achieved using less communication bandwidth while still guaranteeing consistent estimates.

A. Considered Estimation Problem

We consider cooperative nodes complementing each other in the estimation of some true state $\mathbf{x}$, by observing and exchanging filtered estimates of the state. Roughly speaking, the estimation performance of such a setup will depend on properties of the local estimates and the communication bandwidth.

Essentially, three types of estimation problems are implied by the problem formulation:

1) Obtaining estimates by filtering local measurements into a common coordinate frame used for representing the state of the object of interest.
2) A decentralized data fusion problem in which the estimates received via datalink are merged with the local filter estimate. The cross-correlations between the fused estimates are assumed unknown.
3) Estimation of the information that is most valuable for the remaining nodes.

The choice of using CI for solving the decentralized data fusion problem is based upon the guarantee that CI produces consistent fused estimates as long as the estimates to be fused are consistent [17]. In fact, it has been shown that for fusion of two estimates under completely arbitrary cross-correlations CI provides the least conservative bound on the fused covariance [14].

B. Motivating Example

As a motivating example, consider the information matrices $\mathbf{P}_1^{-1}$ and $\mathbf{P}_2^{-1}$ defined according to their information ellipses as illustrated in Fig. 1. The gain of fusing $\mathbf{P}_1^{-1}$ with the full information $\mathbf{P}_2^{-1}$ is in this case not significantly higher than that of fusing with only the information along direction $\mathbf{v}_1$ in $\mathbf{P}_2$. Thus, the entity that has retrieved the information $\mathbf{P}_2^{-1}$ can, instead of transmitting the full information $\mathbf{P}_2^{-1}$, choose to only transmit the information along the $\mathbf{v}_1$. In this way the bandwidth consumption can be reduced, while still generating high information fusion gains.

III. Fusing Correlated Partial Estimates

We here investigate how N unbiased estimates $\{(\hat{\mathbf{x}}_i, \mathbf{P}_i)\}_{i=1}^N$ can be fused into $(\hat{\mathbf{x}}_f, \mathbf{P}_f)$ in a setup where only partial information is available.

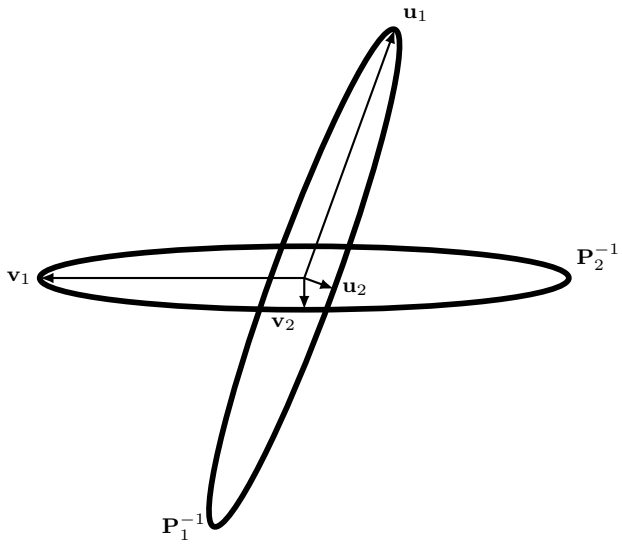

Fig. 1: The information ellipses of two information matrices, $\mathbf{P}_1^{-1}$ and $\mathbf{P}_2^{-1}$. The principal axes of $\mathbf{P}_1^{-1}$ are given by $\mathbf{u}_i$ and the principal axes of $\mathbf{P}_2^{-1}$ are given by $\mathbf{v}_i$, with $i = 1, 2$.

A. Fusion of Arbitrary Projections

An estimate in space $\mathcal{X}$ is given by $(\hat{\mathbf{x}}, \mathbf{P})$. An estimate $(\hat{\mathbf{x}}', \mathbf{P}')$ in an arbitrary space $\mathcal{X}'$ can be constructed, using the mapping $\mathbf{H}\colon \mathcal{X} \to \mathcal{X}'$, according to

$$(\hat{\mathbf{x}}', \mathbf{P}') = (\mathbf{H}\hat{\mathbf{x}}, \mathbf{H}\mathbf{P}\mathbf{H}^\mathsf{T}). \tag{2}$$

We limit ourselves to linear transformations and projections $\mathbf{H}$ with orthogonal basis vectors of unit length, *i.e.* the row vectors of $\mathbf{H}$ are orthonormal. The restriction to orthonormal projections is not necessary, but will nevertheless be useful as we will see later.

Now, the projection of $(\hat{\mathbf{x}}, \mathbf{P})$ along the direction defined by the unit vector $\mathbf{u}$ can be calculated by letting $\mathbf{H} = \mathbf{u}^\mathsf{T}$, *i.e.*

$$(\hat{x}_\mathbf{u}, \sigma_\mathbf{u}^2) = (\mathbf{u}^\mathsf{T}\hat{\mathbf{x}}, \mathbf{u}^\mathsf{T}\mathbf{P}\mathbf{u}), \tag{3}$$

where $\sigma_\mathbf{u}^2$ is the variance of the estimate $\hat{\mathbf{x}}$ in the direction defined by the vector $\mathbf{u}$. Hence, arbitrary components can be extracted of both the state estimates and their associated error covariance by using projections. The fusion of such arbitrary projections are realized by using the projection as the (virtual) measurement model.

B. Weighted Least Squares

N partial uncorrelated estimates can be fused using the *weighted least squares* (WLS) method as [18]

$$\mathbf{P}_f^{-1} = \sum_{i=1}^N \mathbf{H}_i^\mathsf{T}\mathbf{P}_i^{-1}\mathbf{H}_i, \tag{4a}$$

$$\mathbf{P}_f^{-1}\hat{\mathbf{x}}_f = \sum_{i=1}^N \mathbf{H}_i^\mathsf{T}\mathbf{P}_i^{-1}\hat{\mathbf{x}}_i, \tag{4b}$$

where $\hat{\mathbf{x}}_f$ is the fused state estimate and $\mathbf{P}_f$ is the estimated covariance of $\hat{\mathbf{x}}_f$. The matrix $\mathbf{H}_i$ is the mapping of $\hat{\mathbf{x}}_f$ into $\hat{\mathbf{x}}_i$.

C. Covariance Intersection

While (4) fuses uncorrelated estimates consistently, any degree of cross-correlation between estimates implies an inconsistent fused estimate. The CI algorithm fuses N estimates with unknown cross-correlations according to [17]

$$\mathbf{P}_f^{-1} = \sum_{i=1}^{N} \omega_i \mathbf{H}_i^\mathsf{T} \mathbf{P}_i^{-1} \mathbf{H}_i, \tag{5a}$$

$$\mathbf{P}_f^{-1} \hat{\mathbf{x}}_f = \sum_{i=1}^{N} \omega_i \mathbf{H}_i^\mathsf{T} \mathbf{P}_i^{-1} \hat{\mathbf{x}}_i, \tag{5b}$$

where $\sum_{i=1}^{N} \omega_i = 1$. CI produces a consistent fused estimate $(\hat{\mathbf{x}}_f, \mathbf{P}_f)$, for any values of the parameters $\omega_i \in [0, 1]$ fulfilling $\sum_{i=1}^{N} \omega_i = 1$, if each estimate in $\{(\hat{\mathbf{x}}_i, \mathbf{P}_i)\}_{i=1}^{N}$ is consistent [8]. The parameters ω_i are chosen by minimizing a loss function, *e.g.*, the trace or the determinant of $\mathbf{P}_f$.

It should be emphasized that (4) is similar to (5), both in terms of structure and involved matrices. The difference lies in the usage of scalar weights to avoid information from being double counted.

IV. Information Selection Methods

In this section three proposed methods for selecting information are proposed. The information projections are contained as column vectors of unit length in $\mathbf{U}_S$. The number of columns in $\mathbf{U}_S$ are denoted by n_S. The information along the directions defined in $\mathbf{U}_S$ are contained as the diagonal entries in the diagonal matrix $\mathbf{D}_S$. Thus, the partial information is fully described by $\mathbf{D}_S$ and $\mathbf{U}_S$.

A. Selecting Information

When selecting only a subset of projected information $\mathcal{I}_S = \mathbf{U}_S \mathbf{D}_S$, from an estimate $(\hat{\mathbf{x}}, \mathbf{P})$, to be exchanged, the state estimate $\hat{\mathbf{x}}$ must be projected accordingly. The state estimate is projected using (3), *i.e.* $\hat{\mathbf{x}}_S = \mathbf{U}_S^\mathsf{T} \hat{\mathbf{x}}$. An equivalent approach would be to consider projected variances, see, *e.g.*, (3), but here we will focus on information simply because of its additive properties in different algorithms.

To minimize the bandwidth allocation a certain node can choose to transmit only $(\hat{\mathbf{x}}_S, \mathcal{I}_S)$. The node receiving $(\hat{\mathbf{x}}_S, \mathcal{I}_S)$ can derive $\mathbf{U}_S$ and $\mathbf{D}_S$ separately using the fact that each mutually orthogonal eigenvector $\mathbf{u}_i$ of $\mathbf{U}_S$ is of unit length. The considered projections are equivalent to the (virtual) measurement model given by $\mathbf{H} = \mathbf{U}_S^\mathsf{T}$.

Three proposed ISM are introduced below, namely:

- The *largest eigenvalue* (LE) method where only the largest information eigenvalues are exchanged.
- The *transmitted information* (TI) method where the exchanged information is selected based on previously transmitted information.
- The *received information* (RI) method where the exchanged information is selected based on previously received information.

B. Aging Information

So far we have only considered static aspects of estimation, here in terms of fusion formulas. From now on, when the situation requires, we will use argument k for quantities calculated at time k, with the argument $k|k$ being used for filtered quantities.

The TI method and the RI method are based upon maintaining a separate filtered covariance matrix $\bar{\mathbf{P}}(k|k)$, referred to as the *reference covariance*, which is used in the information selection process. $\bar{\mathbf{P}}(k|k)$ is only locally available. When selecting the currently most valuable information, old information should be forgotten. The aging of information can be done in different ways, but is here accomplished by predicting $\bar{\mathbf{P}}(k|k)$ according to

$$\bar{\mathbf{P}}(k+1|k) = \mathbf{F} \bar{\mathbf{P}}(k|k) \mathbf{F}^\mathsf{T} + \mathbf{Q}, \tag{6}$$

where $\mathbf{F}$ is the process model. The process noise $\mathbf{Q}$ basically plays the role of a forgetting factor [18]. Any cross-correlations introduced by (6) are neglected.

C. Largest Eigenvalue Method

The eigendecomposition of the positive definite matrix $\mathbf{P}$ is defined as [19]

$$\mathbf{P} = \mathbf{V} \mathbf{\Sigma} \mathbf{V}^\mathsf{T} = \sum_i \lambda_i \mathbf{v}_i \mathbf{v}_i^\mathsf{T}, \tag{7}$$

where $\mathbf{\Sigma}$ is a diagonal matrix containing the ith eigenvalue λ_i of $\mathbf{P}$ on its ith diagonal entry. $\mathbf{V}$ is an orthogonal matrix, containing the corresponding mutually orthogonal eigenvectors $\mathbf{v}_i$ as column vectors.

Since the information matrix $\mathbf{P}^{-1}$ is positive definite it can also be decomposed according to (7), resulting in

$$\mathbf{P}^{-1} = (\mathbf{V} \mathbf{\Sigma} \mathbf{V}^\mathsf{T})^{-1} = \mathbf{V} \mathbf{\Sigma}^{-1} \mathbf{V}^\mathsf{T} = \sum_i \lambda_i^{-1} \mathbf{v}_i \mathbf{v}_i^\mathsf{T}. \tag{8}$$

The eigenvalues λ_i of $\mathbf{P}$ are assumed to be sorted in ascending order implying that $\mathbf{P}^{-1}$ is the most informative in the $\mathbf{v}_1$ direction since λ_1 comprises the smallest uncertainty.

A trivial and straightforward selection of the subset $\mathcal{I}_S$ is to perform the eigendecomposition of $\mathbf{P}^{-1}$ and selecting the n_S largest eigenvalues. Hence, the *largest eigenvalue* (LE) method can be written compactly as

$$\mathcal{I}_S = \{(\lambda_i^{-1}, \mathbf{v}_i)\}_{i=1}^{n_S}. \tag{9}$$

D. Transmitted Information Method

A more sophisticated approach is to estimate what information the other nodes of the network lack. In the *transmitted information* (TI) method the introduced reference covariance $\bar{\mathbf{P}}$ is the result of fusing only information that has been transmitted from the own node.

The TI procedure is as follows: $\mathcal{I}_S$ which in turn is derived from the covariance $\mathbf{P}$ of a track, is fused with $\bar{\mathbf{P}}$. Denoting the spanning of all subsets $\mathcal{I}_S$ by $\mathcal{F}(\mathcal{I}_S; n_S)$, where the size parameter n_S is explicitly included, the problem of selecting information can then be expressed as

$$\underset{\mathcal{I}_S \in \mathcal{F}}{\arg\min} \quad J(\mathbf{P}_f^c(\bar{\mathbf{P}}, \mathcal{I}_S)), \tag{10}$$

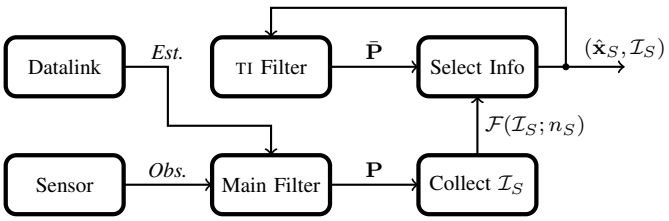

Fig. 2: Schematics over the TI method. The main filter is updating the track estimate $(\hat{\mathbf{x}}, \mathbf{P})$ with own sensor measurements and datalink estimates. Information selected for transmission is fused with $\bar{\mathbf{P}}$ in the TI filter.

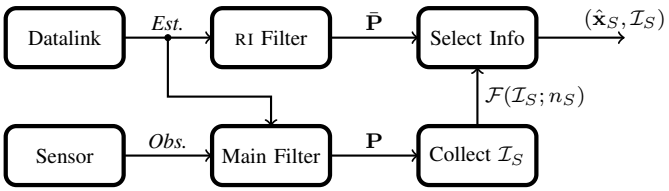

Fig. 3: Schematics over the RI method. The main filter is updating the track estimate $(\hat{\mathbf{x}}, \mathbf{P})$ with own sensor measurements and datalink estimates. Information received via the datalink is fused with $\bar{\mathbf{P}}$ in the RI filter.

where J is a loss function. The fused covariance candidate $\mathbf{P}_f^c$ is hence evaluated with respect to J over the subsets $\mathcal{I}_S \in \mathcal{F}$. The information $\mathcal{I}_S$ minimizing J is selected for transmission and is also fused with $\bar{\mathbf{P}}$.

TI will prefer information that has not yet been communicated. Schematics of TI is given in Fig. 2. Initially, when no information has been transmitted, the TI method selects the largest eigenvalues.

E. Received Information Method

The *received information* (RI) method is similar to the TI method given above but differ in how the reference covariance $\bar{\mathbf{P}}$ is calculated. In the RI method $\bar{\mathbf{P}}$ is fused only with information received from other nodes. RI will favour information not contained in the received information. RI is schematically illustrated in Fig. 3.

V. Experimental Evaluation

In this section the simulation scenarios, the evaluation metrics and the results are presented and discussed. But first, we will begin by mentioning the information projections that are considered in the experimental evaluation.

A. Considered Information Projections

So far we have considered information in arbitrary directions. Even though the methods proposed here hold for arbitrary projections of information we will, to simplify the following analysis, only deal with projections along the eigenvectors of the information matrix.

The eigendecomposition, see (8), allows us to easily break down the information into orthogonal components. In essence, the matrix $\mathbf{V}$ in $\mathbf{P}^{-1} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{V}^\mathsf{T}$ is a rotation matrix, rotating

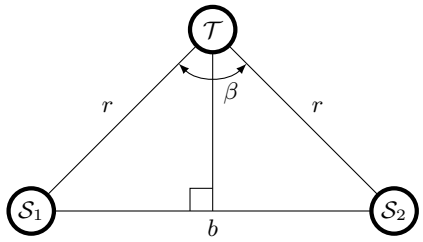

Fig. 4: Two sensors, $\mathcal{S}_1$ and $\mathcal{S}_2$, observing a target $\mathcal{T}$ at range r. The baseline between the sensor nodes is b which is related to r and β.

the diagonal and axis aligned matrix $\mathbf{\Sigma}^{-1}$ into the frame that $\mathbf{P}^{-1}$ is represented in. For example, the information $\mathbf{P}_{\mathbf{v}_i}^{-1}$ can be calculated as

$$\mathbf{P}_{\mathbf{v}_i}^{-1} = \mathbf{v}_i^\mathsf{T}\mathbf{P}^{-1}\mathbf{v}_i = \mathbf{v}_i^\mathsf{T}\mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{V}^\mathsf{T}\mathbf{v}_i = \lambda_i^{-1}, \qquad (11)$$

where λ_i is the ith eigenvalue of $\mathbf{P}$ and where the orthogonality property of the eigenvectors has been used.

Using the eigendecomposition as the basis for information projections reduces the optimization problem in (10) into a grid search among a finite set $\mathcal{F}$ consisting of combinations of information projections.

B. Scenarios and Experimental Setup

Our evaluation will be based on scenarios where the tracking sensors are highly accurate in bearing but are of poor range accuracy, *e.g.*, an *infra-red search and track* (IRST) system. The uncertainty in range for tracking filters deploying such sensors can be enhanced in a setup where the observed target is triangulated using two spatially separated sensors.

The estimation performance when triangulating depends on the geometry. A triangulation setup is illustrated in Fig. 4. Two sensors, denoted $\mathcal{S}_1$ and $\mathcal{S}_2$, are observing a target $\mathcal{T}$ located at distance r from each sensor. The baseline b depends on both r and the baseline angle β according to

$$b = 2r \sin \frac{\beta}{2}. \qquad (12)$$

The baseline angle β will parametrize the geometry in the following.

We will evaluate the considered estimation problem on two extreme triangulation geometries, *i.e.* the following two scenarios:

1) *Scenario 1:* The baseline angle is approximately $90°$.
2) *Scenario 2:* The baseline angle is small but non-zero.

In addition to the proposed ISM a number of reference methods will be used in the evaluation. In total we have the following methods:

1) LE – Largest eigenvalue method (Sec. IV-C).
2) TI – Transmitted information method (Sec. IV-D).
3) RI – Received information method (Sec. IV-E).
4) CRLB – The *Cramér-Rao lower bound* (see below).
5) CKF – A *centralized Kalman filter* utilizing the unprocessed measurements from each sensor.

6) LKF – A *local Kalman filter* having access only to local sensor data.
7) CIF – A decentralized scheme deploying CI and the full filtered estimates from all sensing nodes.

Next, an analysis of the fusion gain dependency on the baseline angle is described. The purpose of the baseline angle analysis is to connect the two extreme geometries evaluated in Scenario 1 and 2.

C. Baseline Angle Analysis

To investigate how the fusion gain more generally depends on the geometry we introduce the covariance matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ according to

$$\mathbf{P}_1 = \mathbf{T}(-\beta/2)\mathbf{A}(a)\mathbf{T}(-\beta/2)^{\mathsf{T}}, \tag{13a}$$

$$\mathbf{P}_2 = \mathbf{T}(\beta/2)\mathbf{A}(a)\mathbf{T}(\beta/2)^{\mathsf{T}}, \tag{13b}$$

respectively, where $\mathbf{A}$ and $\mathbf{T}$ are given by

$$\mathbf{A}(a) = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}, \tag{14a}$$

$$\mathbf{T}(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}. \tag{14b}$$

$\mathbf{P}_1$ and $\mathbf{P}_2$ have the same eigenvalues as $\mathbf{A}$, but with the eigenvectors being rotated by $\mathbf{T}(-\beta/2)$ and $\mathbf{T}(\beta/2)$, respectively.

In the baseline angle analysis $\mathbf{P}_1^{-1}$ and $\mathbf{P}_2^{-1}$ are being fused using CI into $\mathbf{P}_f^{-1}$, with β varied and for different values of a. The weight parameters ω_i of (5a) are optimized using different norms, *i.e.* the determinant $\det(\mathbf{P}_f)$, the trace $\mathrm{tr}(\mathbf{P}_f)$, and the spectral norm $\lambda_{\max}(\mathbf{P}_f)$. The cases where $\mathbf{P}_1^{-1}$ is fused with complete $\mathbf{P}_2^{-1}$, with only the largest eigenvalue of $\mathbf{P}_2^{-1}$ and with only the smallest eigenvalue of $\mathbf{P}_2^{-1}$, are studied. Fig. 5 provides a summary of the baseline angle analysis.

D. Evaluation Metrics

The position component of the *root mean squared error* (RMSE) is used to measure the performance. The position RMSE for each time step over all MC runs are calculated, yielding a time series of position RMSE for each method.

The *parametric* CRLB [20], $\mathbf{P}^0$, is calculated according to the following filter recursion

$$\mathbf{P}^0(k|k) = \left((\mathbf{P}^0(k|k-1))^{-1} + (\mathbf{H}^0)^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{H}^0\right)^{-1}, \tag{15a}$$

$$\mathbf{P}^0(k+1|k) = \mathbf{F}^0\mathbf{P}^0(k|k)(\mathbf{F}^0)^{\mathsf{T}} + \mathbf{Q}^0, \tag{15b}$$

where $\mathbf{F}^0$ is a model of the true dynamics, $\mathbf{H}^0 = \frac{d}{d\mathbf{x}'}\mathbf{h}(\mathbf{x}')|_{\mathbf{x}'=\mathbf{x}}$ is the true measurement model, $\mathbf{h}(\mathbf{x}')$ is the mapping from state coordinates to measurement coordinates, and $\mathbf{R}$ is the measurement covariance. The true process noise $\mathbf{Q}^0$ is derived from the true dynamics of the target. Time indices have been included in (15) since $\mathbf{P}^0$ is a filtered quantity.

Consistency is evaluated using the *normalized estimation error squared* (NEES), defined as [21, 22]

$$\varepsilon(k) = (\mathbf{x}(k) - \hat{\mathbf{x}}(k|k))^{\mathsf{T}}\mathbf{P}(k|k)^{-1}(\mathbf{x}(k) - \hat{\mathbf{x}}(k|k)), \tag{16}$$

where $\mathbf{x}(k)$ is the true state, $\hat{\mathbf{x}}(k|k)$ the state estimate and $\mathbf{P}(k|k)$ the estimated covariance, all given at time k. In a *Monte Carlo* (MC) approach multiple values of $\varepsilon(k)$ are calculated which can be combined into the average NEES (ANEES) according to

$$\bar{\varepsilon}(k) = \frac{1}{n_x M}\sum_{i=1}^{M}\varepsilon_i(k), \tag{17}$$

where $\varepsilon_i(k)$ is the NEES value at time k for MC run i, n_x is the number of dimensions of $\hat{\mathbf{x}}$ and M is the number of MC runs. An estimator yielding ANEES significantly larger than 1 is interpreted as an optimistic and inconsistent estimator. An estimator for which ANEES is significantly smaller than 1 is consistent but conservative.

E. Simulation Specifications

The two simulation scenarios are shown in Fig. 6. In both scenarios there are two cooperating nodes sensing a target. In Scenario 1 the cooperating nodes are well separated, yielding a large baseline angle. In Scenario 2 the baseline angle is small. The simulations are performed in 3D, but the dynamics are constrained to a plane.

The cooperating nodes are synchronized and each tracking sensor generates new measurements at 1 Hz. No clutter is simulated and the detection probability is 100%. The measurement noise is white Gaussian noise, where the uncertainty in range and bearing is set to $5\,000$ m and $0.5°$, respectively. The measurements are filtered using an *extended Kalman filter* (EKF, [23]) and a *constant velocity model* is used to describe the dynamics, *i.e.* the state dimensionality $n_x = 6$. The filter estimates are given in a *Cartesian* frame. The process noise is tuned such that the CKF initially achieves $\bar{\varepsilon}(k) \approx 1$ in Scenario 1.

Also the datalink runs at $1/T_{\text{DL}} = 1$ Hz, where T_{DL} is the transmission period, and is scheduled to exchange filtered estimates immediately after each measurement update.

The simulation environment is implemented in MATLAB®. $10\,000$ (MC) runs are performed for each method with the same noise realization being used for each and every method. The optimization problems implied by CI, TI and RI are solved using the trace of the relevant covariance matrix as the loss function. The fused covariance candidate $\mathbf{P}_f^c(\bar{\mathbf{P}}, \mathcal{I}_S)$ of TI and RI is generated using (4a). It should be noted that the results of using (4a) for optimizing the subset $\mathcal{I}_S$ are in general not the same as using CI for this optimization. However, the results would be the same as using CI constrained to $\omega = 0.5$ since this is equivalent (up to a scaling factor) to (4a).

F. Results and Discussion

The results presented below are derived from only one of the two cooperating nodes. The results generated by the other cooperating node are approximately equivalent. This is true for both scenarios.

The simulation results, RMSE and ANEES, for Scenario 1 are given in Fig. 7. $n_S \leq n_x = 6$ is the number of eigenvalues used in the corresponding ISM, *i.e.* the number of eigenvalues

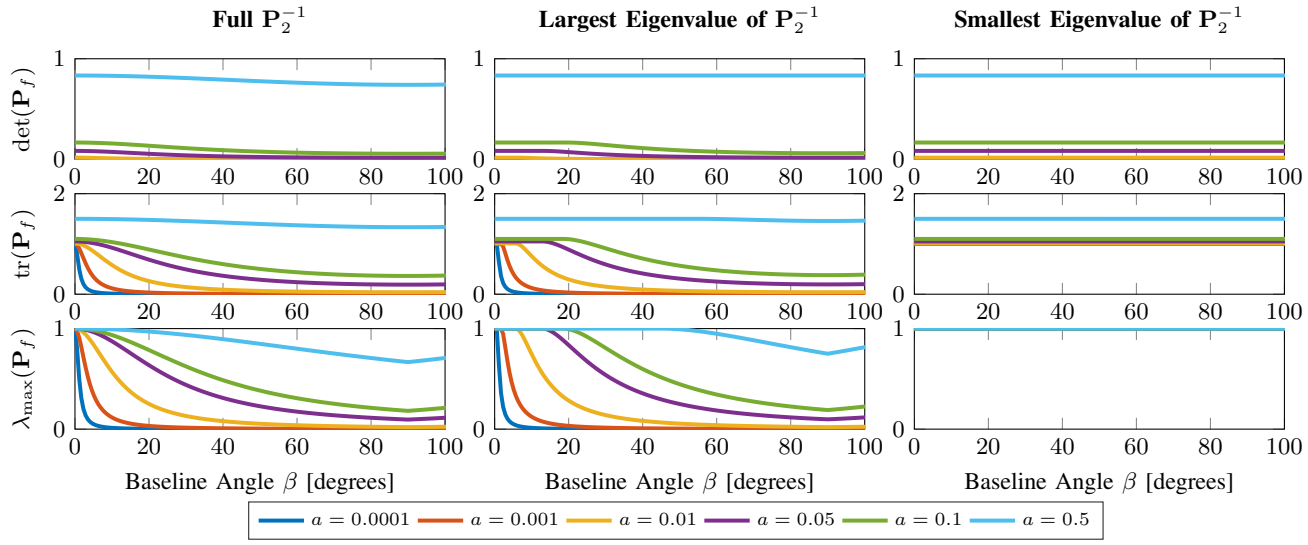

Fig. 5: Analysis of how the fusion gain depends on the baseline angle β for different norms of the covariance $\mathbf{P}_f$ resulting from fusing the information $\mathbf{P}_1^{-1}$ with information given by full $\mathbf{P}_2^{-1}$, the largest eigenvalue of $\mathbf{P}_2^{-1}$ and the smallest eigenvalue of $\mathbf{P}_2^{-1}$. a^{-1} is the largest eigenvalue of each of $\mathbf{P}_1^{-1}$ and $\mathbf{P}_2^{-1}$. All curves are symmetric around 90°.

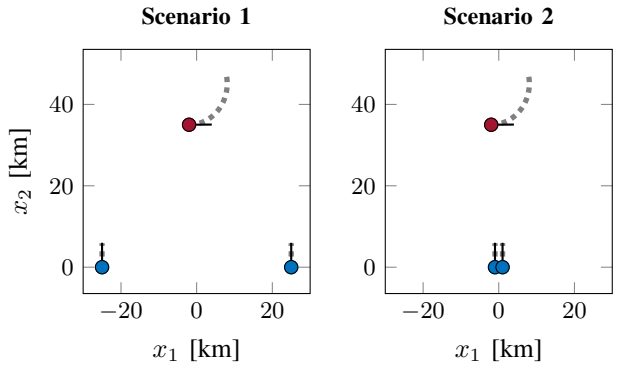

Fig. 6: Simulation scenarios used in the experimental evaluation. The cooperating sensor nodes are marked blue, and the target, located above the sensor nodes each case, is marked red. The black lines from a node indicate its initial orientation. Dotted lines represent the simulated trajectories.

exchanged in each datalink transmission. When $n_S = 1$, only TI performs comparable to CIF, but with a slower transient. As suggested, increasing n_S increases the performance. For TI and RI the performance of CIF is reached when $n_S \geq 3$. Initially, LE performs well since the largest information eigenvalues will initially be the most useful ones, but there is a bump around 10 s for each LE curve. This bump is explained by the fact that the largest information eigenvalues are not always the most informative for the receiving node. In general, LE requires $n_S > 3$ before the performance of CIF is reached. The RMSE for all methods are bounded from below by CKF and from above by LKF. CRLB is represented by the square root of the trace of the position part of $\mathbf{P}^0$.

For Scenario 2, RMSE and ANEES are given in Fig. 8. The RMSE for all methods are bounded from below and above by CKF and LKF, respectively. For this geometry $n_S > 3$ is required for all three ISM before the performance of CIF is approached.

Except for the initial peaks all methods, both in Scenario 1 and 2, are consistent and quickly become conservative due to the accumulated process noise. More eigenvalues exchanged typically means a more conservatively fused estimate.

In Fig. 9 and 10 the performance at $k_e = 12$ s is plotted for different T_{DL} for Scenario 1 and 2, respectively. The geometry provided by Scenario 1 leads to a higher sensitivity to T_{DL} than the geometry provided by Scenario 2, relatively speaking.

The results for the baseline angle analysis is given in Fig. 5. When full information $\mathbf{P}_2^{-1}$ is fused with $\mathbf{P}_1^{-1}$, the gain is strictly monotonically increasing when β goes from 0° to 90°, which is true for each norm and for each a. This is however not true for the case when only the largest eigenvalue of $\mathbf{P}_2^{-1}$ is fused with $\mathbf{P}_1^{-1}$, where there is no gain until a certain threshold angle β_t is reached. The threshold β_t seems to depend on a. As expected, no gain is achieved when fusing $\mathbf{P}_1^{-1}$ with the smallest eigenvalue of $\mathbf{P}_2^{-1}$.

The threshold β_t reflects what is seen in the simulation results. The LE method performs poorly with a small baseline angle since no fusion gain is acquired at the receiving node when transmitting only the most informative projections.

G. Bandwidth Reduction

When considering the exchange of full estimates $n_x(n_x + 1)/2$ parameters must be transmitted. Using the proposed methods with $n_S = 1$ only $n_x + 1$ parameters are required. For arbitrary n_S the number of parameters required in each transmission is calculated as $N_p = n_S + \sum_{i=1}^{n_S}(n_x - i + 1) = n_S(3 + 2n_x - n_S)/2$ since we have assumed mutually orthogonal projections.

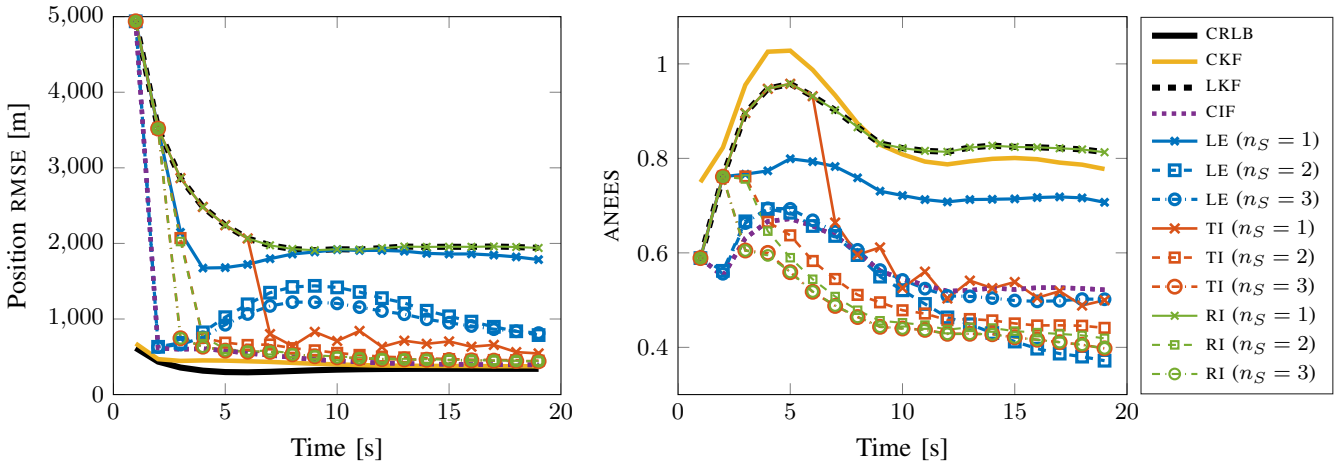

Fig. 7: Scenario 1: Position RMSE to the left and ANEES to the right.

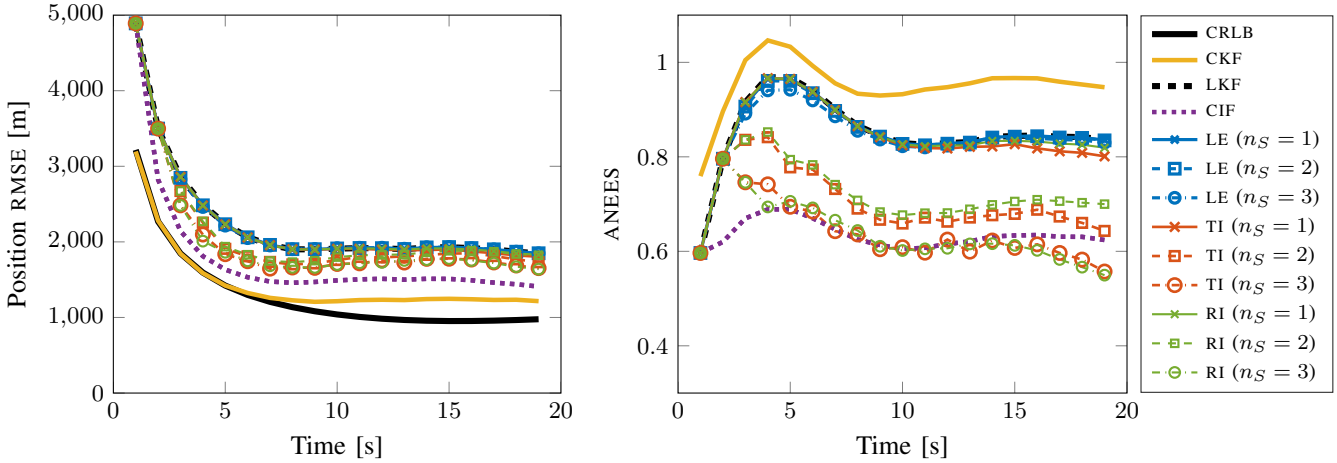

Fig. 8: Scenario 2: Position RMSE to the left and ANEES to the right.

TABLE I: The number of parameters N_p transmitted, when using the proposed ISM, as dependent on the number of projections n_S for different n_x.

n_S	$N_p(n_x = 4)$	$N_p(n_x = 6)$	$N_p(n_x = 9)$
1	5	7	10
2	9	13	19
3	12	18	27
4	14	22	34
6	-	27	45
9	-	-	54

The bandwidth reduction is illustrated in Table I by calculating N_p for different n_x and n_S.

VI. Conclusions and Future Work

Three *information selection methods* (ISM) have been derived, suitable for a decentralized configuration of sensor nodes. The ISM try to select information projections based on what information is the most valuable for the remaining nodes of the sensor network. A comparison study of the ISM has been conducted, regarding their performance, measured as *root mean squared error* (RMSE), and consistency, measured as *average normalized mean squared error* (ANEES). The evaluation was made using two triangulation setups.

The *largest eigenvalue* (LE) method suffers from its inability to prevent the same information projections from being repeatedly exchanged. Once the most informative projections have been transmitted, these will often become the most informative projections at the receiving node, which in turn will transmit the same information projections. This type of *round trip* inhibits complementary information from being extracted.

The *transmitted information* (TI) method and the *received information* (RI) method circumvent the previously discussed round trip issue by trying to maximize the fusion gain at the remaining nodes. Hence, not only the instantaneously most informative projections will be transmitted, but rather the projections that favours the considered fusion problem. This optimization strategy enables complementary information to be extracted in a larger extent.

Both TI and RI outperform the LE method in terms of

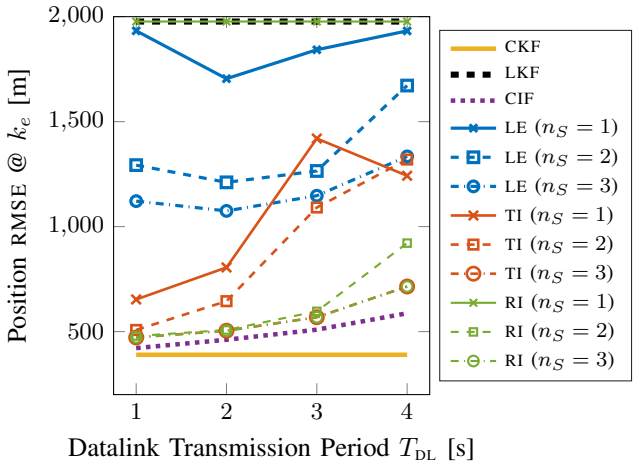

Fig. 9: Scenario 1: Position RMSE evaluated at k_e for different datalink transmission periods T_{DL}.

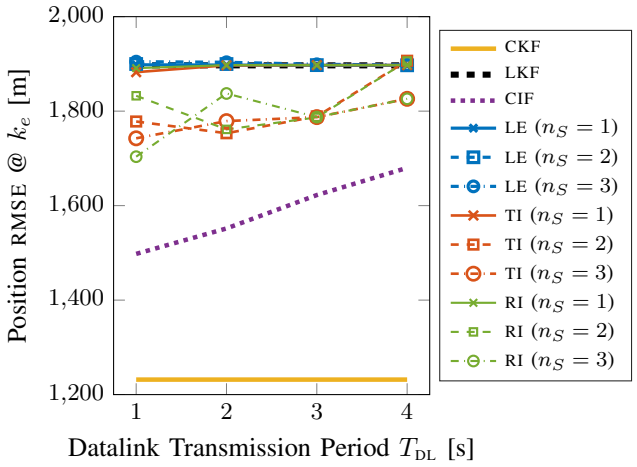

Fig. 10: Scenario 2: Position RMSE evaluated at k_e for different datalink transmission periods T_{DL}.

RMSE. All ISM yield consistent ANEES values. A key feature of the suggested methods is that the bandwidth allocation of the communication link can be significantly reduced. It is also remarkable that none of the proposed ISM require any additional information than the communicated estimates themselves.

A possible future extension of the problem addressed herein is to allow TI and RI to be optimized using arbitrary projections instead of only projections along eigenvectors. The algorithms can be further improved by making them more flexible regarding the number of information projections selected. Another aspect is to analyze how the chosen fusion rule affects the performance. It would also be interesting to combine TI and RI.

REFERENCES

[1] S. J. Julier and J. K. Uhlmann, "A non-divergent estimation algorithm in the presence of unknown correlations," in *Proceedings of the 1997 American Control Conference*, Jun. 1997, pp. 2369–2373.

[2] Y. Bar-Shalom and L. Campo, "The effect of the common process noise on the two-sensor fused-track covariance," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, no. 6, pp. 803–805, Nov. 1986.

[3] K. C. Chang, R. K. Saha, and Y. Bar-Shalom, "On optimal track-to-track fusion," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 33, no. 4, pp. 1271–1276, Oct. 1997.

[4] X. Tian and Y. Bar-Shalom, "Exact algorithms for four track-to-track fusion configurations: All you wanted to know but were afraid to ask," in *Proceedings of 12th IEEE International Conference on Information Fusion*, Seattle, WA, USA, Jul. 2009.

[5] J. Steinbring, B. Noack, M. Reinhardt, and U. D. Hanebeck, "Optimal sample-based fusion for distributed state estimation," in *Proceedings of 19th IEEE International Conference on Information Fusion*, Heidelberg, Germany, Jul. 2016.

[6] S. Radtke, B. Noack, U. D. Hanebeck, and O. Straka, "Reconstruction of cross-correlations with constant number of deterministic samples," in *Proceedings of 21th IEEE International Conference on Information Fusion*, Cambridge, UK, Jul. 2018.

[7] X. Tian and Y. Bar-Shalom, "On algorithms for asynchronous track-to-track fusion," in *Proceedings of 13th IEEE International Conference on Information Fusion*, Edinburgh, Scotland, Jul. 2010.

[8] S. J. Julier and J. K. Uhlmann, *Handbook of Multisensor Data Fusion*. Boca Raton, FL, USA: CRC Press, 2001, ch. General Decentralized Data Fusion with Covariance Intersection.

[9] B. Noack, J. Sijs, and U. D. Hanebeck, "Inverse covariance intersection: New insights and properties," in *Proceedings of 20th IEEE International Conference on Information Fusion*, Xi'an, China, Jul. 2017.

[10] J. Nygårds, V. Deleskog, and G. Hendeby, "Safe fusion compared to established distributed fusion methods," in *Proceedings of IEEE Int. Conf. on Multisensor Fusion*, Baden-Baden, Germany, Sep. 2016.

[11] J. Sijs, M. Lazar, and P. P. J. v. d. Bosch, "State fusion with unknown correlation: Ellipsoidal intersection," in *Proceedings of 2010 American Control Conference*, Baltimore, MD, USA, Jun. 2010, pp. 3992–3997.

[12] J. Sijs and M. Lazar, "Empirical case-studies of state fusion via ellipsoidal intersection," in *Proceedings of 14th IEEE International Conference on Information Fusion*, Salamanca, Spain, Jul. 2014.

[13] S. Radtke, K. Li, B. Noack, and U. D. Hanebeck, "Comparative study of track-to-track fusion methods for cooperative tracking with bearings-only measurements," in *2nd IEEE International Conference on Industrial Cyber Physical Systems*, Taipei, Taiwan, May 2019.

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

[15] J. Ajgl and O. Straka, "Comparison of fusions under unknown and partially known correlations," *IFAC-PapersOnLine*, vol. 51, no. 23, pp. 295–300, Aug. 2018, 7th IFAC Workshop on Distributed Estimation and Control in Networked Systems NECSYS 2018.

[16] R. Forsling, Z. Sjanic, F. Gustafsson, and G. Hendeby, "Consistent distributed track fusion under communication constraints," in *Proceedings of 22th IEEE International Conference on Information Fusion*, Ottawa, Canada, Jul. 2019.

[17] J. K. Uhlmann, "Covariance consistency methods for fault-tolerant distributed data fusion," *Information Fusion*, vol. 4, pp. 201–215, Sep. 2003.

[18] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation theory*, ser. Fundamentals of Statistical Signal Processing. Upper Saddle River, NJ, USA: Prentice Hall, 1993.

[19] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Upper Saddle River, NJ, USA: Prentice Hall, 2000.

[20] J. Taylor, "The Cramér-Rao estimation error lower bound computation for deterministic nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 24, no. 2, pp. 343–344, April 1979.

[21] X. R. Li and Z. Zhao, "Measuring estimator's credibility: Noncredibility index," in *Proceedings of 9th IEEE International Conference on Information Fusion*, Florence, Italy, Jul. 2006.

[22] Y. Bar-Shalom, X. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. New York, NY, USA: John Wiley & Sons, Ltd, 2004.

[23] A. Jazwinski, *Stochastic processes and filtering theory*. New York, NY, USA: Academic Press, 1970.