A linear-complexity second-order multi-object filter via factorial cumulants

Daniel E. Clark and Flávio de Melo

Abstract—**Multi-target tracking solutions with low computational complexity are required in order to address large-scale tracking problems. Solutions based on statistics determined from point processes, such as the PHD filter, CPHD filter, and newer second-order PHD filter are some examples of these algorithms. There are few solutions of linear complexity in the number of targets and number of measurements, with the PHD filter being one exception. However, the trade-off is that it is unable to propagate beyond first-order moment statistics. In this paper, a new filter is proposed with the same complexity as the PHD filter that also propagates second-order information via the second-order factorial cumulant. The results show that the algorithm is more robust than the PHD filter in challenging clutter environments.**

Index Terms—**Multi-target tracking, point processes, factorial cumulants.**

I. Introduction

Due to the practical demand and complexities of multiple target tracking, there have been many different approaches to addressing the problem. Methods for multiple hypothesis tracking have attracted a great deal of attention since they provide an intuitively appealing basis for designing an engineering system, eg. [1]–[4]. Methods for managing joint multi-target probabilities [5]–[7] represent a different way of dealing with data association ambiguities. Methods based on nearest neighbour assignments [8], [9] provide pragmatic solutions to dealing with data association assignment. More recently, closed-form solutions to a class of multi-target tracking problems have been developed [10]–[12], which are gaining attention for scenarios where there is a prioritisation of estimation accuracy over computational complexity.

Methods based on point processes [13]–[17] and random finite sets [18] have also attracted a lot of attention. In particular, the applications of methodology developed in the context of random finite sets have proved to be popular [19]–[25]. Specifically, methods based on propagation of first-order moment statistics of a point process, or intensity function, have been developed for addressing tracking problems, eg. [19].

Methods for propagating higher-order information have been proposed [20], [26], [27]. It has been shown that the popular CPHD filter [20] can have undesirable behaviour [28], which has been shown to be due to strong negative correlations being induced after Bayes' rule [29]. More recently, second-order filtering solutions have been proposed that propagate the mean and variance in target number [29].

In the current paper, a new linear-complexity multi-target filter is derived and implemented that propagates second-order factorial cumulants. It is shown that the filter is more robust than the PHD filter while maintaining the same run-time. This is achieved by making a different approximation on the joint target-measurement process before applying Bayes' rule.

The paper is structured as follows: Point processes are described in the next section in terms of functionals, namely the probability generating functional the factorial cumulant generating functional, and statistics in terms of factorial moments and factorial cumulants and their connections. Section III describes point process models, including the Poisson process and the negative binomial point process via the Laplace-Stieltjes transform, and the extension to the Panjer point process. Section IV presents the algorithm description in terms of propagation of the first two factorial cumulants. Section V shows some experimental scenarios comparing with the PHD filter. The paper concludes in Section VI. The appendix presents the mathematical proofs of the algorithm specification. Algorithm 1 presents the algorithm pseudo-code of a Gaussian mixture implementation based on a similar approach to the Gaussian mixture PHD filter [21].

II. Point processes

In this section we describe point processes [30], [31] and their statistical descriptors, including factorial moments and factorial cumulants. We denote by $\mathcal{X} \subseteq \mathbb{R}^{d_x}$ the d_x-dimensional state space describing the state of an individual object. A point process Φ on $\mathcal{X}$ is a random variable on the process space $\mathfrak{X} = \bigcup_{n\geq 0} \mathcal{X}^n$, i.e., the space of finite sequences of points in $\mathcal{X}$. A realisation of Φ is a sequence $\varphi = (x_1, \dots, x_n) \in \mathcal{X}^n$, representing a population of n objects with states $x_i \in \mathcal{X}$. Point processes can be described using their probability distribution P_Φ on the measurable space $(\mathfrak{X}, \mathcal{B}(\mathfrak{X}))$, where $\mathcal{B}(\mathfrak{X})$ denotes the Borel σ-algebra of the process space $\mathfrak{X}$ [32]. The projection measure $P_\Phi^{(n)}$ of the probability distribution P_Φ on $\mathcal{X}^n$, $n \geq 0$, describes the realisations of Φ with n elements; the projection measures of a point process are always defined as symmetrical functions, so that the permutations of a realisation φ are equally probable.

Daniel Clark is with Telecom SudParis, France. Email: daniel.clark@telecom-sudparis.eu and Flávio de Melo is with Heriot-Watt University, UK. This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) [Grant Number EP/K014227/1]; and the MOD University Defence Research Collaboration (UDRC) in Signal Processing.

A. Generating functional representations

The probability generating functional (p.g.fl.) $\mathcal{G}_\Phi$ of a point process Φ is defined by

$$\mathcal{G}_\Phi(h) = \sum_{n\geq 0} \int \left[\prod_{i=1}^{n} h(x_i)\right] P_\Phi^{(n)}(\mathrm{d}x_{1:n}), \quad (1)$$

respectively for test function $h : \mathcal{X} \to [0,1]$. The *factorial cumulant generating functional (f.g.fl.)*, is defined by

$$\mathcal{C}_\Phi(h) = \ln\left(\mathcal{G}_\Phi(h)\right). \quad (2)$$

We can construct the f.g.fl. through the definition of Khinchin measures as follows. Let $\{K_\Phi^{(n)}\}_{n\geq 1}$ be a sequence of measures and consider a point process defined through the p.g.fl.

$$\mathcal{G}_\Phi(h) = \exp\left(\mathcal{C}_\Phi(h)\right) \quad (3)$$

where

$$\mathcal{C}_\Phi(h) = -K_\Phi^{(0)} + \sum_{n\geq 1} \int h(x_1)\dots h(x_n) K_\Phi^{(n)}(dx_{1:n}), \quad (4)$$

and

$$K_\Phi^{(0)} = \sum_{n\geq 1} \int K_\Phi^{(n)}(dx_{1:n}) \quad (5)$$

ensures that $\mathcal{G}_\Phi(1) = 1$ and therefore G_Φ is a p.g.fl.. The Khinchin measures were recently used to determine results for Bayesian estimation of multi-object systems with independently identically distributed correlations [33] by considering a bivariate version defined in an analogous way. We shall use the bivariate Khinchin process to determine the key result in the paper via application of the Laplace-Stieltjes transform.

B. Factorial moments and factorial cumulants

The *n-th order factorial moment measure* $\nu_\Phi^{(n)}$ of a point process Φ are the measures on $\mathcal{X}^n$ such that [32]

$$\nu_\Phi^{(n)}(B_1\times\cdots\times B_n) = \mathbb{E}\left[\sum_{x_1,\dots,x_n\in\Phi}^{\neq} \mathbb{1}_{B_1}(x_1)\dots\mathbb{1}_{B_n}(x_n)\right], \quad (6)$$

for any regions $B_i \in \mathcal{B}(\mathcal{X})$, $1 \leq i \leq n$. The notation $\mathbb{1}_B$ denotes the indicator function, i.e., $\mathbb{1}_B(x) = 1$ if $x \in B$, and zero otherwise. Factorial moments and cumulants can be determined from the following derivatives using derivatives as described in [38].

$$\nu_\Phi^{(n)}(B_1\times\cdots\times B_n) = \delta^n \mathcal{G}_\Phi(h; \mathbb{1}_{B_1},\dots,\mathbb{1}_{B_n})|_{h=1}. \quad (7)$$

$$c_\Phi^{(n)}(B_1\times\cdots\times B_n) = \delta^n \mathcal{C}_\Phi(h; \mathbb{1}_{B_1},\dots,\mathbb{1}_{B_n})|_{h=1}. \quad (8)$$

Setting the directions to be indicator functions $\mathbb{1}_{B_1},\dots,\mathbb{1}_{B_n}$ and evaluating at $h = 1$, noting that $\mathcal{G}_\Phi(1) = 1$, we find expression for the factorial cumulants in terms of factorial moments,

$$\begin{aligned} c_\Phi^{(n)}(B_1\times\cdots\times B_n) &= \delta^n(\ln\circ\mathcal{G}_\Phi)(h; \mathbb{1}_{B_1},\dots,\mathbb{1}_{B_n})|_{h=1} \\ &= \sum_{\pi\in\Pi(B_1,\dots,B_n)} (-1)^{|\pi|-1}(|\pi|-1)! \prod_{\omega\in\pi} \nu_\Phi^{(|\omega|)}(\omega), \end{aligned} \quad (9)$$

where $\nu_\Phi^{(|\omega|)}(\omega)$ is taken to mean that the factorial moment is evaluated on the product of terms in ω, which are subsets of partition π, in the set of all partitions of the $\Pi(B_1,\dots,B_n)$ of elements $B_1,\dots,B_n$. This result was given in [30, p202] though without explicitly determining via differentiation. Hence, we have the relation for n^{th}-order factorial cumulants in terms of factorial moments determined via derivatives of the generating functional. This will be used to determine the results for different point process parameterisations. The inverse relation is found to be

$$\nu_\Phi^{(n)}(B_1\times\cdots\times B_n) = \sum_{\pi\in\Pi(B_1,\dots,B_n)} \prod_{\omega\in\pi} c_\Phi^{(|\omega|)}(\omega), \quad (10)$$

where $\mathcal{C}_\Phi(1) = 0$. Note that the second-order cumulant is related to the mean and variance with

$$c_\Phi^{(2)}(B,B) = \mathrm{var}_\Phi(B) - \mu_\Phi(B), \quad (11)$$

which will be important for filter assessment.

III. Point process models

This section describes the models that will be used to form approximations in the point process prior used before the update. In particular, we describe the Poisson point process and determine the over-dispersed negative-binomial point process via application of the Laplace-Stieljes transform. We then describe the relation between the negative-binomial process to the Panjer point process [29] which encompasses the negative-binomial, Poisson and binomial processes within the same model.

A. Poisson point process

A Poisson point process with parameter λ and spatial distribution s is a process with spatial distribution s, whose size is Poisson distributed with rate λ. Its Probability Generating Functional (p.g.fl.) is given by

$$\mathcal{G}_{\mathrm{Poisson}}(h) = \exp\left(\int [h(x) - 1]\, \mu(\mathrm{d}x)\right), \quad (12)$$

where the intensity measure μ of the process is such that $\mu(\mathrm{d}x) = \lambda s(\mathrm{d}x)$. It can be shown that the first-order factorial moment and the variance of a Poisson process are equal when evaluated in any region $B \in \mathcal{B}(\mathcal{X})$, i.e., $\mu_\Phi(B) = \mathrm{var}_\Phi(B)$. In other words, the random variable describing the number of objects within B has its mean equal to its variance. The factorial cumulant generating functional (f.g.fl.) is then

$$\mathcal{C}_{\mathrm{Poisson}}(h) = \int (h(x) - 1)\, \mu(\mathrm{d}x) \quad (13)$$

This is a linear, functional so computing cumulants are zero for orders greater than 1, and the first-order cumulant is equal to the first-order factorial moment, i.e.

$$c_\Phi^{(1)}(B) = \mu(B). \quad (14)$$

B. The Laplace-Stieljes transform and the negative binomial point process

This section follows the work by Bates and Neyman [34] on accident proneness, with adaptation from probability generating functions to probability generating functionals. Let us consider the multi-variate probability generating functional describing a mixture of n Poisson processes that are conditionally independent on the same parameter λ, i.e.

$$\mathcal{G}\left(h_1,\ldots,h_n|\lambda\right) = \exp\left(\lambda\sum_{i=1}^{n} a_i \int \left[h_i(x)-1\right] s_i(dx)\right), \tag{15}$$

where, for instance, $\sum_{i=1}^{n} a_i = 1$. Then the unconditional process after margninalisation over random variable Λ is found via the following expectation

$$\begin{aligned}\mathcal{G}\left(h_1,\ \ldots,h_n\right) &= \mathbb{E}\left[\mathcal{G}\left(h_1,\ldots,h_n|\Lambda\right)\right] \qquad (16)\\ &= \int_0^{\infty} \exp\left(\lambda\sum_{i=1}^{n} a_i\int\left[h_i(x)-1\right]s_i(dx)\right)dF(\lambda)\\ &= \mathcal{L}^*\left\{\sum_{i=1}^{n} a_i\int\left[h_i(x)-1\right]s_i(dx)\right\}\end{aligned}$$

where the Laplace-Stieljes transform of a distribution $F(\lambda)$ is defined with

$$\mathcal{L}^*(t) = \int_0^{\infty} e^{-t\lambda} dF(\lambda). \tag{17}$$

If we take Λ to be Gamma distributed with $\alpha, \beta > 0$, i.e.

$$p_\Lambda(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \tag{18}$$

then the unconditional process becomes

$$\mathcal{G}\left(h_1,\ldots,h_n\right) = \left(1+\frac{1}{\beta}\sum_{i=1}^{n} a_i\int\left[1-h_i(x)\right]s_i(dx)\right)^{-\alpha}. \tag{19}$$

This can be viewed as a kind of multi-variate negative binomial distribution. Univariate and bivariate instances of this formula can be found by restricting the number of terms in the summation to be one or two. The reason for introducing this approach here is that the same reasoning will be applied when considering a bivariate p.g.fl. which has an exponential form constructed with Poisson processes to determine an alternative bivariate p.g.fl. that is able to retain second-order information.

The following section describes the Panjer point process [29], based on the Panjer distribution [35], which extends the negative binomial by considering negative α and β.

C. Panjer process

A Panjer point process is a process whose size is Panjer distributed with parameters α and β with spatial distribution s [29], [35]. For finite and positive α and β, the Panjer distribution describes a negative binomial distribution. For finite and negative α and β we obtain a binomial distribution. The limit case $\alpha, \beta \to \infty$ with constant ratio $\lambda = \frac{\alpha}{\beta}$ yields a Poisson process with parameter λ [29], [36]. The p.g.fl. of a Panjer process with parameters α, β is given by [29]

$$\mathcal{G}_{\mathrm{Panjer}}(h) = \left(1+\frac{1}{\beta}\int[1-h(x)]s(\mathrm{d}x)\right)^{-\alpha}. \tag{20}$$

The f.g.fl. of the Panjer process becomes,

$$\mathcal{C}_{\mathrm{Panjer}}(h) = -\alpha\ln\left(1+\frac{1}{\beta}\int[1-h(x)]s(\mathrm{d}x)\right), \tag{21}$$

which takes a very similar form to the f.g.fl. of the Bernoulli process. Hence, the cumulants become

$$c^{(n)}_{\mathrm{Panjer}}(B_1\times\cdots\times B_n) = (n-1)!\frac{\alpha}{\beta^n}s(B_1)\ldots s(B_n). \tag{22}$$

Thus when $B_1 = B_2 = \mathcal{X}$,

$$c^{(1)}_{\mathrm{Panjer}} = \frac{\alpha}{\beta}, \qquad c^{(2)}_{\mathrm{Panjer}} = \frac{\alpha}{\beta^2}, \tag{23}$$

and hence

$$\alpha = \frac{{c^{(1)}_{\mathrm{Panjer}}}^2}{c^{(2)}_{\mathrm{Panjer}}}, \qquad \beta = \frac{c^{(1)}_{\mathrm{Panjer}}}{c^{(2)}_{\mathrm{Panjer}}}. \tag{24}$$

IV. Algorithm Specification

In this section, the new linear-complexity filter and it assumptions are presented. The multi-target model and general assumptions are based on the work [19]. The key difference is that correlations are maintained into the bivariate probability generating functional through the insight in Bates and Neyman [34] that dependencies can be introduced in functionals that take an exponential form through the application of the Laplace-Stieltjes transform. The proofs are given in the Appendix and the algorithm pseudo-code is described in Algorithm 1.

A. Time prediction step (time k)

In the time prediction step, the posterior target process Φ_{k-1} is predicted to $\Phi_{k|k-1}$ based on prior knowledge on the dynamical behaviour of the targets. The assumptions of the time prediction step can be stated as follows:

Assumptions IV.1.

(a) The targets evolve independently from each other;
(b) A target with state $x \in \mathcal{X}$ at time $k-1$ survived to the current time k with probability $p_{\mathrm{s},k}(x)$; if it did so, its state evolved according to a Markov transition kernel $t_{k|k-1}(\cdot|x)$;
(c) New targets entered the scene between time $k-1$ and k, independently of the existing targets and described by a newborn point process $\Phi_{\mathrm{b},k}$ with p.g.fl. $\mathcal{G}_{\mathrm{b},k}$.

The following theorem describes the prediction of the n^{th}-order cumulant.

Theorem IV.2 (Factorial cumulant prediction)**.** *Under Assumptions IV.1, the n^{th}-order factorial cumulant of the predicted target process $\Phi_{k|k-1}$ is given by*

$$c^{(n)}_{k|k-1}(B_1\times\ldots\times B_n) = c^{(n)}_{\mathrm{b},k}(B_1\times\ldots\times B_n)+c^{(n)}_{\mathrm{s},k}(B_1\times\ldots\times B_n), \tag{25}$$

in any $B_i \in \mathcal{B}(\mathcal{X})$, where $c^{(n)}_{\mathrm{s},k}$ is the n^{th}-order factorial cumulant of the process describing the surviving targets

$$c^{(n)}_{\mathrm{s},k}(B_1 \times \ldots \times B_n) = \int \prod_{i=1}^{n} p_{\mathrm{s},k}(x_i) t_{k|k-1}(B_i|x_i) c^{(n)}_{k-1}(\mathrm{d}(x_1, \ldots, x_n)). \tag{26}$$

In the algorithm factorial cumulants are computed as follows. The first-order factorial cumulant density $c^{(1)}_{k|k-1}(x)$ is the same as prediction of the first-order factorial moment density as calculated in [19]. The predicted second-order cumulant is a scalar computed over the whole state space, i.e. $c^{(2)}_{k-1}(\mathcal{X} \times \mathcal{X})$.

B. Data update step (time k)

In the data update step, the predicted process $\Phi_{k|k-1}$ is updated to Φ_k given the current measurement set Z_k, collected from the sensor. This is achieved through the following assumptions:

Assumptions IV.3.

(a) The bivariate target-measurement process is approximated by a Panjer process determined via the Laplace-Stieltjes transform of a bivariate Khinchin process linear in both arguments.

(b) A target with state $x \in \mathcal{X}$ is detected with probability $p_{\mathrm{d},k}(x)$; if so, it produces a measurement whose state is distributed according to a likelihood $l_k(\cdot|x)$.

(c) The clutter process is described by intensity function $\lambda_{c,k}(z)$, and second-order cumulant $c^{(2)}_{c,k}(\mathcal{Z}, \mathcal{Z})$.

Theorem IV.4 (Factorial cumulant update)**.** *Under Assumptions IV.3, the n^{th}-order cumulant of the updated target process Φ_k is given by*

$$c^{(n)}_k(B_1 \times \ldots \times B_n) = (n-1)! \left(\frac{(\alpha_{k|k-1} + |Z|)\mu^{\phi}_k(B_1) \ldots \mu^{\phi}_k(B_n)}{\left(\alpha_{k|k-1} + \mu^d_k(\mathcal{X}, \mathcal{Z}) + \lambda_{c,k}(\mathcal{Z})\right)^n} + (-1)^{n-1} \sum_{z \in Z} \frac{\mu^z_k(B_1) \ldots \mu^z_k(B_n)}{\left(\int_{\mathcal{X}} p_d(x) l(z|x) c^{(1)}_{k|k-1}(\mathrm{d}x) + \lambda_{c,k}(z)\right)^n} \right) \tag{27}$$

in any $B_1, \ldots, B_n \in \mathcal{B}(\mathcal{X})$, where the missed detection term μ^{ϕ}_k is given by

$$\mu^{\phi}_k(B) = \int_B (1 - p_{\mathrm{d},k}(x)) c^{(1)}_{k|k-1}(\mathrm{d}x). \tag{28}$$

and where we have the following terms relating to detection statistics

$$\mu^z_k(B) = \int_B p_{\mathrm{d},k}(x) l(z|x) c^{(1)}_{k|k-1}(\mathrm{d}x) \tag{29}$$

$$\mu^d_k(\mathcal{X}, \mathcal{Z}) = \int_{\mathcal{X}} \int_{\mathcal{Z}} p_{\mathrm{d},k}(x) l(dz|x) c^{(1)}_{k|k-1}(\mathrm{d}x) \tag{30}$$

and the parameter $\alpha_{k|k-1}$ is computed with

$$\alpha_{k|k-1} = \frac{\left(\mu^d_{k|k-1}(\mathcal{X}, \mathcal{Z}) + \lambda_{c,k}(\mathcal{Z})\right)^2}{c^{(2)}_{k|k-1} + c^{(2)}_{c,k}}. \tag{31}$$

Similar to the prediction step, the first-order cumulant density $c^{(1)}_k(x)$ is calculated along with the second-order factorial cumulant computed over the whole space, $c^{(2)}_k(\mathcal{X} \times \mathcal{X})$.

Implementation issues

A closed-form solution to the algorithm developed is presented in Algorithm 1 in the appendix. It is based on the solutions developed for PHD filter [19], CPHD filter [20], and second-order PHD filter [29], eg. [21], [22], [29]. The filter is the same complexity as the PHD filter with target-number variance [37], if the variance is computed over the entire state space. Compared with the Gaussian mixture PHD filter [21], with linear-complexity analysis presented in Section III.C of that work, in the prediction step, there are the following additional calculation to compute the predicted second-order cumulant,

$$c^{(2)}_{k|k-1} = p^2_{s,k} c^{(2)}_{k-1} + c^{(2)}_{b,k} \tag{32}$$

$$\alpha_{k|k-1} = (c^{(1)}_{k|k-1} + \lambda_{c,k})^2 / (c^{(2)}_{k|k-1} + c^{(2)}_{c,k}) \tag{33}$$

and in the update, there is additional calculation of the following terms defined in Algorithm 1,

$$c^{(2)}_k = (\mu^{\phi}_k)^2 l_2(\phi) - \sum_{z \in Z_k} \left(\frac{\mu^z_k}{\mu^z_k + \lambda_{\mathrm{c},k}(z)} \right)^2, \tag{34}$$

$$\ell_1(\phi) = (\alpha_{k|k-1} + |Z|)/(\alpha_{k|k-1} + \mu^d_k) \tag{35}$$

$$\ell_2(\phi) = (\alpha_{k|k-1} + |Z|)/(\alpha_{k|k-1} + \mu^d_k + \lambda_{c,k})^2. \tag{36}$$

V. Numerical experiments

In this section, we present experimental results for an illustrative example, under parametric evaluations that vary either the clutter model, number of targets, average number of false alarms per frame, or probability of detection. These parametric evaluations intend to show differences of performance of the PHD, CPHD and linear-complexity cumulant-based (LC-Cumulant) filters for different ranges of settings.

Let us consider a two-dimensional scenario within the region $[-1000, +1000] \times [-1000, +1000]$ (m $\times$ m). Each target is described by its state vector $x = (p_x, p_y, v_x, v_y)^{\mathrm{T}}$, where (p_x, p_y) is a pair that specifies a position in Cartesian coordinates and (v_x, v_y) is the pair specifying velocity in the same coordinates. Each target moves at nearly-constant velocity, with transition matrix and state process covariance matrix given respectively by

$$F = \begin{pmatrix} \mathbb{I}_2 & \mathbb{I}_2 \Delta t \\ 0_2 & \mathbb{I}_2 \end{pmatrix}, \; Q = \begin{pmatrix} \mathbb{I}_2 \Delta t^3/3 & \mathbb{I}_2 \Delta t^2/2 \\ \mathbb{I}_2 \Delta t^2/2 & \mathbb{I}_2 \Delta t \end{pmatrix} \sigma^2_q,$$

where $\mathbb{I}_2$ and 0_2 are the identity and zero matrices with dimensions 2×2, $\Delta t = 1\,\mathrm{s}$ is the sampling period, and the standard deviation of velocity increments is characterized by $\sigma_q = 1\,\mathrm{m/s}^{\frac{3}{2}}$. Probability of survival is set as $p_s = 0.99$. A sensor collects Cartesian position measurements characterized by the output matrix and measurement noise covariance matrix, $H = \begin{pmatrix} \mathbb{I}_2 & 0_2 \end{pmatrix}$ and $R = \mathbb{I}_2 \sigma^2_r$, respectively, where $\sigma_r = 10\,\mathrm{m}$ is the standard deviation of the measured positions. False alarms can happen according to a Poisson point-process with intensity $\lambda_{c,k}(z) = \lambda_c \cdot s_c(z)$, where λ_c is the average number of false alarms per scan, and $s_c(z)$ is the spatial distribution of clutter, assumed uniform in the surveillance region with "volume" $V = 2000^2\,\text{m}^2$. The example is simulated for $T = 100\,\text{s}$. Targets appear in batches at positions uniformly sampled within the area $[-800, +800] \times [-800, +800]$ (m×m), with random velocities uniformly sampled within the ranges $[-10, +10] \times [-10, +10]$ (m/s × m/s).

By denoting N_t as the total number of targets that appear in the scene, the batches of target appearance are set as follows:

- $0.25N_t$ targets are already in the scene at $t = 0$ and will remain up to $t = 100\,\text{s}$ with exception of 5 targets that are set to disappear at $t = 80\,\text{s}$,
- $0.25N_t$ targets appear at $t = 20\,\text{s}$ and remain,
- $0.25N_t$ of targetsappear at $t = 40\,\text{s}$ and remain,
- $0.25N_t$ of targets appear at $t = 60\,\text{s}$ and remain,
- from N_t targets in the scene, 5 targets disappear at $t = 80\,\text{s}$ and the remaining $N_t - 5$ stay up to $t = 100\,\text{s}$.

The birth process is a Poisson point-process with intensity function $\mu_b(x) = \sum_{i=1}^{4} w_b^{(i)} \mathcal{N}(x; m_b^{(i)}, P_b^{(i)})$, where $w_b^{(i)} = N_t/(4T/\Delta t)$, $m_b^{(1)} = (-500, -500, 0, 0)^{\mathrm{T}}$, $m_b^{(2)} = (-500, +500, 0, 0)^{\mathrm{T}}$, $m_b^{(3)} = (+500, -500, 0, 0)^{\mathrm{T}}$, $m_b^{(4)} = (+500, +500, 0, 0)^{\mathrm{T}}$, $P_b^{(i)} = \mathrm{diag}(500^2\mathbb{I}_2, 10^2\mathbb{I}_2)$, for $i = [1..4]$, where $\mathrm{diag}(A, B)$ is a block diagnoal matrix formed whose blocks are the matrices A and B.

For all filters, pruning of Gaussian components is based on the weight threshold $\tau_{\text{prn}} = 10^{-5}$, merging is performed with threshold of $\tau_{\text{mrg}} = 4\,\text{m}$, and the number of maintained components is limited at $\text{J}_{\max} = 100$ (see [21] for details on the pruning and merging procedure). Measurements are gated with gate-size probability of $p_{\text{gate}} = 0.999$. The cardinality distribution for the CPHD filter is estimated to a maximum of $n_{\max} = 2N_t$ terms. This maximum number of cardinality terms has been chosen to keep the CPHD filter computational effort competitive in relation to the other filters for difficult scenarios. The LC Cumulant filter is evaluated in comparison to the PHD and CPHD filters for four different cases.

Case 1: For $N_t = 50$ targets that appear in the scene, $p_d = 0.80$, $\lambda_c = 10$ number of false alarms per scan on average, all filters are tested for three different clutter models: a) Poisson process with $\mu_c = \lambda_c = 10$ and $\mathrm{var}_c = 10$, b) binomial process with $\mu_c = \lambda_c = 10$ and $\mathrm{var}_c = \lambda_c/20 = 0.5$, c) negative binomial process with $\mu_c = \lambda_c = 10$ and $\mathrm{var}_c = 20\lambda_c = 200$.

Case 2: For $p_d = 0.90$, $\lambda_c = 10$ false alarms per scan on average (Poisson distribution), filters are tested for different numbers of targets that appear in the scene, $N_t \in \{10, 20, 30, 40, 50\}$.

Case 3: For $N_t = 20$ targets, $\lambda_c = 10$ false alarms per scan on average (Poisson distribution), filters are tested for different probabilities of detection, $p_d \in \{0.60, 0.70, 0.80, 0.90, 0.99\}$.

Case 4: For $N_t = 20$ targets, $p_d = 0.90$, filters are tested for different numbers of false alarms per frame (Poisson distribution), $\lambda_c \in \{10, 20, 30, 40, 50\}$.

For each case, 200 Monte Carlo (MC) runs are performed, each with independently generated clutter, and independently generated (target-originated) measurements for each trial. For all filters, performance is evaluated in terms of:

- mean Optimal Subpattern Assignment (OSPA) metric for cut-off $c_{\text{OSPA}} = 100$ and norm order $p_{\text{OSPA}} = 1$,
- root-mean-square error (RMSE) of the estimated number of targets, and
- computation time (per time step).

All indexes are averaged over time steps and consolidated for all values of the varying parameters.

A. Results

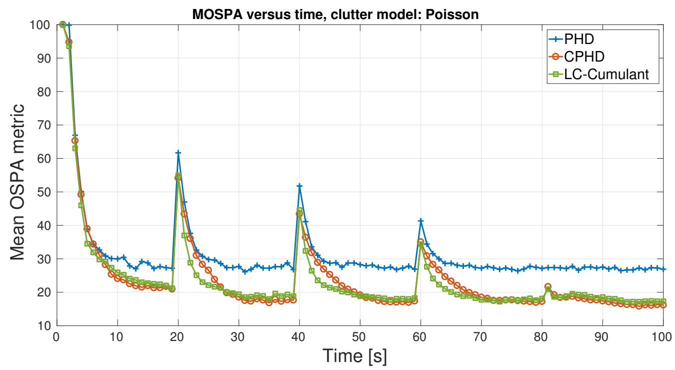
Case 1 : For this case, Figures 1a–3b present, the mean OSPA over time, and the cardinality mean and standard deviation over time for the PHD, CPHD and LC Cumulant filters, where we can perceive the advantage of estimating second-order information on the target number. The LC Cumulant filter maintains second-order information about the target number via the second-order factorial cumulant.

It is clear from the figures that, the performance of the LC Cumulant filter is similar to that of the CPHD filter, but at a computational cost that is practically the same as that of the PHD filter. In all three subcases, with Poisson, binomial and negative-binomial clutter models, the PHD filter underestimates the correct number of targets due to the difficulties imposed by a relatively low probability of detection $p_d = 0.80$, non-Poisson clutter, and closely spaced targets.

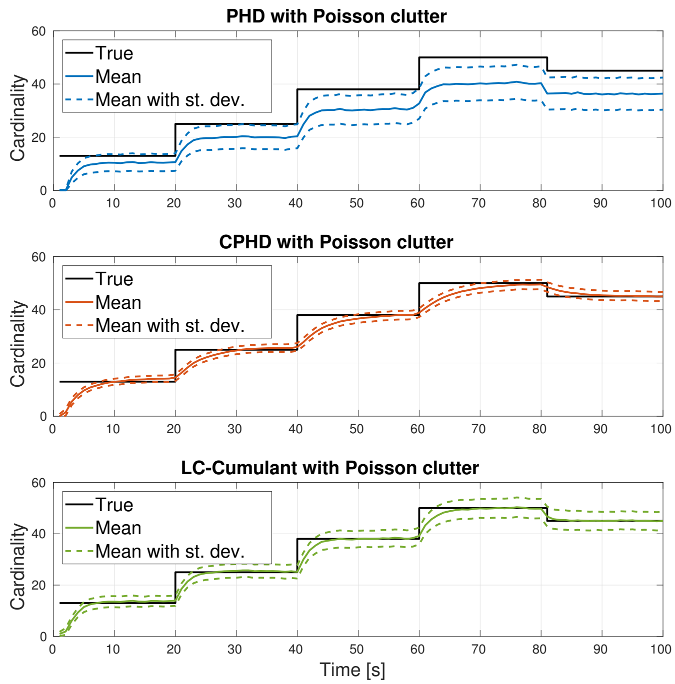
Case 2 : The consolidated performance indexes for case 2, averaged over all time steps for different numbers of targets, $N_t \in \{10, 20, 30, 40, 50\}$, are shown in Figures 4–6. In general, the performance of the LC Cumulant filter is not much different from that of the CPHD filter, but at a lower computational complexity. In terms of average mean OSPA metric, the LC Cumulant filter shows a performance that approaches that of the CPHD filter as the number of targets increases owing to that the LC Cumulant filter seems less sensitive (on average) to the increase in target number. The average cardinality RMSE of the CPHD and LC Cumulant filter seems to increase sub-exponentially with the number of targets, at a small rate than that of the PHD, which is more sensitive to the scenario complexity.

Case 3 : The consolidated performance indexes for case 3, averaged over time for $p_d \in \{0.60, 0.70, 0.80, 0.90, 0.99\}$, are presented in Figures 7–9. Once again, the performance of the LC Cumulant filter is very similar to that of the CPHD filter in terms of average mean OSPA and average cardinality RMSE, but its computational cost is much smaller than that of the CPHD filter, being rather comparable to that of the PHD filter. As expected, the error indexes and cardinality variance decrease as the probability of detection increases, for all filters.

Case 4 : Figures 10–12 present the consolidated performance indexes for case 4, averaged over time for each $\lambda \in \{10, 20, 30, 40, 50\}$. Corroborating with the previous cases, the average cardinality RMSE and average MOSPA of the LC Cumulant filter is close to that of the CPHD

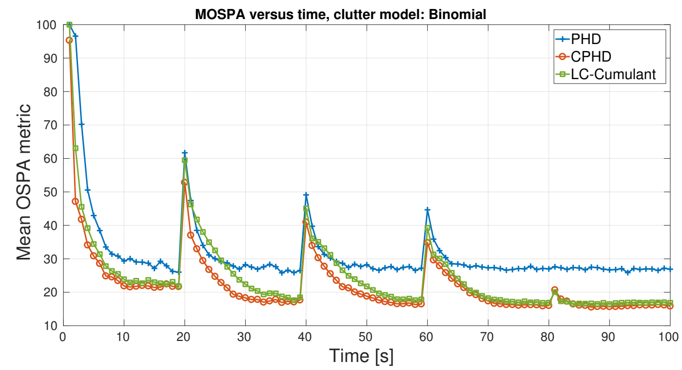


(a) MOSPA vs. time

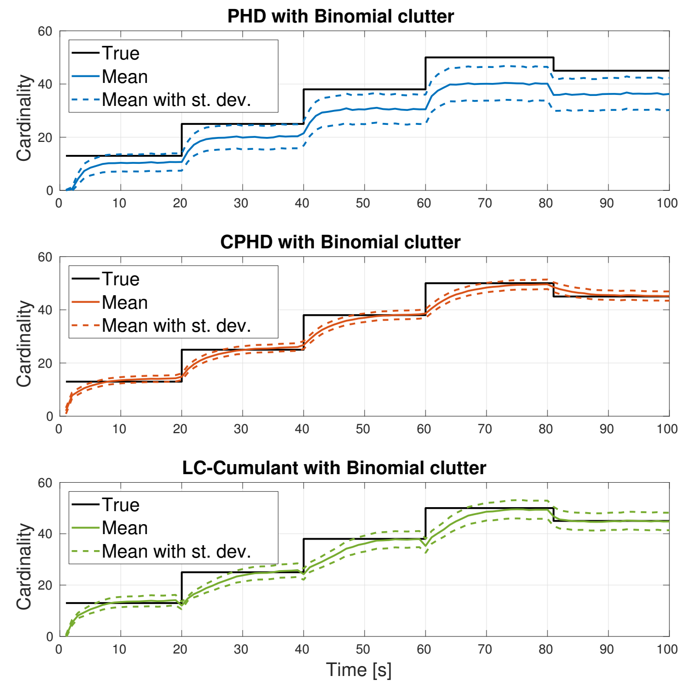


(b) Cardinality vs. time

Fig. 1: Case 1a: MOSPA and cardinality, Poisson clutter



(a) MOSPA vs. time



(b) Cardinality vs. time

Fig. 2: Case 1b: MOSPA and cardinality, binomial clutter

filter overall, with marginally better performance for higher numbers of false alarms. In this case, all filters take almost the same time to perform the computations, but it remains clear that the LC Cumulant filter requires less computational effort than the standard CPHD filter implementation, presenting runtimes that are comparable to the PHD filter. Note that the average cardinality RMSE of all filters seem to increase sub-exponentially with the number of false alarms per frame, suggesting a dependency of the signal-to-noise ratio that is polynomial in the number of measurements.

Appendix: Proofs

B. Proof of Thm. IV.2

In the following, we denote by $\mathcal{G}_{\mathrm{s},k}$ the p.g.fl. of the point process describing the evolution of a target from the previous time step, which might have survived (or not) to the present time step. The p.g.fl. $\mathcal{G}_{k|k-1}$ of the predicted target process takes the form [19]

$$\mathcal{G}_{k|k-1}(h) = \mathcal{G}_{\mathrm{b}}(h)\mathcal{G}_{k-1}(\mathcal{G}_{\mathrm{s}}(h|\cdot)), \tag{37}$$

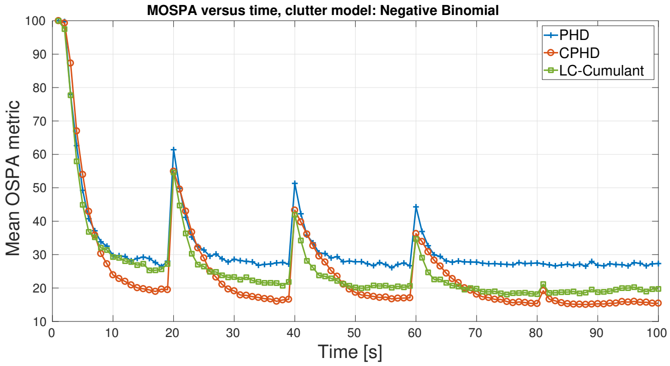
hence the c.g.fl. $\mathcal{C}_{k|k-1}$ of the predicted target process is

$$\begin{aligned} \mathcal{C}_{k|k-1}(h) &= \ln\left(\mathcal{G}_{\mathrm{b}}(h)\mathcal{G}_{k-1}(\mathcal{G}_{\mathrm{s}}(h|\cdot))\right) \\ &= \ln\left(\mathcal{G}_{\mathrm{b}}(h)\right) + \ln\left(\mathcal{G}_{k-1}(\mathcal{G}_{\mathrm{s}}(h|\cdot))\right) \\ &= \mathcal{C}_{\mathrm{b}}(h) + \mathcal{C}_{k-1}(\mathcal{G}_{\mathrm{s}}(h|\cdot)). \end{aligned} \tag{38}$$

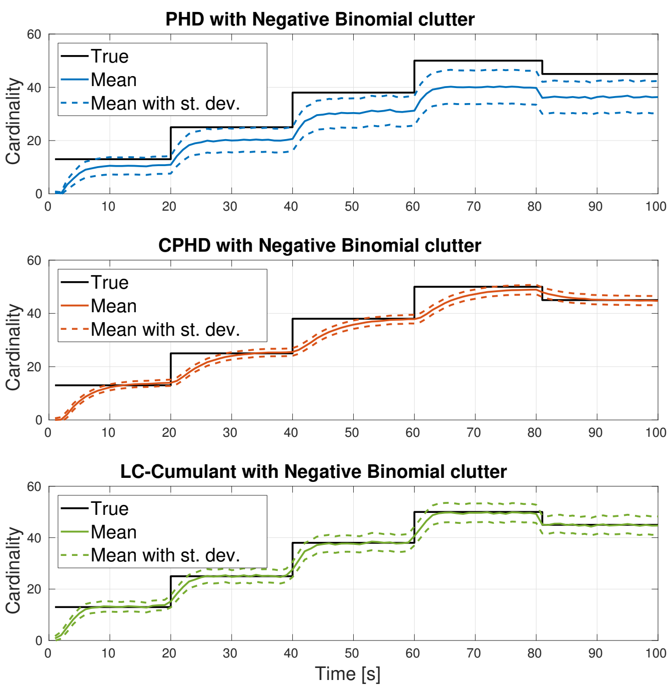
The n^{th}-order factorial cumulant requires the n^{th}-order derivative of $\mathcal{C}_{k|k-1}(h)$, i.e.

$$c^{(2)}_{k|k-1}(B_1 \times \ldots \times B_n) = \delta^n \mathcal{C}_{k|k-1}(h; \mathbb{1}_{B1}, \ldots, \mathbb{1}_{B_n})\big|_{h=1}. \tag{39}$$

The survival process for a target with state x at the previous time step can be described with a Bernoulli point process with



(a) MOSPA vs. time



(b) Cardinality vs. time

Fig. 3: Case 1c: MOSPA and cardinality, negative-binomial clutter

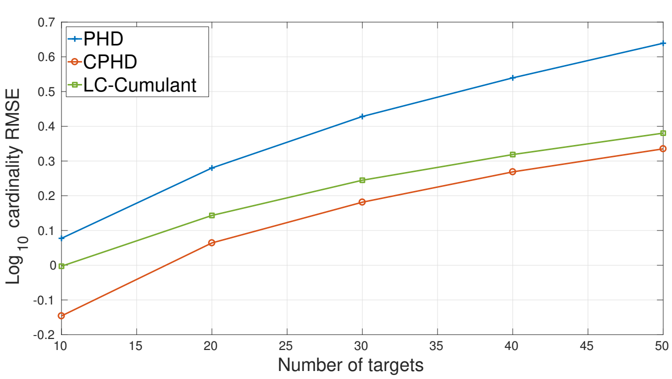


Fig. 4: Case 2: average cardinality RMSE vs. number of targets

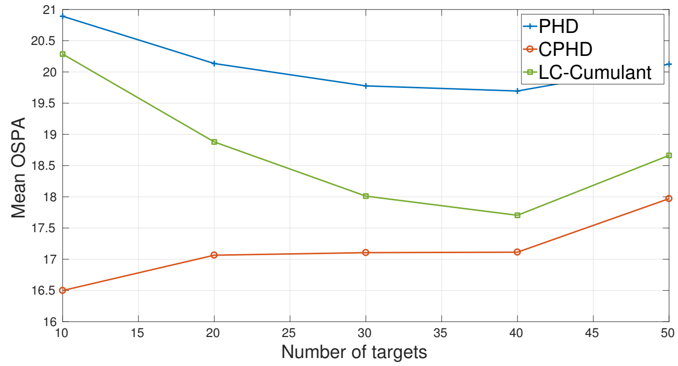


Fig. 5: Case 2: average mean OSPA vs. number of targets

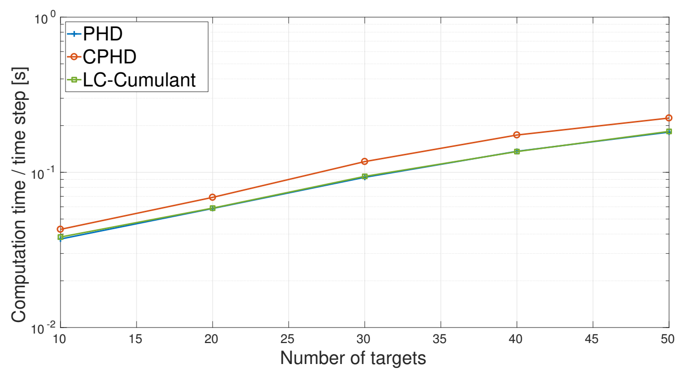


Fig. 6: Case 2: average runtime vs. number of targets

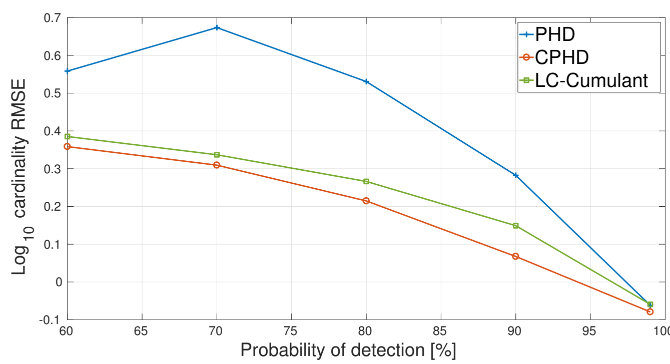


Fig. 7: Case 3: average cardinality RMSE vs. prob. of detection

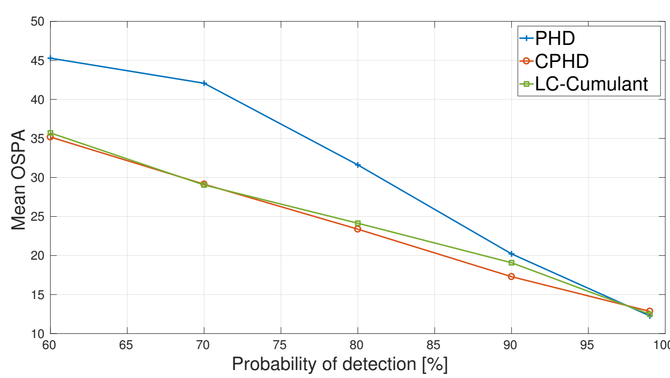


Fig. 8: Case 3: average mean OSPA vs. prob. of detection

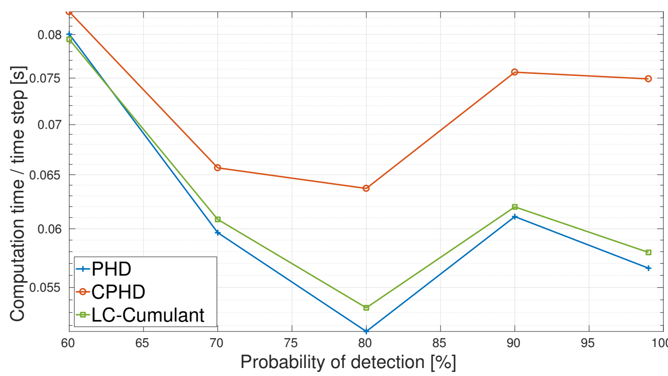


Fig. 9: Case 3: average runtime vs. prob. of detection

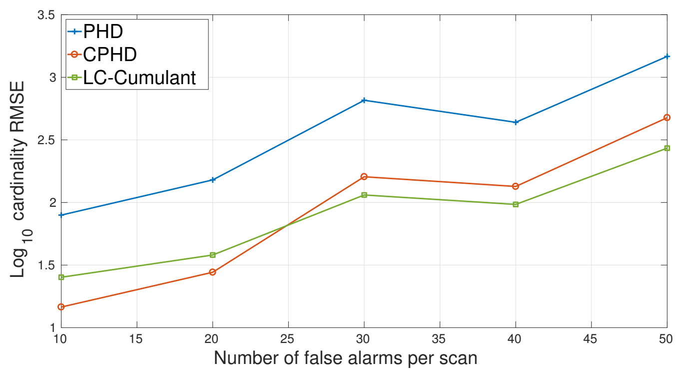


Fig. 10: Case 4: average cardinality RMSE vs. number of false alarms

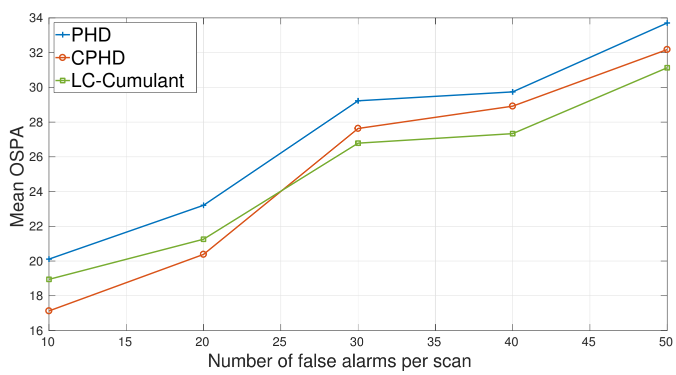


Fig. 11: Case 4: average mean OSPA vs. number of false alarms

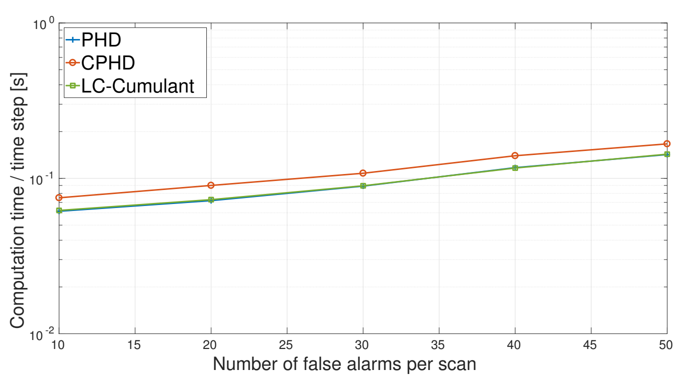


Fig. 12: Case 4: average runtime vs. number of false alarms

parameter $p_{\mathrm{s}}(x)$ and spatial distribution $t(\cdot|x)$,

$$\mathcal{G}_{\mathrm{s}}(h|x) = 1 - p_{\mathrm{s}}(x) + p_{\mathrm{s}}(x) \int h(y) t(\mathrm{d}y|x). \tag{40}$$

It follows that

$$\begin{aligned} \delta(\mathcal{G}_{\mathrm{s}}(h|x); \mathbb{1}_B)\big|_{h=1} &= p_{\mathrm{s}}(x) \int \delta(hy; \mathbb{1}_B) t(\mathrm{d}y|x)\big|_{h=1} \\ &= p_{\mathrm{s}}(x) \int \mathbb{1}_B(y) h(y) t(\mathrm{d}y|x)\big|_{h=1}, \\ &= p_{\mathrm{s}}(x) t(B|x). \end{aligned} \tag{41}$$

The n^{th}-order derivative $\delta^n \mathcal{G}_{\mathrm{s}}(h|\cdot; \mathbb{1}_{B1}, \ldots, \mathbb{1}_{Bn}) = 0$ for second-order and above since $\mathcal{G}_{\mathrm{s}}(h)$ is linear in h. Hence the result follows by Faà di Bruno's formula [38].

C. *Proof of Thm.IV.4*

Now consider the f.g.fl. of a bivariate Khinchin process [33], $\mathcal{C}_{\mathrm{J},k}(g,h)$, i.e.

$$\begin{aligned} \mathcal{C}_{\mathrm{J},k}(g,h) &= \mathcal{C}_{\mathrm{J},k}(0,0) + \mathcal{C}_{\mathrm{J},k}(0,h) + \mathcal{C}_{\mathrm{J},k}(g,0) + \\ &\sum_{n\geq 1, m\geq 1} \int \prod_{j}^{m} g(z_j) \prod_{i}^{n} h(x_i) K_{\mathrm{J},k}^{(n,m)}(dx_{1:n}, dz_{1:m}), \end{aligned} \tag{42}$$

where the zero terms are calculated with

$$\mathcal{C}_{\mathrm{J},k}(0,0) = - \sum_{n,m\geq 0 | n+m\geq 1} \int K_{\mathrm{J},k}^{(n,m)}(dx_{1:n}, dz_{1:m}), \tag{43}$$

$$\mathcal{C}_{\mathrm{J},k}(0,h) = \sum_{m\geq 1} \int h(x_1) \ldots h(x_n) K_{\mathrm{J},k}^{(n,0)}(dx_{1:n}), \tag{44}$$

$$\mathcal{C}_{\mathrm{J},k}(g,0) = \sum_{n\geq 1} \int g(z_1) \ldots g(z_m) K_{\mathrm{J},k}^{(0,m)}(dz_{1:m}). \tag{45}$$

We can determine an alternative bivariate p.g.fl. through the Laplace-Stieltjes transform i.e.

$$\begin{aligned} \mathcal{G}_{\mathrm{J},k}(g,h) &= \mathcal{L}^* \left\{ \frac{\mathcal{C}_{\mathrm{J},k}(g,h)}{\mathcal{C}_{\mathrm{J},k}(0,0)} \right\} \\ &= \left(1 - \frac{\mathcal{C}_{\mathrm{J},k}(g,h)}{\beta_{k|k-1}\, \mathcal{C}_{\mathrm{J},k}(0,0)}\right)^{-\alpha_{k|k-1}}, \end{aligned} \tag{46}$$

The p.g.fl. of the updated target process Φ_k is obtained from the differentiation of the joint p.g.fl. using Bayes' rule [19] with the following

$$\mathcal{G}_k(h) = \frac{\delta^{|Z_k|} \mathcal{G}_{\mathrm{J},k}(g,h; (\delta_z)_{z\in Z_k})|_{g=0}}{\delta^{|Z_k|} \mathcal{G}_{\mathrm{J},k}(g,1; (\delta_z)_{z\in Z_k}))|_{g=0}}. \tag{47}$$

Noting that we can set $\alpha_{k|k-1} = -\beta_{k|k-1}\ \mathcal{C}_{\mathrm{J},k}(0,0)$, and given the assumption that $\mathcal{C}(g,h)$ is a linear functional in g which means that there is only one remaining partition after application of Faà di Bruno's formula [38], the updated p.g.fl. becomes

$$\mathcal{G}_k(h) = \left(\frac{\alpha_{k|k-1} + \mathcal{C}_{\mathrm{J},k}(0,h)}{\alpha_{k|k-1} + \mathcal{C}_{\mathrm{J},k}(0,1)}\right)^{-(\alpha_{k|k-1}+|Z|)} \prod_{z\in Z} \frac{\delta\mathcal{C}_{\mathrm{J},k}(0,h;\delta_z)}{\delta\mathcal{C}_{\mathrm{J},k}(0,1;\delta_z)} \tag{48}$$

and consequently, the f.g.fl. becomes

$$\mathcal{C}_k(h) = -(\alpha_{k|k-1} + |Z|) \ln \left(\frac{\alpha_{k|k-1} + \mathcal{C}_{\mathrm{J},k}(0,h)}{\alpha_{k|k-1} + \mathcal{C}_{\mathrm{J},k}(0,1)} \right) + \sum_{z \in Z} \ln \left(\frac{\delta \mathcal{C}_{J,k}(0,h;\delta_z)}{\delta \mathcal{C}_{J,k}(0,1;\delta_z)} \right). \tag{49}$$

Noting that the joint p.g.fl. in [19, p1174] is a bivariate Khinchin process [33] linear in g and h, with

$$\mathcal{C}_{\mathrm{J},k}(g,h) = \int_{\mathcal{Z}} [g(z) - 1]\, \lambda_{c,k}(\mathrm{d}z) + \int_{\mathcal{X}} \int_{\mathcal{Z}} [h(x)\, (p_d(x) l(dz|x) g(z) + q_d(x)) - 1]\, \mu_{k|k-1}(\mathrm{d}x), \tag{50}$$

we can use this directly in the derivation. Set $\lambda = -\mathcal{C}_{\mathrm{J},k}(0,0) = \mu_{k|k-1}(\mathcal{X}) + \lambda_{c,k}(\mathcal{Z})$ in the Laplace-Stieltjes transform. and hence the f.g.fl. becomes

$$\begin{aligned} &- (\alpha_{k|k-1} + |Z|) \\ &\times \ln \left(\frac{\alpha_{k|k-1} + \int_{\mathcal{X}} [h(x) q_d(x) - 1] \mu_{k|k-1}(\mathrm{d}x) + \lambda_{c,k}(\mathcal{Z})}{\alpha_{k|k-1} + \int_{\mathcal{X}} p_d(x) \mu_{k|k-1}(\mathrm{d}x) + \lambda_{c,k}(\mathcal{Z})} \right) \\ &+ \sum_{z \in Z} \ln \left(\frac{\int_{\mathcal{X}} h(x) p_d(x) l_k(z|x) \mu_{k|k-1}(\mathrm{d}x) + \lambda_{c,k}(z)}{\int_{\mathcal{X}} p_d(x) l(z|x) \mu_{k|k-1}(\mathrm{d}x) + \lambda_{c,k}(z)} \right), \end{aligned} \tag{51}$$

where $\alpha_{k|k-1}$ is defined with the intensity and second-order cumulant of the joint process,

$$\alpha_{k|k-1} = \frac{\left(\mu^d_{k|k-1}(\mathcal{X}, \mathcal{Z}) + \lambda_{c,k}(\mathcal{Z}) \right)^2}{c^{(2)}_{k|k-1} + c^{(2)}_{c,k}}. \tag{52}$$

Differentiating the f.g.fl. with respect to h and setting the argument to be equal to 1, leads to the factorial cumulants as stated.

References

[1] D. Reid, "An algorithm for tracking multiple targets," *IEEE Transactions on Automatic Control*, vol. 24, no. 6, pp. 843–854, Dec 1979.

[2] S. Coraluppi and C. Carthel, "Multi-stage multiple-hypothesis tracking." *J. Adv. Inf. Fusion*, vol. 6(1), pp. 56–67, 2011.

[3] R. L. Streit and T. E. Luginbuhl, "Probabilistic multi-hypothesis tracking," NAVAL UNDERWATER SYSTEMS CENTER NEWPORT RI, Tech. Rep., 1995.

[4] S. Blackman and R. Popoli, "Design and analysis of modern tracking systems(book)," *Norwood, MA: Artech House, 1999.*, 1999.

[5] T. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Sonar tracking of multiple targets using joint probabilistic data association," *IEEE Journal of Oceanic Engineering*, vol. 8, no. 3, pp. 173–184, Jul 1983.

[6] Y. Bar-Shalom, P. K. Willett, and X. Tian, *Tracking and data fusion*. YBS publishing, 2011.

[7] K. Kastella, "Discrimination gain for sensor management in multitarget detection and tracking," in *CESA'96 IMACS Multiconference: computational engineering in systems applications*, 1996, pp. 167–172.

[8] W. D. Blair and M. Brandt-Pearce, "Nnjpda for tracking closely spaced rayleigh targets with possibly merged measurements," in *Signal and Data Processing of Small Targets 1999*, vol. 3809. International Society for Optics and Photonics, 1999, pp. 396–409.

[9] S. Maskell, "Statistical methods for target tracking," *Wiley Encyclopedia of Computer Science and Engineering*, 2008.

[10] J. L. Williams, "Hybrid poisson and multi-bernoulli filters," in *Information Fusion (FUSION), 2012 15th International Conference on*. IEEE, 2012, pp. 1103–1110.

[11] J. Correa and M. Adams, "Estimating detection statistics within a bayes-closed multi-object filter," in *Information Fusion (FUSION), 2016 19th International Conference on*. IEEE, 2016, pp. 811–819.

Input
Posterior target process: $\{w^{(i)}_{k-1}, m^{(i)}_{k-1}, P^{(i)}_{k-1}\}_{i=1}^{N_{k-1}}, c^{(2)}_{k-1}$
Newborn process: $\{w^{(i)}_{\mathrm{b},k-1}, m^{(i)}_{\mathrm{b},k-1}, P^{(i)}_{\mathrm{b},k-1}\}_{i=1}^{N_{\mathrm{b},k-1}}, c^{(2)}_{\mathrm{b},k}$
Survival process
$\mathrm{c}^{(1)}_{k-1} = \sum_{i=1}^{N_{k-1}} w^{(i)}_{k-1}$
for $1 \le i \le n_{k-1}$ **do**
  $w^{(i)}_{k|k-1} = p_{\mathrm{s},k} w^{(i)}_{k-1}$
  $m^{(i)}_{k|k-1} = F_{k-1} m^{(i)}_{k-1}$
  $P^{(i)}_{k|k-1} = F_{k-1} P^{(i)}_{k-1} F^T_{k-1} + Q_{k-1}$
end for
$c^{(2)}_{\mathrm{s},k} = p^2_{\mathrm{s},k} c^{(2)}_{k-1}$
Newborn process
for $1 \le j \le N_{\mathrm{b},k-1}$ **do**
  $w^{(n_{k-1}+j)}_{k|k-1} = w^{(j)}_{\mathrm{b},k-1}$
  $m^{(n_{k-1}+j)}_{k|k-1} = m^{(j)}_{\mathrm{b},k-1}$
  $P^{(n_{k-1}+j)}_{k|k-1} = P^{(j)}_{\mathrm{b},k-1}$
end for
$N_{k|k-1} = N_{k-1} + N_{\mathrm{b},k-1}$
$c^{(2)}_{k|k-1} = c^{(2)}_{\mathrm{b},k} + c^{(2)}_{\mathrm{s},k}$
Predicted process: $\{w^{(i)}_{k|k-1}, m^{(i)}_{k|k-1}, P^{(i)}_{k|k-1}\}_{i=1}^{N_{k|k-1}}, c^{(2)}_{k|k-1}$
Current measurements: $Z_k = \{z_j\}_{j=1}^{M_k}$
Clutter process: $\lambda_{c,k}, c^{(2)}_{c,k}$
Panjer parameters
$c^{(1)}_{k|k-1} = \sum_{i=1}^{N_{k|k-1}} w^{(i)}_{k|k-1}$
$\alpha_{k|k-1} = (c^{(1)}_{k|k-1} + \lambda_{c,k})^2 / (c^{(2)}_{k|k-1} + c^{(2)}_{c,k})$
Missed detection and measurement terms
for $1 \le i \le N_{k|k-1}$ **do**
  $w^{(i)}_{\phi,k} = (1 - p_{\mathrm{d},k}) w^{(i)}_{k|k-1}$
  $m^{(i)}_{\phi,k} = m^{(i)}_{k|k-1}$
  $P^{(i)}_{\phi,k} = P^{(i)}_{k|k-1}$
end for
$\mu^\phi_k = (1 - p_{\mathrm{d},k}) c^{(1)}_{k|k-1}$
$\mu^d_k = p_{\mathrm{d},k} c^{(1)}_{k|k-1}$
for $1 \le j \le M_k$ **do**
  for $1 \le i \le N_{k|k-1}$ **do**
    $y^{(i,j)}_k = z_j - H_k m^{(i)}_{k|k-1}$
    $S^{(i)}_k = H_k P^{(i)}_{k|k-1} H^T_k + R_k$
    $K^{(i)}_k = P^{(i)}_{k|k-1} H^T_k [S^{(i)}_k]^{-1}$
    $w^{(i,j)}_{\mathrm{d},k} = p_{\mathrm{d},k} w^{(i,j)}_{\mathrm{d},k|k-1} \mathcal{N}(z; y^{(i,j)}_k, S^{(i)}_k) / (\mu^z_k + \lambda_{\mathrm{c},k})$
    $m^{(i,j)}_{\mathrm{d},k} = m^{(i)}_{k|k-1} + K^{(i)}_k y^{(i,j)}_k$
    $P^{(i,j)}_{\mathrm{d},k} = (I - K^{(i)}_k H_k) P^{(i)}_{k|k-1}$
  end for
  $\mu^{z_j}_k = \sum_{i=1}^{N_{k|k-1}} w^{(i,j)}_{\mathrm{d},k}$
end for
Corrective terms
$\ell_1(\phi) := (\alpha_{k|k-1} + |Z|) / (\alpha_{k|k-1} + \mu^d_k + \lambda_{\mathrm{c},k})$
$\ell_2(\phi) := (\alpha_{k|k-1} + |Z|) / (\alpha_{k|k-1} + \mu^d_k + \lambda_{\mathrm{c},k})^2$

Algorithm 1: Data update, first part: Compute single target - single measurement updates and corrective terms ℓ_1, ℓ_2

Missed detection terms
for $1 \leq i \leq N_{k|k-1}$ **do**
$w_k^{(i)} = \ell_1(\phi) w_{\phi,k}^{(i)}$
$m_k^{(i)} = m_{\phi,k}^{(i)}$
$P_k^{(i)} = P_{\phi,k}^{(i)}$
Association terms
for $1 \leq j \leq M_k$ **do**
$w_k^{(i \cdot n_{k|k-1}+j)} = w_{\mathrm{d},k}^{(i,j)}$
$m_k^{(i \cdot n_{k|k-1}+j)} = m_{\mathrm{d},k}^{(i,j)}$
$P_k^{(i \cdot n_{k|k-1}+j)} = P_{\mathrm{d},k}^{(i,j)}$
end for
end for
$N_k = N_{k|k-1} + N_{k|k-1} M_k$
$c_k^{(1)} = \sum_{i=1}^{N_k} w_k^{(i)}$

Second-order cumulant update
$c_k^{(2)} = (\mu_k^{\phi})^2 l_2(\phi) - \sum_{z \in Z_k} \left(\frac{\mu_k^z}{\mu_k^z + \lambda_{\mathrm{c},k}(z)} \right)^2$

Output
Posterior target process: $\{w_k^{(i)}, m_k^{(i)}, P_k^{(i)}\}_{i=1}^{N_k}, c_k^{(2)}$

Algorithm 1: Data update, second part: Update component weights.

[12] 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, 2013.

[13] R. B. Washburn, "A random point process approach to multiobject tracking," in *American Control Conference, 1987*. IEEE, 1987, pp. 1846–1852.

[14] S. Mori and C.-Y. Chong, "Point process formalism for multiple target tracking," in *Information Fusion, 2002. Proceedings of the Fifth International Conference on*, vol. 1. IEEE, 2002, pp. 10–17.

[15] K. Gilholm, S. Godsill, S. Maskell, and D. Salmond, "Poisson models for extended target and group tracking," in *Signal and Data Processing of Small Targets 2005*, vol. 5913. International Society for Optics and Photonics, 2005, p. 59130R.

[16] R. L. Streit, "Multiple target tracking," in *Poisson Point Processes*. Springer, 2010, pp. 147–178.

[17] R. Streit, C. Degen, and W. Koch, "The pointillist family of multitarget tracking filters," *arXiv preprint arXiv:1505.08000*, 2015.

[18] I. Goodman, R. Mahler, and H. T. Nguyen, *Mathematics of data fusion*. Springer, 1997.

[19] R. P. S. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 39, no. 4, pp. 1152–1178, 2003.

[20] ——, "PHD filters of higher order in target number," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 43, no. 4, pp. 1523–1543, Oct. 2007.

[21] B.-N. Vo and W.-K. Ma, "The gaussian mixture probability hypothesis density filter," *IEEE Transactions on signal processing*, vol. 54, no. 11, pp. 4091–4104, 2006.

[22] B.-T. Vo, B.-N. Vo, and A. Cantoni, "Analytic implementations of the cardinalized probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3553–3567, 2007.

[23] B.-N. Vo, S. S. Singh, and A. Doucet, "Sequential Monte Carlo methods for multitarget filtering with random finite sets," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 41, no. 4, pp. 1224–1245, 2005.

[24] R. Hoseinnezhad, B.-N. Vo, and B.-T. Vo, "Visual tracking in background subtracted image sequences via multi-bernoulli filtering," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 392–397, 2013.

[25] J. Mullane, B.-N. Vo, M. D. Adams, and B.-T. Vo, "A random-finite-set approach to bayesian SLAM," *IEEE Transactions on Robotics*, vol. 27, no. 2, pp. 268–282, 2011.

[26] R. Mahler, "PHD filters of second order in target number," in *Proc. SPIE Defense and Security Symposium*, vol. 6236. International Society for Optics and Photonics, 2006.

[27] S. S. Singh, B.-N. Vo, A. Baddeley, and S. Zuyev, "Filters for Spatial Point Processes," *SIAM Journal on Control and Optimization*, vol. 48, no. 4, pp. 2275–2295, 2009.

[28] D. Fränken, M. Schmidt, and M. Ulmke, "Spooky Action at a Distance in the Cardinalized Probability Hypothesis Density Filter," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 45, no. 4, pp. 1657–1664, Oct. 2009.

[29] I. Schlangen, E. D. Delande, J. Houssineau, and D. E. Clark, "A second-order phd filter with mean and variance in target number," *IEEE Transactions on Signal Processing*, vol. 66, no. 1, pp. 48–63, Jan 2018.

[30] D. J. Daley and D. Vere-Jones, *An introduction to the theory of point processes. vol. I. , Elementary theory and methods*, ser. Probability and its applications. New York, Berlin, Paris: Springer, 2003.

[31] J. Moyal, "The general theory of stochastic population processes," *Acta mathematica*, vol. 108, no. 1, pp. 1–31, 1962.

[32] D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic geometry and its applications*. John Wiley & Sons, 1997.

[33] J. Houssineau and D. E. Clark, "Bayesian estimation of multi-object systems with independently identically distributed correlations," in *Statistical Signal Processing (SSP), 2014 IEEE Workshop on*. IEEE, 2014, pp. 228–231.

[34] G. Bates and J. Newman, "Contributions to the Theory of Accident Proneness," *Publications in Statistics, University of California*, vol. 1, no. 9, pp. 215–254, 1952.

[35] M. Fackler, "Panjer class united – one formula for the Poisson, Binomial, and Negative Binomial distribution," *ASTIN colloquium*, 2009.

[36] S. A. Klugman, H. H. Panjer, and G. Willmot, *Loss Models: From Data to Decisions*, ser. Wiley Series in Probability and Statistics. Wiley, 2012.

[37] E. D. Delande, M. Üney, J. Houssineau, and D. E. Clark, "Regional Variance for Multi-Object Filtering," *Signal Processing, IEEE Transactions on*, vol. 62, no. 13, pp. 3415–3428, Jul. 2014.

[38] D. E. Clark and J. Houssineau, "Faa di Bruno's formula for chain differentials," *ArXiv e-prints*, Oct. 2013, arXiv:1202.0264v4.