Tiancheng Li, Hongqi Fan, Jesús García, Juan M Corchado. (2019). Second-order statistics analysis and comparison between arithmetic and geometric average fusion: Application to multi-sensor target tracking. *Information fusion*, 51, pp. 233-243.

DOI: https://doi.org/10.1016/j.inffus.2019.02.009

Second Order Statistics Analysis and Comparison between Arithmetic and Geometric Average Fusion

Tiancheng Li, Hongqi Fan, Jesús G. Herrero and Juan M Corchado

***Abstract*—Two fundamental approaches to information averaging are based on linear and logarithmic combination, yielding the arithmetic average (AA) and geometric average (GA) of the fusing initials, respectively. In the context of target tracking, the two most common formats of data to be fused are random variables and probability density functions, for which the corresponding fusion is referred to as v-fusion and f-fusion, respectively. In this work, we analyze and compare the second order statistics (including variance and mean square error) of AA and GA in terms of both v-fusion and f-fusion. The AA and GA fusion of weighted Gaussian mixtures (whose weight sum is not necessarily unit) is also considered. The analysis is given by means of exact derivation or through exemplifying/numerical illustration.**

***Index Terms*—Data fusion, average consensus, Chernoff fusion, arithmetical mean, covariance intersection.**

I. Introduction

The rapid development and extensive deployment of sensor/agent networks, have stemmed remarkable interest in distributed data fusion. Decentralization of information fusion has evident advantages and allows the exchange of information of netted platforms in an efficient, flexible and scalable way. For example, in the context of target tracking using a decentralized sensor network, the sensor cooperation can compensate for the mis-detection and failure of a local sensor and overcoming the limitation of the local field of view, gaining improved estimation accuracy and improved robustness. Particular interest in distributed data fusion has been paid to calculating the "average" over the information owned by local sensors/agents via network communication [1]–[5]. Fundamentally, the average can be defined in two manners including, the *arithmetic average* (AA) and the *geometric average* (GA). Simply put, the former is a type of linear/convex fusion, akin to the linear opinion pool approach, while the latter is nonlinear/logarithmic fusion akin to the logarithmic opinion pool approach [6], [7].

In the context of multi-sensor/multi-agent target tracking, the two most important types of information for fusion among local sensors/agents are random variables (representing parameters such as the number of targets, clutter rate, etc.) and probability density functions (PDFs), for which the fusion is referred to as v-fusion and f-fusion, respectively. In the latter, the GA fusion coincides with the Chernoff fusion [8], [9], which is also known as (generalized) covariance intersection (CI) [10]–[15]. The CI approach was originally proposed for fusing correlated estimates produced at distinct but not necessarily independent sensors when, unfortunately, the correction is unknown. The fundamental property of the CI approach is keeping consistency or conservativeness, avoiding double accounting any information in the fusion [10], [11], [16]. In direct relation with this approach, the Chernoff Information of a pair of distributions quantifies the best achievable exponent in the Bayesian probability of error [11]e. Therefore, the approach (more generally) is also known as Chernoff fusion. Further approaches to the combination of probability distributions in the presence of unknown correlation can be found in [16], [17].

In addition to the Bayesian posterior PDF, the fusing functions can also be the likelihood functions [18]–[20] or the probability hypothesis density (PHD)/intensity functions [21]–[27]. (The PHD [28] differs from the PDF in that its integral over any region gives the expected number of targets in that region which can be any real number.) In comparison, the AA fusion has also been applied for PHD fusion [29]–[35] and for raw data fusion in the means of clustering [36], [37]. Both averaging approaches to data fusion have demonstrated, either theoretically or experimentally, gains in estimation accuracy and/or robustness, whereas weaknesses have also been identified from different viewpoints [6], [7], [22], [26], [32]–[34], [38]–[41].

Despite a few analysis about the variance alone of the GA/CI [11], [14], [21] and of the AA [38] in f-fusion and about the mean square error (MSE) of the AA [35] for uncorrected v-fusion, comprehensive analysis and comparison of the statistics (including the variance and MSE) of both approaches, are still lacking. Therefore, it is actually not so transparent how different they are and how they compare with the fusing estimator.

In this paper, we are not intended to investigate the motivation behind both approaches, or propose any new algorithms. Rather, we analyze and compare the statistics of the GA and of the AA, with respect to v-fusion and f-fusion, respectively. The analysis is based on the classic estimation perspective. In addition to analytic analysis, approximate analysis or exemplifying illustration (mainly based on the most common Gaussian assumption) are also given. For simplicity, we restrict our discussion in the scalar real space $\mathbb{R}$.

The paper is organized as follows. Preliminaries are briefly

Manuscript received on xxx

T. Li is with the BISITE research group, University of Salamanca, Salamanca 37007, Spain, e-mail: t.c.li@usal.es

H. Fan is with ATR laboratory, National University of Defense Technology, Chang-Sha 410073, China. Email:fanhongqi@nudt.edu.cn

J.G. Herrerois is with Computer Science and Engineering, Universidad Carlos III de Madrid, Madrid, Spain, E-mail: jgherrer@inf.uc3m.es

J.M. Corchado is with the BISITE research group, University of Salamanca, Salamanca 37007, Spain and also the Department of EIC, Osaka Institute of Technology, Osaka 535-8585, Japan. E-mail:corchado@usal.es

This work was partially supported by the Marie Skłodowska-Curie Individual Fellowship (H2020-MSCA-IF-2015) under Grant 709267.

introduced in Section II. Our major analysis of the *v*-fusion and of the *f*-fusion is given in sections III and IV, respectively. A hybrid use of both approaches for averaging Gaussian mixtures is discussed in Sec.V, which shows further comparison between *v*-fusion and *f*-fusion. Key findings are summarized by *Remarks* throughout the paper and in Section VI.

II. PRELIMINARIES

A. Definitions

For the unknown parameter θ of interest that takes value in a measurable region $\times \subseteq \mathbb{R}$, the estimator $\hat{\theta}_i$ associated with PDF $f_{\hat{\theta}_i}(x)$, is unbiased if it yields, on average, the true value of the parameter [42, Sec. 2.3], i.e.,

$$\bar{\theta}_i \triangleq \mathrm{E}_{f_{\hat{\theta}_i}}[\hat{\theta}_i] = \int_\times x f_{\hat{\theta}_i}(x)dx = \theta. \tag{1}$$

The variance of the estimator $\hat{\theta}_i$ is given as

$$\Sigma_{\hat{\theta}_i} \triangleq \int_\times \left(x-\bar{\theta}_i\right)^2 f_{\hat{\theta}_i}(x)dx = \int_\times x^2 f_{\hat{\theta}_i}(x)dx - \left(\bar{\theta}_i\right)^2. \tag{2}$$

where $f_{\hat{\theta}_i}(x)$ denotes the PDF of the estimator $\hat{\theta}_i$.

The MSE of an estimator $\hat{\theta}_i$ is given as [42, Sec. 2.4]

$$\mathrm{mse}(\hat{\theta}_i) \triangleq \mathrm{E}_{f_{\hat{\theta}_i}(x)}[(\theta-\hat{\theta}_i)^2] = \int_\times (\theta-x)^2 f_{\hat{\theta}_i}(x)dx. \tag{3}$$

Straightforwardly, an expansion of (3) will lead to

$$\mathrm{mse}(\hat{\theta}_i) = \Sigma_{\hat{\theta}_i} + (\bar{\theta}_i-\theta)^2. \tag{4}$$

That is, the MSE of an estimator equals the sum of its variance and the square of its bias (if any).

Furthermore, suppose that $f : X \to \mathbb{R}$ is a real-valued function whose domain is a set X. The set-theoretic support of f, denoted as $\mathrm{supp}(f)$, is the set of points in X where f is non-zero, i.e.,

$$\mathrm{supp}(f) = \{x \in X | f(x) \neq 0\}. \tag{5}$$

B. Averaging in Terms of Variables: v-fusion

Let us consider estimators $\hat{\theta}_i, i \in \mathcal{I} \subseteq \mathbb{N}$ given in terms of random variables, such as the estimate of the number of targets [35]. There is no PDF or uncertainty information available and so only point estimates that are random variables are involved. Their variable-AA is given as

$$\hat{\theta}_v^{\mathrm{AA}} \triangleq \sum_{i\in\mathcal{I}} \omega_i \hat{\theta}_i. \tag{6}$$

Hereafter, the fusing weights $\omega_i \in (0,1)$, $\sum_{i\in\mathcal{I}} \omega_i = 1$. (Obviously, $\omega_i = 0$ indicates that information i does not really get involved in the fusion.)

In contrast to (6), the variable-GA is given as

$$\hat{\theta}_v^{\mathrm{GA}} \triangleq \prod_{i\in\mathcal{I}} \hat{\theta}_i^{\omega_i}. \tag{7}$$

Note that, the variable-GA fusion may lead to an imaginary number when the fusing variable is negative, which is beyond the consideration of this work.

Obviously, the GA fusion amounts to the AA fusion on the logarithms of the variables, namely, (cf. (6))

$$\log \hat{\theta}_v^{\mathrm{GA}} = \sum_{i\in\mathcal{I}} \omega_i \log \hat{\theta}_i. \tag{8}$$

C. Averaging in Terms of PDFs: f-fusion

When the local estimator–e.g., a Bayesian estimator–is given as a function such as a PDF or a PHD, the f-fusion is involved. Given estimators $\hat{\theta}_i$ with PDFs $f_{\hat{\theta}_i}(x)$, $i \in \mathcal{I}$, their PDF-AA is given as

$$f_{\hat{\theta}_{\mathrm{AA}}}(x) = \sum_{i\in\mathcal{I}} \omega_i f_{\hat{\theta}_i}(x). \tag{9}$$

and the PDF-GA is given as

$$f_{\hat{\theta}_{\mathrm{GA}}}(x) = C^{-1} \prod_{i\in\mathcal{I}} \left(f_{\hat{\theta}_i}(x)\right)^{\omega_i}, \tag{10}$$

where $C \triangleq \int_\times \prod_{i\in\mathcal{I}} \left(f_{\hat{\theta}_i}(x)\right)^{\omega_i} dx$ is a normalization term to ensure the result being a PDF (if possible). (In the generalized GA fusion applied to the PHDs [21], [22], such a normalization is unnecessary. But then, the fused result of two PDFs (that can be viewed as two specific PHDs with unit cardinality) may not be a PDF [22].)

Due to the nonnegative definiteness of the probability, the support of the AA fusion PDF is the union of those of all initial PDFs. In contrast, the support of $f_{\hat{\theta}_{\mathrm{GA}}}(x)$ is the intersection of those of all initial PDFs, which may be empty. We leave any further discussion on this point and we assume all fusing estimators have the same support unless otherwise stated.

III. AVERAGING OVER VARIABLES: *v*-FUSION

A. Variance Analysis

The variance of a weighted sum of multiple variables is given by the weighted sum of their covariances [43], i.e.,

$$\begin{aligned}\Sigma_{\hat{\theta}_v^{\mathrm{AA}}} &= \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{I}} \mathrm{Cov}(\omega_i\hat{\theta}_i, \omega_j\hat{\theta}_j) \\ &= \sum_{i\in\mathcal{I}} \omega_i^2 \Sigma_{\hat{\theta}_i} + \sum_{i<j\in\mathcal{I}} 2\omega_i\omega_j \mathrm{Cov}(\hat{\theta}_i, \hat{\theta}_j).\end{aligned} \tag{11}$$

Here, $\mathrm{Cov}(\hat{\theta}_1, \hat{\theta}_2)$ denotes the correlation between $\hat{\theta}_1$ and $\hat{\theta}_2$.

Let us consider two variables for simplicity and define

$$\rho \triangleq \frac{\mathrm{Cov}(\hat{\theta}_1, \hat{\theta}_2)}{\sqrt{\Sigma_{\hat{\theta}_1}\Sigma_{\hat{\theta}_2}}}. \tag{12}$$

Then, (11) reduces to

$$\Sigma_{\hat{\theta}_v^{\mathrm{AA}}} = \omega_1^2\Sigma_{\hat{\theta}_1} + \omega_2^2\Sigma_{\hat{\theta}_2} + 2\omega_1\omega_2\rho\sqrt{\Sigma_{\hat{\theta}_1}\Sigma_{\hat{\theta}_2}}. \tag{13}$$

Here, $-1 \leq \rho \leq 1$ is known as the correlation coefficient [44, Chapt. 4] between two variables. For two independent variables, $\rho = 0$ and for two identical variables, $\rho = 1$.

We now analyze the bounds of $\Sigma_{\hat{\theta}_v^{\mathrm{AA}}}$. First, it is easy to be verified that $\Sigma_{\hat{\theta}_v^{\mathrm{AA}}} \leq \left(\omega_1\sqrt{\Sigma_1} + \omega_2\sqrt{\Sigma_2}\right)^2 \leq \max(\Sigma_1, \Sigma_2)$. That is, we have the upper bound of $\Sigma_{\hat{\theta}_v^{\mathrm{AA}}}$

$$\Sigma_{\hat{\theta}_v^{\mathrm{AA}}} \leq \max(\Sigma_1, \Sigma_2), \tag{14}$$

where the bound is approached when $w_1 \to 0$ (if $\Sigma_1 \leq \Sigma_2$) or $w_1 \to 1$ (if $\Sigma_1 \geq \Sigma_2$).

To derive the lower bound of $\Sigma_{\hat{\theta}_v^{\mathrm{AA}}}$, we further define

$$\alpha \triangleq \frac{\Sigma_2}{\Sigma_1}.$$

Due to the symmetry of the expression of $\Sigma_{\hat{\theta}_v^{\mathrm{AA}}}$, we only consider $\alpha \geq 1$ in our analysis; the results hold by exchanging Σ_2 with Σ_1 if $\alpha < 1$. We define a convex function of $w \in (0, 1)$ as follows

$$h(w;\alpha,\rho) \triangleq 1 - 2w + w^2 + w^2\alpha + 2\rho\alpha^{\frac{1}{2}}(w - w^2). \quad (15)$$

In what follows, we write $h(w;\alpha,\rho)$ as $h(w)$ for short. So, we have $\Sigma_{\hat{\theta}_v^{\mathrm{AA}}} = h(\omega_2)\Sigma_1$.

Bound analysis of the function $h(w)$ is given in Appendix A. As shown, if $\rho < \alpha^{-\frac{1}{2}}$, the optimal fusing weights that correspond to the minimum $h(w)$ are given by (cf. (47))

$$\omega_1 = \frac{\alpha - \rho\alpha^{\frac{1}{2}}}{1 + \alpha - 2\rho\alpha^{\frac{1}{2}}}, \quad \omega_2 = \frac{1 - \rho\alpha^{\frac{1}{2}}}{1 + \alpha - 2\rho\alpha^{\frac{1}{2}}}, \quad (16)$$

The corresponding lower bound of $\Sigma_{\hat{\theta}_v^{\mathrm{AA}}}$ is given by

$$\Sigma_{\hat{\theta}_v^{\mathrm{AA}}} \geq \frac{\alpha(1 - \rho^2)}{1 + \alpha - 2\rho\alpha^{\frac{1}{2}}}\Sigma_1. \quad (17)$$

Otherwise (if $\rho \geq \alpha^{-\frac{1}{2}}$), the lower bound is given as

$$\Sigma_{\hat{\theta}_v^{\mathrm{AA}}} \geq \min(\Sigma_1, \Sigma_2), \quad (18)$$

where the bound is approached when $w_1 \to 1$ (if $\Sigma_1 \leq \Sigma_2$) or $w_1 \to 0$ (if $\Sigma_1 \geq \Sigma_2$).

Remark 1. *For v-fusion, the variance of the AA is upper bounded by the greatest variance of the fusing estimators. Its lower bound as given in* (17) *is smaller than the smallest variance of the fusing estimators if the correlation coefficient between two fusing estimator satisfies*

$$\frac{\mathrm{Cov}(\hat{\theta}_1, \hat{\theta}_2)}{\sqrt{\Sigma_{\hat{\theta}_1}\Sigma_{\hat{\theta}_2}}} < \left(\frac{\min\left(\Sigma_1, \Sigma_2\right)}{\max\left(\Sigma_1, \Sigma_2\right)}\right)^{\frac{1}{2}}, \quad (19)$$

otherwise, the lower bound is given by the smallest variance of the fusing estimators.

Notably, when these two variables are inversely correlated namely $\mathrm{Cov}(\hat{\theta}_1, \hat{\theta}_2) < 0$, (19) always holds.

On the other hand, to calculate $\Sigma_{\hat{\theta}_v^{\mathrm{GA}}}$, substituting $\hat{\theta}_v^{\mathrm{AA}}, \hat{\theta}_i, i \in \mathcal{I}$ in (11) with $\log\hat{\theta}_v^{\mathrm{GA}}, \log\hat{\theta}_i, i \in \mathcal{I}$ (cf. (8)), respectively, yields

$$\begin{aligned}\Sigma_{\log\hat{\theta}_v^{\mathrm{GA}}} &= \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{I}}\mathrm{Cov}(\omega_i\log\hat{\theta}_i, \omega_j\log\hat{\theta}_j) \\ &= \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{I}}\omega_i\omega_j\mathrm{Cov}(\log\hat{\theta}_i, \log\hat{\theta}_j). \quad (20)\end{aligned}$$

The above formulation involves the calculation of the covariance between (logarithmic) functions of two random variables, which can be addressed in terms of the cumulative distribution function; for the detail the reader is kindly referred to [45]. We omit further analytic analysis on this mathematical problem, but instead, to gain insight and to illustratively compare between the AA and the GA, we study two representative examples by means of the Monte Carlo simulation.

1) Numerical analysis for Gaussian v-fusion: Note that the GA of two Gaussian variables is no longer a Gaussian variable (unless two fusing variables are identical). For numerical illustration, we here consider two approximate Gaussian distributions with $\mu_1 = 50, \Sigma_1 = 100$, and with $\mu_2 = 60, \Sigma_2 = 200$, respectively, in which the negative support of the Gaussian PDF (which is actually ignorable in the given examples as the negative part is far more than 4-sigma to the mean of the distribution) is truncated and so all samples are guaranteed to be positively valued in order to avoid the imaginary number problem of the GA fusion.

Two groups of samples are generated with different correlation coefficients ρ. Correspondingly, the means and variances of the AA and of the GA are given in Fig. 1. As shown in the upper right sub-figure, when $\rho = 0.70846 > \alpha^{-\frac{1}{2}}$ (as $\alpha = \frac{\Sigma_2}{\Sigma_1} = 2$), we obtain dual bounds of $\Sigma_{\hat{\theta}_v^{\mathrm{AA}}}$ as shown in (18) otherwise (as shown in the other sub-figures) the variances of the AA can be smaller than the lowest variance of the fusing estimator. It is further seen that,

Remark 2. *The variances of the AA and of the GA can be either greater or smaller than each other, depending on the choice of the fusing weights. There is a cross-over of their values (namely the smaller becomes the greater) as ω_1 increases from 0 to 1. For certain ρ and α, the lowest AA variance that can be yielded by adjusting the fusing weights is never greater than that of the GA .*

2) Numerical analysis for Poisson v-fusion: We further consider two Poisson variables $\hat{\theta}_1 \sim \mathrm{Poisson}(\lambda_1 = 12)$ and $\hat{\theta}_2 \sim \mathrm{Poisson}(\lambda_2 = 10)$, where λ_1 and λ_2 are the Poisson rates which indicate both the mean and variance of the variable. The Poisson variable is important in the tracking community, e.g., the number of targets or of false alarms that appear at a given time-interval is often modeled as a Poisson variable [28], [46]. Note that both AA and GA of two Poisson variables are no longer Poisson variables. Once more, we use the Monte Carlo method for numerical approximation. The means and variances of the AA and GA of two Poisson random variables under different correlation coefficient ρs and fusing weights are given in Fig. 2. The results are highly consistent to what shown in the Gaussian case (cf. Fig. 1) and the statement given in Remark 2 still holds.

B. MSE Analysis for AA and Numerical Comparison to GA

In this section, we study the MSE of the AA and numerically compare it with that of the GA, based on general variables that may be correlated.

Inserting (6) in (3) yields

$$\begin{aligned}\mathrm{mse}\big(\hat{\theta}_v^{\mathrm{AA}}\big) =& \mathrm{E}_{f_{\hat{\theta}_{\mathrm{AA}}}(x)}\big[\big(\omega_1(\theta - \hat{\theta}_1) + \omega_2(\theta - \hat{\theta}_2)\big)^2\big] \\ =& \omega_1^2\mathrm{mse}(\hat{\theta}_1) + \omega_2^2\mathrm{mse}(\hat{\theta}_2) \\ &+ 2\omega_1\omega_2\beta\sqrt{\mathrm{mse}(\hat{\theta}_1)\mathrm{mse}(\hat{\theta}_2)}, \quad (21)\end{aligned}$$

where $\beta \triangleq \frac{\mathrm{E}_{f_{\hat{\theta}_{\mathrm{AA}}}(x)}\left[(\theta - \hat{\theta}_1)(\theta - \hat{\theta}_2)\right]}{\sqrt{\mathrm{mse}(\hat{\theta}_1)\mathrm{mse}(\hat{\theta}_2)}} \in (-1, 1)$.

As addressed, the fractional order of a Gaussian variable may involve imaginary numbers. Therefore, we cannot simply get the MSE of the GA for v-fusion. To overcome this, once

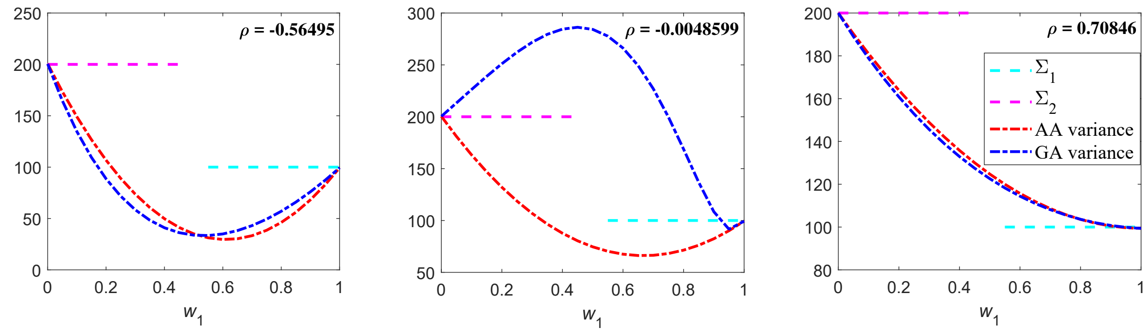

Fig. 1. Comparison of the variances of the AA and GA of two correlated, approximate-Gaussian-distributed variables with mean $\mu_1 = 50$ and variance $\Sigma_1 = 100$, and with mean $\mu_2 = 60$ and variance $\Sigma_2 = 200$, respectively, under three different correlation coefficients ρ and fusing weights w_1.

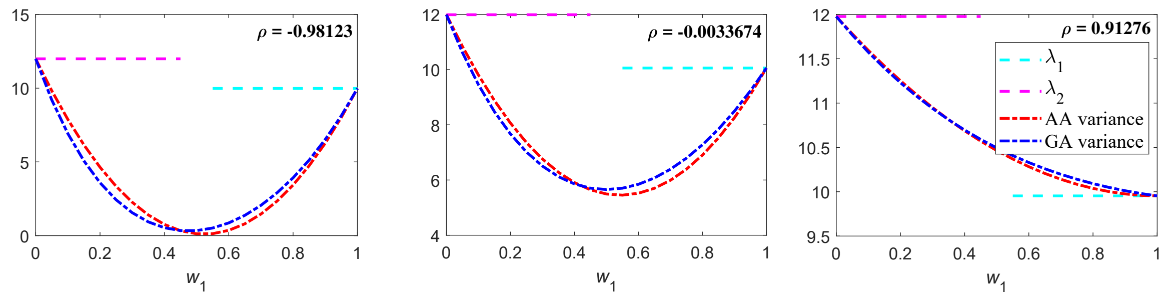

Fig. 2. Comparison of the variances of the AA and GA of two Poisson-distributed variables with rates $\lambda_1 = 10, \lambda_2 = 12$ (and so $\alpha = \frac{\lambda_2}{\lambda_1} = 1.2$), under three different correlation coefficients ρ and fusing weights.

more, by means of the Monte Carlo simulation, we consider two approximate Gaussian variables $\hat{\theta}_1(x) \sim \mathcal{N}(x; 50, 100)$ and $\hat{\theta}_2(x) \sim \mathcal{N}(x; 60, 200)$, for which we simulate three different real variables $\theta = 45, 55, 65$, respectively, for different βs. The v-fusion results based on different fusing weights $\omega_1 \in (0, 1)$ are shown in Fig. 3 for the case of two independent variables and in Fig. 4 for the case of two correlated variable with correlation coefficient $\rho = 0.70736$. The results show that

Remark 3. *The MSE of the AA can be either greater or smaller than that of the GA when different fusing weights are used. The greatest discrepancy between them occurs when the fusing weights are at certain points in the scope (0,1). The lowest bound of the MSE of either the AA or the GA is their corresponding variances, which are obtained when the fused estimates turn out to be unbiased which is only possible when the real parameter lies between two variables. Accordingly, the lower bound of the MSE of the AA is smaller than that of the GA.*

1) Bounds and Comparison: To gain analytical results on the MSE of the AA for v-fusion, we define

$$\gamma \triangleq \frac{\mathrm{mse}(\hat{\theta}_2)}{\mathrm{mse}(\hat{\theta}_1)}.$$

Then, it can be easily varified that $\mathrm{mse}\big(\hat{\theta}_v^{\mathrm{AA}}\big) = h(\omega_2)\mathrm{mse}(\hat{\theta}_1)$, where $h(w)$ is defined in (15) (with ρ and α replaced by β and γ, respectively). Therefore, analogous to our analysis in Sec. III-A, lower and upper bounds of $\mathrm{mse}\big(\hat{\theta}_v^{\mathrm{AA}}\big)$ can be obtained by using the same optimal fusing weights ω_1 and ω_2 as in (47). Akin to Remark 1, we have:

Remark 4. *For v-fusion, the upper bound of the MSE of the AA is given by the greatest MSE of the fusing estimators. The lower bound is smaller than the smallest MSE of the fusing estimators if the correlation between two fusing estimators satisfies*

$$\beta < \left(\frac{\min\left(\mathrm{mse}(\hat{\theta}_1), \mathrm{mse}(\hat{\theta}_2)\right)}{\max\left(\mathrm{mse}(\hat{\theta}_1), \mathrm{mse}(\hat{\theta}_2)\right)} \right)^{\frac{1}{2}}, \tag{22}$$

otherwise, the lower bound is given by the smallest MSE of the fusing estimators.

Notably, when the real parameter θ lies on or between $\bar{\theta}_1$ and $\bar{\theta}_2$ namely $\mathrm{E}_\times\left[(\theta - \hat{\theta}_1)(\theta - \hat{\theta}_2)\right] \leq 0$ and so $\beta \leq 0$, (22) always holds.

2) Unweighted AA: The MSE is a key metric in evaluating an estimator/tracker. However, in practice, γ is often unknown since the MSE of each fusing estimator that is calculated based on the real parameter is practically unknown. (In the literature, e.g., [11]–[13], [15], [24], [33], the most common approach to designing the fusing weights is based on minimizing the (trace or determinant of) variance, which only equals the MSE when the estimator is unbiased.) One may simply choose to use uniform fusing weights $\omega_1 = \omega_2 = 0.5$, namely unweighted averaging. Then, we obtain (cf. (15))

$$h(0.5)|_{\text{unweighted}} = \frac{1 + \gamma + 2\beta\gamma^{\frac{1}{2}}}{4}. \tag{23}$$

In this case, a sufficient and necessary condition for the un-weighted AA fusion to be "better" than the best fus-

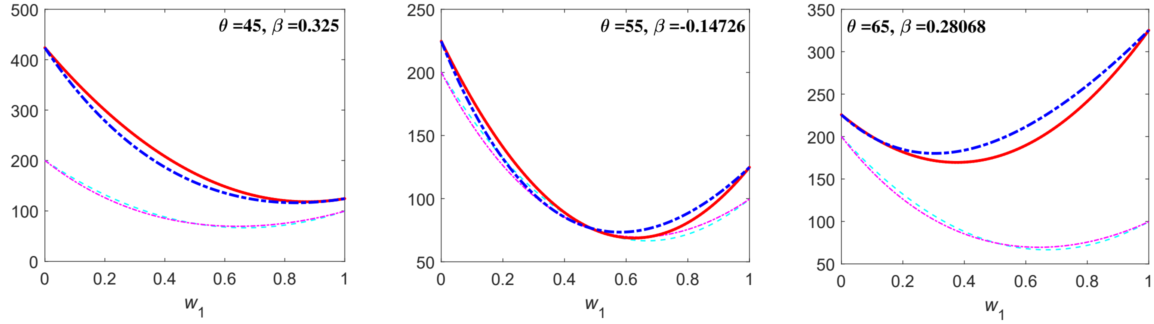

Fig. 3. Comparison of the MSEs of the AA and of the GA of two independent, approximate Gaussian variables $\hat{\theta}_1(x) \sim \mathcal{N}(x; 50, 100)$ and $\hat{\theta}_2(x) \sim \mathcal{N}(x; 60, 200)$, in the case of three different real variables $\theta = 45, 55, 65$, respectively, when different fusing weights are used.

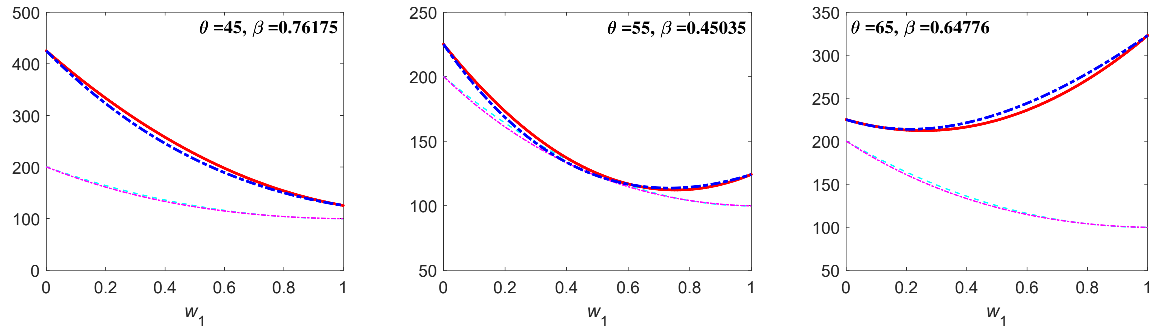

Fig. 4. Comparison of the MSEs of the AA and of the GA of two approximate Gaussian variables $\hat{\theta}_1(x) \sim \mathcal{N}(x; 50, 100)$ and $\hat{\theta}_2(x) \sim \mathcal{N}(x; 60, 200)$ (with correlation coefficient $\rho = 0.70736$), in the case of three different real variables $\theta = 45, 55, 65$, respectively, when different fusing weights are used.

ing estimator in the sense of obtaining smaller MSE (i.e., $h(0.5)|_{\text{unweighted}} < 1$) is given by (if possible)

$$\beta < \frac{3-\gamma}{2\gamma^{\frac{1}{2}}} \triangleq g(\gamma). \tag{24}$$

Calculating the derivative of $g(\gamma)$ with respect to γ yields

$$\frac{\mathrm{d}\, g(\gamma)}{\mathrm{d}\gamma} = \frac{-1 - 3\gamma^{-1}}{4\gamma^{\frac{1}{2}}} < 0, \tag{25}$$

which indicates that $g(\gamma)$ decreases with the increase of γ, and therefore, $g(\gamma) < -1, \forall \gamma > 9$. Since $-1 < \beta$, we therefore assert that $h(0.5)|_{\text{unweighted}} < 1$ is impossible if $\gamma > 9$. In this case, the MSE of the unweighted AA must lie between the best and the worst of the MSEs of the fusing estimators.

IV. Averaging over PDFs: f-Fusion

A. *Variance Analysis (for Two Gaussian PDFs)*

In this section, we analyze the variances of the PDF-AA $f_{\hat{\theta}_{\text{AA}}}(x)$ and PDF-GA $f_{\hat{\theta}_{\text{GA}}}(x)$ for fusing two Gaussian PDFs $f_{\hat{\theta}_1}(x) = \mathcal{N}(x; \mu_1, \Sigma_1)$ and $f_{\hat{\theta}_2}(x) = \mathcal{N}(x; \mu_2, \Sigma_2)$.

1) General Result: In the addressed case, (9) reduces to a GM-PDF $f_{\hat{\theta}_f^{\text{AA}}}(x) = \omega_1 \mathcal{N}(x; \mu_1, \Sigma_1) + \omega_2 \mathcal{N}(x; \mu_2, \Sigma_2)$ whose mean $\bar{\theta}_f^{\text{AA}}$ and variance $\Sigma_{\hat{\theta}_f^{\text{AA}}}$ are

$$\bar{\theta}_f^{\text{AA}} = \omega_1 \mu_1 + \omega_2 \mu_2, \tag{26}$$

$$\Sigma_{\hat{\theta}_f^{\text{AA}}} = \omega_1 \Sigma_1 + \omega_2 \Sigma_2 + \Delta(\omega_1, \omega_2), \tag{27}$$

respectively, where $\Delta(\omega_1, \omega_2) \triangleq \omega_1 \omega_2 (\mu_1 - \mu_2)^2 \geq 0$.

In contrast, the GA of two Gaussian PDFs remains a Gaussian PDF. That is, (10) reduces to a single Gaussian PDF $f_{\bar{\theta}_f^{\text{GA}}}(x) = \mathcal{N}(x; \bar{\theta}_f^{\text{GA}}, \Sigma_{\hat{\theta}_f^{\text{GA}}})$ with [21]

$$\Sigma_{\hat{\theta}_f^{\text{GA}}} = \frac{\Sigma_1 \Sigma_2}{\omega_1 \Sigma_2 + \omega_2 \Sigma_1}, \tag{28}$$

$$\bar{\theta}_f^{\text{GA}} = \frac{\omega_1 \Sigma_1^{-1} \mu_1 + \omega_2 \Sigma_2^{-1} \mu_2}{\omega_1 \Sigma_1^{-1} + \omega_2 \Sigma_2^{-1}}. \tag{29}$$

As shown, both the mean of the AA as in (26) and the mean of the GA as in (29) show a linear combination of the means of the fusing estimators. Differently, the variances of the fusing estimators are also involved in the latter but not in the former. In what follows, we analyze and compare their variances as in (27) and (28).

2) Bounds and Comparison: Given $0 < \omega_1, \omega_2 < 1$, we obtain the dual, tight bounds on $\Sigma_{\hat{\theta}_f^{\text{GA}}}$ from (28)

$$\min\left(\Sigma_1, \Sigma_2\right) \leq \Sigma_{\hat{\theta}_f^{\text{GA}}} \leq \max\left(\Sigma_1, \Sigma_2\right), \tag{30}$$

where the equations hold when and only when $\Sigma_1 = \Sigma_2$ for which $\Sigma_{\hat{\theta}_f^{\text{GA}}} = \Sigma_1 = \Sigma_2$, regardless of the fusing weights. Otherwise, if $\Sigma_1 \neq \Sigma_2$, the bounds are approached when $\omega_1 \to 0, \omega_2 \to 1$ (for one of the dual bounds) or when $\omega_1 \to 1, \omega_2 \to 0$ (for the other bound).

Since $\Delta(\omega_1, \omega_2) \geq 0$, we obtain

$$\Sigma_{\hat{\theta}_{\text{AA}}} > \omega_1 \Sigma_1 + \omega_2 \Sigma_2 \triangleq \text{LB}(\Sigma_{\hat{\theta}_f^{\text{AA}}}), \tag{31}$$

where, the lower bound of $\Sigma_{\hat{\theta}_f^{\mathrm{AA}}}$ is further dually, tightly bounded by (cf. (30))

$$\min\left(\Sigma_1,\Sigma_2\right) \leq \mathrm{LB}(\Sigma_{\hat{\theta}_f^{\mathrm{AA}}}) \leq \max\left(\Sigma_1,\Sigma_2\right).$$

However, $\Sigma_{\hat{\theta}_f^{\mathrm{AA}}}$ can not be upper bounded by the variances alone of the fusion estimators, since it has a component $\Delta(\omega_1,\omega_2)$ related to the discrepancy between the means of two fusing estimators.

Finally, we have the following derivation from (31)

$$\begin{aligned}
\mathrm{LB}(\Sigma_{\hat{\theta}_f^{\mathrm{AA}}}) &= \frac{\left(\omega_1\Sigma_1+\omega_2\Sigma_2\right)\left(\omega_1\Sigma_2+\omega_2\Sigma_1\right)}{\omega_1\Sigma_2+\omega_2\Sigma_1} \\
&= \frac{(\omega_1^2+\omega_1^2)\Sigma_1\Sigma_2+\omega_1\omega_2(\Sigma_1^2+\Sigma_2^2)}{\omega_1\Sigma_2+\omega_2\Sigma_1} \\
&\geq \frac{(\omega_1^2+\omega_1^2)\Sigma_1\Sigma_2+2\omega_1\omega_2(\Sigma_1\Sigma_2)}{\omega_1\Sigma_2+\omega_2\Sigma_1} \\
&= \frac{\Sigma_1\Sigma_2}{\omega_1\Sigma_2+\omega_2\Sigma_1} = \Sigma_{\hat{\theta}_f^{\mathrm{GA}}}.
\end{aligned} \tag{32}$$

In summary, we have the following remark (significantly different to the case of v-fusion given in Remark 2; numerical demonstration will be given in Fig. 5):

Remark 5. *For Gaussian f-fusion, the AA fusion always leads to a greater variance than the GA fusion does when they use the same fusing weights while the variance of the GA, but not that of the AA, is bounded by the smallest and greatest variances of the fusing estimators.*

Inflated variance due to AA or GA has also been pointed out by [14], [38], etc.

B. MSE Analysis

The MSE of $\hat{\theta}_f^{\mathrm{AA}}$ is calculated by (cf. (3))

$$\begin{aligned}
\mathrm{mse}\left(\hat{\theta}_f^{\mathrm{AA}}\right) &= \int_{\mathsf{x}}(\theta-x)^2\sum_{i\in\mathcal{I}}\omega_i f_{\hat{\theta}_i}(x)dx \\
&= \sum_{i\in\mathcal{I}}\omega_i\int_{\mathsf{x}}(\theta-x)^2 f_{\hat{\theta}_i}(x)dx \\
&= \sum_{i\in\mathcal{I}}\omega_i \mathrm{mse}\left(f_{\hat{\theta}_i}(x)\right)
\end{aligned} \tag{33}$$

which simply indicates that (cf. Remark 4 for v-fusion):

Remark 6. *The AA has an MSE that is the linearly weighted average of the MSEs of the fusing estimators and the MSE of the AA is bounded by the smallest and greatest MSEs of the fusing estimators.*

Expression (4) provides an easy way to calculate the MSE of $\hat{\theta}_f^{\mathrm{GA}}$ based on (28) and (29), i.e.,

$$\begin{aligned}
\mathrm{mse}\left(\hat{\theta}_f^{\mathrm{GA}}\right) &= \Sigma_{\hat{\theta}_f^{\mathrm{GA}}} + \left(\bar{\theta}_f^{\mathrm{GA}}-\theta\right)^2 \\
&= \frac{\Sigma_1\Sigma_2}{\omega_1\Sigma_2+\omega_2\Sigma_1} + \left(\frac{\omega_1\Sigma_1^{-1}\mu_1+\omega_2\Sigma_2^{-1}\mu_2}{\omega_1\Sigma_1^{-1}+\omega_2\Sigma_2^{-1}}-\theta\right)^2 \\
&= \underbrace{\frac{\Sigma_1\Sigma_2}{\omega_1\Sigma_2+\omega_2\Sigma_1}}_{\triangleq \mathrm{mse}_1\left(\hat{\theta}_f^{\mathrm{GA}}\right)} + \underbrace{\left(a\xi_1+b\xi_2\right)^2}_{\triangleq \mathrm{mse}_2\left(\hat{\theta}_f^{\mathrm{GA}}\right)}.
\end{aligned} \tag{34}$$

where $a \triangleq \frac{\omega_1\Sigma_1^{-1}}{\omega_1\Sigma_1^{-1}+\omega_2\Sigma_2^{-1}} \in (0,1)$, $b \triangleq \frac{\omega_2\Sigma_2^{-1}}{\omega_1\Sigma_1^{-1}+\omega_2\Sigma_2^{-1}} \in (0,1)$, $\xi_1 \triangleq \mu_1-\theta$ and $\xi_2 \triangleq \mu_2-\theta$.

It is easy to be verified that $\mathrm{mse}_1\left(\hat{\theta}_f^{\mathrm{GA}}\right) \geq \min(\Sigma_1,\Sigma_2)$, $\mathrm{mse}_2\left(\hat{\theta}_f^{\mathrm{GA}}\right) \geq 0$, and so $\mathrm{mse}\left(\hat{\theta}_f^{\mathrm{GA}}\right) \geq \min(\Sigma_1,\Sigma_2)$ where the equation holds when and only when both fusing Gaussian PDFs are unbiased and identical.

We now compare between $\mathrm{mse}\left(\hat{\theta}_f^{\mathrm{GA}}\right)$ and $\mathrm{mse}\left(\hat{\theta}_f^{\mathrm{AA}}\right)$. In the case of two Gaussian PDFs, combining (33) and (4) yields

$$\mathrm{mse}\left(\hat{\theta}_f^{\mathrm{AA}}\right) = \underbrace{\omega_1\Sigma_1+\omega_2\Sigma_2}_{\triangleq \mathrm{mse}_1\left(\hat{\theta}_f^{\mathrm{AA}}\right)} + \underbrace{\omega_1\xi_1^2+\omega_2\xi_2^2}_{\triangleq \mathrm{mse}_2\left(\hat{\theta}_f^{\mathrm{AA}}\right)}. \tag{35}$$

We have the following straightforward derivation

$$\begin{aligned}
&\left(\omega_1\Sigma_1+\omega_2\Sigma_2\right)\left(\omega_2\Sigma_1+\omega_1\Sigma_2\right) \\
&\quad= \left(\omega_1^2+\omega_1^2\right)\Sigma_1\Sigma_2+\omega_1\omega_2\left(\Sigma_1^2+\Sigma_2^2\right) \\
&\quad\geq \left(\omega_1^2+\omega_1^2\right)\Sigma_1\Sigma_2+2\omega_1\omega_2\left(\Sigma_1\Sigma_2\right) \\
&\quad= \Sigma_1\Sigma_2,
\end{aligned} \tag{36}$$

which indicates that

$$\mathrm{mse}_1\left(\hat{\theta}_f^{\mathrm{GA}}\right) \leq \mathrm{mse}_1\left(\hat{\theta}_f^{\mathrm{AA}}\right), \tag{37}$$

as long as both AA and GA fusion uses the same fusing weights.

To compare between $\mathrm{mse}_2\left(\hat{\theta}_f^{\mathrm{GA}}\right)$ and $\mathrm{mse}_2\left(\hat{\theta}_f^{\mathrm{AA}}\right)$, we consider two specific cases: First, if both fusing Gaussian PDFs are unbiased, i.e., $\mu_1=\mu_2=\theta$, we have $\mathrm{mse}_2\left(\hat{\theta}_f^{\mathrm{GA}}\right) = \mathrm{mse}_2\left(\hat{\theta}_f^{\mathrm{AA}}\right)$ and further by using (37),

$$\min\left(\Sigma_1,\Sigma_2\right) \leq \mathrm{mse}\left(\hat{\theta}_f^{\mathrm{GA}}\right) \leq \mathrm{mse}\left(\hat{\theta}_f^{\mathrm{AA}}\right) \leq \max\left(\Sigma_1,\Sigma_2\right)$$

where the bounds are approached when the two fusing weights approach 0 and 1, respectively.

Secondly, if $\Sigma_1=\Sigma_2$, we have $\omega_1=a, \omega_2=b$. Subsequently, the following straightforward derivation is obtained

$$\mathrm{mse}_2\left(\hat{\theta}_f^{\mathrm{GA}}\right) - \mathrm{mse}_2\left(\hat{\theta}_f^{\mathrm{AA}}\right) = -\omega_1\omega_2\left(\xi_1-\xi_2\right)^2 \leq 0, \tag{38}$$

namely $\mathrm{mse}_2\left(\hat{\theta}_f^{\mathrm{GA}}\right) \leq \mathrm{mse}_2\left(\hat{\theta}_f^{\mathrm{AA}}\right)$, as long as they use the same fusing weights. Combining this with (37) yields

$$\mathrm{mse}\left(\hat{\theta}_f^{\mathrm{GA}}\right) \leq \mathrm{mse}\left(\hat{\theta}_f^{\mathrm{AA}}\right).$$

Remark 7. *If both fusing Gaussian PDFs are unbiased or if they have the same variance, the MSE of the GA is smaller than or equals that of the AA and is always greater than the smallest variance of the fusing estimators.*

To gain further insight into their difference in the general case, by means of the Monte Carlo simulation, we consider two Gaussian PDFs $f_{\hat{\theta}_1}(x) = \mathcal{N}(x;\mu_1=50,100)$ and $f_{\hat{\theta}_2}(x) = \mathcal{N}(x;\mu_2=60,200)$ and two Gaussian PDFs $f_{\hat{\theta}_1}(x) = \mathcal{N}(x;\mu_1=50,400)$ and $f_{\hat{\theta}_2}(x) = \mathcal{N}(x;\mu_2=60,200)$, respectively. The results are shown in Fig. 5 and Fig. 6, respectively, for the real parameter $\theta \in [40,80]$ and fusing weight $\omega_1 \in (0,1)$. It is seen that (cf. Remark 3 for v-fusion)

Remark 8. *For Gaussian f-fusion, the MSE of the AA is in most cases greater than that of the GA, unless θ is considerably greater than $\max(\mu_1,\mu_2)$ and the fusing estimator that has a greater mean has a greater variance. Different to the* *case of v-fusion, there is no cross-over of their MSEs when the fusing weights change. That is, for certain PDFs and real parameter, one is always better than the other.*

V. Averaging over Mixtures: PHD-fusion

In practice, the variable θ of interest may regard multiple objects for which the posterior distribution to be fused is "multimodal", whose integral is no more (but usually greater than) unit, which is typically given by means of a mixture of multiple sub-functions (each of which can be referred to as a component) such as a GM. In this case, while the AA of a sum can be straightforwardly expressed as a cascaded sum of the fusing sums (after re-weighting them) that remains in the same form, the fractional order exponential power of a sum does not remain as a sum of the same form, and typically approximation must be resorted to; see, e.g., [23], [31], [47]–[49]. In the following, we briefly review a popular, analytical approximation approach to GA fusion of GMs proposed by [23], etc. and compare it with the AA fusion by means of the Monte Carlo simulation.

A. Approximate GM-GA Fusion

By omitting the cross-products of different Gaussian functions/components(GCs)[1], the fractional order exponential power of a GM consisting of n GCs can be approximated by

$$\left[\sum_{i=1}^{n} w_i \mathcal{N}(x; m_i, P_i)\right]^{\omega} \approx \sum_{i=1}^{n} \left[w_i \mathcal{N}(x; m_i, P_i)\right]^{\omega} \quad (39)$$

where the covariance inflation of CI, for a weighted Gaussian PDF is equivalent to raising the Gaussian function to a power, which remains Gaussian, namely

$$[w\mathcal{N}(x; m, P)]^{\omega} = w^{\omega}\epsilon(\omega, P)\mathcal{N}(x; m, \frac{P}{\omega}) \quad (40)$$

where $\epsilon(\omega, P) = \sqrt{\frac{\det(2\pi P \omega^{-1})}{[\det(2\pi P)]^{\omega}}} = \sqrt{(2\pi P)^{(1-\omega)}\omega^{-1}}$ [21].

In addition, the product of two GCs remains a GC, i.e.,

$$w_1\mathcal{N}(x; m_1, P_1) w_2\mathcal{N}(x; m_2, P_2) = w_{12}\mathcal{N}(x; m_{12}, P_{12}) \quad (41)$$

where $P_{12} = (P_1^{-1} + P_2^{-1})^{-1}, m_{12} = P_{12}(m_1 P_1^{-1} + m_2 P_2^{-1}), w_{12} = w_1 w_2 \mathcal{N}(m_1 - m_2; 0, P_1 + P_2)$ in which the coefficient $\mathcal{N}(m_1 - m_2; 0, P_1 + P_2)$ measures the separation of the two GCs.

By using (39), (40) and (41), the GA-fusion of two GMs can then be obtained; the interested readers are kindly referred to [23] for the detail. We note that

1) The GM-GA fusion requires fusing all pairs of GCs between neighboring sensors, which will result in a multiplied number of GCs. That is, the GA of J_1 GCs and J_2 GCs is a mixture of $J_1 \cdot J_2$ GCs while it is a mixture of $J_1 + J_2$ GCs in the case of AA fusion.
2) To perform the GA fusion as addressed above, the local GM-PHD needs to be normalized to a PDF (cf. (39)). At the end, the resultant GCs need to be properly weighted, such that their sum equals the average of the original weight sums of the fusing GMs. To this end, an extra cardinality consensus scheme may be performed [23], [35], [50]. That is to say, the distribution fusion and the cardinality fusion are performed separately.

[1]This can only make sense in the case where the GCs in the mixture are well distant. In other words, the peaks in the multimodal distribution are well distant. This is quite limited as there are commonly closely-distributed GCs because of closely distributed targets [50].

B. Numerical Comparison between GM-AA and GM-GA

As addressed in Sec. III and Sec. IV, the AA performs better in v-fusion in the sense of yielding smaller bounds on the variance (cf. Remark 2) and MSE (cf. Remark 3) while the GA performs better in f-fusion in the sense of always yielding smaller variance (cf. Remark 5) and smaller MSE in most cases (cf. Remarks 7 and 8). Therefore, we advocate for the PHD fusion in a hrbrid means by using both v-fusion (for cardinality fusion) and f-fusion (for the distribution fusion). This is different to the usual, pure GM-GA fusion [20], [23], [25], [51] and pure GM-AA fusion [33], [50].

To gain the insight into such a hybrid rules for GM-PHD fusion, we consider an example in which two GMs are fused in the manner of unweighted AA fusion and unweighted GA fusion, respectively. Here, unweighted means $\omega_1 = \omega_2 = 0.5$. One GM referred to as GM 1 is given by three GCs (of weight sum 1.8) as follows

$$f_1(x) = 0.7\mathcal{N}(x; 10, 100) + 0.6\mathcal{N}(x; 50, 100) + 0.5\mathcal{N}(x; 90, 200), \quad (42)$$

and the other referred to as GM 2 is given by two GCs (of weight sum 1.7) as follows

$$f_2(x) = 0.9\mathcal{N}(x; 11, 100) + 0.8\mathcal{N}(x; 52, 120). \quad (43)$$

As shown, the two GCs $\mathcal{N}(x; 10, 100)$ and $\mathcal{N}(x; 50, 100)$ in GM 1 match the two GCs $\mathcal{N}(x; 11, 100)$ and $\mathcal{N}(x; 52, 120)$ in GM 2, respectively. They are likely indicating two respective targets. However, there is one extra GC $\mathcal{N}(x; 90, 200)$ in GM 1, which could be either a false alarm (generated in GM 1) or a real detection (and then there is a misdetection in GM 2) - we hereafter refer to this GC as an isolated GC. The fusion results are given in Fig. 7 in which the fused GM-AA or GM-GA is given in the manner of showing each GC or showing the joint distribution of them, where the joint distribution is superimposition of those of each GC distribution along the state space. We obtain the following two remarks (the first of which is consistent with Remark 5):

Remark 9. *The GA fusion generates more significant peaks and lighter tails than the AA fusion.*

Remark 10. *The isolated GC will survive (although its weighted will be reduced) in the AA fusion but will almost vanish in the GA fusion; this indicates that the GA fusion has better capability to suppress false alarm (if the isolated GC is a false alarm in practice) but will also suffer from misdetection (if the isolated GC turns out to be a real detection). This property is a double-edged sword.*

One more comment is in order. As we have addressed earlier in Sec. II.C, the support of the AA is the union of those of all initial functions while the support of the GA is the intersection of those of all initial functions. Therefore, assuming that both

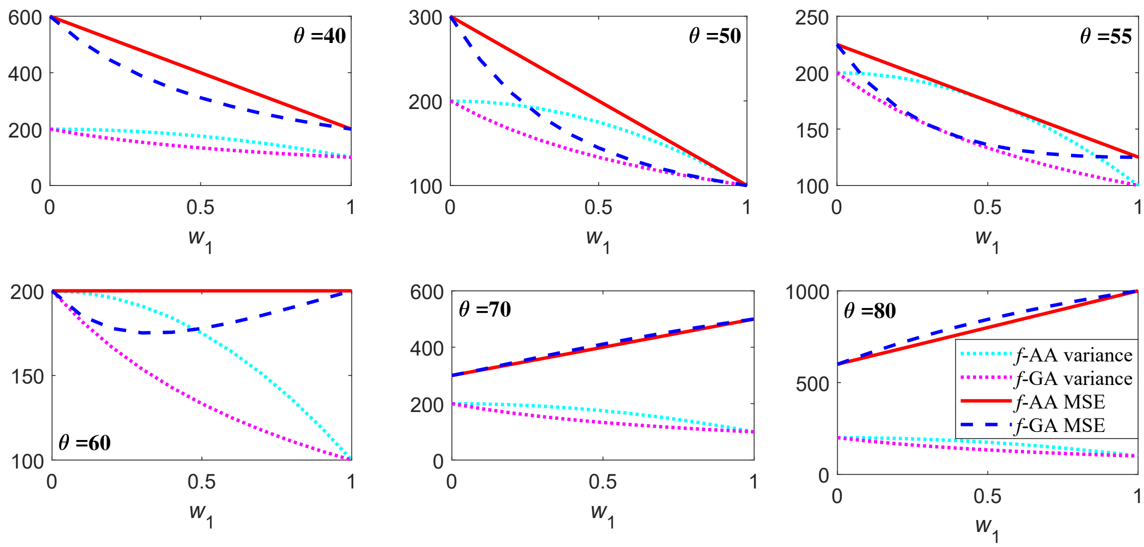

Fig. 5. Comparison of the variances and MSEs of the AA and of the GA of two Gaussian PDFs $f_{\hat{\theta}_1}(x) = \mathcal{N}(x; 50, 100)$ and $f_{\hat{\theta}_2}(x) = \mathcal{N}(x; 60, 200)$ regarding different real variables $\theta \in [40, 80]$, when different fusing weights are used.

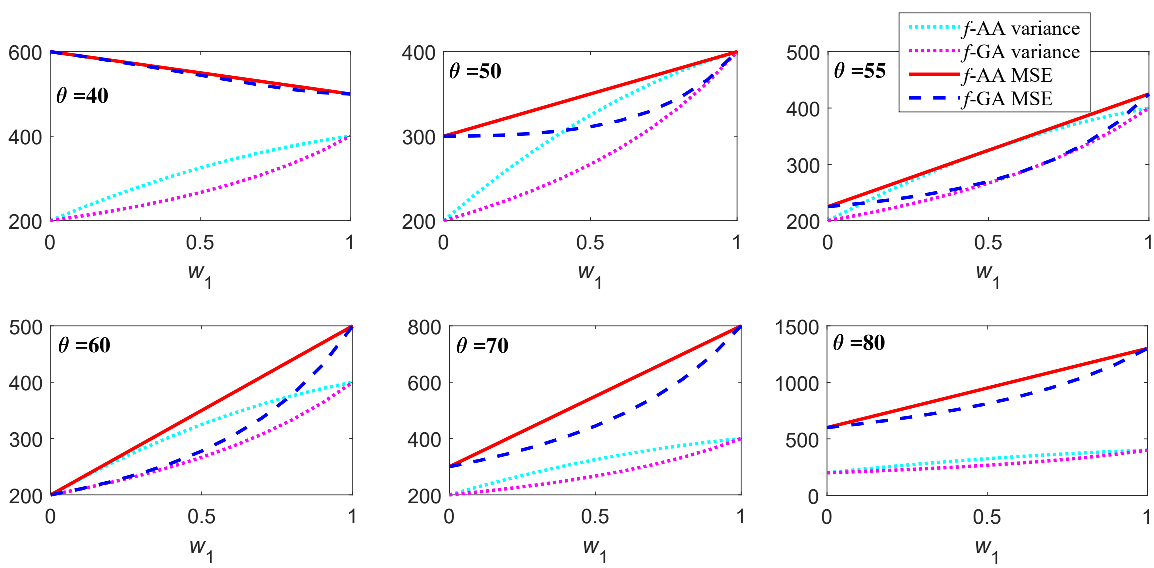

Fig. 6. Comparison of the variances and MSEs of the AA and of the GA of two Gaussian PDFs $f_{\hat{\theta}_1}(x) = \mathcal{N}(x; 50, 400)$ and $f_{\hat{\theta}_2}(x) = \mathcal{N}(x; 60, 200)$ regarding different real variables $\theta \in [40, 80]$, when different fusing weights are used.

misdetection and false alarms are independent across fusing GMs, one complete misdetection occurred in one fusing GM (namely the support of the fusing distribution does not really cover the position of the corresponding target) will "dominate" the final GA result (namely the GA must suffer from the misdetection of that corresponding target), no matter how significant the detections are in the other fusing GMs and even how many GMs there are. In fact, this problem becomes more serious when more sensors/GMs are to be fused in the GA fusion because a missed detection at any single sensor can degrade the performance of GA fusion significantly, and the probability of such a missed detection obviously becomes larger when more sensors are involved. This, however, is not a problem to the AA fusion but instead, the more GMs, the better they compensate for the misdetection occured to a single fusing GM.

To demonstrate this phenomenon, we consider an example in which six GMs are fused, as shown in Fig. 8. There are five targets in total which lie exactly at position 20, 40, 70, 110 and 200, respectively. In our simulation, each target is either detected with probability 0.9 and generates a detection at each sensor (which are given in cyan print) or misdetected with probability 0.1. In addition, at each sensor, false alarms in each GM (which are marked in magenta print) are uniformly distributed in the interval [0, 200] and the number of false alarms is a Poisson random variable with rate 1. Fig. 8 shows

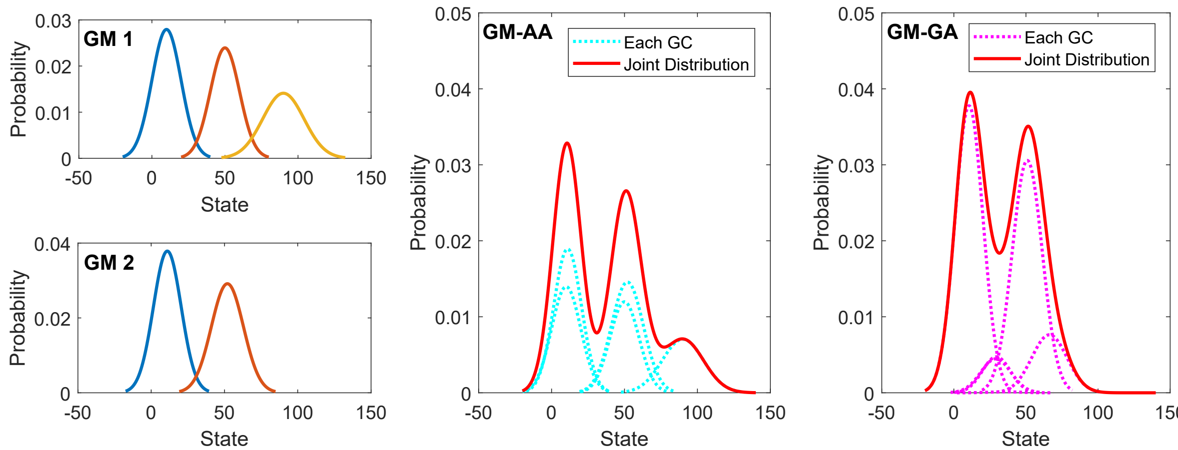

Fig. 7. Unweighted AA and GA of two GMs consisting of components of different significance/weight. The averaging results namely the AA and the GA are shown in components and in joint distribution, respectively.

the result for one trial based on the given statistics. In the result of the GA fusion, it is seen that the target lying at position 70 was mis-detected while the one lying at 200 was almost mis-detected (as only a low-weighted GC is generated). In the meanwhile, the detections of targets that lied at position 20 and 40 were mixed. These problems can be overcomed in the AA fusion which simply reserves all of the peaks (although this can be another problem). We reiterate that, a potential means to ameliorate the AA fusion is applying mixture merging and pruning to reduce the number of GCs/peaks.

Finally, we must stress that it is intractable to analytically compare between the results of the AA and of the GA for the PHD fusion in general due to the following two fundamental issues that remain open:

- *Multiple-target state estimate extraction from the fused mixture/multimodal-PHD* that contains information about a random (unknown) number of targets. Two of the most common solutions are referred to as Threshold and Rank rules [52], [53], respectively. In the former, a threshold is specified in advance and the GCs whose associated weight is greater than that threshold will all be extracted as estimates while in the latter, the number of estimates is determined firstly and the corresponding number of GCs of the highest associated weights are extracted as the estimates.
- *Estimator evaluation metric* that has to take into account the issues of misdetection, false alarms as well as the usual point-to-point estimation errors. A matric that has been widely used in the context of multi-targrt tracking is referred to the optimal sub-pattern assignment (OSPA) metric [54]. It, however, penalizes different numbers of misdetections the same. For example, the OSPA metric does not distinguish between two estimators, one of which misses the detection of one target and the other misses the detection of two (or more) targets, no matter how accurately they detect the other targets.

VI. CONCLUSIONS

We have analyzed and compared the second order statistics of the GA and the AA of a set of estimators, in terms of averaging random variables and the PDFs. The key findings that we have obtained include:

- For v-fusion,
  1) The variance of both AA and GA can be smaller than the smallest variance of the fusing variables given proper fusing weights, when the fusing variables are little or negatively correlated.
  2) For any two variables, the lowest AA variance (namely the lower bound) that can be yielded by adjusting the fusing weights is smaller than that of the GA variance.
  3) The lowest bound of the MSE of either the AA or the GA is their corresponding variance, which is only possible when the real parameter lies between two variables and proper fusing weights are used.
- For Gaussian f-fusion,
  1) The AA fusion always leads to a greater variance than the GA fusion does, for using the same fusion weights,
  2) The AA has an MSE that is the weighted average of the MSEs of the fusing estimators (and so it is bounded by the smallest and greatest MSEs),
  3) The GA fusion tends to perform better than the AA fusion in obtaining smaller MSE in most cases.
- For *PHD*-fusion based on a hybrid use of f-fusion (for localization distribution fusion) and v-fusion (for cardinality fusion),
  1) The GA fusion generates more significant peaks and lighter tails than the AA fusion does; in order words, the GA is comparably more accurate and less robust.
  2) The GA fusion has better capability to suppress false alarm but also suffers from higher risk in causing misdetection as compared to the AA fusion. A greater number of fusing sensors leads to greater problem/gain.

Appendix A: Lower Bound of $h(w)$ as in (15)

Here, we analyze the lower bound of function $h(w)$ as given in (15) for $w \in (0,1)$ and $\alpha \geq 1, \rho \in (-1,1)$. Strightforwardly, the derivative of $h(w)$ with respect to w is

$$\frac{\mathrm{d}\,h(w)}{\mathrm{d}w} = \big(2(\alpha+1-2\rho\alpha^{\frac{1}{2}})w - 2 + 2\rho\alpha^{\frac{1}{2}}\big). \tag{44}$$

Setting it to zero yields

$$w = \frac{1-\rho\alpha^{\frac{1}{2}}}{1+\alpha-2\rho\alpha^{\frac{1}{2}}}, \tag{45}$$

This, however, may not satisfy the rule that $0 < w < 1$ and if not, cannot be used. We discuss two different cases:

1) $\rho < \alpha^{-\frac{1}{2}}$: In this case, (45) satisfies $0 < w < 1$ and yields

$$h(w) = \frac{\alpha(1-\rho^2)}{1+\alpha-2\rho\alpha^{\frac{1}{2}}}. \tag{46}$$

Furthermore, by applying $\rho < \alpha^{-\frac{1}{2}}$, we obtain $\frac{\mathrm{d}^2 h(w)}{\mathrm{d}w^2} = 2(\alpha+1-2\rho\alpha^{\frac{1}{2}}) > 0$. This indicates that the bound given in (46) is indeed the lower bound. That is, if $\rho < \alpha^{-\frac{1}{2}}$, the optimal fusing weights to get the minimal $h(w)$ are given by

$$\omega_1 = \frac{\alpha-\rho\alpha^{\frac{1}{2}}}{1+\alpha-2\rho\alpha^{\frac{1}{2}}}, \;\; \omega_2 = \frac{1-\rho\alpha^{\frac{1}{2}}}{1+\alpha-2\rho\alpha^{\frac{1}{2}}}. \tag{47}$$

2) $\rho \geq \alpha^{-\frac{1}{2}}$: In this case, $w < 0$ and so, (45) can not be used. Considering that $h(w)$ is a convex function of w and $\frac{\mathrm{d}^2 h(w)}{\mathrm{d}w^2} > 0$, we obtain dual bounds of $h(w)$ at the boundaries of the support interval of the fusing weights, namely

$$1 = h(0) < h(w) < h(1) = \alpha.$$

References

[1] C.-Y. Chong and S. P. Kumar, "Sensor networks: evolution, opportunities, and challenges," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1247–1256, Aug. 2003.

[2] M. Wooldridge, *An Introduction to Multi-Agent Systems*, 2nd ed. Wiley Publishing, 2009.

[3] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[4] A. H. Sayed, "Adaptive networks," *Proc. IEEE*, vol. 102, no. 4, pp. 460–497, Apr. 2014.

[5] D. Hall and S. McMullen, *Mathematical Techniques in Multisensor Data Fusion*. London, UK: Artech House, 2004.

[6] T. Heskes, "Selecting weighting factors in logarithmic opinion pools," in *Advances in Neural Information Processing Systems*, M. Kearns, S. Solla, and D. Cohn, Eds. Cambridge, Massachusetts, United States: The MIT Press, 1998, pp. 266–272.

[7] I. Hwang, K. Roy, H. Balakrishnan, and C. Tomlin, "A distributed multiple-target identity management algorithm in sensor networks," in *Proc. CDC 2004*, Atlantis, Paradise Island, Bahamas, 2004, pp. 728–734.

[8] M. B. Hurley, "An information theoretic justification for covariance intersection and its generalization," in *Proc. 5th Int. Conf. Inf. Fusion*, Annapolis, MD, USA, Jul. 2002, pp. 505–511.

[9] F. Nielsen, "Chernoff information of exponential families," *arXiv:1102.2684*, 2011.

[10] J. K. Uhlmann, "Dynamic map building and localization: new theoretical foundations," Ph.D. dissertation, University of Oxford, Oxford, UK, 1995.

[11] S. Julier and J. Uhlmann, *General Decentralized Data Fusion With Covariance Intersection (CI)*. Boca Raton FL, USA: CRC Press, 2001, ch. 12, p. 1211225.

[12] L. Chen, P. O. Arambel, and R. K. Mehra, "Estimation under unknown correlation: covariance intersection revisited," *IEEE Trans. Auto. Control*, vol. 47, no. 11, pp. 1879–1882, Nov. 2002.

[13] Y. Wang and X. R. Li, "Distributed estimation fusion with unavailable cross-correlation," *IEEE Trans. Aero. Ele. Syst.*, vol. 48, no. 1, pp. 259–278, Jan. 2012.

[14] 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, Jul. 2012, pp. 1876–1883.

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

[16] J. K. Uhlmann, "Covariance consistency methods for fault-tolerant distributed data fusion," *Inf. Fusion*, vol. 4, no. 3, pp. 201 – 215, 2003.

[17] C. N. Taylor and A. N. Bishop, "Homogeneous functionals and Bayesian data fusion with unknown correlation," *Information Fusion*, vol. 45, pp. 179 – 189, 2019.

[18] O. Hlinka, O. Sluciak, F. Hlawatsch, P. M. Djuric, and M. Rupp, "Likelihood consensus and its application to distributed particle filtering," *IEEE Trans. Signal Proc.*, vol. 60, no. 8, pp. 4334–4349, Aug. 2012.

[19] O. Hlinka, F. Hlawatsch, and P. M. Djuric, "Distributed particle filtering in agent networks: A survey, classification, and comparison," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 61–81, Jan. 2013.

[20] G. Battistelli, L. Chisci, and C. Fantacci, "Parallel consensus on likelihoods and priors for networked nonlinear filtering," *IEEE Signal Process. Lett.*, vol. 21, no. 7, pp. 787–791, Jul. 2014.

[21] R. P. S. Mahler, "Optimal/robust distributed data fusion: a unified approach," in *Proc. SPIE*, vol. 4052, 2000, pp. 128–138.

[22] D. Clark, S. Julier, R. P. S. Mahler, and B. Ristic, "Robust multi-object sensor fusion with unknown correlations," in *Proc. SSPD 2010*, London, United Kingdom, Sep. 2010, pp. 1–5.

[23] G. Battistelli, L. Chisci, C. Fantacci, A. Farina, and A. Graziano, "Consensus CPHD filter for distributed multitarget tracking," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 3, pp. 508–520, Jun. 2013.

[24] M. Üney, D. E. Clark, and S. J. Julier, "Distributed fusion of PHD filters via exponential mixture densities," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 3, pp. 521–531, Jun. 2013.

[25] G. Battistelli and L. Chisci, "Kullback-Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability," *Automatica*, vol. 50, no. 3, pp. 707 – 718, 2014.

[26] M. Üney, J. Houssineau, E. Delande, S. J. Julier, and D. E. Clark, "Fusion of finite set distributions: Pointwise consistency and global cardinality," *arXiv:1802.06220 [eess.SP]*, 2018.

[27] L. Gao, G. Battistelli, L. Chisci, and P. Wei, "Distributed joint sensor registration and target tracking via sensor network," *Inf. Fusion*, vol. 46, pp. 218 – 230, 2019.

[28] R. P. S. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1152–1178, Oct. 2003.

[29] R. L. Streit, "Multisensor multitarget intensity filter," in *Proc. 11th Int. Conf. Inf. Fusion*, Cologne, Germany, Jun. 2008, pp. 1–8.

[30] R. Streit, "Multisensor traffic mapping filters," in *Proc. SDF 2012*, Bonn Germany, Sep. 2012, pp. 43–48.

[31] M. R. Balthasar and A. M. Zoubir, "Multi-target tracking in distributed sensor networks using particle PHD filters," *arXiv:1505.01668v4*, 2015.

[32] J. Y. Yu, M. Coates, and M. Rabbat, "Distributed multi-sensor CPHD filter using pairwise gossiping," in *Proc. ICASSP 2016*, Shanghai, China, Mar. 2016, pp. 3176–3180.

[33] T. Li, J. Corchado, and S. Sun, "Partial consensus and conservative fusion of Gaussian mixtures for distributed PHD fusion," *arXiv:1711.10783 [cs.SY]*, Nov. 2017, revision under review in IEEE Trans. Aerosp. Electron. Syst.

[34] T. Li, "Distributed SMC-PHD fusion for partial, arithmetic average consensus," *arXiv:1712.06128 [cs.SY]*, Dec. 2017.

[35] T. Li, F. Hlawatsch, and P. M. Djuric, "Cardinality-consensus-based PHD filtering for distributed multitarget tracking," *IEEE Signal Process. Lett.*, pp. 1–1, 2018, dOI:10.1109/LSP.2018.2878064.

[36] T. Li, J. M. Corchado, and H. Chen, "Distributed flooding-then-clustering: A lazy networking approach for distributed multiple target tracking," in *Proc. FUSION 2018*, Cambridge, UK, Jul. 2018, pp. 2415 – 2422.

[37] T. Li, J. Prieto, H. Fan, and J. M. Corchado, "A robust multi-sensor PHD filter based on multi-sensor measurement clustering," *IEEE Comm. Lett.*, vol. 22, no. 10, pp. 2064 – 2067, 2018.

[38] R. P. S. Mahler, "The multisensor PHD filter: II. Erroneous solution via Poisson magic," in *Proc. SPIE*, vol. 7336, 2009, pp. 7336 – 12.

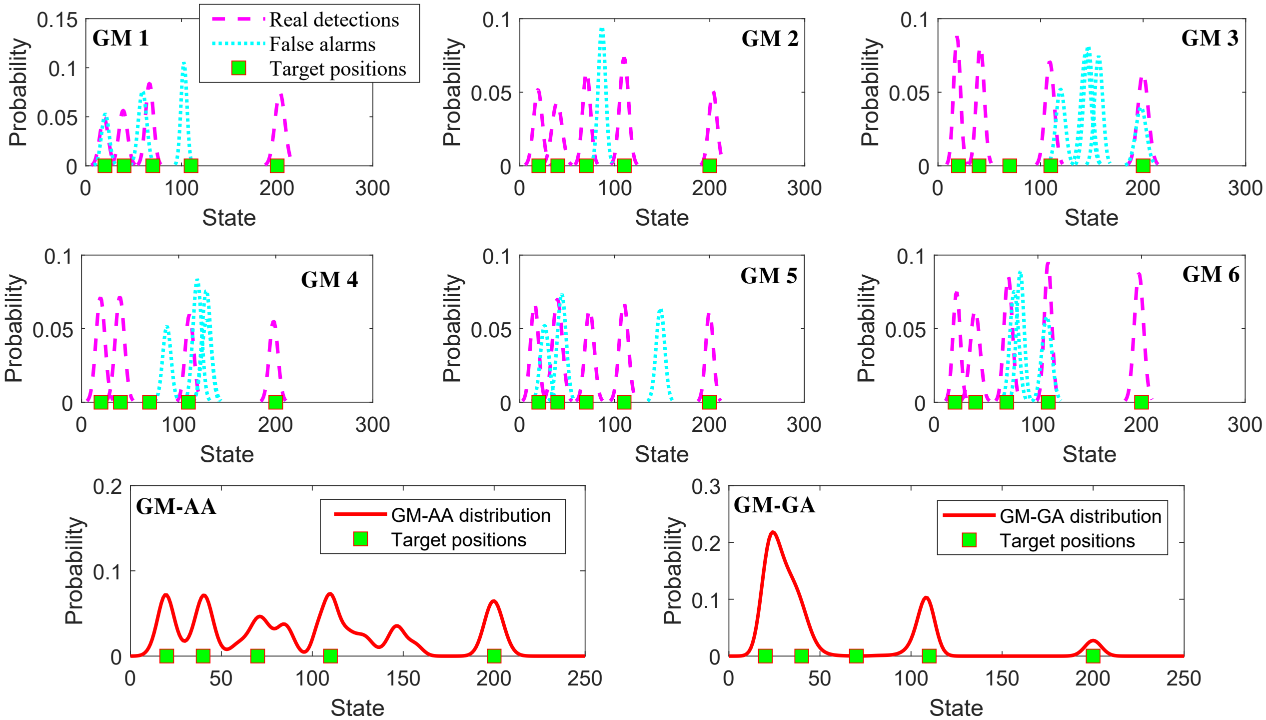

Fig. 8. Unweighted AA and GA of six GMs consisting of both real detections and false alarms, both of which are given by weighted Gaussian distributions: the weights indicate the significance of the detections. The number of false alarms at each GM is Poisson distributed with rate 1 and the position of the false alarm is uniformly distributed in the interval between 0 and 200. There are also potential misdetections in each GM.

[39] S. Mori, K. C. Chang, and C. Y. Chong, "Comparison of track fusion rules and track association metrics," in *Proc. 15th Int. Conf. Inf. Fusion*, Singapore, Jul. 2012, pp. 1996–2003.

[40] W. Yi, M. Jiang, S. Li, and B. Wang, "Distributed sensor fusion for RFS density with consideration of limited sensing ability," in *Proc. FUSION 2017*, Xi'an, China, Jul. 2017, pp. 1–6.

[41] 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 Trans. Signal Proc.*, vol. 65, no. 1, pp. 242–255, Jan. 2017.

[42] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[43] S. Ross, *Introduction to Probability Models (11th Edition)*. Boston, USA: Academic Press, 2014.

[44] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences (Revised Edition)*. Cambridge, Massachusetts, USA: Academic Press, 1977, pp. 109 - 143.

[45] C. Cuadras, "On the covariance between functions," *Journal of Multivariate Analysis*, vol. 81, no. 1, pp. 19 – 27, 2002.

[46] S. S. Singh, B.-N. Vo, A. Baddeley, and S. Zuyev, "Filters for spatial point processes," *SIAM J. Control Optim.*, vol. 48, no. 4, pp. 2275–2295, 2009.

[47] N. T. N. Mariam, "Conservative non-Gaussian data fusion for decentralized networks," Master's thesis, The University of Sydney, Sydney, Australia, Aug. 2007.

[48] M. Gunay, U. Orguner, and M. Demirekler, "Chernoff fusion of Gaussian mixtures based on sigma-point approximation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 6, pp. 2732–2746, Dec. 2016.

[49] J. Li and A. Nehorai, "Distributed particle filtering via optimal fusion of Gaussian mixtures," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 4, no. 2, pp. 280–292, 2018.

[50] T. Li, J. Corchado, and S. Sun, "On generalized covariance intersection for distributed PHD filtering and a simple but better alternative," in *Proc. FUSION 2017*, Xi'an, China, Jul. 2017, pp. 808–815.

[51] G. Battistelli, L. Chisci, G. Mugnai, A. Farina, and A. Graziano, "Consensus-based linear and nonlinear filtering," *IEEE Trans. Autom. control*, vol. 60, no. 5, pp. 1410–1415, May 2015.

[52] T. Li, J. M. Corchado, S. Sun, and H. Fan, "Multi-EAP: Extended EAP for multi-estimate extraction for SMC-PHD filter," *Chin. J. Aeronaut.*, vol. 30, no. 1, pp. 368 – 379, 2017.

[53] B. N. Vo and W. K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4091–4104, Nov. 2006.

[54] D. Schuhmacher, B. T. Vo, and B. N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3447–3457, Aug. 2008.