A multisensor fusion Student’s $t$ filter is proposed for time-series recursive estimation in the presence of heavy-tailed process and measurement noises. It extends the single-sensor Student’s $t$ Kalman filter to the multisensor setup based on the suboptimal arithmetic average (AA) fusion approach which is driven from information-theoretic density fusion optimization and able to deal with unknown correlation among sensors. To ensure computationally efficient, closed-form $t$ density recursion, moment matching approximation has been used for averaging the $t$ densities aggregated from different sensors. Based on the same framework, we also extend the covariance intersection (CI) approach for $t$ density fusion. Simulation demonstrates the strength of the proposed multisensor AA fusion-based $t$ filter in dealing with outliers as compared with the classic Gaussian estimator, and the advantage of the AA fusion in comparison with the CI approach and the augmented measurement fusion.

1. INTRODUCTION

The state-space model (SSM) for state estimation models the uncertainties/randomness of the state dynamics and of the measurement by respective noises. In practice, heavy-tailed noises are involved in many highly uncertain problems, e.g., tracking scenarios with agile/manoeuvering targets and outlier-corrupted measurements [1]. In these situations, even when the SSM is linear, the performance of the predominant Kalman filter (KF) that models the process and measurement noises as Gaussian often deteriorates. This gives rise to outlier-robust estimator that takes advantage of heavy-tailed Student’s $t$ distribution for modeling the noises [2], [3], [4], [5], [6]. The $t$ distribution can be viewed as a generalized Gaussian distribution that has an adjustable parameter referred to as the degree of freedom (dof) to uplift the tail of the distribution [7]. As the key to drive the Bayes recursive Student’s $t$ filter, the joint probability density function (PDF) of the state and noises is assumed to be Student’s $t$, and the prediction/posterior state PDF is then approximated as Student’s $t$ [1], [4], [5] based on linear transformation of $t$ distribution. A variety of strategies have
been further proposed for dealing with nonlinearity using such as the Monte–Carlo method [8], [9], linearization [4], and unscented transform [5], [10].

So far, the majority of existing Student’s t filters are implemented on the base of a single sensor except for few exceptions that seek optimal fusion of Student’s t obtained by multiple sensors in the sense of minimum variance (MV) estimation [11], [12]. First, these approaches rely restrictively on accurate knowledge (or neglect) of the correlation (or to say, the common information [13], [14], [15]) among sensors. Failure to properly account for the correlation will lead to inconsistent estimation (e.g., underestimating the actual squared estimate error) and even filter breakdown. To say the least, even if the sensors are uncorrelated, the multisensor Bayes-optimal Student’s t filter does not admit closed-form recursion and has to be approximated. Second, the key of the MV fusion is to get as accurate an estimate as possible which is lack of congruity with the idea of the heavy-tailed estimator where the robustness-to-outliers stems from a heavy-tailed posterior. These observations motivate us to develop novel fusion approach that does not need to calculate the intersensor correlation yet avoids inconsistent estimate and preserves the heavy tail of the t estimator. Furthermore, when a peer-to-peer sensor network that is highly restrictive in both communication and computation is involved, computation efficiency and tolerance to node faults are also vital factors that need to be taken into account.

In this article, we propose a multisensor Student’s t filter which does not seek optimal fusion like MV but suboptimal, robust fusion in tune with heavy-tailed estimation and peer-to-peer networking. To this end, we employ the arithmetic average (AA) density fusion which, often applied jointly with the distributed consensus/diffusion/flooding approach and case-specific finite mixture reduction strategies, was earlier proposed for multitarget density fusion in the presence of missing and false data [16], [17], [18], [19], [20], [21], [22], [23]; see the cutting-edge reviews [24], [25]. It has recently been advocated for the fundamental KF fusion [26] and belief fusion [27]. In short, the AA fusion has demonstrated high efficiency in computation and tolerance to sensor fault (such as misdetection), and accommodates any degree of intersensor correlation. More relevantly, the AA density fusion is rooted in a dispersive mixture/multimodal expression of the target distribution which is essentially robust to outlier and complies with the motivation for heavy-tailed estimator design.

However, the straightforward application of the AA fusion to the Student’s t distribution leads to a Student’s t mixture which has to be approximated by a single t distribution for the purpose of closed-formed recursive filtering. Furthermore, the optimization of the fusion weights drives a need of Kullback–Leibler (KL) divergence regarding t distributions, which does not admit analytical solution. To solve these challenges, we resort to ideas in dealing with Gaussian-AA fusion [26] through approximating calculation regarding t distributions by that of Gaussian distributions with matched mean and covariance. To concentrate our key contribution on Student’s t AA fusion, the single-sensor filter we employ is the Student’s t filter based on simplified dof choice [1], [4]. All derivations are made from first principles. As an extension of the proposed multisensor Student’s t filter, the covariance intersection (CI) [14] approach advocated originally for Gaussian fusion has also been employed in comparison with the AA fusion.

The remainder of this article is organized as follows. Background is briefly introduced in Section II. The proposed Student’s t filter based on the AA fusion approach is driven in Section III, which is extended to the CI fusion in Section IV. Simulation and comparison study are given in Section V. Section VI concludes the article.

II. BACKGROUND

A. SSM With Student’s t Noises

We consider the following discrete-time SSM with additive noises:

\[ x_k = f_{k-1}(x_{k-1}) + w_{k-1} \]
\[ z_k = h_k(x_k) + v_k \]

where \( k \) is the discrete time index, \( x_k \in \mathbb{R}^n \) is the state vector, \( z_k \in \mathbb{R}^m \) is the measurement vector, and \( f_{k-1}(\cdot) \) and \( h_k(\cdot) \) are the process and measurement functions, respectively, and \( w_k \) and \( v_k \) are the process and measurement noises, respectively. We further denote the Jacobian matrices of \( f_{k-1}(\cdot) \) and \( h_k(\cdot) \) by \( \mathbf{F}_{k-1} \) and \( \mathbf{H}_k \), respectively, which are useful for linearizing the nonlinear models. In this article, we particularly consider the heavy-tailed Student’s t noise, namely,

\[ p(w_k) = \mathcal{S}(w_k; 0, Q_k, v_Q) \]
\[ p(v_k) = \mathcal{S}(v_k; 0, R_k, v_R) \]

where \( \mathcal{S}(\cdot; \mu, \Sigma, v) \) denotes the Student’s t PDF with mean vector \( \mu \), scale matrix \( \Sigma \), and degree of freedom (dof) parameter \( v \), and \( Q_k \) and \( v_Q \) are the scale matrix and dof parameter of the process noise, respectively. The Student’s t PDF \( \mathcal{S}(x; \mu, \Sigma, v) \) can be written as [7]

\[ \mathcal{S}(x; \mu, \Sigma, v) = \frac{\Gamma\left(\frac{v+n}{2}\right)}{(\Gamma\left(\frac{v}{2}\right)|\Sigma|^{1/2}(\pi v)^{n/2})} \left(1 + \frac{(x - \mu)^T \Sigma^{-1} (x - \mu)}{v} \right)^{-\frac{v+n}{2}} \]

where \( \Gamma(\cdot) \) represents the Gamma function.

To note, the mean and covariance of the \( \mathcal{S}(x; \mu, \Sigma, v) \) are \( \mu \) and \( \frac{v}{v+2} \Sigma \), respectively. Hereafter, \( v > 2 \).

B. Student’s t Recursion

In what follows, the initial state vector \( x_0 \) is assumed to have a Student’s t distribution, namely \( p(x_0) = \mathcal{S}(x_0; \hat{x}_0, P_0, v_0) \), and \( x_0, w_k, \) and \( v_k \) are assumed to be mutually uncorrelated. Assume the joint distribution of the state and process noise as Student’s t with joint dof \( v_k^* \) and
parameters $P_k'$ and $Q_k'$, namely,

$$p(x_k, w_k | Z_{1:k}) = \mathcal{S} \left( \begin{bmatrix} x_k \\ w_k \end{bmatrix} ; \begin{bmatrix} \hat{x}_k \\ 0 \end{bmatrix}, \begin{bmatrix} P_k' & 0 \\ 0 & Q_k' \end{bmatrix}, v'_k \right)$$

then from the rules for linear transformation of $t$ vectors [1], one gets

$$p(x_k, x_{k+1} | Z_{1:k+1}) = \mathcal{S} \left( \begin{bmatrix} x_k \\ x_{k+1} \end{bmatrix} ; \begin{bmatrix} \hat{x}_k \\ \hat{x}_{k+1} \end{bmatrix}, \begin{bmatrix} P_k' & F_k P'_k \end{bmatrix}, F_k P'_k, R_{k+1} \right) , v'_k \right). \quad (6)$$

The choices of parameters $\nu', P_k'$, and $Q_k'$ are discussed in [1]. In this article, we use the simple choice that $\nu' = \min(\nu_k, \nu_0)$, $P_k' = P_k$, and $Q_k' = Q_k$. The one-step state prediction results in a $t$ density

$$p(x_{k+1} | Z_{1:k}) = \mathcal{S} \left( x_{k+1} ; \hat{x}_{k+1}, P_{k+1|k}, v'_k \right) \quad (7)$$

where $\hat{x}_{k+1} = f_k(\hat{x}_k)$ and $P_{k+1|k} = F_k P'_k F_k^T + Q_k$.

Similarly, assume the joint distribution of the predicted state and measurement noise as Student’s $t$ with joint dof $\nu'_{k+1}$ and parameters $P_{k+1|k}$ and $R_{k+1|k}$, namely,

$$p(x_{k+1}, v_{k+1} | Z_{1:k}) = \mathcal{S} \left( \begin{bmatrix} x_{k+1} \\ v_{k+1} \end{bmatrix} ; \begin{bmatrix} \hat{x}_{k+1} \\ 0 \end{bmatrix}, \begin{bmatrix} P_{k+1|k} & 0 \\ 0 & R_{k+1} \end{bmatrix}, v'_k \right) \quad (8)$$

Again, a simple choice is $\nu'_{k+1} = \min(\nu'_k, \nu_0)$, $P_{k+1|k} = P_{k+1|k}$, and $R_{k+1|k} = R_{k+1}$. Consequently, the prediction density of the state and output can be written as

$$p(x_{k+1}, z_{k+1} | Z_{1:k}) = \mathcal{S} \left( \begin{bmatrix} x_{k+1} \\ z_{k+1} \end{bmatrix} ; \begin{bmatrix} \hat{x}_{k+1|k} \\ \hat{z}_{k+1} \end{bmatrix}, \begin{bmatrix} P_{k+1|k} & H_{k+1|k} \nu'_{k+1} \end{bmatrix}, v'_k \right) \quad (9)$$

where $\hat{z}_{k+1} = h_{k+1}(\hat{x}_{k+1|k})$ and $S_{k+1} = R_{k+1} + H_{k+1|k} P_{k+1|k} H_{k+1|k}^T$.

Then, the final Student’s $t$ posterior is given by a $t$ density

$$p(x_{k+1} | Z_{1:k+1}) = \mathcal{S} \left( x_{k+1} ; \hat{x}_{k+1}, P_{k+1}, v_{k+1} \right) \quad (10)$$

where

$$\hat{x}_{k+1} = \hat{x}_{k+1|k} + K_{k+1}(z_{k+1} - \hat{z}_{k+1}) \quad (11)$$

$$P_{k+1} = K_{k+1}(P_{k+1|k} - K_{k+1}S_{k+1} K_{k+1}^T) \quad (12)$$

$$v_{k+1} = \nu'_{k+1} + m \quad (13)$$

with

$$\alpha_{k+1} = \frac{\nu'_{k+1} + (z_{k+1} - \hat{z}_{k+1})^T S_{k+1}^{-1} (z_{k+1} - \hat{z}_{k+1})}{\nu'_{k+1} + m} \quad (14)$$

$$K_{k+1} = P_{k+1|k} H_{k+1|k}^T S_{k+1}^{-1} \quad (15)$$

Simply, when $v_0 = v_0 = v_0$, the above Student’s $t$ recursion differs from the classical KF merely in the calculation of the scale matrix. To highlight this similarity, we refer to this basic procedure as Student’s $t$ Kalman filter (StKF), which is the backbone of our proposed multisensor AA fusion StKF. The Monte–Carlo method [8], linearization [4], and unscented transform [5], [10] can be the same applied for versatile nonlinear StKFs as in the nonlinear KFs.

C. Suboptimal AA Density Fusion

Whenever optimal multisensor fusion is sought, the exact correlation between these sensors is needed. Unfortunately, despite ideal cases with a priori information [28], [29], it is practically intractable to do so. When a sensor network is involved, it becomes even more challenging as the correlation becomes more significant and complicated. Then, one may resort to suboptimal, robust fusion which eschews underestimating the actual squared estimate errors [13], [14], [15] and gains robustness. The AA fusion method is one of the fusion methods such motivated at first [15], which was proposed for distributed filter fusion the first in [30, (29)]. In fact, the AA fusion has a variety of important statistic and information-theoretic properties such as the capacity to preserve modes in the fusing sources and to combat false alarm in addition to its high computation efficiency; the reader is referred to [24], [25], [26], [31] for overviews of this method.

Straightforwardly, given a number of probability distributions $f_i(x)$ (of the same family or not) yielded by different estimators $i \in I := \{1, 2, \ldots, I\} \subset \mathbb{N}^+$, the AA fusion approach approximates the target distribution $g(x)$ by their weighted AA, a weighted mixture of distributions

$$f_{AA}(x) = \sum_{i \in I} w_i f_i(x) \quad (16)$$

where $w \in \mathcal{W} := \{w \in \mathbb{R}^I | \mathbf{1}^T w = 1, w_i > 0, \forall i \in I\} \subset \mathbb{R}^I$ are positive, normalized fusing weights. The mixture distribution produced by the AA density fusion is conducive to recursive filtering calculation in two aspects: First, a mixture of conjugate priors is also conjugate and can approximate any kind of prior [32], [33]. Second, the linear fusion of a finite number of mixtures of the same parametric family remains a mixture of the same family [22], [24]. Third, for any target density $g(x)$, the AA ensures a better fit on average than those fusing densities $D_{KL}(f_{AA} || g) \leq \sum_{i \in I} w_i D_{KL}(f_i || g)$, where the equation holds if and only if all densities $f_i$, $i \in I$ are identical [26], [34].

Mathematically, the AA fusion minimizes the weighted sum of the KL divergences of the fused result with relative to the fusing densities as follows [35], [36]:

$$f_{AA}(x) = \arg \min_{g \in \mathcal{F}(x)} \sum_{i \in I} w_i D_{KL}(f_i || g) \quad (17)$$

where $\mathcal{F}(x)$ denotes the set of scalar-valued functions on $x$:

$\mathcal{F}(x) = \{ f : \mathbb{R}^d \to \mathbb{R} \}$. The above minimization holds for any fusing weights that belong to $\mathcal{W}$. Given the true distribution $p(x)$, the optimal fusing weights should minimize...
D_KL(f_{AA}||p), i.e.,
\[
w_{opt} = \arg \min_{w \in W} D_{KL}(f_{AA}||p)
= \arg \min_{w \in W} D_{KL}\left(\sum_{i \in I} w_i f_i || p\right)
= \arg \min_{w \in W} \sum_{i \in I} w_i \left(D_{KL}(f_i || p) - D_{KL}(f_i || f_{AA})\right).
\]

(18)

Notably, there is an essential difference between the two minimization formulations (17) and (18): The former is not directly related with the true distribution as the latter does. Since the true distribution \(p(x)\) is practically unknown, a simplified, practically operable, alternative is ignoring the former part in (18), which will then reduce to the following suboptimal, diversity preference, maximization solution [26]
\[
w_{subopt} = \arg \max_{w \in W} \sum_{i \in I} w_i D_{KL}(f_i || f_{AA}).
\]

(19)

Combining (17) and (19) results in the max–min formulation of the suboptimal AA fusion as follows [26]:
\[
(w_{subopt}, f_{AA}) = \arg \max_{w \in W} \min_{\nu \in F(x)} \sum_{i \in I} w_i D_{KL}(f_i || \nu).
\]

(20)

III. PROPOSED MULTISENSOR STKF BASED ON AA DENSITY FUSION

In this section, we highlight an important statistical property of the AA density fusion. This property facilitates approximating the AA of Student’s \(t\) densities, which is a \(t\) mixture by a single \(t\) distribution via moment matching. The fusing weights are driven from the max–min optimization (20) by means of moment-matching Gaussian-\(t\) approximation, leading to the suboptimal AA fusion-based multisensor STKF.

A. Statistics of AA Density Fusion

With regard to a finite mixture, there are two important properties: First, the moments of the mixture are the convex combination of those of the subposteriors, i.e.,
\[
E_{f_{AA}}[x_k] = \sum_{i \in I} w_i E_{f_i}[x_k]
\]

(21)

where \(E_f[x]\) denotes any moment of the variable \(x\) with distribution \(f(x)\) and (21) can be easily seen from the definition of the \(n\)th-order moment \(E_f^n[x] = \int x^n f(x)dx\). In particular, on the first-order moment \(n = 1\) which is the mean of the density, i.e., \(\hat{x} = \int x f(x)dx\), we have \(\hat{x}_{AA} = \sum_{i \in I} w_i \hat{x}_i\).

Second, the variance of the mixture is driven as
\[
P_{f_{AA}} = \int \left(x - E_{f_{AA}}[x_k]\right) (\cdot)^T f_{AA}(x)d\mathbf{x}
= \sum_{i \in I} w_i \int \left(x - E_{f_{AA}}[x_k]\right) (\cdot)^T f_i(x)d\mathbf{x}
= \sum_{i \in I} w_i \int \left(x - E_{f_i}[x_k] + \hat{x}_i\right) (\cdot)^T f_i(x)d\mathbf{x}
\]

\[
= \sum_{i \in I} w_i \left(\int \left(x - E_{f_i}[x_k]\right) (\cdot)^T f_i(x)d\mathbf{x} + \hat{x}_i^T \hat{x}_i\right)
= \sum_{i \in I} w_i \left(P_{f_i} + \hat{x}_i \hat{x}_i^T\right)
\]

(22)

where \(\hat{x}_i := E_{f_i}[x_k] - E_{f_{AA}}[x_k]\) and \((\cdot)^T := (\cdot)^T\).

B. Averaging Student’s \(t\) Densities

Consider the AA of \(t\) densities \(S_i(x_k; \hat{x}_{i,k}, P_{i,k}, v_{i,k}), i \in I\)
\[
f_{AA}(x_k) = \sum_{i \in I} w_i S_i(x_k; \hat{x}_{i,k}, P_{i,k}, v_{i,k}).
\]

(23)

According to (21) and (22), we have
\[
E_{f_{AA}}[x_k] = \sum_{i \in I} w_i E_{S_i}[x_k]
= \sum_{i \in I} w_i \hat{x}_{i,k}
= \sum_{i \in I} w_i (\frac{v_{i,k}}{v_{i,k} - 2} P_{i,k} + \left(\hat{x}_{i,k} - \sum_{i \in I} w_i \hat{x}_{i,k}\right)(\cdot)^T).
\]

(24)

This paves a way to obtain closed-from recursion of the STKF by means of moment matching, namely the AA of Student’s \(t\) posteriors is approximated as \(t\) distributed.

PROPOSITION 1 The first and second moment matching \(t\) density approximation to the weighted mixture of Student’s \(t\) densities as given in (23) is given by
\[
S_{AA}(x_k) \approx S(x_k; \hat{x}_{AA}, P_{AA}, v_{AA})
\]

(26)

\[
\hat{x}_{AA} = \sum_{i \in I} w_i \hat{x}_{i,k}
= \sum_{i \in I} w_i (\frac{v_{i,k}}{v_{i,k} - 2} P_{i,k} + \left(\hat{x}_{i,k} - \sum_{i \in I} w_i \hat{x}_{i,k}\right)(\cdot)^T).
\]

(27)

\[
P_{AA} = \frac{v_{AA} - 2}{v_{AA}} P_{f_{AA}}^\nu
\]

(28)

where \(P_{f_{AA}}^\nu\) is given in (25) and one may choose \(v_{AA} = \min_{i \in I} v_{i,k}\) to preserve the heaviest tail of all fusing densities or choose their average \(v_{AA} = \frac{1}{|I|} \sum_{i \in I} v_{i,k}\).

C. Suboptimal Weight for Averaging Student’s \(t\) Densities

Substituting Student’s \(t\) densities and the moment-matching approximation (26) into the suboptimal weighting solution (19) yields
\[
w_{subopt} = \arg \max_{w \in W} \sum_{i \in I} w_i D_{KL}(S_i || f_{AA})
\approx \arg \max_{w \in W} \sum_{i \in I} w_i D_{KL}(S_i || S_{AA})
\]

(29)

(30)

Unfortunately, neither (29) nor (30) admits analytical solution as the KL divergence between Student’s \(t\) distributions cannot be analytically expressed. In spite of numerical approximation methods such as the Monte–Carlo
method that is computationally expensive and unsuitable for nonexhaustive optimization, we propose a further simplified alternative based on the similarity between the Student’s t and Gaussian distributions as follows.

**Proposition 2** The KL divergence between Student’s t distributions \( S_i(x) = S(x; \hat{x}_i, \frac{\nu_i-2}{\nu_i} P_i, v_i) \) and \( S_j(x) = S(x; \hat{x}_j, \frac{\nu_j-2}{\nu_j} P_j, v_j) \) is approximated as that between their moment-matching Gaussian distributions \( N_j(x) = N(\hat{x}_j, P_j) \) and \( N_j(x) = N(\hat{x}_j, P_j) \). This leads to an analytically approximate weighting solution

\[
\mathbf{w}_{subopt} \approx \arg\max_{\mathbf{w} \in \mathbf{W}} \sum_{i \in I} w_i D_{KL} (S_i \| S_{AA})
\]

\[
\approx \arg\max_{\mathbf{w} \in \mathbf{W}} \sum_{i \in I} w_i D_{KL} (N_j(\hat{x}_j, P_i) \| N_{AA}(\hat{x}_{AA}, P_{AA}))
\]

\[
= \arg\max_{\mathbf{w} \in \mathbf{W}} \sum_{i \in I} w_i \left[ \frac{v_i}{v_i - 2} \left( P_i^{-1} \hat{x}_j \right) \right] + \log \left( \frac{\det (P_{AA}^\nu)}{\det (P_i)} \right)
\]

\[
+ (\hat{x}_i - \hat{x}_{AA})^T P_{AA}^{-1} (\hat{x}_i - \hat{x}_{AA}) \right]. \tag{33}
\]

There is another alternative approximation rather than the above exact moment matching. That is, ignore the dof of the Student’s t distribution and deal with the scale matrix simply as a covariance, i.e.,

\[ D_{KL}(S_i(\hat{x}_i, P_i, v_i) \| S_j(\hat{x}_j, P_j, v_j)) \approx D_{KL}(N_j(\hat{x}_j, P_i) \| N_j(\hat{x}_j, P_j)). \]

This will lead to the following analytical solution, c.f., (33):

\[
\mathbf{w}_{subopt} \approx \arg\max_{\mathbf{w} \in \mathbf{W}} \sum_{i \in I} w_i \left[ \frac{v_i}{v_i - 2} \left( P_i^{-1} \hat{x}_j \right) \right] + \log \left( \frac{\det (P_{AA}^\nu)}{\det (P_i)} \right)
\]

\[
+ (\hat{x}_i - \hat{x}_{AA})^T P_{AA}^{-1} (\hat{x}_i - \hat{x}_{AA}) \right]. \tag{34}
\]

Obviously, both approximations will converge to the real KL divergence as the dof increases indefinitely \( \nu \to +\infty \). To illustrate the approximation accuracy, we consider some examples for \( \nu = 3 \) as given in Fig. 1. As shown, both methods achieve the minimum value at the same point when the involved two (t or Gaussian) densities have the same means, from which they both monotonously increase faster or slower, as well as the real KL divergence, with the increase of the distance between the means of two densities. This confirms the consistency of the approximation of both methods that a larger approximate divergence indicates a larger real divergence. To gain more insight, note that a sufficient condition to (29) is given by the following “middle distribution” equation [26], [37], of which the proof is given in Appendix A.

\[ D_{KL} (S_i \| f_{AA}^S (\mathbf{w}_{subopt})) = D_{KL} (S_j \| f_{AA}^S (\mathbf{w}_{subopt})) \]  

(35)

and so, c.f., Proposition 1 and (30),

\[ D_{KL} (S_i \| S_{AA} (\mathbf{w}_{subopt})) \approx D_{KL} (S_j \| S_{AA} (\mathbf{w}_{subopt})). \tag{36} \]

Therefore, one can evaluate how much the above equation is satisfied to evaluate whether the solutions provided by (33) or (34) are accurate. We illustrate here the accuracy of (33) again by examples for \( \nu = 3 \) in Fig. 2. It shows that the AA fusion resulted by the proposed suboptimal fusing weights (33) complies well with (35) or (36), especially when the fused distributions have similar covariances. In case two Student’s t distributions have the same scale matrix (as shown in the right-upper subfigure of Fig. 2), the result meets greatly the “middle distribution” equation (36) and confirms the effectiveness of (33). This is an important finding: Although the Gaussian-t approximation/substitution can be inaccurate, the fusing weights yielded by (30) and (33) can be close with each other and so be close to the optimal solution satisfying (35). These being said, more accurate, analytical calculation or approximate method is
desirable. Approaches such as mode approximation [38] and variational inference [3], [39] are notable.

As a final result based on simplified choice of the dof and Gaussian-t approximation, (33) (or (34)) and (26) constitute the overall information-theoretic max–min-optimized AA fusion of Student’s t as follows, c.f. (20),

$$w_{\text{subopt}, S_{\text{AA}}} = \arg \min_{w \in W} \sum_{i \in I} u_i D_{\text{KL}} \left( S_i \parallel g \right)$$ \hspace{1cm} (37)

where $F_S$ denotes the set of all Student’s t functions: $F_S = \{ f : \mathbb{R}^d \rightarrow \mathbb{R} \}$.

D. Algorithm Flow of AA-Fusion StKF

When multiple sensors cooperate with each other, communication will be involved which needs to take into account the intersensor connection topology and constraints. There are considerable relevant research, including those for AA fusion [16], [17], [21], [23], [40]. However, to avoid distracting the readers’ attention, we ignore the communication issue and assume that the sensors have direct access to each other, analogous to the centralized network with feedback. We also assume that the sensors are synchronized and coordinated in the same coordinate system. The overall procedure of the proposed multisensor AA-StKF for totally $S$ interconnected sensors is summarized in Algorithm 1. All parameters of the SSM are assumed known by default.

IV. EXTENDED CI FOR STUDENT’S t FUSION

As addressed so far, the proposed Student’s t AA fusion approach relies on approximating the calculation over $t$ distribution by that over Gaussian. Therefore, many other Gaussian fusion approaches such as the CI [14] and inverse CI [41] can be applied in place of the AA fusion for multisensor Student’s $t$ filter design. Here, we briefly address the extension of the CI fusion. The CI fusion of $t$ densities $S_i(x_k; \hat{x}_{i,k}, P_{i,k}, v_{i,k})$, $i \in I$ can be written by

$$f_{S_i}(x_k) \propto \prod_{i \in I} \left( S_i \left( x_k; \hat{x}_{i,k}, P_{i,k}, v_{i,k} \right) \right)^{w_i}.$$ \hspace{1cm} (38)

The above fusion is not analytically evolvable due to the nonintegrability of the poly-$t$ (product of Student’s $t$ distributions) distribution. By using Gaussian distributions $N_i(x_k; \hat{x}_{i,k}, \frac{v_{i,k}}{v_{i,k} - 2} P_{i,k})$ to approximate the corresponding Student’s $t$ distributions $S_i(x_k; \hat{x}_{i,k}, P_{i,k}, v_{i,k})$, $i \in I$ with the same first and second moments, the CI fusion is given as follows:

$$\hat{x}_{\text{CI}} = P_{\text{CI}} \sum_{i \in I} \frac{u_i}{v_{i,k}} (v_{i,k} - 2) P_{i,k}^{-1} \hat{x}_i$$ \hspace{1cm} (39)

$$P_{\text{CI}} = \left( \sum_{i \in I} \frac{u_i}{v_{i,k}} (v_{i,k} - 2) P_{i,k}^{-1} \right)^{-1}$$ \hspace{1cm} (40)

where

$$w_{\text{CI}} = \arg \min_{w \in W} \text{Tr} \left( P_{\text{CI}} \right).$$ \hspace{1cm} (41)

With regard to Algorithm 1, the above formulation is amount to using (41), (39), and (40) to replace (33)/(34) in

Algorithm 1: One Filtering Iteration of Proposed Multisensor AA-Fusion StKF

Input: $\{ S_i(x_k; \hat{x}_{i,k}, P_{i,k}, v_{i,k}) \}_{i=1}^S$

Output: $\{ S_x(x_{k+1}; \hat{x}_{x,k+1}, P_{x,k+1}, v_{x,k+1}) \}_{x=1}^S$

1: for each sensor $s = 1, \ldots, S$ in parallel do

2: One-step state prediction as in (7) in the following steps:

3: Update the prediction as in (10) using the new measurement $z_{x,k}$, including the following steps:

4: end for

5: for each sensor $s = 1, \ldots, S$ in parallel do

6: Exchange parameters $\{ x_{k+1}, P_{x,k+1}, v_{x,k+1} \}$ among interconnected sensors. Suppose that sensor $s$ finally receives

7: Calculate the AA fusing weights $w_{\text{subopt}}$ via (33) or (34), based on (25) and (27).

8: Calculate the AA-fused Student’s $t$ density as in (26) including the following steps:

9: end for

10: Steps 5–9 may be performed for multiple iterations in the case of decentralized sensor network, like what is done in the context of average consensus [16], [18], [21], [23].

Step 7, Step 8.2), and Step 8.3), respectively. This leads to a multisensor CI fusion-based StKF.

V. SIMULATIONS

We consider a single target tracking problem. Following the literature [4], [5], we simulate abnormal noises of a significant magnitude (much higher than the normal case) which randomly occur with probabilities to assume outliers affecting the state process and the measurement, independently. The state of the target $x_k = [p_{x,k}, p_{y,k}, v_{x,k}, v_{y,k}]^T$ consists of planar position $[p_{x,k}, p_{y,k}]^T$ and velocity $[v_{x,k}, v_{y,k}]^T$. At time $k = 0$, it is randomly initialized as $x_0 \sim N(x_0; \mu_0, P_0)$, where $\mu_0 = [1000 \text{ m}, 20 \text{ m/s}, 1000 \text{ m}, 0 \text{ m/s}]^T$ with $P_0 =$...
where \( \text{diag}\{[500 \, \text{m}^2, 50 \, \text{m}^2/\text{s}^2, 50 \, \text{m}^2, 50 \, \text{m}^2/\text{s}^2]\} \), where \( \text{diag}\{a\} \) represents a diagonal matrix with diagonal \( a \). The target moves following a nearly constant velocity motion given as (with the sampling interval \( \Delta = 1 \, \text{s} \))

\[
x_k = \begin{bmatrix}
1 & \Delta & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & \Delta \\
0 & 0 & 0 & 1
\end{bmatrix} x_{k-1} + \begin{bmatrix}
\Delta^2 \\
\Delta \\
0 \\
0
\end{bmatrix} \begin{bmatrix} 1 \end{bmatrix} u_{k-1}
\] (42)

where the process noise \( u_k \sim \mathcal{N}(0, \mathbf{R}) \). Here, the process noise standard deviation \( r \) is defined with the outlier probability \( p_o \) as follows:

\[
\begin{align*}
    r & = 5, \text{ with probability } 1 - p_o \\
    r & = 50, \text{ with probability } p_o
\end{align*}
\] (43)

We consider only two sensors. Both sensors \( s = 1, 2 \) have the linear measurement model \( z_{i,k} = H_{i,k} x_k + v_{i,k} \) as follows:

\[
H_{i,k} = \begin{bmatrix} 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}, \quad v_{i,k} = \begin{bmatrix} v_{i,k,1} \\
v_{i,k,2}
\end{bmatrix}
\] (44)

with \( v_{i,k,1} \) and \( v_{i,k,2} \) as mutually independent zero-mean Gaussian noise with the same standard deviation \( R_s \).

The measurement noise standard deviations of two sensors are defined with outlier probability too as follows (independent with that of the process noise and with each other sensor):

\[
\begin{align*}
    R_s & = 20 \, \text{m}, \text{ with probability } 1 - p_o \\
    R_s & = 200 \, \text{m}, \text{ with probability } p_o \\
    R_s & = 10 \, \text{m}, \text{ with probability } 1 - p_o \\
    R_s & = 100 \, \text{m}, \text{ with probability } p_o
\end{align*}
\] (45)

We test different outlier probabilities \( p_o \) from 0 (when there is no outlier) to 0.2. In each case of \( p_o \), the simulation is performed for 1000 Monte–Carlo runs, each having 100 filtering steps. The target trajectory is randomly generated according to the process model with a random initial state in each run. The process and measurement noises are approximated by Student’s \( t \) noise with the normal mean and covariance/scale parameters as specified, and with dof \( v_q = v_R = 3 \). So, the resulting posterior PDF is a Student’s \( t \) PDF with fixed degree of freedom parameter. The root mean square error (RMSE) of the position or velocity estimates is used for filter evaluation

\[
\text{RMSE}_k = \sqrt{\frac{1}{M} \sum_{i=1}^{M} (\hat{x}_{i,k} - x_i^k)^T (\hat{x}_{i,k} - x_i^k)}
\] (47)

where \( \hat{x}_{i,k} \) is the position or velocity estimate of the real state \( x_i^k \) at time \( k \) in run \( i \) and \( M \) is the number of runs.

Both StKF and KF are simulated. The KFs are initialized by \( \mathcal{N}(\bf{x}; \mu_0, \mathbf{P}_0) \) and the StKFs are initialized by \( \mathcal{S}(\bf{x}; \mu_0, \mathbf{P}_0) \). Both StKFs and KFs are implemented in both noncooperative manner (using only sensor 1’s measurement) and two-sensor cooperative manner. In the latter, we compare the proposed AA, CI fusion with the augmented measurement (AM) approach [3] (also known as centralized batch fusion [12]). In the AM approach, two sensors’ measurements are cascaded/augmented as a joint \( z_k = H_1 x_k + v_k \) as follows:

\[
z_k = \begin{bmatrix} z_{1,k} \\
z_{2,k}
\end{bmatrix}, \quad H_k = \begin{bmatrix} H_{1,k} & H_{2,k}
\end{bmatrix}, \quad v_k = \begin{bmatrix} v_{1,k} \\
v_{2,k}
\end{bmatrix}
\] (48)

In addition, the AA fusion has been implemented in two means, one using uniform weights \( (w_1 = w_2 = 0.5) \) referred to as “unAA fusion” and the other using the suboptimal weight (referred to as AA fusion by default) as given in (33) in Section III-C. In both cases, the AA fusion of two Gaussian PDFs results in a Gaussian mixture of two components for which a merging scheme based on moment matching is needed to maintain closed-form KF recursion.

The real target trajectory and the estimates of the KFs/StKFs in one trial are given in Fig. 3. The position and velocity RMSEs of these filters are given in Figs. 4 and 5, respectively. The average position and velocity RMSEs over all filtering times against outlier probability are shown in Fig. 3. The method proposed in this section was applied in both noncooperative and cooperative manner in the AM approach.
Fig. 5. Velocity RMSEs of noncooperative and two-sensor-fusion KFs/StKFs when $p_o = 0.05$.

Fig. 6. Average position RMSEs against outlier probability.

Fig. 7. Average velocity RMSEs against outlier probability.

Fig. 8. Variance of the fusion weight assigned to sensor 1 in StKF-AA when $p_o = 0.05$.

Approaches are effective. More specifically, there are several main findings.

1) The StKF-AA using the suboptimal fusing weights (33) performs the best in position estimation accuracy as long as $p_o > 0.03$ and in velocity estimation accuracy as long as $p_o \geq 0.08$. Further on, the variance of the fusion weight assigned to sensor 1 in StKF-AA is given in Fig. 8, with an average level about 0.45. This is reasonable since sensor 1 has a larger measurement noise covariance as compared with sensor 2. Indeed, the default AA fusion using optimized weights performs better than that using uniform weights and also the CI fusion, based on whether KF or StKF.

2) Most StKFs outperform their corresponding KFs using the same fusion approach in position accuracy when $p_o \geq 0.02$ except for the StKF-CI that performs better than KF-CI only when $p_o > 0.04$. This demonstrates the advantage of the Student’s $t$ filter in dealing with outlier/heavy-tailed noises as compared with Gaussian filters. In contrast, when there is no outlier ($p_o = 0$), the KFs outperform the StKFs, regardless of the fusion approach. This is simply because there is no need to maintain a heavy-tailed posterior for a purely linear Gaussian system.

3) When there is no outlier, the KF-AM is optimal for the linear Gaussian system. When there exist outliers, the AM approach performs worse than the AA fusion in position accuracy, using whether KF or StKF. The reasons are twofold: First, the AM approach that seeks MV becomes disadvantageous in dealing with outliers. Second, when the outlier is independent across sensors, the AM approach suffers from a higher actual measurement outlier probability which is calculated by

$$1 - \prod_{i \in I} (1 - p_o^i) > \max_{i \in I} p_o^i$$

where $p_o^i$ is the outlier probability of sensor $i$.

Two final remarks are to be highlighted. First, the performance of the basic StKFs using a simplified choice of the dof as addressed is expected to be improved if the dof can (online) adapt with the probability of outlier. Notably, a recent work proposes to combine Gaussian and $t$ on the base of a probabilistic framework so that to adapt the tail [42].
This remains a valuable, open problem. Second, the CI fusion seeks MV yet covariance consistent fusion which does not have as good fault-tolerant capacity as the AA fusion does [26] and so the forced Gaussian-\(t\) approximation may not suit it. As shown, when the outlier probability is small (\(p_\theta < 0.04\)), the CI-StKF is disappointing although it is good when \(p_\theta > 0.1\). It remains open how to optimally apply Gaussian fusion approaches to the Student’s \(t\) distributions.

VI. CONCLUSION

This article proposes a multisensor Student’s \(t\) filter based on the AA density fusion approach. An information-theoretic optimization formulation based on moment-matching Gaussian-\(t\) approximation is used to drive the suboptimal fusion weights and the \(t\) mixture merging procedure to ensure closed-form Student’s \(t\) recursion. The proposed fusion framework accommodates other Gaussian fusion approaches such as the CI, as well as fusion between Gaussian and Student’s \(t\) distributions. Simulation in a target tracking scenario using various probabilities of outlier has demonstrated the promising performance of the proposed multisensor AA fusion-based Student’s \(t\) filter in dealing with outliers as compared with the multisensor KF/StKF based on either AM or CI.

This work, however, is limited to simplified choice of the dof of the Student’s \(t\) PDF and moment-matching Gaussian-\(t\) approximation for closed-form \(t\)-AA fusion and for fast fusing weight design. Improvement can be expected if these limitations are removed. In particular, a valuable future research is to online adapt the dof of the Student’s \(t\) density according to the significance of the outliers.

APPENDIX A

PROOF OF THE SUFFICIENCY OF (35) FOR (29)

To meet the weight constraint \(w^T I_t = 1\), we define
\[
D_{KL}^i \triangleq \sum_{i \neq j} w_i D_{KL}(S_i \| f_{S_i}^{\text{AA}}) / (1 - w_j) \quad (50)
\]

\[
S_{-j}(x) \triangleq \sum_{i \neq j} \frac{f_{S_i}^{\text{AA}}(x) - w_j S_j(x)}{1 - w_j} \quad (51)
\]

It is obvious that given the condition (35), we have
\[
D_{KL}(S_j \| f_{S_i}^{\text{AA}}) = D_{KL}^i \quad (52)
\]
\[
\int_{\mathbb{R}^d} S_{-j}(x) dx = 1 \quad (53)
\]

Then, the partial derivative of the optimization function given in (29) is
\[
\frac{\partial}{\partial w_j} \sum_{i \in I} w_i D_{KL}(S_i \| f_{S_i}^{\text{AA}}) = D_{KL}(S_j \| f_{S_i}^{\text{AA}}) - D_{KL}^i + \sum_{i \in I} \frac{\partial}{\partial w_j} D_{KL}(S_i \| f_{S_i}^{\text{AA}}) \quad (54)
\]

Further, on,
\[
\frac{\partial^2}{\partial w_j^2} \sum_{i \in I} w_i D_{KL}(S_i \| f_{S_i}^{\text{AA}}) = - \sum_{i \neq j} w_i D_{KL}(S_i \| f_{S_i}^{\text{AA}}) / (1 - w_j)^2 \leq 0 \quad (57)
\]

where equation in (58) holds if and only if \(S_i(x) = S_j(x), \forall i \neq j\).

So, (29) is proven by (56) and (58) based on condition (35).

REFERENCES

[1] M. Roth, T. Ardeshiri, E. Özkan, and F. Gustafsson, “Robust Bayesian filtering and smoothing using Student’s \(t\) distribution,” Mar. 2017, \textit{ArXiv:1703.02428[stat.ME]}.

[2] R. Piché, S. Särkkä, and J. Hartikainen, “Recursive outlier-robust filtering and smoothing for nonlinear systems using the multivariate Student-\(t\) distribution,” in \textit{Proc. IEEE Int. Workshop MLSP}, 2012, pp. 1–6.

[3] H. Zhu, H. Leung, and Z. He, “A variational Bayesian approach to robust sensor fusion based on Student-\(t\) distribution,” \textit{Inf. Sci.}, vol. 221, pp. 201–214, 2013.

[4] M. Roth, E. Özkan, and F. Gustafsson, “A Student’s \(t\) filter for heavy tailed process and measurement noise,” in \textit{Proc. IEEE Int. Conf. Acoust. Speech Signal Process.}, 2013, pp. 5770–5774.

[5] Y. Huang, Y. Zhang, N. Li, S. Mohsen, and J. Chambers, “A robust Student’s \(t\) based cubature filter,” in \textit{Proc. 19th Int. Conf. Inf. Fusion}, 2016, pp. 9–16.

[6] Y. Huang, Y. Zhang, N. Li, Z. Wu, and J. A. Chambers, “A novel robust Student’s \(t\)-based Kalman filter,” \textit{IEEE Trans. Aerosp. Electron. Syst.}, vol. 53, no. 3, pp. 1545–1554, Jun. 2017.

[7] S. Nadarajah and S. Kotz, “Mathematical properties of the multivariate \(t\) distribution,” \textit{Acta Appl. Math.}, no. 2, pp. 53–84, Dec. 2005.

[8] A. Genz and J. Monahan, “Stochastic integration rules for infinite regions,” \textit{SIAM J. Sci. Comput.}, vol. 19, no. 2, pp. 426–439, 1998.

[9] S. Li, H. Wang, and T. Chai, “A \(t\)-distribution based particle filter for target tracking,” in \textit{Proc. Amer. Control Conf.}, Minneapolis, MN, USA, 2006, pp. 1–6.

[10] F. Tronarp, R. Hostettler, and S. Särkkä, “Sigma-point filtering for nonlinear systems with non-additive heavy-tailed noise,” in \textit{Proc. 19th Int. Conf. Inf. Fusion}, 2016, pp. 1859–1866.

[11] C. Xu, S. Zhao, B. Huang, and F. Liu, “Distributed Student’s \(t\) filtering algorithm for heavy-tailed noises,” \textit{Int. J. Adaptive Control Signal Process.}, vol. 32, no. 6, pp. 875–890, 2018.
[12] L. Yan, L. Jiang, and Y. Xia, *Distributed Fusion Estimation for Multisensor Systems With Heavy-Tailed Noises*. Singapore: Springer, 2021, pp. 189–211.

[13] J. K. Uhlmann, “General data fusion for estimates with unknown cross covariances,” *Proc. SPIE*, vol. 2755, pp. 2755–2767, 1996.

[14] S. Julier and J. Uhlmann, “General decentralized data fusion with covariance intersection (CI),” in *Handbook of Data Fusion*, D. Hall and J. Llinas, Eds., ch. 12. Boca Raton, FL, USA: CRC Press, 2001, pp. 1–25.

[15] T. Bailey, S. Julier, and G. Agamennoni, “On conservative fusion of information with unknown non-Gaussian dependence,” in *Proc. 15th Int. Conf. Inf. Fusion*, Singapore, 2012, pp. 1876–1883.

[16] T. Li, J. Corchado, and S. Sun, “Partial consensus and conservative fusion of Gaussian mixtures for distributed fusion,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 5, pp. 2150–2163, Oct. 2019.

[17] T. Li and F. Hlawatsch, “A distributed particle-PHD filter using arithmetic-average fusion of Gaussian mixture parameters,” *Inf. Fusion*, vol. 73, pp. 111–124, 2021.

[18] T. Li, Z. Liu, and Q. Pan, “Distributed Bernoulli filtering for target detection and tracking based on arithmetic average fusion,” *IEEE Signal Process. Lett.*, vol. 26, no. 12, pp. 1812–1816, Dec. 2019.

[19] H. Kim, K. Granström, L. Gao, G. Battistelli, S. Kim, and H. Wymeersch, “5G mmwave cooperative positioning and mapping using multi-model filter and map fusion,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3782–3795, Jun. 2020.

[20] R. K. Ramachandran, N. Fronda, and G. Sukhatme, “Resilience in multi-robot multi-target tracking with unknown number of targets through reconfiguration,” *IEEE Trans. Control Netw. Syst.*, vol. 8, no. 2, pp. 609–620, Jun. 2021.

[21] T. Li, X. Wang, Y. Liang, and Q. Pan, “On arithmetic average fusion and its application for distributed multi-Bernoulli multitarget tracking,” *IEEE Trans. Signal Process.*, vol. 68, pp. 2883–2896, Apr. 2020.

[22] T. Li, Y. Xin, Z. Liu, and K. Da, “Best fit of mixture for computationally efficient poisson multi-Bernoulli mixture filtering,” *Signal Process.*, vol. 202, 2023, Art. no. 108739.

[23] F. Yang, L. Zheng, T. Li, and L. Shi, “A computationally efficient distributed Bayesian filter with random finite set observations,” *Signal Process.*, vol. 194, 2022, Art. no. 108454.

[24] K. Da, T. Li, Y. Zhu, H. Fan, and Q. Fu, “Recent advances in multisensor multitarget tracking using random finite set,” in *Front Inform. Technol. Electron. Eng.*, vol. 22, no. 1, pp. 5–24, 2021.

[25] G. Koliander, Y. El-Laham, P. M. Djuric, and F. Hlawatsch, “Fusion of probability density functions,” *Proc. IEEE*, vol. 110, no. 4, pp. 404–453, 2022.

[26] T. Li, Y. Xin, Y. Song, E. Song, and H. Fan, “Some statistic and information-theoretic results on arithmetic average density fusion,” 2021, arXiv:2110.01440.

[27] M. Kayaalp, Y. Inan, E. Telatar, and A. H. Sayed, “On the arithmetic and geometric fusion of beliefs for distributed inference,” 2022, arXiv:2204.13741.

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

[29] C. Chong, S. Mori, and K. Chang, “Graphical models for nonlinear distributed estimation,” in *Proc. Int. Conf. Inf. Fusion*, Stockholm, Sweden, 2004, pp. 1–8.

[30] T. Li, J. Corchado, and J. Prieto, “Convergence of distributed flooding and its application for distributed Bayesian filtering,” *IEEE Trans. Signal Inf. Process. Netw.*, vol. 3, no. 3, pp. 580–591, Sep. 2017.

[31] T. Li, H. Fan, J. Garcia, and J. M. Corchado, “Second-order statistics analysis and comparison between arithmetic and geometric average fusion: Application to multi-sensor target tracking,” *Inf. Fusion*, vol. 51, pp. 233–243, 2019.

[32] S. R. Dalal and W. J. Hall, “Approximating priors by mixtures of natural conjugate priors,” *J. Roy. Stat. Soc.: Ser. B (Methodological)*, vol. 45, no. 2, pp. 278–286, 1983.

[33] P. Diaconis and Y. D., “Quantifying prior opinion,” Rep. no.: EFS NSF207, Oct. 1983.

[34] R. E. Blahut, *Principles and Practice of Information Theory*. Boston, MA, USA: Addison-Wesley Longman Publishing., 1987.

[35] A. E. Abbas, “A Kullback–Leibler view of linear and log-linear pools,” *Decis. Anal.*, vol. 6, no. 1, pp. 25–37, 2009.

[36] K. Da, T. Li, Y. Zhu, H. Fan, and Q. Fu, “Kullback–Leibler averaging for multitarget density fusion,” in *Proc. Int. Symp. Distrib. Comput. Artif. Intell.*, Avila, Spain, 2019, pp. 253–261.

[37] F. Nielsen, “An information-geometric characterization of Chernoff information,” *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 269–272, Mar. 2013.

[38] J. Loxam and T. Drummond, “Student-t mixture filter for robust, real-time visual tracking,” in *Proc. Eur. Conf. Comput. Vis.*, 2008, pp. 372–385.

[39] Y. Huang, Y. Zhang, and J. A. Chambers, “A novel Kullback–Leibler divergence minimization-based adaptive Student’s t-filter,” *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5417–5432, Oct. 2019.

[40] T. Li, V. Elvira, H. Fan, and J. M. Corchado, “Local-diffusion-based distributed filtering using sensors with limited sensing range,” *IEEE Sensors J.*, vol. 19, no. 4, pp. 1580–1589, Feb. 2019.

[41] B. Noack, J. Sijs, M. Reinhardt, and U. Hanebeck, “Decentralized data fusion with inverse covariance intersection,” *Automatica*, vol. 79, pp. 35–41, 2017.

[42] Y. Huang, Y. Zhang, Y. Zhao, and J. A. Chambers, “A novel robust Gaussian-Student’s t mixture distribution based Kalman filter,” *IEEE Trans. Signal Process.*, vol. 67, no. 13, pp. 3606–3620, Jul. 2019.

---

**CORRESPONDENCE**

3387

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.