Fusion of finite set distributions: Pointwise consistency and global cardinality

Murat Üney, *Member, IEEE*, Jérémie Houssineau, Emmanuel Delande, Simon J. Julier, Daniel Clark

***Abstract*—A recent trend in distributed multi-sensor fusion is to use random finite set filters at the sensor nodes and fuse the filtered distributions algorithmically using their exponential mixture densities (EMDs). Fusion algorithms that extend covariance intersection and consensus based approaches are such examples. In this article, we analyse the variational principle underlying EMDs and show that the EMDs of finite set distributions do not necessarily lead to consistent fusion of cardinality distributions. Indeed, we demonstrate that these inconsistencies may occur with overwhelming probability in practice, through examples with Bernoulli, Poisson and independent identically distributed (IID) cluster processes. We prove that pointwise consistency of EMDs does not imply consistency in global cardinality and vice versa. Then, we redefine the variational problems underlying fusion and provide iterative solutions thereby establishing a framework that guarantees cardinality consistent fusion.**

***Index Terms*—random finite sets, multi-sensor fusion, exponential mixture density, covariance intersection, target tracking**

I. Introduction

In networked sensing, nodes perform local filtering and exchange filtered distributions as opposed to communicating raw measurements [1]. The problem of fusion is to find an estimate for the *a posteriori* distribution over some state space conditioned on two or more (conditionally) independent sensor data streams, given local posteriors computed by local filtering of each data stream individually.

A large body of work utilises exponential mixtures of distributions (EMDs) for fusion. These mixtures are found by taking the weighted geometric mean of their components followed by scaling to ensure integration to unity. They have been widely used for fusion of single object (probability) distributions [2]. A well-known algorithm that utilises EMDs of Gaussian densities is covariance intersection [3]. In covariance intersection (CI), the weights of the components in the mixture are selected using various criteria [4]. The underlying variational problem considers minimising a cost that equals to the weighted sum of Kullback-Leibler divergences [5] of the fused density that is sought with respect to the mixture components. The stationary density and set of weights for this problem specifies an EMD which is deemed as a middle-ground of the components in a way analogous to logarithmic opinion pooling of experts [6].

The EMD form has been adopted for finite set densities in order to address fusion in the case of multiple objects [7]. Following the introduction of tractable recursive filters [8] such as the probability hypothesis density (PHD) filter [9], and, explicit filtering algorithms using Gaussian mixture model (GMM) representations [10] and sequential Monte Carlo (SMC) techniques [11], numerical algorithms that extend CI fusion to Bernoulli, PHD, and cardinalised PHD (C-PHD) were proposed [12]–[14]. These methods have been proved useful in improving localisation accuracy in multi-sensor problems including those involving heterogenous sensors [15].

Another utilisation of EMDs for fusion of finite set distributions has been within the network consensus framework [16]. Briefly, iterative message passing algorithms which asymptotically compute the equally-weighted mixture, i.e., the (unweighted) geometric mean of the components, at all nodes of a sensor network are proposed for C-PHD [17], multi-Bernoulli [18], generalised MB [19], [20], Bernoulli [21], and, labelled [22], [23] finite set filters.

In [24], [25], it has been proved that EMDs have a probability density that at no point in the state space overlooks the density of their components. This property is proposed as a working definition of consistency in the context of fusion [25]. Finite set density EMDs also satisfy this consistency condition pointwise, at every finite collection of points.

Finite set distributions, on the other hand, factorise into a cardinality distribution on the number of objects and a localisation density conditioned on the cardinality [26]. In this article, we show that the cardinality distributions of EMDs are not endowed with such consistency guarantees, in general. Such inconsistencies might result with smaller existence probabilities or estimates on the number of objects when the fused results are used instead of either of the inputs. This phenomena which might undermine the benefits of using diversity in sensing has been empirically observed by other researchers as well (see, e.g., [27]). Here, we provide explicit mathematical formulae specifying conditions under which the cardinality distributions of finite set EMDs are inconsistent. We demonstrate in examples that these inconsistencies are encountered sometimes with overwhelming probability under typical operating conditions and might lead to large discrep-

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) Grant number EP/K014277/1 and the MOD University Defence Research Collaboration (UDRC) in Signal Processing.

Murat Üney is with the Institute for Digital Communications, School of Engineering, University of Edinburgh, EH9 3FB, Edinburgh, UK (e-mail: m.uney@ed.ac.uk).

Jérémie Houssineau is with the National University of Singapore, Department of Statistics and Applied Probability, National University of Singapore, Singapore 119077.

Emmanuel Delande is with the Institute for Computational Engineering and Sciences, University of Texas at Austin (edelande@ices.utexas.edu).

Simon J. Julier is with the Computer Science Department, University College London, London (e-mail: s.julier@ucl.ac.uk).

Daniel E. Clark is with Département CITI, Telecom-SudParis, 9, rue Charles Fourier 91011, EVRY Cedex, France (e-mail: daniel.clark@telecom-sudparis.eu).

ancies in, for example, the estimated number of objects and/or object existence probabilities.

Based on these results, we argue that the variational problem needs to be decoupled for the cardinality and the localisation distributions (i.e., scaled Janossy [26] distributions). Doing so separates the fusion of cardinality distributions and localisation terms. This approach results with the same localisation densities as the direct adoption of the variational problem, and, avoids any inconsistencies in the cardinality distribution. We show that pointwise consistency does not imply consistency in cardinality and vice versa. Then, we derive iterative algorithms for cardinality consistent fusion of finite set distributions.

The outline of the article is as follows: In Section II, we discuss fusion rules that accommodate EMDs in the light of the associated variational problems and pointwise consistency of EMDs. We provide our results regarding the cardinality inconsistencies of finite set EMDs in Section III, together with examples. Then, we redefine the variational problem underpinning fusion and derive solutions for cardinality consistent fusion in Section IV. Conclusions and future directions are provided in Section V.

II. Fusion as a variational problem

A. EMDs as weighted KLD centroids

Given two probability density functions (PDFs) f_i and f_j on a state space $\mathcal{X}$, let us consider finding another density f in the space of PDFs $\mathcal{P}$ such that f captures the information contained in both of the input distributions. An intuitive approach which is geometric in flavour would involve finding the centroid of the input distributions based on a distance/divergence metric. Kullback-Leibler divergence (KLD) is such a divergence metric which is used in information geometry in a way similar to the squared Euclidean distance [28], and, has an established relevance to estimation when $\mathcal{X}$ is a finite alphabet (which is often referred to as hypothesis testing) [5].

The KLD of two distributions with densities f and g is computed as

$$D(f||g) = \int_{\mathcal{X}} f(X) \log \frac{f(X)}{g(X)} \mathrm{d}X, \tag{1}$$

where D is always nonnegative and vanishes for $f = g$.

Let us denote the centroid of f_i and f_j with respect to a weighted sum of KLD by f_ω. This distribution is a solution to the associated variational problem given by

$$\text{(P)} \quad \min_{f\in\mathcal{P}} J_\omega[f]$$
$$J_\omega[f] \triangleq (1-\omega)D(f||f_i) + \omega D(f||f_j) \tag{2}$$

where $\omega \in [0,1]$ is a design parameter selecting the weight of the divergence of each point f_i and f_j in the space of probability distributions $\mathcal{P}$ over $\mathcal{X}$, with respect to f.

The solution to problem (P) with the cost (2) is unique and found as

$$f_\omega(X) = \frac{1}{Z_\omega} f_i^{(1-\omega)}(X) f_j^{\omega}(X) \tag{3}$$

$$Z_\omega = \int_{\mathcal{X}} f_i^{(1-\omega)}(X') f_j^{\omega}(X') \ \mathrm{d}X', \tag{4}$$

which can easily be seen after rearranging the cost in (2) as

$$J_\omega[f] = D(f||f_\omega) - \log \int_{\mathcal{X}} f_i^{(1-\omega)}(X) f_j^{\omega}(X) \ \mathrm{d}X, \tag{5}$$

(see, for example, [29, Eq.(3)]), and, realising that the second term on the right hand side does not depend on f (see Appendix A for a direct proof). In fact, this term is the scaled Rényi divergence [30] of order ω from f_j to f_i, i.e.,

$$J_\omega[f] = D(f||f_\omega) - (\omega-1)R_\omega(f_j, f_i), \tag{6}$$
$$R_\omega(f_j, f_i) \triangleq \frac{1}{\omega-1} \log \int_{\mathcal{X}} f_j^{\omega}(X) f_i^{(1-\omega)}(X) \mathrm{d}X,$$
$$= \frac{1}{\omega-1} \log Z_\omega.$$

Let us consider the weight parameter ω as a free variable, and find the stationary point of J_ω in (2) with respect to ω for $f = f_\omega$. For the case, the KLD term in (6) vanishes and (2) reduces to a cost function for finding the Chernoff information of f_i and f_j [5] which is concave in ω [1]. In [32], it is explained that there is a unique stationary point ω^* which satisfies

$$D(f_\omega||f_i) = D(f_\omega||f_j)\Big|_{\omega=\omega^*}. \tag{7}$$

The density f_ω in (4) is obtained by normalising the weighted geometric mean of f_i and f_j, and, thus referred to as their geometric mean density (GMD), or, exponential mixture density (EMD). In this article we adopt the latter.

B. Covariance intersection and generalisations

The above discussion outlines a fusion algorithm which outputs the pair (ω^*, f_{ω^*}) using (7) and (3) for fusing f_i and f_j. This can be rephrased as a $\max\min$ mathematical programme:

$$\text{(P2)} \quad (\omega^*, f_{\omega^*}) \triangleq \arg \max_{\omega\in[0,1]} \min_{f\in\mathcal{P}} J_\omega[f]. \tag{8}$$

The input densities here are *a posteriori* in nature as they are propagated by local filters, i.e., they are conditioned on the data-streams of sensors i and j, respectively. When $\mathcal{X}$ is $\mathbb{R}^d$, i.e., the d-dimensional space of real vectors, and, the distributions involved are Gaussians, this approach reduces to a set of linear algebraic operations which are known as the "covariance intersection" algorithm [3]. In this setting, how well an approximation the EMD (3) is to the joint posterior[2] is studied in terms of bounds over the uncertainty spread characterised by covariance matrices (see, e.g., [33], [34]). An information geometric characterisation of f_{ω^*} for multivariate Gaussians and other exponential family distributions is provided in [35] where it is proved that f_{ω^*} is the unique intersection point of the exponential geodesic curve joining f_i and f_j (obtained by varying ω from 0 to 1 in (3)) and its dual hyperplane on the induced statistical manifold.

[1] To be specific, in [31], Chernoff introduces $C(f_j, f_i) \triangleq -\log\min Z_\omega$ as a "measure of divergence" between two distributions. This quantity can equivalently be found by $\max -\log Z_\omega$ in which the argument of maximisation is nothing but $J_\omega[f = f_\omega]$.

[2] Here, we refer to the posterior distribution conditioned on the data streams of both sensors which is infeasible to compute given the limited communication and computational resources of the networked setting.

For general distributions, (3), (4) and (7) are still valid as a solution to the variational fusion problem in (8). The optimal weight selected through (7) equates the cost in (2) to the Chernoff information [5] between f_i and f_j [32]. Perhaps for this reason, some authors refer to this fusion rule as Chernoff fusion (see, for example [36] and the references therein).

C. Other fusion rules utilising EMDs

Other fusion methods that use EMDs include consensus based approaches as overviewed in Section I. These methods compute f_ω by iterative message passings between nodes. However, instead of finding stationary weights of the variational problem in (2), this network averaging approach can compute only an equally weighted EMD, and, when the number of iterations tends to infinity. Some other methods differ from the generalised CI approach described above in their weight selection criteria: Some authors argue that it might be more beneficial to select the value of ω in (2) that would maximise the "peakiness" of f_ω [7], or, to minimise the uncertainty captured by f_ω quantified by its Shannon differential entropy [4].

D. A notion of consistency in Fusion

The uncertainty spread in EMDs of arbitrary distributions is characterised in terms of pointwise bounds. In [25], the authors show that the scale factor in (4) is less than or equal to one, i.e., $Z_\omega \leqslant 1$, and, consequently EMDs (3) satisfy the following consistency condition:

$$f_\omega(X) \geqslant \min\{f_i(X), f_j(X)\} \tag{9}$$

for all points $X \in \mathcal{X}$ and $\omega \in [0,1]$. In other words, the fused distribution does not overlook the probability mass assigned by f_i and f_j onto the vicinity of any point in the state space. In this sense, this condition corresponds to a notion of consistency [25], in the context of distributed fusion[3].

In this article, our concern is the consistency properties of EMDs of finite set distributions. These distributions have been commonly used to represent multi-object scenes [8]. The following discussion is valid for any fusion scheme that employs EMDs and random finite set (RFS) distributions in order to quantify uncertainty in, for example, "the number of objects," (e.g., Poisson, i.i.d. cluster RFSs [8]), "existence probabilities" (e.g., Bernoulli, multi-Bernoulli, generalised labelled MB RFSs [37] and MB mixtures [38]) irrespective of their weight selection mechanism. In the next section, we utilise (9) for analysing finite set EMDs and examine the fused global cardinality distributions for inconsistencies and their consequences in estimating object existence probabilities and/or the number of objects.

III. Finite set EMDs and cardinality distributions

In the case of finite set valued random variables, $\mathcal{X}$ is the space of finite subsets of $\mathbb{R}^d$ and the density f is a set function characterised by i) a cardinality distribution with probability mass function (pmf) $p(n)$ over natural numbers $n = 0, 1, \ldots$, and, ii) localisation densities $\rho_n(x_1, ..., x_n)$ for $n = 1, 2, \ldots$ which are symmetric in their arguments [26]. The corresponding density has a set valued argument $X = \{x_1, \ldots, x_n\}$ and is given by

$$\begin{aligned} f(X) &= p\,(n) \sum_{\sigma \in \Sigma_n} \rho_n(x_{\sigma(1)}, ..., x_{\sigma(n)}) \\ &= p\,(n)\, n! \rho_n(x_{\sigma'(1)}, ..., x_{\sigma'(n)}) \end{aligned} \tag{10}$$

where $n = |X|$ and $|.|$ denotes set cardinality. Here, Σ_n is the set of all permutations of $(1, \ldots, n)$, and, $\sigma' \in \Sigma_n$ in the last line is an arbitrary permutation which is selected as the identity permutation in the rest of this article.

Note that p in (10) sums to one and ρ_ns integrate to unity. The finite set density f also integrates to one over $\mathcal{X}$, i.e.,

$$\int_{\mathcal{X}} f(X) \mu(\mathrm{d}X) = 1$$

where μ is an appropriate measure. Let us select μ as

$$\mu(\mathrm{d}X) = \sum_{n=0}^{\infty} \frac{\lambda_n(\mathrm{d}X \cap \mathcal{X}_n)}{n!}$$

where $\mathcal{X}_n$ is the space of n-tuple of points in $\mathbb{R}^d$, and, λ_n is the Lebesgue (volume) measure on $\mathcal{X}_n$ [4]. An alternative form of this integral is referred to as the set integral [8], i.e.,

$$\int_{\mathcal{X}} f(X) \mu(\mathrm{d}X) = \int_{\mathbb{R}^d} f(X) \delta X,$$

where the right hand side is the set integral of f defined as [5]

$$\begin{aligned} \int_{\mathbb{R}^d} f(X) \delta X &\triangleq \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} f(\{x_1, ..., x_n\}) \mathrm{d}x_1 \ldots \mathrm{d}x_n \\ &= \sum_{n=0}^{\infty} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} p(n) \rho_n(x_1, ..., x_n) \mathrm{d}x_1 \ldots \mathrm{d}x_n. \end{aligned} \tag{11}$$

Let us consider the EMD of finite set distributions f_i and f_j. For the case (3) is valid with the scale factor in (4) found using the set integral in (11), i.e.,

$$Z_\omega = \int_{\mathbb{R}^d} f_i^{(1-\omega)}(X') f_j^{\omega}(X') \delta X'. \tag{12}$$

This scale factor is also less than one and consequently the finite set EMD satisfies the pointwise consistency condition in (9) for every finite subset $X \subset \mathbb{R}^d$.

In order to investigate the cardinality distribution of the EMD, let us substitute f_i and f_j in the form given in (10) into (12) and (3), and, obtain the finite set EMD as

$$f_\omega(X) = p_\omega(n) n! \rho_{\omega,n}(x_1, \ldots, x_n)$$

where the localisation density for cardinality n is

[3]Note that the use of the term "consistency" here differs from its use in classical statistics.

[4]Further details on the topic can be found in Section II.B and Appendix B in [11], and, the references therein.

[5]Note that the set integral in (11) is defined for an arbitrary (measurable) function f, but, when f is a finite point process density, (11) is nothing but the total probability theorem applied on (10) [39].

$$\rho_{\omega,n}(x_1,\ldots,x_n) \triangleq \frac{1}{z_\omega(n)}\rho_{i,n}^{(1-\omega)}(x_1,\ldots,x_n) \times \rho_{j,n}^{\omega}(x_1,\ldots,x_n), \quad (13)$$

$$z_\omega(n) = \int_{\mathbb{R}^d}\cdots\int_{\mathbb{R}^d} \rho_{i,n}^{(1-\omega)}(x_1',\ldots,x_n') \times\rho_{j,n}^{\omega}(x_1',\ldots,x_n')\mathrm{d}x_1',\ldots,\mathrm{d}x_n', \quad (14)$$

and, the cardinality pmf is

$$p_\omega(n) = \frac{1}{N_\omega}p_i^{(1-\omega)}(n)p_j^{\omega}(n)z_\omega(n) \quad (15)$$

$$N_\omega = \sum_{n'=0} p_i^{(1-\omega)}(n')p_j^{\omega}(n')z_\omega(n'). \quad (16)$$

where $z_\omega(0) = 1$ by convention, and for $n \neq 0$, $z_\omega(n) < 1$ unless $\rho_{i,n}$ and $\rho_{j,n}$ are identical. The latter is a direct application of Hölder's inequality (see, e.g., Theorem 188 in [40]). It can be shown similarly that $N_\omega < 1$.

Let us focus on the fused cardinality pmf in (15). This distribution is *not an EMD* of the cardinality distributions of the components unlike the fused localisation distributions in (13) that are EMDs of the input localisation densities. In fact, the fused cardinality pmf is the scaled product of the cardinality EMD with the localisation density scale factors $z_\omega(n)$ in (14). As a result, the consistency property of EMDs does not apply to the fused cardinality distribution. Below, we first relate the consistency of the fused cardinality pmf to the sequence of scale factors and give a condition under which the fused cardinality distribution is inconsistent. Then, in the rest of this section, we demonstrate that inconsistent cardinality distributions occur under some typical operating conditions.

Proposition 3.1 (Inconsistency in cardinality distribution): Consider the fused cardinality pmf p_ω in (15), (16). Consider the following *inconsistency* condition for $p_\omega(n)$ obtained by negating the consistency condition:

$$p_\omega(n) < \min\{\ p_i(n),\, p_j(n)\ \}. \quad (17)$$

This condition holds true if

$$z_\omega(n) < \frac{\sum_{n'\neq n} p_i^{(1-\omega)}(n')p_j^{\omega}(n')z_\omega(n')}{\frac{p_i^{(1-\omega)}(n)p_j^{\omega}(n)}{\min\{\ p_i(n),p_j(n)\ \}} - p_i^{(1-\omega)}(n)p_j^{\omega}(n)}, \quad (18)$$

where $z_\omega(n)$ is given in (14).

The proof is given in Appendix B. The above proposition points out that the fused cardinality distribution opts to disagree with local results on the probability of number of objects when the nth localisation scale $z_\omega(n)$ is comparably small. This is in stark contrast with the fused localisation densities in (13) which always satisfy the consistency condition

$$\rho_{\omega,n}(x_1,\ldots,x_n) \geqslant \min\{\rho_{i,n}(x_1,\ldots,x_n),\rho_{j,n}(x_1,\ldots,x_n)\}$$

for all $x_1,\ldots,x_n \in \mathbb{R}^d$ and n, as they are EMDs.

The scale factors modulating the cardinality pmf, i.e., $z_\omega(n)$, are found by taking the inner products of the input localisation densities raised to fractional powers. As explained above, these terms are upper bounded by one with $z_\omega(n)$ equaling unity only when $\rho_{i,n}$ and $\rho_{j,n}$ are equal (see [25] for an alternative proof). In fusion networks, however, one of the main goals is to benefit from sensing diversity which means $\rho_{i,n}$ and $\rho_{j,n}$ will have a comparably small overlap in their confidence regions. As a result, much smaller $z_\omega(n)$ values should be expected in typical operating conditions.

Now, let us consider some particular RFS families and demonstrate the consequences of Proposition 3.1.

A. Bernoulli finite set EMDs and fused existence probabilities

Bernoulli finite set distributions select at most one object from a population. Collections, and mixtures thereof are used to represent multi-object models the fusion of which reduces to EMD fusion of Bernoulli pairs (see, e.g., [19], [20]). For a Bernoulli finite set, the cardinality pmf in (10) is given by

$$p(n) = \begin{cases} 1-\alpha, & n=0, \\ \alpha, & n=1, \\ 0, & \text{otherwise} \end{cases} \quad (19)$$

where the parameter α is referred to as *the existence probability* of the object modelled.

There is also a single localisation density ρ_n for $n = 1$ which we will denote by ρ. Therefore, given two Bernoullis $f_i = (\alpha_i, \rho_i)$ and $f_j = (\alpha_j, \rho_j)$, the sequence $z_\omega(n)$ reduces to

$$z_\omega(n) = \begin{cases} 1, & n=0, \\ z_\omega \triangleq \int_{\mathbb{R}^d} \rho_i^{(1-\omega)}(x)\rho_j^{\omega}(x)\mathrm{d}x, & n=1, \\ 0, & \text{otherwise.} \end{cases} \quad (20)$$

Corollary 3.2: The inconsistency condition given by Proposition 3.1 for Bernoulli finite set distributions reduces to that the existence probability of the EMD given by [14]

$$\alpha_\omega = \frac{\alpha_i^{(1-\omega)}\alpha_j^{\omega}z_\omega}{(1-\alpha_i)^{(1-\omega)}(1-\alpha_j)^{\omega} + \alpha_i^{(1-\omega)}\alpha_j^{\omega}z_\omega} \quad (21)$$

is smaller than either of α_i or α_j if

$$z_\omega < \frac{(1-\alpha_i)^{(1-\omega)}(1-\alpha_j)^{\omega}}{\alpha_i^{(1-\omega)}\alpha_j^{\omega}/\min\{\ \alpha_i\ ,\ \alpha_j\ \} - \alpha_i^{(1-\omega)}\alpha_j^{\omega}}.$$

The proof follows from substituting the sequence (20) in Proposition 3.1, and, in particular in (17) and (18). This condition is very often satisfied in sensing applications as explained before. For example, if α_i and α_j are equal, then this condition reduces to $z_\omega < 1$ which always holds for all practical purposes as ρ_i and ρ_j should not be expected to be identical. For $\alpha_i \neq \alpha_j$, this inconsistency still occurs with overwhelming probability in Bernoulli fusion which is demonstrated in the following example.

Let us consider Bernoulli distributions with Gaussian localisation densities given by

$$\rho_i(x) = \mathcal{N}(x;\mathbf{m}_i,\mathbf{C}_i), \quad \rho_j(x) = \mathcal{N}(x;\mathbf{m}_j,\mathbf{C}_j), \quad (22)$$

where $\mathbf{m}$ is the mean vector and $\mathbf{C}$ is the covariance matrix. The fused localisation density ρ_ω for the case is a Gaussian

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2019.2893083, IEEE Transactions on Aerospace and Electronic Systems

Fig. 1. Localisation densities of the Gauss-Bernoulli finite sets, i.e., ρ_i (solid line) and ρ_j (dash-dotted line), in Example 3.3 for increasing sensing diversity as the covariance condition number is increased as $\kappa = 1, 10, 20, 30, 40$ (left to right).

with mean and covariance given by

$$\mathbf{m}_\omega = \mathbf{C}_\omega\big((1-\omega)\mathbf{C}_i^{-1}\mathbf{m}_i + \omega\mathbf{C}_j^{-1}\mathbf{m}_j\big) \tag{23}$$

$$\mathbf{C}_\omega = \big((1-\omega)\mathbf{C}_i^{-1} + \omega\mathbf{C}_j^{-1}\big)^{-1}. \tag{24}$$

The scale factor z_ω is found using integration rules for Gaussians as

$$z_\omega = \frac{\left|\mathbf{C}_i^{-1}\right|^{(1-\omega)/2}\left|\mathbf{C}_j^{-1}\right|^{\omega/2}}{\left|\mathbf{C}_\omega\right|^{1/2}} \exp\left\{-\frac{1}{2}\Big((1-\omega)\mathbf{m}_i^T\mathbf{C}_i^{-1}\mathbf{m}_i + \omega\mathbf{m}_j^T\mathbf{C}_j^{-1}\mathbf{m}_j - \mathbf{m}_\omega^T\mathbf{C}_\omega^{-1}\mathbf{m}_\omega\Big)\right\}. \tag{25}$$

Let us consider two Bernoullis with existence probabilities $\alpha_i = \alpha_j = 0.8$ with localisation densities of mean vectors $\mathbf{m}_i = [0.25, 0.25]^T$ and $\mathbf{m}_j = [-0.75, -0.25]^T$, respectively, where $(.)^T$ denotes vector transpose. We select the covariance matrices as rotated versions of a diagonal covariance given by

$$\mathbf{C}_i = \mathbf{R}(\pi/4)\mathbf{\Sigma}\mathbf{R}^T(\pi/4),$$
$$\mathbf{C}_j = \mathbf{R}(-\pi/4)\mathbf{\Sigma}\mathbf{R}^T(-\pi/4),$$
$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2, & 0 \\ 0, & \sigma_2^2 \end{bmatrix},$$
$$\mathbf{R}(\phi) = \begin{bmatrix} \cos\phi, & -\sin\phi \\ \sin\phi, & \cos\phi \end{bmatrix}.$$

This covariance structure is typical with sensors placed at different positions and taking their measurements from different aspect angles of the surveillance zone. The condition number of Σ – equivalently, $\mathbf{C}_i$ and $\mathbf{C}_j$ – is given by $\kappa = \sigma_1^2/\sigma_2^2$ and has higher values for sensors with range/cross-range ambiguity such as cameras/radars. We vary this quantity from $\kappa = 1$ to 40. Fig. 1 depicts the uncertainty ellipses of sample Gaussians by using three times the standard deviation along the eigen vector directions.

The behaviours of the fused existence probability in (21) and the scale factor in (25) are our concern. Fig. 2 presents both the z_ω and α_ω values obtained by varying the condition number κ with small steps from 1 to 40 hence increasing the sensing diversity. The exponential mixture weights ω take values from a dense grid over $[0, 1]$. As pointed out in this section, the scale factor values are always smaller than unity, and, can often take very small values. The scale factor monotonically decreases with κ which controls the sensing diversity. It is convex with respect to the mixture weight ω, as pointed out in Section II.

Example 3.3 (Gauss-Bernoulli EMDs): The fused existence probabilities given in Fig. 2 demonstrate the inconsistency in cardinality. In this example, this quantity is always smaller

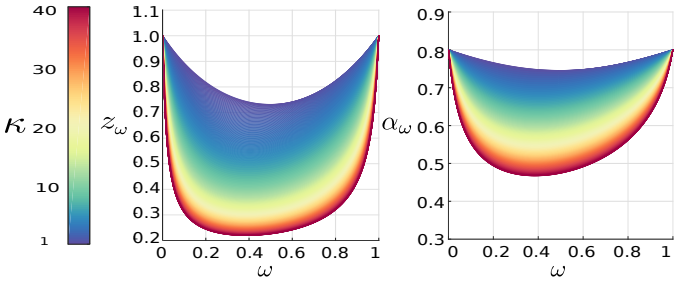


Fig. 2. The scale factor (left) and the fused existence probabilities in the Gauss-Bernoulli Example 3.3. These quantities are calculated for varying sensing diversity κ (equivalently, the covariance condition number of the Gaussians) and (exponential) mixture weight ω values.

than the input existence probabilities admitting inconsistency for all selections of κ and ω. Moreover, the fused existence probability drops below 0.5 for large values of the sensing diversity parameter κ. This threshold is often used as the Bayesian decision boundary for detection and despite that the input sources are fairly confident on the existence of an object with existence probabilities of $\alpha_i = \alpha_j = 0.8$, detection might be missed if based on the fused result instead, thereby undermining the benefits of sensing diversity. As a result, the inconsistency in cardinality may lead to inconsistency in decision making when EMDs of finite set distributions are used. ■

B. EMDs of Poisson finite set distributions

Poisson finite set densities are capable of representing many objects and underpin popular multi-object filters such as the PHD filter [9]. Their cardinality pmf in (10) is given by a Poisson distribution, i.e.,

$$p(n) = \frac{\mathrm{e}^\lambda \lambda^n}{n!} \tag{26}$$

where λ is the expected number of objects. The localisation densities factorise over the density for $n = 1$ as

$$\rho_n(x_1, \ldots, x_n) = \prod_{i=1}^{n} \rho_1(x_i), \tag{27}$$

making it possible to parameterise the entire finite set distribution with a scalar and a single density[6].

For two Poissons $f_i = (\lambda_i, \rho_i)$ and $f_j = (\lambda_j, \rho_j)$, the sequence $z_\omega(n)$ is a geometric sequence found by subsituting from (27) for both i and j into (14). This sequence is found as

$$z_\omega(n) = z_\omega^n \tag{28}$$

$$z_\omega \triangleq z_\omega(1) = \int_{\mathbb{R}^d} \rho_i^{(1-\omega)}(x')\rho_j^\omega(x')\mathrm{d}x, \tag{29}$$

where $z_\omega < 1$ unless ρ_i and ρ_j are identical, as aforementioned.

The expected number of objects with respect to an EMD with weight parameter ω is given by [14]

$$\lambda_\omega = \lambda_i^{(1-\omega)}\lambda_j^\omega z_\omega. \tag{30}$$

[6]We drop the subscript in ρ_1 for the rest of this subsection and denote it by ρ.

Proposition 3.4 (Poisson inconsistency in expectation): Let us consider an inconsistency condition for Poisson cardinality distributions in terms of their expectations:

$$\lambda_\omega < \min\{\lambda_i, \lambda_j\}. \tag{31}$$

This condition holds whenever

$$z_\omega < \frac{\min\{\lambda_i, \lambda_j\}}{\max\{\lambda_i, \lambda_j\}}. \tag{32}$$

The proof follows easily from substituting (32) in (30) and (31). It is instructive to contrast this result with Proposition 3.1. The latter holds for any class of finite set densities and considers their cardinality distributions for different n. The above result is on *the expected value* of n in Poisson finite set densities. The condition in (32) is satisfied with overwhelming probability in practice leading to inconsistencies as observed, for example, in [27]. For example, for $\lambda_i = \lambda_j = \lambda$, this reduces to the common ratio z_ω being less than one which should –as previously discussed– always be expected to be the case in practice.

The inconsistency in decision making for the case is related to the estimation of the number of objects. In Poisson finite set models, the minimum mean squared error (MMSE) estimation principle is used which leads to the use of λ as the estimated number of objects[7]. As a result, the EMD density always underestimates the number of objects despite that the source densities might be consistently suggesting otherwise, in practice. The magnitude of the error stemming from this bias depends on the value of z_ω.

C. EMDs of IID cluster finite set densities

IID cluster finite set distributions relax the Poisson cardinality pmf in (26) and take arbitrary cardinality pmfs underpinning the C-PHD filter [41]. The localisation densities still take the factorised form in (27) leading to the identical geometric series $z_\omega(n)$ in (28). For the case, Proposition 3.1 specialises as follows:

Corollary 3.5 (IID cluster inconsistency): Given two IID cluster finite set distributions $f_i = (p_i(n), \rho_i)$ and $f_j = (p_j(n), \rho_j)$, the fused cardinality distribution p_ω satisfies the inconsistency condition in (17) in Proposition 3.1 for the number of objects n and non-zero $p_i(n), p_j(n)$ if

$$z_\omega < \mathcal{I}_{\omega,n}$$

holds, where the term on the right hand side is

$$\mathcal{I}_{\omega,n} \triangleq \left(N_\omega \frac{\min\{p_i(n), p_j(n)\}}{p_i^{(1-\omega)}(n) p_j^{\omega}(n)} \right)^{1/n}, \tag{33}$$

z_ω is given in (29), and, N_ω is obtained by substituting (28) in (16).

The proof follows from substituting (28) in (18) and using (16) after rearrangement of the terms. The inconsistency condition in (33) depends both on n and ω, and, it is not straightforward to relate the inconsistent bins to object number estimation either in the MMSE or MAP rules. On the other hand, the base of the exponent in (33) is smaller than one and hence $\mathcal{I}$ approaches to one as n grows. Therefore, for some threshold η, the fused object number probabilities will be lower than the input cardinality reports for all $n > \eta$. Such a threshold can

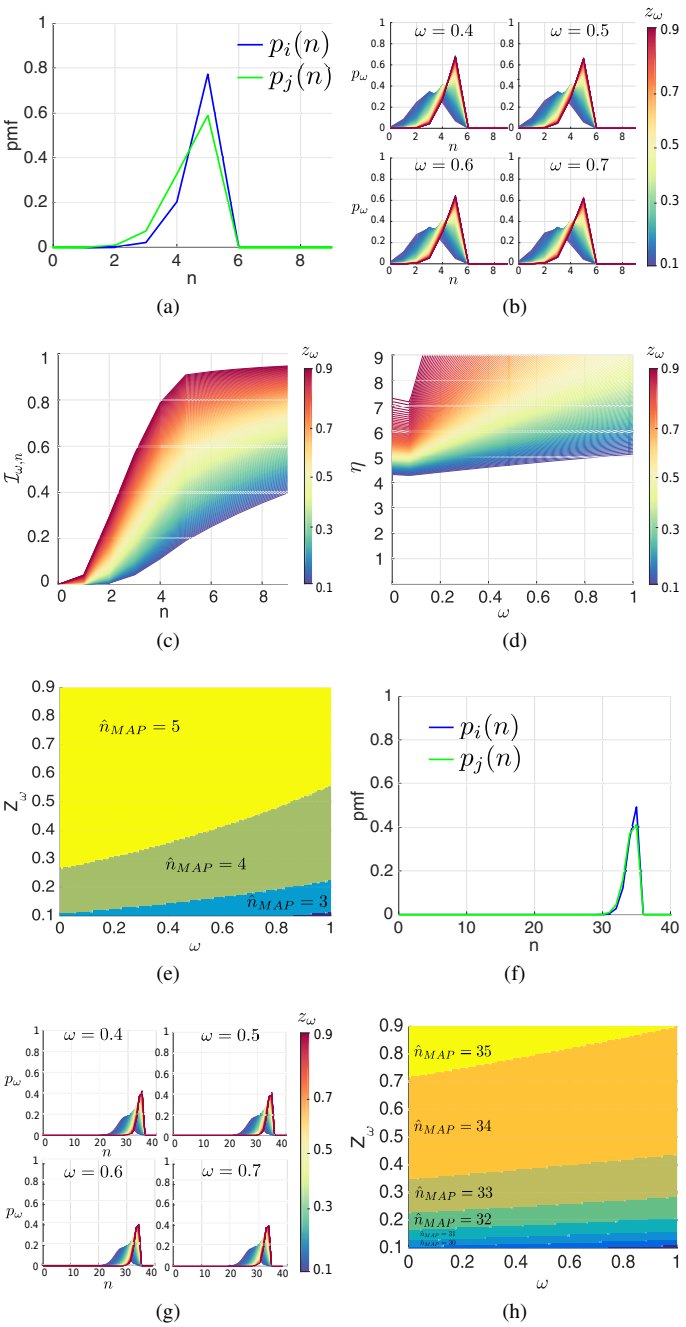


Fig. 3. Illustrations of the results in Example 3.6: (a) Two cardinality distributions peaking at $n = 5$. (b) Fused cardinalities for some intermediate values of ω and $0.1 \leqslant z_\omega \leqslant 0.9$. (c) The inconsistency upper bound in (33). (d) The inconsistency threshold in (34). (e) MAP estimates using the fused cardinalities for varying ω and z_ω. (f) Cardinality distributions peaking at $n = 35$. (g) Fused cardinalities for some intermediate values of ω and $0.1 \leqslant z_\omega \leqslant 0.9$. (h) MAP estimates using the fused cardinalities for varying ω and z_ω.

[7]*Maximum a posteriori* (MAP) estimation is not used with Poisson cardinality distributions as (26) is not guaranteed to have a unique maximum. Notice that, for example, (26) evaluates at the same value for both $n = 0$ and $n = 1$ for $\lambda = 1$.

easily be found from (33) as

$$\eta = \frac{\log\left(N_\omega \gamma_\omega\right)}{\log z_\omega}, \tag{34}$$
$$\gamma_\omega \triangleq \min_{n'} \frac{\min\{p_i(n'), p_j(n')\}}{p_i^{(1-\omega)}(n') p_j^{\omega}(n')}$$

As a result, one should expect estimation biases to become more severe for higher object numbers. For a small number of objects, these effects do not necessarily yield biases in MAP estimations, which also explains the accurate estimates obtained using EMD fusion of C-PHD filters in simulated scenarios, e.g., in [14]. Next, we demonstrate this point in an example involving fusion of two binomial cardinality distributions.

Example 3.6: Let us consider the EMD fusion of two finite set distributions with binomial cardinalities given by $p_i(n) = B(n; k = 5, P = 0.95)$ and $p_j(n) = B(n; k = 5, P = 0.92)$ where these distributions give the probability that n objects exist simultaneously among $k = 5$ possibilities each with an existence probability of P (Fig. 3(a)). Of particular interest is the characteristics of p_ω as z_ω and ω vary in $0.1 \leqslant z_\omega \leqslant 0.9$ and $0 \leqslant \omega \leqslant 1$, respectively. Fig. 3(b) presents fused distributions obtained by varying z_ω and some intermediate values of ω. Note that the cardinality n at which the fused distributions peaks varies with z_ω as suggested in Corollary 3.5. In particular, the inconsistency bound in (3c) is illustrated in Fig. 3(c) which monotonically increases with n as discussed. The inconsistency threshold for n as given in (34) is given in Fig. 3(d). Note that for a large ratio of z_ω and ω values, this threshold is larger than five and the MAP estimate for the cardinality given in Fig. 3(e) agrees with the individual MAP estimates of $\hat{n}_i = \hat{n}_j = 5$. However, there are also MAP estimates that indicate less than five objects caused by the IID inconsistency.

These computations are repeated for cardinality distributions peaking at a higher n value. Specifically, $p_i(n) = B(n; k = 35, P = 0.98)$ and $p_j(n) = B(n; k = 35, P = 0.975)$ are used (see Fig. 3(f)) which have individual map estimates of $\hat{n}_i = \hat{n}_j = 35$. The fused cardinalities in Fig. 3(g) illustrate that for a larger subset of (z_ω, ω) pairs the IID inconsistency occurs, now, as discussed above. The resulting errors in estimating the number of objects is given in Fig. 3(h) which verifies our expectation: Based on that (33) approaches to 1 with increasing n, as the peak cardinality increases, the IID inconsistency detoriates decision making more.

D. Summary of results

As a summary, this section has shown that when EMDs of finite set densities are used for their fusion, the resulting cardinality distribution will bear inconsistencies depending on $z_\omega(n)$. Proposition 3.1 provides a general condition on the fused distribution to be inconsistent with the input distributions at a cardinality value n. This condition is specialised for Bernoulli finite set densities in Corollary 3.2. Example 3.3 has demonstrated that this condition holds with overwhelming probability for Bernoulli EMDs. In Poisson cardinality distributions, there is a single parameter λ that specifies the distribution for all n. Proposition 3.4 provides a condition of inconsistency in this parameter, similarly as an upper bound on $z_\omega(n)$. It is pointed out that because $z_\omega(n)$ is determined by the sensing diversity as well as sensor measurement histories in a sensor network, its value should be expected to be less than one in these settings[8] which in turn shows that Poisson EMDs are very prone to inconsistencies, as well. IID cluster processes have more general cardinality distributions. For the case, Proposition 3.1 specialises to Corollary 3.5 which reveals that inconsistencies should be expected in MAP estimates of the cardinality, when the input densities indicate a high number of objects. These points are demonstrated in Example 3.6.

IV. Cardinality consistent fusion of finite set distributions

In this section, we propose a new approach that accommodates EMD fusion while avoiding the cardinality inconsistencies detailed in Section III. These inconsistencies result from the dependency of the fused cardinality pmf on the scaling factor series $z_\omega(n)$. One way to remove this dependency is to decouple the fusion problem for different cardinalities by asserting a separate variational problem for each cardinality as opposed to using P2 in (8) with finite set distributions as a single entity.

A. Variational problem definitions

Let us first consider finite set distributions as parameterised in (10) and remind that problem P2 is solved with distributions in the form given in (13)–(16). Now, let us consider the following family of variational problems given f_i and f_j:

$$\begin{aligned} &\text{(P3)} \quad \text{For } n = 1, 2, \ldots \\ &(\omega_n^*, \rho_{\omega_n^*,n}) \triangleq \arg \max_{\omega \in [0,1]} \min_{\rho_n \in \mathcal{P}_n} J_{\omega,n}[\rho_n] \\ &J_{\omega,n}[\rho_n] \triangleq (1-\omega) D(\rho_n || \rho_{i,n}) + \omega D(\rho_n || \rho_{j,n}). \end{aligned} \tag{35}$$

Here, $\mathcal{P}_n$ is the space of localisation densities with n arguments which are symmetric in their arguments. Note that P3 is a set of P2 that has the localisation distributions for each cardinality n as the entries, separately. Equivalently, P3 asserts the variational problem of fusion be treated as a conditional problem to be solved given n.

Following our discussion in Section II, solutions of these uncoupled problems have an EMD form given by (13) and (14)[9]. One difference here compared to the solution of problem P2 is that for each n, a different optimal weight ω_n^* will be output, in general, as opposed to a single one. In addition –and, more importantly– problem P2 decouples fusion of the cardinality distributions thus given $(\omega_n^*, \rho_{\omega_n^*,n})$, the fused cardinality distribution becomes an additional degree of freedom in the fused finite set distribution. In other words, the fusion of

[8]The authors at this point would like to conjecture that $z_\omega(n) < 1$ with probability one in a multi-sensor setting in which the finite set densities to be fused are posteriors obtained from recursive Bayesian filtering of local sensor data, i.e., $f_i(X) = f(X|Z_{1:t}^i)$ and $f_j(X) = f(X|Z_{1:t}^j)$ for realisations $Z_{1:t}^i$ and $Z_{1:t}^j$ of (independent) measurement processes associated with sensors i and j, respectively.

[9]It is easy to show that because $\rho_{i,n}$ and $\rho_{j,n}$ are symmetric in their arguments, $\rho_{\omega,n}$ also exhibits this symmetricity.

cardinality distributions can now be carried out in an isolated fashion in addition to problem P2 as a solution to

$$\text{(P4)} \quad (\omega_c^*, \tilde{p}_{\omega_c^*}) \triangleq \arg \max_{\omega\in[0,1]} \min_{p\in\mathcal{P}_c} J_{\omega,c}[p] \tag{36}$$
$$J_{\omega,c}[p] \triangleq (1-\omega)D(p||p_i) + \omega D(p||p_j).$$

Following the discussion in Section II, the solution to problem P4 is the EMD of the cardinality pmfs

$$\tilde{p}_\omega(n) = \frac{1}{\tilde{N}_\omega} p_i^{(1-\omega)}(n) p_j^{\omega}(n) \tag{37}$$
$$\tilde{N}_\omega = \sum_{n'=0} p_i^{(1-\omega)}(n') p_j^{\omega}(n'). \tag{38}$$

evaluated at $\omega = \omega_c^*$.

This distribution differs from the cardinality of the solution to P2 (given in (15) and (16)) in that it does not involve $z_\omega(n)$, and, is an EMD of the input finite set cardinalities. Therefore, the consistency condition (see (9))

$$\tilde{p}_\omega(n) \geqslant \min\{p_i(n), p_j(n)\}$$

is satisfied for all n and for all ω regardless of $z_\omega(n)$. Thus, $\tilde{p}_\omega$ prevents the decision errors stemming from the cardinality inconsistencies of the solutions to P2 as detailed in the previous section.

As a result, P3 and P4 yield a fused finite set density featuring cardinality consistency given by

$$\tilde{f}^*(X) = \tilde{f}_\Omega(X)\Big|_{\Omega=(\omega_c=\omega_c^*,\omega_1=\omega_1^*,\omega_2=\omega_2^*,\ldots)} \tag{39}$$
$$\tilde{f}_\Omega(X) \triangleq p_{\omega_c}(n)\, n! \rho_{\omega_n,n}(x_1,...,x_n) \tag{40}$$

where ω_c^* is found by solving the maximisation in (36) with a cardinality distribution given by (37) and (38). Here, ω_n^* solves the maximisation in (35) with a localisation distribution in (13) and (14). These localisation distributions – similar to the cardinality distribution– are consistent individually, as they are EMDs of the inputs.

The pointwise consistency of $\tilde{f}$ over the space of finite sets, however, is not guaranteed. In order to clarify this point, we provide the following proposition:

Proposition 4.1 (Pointwise inconsistency): Let us consider

$$\tilde{f}_\omega(X) \triangleq \tilde{f}_\Omega(X)\Big|_{\Omega=(\omega_c=\omega,\omega_1=\omega,\omega_2=\omega,\ldots)} \tag{41}$$

for some ω. $\tilde{f}_\omega$ is pointwise inconsistent, i.e.,

$$\tilde{f}_\omega(X) < \min\{f_i(X), f_j(X)\} \tag{42}$$

if

$$\frac{E_{\tilde{p}_\omega}\{z_\omega(n)\}}{z_\omega(n)} < \frac{\min\{f_i(X), f_j(X)\}}{f_\omega(X)} <= 1 \tag{43}$$

where f_ω is the finite set EMD given in (13)–(16).

Proof. By comparing (13)–(16) and (37)–(40), it can be seen that the two finite set densities of concern are related by

$$\tilde{f}_\omega(X) = \frac{E_{\tilde{p}_\omega}\{z_\omega(n)\}}{z_\omega(n)} f_\omega(X). \tag{44}$$

where the expectation is with respect to (37).

Substitution of (44) in (42) yields the first inequality in (43).

Algorithm 1 Newton iterations for solving the nth problem in P3.

Input: $\rho_{i,n}$,$\rho_{j,n}$ ▷ Localisation densities
Input: $\omega^{(0)} \in [0,1]$, ϵ ▷ Initial value, termination threshold
$k \leftarrow 1$, $\omega^{(1)} \leftarrow \infty$
while $|\omega^{(k)} - \omega^{(k-1)}| > \epsilon$ **do** ▷ termination condition
  $z_k \leftarrow z_\omega(n)|_{\omega=\omega^{(k)}}$ using (14)
  $z'_k \leftarrow z'_\omega(n)|_{\omega=\omega^{(k)}}$ using (47)
  $z''_k \leftarrow z''_\omega(n)|_{\omega=\omega^{(k)}}$ using (48)
  $\omega^{(k+1)} \leftarrow \omega^{(k)} - z'_k z_k / \left(z''_k z_k - (z'_k)^2\right)$
  $k \leftarrow k+1$
end while
Return $\omega_n^* \leftarrow \omega^{(k)}$
Return $\rho_{\omega_n^*,n}$ using (13)

Note that, f_ω satistifies the pointwise consistency condition in (9), hence, the right hand side of the inequality is smaller than or equal to one. ■

This proposition points out that pointwise consistency of $\tilde{f}_\omega$ is guaranteed only for those X with cardinality n for which the scaling factor of the localisation density $z_\omega(n) \leqslant 1$ equals to the expectation. If $z_\omega(n)$ is greater than the expectation to the extent that (43) is satisfied, then $\tilde{f}_\omega$ exhibits pointwise inconsistency despite being consistent in the global cardinality and localisation distributions. As a conclusion, pointwise consistency does not imply consistency in global cardinality in fusion of finite set densities and vice versa.

B. Solving the cardinality consistent fusion problems

The variational problems P3 and P4 are $\max\min$ optimisation problems similar to (8). Hence, the minimisations given ω are solved by the EMDs of their argument distributions (see the discussion in Section II and Appendix A). The objective of the outer maximisation in P3 is therefore (see also (6))

$$\begin{aligned} G_n(\omega_n) &\triangleq J_{\omega,n}[\rho_n]\Big|_{\rho_n=\rho_{\omega,n}} \\ &= -(\omega-1)\mathcal{R}_\omega(\rho_{n,i},\rho_{n,j}) \\ &= -\log z_\omega(n) \end{aligned} \tag{45}$$

which is a concave function of its one dimensional argument that takes values from a bounded interval. Newton iterations converge to a solution and have an excellent convergence rate [42]. Starting from an initial value $\omega^{(0)} \in [0,1]$, recursive increments are made by the ratio of the first and second order derivatives, i.e.,

$$\frac{G'_n(\omega_n)}{G''_n(\omega_n)}\Big|_{\omega_n=\omega^{(k)}} = \frac{z'_\omega(n) z_\omega(n)}{z''_\omega(n) z_\omega(n) - (z'_\omega(n))^2}\Big|_{\omega=\omega^{(k)}} \tag{46}$$
$$z'_\omega(n) \triangleq \int \rho_{i,n}^{1-\omega}(x) \rho_{j,n}^{\omega}(x) \log \frac{\rho_{j,n}(x)}{\rho_{i,n}(x)} \mathrm{d}x \tag{47}$$
$$z''_\omega(n) \triangleq \int \rho_{i,n}^{1-\omega}(x) \rho_{j,n}^{\omega}(x) \left(\log \frac{\rho_{j,n}(x)}{\rho_{i,n}(x)}\right)^2 \mathrm{d}x \tag{48}$$

where $\omega^{(k)}$ is the value found in the kth iteration. Here, $z_\omega(n)$ is given in (14) and its derivatives in (47) and (48)

Algorithm 2 Newton iterations solving P4 for consistent cardinality fusion.

1: Input: $p_i(n)$,$p_j(n)$ ▷ Cardinality pmfs
2: Input: $\omega^{(0)} \in [0,1]$, ϵ ▷ Initial value, termination threshold
3: $k \leftarrow 1$, $\omega^{(1)} \leftarrow \infty$
4: **while** $|\omega^{(k)} - \omega^{(k-1)}| > \epsilon$ **do** ▷ termination condition
5: $N_k \leftarrow N_\omega|_{\omega=\omega^{(k)}}$ using (16)
6: $N_k' \leftarrow \sum p_i^{1-\omega}(n) p_j^{\omega}(n) \log\big(p_j(n)/p_i(n)\big)$
7: $N_k'' \leftarrow \sum p_i^{1-\omega}(n) p_j^{\omega}(n) \Big(\log\big(p_j(n)/p_i(n)\big)\Big)^2$
8: $\omega^{(k+1)} \leftarrow \omega^{(k)} - N_k' z_k / \Big(N_k'' N_k - (N_k')^2\Big)$
9: $k \leftarrow k+1$
10: **end while**
11: Return $\omega_c^* \leftarrow \omega^{(k)}$
12: Return $p_{\omega_c^*}$ using (15)

are found in Appendix C. Algorithm 1 explicitly specifies this iterative solution which takes the localisation densities as inputs together with an initial value and termination condition. Upon convergence, the optimal value $\omega_n^* = \omega_n^{(k)}$ is found for which the corresponding EMD $\rho_{\omega_n^{(k)},n}$ is the fused density.

An analogous iterative algorithm for finding the consistently fused cardinality distribution as a solution to P4 is given in Algorithm 2. Note that the computations involved here can be carried out exactly for distributions with finite support, in practice.

Algorithm 1, on the other hand, should accommodate adequate computational schemes for exactly or approximately evaluating the integrals involved. In the latter case, it admits the interpretation of being a stochastic gradient approach [43]. Specification of such procedures is beyond the scope of this work.

There are, nevertheless, structural simplifications in both P3 and P4 for different families of finite set families. For Poisson and IID cluster finite set distributions, the localisation densities are parameterised by a single density over a single state variable –as given in (27)– which is the same for different cardinalities. In addition, the solution to the inner minimisation in P3 (equivalently P in (2)) is also a Poisson and IID cluster, respectively, in the Poisson and IID cluster cases [14]. Thus, $D(\rho_n||\rho_{n,i}) = nD(\rho||\rho_i)$ where ρ parameterises ρ_n, and, the family of problems in P3 satisfy

$$J_{\omega,n}[\rho_n] = nJ_{\omega,1}[\rho]. \tag{49}$$

Consequently, the optimal solution to P3 for cardinality n is parameterised by the optimal solution to $n = 1$ thereby restricting it to the case for only $n = 1$.

Bernoulli finite sets have nonzero cardinality pmf only for $n \leqslant 1$ naturally restricting P3 to $n = 1$. If the parameterising densities ρ_i and ρ_j are Gaussians, then P3 specifies a covariance intersection procedure [4]. For this case, P4 has a closed form solution given by [44]

$$\omega_c^* = \frac{\log\left(\frac{\log(1-\alpha_i)/(1-\alpha_j)}{\log(\alpha_j/\alpha_i)}\right) - \log\left(\frac{\alpha_i}{1-\alpha_i}\right)}{\log\frac{1-\alpha_i}{1-\alpha_j} + \log\frac{\alpha_j}{\alpha_i}} \tag{50}$$

$$\alpha^* = \frac{\alpha_i^{1-\omega_c^*}\alpha_j^{\omega_c^*}}{\alpha_i^{1-\omega_c^*}\alpha_j^{\omega_c^*} + (1-\alpha_i)^{1-\omega_c^*}(1-\alpha_j)^{\omega_c^*}} \tag{51}$$

For Poisson cardinality pmfs, similarly a closed form solution exists for P4 which is given by [44]

$$\omega_c^* = \frac{-\log\big(\log(\lambda_j/\lambda_i)\big) + \log(\lambda_j/\lambda_i - 1)}{\log(\lambda_j/\lambda_i)} \tag{52}$$

$$\lambda^* = \lambda_i^{1-\omega_c^*}\lambda_j^{\omega_c^*} \tag{53}$$

It is worthwhile to notice that both Bernoulli and Poisson distributions are exponential family distributions and the above solutions bear the geometric properties aforementioned in Section II-B and proved in [35].

C. Demonstration of cardinality consistent fusion

In this section, we revisit the examples in Section III involving cardinality inconsistencies and demonstrate the efficacy of the solutions of Problems P3 and P4 in these fusion scenarios.

Example 4.2 (Gauss-Bernoulli case revisited): Let us consider the Gauss-Bernoulli case in Example 3.3 in the light of the discussion above. Fusion of localisation distributions in P3 involve the fusion of only a single pair for $n = 1$, for the case. These distributions ρ_i and ρ_j given by (22) are Gaussians, therefore, Algorithm 1 is equivalently an iterative covariance intersection algorithm that optimises ω to achieve the KLD equality criteria in (7). The Newton update for the parameter ω in Step 8 of Algorithm 1 is carried out as follows: Evaluation of z_ω in Step 5 at $\omega = \omega^{(k)}$ is made using its closed form expression in (25). Given z_ω, the derivative in Step 6 is also found in closed form using the following identity

$$z_\omega' = z_\omega\big(D(\rho_\omega||\rho_i) - D(\rho_\omega||\rho_j)\big), \tag{54}$$

which can easily be verified by dividing both sides of (47) to $z_\omega(n)$. This quantity is computed by evaluating the KLD of multi-variate Gaussian densities given by (see, for example, [45, A.23])

$$D(\rho_\omega||\rho_i) = 1/2\log|\mathbf{C}_i\mathbf{C}_\omega^{-1}| + \mathbf{tr}\{\mathbf{C}_i^{-1}\big((\mathbf{m}_\omega - \mathbf{m}_i)(\mathbf{m}_\omega - \mathbf{m}_i)^T + \mathbf{C}_\omega - \mathbf{C}_i\big)\}, \tag{55}$$

where $\mathbf{tr}\{.\}$ is the trace of its matrix argument. The second derivative in Step 7 on the other hand, is found approximately using the Monte Carlo method [46, Chp.3] targeting the integration in (48) divided by $z_\omega(n)$. Therefore, L samples are generated from ρ_ω for this step, i.e., $x^{(l)} \sim \rho_\omega$ for $l = 1, \ldots, L$. Using these samples, the approximation seeked is given by

$$z_\omega'' \approx z_\omega \times \frac{1}{L}\sum_{l=1}^{L}\Big(\log\frac{\rho_j(x^{(l)})}{\rho_i(x^{(l)})}\Big)^2 \tag{56}$$

where the approximation error decreases with $\mathcal{O}(1/\sqrt{L})$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAES.2019.2893083, IEEE Transactions on Aerospace and Electronic Systems

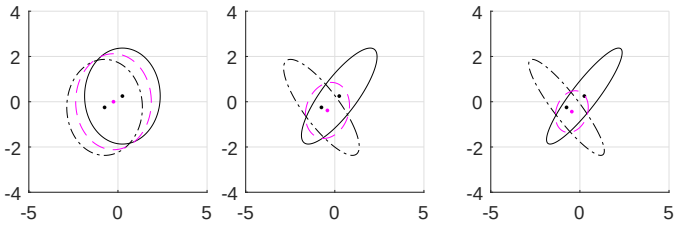


Fig. 4. Optimally weighted EMDs of Gaussian localisation distributions obtained by using Algorithm 1: Here, ρ_i (solid line) and ρ_j (dash-dotted line) are inputs for $\kappa = 1, 10, 20$ (left to right – see Example 3.3 for details) and $\rho_{\omega*}$ (magenta dashed line) are fused outputs with optimal weights found as $\omega^* = 0.500, 0.397$ and 0.387 (left to right).

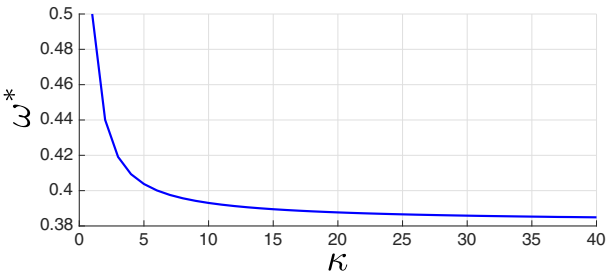


Fig. 5. Optimal weight parameters found using Algorithm 1 as a function of the sensing diversity κ as explained in detail in Example 3.3.

We input the Gaussian pairs in Example 3.3 that are obtained by varying the covariance condition number κ – equivalently the sensing diversity– to Algorithm 1 and use the computational procedures above. Fig. 4 depicts optimally weighted EMDs for $\kappa = 1, 10, 20$. The optimal weight ω^* output by Algorithm 1 as a function of the condition number κ is given in Fig. 5. Here, the termination threshold is set to $\epsilon = 1\text{e}{-4}$ and the number of samples used for the Monte Carlo estimate in Step 7 is $L = 1000$. Convergence is declared after an average of 3.4 and a maximum of 5 iterations. Note that ω^* yields a fused result that bears more influence from ρ_i as κ increases.

Cardinality fusion problem P4 has an analytical solution for the case. Using (50) for $\alpha_i = \alpha_j = 0.8$ (see Example 3.3), we find that $\omega_C^* = 0.5$ and $\alpha^* = 0.8$ regardless of the solution of Problem P3 above, i.e., the optimal weight parameters ω^* of the localisation distributions or the corresponding normalisation constants $z_{\omega*}$s.

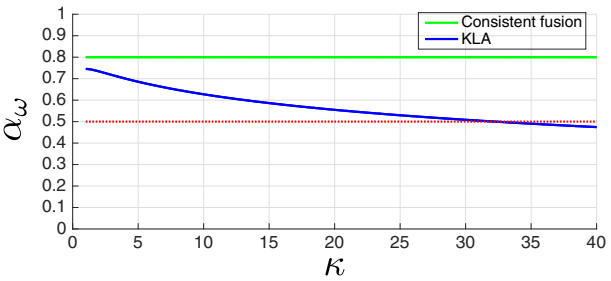


Fig. 6. Fused existence probabilities: Cardinality consistent fusion via Problem P4 (green line) in comparison with fused existence probabilities obtained using KL averaging (blue line–see, e.g., [17]– [23]) depicted as a function of sensing diversity parameter κ as detailed in Example 3.3. The red-dashed line is the canonical Bayesian decision threshold for deciding the existence of an object.

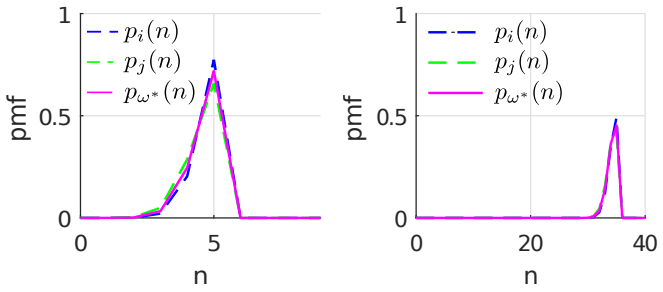
Let us compare this result with conventional EMD fusion with weights selected using "Kullback-Leibler averaging" (KL averaging) as used in, for example, [17]– [23]. In this approach, the fused finite set density is the Bernoulli EMD with weight parameter $\omega = 1/2$. In other words, the fused result solves Problem P for $\omega = 1/2$. The coupling of the fused existence probability with z_ω and the lack of a weight selection mechanism results with the fused existence probabilities given in Fig. 6 which illustrates a monotonically increasing disagreement with the input beliefs on the existence of an object as the sensing diversity increases. This trend results with the fused existence probability falling below the canonical decision threshold of 0.5 [10]. The proposed cardinality consistent fusion, on the other hand, preserves the confidence of input distributions on the existence of an object irrespective of the sensing diversity. The KL averaging fusion outputs a localisation density that is similarly the EMD of the localisation distributions with weight $\omega = 1/2$. This margin between this value and the optimum point found by the proposed algorithm grows significantly with κ, in this example (see Fig. 5).

Note that these results are also relevant for the work in literature on fusion of multi-Bernoulli [20], and, labelled random finite set families as their fusion is often reduced to performing Bernoulli-Bernoulli fusion for multiple pairs using, for example, KL averaging [23].

Example 4.3 (Example 3.6 revisited): Let us demonstrate Algorithm 2 in solving the cardinality fusion problem P4. First, we consider the binomial cardinality distribution pair illustrated in Fig. 3(a). Note that, the algorithm allows for exact computations in all steps. The termination threshold is selected as $\epsilon = 1.0e-4$. In 2 iterations Algorithm 2 declares convergence to the optimal weight parameter $\omega_C^* = 0.5182$. The corresponding fused cardinality pmf is the EMD with this weight and depicted in Fig. 7(a). We repeat the same procedure for the cardinality pair in Fig. 3(g). The proposed algorithm convergences in 2 steps to $\omega_C^* = 0.5090$. The resulting cardinality distribution is depicted in Fig. 7(b). Note that the MAP estimates of the number of objects is in agreement with the inputs. The consistency here is underpinned by that the cardinality fusion here is independent of localisation densities and z_ω.

In order to contrast this result with that obtained by KL averaging (see, e.g., [17]), let us consider the fused cardinalities depicted for $\omega = 1/2$ in the top right panes in Fig. 3(b) and (g) in which the coupling of cardinality fusion with z_ω is demonstrated. Let us remind also that the MAP object number estimate depends on z_ω (Figs. 3(e) and (h)) with an increasing bias towards underestimation as the input distribution peaks shift towards right indicating higher number of objects (Figs. 3(h)). The proposed algorithm, on the other hand, finds a consistent common ground of the input cardinality distributions which is also optimal with respect to Problem P4.

[10]Note that this graph is nothing but the cross-section of the existence probability graph in Fig. 2 along $\omega = 0.5$.



(a) (b)

Fig. 7. Consistent cardinality fusion using Algorithm 2: Inputs (blue and green dashed lines) are the binomial pairs introduced in Example 3.3. The optimal weights converged are $\omega_C^* = 0.5182$ and 0.5090 in (a) and (b), respectively.

V. Conclusion

This work considered the recently growing literature on the use of EMDs –or, weighted geometric means- of finite set densities in multi-sensor fusion for multi-object tracking. EMDs of distributions are pointwise consistent, however, we have proved in this article that they are prone to inconsistency in their cardinality distributions which can lead to serious decision errors related to the number of objects sensed. We have demonstrated that pointwise consistency does not imply consistency in cardinality and vice versa. We remedy this problem by redefining the variational optimisation problem that underlies EMD fusion. Then, we specify iterative solutions and establish a conceptual framework for cardinality consistent fusion of finite set densities which also accommodates EMDs.

Following these results, possible future directions include investigation of numerical computational schemes in order to use within this variational framework. The extension of this variational perspective to accommodate N sources is also anticipated to be a worthwhile direction to pursue.

Appendix

A. Solution of the Problem P in (2)

In this appendix, we provide a direct proof for the assertion that the solution of Problem P in (2) for any ω is the EMD given in (3),(4) when f is constrained to be a density, i.e., to integrate to unity. This constraint on the feasible set of solutions together with the cost functional of the problem are captured in the Lagrangian given by [47]

$$\mathcal{L}[f,\lambda] \triangleq J_\omega[f] + \lambda G[f], \tag{57}$$
$$G[f] \triangleq \left(1 - \int f(X)\mathrm{d}X\right).$$

Here, λ is a free variable referred to as a Lagrange multiplier which –at its stationary point– imposes the constraint of integration to unity on those f which are also stationary. This point together with the convexity of J_ω in f results with the stationary point of (57) f^* being the solution to P in (2) [47].

The necessary (and sufficient) condition of stationarity that f^* should satisfy is given by i) the functional derivative of the Lagrangian, i.e.,

$$\delta\left(J_\omega + \lambda G\right)[f;\delta_X]\Big|_{f=f^*} = 0 \tag{58}$$

for all X, and, ii) the partial differential with respect to the multiplier λ, i.e.,

$$\frac{\partial\left(J_\omega[f] + \lambda G[f]\right)}{\partial\lambda}\Big|_{\lambda=\lambda^*} = 0. \tag{59}$$

The functional derivative in (58) can be expressed in terms of the partial derivative with respect to $f(X)$, i.e.,

$$\delta\left(J_\omega + \lambda G\right)[f;\delta_X] = \frac{\partial\left(J_\omega[f] + \lambda G[f]\right)}{\partial f(X)}$$

for all X. By using the definition of KLD in (1), and rules of differentiation, this expression leads to

$$\frac{\partial\left(J_\omega[f] + \lambda G[f]\right)}{\partial f(X)} = \log f(X) + \frac{f(X)}{f(X)} - (1-\omega)\log f_i(X) - \omega\log f_j(X) - \lambda.$$

The equation above is zero when the multiplier λ takes the value

$$\lambda = 1 + \log\frac{f(X)}{f_i^{(1-\omega)}(X)f_j^{\omega}(X)}, \tag{60}$$

for which the corresponding $f(X)$ is found as

$$f(X) = \exp(\lambda-1)f_i^{(1-\omega)}(X)f_j^{\omega}(X). \tag{61}$$

We substitute from the equality above into the partial differentiation in (59) and obtain

$$\begin{aligned}
&\frac{\partial\left(J_\omega[f] + \lambda G[f]\right)}{\partial\lambda} \\
&= \frac{\partial}{\partial\lambda}\left((\lambda-1)\exp(\lambda-1)\int f_i^{(1-\omega)}(X)f_j^{\omega}(X)\mathrm{d}X\right) \\
&\quad + \frac{\partial}{\partial\lambda}\left(\lambda - \lambda\exp(\lambda-1)\int f_i^{(1-\omega)}(X)f_j^{\omega}(X)\mathrm{d}X\right) \\
&= \frac{\partial}{\partial\lambda}\left(\lambda - \exp(\lambda-1)\int f_i^{(1-\omega)}(X)f_j^{\omega}(X)\mathrm{d}X\right) \\
&= 1 - \exp(\lambda-1)\int f_i^{(1-\omega)}(X)f_j^{\omega}(X)\mathrm{d}X.
\end{aligned} \tag{62}$$

After (62) is set to zero, λ^* is found as

$$\lambda^* = -\log\int f_i^{(1-\omega)}(X)f_j^{\omega}(X)\mathrm{d}X + 1 \tag{63}$$

and, the solution f^* is found by subsituting from (63) into (61) as

$$f^*(X) = \exp\left(-\log\int f_i^{(1-\omega)}(X)f_j^{\omega}(X)\mathrm{d}X\right)f_i^{(1-\omega)}(X)f_j^{\omega}(X).$$

Following a rearrangement of the terms, the expression above takes the form given by

$$f^*(X) = \frac{1}{\int f_i^{(1-\omega)}(X)f_j^{\omega}(X)\mathrm{d}X}f_i^{(1-\omega)}(X)f_j^{\omega}(X), \tag{64}$$

which can be identified as the EMD in (3),(4).

B. Proof of Proposition 3.1

The proof follows from decomposing $p_\omega(n)$ in (15) and (16) as follows

$$p_\omega(n) = \frac{p_i^{(1-\omega)}(n)p_j^{\omega}(n)z_\omega(n)}{p_i^{(1-\omega)}(n)p_j^{\omega}(n)z_\omega(n) + \sum_{n'\neq n} p_i^{(1-\omega)}(n')p_j^{\omega}(n')z_\omega(n')}, \tag{65}$$

and substituting on the left hand side of the inequality in (17). The inequality can easily be solved for $z_\omega(n)$ leading to (18).

C. Derivation of Algorithm 2

We start by finding the first and second order derivatives of G_n in (45):

$$G_n'(\omega_n) \triangleq \left.\frac{dG_n(\omega)}{d\omega}\right|_{\omega=\omega_n} = -\left.\frac{dz_\omega(n)/d\omega}{z_\omega(n)}\right|_{\omega=\omega_n}, \tag{66}$$

$$G_n''(\omega_n) \triangleq \left.\frac{d^2G_n(\omega)}{d\omega^2}\right|_{\omega=\omega_n} = -\left.\frac{d^2z_\omega(n)/d\omega^2 \times z_\omega(n) - (dz_\omega(n)/d\omega)^2}{(z_\omega(n))^2}\right|_{\omega=\omega_n} \tag{67}$$

where $z_\omega(n)$ is given by (14). Newton iterations [42] use recursive increments to the scalar argument of maximisation. These increments are found by evaluating the ratio of (66) and (67) which is found as

$$\frac{G_n'(\omega_n)}{G_n''(\omega_n)} = \left.\frac{z_\omega'(n)z_\omega(n)}{z_\omega''(n)z_\omega(n) - (z_\omega'(n))^2}\right|_{\omega=\omega_n}, \tag{68}$$

$$z_\omega'(n) \triangleq \frac{dz_\omega(n)}{d\omega} \tag{69}$$

$$z_\omega''(n) \triangleq \frac{dz_\omega'(n)}{d\omega} \tag{70}$$

in terms of $z_\omega(n)$ and its first and second order derivatives.

Next, let us find the derivatives of $z_\omega(n)$. The first order derivative follows after substituting (14) in (69) as

$$\begin{aligned}\frac{dz_\omega(n)}{d\omega} &= \int \rho_{i,n}(x)\frac{d}{d\omega}\left(\frac{\rho_{j,n}(x)}{\rho_{i,n}(x)}\right)^{\omega}\mathrm{d}x \\ &= \int \rho_{i,n}^{1-\omega}(x)\rho_{j,n}^{\omega}(x)\log\frac{\rho_{j,n}(x)}{\rho_{i,n}(x)}\mathrm{d}x\end{aligned}$$

which is equivalent to (47). The second order derivative seeked is found by substituting from the above equality into (70) as

$$\begin{aligned}\frac{dz_\omega'(n)}{d\omega} &= \int \rho_{i,n}(x)\log\frac{\rho_{j,n}(x)}{\rho_{i,n}(x)}\frac{d}{d\omega}\left(\frac{\rho_{j,n}(x)}{\rho_{i,n}(x)}\right)^{\omega}\mathrm{d}x \\ &= \int \rho_{i,n}^{1-\omega}(x)\rho_{j,n}^{\omega}(x)\left(\log\frac{\rho_{j,n}(x)}{\rho_{i,n}(x)}\right)^2\mathrm{d}x\end{aligned}$$

which is equivalently given in (48).

REFERENCES

[1] D. Hall, C.-Y. Chong, J. Llinas, and M. L. II, Eds., *Distributed Data Fusion for Network-Centric Operations*. CRC Press, 2013.

[2] S. Julier, T.Bailey, and J. Uhlmann, "Using exponential mixture models for suboptimal distributed data fusion," in *Proc. of the 2006 IEEE Nonlinear Stat. Signal Proc. Workshop (NSSPW'06)*. NSSPW'06, September 2006, pp. 160–163.

[3] S. J. Julier and J. K. Uhlmann, "A Non-divergent Estimation Algorithm in the Presence of Unknown Correlations," in *Proceedings of the IEEE American Control Conference*, vol. 4, Albuquerque NM, USA, June 1997, pp. 2369–2373.

[4] M. Hurley, "An information-theoretic justification for covariance intersection and its generalization," in *Proceedings of the 2002 FUSION Conference*. FUSION 2002, July 2002.

[5] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Second, Ed. John Wiley and Sons, 2006.

[6] T. Heskes, "Selecting weighting factors in logarithmic opinion pools," in *Advances in Neural Information Processing Systems*. The MIT Press, 1998, pp. 266–272.

[7] R. Mahler, "Optimal/robust distributed data fusion: a unified approach," in *Proceedings of the SPIE Defense and Security Symposium 2000*. SPIE Defense and Security Symposium, 2000.

[8] R. P. S. Mahler, *Statistical Multisource Multitarget Information Fusion*. Springer, 2007.

[9] R. Mahler, "Multi-target Bayes filtering via first-order multi-target moments," *IEEE Transactions on Aerospace and Electronic Sysmtes*, vol. 39, no. 4, pp. 1152–1178, October 2003.

[10] B. Vo and W. K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4091–4104, November 2006.

[11] B. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo methods for multi-target filtering with random finite sets." *IEEE Transaction Aerospace and Electronics*, vol. 41, no. 4, pp. 1224–1245, October 2005.

[12] M. Üney, S. Julier, D. Clark, and B. Ristić, "Monte carlo realisation of a distributed multi–object fusion algorithm." SSPD 2010, September 2010.

[13] M. Üney, D. Clark, and S. Julier, "Information Measures in Distributed Multitarget Tracking," in *Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on*, july 2011, pp. 1 –8.

[14] M. Üney, D. E. Clark, and S. Julier, "Distributed fusion of PHD filters via exponential mixture densities," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 7, no. 3, pp. 521 – 531, June 2013.

[15] J. Barr, M. Uney, D. Clark, D. Miller, M. Porter, A. Gning, and S. Julier, "A multi-sensor inference and data fusion method for tracking small, manoeuvrable maritime craft in cluttered regions," in *Proceedings of the 3rd IMA on Mathematics in Defence*, 10 2013.

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

[17] G. Battistelli, L. Chisci, C. Fantacci, A. Farina, and A. Graziano, "Consensus CPHD filter for distributed multitarget tracking," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 508–520, June 2013.

[18] B. Wang, W. Yi, R. Hoseinnezhad, S. Li, L. Kong, and X. Yang, "Distributed fusion with multi-Bernoulli filter based on generalized covariance intersection," *IEEE Transactions on Signal Processing*, vol. 65, no. 1, pp. 242–255, Jan 2017.

[19] M. Jiang, W. Yi, R. Hoseinnezhad, and L. Kong, "Distributed multi-sensor fusion using generalized multi-bernoulli densities," in *19th International Conference on Information Fusion (FUSION)*, July 2016, pp. 1332–1339.

[20] W. Yi, M. Jiang, R. Hoseinnezhad, and B. Wang, "Distributed multi-sensor fusion using generalised multi-bernoulli densities," *IET Radar, Sonar & Navigation*, vol. 11, pp. 434–443(9), March 2017.

[21] M. B. Guldogan, "Consensus bernoulli filter for distributed detection and tracking using multi-static doppler shifts," *IEEE Signal Processing Letters*, vol. 21, no. 6, pp. 672–676, June 2014.

[22] G. Battistelli, L. Chisci, C. Fantacci, A. Farina, and B. N. Vo, "Average Kullback-Leibler divergence for random finite sets," in *2015 18th International Conference on Information Fusion (Fusion)*, July 2015, pp. 1359–1366.

[23] S. Li, W. Yi, R. Hoseinnezhad, G. Battistelli, B. Wang, and L. Kong, "Robust distributed fusion with labeled random finite sets," *IEEE Transactions on Signal Processing*, vol. 66, no. 2, pp. 278–293, Jan 2018.

[24] A. Dabak, "A Geometry for Detection Theory," Ph.D. dissertation, Rice University, Houston, TX, USA, 1992.

[25] S. J. Julier, T. Bailey, and J. K. Uhlmann, "Using exponential mixture models for suboptimal distributed data fusion," in *Proc. of the 2006 IEEE*

Nonlinear Stat. Signal Proc. Workshop. Cambridge, UK: NSSPW'06, September 2006, pp. 160–163.

[26] D. Daley and D. Vere-Jones, *An introduction to the theory of point processes*. Springer, 1988.

[27] M. Gunay, U. Orguner, and M. Demirekler, "Chernoff fusion of Gaussian mixtures based on sigma-point approximation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 6, pp. 2732–2746, December 2016.

[28] I. Csiszar and P. Shields, "Information theory and statistics: A tutorial," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 4, pp. 417–528, 2004. [Online]. Available: http://dx.doi.org/10.1561/0100000004

[29] I. Csiszar and F. Matus, "Information projections revisited," *IEEE Transactions on Information Theory*, vol. 49, no. 6, pp. 1474–1490, June 2003.

[30] A. Rényi, "On Measures of Entropy and Information," in *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, vol. 1, 1960, pp. 547–561.

[31] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *Ann. Math. Statist.*, vol. 23, no. 4, pp. 493–507, 12 1952. [Online]. Available: https://doi.org/10.1214/aoms/1177729330

[32] S. J. Julier, "An empirical study into the use of Chernoff information for robust, distributed fusion of Gaussian mixture models," in *Proceedings of the IEEE ICIF*, 2006.

[33] L. Chen, P. O. Arambel, and R. K. Mehra, "Fusion under unknown correlation - covariance intersection as a special case," in *Proceedings of the Fifth International Conference on Information Fusion. FUSION 2002. (IEEE Cat.No.02EX5997)*, vol. 2, July 2002, pp. 905–912 vol.2.

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

[35] F. Nielsen, "An information-geometric characterization of Chernoff information," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 269–272, March 2013.

[36] K. Chang, C. Y. Chong, and S. Mori, "Analytical and computational evaluation of scalable distributed fusion algorithms," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 4, pp. 2022–2034, Oct 2010.

[37] B. T. Vo and B. N. Vo, "Labeled random finite sets and multi-object conjugate priors," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3460–3475, July 2013.

[38] J. L. Williams, "Marginal multi-Bernoulli filters: RFS derivation of MHT, JIPDA, and association-based MeMBer," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 3, pp. 1664–1687, July 2015.

[39] F. Baccelli and J. O. Woo, "On the entropy and mutual information of point processes," in *2016 IEEE International Symposium on Information Theory (ISIT)*, July 2016, pp. 695–699.

[40] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*. Cambridge University Press, 1934.

[41] R. Mahler, "PHD filters of higher order in target number," *IEEE Transactions on Aerospace and Electronic Sysmtes*, vol. 43, no. 4, pp. 1523–1543, October 2007.

[42] M. S. Bazaraa, H. D. Sherali, and C. Shetty, *Nonlinear Programming*, 2nd ed. John Wiley & Sons, Inc., 1993.

[43] S. Bubeck, "Convex optimization: Algorithms and complexity," *Foundations and Trends in Machine Learning*, vol. 8, no. 3-4, pp. 231–357, 2015. [Online]. Available: http://dx.doi.org/10.1561/2200000050

[44] F. Nielsen, "Chernoff information of exponential families," *CoRR*, vol. abs/1102.2684, 2011. [Online]. Available: http://arxiv.org/abs/1102.2684

[45] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[46] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. Springer, 2004.

[47] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear programming: Theory and algorithms*, 3rd ed. John Wiley & Sons, 2006.