

# Validation Gating for Non-Linear Non-Gaussian Target Tracking

Tim Bailey, Ben Upcroft and Hugh Durrant-Whyte  
Australian Centre for Field Robotics  
University of Sydney, NSW, Australia  
tbailey@acfr.usyd.edu.au

**Abstract** - *This paper develops a general theory of validation gating for non-linear non-Gaussian models. Validation gates are used in target tracking to cull very unlikely measurement-to-track associations, before remaining association ambiguities are handled by a more comprehensive (and expensive) data association scheme. The essential property of a gate is to accept a high percentage of correct associations, thus maximising track accuracy, but provide a sufficiently tight bound to minimise the number of ambiguous associations.*

*For linear Gaussian systems, the ellipsoidal validation gate is standard, and possesses the statistical property whereby a given threshold will accept a certain percentage of true associations. This property does not hold for non-linear non-Gaussian models. As a system departs from linear-Gaussian, the ellipsoid gate tends to reject a higher than expected proportion of correct associations and permit an excess of false ones. In this paper, the concept of the ellipsoidal gate is extended to permit correct statistics for the non-linear non-Gaussian case. The new gate is demonstrated by a bearing-only tracking example.*

**Keywords:** Target tracking, data association, validation gating, Gaussian mixture models.

## 1 Introduction

Target tracking is the problem of estimating the location and velocity of one or more moving targets given a motion model and a set of sensor measurements [1]. Sensed data is imperfect and, moreover, a single measurement might detect only an aspect of a target position (e.g., range-only or bearing-only), so multiple measurements are combined using a stochastic filter. Sensor fusion involving non-distinct targets introduces the problem of data association.

A measurement might originate from one of several targets or from clutter within the environment, and robust data association must deal with this ambiguity. A number of schemes exist in the literature including joint probabilistic data association (JPDA) [3], multi-hypothesis tracking (MHT) [10], and mixture model tracking (MMT) [7]. Each of these systems can be computationally prohibitive if all possible associations are considered, particularly when accounting for non-detection of existing tracks, false alarms, track loss, and new track detection. Furthermore, probabilities

of false positives (clutter) and false negatives (non-detection) are problem specific and may be difficult to model realistically. To reduce complexity for practical implementation *validation gating* is applied [1].

A validation gate is an association threshold, whereby an observation-to-target association is considered if it satisfies a metric of “acceptability”, and is rejected otherwise. Typically this gate is based solely on distance, although other attributes (e.g., colour, radar signature) might also act as discriminants. A data association scheme, such as MHT, becomes more tractable having first culled very unlikely associations.

The standard validation gate is a Gaussian hyper-ellipsoid, and is an optimal metric for linear-Gaussian systems. It remains ubiquitous in the current target tracking literature [12, 8, 5, 11]. However, it ceases to be statistically precise for systems with non-Gaussian probability densities or non-linear observation models. As a system departs from linear-Gaussian, the ellipsoid gate tends to reject a higher than expected proportion of correct associations and permit an excess of false ones.

This paper investigates validation gating for non-linear non-Gaussian systems. It develops a general theory for validation gating as an extension to the linear concept of the ellipsoidal gate. For linear-Gaussian models, the new gate is identical to the ellipsoidal gate. For non-linear non-Gaussian models, it maintains a proper statistical “probability of acceptance” and continues to exhibit sensible gating behaviour. While this paper provides implementation of our theory, the focus is not on computational aspects but on theoretical concepts and intuitions of the key gating properties.

The paper format is as follows. Section 2 defines the properties and calculation of the standard ellipsoidal gate. Section 3 presents a generic Monte Carlo solution for computing a validation threshold. Section 4 introduces the bearing-only scenario used as an example in subsequent sections. Section 5 presents a measure of validation likelihood based on the marginal observation density  $p(\mathbf{z})$ , and examines why it is not sensible. The next section discusses why  $p(\mathbf{z})$  does, however, provide a proper measure of track likelihood. Section 7 presents a second measure of validation likelihood based on the joint target-measurement density  $p(\mathbf{x}, \mathbf{z})$ , and proposes that this is reasonable and general. The final section provides concluding remarks and suggestions for future research.

## 2 Ellipsoidal Validation Gate

The ellipsoidal validation gate is optimal for a linear observation model<sup>1</sup> with additive noise

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{r} \quad (1)$$

where the state *probability density function* (PDF) is Gaussian  $p(\mathbf{x}) = \mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}, \mathbf{P})$ , and the noise is zero-mean Gaussian  $p(\mathbf{r}) = \mathcal{N}(\mathbf{r}; \mathbf{0}, \mathbf{R})$  and independent of  $\mathbf{x}$ . The validity of a measurement  $\mathbf{z}_i$  is determined from its residual (or innovation) with the predicted observation

$$\boldsymbol{\nu} = \mathbf{z}_i - \mathbf{H}\hat{\mathbf{x}} \quad (2)$$

which has covariance  $\mathbf{S} = \mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R}$ . Validation is computed by gating the *normalised innovation squared* (NIS).<sup>2</sup>

$$\boldsymbol{\nu}^T \mathbf{S}^{-1} \boldsymbol{\nu} < M_d \quad (3)$$

The threshold  $M_d$ , for an innovation dimension  $d$ , can be computed efficiently since the NIS follows a chi-square PDF. Thus, for a “probability that  $n\%$  of *true* associations are accepted”,<sup>3</sup>  $M_d$  is obtained from

$$\frac{n}{100} = P\left(\frac{d}{2}, \frac{M_d}{2}\right) \quad (4)$$

where

$$P(a, b) = \frac{1}{\Gamma(a)} \int_0^b e^{-t} t^{a-1} dt \quad (5)$$

is the incomplete gamma function [9].

It is important to realise that a validation gate provides no statistical measure for the rejection of *false* associations. Rather, it defines a region of acceptance such that  $(100 - n)\%$  of *true* measurements are rejected given that each  $\mathbf{z}_i$  is distributed according to

$$p(\mathbf{z}) = \mathcal{N}(\mathbf{z}; \mathbf{H}\hat{\mathbf{x}}, \mathbf{S}) \quad (6)$$

This allows the gate to operate without recourse to models of clutter etc, which may be difficult to define realistically. Effectively, it eliminates very unlikely associations, the bulk of which are presumably other targets and clutter, and the remaining measurements, whether true or false, are plausible candidates.

<sup>1</sup>It is typical for state-space variables to have time-subscripts, such as  $\mathbf{x}_k$ , but, as validation gating is not time dependent, they are omitted here.

<sup>2</sup>The NIS is also commonly known as the Mahalanobis distance.

<sup>3</sup>This statement is a little misleading; the ellipsoidal gate as calculated from (4) does not accept a specified  $n\%$  of true measurements in a real (physical) case. If the true target location  $\mathbf{x}_t$  is somewhere towards the centre of  $p(\mathbf{x})$ , then more than  $n\%$  of measurements, which are distributed according to  $p(\mathbf{z}|\mathbf{x} = \mathbf{x}_t)$ , will be accepted. If, on the other hand,  $\mathbf{x}_t$  is far from the mean of  $p(\mathbf{x})$ , less than  $n\%$  will be accepted. What the ellipsoidal gate *does* achieve is the acceptance of  $n\%$  of true measurements over a series of Monte Carlo simulations, where  $\mathbf{x}_t$  is sampled from  $p(\mathbf{x})$  for each run. This statistical quality is the same for the non-linear validation gates developed in this paper.

## 3 Computing a Validation Gate for Non-Gaussian PDFs

A closed-form validation threshold is not, in general, available for non-Gaussian PDFs. However, an approximation is possible via Monte Carlo sampling.

1. Draw  $N$  samples  $\mathbf{z}_i \sim p(\mathbf{z})$ .
2. Weight each  $\mathbf{z}_i$  according to its likelihood of “association validity”.
3. Select the  $m$ -th largest weight, where  $m = Nn/100$ , to obtain an  $n\%$  acceptance threshold,  $w_n$ . Thus, if  $n$  is large (say 95%), then  $w_n$  is small, (the bottom 5%).

If the validation likelihood of a particular measurement is greater than  $w_n$ , then the association is accepted. For  $p(\mathbf{z})$  Gaussian, this approach becomes equivalent to the ellipsoidal gate as  $N \rightarrow \infty$ .

From the above algorithm, steps 1 and 3 are straightforward. Samples drawn from  $p(\mathbf{z})$  provide a statistic of how many measurements will fall within a certain region of observation-space, and selecting an ordered weight marks a bound for inside or outside that region. Step 2, on the other hand is less intuitive, and the remainder of this paper focuses on defining an appropriate measure for validation likelihood.

Two measures are investigated. The first is the observation PDF  $p(\mathbf{z})$ , such that  $w_i = p(\mathbf{z} = \mathbf{z}_i)$ . This measure, while seemingly the obvious solution, is shown to be an unreasonable validation likelihood. The second is based on the joint PDF of the state and observation  $p(\mathbf{x}, \mathbf{z})$ . Given an observation  $\mathbf{z}_i$ , the validation likelihood is  $w_i = \max p(\mathbf{x}, \mathbf{z} = \mathbf{z}_i)$ ; the mode of a slice along  $\mathbf{z}_i$ . This measure gives sensible gating results. Both measures are equivalent to the standard ellipsoidal gate for linear-Gaussian systems.

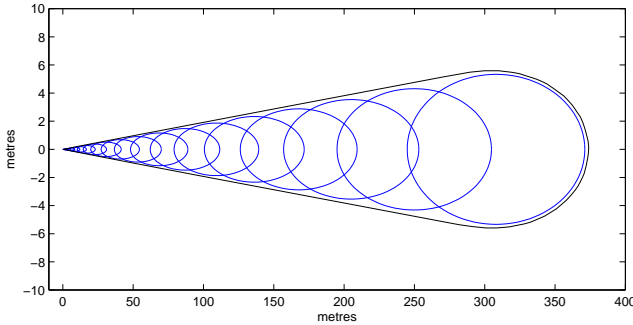
## 4 Bearing-Only Tracking

As an example of a non-linear non-Gaussian problem, we consider a bearing-only target tracking scenario. The bearing observation model is given by

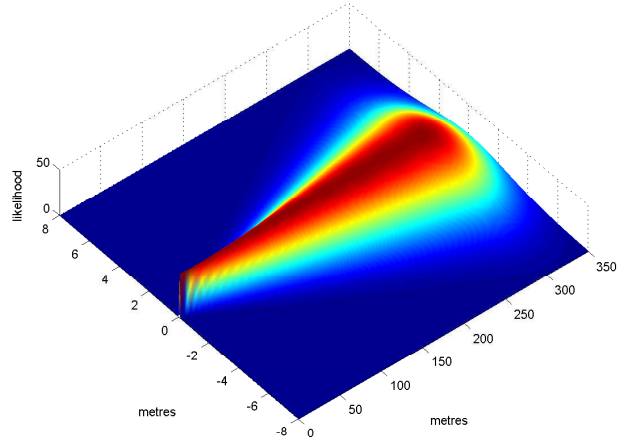
$$z = h(\mathbf{x}, \mathbf{x}_s) + r = \arctan\left(\frac{y - y_s}{x - x_s}\right) + r \quad (7)$$

where  $r$  is zero-mean Gaussian with variance  $R = (\pi/360)^2$ , and the sensor location  $\mathbf{x}_s$  is assumed perfectly known. For a given target location  $\mathbf{x}_i$ , this model defines the conditional PDF  $p(z|\mathbf{x} = \mathbf{x}_i) = \mathcal{N}(z; h(\mathbf{x}_i, \mathbf{x}_s), R)$  and, for a given measurement  $z_i$ , it defines the *likelihood function*  $p(z = z_i|\mathbf{x})$  over the domain of  $\mathbf{x}$ .

A bearing-only likelihood function  $p(z = 0^\circ|\mathbf{x})$ , in the Cartesian plane, is shown in Figure 1. Note, this function differs from the model in (7) in two respects. First, it is a Gaussian mixture model (GMM) approximation, which permits relatively efficient computation. And second, due to the limited number of mixture components, it exhibits a roll-off in likelihood after about



(a) Gaussian components ( $2\sigma$  ellipses) and 95% volume bound



(b) GMM likelihood function

Figure 1: A 20-component Gaussian mixture model of a bearing-only likelihood function in the coordinate frame of the target  $\mathbf{x}$ , where  $\mathbf{x}_s = [0, 0]^T$  and  $z = 0^\circ$ . The individual components are shown in (a), along with an equi-likelihood contour encompassing 95% of the GMM probability mass.

300 metres, whereas (7) defines a constant likelihood along the x-axis from the origin to infinity. Likelihood functions for other bearing angles, taken from different sensor locations, have the same essential shape as in Figure 1, just shifted and rotated to the appropriate position and angle.

Given an uninformative prior, say  $p(\mathbf{x}) = \mathcal{N}(\mathbf{x}; \mathbf{0}, \infty)$ , Bayesian fusion with  $p(z = z_i | \mathbf{x})$  produces a posterior  $p(\mathbf{x} | z_i)$  with the same shape as the likelihood function, but normalised to unit volume. In the remainder of this paper, we use  $p(\mathbf{x} | z = 0^\circ)$  as the prior PDF in our examples; that is, the fusion of the likelihood function in Figure 1 with an uninformative prior over the domain of  $\mathbf{x}$ . For notational convenience, we denote this (informative) prior as simply  $p(\mathbf{x})$ . It is the PDF of the target location based on a single bearing-only measurement.

The bearing-only scenario is interesting because the observation model is very non-linear. In particular,  $p(\mathbf{x})$  has a region of high probability density but little local probability mass near the origin; the bulk of the mass is distant from the origin. This is different to a Gaussian PDF, where the region of most mass is also the region of high likelihood. The existence of high likelihood away from the concentration of mass is fundamental to the derivation of a suitable gating measure.

## 5 $p(\mathbf{z})$ as a Validation Likelihood

The generic gating algorithm in Section 3 draws samples from  $p(\mathbf{z})$ , and so the most intuitive measure of validation likelihood is  $p(\mathbf{z})$  itself. Certainly for the Gaussian case in (6), an  $n\%$  acceptance threshold based on  $p(\mathbf{z})$  is equivalent to the standard chi-square gate, and can be computed directly as

$$w_n = \frac{1}{\sqrt{(2\pi)^d |\mathbf{S}|}} \exp\left(-\frac{1}{2} M_d\right) \quad (8)$$

Thus, a particular  $\mathbf{z}_i$  is accepted if  $p(\mathbf{z} = \mathbf{z}_i) > w_n$ .

In general,  $p(\mathbf{z})$  can be obtained from the observation model and the prior as

$$p(\mathbf{z}) = \int p(\mathbf{z} | \mathbf{x}) p(\mathbf{x}) d\mathbf{x} \quad (9)$$

which is a marginalisation over  $\mathbf{x}$  of the joint PDF  $p(\mathbf{z}, \mathbf{x})$ . If the observation model is linear, as in (1), and the PDFs are represented by GMMs,

$$p(\mathbf{x}) = \sum_{i=1}^N \alpha_i \mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}_i, \mathbf{P}_i) \quad (10)$$

$$p(\mathbf{r}) = \sum_{j=1}^M \beta_j \mathcal{N}(\mathbf{r}; \hat{\mathbf{r}}_j, \mathbf{R}_j) \quad (11)$$

then there is a closed-form solution.

$$p(\mathbf{z}) = \sum_i^N \sum_j^M \alpha_i \beta_j \mathcal{N}(\mathbf{z}; \mathbf{H}\hat{\mathbf{x}}_i + \hat{\mathbf{r}}_j, \mathbf{H}\mathbf{P}_i\mathbf{H}^T + \mathbf{R}_j) \quad (12)$$

However, for a non-linear observation model an approximation is required. In our experiments, we chose a simple, though expensive, Monte Carlo sampling approach.

1. Draw  $N$  samples  $\mathbf{x}_i \sim p(\mathbf{x})$ .
2. The conditional PDF is  $p(\mathbf{z} | \mathbf{x}_i) = \mathbf{h}(\mathbf{x}_i) + p(\mathbf{r})$ .
3. Sum the conditionals to marginalise over  $\mathbf{x}$ ,  $p(\mathbf{z}) \approx \frac{1}{N} \sum_i^N (\mathbf{h}(\mathbf{x}_i) + p(\mathbf{r}))$ .

For the model in (7), with  $p(\mathbf{r})$  Gaussian, we get

$$p(z) \approx \frac{1}{N} \sum_i^N \mathcal{N}(z; h(\mathbf{x}_i, \mathbf{x}_s), R) \quad (13)$$

Figure 2 shows an example validation gate obtained for the prior  $p(\mathbf{x})$  from Section 4 when the sensor location is moved to  $\mathbf{x}_s = [10, 10]^T$ . The resulting  $p(z)$

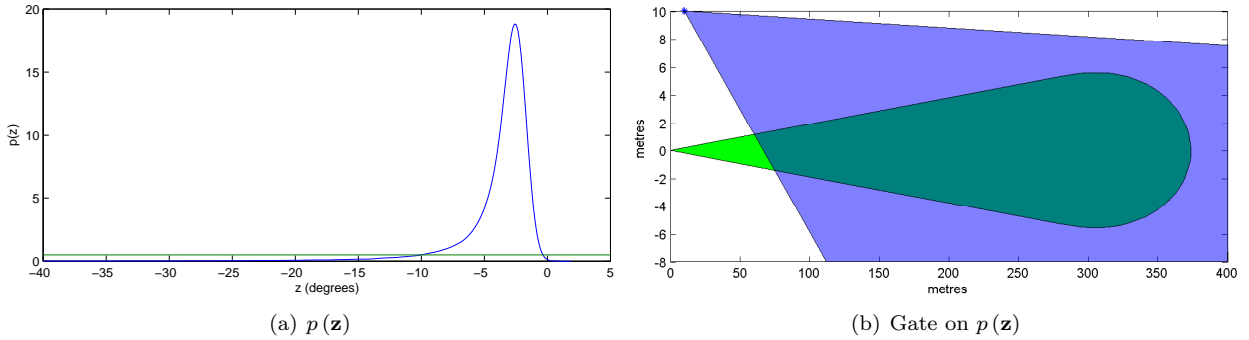


Figure 2: Validation gate for a bearing-only scenario, where  $p(\mathbf{x})$  is depicted in (b) by a shaded green region (marking 95% of its mass), and the sensor location  $\mathbf{x}_s = [10, 10]^T$  by a \*. The threshold weight  $w_{95}$  is shown by a horizontal line in (a), and the resulting region of valid bearing observations radiating from  $\mathbf{x}_s$  are shown in (b) by the shaded blue region.

is shown in Figure 2(a). Sampling from  $p(z)$ , 95% of samples lie between  $-10^\circ$  and  $-0.36^\circ$ , and the threshold weight  $w_{95}$  is 0.5. Thus, observed angles must be  $-10^\circ < z_i < -0.36^\circ$  to be accepted; this is depicted in Figure 2(b).<sup>4</sup>

Clearly, rejecting angles less than  $-10^\circ$ , means rejecting observations that, intuitively, are quite feasible; effectively it rejects the possibility of a target close to the origin. For example, from  $p(\mathbf{x})$ , the probability density of the true target location being at, say,  $[50, 0]^T$  is equal to it being at  $[250, 0]^T$ . This agrees with our intuition where, based on a single bearing-only measurement, we would not be surprised if the target turned out to be near the origin. However, if the true target *was* at  $\mathbf{x}_t = [50, 0]^T$ , so that observations actually received are distributed according to  $p(z|\mathbf{x} = \mathbf{x}_t)$ , 95% of observations would fall between  $-15^\circ$  and  $-13^\circ$ , and all of them would be rejected by this validation gate.

The problem with  $p(\mathbf{z} = \mathbf{z}_i)$  is that its meaning—“the likelihood of obtaining  $\mathbf{z}_i$  when  $\mathbf{x}_t$  could be anywhere in  $p(\mathbf{x})$ ”—does not correspond to our essential requirement for a validation gate. From  $p(\mathbf{z})$  we predict an observation is most likely to come from the direction where concentration of mass in  $p(\mathbf{x})$  is greatest; this is apparent from Figure 2(a). However, the likelihood of an observation is not the same as the *feasibility* of an observation. From the density of  $p(\mathbf{x})$  it is clear that a target location at  $[50, 0]^T$  is as feasible as at  $[250, 0]^T$ .

Thus, we make the following distinction. Before any measurement has been received, the density of  $p(\mathbf{z})$  predicts which observations are most likely, while, having received a particular observation  $\mathbf{z}_i$ , the density of  $p(\mathbf{x})$  indicates whether it might originate from a feasible target. We contend that a validation gate is concerned with the feasibility of an *existing* measurement, not the likelihood of a measurement yet to arrive.

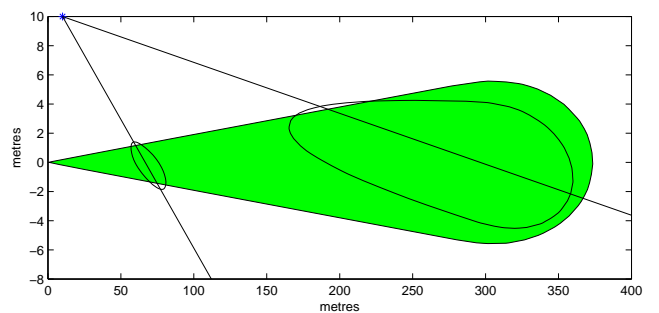


Figure 3: A multiple hypothesis example, where  $p(\mathbf{x})$  and  $\mathbf{x}_s$  are as in Figure 2. The sensor returns two measurements, shown as lines, at  $-10^\circ$  and  $-2^\circ$ . The two possible updates for the target are depicted by contours bounding the 95% mass concentration of the respective posterior PDFs.

## 6 $p(\mathbf{z})$ as a Measure of Track Likelihood

While not sensible as a measure of validation,  $p(\mathbf{z})$  is a reasonable measure of track likelihood when there exist multiple track hypotheses. Consider the example in Figure 3. The prior  $p(\mathbf{x})$  and sensor location  $\mathbf{x}_s$  are the same as in Figure 2, and it is known that there exists a single target only. However, the sensor returns two measurements,  $z_1 = -10^\circ$  and  $z_2 = -2^\circ$ , one of the target and the other of clutter. If the clutter is assumed to be uniformly distributed, then there are two possible hypotheses:

$$\begin{aligned} \mathcal{H}_1 : \quad & z_1 = h(\mathbf{x}, \mathbf{x}_s) + r \\ & z_2 = c, \quad p(c) = \frac{1}{2\pi}, \quad -\pi < c \leq \pi \quad (14) \\ \mathcal{H}_2 : \quad & z_1 = c, \quad p(c) = \frac{1}{2\pi}, \quad -\pi < c \leq \pi \\ & z_2 = h(\mathbf{x}, \mathbf{x}_s) + r \quad (15) \end{aligned}$$

These hypotheses may be ranked, as to one being more likely than the other, by a general mechanism known as *Bayesian model comparison* [6, Chapter 28]. Whenever performing an update using Bayes theorem,

<sup>4</sup>For clarity, the coordinate axes in Figure 2(b), and subsequent figures, are not equal. However, this has the unfortunate effect of exaggerating angles, which distorts the true shape of the validation region, making it appear much larger than it really is.

the models of the system are always an implicit conditioning parameter  $\mathcal{H}$ .

$$p(\mathbf{x}|\mathbf{z}, \mathcal{H}) = \frac{p(\mathbf{z}|\mathbf{x}, \mathcal{H})p(\mathbf{x}|\mathcal{H})}{p(\mathbf{z}|\mathcal{H})} \quad (16)$$

Typically, the posterior is computed as proportional to the numerator of the right-hand-side, and  $p(\mathbf{z}|\mathcal{H})$  is ignored as a normalising constant. When there exist multiple alternative hypotheses  $\{\mathcal{H}_1, \dots, \mathcal{H}_N\}$ , however, the value of  $p(\mathbf{z} = \mathbf{z}_i|\mathcal{H}_m)$  is the ‘‘evidence for  $\mathcal{H}_m$ ’’; it quantifies the relative merit of a hypothesis for predicting the observation  $\mathbf{z}_i$ , and allows different models to be ranked.

In our example, where  $\mathbf{z} = [z_1, z_2]^T$  is the pair of observations,  $p(\mathbf{z}|\mathcal{H}) = p(z_1|\mathcal{H})p(z_2|\mathcal{H})$  as  $z_1$  and  $z_2$  are independent. Given the measurement  $z_1 = -10^\circ$  and the hypothesis  $\mathcal{H}_1$ , the scalar value of  $p(z_1|\mathcal{H}_1)$  is computed as

$$p(z_1|\mathcal{H}_1) = \int p(z_1|\mathbf{x}, \mathcal{H}_1)p(\mathbf{x})d\mathbf{x} \quad (17)$$

For a GMM implementation, the prior and likelihood function are each weighted sums of Gaussians over the domain of  $\mathbf{x}$ .

$$p(\mathbf{x}) = \sum_i^N \alpha_i \mathcal{N}(\mathbf{x}; \hat{\mathbf{p}}_i, \mathbf{P}_i) \quad (18)$$

$$p(z_1|\mathbf{x}, \mathcal{H}_1) = \sum_j^M \beta_j \mathcal{N}(z_1; \hat{\mathbf{q}}_j, \mathbf{Q}_j) \quad (19)$$

Therefore, the value of (17) is found by multiplying (18) and (19) and summing the weights of the resultant GMM. This integral can be computed directly as

$$p(z_1|\mathcal{H}_1) = \sum_i^N \sum_j^M \alpha_i \beta_j \frac{1}{2\pi\sqrt{|\mathbf{S}|}} \exp\left(-\frac{1}{2}\nu^T \mathbf{S}^{-1} \nu\right) \quad (20)$$

where

$$\begin{aligned} \nu &= \hat{\mathbf{p}}_i - \hat{\mathbf{q}}_j \\ \mathbf{S} &= \mathbf{P}_i + \mathbf{Q}_j \end{aligned}$$

Computing  $p(z_2|\mathcal{H}_2)$  is similar. The value of  $p(z_2|\mathcal{H}_1)$  and  $p(z_1|\mathcal{H}_2)$  is simply  $1/2\pi$ . This gives the ranking weights as  $p(\mathbf{z}|\mathcal{H}_1) = 0.06$  and  $p(\mathbf{z}|\mathcal{H}_2) = 2.8$ . Thus, the hypothesis that  $z_2$  observes the target is by far the more likely model.

This result makes sense as  $p(\mathbf{z})$  predicts that targets more distant from the origin are more likely, therefore the hypothesis supporting the more distant target has higher probability.

## 7 $p(\mathbf{x}, \mathbf{z})$ as a Validation Likelihood

Another possible measure of validation likelihood is the joint PDF  $p(\mathbf{x}, \mathbf{z}) = p(\mathbf{z}|\mathbf{x})p(\mathbf{x})$ . Given an observation  $\mathbf{z}_i$ , the product of the likelihood function  $p(\mathbf{z} = \mathbf{z}_i|\mathbf{x})$  and the prior  $p(\mathbf{x})$  produces a slice

through the joint PDF  $p(\mathbf{x}, \mathbf{z} = \mathbf{z}_i)$  along the hyperplane  $\mathbf{z} = \mathbf{z}_i$ .<sup>5</sup> The value of  $p(\mathbf{x}, \mathbf{z} = \mathbf{z}_i)$  at each point  $\mathbf{x}_j$  is the probability that both  $\mathbf{z}_i$  is valid *and* the true target location is  $\mathbf{x}_j$ . The mode

$$\mathbf{x}_{mode} = \arg \max p(\mathbf{x}, \mathbf{z} = \mathbf{z}_i) \quad (21)$$

gives the most likely target location possible for  $\mathbf{z}_i$ . We propose that the mode-weight of  $p(\mathbf{x}, \mathbf{z} = \mathbf{z}_i)$  is the correct measure for likelihood of ‘‘association validity’’ for  $\mathbf{z}_i$ .

Restating the algorithm from Section 3, the full validation gate computation is performed as follows.

1. Draw  $N$  samples  $\mathbf{z}_i \sim p(\mathbf{z})$ . Note, it is not necessary to compute  $p(\mathbf{z})$  in order to draw samples from it. Samples may be generated as  $\mathbf{x}_i \sim p(\mathbf{x})$  then  $\mathbf{z}_i \sim p(\mathbf{z}|\mathbf{x}_i)$ .
2. Weight each  $\mathbf{z}_i$  according to its validation likelihood.
  - (a) Generate the likelihood function  $p(\mathbf{z} = \mathbf{z}_i|\mathbf{x})$ .
  - (b) Compute a slice of the joint PDF along  $\mathbf{z}_i$ ,  $p(\mathbf{x}, \mathbf{z} = \mathbf{z}_i) = p(\mathbf{z} = \mathbf{z}_i|\mathbf{x})p(\mathbf{x})$ .
  - (c) Compute the mode  $w_i = \max p(\mathbf{x}, \mathbf{z} = \mathbf{z}_i)$ .
3. Select the  $m$ -th largest weight, where  $m = Nn/100$ , to obtain an  $n\%$  acceptance threshold,  $w_n$ .

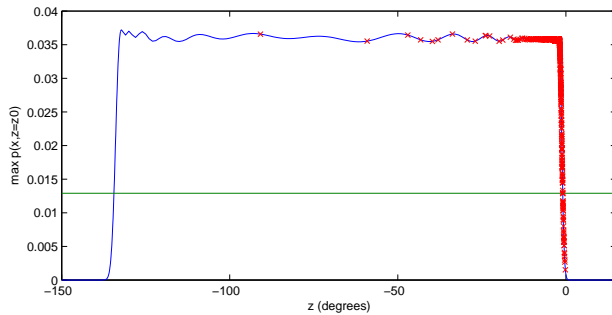
Step 2(b) is the product of the likelihood function and the prior; this operation has a closed-form for both Gaussians and GMMs. Step 2(c) is closed-form for Gaussians, since the mean of  $p(\mathbf{x}, \mathbf{z} = \mathbf{z}_i)$  is also its mode, but there is no analytical solution for GMMs. To find the mode of a general GMM requires a global optimisation algorithm. We implemented the search using a combination of Monte Carlo sampling and local optimisation, but more efficient solutions exist (e.g., [2]).

When  $p(\mathbf{x}, \mathbf{z} = \mathbf{z}_i)$  is Gaussian, its mode is proportional to its volume, i.e.,  $\max p(\mathbf{x}, \mathbf{z} = \mathbf{z}_i) \propto p(\mathbf{z} = \mathbf{z}_i)$  for all  $\mathbf{z}_i$ . Therefore, for linear-Gaussian models, the  $p(\mathbf{x}, \mathbf{z})$  gate is equivalent to the  $p(\mathbf{z})$  gate and, in turn, to the standard ellipsoidal gate.

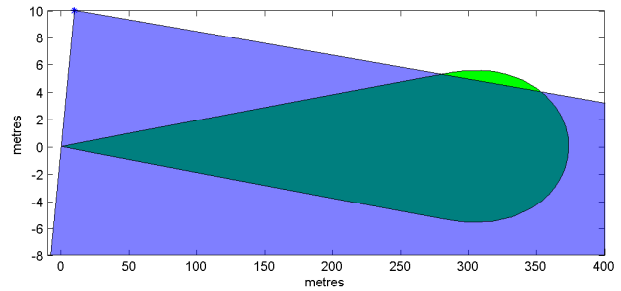
Returning to the bearing-only example from Section 5. The samples drawn from  $p(z)$  are the same as before, but are weighted according to the mode of  $p(\mathbf{x}, z = z_i)$ , as shown in Figure 4(a). (Note, the top of Figure 4(a) should be flat, but a ripple is introduced by the GMM approximation.) Because the samples are from  $p(z)$ , the resultant weight threshold  $w_{95} = 0.013$  accepts 95% of true measurements, the same as for the gate in Section 5. But this time, the 5% rejected are those that intersect with  $p(\mathbf{x})$  along regions of low joint probability. In Figure 4(b) the valid region  $-135^\circ < z_i \leq -1^\circ$  includes locations close to the origin, as is expected intuitively.

A point on the joint PDF  $p(\mathbf{x} = \mathbf{x}_j, \mathbf{z} = \mathbf{z}_i)$  is the probability of  $\mathbf{x}_j$  and  $\mathbf{z}_i$  both being true together.

<sup>5</sup>Be careful to avoid confusion here. A slice through the joint  $p(\mathbf{x}, \mathbf{z} = \mathbf{z}_i)$  is not the same as the conditional  $p(\mathbf{x}|\mathbf{z} = \mathbf{z}_i)$ .



(a) Mode of  $p(\mathbf{x}, z = z_i)$



(b) Gate on  $\max p(\mathbf{x}, z = z_i)$

Figure 4: The values of  $p(x)$  and  $\mathbf{x}_s$  in (b) are the same as in Figure 2(b). In (a), the mode of  $p(\mathbf{x}, z = z_i)$  is shown as a function of  $z_i$ . The weighted samples from  $p(z)$  are red crosses, and the threshold weight  $w_{95}$  is shown by a horizontal line. The valid observation region is shown in (b) by the shaded blue region.

Given a measurement  $\mathbf{z}_i$ , the mode of  $p(\mathbf{x}, \mathbf{z} = \mathbf{z}_i)$  indicates the most likely target location possible for  $\mathbf{z}_i$ . If  $\mathbf{z}_i$  permits a likely target location, it is considered to be feasible, and so  $p(\mathbf{x}, \mathbf{z} = \mathbf{z}_i)$  seems a reasonable measure for a validation gate.

## 8 Conclusion

This paper extends the concept of the ellipsoidal validation gate to non-linear non-Gaussian systems. The proposed gate is based on the joint target-observation PDF  $p(\mathbf{x}, \mathbf{z})$  and exhibits the same “probability of acceptance” characteristics as the ellipsoidal gate.

An alternative gate, based on the marginal observation PDF  $p(\mathbf{z})$ , was found to be unsuitable. It tends to reject observations which intuitively seem possible and, indeed, would be probable if the true target location happens to be in a region of high probability density but low mass concentration. The problem is that  $p(\mathbf{z})$  is a predictive measure for the likelihood of future observations, whereas validation requires a feasibility measure for an observation actually received. Nevertheless,  $p(\mathbf{z})$  is a legitimate Bayesian measure of track likelihood when tracking multiple target hypotheses.

The algorithms in this paper are computationally intensive, and our focus has been on theoretical aspects rather than actual implementation. We expect that more efficient approximations will be found (e.g., applying the unscented transform [4] to non-linear GMM transformations rather than Monte Carlo sampling). The main contribution of this paper is to develop a theory concerning the form of an ideal validation gate. We propose that the mode of  $p(\mathbf{x}, \mathbf{z} = \mathbf{z}_i)$  is the correct measure, giving the right acceptance statistics and permitting feasible associations. As a fundamental measure, it provides a yardstick for comparing the veracity of heuristic or approximate gates developed for practical application.

## Acknowledgement

This work is supported by the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government.

## References

- [1] Y. Bar-Shalom and T.E. Fortmann. *Tracking and Data Association*. Academic Press, 1988.
- [2] M.A. Carreira-Perpiñán. Mode-finding for mixtures of Gaussian distributions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1318–1323, 2000.
- [3] T.E. Fortmann, Y. Bar-Shalom, and M. Scheffe. Sonar tracking of multiple targets using joint probabilistic data association. *IEEE Journal of Oceanic Engineering*, 8(3):173–184, 1983.
- [4] S. Julier, J. Uhlmann, and H.F. Durrant-Whyte. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, 45(3):477–482, 2000.
- [5] T. Kirubarajan and Y. Bar-Shalom. Probabilistic data association techniques for target tracking in clutter. *Proceedings of the IEEE*, 92(3):536–557, 2004.
- [6] D.J.C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.
- [7] L. Pao. Multisensor multitarget mixture reduction algorithms for tracking. *Journal of Guidance, Control, and Dynamics*, 17(6):1205–1211, 1994.
- [8] L.Y. Pao and R.M. Powers. A comparison of several different approaches for target tracking with clutter. In *American Control Conference*, pages 3919–3924, 2003.
- [9] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C*. Cambridge University Press, 2nd edition, 1992.
- [10] D.B. Reid. An algorithm for tracking multiple targets. *IEEE Transactions on Automatic Control*, 24(6):843–854, 1979.
- [11] J. Vermaak, S.J. Godsill, and P. Pérez. Monte Carlo filtering for multi-target tracking and data association. *IEEE Transactions on Aerospace and Electronic Systems*, 41(1):309–332, 2005.
- [12] X. Wang, S. Challa, and R. Evans. Gating techniques for maneuvering target tracking in clutter. *IEEE Transactions on Aerospace and Electronic Systems*, 38(3):1087–1097, 2002.