Estimation of GNSS Atmospheric Delay in Networked Systems

Jean-Bernard Uwineza*

Advisor: Jay A. Farrell

Abstract—This report explores different methods of estimating the atmospheric delay error in Global Navigation Satellite Systems (GNSS) within a network of interacting agents. In the first part, we state the problem and formulate the individual solution when there is no inter-agent interactions, followed by the homogeneous centralized solution when all agents can interact with a central node. We then show that a distributed information-weighted average consensus solution converges to the centralized solution. The second part considers the case of heterogeneous agents by exploring methods of combining estimates with unknown correlations. We conclude by outlining some future directions.

1 Introduction

An important requirement in Intelligent Transportation Systems (ITSs) is the availability of accurate and reliable location information. In navigation systems, and specially for autonomous vehicles (AVs), this is the foremost requirement as AVs require up to centimeter-level navigation accuracy, high availability, and integrity [1], [2]. No individual technology can currently offer that level of performance, and a suite of sensors needs to be used, with the Global Navigation Satellite System (GNSS) sensors [3]–[5] as the main provider of absolute positioning. Hence, to improve overall navigation performance in AVs (and other applications), GNSS-based positioning must be accurate and reliable.

The most prominent framework for achieving high standalone GNSS positioning accuracy is Differential GNSS (DNSS). Local Area DGNSS (LADGNSS) approach uses local networks of base stations to offer corrections to users within a small geographical area. Cited accuracy levels in the literature for LADGNSS are around 1–3 meters [6], [7], which can fall short of the *SAE J2945/1* specifications for autonomous vehicles [8]. Besides accuracy, LADGNSS is limited only to local applicability, making it unfeasible to implement on continental geographical areas. The most prominent alternative that has gained widespread adoption is Precise Point Positioning (PPP), which offers very accurate global corrections derived from external data sources spread across the globe [9], [10]. PPP achieves centimeter-level accuracy by using ionosphere-free pseudorange and carrier-phase observations from dual-frequency receivers, as well as resolved carrier-phase integer ambiguities [11], [12]. The resolution of integer ambiguities depends of factors such as local environmental conditions, the elevation, signal strength, and number of available satellites. As a result, it often takes several epochs for the resolution to converge. However, there are various international organizations that use networks of stations to provide real-time or near real-time PPP corrections (e.g. International GNSS Service (IGS) and European Space Agency (ESA)). In [13], [14], IGS real-time corrections were used in combination with dual-frequency receivers. Due to various constraints (financial and otherwise), consumer markets favor single-frequency receivers.

In [15]–[17], real-time single-frequency implementations of PPP compensate for ionospheric error by combining pseudorange and carrier phase measurements. The solution is reported to take around 1.5–4 hours to converge. The accuracy of 0.30 m and 0.71m is reported for stationary and moving platforms, respectively [17]. The use of GNSS-aided INS approaches significantly reduces convergence time and increases accuracy. A convergence time of 5 minutes and horizontal accuracy of 60 cm is reported in [18], whereas [14] reports a convergence time of 5 minutes and accuracy of 0.05 m. Of the main reasons for long convergence times is that many observations are needed to estimate the residual atmospheric delay in PPP-compensated GNSS measurements. Automotive applications require convergence time on the order of just a few seconds, with high reliability. In this report we study the use of local agents cooperating in a networked system to estimate the residual atmospheric delay error, with the goal of significantly reducing convergence time.

In Section 2, we formulate the state estimation problem, where the full state to be estimated consists of variables unique to each agent and an additional common variable—the local GNSS atmospheric delay error. Section 3 finds both the individual and centralized estimates of the full state. Since we are not interested in cooperating to estimate the unique variables of each agent, Section 4 finds the estimate of only the common variable; first in a centralized non-distributed manner, then in a distributed framework via

* *Email: buwineza@ece.ucr.edu*

information-weighted average consensus. In Section 5, we consider a different distributed framework which addresses the unknown correlations between individual estimates. Sections 6 and 7 offer concluding remarks and explore avenues for future research, respectively.

2 PROBLEM FORMULATION

Suppose we have a network of vehicles navigating in the same *local* environment, each making independent GNSS measurements from its sensors. At time t, the measurement from satellite s available at vehicle v is y_v^s, modeled as

$$y_v^s = \mathbf{1}_v^s \, \mathbf{p}_v + ct_v + M_v^s \, T + \eta_v, \tag{1}$$

where $\mathbf{p}_v$ denotes a position vector, ct_v is clock error term for each vehicle, T is the atmospheric delay error term, and η_v is the measurement noise, assumed to be zero-mean Gaussian. The quantity $\mathbf{1}_v^s$ is the line-of-sight unit vector from s to v, and $M_v^s = 1/\sin(e_v^s)$ is the obliquity coefficient, where e_v^s is the local elevation angle at the signal pierce point relative to vehicle v [19], [20]. The above equation can be expressed as

$$y_v^s = \begin{bmatrix} \mathbf{1}_v^s & 1 & M_v^s \end{bmatrix} \begin{bmatrix} \mathbf{p}_v \\ b_v \\ T \end{bmatrix} + \eta_v \tag{2}$$

where $b_v = ct_v$. Let $\mathbf{x}_v = [\mathbf{p}_v, b_v, T]^\top \in \mathbb{R}^p$, and $\mathbf{h}_v^s = [\mathbf{1}_v^s, 1, M_v^s] \in \mathbb{R}^{1\times p}$. The value of m_v is the number of measurements at vehicle v. Assuming the atmospheric delay T to be constant in a local geographic area, the s superscript is not required. Then, there are elements of the state vector that are common to all vehicles and others that are unique. The state vector becomes $\mathbf{x}_v = [\mathbf{z}_v, T]^\top$, where $\mathbf{z}_v = [\mathbf{p}_v, b_v]^\top$ is the unique part of the state vector for each vehicle.

Hence, combining the measurements from eqn. (2), we obtain the standard form

$$\mathbf{y}_v = \mathbf{H}_v \mathbf{x}_v + \boldsymbol{\eta}_v \tag{3}$$

where $\mathbf{y}_v \in \mathbb{R}^{m_v}$ and $\mathbf{H}_v = [\mathbf{h}_v^1, \ldots, \mathbf{h}_v^{m_v}]^\top \in \mathbb{R}^{m_v \times p}$ is the observation matrix for vehicle v, and is time-varying. The noise $\boldsymbol{\eta}_v \in \mathbb{R}^{m_v}$ is modeled as a Gaussian random variable with $\eta_v \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_v)$ where $\mathbf{R}_v \in \mathbb{R}^{m_v \times m_v}$ is the corresponding covariance matrix. The measurement information matrix at vehicle is $\mathbf{B}_v = \mathbf{R}_v^{-1} \in \mathbb{R}^{m_v \times m_v}$. The goal is to jointly solve the shared atmospheric delay state in a distributed manner.

3 NON-DISTRIBUTED ESTIMATION OF THE FULL STATE

For comparison purposes, this section discusses two particular scenarios regarding the communication between vehicles in the network. The first is when there is zero communication between nodes, and each agent acts alone (has the least information). This is the case when every vehicle is individually solving for its state, using *only* its measurements. In this case, a weighted least-squares estimate provides the Minimum Variance Unbiased (MVU) estimate and could be used every time the measurements are available. The second scenario is when all agents communicate to a central node. Hence, the estimator has maximum information.

3.1 Individual Solution

With no communications in the network, each node can use its own measurements to solve for its state. With the measurement model in eqn. (3), and a prior estimate whose information matrix is $\mathbf{J}_v^-$, the MVU solution is

$$\hat{\mathbf{x}}_v^+ = \left(\mathbf{J}_v^- + \mathbf{H}_v^\top \mathbf{B}_v \mathbf{H}_v\right)^{-1} \left(\mathbf{J}_v^- \mathbf{x}_v^- + \mathbf{H}_v^\top \mathbf{B}_v \mathbf{y}_v\right). \tag{4}$$

The posterior information in the state is

$$\mathbf{J}_v^+ = \mathbf{J}_v^- + \mathbf{H}_v^\top \boldsymbol{B}_v \mathbf{H}_v, \tag{5}$$

where $\mathbf{J}_v$ is the information matrix. The posterior covariance matrix is

$$\mathbf{P}_v^+ = \left(\mathbf{J}_v^- + \mathbf{H}_v^\top \boldsymbol{B}_v \mathbf{H}_v\right)^{-1} = \begin{bmatrix} \mathbf{P}_{\mathbf{z}_v\mathbf{z}_v}^+ & \mathbf{P}_{\mathbf{z}_v T}^+ \\ \mathbf{P}_{T\,\mathbf{z}_v}^+ & \mathbf{P}_{T\,T}^+ \end{bmatrix}. \tag{6}$$

3.2 Centralized MAP Solution

With a network of vehicles $v \in 1, \ldots, N$ where there is communication to a central node, the goal is to estimate the concatenated state vector $\mathbf{x} = [\mathbf{z}_1, \ldots, \mathbf{z}_N, T]^\top \in \mathbb{R}^q$, where $q = (p-1)N + 1$ and N is the number of vehicles (nodes) in the network. For analysis, it is helpful to redefine $\mathbf{x}$ with a permutation P to obtain $\mathbf{x}_P = P\mathbf{x}$ such that $\mathbf{x}_P = [\mathbf{p}_1, \ldots, \mathbf{p}_{N_v}, b_1, \ldots, b_N, T]^\top$. (Subsequently, $\mathbf{x}_P$ will be referred to simply as $\mathbf{x}$.)

The collection of all measurements from all sensor is expressed as

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{H}} \mathbf{x} + \boldsymbol{\nu} \tag{7}$$

where $\boldsymbol{\mathcal{Y}} = \left[\mathbf{y}_1^\top, \mathbf{y}_2^\top, \ldots, \mathbf{y}_N^\top\right]^\top \in \mathbb{R}^m$ is the concatenation of all measurements in the network, with $m = \sum_{v=1}^N m_v$. From eqn. (3), let $\mathbf{H}_v = [\mathbf{A}_v \; \mathbf{1}_v \; \mathbf{M}_v]$, where each $\mathbf{A}_v \in \mathbb{R}^{m_v \times (p-2)}$ is a matrix composed of unit vectors, and both $\mathbf{1}_v \in \mathbb{R}^{m_v}$ and $\mathbf{M}_v \in \mathbb{R}^{m_v}$ are column vectors. Then $\boldsymbol{\mathcal{H}} \in \mathbb{R}^{m \times q}$ is a block matrix of the form:

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{1}_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{M}_1 \\ \mathbf{0} & \mathbf{A}_2 & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{1}_2 & \cdots & \mathbf{0} & \mathbf{M}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_N & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1}_N & \mathbf{M}_N \end{bmatrix}. \tag{8}$$

For the measurement noise vector $\boldsymbol{\nu} = \left[\nu_1^\top, \nu_2^\top, \ldots, \nu_m^\top\right]^\top \in \mathbb{R}^m$, the covariance matrix is $\boldsymbol{\mathcal{R}} \in \mathbb{R}^{m\times m}$ and the corresponding information matrix is $\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{R}}^{-1} \in \mathbb{R}^{m\times m}$. As the measurement noises are assumed to be uncorrelated across all nodes (vehicles), the covariance matrix, $\boldsymbol{\mathcal{R}}$, is diagonal and the information matrix is $\boldsymbol{\mathcal{B}} = diag(\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_N)$.

The value of m can change depending on which satellites are visible at each node. This is because vehicles in the same area do not see the same number of satellites, as some satellites could be obstructed by obstacles such as buildings or trees. Similarly, the value of N can change as vehicles travel in and out of the predefined local area. For simplicity, we will consider that these two values are constant, i.e. all

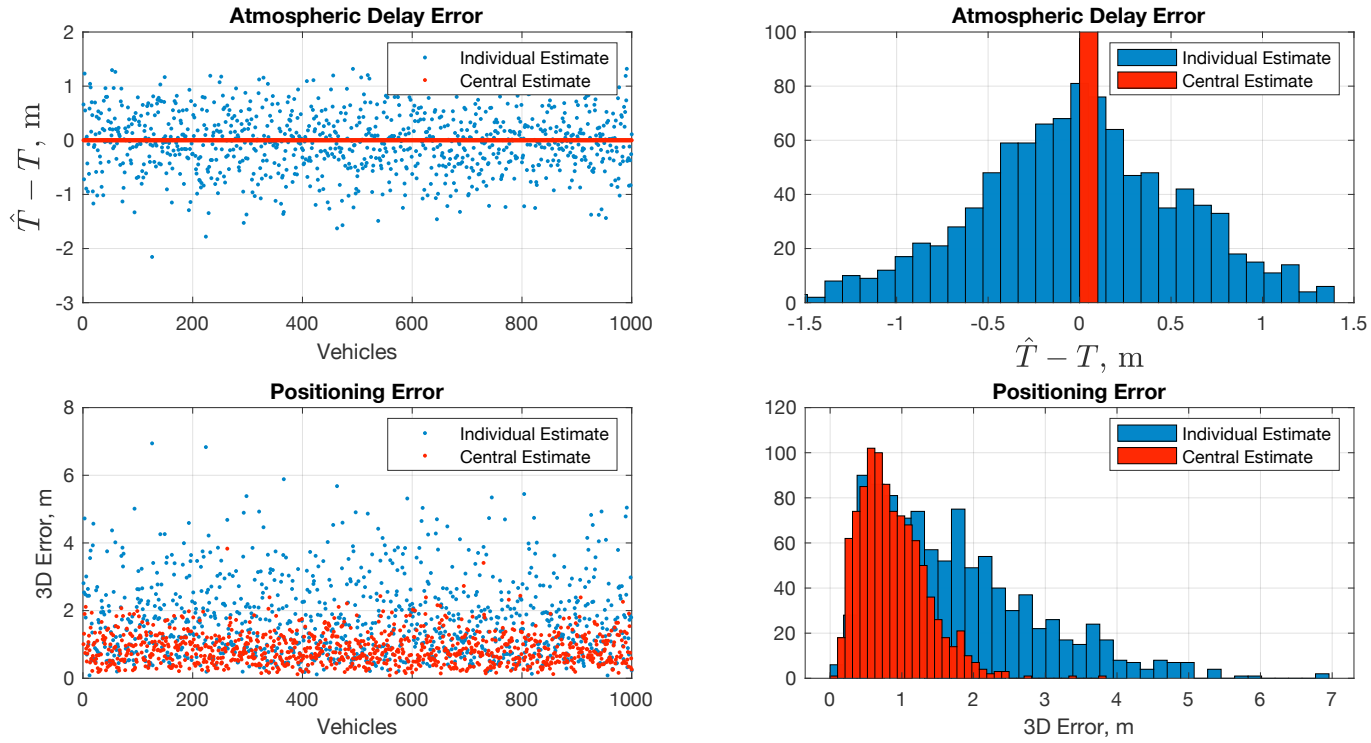


Figure 1: A comparison of simulated individual estimates and the central estimate for $N = 1000$ vehicles.

vehicles remain in the same local area and observe the same number of satellites at all times.

With the model in eqn. (7), we can formulate a centralized maximum-a-posteriori (MAP) solution for an estimator with access to measurements from all nodes. Assume the prior state of $\mathbf{x}$ is available and denote it by $\mathbf{x}_c^- \in \mathbb{R}^q$ with error $\boldsymbol{\eta}_c = \mathbf{x}_c^- - \mathbf{x}$ and $Cov(\eta_c) = \mathbf{P}_c \in \mathbb{R}^{q\times q}$.

Combining eqn. (7) with the prior estimate, we obtain a new model equation

$$\begin{bmatrix} \mathbf{x}_c^- \\ \boldsymbol{\mathcal{Y}} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_q \\ \boldsymbol{\mathcal{H}} \end{bmatrix} \mathbf{x} + \begin{bmatrix} \boldsymbol{\eta}_c \\ \boldsymbol{\nu} \end{bmatrix}, \tag{9}$$

where $\mathbf{I}_q \in \mathbb{R}^{q\times q}$ is the identity matrix.

Let $\boldsymbol{\mathcal{Z}} = \begin{bmatrix} \mathbf{x}_c^- \\ \boldsymbol{\mathcal{Y}} \end{bmatrix} \in \mathbb{R}^{m+q}$, $\boldsymbol{\mathcal{H}}_c = \begin{bmatrix} \mathbf{I}_q \\ \boldsymbol{\mathcal{H}} \end{bmatrix} \in \mathbb{R}^{(m+q)\times q}$, and $\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\eta}_c \\ \boldsymbol{\nu} \end{bmatrix} \in \mathbb{R}^{m+q}$. Then we can express (9) as

$$\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{H}}_c \mathbf{x} + \boldsymbol{\beta}, \tag{10}$$

where $\boldsymbol{\beta} \sim \mathcal{N}(0, \boldsymbol{\mathcal{C}})$. The prior error is assumed to be uncorrelated to the measurement noise and the covariance matrix is $\boldsymbol{\mathcal{C}} = diag(\mathbf{P}_c, \boldsymbol{\mathcal{R}})$. Let $\mathbf{J}_c^- = (\mathbf{P}_c)^{-1} \in \mathbb{R}^{q\times q}$ be the prior information matrix. Then the information matrix combining the prior and the measurements is $\boldsymbol{\mathcal{J}} = \boldsymbol{\mathcal{C}}^{-1} \in \mathbb{R}^{(m+q)\times(m+q)}$ and $\boldsymbol{\mathcal{J}} = diag(\mathbf{J}_c^-, \boldsymbol{\mathcal{B}})$.

The Centralized MAP estimate of $\mathbf{x}$ is

$$\mathbf{x}_c^+ = (\boldsymbol{\mathcal{H}}_c^\top \boldsymbol{\mathcal{J}} \boldsymbol{\mathcal{H}}_c)^{-1}(\boldsymbol{\mathcal{H}}_c^\top \boldsymbol{\mathcal{J}} \boldsymbol{\mathcal{Z}}) \tag{11a}$$

$$= (\mathbf{J}_c^- + \boldsymbol{\mathcal{H}}^\top \boldsymbol{\mathcal{B}} \boldsymbol{\mathcal{H}})^{-1}(\mathbf{J}_c^- \mathbf{x}_c^- + \boldsymbol{\mathcal{H}}^\top \boldsymbol{\mathcal{B}} \boldsymbol{\mathcal{Y}}) \tag{11b}$$

$$\mathbf{J}_c^+ = \mathbf{J}_c^- + \boldsymbol{\mathcal{H}}^\top \boldsymbol{\mathcal{B}} \boldsymbol{\mathcal{H}} \tag{11c}$$

In eqn. (11c), the term $\boldsymbol{\mathcal{H}}^\top \boldsymbol{\mathcal{B}} \boldsymbol{\mathcal{H}}$ represents the information about $\mathbf{x}$ brought by new measurements, and $\mathbf{J}_c^+ = \left(Cov(\mathbf{x}_c^+)\right)^{-1}$ is the information about $\mathbf{x}$ in $\mathbf{x}_c^+$. We note that eqn. (11c) is *similar* to eqn. (5) when performing the individual solution.

Figure 1 shows a comparison of individual solutions and the centralized solution in a simulated environment of $N = 2000$ vehicles. The atmospheric delay estimate significantly improves in the centralized approach. It can be shown that as long as $N > 1$, the information about each vehicle's state is enhanced when using the centralized estimator.

4 MAP Estimation of Atmospheric Delay

In the previous section, to obtain a centralized solution of the full state, each node had so send both its measurement vector and state estimate to a central node, which would then reply to each agent with the corresponding posterior estimate. Since it is impractical to estimate the full state (both the unique and common variables) centrally, this section is concerned with marginalizing and only estimating the common variable.

4.1 Centralized Atmospheric Delay Estimation

In section 3.2, we formulated a centralized MAP solution for the full-state. In this section we formulate a centralized solution for only the atmospheric delay variable. We present the solution in two different parametrizations; namely the covariance form and information form.

4.1.1 Covariance Form

In this centralized framework, each vehicle sends an estimate T_v^+ and a covariance $\mathbf{P}_{T_v}^+$ from eqns. (4)–(5). These can be modeled as $T_v^+ = T^+ + \rho_v$, where $\rho_v \sim \mathcal{N}(0, \mathbf{P}_{T_v}^+)$ is the estimation error with $\mathbf{P}_{T_v}^+$ as the covariance. We can the write the concatenated vector of estimates as

$$\mathbf{T}^+ = \mathbf{H}\,T + \boldsymbol{\rho}, \tag{12}$$

where $\mathbf{T}^+ = [{T_1}^+, \ldots, {T_N}^+]^\top$, $\mathbf{H} = [1, \ldots, 1]^\top$, $\boldsymbol{\rho} = [\rho_1, \ldots, \rho_N]$, and $\boldsymbol{\rho} \sim \mathcal{N}(0, \mathbf{P}_T^+)$ with $\mathbf{P}_T^+ \in \mathbb{R}^{N\times N}$— a block diagonal matrix of $\mathbf{P}_{T_v}^+$ matrices. The MAP optimal solution to (12) is

$$\begin{aligned}\hat{T}^+ &= \big(\mathbf{H}^\top (\mathbf{P}_T^+)^{-1}\mathbf{H}\big)^{-1}\big(\mathbf{H}^\top (\mathbf{P}_T^+)^{-1}\,T^+\big)\\ &= \Big(\sum_{v=1}^{N}(\mathbf{P}_{T_v}^+)^{-1}\Big)^{-1}\Big(\sum_{v=1}^{N}(\mathbf{P}_{T_v}^+)^{-1}T_v^+\Big).\end{aligned} \tag{13}$$

4.1.2 Information Form

We recall that $\mathbf{x}_v = [\mathbf{z}_v, T]^\top$, where $\mathbf{z}_v$ comprises of the unique variables of the state vector for vehicle v, and T is the atmospheric delay variable. Let the submatrices of the individual posterior information matrix for vehicle v from eqn. (5) be written as

$$\mathbf{J}_v^+ = \left[\begin{array}{c:c} \mathbf{B}^+ & b^+ \\ \hdashline (b^+)^\top & c^+ \end{array}\right], \tag{14}$$

where $\mathbf{B}^+ \in \mathbb{R}^{(p-1)\times(p-1)}$, $b^+ \in \mathbb{R}^{p-1}$ and $c^+ \in \mathbb{R}$. (For brevity, the subscript v has been omitted in the matrix components.) The corresponding information vector is

$$\mathbf{j}_v^+ \triangleq \mathbf{J}_v^+\hat{\mathbf{x}}_v^+ = \left[\begin{array}{c} \mathbf{j}_{v_z}^+ \\ \hdashline \mathbf{j}_{v_T}^+ \end{array}\right] = \left[\begin{array}{c} \mathbf{B}^+\mathbf{z}_v^+ + b^+T_v^+ \\ \hdashline (b^+)^\top\mathbf{z}_v^+ + c^+T_v^+ \end{array}\right]. \tag{15}$$

Using this *information form* parametrization, we can perform marginalization via Schur complement [21] to obtain the information submatrix[1] corresponding only to T. That is,

$$\mathbf{J}_{T_v}^+ = c^+ - (b^+)^\top(\mathbf{B}^+)^{-1}b^+. \tag{16}$$

Using (16), the information vector is

$$\mathbf{j}_{T_v}^+ = \mathbf{J}_{T_v}^+ T_v^+. \tag{17}$$

The derivation of the information vector expression is given in Appendix B. We note that the above is the so-called "natural" parametrization of the Gaussian distribution, in which the error in (12) is modeled as $\boldsymbol{\rho} \sim \mathcal{N}^{-1}(0, \mathbf{J}_T^+)$, where $\mathbf{J}_T^+ = \big(\mathbf{P}_T^+\big)^{-1}$.

The solution in (13) can then be written as

$$\hat{T}^+ = \Big(\sum_{v=1}^{N}\mathbf{J}_{T_v}^+\Big)^{-1}\Big(\sum_{v=1}^{N}\mathbf{j}_{T_v}^+\Big), \tag{18}$$

by using the expressions of (17) and (16). In other words, the MAP solution to (12) can be found by averaging the information-weighted estimates from all vehicles.

[1] In this case, both the information "matrix" and "vector" are scalars since T is itself a scalar.

The centralized approach requires sending data to a central node, which could fail; and the computation complexity directly depends on the number of nodes in the network. Furthermore, solution (13) does not take advantage of the fact that messages can be easily broadcast from one node to its neighbors [22].

4.2 Distributed Information-Weighted Average Consensus Estimate

Having found a centralized estimate, we discuss a way to compute it in a distributed framework. We note that the solution in eqns. (13) and (18) can be solved in a distributed manner. Define

$$D = \frac{1}{N}\Big(\sum_{v=1}^{N}\mathbf{J}_{T_v}^+\Big) = \frac{1}{N}\sum_{v=1}^{N}D_v, \tag{19}$$

where $D_v = \mathbf{J}_{T_v}^+$, and $D, D_v \in \mathbb{R}$. Similarly, define

$$d = \frac{1}{N}\Big(\sum_{v=1}^{N}\mathbf{j}_{T_v}^+\Big) = \frac{1}{N}\sum_{v=1}^{N}d_v, \tag{20}$$

where $d_v = \mathbf{j}_{T_v}^+$, and $d, d_v \in \mathbb{R}$. Then (18) can be written as

$$\hat{T}^+ = D^{-1}d. \tag{21}$$

The dimension of D does not depend on the number of nodes N, i.e. as the dimension of the problem grows (more nodes added to the network, for example), the requirements for storing D and D_v do not change. The same applies to d and d_v.

An alternative to the centralized approach is to use the *average consensus algorithm* [23], [24]. With this approach, each node maintains estimates of both D_v and d_v, which we now refer to by the dummy variable x_v, initialized to $x_v(0) = D_v$ and $x_v(0) = d_v$, respectively. Each node can exchange information only with its neighbors, and solve for both D and d by iterating the discrete-time difference equations:

$$x_v(k+1) = x_v(k) + \epsilon\sum_{w\in\aleph_v}a_{vw}\big(x_w(k) - x_v(k)\big), \tag{22}$$

where $0 < \epsilon < 1/\Delta$. For analysis, eqn. (22) can also be written as

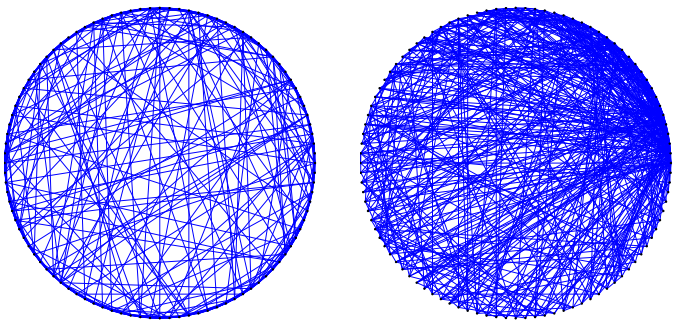
$$x(k+1) = \mathbf{P}x(k), \tag{23}$$

with $\mathbf{P} = \mathbf{I} - \epsilon\mathbf{L}$, where $\mathbf{I}$ is the identity matrix, and $\mathbf{L}$ is the Laplacian matrix defined in eqn. (51) (see Appendix A for graph theory preliminaries). In [23], it is shown that each vehicle converges to the average of initial values, i.e. a consensus is asymptotically reached.

The algorithm for performing an information-weighted consensus (IWC) is presented in Algorithm 1.

4.3 Simulation and Discussion

We set up a network of 100 nodes and three topologies: the small-world topology [25], scale-free topology [26], and the basic ring topology. The small-world and scale-free topologies are shown in Figure 2. These topologies are briefly discussed in Appendix A. The consensus speed parameter



(a) Small-world (b) Scale-free

Figure 2: Network topologies both with 100 nodes, minimum degree of 5, and average degree of around 10. (a): A small-world network with 500 connections, maximum degree of 16, and median degree of 10. (b): A scale-free network with 485 connections, maximum degree of 47, and median degree of 7.

Algorithm 1 Performing Information-Weighted Distributed Average Consensus on T at node v.

Input: Prior estimate, $\mathbf{x}_v^-$, prior information matrix $\mathbf{J}_v^-$, consensus speed, ϵ, tolerance, τ, and maximum consensus iterations, K.

Output: Estimate $\hat{T}_v^+$ and information matrix $\mathbf{J}_{T_v}^+$.

1: Initialize consensus proposals:

$$D_v^0 = c^+ - (b^+)^\top (\mathbf{B}^+)^{-1} b^+,$$
$$d_v^0 = \mathbf{J}_{T_v}^+ T_v^+.$$

2: **while** $\|\hat{T}_v^+(k+1) - \hat{T}_v^+(k)\| > \tau$ & $k < K$ **do**

3: Send D_v^k and d_v^k to all neighbors $w \in \aleph_v$;

4: Receive D_v^k and d_v^k from all neighbors $w \in \aleph_v$;

5: Update using (22):

$$D_v^{k+1} = D_v^k + \epsilon \sum_{w \in \aleph_v} (D_w^k - D_v^k)$$
$$d_v^{k+1} = d_v^k + \epsilon \sum_{w \in \aleph_v} (d_w^k - d_v^k)$$

6: k = k+1

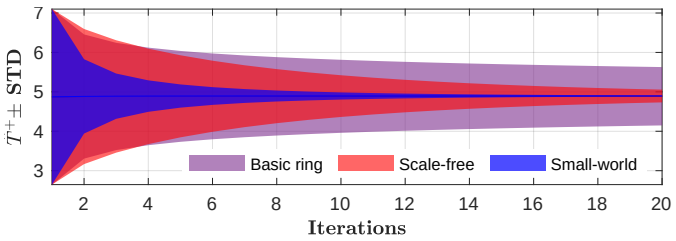
7: **end while**

8: Compute the distributed estimate $\hat{T}_v^+$ by

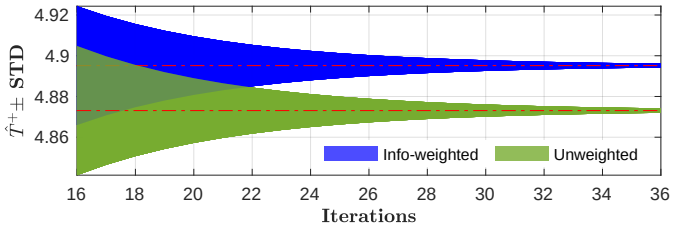
$$\hat{\mathbf{T}}_v^+ = (D)^{-1} d$$

is set to $\epsilon = 1/\Delta$, where Δ is the maximum degree of the network. As shown in Figure 9h, the IWC approach converges to the central estimate for all interaction topologies, but with different convergence speeds. Despite having the largest maximum degree, the scale-free network converges slower than the small-world network. The basic ring network, whose nodes all have the same degree of 2, converges at the slowest rate. This is because even though eqn. (22) guarantees convergence for any undirected connected graph [23], [27], but the convergence speed depends on the nature of connectivity in the graph. Graphs with a smaller average path length converge faster.

We get *unweighted consensus* (UC)—which is the usual average consensus—if we use (22) and modify algorithm 1 to perform consensus on the estimates $\hat{T}_v$ instead of the information vector and matrix. Figure 3b shows that both the information-weighted and unweighted consensus approaches converge at comparable rates for the small-world topology. (Other network topologies should also exhibit the same convergence behavior in relation to information weighting.) This is because the convergence rate depends only on the network characteristics, i.e. the Laplacian matrix $\mathbf{L}$. On the other hand, as shown in Figure 4, IWC produces the smallest convergence error, i.e. it converges closer to the centralized estimate of T compared to UC. As shown in eqns. (18) and (21), the information-weighted consensus is guaranteed to converge at the centralized estimate, whereas the unweighted consensus will not. Notably, IWC converges exponentially, as shown in Figure 4. We note that due to the properties of consensus [23], [24], both IWC and UC algorithms are guaranteed to converge, albeit at different values. We note that in the case when an agent joins the network for the first time, the agent will get the most accurate estimate of the atmospheric delay instantly, thereby significantly decreasing its convergence time. This is in sharp contrast to convergence times of several minutes or hours noted in the introduction [15]–[18].



(a) Consensus for different topologies



(b) Weighted and un-weighted consensus

Figure 3: Distributed average consensus for a network with 100 nodes and different topologies. In all plots, shaded area is within one standard deviation from the mean estimate across all nodes.

Due to various constraints, one would expect real-world communication topologies in navigation networks to be similar to ring lattices, or approaching small-world networks when random connections are possible. The construction of networks with high degrees such as scale-free networks is generally not a good idea [23] since they are not robust to issues such as time-delays and communication failures. This is supported by recent findings that scale-free networks are empirically rare in robust natural systems[2] [28], [30].

[2]Incidentally, this remains a vibrant research topic, but a controversially contended one [28], [29].

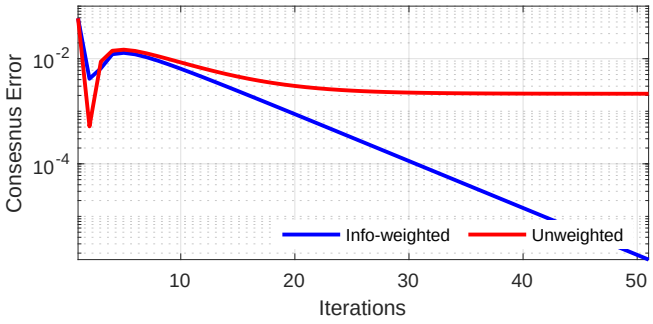


Figure 4: Convergence error for the information-weighted (blue) and the un-weighted (red) consensus algorithms on a vertical log scale for emphasis.

5 HETEROGENEOUS ESTIMATION

In the previous sections it was assumed that all agents are identical in their estimation methods. This led to a further implied assumption that there is some knowledge of the correlation or independence of their priors. In real-world networks, agents are heterogeneous and it is often the case that their prior correlation is unknown. This section removes that assumption and presents algorithms for fusing state estimates with unknown correlation.

Consider two estimates $\hat{\mathbf{x}}_v, \hat{\mathbf{x}}_w$ of $\mathbf{x} \in \mathbb{R}^p$ from two neighboring vehicles, v and w. These estimates are assumed to have a Gaussian distribution of the form

$$\mathbf{x} \sim \mathcal{N}(\hat{\mathbf{x}}_v, \mathbf{P}_v) \quad \text{and} \quad \mathbf{x} \sim \mathcal{N}(\hat{\mathbf{x}}_w, \mathbf{P}_w), \tag{24}$$

respectively, where $\mathbf{P}_v, \mathbf{P}_w \in \mathbb{R}^{p\times p}$ are their respective covariances. For convenience, the distributions of the above estimates will be referred to as $\mathbf{x}_v$ and $\mathbf{x}_w$, respectively. If vehicle w communicates its estimate to v, the goal is to fuse these two estimates *at* v to obtain a fused estimate $\hat{\mathbf{x}}_f$ modeled as

$$\mathbf{x}_f \sim \mathcal{N}(\hat{\mathbf{x}}_f, \mathbf{P}_f). \tag{25}$$

It should be the case that $\hat{\mathbf{x}}_f$ is more accurate than either $\hat{\mathbf{x}}_v$ and $\hat{\mathbf{x}}_w$. In other words,

Property 5.1. *The fusion of estimates $\hat{\mathbf{x}}_v$ and $\hat{\mathbf{x}}_w$ gives an estimate $\hat{\mathbf{x}}_f$ such that $\mathbf{P}_f \leq \mathbf{P}_v$ and $\mathbf{P}_f \leq \mathbf{P}_w$.* △

It is also desired that the fused estimate be *consistent* as defined below.

Definition 5.2 (Consistency of the Fused Estimate [31], [32])**.** *A fused estimate $\mathbf{x}_f \sim \mathcal{N}(\hat{\mathbf{x}}_f, \mathbf{P}_f)$ is consistent if*

$$\mathbf{P}_f \geq \breve{\mathbf{P}}_f,$$

where $\breve{\mathbf{P}}_f = E\langle \tilde{\mathbf{x}}_f \tilde{\mathbf{x}}_f^\top \rangle$ with $\tilde{\mathbf{x}}_f \triangleq \mathbf{x} - \hat{\mathbf{x}}_f$. The actual covariance matrix $\breve{\mathbf{P}}_f$ is unknown and $\mathbf{P}_f$ is its estimate. △

In other words, the estimated covariance is bounded below by the actual error covariance. This implies that the estimate $\hat{\mathbf{x}}_f$ is consistent if its stated level of performance, quantified by the information matrix $\mathbf{J}_f = \mathbf{P}_f^{-1}$ is smaller than the actual information matrix $\breve{\mathbf{J}}_f = \breve{\mathbf{P}}_f^{-1}$. As consequence, a consistent estimate is conservative since it consistently overestimates the uncertainty in the underlying distribution.

The ellipsoids corresponding to $\mathbf{P}_v$ and $\mathbf{P}_w$ are denoted by $\mathbf{\Xi}_v$ and $\mathbf{\Xi}_w$, respectively. For instance, an ellipsoid corresponding to $\mathbf{x} \sim \mathcal{N}(\hat{\mathbf{x}}, \mathbf{P})$ is defined as:

$$\mathbf{\Xi}(\hat{\mathbf{x}}, \mathbf{P}) = \left\{ \mathbf{x} \in \mathbb{R}^n \middle| (\mathbf{x} - \hat{\mathbf{x}})^\top \mathbf{P}^{-1} (\mathbf{x} - \hat{\mathbf{x}}) \leq 1 \right\}. \tag{26}$$

5.1 Combining Uncorrelated Estimates

In the case when two Gaussian estimates are independent, hence uncorrelated, we obtain the optimal fusion via the following lemma whose proof can be found in [33], [34].

Lemma 5.3 (Linear Combination of Two Independent Estimates)**.** *Let $\mathbf{x}_v$ and $\mathbf{x}_w$ be two independent unbiased estimates of $\mathbf{x}$, whose means and covariances are $(\hat{\mathbf{x}}_v, \mathbf{P}_v)$ and $(\hat{\mathbf{x}}_w, \mathbf{P}_w)$, respectively. Then their fused estimate is*

$$\begin{aligned} \mathbf{P}_f &= \left(\mathbf{P}_v^{-1} + \mathbf{P}_w^{-1}\right)^{-1}, \\ \hat{\mathbf{x}}_f &= \mathbf{P}_f(\mathbf{P}_v^{-1}\hat{\mathbf{x}}_v + \mathbf{P}_w^{-1}\hat{\mathbf{x}}_w). \end{aligned} \tag{27}$$

This is the unbiased estimate that minimizes $tr(\mathbf{P}_f)$. For convenience of notation in the text that follows, we will denote $\mathbf{x}_f = \mathbf{\Phi}(\mathbf{x}_v, \mathbf{x}_w)$, where $\mathbf{x}_f \sim \mathcal{N}(\hat{\mathbf{x}}_f, \mathbf{P}_f)$. △

The above is the *optimal* fusion when $\mathbf{x}_v$ and $\mathbf{x}_w$ are independent. Otherwise, and often in practice, the scheme in eqn. (27) is referred to as *naïve fusion*, since it completely disregards possible correlations due to shared information. We now derive some methods of fusing two correlated prior estimates.

5.2 Combining Estimates With Known Correlation

The following considers a situation when the correlation, and hence mutual information, between the estimates is known. To account for the known correlation, we use $\mathbf{\Phi}$ defined in Lemma 5.3 to represent $\mathbf{x}_v$ and $\mathbf{x}_w$ as

$$\mathbf{x}_v = \mathbf{\Phi}(\mathbf{x}_{vv}, \mathbf{x}_{vw}) \text{ and } \mathbf{x}_w = \mathbf{\Phi}(\mathbf{x}_{ww}, \mathbf{x}_{vw}). \tag{28}$$

In this expression,

$$\mathbf{x}_{vv} \sim \mathcal{N}(\boldsymbol{\theta}_v, \mathbf{\Theta}_v), \quad \mathbf{x}_{ww} \sim \mathcal{N}(\boldsymbol{\theta}_w, \mathbf{\Theta}_w), \quad \mathbf{x}_{vw} \sim \mathcal{N}(\boldsymbol{\gamma}, \mathbf{\Gamma}) \tag{29}$$

where $\boldsymbol{\theta}_v, \boldsymbol{\theta}_w, \boldsymbol{\gamma} \in \mathbb{R}^p$ and $\mathbf{\Theta}_v, \mathbf{\Theta}_w, \mathbf{\Gamma} > \mathbf{0} \in \mathbb{R}^{p\times p}$. The quantities $\mathbf{x}_{vv}$ and $\mathbf{x}_{ww}$ represent exclusive information used only in one estimate, whereas $\mathbf{x}_{vw}$ represents the known shared information. Additionally, Assumption 5.4 is required to enable the use of eqn. (27).

Assumption 5.4 (Pair-wise Independence)**.** *The random variables $\mathbf{x}_{vv}, \mathbf{x}_{ww}, \mathbf{x}_{vw}$ are pair-wise independent.* △

Consequently, via eqn. (28), the definition of $\mathbf{\Phi}$ in eqn. (27) and Lemma 5.3, we get

$$\begin{aligned} \mathbf{P}_v^{-1} &= \mathbf{\Theta}_v^{-1} + \mathbf{\Gamma}^{-1}, \\ \hat{\mathbf{x}}_v &= \mathbf{P}_v(\mathbf{\Theta}_v^{-1}\boldsymbol{\theta}_v + \mathbf{\Gamma}^{-1}\boldsymbol{\gamma}) \end{aligned} \tag{30}$$

and

$$\begin{aligned} \mathbf{P}_w^{-1} &= \mathbf{\Theta}_w^{-1} + \mathbf{\Gamma}^{-1}, \\ \hat{\mathbf{x}}_w &= \mathbf{P}_w(\mathbf{\Theta}_w^{-1}\boldsymbol{\theta}_w + \mathbf{\Gamma}^{-1}\boldsymbol{\gamma}). \end{aligned} \tag{31}$$

Because of Assumption 5.4, fusing $\mathbf{x}_v$ and $\mathbf{x}_w$ is equivalent to fusing $\mathbf{x}_{vv}$, $\mathbf{x}_{ww}$, and $\mathbf{x}_{vw}$ via a three-way linear combination due to

Lemma 5.5 (Three-way Linear Combination of Independent Estimates)**.** *Assume $\mathbf{x}_{vv}$, $\mathbf{x}_{ww}$, and $\mathbf{x}_{vw}$ of eqn.* (29) *are independent. Then, the optimal linear combination that minimizes tr$(\mathbf{P}_f)$ is:*

$$\mathbf{P}_f^{-1} = \mathbf{\Theta}_v^{-1} + \mathbf{\Theta}_w^{-1} + \mathbf{\Gamma}^{-1} \tag{32a}$$

$$\hat{\mathbf{x}}_f = \mathbf{P}_f \left(\mathbf{\Theta}_v^{-1}\boldsymbol{\theta}_v + \mathbf{\Theta}_w^{-1}\boldsymbol{\theta}_w + \mathbf{\Gamma}^{-1}\boldsymbol{\gamma}\right). \tag{32b}$$

Proof. Since $\mathbf{x}_{vv}$, $\mathbf{x}_{ww}$, and $\mathbf{x}_{vw}$ are all independent, the result follows by applying Lemma 5.3 recursively. □

Solving (30) and (31) for $\mathbf{\Theta}_v$ and $\mathbf{\Theta}_w$ and substituting into (32) yields:

$$\mathbf{P}_f = (\mathbf{P}_v^{-1} + \mathbf{P}_w^{-1} - \mathbf{\Gamma}^{-1})^{-1} \tag{33a}$$

$$\hat{\mathbf{x}}_f = \mathbf{P}_f\,(\mathbf{P}_v^{-1}\hat{\mathbf{x}}_v + \mathbf{P}_w^{-1}\hat{\mathbf{x}}_w - \mathbf{\Gamma}^{-1}\boldsymbol{\gamma}). \tag{33b}$$

This shows that the fused estimate and its covariance can be determined when the quantities $\boldsymbol{\gamma}$ and $\mathbf{\Gamma}$ are known.

Furthermore, we can use Lemma 5.3 to characterize the covariance of $\mathbf{x}_v$ and $\mathbf{x}_w$, $\mathrm{Cov}(\mathbf{x}_v, \mathbf{x}_w)$. To this end, the estimation errors are

$$\begin{aligned}\tilde{\mathbf{x}}_v = \hat{\mathbf{x}}_v - \mathbf{x}_v &= \mathbf{P}_v\left(\mathbf{\Theta}_v^{-1}\boldsymbol{\theta}_v + \mathbf{\Gamma}^{-1}\boldsymbol{\gamma}\right) - \mathbf{x}_v \\ &= \mathbf{P}_v\left(\mathbf{\Theta}_v^{-1}(\boldsymbol{\theta}_v - \mathbf{x}_v) + \mathbf{\Gamma}^{-1}(\boldsymbol{\gamma} - \mathbf{x}_v)\right) \\ &= \mathbf{P}_v\left(\mathbf{\Theta}_v^{-1}\tilde{\boldsymbol{\theta}}_v + \mathbf{\Gamma}^{-1}\tilde{\boldsymbol{\gamma}}\right)\end{aligned}$$

and similarly

$$\tilde{\mathbf{x}}_w = \mathbf{P}_w\left(\mathbf{\Theta}_w^{-1}\tilde{\boldsymbol{\theta}}_w + \mathbf{\Gamma}^{-1}\tilde{\boldsymbol{\gamma}}\right).$$

Since $E\langle\tilde{\boldsymbol{\theta}}_v\tilde{\boldsymbol{\theta}}_w^\top\rangle = E\langle\tilde{\boldsymbol{\theta}}_v\tilde{\boldsymbol{\gamma}}^\top\rangle = E\langle\tilde{\boldsymbol{\theta}}_w\tilde{\boldsymbol{\gamma}}^\top\rangle = 0$, we obtain the following expression:

$$\begin{aligned}\mathrm{Cov}(\hat{\mathbf{x}}_v, \hat{\mathbf{x}}_w) &= E\langle\tilde{\mathbf{x}}_v\tilde{\mathbf{x}}_w^\top\rangle \\ &= E\langle\mathbf{P}_v(\mathbf{\Gamma}^{-1}\tilde{\boldsymbol{\gamma}} + \mathbf{\Theta}_v^{-1}\tilde{\boldsymbol{\theta}}_v)(\tilde{\boldsymbol{\gamma}}^\top\mathbf{\Gamma}^{-1} + \tilde{\boldsymbol{\theta}}_w^\top\mathbf{\Theta}_w^{-1})\mathbf{P}_w\rangle \\ &= \mathbf{P}_v\,E\langle\mathbf{\Gamma}^{-1}\tilde{\boldsymbol{\gamma}}\tilde{\boldsymbol{\gamma}}^\top\mathbf{\Gamma}^{-1}\rangle\,\mathbf{P}_w \\ &= \mathbf{P}_v\,\mathbf{\Gamma}^{-1}\mathbf{P}_w.\end{aligned} \tag{34}$$

Thus, the covariance of the two estimates $\mathbf{x}_v$ and $\mathbf{x}_w$ can be modeled explicitly in terms of the mutual information, $\mathbf{\Gamma}^{-1}$.

Eqn. (33) is the optimal fusion of two unbiased estimates when mutual covariance $\mathbf{\Gamma}$ and covariance matrices $\mathbf{P}_v$ and $\mathbf{P}_w$ are known. However, in some applications, the mutual covariance $\mathbf{\Gamma}$ is not known. We next consider three methods to handle the situation when the quantities $\boldsymbol{\gamma}$, $\mathbf{\Gamma}$ are not known.

5.3 Combining Estimates With Unknown Correlation

In [31], it shown that $\mathbf{\Xi}_f \subset \mathbf{\Xi}_v \cap \mathbf{\Xi}_w$ for any choice of $\mathbf{\Gamma}$. Section 5.3.1 finds a value of $\mathbf{\Gamma}$ corresponding to the largest $\mathbf{\Xi}_f$ such that $\mathbf{\Xi}_f \subset \mathbf{\Xi}_v \cap \mathbf{\Xi}_w$, and a value of $\boldsymbol{\gamma}$ that minimizes its distance from both $\hat{\mathbf{x}}_v$ and $\hat{\mathbf{x}}_w$. However, this formulation is found not to be consistent for all possible values of $\mathbf{\Gamma}$. Section 5.3.2 finds the optimal consistent fusion scheme for unknown $\mathbf{\Gamma}$ by finding the smallest $\mathbf{\Xi}_f$ such that $\mathbf{\Xi}_v \cap \mathbf{\Xi}_w \subset \mathbf{\Xi}_f$. Finally, Section 5.3.3 finds a consistent fusion scheme such that $\mathbf{\Xi}_v \cap \mathbf{\Xi}_w \subset \mathbf{\Xi}_f$ and $\mathbf{\Xi}_f \subset \mathbf{\Xi}_v \cup \mathbf{\Xi}_w$.

5.3.1 Ellipsoidal Intersection

In some applications, actual values for quantities $\boldsymbol{\gamma}$ and $\mathbf{\Gamma}$ in eqn. (33) are unknown, but could be selected using the previous assumption of pair-wise independence. The goal of this section is to find values of $\boldsymbol{\gamma}$ and $\mathbf{\Gamma}$ satisfying the constraints of the problem.

Selecting $\mathbf{\Gamma}$: The objective is to find a value of $\mathbf{\Gamma}$ that maximizes the covariance $\mathbf{P}_f$ of eqn. (33a). Maximizing $\mathbf{P}_f$ is equivalent to maximizing the mutual information $\mathbf{\Gamma}^{-1}$, and since $\mathbf{\Gamma} > 0$, this is equivalent to minimizing the determinant of the mutual covariance, $|\mathbf{\Gamma}|$. Additionally, since $\mathbf{\Theta}_w^{-1} = \mathbf{P}_w^{-1} - \mathbf{\Gamma}^{-1} \geq 0$, we have $\mathbf{\Gamma} \geq \mathbf{P}_w$. Similarly, we also have $\mathbf{\Gamma} \geq \mathbf{P}_v$. Hence, we define $\check{\mathbf{\Gamma}}$ as the solution to the following optimization problem [33], [34];

$$\begin{aligned}\check{\mathbf{\Gamma}} \doteq \arg\min_{\mathbf{\Upsilon}\in\mathbb{R}^{p\times p}} &\log|\mathbf{\Upsilon}|. \\ \text{subject to} &\mathbf{\Upsilon} \geq \mathbf{P}_v,\ \mathbf{\Upsilon} \geq \mathbf{P}_w.\end{aligned} \tag{35}$$

The ellipsoid $\mathbf{\Xi}_{\check{\mathbf{\Gamma}}}$ is the *Löwner–John* ellipsoid of $\mathbf{\Xi}_v \cup \mathbf{\Xi}_w$, defined as the minimum volume ellipsoid that contains both ellipsoids of $\mathbf{P}_v$ and $\mathbf{P}_w$, i.e. $\mathbf{\Xi}_{\check{\mathbf{\Gamma}}} \subseteq \mathbf{\Xi}_v \cup \mathbf{\Xi}_w$ [35]. We subsequently use the result from [36] to compute $\mathbf{\Xi}_{\check{\mathbf{\Gamma}}}$.

First, we show that eqn. (35) is invariant to linear transformations represented by a non-singular matrix $\mathbf{T} \in \mathbb{R}^{p\times p}$. We use Theorem C.1 in Appendix C.1 to define the diagonalized covariance matrices as[3]

$$\overline{\mathbf{P}}_v = \mathbf{T}^{-1}\mathbf{P}_v\mathbf{T}^{-\top} \quad \text{and} \quad \overline{\mathbf{P}}_w = \mathbf{T}^{-1}\mathbf{P}_w\mathbf{T}^{-\top} \tag{36}$$

where $\mathbf{T}^{-\top} = (\mathbf{T}^{-1})^\top = (\mathbf{T}^\top)^{-1}$. Then, eqn. (35) becomes

$$\begin{aligned}\overline{\mathbf{\Gamma}} \doteq \arg\min_{\overline{\mathbf{\Upsilon}}\in\mathbb{R}^{p\times p}} &\log|\overline{\mathbf{\Upsilon}}|, \\ \text{subject to } \overline{\mathbf{\Upsilon}} \geq \overline{\mathbf{P}}_v,&\quad \overline{\mathbf{\Upsilon}} \geq \overline{\mathbf{P}}_w\end{aligned} \tag{37}$$

where $\overline{\mathbf{\Upsilon}} = \mathbf{T}^{-1}\mathbf{\Upsilon}\mathbf{T}^{-\top}$ and $\overline{\mathbf{\Gamma}} = \mathbf{T}^{-1}\mathbf{\Gamma}\mathbf{T}^{-\top}$. We note that the ellipsoids corresponding to $\overline{\mathbf{P}}_v$ and $\overline{\mathbf{P}}_w$ are aligned, thus making the constraints of (37) easier to satisfy compared to those of (35). Lemma 5.6 shows that the solution to eqn. (35) is equivalent to the inverse linear transformation of the solution to eqn. (37), then Theorem 5.7 shows how to find $\overline{\mathbf{\Gamma}}$.

Lemma 5.6. *For $\mathbf{P}_v$, $\mathbf{P}_w$ and $\overline{\mathbf{P}}_v$, $\overline{\mathbf{P}}_w$, $\overline{\mathbf{\Gamma}}$ satisfying eqn.* (36)*, with some non-singular $\mathbf{T} \in \mathbb{R}^{p\times p}$, the solution to* (35) *via* (37) *is*

$$\check{\mathbf{\Gamma}} = \mathbf{T}\,\overline{\mathbf{\Gamma}}\,\mathbf{T}^\top.$$

Proof. Define $\overline{\mathbf{\Upsilon}} = \mathbf{T}^{-1}\,\mathbf{\Upsilon}\,\mathbf{T}^{-\top}$. Note $\mathbf{P}_v = \mathbf{T}\,\overline{\mathbf{P}}_v\,\mathbf{T}^\top$ and $\mathbf{\Upsilon} = \mathbf{T}\,\overline{\mathbf{\Upsilon}}\,\mathbf{T}^\top$. Then, if $\overline{\mathbf{\Upsilon}} \geq \overline{\mathbf{P}}_v$, we get $(\overline{\mathbf{\Upsilon}} - \overline{\mathbf{P}}_v)^{1/2} \geq 0$, which implies that $\mathbf{T}(\overline{\mathbf{\Upsilon}} - \overline{\mathbf{P}}_v)\mathbf{T}^\top \geq 0$, which further implies $\mathbf{\Upsilon} \geq \mathbf{P}_v$. Similarly, if $\overline{\mathbf{\Upsilon}} \geq \overline{\mathbf{P}}_w$, then $\mathbf{\Upsilon} \geq \mathbf{P}_w$.

Furthermore, for any non-singular matrices $\mathbf{A}$, $\mathbf{B}$, we have $|\mathbf{A}\,\mathbf{B}| = |\mathbf{A}|\,|\mathbf{B}|$ (see Proposition 2.7.3 in [37]). There-

[3]This is to transform both covariance matrices so that the major and minor semi-axes of their corresponding ellipsoids are aligned.

fore, $\log|\boldsymbol{\Upsilon}| = \log(|\mathbf{T}\,\overline{\boldsymbol{\Upsilon}}\,\mathbf{T}^\top|) = \log|\overline{\boldsymbol{\Upsilon}}| + 2\log|\mathbf{T}|$. This implies that

$$\begin{aligned}\overline{\boldsymbol{\Gamma}} &= \arg\min_{\overline{\boldsymbol{\Upsilon}}\in\mathbb{R}^{p\times p}} \log|\overline{\boldsymbol{\Upsilon}}| \\ &\quad \text{subject to } \overline{\boldsymbol{\Upsilon}} \geq \overline{\mathbf{P}}_v,\ \overline{\boldsymbol{\Upsilon}} \geq \overline{\mathbf{P}}_w \\ &= \arg\min_{\boldsymbol{\Upsilon}\in\mathbb{R}^{p\times p}} \log|\boldsymbol{\Upsilon}| \\ &\quad \text{subject to } \boldsymbol{\Upsilon} \geq \mathbf{P}_v,\ \boldsymbol{\Upsilon} \geq \mathbf{P}_w.\end{aligned}$$

Hence,

$$\begin{aligned}\mathbf{T}\,\overline{\boldsymbol{\Gamma}}\,\mathbf{T}^\top &= \arg\min_{\boldsymbol{\Upsilon}\in\mathbb{R}^{p\times p}} \log|\boldsymbol{\Upsilon}| \\ &\quad \text{subject to } \boldsymbol{\Upsilon} \geq \mathbf{P}_v,\ \boldsymbol{\Upsilon} \geq \mathbf{P}_w \\ &= \check{\boldsymbol{\Gamma}}.\end{aligned}$$

This completes the proof. □

Theorem 5.7. *For any positive definite $\mathbf{P}_v, \mathbf{P}_w \in \mathbb{R}^{p\times p}$,*

(i) There exists a non-singular matrix $\mathbf{T} \in \mathbb{R}^{p\times p}$ such that

$$\begin{aligned}\overline{\mathbf{P}}_v &= \mathbf{T}^{-1}\mathbf{P}_v\mathbf{T}^{-\top} = \mathbf{I}_p \\ \overline{\mathbf{P}}_w &= \mathbf{T}^{-1}\mathbf{P}_w\mathbf{T}^{-\top} = \mathbf{D}_w\end{aligned}$$

where $\mathbf{D}_w > 0$ is diagonal, and $\mathbf{I}_p$ is the identity matrix.

(ii) The matrix $\overline{\boldsymbol{\Gamma}}$ defined in (37) satisfies $\overline{\boldsymbol{\Gamma}} = \mathbf{D}_\Gamma$, with $\mathbf{D}_\Gamma = \max\{\overline{\mathbf{P}}_v, \overline{\mathbf{P}}_w\} = \max\{\mathbf{I}_p, \mathbf{D}_w\}$, where $\max$ is the component-wise maximum function.

(iii) The matrix $\check{\boldsymbol{\Gamma}}$ defined in (35) satisfies $\check{\boldsymbol{\Gamma}} = \mathbf{T}\,\mathbf{D}_\Gamma\,\mathbf{T}^\top$.

Note that (i) is the simultaneous diagonalization of $\mathbf{P}_v$ and $\mathbf{P}_w$ (defined in Appendix C.1), whereas (ii) is the solution to (37), and (iii) is obtained using Lemma 5.6. The full proof of Theorem 5.7 is given in Appendix C.2.

To find the appropriate the $\mathbf{T}$, consider the eigenvalue decompositions of $\mathbf{P}_v$ defined by

$$\mathbf{S}_v\mathbf{D}_v\mathbf{S}_v^{-1} = \mathbf{P}_v,$$

where $\mathbf{S}_v, \mathbf{D}_v \in \mathbb{R}^{p\times p}$ are rotational and scaling (diagonal) matrices, respectively. To align $\mathbf{P}_w$ with $\mathbf{P}_v$, we apply a linear transformation (rotation followed by scaling) defined by

$$\mathbf{P}_w^\star = \mathbf{D}_v^{-\frac{1}{2}}\mathbf{S}_v^{-1}\mathbf{P}_w\mathbf{S}_v\mathbf{D}_v^{-\frac{1}{2}}$$

whose eigenvalue decomposition is

$$\mathbf{S}_w\mathbf{D}_w\mathbf{S}_w^{-1} = \mathbf{P}_w^\star.$$

Then, $\mathbf{S}_v^{-1}\mathbf{P}_v\mathbf{S}_v$ and $\mathbf{S}_w^{-1}\mathbf{P}_w^\star\mathbf{S}_w$ are diagonal. Specifically,

$$\mathbf{S}_w^{-1}\mathbf{D}_v^{-\frac{1}{2}}\mathbf{S}_v^{-1}\mathbf{P}_w\mathbf{S}_v\mathbf{D}_v^{-\frac{1}{2}}\mathbf{S}_w = \mathbf{D}_w \tag{38}$$

Using Theorem C.1, and comparing (38) and (54), we set the non-singular matrix $\mathbf{T}$ to

$$\mathbf{T} \doteq \mathbf{S}_v\mathbf{D}_v^{\frac{1}{2}}\mathbf{S}_w.$$

Applying the transformations of eqn. (36), we notice that $\overline{\mathbf{P}}_v$ and $\overline{\mathbf{P}}_w$ are diagonal. In fact, $\overline{\mathbf{P}}_v = \mathbf{I}_p$ and $\overline{\mathbf{P}}_w = \mathbf{D}_w$. Consequently, by Theorem 5.7 (ii), $\overline{\boldsymbol{\Gamma}} = \mathbf{D}_\Gamma = \max(\overline{\mathbf{P}}_v, \overline{\mathbf{P}}_w)$ is a diagonal matrix given by

$$[\mathbf{D}_\Gamma]_{qr} \doteq \begin{cases}\max\{1, [\mathbf{D}_w]_{qr}\} & q = r \\ 0 & q \neq r\end{cases}.$$

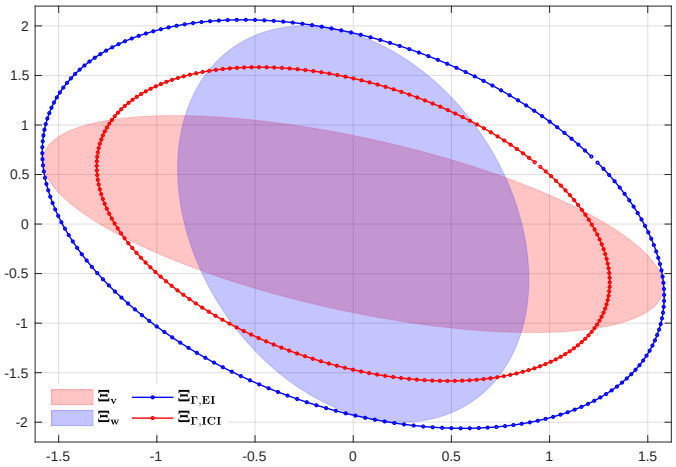


Figure 5: Characterizations of mutual covariances via EI and ICI. Note that $\boldsymbol{\Xi}_{\boldsymbol{\Gamma},EI}$ is the smallest ellipsoid containing $\boldsymbol{\Xi}_v \cup \boldsymbol{\Xi}_w$, whereas $\boldsymbol{\Xi}_{\boldsymbol{\Gamma},ICI} \supset \boldsymbol{\Xi}_v \cap \boldsymbol{\Xi}_w$.

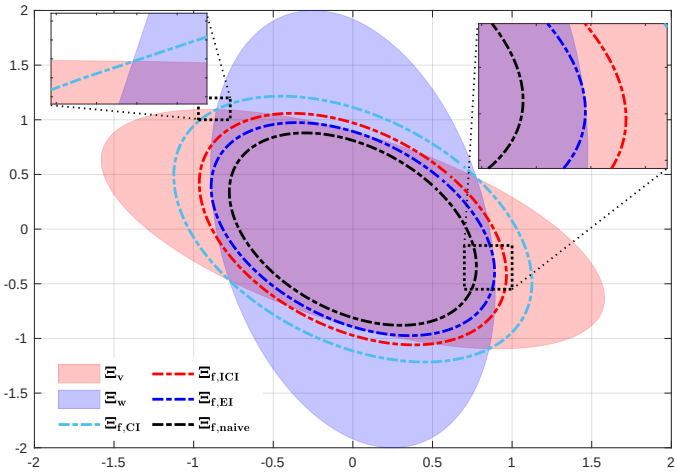


Figure 6: Fused covariance ellipsoids via CI, EI, ICI, and naive fusion.

Therefore, by Theorem 5.7 (iii), the solution to the optimization problem of (35) is

$$\check{\boldsymbol{\Gamma}} = \mathbf{S}_v\mathbf{D}_v^{\frac{1}{2}}\mathbf{S}_w\,\mathbf{D}_\Gamma\,\mathbf{S}_w^{-1}\mathbf{D}_v^{\frac{1}{2}}\mathbf{S}_v^{-1}. \tag{39}$$

As an illustrative example throughout this section, we forgo the previous 1-dimensional example for a 2-dimensional one whose state estimates have covariances

$$\mathbf{P}_v = \begin{bmatrix}2.5 & -1.0\\ -1 & 1.2\end{bmatrix}, \qquad \mathbf{P}_w = \begin{bmatrix}0.8 & -0.5\\ -0.5 & 4.0\end{bmatrix}. \tag{40}$$

Their respective means are: $\hat{\mathbf{x}}_v = [0.5,\ 1]^\top$ and $\hat{\mathbf{x}}_w = [2,\ 1]^\top$. The ellipses corresponding to covariance matrices in (40) are shaded in Figure 6. The resulting $\check{\boldsymbol{\Gamma}}$ using $\mathbf{P}_v$ and $\mathbf{P}_w$ of (40) is:

$$\check{\boldsymbol{\Gamma}} = \begin{bmatrix}2.5 & -1.2\\ -1.2 & 4.3\end{bmatrix}.$$

The corresponding ellipsoid $\boldsymbol{\Xi}_{\check{\Gamma}}$ is shown as the blue curve in Figure 5.

Selecting $\boldsymbol{\gamma}$: Having found an algebraic expression for $\check{\boldsymbol{\Gamma}}$, it is now possible to find an expression for $\check{\boldsymbol{\gamma}}$. The goal is to find $\check{\boldsymbol{\gamma}}$ that minimizes a certain distance of $\boldsymbol{\gamma}$ from both $\hat{\mathbf{x}}_v$ and $\hat{\mathbf{x}}_w$. Since both $\hat{\mathbf{x}}_v$ and $\hat{\mathbf{x}}_w$ have different accuracies, their distances to $\boldsymbol{\gamma}$ should be weighed accordingly. To that end, [33] defines a cost function

$$C(\boldsymbol{\gamma}) \doteq \|\mathbf{W}_v(\boldsymbol{\gamma} - \hat{\mathbf{x}}_v)\|_2^2 + \|\mathbf{W}_w(\boldsymbol{\gamma} - \hat{\mathbf{x}}_w)\|_2^2 \,, \tag{41}$$

where $\mathbf{W}_v, \mathbf{W}_w > \mathbf{0}$ are positive definite weight matrices. Then $\check{\boldsymbol{\gamma}}$ is defined as the value that minimizes the cost function in (41). Specifically,

$$\check{\boldsymbol{\gamma}} \doteq \arg\min_{\boldsymbol{\gamma}\in\mathbb{R}^p} C(\boldsymbol{\gamma}) \tag{42}$$

It is shown in [33] that the unique solution to (42) is

$$\check{\boldsymbol{\gamma}} = (\mathbf{W}_v + \mathbf{W}_w)^{-1}(\mathbf{W}_v\hat{\mathbf{x}}_v + \mathbf{W}_w\hat{\mathbf{x}}_w)\,. \tag{43}$$

Straightforward choices for $\mathbf{W}_v$ and $\mathbf{W}_w$ are the *exclusive information* matrices $\boldsymbol{\Theta}_v^{-1}$ and $\boldsymbol{\Theta}_w^{-1}$, respectively. However, to avoid numerical instabilities, we add $\xi\mathbf{I}_p$ to $\boldsymbol{\Theta}_v^{-1}$ and $\boldsymbol{\Theta}_w^{-1}$, for some $\xi > 0$ to ensure positive definiteness.

Solving (30) and (31) for $\boldsymbol{\Theta}_v^{-1}$ and $\boldsymbol{\Theta}_w^{-1}$ yields:

$$\mathbf{W}_v = \mathbf{P}_v^{-1} - \check{\boldsymbol{\Gamma}}^{-1} + \xi\mathbf{I}_p \text{ and } \mathbf{W}_w = \mathbf{P}_w^{-1} - \check{\boldsymbol{\Gamma}}^{-1} + \xi\mathbf{I}_p. \tag{44}$$

To avoid $\xi\mathbf{I}_p$ potentially having undue effect on $\check{\boldsymbol{\gamma}}$, we choose ξ to be as small as possible. Define $\mathcal{F} = \mathbf{P}_v^{-1} + \mathbf{P}_w^{-1} - 2\check{\boldsymbol{\Gamma}}^{-1}$, and the quantity $\lambda_{0^+}(\mathcal{F})$ to be the smallest positive eigenvalue of $\mathcal{F}$. Then,

$$\xi = \begin{cases} 0 & |\mathcal{F}| \neq 0 \\ c \ll \lambda_{0^+}(\mathcal{F}) & \text{otherwise,} \end{cases}$$

where $c \in \mathbb{R}$ is a small number satisfying $0 < c \ll \lambda_{0^+}(\mathcal{F})$.

When $\hat{\mathbf{x}}_v = [0.5,\ 1]^\top$ and $\hat{\mathbf{x}}_w = [2,\ 1]^\top$, and using the previous result of $\check{\boldsymbol{\Gamma}}$ computed from (39), we obtain $\check{\boldsymbol{\gamma}} = [1.0,\ 1.1]^\top$. These results for $\check{\boldsymbol{\Gamma}}$ and $\check{\boldsymbol{\gamma}}$ are used in eqn. (33) to give the results illustrated in Figure 6 (the blue ellipse). In Figure 6, $\boldsymbol{\Xi}_{f,EI}$ is the largest ellipsoid enclosed by the intersection of $\boldsymbol{\Xi}_v$ and $\boldsymbol{\Xi}_w$.

We note that EI is *not* consistent [32]. This is because the solution in eqn. (39) depends on the decompositions in (30) and (31), which themselves depend on Assumption 5.4. Hence, in cases when Assumption 5.4 is not fulfilled, it is possible that $E\langle\tilde{\mathbf{x}}_f\,\tilde{\mathbf{x}}_f^\top\rangle > \mathbf{P}_f$.

Algorithm 2 summarizes how to perform EI fusion at node v, whose neighbors are $w \in \aleph_v$ and the total number number of neighbors is $L = |\aleph_v|$.

Besides formulating explicit expressions of the mutual information, the fused estimate can also be obtained via convex combinations of the covariances.

5.3.2 Covariance Intersection

Linear combination of covariances using Lemmas 5.3 and 5.5 is the optimal approach for data fusion when the correlation is known (see Section 5.2). For unknown correlation,

Algorithm 2 Performing EI Fusion at node v.

Input: $\hat{\mathbf{x}}_v$, $\mathbf{P}_v$, $\hat{\mathbf{x}}_w$, $\mathbf{P}_w$
Output: $\hat{\mathbf{x}}_f$, $\mathbf{P}_f$.

Initialize:
$\hat{\mathbf{x}}_{v(0)} = \hat{\mathbf{x}}_v, \qquad \mathbf{P}_{v(0)} = \mathbf{P}_v$
for $l = 1, \ldots, L$ **do**
  $\hat{\mathbf{x}}_{w(l)} = \hat{\mathbf{x}}_w, \qquad \mathbf{P}_{w(l)} = \mathbf{P}_w \qquad (w \in \aleph_v)$;
  Compute $\check{\boldsymbol{\Gamma}}_{(l)}$ for $\mathbf{P}_{v(l-1)}$ and $\mathbf{P}_{w(l)}$ using (39);
  Compute $\check{\boldsymbol{\gamma}}_{(l)}$ for $\mathbf{P}_{v(l-1)}, \hat{\mathbf{x}}_{v(l-1)}, \mathbf{P}_{w(l)}, \hat{\mathbf{x}}_{w(l)}$ using (43) and (44);
  $\mathbf{P}_{v(l)} = \left(\mathbf{P}_{v(l-1)}^{-1} + \mathbf{P}_{w(l)}^{-1} - \check{\boldsymbol{\Gamma}}_{(l)}^{-1}\right)^{-1}$;
  $\hat{\mathbf{x}}_{v(l)} = \mathbf{P}_{v(l)}\left(\mathbf{P}_{v(l-1)}^{-1}\hat{\mathbf{x}}_{v(l-1)} + \mathbf{P}_{w(l)}^{-1}\hat{\mathbf{x}}_{w(l)} - \check{\boldsymbol{\Gamma}}_{(l)}^{-1}\check{\boldsymbol{\gamma}}_{(l)}\right)$;
end for
$\hat{\mathbf{x}}_f = \hat{\mathbf{x}}_{v(L)}, \quad \mathbf{P}_f = \mathbf{P}_{v(L)}$;

Julier and Uhlmann [31] proposed the Covariance Intersection (CI) algorithm to fuse estimates via a *convex combination* of their information. Specifically,

$$\mathbf{P}_f = \left(\omega\mathbf{P}_v^{-1} + (1-\omega)\mathbf{P}_w^{-1}\right)^{-1} \tag{45a}$$
$$\mathbf{K} = \omega\mathbf{P}_v^{-1}$$
$$\mathbf{L} = (1-\omega)\mathbf{P}_w^{-1}$$
$$\hat{\mathbf{x}}_f = \mathbf{P}_f\left(\mathbf{K}\,\hat{\mathbf{x}}_v + \mathbf{L}\,\hat{\mathbf{x}}_w\right). \tag{45b}$$

The parameter ω can be determined using conventional methods of optimization by minimizing either the trace or determinant of $\mathbf{P}_f$, formally:

$$\omega \doteq \arg\min_{\omega\in[0,1]} \text{tr}(\mathbf{P}_f). \tag{46}$$

Since the cost function is convex with respect to ω, it will have a unique minimum for $\omega \in [0, 1]$.

A proof of consistency is provided in the appendix of [31].

This consistency is depicted in Figure 6, where the ellipse $\boldsymbol{\Xi}_{f,CI}$ lies *on the outside* of the intersection of $\boldsymbol{\Xi}_v$ and $\boldsymbol{\Xi}_w$. For CI, it is always true that $\boldsymbol{\Xi}_v \cap \boldsymbol{\Xi}_w \subseteq \boldsymbol{\Xi}_{f,CI}$. Hence, it seems possible to find a fusion scheme that is less conservative, but still consistent. This can be done by characterizing and exploiting the fact that both $\mathbf{x}_v$ and $\mathbf{x}_w$ have shared information owing to their correlation. In Figure 6, $\boldsymbol{\Xi}_{f,EI} \subset \boldsymbol{\Xi}_{f,CI}$, and hence EI would be considered more accurate than CI. For EI, it is always true that $\boldsymbol{\Xi}_{f,EI} \subseteq \boldsymbol{\Xi}_v \cap \boldsymbol{\Xi}_w$, compared to CI with $\boldsymbol{\Xi}_v \cap \boldsymbol{\Xi}_w \subseteq \boldsymbol{\Xi}_{f,CI}$. However, we recall that EI is not consistent [33], [34], whereas CI is always consistent.

5.3.3 Inverse Covariance Intersection

EI relies on Assumption 5.4 to select $\check{\boldsymbol{\gamma}}$ and $\check{\boldsymbol{\Gamma}}$, which makes it inconsistent. On the other hand, CI assumes nothing about the structure of the correlation between $\mathbf{x}_v$ and $\mathbf{x}_w$, which allows it to be consistent. However, CI does not explicitly quantify the mutual information between the prior estimates. Noack et al. [32], building on previous work by Sijs et al. [34] proposed Inverse Covariance Intersection (ICI). The goal of ICI is to find a fusion scheme that estimates the mutual information *and* is consistent. As in CI, this fusion scheme is parametrized by $\omega \in (0, 1)$, and as in EI,

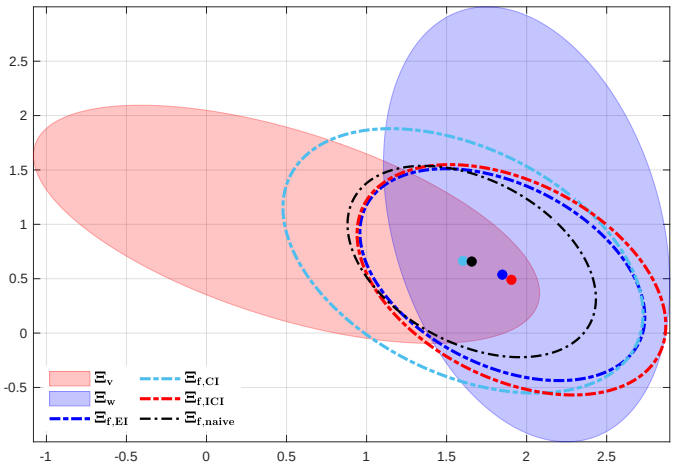


Figure 7: Fused state estimation results via CI, EI, ICI , and naïve fusion. Each ellipse has been centered around its mean, represented by a dot of respective color. It should be noted that naïve fusion is over-confident.

it removes mutual information when computing the fused information matrix. Specifically, [32] derived the following solution:

$$\mathbf{P}_f = \left(\mathbf{P}_v^{-1} + \mathbf{P}_w^{-1} - (\omega\mathbf{P}_v + (1-\omega)\mathbf{P}_w)^{-1}\right)^{-1} \tag{47a}$$
$$\mathbf{K} = \mathbf{P}_v^{-1} - \omega(\omega\mathbf{P}_v + (1-\omega)\mathbf{P}_w)^{-1}$$
$$\mathbf{L} = \mathbf{P}_w^{-1} - (1-\omega)(\omega\mathbf{P}_v + (1-\omega)\mathbf{P}_w)^{-1}$$
$$\mathbf{x}_f = \mathbf{P}_f\left(\mathbf{K}\,\mathbf{x}_v + \mathbf{L}\,\mathbf{x}_w\right) \tag{47b}$$

The parameter ω is found using eqn. (46) for $\mathbf{P}_f$ in (47a).

The formulation in eqn. (47) can also be expressed in terms of selected mutual covariance and mean, as in eqn. (33), where

$$\check{\mathbf{\Gamma}} = \omega\mathbf{P}_v + (1-\omega)\mathbf{P}_w \tag{48a}$$
$$\check{\gamma} = \omega\hat{\mathbf{x}}_v + (1-\omega)\hat{\mathbf{x}}_w. \tag{48b}$$

It should be seen that the solution arising from ICI comprises an *agreement* on the mutual covariance and mean. Consequently, the resulting fused estimate is provably consistent in the same way as CI fusion [32]. The $\check{\mathbf{\Gamma}}$ for ICI is illustrated in Figure 5, where it shown that its ellipsoid is smaller than that of EI.

The ellipsoids corresponding to fused estimation solutions obtained via all presented fusion schemes are shown in Figure 7.

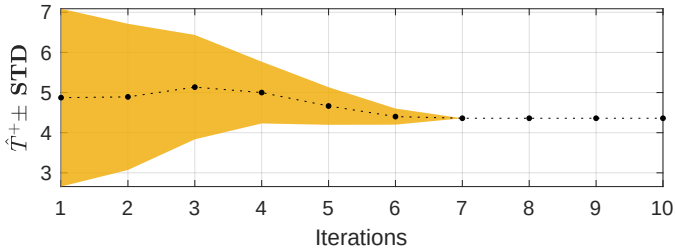
5.4 Simulation and Discussion

Figure 8 compares the convergence of heterogeneous approaches considered in this section and that of the homogeneous consensus approaches considered in Section 4. The simulation consisted of 100 agents with Watts-Strogartz small-world interactions [25] with probability of 0.3. It was found that both CI and ICI converge at the same rate, as shown in Figure 8a. Figure 8b shows that EI converges slower than both CI and ICI, but slightly faster than IWC (Figure 8c). These convergence properties hold in the presence of additional agent interactions in the network, as shown in Figure 9 of Appendix D.

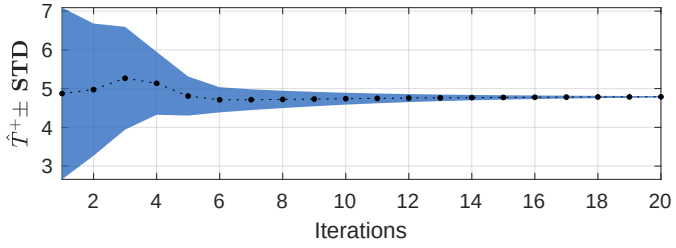
Algorithm 3 Performing CI or ICI Fusion at node v.

Input: $\hat{\mathbf{x}}_v, \mathbf{P}_v, \hat{\mathbf{x}}_w, \mathbf{P}_w$
Output: $\hat{\mathbf{x}}_f, \mathbf{P}_f$.

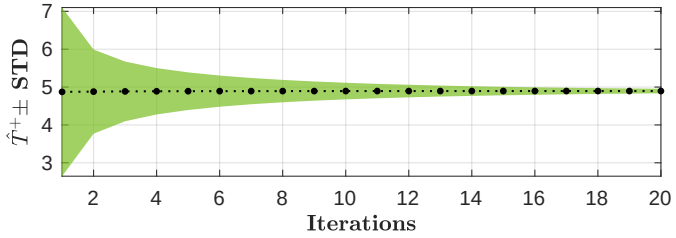
Initialize:
$$\hat{\mathbf{x}}_{v(0)} = \hat{\mathbf{x}}_v, \qquad \mathbf{P}_{v(0)} = \mathbf{P}_v$$
for $l = 1, \ldots, L$ **do**
  $\hat{\mathbf{x}}_{w(l)} = \hat{\mathbf{x}}_w, \qquad \mathbf{P}_{w(l)} = \mathbf{P}_w \qquad w \in \aleph_v$;
  Find ω for $\mathbf{P}_{v(l-1)}$ and $\mathbf{P}_{w(l)}$ by solving (46), using (45a) for CI or (47a) for ICI;
  Compute $\mathbf{P}_{v(l)}$ for $\mathbf{P}_{v(l-1)}, \mathbf{P}_{w(l)}$, and ω using (45a) for CI or (47a) for ICI;
  Compute $\hat{\mathbf{x}}_{v(l)}$ using (45b) for CI or (47b) for ICI;
end for
$\hat{\mathbf{x}}_f = \hat{\mathbf{x}}_{v(L)}, \quad \mathbf{P}_f = \mathbf{P}_{v(L)}$;



(a) Convergence of CI and ICI



(b) Convergence of EI



(c) Convergence of IWC

Figure 8: Comparison of convergence of heterogeneous approaches and IWC for a small-world network with $N = 100$ and $\aleph_v = 2$.

In terms of their convergence point, all these approaches converge at different points. Notably, also as discussed in Section 4.2, IWC is unbiased; it converges to the centralized estimate, albeit slowly. None of the heterogeneous approaches is unbiased. Unlike ICI, CI, and IWC, for EI additional interaction will not only change the convergence rate, but also the convergence point, as shown in Figure 9. It should be noted that this simulation is for the less likely case when all agents simultaneously join a network that had not been operating beforehand.

Table 1: Comparison of Different Estimation Methods

	Distributed	Consistent	Converges	Unbiased	Addresses Unknown Correlation
Centralized	—	✓	—	✓	✗
UC	✓	✗	✓	✗	✗
IWC	✓	✓	✓	✓	✗
EI	✓	✗	✓	✗	✓
CI	✓	✓	✓	✗	✓
ICI	✓	✓	✓	✗	✓

6 CONCLUSION

We have explored methods of estimating the GNSS atmospheric delay error in systems of interacting agents. In the first part, we found optimal solutions for both the individualized and centralized scenarios. Subsequently, we showed that a distributed information-weighted average consensus approach converges to the optimal centralized solution. In a simulation, we briefly explored the differences between prominent interaction topologies in terms of their convergence rate, and showed that the quasi-random small-world topology converges faster than either the scale-free or basic ring topologies. The second part of the report considered the situation of heterogeneous agents and explored methods of fusing estimates. Table 1 summarizes different properties of the methods considered in this report.

7 FUTURE DIRECTIONS

In this section we outline some avenues for future research in the topics discussed above. Some of them necessarily need to be inter-disciplinary within the wider engineering field.

7.1 Real-world Implementation

To the best of our knowledge, none of these algorithms have been implemented in a real-world scenario for distributed estimation of GNSS atmospheric delay error. The ubiquity of GNSS receivers for automotive applications and in devices such as smartphones [38]–[40] presents a real opportunity to implement these algorithms in large-scale settings.

Whereas in this report we have assumed network topologies with static and strongly connected graphs, any real-world implementation will have to consider time-variant dynamic topologies and the possibility of non-connected graphs arising from communication failures, for instance. Additionally, communication delays will also have to be taken into consideration.

7.2 Security and Safety

As mentioned earlier, the biggest challenge in satellite navigation is no longer only about accurate positioning, but about reliable, safe, and robust positioning. Relying on a communications networks allows for higher redundancy and gives an additional layer of reliability, but it also adds a potential vulnerability to hacking attacks, which may compromise other connected sensors and systems [1]. The threats to communication networks, especially the physical layer, are well-known in the literature, but are constantly evolving [41], [42]. However, these threats are beyond the scope of estimation theory as discussed within this report. Nevertheless, there are also emerging threats to GNSS systems in the form of jamming and coordinated spoofing attacks [43]–[45] that have to be detected and mitigated [46]. These attacks not only compromise positioning reliability, but also pose severe risks to personal and collective privacy. In systems with non-interacting agents, spoofing and jamming could be devastating to military and civil infrastructures [47]–[49]. Recent efforts have developed methods detecting spoofing attacks via testing GNSS position solutions against on-board Inertial Navigation System (INS) solution [50], or using machine learning techniques [51]. Future research will have to assess the full potential of coordinated attacks in sensitive networks of interacting agents, and develop novel solutions for detecting and mitigating these attacks.

7.3 Consensus on Signed Graphs

The vast majority of research on cooperative estimation is conducted on standard graphs whose edges have non-negative weights [52]. This is due to the assumption that the agents are mutually cooperative and their coupling facilitates consensus [53]. However, in real-world situations, all agents might not be cooperative; some agents might be antagonistic (or hostile) and could be modeled by assigning negative weights to the corresponding edges. This could be found in social networks, for example, in which hostile interactions could turn consensus into polarization [54], [55]. Altafini gives necessary and sufficient conditions which guarantee that a form of consensus can be achieved in the presence of hostile interactions [54]. Further research could explore the connections between antagonism in networks and the security issues outlined in Section 7.2, and potentially develop techniques to mitigate the security risks while still achieving consensus.

REFERENCES

[1] I. Fernandez-Hernandez, T. Senni, D. Calle, S. Cancela, G. A. Vecchione, and G. Seco-Granados, "Analysis of high-accuracy satellite messages for road applications," *IEEE Intelligent Transportation Systems Magazine*, vol. 12, no. 3, pp. 92–108, 2020.

[2] P.-Y. Gilliéron and P. F. SaPPART, "Satellite positioning performance assessment for road transport," 2014.

[3] G. Blewitt, "Basics of the GPS technique: observation equations," *Geodetic applications of GPS*, pp. 10–54, 1997.

[4] P. Teunissen and O. Montenbruck, *Springer handbook of global navigation satellite systems*. Springer, 2017.

[5] C. J. Hegarty and E. Chatre, "Evolution of the global navigation satellite system (GNSS)," *Proceedings of the IEEE*, vol. 96, no. 12, pp. 1902–1917, 2008.

[6] J. A. Farrell, T. D. Givargis, and M. J. Barth, "Real-time differential carrier phase GPS-aided INS," *IEEE Transactions on Control Systems Technology*, vol. 8, no. 4, pp. 709–721, 2000.

[7] J. Du and M. J. Barth, "Next-generation automated vehicle location systems: Positioning at the lane level," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 1, pp. 48–57, 2008.

[8] Anonymous, "On-Board System Requirements for V2V Safety Communications," Society of Automotive Engineers, Tech. Rep., March, 2016. [Online]. Available: https://saemobilus.sae.org/content/j2945/1_201603

[9] Y. Gao and X. Shen, "A new method for carrier-phase-based precise point positioning," *Navigation*, vol. 49, no. 2, pp. 109–116, 2002.

[10] S. Bisnath and Y. Gao, "Current state of precise point positioning and future prospects and limitations," in *Observing Our Changing Earth*. Springer, 2009, pp. 615–623.

[11] J. Kouba and P. Héroux, "Precise point positioning using IGS orbit and clock products," *GPS solutions*, vol. 5, no. 2, pp. 12–28, 2001.

[12] P. Teunissen and A. Khodabandeh, "Review and principles of PPP-RTK methods," *Journal of Geodesy*, vol. 89, no. 3, pp. 217–240, 2015.

[13] M. Elsobeiey and S. Al-Harbi, "Performance of real-time precise point positioning using igs real-time service," *GPS solutions*, vol. 20, no. 3, pp. 565–571, 2016.

[14] S. Liu, F. Sun, L. Zhang, W. Li, and X. Zhu, "Tight integration of ambiguity-fixed PPP and INS: model description and initial results," *GPS solutions*, vol. 20, no. 1, pp. 39–49, 2016.

[15] Y. Gao, Y. Zhang, and K. Chen, "Development of a real-time single-frequency precise point positioning system and test results," in *Proceedings of Ion GNSS*. Citeseer, 2006, pp. 26–29.

[16] R. J. van Bree and C. C. Tiberius, "Real-time single-frequency precise point positioning: accuracy assessment," *GPS solutions*, vol. 16, no. 2, pp. 259–266, 2012.

[17] V. L. Knoop, P. F. de Bakker, C. C. Tiberius, and B. van Arem, "Lane determination with GPS precise point positioning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 9, pp. 2503–2513, 2017.

[18] M. Elsheikh, W. Abdelfatah, A. Noureldin, U. Iqbal, and M. Korenberg, "Low-cost real-time ppp/ins integration for automated land vehicles," *Sensors*, vol. 19, no. 22, p. 4896, 2019.

[19] F. Rahman, J.-B. Uwineza, and J. A. Farrell, "ECEF position accuracy and reliability: Inertial navigation with GNSS precise point positioning (PPP)," Mar 2020.

[20] F. Rahman, F. O. Silva, Z. Jiang, and J. A. Farrell, "Low-cost real-time PPP-aided INS for CAV applications," Jun 2020.

[21] R. M. Eustice, H. Singh, and J. J. Leonard, "Exactly sparse delayed-state filters for view-based SLAM," *IEEE Transactions on Robotics*, vol. 22, no. 6, pp. 1100–1114, 2006.

[22] R. Tron and R. Vidal, "Distributed computer vision algorithms," *IEEE Signal Processing Magazine*, vol. 28, no. 3, pp. 32–45, 2011.

[23] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and Cooperation in Networked Multi-agent Systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[24] R. Olfati-Saber and R. M. Murray, "Consensus protocols for networks of dynamic agents," in *Proceedings of the 2003 American Control Conference, 2003.*, vol. 2. IEEE, 2003, pp. 951–956.

[25] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[26] A.-L. Barabási and E. Bonabeau, "Scale-free networks," *Scientific American*, vol. 288, no. 5, pp. 60–69, 2003.

[27] W. Ren, R. W. Beard, and E. M. Atkins, "Information consensus in multivehicle cooperative control," *IEEE Control systems magazine*, vol. 27, no. 2, pp. 71–82, 2007.

[28] A. D. Broido and A. Clauset, "Scale-free networks are rare," *Nature Communications*, vol. 10, no. 1, pp. 1–10, 2019.

[29] I. Voitalov, P. van der Hoorn, R. van der Hofstad, and D. Krioukov, "Scale-free networks well done," *Physical Review Research*, vol. 1, no. 3, p. 033034, 2019.

[30] P. Holme, "Rare and everywhere: Perspectives on scale-free networks," *Nature communications*, vol. 10, no. 1, pp. 1–3, 2019.

[31] 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*, vol. 4. IEEE, 1997, pp. 2369–2373.

[32] B. Noack, J. Sijs, M. Reinhardt, and U. D. Hanebeck, "Decentralized Data Fusion with Inverse Covariance Intersection," *Automatica*, vol. 79, pp. 35–41, 2017.

[33] J. Sijs, M. Lazar, and P. Bosch, "State fusion with unknown correlation: Ellipsoidal intersection," in *Proceedings of the 2010 American Control Conference*. IEEE, 2010, pp. 3992–3997.

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

[35] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[36] A. Y. Vazhentsev, "External ellipsoidal estimation of the union of two concentric ellipsoids and its applications," *Computational Mathematics and Modeling*, vol. 15, no. 2, pp. 110–122, 2004.

[37] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas (Second Edition)*. Princeton University Press, 2009.

[38] A. Elmezayen and A. El-Rabbany, "Precise point positioning using world's first dual-frequency GPS/GALILEO smartphone," *Sensors*, vol. 19, no. 11, p. 2593, 2019.

[39] J. Aggrey, S. Bisnath, N. Naciri, G. Shinghal, and S. Yang, "Multi-GNSS precise point positioning with next-generation smartphone measurements," *Journal of Spatial Science*, vol. 65, no. 1, pp. 79–98, 2020.

[40] U. Robustelli, V. Baiocchi, and G. Pugliano, "Assessment of dual frequency GNSS observations from a Xiaomi Mi 8 Android smartphone and positioning performance analysis," *Electronics*, vol. 8, no. 1, p. 91, 2019.

[41] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1727–1765, 2016.

[42] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. Di Renzo, "Safeguarding 5g wireless communication networks using physical layer security," *IEEE Communications Magazine*, vol. 53, no. 4, pp. 20–27, 2015.

[43] Anonymous, "Above Us Only Stars: Exposing GPS Spoofing in Russia and Syria," Online, Tech. Rep., 2019. [Online]. Available: https://www.c4reports.org/aboveusonlystars

[44] B. Bergman, "Systematic GPS manipulation occuring at chinese oil terminals and government installations," Dec 2019, https://skytruth.org/2019/12/systematic-gps-manipulation-occuring-at-chinese-oil-terminals-and-government-installations/.

[45] M. Harris, "Ghost ships, crop circles, and soft gold: A GPS mystery in shanghai," Dec 2019. [Online]. Available: https://www.technologyreview.com/2019/11/15/131940/ghost-ships-crop-circles-and-soft-gold-a-gps-mystery-in-shanghai/

[46] M. L. Psiaki and T. E. Humphreys, "GNSS spoofing and detection," *Proceedings of the IEEE*, vol. 104, no. 6, pp. 1258–1270, 2016.

[47] A. J. Kerns, D. P. Shepard, J. A. Bhatti, and T. E. Humphreys, "Unmanned aircraft capture and control via GPS spoofing," *Journal of Field Robotics*, vol. 31, no. 4, pp. 617–636, 2014.

[48] G. Milner, "How Vulnerable Is G.P.S.?" *The New Yorker*, Aug 2020. [Online]. Available: https://www.newyorker.com/tech/annals-of-technology/how-vulnerable-is-gps

[49] T. Humphreys, "Statement on the vulnerability of civil unmanned aerial vehicles and other systems to civil gps spoofing," *University of Texas at Austin (July 18, 2012)*, pp. 1–16, 2012.

[50] S. Dehnie and R. Ghanadan, "Methods and systems for detecting GPS spoofing attacks," Dec. 30 2014, US Patent 8,922,427.

[51] M. R. Manesh, J. Kenney, W. C. Hu, V. K. Devabhaktuni, and N. Kaabouch, "Detection of GPS spoofing attacks on unmanned aerial systems," in *2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 2019, pp. 1–6.

[52] X. Ge, Q.-L. Han, D. Ding, X.-M. Zhang, and B. Ning, "A survey on recent advances in distributed sampled-data cooperative control of multi-agent systems," *Neurocomputing*, vol. 275, pp. 1684 – 1701, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S092523121731634X

[53] A. V. Proskurnikov, A. S. Matveev, and M. Cao, "Opinion dynamics in social networks with hostile camps: Consensus vs. polarization," *IEEE Transactions on Automatic Control*, vol. 61, no. 6, pp. 1524–1536, 2015.

[54] C. Altafini, "Consensus problems on networks with antagonistic interactions," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 935–946, 2012.

[55] ——, "Dynamics of opinion forming in structurally balanced social networks," *PloS One*, vol. 7, no. 6, p. e38135, 2012.

[56] W. Ren and R. W. Beard, *Distributed consensus in multi-vehicle cooperative control*. Springer, 2008, vol. 27, no. 2.

[57] A. Barrat and M. Weigt, "On the properties of small-world network models," *The European Physical Journal B: Condensed Matter and Complex Systems*, vol. 13, no. 3, pp. 547–560, 2000.

APPENDIX A
GRAPH THEORY PRELIMINARIES

Suppose we have a network represented by an undirected connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \dots, N\}$ represents the nodes and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ represents the agent pairs $(v, w) \in \mathcal{E}$ that are able to communicate directly. The neighbors of node v is a set denoted as $\aleph_v = \{w \in \mathcal{V} : (v, w) \in \mathcal{E}\}$ and its cardinality is denoted by $|\aleph_v|$. The *adjacency matrix* of $\mathcal{G}$, $\mathbf{A} = \{a_{vw}\}$, is defined as

$$\mathbf{A}_{vw} \doteq \begin{cases} 1 & \text{, if } v = w \\ 0 & \text{, if } v \neq w. \end{cases} \tag{49}$$

The *degree matrix* is of $\mathcal{G}$ an $N \times N$ matrix defined as $\mathbf{\Delta} = \mathbf{\Delta}(\mathcal{G}) = \{\Delta_{vw}\}$ where

$$\mathbf{\Delta}_{vw} \doteq \begin{cases} |\aleph_v| & \text{, if } v = w \\ 0 & \text{, if } v \neq w. \end{cases} \tag{50}$$

The Laplacian of graph $\mathcal{G}$ is defined by

$$\mathbf{L} = \mathbf{\Delta} - \mathbf{A} \tag{51}$$

A *path* on the graph is a sequence of nodes $(w_0, w_1, \dots, w_m)$ such that $w_n \in \mathcal{V}, (w_n, w_{n+1} \in \mathcal{E})$ for all n, and m is the *path length*. An undirected graph is *connected* if there is an undirected path between every pair of distinct nodes. A graph is *strongly connected* if there is a directed path connecting any two arbitrary nodes v, w of the graph. In this report, we assume all graphs are connected. The *diameter* of $\mathcal{G}$ is defined as the maximum length of the shortest path between any pair of nodes [22], [27], [56].

In the examples of this report, we consider three undirected communication topologies: ring, scale-free, and small-world. In ring networks, each node node is connected to its *immediate* neighbors in a closed loop so that the corresponding graph is connected. A ring lattice network has the added property that each node is connected to K neighbors. Small world networks have a quasi-random topology consisting of a ring lattice with additional random connections generated by a fixed probability [25], [57]. Scale-free networks have a degree distribution that follows a power law, i.e. few nodes have most of the connections and act as hubs [26]. The small-world and scale-free topologies are illustrated in Figure 2.

APPENDIX B
DERIVATION OF THE INFORMATION VECTOR, $\mathbf{j}_{T_v}^+$

As in the information matrix in eqn. (16), the information vector is also obtained by Schur complement, which leads to

$$\mathbf{j}_{T_v}^+ = \mathbf{J}_{T_v}^+ T_v^+. \tag{52}$$

The above expression is proven below.

Proof. The Schur complement of $\mathbf{j}_{v_z}^+$ in $\mathbf{j}_v^+$ is obtained by using (14) and (15). Specifically,

$$\begin{aligned}
\mathbf{j}_{T_v}^+ &= \mathbf{j}_{v_T}^+ - (b^+)^\top (\mathbf{B}^+)^{-1} \mathbf{j}_{v_z}^+ \\
&= \Big((b^+)^\top \mathbf{z}_v + c^+ T_v \Big) - \Big((b^+)^\top (\mathbf{B}^+)^{-1} \big(\mathbf{B}^+ \mathbf{z}^+ + b^+ T_v \big) \Big) \\
&= (b^+)^\top \mathbf{z}_v + c^+ T_v - (b^+)^\top (\mathbf{B}^+)^{-1} (\mathbf{B}^+) \mathbf{z}_v^+ \\
&\quad - (b^+)^\top (\mathbf{B})^{-1} b^+ T_v \\
&= (b^+)^\top \mathbf{z}_v + c^+ T_v - (b^+)^\top \mathbf{z}_v^+ - (b^+)^\top (\mathbf{B}^+)^{-1} b^+ T_v \\
&= c^+ T_v - (b^+)^\top (\mathbf{B}^+)^{-1} (b^+) T_v \\
&= \big(c^+ - (b^+)^\top (\mathbf{B}^+)^{-1} (b^+) \big) T_v \\
&= \mathbf{J}_{T_v}^+ T_v
\end{aligned}$$

Thus, we have obtained the expression in (17). □

APPENDIX C

C.1 Simultaneous Diagonalization

This appendix states and proves a theorem on simultaneous diagonalization of two real symmetric matrices.

Theorem C.1. *Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{p \times p}$ be two real symmetric matrices, and $\mathbf{A}$ be positive definite. Then there exists a non-singular matrix $\mathbf{T} \in \mathbb{R}^{p \times p}$ such that*

$$\mathbf{T}^{-1} \mathbf{A}\, \mathbf{T}^{-\top} = \mathbf{I} \quad \textit{and} \quad \mathbf{T}^{-1} \mathbf{B}\, \mathbf{T}^{-\top} = \mathbf{D} \tag{53}$$

where $\mathbf{D}$ is a real diagonal matrix and $\mathbf{I}$ is the identity matrix.

Proof. Due to positive definitiveness, we already have $\mathbf{A} = \mathbf{R}^\top \mathbf{R}$, for some non-singular matrix $\mathbf{R}$. Hence $\mathbf{R}^{-\top} \mathbf{B}\, \mathbf{R}^{-1}$ is symmetric and there exists an orthogonal matrix $\mathbf{S}$ such that

$$\mathbf{S}^{-1}\, \mathbf{R}^{-\top} \mathbf{B}\, \mathbf{R}^{-1}\, \mathbf{S} = \mathbf{D}. \tag{54}$$

From (54), define $\mathbf{T}^{-1} = \mathbf{S}^{-1}\, \mathbf{R}^{-\top}$, and observe that

$$\begin{aligned}
&\mathbf{T}^{-1} \mathbf{B}\, \mathbf{T}^{-\top} = \mathbf{D}, \\
&\mathbf{T}^{-1}\, \mathbf{A}\, \mathbf{T}^{-\top} = \mathbf{S}^{-1}\, \mathbf{R}^{-\top} \mathbf{R}^\top \mathbf{R}\, \mathbf{R}^{-1}\, \mathbf{S} = \mathbf{I},
\end{aligned}$$

hence completing the proof. □

C.2 Proof of Theorem 5.7

We restate Theorem 5.7 below.

Theorem. *For any positive definite $\mathbf{P}_v, \mathbf{P}_w \in \mathbb{R}^{p \times p}$,*

(i) There exists a non-singular matrix $\mathbf{T} \in \mathbb{R}^{p \times p}$ such that

$$\begin{aligned}
\overline{\mathbf{P}}_v &= \mathbf{T}^{-1} \mathbf{P}_v \mathbf{T}^{-\top} = \mathbf{I}_p \\
\overline{\mathbf{P}}_w &= \mathbf{T}^{-1} \mathbf{P}_w \mathbf{T}^{-\top} = \mathbf{D}_w
\end{aligned}$$

where $\mathbf{D}_w > 0$ is diagonal, and $\mathbf{I}_p$ is the identity matrix.

(ii) The matrix $\overline{\mathbf{\Gamma}}$ defined in (37) satisfies $\overline{\mathbf{\Gamma}} = \mathbf{D}_\Gamma$, with $\mathbf{D}_\Gamma = \max\{\overline{\mathbf{P}}_v, \overline{\mathbf{P}}_w\} = \max\{\mathbf{I}_p, \mathbf{D}_w\}$, where $\max$ is the component-wise maximum function.

(iii) The matrix $\check{\mathbf{\Gamma}}$ defined in (35) satisfies $\check{\mathbf{\Gamma}} = \mathbf{T}\, \mathbf{D}_\Gamma\, \mathbf{T}^\top$.

An alternative to the following proof can be found in [34].

Proof. (i) This follows directly from Theorem C.1.

(ii) We want to show that $\mathbf{D}^\star = \mathbf{D}_\Gamma$ satisfies

$$\mathbf{D}^\star \doteq \arg\min \log|\boldsymbol{\Upsilon}|, \text{ subject to } \boldsymbol{\Upsilon} \geq \mathbf{D}_v, \boldsymbol{\Upsilon} \geq \mathbf{D}_w$$

We first note two trivial cases. If $\mathbf{I}_p \geqslant \mathbf{D}_w$—that is, $[\mathbf{I}_p]_{qq} \geq [\mathbf{D}_w]_{qq}$ for $q \in \mathbb{Z}_{[1,p]}$—then $\mathbf{D}^\star = \mathbf{D}_\Gamma = \mathbf{I}_p$. Similarly, if $\mathbf{I}_p \leqslant \mathbf{D}_w$, then $\mathbf{D}_\Gamma = \mathbf{D}_w$.
For the remaining cases when $\mathbf{I}_p \nleqslant \mathbf{D}_w$ and $\mathbf{I}_p \ngeqslant \mathbf{D}_w$, we use a result in [36] for general $\mathbf{D}_v, \mathbf{D}_w \geq 0$. But we have to first assume a structure for $\mathbf{D}_v, \mathbf{D}_w$. Specifically,

Assumption C.2. *$[\mathbf{D}_v]_{qq} \geq [\mathbf{D}_w]_{qq}$ for $q \in \mathbb{Z}_{[1,n]}$ and $[\mathbf{D}_v]_{qq} \leq [\mathbf{D}_w]_{qq}$ for $q \in \mathbb{Z}_{[n+1,p]}$, for some $n \in \mathbb{Z}_{[0,p]}$.*

Since $\mathbf{I}_p \nleqslant \mathbf{D}_w$ and $\mathbf{I}_p \ngeqslant \mathbf{D}_w$, this assumption is fulfilled when we apply a suitable transformation to $\mathbf{D}_w$. Also, for each n as in Assumption C.2, we define a family of matrices

$$\mathcal{T} \doteq \left\{ \mathbf{T} \in \mathbb{R}^{n \times p-n} \middle| \mathbf{T}\mathbf{T}^\top \leqslant \mathbf{I}_n \right\},$$

and define $\mathbf{S} \in \mathbb{R}^{p \times p}$ as

$$\mathbf{S}_\mathbf{T} \doteq \begin{bmatrix} \mathbf{I}_n & \mathbf{T} \\ \mathbf{T}^\top & \mathbf{I}_{p-n} \end{bmatrix}. \tag{55}$$

Subsequently, define $\boldsymbol{\Sigma}_\mathbf{T} = \mathbf{Q} + \mathbf{R}\,\mathbf{S}_\mathbf{T}^{-1}\mathbf{R}$, where

$$\mathbf{Q} = \min\{[\mathbf{D}_v]_{qq}[\mathbf{D}_w]_{qq}\}, \tag{56}$$

$$\mathbf{R} = \left| [\mathbf{D}_v]_{qq} - [\mathbf{D}_w]_{qq} \right|^{1/2}. \tag{57}$$

Then, we use Theorem C.3 (Theorem 2 in [36]) that states that $\mathbf{D}_v, \mathbf{D}_w \leq \boldsymbol{\Sigma}$ for any positive definite $\mathbf{D}_v$ and $\mathbf{D}_w$.

Theorem C.3 (From [36])**.** *For $\mathbf{D}_w, \mathbf{D}_v \geq 0$, there does not exist a matrix $\tilde{\mathbf{D}} \neq \boldsymbol{\Sigma}_\mathbf{T}$, for any $\mathbf{T} \in \mathcal{T}$ such that $\mathbf{D}_v \leq \tilde{\mathbf{D}} \leq \boldsymbol{\Sigma}_\mathbf{T}$ and $\mathbf{D}_w \leq \tilde{\mathbf{D}} \leq \boldsymbol{\Sigma}_\mathbf{T}$.*

Using the known property that $0 \leq \log|\mathbf{A}| \leq \log|\mathbf{B}|$ holds for any $\mathbf{0} \leq \mathbf{A} \leq \mathbf{B}$ (Corollary 8.4.10. of [37]), Theorem C.3 implies that $\mathbf{D}^\star = \boldsymbol{\Sigma}_{\mathbf{T}^\star}$. Therefore,

$$\mathbf{D}^\star \doteq \mathbf{Q} + \mathbf{R}\,\mathbf{S}_{\mathbf{T}^\star}^{-1}\mathbf{R}$$

with $\mathbf{T}^\star \doteq \arg\min_{\mathbf{T} \in \mathcal{T}} \log|\boldsymbol{\Sigma}_\mathbf{T}|$.
We notice that $\mathbf{D}_\Gamma = \mathbf{Q} + \mathbf{R}^2$ holds for any $\mathbf{D}_v = \mathbf{I}_p$ and $\mathbf{D}_w > 0$. In that case $\mathbf{D}^\star = \mathbf{Q} + \mathbf{R}^2$, which implies that $\mathbf{S}_{\mathbf{T}^\star} = \mathbf{I}_p$ and $\mathbf{T}^\star = \mathbf{0}_{n \times p-n}$. Hence, we are only left to show that $\mathbf{T}^\star = \mathbf{0}_{n \times p-n}$.
Using the property that $0 \leq \log|\mathbf{A}| \leq \log|\mathbf{B}|$ holds for any $\mathbf{0} \leq \mathbf{A} \leq \mathbf{B}$, we get that $\mathbf{T}^\star$ satisfies

$$\mathbf{Q} + \mathbf{R}\,\mathbf{S}_{\mathbf{T}^\star}^{-1}\mathbf{R} \leq \mathbf{Q} + \mathbf{R}\,\mathbf{S}_\mathbf{T}^{-1}\mathbf{R}$$

which means that $\mathbf{S}_{\mathbf{T}^\star} \geq \mathbf{S}_\mathbf{T}$, for all $\mathbf{T}$. Equivalently, $\mathbf{T}^\star$ is obtained using $\mathbf{T}^\star \doteq \arg\max_{\mathbf{T} \in \mathcal{T}} \log|\mathbf{S}_\mathbf{T}|$. Since $|\mathbf{S}_\mathbf{T}| = \left|\mathbf{I}_n - \mathbf{T}\mathbf{T}^\top\right|$ (Fact 2.14.9. of [37]), we get

$$\begin{aligned} \mathbf{T}^\star &= \arg\max_{\mathbf{T} \in \mathcal{T}} \log\left|\mathbf{I}_n - \mathbf{T}\mathbf{T}^\top\right| \\ &= \arg\max_{\mathbf{T} \in \mathcal{T}} \sum_{q=1}^{n} \lambda_q(\mathbf{I}_n - \mathbf{T}\mathbf{T}^\top) \\ &= \arg\max_{\mathbf{T} \in \mathcal{T}} \sum_{q=1}^{n} 1 - \lambda_q(\mathbf{T}\mathbf{T}^\top) \end{aligned} \tag{58}$$

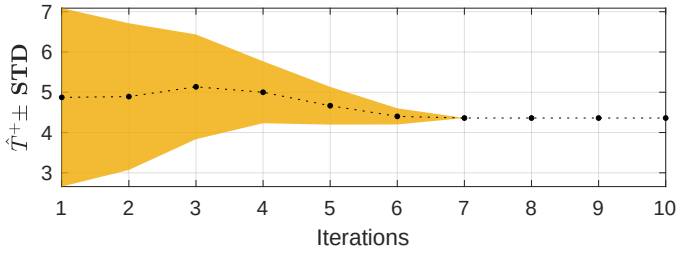
The fact that $\mathbf{0} \leq \mathbf{T}\mathbf{T}^\top \leq \mathbf{I}_n$, for all $\mathbf{T} \in \mathcal{T}$ and $q \in \mathbb{Z}_{[1,n]}$, implies $0 \leq \lambda_q(\mathbf{T}\mathbf{T}^\top) \leq 1$. Hence, the maximum in (58) is reached when $\lambda_q(\mathbf{T}\mathbf{T}^\top) = 1$ for all $q \in \mathbb{Z}_{[1,n]}$. Therefore, $\mathbf{T}^\star = \mathbf{0}_{n \times p-n}$.

(iii) This follows directly from Lemma 5.6.

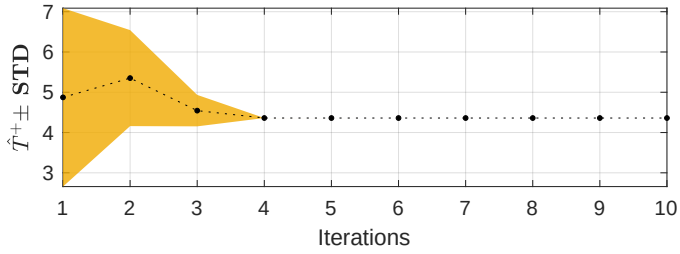
□

APPENDIX D
EFFECT OF NUMBER OF NEIGHBORS

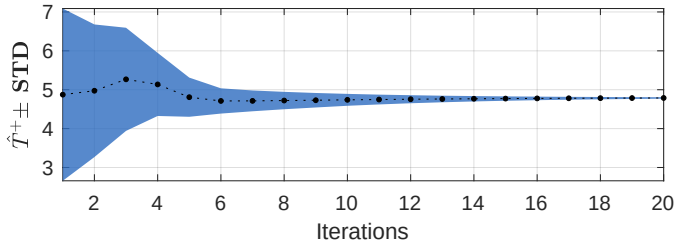
Increasing the average number of interactions for each agent ($\aleph_v$) while holding the size of the network (N) constant decreases the convergence time for all methods considered in this report. However, for EI this also changes its convergence point, as shown in Figures 9c and 9d. This points to the underlying inconsistency of the EI approach.



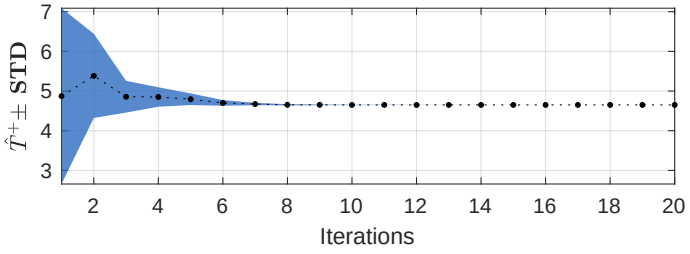
(a) Convergence of CI and ICI for $\aleph_v = 2$



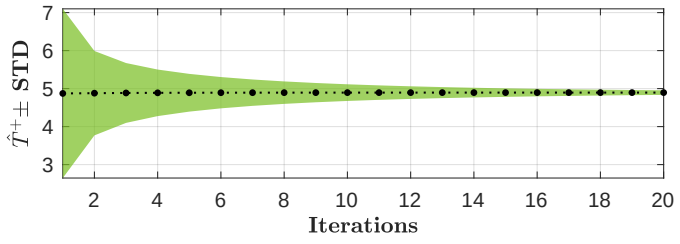
(b) Convergence of CI and ICI for $\aleph_v = 10$



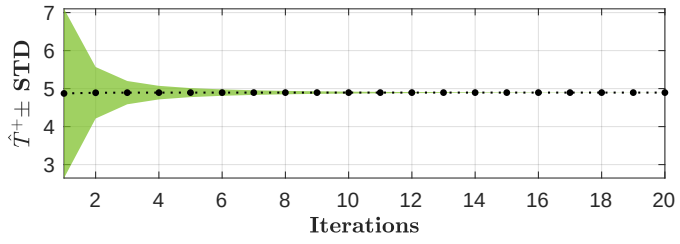
(c) Convergence of EI for $\aleph_v = 2$



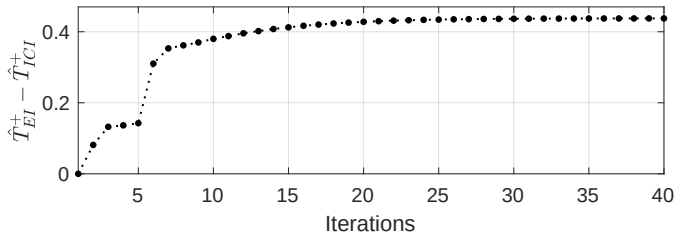
(d) Convergence of EI for $\aleph_v = 10$



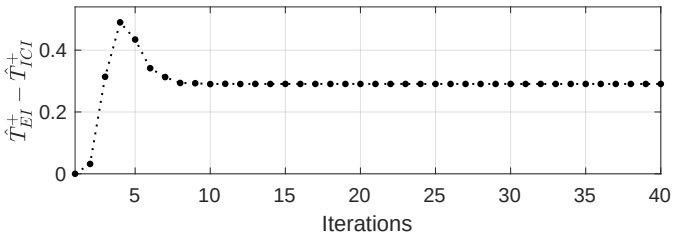
(e) Convergence of IWC for $\aleph_v = 2$



(f) Convergence of IWC for $\aleph_v = 10$



(g) Difference between ICI and EI $\aleph_v = 2$



(h) Difference between ICI and EI $\aleph_v = 10$

Figure 9: Convergence rates for a small-world network of $N = 100$, $\aleph_v = \{2, 10\}$ and $p = 0.3$.