Timing Estimation for a Filtered Poisson Process in Gaussian Noise

Alfred O. Hero, III, *Member, IEEE*

Abstract—We treat the problem of estimation of time shift of an inhomogeneous causally filtered Poisson process in the presence of additive Gaussian noise. Approximate expressions for the likelihood function, the MAP estimator, and the mmse estimator are obtained, which become increasingly accurate as the per-unit-time density of superimposed filter responses becomes small. The optimal MAP estimator takes the form of a cascade of linear and memoryless nonlinear components. For low signal amplitudes the MAP estimator is equivalent to maximizing the output of a linear matched filter arising in optical communications receivers. For smooth point process intensities, the performance of the MAP estimator is studied via local bias and local variance. A rate distortion type lower bound on the mse of any estimator of time delay is then derived by identification of a communications channel that accounts for the mapping from time delay to observation process. Finally, results of numerical studies of estimator performance are presented. Based on the examples considered it is concluded: 1) the small-error mse of the nonlinear MAP estimator can be significantly better than the small-error mse of the optimal linear estimator; 2) the rate distortion lower bound can be significantly tighter than the Poisson limited bounds determined in previous studies.

Index Terms—Shot noise processes, time delay estimation, intensity shift estimation, rate distortion bounds on estimation error.

I. Introduction

We TREAT in this paper the problem of time shift estimation for inhomogeneous, causally filtered, Poisson point processes, equivalently nonstationary "shot noise," in the presence of additive Gaussian noise when the time delay parameter is imbedded in the intensity function of the point process. We call this combined pulse superposition and Gaussian noise signal a Poisson–Gaussian process. Here we consider only the simple case of constant filter gain; an extension to a random gain model is considered elsewhere [15]. Our model is a special case of a random gain model arising in applications including: nuclear particle detection systems [10], optical communications systems [16], neural spike train analysis [6], acoustic echo-localization [5], seismic signal processing [21], and analysis of underwater acoustic reverberation noise [8]. For example, in Positron emission tomography (PET) a nuclear decay event produces gamma rays that are detected and amplified through avalanche production of photons in photo-multiplier (PM) tubes [23]. Each PM tube drives an electronic circuit whose output is a combination of the superposition of single photon filter responses, forming a filtered Poisson process, and thermal Gaussian noise. The objective of the estimator is to give an accurate estimate of the time-of-flight of the gamma rays, from which the approximate spatial position of the nuclear decay can be determined. In optical communications, an optical receiver is to decide between the presence of a logical "1" and a logical "0" based on the number of photons detected over a certain time interval that is specified by slot-synchronization information. In the case of pulse position modulation (PPM) this binary signal is coded in the time shifts of the photon packets. In other modulation formats, such as phase shift keying (PSK) and pulse amplitude modulation (PAM), lack of synchronization of the receiver to the phase/time delay of the photon intensity can significantly degrade detection performance [9].

Previous work on estimation of time shift of point process intensities has focussed on the Poisson limited regime, e.g., [1], [11], [14], or on the Gaussian limited regime [9], [10]. For the Poisson limited regime, the point process can be observed directly and exact analytical forms of the likelihood function can be found. For this regime the major difficulty lies in the manipulation of the likelihood function into a tractible form for maximum likelihood (ML), maximum *a posteriori* (MAP), or minimum mean-square error (mmse) estimation of time shift. For the Gaussian limited regime the time shift parameter is imbedded in the mean and covariance functions of the observation waveforms and an analytical functional form for the likelihood function can be found, see for example [7]. For the mixed Poisson–Gaussian regime considered in this paper, an analytical form for the process density function is more difficult to obtain even though the characteristic function has an analytical form [24]. Here we focus on an approximation to the likelihood function based on a *low-pulse density condition* similar to [22, Ch. 11]: the product of the filter pulse-width and the intensity amplitude is uniformly small over time. This condition can be viewed as complementary to the condition of uniformly large intensity amplitude for which case one can assert that the filtered Poisson process is approximately Gaussian [27].

It is worth pointing out that estimation of point process intensity time shift differs from the related problem of estimation of the amplitudes and arrival times of the superimposed filter response pulses. In particular, in [5, 20] the pulse arrivals are not Poisson and the number of pulses is assumed known. In [21] Poisson arrivals are considered but the point process intensity, assumed constant therein, is only incidental to the pulse amplitude and pulse position estimation problem. The asymptotic likelihood function approximation developed here is obtainable as a special case of an approximation proposed in a paper [18] that appeared a number of

Manuscript received March 23, 1988; revised June 12, 1990. This work was supported in part by the National Institutes of Health (NIH) under Contract R01-CA46622-01. This work was presented in part at the 1988 Conference on Information Sciences and Systems, Princeton, NJ, March 1988.

The author is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109.

IEEE Log Number 9039281.

0018-9448/91/0100-0092$01.00 © 1991 IEEE

Authorized licensed use limited to: University of Michigan Library. Downloaded on July 28, 2009 at 11:45 from IEEE Xplore. Restrictions apply.

months after this paper was submitted. However, in [18] the estimation problem is not considered.

In this paper we use the low pulse density approximation to obtain forms for the likelihood ratio used to detect the time shifted point process for known intensity time shift. The expression is then used to obtain approximations to the MAP estimator and the mmse estimator of time delay. The log-likelihood ratio is seen to have the following structure: After suppression of the Gaussian noise component of the observations via classical matched filtering, an exponential transformation generates a spike train that is subsequently correlated against a time-shifted version of the intensity function. The log-likelihood ratio is maximized over the set of possible times shifts to yield the MAP estimate. It is significant that for low process amplitude, and a narrow superposition filter response, the nonlinear log-likelihood ratio reduces to a classical linear filter receiver that has been previously proposed as a suboptimal detector for optical communications systems. This is a uniformly low intensity extension of the well-known result that the linear matched filter is an approximately optimal linear estimator/detector for uniformly high intensity. While we do not consider random gains on the generated pulses, it is shown in [15] that with random Gaussian distributed gains the resultant likelihood function approximations are quite similar to those in this paper. Specifically, the only difference is that the random gain model introduces a quadratic memoryless nonlinearity between the matched-filter and the exponential transformation in the log-likelihood structure previously described.

To evaluate the performance of the approximation to the MAP estimator, we derive expressions for the local bias and local variance under the assumption that the point process intensity is differentiable. These expressions characterize the small error behavior of the approximate MAP estimator over the full range of low to high pulse density. The local bias and local variance can be related to the asymptotic rates of decay of the actual bias and variance of a MAP estimator obtained from observing successive independent versions of the Poisson–Gaussian process through the mechanism of repeated experiments. Alternatively, the small error variance can be interpreted as an approximate Cramer–Rao lower bound for low-pulse density. To establish fundamental limits on estimation performance, a rate distortion lower bound on estimator mse is then derived. The bound is obtained by identifying a cascaded Poisson and Gaussian channel which maps random time delays to the measurement process. The capacity of the Poisson channel is upper bounded using results in [14]. On the other hand, the capacity of the Gaussian channel is upper bounded using standard results of rate distortion theory [3]. The final form of the rate distortion bound is obtained via the *data processing theorem*. Finally, a numerical evaluation of the expressions for small-error mse and the rate distortion lower bound is presented for the special case of Gaussian intensity and exponential superposition filter response. Based on the examples considered we conclude: 1) the small-error mse of the nonlinear MAP estimator can be significantly better than the small-error mse of the optimal linear estimator; 2) the rate distortion lower bound can be significantly tighter than the Poisson limited bounds studied in [14].

The organization of the paper is as follows. Section II introduces the main assumptions and gives a low-pulse density approximation to the likelihood ratio. Section III reviews conventional linear estimator structures and develops forms for the MAP and mmse estimators. Section IV develops expressions for local bias and local variance of the MAP estimator. Section V presents lower bounds on mse. In Section VI numerical performance comparisons are presented.

II. Problem Statement

A few words about notation are useful. In general, bold faced variables, e.g., X, denote random variables, vectors, or processes. A notable exception is the random variable τ. The probability distribution of X is denoted by the generic "$P(X)$." Observe that $P(X)$ and $P(\tau)$ are different functions, not the evaluation of a function "P" for two different values of its argument. The probability density of X, with respect to some usually unspecified dominating measure μ, is denoted by $f(X) \overset{\text{def}}{=} dP(X)/d\mu$.

We will consider the case of a general random time delay τ. Estimators for the nonrandom case can be treated by specializing to a uniform prior on τ. Conditioned on τ, let $N \overset{\text{def}}{=} \{N(t)\colon t \in [0,T]\}$ be an inhomogeneous Poisson process with intensity function $\{\lambda(t-\tau)\colon t \in [0,T]\}$, where

$$\lambda(t) = \lambda_s(t) + \lambda_o. \tag{1}$$

In view of (1), N corresponds to the sum of an inhomogeneous Poisson signal process, N^s, with intensity $\lambda_s(t-\tau)$, and an independent homogeneous Poisson noise process, N^o, with intensity λ_o. It will be assumed that τ affects λ_s as a shift, without truncation in the sense that the set $\{t\colon \lambda_s(t-\tau) > 0\}$ is contained in the observation interval $[0,T]$ for all τ over the support set $\{\tau\colon f(\tau) > 0\}$.

Let the total number of points of N over $[0,T]$ be denoted by n, specifically $n \overset{\text{def}}{=} N(T)$. Let $\{t_i\}_{i=1}^n$ be the n occurrence times associated with the points of N. The joint probability distribution of $\{t_i\}_{i=1}^n$ and n given τ is specified by the conditional density $f(\{t_i\}_{i=1}^n | n,\tau)$ of $\{t_i\}_{i=1}^n$ and the probability mass function $p(n|\tau)$ of n [27]:

$$f(\{t_i\}_{i=1}^n | n,\tau)\, p(n|\tau) = \begin{cases} e^{-\Lambda} \prod_{i=1}^{n} \lambda(t_i - \tau), & n > 0, \\ e^{-\Lambda}, & n = 0, \end{cases} \tag{2}$$

where

$$\Lambda \overset{\text{def}}{=} \int_0^T \lambda(t-\tau)\, dt = \int_0^T \lambda(t)\, dt = \Lambda_s + \Lambda_o, \tag{3}$$

is the energy or integrated rate of the point process which, by assumption, is functionally independent of τ.

Available for observation over the time interval $t \in [0,T]$ is the sum of a filtered Poisson process and Gaussian noise

$$X(t) = \sum_{i=1}^{N(t)} p(t - t_i) + w(t), \tag{4}$$

where $p(t)$ is a known continuous and square-integrable filter impulse response with 3-dB time-width T_p, W is zero-mean white Gaussian noise independent of τ and $\{t_i\}_{i=1}^n$ with $E[w(t)w(u)] = (N_o/2)\delta(t-u)$. For the purposes of mathematical analysis, the notation in (4) is to be interpreted as shorthand for the equivalent second-order integrated observation $Y(t) = \int^t \sum_{i=1}^{N(u)} p(u - t_i)\, du + W(t)$, where $W(t)$ is a standard Weiner process [28]. We will assume that the filter

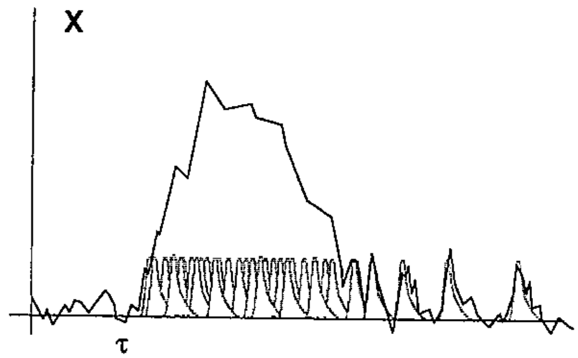

Fig. 1. Sample path of Poisson–Gaussian process $X = X(t)$ is superposition of randomly shifted pulses (indicated in figure by dotted lines), and uncorrelated Gaussian noise.

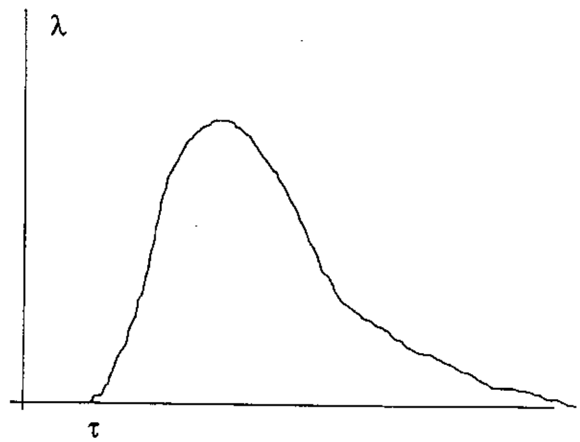

Fig. 2. Intensity function $\lambda = \lambda(t-\tau)$ of time shifts (Poisson occurrence times) of super-imposed pulses in Fig. 1.

p is causal: $p(t) = 0$ for $t < 0$. Under this assumption the process (4) is equivalent to the observation

$$X(t) = \sum_{i=1}^{n} p(t - t_i) + w(t). \quad (5)$$

Fig. 1 illustrates a typical realization of the X process for the intensity plotted in Fig. 2.

Fix a realization of τ. The objective of a signal detector is to decide between the following two hypotheses concerning the presence or absence of the signal intensity λ_s on the basis of the observation X:

$$H_0: \text{ intensity} = \lambda_o,$$
$$H_1: \text{ intensity} = \lambda_s(t-\tau) + \lambda_o. \quad (6)$$

The optimal test of H_1 vs. H_0 is the likelihood ratio test that compares a likelihood ratio statistic

$$L(X|\tau) \overset{\text{def}}{=} \frac{dP(X|H_1,\tau)}{dP(X|H_0)}$$

to a threshold. A general abstract form for this statistic can be obtained using methods given in [19], [28]. Here we take a different approach. Conditioned on the values of the set of occurrence times, $\{t_i\}_{i=1}^n$, and on their number, n, X in (5) is a deterministic signal plus a white Gaussian random process. A form for the conditional density of the observations given $\{t_i\}_{i=1}^n$, n and τ can then be obtained using the Cameron–Martin formula [25]:

$$f(X|\{t_i\}_{i=1}^n, n, \tau) = e^{(2/N_o)\int_0^T X(t)\sum_{i=1}^n p(t-t_i)\,dt - (1/N_o)\int_0^T(\sum_{i=1}^n p(t-t_i))^2\,dt}. \quad (7)$$

The numerator and denominator of the likelihood ratio statistic are obtained by taking the expectation of (7) over $\{t_i\}_{i=1}^n$ and n, denoted $E_{t_1 \cdots t_n}$, given H_1 and H_0 respectively:

$$f(X|H_1,\tau) = E_{t_1\cdots t_n}\left[f(X|\{t_i\}_{i=1}^n, n, \tau)\,\middle|\, H_1, \tau\right],$$
$$f(X|H_0) = E_{t_1\cdots t_n}\left[f(X|\{t_i\}_{i=1}^n, n, \tau)\,\middle|\, H_0\right]. \quad (8)$$

The two conditional density functions of (8) are not in analytic form. In Appendix A, $O(\Lambda T_p \lambda^+)$ approximations to these functions are obtained, where

$$\lambda^+ \overset{\text{def}}{=} \frac{1}{T_p} \max_t \int_t^{t+T_p} \lambda(t)\,dt \le \max_t \lambda(t).$$

Therefore, the approximation is accurate under the assumption $T_p \ll 1/(\Lambda \max_t \lambda(t))$. This assumption is valid when the pulse width T_p is sufficiently small so that the nonoverlapping pulses dominate in the sum $\int(\Sigma)^2$ in (7). Substitution of expressions (A.9) and (A.10) of Appendix A into the likelihood ratio statistic $L(X|\tau) = f(X|H_1,\tau)/f(X|H_0)$ gives

$$L(X|\tau) = \exp\left(\int_0^T \left(e^{(2/N_o)X(t)*p(-t)-\gamma/2} - 1\right)\lambda_s(t-\tau)\,dt\right), \quad (9)$$

where the asterisk denotes convolution and γ is the pulse-to-noise ratio (PNR):

$$\gamma \overset{\text{def}}{=} \frac{2}{N_o}\int_0^T p^2(t)\,dt. \quad (10)$$

III. Estimator Structures

In this section we will investigate the forms of the MAP and mmse estimators for τ, under the approximation (9) to the likelihood ratio. In Section VI a performance analysis is presented for the MAP estimator. Let the MAP estimator and the mmse estimator for τ be denoted $\hat{\tau}_{\text{MAP}}$ and $\hat{\tau}_{\text{mmse}}$ respectively.

A. MAP Estimation

The MAP estimator maximizes the *a posteriori* probability $f(\tau|X, H_1)$, or, equivalently, it maximizes the quantity $\ln[L(X|\tau)f(\tau)]$, over the set of possible τ. From (9) of Section II, it is seen that the maximum occurs at the same point as the maximum of the following function:

$$\hat{\tau}_{\text{MAP}} \overset{\text{def}}{=} \arg\max_\tau \left\{\int_0^T \left[e^{(2/N_o)X(t)*p(-t)-\gamma/2} - 1\right] \cdot \lambda_s(t-\tau)\,dt + \ln f(\tau)\right\}. \quad (11)$$

B. mmse Estimation

The mmse estimator is the conditional mean $E[\tau|X]$ [25] which, using (9), has the form

$$\hat{\tau}_{\text{mmse}} \stackrel{\text{def}}{=} \int \tau \frac{L(X|\tau)f(\tau)}{L(X)}\, d\tau = \frac{\int \tau L(X|\tau) f(\tau)\, d\tau}{\int L(X|\tau) f(\tau)\, d\tau}. \quad (12)$$

Substitution of the right-hand side of (9) into the right-hand side of (12) gives

$$\hat{\tau}_{\text{mmse}} = \frac{\int d\tau\, \tau \exp\left(\int_0^T dt\, e^{(2/N_o)X(t)*p(-t)-\gamma/2} - 1\right)\lambda_s(t-\tau)f(\tau)}{\int d\tau \exp\left(\int_0^T dt\, e^{(2/N_o)X(t)*p(-t)-\gamma/2} - 1\right)\lambda_s(t-\tau)f(\tau)}. \quad (13)$$

C. Linear Estimator Structures

For the purposes of comparison we mention the class of linear estimator structures for time shift. These estimators result from maximization of a linear function of the observed waveform X:

$$\hat{\tau}_L \stackrel{\text{def}}{=} \arg\max_\tau \{X(\tau) * h(\tau)\}, \quad (14)$$

where h is a function of the underlying parameters. Representative examples of h are: the "optical matched filter,"

$$h(t) \stackrel{\text{def}}{=} \frac{\lambda_s(-t)}{\lambda_s(-t) + \lambda_o + \frac{N_o}{2}}, \quad (15)$$

and robust versions of (15), [12]; and the point process domain optimal filter [9]:

$$h(t) = \ln\left(1 + \frac{\lambda_s(-t)}{\lambda_o}\right). \quad (16)$$

Often the estimator $\hat{\tau}$ is implemented by detecting a zero crossing of the right-hand side of (14). Examples include the first electron and the constant fraction timing estimators implemented in scintillation counters [10]. These various filter functions $h(t)$ are substantially different both in their form and in their theoretical motivation. The optical matched filter is an optimal linear filter for signal detection in the white noise limited, voltage waveform domain. The point process domain optimal filter gives a maximum-likelihood estimator structure for ideal photon detection. The first electron timing estimator is a suboptimal filter which has been common in particle detection systems such as those implemented in PET. For the purposes of mse comparisons, we will focus on the optical matched filter (15) estimator since it is the optimal linear estimator for the observation model considered.

D. Discussion

The approximate expression for the likelihood function (9) becomes exact as the pulse width T_p becomes small. In this case the approximate MAP and mmse estimators (11) and (13) become exact. For the remainder of the discussion, we confine our attention to the MAP estimator and the case of uniform τ, $f(\tau) = constant$. For notational convenience, we let $\hat{\tau} = \hat{\tau}_{\text{MAP}}$.

A block diagram of the MAP estimator (11) is given in Fig. 3. Observe that the statistic to be maximized is a simple nonlinear function of the observations. The estimator structure can be divided into three sequential tasks, indicated in Fig. 3, accomplished by subsystems A, B, and C respectively: A) a cascade of a noise prewhitening filter K_w^{-1}, and a classical linear matched filter of the form $h(t) = p(-t)$ extracts a raw signal X_A by smoothing the noise process w and enhancing the superposition process $\Sigma_{i=1}^n p(t - t_i)$ (K_w^{-1} is just a multiplicative factor $2/N_o$ for white noise w); B) the nonlinearity $\exp(\cdot)$ has the effect of producing a spike train, X_B, which emulates a point process by increasing the dynamic range of X_A; C) a correlation of the spike train X_B with $\lambda_s(t - \tau)$ generates the log-likelihood ratio, $\ln L(\{t_i\}_{i=1}^n|\tau)$, and a subsequent maximization over τ extracts the time delay parameter.

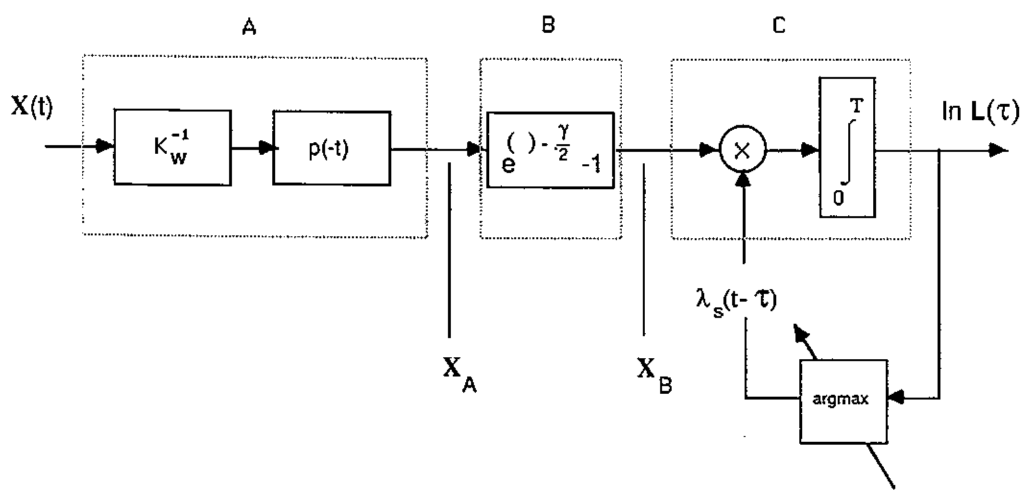

Fig. 3. Block diagram of approximate MAP estimator of time shift parameter for uniform τ.

It is interesting to study the form of the MAP estimator (11) in the limit of small values of the argument of the exponential function: $(2/N_o)X(t)*p(-t)$. In this case, using the standard approximation $e^y \approx 1+y$ and the fact that under our assumptions $\int_0^T \lambda_s(t-\tau)dt$ is functionally independent of τ over the support set of $f(\tau)$, the MAP estimator reduces to the form of a linear estimation structure (14) with

$$h(t) \stackrel{\text{def}}{=} \lambda_s(-t)*p(-t). \tag{17}$$

For a narrow superposition filter response, $p(t)$, and for $\Lambda\gamma \ll 1$, where γ is the PNR defined by (10), $h(t)$ of (17) is equivalent to the "optical matched filter" (15).

IV. Estimator Performance

The relative performance of the MAP and linear estimators is characterized by local estimator bias and local estimator variance. Conditioned on a particular value of τ, let M repeated identical experiments be performed yielding M independent and identically distributed versions $\{X_i\}_{i=1}^M$ of the observation $\{X(t): t\in[0,T]\}$. Let $\hat{\tau}$ be an estimate of τ formed from these M observations.

As a relevant example consider λ_s to be periodic with period T_o, and assume that within each period λ_s is zero except over a small interval of length T_s. This model for the signal λ_s arises in optical telemetry and in optical communications repetition codes for synchronization. Define T_s/T_o the duty cycle of λ_s. If, conditioned on τ, X satisfies suitable asymptotic mixing conditions [2], the trajectory of X over increasingly separated nonoverlapping segments of time are asymptotically conditionally independent. For such a process, the observation of X over M successive periods is asymptotically equivalent to M identical experiments $\{X(t): t\in[0,T_o]\}$ as T_o increases to infinity for fixed duty cycle <1. Furthermore, under such mixing conditions it can be directly shown that the approximate likelihood function $L(X|\tau)$ of (9) decomposes into the product of M approximately independent identically distributed (i.i.d.) factors $L(X_i|\tau)$, $i=1,\cdots,M$.

Let $\xi = z_L$ be a zero of the function $E[(d/d\xi)L(X|\xi)|\tau]$, which is nearest to τ. Conditioned on τ we define the local bias $b(\tau)$ and the local variance $\sigma^2(\tau)$ of the ML estimator $\hat{\tau}$ to be:

$$b(\tau) \stackrel{\text{def}}{=} z_L - \tau,$$

$$\sigma^2(\tau) \stackrel{\text{def}}{=} \frac{1}{M}\frac{E\left[\left(\dfrac{d\ln L(X_i|\tau)}{d\tau}\right)^2\Bigg|\tau\right]}{E^2\left[-\dfrac{d^2\ln L(X_i|\tau)}{d\tau^2}\Bigg|\tau\right]}. \tag{18}$$

Conditioned on τ, the local bias and local variance of $\hat{\tau}$ are equivalent to the asymptotic bias and variance of $\hat{\tau}$ as $M\to\infty$ under certain regularity conditions on the approximate likelihood function $L(X|\xi)$ [26]. Specifically, if $L(X|\xi)$ is twice differentiable at $\xi = z_L$ uniformly in X and the variance of the first derivative is finite, if z_L is an isolated zero of $E[(d/d\xi)L(X|\xi)|\tau]$, if $E[(d^2/d\xi^2)L(X|\xi)|\tau]$ is nonzero at $\xi = z_L$, and if $\hat{\tau}$ is weakly consistent in the sense that it converges to z_L in probability, then $\sqrt{M}(\hat{\tau}-z_L)$ is asymptotically Gaussian with zero mean and variance $\sigma^2(\tau)$ given by (18) [26, Theorem 7.2.2B]. With exception of weak consistency, these technical conditions are easily verified for the function $L(X|\xi)$ displayed in (9) using results derived in Appendix B and using the a.s. continuity of the trajectories of $\int X(t)*p(-t)$. Weak consistency of $\hat{\tau}$ is more difficult and has not been verified. Local weak consistency of $\hat{\tau}$, on the other hand, is simple to verify using a monotone approximation to $(d/d\xi)L(X|\xi)$ in the neighborhood of $\xi = z_L$ [26, Lemma 7.2.2A].

Conditioned on τ the local mse is now naturally defined as

$$\begin{aligned}\text{mse}(\tau) &= \sigma^2(\tau)+b^2(\tau)\\ &= \frac{1}{M}\frac{E\left[\left(\dfrac{d\ln L(X_i|\tau)}{d\tau}\right)^2\Bigg|\tau\right]}{E^2\left[-\dfrac{d^2\ln L(X_i|\tau)}{d\tau^2}\Bigg|\tau\right]}+(\tau-z_L)^2.\end{aligned} \tag{19}$$

If the function $L(X|\tau)$ were the exact probability density function for X given τ, then it could be verified that the local bias is identically zero, and the local variance (18) is identically the Cramer–Rao lower bound on the variance of $\hat{\tau}$ [17]. The above local error analysis applies to any estimator obtained by maximizing an objective function $l(X,\tau)$ over τ which satisfies the additive decomposition property: $l(X,\tau) = \sum_{i=1}^M l(X_i,\tau)$, e.g., the linear estimator structure (14).

A. MAP Estimator Local Bias

We assume that τ is uniformly distributed over its domain. If the local bias is small, i.e., the zero $\xi = z_L$ of $\Psi(\xi) \stackrel{\text{def}}{=} E[(d/d\xi)\ln L(X_i|\xi)|\tau]$ is close to τ, a first-order linear approximation in the neighborhood of $\tau = z_L$ holds: $\Psi(z_L) = 0 = \Psi(\tau)+(z_L-\tau)\Psi'(\tau)$. Assuming exchange of derivative and expectation is justified, this gives an analytical approximation to estimator local bias $b(\tau) = z_L - \tau$:

$$b(\tau) = \frac{E\left[\dfrac{d\ln L(X_i|\tau)}{d\tau}\Bigg|\tau\right]}{E\left[-\dfrac{d^2\ln L(X_i|\tau)}{d\tau^2}\Bigg|\tau\right]}. \tag{20}$$

Expressions for the numerator and denominator of (20) are calculated in Appendix B, ((B.7) and (B.8) respectively). Substitution of these quantities into (20) gives the following expression for the local bias of the approximate MAP estimator $\hat{\tau}$:

$$b(\tau) = \frac{\int_0^{T_o}\exp\left(\int_0^{T_o} e^{(2/N_o)R_p(u-t)}\lambda(u)\,du\right)\lambda'(t)\,dt}{\int_0^{T_o}\exp\left(\int_0^{T_o} e^{(2/N_o)R_p(u-t)}\lambda(u)\,du\right)\lambda''(t)\,dt}, \tag{21}$$

where $R_p(t) \stackrel{\text{def}}{=} \int_0^T p(u+t)p(u)du$ is the pulse autocorrelation function. Note that the local bias is functionally independent of τ, and the expression (21) is thus equivalent to the unconditioned bias $E[b(\tau)] = E[z_L - \tau]$. Observe also that the bias of $\hat{\tau}$ depends on the particular structure of λ and the pulse autocorrelation function R_p. Since autocorrelation functions are symmetric about the origin, it can easily be shown that the local bias is identically zero if λ is symmetric about any point. Hence, for an even signal intensity function λ_s, the approximate MAP estimator is locally unbiased.

For the special case of small PNR γ (10), the local bias (21) reduces to the simple form

$$b(\tau) = \frac{\int_0^{T_o} dt_1 \int_0^{T_o} dt_2 \hat{\lambda}(t_1) \hat{R}_p(t_1 - t_2) \hat{\lambda}'(t_2)}{\int_0^{T_o} dt_1 \int_0^{T_o} dt_2 \hat{\lambda}(t_1) \hat{R}_p(t_1 - t_2) \hat{\lambda}''(t_2)} \tag{22}$$

where we have defined the normalized quantities:

$$\hat{R}_p(t) \stackrel{\text{def}}{=} \frac{R_p(t)}{\overline{p^2}}, \qquad \hat{\lambda}(t) \stackrel{\text{def}}{=} \frac{\lambda(t)}{\Lambda}, \tag{23}$$

and $\overline{p^2} \stackrel{\text{def}}{=} R_p(0)$ is the energy of the pulse $p(t)$. It is important to observe that to a small PNR approximation, the local bias is only dependent on the shapes of the intensity function λ and the pulse autocorrelation function R_p, i.e., bias is independent of the PNR γ and the rate Λ.

B. MAP Estimator Local Variance

Consider the case of uniformly distributed τ again. The numerator and denominator of (18) are calculated in Appendix B, ((B.11) and (B.8) respectively). Thus we have the following expression for the local variance of the approximate MAP estimator conditioned on τ:

$$\sigma^2(\tau) = \frac{\int_0^{T_o} dt_1 \int_0^{T_o} dt_2 \exp\left(\Lambda \int_0^{T_o} [e^{\gamma[\hat{R}_p(u-t_1) + \hat{R}_p(u-t_2)]} - 1] \hat{\lambda}(u)\, du + \gamma \hat{R}_p(t_1 - t_2)\right) \hat{\lambda}'(t_1) \hat{\lambda}'(t_2)}{M\left[\int_0^{T_o} \exp\left(\Lambda \int_0^{T_o} [e^{\gamma \hat{R}_p(u-t)} - 1] \hat{\lambda}(u)\, du\right) \hat{\lambda}''(t)\, dt\right]^2}, \tag{24}$$

where $\hat{\lambda}$ and $\hat{R}_p$ are defined in (23).

We make the following observations based on the local variance (24). Similarly to the local bias, the local variance (24) is functionally independent of τ. Note that, since $R_p(u) = 0$, $|u| > T_p$, $\lambda''(t)$ and $\lambda(u)$ do not jointly affect the local variance for $|u - t| > T_p$. The same property holds for the joint influence of $\lambda'(t)$ and $\lambda(u)$. T_p corresponds to the memory introduced by filtering $\{t_i\}_{i=1}^n$ with a filter with response $p(t)$. Hence memory smears the intensity function over time as far as its influence on the local variance is concerned. This differs from the pure Poisson observations case, where it can be shown that the local variance and the local bias do not depend on the correlation between different time samples of λ (see also the form (27) of the Poisson limited CR bound in Section V).

Due to the presence of integrands of the form $e^{e^{(\cdot)}}$ in the numerator and denominator of (24), the numerical computation of the local variance approximation is impractical for the case of large PNR-rate product $\gamma\Lambda$. For the case of small PNR-rate product, $\gamma\Lambda \ll 1$, it is shown in Appendix B that the expression (24) for local variance reduces to the expression:

$$\sigma^2(\tau) = \frac{1}{\gamma\Lambda^2 M} \frac{-1}{\int_0^{T_o} dt_1 \int_0^{T_o} dt_2 \hat{\lambda}(t_1) \hat{R}_p(t_1 - t_2) \hat{\lambda}''(t_2)}. \tag{25}$$

Under the assumed low $\gamma\Lambda$ conditions, it can easily be shown that the expression (25) is equivalent to the local variance for the linear matched filter estimator (17). For low PNR-rate product, a rate of decay of the variance on the order of Λ^2 as a function of the energy Λ of the Poisson process is indicated by (25). The low PNR expression for local variance (25) is monotonically decreasing in the PNR γ. It is also decreasing in the second derivative of the normalized intensity function $\hat{\lambda}''(t)$ for those values of t such that $\hat{\lambda}(u)$ is large, $|u - t| \le T_p$. On the basis of these observations, the estimator can be expected to have low variance for the cases where: 1) the process $\{t_i\}_{i=1}^n$ is easy to estimate by virtue of high PNR; 2) the intensity function λ is highly resolved in the sense that Λ is high and λ is "sharply peaked," i.e., $\lambda''(t) \gg 0$, in t-regions of high intensity.

V. Lower Bounds on mse

Here we derive lower bounds on estimator error for the estimation problem outlined in Section II. Two bounds will be presented: the Cramer–Rao (CR) lower bound, derived under the optimistic Poisson limited (high PNR) regime; and a rate distortion bound. Both of these bounds specify a lower limit on achievable mse for the general Gaussian–Poisson regime and, unlike the MAP approximation (9), are applicable to arbitrary pulse density conditions.

A. CR Bound

The CR bound on τ is given by the inequality

$$\text{mse} \ge \frac{1}{E\left[\frac{d}{d\tau} \ln L(X|\tau) f(\tau)\right]^2}, \tag{26}$$

where $L(X|\tau)$ is the likelihood ratio for the hypotheses (6). For the case of pure Poisson observations, $X = \{t_i\}_{i=1}^n$, the CR bound is derived in [14]:

$$\begin{aligned} \text{mse} &\ge \frac{1}{\Lambda \int_0^T \left(\frac{d \ln \lambda(t)}{dt}\right)^2 \hat{\lambda}(t)\, dt + E\left[\left(\frac{d \ln f(\tau)}{d\tau}\right)^2\right]} \\ &= \frac{1}{\Lambda \int_0^T \frac{(\lambda_s'(t))^2}{\lambda_o + \lambda_s(t)} \hat{\lambda}(t)\, dt + E\left[\left(\frac{d \ln f(\tau)}{d\tau}\right)^2\right]}, \end{aligned} \tag{27}$$

where it has been assumed that derivatives in (27) exist. Since addition of Gaussian noise to the observations only introduces additional τ-uncoupled nuisance parameters to the estimation problem, the Poisson limited CR bound (27) is less than or equal to the exact but uncomputable CR bound valid for the observation model of Section II [4]. Unfortunately, the right-hand side of (27) fails to be a useful bound for a large class of intensity functions, e.g., bi-exponential, rectangular, or any intensity for which the ratio of the squared first derivative and the intensity magnitude is not absolutely integrable.

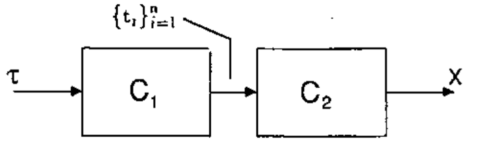

Fig. 4. Effects of mapping random timing parameter τ into observation space as decomposed into cascade of two transformations: 1) channel C_1 accounts for generation of point process through Poisson mechanism, this is point process channel; 2) channel C_2 accounts for finite bandwidth and additive noise effects introduced via detection process, this is continuous observation channel.

B. Rate-Distortion Type Lower Bound

For the case of perfect Poisson observations $X=\{t_i\}_{i=1}^n$, a rate-distortion type lower bound was derived in [14]. The bound in [14] was derived from a special case of Shannon's inequality

$$H(\tau)-\frac{1}{2}\ln(2\pi e\,\mathrm{mse})\le R(\mathrm{mse})\le C, \tag{28}$$

where C is the capacity of a channel taking the source symbols τ to destination symbols X, R is the rate-distortion function associated with a mse distortion measure, and $H(\tau)$ is the entropy of the p.d.f. of τ. Shannon's inequality (28) gives a lower bound on mse in terms of C:

$$\mathrm{mse}\ge\frac{1}{2\pi e}e^{2H(\tau)}e^{-2C}. \tag{29}$$

In the present situation, the channel C can be represented as the cascade of two separate channels C_1 and C_2, where C_1 takes the source symbols τ into the occurrence times $\{t_i\}_{i=1}^n$, and C_2 takes the occurrence times $\{t_i\}_{i=1}^n$ into the observations X (see Fig. 4). Hence, C_1 is an intensity modulated Poisson point process channel with associated intensity $\lambda(t-\tau)$, while C_2 is a Gaussian white noise channel with impulse response $p(-t)$ and Poisson input statistics. An upper bound on the overall capacity C of this cascaded channel is given by the "data processing theorem" [3] as $C\le\min\{C_1,C_2\}$. Furthermore, an explicit upper bound, C_1^*, on C_1 was given in [14]:

$$C_1\le C_1^*\overset{\mathrm{def}}{=}\Lambda\int_0^T\hat{\lambda}(t)\ln\frac{\hat{\lambda}(t)}{\frac{1}{T}}\,dt. \tag{30}$$

In (30) the integral quantity is the nonnegative "information divergence" between $\hat{\lambda}$ and the uniform normalized intensity $1/T$ [3]. This can be interpreted as a (assymetric) distance measure between the two intensities.

An upper bound on C_2 is obtained by recalling the decomposition formula for the capacity of a channel with input Z and output X, which is the sum of a "signal" $S=g(Z)$ plus an independent additive noise w: $X=g(Z)+w$. Specifically:

$$\begin{aligned}C_2&=\sup_{P(X|Z)}\{H(X)-H(X|Z)\}\\&=\sup_{P(X|Z)}\{H(X)-H(w|Z)\}\\&=\sup_{P(X|Z)}\{H(X)\}-H(w),\end{aligned} \tag{31}$$

where $H(X)$ and $H(X|Z)$ are the entropy of X and the conditional entropy of X given the input Z respectively, $P(X|Z)$ is a conditional distribution, and $H(w)$ is the entropy of the noise. In (31) we have used the independence of Z and w to equate the conditional entropy $H(w|Z)$ to the noise entropy $H(w)$, which is independent of $P(X|Z)$. On the other hand, it is well known that, for a fixed output autocovariance function $K_X=K_s+K_w$, the entropy $H(X)$ is maximized for Gaussian X [3]. Hence the channel capacity C_2 is upper bounded by the capacity of a Gaussian channel.

Here Z can be identified with the point process sequence $\{t_i\}_{i=1}^n$ and $S(t)=g(Z)=\sum_{i=1}^n p(t-t_i)$, where $n=N(t)$. The autocovariance $K_s(z_1,z_2)$, at times z_1 and z_2, is computed in Appendix C (C.4):

$$\begin{aligned}K_s(z_1,z_2)=&\int_0^{\min\{z_1,z_2\}}p(z_1-u)p(z_2-u)\bar{\lambda}(u)\,du\\&+\int_0^{z_1}du_1\int_0^{z_2}du_2\,p(z_1-u_1)p(z_2-u_2)\\&\cdot\mathrm{cov}\left[\lambda(u_1-\tau),\lambda(u_2-\tau)\right],\end{aligned} \tag{32}$$

where $\bar{\lambda}(u)\overset{\mathrm{def}}{=}E[\lambda(u-\tau)]$; and the expectation in "cov" is over the random variable τ. For the special case that $K_s(z_1,z_2)$ depends only on z_1-z_2, i.e., S is covariance stationary, C_2 is upper bounded by the capacity of the stationary Gaussian channel with capacity C_2^* [3]:

$$C_2\le C_2^*\overset{\mathrm{def}}{=}\frac{1}{4\pi}\int_{-\infty}^{\infty}\ln\left(1+\frac{2G_s(\omega)}{N_o}\right)d\omega, \tag{33}$$

where G_s is the power spectral density of the signal S, i.e., the Fourier transform of the covariance K_s. For the general case of nonstationary K_s, a more complicated channel capacity formula involving the eigenvalues of the Karhunen Loeve expansion associated with K_s must be used [3].

Combination of (30) and (33) gives, from (29), the rate distortion lower bound:

$$\mathrm{mse}\ge B_{rdlb}\overset{\mathrm{def}}{=}\frac{1}{2\pi e}e^{2H(\tau)}e^{-2\min\{C_1^*,C_2^*\}}. \tag{34}$$

The rate distortion lower bound (34) has some interesting features. Assume that τ is uniformly distributed over $[0,T]$ and that $p(t)$ is a causal impulse response. In this case (see Appendix C) $\bar{\lambda}(t)\approx\Lambda/T$, $t\in[0,T]$ and, as $T\to\infty$,

$$G_s(\omega)=\overline{p^2}\hat{G}_s(\omega), \tag{35}$$

where

$$\begin{aligned}\hat{G}_s(\omega)&\overset{\mathrm{def}}{=}|\hat{P}(\omega)|^2\left[\Lambda+|\mathscr{F}(\lambda(t)-\bar{\lambda}(t))|^2\right]\\&=|\hat{P}(\omega)|^2\left[\Lambda+|\Lambda(\omega)-\Lambda\,\mathrm{sinc}(\omega T/2)|^2\right],\end{aligned} \tag{36}$$

and $\mathrm{sinc}(x)\overset{\mathrm{def}}{=}\sin(x)/x$. In (36), $\Lambda(\omega)=\mathscr{F}(\lambda(t))$ is the Fourier transform of λ, hence $\Lambda(0)=\Lambda$, the integrated rate of the point process. Also the pulse energy normalized Fourier transform of the superposition filter $p(t)$ has been defined: $\hat{P}(\omega)\overset{\mathrm{def}}{=}\mathscr{F}(p(t))/\sqrt{\overline{p^2}}=P(\omega)/\sqrt{\overline{p^2}}$. Note that the minimum value $G_s(\omega)=\overline{p^2}|\hat{P}(\omega)|^2\Lambda$ is attained for the case that $\lambda(t)=\bar{\lambda}(t)$. Using (35) in the bound (34) we obtain the final form:

$$B_{rdlb}=\begin{cases}\dfrac{1}{2\pi e}e^{2H(\tau)}\exp\left(-2\Lambda\displaystyle\int_0^T\hat{\lambda}(t)\ln\frac{\hat{\lambda}(t)}{1/T}dt\right), & \gamma>\gamma_o,\\ \dfrac{1}{2\pi e}e^{2H(\tau)}\exp\left(-\dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty}\ln\left(1+\gamma\hat{G}_s(\omega)\right)d\omega\right), & \gamma\le\gamma_o,\end{cases}\tag{37}$$

where γ_o is a PNR threshold determined by the condition $C_1^*=C_2^*$ in (34). Specifically γ_o is the solution of the equation:

$$\Lambda\int_0^T\hat{\lambda}(t)\ln\frac{\hat{\lambda}(t)}{\frac{1}{T}}dt=\frac{1}{4\pi}\int_{-\infty}^{\infty}\ln\left(1+\gamma_o\left|\hat{G}_s(\omega)\right|^2\right)d\omega,\tag{38}$$

when the solution exists. An important case where the solution of (38) does not exist is when the superposition filter $p(t)$ approaches a delta function, i.e., $|\hat{P}(\omega)|^2=constant$. In this case, since the right-hand side of (38) diverges, $C_1^*<C_2^*$, and hence $\min\{C_1^*,C_2^*\}=C_1^*$, for all $\gamma>0$. Therefore, $\gamma_o=0$ for a zero width pulse $p(t)$ and nonzero PNR, and the lowerbound (37) is identical to the rate distortion lower bound of [14] for the pure Poisson observation $X=\{t_i\}_{i=1}^n$.

The lower bound (37) separates the mse performance into two PNR regimes: the Poisson limited regime ($\gamma>\gamma_o$), and the Gaussian limited regime ($\gamma\le\gamma_o$). In the high PNR Poisson limited regime, we have a PNR independent bound that is independent of the superposition filter $p(t)$ and depends principally on the information divergence between $\hat{\lambda}$ and the worst case uniform intensity $1/T$ over $[0,T]$. The closer $\hat{\lambda}$ is to the uninformative uniform intensity, the poorer becomes the estimator mse. This Poisson limited bound decays to zero at an exponential rate as the point process rate Λ increases. This rate is controlled by the information divergence, which is the magnitude difference between the entropy of the normalized intensity $\hat{\lambda}$ and the normalized maximum entropy uniform intensity. On the other hand, in the low PNR Gaussian limited regime, we have a bound that depends on the pulse shape through its magnitude Fourier transform, and depends on the intensity function through the magnitude squared difference between the Fourier transforms of λ and the uniform intensity $\bar{\lambda}$. Observe that, unlike the Poisson limited case, the decrease of the rate distortion bound for $\gamma<\gamma_o$ is subexponential in Λ for large Λ. In both the Poisson and the Gaussian limited regimes the effect of prior information on τ is manifested through the quantity: $\exp(2H(\tau))/(2\pi e)$, which is the *a priori* "entropy power" of τ [3]. The entropy power of τ is the variance of a Gaussian random variable with equal entropy as τ. Thus the bound (37) provides an indication of the *a posteriori* variance reduction achievable by making use of the observations X.

VI. Numerical Comparisons

In this section the mse performance of the MAP estimator (11), the linear "optical matched filter" (15), and the lower bounds (27) and (37) are investigated numerically.

Using the expressions (21) and (24) with the identification $\ln L(X_i|t)=X(t)*h(t)$, we have the following forms for local bias and local variance of a linear estimator structure of the form (14):

$$b(\tau)=\frac{-\int_0^{T_o}h'(-t)(\hat{\lambda}(t)*\hat{p}(t))\,dt}{\int_0^{T_o}h''(-t)(\hat{\lambda}(t)*\hat{p}(t))\,dt},\tag{39}$$

$$\sigma^2(\tau)=\frac{1}{M}\frac{\int_0^{T_o}dt_1\int_0^{T_o}dt_2\,h'(-t_1)h'(-t_2)\left[K_s'(t_1,t_2)+\frac{N_o}{2}\delta(t_1-t_2)\right]}{\left[\int_0^{T_o}h''(-t)(\hat{\lambda}(t)*\hat{p}(t))\,dt\right]^2},\tag{40}$$

where K_s' is the covariance:

$$K_s'(z_1,z_2)=\int_0^T p(t-z_1)p(t-z_2)\lambda(t)\,dt.$$

A numerical study was performed for the case that the signal intensity function λ_s is a Gaussian pulse with one sided standard width T_λ and $M=1$. The superposition filter response $p(t)$ is a one sided decaying exponential with time constant T_p. Due to the symmetry of λ, both the approximate MAP estimator (11) and the optical matched filter (15) are locally unbiased. Fig. 5 shows the local mse of the MAP estimator (24) and the local mse of the optical matched filter (40) as functions of Poisson rate $\Lambda=\Lambda_s+\Lambda_o$ for the following parameters: Poisson signal-to-noise energy ratio $\rho\overset{\text{def}}{=}\Lambda_s/\Lambda_o=50$ dB; PNR $\gamma=0$ dB; intensity time-width to filter time-width $T_\lambda/T_p=33.3$; *a priori* interval-width to intensity time-width $T/T_\lambda=100$. Observe that for low Poisson rate, $\Lambda\ll 5$ dB, the matched filter and the MAP estimator have similar local mse performance. For higher Λ, however, the MAP estimator has uniformly smaller local mse than the linear matched filter. The improvement of the MAP local mse over the matched filter local mse can be over 10 dB for high Λ. While this dramatic improvement may not occur in the global mse, which takes large errors into account, Fig. 5 is suggestive of performance gains. Also shown for comparison are the CR bound (27) and the rate distortion bound (37). For Λ below a rate threshold of 10 dB, the rate distortion lower bound is a tighter bound. This threshold specifies the lower boundary of the Λ region over which

Authorized licensed use limited to: University of Michigan Library. Downloaded on July 28, 2009 at 11:45 from IEEE Xplore. Restrictions apply.

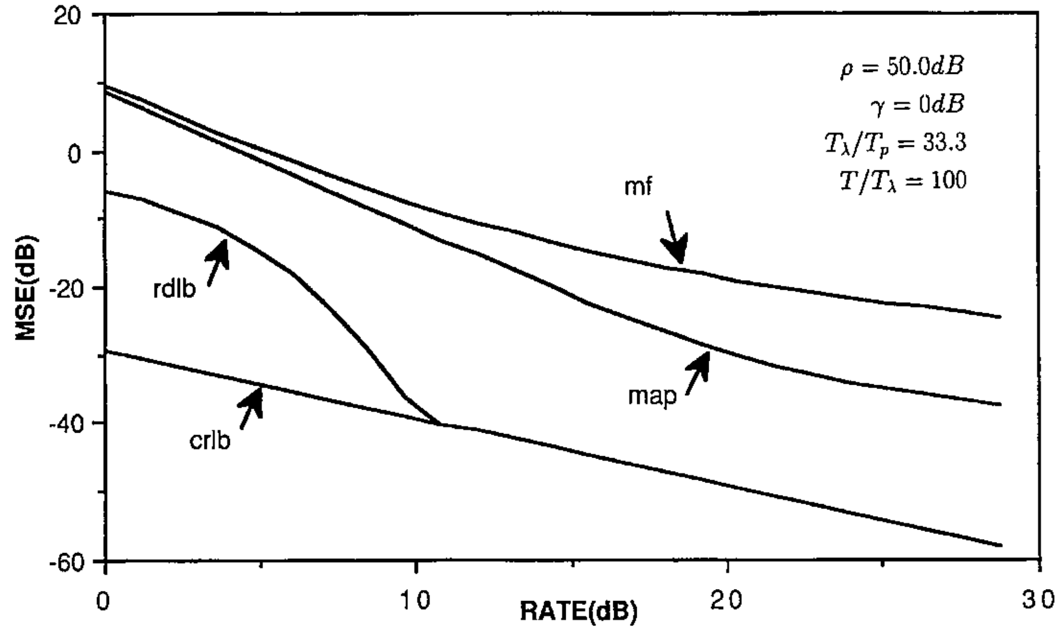

Fig. 5. Local mse approximations for matched filter (mf), MAP (map), rate distortion lower bound (rdlb), and Cramer–Rao lower bound (crlb), as functions of rate Λ of Poisson process. Poisson SNR, denoted ρ, is 50 dB and PNR = γ is 0 dB. mse axis has been normalized so that 0.0 corresponds to the mse of uniform random variable over *a priori* interval [0, 100].

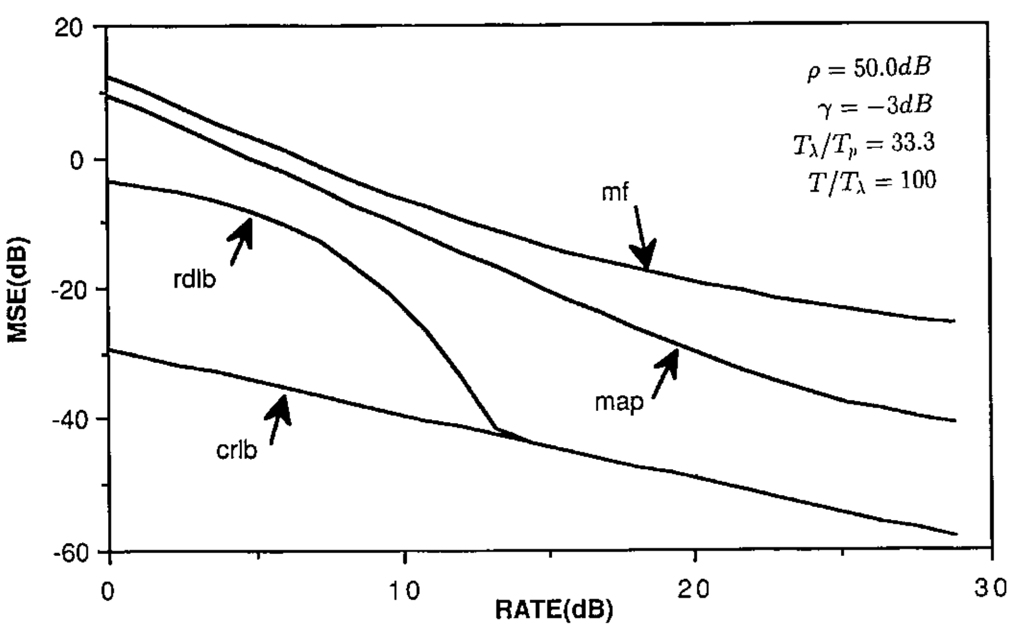

Fig. 6. Local mse approximations and lower bounds as function of rate Λ. Poisson SNR, denoted ρ, is 50 dB and PNR = γ is -3 dB. mse axis is normalized as in Fig. 5.

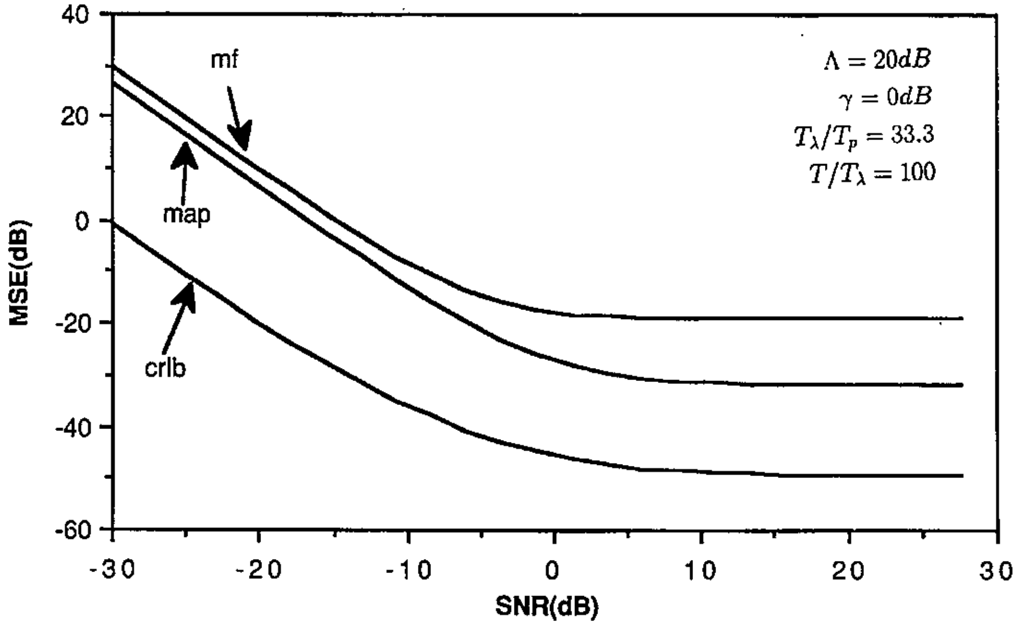

Fig. 7. Local mse approximations and the CR bound as functions of Poisson signal-to-noise ratio $\rho = \Lambda_s/\Lambda_n$. $\Lambda = \Lambda_s + \Lambda_n$ is 20 dB. mse axis is normalized as in Fig. 5.

small errors are theoretically attainable. A decrease of Λ below this threshold implies a sudden and precipitous increase in the achievable mse. In Fig. 6 the PNR has been decreased to -3 dB below the PNR of Fig. 5. The effect of the decreased PNR on the lower bound is a raising of the low Λ threshold that decreases the range of Λ over which small mse is theoretically achievable. The local mse and the CR bound as functions of the Poisson signal-to-noise ratio ρ for a fixed total rate of $\lambda = 20$ dB and PNR $= 0$ dB are shown in Fig. 7. The remaining parameters are identical to the ones used to generate Fig. 5. Note that the local mse of the MAP estimator has similar appearance to the CR bound, both curves displaying abrupt thresholds at approximately the same value of SNR. As before, the local mse performance of the MAP estimator is uniformly lower than that of the matched filter estimator.

For these numerical evaluations, the PNR-rate product $\gamma\Lambda$ had to be held sufficiently low to avoid numerical overflow in the numerator and denominator of the MAP local mse expression (24). For the values of $\gamma\Lambda$ studied, it was observed that in the exponent of the rate distortion bound (34), we have $C_1^* > C_2^*$; that is, the bound was active only for the Gaussian limited regime. At higher values of PNR the Poisson limited rate distortion bound would normally become active. Additional numerical studies of the rate distortion lower bound are presented in [13].

VII. Conclusion

Approximations to the likelihood function, the MAP estimator, and the mmse estimator have been given for the time-shifted intensity function of a causally filtered Poisson process observed in additive Gaussian noise. The approximation becomes more accurate as the per-unit-time density of superimposed filter impulse responses becomes small. The MAP estimator has a simple nonlinear structure as a function of the observations. First, it attempts to enhance the filtered Poisson process via classical matched filtering. Then, a memoryless nonlinear transformation produces a spike train that emulates the underlying point process. Finally, the spike train is correlated against shifted versions of the intensity function. The maximizing shift provides a time delay estimate. For smooth intensities, a local analysis of the bias and mse of the approximate likelihood ratio statistic indicated the conditions under which good estimation and detection performance can be obtained: 1) a high pulse-to-noise power ratio; 2) a high Poisson signal-to-Poisson noise ratio; 3) a large second derivative of the intensity function over regions where the amplitude of the intensity is large.

A rate-distortion type lower bound on the mse of any time estimator was derived. This bound is nontrivial for some important cases where the CR bound gives trivial results. The rate distortion bound indicates the importance of several factors on the inherent estimatibility of the time delay. First, for high pulse-to-noise ratio, the mse bound is Poisson noise limited and it decreases exponentially as a function of Poisson rate, where the rate of decay is the "information divergence" between the intensity function and a uniform intensity over time. The higher the information divergence, i.e., the more the intensity function differs from a uniform intensity, the better the potential mse performance. Second, for low pulse-to-noise power ratio, the mse bound decreases subexponentially as a function of Poisson rate, where the mean-squared difference between the intensity and a uniform intensity over time, and the shape of the superposition filter impulse response govern the rate of decay. The more broadband the filter transfer function, the better is the potential mse performance. Third, there is a pulse-to-noise ratio (PNR) threshold that specifies a PNR boundary between Poisson limited and Gaussian limited mse performance.

A numerical evaluation of the local mse approximations and the lower bounds indicates that the rate distortion bound is tighter than the CR bound for the large error regime, while the opposite is true for the small error regime. A comparison between the local mse of the approximate MAP estimator and the linear "optical matched filter" indicates that the MAP estimator can have better local error performance than the optimal linear filter. A large error sensitivity analysis should be performed to more completely characterize the advantages of the approximate MAP structure for detection and estimation of time shift.

Acknowledgment

The author gratefully acknowledges helpful discussions with W. L. Rogers and N. H. Clinthorne of the Division of Nuclear Medicine, University Hospital, Ann Arbor, MI.

Appendix A
Derivation of Approximate Likelihood Function

We recall the following fact about a Poisson process with intensity λ [27]. Given the total number of points n over $[0,T]$, the (unordered) occurrence times $\{t_i\}_{i=1}^n$ are independent identically distributed random variables over $[0,T]$ with marginal probability density functions $f(t_i)=\lambda(t_i)/\Lambda$, where $\Lambda=\int_0^T\lambda(t)\,dt$. The following identity will be needed [27]:

$$E_{t_1\cdots t_n}\left[\prod_{i=1}^n Q(t_i)\right]=e^{\int_0^T\lambda(t)(Q(t)-1)\,dt},\qquad \text{(A.1)}$$

where $Q(t)$ is an arbitrary (integrable) function. If Q is independent of t, (A.1) specializes to the useful formula:

$$E_{t_1\cdots t_n}[Q^n]=e^{\Lambda(Q-1)}.\qquad \text{(A.2)}$$

Another identity that will be useful is

$$E_{t_1\cdots t_n}\left[\sum_{i=1}^n p(t-t_i)\right]=\int_0^T\lambda(u)p(t-u)\,du.\qquad \text{(A.3)}$$

In this Appendix approximations to the numerator and denominator in (8) of the likelihood ratio $L(X|\tau)=f(X|H_1,\tau)/f(X|H_0)$ are derived. In the following $I(A)$ denotes the indicator function of the set A, $R_p=\int_0^T p(u+t)p(u)\,du$ is the pulse auto-correlation function, and $\overline{p^2}=R_p(0)$ is the energy of the pulse $p(t)$. We start from (7) by expanding the double sum $(\Sigma)^2$ and using, by definition of T_p, $R_p(t)=0$ for $|t|>T_p$:

$$\begin{aligned}
f(X|H_1,\tau)&\\
&=E_{t_1\cdots t_n}\left[e^{(2/N_o)\int_0^T X(t)\sum_{i=1}^n p(t-t_i)\,dt-(1/N_o)\int_0^T(\sum_{i=1}^n p(t-t_i))^2\,dt}\Big|\tau\right]\\
&=E_{t_1\cdots t_n}\left[e^{(2/N_o)\sum_{i=1}^n(\int_0^T X(t)p(t-t_i)\,dt-\overline{p^2}/2)-(1/N_o)\sum_{i\neq j}^n R_p(t_i-t_j)}\Big|\tau\right]\\
&=E_{t_1\cdots t_n}\left[e^{(2/N_o)\sum_{i=1}^n(\int_0^T X(t)p(t-t_i)\,dt-\overline{p^2}/2)}\prod_{i\neq j}I\left[|t_i-t_j|>T_p\right]\Big|\tau\right]\\
&\quad+E_{t_1\cdots t_n}\left[e^{(2/N_o)\sum_{i=1}^n(\int_0^T X(t)p(t-t_i)\,dt-\overline{p^2}/2)-(1/N_o)\sum_{i\neq j}^n R_p(t_i-t_j)}\left[1-\prod_{i\neq j}I\left[|t_i-t_j|>T_p\right]\right]\Big|\tau\right]\\
&=E_{t_1\cdots t_n}\left[e^{(2/N_o)\sum_{i=1}^n(\int_0^T X(t)p(t-t_i)\,dt-\overline{p^2}/2)}\Big|\tau\right]\\
&\quad+E_{t_1\cdots t_n}\left[G\left(X,\{t_i\}_{i=1}^n\right)\left[1-\prod_{i\neq j}^n I\left[|t_i-t_j|>T_p\right]\right]\Big|\tau\right],
\end{aligned}\qquad \text{(A.4)}$$

where

$$G\left(X,\{t_i\}_{i=1}^n\right)\stackrel{\text{def}}{=}e^{(2/N_o)\sum_{i=1}^n(\int_0^T X(t)p(t-t_i)\,dt-\overline{p^2}/2)}\cdot\left[e^{-(1/N_o)\sum_{i\neq j}^n R_p(t_i-t_j)}-1\right].\qquad \text{(A.5)}$$

The random variable $\prod_{i\neq j}I[|t_i-t_j|>T_p]$ is equivalent to the random variable $\prod_{i=1}^{n-1}I[z_i>T_p]$, where z_k is the increment from the kth largest occurrence time to the $(k+1)$st largest occurrence time. Observe the following chain of inequalities:

$$\begin{aligned}
E&\left\{1-\prod_{i=1}^{n-1}I\left[z_i>T_p\right]\Big|\tau\right\}\\
&\le E\left\{\sum_{i=1}^{n-1}E\left[I\left[z_i\le T_p\right]\Big|\tau,n\right]\right\}\\
&=E\left\{\sum_{i=1}^{n-1}P\left(N(t_i,t_i+T_p)>0|\tau,n\right]\right\}\\
&\le E\left\{\sum_{i=1}^{n-1}\max_t P\left(N(t,t+T_p)>0|\tau,n\right)\right\}\\
&\le E\left\{n\left[1-\min_t P\left(N(t,t+T_p)=0|\tau,n\right)\right]\right\}\\
&=E\left\{n\left[1-\min_t\left(1-\frac{\Lambda(t,t+T_p)}{\Lambda}\right)^n\right]\right\}\\
&=\Lambda\left[1-\min_t\left(1-\frac{\Lambda(t,t+T_p)}{\Lambda}\right)e^{-\Lambda(t,t+T_p)}\right]\\
&=O\left(\max_t\Lambda(t,t+T_p)\right),
\end{aligned}\qquad \text{(A.6)}$$

where $\Lambda(t,v)\stackrel{\text{def}}{=}\int_t^v\lambda(u-\tau)\,du$ and $\lim_{u\to\infty}O(u)=0$. Defining

$$\lambda^+\stackrel{\text{def}}{=}\max_t\frac{1}{T_p}\int_t^{t+T_p}\lambda(u-\tau)\,du=\max_t\frac{1}{T_p}\int_t^{t+T_p}\lambda(u)\,du,$$

Authorized licensed use limited to: University of Michigan Library. Downloaded on July 29, 2009 at 11:45 from IEEE Xplore. Restrictions apply.

after applying the Schwarz inequality and (A.6) to the second term on the right-hand side of (A.4) it is seen that

$$E^2_{t_1\cdots t_n}\left[G(X,\{t_i\}_{i=1}^n)\left[1-\prod_{i\neq j}^n I[|t_i-t_j|>T_p]\right]\Big|\tau\right]$$
$$\le E_{t_1\cdots t_n}\left[G^2(X,\{t_i\}_{i=1}^n)|\tau\right]$$
$$\cdot E_{t_1\cdots t_n}\left[1-\prod_{i=1}^{n-1} I[z_i>T_p]\Big|\tau\right]$$
$$=O(\Lambda T_p\lambda^+). \tag{A.7}$$

Hence, using (A.7) in (A.4):

$$f(X|H_1,\tau)$$
$$=E_{t_1\cdots t_n}\left[\exp\left(\frac{2}{N_o}\sum_{i=1}^n\left(\int_0^T X(t)p(t-t_i)\,dt-\frac{1}{2}\overline{p^2}\right)\right)\Big|\tau\right]$$
$$+O(\Lambda T_p\lambda^+). \tag{A.8}$$

Thus to order $O(\Lambda T_p\lambda^+)$ we have, using the identity (A.1),

$$f(X|H_1,\tau)$$
$$=E_{t_1\cdots t_n}\left[\prod_{i=1}^n e^{(2/N_o)(\int_0^T X(t)p(t-t_i)dt-\overline{p^2}/2)}\Big|\tau\right]$$
$$=\exp\left(\int_0^T\left(e^{(2/N_o)(\int_0^T X(u)p(u-t)du-\overline{p})}-1\right)\lambda(t-\tau)\,dt\right)$$
$$=\exp\left(\int_0^T\left(e^{(2/N_o)\int_0^T X(u)p(u-t)du-\gamma/2}-1\right)\lambda(t-\tau)\,dt\right). \tag{A.9}$$

To obtain an order $O(\Lambda T_p\lambda^+)$ approximation to $f(X|H_0)$, simply replace $\lambda(t-\tau)$ in (A.9) by the constant Poisson intensity λ_o:

$$f(X|H_0)=\exp\left(\int_0^T\left(e^{(2/N_o)\int_0^T X(u)p(u-t)du-\gamma/2}-1\right)\lambda_o\,dt\right). \tag{A.10}$$

Recalling the relation (1), $\lambda(t-\tau)=\lambda_s(t-\tau)+\lambda_o$, the ratio of (A.9) and (A.10) reduces to the expression (9).

Appendix B
Derivation of Moments of Log-Likelihood Derivatives

In this Appendix we calculate the numerator and denominator quantities on the right-hand sides of (20) and (19) using the approximate form for the log-likelihood given in (9):

$$\ln L(X|\tau)$$
$$=\int_0^T\left(e^{(2/N_o)\int_0^T X(u)p(u-t)du-(1/N_o)\overline{p^2}}-1\right)\lambda_s(t-\tau)\,dt$$
$$=e^{-\overline{p^2}/N_o}\int_0^T e^{(2/N_o)\int_0^T X(u)p(u-t)du}\lambda_s(t-\tau)\,dt-\Lambda. \tag{B.1}$$

For notational simplicity, we will denote $T_i=T$ and $X_i=X$ where convenient.

Take the first derivative with respect to τ of the log-likelihood function approximation (B.1) and take the expectation

$$E\left[\frac{d\ln L(X_i|\tau)}{d\tau}\Big|\tau\right]$$
$$=-e^{-(\overline{p^2}/N_o)}\int_0^T E\left[e^{(2/N_o)\int_0^T X(u)p(u-t)du}|\tau\right]\lambda'(t-\tau)\,dt, \tag{B.2}$$

where we have used the fact $\lambda_s'=\lambda'$. Consider the expectation within the $\int$ on the right-hand side of (B.2). We have by definition, $X=\sum_{i=1}^n p(t-t_i)+w$, and hence:

$$E\left[e^{(2/N_o)\int_0^T X(u)p(u-t)du}|\tau\right]$$
$$=E\left[e^{(2/N_o)\sum_{i=1}^n\int_0^T p(u-t_i)p(u-t)du}\right.$$
$$\left.\cdot E\left[e^{(2/N_o)\int_0^T w(u)p(u-t)du}\log\{t_i\}_{i=1}^n,n,\tau\right]\Big|\tau\right]. \tag{B.3}$$

Now since, by assumption, the Gaussian noise w is independent of the Poisson occurrence times $\{t_i\}_{i=1}^n$ and τ, we have

$$E\left[e^{(2/N_o)\int_0^T w(u)p(u-t)du}\big|\{t_i\}_{i=1}^n,n,\tau\right]$$
$$=E\left[e^{(2/N_o)\int_0^T w(u)p(u-t)du}\right]. \tag{B.4}$$

Now the exponent of the argument of "E" in (B.4) is a zero-mean Gaussian random variable with variance $\sigma^2=\overline{p^2}/N_o$. Using the well-known form of the characteristic function of such a Gaussian random variable, $\exp(\sigma^2/2)$, we have from (B.4)

$$E\left[e^{(2/N_o)\int_0^T w(u)p(u-t)du}\big|\{t_i\}_{i=1}^n,n,\tau\right]=e^{\overline{p^2}/N_o}. \tag{B.5}$$

Substitution of (B.5) into (B.3) gives:

$$E\left[e^{(2/N_o)\int_0^T X(u)p(u-t)du}|\tau\right]$$
$$=e^{\overline{p^2}/N_o}E\left[e^{(2/N_o)\sum_{i=1}^n\int_0^T p(u-t_i)p(u-t)du}|\tau\right]$$
$$=e^{\overline{p^2}/N_o}E\left[\prod_{i=1}^n e^{(2/N_o)\int_0^T p(u-t_i)p(u-t)du}\Big|\tau\right]$$
$$=e^{\overline{p^2}/N_o}\exp\left(\int_0^T\left(e^{(2/N_o)\int_0^T p(u-z)p(u-t)du}-1\right)\right.$$
$$\left.\cdot\lambda(z-\tau)\,dz\right)$$
$$=e^{\overline{p^2}/N_o}\exp\left(\int_0^T\left(e^{(2/N_o)R_p(t-z)}-1\right)\lambda(z-\tau)\,dz\right)$$
$$=e^{\overline{p^2}/N_o-\Lambda}\exp\left(\int_0^T e^{(2/N_o)R_p(t-u)}\lambda(u-\tau)\,du\right), \tag{B.6}$$

where in the fourth line of (B.6) the identity (A.1) has been used. Substitution of (B.6) into (B.2) gives the following:

$$E\left[\frac{d\ln L(X_i|\tau)}{d\tau}\Big|\tau\right]$$
$$=-e^{-\Lambda}\int_0^T\exp\left(\int_0^T e^{(2/N_o)R_p(u-t)}\lambda(u-\tau)\,du\right)\lambda'(t-\tau)\,dt$$
$$=-e^{-\Lambda}\int_0^T\exp\left(\int_0^T e^{(2/N_o)R_p(u-t)}\lambda(u)\,du\right)\lambda'(t)\,dt, \tag{B.7}$$

The calculation of

$$E\left[-\frac{d^2\ln L(X_i|\tau)}{d\tau^2}\Big|\tau\right]$$

Authorized licensed use limited to: University of Michigan Library. Downloaded on July 28, 2009 at 11:45 from IEEE Xplore. Restrictions apply.

is similar to the calculation of

$$E\left[\frac{d\ln L(X_i|\tau)}{d\tau}\bigg|\tau\right];$$

the details are omitted. The result is

$$E\left[-\frac{d^2\ln L(X_i|\tau)}{d\tau^2}\bigg|\tau\right]$$
$$= -e^{-\Lambda}\int_0^T \exp\left(\int_0^T e^{(2/N_o)R_p(u-t)}\lambda(u)\,du\right)\lambda''(t)\,dt. \quad \text{(B.8)}$$

Next the numerator of (19) is derived. From (B.2):

$$E\left[\left(\frac{d\ln L(X_i|\tau)}{d\tau}\right)^2\bigg|\tau\right]$$
$$= e^{-2\overline{p^2}/N_o}\int_0^T dz_1\int_0^T dz_2$$
$$\cdot E\left[e^{(2/N_o)(\int_0^T X(u)p(u-z_1)\,du+\int_0^T X(u)p(u-z_2)\,du)}\Big|\tau\right]$$
$$\cdot\lambda'(z_1-\tau)\lambda'(z_2-\tau). \quad \text{(B.9)}$$

Consider the expectation on the right-hand side of (B.9). In a manner analogous to the derivation of

$$E[d\ln L(X_i|\tau)/d\tau|\tau]$$

find

$$E\left[e^{(2/N_o)(\int_0^T X(u)p(u-z_1)\,du+\int_0^T X(u)p(u-z_2)\,du)}\Big|\tau\right]$$
$$= E\left[e^{(2/N_o)\sum_{i=1}^n\int_0^T p(u-t_i)[p(u-z_1)+p(u-z_2)]\,du}\right.$$
$$\left.\cdot E\left[e^{(2/N_o)\int_0^T w(u)[p(u-z_1)+p(u-z_2)]\,du}\Big|\{t_i\}_{i=1}^n, n, \tau\right]\Big|\tau\right]$$
$$= E\left[e^{(2/N_o)\sum_{i=1}^n[R_p(t_i-z_1)+R_p(t_i-z_2)]}\right.$$
$$\left.\cdot E\left[e^{(2/N_o)\int_0^T w(u)[p(u-z_1)+p(u-z_2)]\,du}\right]\Big|\tau\right]$$
$$= e^{2\overline{p^2}/N_o+(1/N_o)R_p(z_1-z_2)}$$
$$\cdot E\left[e^{(2/N_o)\sum_{i=1}^n[R_p(t_i-z_1)+R_p(t_i-z_2)]}\Big|\tau\right]$$
$$= e^{2\overline{p^2}/N_o+(1/N_o)R_p(z_1-z_2)}$$
$$\cdot\exp\left(\int_0^T\left(e^{(2/N_o)[R_p(u-z_1)+R_p(u-z_2)]}-1\right)\lambda(u-\tau)\,du\right)$$
$$= e^{2\overline{p^2}/N_o-\Lambda+(1/N_o)R_p(z_1-z_2)}$$
$$\cdot\exp\left(\int_0^T e^{(2/N_o)[R_p(u-z_1)+R_p(u-z_2)]}\lambda(u)\,du\right). \quad \text{(B.10)}$$

Substitution of (B.10) into (B.9) gives, after some rearrangement of terms,

$$E\left[\left(\frac{d\ln L(X_i|\tau)}{d\tau}\right)^2\bigg|\tau\right]$$
$$= e^{-\Lambda}\int_0^T dt_1\int_0^T dt_2$$
$$\cdot\exp\left(\int_0^T e^{(2/N_o)[R_p(u-t_1)+R_p(u-t_2)]}\lambda(u)\,du\right.$$
$$\left.+\frac{1}{N_o}R_p(t_1-t_2)\right)\lambda'(t_1)\lambda'(t_2). \quad \text{(B.11)}$$

Finally, it is shown that the local mse approximation (24) reduces to the expression (25) of Section IV-B under the small PNR assumption. For convenience we repeat the expression (24):

$$E[(\hat{\tau}-\tau)^2] = \frac{\int_0^{T_o}dt_1\int_0^{T_o}dt_2\exp\left(\Lambda\int_0^{T_o}[e^{\gamma[\hat{R}_p(u-t_1)+\hat{R}_p(u-t_2)]}-1]\hat{\lambda}(u)\,du+\gamma\hat{R}_p(t_1-t_2)\right)\hat{\lambda}'(t_1)\hat{\lambda}'(t_2)}{M\left[\int_0^{T_o}\exp\left(\Lambda\int_0^{T_o}[e^{\gamma\hat{R}_p(u-t)}-1]\hat{\lambda}(u)\,du\right)\hat{\lambda}''(t)\,dt\right]^2}, \quad \text{(B.12)}$$

where γ is the PNR defined in (10). Make the following low PNR substitutions in the numerator and denominator of (B.12):

$$\exp\left(\Lambda\int_0^{T_o}[e^{\gamma[\hat{R}_p(u-t_1)+\hat{R}_p(u-t_2)]}-1]\hat{\lambda}(u)\,du+\gamma\hat{R}_p(t_1-t_2)\right) \quad \text{(B.13)}$$
$$= 1+\gamma\Lambda\int_0^{T_o}\left[\hat{R}_p(u-t_1)+\hat{R}_p(u-t_2)\right]\hat{\lambda}(u)\,du$$
$$+\gamma\hat{R}_p(t_1-t_2), \quad \text{(B.14)}$$

and:

$$\exp\left(\Lambda\int_0^{T_o}[e^{\gamma\hat{R}_p(u-t)}-1]\hat{\lambda}(u)\,du\right)$$
$$= 1+\gamma\Lambda\int_0^{T_o}\hat{R}_p(u-t)\hat{\lambda}(u)\,du. \quad \text{(B.15)}$$

Since the integrals of λ' and λ'' are identically zero the substitution of the approximations (B.13) and (B.15) into (B.12) gives:

$$E[(\hat{\tau}-\tau)^2]$$
$$= \frac{1}{\gamma\Lambda^2}\,\frac{\int_0^{T_o}dt_1\int_0^{T_o}dt_2\hat{\lambda}'(t_1)\hat{R}_p(t_1-t_2)\hat{\lambda}'(t_2)}{M\left[\int_0^{T_o}dt_1\int_0^{T_o}dt_2\hat{\lambda}(t_1)\hat{R}_p(t_1-t_2)\hat{\lambda}''(t_2)\right]^2}. \quad \text{(B.16)}$$

Next note that, due to the assumption $T = T_o \gg T_\lambda$, and the differentiability of λ:

$$
\begin{aligned}
0 &= \frac{d^2}{d\tau^2}\int dt_1 \int dt_2\, \lambda(t_1) R_p(t_1 - t_2) \lambda(t_2) \\
&= \frac{d^2}{d\tau^2}\int dt_1 \int dt_2\, \lambda(t_1 - \tau) R_p(t_1 - t_2) \lambda(t_2 - \tau) \\
&= \int dt_1 \int dt_2\, \lambda'(t_1 - \tau) R_p(t_1 - t_2) \lambda'(t_2 - \tau) \\
&\quad + \int dt_1 \int dt_2\, \lambda(t_1 - \tau) R_p(t_1 - t_2) \lambda''(t_2 - \tau) \\
&= \int dt_1 \int dt_2\, \lambda'(t_1) R_p(t_1 - t_2) \lambda'(t_2) \\
&\quad + \int dt_1 \int dt_2\, \lambda(t_1) R_p(t_1 - t_2) \lambda''(t_2). \qquad \text{(B.17)}
\end{aligned}
$$

Finally, using the identity (B.17) in (B.18), we obtain

$$
\begin{aligned}
E\left[(\hat{\tau} - \tau)^2\right] &= \frac{1}{\gamma \Lambda^2 M} \frac{1}{\int_0^{T_o} dt_1 \int_0^{T_o} dt_2\, \hat{\lambda}'(t_1) \hat{R}_p(t_1 - t_2) \hat{\lambda}'(t_2)} \\
&= \frac{1}{\gamma \Lambda^2 M} \frac{-1}{\int_0^{T_o} dt_1 \int_0^{T_o} dt_2\, \hat{\lambda}(t_1) \hat{R}_p(t_1 - t_2) \hat{\lambda}''(t_2)}. \qquad \text{(B.18)}
\end{aligned}
$$

Appendix C
Computation of Covariance of Filtered Poisson Process

Here the autocovariance of the superposition $S(t) = \sum_{i=1}^{N(t)} p(t - t_i)$ is derived.

Since, conditioned on τ, $\{t_i\}_{i=1}^n$ is Poisson, application of the identity (A.3) gives

$$
E\left[\sum_{i=1}^{n} p(t - t_i)\right] = E\left[E\left[\sum_{i=1}^{n} p(t - t_i) \middle| \tau\right]\right]
$$

$$
E\left[\int_0^t p(t-u)\lambda(u-\tau)\, du\right] = \int_0^t p(t-u)\bar{\lambda}(u)\, du, \qquad \text{(C.1)}
$$

where $\bar{\lambda}(t) = E[\lambda(u - \tau)]$ and $n = N(t)$. Next consider the autocovariance function $K_\lambda(z_1, z_2)$:

$$
\begin{aligned}
&E\left[\sum_{i=1}^{N(z_1)} p(z_1 - t_i) \sum_{j=1}^{N(z_2)} p(z_2 - t_j)\right] \\
&\quad - \int_0^{z_1} p(z_1 - u)\bar{\lambda}(u)\, du \int_0^{z_2} p(z_2 - u)\bar{\lambda}(u)\, du \\
&= E\left[E\left[\sum_{i=1}^{N(z_1)} p(z_1 - t_i) \sum_{j=1}^{N(z_2)} p(z_2 - t_j) \middle| \tau\right]\right] \\
&\quad - \int_0^{z_1} p(z_1 - u)\bar{\lambda}(u)\, du \int_0^{z_2} p(z_2 - u)\bar{\lambda}(u)\, du. \qquad \text{(C.2)}
\end{aligned}
$$

The inner expectation of the first term on the right-hand side of (C.2) decomposes into a sum of two terms due to the independent increment property of $\{t_i\}_{i=1}^n$ [27]:

$$
\begin{aligned}
&E\left[\sum_{i=1}^{N(z_1)} p(z_1 - t_i) \sum_{j=1}^{N(z_2)} p(z_2 - t_j) \middle| \tau\right] \\
&= E\left[\sum_{i=1}^{N(\min\{z_1, z_2\})} p(z_1 - t_i) p(z_2 - t_i) \middle| \tau\right] \\
&\quad + \int_0^{z_1} p(z_1 - u)\lambda(u - \tau)\, du \cdot \int_0^{z_2} p(z_2 - u)\lambda(u - \tau)\, du \\
&= \int_0^{\min\{z_1, z_2\}} p(z_1 - u) p(z_2 - u) \lambda(u - \tau)\, du \\
&\quad + \int_0^{z_1} p(z_1 - u)\lambda(u - \tau)\, du \cdot \int_0^{z_2} p(z_2 - u)\lambda(u - \tau)\, du. \qquad \text{(C.3)}
\end{aligned}
$$

Finally, taking the expectation of (C.3) with respect to τ gives the covariance function:

$$
\begin{aligned}
K_s(z_1, z_2) &= \operatorname{cov}\left(\sum_{i=1}^{N(z_1)} p(z_1 - t_i) \sum_{i=1}^{N(z_2)} p(z_2 - t_i)\right) \\
&= \int_0^{\min\{z_1, z_2\}} p(z_1 - u) p(z_2 - u) \bar{\lambda}(u)\, du \\
&\quad + \int_0^{z_1} du_1 \int_0^{z_2} du_2\, p(z_1 - u_1) p(z_2 - u_2) \cdot E[\lambda(u_1 - \tau)\lambda(u_2 - \tau)] \\
&\quad - \int_0^{z_1} p(z_1 - u)\bar{\lambda}(u)\, du \int_0^{z_2} p(z_2 - u)\bar{\lambda}(u)\, du \\
&= \int_0^{\min\{z_1, z_2\}} p(z_1 - u) p(z_2 - u) \bar{\lambda}(u)\, du \\
&\quad + \int_0^{z_1} du_1 \int_0^{z_2} du_2\, p(z_1 - u_1) p(z_2 - u_2) \cdot \operatorname{cov}[\lambda(u_1 - \tau), \lambda(u_2 - \tau)], \qquad \text{(C.4)}
\end{aligned}
$$

where

$$
\operatorname{cov}[\lambda(u_1 - \tau), \lambda(u_2 - \tau)] = E\left[\lambda(u_1 - \tau) - \bar{\lambda}(u_1)\right]\left[\lambda(u_2 - \tau) - \bar{\lambda}(u_2)\right].
$$

Next it is shown that the covariance $K_s(z_1, z_2)$ of (C.4) depends only upon the time difference $z_1 - z_2$ under the following conditions: 1) $T \to \infty$; 2) uniformly distributed τ over $[0, T]$; 3) causal impulse response $p(t)$. Under these conditions we have the following results:

$$
\begin{aligned}
\bar{\lambda}(t) &= E[\lambda(t - \tau)] \\
&= \frac{1}{T}\int_0^T \lambda(t - \tau)\, d\tau \\
&= \frac{1}{T}\int_0^T (\lambda_s(t - \tau) + \lambda_o)\, d\tau \\
&= \frac{1}{T}\int_0^T (\lambda_s(\tau) + \lambda_o)\, d\tau \\
&= \frac{1}{T}\int_0^T \lambda(\tau)\, d\tau \\
&= \frac{\Lambda}{T}, \qquad 0 \le t \le T, \qquad \text{(C.5)}
\end{aligned}
$$

Authorized licensed use limited to: University of Michigan Library. Downloaded on July 28, 2009 at 11:45 from IEEE Xplore. Restrictions apply.

and

$$\begin{aligned}\operatorname{cov}[\lambda(u_1-\tau),\lambda(u_2-\tau)] &= E[(\lambda(u_1-\tau)-\bar{\lambda}(u_1))(\lambda(u_2-\tau)-\bar{\lambda}(u_2))] \\ &= \frac{1}{T}\int_0^T[\lambda(u_1-\tau)-\bar{\lambda}(u_1)][\lambda(u_2-\tau)-\bar{\lambda}(u_2)]\,d\tau.\end{aligned} \tag{C.6}$$

Hence, from the causality assumption,

$$\begin{aligned}K_s(z_1,z_2) &= \operatorname{cov}\left(\sum_{i=1}^{N(z_1)} p(z_1-t_i), \sum_{i=1}^{N(z_2)} p(z_2-t_i)\right) \\ &= \int_0^T p(z_1-u)p(z_2-u)\bar{\lambda}(u)\,du \\ &\quad + \int_0^T du_1\int_0^T du_2\, p(z_1-u_1)p(z_2-u_2) \\ &\quad \cdot\operatorname{cov}[\lambda(u_1-\tau),\lambda(u_2-\tau)].\end{aligned} \tag{C.7}$$

Using (C.6) in (C.4) and taking a two-dimensional Fourier transform over the arguments z_1 and z_2 gives

$$\mathscr{F}(K_s(z_1,z_2)) = \int_{-\infty}^{\infty} dz_1\int_{-\infty}^{\infty} dz_2 K_s(z_1,z_2)e^{-j\omega z_1}e^{-j\nu z_2} \tag{C.8}$$

$$\begin{aligned}&= P(\omega)P(\nu)[\Lambda(0)\operatorname{sinc}([\omega+\nu]T/2) \\ &\quad + \Lambda(\omega)\Lambda(\nu)\operatorname{sinc}([\omega+\nu]T/2) \\ &\quad - \Lambda^2(0)\operatorname{sinc}(\omega T/2)\operatorname{sinc}(\nu T/2)],\end{aligned} \tag{C.9}$$

where $\Lambda(\omega)$ and $P(\omega)$ are the Fourier transforms of $\lambda(t)$ and $p(t)$, and $\operatorname{sinc}(x) \stackrel{\text{def}}{=} \sin(x)/x$. As $T\to\infty$, $\operatorname{sinc}([\omega+\nu]T/2)=0$, unless $\omega=-\nu$, and $\operatorname{sinc}(\omega T/2)\operatorname{sinc}(\nu T/2)=\operatorname{sinc}^2(\omega T/2)$ if $\omega=\nu$ and zero otherwise. Therefore in the limit, over its nonzero region of definition (the diagonal $\omega=-\nu$), the two-dimensional Fourier transform (C.8) reduces to the one-dimensional Fourier transform in the frequency variable ω. Hence to an $O(1/T)$ approximation, the covariance function $K_s(z_1,z_2)$ has a one-dimensional Fourier transform over the difference z_1-z_2:

$$G_s(\omega) = |P(\omega)|^2\left[\Lambda(0)+|\Lambda(\omega)-\Lambda(0)\operatorname{sinc}(\omega T/2)|^2\right], \tag{C.10}$$

and hence K_s is a function of z_1-z_2. The final form for the Fourier transform of K_s used in Section V is obtained from (C.10) by definition of the energy normalized Fourier transform of $p(t)$, denoted $\hat{P}(\omega)$, and recognition of the D.C. value $\Lambda(0)$ as the energy Λ of the point process. The result is

$$G_s(\omega) = \overline{p^2}|\hat{P}(\omega)|^2\left[\Lambda+|\Lambda(\omega)-\Lambda\operatorname{sinc}(\omega T/2)|^2\right]. \tag{C.11}$$

References

[1] I. Bar-David, "Minimum-mean-square-error estimation of photon pulse delay," *IEEE Trans. Inform. Theory*, vol. IT-21, no. 3, pp. 326–330, May 1975.

[2] P. Billingsley, *Convergence of Probability Measures*. New York: Wiley, 1968.

[3] R. Blahut, *Applications of Information Theory*. Englewood Cliffs, NJ: Prentice Hall, 1987.

[4] B. Z. Bobrovsky, E. Mayer-Wolf, and M. Zakai, "Some classes of global Cramer-Rao bounds," *Ann. Statist.*, vol. 15, no. 4, pp. 1421–1438, 1987.

[5] A. M. Bruckstein, T. J. Shan, and T. Kailath, "The resolution of overlapping echos," *IEEE Trans. Acoust. Speech Signal Processing*, vol. ASSP-33, pp. 1357–1368, Dec. 1985.

[6] L. J. Dorfman, "The distribution of conduction velocities (DCV) in peripheral nerves: A review," *Muscle Nerve*, vol. 7, pp. 2–11, 1984.

[7] M. T. El-Hadidi and B. Hirosaki, "The Bayes optimal receiver for digital fibre optic communications systems," *Opt. Quantum Electron.*, vol. 13, pp. 469–486, 1981.

[8] P. Faure, "Theoretical models of reverberation noise," *J. Acoust. Soc. Am.*, vol. 36, pp. 259–268, 1964.

[9] R. M. Gagliardi and S. Karp, *Optical Communications*. New York: Wiley, 1976.

[10] F. Gatti and V. Svelto, "Review of theories and experiments of resolving time with scintillation counters," *Nucl. Instruments Meth.*, vol. 43, pp. 248–268, 1966.

[11] C. N. Georghiades and D. L. Snyder, "Locating data frames in direct-detection optical communications," *IEEE Trans. Commun.*, vol. COM-32, pp. 118–123, Feb. 1984.

[12] E. A. Geraniotis and H. V. Poor, "Robust matched filters for optical receivers," *IEEE Trans. Commun.*, vol. 35, pp. 1289–1296, Dec. 1987.

[13] A. O. Hero, "A rate distortion lower bound on phase errors for optical receivers," in *Proc. Conf. Inform. Sci. Syst.*, Princeton, NJ, 1988, pp. 140–145.

[14] ——, "Lower bounds on estimator performance for energy invariant parameters of multidimensional Poisson processes," *IEEE Trans. Inform. Theory*, vol. 35, no. 4, pp. 843–858, July 1989.

[15] A. O. Hero, N. Antoniadis, N. H. Clinthorne, W. L. Rogers, and G. D. Hutchins, "Optimal and sub-optimal post-detection timing estimators for PET," *IEEE Trans. Nucl. Sci.*, vol. 37, no. 2, pp. 725–736, Apr. 1990.

[16] E. V. Hoversten, D. L. Snyder, R. O. Harger, and K. Kurimoto, "Direct-detection optical communications receivers," *IEEE Trans. Commun.*, vol. COM-22, pp. 17–27, 1974.

[17] I. A. Ibragimov and R. Z. Has'minskii, *Statistical Estimation: Asymptotic Theory*. New York: Springer-Verlag, 1981.

[18] T. T. Kadota, "Approximately optimum detection of deterministic signals in Gaussian and compound Poisson noise," *IEEE Trans. Inform. Theory*, vol. 34, no. 6, pp. 1517–1527, Nov. 1988.

[19] T. Kailath, "A general likelihood ratio formula for random signals in Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-15, no. 3, pp. 350–361, 1969.

[20] H. Kwakernak, "Estimation of pulse heights and arrival times," *Automatica*, vol. 16, pp. 367–377, 1980.

[21] J. M. Mendel, "White noise estimators for seismic data processing in oil exploration," *IEEE Trans. Automat. Contr.*, vol. AC-22, no. 5, pp. 694–706, Oct. 1977.

[22] D. Middleton, *An Introduction to Statistical Communication Theory*. New York: McGraw-Hill, 1960.

[23] N. M. Moody, W. Paul, and M. L. Joy, "A survey of medical gamma-ray cameras," *IEEE Proc.*, vol. 58, Feb. 1970, pp. 217–242.

[24] J. J. O'Reilly, "Generating functions and bounds in optical communications," in *Problems of Randomness in Communications Engineering*, Ch. 7, K. Cattermole and J. O'Reilly, Eds. New York: Wiley, Nov. 1987, pp. 119–133.

[25] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1988.

[26] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*. New York: Wiley, 1980.

[27] D. L. Snyder, *Random Point Processes*. New York: Wiley, 1975.

[28] E. Wong, *Stochastic Processes in Information and Dynamical Systems*. Huntington, NY: Krieger, 1979.

Authorized licensed use limited to: University of Michigan Library. Downloaded on July 28, 2009 at 11:46 from IEEE Xplore. Restrictions apply.