

# Non-parametric Maximum Likelihood Estimation of Pulsed Interference on Wireless Links

Douglas J. Leith, Brad Zarikoff  
Hamilton Institute, NUI Maynooth, Ireland

**Abstract**—In this paper we consider a wireless link subject to on-off interference. We show that by using different size packets and by measuring the packet error rate as a function of packet transmit duration on the air, the distribution of the off times between interference bursts can be accurately estimated. In the context of white space usage, this directly corresponds to estimation of the time duration of the spectrum hole usable by a secondary station. We present simulation results demonstrating the effectiveness of the proposed technique for a range of interference patterns, and illustrate its ability to distinguish between on-off interference and channel noise and provide quantitative measures of both. By being able to carry out sensing without requiring the transmitter to remain silent, this creates the potential for significant throughput gains.

## I. INTRODUCTION

There is currently a strong industrial and governmental push to devise intelligent methods for reusing lightly-used spectrum allocations. Such spectrum is referred to as “white space”, where the licensed user of the spectral band is referred to as the primary user (PU). The PU does not occupy its allocated spectrum in all geographic or temporal space, and so leaves room for spectrum reuse. Evidence of the potential gain for efficient cognitive radio solutions are seen in recent measurement studies [1], [2], which identify significant spectrum holes even in urban areas one would suspect to be moderately active. Cognitive radio techniques that aim to capitalize on white space are commonly referred to as dynamic spectrum access (DSA), while the DSA radios are referred to as secondary users (SUs).

In order for DSA techniques to function, advances in physical, media access and network layers are required [3]. Specifically, reliable methods for spectrum sensing in the physical layer and sensing schedules in the MAC layer must be capable of 1) providing an advantage to utilizing the white space for the SUs in terms of standard metrics of throughput and delay, and 2) operating in such

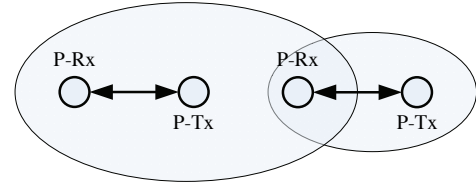


Fig. 1. The primary user (PU) link acts as a “hidden node” to the secondary user (SU), where transmissions from the SU are unheard at the PU receiver. The task we consider is for the SU is to distinguish between channel noise and the pulsed interference generated by the PU, and to do this without requiring the SU to remain silent.

a manner that the PU is unaware of the presence of DSA. In terms of the MAC layer, a balance must be struck between the time allocated to spectrum sensing and access.

This work presents a step in the direction of joint spectrum sensing and access. In this way, we can optimize the inefficiency resulting from spectrum sensing, during which no data transmission can occur. We propose that by using different size packets and by measuring the packet error rate as a function of packet transmit duration on the air, the distribution of the off times between interference bursts can be accurately estimated. Essentially, we exploit the fact that as the packet transmit duration increases, the likelihood of it overlapping with an interference burst increases in a way that is dependent on the off times between interference bursts. Note that these off times directly correspond to the time duration of the spectrum hole useable by the SU. By being able to carry out sensing without requiring the transmitter to remain silent, this creates the potential for significant throughput gains.

We consider a SU transmitter and receiver operating in the PU-induced interference-limited region. The SU transmit power and the distance between the SU transmitter and PU receiver are such that the SUs cannot impact the performance on the PU link (Fig. 1). However, the PU transmitter is of strong enough power to induce packet loss at the SU receiver.

## II. RELATED WORK

This paper builds on the recent work in [4], [5] that proposes methods for separately identifying the sources of packet errors in 802.11 links in unlicensed bands. There also exists a relevant body of work in the statistics literature relating to the estimation of renewal process parameters and residual lifetimes. The seminal work in [6] uses a similar setup to that analyzed in Section III here, but with the important difference that direct observations of the *on* and *off* times within each measurement window are assumed. The approach in [6] is extended in [7] to situations where observations are restricted to the number of *on* and *off* periods during each window (and so there are no direct observations of the durations of the *on* and *off* periods). In the present paper we further extend this line of work by imposing the still stronger restriction to binary observations; namely, where we can only observe whether there was no *on* period or whether one or more *on* periods occurred during each measurement window. It is this extension which allows us to make the connection with packet loss rate measurements.

## III. ESTIMATION USING WINDOWED DATA

We begin by considering a setup consisting of an ensemble of  $K$  identical on-off processes with iid *off* times and *on* times. We shall assume, for the moment, that the *on* times are of zero duration but will relax this later. Let the *off* times  $\Delta$  have distribution function  $\text{Prob}[\Delta \leq x] = F(x)$ . Also assume, for the moment, that the *off*-times can only take one of a finite set of values  $\mathcal{X} = \{x_j, j = 1, 2, \dots, n\}$  and let  $\text{Prob}[\Delta = x_j] := f_j$ . We will assume, without loss, that the  $x_i$  are distinct and sorted into ascending order so that  $x_i < x_{i+1}$  with  $0 < x_1$ . Suppose we observe each process over fixed time windows  $[0, x_j]$ ,  $x_j \in \mathcal{X}$  and record a 1 if one or more *on* events occur during the window and 0 otherwise. Let  $N(x_j)$  denote the number of 1's recorded for the window of duration  $x_j$ . Our task is to estimate the  $f_j$ ,  $j = 1, 2, \dots, n$  using this data; alternatively, given a set of intervals  $\mathcal{X}$  we can think of our task as being to estimate  $F(x)$  at sample points  $x \in \mathcal{X}$  (this is discussed in more detail in Section III-E).

### A. Likelihood Function

The likelihood of the observed count  $N(x_j)$  is given by

$$\begin{aligned} L(N(x_j)|K, \mathcal{X}, f) \\ = \binom{K}{N(x_j)} p^{N(x_j)}(x_j)(1 - p(x_j))^{K - N(x_j)} \quad (1) \end{aligned}$$

where  $f$  denote the vector with elements  $f_i$  and  $p(x_j)$  is the probability of observing an *on* event in the window and we have

$$p(x_j) = 1 - \frac{\sum_{i: x_i > x_j} (x_i - x_j) f_i}{\sum_{i=1}^n x_i f_i} \quad (2)$$

Note that we do not explicitly indicate the dependence of  $p(x_j)$  on  $\mathcal{X}$  and  $f$  in order to avoid notational clutter. When we make count observations  $\mathcal{N} = \{N(x_j) : x_j \in \mathcal{X}\}$  for a set of window durations  $\mathcal{X}$  the joint likelihood function is

$$L(\mathcal{N}|K, \mathcal{X}, f) = \prod_{x_j \in \mathcal{X}} L(N(x_j)|K, \mathcal{X}, f)$$

### B. Maximum Likelihood Estimation is Convex Optimisation

An estimate  $\hat{f}$  of  $f$  which maximizes the likelihood is a solution to the following optimisation problem

$$\begin{aligned} \max_{f_i} L(\mathcal{N}|K, \mathcal{X}, \hat{f}) \\ \text{s.t.} \sum_{i=1}^n \hat{f}_i = 1, \quad \hat{f}_i \geq 0 \end{aligned}$$

It will be useful to instead consider the transformed problem

$$\max_{\tilde{f}_i} \log \tilde{L}(\mathcal{N}|K, \mathcal{X}, \tilde{f}) \quad (3)$$

$$\text{s.t.} \sum_{i=1}^n x_i \tilde{f}_i = 1, \quad \tilde{f}_i \geq 0 \quad (4)$$

where

$$\tilde{L}(\mathcal{N}|K, \mathcal{X}, \tilde{f}) = L(\mathcal{N}|K, \mathcal{X}, \frac{\tilde{f}_i}{\sum_{j=1}^n \tilde{f}_j})$$

with  $\tilde{f}_j = \frac{\tilde{f}_j}{\sum_{i=1}^n x_i \tilde{f}_i}$  just a rescaling of the estimates  $\hat{f}_j$  and

$$\hat{f}_j = \frac{\tilde{f}_j}{\sum_{i=1}^n \tilde{f}_i}$$

and taking the logarithm in the objective leaves the location of the optimum unchanged since the logarithm is monotonically increasing. Our interest in this transformed problem is that the resulting optimisation problem is convex and has a unique solution.

*Lemma 1:*  $\log \tilde{L}(\mathcal{N}|K, \mathcal{X}, \tilde{f})$  is concave in  $\tilde{f}$  (but not strictly concave) and the optimisation problem (3)-(4) is convex.

*Proof:* Observe that

$$\begin{aligned} & \log \tilde{L}(\mathcal{N}|K, \mathcal{X}, \tilde{f}) \\ &= \sum_{x_j \in \mathcal{X}} \left( \log \binom{K}{N(x_j)} \right. \\ & \quad + N(x_j) \log \left( 1 - \sum_{i: x_i > x_j} (x_i - x_j) \tilde{f}_i \right) \\ & \quad \left. + (K - N(x_j)) \log \left( \sum_{i: x_i > x_j} (x_i - x_j) \tilde{f}_i \right) \right) \end{aligned}$$

Now

$$\begin{aligned} \nabla_{\tilde{f}_k} \log \tilde{L} &= \sum_{\substack{x_j \in \mathcal{X}, \\ x_j < x_k}} \left( -\frac{N(x_j)(x_k - x_j)}{1 - \sum_{i: x_i > x_j} (x_i - x_j) \tilde{f}_i} \right. \\ & \quad \left. + \frac{(K - N(x_j))(x_k - x_j)}{\sum_{i: x_i > x_j} (x_i - x_j) \tilde{f}_i} \right) \\ \nabla_{\tilde{f}_i \tilde{f}_k}^2 \log \tilde{L} &= - \sum_{\substack{x_j \in \mathcal{X}, \\ x_j < x_i, x_k}} \alpha(x_j)(x_i - x_j)(x_k - x_j) \end{aligned}$$

where

$$\begin{aligned} \alpha(x_j) &= \frac{N(x_j)}{\left( 1 - \sum_{i: x_i > x_j} (x_i - x_j) \tilde{f}_i \right)^2} \\ & \quad + \frac{K - N(x_j)}{\left( \sum_{i: x_i > x_j} (x_i - x_j) \tilde{f}_i \right)^2} \\ & \geq 0 \end{aligned}$$

Hence, the Hessian of  $\log \tilde{L}(\mathcal{N}|K, \mathcal{X}, \tilde{f})$  is

$$\nabla^2 \log \tilde{L}(\mathcal{N}|K, \mathcal{X}, \tilde{f}) = - \sum_{x_j \in \mathcal{X}} \alpha(x_j) B(x_j)$$

where matrix  $B(x_j) = [x - \mathbf{1}x_j]^+ ([x - \mathbf{1}x_j]^+)^T$ ,  $[y]^+ = y$  when  $y \geq 0$  and 0 otherwise (we apply this in an element-wise manner when  $y$  is a vector),  $x$  is the column vector with elements  $x_i$  and  $\mathbf{1}$  is the all ones column vector. The matrices  $B(x_j D)$ ,  $x_j \in \mathcal{X}$  are non-negative definite since for any vector  $y$  we have  $y^T B(x_j) y = z^2 \geq 0$  where  $z = y^T [x - \mathbf{1}x_j]^+$ . The Hessian is therefore the negative weighted sum of non-negative definite matrices and so is non-positive definite. It follows that  $\log \tilde{L}(\mathcal{N}|K, \mathcal{X}, \tilde{f})$  is concave in  $\tilde{f}$ , although not strictly concave (since the Hessian only non-negative definite rather than positive definite). Since the constraints are linear, it follows that the optimisation problem (3)-(4) is convex. ■

The lack of strict concavity of  $\log \tilde{L}(\mathcal{N}|K, \mathcal{X}, \tilde{f})$  means that additional work is needed to show that the optimization has a unique maximum.

*Lemma 2:* Optimisation problem (3)-(4) has a unique maximum.

*Proof:* The level sets of  $\log \tilde{L}(\mathcal{N}|K, \mathcal{X}, \tilde{f})$  are defined by the linear constraints  $\sum_{i: x_i > x_j} (x_i - x_j) \tilde{f}_i = c_j$ ,  $x_j \in \mathcal{X}$  where  $c_j \geq 0$  is constant. Taking these linear constraints together with equality constraint  $\sum_{i=1}^n x_i \tilde{f}_i = 1$  we have

$$A \tilde{f} = c$$

where  $\tilde{f}$  is the column vector with elements  $\tilde{f}_i$ , and  $c^T = [1 \ c_1 \ c_2 \ \dots \ c_{n-1}]$  and

$$A = \begin{bmatrix} x_1 - x_0 & x_2 - x_0 & x_3 - x_0 & \dots & x_n - x_0 \\ 0 & x_2 - x_1 & x_3 - x_1 & \dots & x_n - x_1 \\ 0 & 0 & x_3 - x_2 & \dots & x_n - x_2 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & x_n - x_{n-1} \end{bmatrix} \quad (5)$$

and we formally define  $x_0 = 0$  and  $c_0 = 1$  since this will prove useful later. Since the  $x_i$  are distinct,  $A$  is rank  $n$  and the level sets consist of single points. It follows immediately that the maximum is unique. ■

### C. Maximum Likelihood Estimator

*Theorem 1 (ML Estimator):* The estimator

$$\hat{f} = \frac{\tilde{f}}{\sum_{j=1}^n \tilde{f}_j} \quad (6)$$

with

$$\tilde{f} = A^{-1} b(K) \quad (7)$$

where  $b(K)$  is the vector with elements  $b_1 = 1$ ,  $b_j = 1 - \frac{N(x_{j-1})}{K}$ ,  $j = 2, \dots, n$  and  $A$  is the upper triangular matrix (5), is a maximum likelihood estimator solving optimisation (3)-(4) whenever the elements of  $A^{-1} b(K)$  are non-negative.

*Proof:* By Lemma 1 the optimisation (3)-(4) is convex and so the KKT conditions are necessary and sufficient for optimality. Ignoring, for the moment, the constraints (4), the KKT conditions for the unconstrained optimization are

$$-\nabla_{\tilde{f}_k} \log \tilde{L}(\mathcal{N}|K, \mathcal{X}, \tilde{f}) = 0, \quad k = 1, 2, \dots, n$$

That is,

$$\begin{aligned} & \sum_{\substack{x_j \in \mathcal{X}, \\ x_j < x_k}} N(x_j) \frac{x_k - x_j}{1 - \sum_{i: x_i > x_j} (x_i - x_j) \tilde{f}_i} \\ &= \sum_{\substack{x_j \in \mathcal{X}, \\ x_j < x_k}} (K - N(x_j)) \frac{x_k - x_j}{\sum_{i: x_i > x_j} (x_i - x_j) \tilde{f}_i}, \\ & \quad k = 1, 2, \dots, n \end{aligned}$$

It can be verified that solutions satisfy the following linear equations

$$\sum_{i:x_i > x_j} (x_i - x_j) \tilde{f}_i = 1 - \frac{N(x_j)}{K}, \quad x_j : x_j \in \mathcal{X} \setminus \{x_1\}$$

Note that there are only  $n - 1$  equations here but  $n$  unknowns. Using the equality constraint  $\sum_{i=1}^n x_i \tilde{f}_i = 1$  to obtain an additional linear equation and taking these equations together, they can be rewritten in matrix form as

$$A\tilde{f} = b(K)$$

Since  $A$  is rank  $n$ , we can solve for  $\tilde{f}$  using (7) and obtain an estimate  $\hat{f}$  for  $f$  by renormalising using (6). Since (6) maximizes an optimization with the same objective but fewer constraints than (3)-(4), if its solution satisfies the constraints (4) then it also maximizes (3)-(4). Hence, provided the elements on the RHS of (7) are non-negative then the solution (6) solves the constrained Maximum Likelihood optimisation (3)-(4) and equations (6)-(7) together define a maximum likelihood estimator. ■

When some elements on the RHS of (6) are negative, the non-negativity constraints (4) need to be enforced. Since the optimisation problem is convex, many standard solvers exist that can efficiently find the constrained maximum likelihood estimate solution. However, we do not pursue this further here since provided  $K$  is sufficiently large, this situation is rare – this is discussed in more detail in the next section.

#### D. Consistency

*Theorem 2 (Consistency of ML Estimator):* As the number of observations  $K \rightarrow \infty$ , the estimate  $\hat{f}$  obtained using (6)-(7) equals the true parameter vector  $f$  with probability one. That is, the estimator (6)-(7) is consistent.

*Proof:* Observe that

$$\begin{aligned} \text{Prob}[|b_i(K) - \frac{A_i f}{\mu}| \geq \epsilon] &= \text{Prob}[|\frac{N(x_{i-1})}{K} - p(x_i)| \geq \epsilon] \\ &\stackrel{(a)}{\leq} 2e^{-2K\epsilon^2} \end{aligned}$$

for any  $\epsilon > 0$  and  $i = 1, 2, \dots, n$ , where  $\mu = \sum_{i=1}^n x_i f_i$ ,  $A_i$  is the  $i$ 'th row of matrix  $A$  and (a) follows from Hoeffding's inequality. Hence,  $\lim_{K \rightarrow \infty} \text{Prob}[|b_i(K) - \frac{A_i f}{\mu}| \geq \epsilon] = 0$  for any  $\epsilon > 0$  and  $i = 1, 2, \dots, n$ . It follows that  $\lim_{K \rightarrow \infty} \text{Prob}[|A_i \tilde{f} - \frac{A_i f}{\mu}| \geq \epsilon] = 0$  for  $i = 1, 2, \dots, n$  and, since matrix  $A$  is full rank, that  $\hat{f} = \frac{f/\mu}{\sum_{j=1}^n f_j/\mu} = f$  with probability one. ■

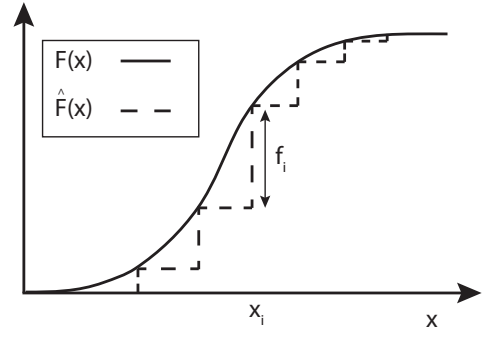


Fig. 2. Illustrating approximation of distribution function  $F(x)$  by staircase function  $\hat{F}(x)$ .

#### E. Extension to Off-Time Distributions With General Support

The analysis can be extended to encompass situations where the *off* times are not restricted to the set of values  $\mathcal{X}$ . Relaxing this restriction, we now have

$$p(x_j) = 1 - \frac{\int_{x_j}^{\infty} (x - x_j) dF(x)}{\int_0^{\infty} x dF(x)}$$

which can be rewritten as

$$p(x_j) = 1 - \frac{\sum_{i:x_i > x_j} (x_i - x_j) f_i}{\sum_{i=1}^n x_i f_i} + \xi(x_j)$$

where

$$\xi(x_j) = \frac{\sum_{i:x_i > x_j} (x_i - x_j) f_i}{\sum_{i=1}^n x_i f_i} - \frac{\int_{x_j}^{\infty} (x - x_j) dF(x)}{\int_0^{\infty} x dF(x)}$$

Importantly, there exists an  $f_i$  such that  $\xi(x_j) = 0$ ; namely,

$$\tilde{f} = A^{-1}d, \quad f = \frac{\tilde{f}}{\sum_{j=1}^n \tilde{f}_j} \quad (8)$$

where  $d$  is a column vector with  $d_0 = 1$  and element  $d_j = \int_{x_{j-1}}^{\infty} (x - x_{j-1}) dF(x) / \int_0^{\infty} x dF(x)$ ,  $j = 2, \dots, n$ , and  $A$  is given by (5).

Now consider estimator (6)-(7). Revisiting the consistency analysis in Section III-D, we have

$$\begin{aligned} \text{Prob}[|b_i(K) - \frac{A_i f}{\mu}| \geq \epsilon] \\ = \text{Prob}[|\frac{N(x_{i-1})}{K} - p(x_i)| \geq \epsilon] \leq 2e^{-2K\epsilon^2} \end{aligned}$$

for any  $\epsilon > 0$  and  $i = 1, 2, \dots, n$ , where  $f$  is the solution to (8) and  $\mu = \sum_{i=1}^n x_i f_i$ . Hence,  $\lim_{K \rightarrow \infty} \text{Prob}[|b_i(K) - \frac{A_i f}{\mu}| \geq \epsilon] = 0$  for any  $\epsilon > 0$  and  $i = 1, 2, \dots, n$ . It follows that  $\lim_{K \rightarrow \infty} \text{Prob}[|A_i \tilde{f} - \frac{A_i f}{\mu}| \geq \epsilon] = 0$  for  $i = 1, 2, \dots, n$  and, since matrix  $A$  is full rank, that with probability one  $\hat{f} = f$ .

In other words, the estimates returned by (6)-(7) are estimates of the vector  $f$  which makes  $\xi(x_j) = 0$ . It is easy to verify that the elements of this vector  $f_i = F(x_i) - F(x_{i-1})$ ,  $i = 1, 2, \dots, n$ . With this in mind, we can alternatively think of (6)-(7) as estimating the distribution function  $\hat{F}(x)$  where  $\hat{F}(x_i) = F(x_i)$ ,  $i = 1, 2, \dots, n$ , samples the true distribution function  $F(x)$  at sample points  $x \in \mathcal{X}$ . This is illustrated in Fig. 2.

#### F. Extension to Non-zero On Times

The analysis can be readily extended to encompass situations where the duration of the *on* times can be non-zero. Suppose the *on* times are iid with finite mean  $\mathbb{E}[S]$ . Then, in the foregoing analysis:

1)  $p(T_D)$  changes to

$$p(x_j) = 1 - \frac{\sum_{i: x_i \geq x_j} (x_i - x_j) f_i}{\mathbb{E}[S] + \sum_{i=1}^n x_i f_i};$$

2) Transformed variable  $\tilde{f}_j$  becomes  $\tilde{f}_j = \frac{f_j}{\mathbb{E}[S] + \sum_{i=1}^n x_i f_i}$  but we still have  $\hat{f} = \frac{\tilde{f}}{\sum_{j=1}^n \tilde{f}_j}$ .

When  $\mathbb{E}[S]$  is known, then we can insert this information into the foregoing analysis (the resulting ML estimator is unchanged aside from modifying element  $b_1(K)$  of vector  $b(K)$  to include  $\mathbb{E}[S]$ ). Alternatively, if we are able to take measurements at the additional data point  $x_0 = 0$ , then this adds the additional linear equation

$$\sum_{i: x_i > 0} x_i \tilde{f}_i = 1 - \frac{N(0)}{K}$$

Note that, by assumption,  $f_0 = 0$  (*off* periods of zero duration occur with probability zero). Hence,  $\sum_{i=0}^n x_i \tilde{f}_i = \sum_{i=1}^n x_i \tilde{f}_i$  and renormalization (6) remains valid – this is the reason why we choose  $x_0 = 0$ . Using  $b_1(K) = 1 - \frac{N(0)}{K}$  in (6)-(7) now yields the ML estimate. Observe that asymptotically as  $K \rightarrow \infty$ ,  $N(0)/K \rightarrow \frac{\mathbb{E}[S]}{\mathbb{E}[S] + \mu}$  and so we can use this to also estimate  $\mathbb{E}[S]$ . In practice, it is usually not possible to make observations for  $x_0 = 0$ . Nevertheless, it is also possible to use a window size  $0 \leq x_0 < x_1$ . In this case the estimate  $\hat{f}$  will continue to be consistent but a bias will be introduced into  $\hat{f}$  due to normalizing in (6) by  $\sum_{i=1}^n x_i \tilde{f}_i$  rather than  $\sum_{i=0}^n x_i \tilde{f}_i$  (recall that we cannot estimate  $f_0$  from (7) since matrix  $A$  is rank  $n$  not rank  $n + 1$ ). However, provided  $x_0$  is sufficiently small, then we can usually expect that this bias will be small.

#### G. Extension to Include Channel Noise

The analysis can be extended to include channel noise as well as on-off interference. That is, to situations where

even during an interference *off* period there is still a probability  $p_G > 0$  that a packet is lost. In this case (2) becomes

$$p(x_j D) = 1 - (1 - p_G) \frac{\sum_{i: x_i > x_j} (x_i - x_j) f_i}{\sum_{i=1}^n x_i f_i}$$

and so the likelihood is now a function own both  $p_G$  and the  $f_i$ 's.

The KKT conditions become

$$\begin{aligned} & \sum_{\substack{x_j \in \mathcal{X}, \\ x_j < x_k}} N(x_j) \frac{(1 - p_G)(x_k - x_j)}{1 - (1 - p_G) \sum_{i: x_i > x_j} (x_i - x_j) \tilde{f}_i} \\ &= \sum_{\substack{x_j \in \mathcal{X}, \\ x_j < x_k}} (K - N(x_j)) \frac{x_k - x_j}{\sum_{i: x_i > x_j} (x_i - x_j D) \tilde{f}_i}, \quad (9) \\ & k = 1, 2, \dots, n \end{aligned}$$

$$\begin{aligned} & \sum_{\substack{x_j \in \mathcal{X}, \\ x_j < x_k}} \frac{K - N(x_j)}{1 - p_G} \\ &= \sum_{\substack{x_j \in \mathcal{X}, \\ x_j < x_k}} \frac{N(x_j) \sum_{i: x_i > x_j} (x_i - x_j) \tilde{f}_i}{1 - (1 - p_G) \sum_{i: x_i > x_j} (x_i - x_j) \tilde{f}_i} \quad (10) \\ & k = 1, 2, \dots, n \end{aligned}$$

It can be verified that solutions satisfy the following linear equations

$$\sum_{i: x_i > x_j} (x_i - x_j) \tilde{f}_i = \frac{1 - \frac{N(x_j)}{K}}{1 - p_G}, \quad x_j \in \mathcal{X} \setminus \{x_1\}$$

which can be rewritten as

$$\sum_{i: x_i > x_j} (x_i - x_j) \tilde{g}_i = 1 - \frac{N(x_j)}{K}, \quad x_j \in \mathcal{X} \setminus \{x_1\}$$

where  $\tilde{g} = (1 - p_G) \tilde{f}$ . Note that there are only  $n - 1$  equations here. If we are able to take measurements at the additional data point  $x_0 = 0$ , then this adds the additional linear equation

$$\sum_{i: x_i > 0} x_i \tilde{g}_i = 1 - \frac{N(0)}{K}$$

Taking these equations together, they can be rewritten in matrix form as

$$A \tilde{g} = b(K) \quad (11)$$

where  $b_j = 1 - \frac{N(x_{j-1})}{K}$ ,  $j = 1, \dots, n$  and  $A$  is rank  $n$  the upper triangular matrix (5). We can then renormalise as before to obtain

$$\hat{f} = \frac{\tilde{g}}{\sum_{j=1}^n \tilde{g}_j}$$

Note that as well as obtaining estimate  $\hat{f}$  of  $f$ , we also have estimate  $\hat{p}_G = \frac{N(0)}{K}$  of  $p_G$ . As noted previously, is usually impractical to take observations of the point  $x_0 = 0$ . Nevertheless, it is also possible to use a window size  $0 \leq x_0 < x_1$  in which case we have that the ML estimate is

$$\sum_{i:x_i > x_0} x_i \tilde{g}_i = 1 - \frac{N(x_0)}{K}$$

and using this in (11) is enough to make  $A$  full rank and thereby allow  $\tilde{g}$  to be estimated. The estimate  $\hat{f}$  will be biased due to normalizing in (6) by  $\sum_{i=1}^n x_i \tilde{g}_i$  rather than  $\sum_{i=0}^n x_i \tilde{g}_i$ , but provided  $x_0$  is sufficiently small then we can usually expect that this bias will be small.

#### IV. ESTIMATION USING MEASURED PACKET LOSS RATES

The foregoing analysis is for an ensemble of on-off processes which is observed via a set of fixed windows. On a wireless link we have instead a single on-off interference process, and we can only perform observations indirectly via the success/failure of packet transmissions. Nevertheless, the insight gained in the previous analysis applies directly to the wireless context.

To see this, we begin by observing that a packet transmission of duration  $T_D$  can be thought of as sampling the channel conditions over a window of length  $T_D$ . If we assume that a packet is lost whenever it overlaps with an interference pulse, then we have a binary observation over this window, observing a 1 if there were one or more interference pulses and a 0 otherwise. Note that this assumption involves no real loss of generality since, if an interference pulse occurs but we do not observe a packet loss, then we cannot distinguish between this event and when no pulse occurred.

It remains to ensure that our packets sample uniformly from the on-off process. We can ensure this using the PASTA property. Namely, assume the intervals between packet transmissions are exponentially randomly distributed and are independent of the interference process. Provided the packet duration  $T_D$  is sufficiently small relative to the mean time between packets, the transmit times  $\{t_j\}$  effectively possess the Lack of Anticipation property (the number of packet transmissions in any interval  $[t, t+u]$ ,  $u \geq 0$ , is independent of the interference process at times  $s \leq t$  [8]). When this property holds, by [8, Theorem 1] we almost surely have

$$\lim_{t \rightarrow \infty} 1 - \frac{1}{K(t)} \sum_{j=1}^{K(t)} U_{T_D}(t_j) = p(T_D)$$

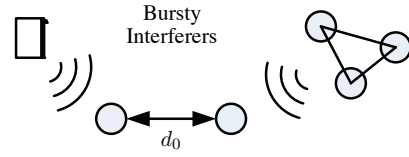


Fig. 3. Simulation setup. Wireless link  $d_0$  is subject to interference from pulsed transmitters. These transmitters might, for example, be a network of similar wireless radios that are operating as hidden nodes or a microwave oven.

where transmit  $K(t)$  packets in interval  $[0, t]$ , each packet being of duration  $T_D$ , and  $U_{T_D}(t_j) = 1$  if the  $j$ 'th packet is lost and 0 otherwise. The left-hand side is the measured packet loss rate. Hence, we have that the packet loss rate is an asymptotically unbiased estimate of  $p(T_D)$ . Substituting the measured packet loss rates as the elements of the vector  $b$  in estimator (6)-(7), instead of the window counts  $N(T_D)$ , it follows that the resulting estimate  $\hat{f}$  will also be consistent and asymptotically unbiased (by a similar argument to that in Section III-D).

#### V. SIMULATION RESULTS

We present simulation results for a wireless link operating in an unlicensed band or white space and subject to pulsed interference, illustrated schematically in Fig. 3.

##### A. Off-Time Distributions with Discrete Support

Fig. 4(c) plots the estimation error  $\max_{x_i \in \mathcal{X}} |\hat{f}_i - f_i|$  versus the number of observed packets  $K$  for both periodic and exponentially distributed on-off interference (the distribution functions are shown in Fig. 4(a)-(b)). It can be seen that the estimation error converges roughly as  $O(1/\sqrt{K})$ , as might be expected, and that an estimation error of less than 5% requires observations for  $K \geq 2000$  packets.

##### B. Off-Time Distributions with Continuous Support

Fig. 5 plots corresponding convergence measurements for Poisson on/off interference. The distribution function  $F(x)$  now has continuous rather than finite support, and the estimator can be thought of as estimating the set points on  $\{F(x) : x \in \mathcal{X}\}$ . It can be seen that the convergence behaviour is similar to the examples with finite support. The estimates are shown in more detail in Fig. 6 for  $K = 1000$  packets. The figure plots the true distribution function  $F(x)$  and the estimate  $\hat{F}(x) = \sum_{i=1}^n \delta(x-x_i) \hat{f}_i$ . It can be seen that the mean estimate is close to the true value and that the standard deviation of the estimates tends to higher to the left of the plot. This is to be expected since as  $F(x)$  decreases the probability

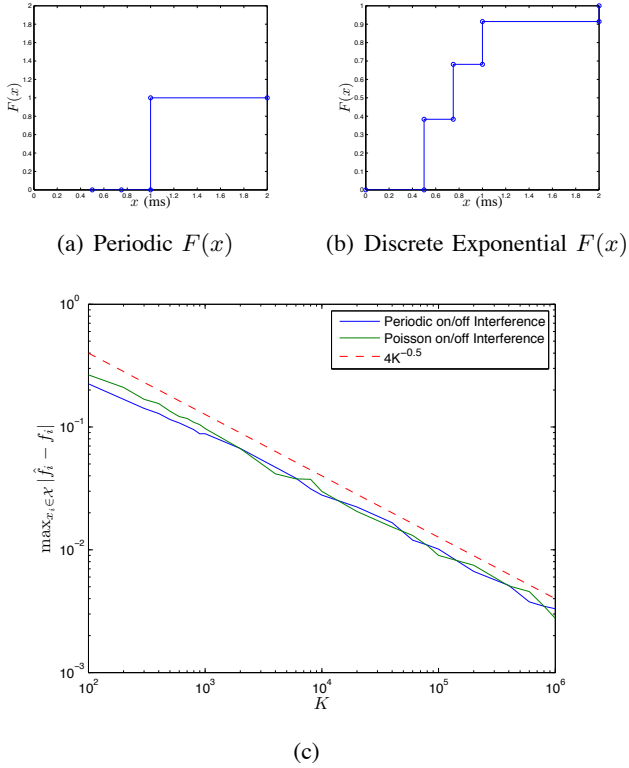


Fig. 4. Convergence of estimates of  $F(x)$  versus the number  $K$  of packets observed when  $F(x)$  has finite support. For each  $K$ , we calculate  $\max_{x_i \in \mathcal{X}} |\hat{f}_i - f_i|$  and average this value over 100 runs to obtain the curves shown. Data is shown for periodic on-off interference with period 10ms and for exponential interference with mean rate 1 pulse per 10ms,  $n = 4$ . Also marked is the line  $4/\sqrt{K}$ , with the data indicating that the estimation error is converging roughly as  $O(1/\sqrt{K})$ .

of observing an interference pulse also decreases and so either more observations are needed to estimate the packet error rate or the estimation errors tend to increase.

### C. Distinguishing On-Off Interference and Channel Noise

Fig. 7 shows results for a wireless link that is subject to both channel noise and to on-off interference. It can be seen that we can accurately estimate both the interference distribution function  $F(x)$  and the channel loss rate  $p_G$ . That is, we are able to successfully and quantitatively distinguish between these two sources of packet loss affecting the link.

## VI. CONCLUSION

We show that by using different size packets and by measuring the packet error rate as a function of packet transmit duration on the air, the distribution of the off times between interference bursts can be accurately estimated. This directly corresponds to the time

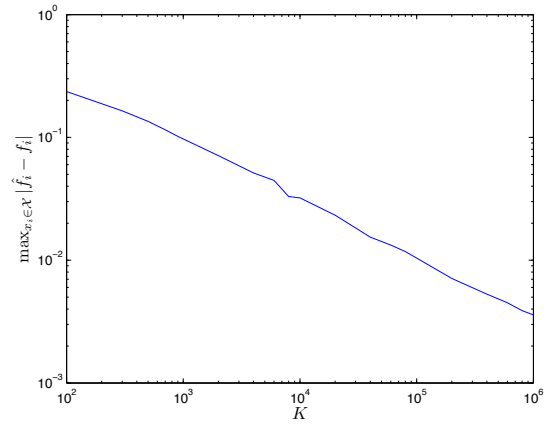


Fig. 5. Convergence of estimates of  $F(x)$  versus the number  $K$  of packets observed when  $F(x)$  has continuous support. Data is shown for Poisson on/off interference, with distribution function shown in Fig. 6.

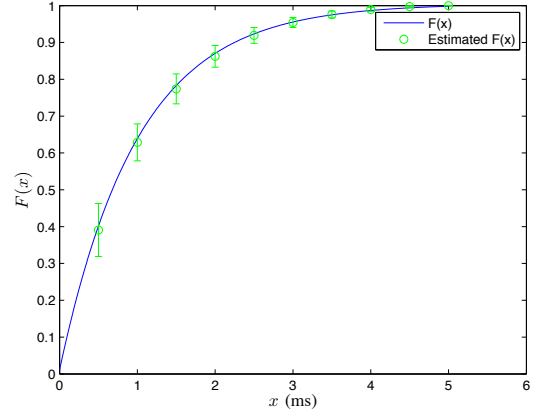


Fig. 6. Distribution function  $F(x)$  and estimate  $\hat{F}(x) = \sum_{i=1}^n \delta(x - x_i) \hat{f}_i$  for  $K = 1000$  packets. Error bars indicate one standard deviation level, estimated empirically from 1000 estimates.

duration of the spectrum hole useable by a secondary station. We present simulation results demonstrating the effectiveness of the proposed technique for a range of interference patterns, and illustrate its ability to distinguish between on-off interference and channel noise and provide quantitative measures of both. By being able to carry out sensing without requiring the transmitter to remain silent, this creates the potential for significant throughput gains.

## REFERENCES

- [1] "Spectrum Occupancy Measurements," Waterford Institute of Technology, Tech. Rep.
- [2] "Spectrum survey in Singapore: Occupancy measurements and analysis," in *Proc. 3rd IEEE Int. Conf. on Cognitive Radio Oriented Wireless Networks and Commun. (CROWNCOM)*, May 2008, pp. 1–7.

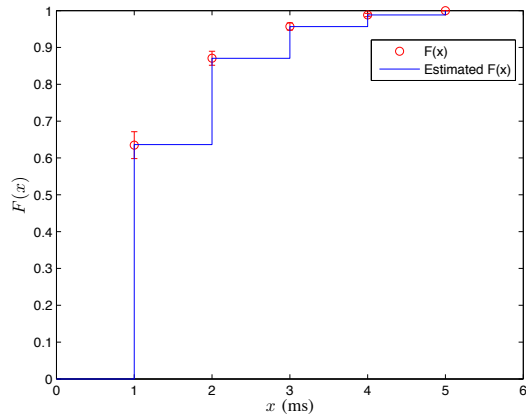


Fig. 7. Example with both channel noise and on-off interference,  $p_G = 0.2$  and  $F(x)$  is discrete exponential. Plot shows distribution function  $F(x)$  and the estimate  $\hat{F}(x) = \sum_{i=1}^n \delta(x - x_i) \hat{f}_i$  for  $K = 1000$  packets. Error bars indicate one standard deviation level, estimated empirically from 1000 estimates. Mean  $\hat{p}_G$  estimate of  $p_G$  is 0.200, with standard deviation 0.012.

- [3] Y.-C. Liang, K.-C. Chen, G. Y. Li, and P. Mahonen, "Cognitive Radio Networking and Communications: An Overview," *IEEE Trans. on Vehic. Tech.*, vol. 60, no. 7, pp. 3386–3407.
- [4] D. Giustiniano, D. Malone, D. J. Leith, and K. Papagianak, "Measuring Transmission Opportunities in 802.11 Links," *IEEE/ACM Trans. on Networking*, vol. 18, no. 5, pp. 1516–1529, 2010.
- [5] B. W. Zarikoff and D. J. Leith, "Measuring Pulsed Interference in 802.11 Links," *submitted to IEEE/ACM Trans. on Networking*, 2011, <http://arxiv.org/abs/1107.5468>.
- [6] Y. Vardi, "Nonparametric Estimation in Renewal Processes," *Annals of Statistics*, vol. 10, no. 3, pp. 772–785, 1982.
- [7] Y. Guedon and C. Coccozza-Thivent, "Nonparametric Estimation in Renewal Processes from Count Data," *Canadian J. Statistics*, vol. 31, no. 2, pp. 191–223, 2003.
- [8] R. W. Wolff, "Poisson Arrivals See Time Averages," *Operations Research*, vol. 30, no. 2, pp. 223–231, 1982.