Continuous-discrete multiple target filtering: PMBM, PHD and CPHD filter implementations

Ángel F. García-Fernández, Simon Maskell

***Abstract*—This paper develops models and algorithms for continuous-discrete multiple target filtering, in which the multi-target system is modelled in continuous time and measurements are available at discrete time steps. In order to do so, this paper first proposes a statistical model for multi-target appearance, dynamics and disappearance in continuous time, based on continuous time birth/death processes and stochastic differential equations. The multitarget state is observed at known time instants based on the standard measurement model, and the objective is to compute the distribution of the multi-target state at these time steps. For the Wiener velocity model, we derive a closed-form formula to obtain the best Gaussian Poisson point process fit to the birth density based on Kullback-Leibler minimisation. The resulting discretised model gives rise to the continuous-discrete Gaussian Poisson multi-Bernoulli mixture (PMBM) filter, the continuous-discrete Gaussian mixture probability hypothesis density (PHD) filter and the continuous-discrete Gaussian mixture cardinality PHD (CPHD) filter. The proposed filters are specially useful for multi-target estimation when the time interval between measurements is non-uniform.**

***Index Terms*—Multiple target filtering, stochastic differential equations, Poisson multi-Bernoulli mixture filter.**

I. Introduction

Multiple target filtering plays a central role in many disciplines such as autonomous driving, computer vision, and defense [1], [2]. In this problem, we take noisy measurements from an unknown and variable number of targets that appear, move and disappear from the scene of interest and the objective is to infer the locations and velocities of the targets at each time step.

Multi-target filtering is usually posed in a Bayesian framework, which enables the inclusion of prior knowledge in the system regarding target appearance and disappearance, and target dynamics. The standard approach is to model target appearance, disappearance and dynamics directly in discrete time, at the time steps when measurements are taken [3]–[15].

Even if the measurements are taken at discrete time steps, continuous time models for target dynamics are important in many applications, for instance, ballistic tracking, satellite orbit determination [16] and when the time interval between measurements is nonuniform [17]–[19], as in event-driven tracking [20], [21]. Single target dynamics can be modelled in continuous time using stochastic differential equations (SDEs) [22]–[24], as in [17], [25]–[27]. One option to solve a continuous-discrete single target filtering problem, in which we aim to obtain the posterior density of the target, with continuous-time dynamics, at the known discrete time steps when the measurements are taken, is to obtain the corresponding single target transition density between any two time steps using the Fokker-Planck-Kolmogov forward partial differential equation [23]. Once we have the corresponding transition density or an approximation, one can solve the continuous-discrete problem using standard Bayesian filtering techniques on discrete time. For example, the widely used nearly constant velocity model, also called Wiener velocity model, uses a transition density that arises from an SDE [23], [25].

Unified continuous-discrete models for multi-target systems, which include continuous time models for target appearance, dynamics, disappearance, and the corresponding discretisation at the time steps when measurements are received, are not available in the multi-target filtering literature. One benefit of a continuous-discrete multi-target model is that it allows us to perform multi-target filtering with nonuniform time intervals between measurements in an elegant and mathematically sound manner. That is, different time intervals between measurements imply a different distribution of number of targets entering and leaving the scene of interest and their corresponding locations. Continuous-discrete models also allow us to analyse the performance of measurement systems with different sampling times, with the same underlying continuous time multi-target model, e.g., two radar systems that differ in the sampling times to track the same types of airplanes.

In this paper, we fill this gap in the literature and propose a continuous time multi-target model and its discretisation. The proposed model uses an $\mathrm{M}/\mathrm{M}/\infty$ queuing system[1] [28] to model the distribution of the time of appearance and disappearance of the targets. In the $\mathrm{M}/\mathrm{M}/\infty$ system, the time of appearance of the targets is modelled by a Poisson process in time, and their life span is exponentially distributed. Poisson processes in time are commonly used in different applications, for example, airport departures [29], customer arriving at service stations, and traffic [30]. In the proposed model, while a target is alive, its dynamics are governed by an SDE. The resulting continuous-discrete multi-target system is Markovian with time-dependent parameters: new born targets are distributed as a Poisson point process (PPP), and targets move independently with a certain single target transition

A. F. García-Fernández and S. Maskell are with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3GJ, United Kingdom (emails: {angel.garcia-fernandez, s.maskell}@liverpool.ac.uk). A. F. García-Fernández is also with the ARIES Research Centre, Universidad Antonio de Nebrija, Madrid, Spain.

[1] $\mathrm{M}/\mathrm{M}/\infty$ denotes a type of queuing system using queuing system notation [28]. The first letter denotes the type of arrival process, where M means Markovian (Poisson). The second letter denotes the service time distribution (life span distribution), where M means Markovian (exponential). The last entry denotes the number of servers. In this paper, it is ∞, as newly appearing targets do not have to wait for other targets to leave the scene of interest.

density and a probability of survival.

The discretised multi-target model gives rise to continuous-discrete versions of multi-target filters that can use a PPP as the birth process, for example, the probability hypothesis density (PHD) filter [31], the cardinality PHD (CPHD) filter [32], and the Poisson multi-Bernoulli mixture (PMBM) filter [8]. The Gaussian mixture implementations of the PHD, CPHD and PMBM filters [33]–[35] require that the birth PPP has an intensity represented as a Gaussian mixture. However, the birth intensity of the continuous-discrete (CD) PHD, CD-CPHD and CD-PMBM is not a Gaussian mixture so these implementations are not directly applicable. An additional contribution of this paper, is that, for the Wiener velocity model, we derive the best Gaussian PPP fit to the discretised birth model by minimising the Kullback-Leibler divergence, which is also inherent in the PHD/CPHD filtering recursion [36], [37]. The best Gaussian PPP fit allows us to extend the Gaussian mixture implementations of the PHD, CPHD and PMBM filters to their continuous-discrete counterparts. We provide simulation results comparing the performance of the different filters. Importantly, we demonstrate that, when the time interval between measurements is non-uniform, the proposed filters can show a remarkable decrease in computational time compared with filters with fixed sampling time, as commonly adopted in the multi-target filtering literature, without affecting performance.

The rest of the paper is organised as follows. The problem formulation and the continuous time multi-target dynamic model are explained in Section II. The discretisation of the continuous time multi-target model is provided in Section III. The derivation of the best Gaussian PPP to the discretised birth distribution and its extension to Gaussian mixtures are given in Section IV. The CD-PHD, CD-CPHD and CD-PMBM filters, along with a complete description of the Gaussian implementation of the CD-PMBM filter, are explained in Section V. Simulation results comparing the CD-PHD, CD-CPHD and CD-PMBM filters are provided in Section V. Finally, conclusions are drawn in Section VI.

II. Problem formulation

We consider the problem of multi-target filtering when target appearance, dynamics and disappearance are governed by continuous time models. The multi-target state at time $t \in [0, \infty)$ can be represented as a set $X(t) \in \mathcal{F}(\mathbb{R}^{n_x})$, where $\mathbb{R}^{n_x}$ is the single target space, and $\mathcal{F}(\mathbb{R}^{n_x})$ is the space of all finite subsets of $\mathbb{R}^{n_x}$. At any time instant, new targets may appear and added to $X(t)$, and targets may die and be removed from $X(t)$. The specifics of this model will be explained in Section II-A.

The system state $X(t)$ is measured at known time instants t_k, $k \in \mathbb{N}$, and we denote $X_k = X(t_k)$. We consider the standard point target detection model, such that the measurement at time step k is a set $Z_k = \{z_k^1, ..., z_k^{m_k}\} \in \mathcal{F}(\mathbb{R}^{n_z})$ where m_k is the number of measurements at time step k. Given X_k, each target state $x \in X_k$ is either detected with probability $p_D(x)$ and generates one measurement with density $l(\cdot|x)$, or missed with probability $1 - p_D(x)$. Set Z_k is then the union of the target-generated measurements and Poisson clutter with intensity $\kappa(\cdot)$.

The objective is to compute the (multi-target) density of the multi-target state X_k given the sequence of measurements $(Z_1, ..., Z_k)$ [4]. In order to do so, in this paper, we discretise the continuous time multi-target model to obtain the corresponding discrete birth, dynamics and death models, as is standard in the literature. With the discretised models, the filtering density can be computed recursively by the prediction and the update steps of the multi-target filtering recursion in discrete time [4].

In order to distinguish between target births and deaths in continuous and discrete time, we refer to target births and deaths in continuous time as target appearances and disappearances, and keep the terms births and deaths for the discretised system. It should be noted that a target appearance in continuous time does not imply a target birth in a discretised time, as the target may have already disappeared before a measurement was taken.

A. Continuous time multi-target dynamic model

The continuous time multi-target dynamic model we propose is characterised by the following assumptions

- A1 Target appearance times follow a Poisson process (in time) with rate λ [28].
- A2 When a target appears, its single target distribution is Gaussian with mean $\overline{x}_a$ and covariance matrix P_a, and is independent of the rest of the targets.
- A3 The life span τ of a target, time from appearance to disappearance, is an independent random variable that is exponentially distributed with rate μ.
- A4 Targets move independently according to a linear time invariant SDE [23]. For a target state $x(t) \in \mathbb{R}^{n_x}$, its dynamics are given by

$$dx(t) = Ax(t)\,dt + Ld\beta(t) \tag{1}$$

  where A and L are matrices, $dx(t)$ is the differential of $x(t)$, and $\beta(t) \in \mathbb{R}^{n_\beta}$ is a Brownian motion, also called Wiener process, with diffusion matrix Q_β. The size of matrices A and L are $n_x \times n_x$ and $n_x \times n_\beta$, respectively.

It should be noted that the birth/death process, which models the number of alive objects in the continuous time multi-target process and is characterised by A1 and A3, corresponds to an $\mathrm{M/M/\infty}$ queuing system [28]. Assumption A1 implies that the number of target arrivals in any interval $(t, t + \Delta t]$ is independent of the number of arrivals in other non-overlapping intervals. The distribution of the number of target arrivals in $(t, t + \Delta t]$ is $\mathcal{P}(\cdot; \lambda\Delta t)$, where

$$\mathcal{P}(n; \lambda\Delta t) = \frac{(\lambda\Delta t)^n}{n!} e^{-\lambda\Delta t}$$

represents the probability mass function of a Poisson distribution with parameter $\lambda\Delta t$ evaluated at n. It should be noted that λ corresponds the mean number of target appearances per unit of time. In addition, A1 implies that given that there are n targets that have arrived in the interval $(t, t + \Delta t]$, their arrival times are independent and uniformly distributed in the interval $(t, t + \Delta t]$ [28].

Given A3, the distribution of the life span of a target is

$$p(\tau) = \begin{cases} \mu e^{-\mu\tau} & \tau \geq 0 \\ 0 & \tau < 0, \end{cases} \tag{2}$$

and the average life span of a target is $1/\mu$. In a practical problem, both λ and μ can be chosen according to previous knowledge of target appearance and disappearance in a scenario.

III. Discretisation of the multi-target dynamic model

In this section, we explain how to discretise the continuous time multi-target model in Section II to obtain a standard multi-target dynamic model that captures the evolution of the system from time t_{k-1} to t_k. This model is characterised by probability $p_{S,k}$ of survival at time step k, which is obtained in Section III-A, single target transition density $g_k(\cdot|\cdot)$ at time step k, which is obtained in Section III-B, and multi-target birth density at time step k, which is obtained in Section III-C. We use the notation $\Delta t_k = t_k - t_{k-1}$.

A. Probability of survival

The exponential distribution in A3 is the only distribution of the life span of a target in continuous time that is memoryless [30]. Memoryless means that, given that the target is alive at a certain time step, the distribution of the rest of the life span of a target is also exponential with the same parameter μ. In probabilistic terms, the memoryless property can be written as

$$p(\tau | \tau > t) = p(\tau - t). \tag{3}$$

The memoryless property also implies that the probability that a target survives from time t_{k-1} to t_k does not depend on how long the target has been alive before t_{k-1}, only on the length Δt_k of the time interval. Therefore, this model keeps the Markovian property of standard, discrete multi-target dynamic systems in which the probability of survival of a target does not depend on how long the target has been alive so far. Using (2), the probability of survival of a target in the time interval from t_{k-1} to t_k is

$$p_{S,k} = p(\tau | \tau > \Delta t_k) \tag{4}$$

$$= e^{-\mu \Delta t_k}. \tag{5}$$

B. Single target transition density

Using the SDE (1), we can directly obtain an expression for the transition density from time t_{k-1} to t_k [23]

$$g_k(x(t_k)|x(t_{k-1})) = \mathcal{N}(x(t_k); F_k x(t_{k-1}), Q_k) \tag{6}$$

$$F_k = \exp(A\Delta t_k) \tag{7}$$

$$Q_k = \int_0^{\Delta t_k} \exp(A(\Delta t_k - \xi)) L Q_\beta L^T \times \exp(A(\Delta t_k - \xi))^T d\xi \tag{8}$$

where $\exp(A)$ denotes the matrix exponential of A and $\mathcal{N}(x; \overline{x}, Q)$ denotes a Gaussian density with mean $\overline{x}$ and covariance matrix Q evaluated at x. Methods to compute (8) are discussed in [23], [38].

C. Birth model

From time t_{k-1} to time t_k, targets appear according to A1. However, A1 does not correspond to the distribution of the number of new born targets at time step k in the discretised system. The reason is that some of the targets that may appear, may also disappear before time step k, according to the model in A3, so they are not included in the set X_k. The new targets that are included in set X_k are therefore those that appeared between time t_{k-1} and time t_k, and survived up to time step t_k. We proceed to calculate the distribution of the new born targets at time step k, which includes the cardinality distribution and the spatial distribution.

At time step step t_{k-1}, the number of appearing targets between t_{k-1} and time t_k is zero. Then, targets appear according to A1, but they also disappear according to A3. The distribution of the number of appearing targets in the interval t_{k-1} and t_k that are still alive at time step t_k is given by the transient solution of an $\mathrm{M}/\mathrm{M}/\infty$ system. Therefore, the cardinality distribution $\rho_k^b(\cdot)$ of the new born targets at time step k is [30]

$$\rho_k^b(n) = \mathcal{P}\left(n; \frac{\lambda}{\mu}\left(1 - e^{-\mu\Delta t_k}\right)\right). \tag{9}$$

We proceed to calculate the spatial distribution of new born targets. A1 implies that, given that there are n targets that appear in $(t_{k-1}, t_k]$, their appearance times are independent and uniformly (IID) distributed in the interval $(t_{k-1}, t_k]$ [28]. In addition, the spatial distribution and target dynamics are both IID, see A2 and A4, which implies that the spatial distribution of new born targets given the cardinality is IID.

If a target appears at time step $t_k - t$ with $t \in [0, \Delta t_k)$ and t denoting the time lag between appearing time and t_k, its spatial density at time $t_k - t$ is independent of the rest of the targets and is given by A2

$$p(x(t_k - t)|t) = \mathcal{N}(x(t_k - t); \overline{x}_a, P_a). \tag{10}$$

Using the single target transition density (6), we have that the transition density up to time t_k is

$$g(x_k | x(t_k - t)) = \mathcal{N}\left(x_k; F_{(t)} x(t_k - t), Q_{(t)}\right) \tag{11}$$

$$F_{(t)} = \exp(At) \tag{12}$$

$$Q_{(t)} = \int_0^t \exp(A(t-\xi)) L Q_\beta L^T \times \exp(A(t-\xi))^T d\xi. \tag{13}$$

where we use subindex with a parenthesis in $F_{(t)}$ and $Q_{(t)}$ to highlight that we use a time lag t as a parameter, not a discrete time index, as in (7) and (8).

Given the time lag t, the spatial density at appearing time (10) and the transition density (11), the birth density $p(x_k|t)$ is Gaussian with parameters

$$p_k(x_k|t) = \mathcal{N}\left(x_k; F_{(t)}\overline{x}_a, F_{(t)} P_a F_{(t)}^T + Q_{(t)}\right), \tag{14}$$

where we have applied the Kalman filter prediction step [39].

In Appendix A, we show that the distribution of the time lag t of new born targets is a truncated exponential distribution with parameter μ in the interval $[0, \Delta t_k)$, whose density is

$$p_k(t) = \frac{\mu}{1 - e^{-\mu\Delta t_k}} e^{-\mu t} \chi_{[0,\Delta t_k)}(t) \tag{15}$$

where $\chi_A(t)$ is the indicator function of set A evaluated at t: $\chi_A(t) = 1$ if $t \in A$ and $\chi_A(t) = 0$ otherwise.

Using (14) and (15), the single target birth density is

$$p_k(x_k) = \int_0^{\Delta t_k} p_k(x_k|t)\, p_k(t)\, dt. \tag{16}$$

To sum up, the distribution of the new born targets is a PPP as the distribution over the number of targets is Poisson, see (9), and given a cardinality, targets are IID with density (16). The corresponding multi-target density is

$$b_k\left(\left\{x_k^1, ..., x_k^n\right\}\right) = n!\mathcal{P}\left(n; \frac{\lambda}{\mu}\left(1 - e^{-\mu\Delta t_k}\right)\right) \prod_{i=1}^{n} p_k\left(x_k^i\right). \tag{17}$$

This implies that the intensity (also called probability hypothesis density) of the new born targets at time step k is $D_{b_k}(x_k) = \frac{\lambda}{\mu}\left(1 - e^{-\mu\Delta t_k}\right) p_k(x_k)$ [4].

Finally, we provide the steps for sampling the multi-target states $X_1, ..., X_K$ up to time step K using the proposed continuous-discrete multi-target system in Algorithm 1.

Algorithm 1 Sampling the continuous-discrete multi-target system

Input: Time instants $t_1, ..., t_K$, models in A1-A4.
Output: Multi-target samples $X_1, ..., X_K$ at each time step.
- Initialisation $X_0 = \emptyset$, $t_0 = 0$.
for $k = 1$ to K **do**
- $\Delta t_k = t_k - t_{k-1}$.
- $X_k = \emptyset$. ▷ Initialisation at time step k.
for all $x_{k-1} \in X_{k-1}$ **do** ▷ Go through previous targets.
- Sample $s = 1$ with probability $p_{S,k}$, see (5), $s = 0$ with probability $1 - p_{S,k}$,
if $s = 1$ **then** ▷ Target x_{k-1} survives.
- Sample $x(t_k)$ from $g_k(x(t_k)|x_{k-1})$, see (6).
- Set $X_k = X_k \cup \{x(t_k)\}$.
end if
end for
- Sample n, number of new born targets, from (9).
for $i = 1$ to n **do** ▷ Go through new born targets.
- Sample t from (15).
- Sample x_k from $p_k(x_k|t)$, see (14).
- Set $X_k = X_k \cup \{x_k\}$.
end for
end for

IV. Best Gaussian PPP fit to the birth process

Gaussian implementations of multi-target filters with PPP birth typically require that the intensity is Gaussian or Gaussian mixture, for example, the PMBM, PHD and CPHD filters [33]–[35]. However, the intensity of the birth density (17) is not Gaussian or Gaussian mixture in general.

In this section, we obtain the PPP with Gaussian intensity, which is referred to as Gaussian PPP, that most closely approximates the PPP birth process (17) by minimising the Kullback-Leibler divergence (KLD). This allows us to extend the Gaussian implementation of the PMBM, PHD and CPHD filters to the continuous-discrete case. In this section, we first introduce a KLD minimisation result in Section IV-A. We introduce the Wiener velocity model for single target dynamics in Section IV-B, and obtain the corresponding solution to the KLD minimisation problem in closed-form Section IV-C. Finally, we address the Gaussian mixture appearance model in Section IV-D.

A. Kullback-Leibler divergence minimisation

KLD minimisation is a widely used criterion to perform approximate Bayesian inference in an optimal way [36], [37], [40], [41]. In this section, we obtain the best Gaussian PPP fit to the birth density (17) by minimising the KLD.

The KLD $\mathrm{D}(p\|q)$ between multi-target densities $p(\cdot)$ and $q(\cdot)$ is [4]

$$\mathrm{D}(p\|q) = \int p(X) \log \frac{p(X)}{q(X)} \delta X \tag{18}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \int p\left(\left\{x^1, ..., x^n\right\}\right) \log \frac{p\left(\left\{x^1, ..., x^n\right\}\right)}{q\left(\left\{x^1, ..., x^n\right\}\right)} dx^{1:n}. \tag{19}$$

We denote the Gaussian PPP approximation of the birth density (17) as

$$\widetilde{b}_k\left(\left\{x_k^1, ..., x_k^n\right\}\right) = n!\mathcal{P}(n; \lambda_{b,k}) \prod_{i=1}^{n} \mathcal{N}\left(x_k^i; \overline{x}_{b,k}, P_{b,k}\right) \tag{20}$$

where $\lambda_{b,k}$ if the parameter of the Poisson distribution of the cardinality, and $\overline{x}_{b,k}$ and $P_{b,k}$ denote the mean and covariance matrix of the single target density, respectively.

Proposition 1. *Given $b_k(\cdot)$ in (17), the Gaussian PPP $\widetilde{b}_k(\cdot)$ in (20) that minimises the KLD $\mathrm{D}\left(b_k \middle\| \widetilde{b}_k\right)$ is parameterised by*

$$\lambda_{b,k} = \frac{\lambda}{\mu}\left(1 - e^{-\mu\Delta t_k}\right) \tag{21}$$

$$\overline{x}_{b,k} = \mathrm{E}_{p_k}[x] = \int x p_k(x)\, dx \tag{22}$$

$$P_{b,k} = \mathrm{C}_{p_k}[x] = \int (x - \overline{x}_{b,k})(x - \overline{x}_{b,k})^T p_k(x)\, dx, \tag{23}$$

where $\mathrm{E}_{p_k}[x]$ and $\mathrm{C}_{p_k}[x]$ denote the mean and covariance matrix of a random variable x distributed according to the single target density $p_k(\cdot)$.

Proposition 1 is proved in Appendix B.

B. Wiener velocity model

We consider the widely-used Wiener velocity model [23] for single target dynamics, for which the solution to Proposition 1 is closed-form, as derived in the next section. We proceed to describe the resulting $F_{(t)}$ and $Q_{(t)}$ in (12) and (13).

We consider a state

$$x(t_k) = [p_1(t_k), ..., p_d(t_k), v_1(t_k), ..., v_d(t_k)]^T \tag{24}$$

where $d = n_x/2$ is the dimension of the space where the target moves, typically $d \in \{1, 2, 3\}$, $p_i(t_k)$ is the position in the ith dimension at time t_k and $v_i(t_k)$ is the velocity in the ith dimension at time t_k. The target moves according to a Wiener velocity model [23]–[25] in each direction, which is characterised by the SDE (1) with $n_\beta = d$, $Q_\beta = qI_d$ and

$$A = \begin{pmatrix} 0_d & I_d \\ 0_d & 0_d \end{pmatrix}, \; L = \begin{pmatrix} 0_d \\ I_d \end{pmatrix} \tag{25}$$

where 0_d and I_d represent the zero and identity matrices of size d, respectively.

The solutions to (12) and (13) are [23]

$$F_{(t)} = \begin{pmatrix} I_d & tI_d \\ 0_d & I_d \end{pmatrix} \tag{26}$$

$$Q_{(t)} = q \begin{pmatrix} \frac{t^3}{3} I_d & \frac{t^2}{2} I_d \\ \frac{t^2}{2} I_d & tI_d \end{pmatrix}. \tag{27}$$

This implies that $F_k = F_{(\Delta t_k)}$ and $Q_k = Q_{(\Delta t_k)}$ in (7) and (8).

C. KLD minimisation for the Wiener velocity model

In this section, we provide a closed-form expression for the Gaussian PPP that minimises the KLD in Proposition 1 for the Wiener velocity model. First, we write the mean and covariance at the time of appearance, see A2, as block matrices whose blocks correspond to position and velocity elements. That is,

$$\overline{x}_a = \left[\overline{p}_a^T, \overline{v}_a^T\right]^T \tag{28}$$

$$P_a = \begin{pmatrix} P_a^{pp} & P_a^{pv} \\ \left(P_a^{pv}\right)^T & P_a^{vv} \end{pmatrix} \tag{29}$$

where $\overline{p}_a$ is the mean position, $\overline{v}_a$ is the mean velocity, P_a^{pp} is the covariance matrix of the position, P_a^{vv} is the covariance matrix of the velocity and P_a^{pv} is the covariance matrix between the position and velocity.

Proposition 2. *For the Wiener velocity model, the Gaussian PPP density $\widetilde{b}_k(\cdot)$, see (20), that minimises the KLD $\mathrm{D}\left(b_k \middle\| \widetilde{b}_k\right)$ is characterised by*

$$\lambda_{b,k} = \frac{\lambda}{\mu}\left(1 - e^{-\mu \Delta t_k}\right) \tag{30}$$

$$\overline{x}_{b,k} = \begin{pmatrix} I_d & \mathrm{E}[t] I_d \\ 0_d & I_d \end{pmatrix} \overline{x}_a \tag{31}$$

$$P_{b,k} = \begin{pmatrix} P_{b,k}^{pp} & P_{b,k}^{pv} \\ \left(P_{b,k}^{pv}\right)^T & P_{b,k}^{vv} \end{pmatrix} \tag{32}$$

where

$$P_{b,k}^{pp} = \mathrm{C}[t]\, \overline{v}_a \overline{v}_a^T + q\frac{\mathrm{E}\left[t^3\right]}{3} I_d + P_a^{pp} + \mathrm{E}[t]\left(P_a^{pv} + \left(P_a^{pv}\right)^T\right) + \mathrm{E}\left[t^2\right] P_a^{vv} \tag{33}$$

$$P_{b,k}^{pv} = q\frac{\mathrm{E}\left[t^2\right]}{2} I_d + P_a^{pv} + \mathrm{E}[t]\, P_a^{vv} \tag{34}$$

$$P_{b,k}^{vv} = q\mathrm{E}[t]\, I_d + P_a^{vv} \tag{35}$$

and the moments of t, whose density is (15), are

$$\mathrm{E}[t] = \frac{1}{\mu} - \frac{\Delta t_k e^{-\mu \Delta t_k}}{1 - e^{-\mu \Delta t_k}} \tag{36}$$

$$\mathrm{E}\left[t^2\right] = \frac{1}{1 - e^{-\mu \Delta t_k}} \times \left[\frac{2}{\mu^2} - e^{-\mu \Delta t_k}\left((\Delta t_k)^2 + \frac{2\Delta t_k}{\mu} + \frac{2}{\mu^2}\right)\right] \tag{37}$$

$$\mathrm{C}[t] = \mathrm{E}\left[t^2\right] - (\mathrm{E}[t])^2 \tag{38}$$

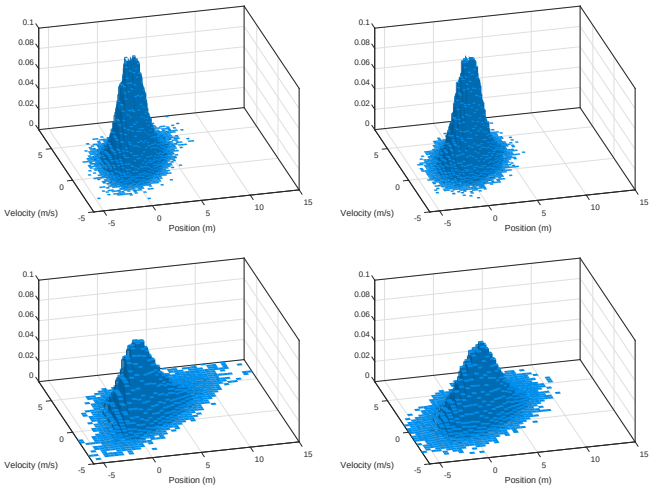


Figure 1: Normalised histograms of Example 3 with 10^5 samples from the true single target birth density $\Delta t_k = 1\,\mathrm{s}$ (top left), best Gaussian fit $\Delta t_k = 1\,\mathrm{s}$ (top right), true birth density $\Delta t_k = 2\,\mathrm{s}$ (bottom left), and best Gaussian fit $\Delta t_k = 2\,\mathrm{s}$ (bottom right). The best Gaussian fit is more accurate for $\Delta t_k = 1\,\mathrm{s}$ than for $\Delta t_k = 2\,\mathrm{s}$.

$$\mathrm{E}\left[t^3\right] = \frac{1}{1 - e^{-\mu \Delta t_k}}\left[\frac{6}{\mu^3} - e^{-\mu \Delta t_k} \times \left((\Delta t_k)^3 + \frac{3(\Delta t_k)^2}{\mu} + \frac{6\Delta t_k}{\mu^2} + \frac{6}{\mu^3}\right)\right]. \tag{39}$$

Proposition 2 is proved in Appendix C. For clarity of presentation, we have considered that the single target distribution at appearance time is Gaussian, see A2. Nevertheless, we would like to remark that Proposition 2 holds for any single target distribution at appearance time such that its mean and covariance matrix are (28) and (29), as the KLD minimisation only requires the first two moments of this distribution. In addition, note that, even though, we can minimise the KLD in closed-form, the resulting value of the KLD does not admit a closed-form expression, as (16) does not have a closed-form expression. We proceed to illustrate Proposition 2 with an example.

Example 3. We consider a one dimensional target ($d = 1$) with $\overline{x}_a = [0\,(\mathrm{m}), 2\,(\mathrm{m/s})]^T$ and $P_a = I_2$, with units in the international system, that moves with the Wiener velocity model with $q = 1\ \left(\mathrm{m}^2/\mathrm{s}^3\right)$. Its life span rate is $\mu = 0.01\,\mathrm{s}^{-1}$. We draw 10^5 samples from the true single target birth density (16) considering $\Delta t_k = 1\,\mathrm{s}$ and $\Delta t_k = 2\,\mathrm{s}$. The normalised histograms are shown in Figure 1. In this figure, we also show the normalised histograms for 10^5 samples drawn from the best Gaussian fits, whose mean and covariance are given by (31) and (32). The tails of the true density w.r.t. the best Gaussian fit are larger for $\Delta t_k = 2\,\mathrm{s}$ than for $\Delta t_k = 1\,\mathrm{s}$, as the target can potentially move longer distances in a time interval $\Delta t_k = 2\,\mathrm{s}$ than in $\Delta t_k = 1\,\mathrm{s}$, for its considered life span. Therefore, the best Gaussian fit is more accurate for $\Delta t_k = 1\,\mathrm{s}$ than for $\Delta t_k = 2\,\mathrm{s}$. □

D. Gaussian mixture appearance model

In this paper, we mainly consider an appearance model based on Gaussian densities, see A2. In some cases, it can be convenient to consider Gaussian mixtures, which can approximate any density of interest with negligible accuracy with a sufficient number of terms [42]. That is, we substitute A2 by

- A5 When a target appears, its single target distribution is a Gaussian mixture

$$\sum_{j=1}^{N_a} w_a^j \mathcal{N}\left(x; \overline{x}_a^j, P_a^j\right)$$

where N_a are the number of components, and w_a^i, $\overline{x}_a^j$ and P_a^j represent the weight, mean, and covariance of the j-th Gaussian component, respectively.

In this section, we indicate how to obtain a suitable discretisation of this birth model with the Wiener velocity model based on the previous results for the Gaussian case. Considering A5, the single target birth density (16) becomes

$$p_k(x_k) = \sum_{j=1}^{N_a} w_a^j \int_0^{\Delta t_k} \mathcal{N}\left(x_k; F_{(t)}\overline{x}_a^j, F_{(t)}P_a^j F_{(t)}^T + Q_{(t)}\right) \times p_k(t)\, dt. \quad (40)$$

While one can fit the best Gaussian PPP applying Proposition 1 to (40), it is more convenient to consider a Gaussian mixture single target birth density that performs moment matching for each mixture component. That is, the approximated birth distribution becomes

$$\widetilde{b}_k\left(\{x_k^1, ..., x_k^n\}\right) = n!\mathcal{P}(n; \lambda_{b,k}) \times \prod_{i=1}^{n}\left(\sum_{j=1}^{N_a} w_a^j \mathcal{N}\left(x_k^i; \overline{x}_{b,k}^j, P_{b,k}^j\right)\right) \quad (41)$$

where $\lambda_{b,k}$ is given by (30), and $\overline{x}_{b,k}^j$ and $P_{b,k}^j$ are obtained by substituting $\overline{x}_a^j$ and P_a^j into (31) and (32). In Appendix D, we prove that (41) corresponds to a KLD minimisation that considers auxiliary variables, not the KLD minimisation in Proposition 1.

V. Continuous-discrete PMBM filter

In this section, we explain the continuous-discrete PHD (CD-PHD) filter, the continuous-discrete CPHD (CD-CPHD) filter and the continuous discrete PMBM (CD-PMBM) filter. We first explain how to obtain these continuous-discrete filters in Section V-A. We then explain the structure of the CD-PMBM filter in Section V-B and its Gaussian implementation in Section V-C. Practical considerations are discussed in Section V-D.

A. Continuous-discrete multi-target filters

As indicated in Section III, the discretised dynamic model has a PPP birth, a probability of survival and a single target dynamic density that depend on the time interval between measurements. We can then directly extend discrete-time multi-target filters that can work with PPP birth, such as PMBM, PHD and CPHD filters to the continuous-discrete case. That is, the CD-PMBM, the CD-PHD and CD-CPHD filter correspond to the PMBM, PHD and CPHD recursions with the following models

- The probability of survival is given by (5).
- The PPP birth process is given by (17).
- The single target transition density is given by (6).

It is important to note that the discretisation of the continuous dynamics only affects the dynamic model and the prediction step. The measurement model and the update step are not affected by the discretisation.

We would like to recall several properties of the CD-PMBM, CD-PHD and CD-CPHD filtering recursions. The CD-PMBM recursion provides us with the exact posterior density of the multi-target state given all past measurements. The CD-PHD filtering recursion propagates a PPP through the filtering recursion by minimising the KLD after each update step [31], [36]. A simplification of the CD-PHD filter with respect to the CD-PMBM is that it only needs to propagate the intensity of the PPP, which is computationally efficient, but implies a loss in performance. The CD-CPHD filtering recursion propagates an IID cluster density, which considers an arbitrary cardinality distribution and IID target states given each cardinality. The IID cluster approximation is performed after each prediction and update by minimising the KLD [36], [37]. Therefore, the CD-CPHD has a higher degree of flexibility than the CD-PHD, as it considers an arbitrary cardinality, with an increase in the computational burden.

The Gaussian/Gaussian mixture implementations of the CD-PHD, CD-CPHD and CD-PMBM filters for Wiener velocity model can be obtained in closed form by considering

- The probability of survival is given by (5).
- The PPP birth process is given by Proposition 2, or Eq. (41).
- The single target transition density is $g(x_k|x_{k-1}) = \mathcal{N}(x_k; F_k x_{k-1}, Q_k)$ where $F_k = F_{(\Delta t_k)}$ and $Q_k = Q_{(\Delta t_k)}$, see (26) and (27).
- Constant probability p_D of detection and a linear and Gaussian single measurement density $l(z|x) = \mathcal{N}(z; Hx, R)$.

The Gaussian mixture implementation of the CD-PHD and CD-CPHD filters correspond to the implementations in [33], [34] using the above time-varying parameters for the dynamic model. As these filters are well-described in [33], [34], we do not provide the specific details in this paper. The Gaussian implementation of the CD-PMBM corresponds to the implementation in [35] using the above time-varying parameters. In the following, we provide a detailed description of the Gaussian implementation of the CD-PMBM filter in [35], explicitly addressing the continuous-discrete aspects.

B. CD-PMBM filter structure

In the CD-PMBM filter, the posterior is a PMBM density, which is the (multi-target) density of the union of a PPP and a multi-Bernoulli mixture. We proceed to give an overview of the structure of this density before introducing the mathematical formulation. The PPP represents the information on the

undetected targets, which are targets that exist at the current time but have never been detected. Each measurement can be the first detection of a target, a detection from a previously detected target or clutter. Therefore, as the target whose first detection is a particular measurement may exist or not, its resulting distribution is Bernoulli and is referred to as "potentially detected target". Then, each potentially detected target has single target hypotheses that represent possible histories of measurement associations, including misdetections. Each term in the multi-Bernoulli mixture represents a global hypothesis that includes a single target hypothesis for each Bernoulli component.

In mathematical terms, the density of $X_{k'}$ with $k' \in \{k, k+1\}$ given the measurements up to time step k is a PMBM density [35]

$$f_{k'|k}\left(X_{k'}\right) = \sum_{Y \uplus W = X_{k'}} f^{p}_{k'|k}\left(Y\right) f^{mbm}_{k'|k}\left(W\right) \tag{42}$$

where $f^{p}_{k'|k}\left(\cdot\right)$ and $f^{mbm}_{k'|k}\left(\cdot\right)$ are the densities of the PPP and the multi-Bernoulli mixture (MBM). Also, $\uplus$ denotes the disjoint union and the summation in (42) is taken over all mutually disjoint (and possibly empty) sets Y and W whose union is $X_{k'}$. The density of the PPP in terms of its intensity $D_{k'|k}\left(\cdot\right)$ can be written as

$$f^{p}_{k'|k}\left(Y\right) = e^{-\int D_{k'|k}(x)dx} \prod_{y \in Y} D_{k'|k}\left(y\right). \tag{43}$$

The MBM can be written as

$$f^{mbm}_{k'|k}\left(W\right) \propto \sum_{a \in \mathcal{A}_{k'|k}} \sum_{\uplus_{l=1}^{n_{k'|k}} X^l = W} \prod_{i=1}^{n_{k'|k}} \left[w^{i,a^i}_{k'|k} f^{i,a^i}_{k'|k}\left(X^i\right) \right] \tag{44}$$

where $\propto$ stands for "proportional to". We proceed to explain (44). First, $n_{k'|k}$ is the number of Bernoulli components, i is an index over the Bernoulli components and a global hypothesis $a = \left(a^1, ..., a^{n_{k'|k}}\right)$, where $a^i \in \left\{1, ..., h^i_{k'|k}\right\}$ is an index over the $h^i_{k'|k}$ single target hypotheses for the i-th Bernoulli component. The density $f^{i,a^i}_{k'|k}\left(\cdot\right)$ of the i-th Bernoulli component with single target hypothesis a^i is

$$f^{i,a^i}_{k'|k}\left(X\right) = \begin{cases} 1 - r^{i,a^i}_{k'|k} & X = \emptyset \\ r^{i,a^i}_{k'|k} p^{i,a^i}_{k'|k}\left(x\right) & X = \{x\} \\ 0 & \text{otherwise} \end{cases} \tag{45}$$

where $r^{i,a^i}_{k'|k}$ is the existence probability and $p^{i,a^i}_{k'|k}\left(\cdot\right)$ is the single target density. The variable $w^{i,a^i}_{k'|k}$ is the weight of the a^i single target hypothesis of the i-th Bernoulli component.

The set $\mathcal{A}_{k'|k}$ of global hypotheses and single target hypotheses are built as follows [8]. As explained before, the measurement set at time step k is denoted as $Z_k = \left\{z^1_k, ..., z^{m_k}_k\right\}$. We refer to measurement z^j_k using the pair (k, j) and the set of all these measurement pairs up to (and including) time step k is denoted by $\mathcal{M}_k$. Then, a single target hypothesis a^i for the i-th Bernoulli component has a set of measurement pairs denoted as $\mathcal{M}^{i,a^i}_k \subseteq \mathcal{M}_k$, where there is the constraint that there can be at most one measurement index per each time step in $\mathcal{M}^{i,a^i}_k$. For example, for $k = 3$, we can have $\mathcal{M}^{i,a^i}_3 = \{(1,2), (3,1)\}$, which indicates that under the single target hypothesis a^i, the i-th Bernoulli component is associated with measurement z^2_1 at time step 1, misdetected at time step 2, and associated with measurement z^1_3 at time step 3. In a global hypothesis, each received measurement must belong to only one single target hypothesis. Then, the set $\mathcal{A}_{k'|k}$ of all global hypotheses meets [8].

$$\mathcal{A}_{k'|k} = \Bigg\{ \left(a^1, ..., a^{n_{k'|k}}\right) : a^i \in \left\{1, ..., h^i_{k'|k}\right\} \forall i, \bigcup_{i=1}^{n_{k'|k}} \mathcal{M}^{i,a^i}_k = \mathcal{M}_k, \mathcal{M}^{i,a^i}_k \cap \mathcal{M}^{j,a^j}_k = \emptyset \, \forall i \neq j \Bigg\}.$$

C. Gaussian implementation

In the Gaussian implementation of the CD-PMBM filter, the intensity of the PPP is a Gaussian mixture and the density of each Bernoulli component is also Gaussian. That is, the intensity of the PPP (43) is

$$D_{k'|k}\left(x\right) = \sum_{j=1}^{n^p_{k'|k}} w^{p,j}_{k'|k} \mathcal{N}\left(x; \overline{x}^{p,j}_{k'|k}, P^{p,j}_{k'|k}\right) \tag{46}$$

where $n^p_{k'|k}$, is the number of components, and $w^{p,j}_{k'|k}$, $\overline{x}^{p,j}_{k'|k}$ and $P^{p,j}_{k'|k}$ are the weight, mean and covariance matrix of the j-th component.

The single target density of each Bernoulli component (45) is Gaussian with

$$p^{i,a^i}_{k'|k}\left(x\right) = \mathcal{N}\left(x; \overline{x}^{i,a^i}_{k'|k}, P^{i,a^i}_{k'|k}\right) \tag{47}$$

where $\overline{x}^{i,a^i}_{k'|k}$ is the mean and $P^{i,a^i}_{k'|k}$ is the covariance matrix.

The prediction and update for the continuous-discrete Gaussian PMBM filter for the Wiener velocity model are given in the following lemmas.

Lemma 4 (Prediction)**.** *We consider that the filtering density at time step k (time t_k) is a PMBM $f_{k|k}\left(\cdot\right)$ of the form (42), with PPP intensity (46) and single target densities given by (47). At time step $k+1$ (time t_{k+1}), the probability $p_{S,k+1}$ of survival is given by (5), the transition matrix F_{k+1} by (26) and the covariance matrix Q_{k+1} by (27). Then, the predicted density is a PMBM with PPP parameters*

$$D_{k+1|k}\left(x\right) = p_{S,k+1} \sum_{j=1}^{n^p_{k|k}} w^{p,j}_{k|k} \mathcal{N}\left(x; \overline{x}^{p,j}_{k+1|k}, P^{p,j}_{k+1|k}\right) + \lambda_{b,k+1} \mathcal{N}\left(x; \overline{x}_{b,k+1}, P_{b,k+1}\right) \tag{48}$$

where $\lambda_{b,k+1}$, $\overline{x}_{b,k+1}$ and $P_{b,k+1}$ are given by Proposition 2, and

$$\overline{x}^{p,j}_{k+1|k} = F_{k+1} \overline{x}^{p,j}_{k|k} \tag{49}$$

$$P^{p,j}_{k+1|k} = F_{k+1} P^{p,j}_{k|k} F^T_{k+1} + Q_{k+1}. \tag{50}$$

The predicted multi-Bernoulli mixture parameters are

$$r^{i,a^i}_{k+1|k} = p_{S,k+1} r^{i,a^i}_{k|k} \tag{51}$$

$$\overline{x}_{k+1|k}^{i,a^i} = F_{k+1}\overline{x}_{k|k}^{i,a^i} \tag{52}$$

$$P_{k+1|k}^{i,a^i} = F_{k+1}P_{k|k}^{i,a^i}F_{k+1}^T + Q_{k+1} \tag{53}$$

and $n_{k+1|k} = n_{k|k}$, $h_{k+1|k}^i = h_{k|k}^i$, $w_{k+1|k}^{i,a^i} = w_{k|k}^{i,a^i}$ *for all* $i \in \left\{1, ..., n_{k|k}\right\}$ *and* a^i.

Lemma 5 (Update)**.** *We consider that the predicted density at time step* k *(time* t_k*) is a PMBM* $f_{k|k-1}(\cdot)$ *of the form (42), PPP intensity (46) and single target densities given by (47). The updated PMBM at time step* k *with measurement set* $Z_k = \left\{z_k^1, ..., z_k^{m_k}\right\}$ *is of the form (42) with PPP intensity*

$$D_{k|k}(x) = (1 - p_D)\sum_{j=1}^{n_{k|k-1}^p} w_{k|k-1}^{p,j}\mathcal{N}\left(x; \overline{x}_{k|k-1}^{p,j}, P_{k|k-1}^{p,j}\right). \tag{54}$$

The number of updated Bernoulli components is $n_{k|k} = n_{k|k-1} + m_k$*. For each single target hypothesis of the Bernoulli components existing at previous time steps* $i \in \left\{1, ..., n_{k|k-1}\right\}$*, the update creates* $(m_k + 1)$ *new single target hypotheses corresponding to a missed detection and an update with one of the received measurements, which implies* $h_{k|k}^i = h_{k|k-1}^i(m_k + 1)$*. For missed detection hypotheses,* $i \in \left\{1, ..., n_{k|k-1}\right\}$, $a^i \in \left\{1, ..., h_{k|k-1}^i\right\}$*, the parameters are*

$$\mathcal{M}_k^{i,a^i} = \mathcal{M}_{k-1}^{i,a^i} \tag{55}$$

$$w_{k|k}^{i,a^i} = w_{k|k-1}^{i,a^i}\left(1 - r_{k|k-1}^{i,a^i} + r_{k|k-1}^{i,a^i}(1 - p_D)\right) \tag{56}$$

$$r_{k|k}^{i,a^i} = \frac{r_{k|k-1}^{i,a^i}(1 - p_D)}{1 - r_{k|k-1}^{i,a^i} + r_{k|k-1}^{i,a^i}(1 - p_D)} \tag{57}$$

$$\overline{x}_{k|k}^{i,a^i} = \overline{x}_{k|k-1}^{i,a^i} \tag{58}$$

$$P_{k|k}^{i,a^i} = P_{k|k-1}^{i,a^i}. \tag{59}$$

For a previous Bernoulli component $i \in \left\{1, ..., n_{k|k-1}\right\}$ *and previous single target hypothesis* $\widetilde{a}^i \in \left\{1, ..., h_{k|k-1}^i\right\}$*, the new hypothesis generated by measurement* z_k^j *has* $a^i = \widetilde{a}^i + h_{k|k-1}^i j$, $r_{k|k}^{i,a^i} = 1$*, and*

$$\mathcal{M}_k^{i,a^i} = \mathcal{M}_{k-1}^{i,\widetilde{a}^i} \cup \{(k,j)\} \tag{60}$$

$$w_{k|k}^{i,a^i} = w_{k|k-1}^{i,\widetilde{a}^i} r_{k|k-1}^{i,\widetilde{a}^i} p_D \mathcal{N}\left(z_k^j; H\overline{x}_{k|k-1}^{i,\widetilde{a}^i}, S_{k|k-1}^{i,\widetilde{a}^i}\right) \tag{61}$$

$$\overline{x}_{k|k}^{i,a^i} = \overline{x}_{k|k-1}^{i,\widetilde{a}^i} + K_{k|k-1}^{i,\widetilde{a}^i}\left(z_k^j - H\overline{x}_{k|k-1}^{i,\widetilde{a}^i}\right) \tag{62}$$

$$P_{k|k}^{i,a^i} = P_{k|k-1}^{i,\widetilde{a}^i} - K_{k|k-1}^{i,\widetilde{a}^i} H P_{k|k-1}^{i,\widetilde{a}^i} \tag{63}$$

$$K_{k|k-1}^{i,\widetilde{a}^i} = P_{k|k-1}^{i,\widetilde{a}^i} H^T \left(S_{k|k-1}^{i,\widetilde{a}^i}\right)^{-1} \tag{64}$$

$$S_{k|k-1}^{i,\widetilde{a}^i} = H P_{k|k-1}^{i,\widetilde{a}^i} H^T + R. \tag{65}$$

For the new Bernoulli component initiated by z_k^j*, whose index is* $i = n_{k|k-1} + j$*, we have two single target hypotheses, which imply that* $h_{k|k}^i = 2$*, one corresponding to a non-existent Bernoulli*

$$\mathcal{M}_k^{i,1} = \emptyset,\ w_{k|k}^{i,1} = 1,\ r_{k|k}^{i,1} = 0 \tag{66}$$

and the other representing that the measurement z_k^j *can have been originated by clutter or by a new potential target*

$$\mathcal{M}_k^{i,2} = \{(k,j)\} \tag{67}$$

$$w_{k|k}^{i,2} = \kappa\left(z_k^j\right) + e\left(z_k^j\right) \tag{68}$$

$$e\left(z_k^j\right) = p_D \sum_{j=1}^{n_{k|k-1}^p} w_{k|k-1}^{p,j}\mathcal{N}\left(z_k^j; H\overline{x}_{k|k-1}^{p,j}, S_{k|k-1}^{p,j}\right) \tag{69}$$

$$r_{k|k}^{i,2} = \frac{e\left(z_k^j\right)}{\kappa\left(z_k^j\right) + e\left(z_k^j\right)} \tag{70}$$

$$\overline{x}_{k|k}^{i,2} = \sum_{l=1}^{n_{k|k-1}^p} w_l^j \overline{x}_l^j \tag{71}$$

$$P_{k|k}^{i,2} = \sum_{l=1}^{n_{k|k-1}^p} w_l^j\left[P_l + \left(\overline{x}_l^j - \overline{x}_{k|k}^{i,2}\right)\left(\overline{x}_l^j - \overline{x}_{k|k}^{i,2}\right)^T\right] \tag{72}$$

$$w_l^j \propto w_{k|k-1}^{p,l}\mathcal{N}\left(z_k^j; H\overline{x}_{k|k-1}^{p,l}, S_{k|k-1}^{p,l}\right) \tag{73}$$

$$\overline{x}_l^j = \overline{x}_{k|k-1}^{p,l} + K_l\left(z_k^j - H\overline{x}_{k|k-1}^{p,l}\right) \tag{74}$$

$$P_l = P_{k|k-1}^{p,l} - K_l H P_{k|k-1}^{p,l} \tag{75}$$

$$K_l = P_{k|k-1}^{p,l} H^T \left(S_{k|k-1}^{p,l}\right)^{-1} \tag{76}$$

$$S_{k|k-1}^{p,l} = H P_{k|k-1}^{p,l} H^T + R. \tag{77}$$

It should be noted that the single target density of new Bernoulli components is actually a Gaussian mixture [35, Sec. V.C], but it is approximated as Gaussian with moments (71) and (72) by moment matching (KLD minimisation). The reason for this approximation is that it enables us to have Gaussian densities in each Bernoulli component, see (47).

D. Practical considerations

Practical implementations must prune hypotheses and components in the PMBM as these grow boundlessly in time. In practice, we represent the MBM part of the PMBM as

$$f_{k'|k}^{mbm}(W) = \sum_{a \in \mathcal{A}_{k'|k}} w_{k'|k}^a \sum_{\uplus_{l=1}^{n_{k'|k}} X^l = W} \prod_{i=1}^{n_{k'|k}} f_{k'|k}^{i,a^i}\left(X^i\right) \tag{78}$$

where the weight of global hypothesis a is $w_{k'|k}^a \propto \prod_{i=1}^{n_{k'|k}} w_{k'|k}^{i,a^i}$, and the factor

$$\sum_{\uplus_{l=1}^{n_{k'|k}} X^l = W} \prod_{i=1}^{n_{k'|k}} f_{k'|k}^{i,a^i}\left(X^i\right)$$

represents the multi-Bernoulli density [35, Eq. (11)] for global hypothesis a. Pruning consists of approximating the weights $w_{k'|k}^a$ that are considered to be close to zero as zero, followed by a weight normalisation. As a result, the global hypotheses and their corresponding multi-Bernoullis with negligible weight are removed from (78) and the pruned posterior is also a PMBM. Pruning is also performed by approximating probabilities of existence of Bernoulli components that are

close to zero as zero. We proceed to explain how we perform pruning in the update step and after estimation.

In the update step, we first use ellipsoidal gating [43], though other types of gating are also possible [44], to disregard single target hypotheses which have a negligible weight. Then, for each previous global hypothesis, we select the $k_u = \left\lceil N_h \cdot w^a_{k|k-1} \right\rceil$ newly generated global hypotheses with highest weight, where N_h is the maximum number of global hypotheses, by solving the corresponding ranked assignment problem via Murty´s algorithm [45] [35, Sec. V.C], used in conjunction with the Hungarian algorithm [46].

After estimation, we prune the mixture of components of the PPP, the Bernoulli components and the global hypotheses as follows:

1) Discard the Gaussian components in $D_{k|k}(\cdot)$ whose weight is below a threshold Γ_p.
2) Keep the global hypotheses with N_h highest weights, and whose weight is higher than a threshold Γ_{mbm}.
3) Remove the single target hypotheses of the Bernoulli components that do not take part in any of the considered global hypotheses.
4) Remove the Bernoulli components whose existence is lower than a threshold Γ_b for all its single target hypotheses.

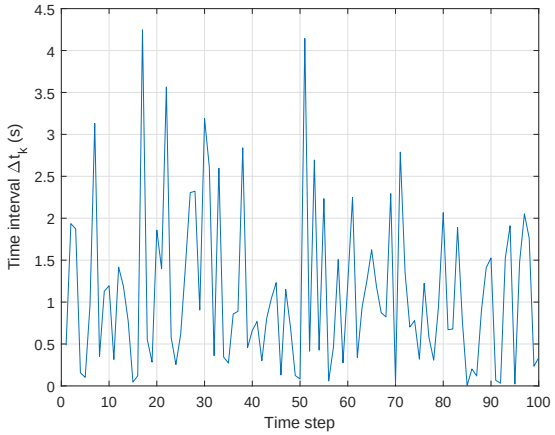
After pruning the global hypotheses, a weight normalisation is performed. Due to pruning, there can be two or more global hypotheses that have the same associated multi-Bernoulli density. These similar global hypotheses are merged into one global hypothesis by adding their weights and keeping the same multi-Bernoulli. It should also be mentioned that the resulting CD-PMBM filter has a track-oriented form [43]: there is a hypothesis tree for each Bernoulli component, representing single target hypotheses and their Bernoulli distributions, and a global hypothesis a consist of a weight $w^a_{k'|k}$ and a pointer to single target hypotheses for each Bernoulli component. Finally, multi-target estimation in the CD-PMBM filter can be done as in the PMBM filter [35, Sec. VI]. A pseudocode of the Gaussian CD-PMBM filter is given in Algorithm 2.

Algorithm 2 Gaussian CD-PMBM filter pseudocode

- Initialisation: $t_0 = 0$, $D_{0|0}(\cdot) = 0$, $n_{0|0} = 0$.
for $k = 1$ to *final time step* **do**
- $\Delta t_k = t_k - t_{k-1}$.
- Compute $p_{S,k}$ using (5).
- Compute $F_k = F_{(\Delta t_k)}$ and $Q_k = Q_{(\Delta t_k)}$ using (26) and (27).
- Compute $\lambda_{b,k}$, $\overline{x}_{b,k}$ and $P_{b,k}$ using Proposition 2.
- Prediction: use Lemma 4.
- Update: use Lemma 5 with ellipsoidal gating and Murty's algorithm to obtain updated global hypotheses, see [35, Alg. 1].
- Estimate the current set of targets, see Section V-D.
- Perform pruning, see Section V-D.
end for

VI. Simulations

In this section, we evaluate via simulations the proposed CD-PMBM, CD-PHD and CD-CPHD filters in a two-dimensional scenario with the Wiener velocity model[2].



(a)

Figure 2: Time interval Δt_k between consecutive measurements for the simulations. There are 100 time steps and the total length of the simulation is 110.63 s.

The continuous-time parameters of the multi-target dynamic system are: $\lambda = 0.08\,\mathrm{s}^{-1}$, $\mu = 0.01\,\mathrm{s}^{-1}$, $q = 0.2\,\mathrm{m}^2/\mathrm{s}^3$, $d = 2$, $\overline{p}_a = [200, 200]^T$ (m), $\overline{v}_a = [3, 0]^T$ (m/s), $P_a^{pp} = \mathrm{diag}\left(\left[50^2, 50^2\right]\right)$ $\left(\mathrm{m}^2\right)$, $P_a^{vv} = \mathrm{diag}\left([1, 1]\right)$ $\left(\mathrm{m}^2/\mathrm{s}^2\right)$ and $P_a^{pv} = 0_2$ $\left(\mathrm{m}^2/\mathrm{s}\right)$. We also consider that, at time $t_0 = 0$, the probability that there are zero targets in the scene is one. It should be noted that, according to μ, the average life span of a target is 100 s, and, in stationary regime, the number of targets in the scene follows a Poisson distribution with parameter $\frac{\lambda}{\mu} = 8$, so the average number of alive targets is 8 [30].

We consider that the sensor system takes 100 measurements and that the time interval Δt_k between measurements is drawn from an exponential distribution with parameter $\mu_m = 1\,\mathrm{s}^{-1}$. The resulting time intervals used in the simulation are shown in Figure 2. The true target trajectories considered in this scenario are drawn from the true continuous time model sampled at the time steps when measurements are taken, which is done by running Algorithm 1 with the parameters $F_{(t)}$ and $Q_{(t)}$ of (6) given by (26) and (27). The resulting target trajectories are shown in Figure 3 (top). There are 10 targets in the scene and the maximum number of targets at a given time step is 8, see Figure 3 (bottom).

The sensor measures the position of the targets with

$$H = \begin{bmatrix} I_2 & 0_2 \end{bmatrix}, R = \sigma_r^2 I_2$$

where $\sigma_r^2 = 4\,\mathrm{m}^2$. In addition, the probability p_D of detection is 0.9 and the Poisson clutter is uniformly distributed in the region of interest $A = [0, 600] \times [0, 400]$ $\left(\mathrm{m}^2\right)$ with intensity $\kappa(z) = \lambda_c \cdot u_A(z)$ where $u_A(z)$ is a uniform density in the area A and $\lambda_c = 10$, which implies 10 expected false alarms per time step.

The CD-PHD and CD-CPHD filters have the following parameters: maximum number of components is 30, pruning

[2]Matlab implementations of these continuous-discrete filters will be available at https://github.com/Agarciafernandez/MTT.

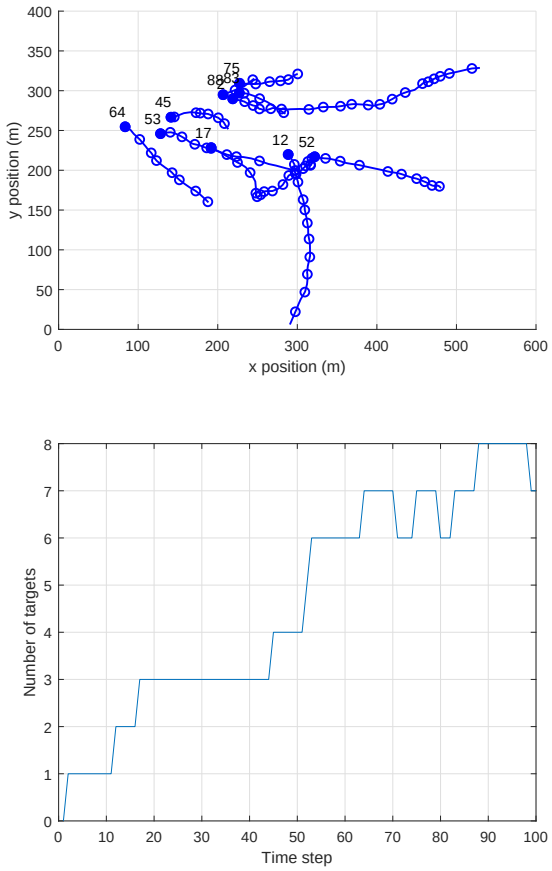


Figure 3: Trajectories considered in the simulation (top). There are ten trajectories. The positions of a trajectory every ten time steps is marked with circles, and the circle for the initial position is filled. The number next to the initial position indicates time step of birth. The resulting number of targets in the scene at each time step is shown in the the bottom figure.

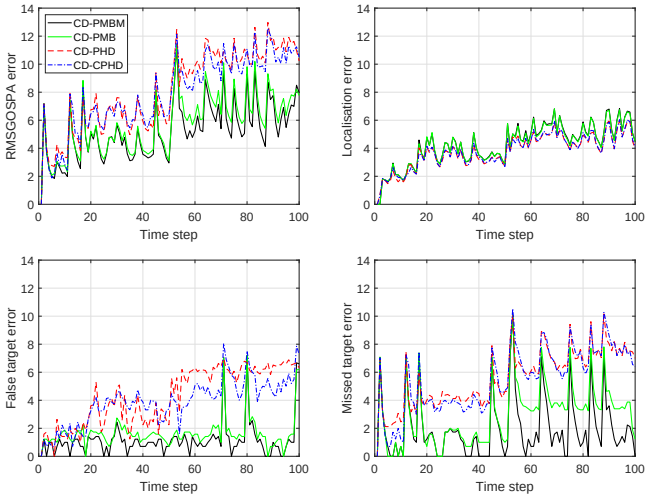


Figure 4: RMS-GOSPA error (m) for the position elements and its decomposition for the continuous-discrete filters. The CD-PMBM filter outperforms CD-PMB, CD-PHD and CD-CPHD filters. The difference in performance is mainly due to a lower number of missed and false targets.

threshold 10^{-5} and merging threshold 0.1. These filters use the estimators in [4, Sec. 9.5.4.4] and [4, Sec. 9.5.5.4], respectively. The CD-PMBM filter has $N_h = 200$, $\Gamma_p = 10^{-5}$, $\Gamma_{mbm} = 10^{-4}$, and $\Gamma_b = 10^{-5}$ and we also use ellipsoidal gating with threshold 20. We use Estimator 1 in [35], which selects the global hypotheses with highest weight and reports the mean of Bernoulli components whose existence is above a threshold set to 0.4. We also test the CD-PMBM filter with $N_h = 1$, which is referred to as CD-PMB filter, as it only propagates one multi-Bernoulli component. This filter can be considered as a version of the global nearest neighbour approach [47], as the filter only retains the most likely global hypothesis.

We evaluate the performance of the filters using the generalised optimal subpattern assignment (GOSPA) metric [48]. We use GOSPA on the position elements with the Euclidean metric as base metric, and parameters $p = 2$, $c = 10\,\mathrm{m}$ and $\alpha = 2$. Setting $\alpha = 2$ enables the decomposition of this metric into costs for localisation error for properly detected targets, missed and false targets [48, Prop. 1]. The resulting root mean square GOSPA (RMS-GOSPA) is shown in Figure 4. The CD-PMBM is the best performing filter, followed by the CD-PMB filter. From the GOSPA decomposition, we can observe that the main difference in performance between these two filters arises due to a lower number of missed targets in the CD-PMBM filter. The CD-PMBM filter is slightly better at false target errors than the CD-PMB filter, and quite similar in localisation errors. The CD-PHD and CD-CPHD filters fare worse than the CD-PMBM and CD-PMB filters. In particular, they have a higher number of false and missed targets. Also, the CD-CPHD filter has lower error than the CD-PHD filter mainly due to a lower number of false targets.

The running times of the Matlab implementations of the algorithms in computer with a 3.5 GHz Intel Xeon E5 processor are: 9.9 s (CD-PMBM), 0.7 s (CD-PMB), 1.4 s (CD-PHD) and 1.7 s (CD-CPHD). The CD-PMBM is the filter with best performance, but also with highest computational burden. The CD-PMB filter is the second best in performance and fastest algorithm in our implementations. CD-PHD and CD-CPHD filters are also fast algorithms, though their performance is far from the performance of the CD-PMBM and CD-PMB filters.

An additional benefit of the continuous-discrete multi-target model is that it can significantly speed up processing compared to a discrete time equivalent model that considers a fixed sampling time. In order to analyse this effect, we consider a discrete time model in which the sampling time is the minimum interval Δt_k considered in the previous simulation, which corresponds to the minimum Δt_k in Figure 2. The parameters of the discretised dynamic model, including birth, are obtained as indicated in this paper, but they do not change with time. The time steps at which we receive measurements are rounded to the nearest integer. In the resulting filtering recursions for the different multi-target filters, there is no update step at the time steps without measurements, but there is a prediction at all time steps. This discretisation applied to our scenario has 29374 time steps, which means that we have to perform a large number of prediction steps compared to the 100 predictions required with the use of the continuous-

Table I: RMS-GOSPA errors and their decompositions for the filters and different parameters

		CD-PMBM				CD-PMB				CD-PHD				CD-CPHD			
λ	λ_c	Tot.	Loc.	Fal.	Mis.	Tot.	Loc.	Fal.	Mis.	Tot.	Loc.	Fal.	Mis.	Tot.	Loc.	Fal.	Mis.
0.12	5	6.25	4.76	2.56	3.13	6.58	4.75	2.72	3.64	9.27	4.28	5.48	6.15	8.86	4.32	4.87	6.01
	10	6.38	4.81	2.55	3.35	7.19	4.75	2.83	4.60	9.38	4.28	5.37	6.39	9.14	4.30	4.98	6.34
	20	6.60	4.81	2.71	3.62	8.56	4.58	3.21	6.48	9.58	4.22	5.27	6.80	9.51	4.23	5.13	6.81
0.08	5	5.50	4.53	1.61	2.66	5.58	4.54	1.81	2.70	8.52	4.04	4.85	5.71	8.00	4.10	4.14	5.49
	10	5.63	4.50	1.72	2.92	6.12	4.46	2.01	2.68	8.62	3.99	4.77	5.97	8.35	4.02	4.24	5.96
	20	5.82	4.53	1.82	3.17	7.50	4.33	2.13	5.73	8.98	3.94	4.63	6.60	8.72	3.96	4.20	6.54
0.04	5	3.46	2.74	1.17	1.76	3.52	2.75	1.32	1.75	4.40	2.48	1.53	3.29	4.59	2.56	2.36	2.99
	10	3.48	2.67	1.21	1.85	3.62	2.70	1.42	1.95	4.48	2.47	1.67	3.34	4.69	2.52	2.39	3.16
	20	3.65	2.70	1.37	2.03	5.23	2.48	1.60	4.31	4.71	2.45	1.99	3.50	4.88	2.52	2.44	3.39

discrete multi-target model. The corresponding GOSPA errors for the PMBM, PMB, PHD, and CPHD filters applied in this form are basically unaltered compared to the results of the filters with time varying sampling time, shown in Figure 4, so the results are not plotted. The running times of the algorithms are: 178.8 s (PMBM), 111.2 s (PMB), 76.7 s (PHD) and 85.8 s (CPHD). By comparing these running times with the running times of the continuous-discrete versions discussed above, we can see that continuous-discrete multi-target filters with variable sampling time show a remarkable decrease in the computational burden (over 18 times faster) compared to multi-target filters with fixed sampling time, which are the standard approach in the multi-target filtering literature, without affecting performance. This implies that the continuous-discrete filters are more suitable for developing real-time implementation than the discrete versions if measurements are asynchronous.

In order to provide more thorough results, we also show the RMS-GOSPA errors, along with the GOSPA error decomposition, considering all time steps for different values of the target arrival rate λ and clutter rate λ_c in Table I. In this table, "Tot.", "Loc.", "Fal." and "Mis." stand for total cost, localisation cost, false target cost and missed target cost, respectively. Simulations with different λ consider different ground truth drawn from the true multi-target system, see Algorithm 1. The CD-PMBM filter is always the best performing algorithm. The CD-PMB filter is the second best algorithm except for $\lambda = 0.04$ and $\lambda_c = 20$, in which it is outperformed by the CD-PHD and CD-CPHD filters. The CD-CPHD filter provides lower errors than the CD-PHD filter except for $\lambda = 0.04$, where the CD-PHD filter is slightly better.

VII. Conclusions

In this paper, we have proposed a probabilistic model for multiple target appearance, dynamics and disappearance in continuous time and we have derived the corresponding discretised multi-target model. We have also obtained the best fitting Gaussian PPP fit to the discretised birth density for the Wiener velocity model by minimising the Kullback-Leibler divergence.

The discretised multiple target model and the Kullback-Leibler divergence minimisation of the birth process provide the required models for performing continuous-discrete multiple target filtering using Gaussian distributions. In particular, in this paper, we present and analyse the Gaussian implementations of the continuous-discrete PMBM, PHD and CPHD filters. An important benefit of the continuous-discrete multi-target filters is that they can significantly lower the computational complexity compared to multi-target filters with fixed sampling interval without affecting performance when the time interval between measurements is non-uniform.

We think there are many lines of future work. The derived continuous-discrete multi-target models can also be used to extend the CD-PMBM, CD-PHD and CD-CPHD to sets of trajectories [49]–[51] to obtain complete trajectory information of the targets at the sampled time steps. It is also interesting to consider Gaussian implementations of these filters with dynamic/measurement models that are nonlinear/non-Gaussian, which in practice requires suitable linearisations [23], [39], [52]. More generally, with the obtained general birth process equations in Section III, it is also possible to develop sequential Monte Carlo implementations [53] of the CD-PMBM, CD-PHD and CD-CPHD filters, which do not require Gaussian single-target densities and KLD minimisations.

Appendix A

In this appendix, we prove (15), which is the distribution of the time lags of appearance for new born targets at time step k, which denotes the difference between time t_k and the time of appearance. From Assumption A1, we know that the time lags of appearing targets are IID distributed in the interval $[0, \Delta t_k)$ [28]. However, not all of the appearing targets in the interval $(t_{k-1}, t_k]$ are born at time step k, as some of them may disappear before time t_k. Using A1 and A3, the joint distribution of the time lag t and life span τ of an appearing target, which denotes the time from appearance to disappearance, is

$$p(t, \tau) = \frac{1}{\Delta t_k} \chi_{[0,\Delta t_k)}(t)\, \mu e^{-\mu\tau}. \tag{79}$$

A target that appeared in $(t_{k-1}, t_k]$ is born at time step k if $\tau > t$. Therefore, the distribution of the time lag for a new born target at time k is

$$\begin{aligned} p_k(t) &\triangleq p(t \mid \tau > t) \\ &= \int p(t, \tau \mid \tau > t)\, d\tau \end{aligned} \tag{80}$$

where

$$p(t, \tau \mid \tau > t) = \frac{p(t, \tau)}{p(\tau > t)}, \quad \tau > t. \tag{81}$$

We first calculate the probability

$$p\left(\tau > t\right) = \int_0^{\Delta t_k} \left[\int_t^{\infty} \frac{1}{\Delta t_k} \mu e^{-\mu\tau} d\tau \right] dt \tag{82}$$

$$= \int_0^{\Delta t_k} \frac{1}{\Delta t_k} e^{-\mu t} dt \tag{83}$$

$$= \frac{1 - e^{-\mu \Delta t_k}}{\mu \Delta t_k}. \tag{84}$$

Substituting this result into (81), we obtain

$$p\left(t, \tau \mid \tau > t\right) = \frac{\mu}{1 - e^{-\mu \Delta t_k}} \chi_{[0, \Delta t_k)}\left(t\right) \mu e^{-\mu\tau} \tag{85}$$

for $\tau > t$. Then, (80) becomes

$$p_k\left(t\right) = \int_t^{\infty} p\left(t, \tau \mid \tau > t\right) d\tau \tag{86}$$

$$= \frac{\mu}{1 - e^{-\mu \Delta t_k}} e^{-\mu t} \chi_{[0, \Delta t_k)}\left(t\right), \tag{87}$$

which completes the proof of (15).

Appendix B

In this appendix, we prove Proposition 1. Using (18), we obtain

$$\mathrm{D}\left(b_k \middle\| \widetilde{b}_k\right)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \int b_k\left(\left\{x^1, ..., x^n\right\}\right) \log \frac{b_k\left(\left\{x^1, ..., x^n\right\}\right)}{\widetilde{b}_k\left(\left\{x^1, ..., x^n\right\}\right)} dx^{1:n} \tag{88}$$

$$= \sum_{n=0}^{\infty} \mathcal{P}\left(n; \frac{\lambda}{\mu} \Delta t_k\right) \int \prod_{i=1}^{n} p_k\left(x^i\right) \times \log \frac{\mathcal{P}\left(n; \frac{\lambda}{\mu} \Delta t_k\right) \prod_{i=1}^{n} p_k\left(x^i\right)}{\mathcal{P}\left(n; \lambda_{b,k}\right) \prod_{i=1}^{n} \mathcal{N}\left(x^i; \overline{x}_{b,k}, P_{b,k}\right)} dx^{1:n} \tag{89}$$

$$= \sum_{n=0}^{\infty} \mathcal{P}\left(n; \frac{\lambda}{\mu} \Delta t_k\right) \log \frac{\mathcal{P}\left(n; \frac{\lambda}{\mu} \Delta t_k\right)}{\mathcal{P}\left(n; \lambda_{b,k}\right)} + \sum_{n=0}^{\infty} \mathcal{P}\left(n; \frac{\lambda}{\mu} \Delta t_k\right) \times \int \prod_{i=1}^{n} p_k\left(x^i\right) \log \frac{\prod_{i=1}^{n} p_k\left(x^i\right)}{\prod_{i=1}^{n} \mathcal{N}\left(x^i; \overline{x}_{b,k}, P_{b,k}\right)} dx^{1:n}. \tag{90}$$

The optimal value of $\lambda_{b,k}$ can be found by minimising the first term in the previous equation and the optimal values of $\overline{x}_{b,k}, P_{b,k}$ by minimising the second term. The first term coincides with the KLD between two Poisson distributions on variable n. The minimum is obtained by setting $\lambda_{b,k} = \frac{\lambda}{\mu} \Delta t_k$, as the two distributions are alike and the KLD is zero, which is the minimum attainable value for the KLD.

Now, we consider the second term, which we can write as

$$\sum_{n=0}^{\infty} \mathcal{P}\left(n; \frac{\lambda}{\mu} \Delta t_k\right) \int \prod_{i=1}^{n} p_k\left(x^i\right) \times \log \frac{\prod_{i=1}^{n} p_k\left(x^i\right)}{\prod_{i=1}^{n} \mathcal{N}\left(x^i; \overline{x}_{b,k}, P_{b,k}\right)} dx^{1:n}$$

$$= \left[\sum_{n=0}^{\infty} \mathcal{P}\left(n; \frac{\lambda}{\mu} \Delta t_k\right) n \right] \int p_k\left(x\right) \log \frac{p_k\left(x\right)}{\mathcal{N}\left(x; \overline{x}_{b,k}, P_{b,k}\right)} dx. \tag{91}$$

This term is proportional to the KLD between $p_k\left(\cdot\right)$ and $\mathcal{N}\left(x; \overline{x}_{b,k}, P_{b,k}\right)$. This KLD is minimised by moment matching [40], which implies that the optimal $\overline{x}_{b,k}$ and $P_{b,k}$ are given by (22) and (23). It should be noted that the result of Proposition 1, and also Lemma 6, are to be expected as both the Poisson and the Gaussian distribution belong to the exponential family of distributions for which this type of KLD minimisation is achieved by matching the sufficient statistics, which is the mean for the Poisson, and the mean and covariance matrix for the Gaussian [40].

Appendix C

In this appendix, we prove Proposition 2 by applying Proposition 1 to the Wiener velocity model, explained in Section IV-B. The parameter $\lambda_{b,k}$ is directly obtained from Proposition 1, so we proceed to compute the mean $\overline{x}_{b,k}$ and the covariance matrix $P_{b,k}$. The moments (36)-(39) of a truncated exponential distribution with density (15) can be computed analytically by solving the corresponding integrals [54], but details are not provided here.

In order to calculate the mean (31), we apply the law of total expectation [55] to (22) on the time lag t to yield

$$\overline{x}_{b,k} = \mathrm{E}\left[\mathrm{E}\left[x_k \left| t \right.\right]\right] \tag{92}$$

$$= \mathrm{E}\left[F_{(t)} \overline{x}_a\right] \tag{93}$$

$$= \mathrm{E}\left[\begin{pmatrix} I_d & tI_d \\ 0_d & I_d \end{pmatrix} \overline{x}_a\right] \tag{94}$$

$$= \begin{pmatrix} I_d & \mathrm{E}\left[t\right] I_d \\ 0_d & I_d \end{pmatrix} \overline{x}_a. \tag{95}$$

It should be noted in the previous equations, we have used that $\mathrm{E}\left[x_k \left| t \right.\right]$ is obtained from (11), and that $F_{(t)}$ for the Wiener velocity model is given by (26).

We apply the law of total covariance [55] to (23)

$$P_{b,k} = \mathrm{C}\left[\mathrm{E}\left[x_k \left| t \right.\right]\right] + \mathrm{E}\left[\mathrm{C}\left[x_k \left| t \right.\right]\right]. \tag{96}$$

We first calculate $\mathrm{C}\left[\mathrm{E}\left[x_k \left| t \right.\right]\right]$ as follows

$$\mathrm{C}\left[\mathrm{E}\left[x_k \left| t \right.\right]\right] = \mathrm{C}\left[\begin{pmatrix} I_d & tI_d \\ 0_d & I_d \end{pmatrix} \overline{x}_a\right] \tag{97}$$

$$= \mathrm{C}\left[\begin{pmatrix} \overline{p}_a + t\overline{v}_a \\ \overline{v}_a \end{pmatrix}\right] \tag{98}$$

$$= \mathrm{E}\left[\begin{pmatrix} \left(t - \mathrm{E}\left[t\right]\right) \overline{v}_a \\ 0_d \end{pmatrix} \begin{pmatrix} \left(t - \mathrm{E}\left[t\right]\right) \overline{v}_a^T & 0_d \end{pmatrix}\right] \tag{99}$$

$$= \begin{pmatrix} \mathrm{C}\left[t\right] \overline{v}_a \overline{v}_a^T & 0_d \\ 0_d & 0_d \end{pmatrix}. \tag{100}$$

We would like to clarify that, in order to obtain (100), we have used (28) and the definition

$$\mathrm{C}\left[g\left(t\right)\right] = \mathrm{E}\left[\left(g\left(t\right) - \mathrm{E}\left[g\left(t\right)\right]\right)\left(g\left(t\right) - \mathrm{E}\left[g\left(t\right)\right]\right)^T\right] \tag{101}$$

for any function $g\left(\cdot\right)$.

We proceed to calculate $\mathrm{E}\left[\mathrm{C}\left[x_k \left| t \right.\right]\right]$, which can be written as

$$\mathrm{E}\left[\mathrm{C}\left[x_k \left| t \right.\right]\right] = \mathrm{E}\left[F_{(t)} P_a F_{(t)}^T + Q_{(t)}\right] \tag{102}$$

$$= \mathrm{E}\left[F_{(t)} P_a F_{(t)}^T\right] + \mathrm{E}\left[Q_{(t)}\right]. \tag{103}$$

The first term in (103) is given by

$$\mathrm{E}\left[F_{(t)} P_a F_{(t)}^T\right]$$
$$= \mathrm{E}\left[\begin{pmatrix} I_d & tI_d \\ 0_d & I_d \end{pmatrix} P_a \begin{pmatrix} I_d & 0_d \\ tI_d & I_d \end{pmatrix}\right] \tag{104}$$
$$= \mathrm{E}\left[\begin{pmatrix} P_a^{pp} + t\left(P_a^{pv} + \left(P_a^{pv}\right)^T\right) + t^2 P_a^{vv} & P_a^{pv} + tP_a^{vv} \\ \left(P_a^{pv}\right)^T + tP_a^{vv} & P_a^{vv} \end{pmatrix}\right] \tag{105}$$
$$= \begin{pmatrix} A_{11} & P_a^{pv} + \mathrm{E}[t] P_a^{vv} \\ \left(P_a^{pv}\right)^T + \mathrm{E}[t] P_a^{vv} & P_a^{vv} \end{pmatrix} \tag{106}$$

where $A_{11} = P_a^{pp} + \mathrm{E}[t]\left(P_a^{pv} + \left(P_a^{pv}\right)^T\right) + \mathrm{E}\left[t^2\right] P_a^{vv}$ and we have used the decomposition of the covariance P_a in (29). The second term in (103) is given by

$$\mathrm{E}_t\left[Q_{(t)}\right] = \mathrm{E}_t\left[q\begin{pmatrix} \frac{t^3}{3} I_d & \frac{t^2}{2} I_d \\ \frac{t^2}{2} I_d & tI_d \end{pmatrix}\right] \tag{107}$$
$$= q\begin{pmatrix} \frac{\mathrm{E}[t^3]}{3} I_d & \frac{\mathrm{E}[t^2]}{2} I_d \\ \frac{\mathrm{E}[t^2]}{2} I_d & \mathrm{E}[t] I_d \end{pmatrix}. \tag{108}$$

Substituting (100), (103), (106) and (108) into (96) completes the proof of Proposition 2.

APPENDIX D

In this appendix we show the KLD minimisation properties of the Gaussian mixture birth model explained in (41). From (40), we consider the joint distribution of the state and the index $j \in \{1, ..., N_a\}$ over the mixture components to yield

$$p_k(x_k, j) = w_a^j p_k(x_k | j) \tag{109}$$

where

$$p_k(x_k | j) = \int_0^{\Delta t_k} \mathcal{N}\left(x_k; F_{(t)} \overline{x}_a^j, F_{(t)} P_a^j F_{(t)}^T + Q_{(t)}\right) \times p_k(t)\, dt. \tag{110}$$

This approach of considering a joint variable over the state and the index over a mixture component is standard in auxiliary particle filtering [7], [56]. Note that if we sum over variable j in (109), we obtain $p_k(x_k)$ in (40).

The birth PPP (17) considering the auxiliary variable becomes

$$b_k\left(\left\{\left(x_k^1, j^1\right), ..., \left(x_k^n, j^n\right)\right\}\right) = n! \mathcal{P}\left(n; \frac{\lambda}{\mu}\left(1 - e^{-\mu \Delta t_k}\right)\right) \times \prod_{i=1}^n \left[w_a^{j^i} p_k\left(x_k^i | j^i\right)\right]. \tag{111}$$

We seek the best Gaussian PPP approximation $\widetilde{b}_k(\cdot)$ of the form

$$\widetilde{b}_k\left(\left\{\left(x_k^1, j^1\right), ..., \left(x_k^n, j^n\right)\right\}\right) = n! \mathcal{P}(n; \lambda_{b,k}) \prod_{i=1}^n \left[w_b^{j^i} \mathcal{N}\left(x_k^i; \overline{x}_{b,k}^{j^i}, P_{b,k}^{j^i}\right)\right]. \tag{112}$$

The resulting PPP is given by the following lemma.

Lemma 6. *The Gaussian PPP density $\widetilde{b}_k(\cdot)$ of the form (112) that minimises the KLD $\mathrm{D}\left(b_k \middle\| \widetilde{b}_k\right)$, with $b_k(\cdot)$ given by (111), is parameterised by*

$$\lambda_{b,k} = \frac{\lambda}{\mu}\left(1 - e^{-\mu \Delta t_k}\right) \tag{113}$$
$$w_b^j = w_a^j \tag{114}$$
$$\overline{x}_{b,k}^j = \int x p_k(x | j)\, dx \tag{115}$$
$$P_{b,k}^j = \int \left(x - \overline{x}_{b,k}^j\right)\left(x - \overline{x}_{b,k}^j\right)^T p_k(x | j)\, dx \tag{116}$$

for $j \in \{1, ..., N_a\}$.

This lemma is proved in Appendix D-A. For the Wiener velocity model, we can compute the moments (115) and (116) using Proposition 2 for each j. Once we obtain (115) and (116), one can integrate out the auxiliary variables in (112) to yield (41).

A. Proof of Lemma 6

Taking into account that the auxiliary variable j is discrete, the KLD is given by

$$\mathrm{D}\left(b_k \middle\| \widetilde{b}_k\right)$$
$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{j^1, ..., j^n} \int b_k\left(\left\{\left(x^1, j^1\right), ..., \left(x^n, j^n\right)\right\}\right) \times \log \frac{b_k\left(\left\{\left(x^1, j^1\right), ..., \left(x^n, j^n\right)\right\}\right)}{\widetilde{b}_k\left(\left\{\left(x^1, j^1\right), ..., \left(x^n, j^n\right)\right\}\right)} dx^{1:n} \tag{117}$$
$$= \sum_{n=0}^{\infty} \mathcal{P}\left(n; \frac{\lambda}{\mu}\left(1 - e^{-\mu \Delta t_k}\right)\right) \log \frac{\mathcal{P}\left(n; \frac{\lambda}{\mu}\left(1 - e^{-\mu \Delta t_k}\right)\right)}{\mathcal{P}(n; \lambda_{b,k})} + \sum_{n=0}^{\infty} \mathcal{P}\left(n; \frac{\lambda}{\mu}\left(1 - e^{-\mu \Delta t_k}\right)\right) \sum_{j^1, ..., j^n} \int \prod_{i=1}^n \left[w_a^{j^i} p_k\left(x^i | j^i\right)\right] \times \log \frac{\prod_{i=1}^n \left[w_a^{j^i} p_k\left(x^i | j^i\right)\right]}{\prod_{i=1}^n \left[w_b^{j^i} \mathcal{N}\left(x^i; \overline{x}_{b,k}^{j^i}, P_{b,k}^{j^i}\right)\right]} dx^{1:n} \tag{118}$$
$$= \sum_{n=0}^{\infty} \mathcal{P}\left(n; \frac{\lambda}{\mu}\left(1 - e^{-\mu \Delta t_k}\right)\right) \log \frac{\mathcal{P}\left(n; \frac{\lambda}{\mu}\left(1 - e^{-\mu \Delta t_k}\right)\right)}{\mathcal{P}(n; \lambda_{b,k})} + \left[\sum_{n=0}^{\infty} \mathcal{P}\left(n; \frac{\lambda}{\mu}\left(1 - e^{-\mu \Delta t_k}\right)\right) n\right] \times \sum_j \int w_a^j p_k(x | j) \log \frac{w_a^j p_k(x | j)}{w_b^j \mathcal{N}\left(x; \overline{x}_{b,k}^j, P_{b,k}^j\right)} dx \tag{119}$$
$$= \sum_{n=0}^{\infty} \mathcal{P}\left(n; \frac{\lambda}{\mu}\left(1 - e^{-\mu \Delta t_k}\right)\right) \log \frac{\mathcal{P}\left(n; \frac{\lambda}{\mu}\left(1 - e^{-\mu \Delta t_k}\right)\right)}{\mathcal{P}(n; \lambda_{b,k})} + \left[\sum_{n=0}^{\infty} \mathcal{P}\left(n; \frac{\lambda}{\mu}\left(1 - e^{-\mu \Delta t_k}\right)\right) n\right]$$

$$\times \left[\sum_j w_a^j \log \frac{w_a^j}{w_b^j} + \int p_k\left(x \,|j\,\right) \log \frac{p_k\left(x \,|j\,\right)}{\mathcal{N}\left(x; \overline{x}_{b,k}^j, P_{b,k}^j\right)} dx \right]. \tag{120}$$

By standard KLD minimisations, for each term considering $\lambda_{b,k}$, $w_b^{j^i}$,$\overline{x}_{b,k}^j$, and $P_{b,k}^j$, we get the result in Lemma 6.

REFERENCES

[1] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Artech House, 1999.

[2] K. Granström, L. Svensson, S. Reuter, Y. Xia, and M. Fatemi, "Likelihood-based data association for extended object tracking using sampling methods," *IEEE Transactions on Intelligent Vehicles*, vol. 3, no. 1, pp. 30–45, March 2018.

[3] F. Meyer, T. Kropfreiter, J. L. Williams, R. Lau, F. Hlawatsch, P. Braca, and M. Z. Win, "Message passing algorithms for scalable multitarget tracking," *Proceedings of the IEEE*, vol. 106, no. 2, pp. 221–259, Feb. 2018.

[4] R. P. S. Mahler, *Advances in Statistical Multisource-Multitarget Information Fusion*. Artech House, 2014.

[5] S. Challa, M. R. Morelande, D. Musicki, and R. J. Evans, *Fundamentals of Object Tracking*. Cambridge University Press, 2011.

[6] M. R. Morelande, C. M. Kreucher, and K. Kastella, "A Bayesian approach to multiple target detection and tracking," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1589–1604, May. 2007.

[7] A. F. García-Fernández, J. Grajal, and M. R. Morelande, "Two-layer particle filter for multiple target detection and tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 3, pp. 1569–1588, July 2013.

[8] 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.

[9] S. Särkkä, A. Vehtari, and J. Lampinen, "Rao-Blackwellized particle filter for multiple target tracking," *Information Fusion*, vol. 8, no. 1, pp. 2–15, Jan. 2007.

[10] 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.

[11] K. Granström, C. Lundquist, and O. Orguner, "Extended target tracking using a Gaussian-mixture PHD filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 4, pp. 3268–3286, Oct. 2012.

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

[13] J. Houssineau and D. E. Clark, "Multitarget filtering with linearized complexity," *IEEE Transactions on Signal Processing*, vol. 66, no. 18, pp. 4957–4970, Sept. 2018.

[14] F. E. De Melo and S. Maskell, "A CPHD approximation based on a discrete-Gamma cardinality model," *IEEE Transactions on Signal Processing*, vol. 67, no. 2, pp. 336–350, Jan. 2019.

[15] K. Granström, M. Fatemi, and L. Svensson, "Poisson multi-Bernoulli mixture conjugate prior for multiple extended target filtering," *accepted in IEEE Transactions on Aerospace and Electronic Systems*, 2019.

[16] M. Mallick, M. Morelande, and L. Mihaylova, "Continuous-discrete filtering using EKF, UKF, and PF," in *15th International Conference on Information Fusion*, July 2012, pp. 1087–1094.

[17] L. M. Millefiori, P. Braca, and P. Willett, "Consistent estimation of randomly sampled Ornstein-Uhlenbeck process long-run mean for long-term target state prediction," *IEEE Signal Processing Letters*, vol. 23, no. 11, pp. 1562–1566, Nov. 2016.

[18] A. F. García-Fernández and J. Grajal, "Asynchronous particle filter for tracking using non-synchronous sensor networks," *Signal Processing*, vol. 91, no. 10, pp. 2304–2313, Oct. 2011.

[19] G. Li, W. Yi, S. Li, B. Wang, and L. Kong, "Asynchronous multi-rate multi-sensor fusion based on random finite set," *Signal Processing*, vol. 160, pp. 113–126, 2019.

[20] M. Üney, L. M. Millefiori, and P. Braca, "Data driven vessel trajectory forecasting using stochastic generative models," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 8459–8463.

[21] X. Lagorce, C. Meyer, S. Ieng, D. Filliat, and R. Benosman, "Asynchronous event-based multikernel algorithm for high-speed visual features tracking," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 8, pp. 1710–1720, Aug. 2015.

[22] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. Academic Press, 1970.

[23] S. Särkkä and A. Solin, *Applied Stochastic Differential Equations*. Cambridge University Press, 2019.

[24] S. Särkkä, "Recursive Bayesian inference on stochastic differential equations," Ph.D. dissertation, 2006.

[25] Y. Bar-Shalom, T. Kirubarajan, and X. R. Li, *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc., 2001.

[26] P. Coscia, P. Braca, L. M. Millefiori, F. A. N. Palmieri, and P. Willett, "Multiple Ornstein-Uhlenbeck processes for maritime traffic graph representation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 5, pp. 2158–2170, Oct. 2018.

[27] S. Särkkä, "On unscented Kalman filtering for state estimation of continuous-time nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1631–1641, Sep. 2007.

[28] L. Kleinrock, *Queueing Systems*. John Wiley & Sons, 1976.

[29] K. Andersson, F. Carr, W. D. Hall, N. Pujet, and E. Feron, "Analysis, modeling, and ccontrol of ground operations at hub airports," in *Air transportation systems engineering*, 2001.

[30] V. G. Kulkarni, *Modeling and analysis of stochastic systems*. Chapman & Hall/CRC, 2016.

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

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

[33] 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, Nov. 2006.

[34] 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, July 2007.

[35] A. F. García-Fernández, J. L. Williams, K. Granström, and L. Svensson, "Poisson multi-Bernoulli mixture filter: direct derivation and implementation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1883–1901, Aug. 2018.

[36] A. F. García-Fernández and B.-N. Vo, "Derivation of the PHD and CPHD filters based on direct Kullback-Leibler divergence minimization," *IEEE Transactions on Signal Processing*, vol. 63, no. 21, pp. 5812–5820, Nov. 2015.

[37] J. L. Williams, "An efficient, variational approximation of the best fitting multi-Bernoulli filter," *IEEE Transactions on Signal Processing*, vol. 63, no. 1, pp. 258–273, Jan. 2015.

[38] P. Axelsson and F. Gustafsson, "Discrete-time solutions to the continuous-time differential Lyapunov equation with applications to Kalman filtering," *IEEE Transactions on Automatic Control*, vol. 60, no. 3, pp. 632–643, March 2015.

[39] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge University Press, 2013.

[40] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[41] M. Raitoharju, A. F. García-Fernández, and R. Piché, "Kullback-Leibler divergence approach to partitioned update Kalman filter," vol. 130, pp. 289–298, Jan. 2017.

[42] H. W. Sorenson and D. L. Alspach, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, vol. 7, no. 4, pp. 465–479, July 1971.

[43] T. Kurien, "Issues in the design of practical multitarget tracking algorithms," in *Multitarget-Multisensor Tracking: Advanced Applications*, Y. Bar-Shalom, Ed. Artech House, 1990.

[44] R. O. Lane, M. Briers, T. M. Cooper, and S. R. Maskell, "Efficient data structures for large scale tracking," in *17th International Conference on Information Fusion*, July 2014, pp. 1–8.

[45] K. G. Murty, "An algorithm for ranking all the assignments in order of increasing cost." *Operations Research*, vol. 16, no. 3, pp. 682–687, 1968.

[46] H. W. Kuhn, "The Hungarian method for the assignment problem," vol. 2, pp. 83–97, 1955.

[47] S. S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *IEEE Aerospace and Electronic Systems Magazine*, vol. 19, no. 1, pp. 5–18, Jan. 2004.

[48] A. S. Rahmathullah, A. F. García-Fernández, and L. Svensson, "Generalized optimal sub-pattern assignment metric," in *20th International Conference on Information Fusion*, 2017.
[49] A. F. García-Fernández, L. Svensson, and M. R. Morelande, "Multiple target tracking based on sets of trajectories," *accepted for publication in IEEE Transactions on Aerospace and Electronic Systems*, 2019. [Online]. Available: https://arxiv.org/abs/1605.08163
[50] K. Granström, L. Svensson, Y. Xia, J. L. Williams, and A. F. García-Fernández, "Poisson multi-Bernoulli mixture trackers: continuity through random finite sets of trajectories," in *21st International Conference on Information Fusion*, 2018.
[51] A. F. García-Fernández and L. Svensson, "Trajectory PHD and CPHD filters," *IEEE Transactions on Signal Processing*, vol. 67, no. 22, pp. 5702–5714, Nov 2019.
[52] A. F. García-Fernández, L. Svensson, M. R. Morelande, and S. Särkkä, "Posterior linearization filter: principles and implementation using sigma points," *IEEE Transactions on Signal Processing*, vol. 63, no. 20, pp. 5561–5573, Oct. 2015.
[53] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb. 2002.
[54] A. Jeffrey, *Handbook of Mathematical Formulas and Integrals*. Elsevier, 1995.
[55] S. Ross, *A First Course in Probability*, 8th ed. Prentice-Hall, 2010.
[56] M. K. Pitt and N. Shephard, "Filtering via simulation: Auxiliary particle filters," *Journal of the American Statistical Association*, vol. 94, no. 446, pp. 590–599, Jun. 1999.

Ángel F. García-Fernández received the telecommunication engineering degree (with honours) and the Ph.D. degree from Universidad Politécnica de Madrid, Madrid, Spain, in 2007 and 2011, respectively.
He is currently a Lecturer in the Department of Electrical Engineering and Electronics at the University of Liverpool, Liverpool, U.K. He previously held postdoctoral positions at Universidad Politécnica de Madrid, Chalmers University of Technology, Gothenburg, Sweden, Curtin University, Perth, Australia, and Aalto University, Espoo, Finland. His main research activities and interests are in the area of Bayesian inference, with emphasis on nonlinear dynamic systems and multiple target tracking. He has received paper awards at the International Conference on Information Fusion in 2017 and 2019.

Simon Maskell received the M.A., M.Eng., and Ph.D. degrees in engineering from the University of Cambridge, Cambridge, U.K., in 1998, 1999, and 2003, respectively.
Prior to 2013, he was a Technical Manager for Command, Control and Information Systems with QinetiQ, U.K. Since 2013, he has been a Professor of Autonomous Systems with the University of Liverpool, Liverpool, U.K. His research interests include Bayesian inference applied to signal processing, multitarget tracking, data fusion, and decision support with particular emphasis on the application of sequential Monte Carlo methods in challenging data science contexts.