

# PARTIAL-STATE FUSION FOR DISTRIBUTED SIMULTANEOUS LOCALIZATION AND TRACKING

Simone Semeraro\*, Keith A. LeGrand†, and Jackson Kulik‡

In target tracking problems, an agent’s knowledge of its own state and its ability to accurately process the relative measurements of the target it senses are strongly interdependent. In simultaneous localization and tracking algorithms, an autonomous agent, such as a satellite, jointly estimates the target and its own ego state. By this approach, incoming navigation data may refine existing target estimates, and vice versa. In distributed networks of autonomous agents tracking a common target, new target-related information can be disseminated throughout the network in the form of probabilistic beliefs. Optimal fusion of these beliefs requires full knowledge of the underlying data pedigree and inter-agent correlations, which is typically unavailable. On the other hand, conservative fusion strategies, while resulting in sub-optimal information gain, require no knowledge of network topology or data pedigree and inherently prevent information double-counting. This paper discusses a new mixture based approximation of non-Gaussian partial state fusion suitable for nonlinear distributed simultaneous localization and tracking. In this framework, each agent operates without knowledge of other agents’ ego-states, and thus only target-related beliefs are exchanged and fused. By this agnosticism to other agents’ ego-states, the resulting algorithm is scalable to large networks of heterogeneous agents. It is shown that the exchange of target-related beliefs improves not only target state estimates, but also the ego states. New mixture reduction techniques are explored to reduce the computational complexity while mitigating information loss. The overall framework is applied to an autonomously navigating spaceborne constellation of agents tracking an uncooperative target satellite.

## INTRODUCTION

Multi-agent networks are becoming more prominent in real-world application scenarios spanning many fields, such as space situational awareness (SSA),<sup>1</sup> bio-inspired nanotechnology,<sup>2</sup> and traffic management.<sup>3</sup> In many cases, such systems are centralized, meaning that network information is routed to a single node for processing by a single unit. Disadvantages of the centralized architecture include its inherent single point of failure and communication bottlenecks as the network size grows. Both issues can be mitigated in distributed architectures, which promote agent autonomy and responsibility. In cooperative scenarios, the agents can maintain autonomy and improve their awareness by exchanging information across the network, in the form of raw measurements or statistics. Statistics dissemination offers distinct advantages in large-network or communication-limited scenarios.<sup>4,5</sup> Additionally, cooperation can be achieved even between heterogeneous agent networks by correctly handling the diverse state space models the agents operate on. Among several information exchange protocols, *consensus* involves all agents “agreeing” on the information via an iterative process.<sup>4</sup> Optimal decentralized information exchange comes, however, at the cost of maintaining detailed data history and pedigree to avoid information double-counting, which is often prohibitively expensive.<sup>6</sup>

Conservative fusion algorithms, on the other hand, require no data history or knowledge of network topology. These approaches, while inherently suboptimal, are designed to produce the best possible fusion result

\*Ph.D. Student, School of Aeronautics and Astronautics, Purdue University, 701 W Stadium Ave, West Lafayette, IN 47907, United States.

†Assistant Professor, School of Aeronautics and Astronautics, Purdue University, 701 W Stadium Ave, West Lafayette, IN 47907, United States.

‡Assistant Professor, Mechanical and Aerospace Engineering, Utah State University, 4130 Old Main Hill Logan, UT 84322-4130, United States.

assuming that the input estimates are fully correlated. The first widely adopted technique for handling unknown correlations in state estimate fusion was covariance intersection (CI), proposed by Uhlmann.<sup>7</sup> Since then, additional conservative information fusion algorithms have been developed to accommodate more general settings. Conservative fusion of probability densities can be obtained through exponentiated products or weighted summation in techniques respectively known as exponential mixture density (EMD) and arithmetic mean density (AMD) fusion.<sup>8</sup> These conservative fusion strategies have been extended to multi-sensor multi-object environments using random finite set (RFS) theory.<sup>9–16</sup> Ahmed et al.<sup>17,18</sup> derived a *partial state* update from the general EMD strategy, where only a common subset of estimated states are communicated between agents to reach consensus. Information about the non-common states is then obtained by leveraging their cross correlations, and the approach is applied to a linear, Gaussian system. Semeraro et al.<sup>19</sup> expanded the partial state fusion framework to non-Gaussian beliefs by deriving Gaussian mixture (GM) approximations of varying computational complexity and accuracy.

The partial state fusion GM approximations involve several GM multiplications that, if computed directly, can sharply inflate the fusion mixture size (as measured by the number of mixands). For online estimation applications, where the proposed fusion algorithm would be performed within a filter, this undesirable mixture inflation creates computational tractability challenges that must be addressed. Furthermore, the mixture growth, if unaddressed, practically precludes the naive application of existing mixture reduction techniques post-fusion, motivating new reduction algorithms.

This paper extends the authors' prior work<sup>19</sup> by introducing mixture reduction methods that are tightly integrated with the partial-state fusion approximation, making the approach tractable in online estimation problems. The overall approach is demonstrated in a spaceborne simultaneous localization and tracking (SLAT) multi-agent, single target scenario. The problem formulation section introduces the framework of interest, followed by a background section that reviews the main properties of GM probability density functions (pdfs) and the general partial-state fusion formulation. The methodology section presents the GM solution to fusion via three approximation techniques and several reduction methods, including two newly proposed approaches. The results section evaluates the accuracy of the approximations and reductions and applies them to a tracking scenario in an astrodynamics setting. The article concludes with a summary of the main findings.

## PROBLEM FORMULATION

Consider a fully connected network of agents discerned by the index set  $\mathcal{I} = \{1, 2, 3, \dots\} \subset \mathbb{N}$ . Each agent  $i \in \mathcal{I}$  estimates states subject to the nonlinear system

$$\mathbf{x}_{k+1}^{(i)} = \mathbf{f}_k^{(i)}(\mathbf{x}_k^{(i)}) + \mathbf{G}\mathbf{w}_k^{(i)} \quad (1)$$

where the superscript  $(\cdot)^{(i)}$  and subscript  $(\cdot)_k$  indicate quantities associated with the  $i^{\text{th}}$  agent and the time  $t_k$ , respectively. Furthermore, in (1), state and dynamics function, are respectively represented by  $\mathbf{x}^{(i)} \in \mathbb{R}^{n_x}$ , and  $\mathbf{f}_k^{(i)} : \mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_x}$ , the Gaussian process noise  $\mathbf{w}^{(i)} \in \mathbb{R}^{n_w}$  is white, zero-mean, with covariance  $\mathbf{Q}_k^{(i)} = \mathbb{E}[\mathbf{w}_k^{(i)}(\mathbf{w}_k^{(i)})^\top]$ , and  $\mathbf{G} \in \mathbb{R}^{n_x \times n_w}$ . The  $i^{\text{th}}$  agent takes target measurements

$$\mathbf{z}_k^{(i)} = \mathbf{h}_k^{(i)}(\mathbf{x}_k^{(i)}) + \boldsymbol{\nu}_k^{(i)} \quad (2)$$

with measurement map  $\mathbf{h}_k^{(i)} : \mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_z}$  and white, zero-mean, Gaussian measurement noise with covariance  $\mathbf{R}_k^{(i)} = \mathbb{E}[\boldsymbol{\nu}_k^{(i)}(\boldsymbol{\nu}_k^{(i)})^\top]$ .

Only a subset of the estimated states are common across all agents of the network, henceforth defined as  $\mathbf{t} \in \mathbb{R}^{n_t}$ , while the remainder are of no interest to any other agent, represented by  $\mathbf{e}^{(i)} \in \mathbb{R}^{n_e}$ , such that  $n_t + n_e = n_x$ . The state vector is thus partitioned as

$$\mathbf{x}^{(i)} = \begin{bmatrix} \mathbf{t} \\ \mathbf{e}^{(i)} \end{bmatrix} \quad (3)$$

In the context of the SLAT problem,  $\mathbf{t}$  represents the target state and  $\mathbf{e}^{(i)}$  represents the ego- or navigation state of agent  $i$ . More generally, the presented framework applies to any setting in which the state can be partitioned into common and non-common states. The agents communicate only the common target information in the form of pdfs following a *consensus* protocol, where the exchange happens bidirectionally. According to this strategy, the goal is to allow all the connected agents to reach a certain probabilistic agreement over the target states.

## BACKGROUND

### Gaussian Mixtures

In nonlinear filtering, Gaussian assumptions are often violated, making standard Kalman filters inadequate for representing non-Gaussian uncertainty. GMs can approximate arbitrary distributions with arbitrary accuracy while retaining many desirable properties of Gaussian pdfs. A GM pdf of  $L$  mixands is compactly represented by the weighted sum of Gaussian distributions

$$p(\mathbf{x}) = \sum_{\ell=1}^L w^{(\ell)} \mathcal{N}(\mathbf{x}; \mathbf{m}^{(\ell)}, \mathbf{P}^{(\ell)}) \quad (4)$$

with  $w^{(\ell)}$ ,  $\mathbf{m}^{(\ell)}$ , and  $\mathbf{P}^{(\ell)}$  the  $\ell^{\text{th}}$  Gaussian mixand's weight, mean, and covariance, respectively; moreover, each weight must belong to the interval  $[0, 1]$  and respect the property  $\sum_{\ell=1}^L w^{(\ell)} = 1$ . GM filters present an attractive opportunity to estimate states mapped by nonlinear dynamics and/or measurement functions. Following the state partition in (3), the  $i^{\text{th}}$  agent's statistical information can also be partitioned as

$$p^{(i)}(\mathbf{x}^{(i)}) = \sum_{\ell=1}^{L^{(i)}} w^{(i,\ell)} \mathcal{N}(\mathbf{x}^{(i)}; \mathbf{m}^{(i,\ell)}, \mathbf{P}^{(i,\ell)}) \quad (5)$$

$$= \sum_{\ell=1}^{L^{(i)}} w^{(i,\ell)} \mathcal{N}\left(\begin{bmatrix} \mathbf{t} \\ \mathbf{e}^{(i)} \end{bmatrix}; \begin{bmatrix} \mathbf{m}_t^{(i,\ell)} \\ \mathbf{m}_e^{(i,\ell)} \end{bmatrix}, \begin{bmatrix} \mathbf{P}_t^{(i,\ell)} & \boldsymbol{\Psi}^{(i,\ell)} \\ \left(\boldsymbol{\Psi}^{(i,\ell)}\right)^\top & \mathbf{P}_e^{(i,\ell)} \end{bmatrix}\right) \quad (6)$$

where the superscripts  $(\cdot)^{(i,\ell)}$  refer to the  $\ell^{\text{th}}$  mixand of the  $i^{\text{th}}$  agent, and the subscripts  $(\cdot)_t$  and  $(\cdot)_e$  discern quantities of target states and ego-states, respectively. The cross correlations between uncertainties in the target and ego states  $\boldsymbol{\Psi}^{(i,\ell)}$  are, in general, nonzero.

In *adaptive* GM filters, the number of GM mixands is systematically increased via splitting to better approximate the nonlinearly mapped mixture approximations.<sup>20–27</sup> Conversely, the mixture resolution is reduced as needed to decrease computational load by merging statistically similar and/or probabilistically insignificant mixands.<sup>28,29</sup>

### Conditionally Factorized Fusion

When agents do not share the same estimated state, full-state fusion approaches cannot be applied directly without marginalizing out the non-common portion of the state and, consequently, discarding known correlations with the target vector. Consider the EMD fusion rule between two distinct agents  $i, j \in \mathcal{I}$ ,

$$p_f^{(i)}(\mathbf{x}^{(i)}) \propto \left(p^{(i)}(\mathbf{x}^{(i)})\right)^{\tilde{\omega}} \left(p^{(j)}(\mathbf{x}^{(i)})\right)^{1-\tilde{\omega}} \quad (7)$$

where  $\tilde{\omega} \in [0, 1]$  is the fusion weight and  $p_f^{(i)}$  represents the fused belief as performed by agent  $i$ . Ahmed<sup>17</sup> showed that conditional factorization can be exploited to extract target state beliefs as

$$p_f^{(i)}(\mathbf{x}^{(i)}) \propto \left[p^{(i)}(\mathbf{e}^{(i)}|\mathbf{t}) p^{(i)}(\mathbf{t})\right]^{\tilde{\omega}} \left[p^{(j)}(\mathbf{e}^{(i)}|\mathbf{t}) p^{(j)}(\mathbf{t})\right]^{1-\tilde{\omega}} \quad (8)$$

$$= \left[p^{(i)}(\mathbf{t})\right]^{\tilde{\omega}} \left[p^{(j)}(\mathbf{t})\right]^{1-\tilde{\omega}} \left[p^{(i)}(\mathbf{e}^{(i)}|\mathbf{t})\right]^{\tilde{\omega}} \left[p^{(j)}(\mathbf{e}^{(i)}|\mathbf{t})\right]^{1-\tilde{\omega}} \quad (9)$$

Examination of (9) reveals that a more granular fusion is possible, where each fusion operation can be performed with a different weight that is commensurate to the knowledge each agent possesses of the state in question. To that end, two different weights are considered: one for the target marginal spaces pdfs,  $\omega \in [0, 1]$ , and one affecting the conditional distributions,  $\omega_c \in [0, 1]$ , leading to<sup>17</sup>

$$p_f^{(i)}(\mathbf{x}^{(i)}) \propto [p^{(i)}(\mathbf{t})]^\omega [p^{(j)}(\mathbf{t})]^{1-\omega} [p^{(i)}(\mathbf{e}^{(i)}|\mathbf{t})]^{\omega_c} [p^{(j)}(\mathbf{e}^{(i)}|\mathbf{t})]^{1-\omega_c} \quad (10)$$

and noting that agents do not estimate each other's ego-states, an obvious choice for  $\omega_c$  is unity. Thus, rearranging the terms in (10), the partial state fusion update becomes

$$p_f^{(i)}(\mathbf{x}^{(i)}) \propto [p^{(i)}(\mathbf{t})]^\omega [p^{(j)}(\mathbf{t})]^{1-\omega} [p^{(i)}(\mathbf{e}^{(i)}|\mathbf{t})] \quad (11)$$

$$= [p^{(i)}(\mathbf{t})]^{\omega-1} [p^{(j)}(\mathbf{t})]^{1-\omega} p^{(i)}(\mathbf{e}^{(i)}|\mathbf{t}) p^{(i)}(\mathbf{t}) \quad (12)$$

$$= p^{(i)}(\mathbf{x}^{(i)}) [p^{(i)}(\mathbf{t})]^{\omega-1} [p^{(j)}(\mathbf{t})]^{1-\omega} \quad (13)$$

By this factorization, the posterior fused density can be written as the product of the prior joint distribution and a partial (target) state update factor  $g^{(i)}(\mathbf{t}, \omega)$  as

$$p_f^{(i)}(\mathbf{x}^{(i)}) \propto p^{(i)}(\mathbf{x}^{(i)}) g^{(i)}(\mathbf{t}, \omega) \quad (14)$$

$$g^{(i)}(\mathbf{t}, \omega) = [p^{(i)}(\mathbf{t})]^{\omega-1} [p^{(j)}(\mathbf{t})]^{1-\omega} \quad (15)$$

Moreover, Ahmed et al.<sup>18</sup> applied (14) to a system involving Gaussian beliefs, as well as introducing an approach to obtaining a valid Gaussian representation of the quotient in (15).

## METHODOLOGY

### Exponential Quadratic Mixture

An exact solution for the general exponentiated product (15) is not available in closed form for GMs, and the computation of the update factor  $g^{(i)}$  is aggravated by its expression as a quotient of distributions. To reduce the complexity of the operations involved, a more flexible representation is introduced, called an exponential quadratic (EQ) function:<sup>19</sup>

$$\mathcal{E}(\mathbf{x}; \mathbf{m}, \mathbf{P}) = \exp \left\{ -\frac{1}{2}(\mathbf{x} - \mathbf{m})^\top \mathbf{P}^{-1}(\mathbf{x} - \mathbf{m}) \right\} \quad (16)$$

A Gaussian pdf is related to an EQ via the scaling

$$\mathcal{N}(\mathbf{x}; \mathbf{m}, \mathbf{P}) = |2\pi\mathbf{P}|^{-1/2} \mathcal{E}(\mathbf{x}; \mathbf{m}, \mathbf{P}) \quad (17)$$

and a GM pdf is translated to an exponential quadratic mixture (EQM) as

$$\sum_{\ell=1}^L w^{(\ell)} \mathcal{N}(\mathbf{x}^{(\ell)}; \mathbf{m}^{(\ell)}, \mathbf{P}^{(\ell)}) = \sum_{\ell=1}^L \bar{w}^{(\ell)} \mathcal{E}(\mathbf{x}^{(\ell)}; \mathbf{m}^{(\ell)}, \mathbf{P}^{(\ell)}), \quad \bar{w}^{(\ell)} = \frac{w^{(\ell)}}{\sqrt{|2\pi\mathbf{P}^{(\ell)}|}} \quad (18)$$

Now, as the exponentiation of pdfs in (15) is not available in exact, closed form, it can be obtained by a first-order power approximation (FOPA),<sup>30</sup> which can be directly applied to GMs and their EQMs counterparts as

$$\left[ \sum_{\ell=1}^L w^{(\ell)} \mathcal{N}(\mathbf{x}^{(\ell)}; \mathbf{m}^{(\ell)}, \mathbf{P}^{(\ell)}) \right]^\alpha \approx \sum_{\ell=1}^L \left[ \bar{w}^{(\ell)} \mathcal{E}(\mathbf{x}^{(\ell)}; \mathbf{m}^{(\ell)}, \mathbf{P}^{(\ell)}) \right]^\alpha \quad (19)$$

$$= \sum_{\ell=1}^L \left[ \bar{w}^{(\ell)} \right]^\alpha \mathcal{E}(\mathbf{x}^{(\ell)}; \mathbf{m}^{(\ell)}, \mathbf{P}^{(\ell)}/\alpha) \quad (20)$$

The GM representation is unsuitable for this approximation in the case where the exponent  $\alpha$  is negative, as the covariance matrices become negative definite and weights imaginary. Instead, adoption of the EQM functions obviates such issues, becoming a desirable conduit for the operations that follow.

At a certain fusion instance, consider two different agents  $i, j \in \mathcal{I}$ , each of which possesses a probabilistic belief  $p(\mathbf{x}^{(\cdot)})$  in the form of GMs, like in (4), and with the same state partitioning from (3). Assuming without loss of generality agent  $i$  receives  $j$ 's target state, the FOPA is used to compute the update factor EQM mixands' means and covariances<sup>19</sup>

$$g^{(i)}(\mathbf{t}, \omega) \approx g_{\text{FOPA}}^{(i)}(\mathbf{t}, \omega) = \sum_{\ell=1}^L \bar{w}_{\text{FOPA}}^{(i,\ell)}(\omega) \mathcal{E}\left(\mathbf{t}; \mathbf{m}_{\text{t,FOPA}}^{(i,\ell)}(\omega), \mathbf{P}_{\text{t,FOPA}}^{(i,\ell)}(\omega)\right) \quad (21)$$

$$\mathbf{P}_{\text{t,FOPA}}^{(i,\ell)}(\omega) = \left( (\omega - 1) \left( \mathbf{P}_{\text{t}}^{(i,q)} \right)^{-1} + (1 - \omega) \left( \mathbf{P}_{\text{t}}^{(j,r,t)} \right)^{-1} \right)^{-1} \quad (22)$$

$$\mathbf{m}_{\text{t,FOPA}}^{(i,\ell)}(\omega) = \mathbf{P}_{\text{t,FOPA}}^{(i,\ell)}(\omega) \left( (\omega - 1) \left( \mathbf{P}_{\text{t}}^{(i,q)} \right)^{-1} \mathbf{m}_{\text{t}}^{(i,q)} + (1 - \omega) \left( \mathbf{P}_{\text{t}}^{(j,r)} \right)^{-1} \mathbf{m}_{\text{t}}^{(j,r)} \right) \quad (23)$$

$$\ell = (q - 1)L^{(j)} + r, \quad q = 1, \dots, L^{(i)}, \quad r = 1, \dots, L^{(j)} \quad (24)$$

A FOPA-based expression for the weights  $\bar{w}_{\text{FOPA}}^{(i,\ell)}(\omega)$  exists, but is omitted here for brevity, as the resulting weights might return too coarse of an approximation of the fusion output (14)<sup>19,30</sup> and a more detailed discussion is presented later on. To reduce clutter, the dependency on  $\omega$  of the update factor (15) and its parameters in (21)-(23) is dropped. Once the update factor approximation (21) is obtained, the fusion output (14) is computed by carrying out another mixand-wise multiplication between agent  $i$ 's distribution and  $g_{\text{FOPA}}^{(i)}(\mathbf{t})$ , a GM with lower dimension:

$$p_f^{(i)}(\mathbf{x}^{(i)}) \propto \sum_{s=1}^{L_f^{(i)}} \bar{w}_f^{(i,s)} \mathcal{E}\left(\mathbf{x}^{(i)}; \mathbf{m}_f^{(i,s)}, \mathbf{P}_f^{(i,s)}\right) \quad (25)$$

$$\mathbf{P}_f^{(i,s)} = \left( \left( \left[ \begin{array}{cc} \mathbf{P}_{\text{t}}^{(i,q)} & \boldsymbol{\Psi}^{(i,q)} \\ \left( \boldsymbol{\Psi}^{(i,q)} \right)^\top & \mathbf{P}_{\text{e}}^{(i,q)} \end{array} \right] \right)^{-1} + \left[ \begin{array}{cc} \left( \mathbf{P}_{\text{t,FOPA}}^{(i,\ell)} \right)^{-1} & 0 \\ 0 & 0 \end{array} \right] \right)^{-1} \quad (26)$$

$$\mathbf{m}_f^{(i,s)} = \mathbf{P}_f^{(i,s)} \left( \left( \mathbf{P}^{(i,s)} \right)^{-1} \mathbf{m}^{(i,s)} + \left[ \begin{array}{c} \left( \mathbf{P}_{\text{t,FOPA}}^{(i,\ell)} \right)^{-1} \mathbf{m}_{\text{t,FOPA}}^{(i,\ell)} \\ 0 \end{array} \right] \right) \quad (27)$$

$$s = (\ell - 1)L^{(i)} + q, \quad \ell = 1, \dots, L, \quad q = 1, \dots, L^{(i)} \quad (28)$$

The bracketed vector in (27) and second bracketed matrix in (26) have a structure similar to the information vector and information matrix, respectively. The zero rows and columns corresponding to the ego-state variables indicate that the update factor carries no explicit information about the ego-state. However, correlations between target and ego states still indirectly yield an updated ego-state covariance, allowing agent  $i$  to refine its ego-state estimate by fusing only target state information with the network. The fused mixand covariance  $\mathbf{P}_f^{(i,s)}$  in (26) is guaranteed to be positive definite.<sup>19</sup>

The FOPA approach is generally unsuitable to compute weights of the approximated GM (14). Instead, they can be chosen to reduce the overall error of  $p_f^{(i)}(\mathbf{x}^{(i)})$  induced by update factor approximation errors.<sup>19</sup> In particular, weights can be chosen evaluating the *exact* update factor (15) at strategically chosen collocation points. Three update methods are discussed in the following. The first weight update method, collocation 1 (C1), assumes the mixands of the fusion outcome  $p_f^{(i)}$  are well separated and uses as a collocation point the mean of the prior mixand index

$$q' = \text{mod}\left(s + 1, L^{(i)}\right) - 1 \quad (29)$$

from which the posterior mixand was formed in (26) and (27), thus returning

$$\bar{w}_f^{(i,s)} = \frac{1}{c} \bar{w}^{(i,s)} \mathcal{E} \left( \mathbf{m}_f^{(i,s)}; \mathbf{m}^{(i,q')}, \mathbf{P}^{(i,q')} \right) g^{(i)} \left( \mathbf{m}_{t,f}^{(i,s)} \right) \quad (30)$$

where  $c$  is an unspecified normalization constant.

The second approach, collocation 2 (C2), retains the simplifying assumption that the fused GM mixands are well separated, but integrates the contribution from the full input distribution, thus refining the weight calculation to

$$\bar{w}_f^{(i,s)} = \frac{1}{c} g^{(i)} \left( \mathbf{m}_{t,f}^{(i,s)} \right) \sum_{q=1}^{L^{(i)}} \bar{w}^{(i,s)} \mathcal{E} \left( \mathbf{m}_f^{(i,s)}; \mathbf{m}^{(i,q)}, \mathbf{P}^{(i,q)} \right) \quad (31)$$

In contrast to the localized approximations of the previous methods, collocation 3 (C3) treats the weight selection as a joint optimization problem, seeking to reduce the residual error across the entire set of collocation points simultaneously. Defining the vector  $\bar{\mathbf{w}}^\top = c[\bar{w}_f^{(i,1)} \dots \bar{w}_f^{(i,L_f^{(i)})}]$ , the weights are optimized as

$$\arg \min_{\bar{\mathbf{w}} \geq 0} \|\mathbf{E}\bar{\mathbf{w}} - \mathbf{p}\|_2^2 \quad (32)$$

where

$$\mathbf{E} = \begin{bmatrix} \mathcal{E} \left( \mathbf{m}_f^{(i,1)}; \mathbf{m}_f^{(i,1)}, \mathbf{P}_f^{(i,1)} \right) & \dots & \mathcal{E} \left( \mathbf{m}_f^{(i,1)}; \mathbf{m}_f^{(i,L_f^{(i)})}, \mathbf{P}_f^{(i,L_f^{(i)})} \right) \\ \vdots & \ddots & \vdots \\ \mathcal{E} \left( \mathbf{m}_f^{(i,L_f^{(i)})}; \mathbf{m}_f^{(i,1)}, \mathbf{P}_f^{(i,1)} \right) & \dots & \mathcal{E} \left( \mathbf{m}_f^{(i,L_f^{(i)})}; \mathbf{m}_f^{(i,L_f^{(i)})}, \mathbf{P}_f^{(i,L_f^{(i)})} \right) \end{bmatrix} \quad (33)$$

and

$$\mathbf{p}^\top = \left[ p^{(i)} \left( \mathbf{m}_f^{(i,1)} \right) g^{(i)} \left( \mathbf{m}_{t,f}^{(i,1)} \right) \quad \dots \quad p^{(i)} \left( \mathbf{m}_{t,f}^{(i,L_f^{(i)})} \right) g^{(i)} \left( \mathbf{m}_{t,f}^{(i,L_f^{(i)})} \right) \right] \quad (34)$$

When mixands take significantly different values at the various collocation points, the matrix  $\mathbf{E}$  can become ill-conditioned, thus complicating the numerical solution of the non-negative linear least squares problem (32). To mitigate convergence issues, Tikhonov regularization is applied to augment the system. As the augmentation introduces a bias in the solution, it is designed to only maintain the conditioning number below  $10^{10}$ . A full description of the method is available in appendix.

## Mixture Reduction

In the context of collaborative estimation, the pdfs the agents exchange are their respective filter posteriors. The product of mixtures is carried out by multiplying each mixand of the input densities in pairs. This combination returns a density with a number of mixands equal to the product of the number of mixands of the inputs. For instance, the non-reduced update factor approximation (21) is comprised of  $L_{k|k}^{(i)} \cdot L_{k|k}^{(j)}$  mixands. Thus, the product of the update factor and the prior in (14) produces a mixture of size  $L_{k|k}^{(i)} \cdot L_{k|k}^{(i)} \cdot L_{k|k}^{(j)} = (L_{k|k}^{(i)})^2 \cdot L_{k|k}^{(j)}$ . State-of-the-art reduction algorithms present a complexity on the order of the input number of mixands squared,<sup>29</sup> rendering direct reduction impractical for real-time usage.

To obtain tractable partial-state fusion, this paper proposes *preemptive* reduction operations employed at intermediate stages of the partial-state fusion computation. Reductions earlier in the fusion process has the potential to bring the highest computational savings. A challenge immediately arises in this approach in that

many of the intermediate functions are quotients represented by EQMs. Thus, existing reduction algorithms are not compatible with these functions, because they are not formal densities and lack moments in the general case.<sup>19</sup> This leaves three potential locations for reduction in the fusion process:

1. reduction of the common target state-marginal distributions before the computation of the update factor (21);
2. reduction of the less significant mixands during the application of the collocation condition, after the fusion output's mixands means (27) and covariances (26) are computed but before the final weights computations (30), (31), or (32); and
3. reduction of the fusion posterior output (25).

The first reduction location targets the update factor (21) to limit its mixand size, resulting in  $\tilde{L}_{k|k}^{(i)} \cdot \tilde{L}_{k|k}^{(j)}$  mixands. This process can be enacted by the same reduction algorithm used during filtering, where the pdfs would be reduced individually. Alternatively, this paper proposes identifying merge candidates based on their contributions to the exponentiated product. A desirable candidate metric is the Kullback-Leibler divergence (KLD), as it provides direct insight on the information discrepancy the mixands possess. In this sense, with little modification, the KLD can be used to infer the mixands of  $p_{k|k}^{(i)}$  that do not have excessive surprisal compared to those of  $p_{k|k}^{(j)}$ . Denote for brevity  $w_{k|k}^{(i,q)} \mathcal{N}_{k|k}^{(i,q)}$  the  $q^{\text{th}}$  mixand of agent  $i$ 's filter posterior pdf at time  $t_k$ , and, similarly,  $w_{k|k}^{(j,r)} \mathcal{N}_{k|k}^{(j,r)}$  the  $r^{\text{th}}$  mixand of agent  $j$ . Consider the log-likelihood ratio between  $\mathcal{N}_{k|k}^{(i,q)}$  and  $\mathcal{N}_{k|k}^{(j,r)}$ , averaged with respect to  $\mathcal{N}_{k|k}^{(i,q)}$  and weighted by its probability mass  $w_{k|k}^{(i,q)}$

$$\mathbb{E}_{\mathcal{N}_{k|k}^{(i,q)}} \left[ w_{k|k}^{(i,q)} \log \left( \frac{\mathcal{N}_{k|k}^{(i,q)}}{\mathcal{N}_{k|k}^{(j,r)}} \right) \right] = \int_{\mathbb{R}^{n_t}} \mathcal{N}_{k|k}^{(i,q)} w_{k|k}^{(i,q)} \log \left( \frac{\mathcal{N}_{k|k}^{(i,q)}}{\mathcal{N}_{k|k}^{(j,r)}} \right) d\mathbf{t}_{k|k} \quad (35)$$

$$= w_{k|k}^{(i,q)} \int_{\mathbb{R}^{n_t}} \mathcal{N}_{k|k}^{(i,q)} \log \left( \frac{\mathcal{N}_{k|k}^{(i,q)}}{\mathcal{N}_{k|k}^{(j,r)}} \right) d\mathbf{t}_{k|k} \quad (36)$$

$$= w_{k|k}^{(i,q)} D_{\text{KL}} \left( \mathcal{N}_{k|k}^{(i,q)} \parallel \mathcal{N}_{k|k}^{(j,r)} \right) \quad (37)$$

The KLD between Gaussian pdfs is available in closed form [31, Ch. 9]. This metric evaluates the log-likelihood information loss incurred by replacing  $\mathcal{N}_{k|k}^{(i,q)}$  with  $\mathcal{N}_{k|k}^{(j,r)}$ , scaled by mixand  $q$ 's mass. It presents similarities with the Runnalls' information loss upper bound.<sup>28</sup> While Runnalls' upper bound measures the cost of merging  $q$  and  $r$  into one, the present metric is the unilateral cost of substituting  $q$  with  $r$ . This distinction reflects the differing objectives of the measures: Runnalls' method seeks to reduce mixands of the same GM, whereas the present strategy aims to predict the contributions of a given mixand to the fusion product, as measured by the weighted cross-agent divergence. Thus

$$D_k^{(q,r)} = w_{k|k}^{(i,q)} D_{\text{KL}} \left( \mathcal{N}_{k|k}^{(i,q)} \parallel \mathcal{N}_{k|k}^{(j,r)} \right) \quad (38)$$

can be used to determine a threshold for mixand merging. The pairwise costs are used to populate the cross-relevance matrix

$$\mathbf{D}_k^{(i,j)} = \begin{bmatrix} D_k^{(1,1)} & \dots & D_k^{(1,L_k^{(j)})} \\ \vdots & \ddots & \vdots \\ D_k^{(L_k^{(i)},1)} & \dots & D_k^{(L_k^{(i)},L_k^{(j)})} \end{bmatrix} \quad (39)$$

As this reduction method seeks to merge mixands of lower stochastic significance, observe that mixand  $q$  is redundant for the fusion result if it presents negligible overlap with distribution  $p_{k|k}^{(j)}$ . Define the minimum

cross-agent discrepancy for each mixand  $q \in \{1, \dots, L_{k|k}^{(i)}\}$  as the minimum discrepancy relative to the target distribution  $p_{k|k}^{(j)}$  as

$$\delta_k^{(q)} = \min_{r \in \{1, \dots, L_{k|k}^{(j)}\}} \left\{ D_k^{(q,r)} \right\} \quad (40)$$

Then, the set of high-surprisal mixands  $\mathcal{Q}_k^{(i)}$  is identified as the collection of all the mixands of  $p_{k|k}^{(i)}$  whose nearest-neighbor divergence exceed a threshold proportional to the minimum surprisal between any pair of mixands

$$\mathcal{Q}_k^{(i)} = \left\{ q \in \{1, \dots, L_{k|k}^{(i)}\} : \delta_k^{(q)} > \epsilon \cdot \min_{q'} \left\{ \delta_k^{(q')} \right\} \right\} \quad (41)$$

with  $\epsilon > 1$ . By this construction, the set  $\mathcal{Q}_k^{(i)}$  contains the indices associated with the mixands of  $p_{k|k}^{(i)}$  that possess no significant overlap with any of the mixands of  $p_{k|k}^{(j)}$ , and, as such, can be prioritized for intra-distribution merging. The mixands whose indices belong to  $\mathcal{Q}_k^{(i)}$  are moment-matched by a single mixand.<sup>28</sup> As the KLD metric in (38) is asymmetric, the set  $\mathcal{Q}_k^{(j)}$  is constructed to identify merging opportunities within the mixands  $r \in \{1, \dots, L_{k|k}^{(j)}\}$  of  $p_{k|k}^{(j)}$  from (40) and (41).

The second integrated reduction stage operates on the outcome of C3. Consider the pivoted QR decomposition [32, Ch. 5] of the EQMs values matrix in (33)

$$\mathbf{E} = \mathbf{V}\mathbf{U}\mathbf{\Pi}^\top \quad (42)$$

where  $\mathbf{V}$  is an orthonormal matrix,  $\mathbf{U}$  an upper-triangular matrix, and  $\mathbf{\Pi}$  a permutation matrix ordering the absolute value of the diagonal elements of  $\mathbf{U}$  in non-increasing fashion (so that the first element  $|\mathbf{U}_{11}|$  is greater or equal to all others). Through this property of  $\mathbf{U}$ , the decomposition reveals which mixands are numerically redundant and cause  $\mathbf{E}$  to be near rank-deficient. In this sense, pivoted QR truncation can be interpreted as a greedy factor selection algorithm. For a tolerance  $0 < \delta < 1$ , all the mixands that correspond to a diagonal element lower than threshold

$$s_{\text{pQR}} = \delta \cdot \max \{ |\text{diag } \mathbf{U}| \} = \delta \cdot |\mathbf{U}_{11}| \quad (43)$$

can be merged together using the same moment-matching reduction algorithm. The proportional tolerance  $\delta$  should be chosen to balance numerical cost and approximation accuracy. Rank-revealing QR factorization methods can provide an insight on said trade-offs.

The third and last reduction concerns  $p_{k|k,f}^{(i)}$ , after fusion has occurred. The operation can be performed using, once again, the same reduction algorithms employed by the filter to directly cut down on the GM mixands  $\tilde{L}_{k|k,f}^{(i)} < (L_{k|k}^{(i)})^2 \cdot L_{k|k}^{(j)}$ . Due to the limitations outlined above, it would be more effective to apply this strategy in tandem with one or more of the previous methods.

Additionally, mixands means produced by (27) may result in near-zero likelihoods when used to evaluate the input distributions for the collocation conditions, indication of little statistical significance, and, as such, can be pruned before being processed. Also, although the fused GM covariances in (26) are analytically guaranteed to be positive definite, numerical instabilities may occasionally violate this property. In such instances, mixands with non-positive definite covariances are treated as numerical artifacts and discarded.

Algorithm 1 summarizes the process of fusing input distribution while enacting the reduction methods outlined in this section.

---

**Algorithm 1: Partial-State Gaussian Mixture Fusion and Reduction**

---

**Input:**  $p_{k|k}^{(i)}(\mathbf{x}_{k|k}^{(i)})$ ;  $p_{k|k}^{(j)}(\mathbf{t}_{k|k})$ ;  $\omega$ ; collocation condition; reduction parameters

**Output:**  $\tilde{p}_{k|k,f}^{(i)}(\mathbf{x}_{k|k}^{(i)})$

- 1  $p_{k|k}^{(i)}(\mathbf{t}_{k|k}) \leftarrow$  Obtain common target state marginal pdf from  $p_{k|k}^{(i)}(\mathbf{x}_{k|k}^{(i)})$
- 2  $\tilde{p}_{k|k}^{(i)}(\mathbf{t}_{k|k}), \tilde{p}_{k|k}^{(j)}(\mathbf{t}_{k|k}) \leftarrow$  Reduce common-state pdfs using Runnalls' algorithm and/or (41)
- 3  $g_{k|k,\text{FOPA}}^{(i)}(\mathbf{t}_{k|k}, \omega) \leftarrow$  Approximate fusion update factor with mixture according to (22)-(24)
- 4  $\left\{ \mathbf{m}_{k|k,f}^{(i,\ell)}, \mathbf{P}_{k|k,f}^{(i,\ell)} \right\}_{\ell=1}^{L_{k|k,f}^{(i)}} \leftarrow$  Compute fusion posterior GM  $p_{k|k,f}^{(i)}(\mathbf{x}_{k|k}^{(i)})$  mixand means and covariances via (27) and (26), respectively
- 5  $\left\{ \underline{\mathbf{m}}_{k|k,f}^{(i,\ell)}, \underline{\mathbf{P}}_{k|k,f}^{(i,\ell)} \right\}_{\ell=1}^{L_{k|k,f}^{(i)}} \leftarrow$  Prune statistically insignificant and numerically unstable mixands
- 6 **if** collocation condition is C3 (32) **then**
  - 7  $\left\{ \tilde{\mathbf{m}}_{k|k,f}^{(i,\ell)}, \tilde{\mathbf{P}}_{k|k,f}^{(i,\ell)} \right\}_{\ell=1}^{\tilde{L}_{k|k,f}^{(i)}} \leftarrow$  Reduce mixands using (43)
  - 8  $\left\{ w_{k|k,f}^{(i,\ell)} \right\}_{\ell=1}^{L_{k|k,f}^{(i)}} \leftarrow$  Compute fusion posterior GM  $p_{k|k,f}^{(i)}(\mathbf{x}_{k|k}^{(i)})$  mixands weights using C3 (32) and (18)
- 9 **else**
  - 10  $\left\{ w_{k|k,f}^{(i,\ell)} \right\}_{\ell=1}^{L_{k|k,f}^{(i)}} \leftarrow$  Compute fusion posterior GM  $p_{k|k,f}^{(i)}(\mathbf{x}_{k|k}^{(i)})$  mixands weights using C1 (30) or C2 (31) and (18)
- 11 **end**
- 12  $\tilde{p}_{k|k,f}^{(i)}(\mathbf{x}_{k|k}^{(i)}) \leftarrow$  Reduce  $p_{k|k,f}^{(i)}(\mathbf{x}_{k|k}^{(i)})$  using Runnalls' algorithm

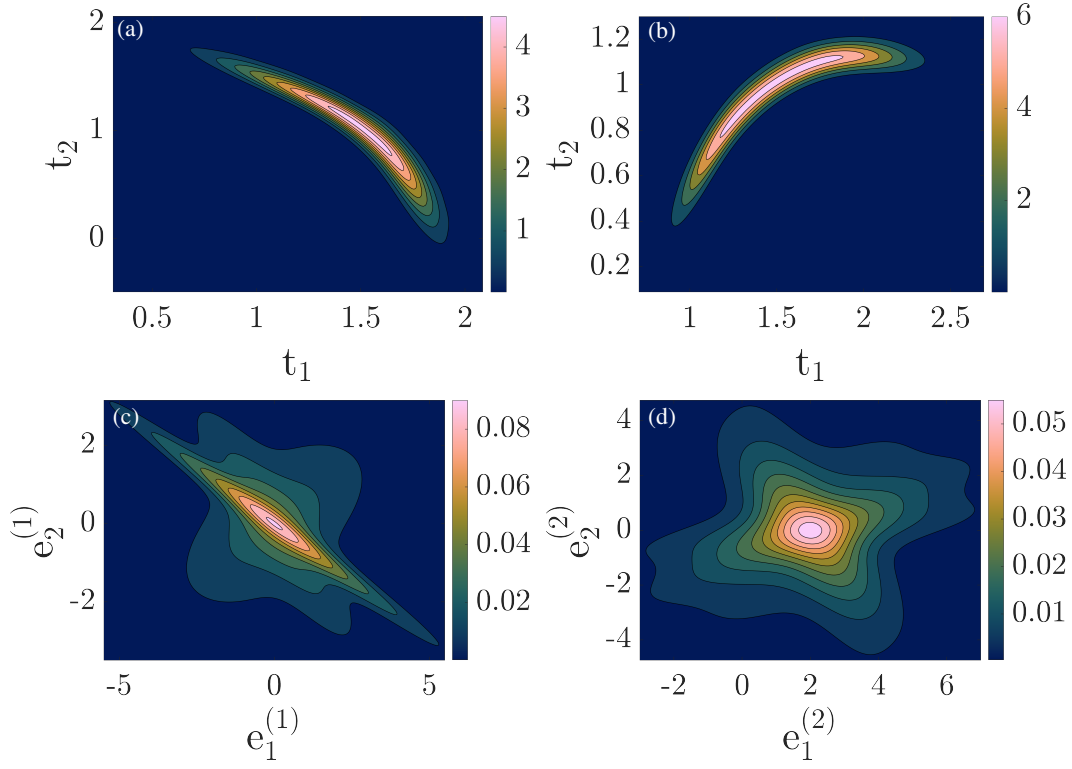
---

## RESULTS

In this section, the performance of collocation conditions and reduction criteria are analyzed separately to isolate the effects that each has on the overall fusion performance. Then, the framework is applied to a space-based SLAT problem. While sophisticated methods exist for dynamically choosing the fusion weight, this paper adopts a fixed fusion weight of  $\omega = 1/2$ , which is known as the Bhattacharyya exponent.<sup>6</sup> This choice avoids additional algorithmic complexity beyond the scope of this paper and, due to its constant and symmetric nature, promotes faster consensus among agents.

### Collocation Conditions Impact

To highlight the effect that collocation conditions have on fusion, a four-dimensional joint target and ego-state representation is examined. In this two-agent network ( $|\mathcal{I}| = 2$ ) toy-problem, the first two states,  $t_1$  and  $t_2$ , are the common-state beliefs, while the last two,  $e_1^{(\cdot)}$  and  $e_2^{(\cdot)}$ , are not shared between agents. The agents' marginal pdfs are constructed to emulate spatial uncertainty with low radial uncertainty but high directional uncertainty and are shown in Fig. 1.



**Figure 1:** Agent 1’s target (a) and ego (c) state marginal pdfs, before fusion; and Agent 2’s target (b) and ego (d) state marginal distributions.

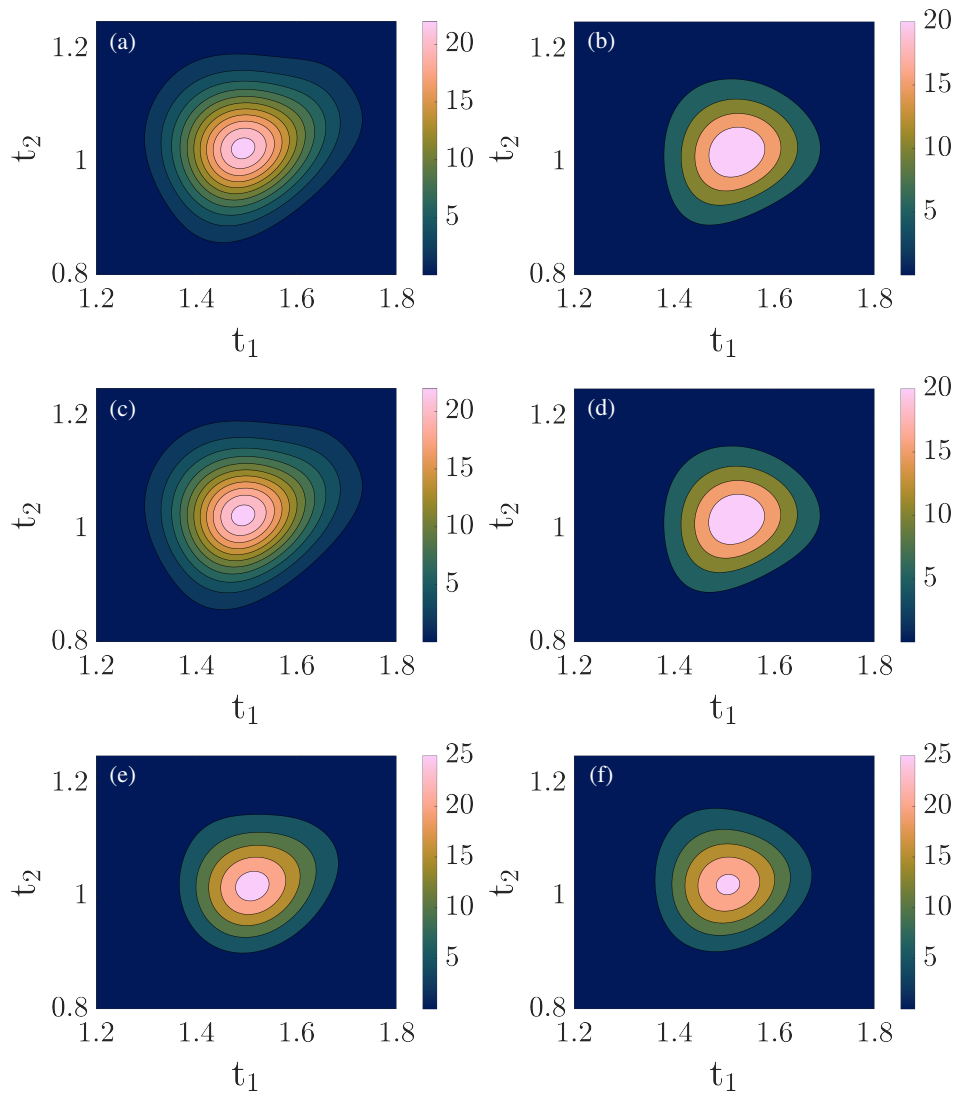
The accuracy of the GM approximations of (14), hereon denoted as  $\hat{p}_f^{(i)}$ , is measured by Jeffreys’ divergence

$$d(p, q) = \frac{1}{2} (\text{D}_{\text{KL}}(p \parallel q) + \text{D}_{\text{KL}}(q \parallel p)) \quad (44)$$

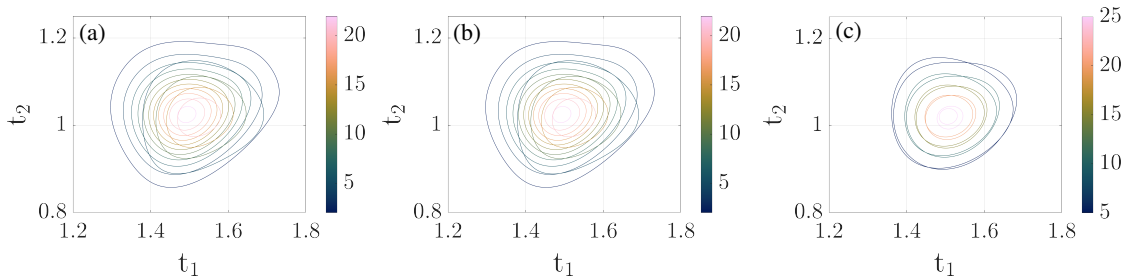
as, unlike KLD, it is a symmetric measure. The proposed distance is used to measure the disagreement between approximations and true fusion outcomes  $d(\hat{p}_f^{(i)}(\mathbf{x}^{(i)}), p_f^{(i)}(\mathbf{x}^{(i)}))$ ,  $i = 1, 2$ . While not available in closed form,  $p_f^{(i)}(\mathbf{x}^{(i)})$  in its general form (14) can be evaluated at arbitrary points. As the ultimate goal of the framework is to reach consensus over the common state, inter-agent target-marginal disagreement  $d(\hat{p}_f^{(1)}(\mathbf{t}), \hat{p}_f^{(2)}(\mathbf{t}))$  is also measured. The KLD integrals are computed using trapezoidal quadrature rule with 50 uniformly-spaced points per dimension, spanning the bounds shown in Fig. 2. Applying the mixture weight approximation C1 yields the target-marginal distributions in Fig. 2a, Fig. 2b, while C2 and C3 produce the pdfs in Fig. 2c, Fig. 2d, and Fig. 2e, Fig. 2f, respectively. The quantitative fusion results summarized in Table 1 show that the approximation error decreases from C1 to C2 to C3, with C3 significantly outperforming the other two methodologies. The same trend is visible also from Fig. 3, where the two inspectors’ target-marginal pdfs are overlapped; update C3 demonstrates the highest degree of overlap.

### Reduction Impact

Using the C3 collocation condition-based approximation, the effectiveness of each discussed reduction method is tested. Jeffrey’s divergence in (44) is again used to measure how reduction affects the fusion results. To maintain a fair comparison, the fusion outputs are reduced to  $\tilde{L}_f^{(1)} = 36$  and  $\tilde{L}_f^{(2)} = 37$  in all cases. Table 2 shows that information-based reduction methods do not significantly impact the distributions, while the computationally-oriented pivoted QR method causes the most discrepancy, mostly impacting the ego-state



**Figure 2:** C1 (top row), C2 (mid row), and C3 (bottom row) target-marginal pdfs approximations of the fused test distributions.



**Figure 3:** Agents' overlapped target-marginal distributions for all collocation conditions.

**Table 1:** Divergence values of different collocation conditions (lower is better).

Colloc. Cond.	$d(\hat{p}_f^{(1)}(\mathbf{x}^{(1)}), p_f^{(1)}(\mathbf{x}^{(1)}))$	$d(\hat{p}_f^{(2)}(\mathbf{x}^{(2)}), p_f^{(2)}(\mathbf{x}^{(2)}))$	$d(\hat{p}_f^{(1)}(\mathbf{t}), \hat{p}_f^{(2)}(\mathbf{t}))$
C1	0.089615	0.032246	0.077942
C2	0.089541	0.031239	0.077587
C3	0.0053904	0.015156	0.014354

marginal pdfs. In particular, the KLD method in (41) demonstrates maintenance of similar divergence levels to collocation C3 without reduction, even improving the target-marginal pdfs agreement.

**Table 2:** Divergence values of different reduction methods (lower is better).

Reduction	$d(\tilde{p}_f^{(1)}(\mathbf{x}^{(1)}), p_f^{(1)}(\mathbf{x}^{(1)}))$	$d(\tilde{p}_f^{(2)}(\mathbf{x}^{(2)}), p_f^{(2)}(\mathbf{x}^{(2)}))$	$d(\tilde{p}_f^{(1)}(\mathbf{t}), \tilde{p}_f^{(2)}(\mathbf{t}))$
Marginal Runnalls	0.014412	0.024011	0.02397
KLD (41)	0.0082481	0.021106	0.011728
Pivoted QR (43)	3.9933	1.8573	0.10516

### Application to Space-Based Simultaneous Localization and Tracking (SLAT)

The performance of the proposed setup is tested in a two-agent constellation SLAT scenario. The agents are satellites phased 30 [deg] along the same initial geostationary orbit (GEO) collaboratively estimating a GPS-like target satellite in medium Earth orbit (MEO). The agents estimate the target Earth-centered inertial (ECI) Cartesian position and velocity, here corresponding to their common state, alongside their own navigation ego-state, so that the state partitions in (3) become

$$\mathbf{t}_k = [\mathbf{r}_{t,k}^\top \quad \dot{\mathbf{r}}_{t,k}^\top]^\top = [x_{t,k} \quad y_{t,k} \quad z_{t,k} \quad \dot{x}_{t,k} \quad \dot{y}_{t,k} \quad \dot{z}_{t,k}]^\top \quad (45)$$

$$\mathbf{e}_k^{(i)} = \left[ \left( \mathbf{r}_{e,k}^{(i)} \right)^\top \quad \left( \dot{\mathbf{r}}_{e,k}^{(i)} \right)^\top \right]^\top = [x_{e,k}^{(i)} \quad y_{e,k}^{(i)} \quad z_{e,k}^{(i)} \quad \dot{x}_{e,k}^{(i)} \quad \dot{y}_{e,k}^{(i)} \quad \dot{z}_{e,k}^{(i)}]^\top \quad (46)$$

For simplicity, the motion of all satellites (observers and target) is modeled by simple two-body dynamics with perturbations from  $J_2$ , such that the discrete-time system dynamics map is the solution flow of [33, Ch. 8]

$$\frac{d}{dt} \begin{bmatrix} \mathbf{r}_{(\cdot),k}^{(\cdot)} \\ \dot{\mathbf{r}}_{(\cdot),k}^{(\cdot)} \end{bmatrix} = \begin{bmatrix} \dot{\mathbf{r}}_{(\cdot),k}^{(\cdot)} \\ -\mu \frac{\mathbf{r}_{(\cdot),k}^{(\cdot)}}{\|\mathbf{r}_{(\cdot),k}^{(\cdot)}\|^3} + \frac{1}{2} J_2 \mu R^2 \left( 15 \mathbf{r}_{(\cdot),k}^{(\cdot)} \frac{(z_{(\cdot),k}^{(\cdot)})^2}{\|\mathbf{r}_{(\cdot),k}^{(\cdot)}\|^2} - 3 \mathbf{r}_{(\cdot),k}^{(\cdot)} \frac{(z_{(\cdot),k}^{(\cdot)})^2}{\|\mathbf{r}_{(\cdot),k}^{(\cdot)}\|^2} - \begin{bmatrix} 0 \\ 0 \\ 6z_{(\cdot),k}^{(\cdot)} \end{bmatrix} \right) \end{bmatrix} \quad (47)$$

where  $\mu$ ,  $R$ , and  $J_2$  are, respectively, Earth's gravitational parameter, equatorial radius, and harmonic factor due to its oblateness. Then, as an abuse of notation for simplicity,  $\mathbf{f}_k$  is used to denote both the individual and stacked solution flows of (47). Under this convention, the dynamics in (1) estimated by each agent can be written as

$$\mathbf{f}_k(\mathbf{x}_k^{(i)}) = \begin{bmatrix} \mathbf{f}_k(\mathbf{t}_k) \\ \mathbf{f}_k(\mathbf{e}_k^{(i)}) \end{bmatrix} \quad (48)$$

where the specific function is determined by the nature of its argument. The observers take topocentric angular measurements of the target, with measurement map (2)

$$\mathbf{h}_k^{(i)}(\mathbf{x}_k^{(i)}) = \left[ \arcsin \left( \frac{y_{t,k} - y_{e,k}^{(i)}}{\|\mathbf{r}_{t,k} - \mathbf{r}_{e,k}^{(i)}\|_2} \right) \quad \arcsin \left( \frac{z_{t,k} - z_{e,k}^{(i)}}{\|\mathbf{r}_{t,k} - \mathbf{r}_{e,k}^{(i)}\|_2} \right) \right]^\top \quad (49)$$

While not strictly required by the distributed estimation formulation, the agents are assumed to be identical to each other, in terms of sensors specifications and filtering approach. Each agent is equipped with an optical instrument similar to OSIRIS-REx' PolyCam<sup>34</sup> and performs joint tracking and navigation using an adaptive Gaussian mixture square-root unscented Kalman filter (AGMSRUKF). The simulation spans 24 hours, during which agents acquire target angle-only measurements every 2 hours when visibility permits. To model measurement unavailability, an observation gap is imposed over the interval from hour 10 to hour 20. The onboard camera has a  $0.8^\circ \times 0.8^\circ$  field-of-view (FoV) and a  $1024 \times 1024$  pixel focal plane array. The corresponding pixel scale is 2.8 [arcsecond/pixel]. Assuming 5 pixels standard deviation of error, the corresponding measurement error standard deviation is 14 [arcsecond] per angle. Reflecting the sparsity of effective windows, a single precise measurement consisting of target full-state ephemeris information of covariance  $\mathbf{R}_{\text{precise}}^{(i)}$  is made available to both agents at time  $t = 22$  [hr]. The process noise covariance values are also the same across the constellation and constant over time and defined as

$$\mathbf{Q}^{(i)} = \text{blockdiag} \left( \begin{bmatrix} \mathbf{Q}_t & \mathbf{Q}_e^{(i)} \end{bmatrix} \right) \quad (50)$$

$$\mathbf{Q}_t = \mathbf{Q}_e^{(i)} = \text{diag} \left( [q_r^2 \quad q_r^2 \quad q_r^2 \quad q_r^2 \quad q_r^2 \quad q_r^2] \right) \quad (51)$$

The initial uncertainties are provided in RSW frame

$$\mathbf{P}_0^{(i)} = \text{blockdiag} \left( \begin{bmatrix} \mathbf{P}_{t,0} & \mathbf{P}_{e,0}^{(i)} \end{bmatrix} \right) \quad (52)$$

$$\mathbf{P}_{t,0} = \mathbf{T}_t \boldsymbol{\Sigma}_t (\mathbf{T}_t \boldsymbol{\Sigma}_t)^\top \quad (53)$$

$$\boldsymbol{\Sigma}_t = \text{diag} \left( \begin{bmatrix} (\boldsymbol{\sigma}_{t,r})^\top & (\boldsymbol{\sigma}_{t,\dot{r}})^\top \end{bmatrix} \right) \quad (54)$$

$$\mathbf{P}_{e,0}^{(i)} = \mathbf{T}_e^{(i)} \boldsymbol{\Sigma}_e^{(i)} (\mathbf{T}_e^{(i)} \boldsymbol{\Sigma}_e^{(i)})^\top \quad (55)$$

$$\boldsymbol{\Sigma}_e^{(i)} = \text{diag} \left( \begin{bmatrix} (\boldsymbol{\sigma}_{e,r}^{(i)})^\top & (\boldsymbol{\sigma}_{e,\dot{r}}^{(i)})^\top \end{bmatrix} \right) \quad (56)$$

where

$$\mathbf{T}_{(\cdot)}^{(\cdot)} = \text{blockdiag} \left( \begin{bmatrix} \mathbf{T}_{(\cdot),r}^{(\cdot)} & \mathbf{T}_{(\cdot),\dot{r}}^{(\cdot)} \end{bmatrix} \right) \quad (57)$$

$$\mathbf{T}_{(\cdot),r}^{(\cdot)} = \mathbf{T}_{(\cdot),\dot{r}}^{(\cdot)} = \begin{bmatrix} \frac{\mathbf{r}_{(\cdot)}^{(\cdot)}}{\|\mathbf{r}_{(\cdot)}^{(\cdot)}\|_2} & \frac{\mathbf{r}_{(\cdot)}^{(\cdot)} \times \dot{\mathbf{r}}_{(\cdot)}^{(\cdot)}}{\|\mathbf{r}_{(\cdot)}^{(\cdot)} \times \dot{\mathbf{r}}_{(\cdot)}^{(\cdot)}\|_2} \times \frac{\mathbf{r}_{(\cdot)}^{(\cdot)}}{\|\mathbf{r}_{(\cdot)}^{(\cdot)}\|_2} & \frac{\mathbf{r}_{(\cdot)}^{(\cdot)} \times \dot{\mathbf{r}}_{(\cdot)}^{(\cdot)}}{\|\mathbf{r}_{(\cdot)}^{(\cdot)} \times \dot{\mathbf{r}}_{(\cdot)}^{(\cdot)}\|_2} \end{bmatrix} \quad (58)$$

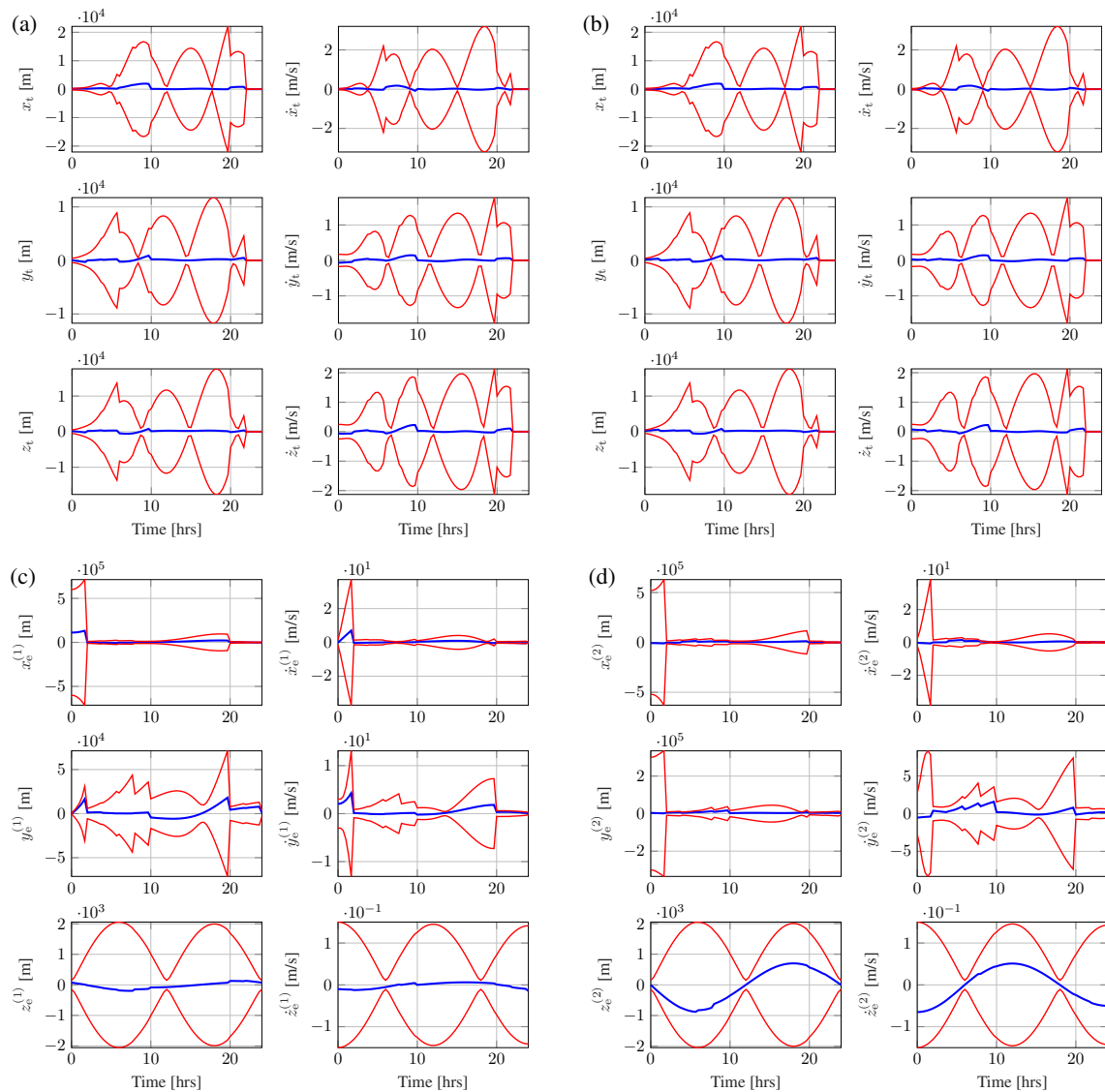
is the transformation matrix from RSW frame to inertial frame. The ECI initial states of all satellites and model uncertainty values are reported in Table 3.

Fusion between agents is performed every time a measurement is available, and reduction is enacted on the target-marginal distributions via the novel KLD method (41) with  $\epsilon = 40$  and on the fusion posterior using Runnalls' algorithm.<sup>28</sup> The metric used to split for the dynamics is uncertainty-scaled first-order stretching (US-FOS) for its simplicity and accuracy for the problem in consideration.<sup>23</sup>

As the consensus fusion is an agreement of pdfs, the loss of accuracy in one agent might correspond to an increase in the other. For this reason, the root-mean square error (RMSE) is computed across the constellation, separately on target and ego-state positions and velocities. The filters show consistent results in Fig. 4, with major reductions in target estimation uncertainty when the precise ephemeris measurement is assimilated. Comparing the outcome of the same scenario without fusion, Fig. 5 highlights the balance in information exchange. For example, at the 8 [hr] mark, agent 1's target estimate loses accuracy in favor of agent 2's improvement. As a general trend, when looking at the RMSEs in Fig. 6, fusing brings a net benefit for the constellation estimation performance.

## CONCLUSIONS

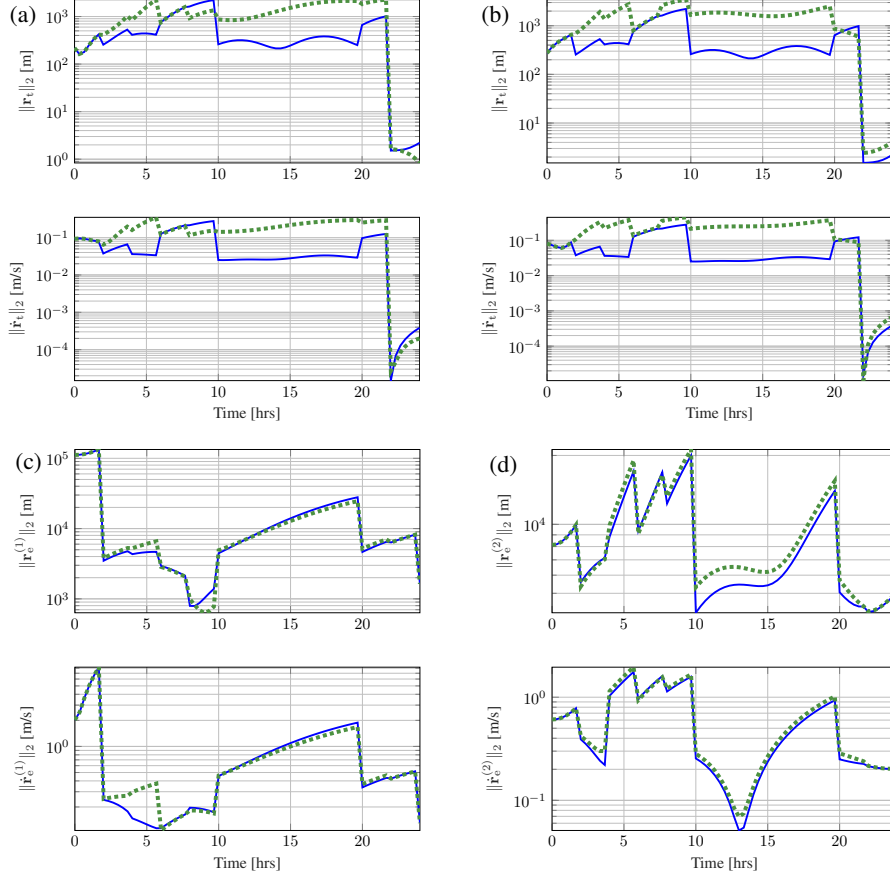
This paper presents a tractable Gaussian mixture based algorithm for distributed simultaneous localization and tracking in nonlinear non-Gaussian settings. Algorithmic variants of varying complexity and accuracy



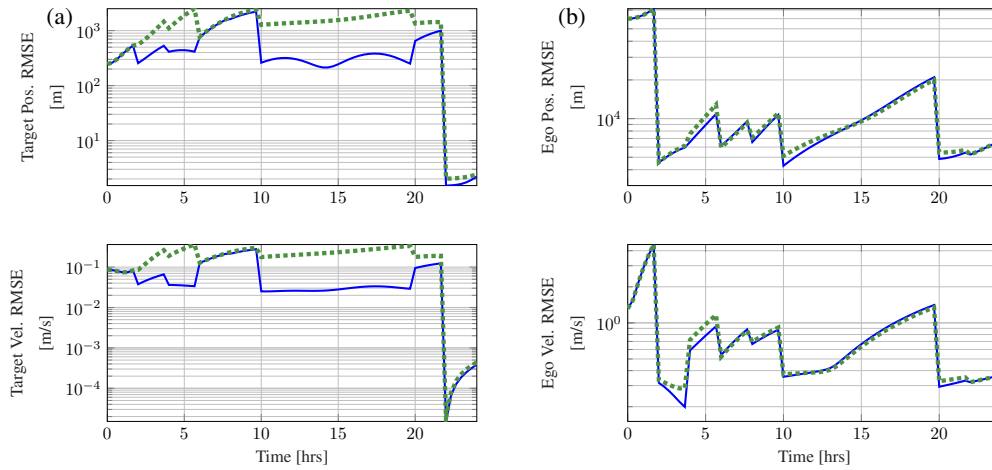
**Figure 4:** Estimation error (blue) and  $3\sigma$  uncertainty (red); agent 1's estimation is portrayed in plots (a) and (c), agent 2's in (b) and (d), top is target states and bottom is ego-states.

**Table 3:** Simulation initial variables and model assumptions.

Parameter	Value
$\mathbf{r}_{t,0}$	$[26491.58 \ 2980.266 \ -357.1760]^\top$ [km]
$\dot{\mathbf{r}}_{t,0}$	$[-0.173028 \ 2.191596 \ 3.17845]^\top$ [km/s]
$\mathbf{r}_{e,0}^{(1)}$	$[-42241.10 \ 0 \ 0]^\top$ [km]
$\dot{\mathbf{r}}_{e,0}^{(1)}$	$[0 \ 3.071859 \ 0]^\top$ [km/s]
$\mathbf{r}_{e,0}^{(2)}$	$[36581.86 \ -21120.55 \ 0]^\top$ [km]
$\dot{\mathbf{r}}_{e,0}^{(2)}$	$[1.535930 \ 2.660308 \ 0]^\top$ [km/s]
$\mathbf{R}_{\text{precise}}^{(1)} = \mathbf{R}_{\text{precise}}^{(2)}$	$\text{diag}([1 \ 1 \ 1 \ 10^{-10} \ 10^{-10} \ 10^{-10}])$ [ $\text{m}^2, [\text{m}^2/\text{s}^2]$ ]
$q_r$	$10^{-8}$ [m]
$q_{\dot{r}}$	$10^{-6}$ [m/s]
$\boldsymbol{\sigma}_{t,r}$	$[30000 \ 200 \ 1000]^\top$ [m]
$\boldsymbol{\sigma}_{t,\dot{r}}$	$[1 \ 1 \ 0.05]^\top$ [m/s]
$\boldsymbol{\sigma}_{e,r}^{(1)} = \boldsymbol{\sigma}_{e,r}^{(2)}$	$[500 \ 1000 \ 50]^\top$ [m]
$\boldsymbol{\sigma}_{e,\dot{r}}^{(1)} = \boldsymbol{\sigma}_{e,\dot{r}}^{(2)}$	$[0.5 \ 1 \ 0.05]^\top$ [m/s]



**Figure 5:** Target (top row) and ego-state (bottom row) position and velocity log-scaled norm with fusion (solid blue) and without fusion (dashed green) errors over time; agent 1's results on the left and agent 2's on the right.



**Figure 6:** Constellation target (a) and ego state (b) position and velocity RMSEs, fusion in solid blue, without fusion in dashed green.

are discussed and analyzed in the context of a spacecraft-to-spacecraft joint navigation and tracking scenario. Particular focus is dedicated to mixture management, and novel mixture reduction algorithms are proposed for tractable application in online estimation applications. The overall approach shows great promise for large scale heterogeneous networks, as by construction, no agent-specific navigation information is exchanged within the network. Yet, the approach preserves target and ego-state correlations and thus derives implicit navigation information through the explicit exchange of target-marginal beliefs. Simulations demonstrate that the proposed distributed fusion algorithm improves overall constellation-wide estimation performance, occasionally showing that a reduction in estimation performance for one agent improves the performance of the other.

## ACKNOWLEDGMENT

The authors would like to thank the Boeing Company’s Enterprise Technology Office for their support under contract SSOB-BRT-Z0722-5045. This material is also based upon work supported by the Air Force Office of Scientific Research under award number FA9550-25-1-0101.

## REFERENCES

- [1] A. Oliveira, K. Dyreby, F. Caldas, and C. Soares, “OrbitZoo: Real Orbital Systems Challenges for Reinforcement Learning,” Oct. 2025, 10.48550/arXiv.2504.04160.
- [2] Z. Chen and Y. Zheng, “Persistent and Responsive Collective Motion with Adaptive Time Delay,” *Science Advances*, Vol. 10, Apr. 2024, p. eadk3914, 10.1126/sciadv.adk3914.
- [3] S. Z. Zhao, H. Xiang, C. Xu, X. Xia, B. Zhou, and J. Ma, “CooPre: Cooperative Pretraining for V2X Cooperative Perception,” June 2025, 10.48550/arXiv.2408.11241.
- [4] O. Hlinka, F. Hlawatsch, and P. M. Djuric, “Distributed Particle Filtering in Agent Networks: A Survey, Classification, and Comparison,” *IEEE Signal Processing Magazine*, Vol. 30, Jan. 2013, pp. 61–81, 10.1109/MSP.2012.2219652.
- [5] A. Buonviri, M. York, K. LeGrand, and J. Meub, “Survey of Challenges in Labeled Random Finite Set Distributed Multi-Sensor Multi-Object Tracking,” *2019 IEEE Aerospace Conference*, Mar. 2019, pp. 1–12, 10.1109/AERO.2019.8742216.
- [6] C.-Y. Chong, K.-C. Chang, and S. Mori, “Fundamentals of Distributed Estimation,” *Distributed Data Fusion for Network-Centric Operations*, CRC Press, 2013.
- [7] J. K. Uhlmann, “General Data Fusion for Estimates with Unknown Cross Covariances,” *Signal Processing, Sensor Fusion, and Target Recognition V*, Vol. 2755, SPIE, June 1996, pp. 536–547, 10.1117/12.243195.

- [8] S. Lubold and C. N. Taylor, “Formal Definitions of Conservative Probability Distribution Functions (PDFs),” *Information Fusion*, Vol. 88, Dec. 2022, pp. 175–183, 10.1016/j.inffus.2022.07.014.
- [9] R. P. S. Mahler, “Optimal/Robust Distributed Data Fusion: A Unified Approach,” *Signal Processing, Sensor Fusion, and Target Recognition IX*, Vol. 4052, SPIE, Aug. 2000, pp. 128–138, 10.1117/12.395064.
- [10] B. Wang, W. Yi, R. Hoseinnezhad, S. Li, L. Kong, and X. Yang, “Distributed Fusion With Multi-Bernoulli Filter Based on Generalized Covariance Intersection,” *IEEE Transactions on Signal Processing*, Vol. 65, Jan. 2017, pp. 242–255, 10.1109/TSP.2016.2617825.
- [11] S. Li, W. Yi, R. Hoseinnezhad, G. Battistelli, B. Wang, and L. Kong, “Robust Distributed Fusion With Labeled Random Finite Sets,” *IEEE Transactions on Signal Processing*, Vol. 66, Jan. 2018, pp. 278–293, 10.1109/TSP.2017.2760286.
- [12] L. Gao, G. Battistelli, and L. Chisci, “Fusion of Labeled RFS Densities With Minimum Information Loss,” *IEEE Transactions on Signal Processing*, Vol. 68, 2020, pp. 5855–5868, 10.1109/TSP.2020.3028496.
- [13] W. Yi, S. Li, B. Wang, R. Hoseinnezhad, and L. Kong, “Computationally Efficient Distributed Multi-Sensor Fusion With Multi-Bernoulli Filter,” *IEEE Transactions on Signal Processing*, Vol. 68, 2020, pp. 241–256, 10.1109/TSP.2019.2957638.
- [14] W. Wu, H. Sun, W. Huang, M. Zheng, and X. Feng, “Multi-GMTI Decentralized Tracking via Consensus LMB Density Fusion,” *2021 International Conference on Control, Automation and Information Sciences (ICCAIS)*, Oct. 2021, pp. 122–129, 10.1109/ICCAIS52680.2021.9624574.
- [15] T. Li, R. Yan, K. Da, and H. Fan, “Arithmetic Average Density Fusion-Part III: Heterogeneous Unlabeled and Labeled RFS Filter Fusion,” *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 60, Feb. 2024, pp. 1023–1034, 10.1109/TAES.2023.3334223.
- [16] T. Li, H. Liang, G. Li, J. García Herrero, and Q. Pan, “Arithmetic Average Density Fusion—Part IV: Distributed Heterogeneous Fusion of RFS and LRFS Filters via Variational Approximation,” *IEEE Transactions on Signal Processing*, Vol. 73, 2025, pp. 1454–1469, 10.1109/TSP.2025.3550157.
- [17] N. Ahmed, “Conditionally Factorized DDF for General Distributed Bayesian Estimation,” *2014 International Conference on Multisensor Fusion and Information Integration for Intelligent Systems (MFI)*, Sept. 2014, pp. 1–7, 10.1109/MFI.2014.6997717.
- [18] N. R. Ahmed, W. W. Whitacre, S. Moon, and E. W. Frew, “Factorized Covariance Intersection for Scalable Partial State Decentralized Data Fusion,” *2016 19th International Conference on Information Fusion (FUSION)*, July 2016, pp. 1049–1056.
- [19] S. Semeraro, K. A. LeGrand, and J. Kulik, “Partial State Gaussian Mixture Fusion for Distributed Navigation and Tracking,” *2025 AAS/AIAA Astrodynamics Specialist Conference*, Boston, Aug. 2025.
- [20] M. F. Huber, T. Bailey, H. Durrant-Whyte, and U. D. Hanebeck, “On Entropy Approximation for Gaussian Mixture Random Vectors,” *2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Aug. 2008, pp. 181–188, 10.1109/MFI.2008.4648062.
- [21] V. Vittaldev and R. Russell, “Multidirectional Gaussian Mixture Models for Nonlinear Uncertainty Propagation,” *Computer Modeling in Engineering & Sciences*, Vol. 111, No. 1, 2016, pp. 83–117, 10.3970/cmescs.2016.111.083.
- [22] M. Losacco, A. Fossà, and R. Armellin, “Low-Order Automatic Domain Splitting Approach for Nonlinear Uncertainty Mapping,” *Journal of Guidance, Control, and Dynamics*, Vol. 47, Feb. 2024, pp. 291–310, 10.2514/1.G007271.
- [23] J. Kulik and K. A. LeGrand, “Nonlinearity and Uncertainty Informed Moment-Matching Gaussian Mixture Splitting,” *IEEE Transactions on Aerospace and Electronic Systems*, 2025, pp. 1–21, 10.1109/TAES.2025.3632242.
- [24] K. Tuggle and R. Zanetti, “Automated Splitting Gaussian Mixture Nonlinear Measurement Update,” *Journal of Guidance, Control, and Dynamics*, Vol. 41, No. 3, 2018, pp. 725–734, 10.2514/1.G003109.
- [25] U. D. Hanebeck, K. Briechle, and A. Rauh, “Progressive Bayes: A New Framework for Nonlinear State Estimation,” *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2003*, Vol. 5099, SPIE, Apr. 2003, pp. 256–267, 10.1117/12.487806.
- [26] M. F. Huber and U. D. Hanebeck, “Progressive Gaussian Mixture Reduction,” *2008 11th International Conference on Information Fusion*, June 2008, pp. 1–8.
- [27] K. J. DeMars, R. H. Bishop, and M. K. Jah, “Entropy-Based Approach for Uncertainty Propagation of Nonlinear Dynamical Systems,” *Journal of Guidance, Control, and Dynamics*, Vol. 36, July 2013, pp. 1047–1057, 10.2514/1.58987.
- [28] A. Runnalls, “Kullback-Leibler Approach to Gaussian Mixture Reduction,” *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 43, July 2007, pp. 989–999, 10.1109/TAES.2007.4383588.
- [29] D. F. Crouse, P. Willett, K. Pattipati, and L. Svensson, “A Look at Gaussian Mixture Reduction Algorithms,” *14th International Conference on Information Fusion*, July 2011, pp. 1–8.

- [30] M. Gunay, U. Orguner, and M. Demirekler, “Chernoff Fusion of Gaussian Mixtures Based on Sigma-Point Approximation,” *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 52, Dec. 2016, pp. 2732–2746, 10.1109/TAES.2016.150403.
- [31] Solomon. Kullback, *Information Theory and Statistics*. Dover Books on Mathematics, Mineola, N.Y: Dover Publications, dover ed. ed., 1997–1968.
- [32] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences, Philadelphia, PA: The Johns Hopkins University Press, fourth edition ed., 2013, 10.1137/1.9781421407944.
- [33] D. A. Vallado and W. D. McClain, *Fundamentals of Astrodynamics and Applications*. Space Technology Library ; v. 21, Hawthorne: Microcosm, 4th ed. ed., 2013.
- [34] B. Rizk et al., “OCAMS: The OSIRIS-REx Camera Suite,” *Space Science Reviews*, Vol. 214, Jan. 2018, p. 26, 10.1007/s11214-017-0460-7.

## APPENDIX

### Tikhonov Regularization

One of the most common ways to represent the conditioning number  $\kappa$  of a matrix  $\mathbf{A}$  is expressing it as the ratio of its highest and smallest nonzero singular values. When solving the least-square problem

$$\arg \min_{\mathbf{x}} \|\mathbf{Ax} - \mathbf{b}\|_2^2 \quad (59)$$

if  $\mathbf{A}$  is ill-conditioned, numerical solutions might yield unreliable results. Tikhonov regularization offers to augment the system as  $\bar{\mathbf{A}}^\top = [\mathbf{A}^\top \ \alpha \mathbf{I}]^\top$  and  $\bar{\mathbf{b}}^\top = [\mathbf{b}^\top \ \mathbf{0}^\top]$ , with  $\alpha \geq 0$  the Tikhonov factor. The tunable factor can be directly computed from a user-requested tolerance on the augmented matrix conditioning number  $\kappa(\bar{\mathbf{A}})$

$$\bar{\mathbf{A}} = \begin{bmatrix} \mathbf{A} \\ \alpha \mathbf{I} \end{bmatrix} \quad (60)$$

$$\bar{\mathbf{A}}\bar{\mathbf{A}}^\top = \begin{bmatrix} \mathbf{A} \\ \alpha \mathbf{I} \end{bmatrix} [\mathbf{A}^\top \ \alpha \mathbf{I}^\top] = \mathbf{AA}^\top + \alpha^2 \mathbf{I} \quad (61)$$

$$\lambda_i(\bar{\mathbf{A}}\bar{\mathbf{A}}^\top) = \lambda_i(\mathbf{AA}^\top) + \alpha^2 \quad (62)$$

$$\sigma_i^2(\bar{\mathbf{A}}) = \sigma_i^2(\mathbf{A}) + \alpha^2 \quad (63)$$

$$\sigma_i(\bar{\mathbf{A}}) = \sqrt{\sigma_i^2(\mathbf{A}) + \alpha^2} \quad (64)$$

$$\kappa(\bar{\mathbf{A}}) = \sqrt{\frac{\sigma_i^2(\mathbf{A}) + \alpha^2}{\sigma_n^2(\mathbf{A}) + \alpha^2}} \quad (65)$$

$$\alpha = \sqrt{\frac{\sigma_i^2 - \kappa^2(\bar{\mathbf{A}})\sigma_n^2(\mathbf{A})}{\kappa^2(\bar{\mathbf{A}}) - 1}} \quad (66)$$

This strategy makes the problem well-posed, albeit biasing the solution of (59) towards the origin

$$\arg \min_{\mathbf{x}} \|\mathbf{Ax} - \mathbf{b}\|_2^2 + \|\alpha \mathbf{Ix}\|_2^2 \quad (67)$$

For this reason, the value of  $\alpha$  should be selected to balance concerns about conditioning and biasing.