Performance Analysis of a Dynamic Programming Track Before Detect Algorithm

LEIGH A. JOHNSTON, Member, IEEE

VIKRAM KRISHNAMURTHY, Senior Member, IEEE
The University of Melbourne

We analyze a dynamic programming (DP)-based track before detect (TBD) algorithm. By using extreme value theory we obtain explicit expressions for various performance measures of the algorithm such as probability of detection and false alarm. Our analysis has two advantages. First the unrealistic Gaussian and independence assumptions used in previous works are not required. Second, the probability of detection and false alarm curves obtained fit computer simulated performance results significantly more accurately than previously proposed analyses of the TBD algorithm.

Manuscript received January 10, 2001; revised September 17, 2001; released for publication September 17, 2001.

IEEE Log No. T-AES/38/1/02586.

Refereeing of this contribution was handled by P. Willett.

This research was partially funded by the Centre of Expertise in Networked Decision Systems (CENDS) and the CRC for Sensor Signal and Information Processing (CSSIP).

Parts of this paper appeared in the Proceedings of ICASSP2000, Istanbul, Turkey, June 2000.

Authors' address: Department of Electrical and Electronic Engineering, The University of Melbourne, Victoria, 3010, Australia, E-mail: (l.johnston,v.krishnamurthy@ee.mu.oz.au).

I. INTRODUCTION

In high noise and clutter, target tracking techniques, such as the Kalman filter, probabilistic data association (PDA) and multiple hypothesis testing (MHT), declare detections at each measurement time. The detections are then used to estimate the target trajectory. A limitation of these methods is that much of the information contained in the measurements is discarded due to the application of a detection threshold. This is akin to the information loss when hard rather than soft decisions are made in communication system receivers [1].

Track before detect (TBD) is a technique for target detection and tracking that is useful when the signal-to-noise ratio (SNR) is low. Unlike the above approaches, in TBD detections are not declared at each frame. Instead, a number of frames of data are processed, after which the estimated target track is returned when the detection is declared.

The literature on TBD is limited. In the early 1980s, unthresholded CCD images were used to track satellites against background noise. In [2], 3-D matched filtering is applied to moving target detection using optical images, while in [3] a recursive moving-target indication algorithm is used. These methods assume that the target velocity is known. Their performance degrades in the presence of a velocity mismatch or a target maneuver.

More recently, dynamic programming (DP)-based TBD methods have been proposed in [4–7] for detecting and tracking targets in low SNR. Indeed [6] is one of the first papers in the literature to analyze the detection performance of a DP-based tracking algorithm. The DP approach originally proposed in [5] avoids problems with velocity mismatch and can handle slowly maneuvering targets. The analysis in [6] adequately describes the detection performance of the DP algorithm in [5] but does not address the tracking performance of the algorithm. In [4], the DP algorithm of [5] is modified and extended to improve efficiency, enhance performance in non-Gaussian noise and allow velocity transitions.

This paper is based on the DP-based TBD algorithm proposed in [7]. The DP algorithm in [7] effectively integrates the measurements along possible target trajectories, returning as possible targets those trajectories for which the measurement sum (merit function) exceeds a given threshold. The main contribution of this work is to analyze the DP-based TBD algorithm presented in [7]. By using extreme value theory (EVT), we obtain explicit expressions for the asymptotic false alarm probability P_{FA} and track detection probability P_{D}. Our performance analysis provides two major advantages as follows over the heuristic analysis presented in [7].

1) The asymptotic expressions for the P_{FA} and P_{D} obtained via EVT fit computer-simulated results

significantly better (up to an order of magnitude) than existing expressions in the TBD literature such as those in [7].

2) In [7] (and [6]), to simplify the analysis it is assumed that the merit functions at each time instant in the DP algorithm are statistically independent and have Gaussian densities. These assumptions are unrealistic since the merit functions in a DP recursion are usually dependent and non-Gaussian. The max operator in the DP algorithm clearly means that the densities are non-Gaussian. We show that by using EVT neither the independence nor Gaussianity of the merit function is required to derive expressions for P_{D} and P_{FA}.

EVT is a subclass of weak convergence theory, concerned with limiting distributions of extremal events. The limiting distributions for sums of random variables are characterized by the Central Limit Theorem (CLT). In the same way, maxima of random variables are characterized by EVT as one of three possible invariant *max-stable* distributions. These are the Weibull, Gumbel, and Frechet distributions [8]. EVT can be applied to approximate the tail of a probability distribution. The curve is defined by two parameters, location and scale, in the case of classical EVT, or three parameters, location, scale and a shape exponent, when applying generalized EVT (GEVT).

Although analytic expressions for the EVT and GEVT curves can be calculated for certain known distributions from which the maxima are sampled, in most cases of interest the parameters must be estimated numerically. While there is inherent error in applying extremal analysis due to the need to estimate curve parameters, this is compensated for by the accuracy attainable with relatively little computational effort. Here we estimate the parameters of the extremal curves in two ways—numerically and by an approximate analytical procedure—both of which provide superior performance curves to those achieved by the Gaussian approximation in [7]. The approximate analytical procedure we present is based on viewing the TBD problem as a highly interconnected stochastic network to which the DP algorithm is applied. Extremal analysis of much simpler network configurations is considered in [9]. The analytical expressions that we derive readily yield accurate analytical performance bounds of the TBD algorithm as a function of the target parameters and can be used in developing "rules of thumb" in the design of TBD systems.

Traditionally, EVT has been applied in areas of science and econometrics to analyze both naturally occurring rare event data, such as river floodings and drought cycles [10], and extremal events in the finance and insurance industries such as stock market crashes [8]. Recently EVT has found application in communications engineering, having been used in bit error rate estimation for digital communication systems [11], ATM buffer dimensioning [12], interference parameter estimation [13], parameter estimation in hidden Markov models [14], and some radar system design problems [15].

The performance analysis of the TBD problem we present here is both useful and necessary to reduce the number of costly simulations required to obtain accurate false alarm and detection probabilities via a classical counting procedure. The threshold level V_T (see Section IIB) is an important factor in system design, and theoretical curves are required to determine appropriate values, without having to resort to resimulation of the system. While false alarm performance measures are analyzed in [15], their radar systems are assumed to handle independent identically distributed (IID) random variables, for which the application of EVT is straightforward. The dependencies introduced in the DP algorithm render a fundamentally different problem here.

This paper is organized as follows. In Section II, the TBD problem and DP solution are stated. This is followed by an outline of relevant EVT results in Section III. The evaluation of TBD system performance via EVT and numerical parameter estimation is then presented in Section IV, and for completeness, a review of the Gaussian approximation used in [7] follows in Section IIE. A comparison of the EVT and Gaussian approximation performance analysis techniques is carried out via numerical simulations in Section V. Lastly, in Section VI, an approximate analytical method for obtaining the EVT curve parameters is demonstrated. This too is compared with the Gaussian approximation via numerical simulations.

II. TRACK BEFORE DETECT PROBLEM FORMULATION AND ALGORITHM

A. Problem Formulation

We consider the same problem as in [7]. Consider the problem of tracking a point target that moves with constant velocity in the x-y plane. An extended region surrounding the target is monitored by a sensor, consisting of an $L \times L$ grid of square resolution cells of side length Δ. At time k, the matrix of measured intensities recorded by the sensor is $\mathbf{Z}_k = \{z_k(i,j)\}$, $i,j \in [1,L]$, where

$$z_k(i,j) = \begin{cases} A_k + w_k(i,j) & \text{target in cell } (i,j) \\ w_k(i,j) & \text{no target in } (i,j) \end{cases}. \tag{1}$$

Here A_k denotes the target amplitude which is assumed to be a constant for simplicity, i.e., $A_k = A$. The additive noise, $w_k(i,j) \sim N(0,\sigma_w^2)$, is assumed to be IID.

REMARK The above model is obtained by discretization of the target dynamics. Since our aim is to analyze the TBD algorithm, we only consider basic linear dynamics as given in [7]. More sophisticated versions of the algorithm are given in [6].

The sensor grid provides the necessary structure to model the target motion by the discrete process

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k, \qquad k = 1,\ldots,K \tag{2}$$

$$\mathbf{x}_k = \begin{bmatrix} x_k \\ y_k \\ u_k \\ v_k \end{bmatrix}, \qquad \mathbf{F} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

where $(x_k, y_k) \in [(1,1),(L,L)]$ and $(u_k, v_k) \in [(-M/2,-M/2),(M/2,M/2)]$ are the discrete target positions and velocities, respectively. T is the time between successive sensor readings, termed *frames*. For simplicity we assume that $T = 1$. M is a design parameter, chosen to encompass the velocity range of the target motion.

The TBD estimation objective is: Given the measurement sequence of K frames, determine the trajectories (state sequences) most likely to have originated from the actual target. The DP algorithm outlined in Section IIB is applied to solve the TBD problem.

B. Track Before Detect Dynamic Programming Algorithm

The TBD problem formulation (1)–(3) enables application of the following DP (modified Viterbi) algorithm, as in [7].

Let $\mathbf{X}_k$ be the set of all possible discrete states of the target at time k:

$$\mathbf{X}_k = \{[i,j,r,s]'\}, \qquad i,j \in [1,L], \quad r,s \in [-M/2,M/2]. \tag{4}$$

1) *Initialization*: For all $\mathbf{x}_1 = [i,r,j,s]' \in \mathbf{X}_1$,

$$I(\mathbf{x}_1) = z_1(i,j) \tag{5}$$

$$\Psi_1(\mathbf{x}_1) = 0. \tag{6}$$

2) *Recursion*: For $2 \le k \le K$,
For all $\mathbf{x}_k = [i,r,j,s] \in \mathbf{X}_k$,

$$I(\mathbf{x}_k) = \max_{\mathbf{x}_{k-1}}[I(\mathbf{x}_{k-1})] + z_k(i,j) \tag{7}$$

$$\Psi_k(\mathbf{x}_k) = \arg\max_{\mathbf{x}_{k-1}}[I(\mathbf{x}_{k-1})] \tag{8}$$

where the maximization is performed over the $\mathbf{x}_{k-1}$ for which a transition to $\mathbf{x}_k$ is possible (see below).

3) *Termination*: For threshold V_T find,

$$\{\hat{\mathbf{x}}_K\} = \{\mathbf{x}_K : I(\mathbf{x}_K) > V_T\}. \tag{9}$$

4) *Backtracking*: For all $\hat{\mathbf{x}}_K$, for $k = K-1,\ldots,1$

$$\hat{\mathbf{x}}_k = \Psi_{k+1}(\hat{\mathbf{x}}_{k+1}). \tag{10}$$

Notation: At any stage k, there are $n = L^2M^2$ recursions being followed through in the algorithm. We denote this set as $\mathbf{x}_k^{(1)},\ldots,\mathbf{x}_k^{(n)}$, in no particular order.

C. Valid Transitions

Let $q \in \{1,4,9,16,\ldots\}$ denote a fixed squared integer which defines the valid state transitions on the x-y plane. We assume that for every state at time k, there are q possible (valid) states at time $k+1$. From state $\mathbf{x}_k = [i,r,j,s]'$, the sets of possible states at time $k+1$ for $q = 4,9,16,25$ are

$$q = 4 \qquad \mathbf{x}_{k+1} \in \{[i+r-\delta_i, j+s-\delta_j, r-\delta_i, s-\delta_j] : \delta_i,\delta_j = -1,0\}$$

$$q = 9 \qquad \mathbf{x}_{k+1} \in \{[i+r-\delta_i, j+s-\delta_j, r-\delta_i, s-\delta_j] : \delta_i,\delta_j = -1,0,1\}$$

$$q = 16 \qquad \mathbf{x}_{k+1} \in \{[i+r-\delta_i, j+s-\delta_j, r-\delta_i, s-\delta_j] : \delta_i,\delta_j = -2,-1,0,1\}$$

$$q = 25 \qquad \mathbf{x}_{k+1} \in \{[i+r-\delta_i, j+s-\delta_j, r-\delta_i, s-\delta_j] : \delta_i,\delta_j = -2,-1,0,1,2\}.$$

Fig. 1 depicts the valid state transitions, for each q value, from the state $\mathbf{x}_k = [3,2,2,3]$.

This definition of valid transition is more precise than that given in [7]. Through the above definition, $q = 1$ is of no use, due to the discretization inherent in the algorithm. For example, if the signal was initially at position (1.5,2.3), and moved with a velocity of (1.2,0.8), $q = 1$ would lose the signal due to the noninteger position and velocity. It is for this reason that $q = 4,\ldots,25$ are considered.

The performance of the TBD algorithm is affected by the choice of q. As seen via the numerical simulations in Section V, $q = 4$ does not allow enough target maneuverability, even for constant velocity models. This stands to reason, for as Fig. 1 depicts, the valid transitions do not allow for a decceleration (which can be present for a constant velocity target due to the discretization), making this choice of q little better than $q = 1$. In Section VI it is demonstrated that satisfactory analytic extremal curves approximations are obtained for larger q (increased target maneuverability).

D. Performance Measures

In [7] the following two performance measures are used for evaluating the performance of the TBD algorithm. In Section IV, we use EVT to characterize these two performance measures.

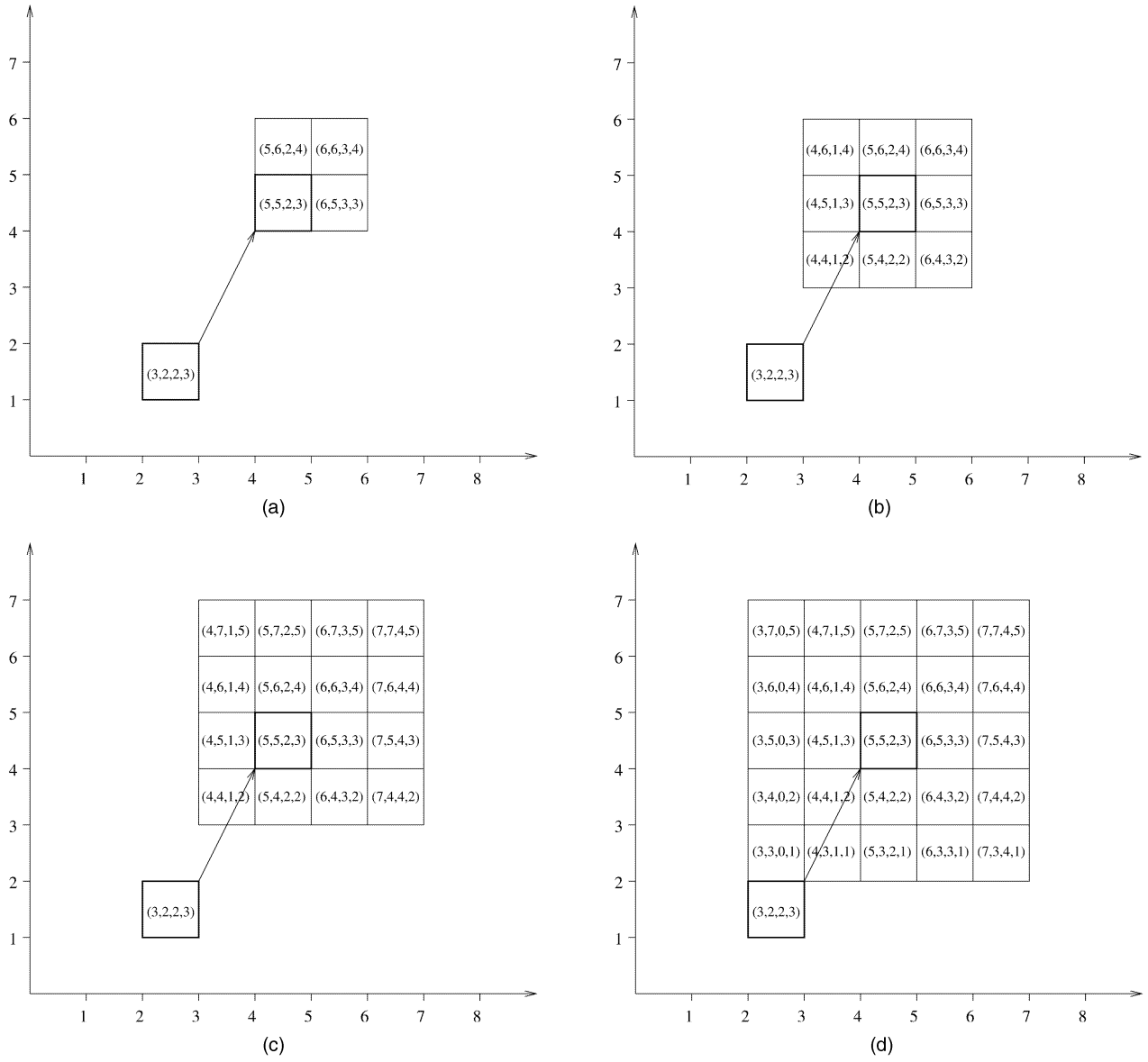


Fig. 1. Examples of valid transitions. (a) $q = 4$. (b) $q = 9$. (c) $q = 16$. (d) $q = 25$.

1) *Probability of false alarm* P_{FA} is defined as the probability of detecting at least one false track. This is equal to the probability of the maximum noise state merit function exceeding the threshold V_T. Hence

$$P_{\mathrm{FA}} = \Pr\left(\max_{\mathbf{x}_K} I(\mathbf{x}_K) > V_T\right) \tag{11}$$

where $\mathbf{x}_K \in \{\text{noise states}\}$.

2) *Probability of target detection* P_{D} is defined as the probability of the merit function $I(\mathbf{x}_K)$ exceeding the threshold V_T for at least one state for which the final (x, y) cell is within 2 cells of the actual target cell, and the velocity cell is equal to the actual target velocity cell. This is given by

$$P_{\mathrm{D}} = \Pr\left(\max_{\mathbf{x}_K} I(\mathbf{x}_K) > V_T\right) \tag{12}$$

where $\mathbf{x}_K \in \{\text{states within 2 cells of the target cell, with correct velocity cell}\}$. The limit of 2 cells has been chosen to replicate the scenario in [7].

Approximate evaluation of (11)–(12) is carried out in [6, 7] by assuming that the merit functions are independent and Gaussian distributed, a limited approximation due to the dependency introduced by the max operators in the DP algorithm (7). We approach the analysis using EVT.

E. Review of Gaussian Approximation of P_{FA} and P_{D}

So as to compare the EVT approach of Section III with the Gaussian approximation method of [7], in this section we briefly review this latter method. The approximate analysis of P_{FA} and P_{D} presented in [7] is based on the following two assumptions.

A1. The probability densities of the merit functions are assumed Gaussian.

A2. The merit functions at each frame k are assumed IID.

Under these assumptions, the performance measures, derived using order statistics, are

$$P_{\mathrm{FA}} = 1 - \Phi^n\left(\frac{V_T - \mu_K}{\sigma_K}\right) \tag{13}$$

$$P_{\mathrm{D}} = 1 - \Phi\left(\frac{V_T - \overline{\mu}_K}{\overline{\sigma}_K}\right)\Phi^{L^2M^2-1}\left(\frac{V_T - \mu_K}{\sigma_K}\right) \tag{14}$$

where $\Phi(x)$ is the Gaussian cumulative distribution function, with corresponding probability density function $\phi(x)$. The two sets of means and covariances, $\{\mu_K, \sigma_K\}$ and $\{\overline{\mu}_K, \overline{\sigma}_K\}$, are computed according to the following recursions (for derivation details refer to [7]).

1) *Initialization*: For $p = q, q-1$, compute:

$$\begin{aligned}\mu_{\max}(p) &= \int_{-\infty}^{\infty} px\phi(x)\Phi^{p-1}(x)\,dx \\ \sigma^2_{\max}(p) + \mu^2_{\max}(p) &= \int_{-\infty}^{\infty} px^2\phi(x)\Phi^{p-1}(x)\,dx\end{aligned} \tag{15}$$

and initialize:

$$\mu_1 = 0, \qquad \sigma_1^2 = \sigma_w^2, \qquad \overline{\mu}_1 = A, \qquad \overline{\sigma}_1^2 = \sigma_w^2.$$

2) *Recursion*: For $k = 1, \ldots, K-1$

$$\mu_{k+1} = \mu_k + \sigma_k\mu_{\max}(q) \tag{16}$$

$$\sigma^2_{k+1} = \sigma_w^2 + \sigma_k^2\sigma^2_{\max}(q) \tag{17}$$

$$\overline{\mu}_{k+1} = A + \Phi(\mu_k + \sigma_k\mu_{\max}(q-1)) + \overline{\mu}_k\overline{\Phi} + \theta\phi \tag{18}$$

$$\begin{aligned}\overline{\sigma}^2_{k+1} = {} & \sigma_w^2 + \Phi\sigma_k^2\sigma^2_{\max}(q-1) \\ & + \overline{\sigma}_k\overline{\Phi} + (\Delta\Phi + \theta\phi)(\Delta\overline{\Phi} - \theta\phi)\end{aligned} \tag{19}$$

where

$$\begin{aligned}\Delta &= \mu_k + \sigma_k\mu_{\max}(q-1) - \overline{\mu}_k \\ \theta &= \sqrt{\sigma_k^2\sigma^2_{\max}(q-1) + \overline{\sigma}_k} \\ \phi &= 1/\sqrt{2\pi}\exp\left(-\frac{\Delta^2}{2\theta^2}\right) \\ \Phi &= \int_{-\infty}^{\Delta/\theta} 1/\sqrt{2\pi}\exp\left(-\frac{x^2}{2\sigma_w^2}\right)dx \\ \overline{\Phi} &= 1 - \Phi.\end{aligned}$$

Equations (15)–(19) for obtaining P_{FA} and P_{D} via Gaussian and independence assumptions are complicated and involve expensive numerical integration. This is compared with the straight-forward EVT/GEVT parameter estimation methods presented in the following two sections.

III. EXTREME VALUE THEORY

The occurrence of $\max_{\mathbf{x}_K} I(\mathbf{x}_K)$ functions in (11)–(12) suggests application of EVT. In this section we review a key result in EVT called the Fisher Tippett Theorem. We show in Section IV how this result can be used to compute approximate values of P_{FA} and P_{D}.

A. Fisher Tippett Theorem

In the first instance, suppose $\{y_1, \ldots, y_n\}$ is a sequence of IID random variables with common cumulative distribution function

$$F(x) = \Pr(Y_k \le x), \qquad k = 1, 2 \ldots, n.$$

Let M_n denote the maximum of the sequence, i.e.,

$$M_n = \max\{y_1, \ldots, y_n\}.$$

Classical EVT is concerned with the asymptotic distribution of M_n as $n \to \infty$. If $F(x)$ is known, clearly the distribution function of M_n is given by

$$\begin{aligned}\Pr(M_n \le x) &= \Pr(y_1 \le x, \ldots, y_n \le x), \qquad x \in \mathbb{R} \\ &= F^n(x).\end{aligned}$$

According to EVT, the asymptotic distribution of M_n must belong to one of three possible distributions if it exists. The particular asymptotic distribution can be determined from only limited knowledge of the tail distribution of y_k, thereby making the knowledge of $F(x)$ for all x unnecessary. In particular, for a sequence of IID random variables, $\{y_n\}$, whose probability density function has support $(-\infty, \infty)$, the Fisher–Tippett Theorem [8, p. 121] states that

$$F^n(a_n + b_n x) \xrightarrow{n\to\infty} \Lambda(x) \tag{20}$$

where $\Lambda(x)$ denotes the Gumbel distribution

$$\Lambda(x) \stackrel{\Delta}{=} \exp\{-\exp(-x)\}, \qquad x \in \mathbb{R}$$

and parameters $b_n > 0$ and $a_n \in \mathbb{R}$ are normalizing scale and location constants, respectively. That is, the normalized distribution of M_n converges in distribution (weakly) to the Gumbel distribution.

The approximate distribution of the maxima $F_n(x)$ for sufficiently large finite n is derived from (20),

$$F_n(x) = \exp\{-\exp[-(x - a_n)/b_n]\}. \tag{21}$$

GEVT generalizes the attraction to the extremal distribution (20) by incorporating a shape parameter $\nu > 0$,

$$F^n([a_n^\nu + c_n x]^{1/\nu}) \xrightarrow{n\to\infty} \exp\{-\exp(-x)\}, \qquad x \in \mathbb{R}. \tag{22}$$

Equation (22) holds for all ν providing (20) holds. As sample size n approaches infinity, the sequence of

distributions in (22) does not differ from the sequence in (20), due to the disappearance of higher order terms [16]. From (22), the approximate distribution of the maxima (M_n), denoted as $\overline{F}_n(x)$, for sufficiently large finite n is derived as

$$\overline{F}_n(x) = \exp\{-\exp[-(x^\nu - a_n^\nu)/c_n]\}. \quad (23)$$

It is clear from the recursion (7) that $I(\mathbf{x}_K)$ cannot have a Gaussian density—hence Assumption A1 of Section IIE is clearly untrue. Moreover, the sequence of merit functions at the last frame K, $I(\mathbf{x}_K^{(1)}),\ldots,I(\mathbf{x}_K^{(n)})$ are actually correlated since two states at frame K may have one of more possible previous states in common.

Consider now applying the EVT Fisher Tippett Theorem to the sequence $I(\mathbf{x}_K^{(1)}),\ldots,I(\mathbf{x}_K^{(n)})$. Because the noise $w_k(i,j)$ in (1) is Gaussian, it is clear from (7) that the merit functions $I(\mathbf{x}_K^{(i)})$ have probability densities with support in $(-\infty,\infty)$. Thus with the IID Assumption A2 of Section IIE, the Fisher Tippett Theorem directly applies. In particular, the Gaussianity Assumption A1 used in [7] is not required at all.

EVT also allows the IID Assumption A2 to be relaxed somewhat as follows. Suppose $\{y_1,\ldots,y_n\}$ is a stationary dependent sequence of random variables. EVT for dependent sequences [8, p. 211] states that subject to certain weak distributional mixing conditions (i.e., rapid decay in dependence between successive elements) of $\{y_k\}$, it follows that (20) holds. The weak distributional conditions required are the conditions $D(y_n)$ and $D'(y_n)$ detailed in [8, p. 211]. Verifying these conditions for the sequence $\{I(\mathbf{x}_K^{(1)}),\ldots,I(\mathbf{x}_K^{(n)})\}$ is a difficult task. However, heuristically these conditions seem plausible since the dependency between merit function elements $I(\mathbf{x}_K^{(i)})$ and $I(\mathbf{x}_K^{(j)})$ decreases for $\mathbf{x}_K^{(i)}$ and $\mathbf{x}_K^{(j)}$ spaced further apart on the $L^2 \times M^2$ grid of states at frame K. In Section V we show that the Gumbel distribution yields a remarkably accurate fit to the empirical distribution of the maximum of the sequence $\{I(\mathbf{x}_K^{(1)}),\ldots,I(\mathbf{x}_K^{(n)})\}$. We also note that in [17] it is proven that the maxima of dependent random variables over a square grid, similar to the situation arising here from the TBD problem, will be attracted to one of the three extremal distributions.

For known underlying distributions, the parameters a_n, b_n, and/or c_n and ν of the GEVT distribution (23) can be analytically evaluated. Due to the unknown nature of the merit function distribution in the TBD problem, these parameters must be determined numerically by simulation, or approximated analytically. We demonstrate in Section V that the Gumbel limiting distributions (20)–(22) with parameters determined by simulation, provide far superior distributions of P_{FA} and P_{D} than does the approximate analysis based on a Gaussian approximation (see Section IIE). The GEVT and EVT parameter estimation methods are outlined in Section IV. Section VI then demonstrates the estimation of EVT curve parameters via an approximate analytical method, shown also to provide a far superior fit compared with the approximate Gaussian approach of [7].

IV. EVALUATION OF P_{FA} AND P_{D} VIA EXTREME VALUE THEORY

The false alarm probability (11) can be reexpressed via the tail of the distribution function $F^n(x)$

$$P_{\mathrm{FA}} = 1 - \Pr\left(\max_{\mathbf{x}_K} I(\mathbf{x}_K) \le V_T\right) = 1 - F^n(V_T), \qquad \mathbf{x}_K \in \{\text{noise states}\} \quad (24)$$

where there are of the order of $n \approx L^2 \times M^2$ noise (nonsignal) states at the endpoint K.

In a similar vein, the detection probability is

$$P_{\mathrm{D}} = 1 - \Pr\left(\max_{\mathbf{x}_K} I(\mathbf{x}_K) \le V_T\right) = 1 - F^n(V_T), \qquad \mathbf{x}_K \in \{\text{signal states}\}. \quad (25)$$

For example if $L = 5, M = 6$, there are $n = 25 * 36$ valid signal states, these being the set of cells within 2 cells of the true signal state.

In previous applications of EVT and GEVT, the procedure for estimating the tails of the distributions has been to collect $C = nN$ samples, which are then partitioned into N groups of n samples. The maximum of each of the N groups is selected, producing the sample set of maxima, $\mathbf{X} = [X_1,\ldots,X_N]$, from which the parameters of the distributions (21) and (23) are estimated. The choice of n and N has been subject to much discussion, as n must be large enough to ensure the equalities in (21) and (23) valid, while N must be large enough to ensure the parameter estimates are accurate [11, 15, 18].

For the analysis of the DP algorithm attempted here, the size of the groups n is known when calculating (24)–(25). This leaves only N, the number of simulation repetitions, to be chosen.

To evaluate (24) and (25), estimates of the location, scale and (if using GEVT) shape parameters, a_n, b_n, and ν, must be computed. A description of the least squares (LS) and maximum likelihood (ML) parameter estimation methods to compute the GEVT parameters follows.

Least Squares Parameter Estimation: While the GEVT distribution (22) will hold for all values of ν, providing the EVT distribution is applicable, the accuracy of the approximate distribution (23) depends on the chosen value of ν. Therefore, the parameter

estimation method proceeds by obtaining the optimal ν before estimation of a_n and c_n occurs.

Rearranging (23) produces the linear relationship,

$$-\ln(-\ln(\overline{F}_n(x))) = \frac{x^\nu - a_n^\nu}{c_n}. \tag{26}$$

The array, $Y(\nu) = [\tilde{X}_1^\nu, \ldots, \tilde{X}_N^\nu]$ is constructed for the ordered sequence of sample maxima $\tilde{\mathbf{X}}$ and exponent value ν. For an array of equi-spaced points between 1 and N, $U = [u_1, \ldots, u_N]$, evaluate the left-hand side of (26). Denote the resulting array $Z = [Z_1, \ldots, Z_N]$ where $Z_i = \ln(-\ln(\overline{F}_n(u_i))$. The correlation between Z and $Y(\nu)$,

$$f(\nu) = \frac{\sum_{i=1}^N (Y_i(\nu) - \overline{Y}_\nu)(Z_i - \overline{Z})}{\left[\sum_{i=1}^N (Y_i(\nu) - \overline{Y}_\nu)^2 \sum_{i=1}^N (Z_i - \overline{Z})^2\right]^{1/2}} \tag{27}$$

will be maximized at some $\hat{\nu}$. In (27), $\overline{Z}$ and $\overline{Y}_\nu$ are the array means.

With the optimal value $\hat{\nu}$ in hand, the parameters a_n and c_n can be estimated by application of least squares (LS) fit to the linear relationship (26). The LS objective is

$$\text{minimize} \left(Z - \frac{1}{c_n} Y(\hat{\nu}) - \frac{d_n}{c_n}\mathbf{1}\right)' \left(Z - \frac{1}{c_n} Y(\hat{\nu}) - \frac{d_n}{c_n}\mathbf{1}\right) \tag{28}$$

where $d_n = a_n^{\hat{\nu}}$, and $\mathbf{1}$ is a vector of ones of length N. Taking the partial derivatives of (28) with respect to d_n and c_n, and solving for these parameters results in estimates,

$$\hat{c}_n^{\text{LS}} = \frac{Y(\hat{\nu})' Y(\hat{\nu}) - N\overline{Y}_{\hat{\nu}}^2}{Z' Y(\hat{\nu}) - N\overline{Y}_{\hat{\nu}}\overline{Z}} \tag{29}$$

$$\hat{d}_n^{\text{LS}} = \overline{Y}_{\hat{\nu}} - \hat{c}_n^{\text{LS}}\overline{Z}. \tag{30}$$

The LS EVT parameter estimates results from taking $\hat{\nu} = 1$ in (26)–(30).

Maximum Likelihood Parameter Estimation: While the LS parameter estimation method is constructed to estimate the three parameters of the GEVT distribution (23), we use the ML method to estimate only the parameters of the EVT distribution (21), due to inherent numerical instability in the ML method when attempting to solve iteratively for the three GEVT parameters. The numerical simulations of Section V demonstrate that the performance of the LS GEVT parameter estimator is high enough not to require an alternative ML estimator.

The log likelihood function of $\mathbf{X}$ under EVT distribution (21),

$$L(a_n, b_n, \mathbf{X}) = -N\ln b_n - \sum_{i=1}^N \frac{X_i - a_n}{b_n} - \sum_{i=1}^N \exp\left\{-\frac{X_i - a_n}{b_n}\right\} \tag{31}$$

is differentiated with respect to the EVT parameters b_n and a_n. The resulting partial derivatives are set to zero and rearranged, producing the relations,

$$\hat{b}_n = \overline{\mathbf{X}} - \frac{\sum_{i=1}^N X_i \exp\left\{-\frac{X_i}{\hat{b}_n}\right\}}{\sum_{i=1}^N \exp\left\{-\frac{X_i}{\hat{b}_n}\right\}} \tag{32}$$

$$\hat{a}_n = \hat{b}_n \ln\left[\frac{N}{\sum_{i=1}^N \exp\left\{\frac{X_i}{\hat{b}_n}\right\}}\right]. \tag{33}$$

The equations (32)–(33) can be iteratively solved for optimal values $\hat{a}_n^{\text{ML}}$ and $\hat{b}_n^{\text{ML}}$. Such an iterative procedure is described in [19, p. 231–2], where the lth estimate $\hat{b}_n^{(l)}$ is computed by (32) and is then averaged according to $\hat{b}_n^{l+1} = \hat{b}_n^{(l)} + (\hat{b}_n^{(l-1)} - \hat{b}_n^{(l)})/3$, prior to computing the next value of $\hat{a}_n$. In the numerical simulations in Section V, initializing with $\hat{b}_n^{(1)} = 1$ was found to be completely adequate and avoided unnecessary computation introduced in the procedure in [19].

V. NUMERICAL EXAMPLES

Computer simulations of the TBD problem are carried out to demonstrate the superiority of the EVT and GEVT analysis of Section IV as compared with the approximate analysis in [7], which was briefly described in Section IIE.

The simulated TBD scenario is as follows. The measurement sensor has side length of $L = 30$ cells, and the model considers $M = 6$ velocity cells, a total number of $L^2 \times M^2 = 32400$ discrete states at each time k. The initial target state is chosen randomly with $3.0 \le x_1, y_1 \le 6.0$ and $1.5 \le u_1, v_1 \le 2.5$. The signal amplitude A is varied between sets of simulations, as is the transition size parameter q. The noise is kept at the constant level of $\sigma_w^2 = 0.8$. $N = 1000$ independent simulation runs are performed for each parameter set, and estimates of P_{FA} and P_{D} are obtained by 1) a classical counting procedure, 2) the approximate Gaussian analysis [7], and 3) EVT/GEVT analysis.

Fig. 2 shows the LS GEVT, EVT (ML and LS), and Gaussian approximation curves plotted with the curves based on the classical counting procedure (henceforth called the "simulated" curves), for the probability of false alarm. The simulations carried out are for no signal present ($A = 0$), and hence a false alarm is declared if the maximum end-state merit function $I(\mathbf{x}_K)$ exceeds the threshold. The GEVT curves in Fig. 2 provide a remarkably accurate fit. They are nearly indistinguishable from the simulated curves. In comparison the Gaussian approximation from [7] underestimates the P_{FA} in Fig. 2, with an inaccurate distributional shape across the range of

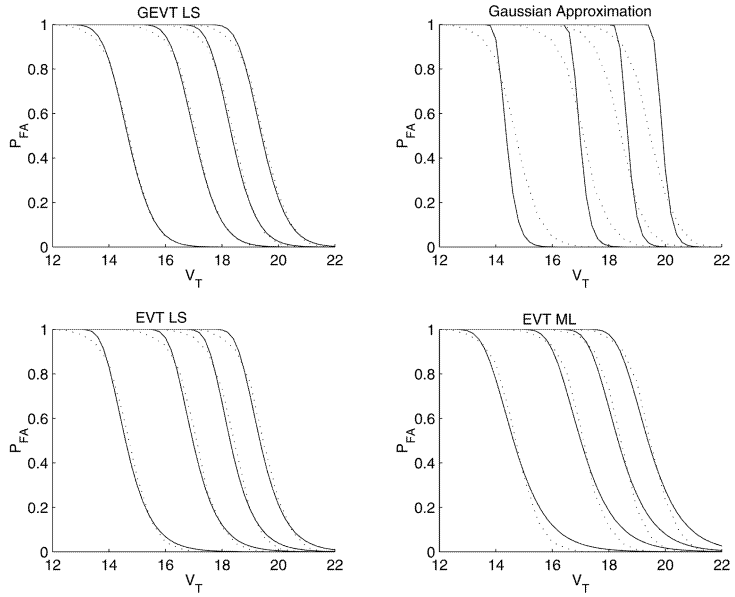


Fig. 2. P_{FA} curves for methods compared in Section V ($\cdots$ actual P_{FA}, ——— estimated curve). Within each plot, curves are left-to-right $q = \{4, 9, 16, 25\}$.

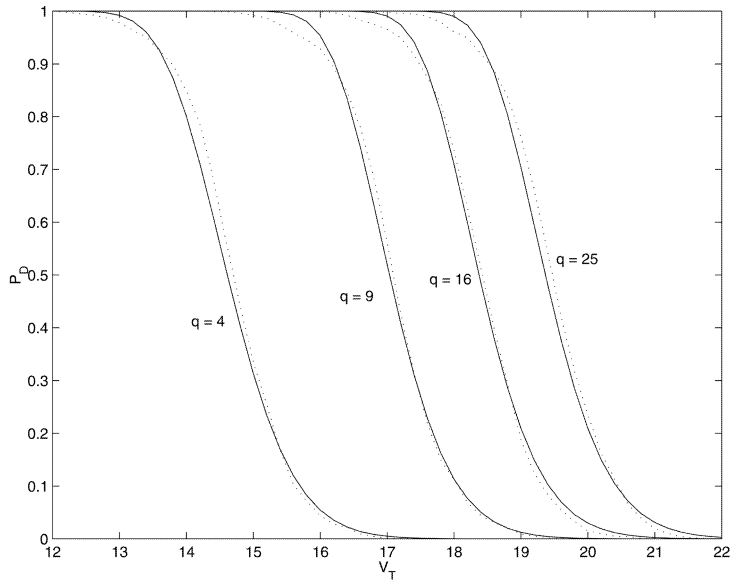


Fig. 3. P_{FA} curves obtained by LS estimates of GEVT parameters, using reduced sample size $N = 100$. ($\cdots$ actual P_{FA}, ——— GEVT-based estimate of P_{FA}).

q values. Neither set of EVT curves (LS and ML) in Fig. 2 is as accurate as the GEVT, but each is significantly superior to the Gaussian approximation curves.

To demonstrate the usefulness of the extremal analysis, Fig. 3 presents the LS GEVT curves for P_{FA}, with parameter estimates taken from only $N = 100$ sample maxima, rather than $N = 1000$ as in the previous plots. The accuracy of the GEVT curves is still better than even that obtained by the EVT curves in Fig. 2.

Figs. 4–7 compare the simulated P_{D} curves with GEVT, EVT (ML and LS), and Gaussian approximation curves for $q = 1, 4, 9, 16$. The accuracy of the GEVT Gumbel distributions across the entire range of signal amplitudes is evident from these figures. The EVT based plots are less accurate, but still provide a good estimate of the distributional shape, particularly at low signal amplitude A.

In comparison, the Gaussian approximation analysis of [7] in these figures provide a close distributional estimate only for large A (high SNR). This is to be expected, given that in high SNR the DP is redundant and the problem becomes one of choosing a strong signal from a noisy region, the tail performance of which is known to approach a Q

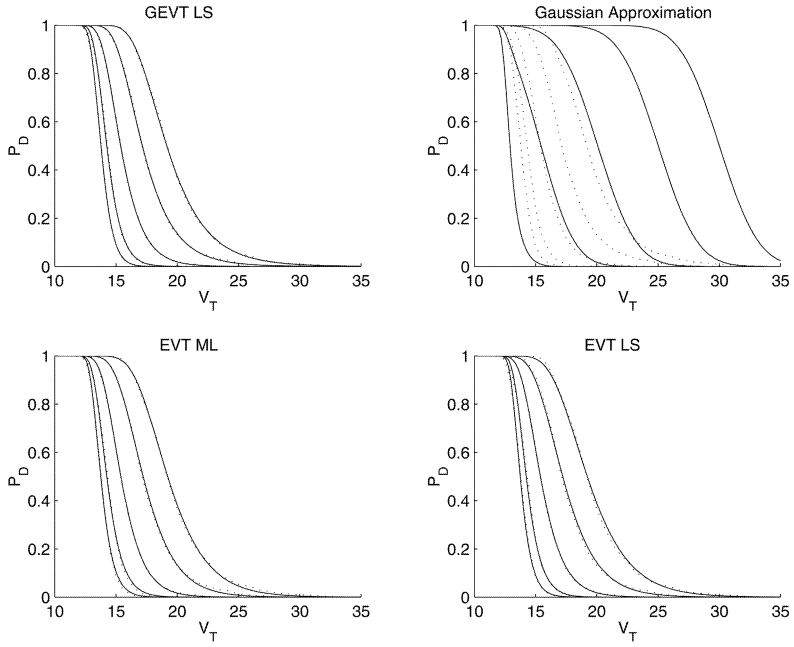


Fig. 4. P_D curves for methods compared in Section V and valid transition parameter, $q = 4$. ($\cdots$ actual P_D, ——— estimated curve). Within each plot, curves are left-to-right $A = \{1.0, 1.5, 2.0, 2.5, 3.0\}$.

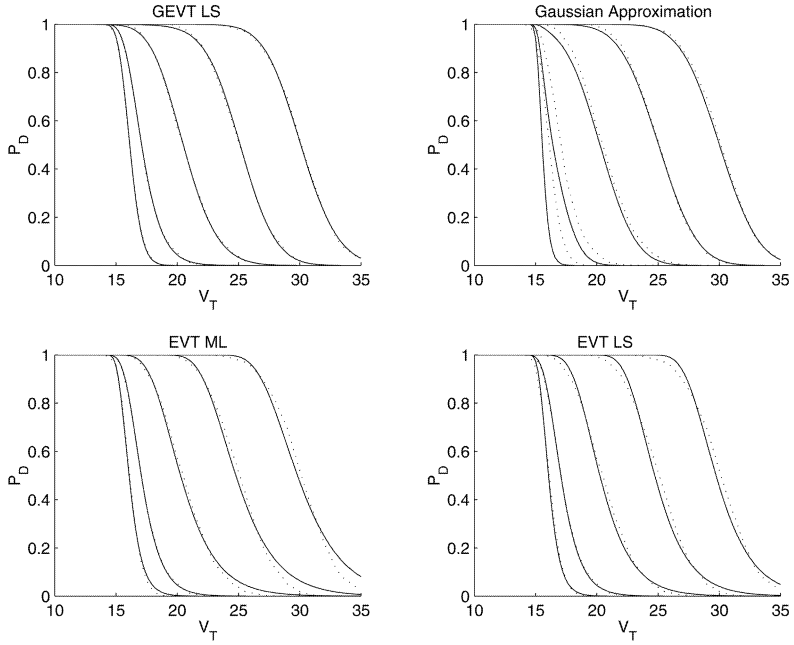


Fig. 5. P_D curves for methods compared in Section V and valid transition parameter, $q = 9$. ($\cdots$ actual P_D, ——— estimated curve). Within each plot, curves are left-to-right $A = \{1.0, 1.5, 2.0, 2.5, 3.0\}$.

function. Given that the TBD method is proposed as a method to be used when SNR is low (see Section I), the Gaussian approximation analysis is clearly not sufficient, while EVT and GEVT analyses provide results of high accuracy.

As with the probability of false alarm above, the GEVT curves can be estimated using a fraction of the sampled data, $N = 100$ rather than $N = 1000$. The results for the four q values across the range of signal amplitudes are plotted in Fig. 8. The curves remain accurate, particularly at low levels of A, where the Gaussian approximation in particular fails.

VI. ANALYTICAL DETERMINATION OF EXTREME VALUE THEORY CURVE PARAMETERS

Thus far, we have demonstrated that by using EVT, one can obtain performance curves that fit the actual performance of the TBD algorithm extremely accurately. In particular, we presented simulation

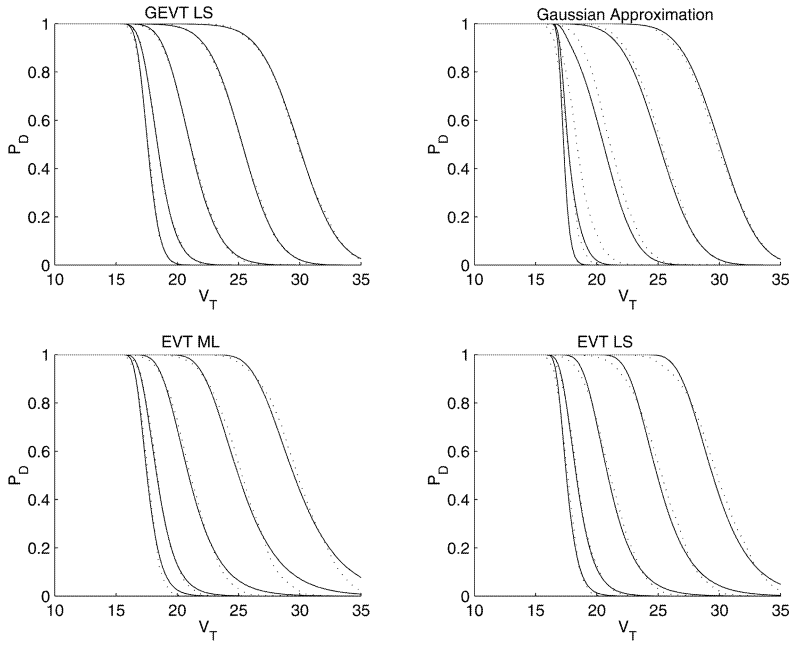


Fig. 6. P_D curves for methods compared in Section V and valid transition parameter, $q = 16$. ($\cdots$ actual P_D, ——— estimated curve). Within each plot, curves are left-to-right $A = \{1.0, 1.5, 2.0, 2.5, 3.0\}$.

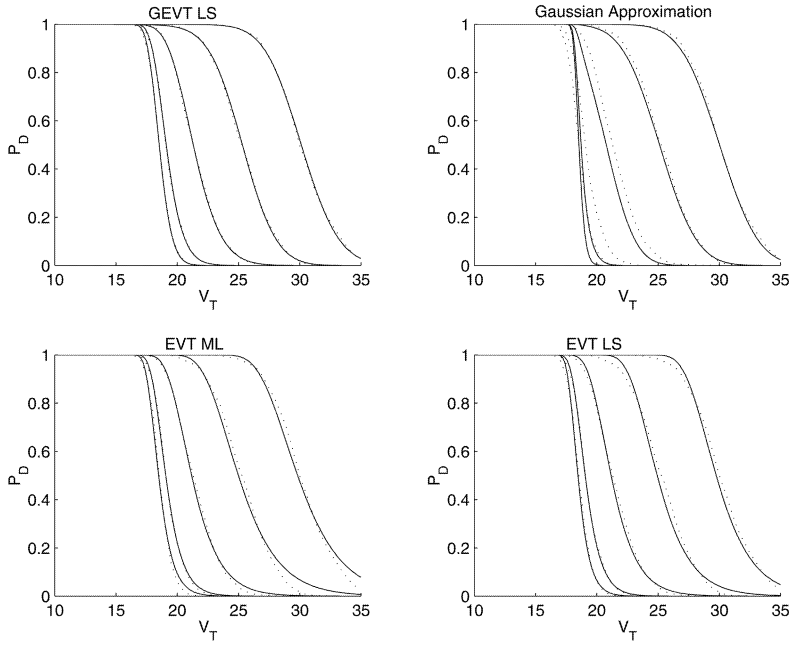


Fig. 7. P_D curves for methods compared in Section V and valid transition parameter, $q = 25$. ($\cdots$ actual P_D, ——— estimated curve). Within each plot, curves are left-to-right $A = \{1.0, 1.5, 2.0, 2.5, 3.0\}$.

based numerical methods to estimate the curve parameters. In this section, by considering the TBD problem to be analogous to the determination of a **dominating path** in a stochastic network, we derive approximate analytical expressions for the curve parameters. These expressions readily yield accurate analytical performance bounds (e.g. probability of false alarm or probability of detection) of the TBD algorithm as a function of the target parameters (target aamplitude A, noise variance σ_w^2, valid transition horizon q).

We consider the analytical estimation of only the EVT curve parameters, as the exponent in the GEVT case (see (22)) renders the task too difficult.

Consider first the network view of the TBD problem, briefly discussed in Section I. At each time frame in the TBD algorithm, there are $L^2 \times M^2$ possible cells in which the target may reside. Thus the

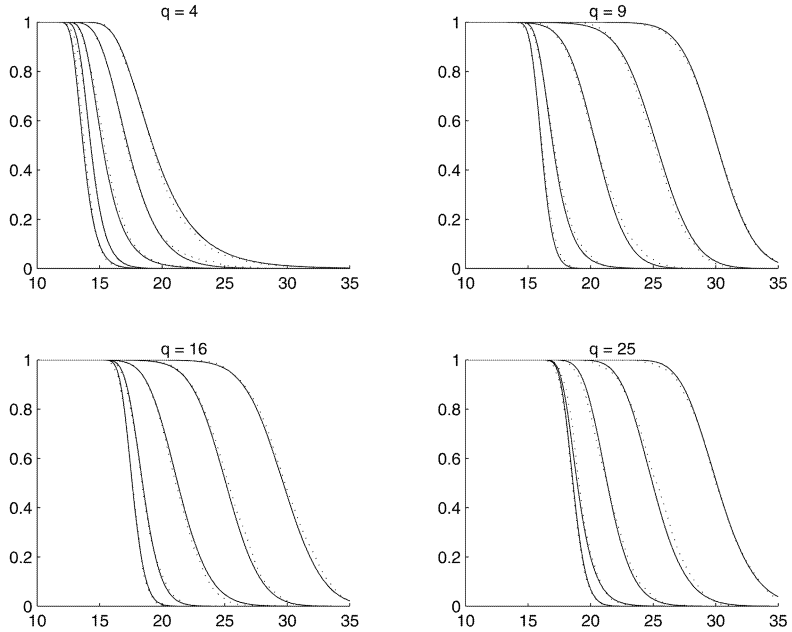


Fig. 8. P_D curves obtained by LS estimates of GEVT parameters, using reduced sample size $N = 100$. ($\cdots$ actual P_D, ——— GEVT-based estimate of P_D). Within each plot, curves are left-to-right $A = \{1.0, 1.5, 2.0, 2.5, 3.0\}$.

target maps out a path as it traverses the K frames. The task of the TBD algorithm is to determine which of the paths in this highly connected, stochastic network exceeds a given threshold value (and hence is thought to have originated from the target).

The DP algorithm results in $L^2 \times M^2$ potential target paths. On each branch (i.e., arc or frame, $k = 1, \ldots, K$) is a measurement that is either $N(0, \sigma_w^2)$ or $N(A, \sigma_w^2)$ according to the signal model (1).

As presented in [9], EVT can be applied to stochastic networks using the concept of a **dominating path**. A dominating path is one that dominates a network. An example is given in [9] that Kleindorfer's network, consisting of 20 nodes, 38 arcs and 51 paths, has 3 paths that are 9 arcs long, longer than the other paths. Therefore $n = 3$ paths are said to be *dominating paths*.

If each dominating path has mean, μ, and variance, σ^2, then the EVT parameters can be approximately determined by the following relationships (refer to [9] for details on derivation):

$$a_n = \mu + \sigma \left[(2 \log n)^{1/2} - \frac{1}{2} \frac{(\log \log n + \log 4\pi)}{(2 \log n)^{1/2}} \right] \tag{34}$$

$$b_n = \frac{(2 \log n)^{1/2}}{\sigma}. \tag{35}$$

This theory relies on the assumption that the network paths are IID. Obviously this is not the case in any nontrivial (nonparallel) network, as certain branches are common to multiple paths. Thus the above method assumes that the performance degradation incurred by violating the IID assumption is not too severe.

We proceed now to apply the concept of dominating paths to the two performance measures of interest in this work. The numerical data used to test the method's accuracy arise from the simulations presented in Section V.

A. False Alarm Probability

As the probability of false alarm curves are determined via simulation with no signal present ($A = 0$) (see Section V), no path can be said to dominate the network. Each path is K frames/branches long, with a measurement on each branch that is $N(0, \sigma_w^2)$. It is a natural application of the method (34)–(35) above, therefore, to consider $\mu = 0$, $\sigma^2 = K\sigma_w^2$, and $n = L^2M^2$. Most problematic with this is that it does not take into account the parameter q. This can be rectified by noting that in the Kth step of the DP algorithm, sets of q measurements will be tracked back to the same cell in the $(K-1)$th frame, due to the max operation. Therefore an initial attempt with $n = (L^2M^2/q)$ as the number of dominating paths in the network is justified. With these three parameters chosen, a_n and b_n can be determined from (34)–(35), fully determining the desired EVT curve.

EXAMPLE Consider the case where $q = 16$, $L = 30$, $M = 6$. Fig. 9 shows the P_{FA} curve obtained via empirical estimation from simulations and the curve using the above analytical method. The latter curve has the **correct shape**, but a completely **incorrect shift**.

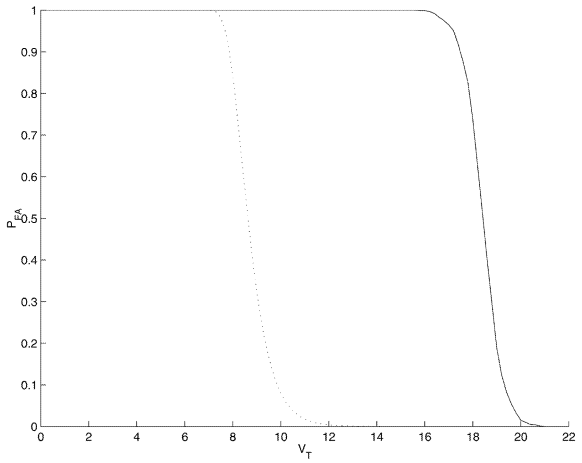


Fig. 9. P_{FA} curve obtained via unshifted analytical EVT method of Section VIA. (——— actual P_{FA}, $\cdots$ unshifted analytical curve).

As the curve shape is correct, the parameters σ and n are good estimates of the true parameters (see (35) above). The incorrect shift is explained by the approximation made in assuming that there are $n = 32400/q$ IID paths in the network. Missing is the upwards shift in the mean incurred at each stage of the algorithm via the max operations (see (31)–(32) for a mathematical demonstration of this shift).

Values for the shift in mean for each q are obtained using a Gaussian approximation as follows. At each time $k = 2,\ldots,K$, a max operation is applied in the DP algorithm (see (8)). The recursion for the shifting mean and variance obtained via the Gaussianity approximation is given in (31)–(32) in Section V, repeated here for clarity:

$$\mu_k = \mu_{k-1} + \sigma_{k-1}\mu_{\max}(q) \tag{36}$$

$$\sigma_k^2 = \sigma_w^2 + \sigma_{k-1}^2\sigma_{\max}^2(q) \tag{37}$$

for $k = 2,\ldots,K-1$, with $\mu_1 = 0$ and $\sigma_1^2 = \sigma_w^2$. Here we computed the recursion until frame $K-1$ not K, as the Kth maximum operation is taken into account in the choice, $n = 32400/q$. The recursion produces shifts, defined as $\mu_{\mathrm{shift}}(q) = \mu_{K-1}$ for each q. The shift, $\mu_{\mathrm{shift}}(q)$, is then incorporated into the analytic method from determining the curve parameters, by setting $\mu = \mu_{\mathrm{shift}}(q)$ in (34).

Fig. 10 shows the result of the analytic parameter estimation that incorporates the shift in mean due to the max operations. For $q = 4$, the shift overcompensates for the max operations. The use of $q = 4$ has already been discredited as being far too restrictive on the input signal model (see Section IIC). It is observed in these plots that as q increases, the agreement between the analytic and simulated EVT curves improves greatly.

B. Detection Probability

Determining the probability of detection curves analytical method is a more difficult task (than the probability of false alarm), owing to the fact that some branches in the network are distributed according to $N(0,\sigma_w^2)$, while others are $N(A,\sigma_w^2)$. The results presented here do not provide a completely analytical method for P_{D} parameter estimation. What is shown, however, is that the parameters can be estimated through comparison to the simulated curves, which sheds light onto the performance analysis task at hand.

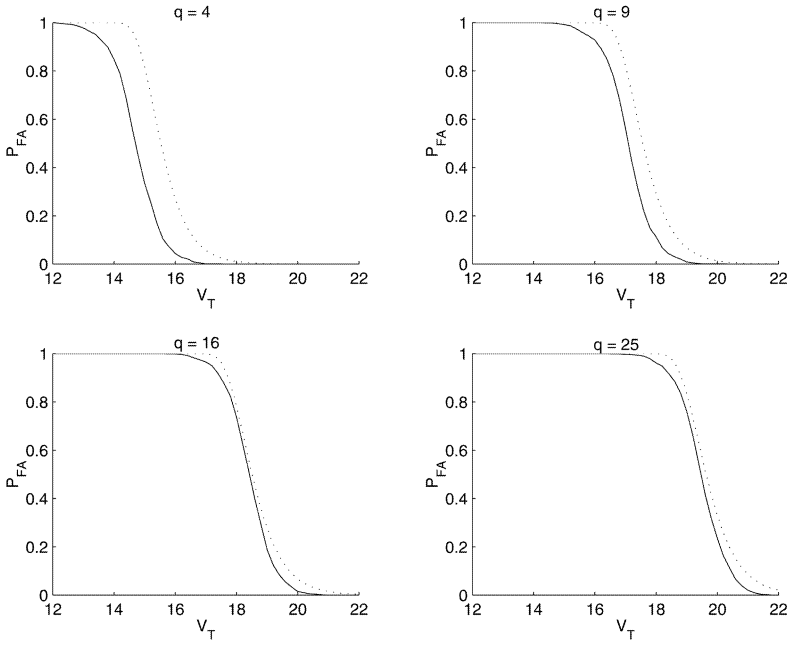


Fig. 10. P_{FA} curves via analytical EVT method (——— actual P_{FA}, $\cdots$ shifted analytical curve).

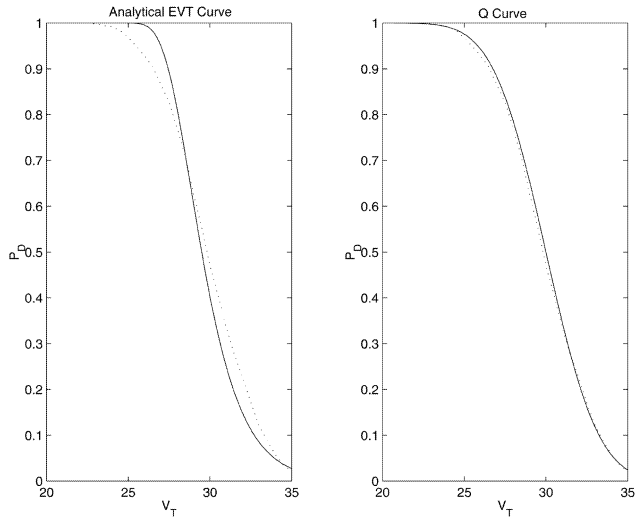


Fig. 11. P_D curve comparison for $q = 16$, $A = 3.0$, between analytical EVT method of Section VIB and standard Q curve. ($\cdots$ actual P_D, ——— estimated curve.)

Signal detection via the TBD algorithm can be broken into two classes, those of detection in high A (high SNR) and low A (low SNR). In the latter category, the signal is more likely to be masked by the noise. In the former category, should A be high enough, the DP algorithm is rendered unnecessary, as the task reduces to selecting a strong Gaussian signal in white noise.

To demonstrate this fact, the case $q = 16$ and $A = 3.0$ (high SNR, given that $\sigma_w^2 = 0.64$) is considered. Fig. 11 demonstrates that simply fitting a Q function (erf function) with mean $A*K = 30$ and variance $K*\sigma_w^2 = 6.4$, estimates the empirical curve extremely well. The analytic EVT curves (discussed later) fits quite well, but not to the same degree as the Q function curves. Similar results were noticed for other q values. Thus in high SNR, a single path dominates the network, rendering EVT both unnecessary and unwarranted.

In the low SNR region, the network is not dominated by only a few paths. Two low signal amplitudes, $A = 1.5$ and $A = 1.0$, were considered. For the former case, it was found that $n = 25$ paths dominate. Practically this means that each position in the 25-position-cell region being searched for exceedances of V_T (see (12)) is liable to contain a dominating path. This path may not hold the true signal, but because of the low SNR the noise may give the appearance of a signal existing there. For the weaker $A = 1.0$, even more paths dominate the network. Indeed, all $25 \times M^2 = 25*36$ paths (region including the velocity cells) are liable to contain a dominating path. Together with $\sigma^2 = K\sigma_w^2$, this consistently gave the true shape of the EVT curve.

The harder task is determining the mean, μ, of the dominating paths in the network. Table I displays the mean value of μ. It is not possible to formulate an expression to obtain these values without comparison with the simulated curves. The $q = 9$ results are explained by the fact that only $(K-2)$ frames of the dominating paths can be considered to be definitely signal states, and in the last two frames, on average only one of the cells contained a measurement that looked like the signal. The $q = 25$, $A = 1.0$ value for μ is harder to explain, as it is actually greater than $K \times A$! However, one can view this as the noise completely dominating the paths, due to the flexibility in $q = 25$ for the paths to wander widely to achieve the highest merit function value.

TABLE I
Mean Values, μ, for Low SNR Signals

	$A = 1.0$	$A = 1.5$
$q = 9$	$(K-2)A + 2A/25$	$(K-2)A + 2A/25$
$q = 16$	$(K-1)A + A/2$	$(K-1)A + A/25$
$q = 25$	$KA + A/2$	$(K-1)A + A/2$

Fig. 12 demonstrates the high level of accuracy achieved semi-analytically using the parameter values from Table I in the low SNR region of operation of the TBD algorithm.

C. Summary of Analytical Evaluation of Extremal Curve Parameters

Obtain **probability of false alarm curve** given the number of valid state transitions q and the variance of the sensor background noise σ_w^2.

1) Set $n = L^2M^2/q$ and $\sigma^2 = K\sigma_w^2$, where L is the side length of the sensor, M the number of velocity cells, and K the number of processed frames.
2) Determine shift in mean, μ, by evaluating (36)–(37) for $k = 2, \ldots, K-1$.
3) Evaluate a_n and b_n in (34)–(35) using n, σ, μ.
4) The EVT curve is given by substitution of a_n and b_n into (21).

Obtain **probability of detection curve** given values of the target intensity A, and q and σ_w^2 as above.

1) For $A = 1.0$, set $n = 25$. For $A = 1.5$, set $n = 25*M^2$.
2) Set $\sigma^2 = K\sigma_w^2$.
3) Determine shift in mean, μ, by reference to Table I.
4) Evaluate a_n and b_n in (34)–(35) using n, σ, μ.
5) The EVT curve is given by substitution of a_n and b_n into (21).

These semi-analytic results can be used in TBD system design. With desired probabilities of detection and false alarm specified, the appropriate values of transition size parameter q can be found across the range of SNR values the system is likely to operate within via reference to the appropriate graphs. Other

low SNR values to the ones considered here may be taken into account via interpolation of the given results.

VII. CONCLUSION

The analysis of the TBD problem has been approached using EVT and GEVT, resulting in significantly more accurate results than previous approaches that were based on assumptions of independence and Gaussianity [6, 7]. This novel application of EVT/GEVT to DP algorithm performance demonstrates the power of distributional convergence theory, given that DP introduces dependencies in the sequences of random variables not commonly considered in the literature. Work is continuing in the investigation of EVT applied to the theoretical analysis of low SNR uses of DP (Viterbi) algorithms.

REFERENCES

[1] Hagenauer, J., Offer, E., and Papke, L. (1996)
Iterative decoding of binary block and convolutional codes.
IEEE Transactions on Information Theory, **42** (Mar. 1996), 429–445.

[2] Reed, I. S., Gagliardi, R. M., and Stotts, L. B. (1988)
Optical moving target detection with 3-D matched filtering.
IEEE Transactions on Aerospace and Electronic Systems, **24** (July 1988), 327–336.

[3] Reed, I. S., Gagliardi, R. M., and Stotts, L. B. (1990)
A recursive moving-target-indication algorithm for optical image sequences.
IEEE Transactions on Aerospace and Electronic Systems, **26** (May 1990), 434–440.

[4] Arnold, J., Shaw, S., and Pasternack, H. (1993)
Efficient target tracking using dynamic programming.
IEEE Transactions on Aerospace and Electronic Systems, **29** (Jan. 1993), 44–56.

[5] Barniv, Y. (1985)
Dynamic programming solution for detecting dim moving targets.
IEEE Transactions on Aerospace and Electronic Systems, **21** (Jan. 1985), 144–156.

[6] Barniv, Y., and Kella, O. (1987)
Dynamic programming solution for detecting dim moving targets, Part II: Analysis.
IEEE Transactions on Aerospace and Electronic Systems, **23** (Nov. 1987), 776–788.

[7] Tonissen, S. M., and Evans, R. J. (1996)
Performance of dynamic programming techniques for track-before-detect.
IEEE Transactions on Aerospace and Electronic Systems, **32** (Oct. 1996), 1440–1451.

[8] Embrechts, P., Kluppelberg, C., and Mikosch, T. (1997)
Modeling External Events for Insurance and Finance.
Berlin Heidelberg: Springer-Verlag, 1997.

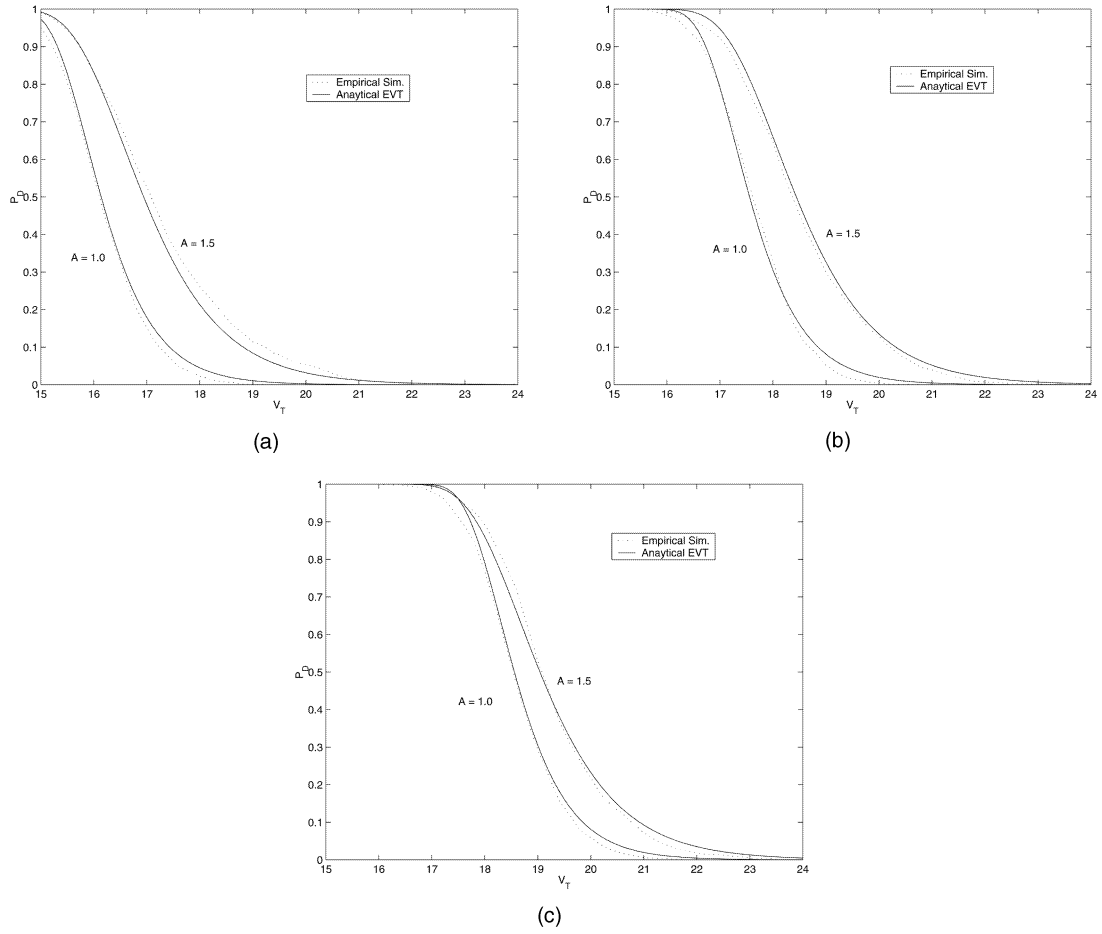


Fig. 12. P_D curves obtained by analytical EVT method of Section VIB. (a) $q = 4$. (b) $q = 16$. (c) $q = 25$.

[9] Dodin, B., and Sirvanci, M. (1990)
Stochastic networks and the extreme value distribution.
Computers in Operations Research, **17**, 4 (1990), 397–409.

[10] Galambos, J., Lechner, J., and Simui, E. (Eds.) (1993)
In *Proceedings of the Conference on Extreme Value Theory and Applications*.
Gaithersburg, MD, 1993.

[11] Jeruchim, M. C. (1984)
Techniques for estimating the bit error rate in the simulation of digital communication systems.
IEEE Journal on Selected Areas in Communications, **2** (Jan. 1984), 153–170.

[12] Choe, J., and Shroff, N. B. (1998)
A central-limit-theorem-based approach for analyzing queue behavior in high-speed networks.
IEEE/ACM Transactions on Networking, **6** (Oct. 1998), 659–671.

[13] Tsihrintzis, G. A., and Nikias, C. L. (1996)
Fast estimation of the parameters of alpha-stable impulsive interference.
IEEE Transactions on Signal Processing, **44** (June 1996), 1492–1503.

[14] Dogancay, K., and Krishnamurthy, V. (2000)
Application of extreme value theory to level estimation in nonlinearly distorted hidden Markov models.
IEEE Transactions on Signal Processing (2000).

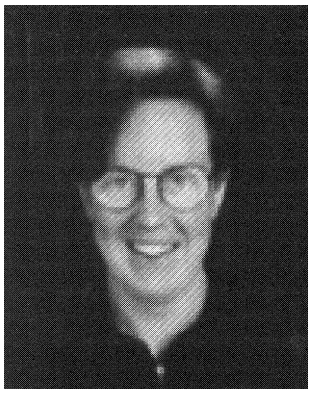
[15] Guida, M., Iovino, D., and Longo, M. (1988)
Comparative performance analysis of some extrapolative estimators of probability tails.
IEEE Journal on Selected Areas in Communications, **6** (Jan. 1988), 76–84.

[16] Weinstein, S. B. (1973)
Theory and application of some classical and generalized asymptotic distributions of extreme values.
IEEE Transactions on Information Theory, **19** (Mar. 1973), 148–154.

[17] Galambos, J. (1978)
The Asymptotic Theory of Extreme Order Statistics.
Series in Probability and Mathematical Statistics.
New York: Wiley, 1978.

[18] Bernabei, F., Listanti, M., Matrullo, D., and Zingrillo, G. (1992)
A combined montecarlo/gevt extrapolating method for the estimate of buffer length distribution tails.
In *Proceedings of the 1992 International Conference on Communications* (1992), 1206–1211.

[19] Gumbel, E. J. (1958)
Statistics of Extremes.
New York: Columbia University Press, 1958.



Leigh A. Johnston received B.Sc. and B.E. (Elec.) degrees from the University of Melbourne, Australia in 1997, and a Ph.D. from the same institution in 2000.

Upon completion of the Ph.D. degree, she spent a year as a research fellow in the Centre for Systems Engineering and Applied Mechanics, at the Université catholique de Louvain, Belgium. Her interests lie in network optimisation and iterative estimation methods, currently the focus of her research activities at the University of Melbourne.

Vikram Krishnamurthy (S'90—M'91—SM'99) was born in 1966. He received his bachelor's degree in electrical engineering from the University of Auckland, New Zealand in 1988, and doctoral degree from the Australian National University, Canberra in 1992.

He is currently a Professor at the Department of Electrical Engineering, University of Melbourne, Australia and also serves as deputy head of department. His research interests span several areas including stochastic scheduling and network optimization, time-series analysis, and statistical signal processing.

Dr. Krishnamurthy is currently an associate editor for *IEEE Transactions on Signal Processing* and *Systems and Control Letters*. He has served on the technical program committee of several conferences including the 37th IEEE Conference on Decision and Control, Tampa, FL, 1998 and IFAC Symposium on System Identification in 2000 (SYSID'2000), Santa Barbara, CA.