Defence Research and Development Canada   Recherche et développement pour la défense Canada

DEFENCE R&D DÉFENSE

Statistical Decision Thresholds for Pulse Template Cross-Correlators

Fred A. Dilkes

Defence R&D Canada – Ottawa

TECHNICAL MEMORANDUM

DRDC Ottawa TM 2005-167

November 2005

Canada

Statistical Decision Thresholds for Pulse Template Cross-Correlators

Fred A. Dilkes

Defence R&D Canada – Ottawa

Technical Memorandum

DRDC Ottawa TM 2005-167

November 2005

Author

Original signed by Fred A. Dilkes

Fred A. Dilkes

Approved by

Original signed by Jean-François Rivest

Jean-François Rivest
Acting Section Head/Radar Electronic Warfare

Approved for release by

Original signed by Anthony Ashley

Anthony Ashley
Chair/Document Review Panel

Abstract

A statistical formalism is proposed to evaluate the performance of waveform-based signal detection techniques for Electronic Support. The formalism applies to a large class of cross-correlation methods that attempt to match the arrival times of a sequence of pulses against an *a priori* pulse template. In this framework, the detection of a particular signal of interest is treated as a problem from binary decision theory and the observations are treated as outcomes of a point process. The presence of a signal is modeled using a pulse template with non-cumulative Gaussian jitter while background noise is modeled using a stationary Poisson point process. Two specific cross-correlation techniques are evaluated and their performances are compared in terms of false-alarm and detection probabilities. The capability of these techniques to discern between the presence and absence of the signal is measured as a function of a decision threshold.

Résumé

Un formalisme statistique est proposé pour l'évaluation de la performance de techniques de détection de signaux fondées sur la forme d'onde pour le soutien électronique. Le formalisme s'applique à un vaste groupe de méthodes de corrélation croisée qui tentent de faire concorder les heures d'arrivée d'une séquence d'impulsions avec un gabarit d'impulsion a priori. Dans ce cadre, la détection d'un signal particulier présentant de l'intérêt est traitée comme un problème découlant de la théorie des décisions binaires, et les observations sont traitées comme les résultats d'un processus ponctuel. La présence d'un signal est modélisée à l'aide d'un gabarit d'impulsion avec instabilité gaussienne non cumulative, tandis que le bruit de fond est modélisé au moyen d'un processus ponctuel de Poisson stationnaire. Deux techniques particulières de corrélation croisée sont évaluées, et leurs rendements sont comparés en termes de probabilités de détection et de fausse alarme. Leur capacité à discerner entre la présence et l'absence d'un signal est mesurée en fonction d'un seuil de décision.

This page intentionally left blank.

Executive summary

Statistical Decision Thresholds for Pulse Template Cross-Correlators

Fred A. Dilkes; DRDC Ottawa TM 2005-167; Defence R&D Canada - Ottawa; November 2005.

The field of radar "Electronic Support" (ES) depends heavily on the availability of a robust mechanism with which to analyze sequences of pulses. The complexity of the complete ES problem has lead to the development of numerous techniques designed to separate and identify interleaved signals.

An emerging approach to pulse-train analysis involves searching for specific sequences of pulses that are known to be associated with particular radar systems, using *a priori* patterns and parameters. This approach has the advantage over some of the more traditional approaches, such as delta-τ autocorrelation histograms, in that the patterns of interest can be quite complicated, and need not be periodic. The most straightforward implementation of such an approach is to perform a cross-correlation between the observed pulse sequence and some specified template sequence in order to detect a good match. However, unlike in applications of digital signal processing, there does not seem to be a unique or canonical approach to pulse-train cross-correlation.

The output of these cross correlation techniques is a scoring profile that depends on the value of some continuous-valued time-offset. High values of the score are associated with some elevated confidence that the signal of interest is present at the indicated time. Low values of the score are expected to indicate the absence of such a signal.

Naturally, channel impairments conspire to make precise error-free signal detection an impossibility. The objective of this memo is to present an analysis based on decision-theoretic principles to evaluate candidate cross-correlation techniques and to quantify their ability to distinguish between the absence or presence of a specified signal. Specifically, the following questions will be addressed:

- How high/low does the score need to be in order to decide that the signal is present/absent at some specified time?
- How likely is it that such a decision is actually incorrect?
- Are some scoring techniques and threshold values likely to lead to fewer incorrect decisions than others?

These issues are addressed in a probabilistic manner based on two competing hypotheses: if the signal is absent (noise-only) then the pulse train is modeled as a stationary Poisson point process; if the signal is present (signal-in-noise) then the pulse train is a point process consisting of the union of Poisson noise events and a Gaussian signal with non-cumulative jitter.

A mathematical decision framework is presented for two different types of cross-correlation implementations. The formalism includes false-alarm and detection probabilities along with the notion of a "Receiver Operating Characteristic". All of the proposed methods are illustrated with two examples.

Sommaire

Statistical Decision Thresholds for Pulse Template Cross-Correlators

Fred A. Dilkes; DRDC Ottawa TM 2005-167; R & D pour la défense Canada - Ottawa; novembre 2005.

Le soutien électronique (SE) radar dépend fortement de la disponibilité d'un mécanisme solide permettant d'analyser des séquences d'impulsions. La complexité du problème de SE a mené à la mise au point de nombreuses techniques pour séparer et identifier des signaux entrelacés.

Une approche émergente de l'analyse des trains d'impulsions comporte la recherche de séquences particulières d'impulsions que l'on sait être associées à des systèmes radars particuliers, au moyen de paramètres et de diagrammes a priori. Cette approche est avantageuse par rapport à certaines approches classiques, comme les histogrammes d'autocorrélation delta-τ, dans le sens que les diagrammes présentant de l'intérêt peuvent être assez complexes et n'ont pas besoin d'être périodiques. La mise en œuvre la plus directe de cette approche consiste à effectuer une corrélation croisée entre la séquence d'impulsions observée et une séquence particulière de gabarit pour détecter une bonne concordance. Toutefois, contrairement à ce qui se passe dans les applications de traitement numérique des signaux, il ne semble pas y avoir d'approche unique ou canonique en matière de corrélation croisée des trains d'impulsions.

Ces techniques de corrélation croisée donnent un profil de notation qui dépend de la valeur d'un certain décalage temporel à valeur continue. Des valeurs élevées de la notation sont associées à un degré de certitude assez élevé que le signal d'intérêt est présent à l'heure indiquée. On s'attend à de faibles valeurs de la notation pour indiquer l'absence de signal.

Naturellement, les dégradations des canaux contribuent à rendre impossible la détection précise de signal sans erreur. L'objectif de la présente note est de présenter une analyse fondée sur des principes théoriques de décision pour évaluer les techniques possibles de corrélation croisée et quantifier leur capacité à discerner entre l'absence et la présence d'un signal donné. Plus précisément, les questions qui suivent seront étudiées :
- À quel point la notation doit-elle être élevée/faible pour que l'on décide que le signal est présent/absent à une heure donnée ?
- Quelles sont les probabilités que la décision soit, en fait, incorrecte ?
- Y a-t-il des techniques de notation et des valeurs de seuil susceptibles de mener à moins de décisions incorrectes que d'autres ?

Ces questions sont étudiées de façon probabiliste à partir de deux hypothèses concurrentes : si le signal est absent (il n'y a que du bruit), le train d'impulsions est alors

modélisé comme processus ponctuel de Poisson stationnaire ; si le signal est présent (signal mêlé au bruit), le train d'impulsions est un processus ponctuel composé de l'union d'événements de bruit de Poisson et d'un signal gaussien avec de l'instabilité non cumulative.

Un cadre de décision mathématique est présenté pour deux types de mise en œuvre de corrélation croisée. Le formalisme comprend les probabilités de détection et de fausse alarme, de même que la notion d'une *caractéristique de fonctionnement du récepteur*. Toutes les méthodes proposées sont illustrées par deux exemples.

Table of contents

List of figures

Acknowledgements

The author would like to thank Stephen Sung for valuable discussions related to the distribution of the optimal decision score and specifically for recognizing Equation (A.8).

This page intentionally left blank.

1 Introduction

The field of "Electronic Support" (ES) depends heavily on the availability of a robust mechanism with which to analyze sequences of pulses. In general, the complete stream of radar pulses received by an ES system is composed of multiple signals exhibiting diverse features, both periodic and aperiodic, that are observed through a common, imperfect channel. The complexity of the complete ES problem has lead to the development of numerous techniques designed to separate and identify the constituent signals.

An emerging approach to pulse-train analysis involves searching for specific sequences of pulses that are known to be associated with particular systems, using *a priori* patterns and parameters. This approach has the advantage over some of the more traditional approaches, such as delta-τ autocorrelation histograms [1, 2, 3] in that the patterns of interest can be quite complicated, and need not be periodic. The most straightforward implementation of such an approach is to perform some form of cross-correlation between the observed pulse sequence a some specified template sequence in order to detect a good match. Unlike in applications of digital signal processing, there does not seem to be a unique or canonical approach to pulse-train cross-correlation. Proposed methods include "cross-correlation histograms" [4] and "$q_{\mathcal{N}}$-scores" [5], both of which deal exclusively with the arrival times of a sequence of pulses, to the exclusion of other pulse descriptor words. Other related techniques have been explored as part of a collaboration under The Technical Cooperation Panel [6].

The output of these cross-correlation techniques is a scoring profile that depends on the value of some continuous-valued time-offset. High values of the score are associated with some elevated confidence that the signal of interest is present at the indicated time. Low values of the score are expected to indicate the absence of such a signal.

Naturally, when the ES receiver is faced with such realities as channel interference, imperfect receiver characteristics, and signal jitter, there is no detection algorithm that can claim to be a perfect indicator for the presence of a signal. These impairments conspire to make precise error-free signal detection an impossibility. The objective of this memo is to present an analytical framework based on decision-theoretic principles [7] to evaluate candidate cross-correlation techniques and to quantify their ability to distinguish between the absence or presence of a specified signal. Specifically, the following questions will be addressed:

- How high/low does the score need to be in order to decide that the signal is present/absent at some specified time?
- How likely is it that such a decision is actually incorrect?

- Are some scoring techniques and threshold values likely to lead to fewer incorrect decisions than others?

These issues are addressed in a probabilistic manner based on two competing hypotheses: if the signal is absent (noise-only) then the pulse train is modeled as a stationary Poisson point process; if the signal is present (signal-in-noise) then the pulse train is a point process consisting of the union of Poisson noise events and a Gaussian signal with non-cumulative jitter. The methods presented bare some resemblance to those of [8] where point processes and statistical decision criteria were used to distinguish pulse train jitter models.

The memo is arranged as follows. Section 2 presents a general quantitative formalism for the problem and defines the statistical models of the two hypotheses. Sections 3 and 4 present statistical performance analyses of the $q_{\mathcal{N}}$-score and histogram scores. Section 5 presents an example and discussion. Concluding remarks appear in Section 6. Finally, Annex A attempts to develop a theory of optimal cross-correlation scoring by drawing a connection between pulsed signals and the theory of point processes [9].

2 Decision framework

Suppose that some discrete set of observed pulse arrival times, denoted by[1] $\mathbf{X}$, is observed by some Electronic Support receiver. For the sake of simplicity, only the arrival times of the individual pulses is considered in this memo so that $\mathbf{X} \subset \mathbb{R}$ where the real line $\mathbb{R}$ represents the space of all possible times. Additional pulse descriptor words, such as pulse-width and carrier frequency may be integrated into this formalism with minor modifications.

The pulses contained in $\mathbf{X}$ may originate from multiple radar signals, all sharing the same noisy propagation channel. In an effort to identify the individual constituents, one attempts to match this observation against some *a priori* expected behaviour of particular signal of interest. In this investigation, it is assumed that the signal of interest can be represented over some period of time by a template $T \subset \mathbb{R}$. The template consists of a finite set of discrete times describing the leading edges of various pulses, relative to some unknown time offset denoted by τ. Roughly speaking, if the template "starts at time-offset τ", then for each time element $t \in T$, one should expect to find a pulse in $\mathbf{X}$ that is near to $t + \tau$. Whether or not each such pulse is actually detected, and its exact time-of-arrival, depends the detailed statistics of the signal and channel.

Suppose that the degree with which $\mathbf{X}$ matches T at time-offset τ is measured by

[1] Bold-faced symbols indicate random processes and/or random variables.

some scoring profile, that is generically denoted by

$$\mathbf{s}_\tau \equiv s(\mathbf{X}, T; \tau)\,. \tag{1}$$

For each fixed value of τ, the value of the score $\mathbf{s}_\tau$ is determined by a specified scoring functional $s(\cdot, \cdot; \tau)$. This functional is taken to implement some kind of cross-correlation technique that may be defined by either rigorous or *ad hoc* reasoning. In any case, one assumes that it has been designed in such a way so that that larger values of the score are intended to indicate some degree of confidence that the signal of interest is actually present in the observation; smaller values express some lack of confidence.

A statistical measure of the performance of these functionals requires the use of some statistical model for the pulse observations $\mathbf{X}$. For simplicity, we suppose a decision needs to be made between two possible hypotheses: either,

1. the observation $\mathbf{X}$ is explained by a random noise process, or
2. $\mathbf{X}$ has been generated by the union of a random noise process and a signal.

A reasonable but very simple statistical model for each of these two cases is described in Subsections 2.1 and 2.2 respectively.

2.1 Noise-only model

If the signal of interest is not present in the observation then it is simply assumed that the observation is given by

$$\mathbf{X} = \mathbf{X}_\rho \tag{2}$$

where $\mathbf{X}_\rho$ consists of a set of events drawn from a stationary Poisson process with an average density of ρ pulses per unit time. Some aspects of the statistics of such a process are reviewed in Appendix B.2.

2.2 Signal-in-noise model

If the signal of interest is present then the observation is modeled by the union

$$\mathbf{X} = \mathbf{X}_\rho \cup \mathbf{X}_T \tag{3}$$

where the Poisson noise process $\mathbf{X}_\rho$ is described in the preceding subsection, and $\mathbf{X}_T$ represents the set of detected pulses arising from the signal.

The statistics of signal $\mathbf{X}_T$ are based on the hypothesis that the template T starts at time offset τ. Each element of the template $t \in T$ is associated with a corresponding "transmitted pulse time" denoted by $\mathbf{x}_t$. These are taken to be independent Gaussian variables whose distributions are determined by τ and some measure of jitter variance denoted by σ^2,

$$\mathbf{x}_t \sim \mathcal{N}(t+\tau, \sigma^2)\ ,\ t \in T\ . \tag{4}$$

The distribution in (4) represents a Gaussian non-cumulative jitter model [8]. The set of transmitted pulses $\{\mathbf{x}_t | t \in T\}$ is then randomly mapped to the detected pulse subset $\mathbf{X}_T$ by a binary erasure channel. Each pulse $\mathbf{x}_t$ has probability P_p of appearing in the detected pulse set $\mathbf{X}_T$; conversely each has a probability of $1 - P_\mathrm{p}$ of being missed and not appearing in $\mathbf{X}_T$.

Note that real applications are such that some pulses may be unobservable simply because they lie outside of the intercept opportunity of the receiver. To accommodate such effects, one may define an "observation window" denoted $\mathcal{X} \subseteq \mathbb{R}$ that represents the period of time during which the receiver is open to the reception of radar pulses. In any real-world application $\mathcal{X}$ must be a finite union of bounded time intervals. The observed pulses would then be required to lie in this window ensuring that $\mathbf{X} \subset \mathcal{X}$. Events of the Poisson noise process $\mathbf{X}_\rho$ would be drawn from the set $\mathcal{X}$ and the signal $\mathbf{X}_T$ would exclude any transmitted pulse for which $\mathbf{x}_t \notin \mathcal{X}$. Under these conditions, $\mathbf{X}$ becomes a "finite point process" [9]. A more complete definition of the probability space of pulse sequence observations on $\mathcal{X}$ has been deferred to Annex A.1, along with its probability measures under models (2) and (3). The effect of a finite window is a very important and interesting topic for future investigation. Hereafter, the simpler but less realistic assumption $\mathcal{X} = \mathbb{R}$ will be adopted.

2.3 Measures of performance

A performance measure for the generic scoring functional (1) should reflect its ability to distinguish between the presence of the signal with template T and its absence. Given an observation $\mathbf{X}$ one is required to determine which of the models described (2) or (3) best describes those observations. For simplicity, it is assumed here that one is only interested in making a hard decision. That is, one chooses a threshold value, denoted by γ_s, and then determines whether or not the actual score exceeds the threshold

$$\mathbf{s}_\tau \gtrless \gamma_s\ . \tag{5}$$

If $\mathbf{s}_\tau \geq \gamma_s$ then one estimates that the signal is present; if $\mathbf{s}_\tau < \gamma_s$ then the signal is estimated to be absent.

For the purpose of this paper, it is a assumed that template T used to evaluate the score (1) is accurate and agrees with template used to determine the statistics of $\mathbf{X}_T$, as described in (4). The robustness of such scores against incorrect templates is outside of the scope of the present contribution, but could be an interesting topic for future study.

More importantly, it shall be assumed that the time-offset τ used to score the observation in (1) agrees with the time-offset in the signal model (4). This subtlety is potentially very important since one is generally faced with a complete scoring profile $\{\mathbf{s}_\tau | \tau \in \mathbb{R}\}$ without prior knowledge the actual time at which the signal may have started. Such a problem falls into the category of "composite hypothesis" decisions [7]. Many scoring functions may tend to exhibit large values for some range of offset values in the neighbourhood of the true value. It is not possible to perform a meaningful analysis of the effects of this dispersion without further details concerning how the output of (1) will be used in further processing. This presents a complication that lies outside of the scope of this memo and must be evaluated on a case-by-case basis that cannot be exhaustively discussed here.

Naturally, there is no guarantee that the decision made will be correct. Indeed model (2) may lead to a score that exceeds the threshold (5) and model (3) may lead to a score that falls short of the threshold. Both of these undesirable situations can be represented by finite probabilities. Since $\mathbf{X}$ is a random process the value of the score $\mathbf{s}_\tau \equiv s(\mathbf{X}, T; \tau)$ represents a random scalar variable, for fixed T and τ. Under each of the hypotheses (2) and (3), the distribution for $\mathbf{s}_\tau$ can be determined. Suppose that these distributions are described by conditional probability density functions denoted respectively by

$$p_{\mathbf{s}_\tau | \mathbf{X}_\rho}(s) \quad \text{and} \quad p_{\mathbf{s}_\tau | \mathbf{X}_\rho \cup \mathbf{X}_T}(s) \, .$$

For consistency with commonly used terminology, the situation in which the signal is absent, $\mathbf{X} = \mathbf{X}_\rho$, and yet the score exceeds the threshold $\mathbf{s}_\tau \geq \gamma_s$ is known as a "false alarm". Using the decision criterion (5), the probability of a false alarm at any time offset τ, conditioned on the noise model (2), is given by the integral

$$P_{\mathrm{fa}}(\gamma_s) = \int_{\gamma_s}^{\infty} p_{\mathbf{s}_\tau | \mathbf{X}_\rho}(s) \, \mathrm{d}s \, .$$

Conversely, the situation in which the signal is present $\mathbf{X} = \mathbf{X}_\rho \cup \mathbf{X}_T$, and the score exceeds the threshold can be represented by a detection probability

$$P_{\mathrm{det}}(\gamma_s) = \int_{\gamma_s}^{\infty} p_{\mathbf{s}_\tau | \mathbf{X}_\rho \cup \mathbf{X}_T}(s) \, \mathrm{d}s \, .$$

In general, $P_{\mathrm{det}}(\gamma_s)$ falls short of unity; its compliment $1 - P_{\mathrm{det}}(\gamma_s)$ represents the probability that the signal was not detected ($\mathbf{s}_\tau < \gamma_s$), given that it was actually present.

The parameter γ_s should be chosen to make $P_{\text{det}}(\gamma_s)$ as close as possible to unity, while keeping $P_{\text{fa}}(\gamma_s)$ as small as possible. However, since both the detection and false-alarm probabilities are non-decreasing functions of the threshold γ_s, one must must generally make a compromise between the two. By allowing γ_s to be a free parameter, one can plot the false alarm probability against the detection probability. Consistent with popular terminology, such a plot is known as a "Receiver Operating Characteristic" (ROC) [7] and will become a central theme in the remainder of this memo.

These concepts are illustrated in the following sections.

3 Performance analysis of $q_{\mathcal{N}}$-score

In this section, the $q_{\mathcal{N}}$-score introduced in [5] is evaluated to determine its ability to discriminate between the presence and absence of a signal of interest.

Consistent with the general framework described above, the $q_{\mathcal{N}}$-scoring functional makes use of a pulse template T, and some time offset parameter τ; both are taken to match the corresponding values in the signal model (3). The basis of this score can be understood by noting that the likelihood ratio for the binary decision between the two models (2) and (3) can be computed exactly, as discussed in Annex A. Over a long observation interval, the likelihood ratio reduces to an optimal statistic given in (A.3). This statistic is of the form

$$\mathbf{\Lambda} \propto \sum_{K \in \mathbb{K}(T,\mathbf{X})} \exp\Big(- \sum_{(t,x)\in K} \left[\frac{(t+\tau-x)^2}{2\sigma^2} + c_0 \right] \Big) \,,$$

where

$$c_0 \equiv \ln\left(\frac{(1-P_{\text{det}})\sigma\rho}{P_{\text{det}}} \right) + \frac{1}{2}\ln(2\pi) \,.$$

Each object $K \in \mathbb{K}(T,\mathbf{X})$ represents an "association relation" between the sets $\mathbf{X}$ and T. The class of association relations between two sets is defined to be the collection of all relations between the sets such that no element of either set is associated more than once. More precisely, for any two sets T and X, we can define

$$\mathbb{K}(T,X) \equiv \left\{ \begin{array}{c} K \,|\, K \subseteq T \times X \,, \\ (t,x),(t',x') \in K \Longrightarrow [t=t', x=x'] \text{ or } [t \neq t', x \neq x'] \end{array} \right\} \,. \tag{6}$$

An association relation need not include every element of either set.

The presence of a sum over associations $K \in \mathbb{K}(T,X)$ makes $\mathbf{\Lambda}$ a computationally expensive scoring statistic. A more practical score could exploit a truncation of the summation. The score proposed in [5] simply replaces the summation by a maximization. The $\mathbf{q}_{\mathcal{N}}$-score is thus defined as

$$\ln \mathbf{q}_{\mathcal{N}} \equiv \ln q_{\mathcal{N}}(\mathbf{X}|T, \hat{\sigma}^2, \hat{c}_0; \tau) = -\min_{K \in \mathbb{K}(T,\mathbf{X})} \sum_{(t,x)\in K} c(t,x,\tau) \tag{7}$$

where

$$c(t,x,\tau) \equiv \frac{(t+\tau-x)^2}{2\hat{\sigma}^2} + \hat{c}_0 \,.$$

Here, the score depends on two parameters, $\hat{\sigma}^2$ and $\hat{c}_0$ which, for the sake of generality, are not necessarily taken to coincide with the parameters σ^2 and c_0 of the underlying statistical model. The effect of parameter mismatch is a problem of some interest for practical implementations.

The form of (7) unnecessarily general. Since the empty set is a valid association relation, $\emptyset \in \mathbb{K}(T, X)$, the $q_{\mathcal{N}}$-score in (7) is always non-negative and may exceed zero only if $\hat{c}_0 < 0$. Moreover, the combined dependence of the score on the two parameters has an obvious scale symmetry so that

$$\ln q_{\mathcal{N}}(\mathbf{X}|T, \eta^{-2}\hat{\sigma}^2, \eta^2\hat{c}_0; \tau) = \eta^2 \ln q_{\mathcal{N}}(\mathbf{X}|T, \hat{\sigma}^2, \hat{c}_0; \tau)$$

for any $\eta^2 > 0$. Since such a scaling has no impact on the resulting ROC, it is convenient fix the scale symmetry and consider only the case in which

$$\hat{c}_0 = -1 \,.$$

This parameter value is used implicitly hereafter.

The $q_{\mathcal{N}}$-score may be treated as a particular instance of the generic score (1). The distribution for the random variable $\ln \mathbf{q}_{\mathcal{N}}$, conditioned on either model (2) or (3) can be computed using the following reasoning.

For any particular template element $t \in T$, it is not difficult to see that the minimizing relation K in (7) may include an association involving t only if there is at least one element of $\mathbf{X}$ that lies in the interval

$$I_t \equiv (t+\tau-\sqrt{2}\hat{\sigma}, t+\tau+\sqrt{2}\hat{\sigma}) \,.$$

If the minimum time interval between template elements

$$\Delta_T \equiv \min\left\{|t_1 - t_2| \middle| t_1 \in T, t_2 \in T, t_1 \neq t_2\right\} \,, \tag{8}$$

is such that $\Delta_T > 2\sqrt{2}\hat{\sigma}$, then the intervals $\{I_t | t \in T\}$ must be pairwise disjoint making it impossible for any element of $\mathbf{X}$ to appear in two different intervals I_{t_1} and

I_{t_2} for $t_1 \neq t_2$. In this case, the scoring statistic (7) can be separated into a sum of the form

$$\ln \mathbf{q}_{\mathcal{N}} = -\sum_{t \in T} \mathbf{C}_t \tag{9a}$$

where

$$\mathbf{C}_t = \min \left\{ \frac{\mathbf{\Delta}_t^2}{2\hat{\sigma}^2} - 1, 0 \right\} \tag{9b}$$

and

$$\mathbf{\Delta}_t \equiv \min \left\{ |t + \tau - x| \middle| x \in \mathbf{X} \right\} .$$

The variable $\mathbf{\Delta}_t$ is the absolute duration of the time interval between $t + \tau$ and the nearest pulse time in $\mathbf{X}$. If there are no observed pulses inside of the interval I_t then it is not beneficial to associate t with any pulse in $\mathbf{X}$ so that $\mathbf{C}_t$ is taken to vanish. As a result, $\mathbf{C}_t$ depends only on the sub-process $\mathbf{X} \cap I_t$.

The distribution for the $q_{\mathcal{N}}$-score depends on the distributions for $\mathbf{C}_t$ and ultimately on the distribution for $\mathbf{\Delta}_t$. It will be shown in the following subsections that, using either model (2) or (3), $\mathbf{\Delta}_t$ is a continuous non-negative random variable whose distribution can be described by a probability density function, denoted by $p_{\mathbf{\Delta}_t}(\delta)$ for $\delta \geq 0$. Then, applying elementary probability theory to (9b), one finds that the distribution for $\mathbf{C}_t$ is described by a continuous probability density function

$$p_{\mathbf{C}_t}(c) = \frac{\hat{\sigma}}{\sqrt{2}\sqrt{c+1}} p_{\mathbf{\Delta}_t} \left(\sqrt{2}\hat{\sigma}\sqrt{c+1} \right) \quad \text{for } -1 < c < 0 \tag{10a}$$

and a finite probability that $\mathbf{C}_t$ vanishes

$$P(\mathbf{C}_t = 0) = \int_{\sqrt{2}\hat{\sigma}}^{\infty} p_{\mathbf{\Delta}_t}(\delta) \, \mathrm{d}\delta \,. \tag{10b}$$

The distribution for $\mathbf{C}_t \in [-1, 0]$ is completely described by the two equations in (10).

Under many conditions, the sub-processes $\{\mathbf{X} \cap I_t | t \in T\}$ are mutually independent so that the the set $\{\mathbf{C}_t | t \in T\}$ represents a finite collection of independently distributed random variables. If this is the case, then the distribution for the sum in (9a) can be computed by a $|T|$-fold self-convolution of (10) where $|T|$ represents the number of template elements in the set T. The continuous component can be written

$$p_{\ln \mathbf{q}_{\mathcal{N}}}(c) = \sum_{n=0}^{|T|} \binom{|T|}{n} \left[P(\mathbf{C}_t = 0) \right]^{|T|-n} p_{\mathbf{C}_t}^{*n}(-c) \,, \quad \text{for } 0 < c \leq |T| \,, \tag{11a}$$

where $p_{\mathbf{C}_t}^{*n}$ refers to the convolution of n instances of (10a). The probability that $\ln \mathbf{q}_{\mathcal{N}}$ vanishes is

$$P(\ln \mathbf{q}_{\mathcal{N}} = 0) = \left[P(\mathbf{C}_t = 0) \right]^{|T|} \,. \tag{11b}$$

In the following two subsections, this distribution will be calculated respectively for the cases in which $\mathbf{X}$ consists of a noise-only process and a signal-in-noise process.

3.1 Noise-only model

In this section, it is of interest to compute some of the preceding distributions for the case in which $\mathbf{X}$ follows the stationary Poisson model in (2) with expected pulse density ρ.

Let $\mathbf{\Delta}_{t;\rho}$ represent the absolute time duration between $t+\tau$ and the nearest pulse in $\mathbf{X}_\rho$,

$$\mathbf{\Delta}_{t;\rho} \equiv \min\left\{|t+\tau-x|\big|x \in \mathbf{X}_\rho\right\} \,. \tag{12}$$

In the case of the noise-only model (2), there is no distinction between $\mathbf{X}$ and $\mathbf{X}_\rho$ so that $\mathbf{\Delta}_t = \mathbf{\Delta}_{t;\rho}$. As discussed in Appendix B.2.1, the distribution for this interval is described by the probability density function

$$p_{\mathbf{\Delta}_t|\mathbf{X}_\rho}(\delta) = p_{\mathbf{\Delta}_{t;\rho}}(\delta) = 2\rho e^{-2\delta\rho} \text{ for } \delta \geq 0\,. \tag{13}$$

Substituting (13) into (10), one finds that the distribution for $\mathbf{C}_t$ is described by a continuous probability density function

$$p_{\mathbf{C}_t|\mathbf{X}_\rho}(c) = \frac{\sqrt{2}\hat{\sigma}\rho}{\sqrt{c+1}} e^{-2\sqrt{2}\rho\hat{\sigma}\sqrt{c+1}} \text{ for } 0 > c > -1 \tag{14a}$$

along with a finite probability that $\mathbf{C}_t = 0$,

$$P(\mathbf{C}_t = 0|\mathbf{X}_\rho) = e^{-2\sqrt{2}\rho\hat{\sigma}}\,. \tag{14b}$$

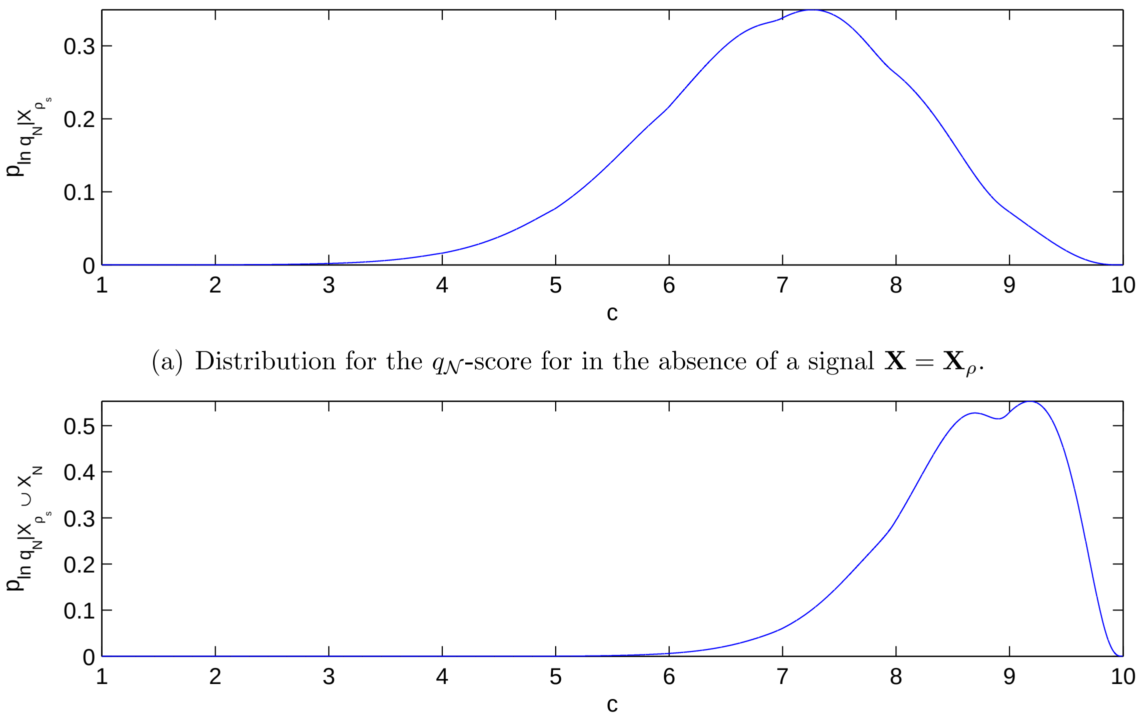
If the elements of the pulse template T satisfy the separation criterion $\Delta_T > 2\sqrt{2}\hat{\sigma}$ then the variables $\{\mathbf{C}_t|t \in T\}$ must be independent and the distribution for the logarithm of the $q_\mathcal{N}$-score can be computed using (11). However, due to the algebraic structure of (14), the continuous convolutions $p^{*n}_{\mathbf{C}_t|\mathbf{X}_\rho}(-c)$ are challenging to compute exactly and one must rely on some other technique or numerical approximation.

An illustration of the continuous distribution of the logarithmic $q_\mathcal{N}$-score (11a), conditioned on this noise model is shown in Figure 1(a). The convolutions that lead to this distribution have been calculated using numerical integration. The distribution is plotted for score values in the interval from 1 to $|T|$. (The distribution of scores in the interval $c \in [0,1]$ has not been illustrated but note that (10a) indicates the presence of a singularity as $c \to 1^-$.)

3.2 Signal-in-noise model

The preceding analysis can be contrasted against the distribution for the $q_\mathcal{N}$-score realized using the statistical process (3). When a signal is present, the minimum absolute time parameter is given by

$$\mathbf{\Delta}_t = \begin{cases} \min\{\mathbf{\Delta}_{t;\rho}, |t+\tau-\mathbf{x}_t|\} & \text{if the pulse } \mathbf{x}_t \text{ is detected}\,. \\ \mathbf{\Delta}_{t;\rho} & \text{if the pulse was undetected}\,. \end{cases}$$



(a) Distribution for the $q_{\mathcal{N}}$-score for in the absence of a signal $\mathbf{X} = \mathbf{X}_\rho$.

(b) Distribution for the $q_{\mathcal{N}}$-score in the presence of a signal $\mathbf{X} = \mathbf{X}_\rho \cup \mathbf{X}_T$.

Figure 1: *Distribution for the $q_{\mathcal{N}}$-score using $\hat{\sigma} = 17.4\mu\text{s}$ with noise parameter $\rho = 0.04/\mu\text{s}$ and signal parameters $\sigma = 12\mu\text{s}$, $P_\text{p} = 0.8$, $|T| = 10$.*

where $\mathbf{\Delta}_{t;\rho}$ given in (12). Note that it has been assumed here that Δ_T is large enough so that one may neglect the remote possibility that the closest observed pulse to $t+\tau$ is actually a detected pulse $\mathbf{x}_{t'}$ arising from some other template element $t' \neq t$. The distribution for $\mathbf{\Delta}_t \geq 0$ can be computed by noting that the probability of detecting an individual pulse is P_p. The distributions (4) and (13) can be combined with the identity (B.1), resulting in the probability density function

$$p_{\mathbf{\Delta}_t|\mathbf{X}_\rho \cup \mathbf{X}_T}(\delta) = \left(2\rho e^{-2\delta\rho}\left[1 - \text{erf}(\delta/\sqrt{2}\sigma)\right] + \frac{2}{\sqrt{2\pi\sigma^2}} e^{-\delta^2/2\sigma^2} e^{-2\delta\rho}\right) P_\text{p} + \left(2\rho e^{-2\rho\delta}\right)(1 - P_\text{p}) \quad (15)$$

for $\delta \geq 0$. Here, the error function follows the convention $\text{erf}(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} \text{d}u$. This result, in turn, can be combined with (10) in order to evaluate a distribution function for the cost variables $\mathbf{C}_t$. The discrete component of this distribution is given by

$$P(\mathbf{C}_t = 0|\mathbf{X}_\rho \cup \mathbf{X}_T) = [1 - \text{erf}(\hat{\sigma}/\sigma)]\, e^{-2\sqrt{2}\rho\hat{\sigma}} P_\text{p} + e^{-2\sqrt{2}\rho\hat{\sigma}}(1 - P_\text{p})\,.$$

The continuous component of this distribution $p_{\mathbf{C}_t|\mathbf{X}_\rho\cup\mathbf{X}_T}(c)$ is algebraically complicated for $c \in (-1, 0)$ and need not be reproduced here. If $\Delta_T > 2\sqrt{2}\hat{\sigma}$ then the distribution for $\ln \mathbf{q}_\mathcal{N}$ can be calculated using (11) but, once again, a numerical or statistical approximation is required to compute the continuous convolutions. From a numerical point of view, (15) presents no more of a challenge than does the corresponding noise-only distribution (13).

A distribution for a $q_\mathcal{N}$-score in the presence of a signal with noise is illustration in Figure 1(b). The seemingly multi-modal behaviour of the probability density in Figure 1(b) can traced back to the algebraic structure of (11). For sufficiently large values of n, the distributions $p^{*n}_{\mathbf{C}_t}(c)$ are expected to be very close to Gaussian, as dictated by the central limit theorem. The linear combination of such distributions exhibits some local maxima near to the peaks of the constituent distributions.

The noise and scoring parameters in Figures 1(a) and 1(b) have been chosen to agree so that the two distributions can be compared directly.

3.3 Measures of performance

By choosing a particular threshold value of the score, denoted by $\gamma_{\ln q_\mathcal{N}}$, one can define decision regions for the presence or absence of the signal;

$$\ln \mathbf{q}_\mathcal{N} \gtrless \gamma_{\ln q_\mathcal{N}} .$$

For each value of the threshold $\gamma_{\ln q_\mathcal{N}}$, one can compute a detection and false alarm probability. A Receiver Operating Characteristic (ROC) is obtained by plotting the detection probability against the false alarm probability for various values of the threshold.

Some representative ROCs for the $q_\mathcal{N}$-score have been plotted in Figure 2 where the values of the the noise and signal parameters have been chosen to match those of Figure 1. The illustration shows various values of $\hat{\sigma}$ in uniform multiplicative increments.

4 Performance analysis of cross-correlation histograms

In this section, the performance characteristics of the cross-correlation histogram [4] are evaluated.

Here, one selects a particular value of the histogram width parameter Δ and then partitions the time offset values of τ into intervals of the form $((j-1)\Delta, j\Delta]$ for integer

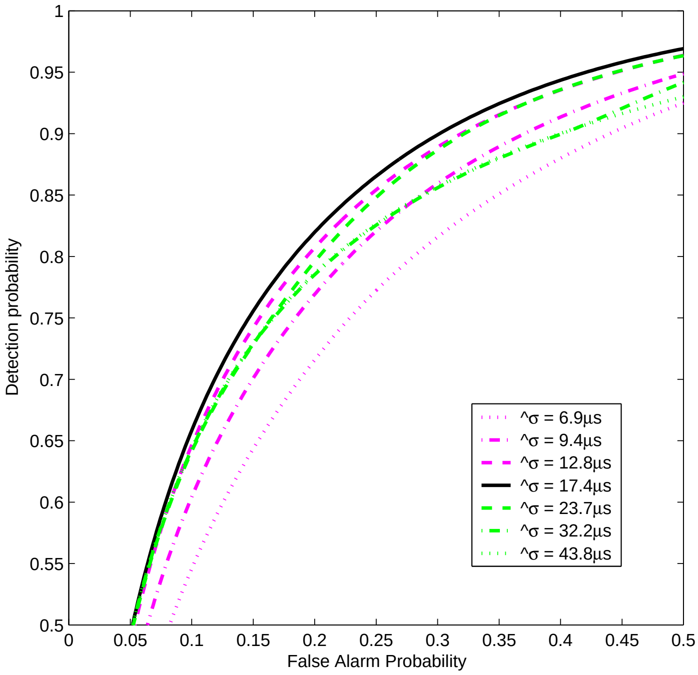


Figure 2: *ROC for the $q_{\mathcal{N}}$-score with noise parameter $\rho = 0.04/\mu\text{s}$ and signal parameters $\sigma = 12\mu\text{s}$, $P_{\text{p}} = 0.8$, $|T| = 10$.*

values of the index j. For each such interval, one counts the number of distinct pairs (t, x) such that $x \in \mathbf{X}$, $t \in T$ and $x - t \in ((j-1)\Delta, j\Delta]$. The set of such pairs may be denoted by

$$R_{T,\mathbf{X},\Delta;j} \equiv \{(t, x)|t \in T, x \in \mathbf{X}, (j-1)\Delta < x - t \leq j\Delta\} \ . \tag{16}$$

For each value of τ, let $j_\tau \equiv \lceil \tau/\Delta \rceil$ be the associated partition index so that the histogram scoring functional, analogous to (1), is given by

$$n(\mathbf{X}, T, \Delta; \tau) \equiv |R_{T,\mathbf{X},\Delta;j_\tau}|$$

where $|R_{T,\mathbf{X},\Delta;j_\tau}|$ is number of pairs in $R_{T,\mathbf{X},\Delta;j_\tau}$. Now, assuming that the minimum template separation (8) satisfies $\Delta_T > \Delta$, the value of this function may be expressed as

$$\mathbf{n} \equiv n(\mathbf{X}, T, \Delta; \tau) = \sum_{t \in T} \mathbf{n}_t \tag{17}$$

where, for each $t \in T$, the random integer-valued variable $\mathbf{n}_t$ represents the number of pulses in $\mathbf{X}$ that lie in the interval $(j_\tau - 1)\Delta + t < x \leq j_\tau\Delta + t$.

4.1 Noise-only model

For each $t \in T$, let $\mathbf{n}_{t;\rho}$ be the number of spurious pulses $x \in \mathbf{X}_\rho$ that lie in the interval $(j_\tau - 1)\Delta + t < x \le j_\tau\Delta + t$. Since $\mathbf{X}_\rho$ is a Poisson process, the number of elements of $\mathbf{X}$ lying in any interval of length Δ is simply a Poisson variable whose expected value is $\rho\Delta$. Following (B.2), the distribution for $\mathbf{n}_{t;\rho}$ must be

$$P(\mathbf{n}_{t;\rho} = n) = e^{-\rho\Delta}\frac{(\rho\Delta)^n}{n!} \text{ for } n \ge 0\,.$$

In the case of the noise-only model (2), it must be that $\mathbf{n}_t = \mathbf{n}_{t;\rho}$. Moreover, if $\Delta_T > \Delta$ then the intervals do not overlap, and the set $\{\mathbf{n}_t | t \in T\}$ is a collection of $|T|$ independent, and identically distributed random variables. Consequently, $\mathbf{n}$ is a Poisson random variable with expected value $|T|\rho\Delta$

$$P(\mathbf{n} = n|\mathbf{X}_\rho) = P\Big(\sum_t \mathbf{n}_{t;\rho} = n\Big) = e^{-|T|\rho\Delta}\frac{(|T|\rho\Delta)^n}{n!} \text{ for } n \ge 0\,. \tag{18}$$

Illustrations of such a distribution are shown in Figures 3(a) and 4(a) for a finite set of discrete values n. These two figures show two different values of the scoring parameter Δ for reasons that should become apparent in the following sections.

4.2 Signal-in-noise model

In the case of model (3), the value of the histogram score function may be written

$$\mathbf{n} \equiv n(\mathbf{X}, T, \Delta; \tau) = \sum_{t \in T} (\mathbf{n}_{t;\rho} + \mathbf{n}_{t;T})\ . \tag{19}$$

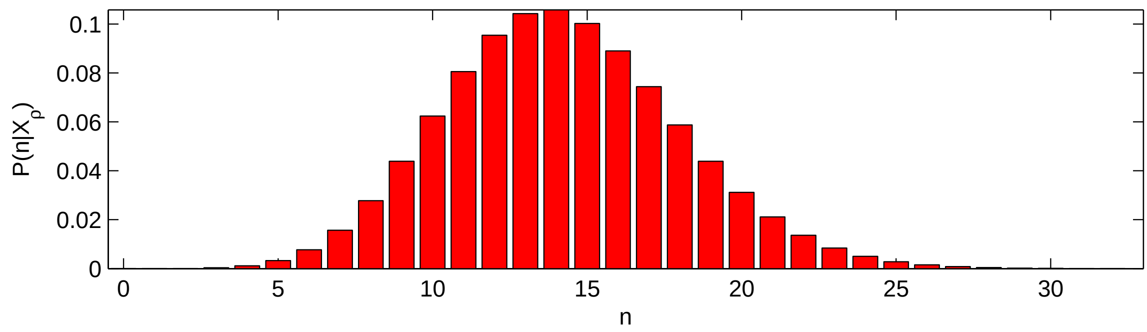
Here the noise contribution $\sum_t \mathbf{n}_{t;\rho}$ is distributed according to (18) and the signal contribution is represented by independent variables $\mathbf{n}_{t;T}$

$$\mathbf{n}_{t;T} = \begin{cases} 1 & \text{if the pulse } \mathbf{x}_t \text{ is detected and } (j_\tau - 1)\Delta < \mathbf{x}_t - t \le j_\tau\Delta, \\ 0 & \text{otherwise.} \end{cases}$$

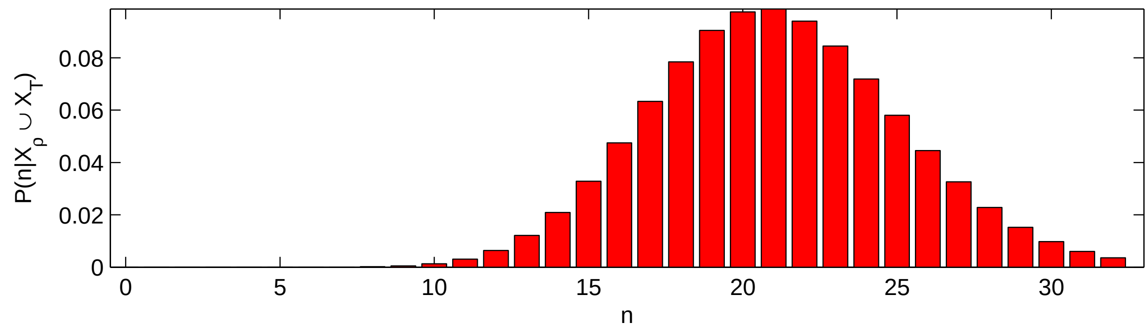
It has been assumed here that the τ value used to evaluate (19) corresponds with that in the signal model (3) and that $\Delta_T \gg \sigma$. In this case, one may safely neglect the unlikely possibility that $(j_\tau - 1)\Delta < \mathbf{x}_{t'} - t \le j_\tau\Delta$ for any other template element $t' \ne t$.

Following the Gaussian model (4) the probability that $\mathbf{x}_t$ is detected *and* falls into the interval $(j_\tau - 1)\Delta + t < \mathbf{x}_t \le j_\tau\Delta + t$ is

$$\tilde{P}_{\rm p} = P_{\rm p}\frac{1}{\sqrt{2\pi\sigma^2}}\int_{t+(j_\tau-1)\Delta}^{t+j_\tau\Delta} e^{-(t+\tau-x)^2/2\sigma^2}\,{\rm d}x\,.$$



(a) Distribution for the histogram score in the absence of a signal $\mathbf{X} = \mathbf{X}_\rho$.



(b) Distribution for the histogram score in the presence of a signal $\mathbf{X} = \mathbf{X}_\rho \cup \mathbf{X}_T$.

Figure 3: *Distribution for the histogram score using* $\Delta = 35.5\mu\text{s}$ *with noise parameter* $\rho = 0.04/\mu\text{s}$*, signal parameters* $\sigma = 12\mu\text{s}$, $P_\text{p} = 0.8$, $|T| = 10$ *and phase* $u = 1/2$.

Since $j_\tau \equiv \lceil \tau/\Delta \rceil$, the time-offset must be of the form $\tau = (j_\tau - u)\Delta$ where $u \in [0, 1)$ is some parameter representing the phase of the signal with respect to the histogram partition and j_τ is an integer. The preceding probability may then be viewed as a function of u,

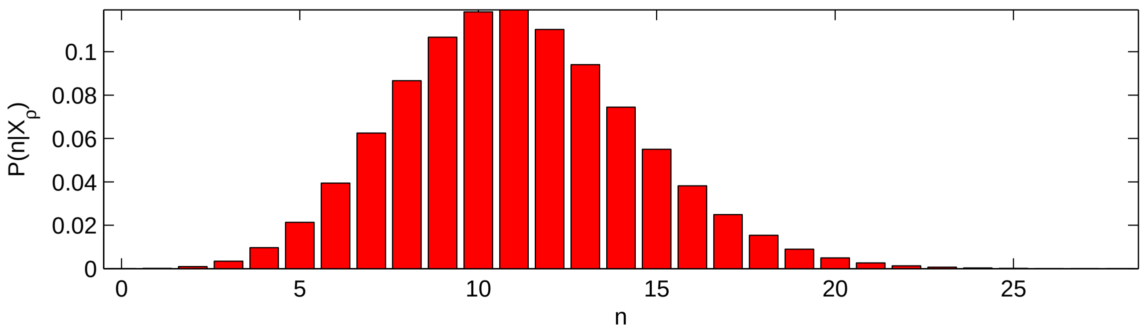
$$\tilde{P}_\text{p}(u) = \frac{P_\text{p}}{2}\left[\text{erf}\Big(\frac{\Delta}{\sqrt{2}\sigma}(1-u)\Big) + \text{erf}\Big(\frac{\Delta}{\sqrt{2}\sigma}u\Big)\right] .$$

For any particular value of the phase u, the conditional signal contribution to (19) follows a binomial distribution

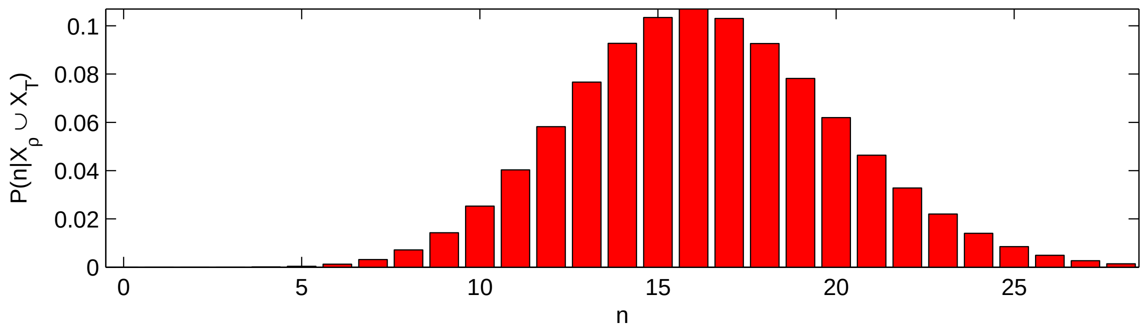
$$P\Big(\sum_t \mathbf{n}_{t;T} = n \big| u\Big) = \binom{|T|}{n}\big(\tilde{P}_\text{p}(u)\big)^n\big(1 - \tilde{P}_\text{p}(u)\big)^{|T|-n} \tag{20a}$$

for $0 \leq n \leq |T|$. Since the starting phase between the receiver and transmitter are essentially random on small time scales, the phase should be treated as a random variable distributed uniformly as $u \sim \mathcal{U}[0, 1)$. If this is the case then the distribution for the signal contribution should properly be given by

$$P\Big(\sum_t \mathbf{n}_{t;T} = n\Big) = \int_0^1 P\Big(\sum_t \mathbf{n}_{t;T} = n \big| u\Big)\, \text{d}u\,. \tag{20b}$$



(a) Distribution for the histogram score in the absence of a signal $\mathbf{X} = \mathbf{X}_\rho$.



(b) Distribution for the histogram score in the presence of a signal $\mathbf{X} = \mathbf{X}_\rho \cup \mathbf{X}_T$ with $u \sim \mathcal{U}[0, 1)$.

Figure 4: *Distributions for the histogram score using* $\Delta = 27.7\mu\text{s}$ *with noise parameter* $\rho = 0.04/\mu\text{s}$*, signal parameters* $\sigma = 12\mu\text{s}$*,* $P_\text{p} = 0.8$*,* $|T| = 10$ *and phase* $u \sim \mathcal{U}[0, 1)$*.*

For a particular set of parameters, the latter distribution can be computed using straightforward numerical integration techniques.

A shortcoming of the raw histogram is the conditional dependence on the phase of the signal of with respect to the histogram partition. The probability $\tilde{P}_\text{p}(u)$ is achieves a maximum value at $u = 1/2$ and exhibits a performance degradation for other values of u. The author expects that it is possible to mitigate the effects of an unfortunate phase value u by designing a more complex signal processing technique such as some form of digital filter applied to histograms with smaller values of Δ. Since such techniques would constitute a complete study on their own, they will not be addressed here. Instead, the present memo will illustrate separate performances of the histogram based firstly on the model (20a) with the most favourable phase value $u = 1/2$, and secondly based on (20b) where $u \sim \mathcal{U}[0, 1)$.

Since the signal and noise are independent processes, the distribution for (19) can be obtained by convolving the distribution (18) with either (20a) or (20b).

An example of the $u = 1/2$ scoring distribution is shown in Figure 3(b) and an example of the $u \sim \mathcal{U}[0, 1)$ scoring distribution has appears in Figure 4(b). In either case, score and noise parameters have been chosen to match those of Figures 3(a) and 4(a), respectively, to facilitate direct comparison between the noise-only and signal-in-noise cases.

4.3 Measures of performance

By choosing a particular threshold value of of the score, denoted by γ_n, one can define decision regions for the presence or absence of the signal;

$$\mathbf{n} \gtrless \gamma_n \,.$$

Note that, since the histogram score $\mathbf{n}$ is integer-valued, it is sufficient to consider only integer values of the threshold $\gamma_n \in \mathbb{Z}$. Any outcome with $\mathbf{n} \geq \gamma_n$ is taken to indicate a decision in favour of the presence of the signal while $\mathbf{n} < \gamma_n$ results in a decision for its absence. With this convention, the detection and false-alarm probabilities are given by the discrete sums

$$\begin{aligned} P_{\mathrm{det}}(\gamma_n) &= \sum_{n=\gamma_n}^{\infty} P(\mathbf{n} = n | \mathbf{X}_\rho \cup \mathbf{X}_T) \,, \\ P_{\mathrm{fa}}(\gamma_n) &= \sum_{n=\gamma_n}^{\infty} P(\mathbf{n} = n | \mathbf{X}_\rho) \,. \end{aligned}$$

By plotting the false alarm probability against the detection probability, one obtains a ROC consisting of a set of discrete points.

Figure 5 illustrates characteristic curves for the histogram score using the best-case phase $u = 1/2$ with particular choices for the parameters ρ, σ, P_{p} and $|T|$. Figure 6 shows the same plot in which the phase is taken to be a uniformly distributed variable $u \sim \mathcal{U}[0, 1)$. In order to facilitate comparison between the two scoring mechanisms, the three Figures 2, 5, and 6 show the a common reference curve corresponding to the best $q_{\mathcal{N}}$-score ROC with the same parameter set. Comparing the latter two figures, it is clear that the uniform phase distribution has had a significant negative impact on the performance of the histogram.

Some discrete values of the parameter Δ admit thresholds $\gamma_n \in \mathbb{Z}$ for which the false-alarm and detection probabilities satisfy $P_{\mathrm{fa}} = 1 - P_{\mathrm{det}}$. This constraint has no fundamental significance, except that it reduces the ROC shown in Figures 5 and 6 to one-dimensional plots of the false-alarm probability P_{fa} versus Δ, shown in Figure 7.

5 Examples and discussion

Figures 1-7 illustrate some aspects of the statistical behaviour of the two scoring mechanisms using the signal and noise parameters $\rho = 0.04/\mu\mathrm{s}$, $\sigma = 12\mu\mathrm{s}$, $P_{\mathrm{p}} = 0.8$

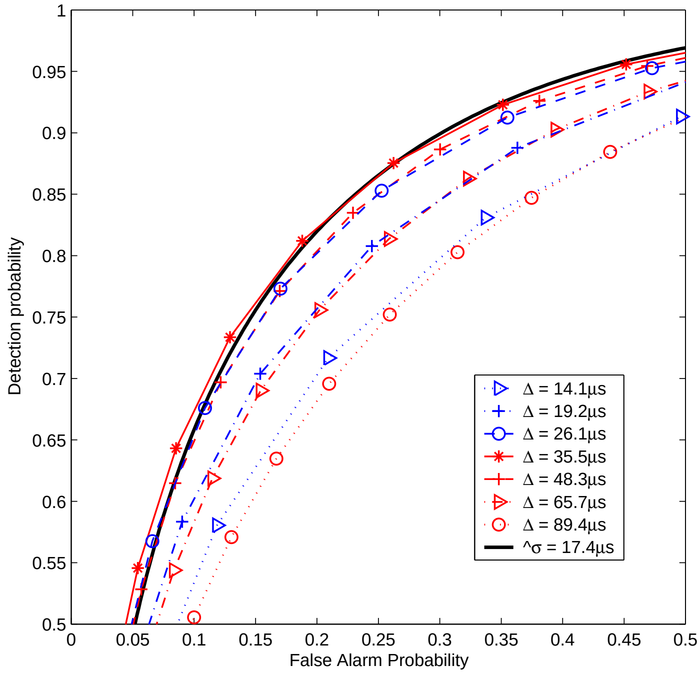


Figure 5: *ROC for the histogram score using noise parameter* $\rho = 0.04/\mu\mathrm{s}$*, signal parameters* $\sigma = 12\mu\mathrm{s}$*,* $P_{\mathrm{p}} = 0.8$*,* $|T| = 10$ *and phase* $u = 1/2$*.*

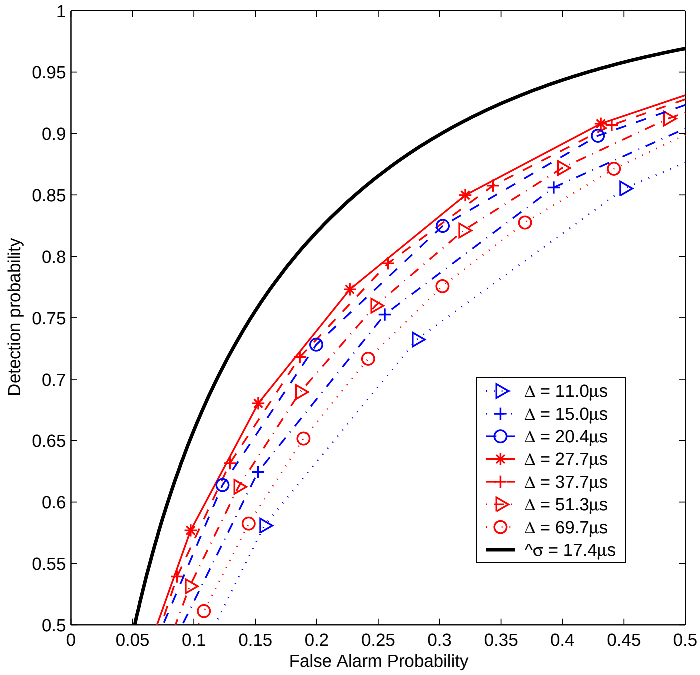


Figure 6: *ROC for the histogram score using noise parameter* $\rho = 0.04/\mu\mathrm{s}$*, signal parameters* $\sigma = 12\mu\mathrm{s}$*,* $P_{\mathrm{p}} = 0.8$*,* $|T| = 10$ *and phase* $u \sim \mathcal{U}[0, 1)$*.*

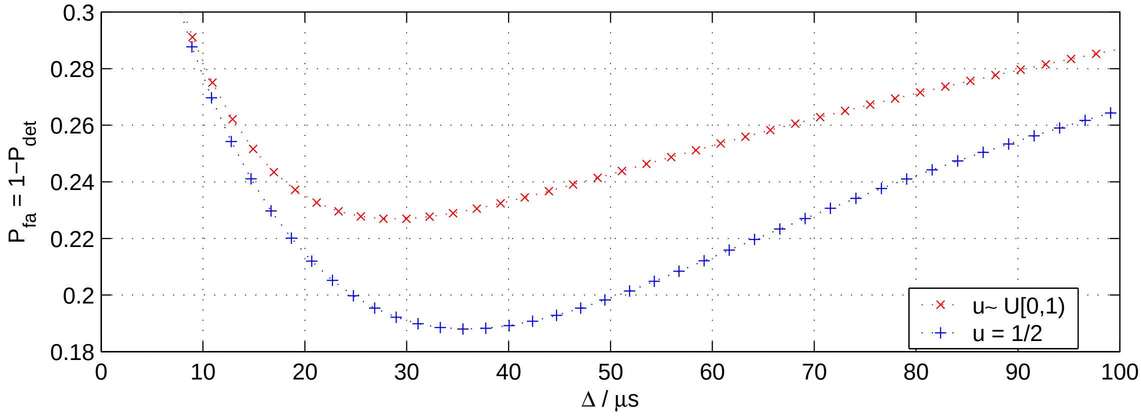


Figure 7: *Constrained probability $P_{\text{fa}} = 1 - P_{\text{det}}$ shown as a function of Δ using noise parameter $\rho = 0.04/\mu\text{s}$, signal parameters $\sigma = 12\mu\text{s}$, $P_{\text{p}} = 0.8$, $|T| = 10$.*

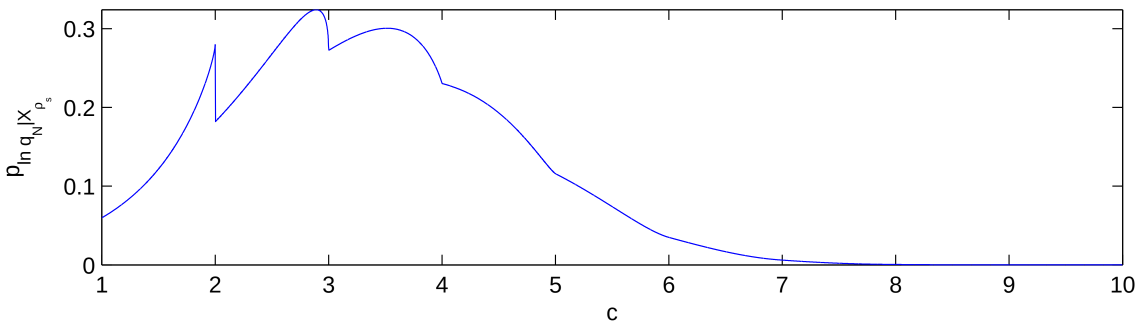
and $|T| = 10$. It can be shown that the best $q_{\mathcal{N}}$-score ROC is achieved in this case using $\hat{\sigma} \approx 17.4\mu\text{s}$. Using the $P_{\text{fa}} = 1 - P_{\text{det}}$ criterion shown in Figure 7, the best histogram score, conditional on the phase $u = 1/2$ is at $\Delta = 35.5\mu\text{s}$ and the best unconditional histogram score $u \sim \mathcal{U}[0, 1)$ occurs at $\Delta = 27.7\mu\text{s}$.

If some mechanism were available to arrange for an optimal phase value $u \approx 1/2$ then, for these parameters and a suitable choice for the bin-width Δ, the histogram score could be made to have a similar or marginally superior performance than that of the $q_{\mathcal{N}}$-score. However, in the absence of additional processing, the phase is uniformly distributed on $u \sim \mathcal{U}[0, 1)$ and the $q_{\mathcal{N}}$-score exhibits better performance.

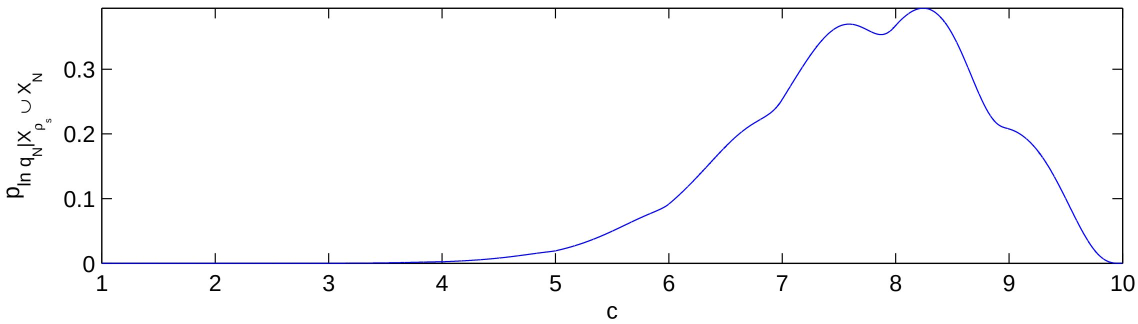
The preceding example corresponds to a fairly high pulse density and it is shown below that the situation changes significantly as ρ decreases.

Consider a second example in which $\rho = 0.01/\mu\text{s}$, $\sigma = 12\mu\text{s}$, $P_{\text{p}} = 0.8$ and $|T| = 10$. The situation is illustrated in Figures 8-14. Figures 8 and 9 show the distributions and ROC for the $q_{\mathcal{N}}$-score. Figures 10 and 11 show the properties of the histogram score in the $u = 1/2$ best-case. Figures 12 and 13 show the properties of the histogram score with a uniformly distributed phase $u \sim [0, 1)$. Figure 14 shows the performance curve subject to the $P_{\text{fa}} = 1 - P_{\text{det}}$ constraint.

In this case, the best $q_{\mathcal{N}}$-score parameter was found to be near $\hat{\sigma} \approx 21.6\mu\text{s}$. The best choice for the histogram using the $P_{\text{fa}} = 1 - P_{\text{det}}$ criterion is $\Delta \approx 40.3\mu\text{s}$ for $u = 1/2$, and $\Delta \approx 31.2\mu\text{s}$ for $u \sim \mathcal{U}[0, 1)$. Not surprisingly, both scoring techniques perform better under these conditions than they do in the higher noise density environment. However, in this case of lower noise density, it is seen that the $q_{\mathcal{N}}$-score has a generally better performance than the histogram, regardless of the model used for the phase.



(a) Probability density function for the $q_{\mathcal{N}}$-score for in the absence of a signal $\mathbf{X} = \mathbf{X}_\rho$.



(b) Probability density function for the $q_{\mathcal{N}}$-score in the presence of a signal $\mathbf{X} = \mathbf{X}_\rho \cup \mathbf{X}_{\mathcal{N}}$.

Figure 8: *Distribution for the $q_{\mathcal{N}}$-score using $\hat{\sigma} = 21.6\mu$s with noise parameter $\rho = 0.01/\mu$s and signal parameters $\sigma = 12\mu$s, $P_\mathrm{p} = 0.8$, $|T| = 10$.*

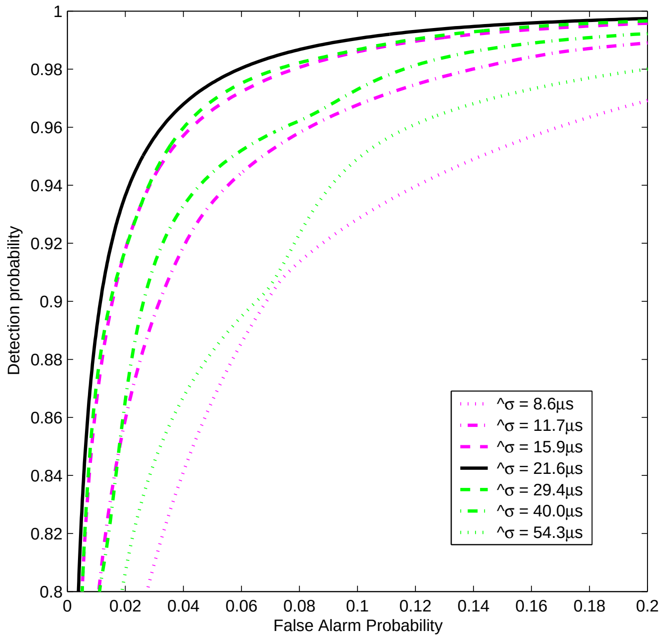
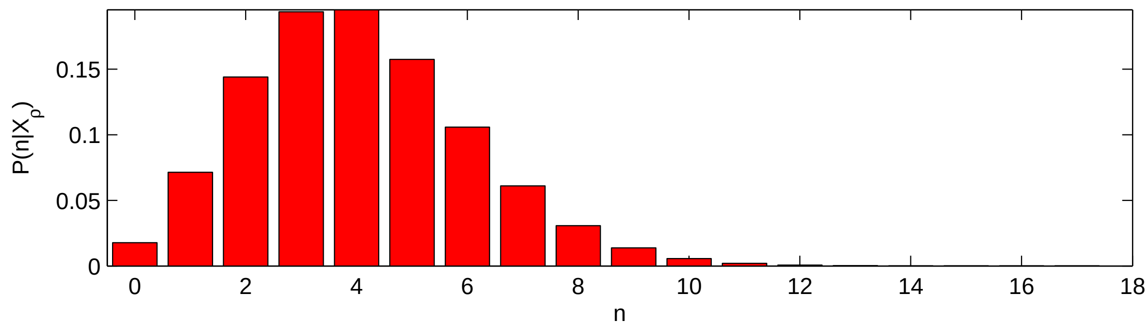
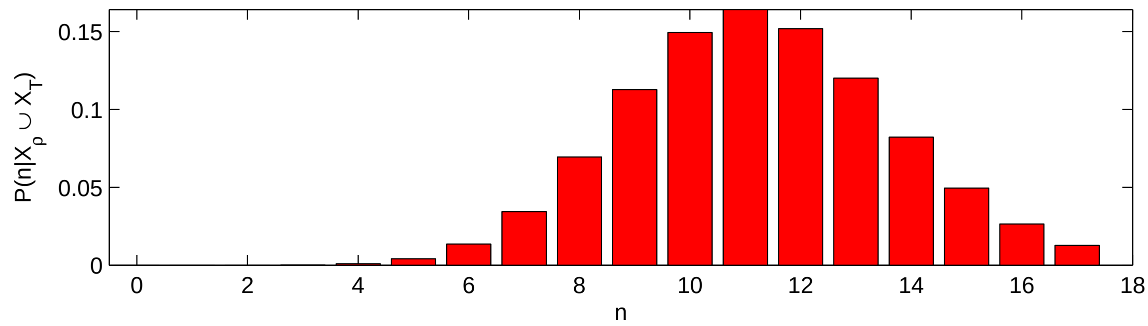


Figure 9: *ROC for the $q_{\mathcal{N}}$-score with noise parameter $\rho = 0.01/\mu$s and signal parameters $\sigma = 12\mu$s, $P_\mathrm{p} = 0.8$, $|T| = 10$.*



(a) Distribution for the histogram score in the presence of a signal $\mathbf{X} = \mathbf{X}_\rho \cup \mathbf{X}_T$.



(b) Distribution for the histogram score in the absence of a signal $\mathbf{X} = \mathbf{X}_\rho$.

Figure 10: *Distributions for the histogram scores using* $\Delta = 40.3\mu$s *with noise parameter* $\rho = 0.01/\mu$s*, signal parameters* $\sigma = 12\mu$s*,* $P_\mathrm{p} = 0.8$*,* $|T| = 10$ *and phase* $u = 1/2$*.*

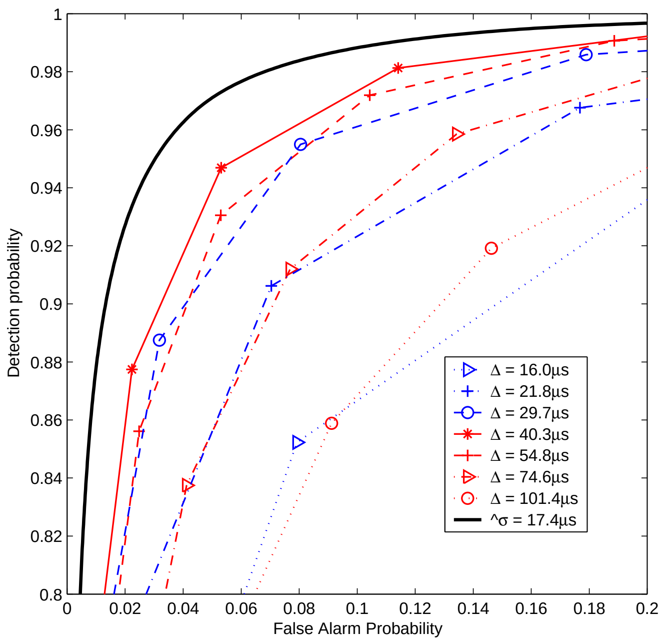
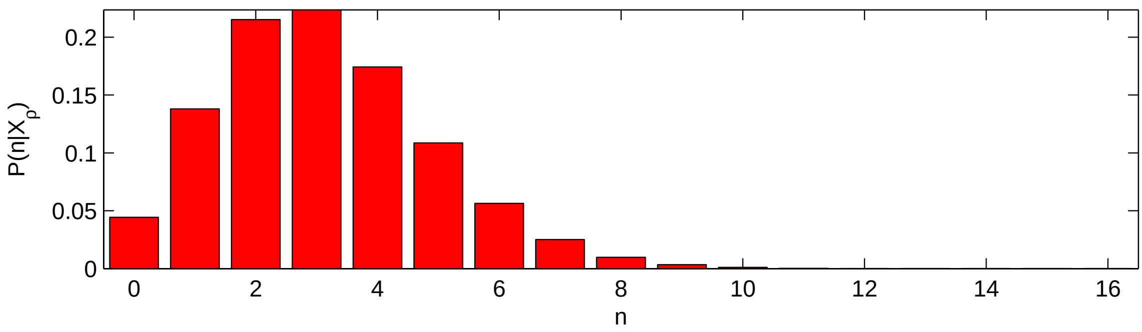
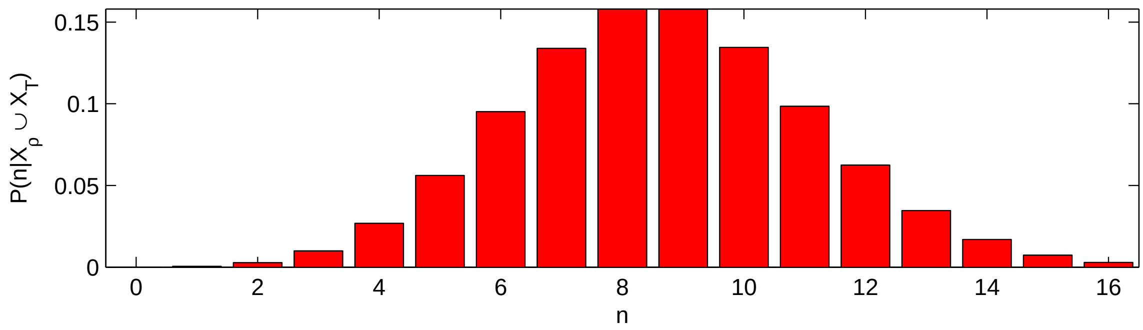


Figure 11: *ROC for the histogram scores using noise parameter* $\rho = 0.01/\mu$s *with signal parameters* $\sigma = 12\mu$s*,* $P_\mathrm{p} = 0.8$*,* $|T| = 10$ *and phase* $u = 1/2$*.*



(a) Distribution for the histogram score in the absence of a signal $\mathbf{X} = \mathbf{X}_\rho$.



(b) Distribution for the histogram score in the presence of a signal $\mathbf{X} = \mathbf{X}_\rho \cup \mathbf{X}_T$.

Figure 12: *Distribution for the histogram score using $\Delta = 31.2\mu s$ with noise parameter $\rho = 0.01/\mu s$, signal parameters $\sigma = 12\mu s$, $P_p = 0.8$, $|T| = 10$ and phase $u \sim \mathcal{U}[0,1)$.*

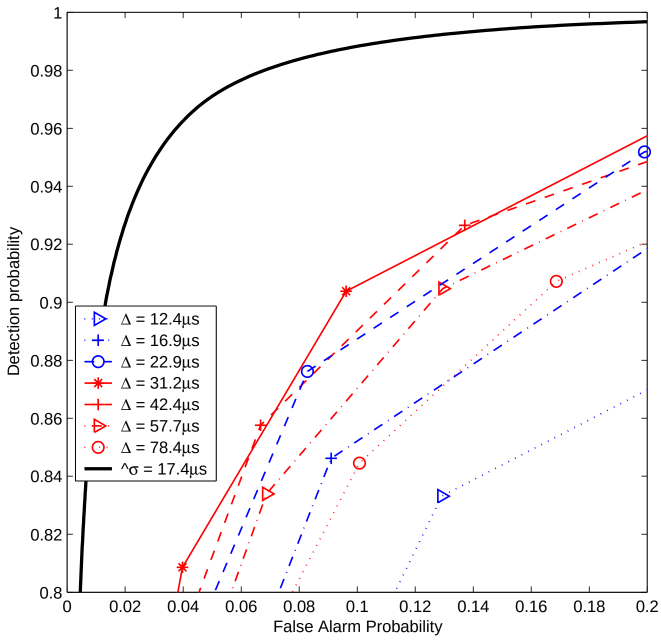


Figure 13: *ROC for the histogram score using noise parameter $\rho = 0.01/\mu s$ with signal parameters $\sigma = 12\mu s$, $P_p = 0.8$, $|T| = 10$ and phase $u \sim \mathcal{U}[0,1)$.*

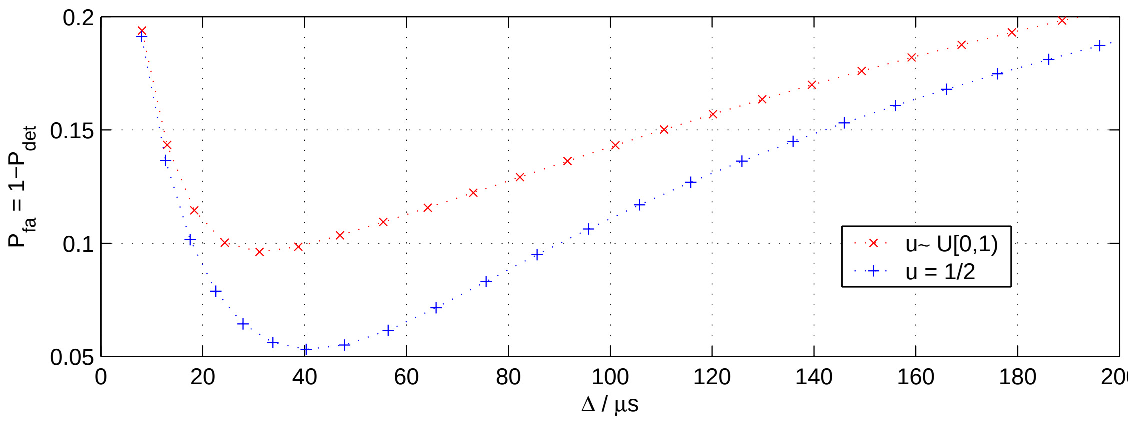


Figure 14: *Constrained probability $P_{\mathrm{fa}} = 1 - P_{\mathrm{det}}$ shown as a function of Δ using noise parameter $\rho = 0.01/\mu\mathrm{s}$, signal parameters $\sigma = 12\mu\mathrm{s}$, $P_{\mathrm{p}} = 0.8$, $|T| = 10$.*

The plots discussed above may lead one to wonder whether there exists a performance bound on the ROC that cannot be exceeded by any functional of the type (1). Such a bound is discussed in Annex A where it is argued that the optimal scoring functional is a likelihood ratio.

6 Conclusions

Cross-correlation techniques have been identified as potentially valuable tools with which to make use of *a priori* information for pulse sequence de-interleaving and signal identification [4, 5, 6]. This memo has presented some concepts and techniques with which to measure the ability of various cross-correlation techniques to discern between the presence and absence of a signal in a noisy environment.

The proposed formalism defines the concepts of detection and false alarm probabilities along with "Receiver Operator Characteristics" (ROCs) in the context of radar pulse sequence detection. These ideas are well known from decision theory [7] and have been applied in numerous applications, including Electronic Warfare [10]. It is shown how the ROC can be determined for the various scoring mechanisms by theoretical means using statistical models of the noise and signal.

These concepts help to answer some important questions related to the application and performance of detection techniques. Firstly, they help to determine an optimal threshold for any particular technique. This is an extremely important issue since one is always faced with the problem determining of whether or not the peaks of the cross-correlation profile are really high enough to declare signal detections. Secondly, they help to quantify the degree of confidence that such declarations will actually be correct. That is, they measure the probability that the cross-correlation will exceed a specified threshold crossing given that a signal is present and, conversely, the probability that the cross-correlation will not exceed the threshold, given that the signal is absent. Finally, this technique can be used to compare the performances of various candidate cross-correlation techniques and to determine under which circumstances some techniques may be more suitable than others.

The statistical models proposed to measure these techniques are ones in which the signal is modeled with a Gaussian non-cumulative jitter and a stationary binary erasure channel, while the noise is a stationary Poisson process. Under these conditions, it is found that the $q_{\mathcal{N}}$-score proposed in [5] is expected to perform better than the basic cross-correlation histogram proposed in [8]. However, for high pulse densities, it is conjectured that a digitally filtered variant of the histogram may yield similar or marginally better results under very dense noise conditions. This conjecture is left as a topic for future investigation.

Other potential areas for future consideration include the effects of finite observation windows and signal clipping, the requirement for an estimate of the noise parameter ρ, the effects of non-stationary noise $\rho = \rho(t)$, and competing signal models that exhibit similar or harmonically related temporal characteristics.

This page intentionally left blank.

Annex A: Optimal decisions and performance bounds

Motivated by the results presented above, this annex attempts to address the following question: Does there exists a fundamental performance bound that cannot be surpassed by any scoring functional of the form (1), and if so, can the bound be saturated? It seems unlikely that either of the scoring functionals evaluated in the body of this memo would lead to an optimal decision criterion.

It is argued below that an optimal scoring functional is a likelihood ratio. Several properties of the distributions associated with this score, including all moments, can be evaluated in a straightforward manner. However, the discussion falls short of computing the bound itself since an acceptable solution to evaluate the optimal distribution has yet to be be found.

This content of this annex is arranged as follows. In Sub-Annex A.1 the notion of a probability space for a point process [9] is developed and the associated measures for the noise-only and signal-in-noise models are presented. In Sub-Annex A.2, the well-known classical theory of detections and decisions [7] is briefly reviewed and applied to the point process. Sub-Annex A.3 then presents some properties that may prove useful for evaluating the distribution of the likelihood ratio score.

A.1 Probability spaces and measures of point processes

In this memo, the observed random process is a discrete set of points drawn from some space $\mathcal{X}$. Such a processes are known in the literature as "point processes" [9] and have been studied in a variety of contexts. Poisson processes, in particular, have been used extensively to model optical communication channels [11].

The distribution of outcomes in an experiment from a point process may be represented using the following definition using notation and terminology adapted from [9].

Definition 1 *A "finite point process" is described by a 3-tuple* $(\mathcal{X}, \{p_n\}, \{\Pi_n(\cdot)\})$ *with the following components. The set* $\mathcal{X}$ *is some complete separable metric space. The numbers* $\{p_n \geq 0 | n = 0, 1, \ldots\}$ *represent a discrete probability distribution for a non-negative integer variable with* $\sum_{n=0}^{\infty} p_n = 1$*. For each non-negative integer* $n \geq 0$*,* $\Pi_n(\cdot)$ *represents a probability measure on the space of Borel sets of* $\mathcal{X}^{(n)} \equiv \mathcal{X} \times \cdots \times \mathcal{X}$*. (By convention,* $\mathcal{X}^{(0)} \equiv \{\varepsilon\}$ *consists of a single point known as the "empty outcome"* ε*. Its measure is* $\Pi_0(\{\varepsilon\}) \equiv 1$*.)*

This definition has a very straightforward interpretation. The outcome of an experiment is taken to be a finite sequence of points $(x_1, \cdots, x_n) \in \mathcal{X}^{(n)}$ for some integer $n \geq 0$. The length of the sequence is a random variable whose distribution is determined by the probabilities $\{p_n | n = 0, 1, \ldots\}$. Conditional on some particular sequence length n, the distribution for the sequences themselves is determined from $\Pi_n(\cdot)$.

This construction can be stated rigorously by defining a new probability space whose sample space is taken to be union $\mathcal{X}^{\cup} \equiv \cup_{n=0}^{\infty} \mathcal{X}^{(n)}$, whose event space given naturally by its Borel algebra and whose probability measure is denoted by $\Pi(\cdot)$. The measure of any event $\omega \subseteq \mathcal{X}^{(n)}$ that consists of outcomes containing exactly n points is simply given by $\Pi(\omega) \equiv p_n \Pi_n(\omega)$.

It is then convenient to define the so-called "Janossy measure" $J_n(\cdot)$ over the Borel sets of $\mathcal{X}^{(n)}$. For $n \geq 1$, the measure of a a cartesian product of Borel sets $A_1 \times \cdots \times A_n \subseteq \mathcal{X}^{(n)}$, is defined by

$$J_n(A_1 \times \cdots \times A_n) \equiv p_n \sum_{(i_1, \cdots, i_n)} \Pi_n(A_{i_1} \times \cdots \times A_{i_n}) \, .$$

Here, the sum extends over all $n!$ permutations of the n indices $(1, \cdots, n)$. The significance of this measure can be appreciated by noting that if $A_1, \cdots, A_n$ are mutually disjoint then $J_n(A_1 \times \cdots \times A_n)$ represents the probability that the outcome of the experiment results in n points such that each of the sets $A_1, \cdots, A_n$ contains exactly one point. The sum over permutations amounts to a symmetrization of the measure and is relevant if the points in the outcome are indistinguishable. The Janossy measure on $\mathcal{X}^{(0)}$ is defined by $J_0(\{\varepsilon\}) \equiv p_0$.

In the case of interest, the points represent received radar pulses and the ES observation window $\mathcal{X} \subset \mathbb{R}$ is a some finite union of bounded intervals.[2] Under these circumstances, each of the statistical models (2) and (3) defines a finite point process that can be described by Janossy measures over the spaces $\mathcal{X}^{(n)}$. For $n \geq 1$, each of these can be described by a density function known as a "Janossy density" denoted, respectively, by $j_{n|\mathbf{X}_\rho}(\cdot)$ and $j_{n|\mathbf{X}_\rho \cup \mathbf{X}_T}(\cdot)$.

For the Poisson noise process (2), the Janossy density function is simply given by

$$j_{n|\mathbf{X}_\rho}(x_1, \cdots, x_n) = e^{-\rho D} \rho^n \tag{A.1}$$

where $D = \int_{\mathcal{X}} \mathrm{d}x$ represents the duration, or Lebesgue measure, of the observation window. Not surprisingly, this function depends on the observation only through the number of pulses n. When this number is fixed, the events themselves are distributed

[2] This is in contrast to the less realistic approximation $W = \mathbb{R}$ used in the body of this memo. A stationary Poisson process over $\mathbb{R}$ a point process but it is not finite.

uniformly throughout the window. The Janossy measure of the empty outcome is $J_{0|\mathbf{X}_\rho}(\{\varepsilon\}) = e^{-\rho D}$.

In the case of the signal-in-noise model (3), the density must account for the normal distribution (4) of the transmitted pulses. If the window $\mathcal{X}$ a sufficiently long interval, then the effect of transmitted pulses falling outside of $\mathcal{X}$ can be neglected and the density function is approximately given by

$$j_{n|\mathbf{X}_\rho \cup \mathbf{X}_T}(x_1, \cdots, x_n) \approx e^{-\rho D} \sum_{K \in \mathbb{K}(T,X)} \Big[\rho^{n-|K|}(1-P_\mathrm{p})^{|T|-|K|} \times P_\mathrm{p}^{|K|} \prod_{(t,x)\in K} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(t+\tau-x)^2/2\sigma^2}\Big] . \tag{A.2}$$

For shorthand, the symbol $X \equiv \{x_1, \cdots, x_n\}$ has been used to represent the multiset of points from the process. The sum $\sum_K$ extends over all association relations (6) between the signal template T and the multiset X. The Janossy measure of the empty outcome is $J_{0|\mathbf{X}_\rho \cup \mathbf{X}_T}(\{\varepsilon\}) \approx e^{-\rho D}(1-P_\mathrm{p})^{|T|}$.

The two density functions (A.1) and (A.2) are expected to agree in the special case when $P_\mathrm{p} = 0$, since it is almost surely the case that $\mathbf{X}_T = \emptyset$. This agreement is realized since, as $P_\mathrm{p} \to 0$, only the $K = \emptyset$ relation contributes to (A.2).

A.2 Optimal decision criteria

The objective of binary decision theory is to partition the complete sample space $\cup_{n=0}^{\infty} \mathcal{X}^{(n)}$ into two complimentary events that satisfy some optimized criterion. For the finite point process over $\mathcal{X}$, let $\omega^{(n)} \subseteq \mathcal{X}^{(n)}$ be some Borel set defined for each $n \geq 0$, and let $\mathcal{X}^{(n)} \backslash \omega^{(n)}$ denote its compliment. Once such a partition is established, it can be used to make decisions. If an observed outcome has n points and belongs to $\omega^{(n)}$ then one should estimate that the signal is present and the process is governed by (3). If the outcome belongs to $\mathcal{X}^{(n)} \backslash \omega^{(n)}$, then one should estimate that (2) is governing process. The partition may be seen as a generalization of the simple scoring threshold (5).

Since the ES system is not generally provided with a labeling mechanism to distinguish individual pulses, the decision region $\omega^{(n)}$ is required to be symmetric under the interchange of pulse indices. For each n-tuple that appears in the decision region $(x_1, \cdots, x_n) \in \omega^{(n)}$, each of its permutations must also be present, $(x_{i_1}, \cdots, x_{i_n}) \in \omega^{(n)}$, since there is no mechanism to tell the difference. In this case, the total probability measure of $\cup_n \omega^{(n)}$ can be written in terms of the Janossy measures

$$\Pi\Big(\bigcup_{n=0}^{\infty} \omega^{(n)}\Big) \equiv \sum_{n=0}^{\infty} p_n \Pi_n\big(\omega^{(n)}\big) = \sum_{n=0}^{\infty} \frac{1}{n!} J_n\big(\omega^{(n)}\big) .$$

Applying the noise-only and signal-in-noise models to this result, and writing the measures in terms of densities one obtains

$$P_{\rm fa} = J_{0|\mathbf{X}_\rho}\big(\omega^{(0)}\big) + \sum_{n=1}^{\infty} \frac{1}{n!} \int \cdots \int_{\omega^{(n)}} j_{n|\mathbf{X}_\rho}(x_1, \cdots, x_n)\, {\rm d}x_1 \cdots {\rm d}x_n$$

and

$$P_{\rm det} = J_{0|\mathbf{X}_\rho \cup \mathbf{X}_T}\big(\omega^{(0)}\big) + \sum_{n=1}^{\infty} \frac{1}{n!} \int \cdots \int_{\omega^{(n)}} j_{n|\mathbf{X}_\rho \cup \mathbf{X}_T}(x_1, \cdots, x_n)\, {\rm d}x_1 \cdots {\rm d}x_n \,.$$

Note that the factors $\frac{1}{n!}$ can be interpreted as corrections for over-counting indistinguishable regions in the sample space. Since both the Janossy densities and the integration region $\omega^{(n)}$ are defined to be symmetric, some authors may elect to eliminate these factors by restricting the integration region to the subspace in which $x_1 \leq x_2 \leq \cdots \leq x_n$.

A typical criterion for determining the partitions $\omega^{(n)}$ is the requirement that some linear cost function be minimized [7]. For example, if $c_{\rm fa}$ and $c_{\rm miss}$ are specified positive cost coefficients then the combination

$$c_{\rm fa} P_{\rm fa} + c_{\rm miss}(1 - P_{\rm det})$$

can be precisely minimized by the decision region determined by a likelihood ratio threshold

$$\omega^{(n)} = \big\{(x_1, \cdots, x_n) \in \mathcal{X}^{(n)} \big| \Lambda_n(x_1, \cdots, x_n) \geq \gamma_\Lambda \big\}$$

where $\gamma_\Lambda \equiv c_{\rm fa}/c_{\rm miss}$ and, for $n \geq 1$,

$$\begin{aligned} \Lambda_n(x_1, \cdots, x_n) &\equiv \frac{j_{n|\mathbf{X}_\rho \cup \mathbf{X}_T}(x_1, \cdots, x_n)}{j_{n|\mathbf{X}_\rho}(x_1, \cdots, x_n)} \\ &\approx \sum_{K \in \mathbb{K}(T,X)} \left[\rho^{-|K|} (1 - P_{\rm p})^{|T|-|K|} P_{\rm p}^{|K|} \prod_{(t,x) \in K} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(t+\tau-x)^2/2\sigma^2} \right]. \end{aligned} \tag{A.3}$$

Here $X = \{x_1, \cdots, x_n\}$ is the multiset of pulse points. The above criterion is equally valid in the $n = 0$ case where the likelihood ratio of the empty outcome reduces to

$$\Lambda_0(\varepsilon) \equiv \frac{J_{0|\mathbf{X}_\rho \cup \mathbf{X}_T}(\{\varepsilon\})}{J_{0|\mathbf{X}_\rho}(\{\varepsilon\})} \approx (1 - P_{\rm p})^{|T|} \,.$$

A.3 Distribution of the likelihood score

The collection of functions $\Lambda_n : \mathcal{X}^{(n)} \to \mathbb{R}$ for $n \geq 0$ defines a likelihood ratio function over the larger sample space denoted by $\Lambda : \cup_n \mathcal{X}^{(n)} \to \mathbb{R}$. Since this function is continuous over the entire sample space, it may be identified with with a random variable, denoted by $\mathbf{\Lambda}$. Distributions for $\mathbf{\Lambda}$ conditioned on the two models (2) and (3), or approximations thereof, are required to determine the performance bounds on scoring functionals of the type (1).

Several aspects of this problem are discussed hereafter, although they fall short of a complete solution.

A.3.1 Separation of the likelihood score

The structure of the likelihood ratio in (A.3) is complicated by the fact that the association relations K may associate no more than one template element from T with each pulse from X, in accordance with (6). Under most circumstances, however, this restriction does not significantly impact on the value of $\Lambda_n(x_1, \cdots, x_n)$ for any given outcome $(x_1, \cdots, x_n)$, and the sum can be extended to relations in the class

$$\tilde{\mathbb{K}}(T, X) \equiv \left\{ \begin{array}{c} K \,|\, K \subset T \times X \,, \\ (t, x), (t', x') \in K \Longrightarrow [t = t', x = x'] \text{ or } t \neq t' \end{array} \right\} . \tag{A.4}$$

To see this suppose that the template separation (8) is large compared with the jitter so that $\Delta_T \gg \sigma$. Suppose that $K \in \mathbb{K}(T, X)$ is some relation and $\tilde{K} = K \cup \{(t_1, x), (t_2, x)\}$ is an augmentation of K generated by associating $x \in X$ with two distinct template elements $t_1 \neq t_2$, in violation of (6), but still satisfying (A.4). If K and $\tilde{K}$ were both allowed to contribute to the sum in (A.3), then the ratio of their contributions would be equal to

$$\frac{P_{\mathrm{p}}^2}{(1-P_{\mathrm{p}})^2} \frac{1}{2\pi\sigma^2\rho^2} e^{-(t_1+\tau)^2/2\sigma^2} e^{-(t_2+\tau)^2/2\sigma^2} < \frac{P_{\mathrm{p}}^2}{(1-P_{\mathrm{p}})^2} \frac{1}{2\pi\sigma^2\rho^2} e^{-(t_1-t_2)^2/4\sigma^2} \,.$$

For the parameter values considered in this memo, with say $\Delta_T \approx 100\mu$s, the bound on this ratio is on the order of 10^{-6}. The inclusion of all relations that satisfy (A.4) is expected to have a negligible effect on (A.3).

The advantage of this approximation is that it allows the likelihood ratio $\Lambda(X)$ to be factored into distinct contributions for each element of T resulting in the product

$$\Lambda(\cdot) \approx \prod_{t \in T} \Lambda_t(\cdot)$$

where

$$\Lambda_t(x_1, \cdots, x_n) \equiv (1 - P_{\mathrm{p}}) + \frac{P_{\mathrm{p}}}{\sqrt{2\pi}\sigma\rho} \sum_{x \in X} e^{-(t+\tau-x)^2/2\sigma^2} \tag{A.5}$$

and the multiset notation $X \equiv \{x_1, \cdots, x_n\}$ has been used. The function $\Lambda_t(\cdot)$ is likewise a continuous function on the sample space and may be denoted by the random variable $\mathbf{\Lambda}_t$. Strictly speaking, the variables $\mathbf{\Lambda}_{t_1}$ and $\mathbf{\Lambda}_{t_2}$ are not independent for any two template elements $t_1, t_2 \in T$ since they both depend on the entire underlying point process. However, when the template elements are well separated so that $\Delta_T \gg \sigma$, then, to a very good approximation, the significant contributions to the sum in $\sum_{x\in X}$ in (A.5) come only from those points x that lie in the interval $x \in (t+\tau-\Delta_T/2, t+\tau+\Delta_T/2) \cap \mathcal{X}$. Since such intervals do not overlap for different $t \in T$, the variables $\mathbf{\Lambda}_t$ must be approximately independent.

A.3.2 Distribution of likelihood factors

The distribution for the variables $\mathbf{\Lambda}_t$ is difficult to estimate, although some aspects can be computed in closed form. In order to analyze this problem, it is convenient to consider two statistically independent random variables representing the contributions to $\mathbf{\Lambda}_t$ from noise and signal

$$\mathbf{\Lambda}_{t;\rho} = (1 - P_{\mathrm{p}}) + \frac{P_{\mathrm{p}}}{\sqrt{2\pi}\sigma\rho} \sum_{x\in\mathbf{X}_\rho} e^{-(t+\tau-x)^2/2\sigma^2}$$
$$\mathbf{\Lambda}_{t;T} = \frac{P_{\mathrm{p}}}{\sqrt{2\pi}\sigma\rho} \sum_{x\in\mathbf{X}_T} e^{-(t+\tau-x)^2/2\sigma^2} .$$

In this way, the hypotheses $\mathbf{X} = \mathbf{X}_\rho$ implies that $\mathbf{\Lambda}_t = \mathbf{\Lambda}_{t;\rho}$; conversely, $\mathbf{X} = \mathbf{X}_\rho \cup \mathbf{X}_T$ implies that $\mathbf{\Lambda}_t = \mathbf{\Lambda}_{t;\rho} + \mathbf{\Lambda}_{t;T}$.

The characteristic function for the variable $\mathbf{\Lambda}_t$ is given by

$$\phi_{\mathbf{\Lambda}_t}(z) \equiv E\Big\{\exp\big(iz\mathbf{\Lambda}_t\big)\Big\} .$$

In the $\mathbf{X} = \mathbf{X}_\rho$ and $\mathbf{X} = \mathbf{X}_\rho \cup \mathbf{X}_T$ cases, respectively, this characteristic function becomes

$$\begin{aligned} \phi_{\mathbf{\Lambda}_t|\mathbf{X}_\rho}(z) &= \phi_{\mathbf{\Lambda}_{t;\rho}}(z) \\ \phi_{\mathbf{\Lambda}_t|\mathbf{X}_\rho\cup\mathbf{X}_T}(z) &= \phi_{\mathbf{\Lambda}_{t;\rho}}(z)\phi_{\mathbf{\Lambda}_{t;T}}(z) \end{aligned} \tag{A.6}$$

where the characteristic functions for $\mathbf{\Lambda}_{t;\rho}$ and $\mathbf{\Lambda}_{t;T}$ are defined by

$$\phi_{\mathbf{\Lambda}_{t;\rho}}(z) \equiv e^{iz\beta} E\Big\{\prod_{x\in\mathbf{X}_\rho} \exp\Big(iz\alpha e^{-(t+\tau-x)^2/2\sigma^2}\Big)\Big\}, \text{ and}$$
$$\phi_{\mathbf{\Lambda}_{t;T}}(z) \equiv E\Big\{\prod_{x\in\mathbf{X}_T} \exp\Big(iz\alpha e^{-(t+\tau-x)^2/2\sigma^2}\Big)\Big\}.$$

Here the parameters

$$\alpha \equiv \frac{P_{\mathrm{p}}}{\sqrt{2\pi}\rho\sigma} ,$$
$$\beta \equiv 1 - P_{\mathrm{p}}$$

have been used for notational convenience.

Since $\mathbf{X}_\rho$ is a Poisson process, the result cited in Annex B.2.2 can be used to represent the first characteristic function $\phi_{\mathbf{\Lambda}_{t;\rho}}(z)$ as an integral expression:

$$\ln \phi_{\mathbf{\Lambda}_{t;\rho}}(z) = iz\beta + \int_{\mathcal{X}} \rho \Big[\exp\big(iz\alpha e^{-(t+\tau-x)^2/2\sigma^2} \big) - 1 \Big] \,\mathrm{d}x \,. \qquad \text{(A.7a)}$$

The distribution for $\mathbf{\Lambda}_{t;\rho}$ can be represented by a density function denoted by

$$p_{\mathbf{\Lambda}_{t;\rho}}(\lambda)\,, \quad \text{for } \lambda > \beta \,.$$

The density function $p_{\mathbf{\Lambda}_{t;\rho}}(\lambda)$ is related to the Fourier transform of the characteristic function $\phi_{\mathbf{\Lambda}_{t;\rho}}(z)$. Neither representation is straightforward to evaluate directly.

The second characteristic function $\phi_{\mathbf{\Lambda}_{t;T}}(z)$ can be computed by invoking the large template separation approximation $\Delta_T \gg \sigma$. In this case, most of the pulses in $\mathbf{X}_T$ are likely contribute factors that are very close to unity. Only the pulse $\mathbf{x}_t$ is likely to have a non-trivial contribution, if it is detected. As a result, one may use the Gaussian probability distribution (4) to write

$$\phi_{\mathbf{\Lambda}_{t;T}}(z) \approx \beta + \alpha\rho \int_{\mathcal{X}} e^{-(t+\tau-x)^2/2\sigma^2} \exp\big(iz\alpha e^{-(t+\tau-x)^2/2\sigma^2} \big) \,\mathrm{d}x \,. \qquad \text{(A.7b)}$$

The distribution for $\mathbf{\Lambda}_{t;T}$ is well approximated in terms of a finite miss probability and a continuous density function defined respectively by

$$P(\mathbf{\Lambda}_{t;T} = 0) = \beta\,, \text{ and}$$
$$p_{\mathbf{\Lambda}_{t;T}}(\lambda) = \frac{\sqrt{2}\rho\sigma}{\sqrt{\ln(\alpha/\lambda)}}\,, \quad \text{for } 0 < \lambda < \alpha \,.$$

A.3.3 An integral equation

It is interesting to note that the derivative of (A.7a) is exactly equal to (A.7b), demonstrating that the two characteristic functions satisfy the equation

$$\frac{\mathrm{d}}{\mathrm{d}(iz)} \phi_{\mathbf{\Lambda}_{t;\rho}}(z) = \phi_{\mathbf{\Lambda}_{t;\rho}}(z) \phi_{\mathbf{\Lambda}_{t;T}}(z) \,. \qquad \text{(A.8)}$$

This differential relationship between the characteristic functions implies that exists an integral relationship between the density functions. After accounting for the finite probability that $\mathbf{\Lambda}_{t;T}$ vanishes, this relationship can be expressed in terms of a convolution between the continuous components of the probability density functions

$$(\lambda - \beta) p_{\mathbf{\Lambda}_{t;\rho}}(\lambda) = \int_{\max\{\lambda-\alpha,\beta\}}^{\lambda} p_{\mathbf{\Lambda}_{t;T}}(\lambda - \nu) p_{\mathbf{\Lambda}_{t;\rho}}(\nu) \,\mathrm{d}\nu \,, \qquad \text{for } \lambda \geq \beta \,.$$

The integration limits on the convolution have been chosen to respect the support region of the constituent densities. Since $p_{\mathbf{\Lambda}_{t;T}}(\lambda)$ is a known function, this result may be treated as a linear homogenous integral equation for the unknown density function $p_{\mathbf{\Lambda}_{t;\rho}}(\lambda)$.

A.3.4 Moments

The moments of these distributions can be computed in a straightforward manner. Suppose that the cumulants of the known distribution for $\mathbf{\Lambda}_{t;T}$ are denoted by χ_n so that

$$\chi_n \equiv \left.\frac{\mathrm{d}^n \phi_{\mathbf{\Lambda}_{t;T}}(z)}{\mathrm{d}^n(iz)}\right|_{z=0} \quad \text{for } n = 0, 1, \cdots .$$

Then, using (A.7b), it is straightforward to see that these constants are given by $\chi_0 = 1$ and

$$\chi_n = \rho\alpha^{n+1} \int_{\mathcal{X}} e^{-(n+1)(t+\tau-x)^2/2\sigma^2} \mathrm{d}x \approx \frac{P_{\mathrm{p}}\alpha^n}{\sqrt{n+1}} \quad \text{for} \;\; n = 1, 2, \cdots .$$

In estimating the values of these constants, it has been noted that the integral can be approximately evaluated in the case when the observation interval is a sufficiently long interval. The approximation becomes exact if $\mathcal{X} = \mathbb{R}$.

The differential relationship (A.8) implies that these same constants are related to the Taylor coefficients of $\ln \phi_{\mathbf{\Lambda}_{t;\rho}}(z)$ by

$$\left.\frac{\mathrm{d}^n \ln \phi_{\mathbf{\Lambda}_{t;\rho}}(z)}{\mathrm{d}^n(iz)}\right|_{z=0} = \chi_{n-1} \quad \text{for} \;\; n = 1, 2, 3, \cdots .$$

These results can be used to compute any finite number of cumulants and moments of the distribution for $\mathbf{\Lambda}_t$ conditioned either the noise or signal-in-noise model. In principal this may provide a means to estimate the distributions themselves through the application of an Edgeworth or Gram-Charlier series [12].

Annex B: Relevant statistics and distributions

B.1 Statistics of minima

It $\mathbf{x}$ and $\mathbf{y}$ are two independent continuous random variables whose probability densities are given by $p_{\mathbf{x}}(x)$ and $p_{\mathbf{y}}(y)$ respectively then it is straightforward to show that the probability density for $\mathbf{z} = \min\{\mathbf{x}, \mathbf{y}\}$ is given by

$$\begin{aligned} p_{\min\{\mathbf{x},\mathbf{y}\}}(z) &= p_{\mathbf{x}}(z)\left[\int_z^\infty p_{\mathbf{y}}(y)\,\mathrm{d}y\right] + p_{\mathbf{y}}(z)\left[\int_z^\infty p_{\mathbf{x}}(x)\,\mathrm{d}x\right] \\ &= p_{\mathbf{x}}(z)P(\mathbf{y} \geq z) + p_{\mathbf{y}}(z)P(\mathbf{x} \geq z)\,. \end{aligned} \tag{B.1}$$

B.2 Poisson statistics

The noise $\mathbf{X}_\rho$ process introduced in Subsection 2.1 has been modeled to follow the distribution of events in a Poisson point process on some window $\mathcal{X} \subseteq \mathbb{R}$. It has been assumed that the noise is stationary and can be described by a constant event density ρ.

A standard property of such a process is that the expected number of observed pulses contained within any interval of the form $(t_-, t_+) \subset \mathcal{X}$ is given by

$$E\left(|\mathbf{X}_\rho \cap (t_-, t_+)|\right) = (t_+ - t_-)\rho\,.$$

The exact distribution for the number of pulses in a finite interval is

$$P\left(|\mathbf{X}_\rho \cap (t_-, t_+)| = n\right) = e^{-(t_+ - t_-)\rho}\frac{((t_+ - t_-)\rho)^n}{n!}\ ,\ n \geq 0\,, \tag{B.2}$$

for any non-negative integer n.

Some useful properties and statistics of the Poisson process are summarized below.

B.2.1 Avoidance probability for Poisson processes

Suppose that x_0 is some fixed time and one is interested in the interval between x_0 and the nearest element of $\mathbf{X}_\rho$. Let the length of this interval be denoted by

$$\mathbf{\Delta} = \min\{|x - x_0|\big|x \in \mathbf{X}_\rho\}\,.$$

In the case where the observation window is the entire real line $\mathcal{X} = \mathbb{R}$, then $\mathbf{\Delta}$ almost certainly exists, and its statistics can be represented by an "avoidance probability" [9], given by (B.2) with $n = 0$,

$$P(\mathbf{\Delta} > \delta) = P(\mathbf{X}_\rho \cap (x_0 - \delta, x_0 + \delta) = \emptyset) = e^{-2\delta\rho}\,.$$

The density function for this variable is obtained by differentiating the complement

$$p_{\boldsymbol{\Delta}}(\delta) = \begin{cases} 2\rho e^{-2\rho\delta} & \text{if } \delta \geq 0 \\ 0 & \text{if } \delta < 0 \,. \end{cases}$$

B.2.2 Product functionals on Poisson processes

In some cases it is of interest to compute the statistics of some functional of the random point process. One example of such a functional is

$$\mathbf{F} \equiv \prod_{x \in \mathbf{X}_\rho} f(x)$$

where $\mathbf{X}_\rho$ is the Poisson point process with density ρ over the window $\mathcal{X}$ and the function $f : \mathcal{X} \to \mathbb{R}$ specifies the point-wise contributions.

Although the author is not aware of a general formalism for computing arbitrary statistics of such a functional, the expectation can be computed in closed form:

$$E\big\{\mathbf{F}\big\} = \exp\Big(-\int_{\mathcal{X}} \rho \, \mathrm{d}x + \int_{\mathcal{X}} \rho f(x) \, \mathrm{d}x\Big) \,.$$

This result applies in the case where $\mathcal{X}$ is a finite union of bounded intervals, and the function $f(x)$ is integrable over $\mathcal{X}$, ensuring that the preceding integrals converge. Interestingly [11], this result also applies in the case of non-stationary Poisson processes where ρ is a known function of $x \in \mathcal{X}$.

References

[1] Wiley, R. G. (1993). Electronic intelligence: the analysis of radar signals, Second ed. Norwood, MA: Artech House.

[2] Mardia, H. K. (1989). New techniques for the deinterleaving of repetitive sequences. *IEE Proc. F, Commun., Radar & Signal Process.*, **136**(4), 149–154.

[3] Milojević, D. J. and Popović, B. M. (1992). Improved algorithm for the deinterleaving of radar pulses. *IEE Proc. F, Commun., Radar & Signal Process.*, **139**(1), 98–104.

[4] Elton, S. D. (2001). A cross-correlation technique for radar emitter recognition using pulse time of arrival information. In *Proceedings of 2001 Workshop on Defence Applications of Signal Processing*, pp. 19–23. Borossa Valley, South Australia.

[5] Dilkes, F. A. (2004). Statistical Cross-correlation Algorithms for Temporal Pulse Profiles. (DRDC Ottawa TM 2004-220). Defence R&D Canada – Ottawa. UNCLASSIFIED.

[6] Roe, J., Elton, S., Huntington-Thresher, I., Noone, G., Lavoie, P., Dworkin, K., and Reid, C. (2002). Emitter deinterleaving and identification using *a priori* temporal signal profiles - Interim Report. TTCP Technical Report, TR-EWS-TP4-1-2002. UNCLASSIFIED.

[7] Van Trees, H. L. (1968). Detection, Estimation and Modulation Theory, Vol. I Detection, Estimation and Linear Modulation Theory. John Wiley & Sons.

[8] Elton, S. D. (1999). Jitter identification techniques for a regular event-based process. In *Proceedings of the 1999 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'99)*, Vol. 5, pp. 2801–2804. Pheonix, Arizona, USA.

[9] Daley, D. J. and Vere-Jones, D. (2003). An Introduction to the Theory of Point Processes, Second ed. Vol. I: Elementary Theory and Methods. New York: Springer.

[10] Wiley, R. G. (1985). Electronic intelligence: the interception of radar signals, Norwood, MA: Artech House.

[11] Bar-David, I. (1969). Communication under the Poisson Regime. *IEEE Trans. Information Theory*, **15**(1), 31–37.

[12] Stuart, A. and Ord, J. K. (1987). Kendall's Advanced Theory of Statistics, 5th ed. Vol. 1: Distribution Theory. New York: Oxford University Press.

This page intentionally left blank.

Distribution list

DRDC Ottawa TM 2005-167

External distribution

General

1 ADM (S&T)

1 DRDKIM 3

1 CISTI

2 Library and Archives Canada

Maritime

1 DSTM (Lavoie, Pierre)

Total external copies: 6

Internal distribution

1 Dilkes, Fred / REW

1 Rivest, Jean-François / REW

1 Zhou, Yifeng / REW

4 DRDC Ottawa Library

Total internal copies: 7

Total copies: 13

This page intentionally left blank.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.)

   Defence R&D Canada - Ottawa
   3701 Carling Avenue, Ottawa ON, CANADA, K1A 0Z4

2. SECURITY CLASSIFICATION (overall security classification of the document including special warning terms if applicable).

   UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C,R or U) in parentheses after the title).

   Statistical Decision Thresholds for Pulse Template Cross-Correlators (U)

4. AUTHORS (last name, first name, middle initial)

   Dilkes, Fred A.

5. DATE OF PUBLICATION (month and year of publication of document)

   November 2005

6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc).

   52

6b. NO. OF REFS (total cited in document)

   12

7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered).

   Technical Memorandum

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include address).

   Defence R&D Canada - Ottawa
   3701 Carling Avenue, Ottawa ON, CANADA, K1A 0Z4

9a. PROJECT NO. (the applicable research and development project number under which the document was written. Specify whether project).

   13EL13

9b. GRANT OR CONTRACT NO. (if appropriate, the applicable number under which the document was written).

10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique.)

   DRDC Ottawa TM 2005-167

10b. OTHER DOCUMENT NOs. (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

   (X) Unlimited distribution
   (   ) Defence departments and defence contractors; further distribution only as approved
   (   ) Defence departments and Canadian defence contractors; further distribution only as approved
   (   ) Government departments and agencies; further distribution only as approved
   (   ) Defence departments; further distribution only as approved
   (   ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution beyond the audience specified in (11) is possible, a wider announcement audience may be selected).

   Full unlimited announcement

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual).

A statistical formalism is proposed to evaluate the performance of waveform-based signal detection techniques for Electronic Support. The formalism applies to a large class of cross-correlation methods that attempt to match the arrival times of a sequence of pulses against an $a\ priori$ pulse template. In this framework, the detection of a particular signal of interest is treated as a problem from binary decision theory and the observations are treated as outcomes of a point process. The presence of a signal is modeled using a pulse template with non-cumulative Gaussian jitter while background noise is modeled using a stationary Poisson point process. Two specific cross-correlation techniques are evaluated and their performances are compared in terms of false-alarm and detection probabilities. The capability of these techniques to discern between the presence and absence of the signal is measured as a function of a decision threshold.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title).

Electronic Support
Electronic Warfare
Pattern Recognition
Pulse Processing
Radar
Cross-correlation
Detection
Point process
Receiver Operating Characteristic

Defence R&D Canada

Canada's leader in Defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale

DEFENCE DÉFENSE

www.drdc-rddc.gc.ca