A filter for distinguishable and independent populations

E. D. Delande, J. Houssineau, D. E. Clark

Abstract—This article introduces a multi-object filter for the resolution of joint detection/tracking problems involving multiple targets, derived from the novel Bayesian estimation framework for stochastic populations. Fully probabilistic in nature, the filter for Distinguishable and Independent Stochastic Populations (DISP) exploits two exclusive probabilistic representations for the potential targets. The distinguishable targets are those for which individual information is available through past detections; they are represented by individual tracks. The indistinguishable targets are those for which no individual information is available yet; they are represented collectively by a single stochastic population. Assuming that targets are independent, and adopting the “at most one measurement per scan per target” rule, the DISP filter propagates the set of all possible tracks, with associated credibility, based on the sequence of measurement sets collected by the sensor so far. A few filtering approximations, aiming at curtailing the computational cost of a practical implementation, are also discussed.

I. INTRODUCTION

Multi-target tracking became a topic of interest and a field of active research in the 1970s, when aerospace or ground-based surveillance scenarios required the estimation of the states (position, velocity, etc.) of an unknown and time-varying number of targets using noisy sensor observations [1], [2]. Since then, the solving of multi-target estimation problems within the Bayesian paradigm has been enjoying considerable developments, spanning from advances in the general estimation framework to the derivation of practical detection and tracking algorithms. The track-based and population-based approaches, detailed in the following paragraphs, have distinct and exclusive advantages, and thus enjoy popularity of their own in the resolution of practical surveillance scenarios.

Initial contributions [1]–[3] focussed on track-based approaches, in which individual information about a potential target is conserved in a track, updated across time with the stream of sensor observations. An explicit data association scheme is usually necessary to link the current observations to the potential targets maintained by the tracks, and heuristic approaches are required to create and discard tracks based on the operator’s knowledge of the target behaviour and the sensor characteristics. Well-established methods include Multiple Hypothesis Tracking (MHT) and Joint Probabilistic Density Association (JPDA) approaches [4]. The proper tuning of these heuristics is a well-known issue in the practical implementation stage, for it usually depends on parameters that are heavily scenario-dependent and may not be readily expressible from the operator’s wishes. For example, “m-out-of-n techniques” discard tracks that have not been associated to any observations during m of the last n time steps and must be tuned accordingly to the specifics of the sensor system; however, operator should be able to conserve or discard a track based on an objective assessment of the track credibility – that is, how likely this track is to account for an underlying target.

The more recent Finite Set Statistics (FISST) methodology [5] provides an alternative approach relying on a full probabilistic description of all the system uncertainties. A sound and well-developed mathematical framework underpinned by the point process theory [6]–[8], it shifts the focus from individual targets to the target population, describing the uncertainties on the number of targets and their states are encapsulated in a single random object, a (simple) point process or Random Finite Set (RFS). The derivation of filtering solutions follows a rigorous and principled approach which naturally incorporates target birth and death, i.e. without further need for ad hoc solutions, and avoids explicit data associations. During the last decade, it has produced inexpensive algorithms such as the Probability Hypothesis Density (PHD) [9], the Cardinalized Probability Hypothesis Density (CPHD) [5], or the multi-Bernoulli [10] filters, which perform remarkably well for the tracking of multiple targets in challenging environments and are receiving a growing attention from the tracking community. However, as point processes were essentially developed for the stochastic description of a collection of indiscernible objects [8] – in the sense that they rely on symmetric probabilities for which the order of the objects in the collection is irrelevant – the multi-target representation maintained by the usual RFS-based filters lacks a natural ordering as well. In particular, the filtering framework does not provide a natural way to relate two individuals from successive populations as representatives of a single target. As a consequence, RFS-based filters do not maintain an inherent track history for the identified targets, although recent developments aims at introducing target labelling within the FISST framework with the construction of the Labeled Multi-Bernoulli filter [11].

The multi-target filter presented in this paper is derived from the very recent Bayesian estimation framework for stochastic populations [12], which has been previously applied in [13], [14] and aims at addressing some of the shortcomings of the traditional track-based and RFS-based approaches. Fully probabilistic in nature, the DISP filter does not rely on heuristics and ad hoc techniques; by propagating the right amount of information on each (potential) target, it does maintain specific information on a target (e.g. track) whenever appropriate. More specifically, the level of information maintained by the filter on any particular target depends on its status:

- **distinguishable** individuals are those for which specific information is available, usually through the sensor observation process (past or present). For example, if any observation is assumed to originate from at most one tar-
get, “the individual that produced observation \( z_3 \) at time \( t = 3 \), \( z_4 \) at time \( t = 4 \) and has been miss-detected ever since” characterises a **unique** (potential) target without ambiguity. Such an individual is adequately described by a *track*, with a specific probability distribution describing its current state; • *indistinguishable* individuals, on the other hand, are those that are known only as members of a larger population whose individuals share a *common* description. For example, “the individuals that arrived through the North-West ten time steps ago, and have been miss-detected ever since” characterises a *population* of (potential) targets for which nothing is known on a particular individual so far. Such a population is adequately described by a cardinality distribution describing its size, and a *single* probability distribution describing the state of each of its members.

The paper is organized as follows. Section II introduces some terminology of the novel estimation framework [12], describes a generic multi-object estimation problem in the context of stochastic populations, and lists the *modelling assumptions* that shape the restricted class of problems – albeit covering a broad range of practical multi-target surveillance scenarios – that can be addressed by the DISP filter. General concepts and notations exploited in the paper follow in Section III. Section IV then explains the nature of the information propagated by the DISP filter, drawing from the representation of stochastic populations. The Bayesian mechanisms of (time) prediction and (data) update for the DISP filter are successively detailed in Sections V and VI. Section VII then discusses the filter’s output and its exploitation in surveillance scenarios. Section VIII focuses on the implementation of the DISP filter; in particular, it proposes several *filtering approximations* that may be implemented or not on an individual basis, depending on the operator’s computational resources and performance requirements. Proofs and algorithms are given in the appendices.

**II. BAYESIAN ESTIMATION WITH STOCHASTIC POPULATIONS: GENERAL CONCEPTS AND MODELLING ASSUMPTIONS**

In this section, we provide a thorough description of a typical multi-target surveillance scenario for which the DISP filter has been developed, and how it relates to the Bayesian estimation framework for stochastic populations [12], through (a) the interactions of the individuals of interest – e.g. the *targets* – with the surveillance scene (Section II-A), (b) the interactions of the sensor system with the surveillance scene (Section II-B), and (c) the nature of the *a priori* information available to the operator (Section II-C). In essence, the DISP filter was designed to address detection and tracking problems involving *independent* targets, and with *modelling assumptions* that are (a) well-identified and interpretable in a physical sense, (b) as few as possible, and (c) kept well-separated from the *filtering approximations*.

**A. Surveillance activity**

In the context of this paper, a surveillance activity is an estimation problem in which an operator wishes to gather information about the individuals of some population \( \mathcal{X} \) of interest (e.g. enemy vehicles) while they lie in some defined region of the physical space called the surveillance scene (e.g. the surroundings of a military facility of strategic importance) and over a given period of time, usually with a finite horizon\(^1\). While the existence of each individual is assumed certain throughout the period of time covered by the scenario, its presence in the surveillance scene *is not*; that is, an individual may possibly enter/leave the surveillance scene any number of times during the scenario, such that its presence or absence at any time is unknown to the operator and is part of the estimation problem.

The time flow is indexed by some integer \( t \), and the surveillance activity starts at \( t = 0 \). At any time \( t \geq 0 \), each individual within the surveillance scene is represented via a *state* \( x \), belonging to some target state space \( \mathbf{X}_t \), describing physical and measurable characteristics which are unknown and of interest to the operator. The nature of the target state space \( \mathbf{X}_t \) varies considerably depending on the nature of the surveillance activity. It is assumed to be a bounded subset of \( \mathbb{R}^d_t \), such that any state \( x \in \mathbf{X}_t \) is a real vector of (finite) dimension \( d_t \), with a mixture of continuous components such as position and/or velocity coordinates, and discrete components such as a vehicle type, a level of threat, etc. Conversely, the individuals currently absent from the surveillance scene assume the “empty state” \( \psi \); in consequence, each individual of the population \( \mathcal{X} \) is represented by some point in the augmented state space \( \mathbf{X}_t = \{ \psi \} \cup \mathbf{X}_t \). Note that the state space may evolve across time, for the operator may wish, for example, to estimate additional characteristics or modify the physical boundaries of the surveillance scene at some point during the scenario.

The goal of the surveillance activity is then for the operator to estimate, at any time \( t \geq 0 \) relevant to the scenario, the state of *every* individual of the population \( \mathcal{X} \) in the augmented state space, i.e.: (a) whether the individual is currently in the scene, and (b) if so, its characteristics. A *potential* individual of the population, as identified by the DISP filter, is also called a *target*.

In the context of the DISP filter, we shall consider the following assumptions.

**Modelling assumptions.** *Individuals in the population \( \mathcal{X} \) of interest:*

\( \text{(M1) behave independently;} \)
(\( \text{(M2) enter the scene at most once during the scenario;} \)
(\( \text{(M3) all have state } \psi \text{ before } t = 0.} \)

Assumption (M1), while fairly usual in multi-target detection and tracking problems, imposes significant restrictions on the class of surveillance problems that can be adequately addressed with the DISP filter. Assumption (M2) has more limited consequences, for it implies that if the same individual enters twice in the surveillance scene during the scenario, it will be treated as two different individuals for estimation purposes. Assumption (M3), with obvious implications on

\(^{1}\text{The surveillance scene and the sensor’s field of view (Section II-B) are distinct concepts.} \)
the initialisation of the DISP filter, will be discussed in Section VIII.

Following these assumptions, the population $X$ shall be decomposed at any time $t \geq 0$ as

$$X = X_{t}^\psi \cup X_{t}^\bullet \cup X_{t}^{\dagger},$$  \hspace{1cm} (1)

where $a) X_{t}^\psi$ are the individuals who have not entered the scene yet (and may never do so), $b) X_{t}^\bullet$ are the appearing individuals, i.e. those who have just entered the scene, and $c) X_{t}^{\dagger}$ are the remaining individuals (they may still be in the scene or may have already left it). In addition, the individuals in $X$ who have already entered the scene at time $t$ are denoted by $X_{t}^0$, i.e.

$$X_{t}^0 = X_{t}^\bullet \cup X_{t}^{\dagger}. \hspace{1cm} (2)$$

B. Sensor system and observation process

At any time $t \geq 0$, the surveillance scene is observed by some sensor system, providing information on the individuals in $X$, through a collection of observations immediately collected by the operator. An observation (or measurement) $z$ belongs to some observation space $Z_t$ and describes physical quantities measurable by the sensor system and relevant to an individual’s characteristics of interest (e.g. range, bearing, angle velocity, etc.). The observation space is assumed to be some bounded subset of $\mathbb{R}^{d_z}$, such that any observation $z \in Z_t$ is a real vector of (finite) dimension $d_z$. As for the target states, the observations may have a mixture of continuous and discrete components in the most general form, although a discrete observation space is sufficient to interpret the output of many sensors (e.g. pixels for a camera or resolution cells for a radar). Note that the observation space may not be constant throughout the scenario depending on the context, e.g. if sensors of different nature are exploited at different times. The measurements collected by the operator may also contain spurious observations or false alarms.

Modelling assumptions. The observation process at any time $t \geq 0$ is such that:

(M4) an individual produces at most one observation (if not, it is miss-detected);

(M5) an observation originates from at most one individual (if not, it is a false alarm);

(M6) individuals outside the scene produce no observations;

(M7) observations are produced independently;

(M8) observations are distinct;

(M9) the number of observations is finite.

Assumption (M8) guarantees that the observations collected by the operator at any time $t \geq 0$ can be described by a (possibly empty) observation set $Z_t = \{z_1, \ldots, z_n\}$. In order to account for the miss-detection of individuals by the sensor, it is convenient to introduce the “empty observation” $\phi$ and define, at any time $t \geq 0$, a) the augmented observation set $\bar{Z}_t = \{\phi\} \cup Z_t$, and b) the augmented observation space $\bar{Z}_t = \{\phi\} \cup Z_t$.

While the assumptions above are fairly usual in multi-target detection and tracking problems, another one is necessary in the context of the DISP filter and deserves special attention.

Modelling assumptions. First detection

(M10) an individual is detected upon entering the scene.

Assumption (M10) has important consequences on the structure of the DISP filter, for it simplifies significantly the nature of the propagated information (see Section IV-A). It also implies that, for estimation purposes, a newly detected target is considered as the representation of an individual that has just appeared in the scene (i.e. a member of $X_{t}^\bullet$ in Eq. (1)); therefore, the duration between the entrance of an individual in the scene and its first detection cannot be estimated by the DISP filter.

Note that, given Assumption (M6), no information is acquired at time $t \geq 0$ on the population $X_{t}^\psi$, i.e. the individuals that have not entered the scene yet (see Eq. 1). The DISP filter aims at estimating the remaining population $X_{t}^0$ (see Eq. 2).

C. Prior information

In addition to the observations collected from the sensor, the operator may possess and exploit some prior information on the population $X$ to produce and maintain an up-to-date estimation.

Modelling assumptions. Prior information (time $t \geq 0$)

(M11) global information on the population $X_{t}^\bullet$ is available, but no specific information is available on any of its individuals.

Assumption (M11) states that the operator only possesses some global knowledge, however informative, related to the population of appearing individuals as a whole – for example, no more than two individuals are expected to enter the scene between two steps, the individuals are more likely to appear at the physical boundaries of the surveillance scene, etc. They do not, however, have any information specific to any of the appearing individuals – for example, that one individual is likely to enter the scene from the North, another one from the South, etc.

D. Information propagation

In the context of Bayesian filtering, the information on the population $X$ acquired by the operator is successively $a)$ propagated across time (prediction step, see Section V), and $b)$ updated with the sensor observations (update step, see Section VI).

In order to propagate the acquired information, the knowledge of the operator on the propagation of the system uncertainties – i.e. how do the individuals evolve across time, and how do the sensor observations relate to the individuals’s state – must be fed to the filter. We shall therefore assume that:

Modelling assumptions. At any time $t \geq 0$, the knowledge of the operator about:
(M12) the evolution of the individuals in $X$ since time $t-1$ is described by a Markov kernel $m_{t-1,t}$ (detailed later);

(M13) the observation process is described by a likelihood $g_t(z,\cdot)$ for any $z \in \mathcal{Z}_t$, and a probability of false alarm $p_{f,a,t}$ (detailed later).

Assumptions (M12) and (M13) will be exploited and further discussed in Sections V and VI, respectively.

III. General notations

In this section, we introduce general notations that will play an important role in the transformation of the single-target probability distributions within the Bayesian paradigm.

A. Measure theory

Henceforth, $\mathcal{M}(E)$ (resp. $\mathcal{P}(E)$) will stand for the set of finite positive measures (resp. probability measures) on a given measurable space $(E, \mathcal{E})$. Additionally, the Banach space of all bounded and measurable functions equipped with the uniform norm $\| \cdot \|$ will be denoted $B(E)$, and we write $\mu(f) = \int \mu(dx)f(x)$ for any $\mu \in \mathcal{M}(E)$ and any test function $f \in B(E)$.

We shall also use the notation $dx$, defined for any point $x \in X_t$, where $\mathcal{B}_X$ is the Borel $\sigma$-algebra on $X_t$. If $x \in X_t$, $dx$ denotes some Borel set in $\mathcal{B}_X$, containing $x$; if $x = \psi$, $dx$ denotes the Borel set $\{\psi\}$.

Finally, for some region $B \subset \mathcal{B}_X$, the indicator function $1_B$ is the function on $X$ defined by

$$1_B(x) = \begin{cases} 1, & x \in B, \\ 0, & x \notin B. \end{cases}$$

(3)

B. Markov kernel

A bounded positive integral operator $Q$ from a measurable space $E$ into a measurable space $E'$ is an operator $f \mapsto Q(f)$ from $B(E)$ to $B(E')$ such that the functions

$$x \mapsto Q(f)(x) = \int_{E'} Q(x,dy) f(y)$$

are bounded and measurable for some measure $Q(x,\cdot) \in \mathcal{M}(E')$. If $Q(1)(x) = 1$ for any $x \in E$, then $Q$ is referred to as a Markov kernel from $E$ to $E'$. In this case, let $\Gamma_Q$ be the following change of probability measure:

$$\Gamma_Q : \mathcal{P}(E) \to \mathcal{P}(E') \quad \mu \mapsto \Gamma_Q(\mu),$$

(4)

where, for any $x' \in E'$, $\Gamma_Q(\mu)(dx') = \mu(Q(\cdot,dx'))$.

For example, if $Q$ denotes the Markov transition kernel for the targets from $X_{t-1}$ to $X_t$, then $\Gamma_Q$ transforms the information maintained by the operator on some target’s state from time $t-1$ to $t$ (see details on the Bayesian prediction step in Section V). Note that in a linear Gaussian problem, it describes the usual single-target prediction step of the Kalman filter [15].

C. Boltzmann-Gibbs transformation

Let $G : E \to (0,\infty)$ be a bounded positive measurable function. The following change of probability measure is referred to as Boltzmann-Gibbs transformation [16]:

$$\Psi_G : \mathcal{M}(E) \to \mathcal{P}(E) \quad \eta \mapsto \Psi_G(\eta)$$

(5)

where, for any $x \in E$ and assuming $\eta(G(x)) > 0$,

$$\Psi_G(\eta)(dx) = \frac{1}{\eta(G(x))} G(x)\eta(dx).$$

(6)

For example, if $G$ denotes the likelihood function for a particular observation $z \in \mathcal{Z}_t$ produced by the sensor, then $\Psi_G$ transforms the information maintained by the operator on some target’s state at time $t$ once it is associated to $z$ (see details on the Bayesian update step in Section VI). Note that in a linear Gaussian problem, it describes the usual single-target/single-observation update step of the Kalman filter.

IV. DISP filter: Target representation

In this section, $t \geq 0$ represents an arbitrary time step relevant to the scenario. Following the general principles of the surveillance activity established in Section II, the operator aims to estimate the population of the individuals that have already entered the surveillance at time $t$, i.e. $X_t^\circ$, through the following sources: a) the sequences of observation sets collected so far, i.e. $Z_0, \ldots, Z_t$, and b) any prior information on $X_t^\circ$. Fed by these sources, the DISP filter maintains and propagates a probabilistic description of the population $X_t^\circ$ through a stochastic population $\mathcal{Y}_t$. The estimation framework for stochastic populations [12] proposes two different and exclusive probabilistic representations for targets depending on their nature, i.e. whether they are indistinguishable or distinguishable. The two cases are covered in Sections IV-A and IV-B, respectively.

A. Indistinguishable targets

Given Assumption (M5) an observation produced by the sensor provides specific information on one individual of the population $X$ (unless it is a false alarm). Therefore, the (potential) individuals identified by the filter so far, and for which no specific information is yet available, must be found among:

1) those who have not yet entered the scene ($X_t^{\psi}$);
2) those who have entered in a previous time $t' < t$, and have been miss-detected ever since;
3) those who have just entered the scene, before the sensor operates at the current time step ($X_t^{\bullet}$).

Assumption (M10) states that the population described in 2 is empty. Then, given Assumption (M11), some prior information is available on the whole population $X_t^\circ$, but nothing specific is available on any individual.

The DISP filter thus maintains a representation of the appearing individuals $X_t^\circ$ as a stochastic population of indistinguishable targets $\mathcal{Y}_t$, composed of:

- a cardinality distribution $c_t^\circ$ on $\mathbb{N}$, describing the number of appearing targets;
• a single probability distribution $\tilde{p}_t^a \in \mathcal{P}(\tilde{X}_t)$ collectively describing the initial state of any appearing targets.

Note that, per definition, the appearing targets are currently in the surveillance scene; therefore, the probability distribution $\tilde{p}_t^a$ is such that

$$\tilde{p}_t^a(1X_t) = 1,$$

where $1X_t$ is the indicator function of the subset $X_t$ of $\tilde{X}_t$, that is, an appearing target is present in the scene almost surely. The scalar $\tilde{p}_t^a(1X_t)$ is called the probability of presence of the appearing targets, and the restriction of the probability distribution $\tilde{p}_t^a$ to the state space $X_t$, i.e. the probability distribution on $\mathcal{P}(X_t)$ defined by

$$\tilde{p}_t^a(dx) = \frac{1}{\tilde{p}_t^a(1X_t)} \tilde{p}_t^a(dx), \ \ x \in X_t,$$

is called the spatial distribution of the appearing targets. The same notions will apply to distinguishable targets in Section IV-A.

Note that the availability of prior information on appearing targets is inherent to the Bayesian update step detailed in Section VI, but the cardinality distribution $c_t^a$ and the probability distribution $\tilde{p}_t^a$ are not necessarily informative. If the operator is clueless about the whereabouts of the incoming individuals, for example, $\tilde{p}_t^a$ is uniformly distributed over $X_t$.

Note also that the estimation framework for stochastic populations [12] is not, in the general case, constrained by Assumptions (M10) or (M11). If Assumption (M10) was dropped, for example, the yet-to-be-detected individuals appearing at different times in the past would be represented by different stochastic populations of indistinguishable targets, to be maintained separately by the filter – i.e. the yet-to-be-detected targets that appeared at time $t = 0$, those that appeared at time $t = 1$, etc. If Assumption (M11) was dropped, several stochastic populations of indistinguishable targets could be maintained by the filter to represent distinct subpopulations of $\tilde{X}_t^*$, for individuals appearing at the same time $t \geq 0$ but with distinct behaviours – for example, if the operator possesses distinct information for individuals entering in the surveillance scene from the North than from those entering from the South. These modifications, however, would lead to more involved filtering solutions than are left out of the scope of this paper.

In the context of the DISP filter, the stochastic population of appearing targets $\tilde{Y}_t^a$ is thus characterised by the product measure

$$\tilde{P}_t^a = \sum_{n \geq 0} c_t^n (n) \tilde{p}_t^n \mathbb{1}^\otimes n,$$

where $\tilde{P}_t^n \mathbb{1} = (\tilde{p}_t) \mathbb{1}^\otimes n$.

B. Distinguishable targets

Conversely, each (potential) individual identified by the filter so far, and for which specific information is already available, is represented by a target that has been associated at least once to a past observation collected from the sensor, that is, by a previously detected target. Altogether, they are represented by a stochastic population of distinguishable targets $\tilde{Y}_t^d$.

1) Tracks: A distinguishable target which entered the scene at some birth time $0 \leq t_\bullet \leq t$ is characterised by the following 3-tuple or track

$$(t_\bullet, y, p_t^y),$$

where $0 \leq t_\bullet \leq t$ is its time of birth, $y$ its observation path, and $p_t^y \in \mathcal{P}(X_t)$ a probability distribution describing its current state. The observation path $y$ is a sequence of observations (see [17] and [13])

$$y = (\phi, z_{t_\bullet}, z_{t_\bullet+1}, \ldots, z_t),$$

such that $z_{t_\bullet} \in Z_{t_\bullet}$ and $z_{t'} \in Z_{t'}$ for $t_\bullet < t' \leq t$, that identifies the observations produced by the target up to the current time.

Note that, subsequently to its time of first detection $t_{d_\bullet}$, a target may be miss-detected any number of times – in the extreme case, $y$ may contain only empty observations save $z_{d_\bullet}$. The notions of probability of presence and spatial distribution defined for appearing targets naturally extend to a distinguishable target, through the scalar $p_t^y(1X_t)$ and the restriction of the probability distribution $p_t^y$ to the state space $X_t$, respectively.

Note also that Assumption (M6) implies that $t_{d_\bullet} \geq t_\bullet$ (i.e. a target is never detected before it enters the surveillance scene), and Assumption (M10) further implies that $t_{d_\bullet} = t_\bullet$ (i.e. a target is detected upon entering the surveillance scene). A distinguishable target is thus characterised by its observation path $y$ and its probability distribution $p_t^y$. From now on, $y$ may denote a distinguishable target, the corresponding track or the corresponding observation path depending on the context.

2) Hypotheses: The current set of all the possible observation paths is denoted by $Y_t$. Some examples of observation paths in $Y_4$ are illustrated in Figure 1.

![Fig. 1: A few observation paths at time t = 4. The target characterised by $y'_4$ entered the scene at time 2 and was detected through some observation $z'_2$, was miss-detected at time 3 and produced some observation $z'_4$ at time 4.](image)

An hypothesis $h$ is then defined as a given subset of observation paths in $Y_t$ that represents a realisation of the stochastic population $\tilde{Y}_t^d$, i.e. a possible representation of the estimated population $\tilde{X}_t^d$, described by the product measure

$$P_t^h = \bigotimes_{y \in h} p_t^y.$$

The DISP filter maintains various representations of the estimated population through a weighted set of hypotheses $H_t$;
whose weights are given by a distribution \(c_t \) on \(H_t \) satisfying
\[
\sum_{h \in H_t} c_t(h) = 1. \tag{13}
\]

For any hypothesis \( h \in H_t \) the scalar \( c_t(h) \) assesses its credibility, i.e. the likelihood that the tracks in \( h \) represent the individuals from the estimated population \( X^0_t \). It is called the \textit{probability of existence} of hypothesis \( h \). Note that the set of hypotheses \( H_t \) may – and always does, see discussion in Section VII-A – contain the special “empty” hypothesis \( \emptyset^d \), corresponding to the empty subset in \( Y_t \) representing an empty population \( X^\emptyset_p \).

From Eqs (12) and (13), the stochastic population of distinguishable targets \( \mathcal{Y}^d_t \) is described by the product measure
\[
P^d_t = \sum_{h \in H_t} c_t(h) P^d_h, \tag{14}
\]

\textbf{C. Data flow}

It holds from Sections IV-A and IV-B that, over a time step, the only stochastic population of indistinguishable targets maintained by the DISP filter represents the appearing individuals, described by the probability measure \( \hat{P}_t \) and given as a parameter model. Meanwhile, the filter maintains and propagates a stochastic population of distinguishable targets representing the previously detected individuals, described by the probability measures \( P^d_{t-1} \), \( \hat{P}^d_t \), and \( P^d_t \), before the prediction step, after the prediction step, and after the update step, respectively. The data flow of the DISP filter is depicted in Figure 2. Note that this filter does not propagate stochastic populations of indistinguishable targets, hence its name.

\textbf{V. DISP FILTER: TIME PREDICTION}

In this section we provide a detailed construction of the Bayes prediction step of the DISP filter, i.e. we describe how the probability measure propagated from the previous step \( P^d_{t-1} \) is updated to \( \hat{P}^d_t \) (see Figure 2). In this section, \( t \geq 0 \) designs an arbitrary time step relevant to the scenario.

\textbf{A. Input}

The current input is a probability measure \( P^d_{t-1} \) (14) representing the individuals in the population \( X^\emptyset_{t-1} \), i.e. those who have entered the scene no later than at time \( t - 1 \).

The case \( t = 0 \) deserves special mention. Given Assumption (M3) the population \( X^\emptyset_0 \) is empty, and the corresponding stochastic population \( \mathcal{Y}^d_0 \) is therefore almost surely empty. That is, the set of hypotheses \( H_{-1} \) is reduced to the singleton
\[
H_{-1} = \{ \emptyset^d \}, \tag{15}
\]

where the “empty” hypothesis \( \emptyset^d \) represents an empty population \( X^\emptyset_{t-1} \) (see Section IV-B2), and (13) becomes
\[
c_{-1}(\emptyset^d) = 1. \tag{16}
\]

\textbf{B. Modelling}

Given Assumption (M12) the knowledge of the operator regarding the evolution of the individuals since time \( t - 1 \) is given by a Markov kernel \( m_{t-1,t} \) from the previous target state space \( \mathbf{X}_{t-1} \) to the current one \( \mathbf{X}_t \). In particular, the function \( m_{t-1,t}(\cdot, \psi) \) describes the transition for an individual to the empty state \( \psi \); since Assumption (M2) implies that no individual may re-enter the surveillance scene, it holds that
\[
m_{t-1,t}(\psi, \psi) = 1. \tag{17}
\]

In target tracking applications it is then customary to rewrite \( m_{t-1,t} \) through a reduced Markov kernel \( \hat{m}_{t-1,t} \) from \( \mathbf{X}_{t-1} \) to \( \mathbf{X}_t \) and a probability of survival \( p_{s,t} \) on \( \mathbf{X}_{t-1} \) such that
\[
\begin{aligned}
m_{t-1,t}(x, dx') &= p_{s,t}(x)s_{t-1,t}(x, dx'), \quad x \in \mathbf{X}_{t-1}, x' \in \mathbf{X}_t, \\
m_{t-1,t}(x, \psi) &= 1 - p_{s,t}(x), \quad x \in \mathbf{X}_{t-1}, \\
m_{t-1,t}(\psi, dx') &= 0, \quad x' \in \mathbf{X}_t, \\
m_{t-1,t}(\psi, \psi) &= 1.
\end{aligned} \tag{18}
\]

The reduced Markov kernel \( \hat{m}_{t-1,t} \) describes the transition for an individual within the surveillance scene. In most practical problems it accounts for the operator’s knowledge about the motion of individuals in the physical space – presence of roads and/or obstacles, maximum speed per vehicle type, etc. – but it can cover more general evolution processes based on the nature of the target state space \( \mathbf{X}_t \). The probability of survival \( p_{s,t}(x) \) is a scalar that describes how likely is an individual, with state \( x \) at time \( t - 1 \), to be still in the scene at time \( t \).

\textbf{C. Track prediction}

The information gathered so far by the operator on any target \( y \in Y_{t-1} \), described by some probability distribution \( p^y_{t-1} \) on the former state space \( \mathbf{X}_{t-1} \), is then transferred to the current state space \( \mathbf{X}_t \) through the Markov kernel \( m_{t-1,t} \) (18). The resulting probability distribution \( \hat{p}^y_t \in \mathcal{P}(\mathbf{X}_t) \) is found to be [14]
\[
\hat{p}^y_t = \Gamma m_{t-1,t}(p^y_{t-1}), \tag{19}
\]

as detailed in Section III-B.

It is formative to write explicitly the evolution of the probability of presence of a target \( y \in Y_{t-1} \) in the scene:

\textbf{Property 1.} Let \( y \in Y_{t-1} \). The probability of presence of the predicted track is found to be
\[
\hat{p}^y_t(1_{\mathbf{X}_t}) = p^y_{t-1}(1_{\mathbf{X}_{t-1}}) - \int_{\mathbf{X}_{t-1}} [1 - p_{s,t}(x)]p^y_{t-1}(dx). \tag{20}
\]

The proof is given in Appendix A-A. Since \( a) \) no new observation is available to confirm the presence of the target in the scene, and \( b) \) given Assumption (M2) no individual can re-enter the scene, the probability of presence of a target
y ∈ Y_{t-1} is non-increasing during the prediction step, as confirmed by Eq. (20). Furthermore, it is monotonous if and only if (iff) the probability of survival is one over the support of \( p^y_{t-1} \) – that is, iff the target \( y \) stays in the scene almost surely – and it drops to zero iff the probability of survival is zero over the support of \( p^y_{t-1} \) – that is, iff the target \( y \) exits the scene almost surely.

### D. Hypothesis prediction

Since the observation set \( Z_t \) is not available yet (see Figure 2), neither the observation paths in \( Y_{t-1} \) nor the composition of the hypotheses in \( H_{t-1} \) are modified by the prediction step. The probabilities of existence \( c_{t-1}(h) \) remained unchanged as well, for there is no new information to reassess the credibility of the hypotheses \( h ∈ H_{t-1} \) currently maintained by the filter.

#### E. Output

Following Eq. (19) an hypothesis \( h ∈ H_{t-1} \) is now described by the product measure

\[
\hat{P}_t^h = \bigotimes_{y ∈ h} \hat{p}_t^y,
\]

and the predicted stochastic population \( \hat{Y}_t^{\text{pl}} \) is described by the probability measure

\[
\hat{P}_t^{\text{pl}} = \sum_{h ∈ H_{t-1}} c_{t-1}(h) \hat{P}_t^h.
\]

### VI. DISP filter: data update

In this section, we provide a detailed construction of the Bayes update step for the DISP filter, i.e. how the predicted probability measure \( \hat{P}_t^{\text{pl}} \) is updated to \( P_t^{\text{pl}} \) once the current observation set \( Z_t \) is collected by the operator (see Figure 2). In this section, \( t ≥ 0 \) designs an arbitrary time step relevant to the scenario.

#### A. Input

Some information about the appearing individuals is represented by the stochastic population \( \hat{Y}_t^{\text{pl}} \), whose probability measure given in Eq. (9) is known to the operator (see Section IV-A), while the information from the previously detected individuals propagated from the prediction step is given by Eq. (22). The product measure describing the stochastic population \( \hat{Y}_t^{\text{pl}} \), representing the previously detected individuals and the appearing ones, is thus

\[
\hat{P}_t = \hat{P}_t^{\text{pl}} ⊗ \hat{P}_t^{\text{a}}.
\]

A couple \((h ∈ H_{t-1}, n ∈ \mathbb{N})\) can be seen as a realisation of the population \( \hat{Y}_t^{\text{pl}} \), with a) \(|h|\) previously detected individuals described by the targets \( y ∈ h \), and b) \( n \) appearing individuals.

#### B. Modelling

Given Assumption (M13) the knowledge of the operator about the observation process, and in particular the imperfections of the sensor (measurement noise, false alarms, misdetections) is described by a) a non-negative real-valued function \( g_t(z, ·) ∈ B(X_t) \), interpreted as a likelihood and given for any \( z ∈ Z_t \), and b) a probability of false alarm \( p_{fa,t} \) on \( Z_t \).

In particular, the function \( g_t(·, ψ) \) describes the observation from an individual absent from the scene; since Assumption (M6) implies that no individual cannot be detected unless it is in the scene, it holds that

\[
\begin{align*}
g_t(z, ψ) &= 0, \quad z ∈ Z_t, \\
g_t(\phi, ψ) &= 1.
\end{align*}
\]

For any observation \( z ∈ Z_t \), it is then customary to write \( g_t(z, ·) \) through a restricted likelihood \( \ell_t(z, ·) ∈ B(X_t) \) and a probability of detection \( p_{d,t} \) on \( X_t \) such that

\[
\begin{align*}
g_t(z, x) &= p_{d,t}(x) \ell_t(z, x), \quad z ∈ Z_t, x ∈ X_t, \\
g_t(\phi, x) &= 1 - p_{d,t}(x), \quad x ∈ X_t, \\
g_t(z, ψ) &= 0, \quad z ∈ Z_t, \\
g_t(\phi, ψ) &= 1.
\end{align*}
\]

The restricted likelihood describes the accuracy of the sensor – \( \ell_t(z, x) \) denotes the likelihood that an observation \( z \) originates from an individual with state \( x \). The probability of detection \( p_{d,t}(x) \) is a scalar that describes how likely an individual with state \( x \) is to be detected by the sensor at the current time \( t \).

The support of the function \( p_{d,t} \), i.e. the subset of the target state space \( X_t \) where the probability of detection is non-zero, is called the (current) field of view of the sensor. Note that the field of view is included, but does not necessarily equal, the surveillance scene: if the sensor coverage is limited, there are “blind zones” in which individuals cannot be observed.

In the context of this paper, the false alarms are spurious observations stemming from the sensor system itself, independently of the population of interest \( X' \) (e.g. malfunctions from the sensor device, transmission errors, etc.). For any collected observation \( z ∈ Z_t \), the scalar \( p_{fa,t}(z) \) is the probability that \( z \) is a spurious observation; conversely, the scalar \( 1 - p_{fa,t}(z) \) is the probability that \( z \) originates from an individual of the population \( X' \).

#### C. Data association

The core of the update step consists in the data association, where potential sources of observations are matched with the new observation set \( Z_t \), and every resulting association is assessed. The potential sources of observations are:

1. the individuals in \( X'_t \) (i.e. those who entered the scene in a previous time step);
2. the individuals in \( X'^o_t \) (i.e. those who have just entered the scene);
3. the clutter, generating the false alarms.

For the purpose of data association, a clutter generator is modelled for each collected observation \( z ∈ Z_t \); as discussed in Section VI-B, it produced the observation \( z \) with probability
where \( p_{fa,t}(z) \) (in which case \( z \) is a false alarm), or it produced
no observation with probability \( 1 - p_{fa,t}(z) \) (in which case \( z \) originates from an individual of the population \( X \)).

Let us fix an observation set \( Z_t \) collected by the operator, and a realisation \((h \in H_{t-1}, n \in \mathbb{N})\) of the predicted population \( \bar{Y} \) (see Section VI-A). A valid data association \( h \) then associates the three sources of observations – the tracks \( y \in h \), the \( n \) appearing targets, and the \( |Z_t| \) clutter generators – to the collected observations \( z \in Z_t \) and the empty observation \( \phi \) (see Figure 3). Given Assumptions (M4), (M5), and (M10), a valid association \( h \) is an element of the set [14]

\[
\text{Adm}_{Z_t}(h, n) = \left\{ \left( h_d, Z_d, Z_a, \nu \right) \mid h_d \subseteq h, \ Z_d \subseteq Z_t, \ Z_a \subseteq Z_t \setminus Z_d, \ |Z_a| = n, \ \nu \in \mathcal{S}(h_d, Z_d) \right\},
\]

(26)

where \( h_d \) designs the tracks that are detected (at the current time step), \( Z_d \) the observations associated to these detected targets, \( Z_a \) the observations associated to the \( n \) appearing targets, and \( \nu \) the bijective function associating detected tracks to observations in \( Z_d \).

![Fig. 3: Data association, for a given population \((h, n)\).](image)

Each triplet \( a = (h, n, h) \), where \( h \in \text{Adm}_{Z_t}(h, n) \), then represents a specific association scheme described by the scalar

\[
P_t^a = P_{d,t}^a \times P_{md,t}^a \times P_{fa,t}^a,
\]

(27)

where

\[
P_{d,t}^a = \prod_{z \in Z_a} p_d^a(g_t(z, \cdot)) \prod_{y \in h_d} p_d^a(g_t(\nu(y), \cdot)),
\]

\[
P_{md,t}^a = \prod_{y \in h_d} p_d^a(g_t(\phi, \cdot)),
\]

\[
P_{fa,t}^a = \prod_{z \in Z_d \cup Z_a} \left( 1 - p_{fa,t}(z) \right) \prod_{z \in Z_t \setminus (Z_d \cup Z_a)} p_{fa,t}(z),
\]

(28)

where \( P_{d,t}^a, P_{md,t}^a \) and \( P_{fa,t}^a \) denote the joint probability of the data association of the detected targets, the miss-detected targets, and the clutter generators, respectively.

D. Hypothesis update

A given realisation \((h, n)\) of the predicted population, under some valid association \( h \in \text{Adm}_{Z_t}(h, n) \), leads to the construction of a unique updated hypothesis \( \hat{h} \), composed of the updated observation paths of the form

\[
\hat{h} = \bigcup_{y \in h} \{ y : \nu(y) \} \cup \bigcup_{y \in h \setminus h_d} \{ y : \phi \} \cup \bigcup_{z \in Z_a} \{ \phi_{t-1} : z \},
\]

(29)

where \( \phi_{t-1} \) denotes the "empty" observation path, i.e. the sequence of \( t-1 \) empty observations\(^2\). Using Bayes’ rule its probability of existence \( c_t(h) \) is found to be

\[
c_t(h) = \frac{c_t^a(n) c_{t-1}(h)}{E_{\phi, z]} \sum_{h \in \text{Adm}_{Z_t}(h, n)} P_t^a},
\]

(30)

where the expectation in the denominator is the sum

\[
\sum_{h \in H_{t-1}} c_t^a(n) c_{t-1}(h) \sum_{h \in \text{Adm}_{Z_t}(h, n)} P_t^a.
\]

(31)

As shown in Eq. (31), the set of updated hypotheses \( H_t \) is constructed by considering all the possible admissible association schemes \( a = (h, n, h) \), where \( h \in H_{t-1}, n \in \mathbb{N} \), and \( h \in \text{Adm}_{Z_t}(h, n) \). Following Eq. (29), the set of updated observation paths produced by some data association \( h = (h_d, Z_d, Z_a, \nu) \) can be written as the union of \( S_{md}(h_d, \nu) \) and \( S_a(Z_a) \) with

\[
S_{md}(h_d, \nu) = \{ y : \nu(y) \mid y \in h_d \}
\]

\[
S_a(Z_a) = \{ \phi_{t-1} : z \mid z \in Z_a \}
\]

so that the set of updated hypotheses \( H_t \) is found to be

\[
H_t = \left\{ S_{md}(h_d, \nu) \cup S_{md}(h \setminus h_d) \cup S_a(Z_a) \mid (h_d, Z_d, Z_a, \nu) \in \text{Adm}_{Z_t}(h, n), \ h \in H_{t-1}, n \in \mathbb{N} \right\}.
\]

(32)

Note that the number of updated hypotheses \( |H_t| \) if finite, for the number of predicted hypotheses \( |H_{t-1}| \) is finite, the number of observations \( |Z_t| \) is finite given Assumption (M9), and the set of admissible associations \( \text{Adm}_{Z_t}(h, n) \) in Eq. (26) is empty if \( n > |Z_t| \).

E. Track update

An updated observation path of the form \( y : z \) in Eq. (29), where \( y \in Y_{t-1} \) and \( z \in Z_t \), corresponds to a predicted track \( y \) which was \( a) \) detected and produced observation \( z \), if \( z \in Z_t \), or \( b) \) miss-detected, if \( z = \phi \). In both cases, the state of the updated track \( y : z \) is described by the probability distribution \( p_{t}^{y:z} \in \mathcal{P}(X_t) \) defined through the Boltzmann-Gibbs transformation [16]

\[
p_{t}^{y:z} = \Psi_{g_t(z, \cdot)}(p_{t}^{y}),
\]

(33)

as detailed in Section III-C.

It is formative to write explicitly the evolution of the probability of presence from the predicted track \( y \) to the updated track \( y : z \):

\[\text{Property 2. Let } y \in Y_{t-1} \text{ and } z \in Z_t. \text{ The probability of presence of the updated track } y : z \text{ in } Y_t \text{ is found to be}\]

\[
p_{t}^{y:z}(1X_z) = \begin{cases} 1, & z \in Z_t, \\ \frac{p_{t}^{y}(g_t(\phi, \cdot)1X_z)}{1 - p_{t}^{y}(1X_z) + p_{t}^{y}(g_t(\phi, \cdot)1X_z)}, & z = \phi. \end{cases}
\]

\[\text{as } \cdot \text{ is the concatenation operator, i.e. } (e_1, \ldots, e_n) \cdot e = (e_1, \ldots, e_n, e).\]
The proof is given in Appendix A-B. From Eq. (34) we see that the probability of presence of some target \( y \) “bursts” to 1 upon detection: given Assumption (M6) only individuals in the scene can indeed be detected, and therefore a detected target lies in the scene almost surely. On the other hand, if \( y \) is assumed miss-detected, its updated probability of presence is non-increasing: indeed, no fresh evidence on the presence of the corresponding individual in the scene is available.

F. Track creation

An updated observation path of the form \( \phi_{t-1}:z \) in Eq. (29), where \( z \in Z_t \), corresponds to an appearing target, which has become distinguishable with the detection that produced observation \( z \). The state of the newborn track is described by the probability distribution \( p^{\phi_{t-1}:z} \in \mathcal{P}(\hat{X}_t) \) defined through the Boltzmann-Gibbs transformation [16]

\[
p^{\phi_{t-1}:z}_t = \Psi_{g_t(z \cdot)}(p^0_t),
\]

as detailed in Section III-C.

It is formative to write explicitly the value of the probability of presence from a newborn track \( \phi_{t-1}:z \):

**Property 3.** Let \( z \in Z_t \). The probability of presence of the newborn track \( \phi_{t-1}:z \) is found to be

\[
p^{\phi_{t-1}:z}_t(1|X_t) = 1.
\]

The proof is given in Appendix A-C. Per definition, appearing individuals are currently in the scene and the corresponding newborn targets are thus present in the scene almost surely.

G. Output

Following Eqs (33) and (35), the updated hypothesis \( \hat{h} \in H_t \) characterised by Eq. (29) is produced by the product measure

\[
P^h_t = \bigotimes_{z \in Z_t} p^{\phi_{t-1}:z}_t \bigotimes_{y \in h_d} p_y^y(u(y)),
\]

and the updated stochastic population \( \mathcal{Y}_t^d \) is described by the probability measure

\[
P^d_t = \sum_{h \in H_t} c_t(h)P^h_t,
\]

where the updated probabilities of existence \( c_t(h) \) are given by Eq. (30).

VII. DISCUSSION

This section discusses the nature of the information propagated by the DISP filter at each time step, namely the probability measure \( P^d_t \) describing the stochastic population \( \mathcal{Y}_t^d \) (see Figure 2). It then provides leads on its exploitation by the operator to produce meaningful statistics for detection tracking purposes; most notably, it shows how to extract a Maximum a Posteriori (MAP) estimate of the estimated population \( X_t^a \).

A. Is the stochastic population \( \mathcal{Y}_t^d \) a “consistent” representation of the estimated population \( X_t^a \)?

Per construction, the set \( Y_t \) contains all the possible observation paths, constructed from the sequence of past observation sets \( Z_0, \ldots, Z_t \) collected from the sensor system and the prior information available to the operator (see Section VI). The possible representations of the estimated population \( X_t^a \) are thus to be found among the subsets of \( Y_t \).

However, not every subset \( Y \subseteq Y_t \) constitute a valid representation. Indeed, Assumption (M5) implies that two tracks \( y, y' \in Y_t \) sharing a common observation in their observation paths cannot characterise two distinct individuals of the population \( X_t^a \); in other words, at most one of these two targets represent an individual of the population, and \( y, y' \) are incompatible. More formally, we define the symmetric binary relation \( \sim \) (“compatibility”) on the set of observation paths \( Y_t \) as follows:

\[
\forall y, y' \in Y_t, y = (z_0, \ldots, z_t), y' = (z_0', \ldots, z_t'),
\]

\[
y \sim y' \iff \left[ [z_1 = z_1', 0 \leq t \leq t] \Rightarrow z_t = \phi \right].
\]

That is, two distinct observation paths are compatible (between each other) iff they share no observations save empty ones. We define the set of consistent subsets of observation paths as

\[
\text{Const}(Y_t) = \left\{ Y \subseteq Y_t \mid \forall y, y' \in Y, y \neq y' \Rightarrow y \sim y' \right\}.
\]

That is, a subset of observation paths is consistent iff any two distinct observation paths it contains are compatible. In the situation illustrated in Figure 1, for example, \( \{ y, y' \} \) is consistent, but \( \{ y, y'' \} \) is not since \( y \) and \( y'' \) share the observation \( z_1 \in Z_1 \) and are therefore incompatible. Because the consistent subsets of tracks \( Y \in \text{Const}(Y_t) \) are those which do not violate the modelling assumptions, they constitute the valid representations of the estimated population \( X_t^a \).

We must now establish how good is the match between the set of hypotheses \( H_t \) produced by the DISP filter and the consistent subsets \( \text{Const}(Y_t) \). One can show that

**Theorem 1.** If the initial set of hypotheses \( H_{-1} \) is reduced to the “empty” hypothesis \( \emptyset^d \), then for any time \( t \geq -1 \) relevant to the scenario it holds that

\[
H_t = \text{Const}(Y_t),
\]

that is, the consistent subsets of tracks in \( Y_t \) are the hypotheses in \( H_t \).

The proof is given in Appendix A-D. In addition, the hypothesis weights \( c_t(\cdot) \) are updated by exploiting all the prior information available to the operator and the sequences of observations collected so far. These remarks suggest that the description of the stochastic population \( \mathcal{Y}_t^d \) maintained by the DISP filter, under the proviso that the modelling assumptions in Section II are valid and that the initial set of hypotheses \( H_{-1} \) is initialised to the “empty” hypothesis \( \emptyset^d \), can be seen as the optimal representation of the estimated population \( X_t^a \), within the Bayesian paradigm, available to the operator.
B. Tracks: probability of presence and probability of existence

The probability presence and probability of existence are two well-defined scalars that can be produced from any track \( y \in Y_t \). While very different in nature, both can be exploited by the operator depending on their objectives – most notably, to assess the multi-target configuration in the surveillance scene, or to identify candidates for pruning (this last point will be discussed later in Section VIII).

1) Probability of presence: As explained in Section IV, the probability of presence \( p^y_t(1|x_y) \) of some target \( y \in Y_t \) assesses the likelihood that the (potential) individual represented by \( y \) has not left the surveillance scene yet. Figure 4a depicts a typical example of a probability distribution \( p^y_t \) that suggests a low probability of presence, because the potential individual has probably reached the edge of the surveillance scene and is already gone.

2) Probability of existence: It is important to note that the probability of presence \( p^y_t(1|x_y) \) does not assess the credibility of target \( y \), i.e. how likely is \( y \) to represent an individual of the estimated population \( X^y_t \). Recall from Section IV-B2 that the credibility of the hypotheses is described by the their probability of existence \( (13) \), from which we shall define the probability of existence \( \alpha^y_t \) of any track \( y \in Y_t \) as

\[
\alpha^y_t = \sum_{h \in H_t, h \ni y} c_t(h). 
\]  

As for the hypotheses, the probability of existence \( \alpha^y_t \) is a scalar between 0 and 1 that assesses the credibility of target \( y \). Figure 4b depicts an observation path, stretched over the whole state space, that would typically describe a target with low probability of existence since a single individual is unlikely to have produce such a sequence of observations across time.

C. How to extract a MAP estimate?

The operator may wish to produce a MAP estimate of the population \( X^y_t \) in order to display the most probable multi-target configuration from the output of the DISP filter. Since the hypotheses \( H_t \) are the valid representations of the estimated population \( X^y_t \) (see Section VII-A), the MAP estimate of the population can be produced through the following procedure:

1) select the hypothesis \( h^*_t \in H_t \) with the highest probability of existence \( c_t(h) \);

2) extract the MAP estimate of the states of each target \( y \in h^*_t \).

The first step above is straightforward and independent of the design choices regarding the implementation of single-object probability distributions, but the second is not and shall be discussed later in Section VIII.

Depending on the operator’s wishes, not all the targets contained in the most credible hypothesis \( h^*_t \) are worth extracting to produce the MAP estimate of the population. For example, a “cautious” operator may wish to focus on “certain” targets that are reliable with a high degree of confidence – say, 98%. In this case, only the targets \( y \in h^*_t \) with a probability of existence \( \alpha^y_t \) exceeding 0.98 will be transmitted from the filter’s output for displaying purposes.

Recall that the DISP filter represents the population \( X^y_t \) of the individuals that have entered the surveillance scene since the beginning of the scenario but may have already left it (see Section II-A). In many practical surveillance activities, the operator may wish to ignore the targets that are gone from the scene with a high degree of confidence – say, 98%. In this case, only the targets \( y \in h^*_t \) with a probability of presence \( p^y_t(1|x_y) \) exceeding 0.02 will be displayed.

More advanced criteria for the display of targets are illustrated in Section VIII.

VIII. IMPLEMENTATION

This section focuses on the practical implementation of the DISP filter. Section VIII-A provides the outlines of a generic implementation of the “exact” filter, in the sense that it transcribes the prediction (see Section V) and update (see Section VI) steps without loss of information. For a practical exploitation of the DISP filter, however, it is necessary to consider filtering approximations in order to curtail the memory and/or computational power requirements at the expense of a reasonable information loss. We propose a list, which is not meant to be exhaustive, of such approximations that the algorithm designer may wish to consider or not, on an individual basis, depending on the available resources and the specific requirements of the surveillance activity that must be addressed. The proposed approximations are categorized into three classes:

- pruning techniques discard some information maintained by the filter (see Section VIII-B);
- modelling restrictions curtail the growth of the amount of information stored by the filter (see Section VIII-C);
- merging techniques replace a piece of information maintained by the filter by a smaller one (see Section VIII-D).
A. Exact DISP filter: generic implementation

The implementation of the exact DISP filter requires few design choices, save the selection of an implementation approach for single-object probability distributions in order to proceed with the following operations:

- compute the predicted probability distributions, i.e. resolve Eq. (19);
- compute the updated probability distributions, i.e. resolve Eqs (33) and (35);
- compute the joint probability of data associations, i.e. resolve Eq. (28).

Note that all these operations involve the handling of single-object probability distributions and can be addressed through either one of two widely popular and well-established approaches in single-target estimation problems for which the relevant literature is abundant: a) Gaussian Mixture (GM) techniques (see for example [18]), or b) Sequential Monte Carlo (SMC) techniques (see for example [19]). Algorithm 1 provides a pseudo-code for a generic implementation of the DISP filter which can be adapted to either techniques.

Algorithm 2 proposes a solution to extract a MAP estimate from the DISP filter’s output. Candidate targets in the most credible hypothesis \( h_t^* \in H_t \) are not considered in the global MAP estimate if their probability of existence (see discussion in Section VII-C) fails to reach a certain level. In order to prevent flickering in the display of targets, each candidate target \( y \in h_t^* \) is tested through the following procedure

1) if its parent track was not displayed, do not display \( y \) unless its probability of existence \( \alpha_t^y \) exceeds some confirmation threshold \( t_c \);
2) if its parent track \( y \) was displayed, do not display \( y \) unless its probability of existence \( \alpha_t^y \) exceeds some lower de-confirmation threshold \( t_d < t_c \).

Note that Algorithm 1 must be slightly amended so that newborn tracks have a display status initialised at false, and the updated tracks inherit their parent track’s display status.

B. Approximation: pruning techniques

In this context, pruning techniques aim at discarding a number of tracks and/or hypotheses maintained by the DISP filter. For the sake of simplicity, these solutions are applied here to the filter’s output at some time \( t \geq 0 \), i.e. following an update step, but they can be applied following a prediction step as well (see Figure 2).

Recall from Section II that the estimated population \( \lambda_t \) may contain individuals that have already left the scene and for which the operator may have little interest. In addition to excluding candidate tracks with a low probability of presence from display as discussed in Section VII-C, one may choose to discard them completely from the propagated information, i.e. remove them from the set of tracks \( Y_t \):

Filtering approximations. Target presence

(A1) tracks \( y \in Y_t \) with probability of presence \( p_t^y(1|X_t) \) below some threshold are discarded.

On another plan, some hypotheses in \( H_t \) and/or some targets in \( Y_t \) may fit poorly with the collected observations and the prior information collected by the operator, and one may choose to discard hypotheses and/or targets with low credibility, i.e.:

Filtering approximations. Target and hypothesis existence

(A2) tracks \( y \in Y_t \) with probability of existence \( \alpha_t^y \) below some threshold are discarded.

(A3) hypotheses \( h \in H_t \) with probability of existence \( c_t(h) \) below some threshold are discarded.

Note that if Approximations (A1) and/or (A2) are implemented, the hypotheses in \( H_t \) must be marginalized over the discarded targets. Indeed, if two hypotheses \( h, h' \in H_t \) differ only through some track \( y \in Y_t \) – say \( h' = h \cup \{y\} – \) that is bound to be discarded, then a single hypothesis \( h = h' \setminus \{y\} \) with probability of existence \( c_t(h) = c_t(h) + c_t(h') \) must be substituted to the pair \( h, h' \setminus \{y\} \) in \( H_t \), for \( H_t \) is defined as a set and may not contain identical elements \( h \) and \( h' \setminus \{y\} \).

If Approximation (A3) is implemented, some weight is lost in the distribution of the hypotheses and Eq. (13) does not hold following the discard. An additional normalization can be spared, however, for the hypothesis probabilities of existence will naturally sum to one following the next update step (30). Note also that some targets in \( Y_t \) may not belong to any hypothesis following the discard of some hypotheses in \( H_t \) with Approximation (A3). These targets will not appear in the construction of the propagated information (38) any more; thus, they can be discarded from the set of targets as well without additional loss of information.

Alternatively (or in complement) to Approximations (A2) and/or (A3), if the available memory is limited, one may consider discarding a number targets and/or hypotheses when their number exceeds some threshold:

Filtering approximations. Target and/or hypothesis likelihood (cont.)

(A4) if the number of targets \( |Y_t| \) exceeds some threshold, those with the lowest probability of existence are discarded.

(A5) if the number of hypotheses \( |H_t| \) exceeds some threshold, those with the lowest probability of existence are discarded.

If computational resources allow, Approximations (A2) and/or (A3) should be favoured over Approximations (A4) and/or (A5), for the incurred information loss is easier to quantify and to describe in the general context of a surveillance activity – e.g. “the targets with a credibility of less than 5 percent are systematically discarded”.

C. Approximation: modelling restrictions

In this context, modelling restrictions aim at avoiding the creation of a number of well-defined tracks and/or hypotheses during the data update step (see Section VI).
Widely used in data association problems, measurement gating consists in discarding unlikely associations between targets and observations that would produce updated tracks with negligible probability of existence.

Filtering approximations. Measurement gating (time $t \geq 0$)

(A6) if some distance between $p_i^t$ and the probability distribution on $\hat{X}_i$ induced by $g_t(z, \cdot)$ exceeds some threshold, associations $h = (h_3, Z_3, Z_{\alpha}, \nu) \in \text{Adm}_Z(h, n)$ such that $z = \nu(\tilde{y})$ are discarded.

(A7) if some distance between $p_i^t$ and the probability distribution on $\hat{X}_i$ induced by $g_t(z, \cdot)$ exceeds some threshold, associations $h = (h_3, Z_3, Z_{\alpha}, \nu) \in \text{Adm}_Z(h, n)$ such that $z \in Z_\alpha$ are discarded.

Note that the implementation of Assumptions (A6) and (A7) depends whether single-object spatial distributions are implemented with GM or SMC techniques.

Even though the number of hypotheses generated by the data association is always finite, as discussed in Section VI-D, it is clear from Eqs (26) and (32) that it grows considerably with the number of observations $|Z_t|$ collected from the sensor. If the time span between two successive time steps is such that the simultaneous appearance of more than a few individuals in the scene is unlikely to occur, capping the number of appearing individuals in the data association step can reduce the computational load very significantly for a negligible degradation in filtering performances.

Filtering approximations. Appearing targets

(A8) the initial cardinality distribution $\hat{c}_0^t$ has a finite support and its upper bound is set to some threshold.

(A9) except for the initial time, cardinality distributions $\hat{c}_t^a$ have a finite support and their upper bound is set to some threshold.

If the validity of Assumption (M3) is doubtful because the sensor was “switched on” in a busy environment, it may be prudent not to cap the number of appearing individuals at the initial time, for the corresponding distribution $\hat{P}_0^a$ can be exploited to describe the individuals genuinely appearing at the initial time and those that were already in the scene when the scenario began. In this case, Approximation (A9) would be implemented, but Approximation (A8) would not.

D. Approximation: merging techniques

In this context, merging aims at replacing several close tracks $y \in Y_t$, according to a well-defined distance criteria, by a single track $\tilde{y}$ whose characteristics – probability distribution $p_t^\tilde{y}$, probability of existence $\alpha_t^\tilde{y}$ – collate the information of the discarded tracks. For the sake of simplicity, we only describe in this section the merging of two tracks, but the method easily extends to an arbitrary number of tracks.

The purpose of the merging technique is to identify targets which are representing the same individual of the estimated population $X_t^n$ through close, albeit different, probabilistic descriptions. A typical situation where merging is beneficial occurs when two tracks $y, y' \in Y_t$ have almost identical observation paths (see Figure 5). Determining which track among $y, y'$ (if any) proposes the “right” observation path is irrelevant; the merging procedure essentially substitutes a “compromise track” $\tilde{y}$ to the candidates $y, y'$ in the set of tracks $Y_t$.

Fig. 5: Two close tracks, candidates for merging.

Per construction, two distinct tracks $y, y' \in Y_t$ belonging to a common hypothesis $h \in H_t$ may not represent the same individual of the population (see Section VII-A) – even if their probability distributions $p_t^y, p_t^{y'}$ are close, in which case they are likely to represent two distinct individuals criss-crossing in the state space. A simple test to identify candidates tracks $y, y'$ for merging can thus be designed as follows:

1) their probability distributions must be close according to some appropriate distance, and
2) they cannot belong to a common hypothesis.

The first condition is of course dependent to the design choice for the implementation of the single-object probability distributions with GM or SMC techniques.

Filtering approximations. Merging (two targets, time $t \geq 0$)

(A10) any two distinct targets $y, y' \in Y_t$, provided that they do not belong to a common hypothesis in $H_t$, are replaced by a merged track $\tilde{y}$ in $Y_t$ if the distance between their probability distributions $p_t^y, p_t^{y'}$ falls below some threshold.

The merging mechanism itself, i.e. the construction of the merged track $\tilde{y}$, is also dependent on whether single-object probability distributions are implemented with GM or SMC techniques. An efficient merging procedure adapted to GM approaches can be inspired from Vo’s approach on multi-object probability distributions in [20]. Note also that, for the same reasons as exposed in Section VIII-B, hypotheses must be marginalized over the discarded targets whenever a merging occurs.

While similar in spirit, the merging of hypotheses is far more complex in nature. It indeed requires, among other challenges, the design of a distance criteria between hypotheses of (possibly) different sizes and of an appropriate policy to determine the probability of existence of the merged hypothesis. This is left out of the scope of this article.
APPENDIX A

Proofs

A. Property 1

Proof. Let \( y \in Y_{t-1} \), and let \( f \in \mathcal{B}(X_t) \) be an arbitrary test function. Substituting the expression of the transformation (4) into the definition of the predicted distribution (19) yields

\[
\hat{p}_t^y(f) = p_{t-1}^y(m_{t-1}, f(x))
\]

The probability of presence of the predicted track is then given by selecting \( f = 1_{x_t} \) in Eq. (43b), i.e.

\[
\hat{p}_t^y(1_{x_t}) = \int_{X_t} \left[ \int f(x, t) \mu_{t-1}(x, dx') \right] p_{t-1}^y(dx).
\]

Now, considering the expression of the Markov transition kernel (18), Eq. (44) simplifies to

\[
\hat{p}_t^y(1_{x_t}) = \int_{X_t} \left[ \int p_{t}(x) \mu_{t-1}(x, dx') \right] p_{t-1}^y(dx).
\]

B. Property 2

Proof. Let \( y \in Y_{t-1} \) and \( z \in \tilde{Z}_t \), and let \( f \in \mathcal{B}(X_t) \) be an arbitrary test function. The probability distribution of the predicted track \( \hat{p}_t^y \in \mathcal{P}(X_t) \) can be decomposed as follows:

\[
\hat{p}_t^y(f) = \hat{p}_t^y(1_{\psi})f + \hat{p}_t^y(1_{x_t}).
\]

Now, substituting the expression of the Boltzmann-Gibbs transformation (6) into the definition of the posterior distribution (33) yields

\[
\hat{p}_t^{y,z}(f) = \frac{p_t^{y,z}(g_t(z, \cdot))}{p_t^{y,z}(g_t(\cdot))},
\]

where we substitute Eq. (46) into Eq. (47a) to yield the result (47b). The probability of presence of the updated track \( y:z \) is then given by selecting \( f = 1_{x_t} \) in Eq. (47), i.e.

\[
\hat{p}_t^{y,z}(1_{x_t}) = \frac{p_t^{y,z}(g_t(z, \cdot))}{1 - p_t^{y,z}(1_{x_t})}.
\]

C. Property 3

Proof. Let \( z \in Z_t \). Substituting the expression of the Boltzmann-Gibbs transformation (6) into the definition of the posterior distribution (35), and following the exact same reasoning as in the proof given in Section A-B lead to a similar result than Eq. (48), i.e.

\[
\hat{p}_t^{o-1,z}(1_{x_t}) = \frac{p_t^{o-1,z}(g_t(z, \cdot))}{1 - p_t^{o-1,z}(1_{x_t})}.
\]

D. Theorem 1

Proof. The proof follows an inductive approach. By definition of the original time, no observation were produced before \( t = 0 \); therefore, the initial set of observation paths \( Y_0 \) is empty. It follows from the definition of consistent populations (40) that \( \text{Const}(Y_{t-1}) \) is reduced to the “empty” hypothesis, i.e. \( \text{Const}(Y_{t-1}) = \{0 \} \). Since, by assumption, the set of hypotheses is initialized to the “empty” hypothesis, it follows that the base case

\[
H_{-1} = \text{Const}(Y_{-1})
\]

is true. Let us now suppose that the case is true at some rank \( t - 1 \geq 1 \), i.e. \( H_{t-1} = \text{Const}(Y_{t-1}) \), and let us prove that it is true at rank \( t \).

1) \( \text{Const}(Y_t) \subseteq H_t \). Let \( Y \in \text{Const}(Y_t) \), and let \( Y_p \) be the set of parent tracks from \( Y \), i.e. the set

\[
Y_p = \{ y \in Y_{t-1} \mid y : z \in Y, z \in \tilde{Z}_t \}.
\]

Per construction \( Y_p \subseteq Y_{t-1} \), and if \( |Y_p| \leq 1 \) we have immediately \( Y_p \in \text{Const}(Y_{t-1}) \). Let us suppose that \( |Y_p| > 2 \), then let \( y, y' \in Y_p \) with \( y \neq y' \), where \( y = (z_0, \ldots, z_{t-1}) \) and \( y' = (z'_0, \ldots, z'_{t-1}) \). Since \( Y \in \text{Const}(Y_t) \), there exist \( z_t, z'_t \in \tilde{Z}_t \) such that \( y : z, y' : z' \in Y \) and \( y : z \sim y' : z' \). Then by definition of the track compatibility (39):

\[
[z_t = z'_t, 0 \leq t \leq t] \Rightarrow z_t = \phi,
\]

which implies that \( y \sim y' \). Thus, using the definition of consistent populations (40), \( Y_p \in \text{Const}(Y_{t-1}) \). Using the case at rank \( t - 1 \), we conclude that \( Y_p \in H_{t-1} \). Now, let us define:

- \( Y_2 = \{ y \in Y_p \mid y : z \in Y, z \in \tilde{Z}_t \} \);
- \( Z_d = \{ z \in \tilde{Z}_t \mid y : z \in Y, y \in Y_d \} \);
- \( \nu \in S(Y_d, Z_d) \) such that \( \forall y \in Y_d, y : \nu(y) \in Y \);
Then, since \( Y_p \in H_{t-1} \) and using the definition of the valid data associations (26):

\[
h = (Y_d, Z_d, Z_a, \nu) \in \text{Adm}_{Z_t}(Y_p, n),
\]

and the association \( a = (Y_p, n, h) \) leads to the construction of hypothesis \( Y \); in other words, \( Y \in H_t \).

2) \( H_t \subseteq \text{Const}(Y_t) \): Let \( h \in H_t \), and let \( a = (h_p, n, h) \), where \( h_p \in H_{t-1}, n \in \mathbb{N}, \) and \( h = (h_d, Z_d, Z_a, \nu) \in \text{Adm}_{Z_t}(h_p, n) \), be the association scheme that produced hypothesis \( h \) (see Eq. (29)). Per construction \( h \subseteq Y_t \), and if \( |h| \leq 1 \) we have immediately \( h \subseteq \text{Const}(Y_t) \). Let us suppose that \( |h| > 2 \), and let \( y, y' \in h \) with \( y \neq y' \).

1. If \( y = \psi_{t-1}:z \) and \( y' = \phi_{t-1}:z' \), with \( z, z' \in Z_a \), then since \( y 
eq y' \) we have \( z \neq z' \) and therefore \( y \sim y' \).

2. If \( y = \phi_{t-1}:z \) and \( y' = y_p', \) with \( z \in Z_a \) and \( y_p' \in h_p \). Then either \( z' = \phi \) or \( z \in Z_d \), in both options \( z \neq z' \) and therefore \( y \sim y' \). The same reasoning applies if \( y = y_p : z \) and \( y' = \phi_{t-1}:z' \).

3. Let us now suppose that \( y = y_p : z \) and \( y' = y_p' : z' \), with \( y_p, y_p' \in h_p \). Let us prove that \( y_p \neq y_p' \); for that, let us suppose that \( y_p = y_p' \). If \( y_p \in h_d \), then \( z = \nu(y_p) = \nu(y_p') = z' \); if \( y_p \in h_p \setminus h_d \), then \( z = \phi = z' \); in both options \( y = y' \), which contradicts the assumption \( y \neq y' \). We thus have \( y_p \neq y_p' \). Using the case at rank \( t - 1 \) we have also \( \{y_p, y_p'\} \subseteq h_p \subseteq H_{t-1} = \text{Const}(Y_{t-1}) \) and the definition of the track compatibility (39) therefore yields

\[
y_p \sim y_p'.
\]

Let us now suppose that \( z_t = z_t' \). If \( z_t \neq \phi \), then \( y_p = \nu^{-1}(z_t) = \nu^{-1}(z_t') = y_p' \), which contradicts the fact that \( y_p \neq y_p' \). Therefore

\[
[z_t = z_t'] \Rightarrow z_t = \phi,
\]

and from (57) and (58) we conclude that \( y \sim y' \).

In all three options above we have \( y \sim y' \), and therefore \( h \in \text{Const}(Y_t) \).

3) \( H_t = \text{Const}(Y_t) \): Since \( \text{Const}(Y_t) \subseteq H_t \) and \( H_t \subseteq \text{Const}(Y_t) \) it follows that \( H_t = \text{Const}(Y_t) \), which proves the case at rank \( t \).
Algorithm 2 Track extraction (time $t$)

**Input**
- Set of tracks: $Y_t$
- Set of hypotheses: $H_t$

**Parameters**
- Confirmation threshold: $t_c$
- Unconfirmation threshold: $t_u$

**Hypothesis selection**
Select best hypothesis: $h^*_t \leftarrow \arg \max_{h \in H_t} \alpha_t(h)$

**Track extraction**
Init. extracted tracks: $Y^*_t \leftarrow \emptyset$

for $y \in h^*_t$ do
  Compute $\alpha^y_t$ using (42)
  if $\alpha^y_t > t_c$ then
    Extract current track: $Y^*_t \leftarrow Y^*_t \cup \{y\}$
    Set display status: $t^y_t \leftarrow \text{true}$
  else if $\alpha^y_t > t_u$ then
    if $t^y_t = \text{true}$ then
      Extract current track: $Y^*_t \leftarrow Y^*_t \cup \{y\}$
    end if
  else
    Set display status: $t^y_t \leftarrow \text{false}$
  end if
end for

**Track status update**
for $y \in Y_t \setminus h^*_t$ do
  Set display status: $t^y_t \leftarrow \text{false}$
end for

**Output**
Set of extracted tracks: $Y^*_t$

**REFERENCES**

[1] Y. Bar-Shalom, “Tracking methods in a multitarget environment,” *Automatic Control, IEEE Transactions on*, vol. 23, no. 4, pp. 618–626, Aug. 1978.

[2] D. Reid, “An Algorithm for Tracking Multiple Targets,” *Automatic Control, IEEE Transactions on*, vol. 24, no. 6, pp. 843–854, Dec. 1979.

[3] S. S. Blackman and R. Popoli, *Design and analysis of modern tracking systems*. Artech House, 1999.

[4] W. D. Blair and Y. Bar-Shalom, Eds., *Multitarget-Multisensor Tracking: Applications and Advances (Volume III)*. Artech House, 2000.

[5] R. P. S. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Artech House, 2007.

[6] D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic geometry and its applications*, 2nd ed. Wiley, Sep. 1995.

[7] D. Vere-Jones and D. J. Daley, *An Introduction to the Theory of Point Processes*, 2nd ed., ser. Statistical Theory and Methods, D. Vere-Jones and D. J. Daley, Eds. Springer Series in Statistics, 2003, vol. 1.

[8] ——, *An Introduction to the Theory of Point Processes*, 2nd ed., ser. Statistical Theory and Methods, D. Vere-Jones and D. J. Daley, Eds. Springer Series in Statistics, 2008, vol. 2.

[9] 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, Oct. 2003.

[10] ——, “PHD Filters of Higher Order in Target Number,” *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 43, no. 4, pp. 1523–1543, Oct. 2007.

[11] B.-T. Vo and B.-N. Vo, “Labeled Random Finite Sets and Multi-Object Conjugate Priors,” *Signal Processing, IEEE Transactions on*, vol. 61, pp. 3460–3475, Jul. 2013.

[12] J. Houssineau, “Representation and estimation of stochastic populations,” Ph.D. dissertation, 2015, to appear.

[13] J. Houssineau, P. Del Moral, and D. E. Clark, “General multi-object filtering and association measure,” in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2013 IEEE International Workshop on*, 2013.

[14] J. Houssineau and D. E. Clark, “Hypothesised filter for independent stochastic populations,” 2014, arXiv:1404.7408.

[15] R. E. Kalman, “A New Approach to Linear Filtering and Prediction Problems,” *ASME–Journal of Basic Engineering*, vol. 82, no. D, pp. 32–45, 1960.

[16] P. Del Moral and A. Doucet, “Particle methods: An introduction with applications,” INRIA, Tech. Rep., 2009.

[17] M. Pace and P. Del Moral, “Mean-field PHD filters based on generalized Feynman-Kac flow,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 7, no. 3, pp. 484–495, Jun. 2013.

[18] D. L. Alspach and H. W. Sorenson, “Nonlinear Bayesian estimation using Gaussian sum approximations,” *Automatic Control, IEEE Transactions on*, vol. 17, no. 4, pp. 439–448, Aug. 1972.

[19] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*, ser. Statistics for Engineering and Information Science. Springer, 2001.

[20] B.-N. Vo and W.-K. Ma, “The Gaussian Mixture Probability Hypothesis Density Filter,” *Signal Processing, IEEE Transactions on*, vol. 54, no. 11, pp. 4091–4104, Nov. 2006.