Optimal Bernoulli point estimation with applications

Alexey Narykov, Murat Üney, Jason F. Ralph
Dept. of Electrical Engineering and Electronics
University of Liverpool, L69 3BX, Liverpool, UK
Emails: {a.narykov, m.uney, jfralph}@liverpool.ac.uk

***Abstract*—This paper develops optimal procedures for point estimation with Bernoulli filters. These filters are of interest to radar and sonar surveillance because they are designed for stochastic targets that can enter and exit the surveillance region at random instances. Because of this property they are not served by the minimum mean square estimator, which is the most widely used approach to optimal point estimation. Instead of the squared error loss, this paper proposes an application-oriented loss function that is compatible with Bernoulli filters, and it develops two significant practical estimators: the minimum probability of error estimate (which is based on the rule of ideal observer), and the minimum mean operational loss estimate (which models a simple defence scenario).**

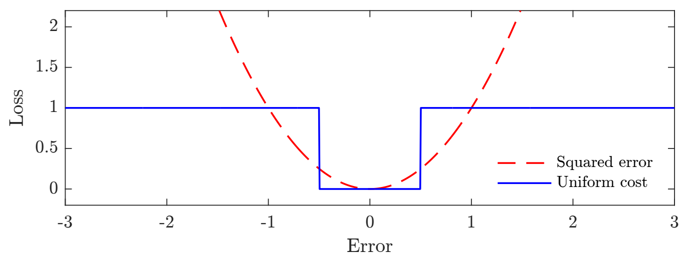


Fig. 1. The SE loss (solid) and the UC loss [13], [14] (dashed) with tolerance r_0 ($r_0 = 0.5$ shown). These losses are symmetric, i.e., they equally penalise errors of over- and underestimation. Note that only the UC loss is bounded.

I. Introduction

Radar and sonar processing chains often use a Bayesian filter that outputs a probability distribution describing the state of a time-varying stochastic world. Such a probabilistic representation is unintelligible in many practical applications and to human decision makers alike. More interpretable results are obtained by collapsing the full distribution into the best possible estimate (called an *optimal point estimate*), which is then used by the dependent application as if it were the true state of the world. The best estimate, from the perspective of Bayesian decision theory, is the one which minimises the expected amount of loss in the application. This loss emerges due to the discrepancy between the revealed true state of the world and its estimate, and is typically quantified by the squared error (SE) (loss) function (shown on Fig. 1). This function leads to the minimum mean SE (MMSE) estimate, which happens to coincide with the expected value of the random variable, and is often easy to compute.

This paper studies optimal point estimation for Bernoulli filters [1]. These filters are designed for stochastic dynamic systems that randomly switch *on* and *off* and of interest to radar and sonar surveillance [2], [3] where the target of interest may not always exist in the surveillance region. The MMSE estimator is known to be incompatible with such random finite set filters [4] since in the SE loss the underlying definition of error is based on the Euclidean distance, which does not extend to the cardinality errors, i.e., errors in the number of targets. In Bernoulli filtering, this latter errors are the equivalents of false alarm and missed detections.

Nevertheless, there have been efforts to adopt the SE loss regardless. Some authors have proposed using alternative set distance definitions (such as the optimal subpattern assignment (OSPA) distance), which combine errors in location and cardinality after redefining the SE loss [5], [6] (see also [7]). To the best of our knowledge, these approaches have not reached widespread use (see, e.g., [8], [9] for relevant developments). One difficulty for their use is that the resulting estimates are not as easy to compute as other sub-optimal ones. For example, a commonly used alternative is to test the target's probability of existence against a pre-defined threshold, and, only if it exceeds, use the SE loss to extract an estimate of the kinematic state from the localisation density [10], [11], [12]. This approach is ad hoc in its nature as there is no criterion to uniquely select the threshold, and the resulting estimate is not endowed with properties of optimality in some prespecified sense.

This paper proposes a loss function that is directly compatible with Bernoulli filters. The proposed approach can be configured to model losses in different applications. In particular, the loss function is constructed to integrate the loss resulting from the error in the target's kinematic state (quantified with the uniform cost (UC) loss shown on Fig. 1) and the loss due to the error in cardinality. The approach is validated with two examples, which are irreducible to each other, and yield practical optimal point estimates:

- the minimum probability of error (MPE) estimate, and
- the minimum mean operational loss (MMOL) estimate.

To the best of our knowledge, precursors of our approach combine other loss functions, and do that in different contexts, such as joint signal detection and estimation [15, Ch.6], [16], or joint tracking of a target and its classification [17].

The developments in this paper are for a single Bernoulli variable. Technically, the resulting expressions can be applied to individual Bernoulli components of a multi-Bernoulli process output by a multi-object filter (e.g., [12], [18], [19]). However, investigation of optimality in, and extension of, our optimal estimation approach to multi-Bernoulli systems is the subject of future work.

The paper is organised as follows. Section II introduces a Bernoulli point process, which models a stochastic target, and outlines the procedure of Bayes-optimal point estimation. Section III proposes an application-oriented loss function, and develops an optimal point estimator that thresholds the probability of existence to declare a target. Section IV develops two practical estimators and studies their thresholds.

II. Background

A. Bernoulli point process

In this article, the objects of interest, i.e., the targets, have individual states x in some d_x-dimensional state space $\mathcal{X} \subset \mathbb{R}^{d_x}$, typically consisting of position, velocity and class variables. A point process (p.p.) Φ on $\mathcal{X}$ is a random variable on the process space $\mathfrak{X} = \bigcup_{n=0}^{\infty} \mathcal{X}^n$, i.e. the space of all finite sequences of points in $\mathcal{X}$, whose number of elements *and* element states are unknown and (possibly) time-varying. A realisation of Φ is a sequence $x_{1:n} \in \mathcal{X}^n$, representing a *population* of n objects with states $x_i \in \mathcal{X}$, $1 \leq i \leq n$, where $n \in \mathbb{N}$. A more formal definition can be found in [20]. In the context of Bayesian filtering, this sequence depicts a specific multi-object configuration.

As for regular real-valued random variables, a p.p. is described by its probability distribution P_Φ on $\mathfrak{X}$; the projection measure $P_\Phi^{(n)}$ describes the realisations of Φ with n elements, $n \geq 0$. The projection measures are assumed to be symmetrical functions, so that the order of points in a realisation is irrelevant for statistical purposes and the permutations of a realization of the p.p.—such as (x_1, x_2) and (x_2, x_1)—are equally probable. In addition, a p.p. is called *simple* if the probability distribution is such that realisations are sequences of points that are pairwise distinct almost surely, i.e., a realization does not contain repetitions. For the rest of the paper, all of the point processes are assumed to be simple. The density of the projection measure $P_\Phi^{(n)}$, $n \geq 0$, is then denoted by $p_\Phi^{(n)}$.

Definition II.1 (Bernoulli point process [21], [12])**.** *A Bernoulli point process Φ on $\mathcal{X}$ with parameter $0 \leq p \leq 1$ and spatial distribution s is an i.i.d. cluster process with spatial distribution s, whose size is 1 with probability p and 0 with probability $1-p$. Its probability density is given by:*

$$p_\Phi^{(n)}(\varphi) = \begin{cases} 1-p, & \text{if } \varphi = \emptyset, \\ p \cdot s(x), & \text{if } \varphi = \{x\}, \\ 0, & \text{otherwise}. \end{cases}, \quad (1)$$

where $n = |\varphi|$ is the set cardinality, and $\emptyset$ is the empty set. Its probability generating functional (p.g.fl.) is given by[1]

$$\mathcal{G}_\Phi[h] = 1 - p + p \int h(x) s(\mathrm{d}x), \quad (2)$$

where $h : \mathcal{X} \to [0,1]$ is a test function.

In the context of target tracking, the parameter p is typically referred to as the target's *probability of existence*. Note that the definition above describes a system with at most a single target, and does not cover more complex systems.

[1]Here and in the following, notation $s(\mathrm{d}x) = s(x)\mathrm{d}x$ is used for the sake of compactness.

B. Bayes-optimal point estimation

In the Bayesian framework, the optimal solution to a point estimation problem is obtained following the minimum expected loss principle [22], [23], where a loss function

$$L : \mathfrak{X} \times \mathfrak{X} \to \mathbb{R}_0^+ \quad (3)$$

assigns a non-negative real number to every possible pair of an estimate and the true state on the state space $\mathfrak{X}$.

Proposition II.2 (Optimal Bernoulli point estimation)**.** *For a Bernoulli p.p. Φ from Definition II.1, the solution to the optimal point estimation problem is a pair $(\alpha_\Phi^*, \rho_\Phi^*)$ of, respectively, the optimal estimate and associated expected loss*

$$\alpha_\Phi^* = \underset{\alpha \in \mathfrak{X}}{\arg\min}\, \mathbb{E}\left[L(\alpha, \Phi)\right], \quad (4)$$

$$\rho_\Phi^* = \mathbb{E}\left[L(\alpha_\Phi^*, \Phi)\right], \quad (5)$$

where L is defined in (3)*, and its expected value for some $\alpha \in \mathfrak{X}$ is computed as*

$$\mathbb{E}\left[L(\alpha, \Phi)\right] = \sum_{n \geq 0} \int L(\alpha, \varphi) P_\Phi^{(n)}(\mathrm{d}x_{1:n}) \quad (6a)$$

$$= p_\Phi^{(0)}(\emptyset) L(\alpha, \emptyset) + \int L(\alpha, \{x\}) p_\Phi^{(1)}(x) \mathrm{d}x + \sum_{n \geq 2} \int L(\alpha, \{x_{1:n}\}) p_\Phi^{(n)}(x_{1:n}) \mathrm{d}x_{1:n} \quad (6b)$$

$$= (1-p) L(\alpha, \emptyset) + p \int L(\alpha, \{x\}) s(\mathrm{d}x). \quad (6c)$$

III. Application-oriented point estimation

A. Proposed loss function

We propose an estimation loss compatible with Bernoulli p.p. that, as will be shown in Section IV, can be configured to model loss in particular applications.

Definition III.1 (Application-oriented loss function)**.** *The estimation loss function is defined as*

$$L(\alpha, \varphi) := \begin{cases} c_{00}, & \text{if } \alpha = \emptyset, \varphi = \emptyset, \\ c_{01}, & \text{if } \alpha = \emptyset, \varphi = \{x\}, \\ c_{10}, & \text{if } \alpha = \{a\}, \varphi = \emptyset, \\ c_{11} + c \cdot \mathbb{1}_{B_a}(x), & \text{if } \alpha = \{a\}, \varphi = \{x\}. \end{cases}, \quad (7)$$

where $c_{00}, c_{01}, c_{10}, c_{11}, c \in \mathbb{R}^+$, and $\mathbb{1}_{B_a}$ is the indicator function on a region $B_a \subset \mathcal{X}$ such that

$$\mathbb{1}_{B_a}(x) := \begin{cases} 1, & \text{if } x \in B_a, \\ 0, & \text{if } x \notin B_a. \end{cases}, \quad (8)$$

and where B_a is the rejection *region*

$$B_a := \{x \mid d(a, x) > r_0\}, \; \forall\, x \in \mathcal{X}, \quad (9)$$

where d is the Euclidean distance.[2]

The proposed loss function is a combination of a set of coefficients $\{c_{00}, c_{01}, c_{10}, c_{11}\}$ and the loss in (8). The set encodes a cost matrix (hence the subscripts), which is essentially a loss

[2]We note that here, B_a does *not* denote the ball of radius r_0 around a, but rather it indicates the complement of the ball in $\mathcal{X}$.

TABLE I
COST ASSIGNMENT IN (7) FOR THE MPE AND MMOL ESTIMATORS.

Estimator	c_{00}	c_{01}	c_{10}	c_{11}	c
MPE	0	1	1	0	1
MMOL	0	c_A	c_M	c_M	c_A

function on the state space comprising just two points [24, Ch. 8.11]; (8) is effectively the UC loss function (Fig. 1) [13], [14]. The UC loss assigns cost 1 to every pair of a and x with distance between them higher than the tolerance parameter r_0, and 0 otherwise. In principle, the UC loss is appropriate for modelling effectors with a limited impact region, e.g., pencil-beam radars, low-power sensing nodes in a network [25], [26], [27], precise defensive countermeasures [28], [29], or rescue supplies [30, p. 40].

Remark III.2. *The loss in* (7) *and the squared OSPA error loss (discussed in the introduction) are distinct and not reducible to each other as, in the latter,* c_{00} *and* c_{11} *are set to* 0*,* c_{01} *and* c_{10} *to* c^2*, and,* $c \cdot \mathbb{1}_{B_a}(x)$ *is replaced by* $\min(c, d(a,x))^2$ *for each* $a, x \in \mathcal{X}$*, where* $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_0^+$ *is the Euclidean or other distance metric.*

B. Bayes-optimal point estimation

Theorem III.3 (Bayes-optimal Bernoulli point estimation)**.** *For a Bernoulli p.p.* Φ *with parameter* p *and spatial distribution* s*, a Bayes-optimal solution to the point estimation problem under loss* (7) *is a pair* (α^*, ρ^*)*, respectively, of the optimal point estimate and associated expected loss given by*

$$(\alpha^*, \rho^*) = \begin{cases} (\{a^*_{\mathrm{MMUC}}\}, \rho_{\{a^*_{\mathrm{MMUC}}\}}), & \text{if } p > \Gamma, \\ (\emptyset, \rho_\emptyset), & \text{if } p < \Gamma. \end{cases}, \tag{10}$$

where Γ *is the reporting threshold obtained as*

$$\Gamma = \frac{c_{00} - c_{10}}{c_{00} - c_{01} - c_{10} + c_{11} + c \int \mathbb{1}_{B_{a^*_{\mathrm{MMUC}}}}(x) s(\mathrm{d}x)}, \tag{11}$$

and a^*_{MMUC} *is the minimum mean uniform cost (MMUC) estimate*

$$a^*_{\mathrm{MMUC}} = \arg\min_{a \in \mathcal{X}} \int \mathbb{1}_{B_a}(x) s(\mathrm{d}x), \tag{12}$$

with its corresponding expected loss

$$\rho_{\{a^*_{\mathrm{MMUC}}\}} = (1-p)c_{10} + p\left[c_{11} + c \int \mathbb{1}_{B_{a^*_{\mathrm{MMUC}}}}(x) s(\mathrm{d}x)\right]. \tag{13}$$

In (10)*, the expected loss of the empty set* $\emptyset$ *is given by*

$$\rho_\emptyset = (1-p)c_{00} + p c_{01}. \tag{14}$$

The proof is given in Appendix. This estimator is a test over the Bernoulli parameter p in (1) against a threshold Γ. There are sub-optimal procedures [10], [11], [12] with a similar test structure. Our approach differs in that the threshold Γ optimally adapts to the spatial distribution s and the loss function (7) that models the application at hand (cf. the estimator in [31, Ch. 14.7.5.2]).

The Bayes-optimal estimator in Theorem III.3 requires the solution to (12) which can be computationally expensive. Nevertheless, from Sherman's theorem [32], if s is unimodal and symmetric around its mean, the optimal estimate in (12) is the mean of s, i.e., $\int x s(\mathrm{d}x)$, which is easier to compute. When s is multimodal, using the SE loss and the corresponding mean estimate is often discouraged in practice [33], [34], and an estimator based on a bounded loss function, e.g., (8), is preferable [29].

Finally, since s and p are the quantities that are common both to Bernoulli filters and the integrated probabilistic data association (IPDA) filter [35], [36], this estimator is compatible with both algorithms.

IV. APPLICATION-SPECIFIC POINT ESTIMATES

This section develops two examples of application-specific estimators that are based on the proposed loss function in (7) and use the cost relations in Table I. We study the estimators for Bernoulli-Gaussian processes, i.e., Bernoullis with $s(\cdot) = \mathcal{N}(\cdot; \mu, \sigma^2)$, where μ and σ are, respectively, the mean and standard deviation. The focus is primarily on the behaviour of Γ, which tests p to determine whether the empty set $\emptyset$ or a singleton $\{\mu\}$ should be reported.

A. Minimum probability of error estimate

Cost assignment in this estimator is inspired by the MPE decision rule in detection theory [37, p. 8], which is sometimes called *the rule of ideal observer* [15, p. 51] or *the Siegert-Kotelnikov rule* [38, p. 65]. It assigns the cost values such that correct decisions incur no penalties, and incorrect decisions are penalised equally with the unit cost. Such cost assignment appears to be compatible with the UC loss function within loss (7) when costs from Table I are used, and provided that correct detection can also be penalised if the true target kinematic state falls inside the rejection region.

Corollary IV.1 (Minimum probability of error estimation)**.** *Under the MPE cost assignment from Table I, the MPE estimator is a pair* $(\alpha^*_{\mathrm{MPE}}, \rho^*_{\mathrm{MPE}})$ *that is obtained from* $(\alpha^*_\Phi, \rho^*_\Phi)$ *in Theorem III.3 with*

$$\Gamma = \left[2 - \int \mathbb{1}_{B_{a^*_{\mathrm{MMUC}}}}(x) s(\mathrm{d}x)\right]^{-1}, \tag{15}$$

$$\rho_{\{a^*_{\mathrm{MMUC}}\}} = \mathcal{G}_\Phi\left[\mathbb{1}_{B_{a^*_{\mathrm{MMUC}}}}\right], \tag{16}$$

$$\rho_\emptyset = p, \tag{17}$$

where $\mathcal{G}_\Phi[\cdot]$ *is defined in* (2)*.*

Proof. The result is obtained by substituting the MPE costs from Table I into Theorem III.3. For (13), this leads to

$$\rho_{\{a^*_{\mathrm{MMUC}}\}} = 1 - p + p \int \mathbb{1}_{B_{a^*_{\mathrm{MMUC}}}}(x) s(\mathrm{d}x), \tag{18}$$

which is equivalent to (16) when notations (2) are used. □

The result in (16) highlights the utility of p.g.fl.s in practical applications, in addition to filter derivations, see, e.g., [21]. Another relevant example is the statistics of the stochastic adversarial risk in [39, Thm. IV.2].

Fig. 2a compares the quality of MPE and conventional estimates that are produced, respectively, using $\Gamma_1 = 0.5943$ and $\Gamma_2 = 0.5$. The MPE threshold yields estimates with lower probability of error for Bernoullis with $p \in [\Gamma_2, \Gamma_1]$. The MPE threshold is further studied on Fig. 3a: Bernoullis with higher spatial uncertainty require higher thresholds for a target to be declared. A Bernoulli with $p < 0.5$ is never declared as a target (i.e., the threshold values are bounded from below), whereas when $p > 0.5$ it may be estimated as no target in case the uncertainty is high with respect to r_0.

B. Minimum mean operational loss estimate

Cost assignment in the MMOL estimator is inspired by a textbook example of decision making under uncertainty, which is typically called the umbrella problem [40, p. 24] or cost-loss model [41]. This model is used to determine how the probability of adverse events affects the decision of whether to take a costly precautionary measure for protection against losses from that event. We consider an operational scenario with one potential target aiming to destroy the asset of cost c_A (see, e.g., [39]), when we control a countermeasure of cost c_M. What distinguishes it from the rule of ideal observer is that cardinality errors are penalised in different ways (see Table I): $c_{01} \neq c_{10}$ (losing the asset is commonly more damaging than wasting the countermeasure), and $c_{11} > 0$ (countermeasure is committed to prevent losing the asset). We also extend the cost-loss situation within loss (7) by considering that the countermeasure has limited impact centred around the point of its application with radius r_0: failure to apply it sufficiently close to the target is then modelled by the UC loss in (8), and, naturally, is penalised by c_A in addition to c_M.

Although the original model is designed for studying decisions about what course of action is to implement, it permits dual interpretation, see e.g. [42], and thus communicates a statement about the state of stochastic world. For example, if the optimal action is to preserve the countermeasure, it is equivalent to acting as if *there were no target*. And similarly, applying the countermeasure in a certain location is equivalent to acting as if *there were a target in that point*.

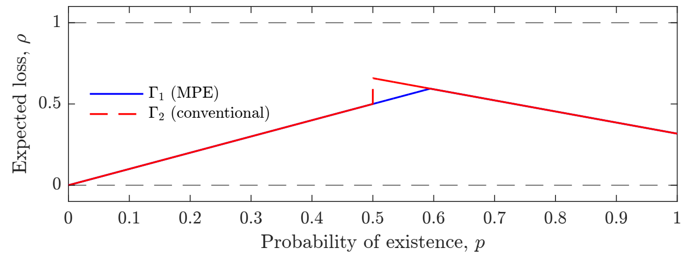
Corollary IV.2 (Minimum mean operational loss estimation)**.** *Under the MMOL cost assignment from Table I, the MMOL estimator is a pair* $(\alpha^*_{\text{MMOL}}, \rho^*_{\text{MMOL}})$ *that is obtained from* $(\alpha^*_\Phi, \rho^*_\Phi)$ *in Theorem III.3 with*

$$\Gamma = \frac{c_M}{c_A}\left[1 - \int \mathbb{1}_{B_{a^*_{\text{MMUC}}}}(x)s(\mathrm{d}x)\right]^{-1} \tag{19}$$

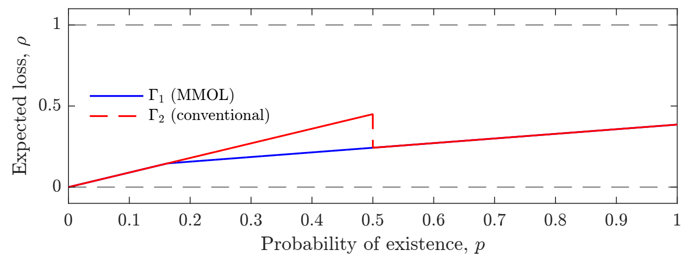
$$\rho_{\{a^*_{\text{MMUC}}\}} = c_M + p \cdot c_A \cdot \int \mathbb{1}_{B_{a^*_{\text{MMUC}}}}(x)s(\mathrm{d}x), \tag{20}$$

$$\rho_\emptyset = p \cdot c_A. \tag{21}$$

The result is obtained by substituting the MMOL costs from Table I into Theorem III.3. Fig. 2b compares the quality of the MMOL and conventional estimates, which are produced, respectively, with $\Gamma_1 = 0.1628$ and $\Gamma_2 = 0.5$. The MMOL threshold yields lower mean operational loss for Bernoullis with $p \in [\Gamma_1, \Gamma_2]$. The MMOL threshold is studied on Fig. 3b: it is bounded from below by the value coinciding with c_M/c_A, which can generally be smaller than 0.5. When the threshold is studies for various c_M/c_A (dotted), it reveals its characteristic behaviour: if $c_M = 0$, the target is always declared as there

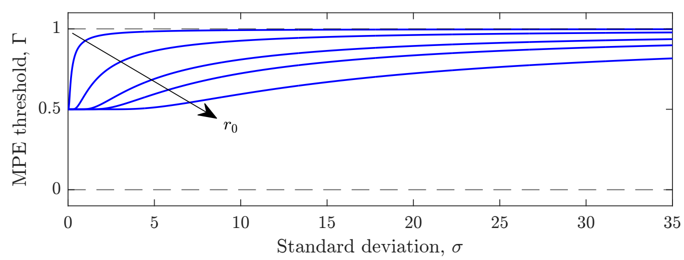


(a) Expected loss as *probability of error* ($r_0 = 1$).

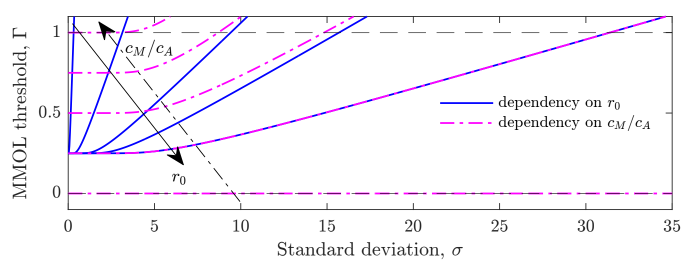


(b) Expected loss as *mean operational loss* ($r_0 = 1$, $c_A = 0.9$, $c_M = 0.1$).

Fig. 2. The quality of point estimates as perceived by the respective application, which are produced with optimal (Γ_1) and conventional ($\Gamma_2 = 0.5$) thresholds. The estimates are of Bernoulli-Gaussians with distinct p values, $0 \leq p \leq 1$, and spatial distributions with the same $\mu = 0$ and $\sigma = 1$. The quality is naturally quantified by the expected loss in the application.



(a) Dependency on the tolerance r_0 ($r_0 = 0.1$, 1, 3, 5, and 10 shown).



(b) Dependencies on r_0 for $c_M/c_A = 0.25$ ($r_0 = 0.1$, 1, 3, 5, and 10 shown), and on c_M/c_A for $r_0 = 10$ ($c_M/c_A = 0$, 0.25, 0.5, 0.75, and 1 shown).

Fig. 3. Reporting threshold Γ as a function of spatial uncertainty, which is characterized by the standard deviation σ in the Bernoulli-Gaussian case.

is no cost of committing the countermeasure; if $c_M \geq c_A$, the target is never declared since it is always better to preserve a costly countermeasure.

V. Conclusion

In this paper we have proposed an application-oriented loss function for Bernoulli filters, and developed two examples of optimal application-specific point estimators. Similar to the conventional estimators, they involve the step of thresholding of the target's probability of existence. However, this threshold is not a constant, but a function of specific parameterization in the loss function as well as certain features of the spatial probability density. A critical difference of the resulting estimators is that a Bernoulli with high probability of existence may still be declared as absent if the spatial uncertainty is high, or if committing costly measures brings unjustified expected losses.

Appendix: Proof of Theorem III.3

Proof. Let us first obtain expressions of the expected loss $\rho_\emptyset$ for the empty set and $\rho_{\{a\}}$ for a singleton containing an arbitrary kinematic state a. For $\alpha = \emptyset$, the expected loss $\rho_\emptyset = \mathbb{E}\left[L(\emptyset, \Phi)\right]$ is given by

$$\rho_\emptyset = (1-p)\cdot L(\emptyset,\emptyset) + p\int L(\emptyset,\{x\})s(\mathrm{d}x), \qquad (22a)$$

and substituting from (7) in the above equation yields (14).

For $\alpha = \{a\}$, the expected loss is

$$\begin{aligned}\rho_{\{a\}} &= \mathbb{E}\left[L(\{a\},\Phi)\right] && (23a)\\ &= (1-p)\cdot L(\{a\},\emptyset) + p\int L(\{a\},\{x\})s(\mathrm{d}x) && (23b)\\ &= (1-p)\cdot c_{01} + p\cdot\left[c_{11} + c\int \mathbb{1}_{B_a}(x)s(\mathrm{d}x)\right]. && (23c)\end{aligned}$$

The minimum of (23c) is obtained for $a = a^*_{\mathrm{MMUC}}$ given by (12). Substituting (12) into (23c) yields (13), i.e., the minimum expected loss $\rho_{\{a^*_{\mathrm{MMUC}}\}}$ for a singleton. The optimal estimate (and associated minimum expected loss) is then obtained by comparing the resulting values of expected loss as

$$\alpha^* = \underset{\alpha\in\{\emptyset,\{a^*_{\mathrm{MMUC}}\}\}}{\arg\min}\ \rho_\alpha, \qquad (24)$$

which is written as a test for p in (10), where the threshold Γ in (11) is obtained by solving $\rho_\emptyset = \rho_{\{a^*_{\mathrm{MMUC}}\}}$ w.r.t. p. □

Acknowledgements

The authors would like to thank the UK Defence Science and Technology Labs (Dstl) grant no. 1000143726 (Project BLUE) for financial support.

References

[1] B. Ristic, B.-T. Vo, B.-N. Vo, and A. Farina, "A tutorial on Bernoulli filters: theory, implementation and applications," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3406–3430, 2013.

[2] D. Cormack and D. Clark, "Tracking small UAVs using a Bernoulli filter," in *2016 SSPD*, pp. 1–5, IEEE, 2016.

[3] A. Gunes and M. B. Guldogan, "Joint underwater target detection and tracking with the Bernoulli filter using an acoustic vector sensor," *Digital Signal Processing*, vol. 48, pp. 246–258, 2016.

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

[5] M. Baum, P. Willett, and U. D. Hanebeck, "Polynomial-time algorithms for the exact MMOSPA estimate of a multi-object probability density represented by particles," *IEEE Transactions on Signal Processing*, vol. 63, no. 10, pp. 2476–2484, 2015.

[6] Á. F. García-Fernández and L. Svensson, "Spooky effect in optimal OSPA estimation and how GOSPA solves it," in *2019 22th FUSION Conference*, pp. 1–8, IEEE, 2019.

[7] T. Vu, "A complete optimal subpattern assignment (COSPA) metric," in *2020 IEEE 23rd FUSION Conference*, pp. 1–8, IEEE, 2020.

[8] M. Rezaeian and B.-N. Vo, "Error bounds for joint detection and estimation of a single object with random finite set observation," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1493–1506, 2009.

[9] Á. F. García-Fernández, M. Hernandez, and S. Maskell, "An analysis on metric-driven multi-target sensor management: GOSPA versus OSPA," in *2021 IEEE 24th FUSION Conference*, pp. 1–8, IEEE, 2021.

[10] B.-T. Vo, D. Clark, B.-N. Vo, and B. Ristic, "Bernoulli forward-backward smoothing for joint target detection and tracking," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4473–4477, 2011.

[11] B. Ristic and S. Arulampalam, "Bernoulli particle filter with observer control for bearings-only tracking in clutter," *IEEE Transactions on Aerospace and Electronic systems*, vol. 48, no. 3, pp. 2405–2415, 2012.

[12] Á. F. García-Fernández, J. L. Williams, K. Granström, and L. Svensson, "Poisson multi-Bernoulli mixture filter: direct derivation and implementation," *IEEE TAES*, vol. 54, no. 4, pp. 1883–1901, 2018.

[13] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Pt. I: Detection, Estimation, and Linear Modulation Theory*. JWS, 2004.

[14] A. H. Jazwinski, *Stochastic processes and filtering theory*. Dover, 2007.

[15] D. Middleton, *Non-Gaussian statistical communication theory*, vol. 22. JWS, 2012.

[16] G. V. Moustakides, G. H. Jajamovich, A. Tajer, and X. Wang, "Joint detection and estimation: Optimum tests and applications," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4215–4229, 2012.

[17] X. R. Li, "Optimal Bayes joint decision and estimation," in *2007 10th International Conference on Information Fusion*, pp. 1–8, IEEE, 2007.

[18] B. Vo, B. Vo, and D. Phung, "Labeled random finite sets and the Bayes multi-target tracking filter," *IEEE TSP*, vol. 62, no. 24, pp. 6554–6567, 2014.

[19] D. Musicki and R. Evans, "Joint integrated probabilistic data association: JIPDA," *IEEE TAES*, vol. 40, no. 3, pp. 1093–1099, 2004.

[20] D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic geometry and its applications*. John Wiley & Sons, 1995.

[21] I. Schlangen, E. D. Delande, J. Houssineau, and D. E. Clark, "A Second-Order PHD Filter With Mean and Variance in Target Number," *IEEE Transactions on Signal Processing*, vol. 66, pp. 48–63, Jan 2018.

[22] J. O. Berger, *Statistical decision Theory and Bayesian Analysis*. Springer Science & Business Media, 1985.

[23] H. Raiffa and R. Schlaifer, *Applied statistical decision theory*. Harvard Univ., 1961.

[24] M. H. DeGroot, *Optimal statistical decisions*, vol. 82. JWS, 2005.

[25] Y. Boers, H. Driessen, and L. Schipper, "Particle filter based sensor selection in binary sensor networks," in *2008 11th International Conference on Information Fusion*, pp. 1–7, IEEE, 2008.

[26] W. Koch, "Adaptive parameter control for phased-array tracking," in *Signal and data Processing of Small Targets 1999*, vol. 3809, pp. 444–455, International Society for Optics and Photonics, 1999.

[27] A. S. Narykov, O. A. Krasnov, and A. Yarovoy, "Effectiveness-based radar resource management for target tracking," in *2014 International Radar Conference*, pp. 1–5, IEEE, 2014.

[28] M. Guerriero, L. Svensson, D. Svensson, and P. Willett, "Shooting two birds with two bullets: How to find minimum mean OSPA estimates," in *2010 13th FUSION Conference*, pp. 1–8, IEEE, 2010.

[29] D. Salmond, N. Everett, and N. Gordon, "Target tracking and guidance using particles," in *Proceedings of the 2001 American Control Conference.*, vol. 6, pp. 4387–4392, IEEE, 2001.

[30] J. L. Williams, *Information theoretic sensor management*. PhD thesis, Massachusetts Institute of Technology, 2007.

[31] R. P. Mahler, *Statistical multisource-multitarget information fusion*. Artech House, Inc., 2007.

[32] S. Sherman, "A theorem on convex sets with applications," *The Annals of Mathematical Statistics*, pp. 763–767, 1955.

[33] D. Dionne, H. Michalska, and C. Rabbath, "Predictive guidance for pursuit-evasion engagements involving multiple decoys," *Journal of guidance, control, and dynamics*, vol. 30, no. 5, pp. 1277–1286, 2007.

[34] S. Saha, Y. Boers, H. Driessen, P. K. Mandal, and A. Bagchi, "Particle based MAP state estimation: A comparison," in *2009 12th FUSION Conference*, pp. 278–283, IEEE, 2009.

[35] D. Musicki, R. Evans, and S. Stankovic, "Integrated probabilistic data association," *IEEE TAC*, vol. 39, no. 6, pp. 1237–1241, 1994.

[36] E. Brekke, O. Hallingstad, and J. Glattetre, "The signal-to-noise ratio of human divers," in *OCEANS'10 IEEE SYDNEY*, pp. 1–10, IEEE, 2010.

[37] H. V. Poor, *An introduction to signal detection and estimation*. Springer Science & Business Media, 2013.

[38] Y. Tsypkin, *Foundations of the Theory of Learning Systems*. New York: Academic Press, 1973.

[39] A. Narykov, E. Delande, and D. E. Clark, "A formulation of the adversarial risk for multiobject filtering," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 4, pp. 2082–2092, 2021.

[40] R. L. Winkler, *An introduction to Bayesian inference and decision*. Probabilistic Publishing, 2003.

[41] A. H. Murphy, "The value of climatological, categorical and probabilistic forecasts in the cost-loss ratio situation," *Monthly Weather Review*, vol. 105, no. 7, pp. 803–816, 1977.

[42] J. W. Tukey, "Conclusions vs decisions," *Technometrics*, vol. 2, no. 4, pp. 423–433, 1960.