On the statistics of binned neural point processes: the Bernoulli approximation and AR representation of the PST histogram

B. W. Edwards and G. H. Wakefield

Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-1102, USA

Abstract. Neural point processes are often approximated by partitioning time into bins, each with a Bernoulli distribution of firing, in order to simplify the mathematical description of their properties. Some of the basic statistics of a neural process are compared using the Bernoulli approximation and the actual Poisson representation. It is seen that in general the Bernoulli approximation is an accurate model only for small $\lambda\Delta$ where λ is the intensity and Δ is the width of the time bin. This discrete representation leads to a model of the PST histogram as an AR system, where the parameters depend upon the driving signal $s(t)$, the refractory effect $r(t)$ and the binwidth Δ. This AR representation is used to predict the PST histogram given $s(t)$, $r(t)$ and Δ. Estimates of $s(t)$ and $r(t)$ are derived within this parameterization and results discussed for several types of recovery functions given a constant $s(t)$. AR techniques are used to estimate the AR parameters from the PST histogram of a simulated point process, from which both $s(t)$ and $r(t)$ are estimated.

1 Introduction

Point process theory has become prominent in the study of neural systems because information is conveyed through the timing of discharge activity. Different models have been proposed for the generation of nerve discharge (Gerstein and Mandelbrot 1964; Geisler and Goldberg 1966), different mathematical structures for the temporal occurrence of discharge have been presented (Brillinger 1988; Johnson and Swami 1983; Chornoboy et al. 1988), as have different means of displaying specific properties of the process (Perkel et al. 1967; van Stokkum et al. 1986; Palm et al. 1988). These descriptions have been used when applying techniques such as the maximum likelihood estimator or the EM algorithm to estimate characteristics of the process (Miller 1985).

Frequently in such analysis the process under inves-

tigation is discretized in time by partitioning time into equally sized bins. A variable is assigned to each bin which is equal to the number of firings that occurred within that bin. The bins are often assumed to be sufficiently small such that the probability of the process firing more than once within the bin is negligible. This simplification will be termed the *Bernoulli approximation*, since each bin variable is assumed to be binary, and results in a Bernoulli process that is much easier to analyze mathematically.

The extent to which the simplifications introduced by the Bernoulli approximation lead to disparities between the derived statistics and those for the continuous-time case is unclear. For example, biases in the mean and higher moments of the number of firings within a bin may be introduced. Such biases might subsequently affect any estimator based on the discrete representation. The accuracy of the assumption is linked to the size of the bin since the larger the bin the less the Bernoulli distribution resembles the true distribution, in part because the probability of more than one firing within a bin becomes non-negligible. On the other hand, infinitesimally small bins will drastically reduce the number of firings within a bin which will cause the variance of any estimator that is dependent on the number of firings within the bin to be unacceptably large.

In Sect. 2.1, the effects of assuming a Bernoulli instead of a Poisson distribution will be examined. Some basic properties of a point process will be derived with and without the Bernoulli approximation. It will be seen under which conditions the Bernoulli approximation is valid, i.e., provides the same statistics as those derived using the exact Poisson distribution. It will also be seen that the expected number of firings within a bin is correctly predicted using the Bernoulli approximation regardless of bin size, but that higher moments and the correlation between the bins are not well approximated if the binwidths are too large.

For a stimulus that is presented repeatedly, the PST histogram indicates the relative time at which a discharge occurs with respect to the onset of the stimulus

(Gerstein and Kiang 1960). It is a simple and useful method of examining the first-order properties of the firing process from which the stimulus can be estimated. The PST histogram is an unbiased estimate of the *average* intensity $\lambda(t)$ of the point process during the stimulus presentation (Johnson 1978) and can be used to estimate the driving stimulus when the the form of an inhibitory refractory effect is known (Miller 1985). When using the PST histogram as an estimator, the Bernoulli approximation is often made, i.e., that the binwidths are small enough so that the process can only fire once within a bin during one stimulus presentation. The Bernoulli analysis of Sect. 2.1 cannot be extended to the PST histogram, however, because possible refractory effects destroy necessary stationarity assumptions across presentations. In Sect. 2.2 the unconditional probability of a process firing at any point in the PST histogram will be derived. It will once again be seen that the Bernoulli approximation holds if the binwidth is sufficiently small. It will also be seen that the Bernoulli approximation produces the exact asymptotic probability of firing with respect to time regardless of the binsize.

The analysis leads to a useful property of the PST histogram. It is shown that if each bin in the histogram represents a random variable in a discrete random process, then the structure of the histogram is that of an auto-regressive (AR) process. The exact form of the parameters using the Bernoulli approximation is derived in Sect. 3. Typically, the stimulus can only be estimated from the PST histogram when the exact form of the recovery function is known, either a priori or by estimates from the interval histogram. Using the AR representation of the PST histogram, AR techniques can be applied to estimate the coefficients of the AR process and from these both the stimulus and recovery functions can be estimated. The AR representation can also be used to determine the expected value of the PST histogram given the stimulus and the recovery function. Examples of these applications will be shown in Sect. 3.

2 Bernoulli approximation

2.1 Statistics

Consider a nerve which, when stimulated by $s(t)$, has an intensity function $\lambda(t)$. This means that

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\Pr\{\text{nerve fires in } [t, t + \Delta t)\}}{\Delta t}. \tag{1}$$

Partition the timeline into bins of equal width Δ. As stated in the introduction, a typical simplifying assumption for analysis is that Δ is small enough so that the intensity $\lambda(t)$ is approximately constant within the bin and, more importantly, no more than one point lies in each bin. The process is then considered to consist of discrete random variables, one assigned to each bin, each with a Bernoulli distribution, i.e., with a probability of firing once of $\lambda(t)\Delta$ and a probability of not firing of $1 - \lambda(t)\Delta$. Estimators can be easily implemented us-

ing this Bernoulli representation since the probability of a specific spike-train realization is the product of these simple terms. For multiple presentations of the simulus under the condition of stationarity across presentations, the distribution of firings within each bin is binomial with parameter $\lambda(t)\Delta$ and again the form of the estimator has been simplified. If these simplified distributions are not accurate, however, then the estimators based on the simplified distributions may also be inaccurate. It is therefore important to consider how well the Bernoulli approximation models the actual statistics of the point process.

Assume that $\lambda(t)$ is in fact constant over a binwidth, and describe it as λ for the binwidth in question. Under the true Poisson distribution for firing time, which is described by

$$\lambda \exp(-\lambda t), \tag{2}$$

the probability of firing at least once within the bin is

$$\int_0^\Delta \lambda \exp(-\lambda t) \, dt = 1 - \exp(-\lambda\Delta). \tag{3}$$

The relation between the actual statistics and those derived from the Bernoulli approximation rely on the approximation

$$1 - \exp(-\lambda\Delta) = 1 - [1 - \lambda\Delta + (\lambda\Delta)^2/2! - (\lambda\Delta)^3/3! + \cdots]$$

$$\approx \lambda\Delta. \tag{4}$$

As $\lambda\Delta$ becomes large, however, the approximation becomes worse and will result in an increasingly poorer estimate of firing probability.

Considering the probability of exactly one firing within a bin, the difference between the probability derived from the Bernoulli assumption and that derived from the actual Poisson distribution is

$$\Pr\{\text{one firing} \mid \text{Bernoulli}\} - \Pr\{\text{one firing} \mid \text{Poisson}\}$$

$$= \lambda\Delta - \lambda\Delta \exp(-\lambda\Delta)$$

$$= \lambda\Delta(1 - \exp(-\lambda\Delta)). \tag{5}$$

The difference can be seen in Fig. 1, where the probabilities for both distributions are shown as a function of $\lambda\Delta$. The solid line shows the probability derived from the Bernoulli approximation while the dashed line shows that from the Poisson distribution. The probabilities are seen to significantly differ for $\lambda\Delta$ greater than 0.1.

The expected number of firings within a bin using the Bernoulli approximation is $\lambda\Delta$ which is also the expected number using the Poisson distribution, regardless of the size of $\lambda\Delta$. Also, given the multiple presentations of the stimulus, the maximum likelihood estimator for λ given the number of firings within the bin is the same using both distributions. Thus, although the accuracy of the probability of firing varies with the magnitude of $\lambda\Delta$, the Bernoulli approximation produces an accurate mean regardless of the size of $\lambda\Delta$. This might explain why the potential problem of bin size is often not encountered when estimating λ.
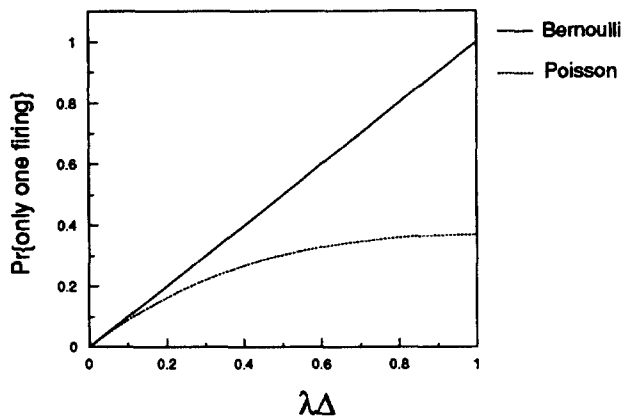
Fig. 1. Probability of exactly 1 firing within a bin as a function of $\lambda\Delta$. The solid line represents the probability derived using the Bernoulli approximation, and the dashed line represents the Poisson derived probability

In general, the actual moments of the firings within a bin differ from those based on the Bernoulli assumption. The general forms of the nth moment about the mean are

(nth moment | Binomial)

$$= \frac{d^n}{(dt)^n}[\exp(-\lambda\Delta t)(\lambda\Delta \exp(t) + (1 - \lambda\Delta))]^n|_{t=0},$$

(nth moment | Poisson)

$$= \frac{d^n}{(dt)^n}[\exp(\lambda\Delta(\exp(t) - 1 - t))]|_{t=0}. \quad (6)$$

For example, this gives a difference in variance of

$$\text{var}\{\text{firing} \mid \text{Poisson}\} - \text{var}\{\text{firing} \mid \text{Bernoulli}\}$$

$$= \lambda\Delta - \lambda\Delta(1 - \lambda\Delta)$$

$$= (\lambda\Delta)^2. \quad (7)$$

This difference is shown in Fig. 2 where the variances for the two distributions are plotted. Again, the as-
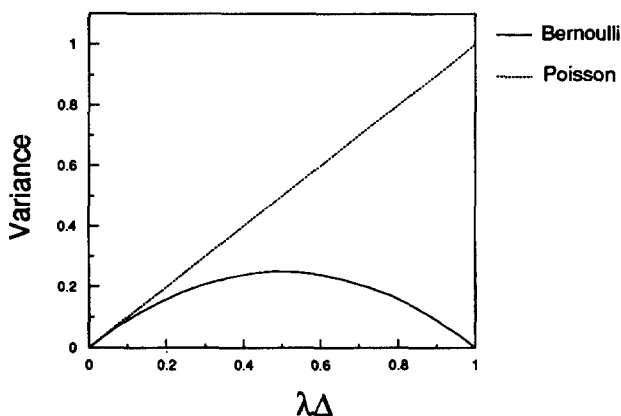

Fig. 2. The variance of the number of firings within a bin as a function of $\lambda\Delta$. The lines are formatted as in Fig. 1

sumption's results differ from the Poisson's results for large $\lambda\Delta$ and may greatly affect whatever distribution-based estimator is used.

To illustrate the problems that may arise when the Bernoulli approximation is used incorrectly, consider a statistical test for constant intensity based on a PST histogram of the process. If $\lambda\Delta = 0.3$, say, then Fig. 2 shows that the variance under the Bernoulli approximation is significantly smaller than the actual variance of the Poisson process. Therefore, a test of significance based on estimated variance under the Bernoulli assumption will be biased towards incorrectly rejecting the hypothesis of constant threshold.

2.2 Structure of PST histogram

The structure of the PST histogram under both Poisson and Bernoulli distributions will now be considered. To simplify the derivation, let the intensity function be a constant λ and let the absolute refractory period exist with length $n\Delta$, where n is an integer. Without the refractory effect, independence of firing realizations from different stimulus presentations could be used to extend the results of the binned process in the previous section to the binned PST histogram. The absolute refractory period, however, inhibits the nerve from firing for a time of $n\Delta$ seconds immediately after the nerve fires. The intensity during a bin could be either 0 or $\lambda\Delta$, depending upon the history of the nerve during that presentation, thereby violating the stationarity assumption. Thus, the bins previous to the bin being analyzed must be considered in order to determine the statistics of that particular bin.

Firstly, the PST histogram under Bernoulli approximation will be examined. Defining the kth bin in the PST histogram by B_k, the PST histogram is described as a discrete process where B_k, at each stimulus presentation, has an unconditional probability of firing of p_k. Note that this assumes that, for a specific bin B_k, the histogram process is stationary since the probability of firing is always p_k, as opposed to the actual nerve process which has a probability of either $\lambda\Delta$ or 0, depending upon the process's history during that presentation.

Under the Bernoulli approximation, the probability of firing in bin B_k is

$$p_k = \text{Pr}\{\text{fires in } B_k \mid \text{doesn't fire in } [B_{k-n}, B_{k-1}]\}$$

$$\times \text{Pr}\{\text{doesn't fire in } [B_{k-n}, B_{k-1}]\}$$

$$= \lambda\Delta \left(1 - \sum_{i=1}^{n} p_{k-i}\right)$$

$$= \lambda\Delta - \lambda\Delta \sum_{i=1}^{n} p_{k-i}. \quad (8)$$

Given the form of the probability of firing in B_k during one presentation, the expected value of $B(k)$ for N presentations can be derived, where $B(k)$ is the value of the PST histogram in bin B_k. Estimating p_{k-i} with

$B(k - i)/N$ – the MLE of p_{k-i} – one obtains

$$p_k = \lambda\Delta - \lambda\Delta \sum_{i=1}^{n} B(k - i)/N,$$

$$p_k N = \lambda\Delta N - \lambda\Delta \sum_{i=1}^{n} B(k - i),$$

$$E\{B(k)\} = \lambda\Delta N - \lambda\Delta \sum_{i=1}^{n} B(k - i). \tag{9}$$

If the steady-state segment of the PST histogram where the transients have disappeared is considered, i.e., for large k (say for $k > K$, where K is large), then

$$p_k = p, \qquad \forall k > K. \tag{10}$$

Substituting (10) into (8) and solving for p, the asymptotic probability is obtained:

$$p = \lambda\Delta - \lambda\Delta np,$$

$$p = \lambda\Delta/(1 + \lambda\Delta n). \tag{11}$$

So, for N presentations, the expected number in each bin is

$$E\{B(k)\} = \lambda\Delta N/(1 + \lambda\Delta n). \tag{12}$$

This result concurs with results in which (12) has been derived using other methods (Johnson and Swami 1983). Note that the denominator is independent of Δ since the product Δn is equal to the length of the refractory effect. Thus, $E\{B(k)\}$ is proportional to Δ.

The structure of the PST histogram is now derived using the actual Poisson distribution. The analysis considers three separate cases for each bin. Each case is distinguished by whether the refractory period is still in effect at the beginning of the bin and, if so, when the previous firing occurred. These cases are analyzed in Appendix A and result in the following description:

$$p_k = (1 - e^{-\lambda\Delta}) - (1 - e^{-\lambda\Delta}) \sum_{i=1}^{n-1} p_{k-i}$$

$$+ [e^{-\lambda\Delta} - (1 - e^{-\lambda\Delta})/(\lambda\Delta)]p_{k-n}. \tag{13}$$

This result holds for all values of $\lambda\Delta$. As stated in the previous section, the Bernoulli approximation relies on the approximation

$$1 - e^{-\lambda\Delta} \approx \lambda\Delta, \tag{14}$$

and if this approximation is made in (13), the recursion reduces to (8), the PST description derived using the Bernoulli approximation. Equation (8) is then an accurate description of the histogram's structure if $\lambda\Delta$ is small enough so that the approximation in (14) holds.

Assuming steady-state once again,

$$p = (1 - np)(1 - e^{-\lambda\Delta}) + p[1 - (1 - e^{-\lambda\Delta})/(\lambda\Delta)] \tag{15}$$

reduces to

$$p = \lambda\Delta/(1 + \lambda\Delta n), \tag{16}$$

such that the expected number of firings is

$$E\{B(k)\} = \lambda\Delta N/(1 + \lambda\Delta n) \tag{17}$$

for N presentations, which is the same as that obtained from the Bernoulli approximation in (12). Thus, the Bernoulli approximation exactly predicts the asymptotic expected number of firings within a bin, regardless of the size of $\lambda\Delta$.

It should be noted that because the absolute refractory effect is longer than one binwidth, there can never be more than one firing in a bin during a single presentation. Thus, the PST histogram is a true binomial system with binomial statistics since it cannot fire more than once per bin presentation. This is one of the few examples when a self-exciting point process results in a simpler formulation of an estimation problem.

Note that the exact analysis produces the same steady-state results as that derived using the Bernoulli approximation and they are the same even if $\lambda\Delta$ is large. The Bernoulli approximation does not give the correct recursion, however, and thus the higher-order moments of the bins as determined under the Bernoulli approximation are incorrect for large $\lambda\Delta$, as are the cross-moments between the different bins. This will be important in the next section where the correlation between bins is used to estimate refractory effects.

Johnson and Swami (1983) used a multiplicative relationship between the stimulus and the nerve's refractory effects to model the instantaneous intensity of the firing process. If the stimulus is $s(t)$, then the instantaneous intensity $\mu(t;w_t)$ is defined by

$$\mu(t;w_t) = s(t)r(t - w_t), \tag{18}$$

where $r(t)$ is termed the recovery function and w_t is the time of the most recent firing. $r(t)$ is the recovery function which varies monotonically over time from 0 to 1, thereby reducing its inhibitory effect on the firing process as its argument (the time since the most recent firing) increases. In the previous discussion, $s(t)$ was a constant λ and $r(t)$ was a step function that was 0 over $[0, n\Delta)$. The case of time-varying $s(t)$ and arbitrary $r(t)$ will now be considered.

Allowing $r(t)$ to have values other than 0 and 1 automatically introduces ambiguity into the solution to p_k. Consider the case discussed previously where

$$r(t) = 0, \quad t < n\Delta,$$

$$= 1, \quad \text{otherwise}. \tag{19}$$

If the n bins before B_k contained M firings and the stimulus was presented N times, then the probability of firing in B_k was 0 for M presentations and the probability of firing was $\lambda\Delta$ for $N - M$ presentations. For example, if the nerve fires in bin B_{n-3}, then the recovery function will be 0 during B_k. If, however,

$$r(t) = t/(n\Delta), \quad t < n\Delta,$$

$$= 1, \quad \text{otherwise}, \tag{20}$$

then all absolute knowledge about the time since the most recent firing is lost. If the nerve fired in bin B_{k-3}, the recovery function might be $3/(n\Delta)$ at B_k but the nerve could also fire in bin B_{k-2} or B_{n-1} such that B_k would have a recovery function of $2/(n\Delta)$ or $1/(n\Delta)$, respectively. Therefore, the *probability* that a firing in a

bin was the most recent firing must now be considered to resolve this ambiguity.

The derivation of the solution to p_k is given in Appendix B. The assumptions made are that $s(t)$ and $r(t)$ are constant over a binwidth, and that $s(t)\Delta$ is small enough such that the Bernoulli approximation holds. The result is

$$p_k = s(k\Delta)\Delta - s(k\Delta)\Delta \sum_{i=1}^{n} p_{k-i}q(i\Delta)$$

$$\times \exp\left(-\int_0^{(i-1)\Delta} s([k-1]\Delta - y)r(y)\,dy\right), \qquad (21)$$

where $q(t)$ is a term defined by Qi Bi (1989) as the *shape function* and is equal to $1 - r(t)$. When $q(t)$ is positive, inhibitory effects are present and when $q(t)$ is 0 the inhibitory effect is absent. If the p_k is assumed to be $\lambda(t)/\Delta$, then (21) is identical to that derived by Bi using much more complicated methods.

The exponential in the above recursion is due to the uncertainty about which firings in the n bins prior to B_k are the most recent ones. Bi makes the following argument for ignoring the exponential term. When $r(t)$ is small, the exponential term is close to one, as is $q(t)$. As $r(t)$ approaches 1, the exponential becomes smaller, approaching 0 and $q(t)$ also approaches 0. So, the terms in the summation in which the exponential is significantly different from 1 is a multiplier for the small $q(t)$ which has very little effect on the total summation. Thus, Bi argues that it is reasonable to ignore the exponential and simplify the recursion to

$$p_k = s(k\Delta)\Delta - s(k\Delta)\Delta \sum_{i=1}^{n} p_{k-i}q(i\Delta). \qquad (22)$$

Using (22), the expected value of the firings in bin B_k can be derived as was done in (9):

$$E\{B(k)\} = s(k\Delta)N - s(k\Delta)\Delta \sum_{i=1}^{n} B(k-i)q(i\Delta). \qquad (23)$$

The simplification of (22) is, in fact, exactly equal to (21) when the recovery function contains only an absolute refractory effect – i.e., is a step delay – because the exponential is unity in each term. The effect of ignoring the exponential for a recovery function that is not just a step delay is shown in Fig. 3. The recovery function is a linear ramp with a 10 ms rise-time and is shown in Fig. 3a. The derived estimations for $s(k\Delta)\Delta q(i\Delta)$, $i = 1, \ldots, 10$ using the approximation is compared to the actual parameters in Fig. 3b, and the resulting derived probabilities using the respective summation factors is shown in Fig. 3c.

From this example, it is clear that the effect of ignoring the exponential in (21) is minimal. The summation parameters shown in Fig. 3b are approximately the same and the resulting PST histograms differ by 2.1%. Thus, from now on we will adopt Bi's suggestion and use (22) as the approximate recursion for p_k.
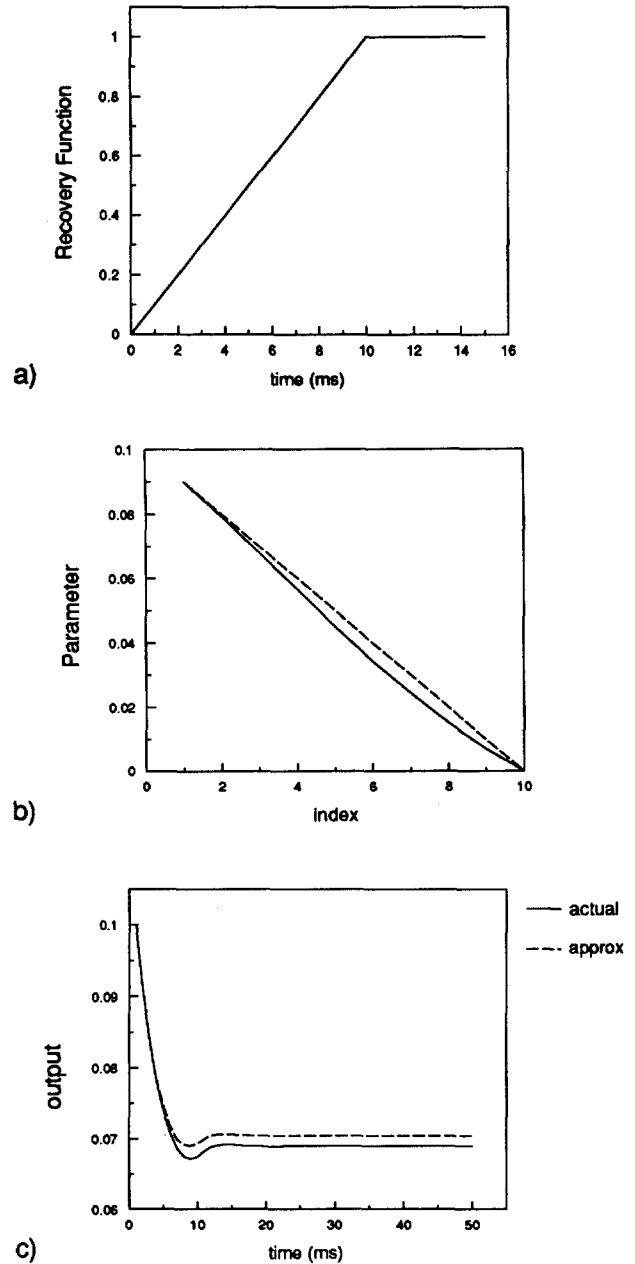

a)

b)

c)

Fig. 3a–c. Comparing the effect of ignoring the exponential in (21) a The recovery function $r(t)$ used; b The resulting summation parameters using (21) for the actual and (22) for the approximation; c the derived histograms using (21) and (22). Note that equation (21) is the actual representation and (22) is the approximation

3 PST histogram as an AR system

In the previous sections, the discussion has been strictly with regard to the firing probabilities. The conditions under which the Bernoulli approximation is valid have been examined and an expression for the probability of firing within each bin in a PST histogram has been derived. This information will now be used to gain knowledge about the driving stimulus $s(t)$ and any refractory effects that may exist.

As shown by (22), the PST histogram can be modeled as an auto-regressive (AR) process. This AR system is time-varying if $s(t)$ is time varying and stationary if $s(t)$ is a constant. Because the histogram can be modeled as an AR system, AR techniques can be used to estimate the AR parameters and, from these, $s(t)$ and $r(t)$.

From (23), the histogram's process can be modeled as

$$B(k) = s(k\Delta)\Delta N$$

$$- s(k\Delta)\Delta \sum_{i=1}^{n} B(k - i)q(i\Delta) + u(k), \qquad (24)$$

where $u(k)$ is noise. Thus, $B(k)$ is an AR process with parameters determined by $s(t)$, $r(t)$ and Δ.

The noise $u(k)$ has the following properties:

i) Define

$$\text{fl}M(k): = s(k\Delta)\Delta N - s(k\Delta)\Delta \sum_{i=1}^{n} B(k - i)q(i\Delta). \qquad (25)$$

Then, the probability mass function for $u(k)$ is

$$P_{u(k)}(U) = \binom{N}{U + M(k)} p_k^{U + M(k)} (1 - p_k)^{N - U - M(k)}. \qquad (26)$$

ii) $E\{u(k)\} = 0$.
iii) $\text{var}\{u(k)\} = p_k(1 - p_k)N$. $\qquad (27)$

iv) $E\{u(k)u(k + 1)\} = 0, \forall 1 \neq 0$, i.e., its autocorrelation function is zero except at the origin, so $u(k)$ is white noise and the noise components are uncorrelated.

Of the four properties, (iv) is the most important because most AR estimation techniques require that the driving signal be uncorrelated.

Some of the uses of the AR model will now be demonstrated. Suppose that $s(k) = \lambda$ and

$$r(k) = 0, \quad k \leq 5,$$

$$= (k - 5)/5, \quad 6 \leq k \leq 10,$$

$$= 1, \quad \text{else}. \qquad (28)$$

as shown in Fig. 4.

The expected value of the PST histogram can be considered as the output of a linear time-invariant filter with transfer function

$$H(z) = 1/\left(1 + \sum_{i=1}^{5} \lambda\Delta z^{-i} + \sum_{i=6}^{10} \lambda\Delta \left(\frac{10 - i}{5}\right) z^{-i}\right), \qquad (29)$$

where $\lambda\Delta$ is the input. Figure 5 shows the output of this filter for $\lambda = 10$ and $\Delta = 5$ ms, along with a PST histogram from a point process modeled from the parameters using the method described by Snyder (1975). The expected value of the PST histogram could also be produced applying the recursion of (23):

$$E\{B(k)\} = .05N - .05 \sum_{i=1}^{5} B(k - i)$$

$$- .05 \sum_{i=6}^{10} ((10 - i)/5)B(k - i). \qquad (30)$$

The PST histogram can be used to estimate $r(t)$ when $s(t)$ is known to be a constant. Assume that a nerve is known to be stimulated by a constant signal of
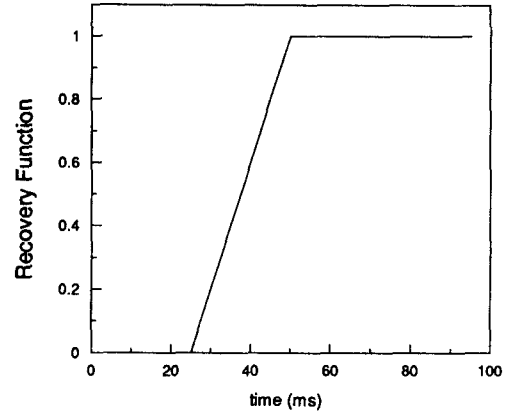

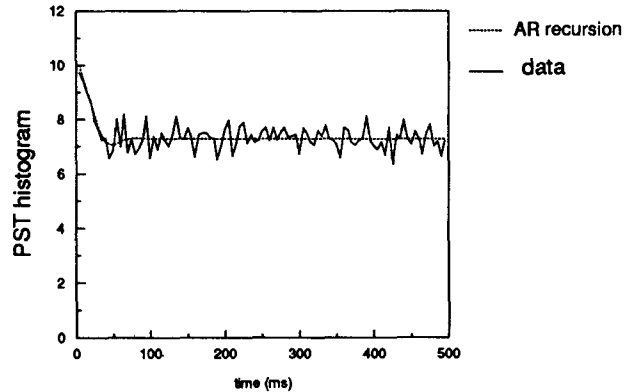Fig. 4. The recovery function used for the point process simulation and the AR recursion

Fig. 5. Demonstrating the performance of the AR model for a PST histogram. The solid line represents data obtained from a simulated point process with a $\lambda = 10$ and a recovery function shown in Fig. 4; the dashed line shows the expected value of the histogram using the filter described in (29) and an input of $\lambda\Delta = .05$

unknown intensity λ. Since AR estimation techniques require that the data used for estimation be zero mean and stationary, the PST histogram data are normalized by subtracting the ensemble mean (Maragos et al. 1984) and the steady-state portion of the PST histogram is considered in which the transients have decayed.

Two sample histograms without the transients are shown in Fig. 6 with $\lambda = 100$, $\Delta = .001$, and $N = 50000$. Since the expected number of firings in each bin is

$$E\{B(k)\} = \frac{\lambda\Delta N}{1 + \sum_{i=1}^{n} \lambda\Delta q(i\Delta)},$$

the recovery function is needed to estimate λ from the PST histogram. The recovery functions for each histogram are

$$r_a(t) = 0, \quad t \leq .004,$$

$$= 1, \quad \text{otherwise}, \qquad (31a)$$

$$r_b(t) = t/.004, \quad t \leq .004,$$

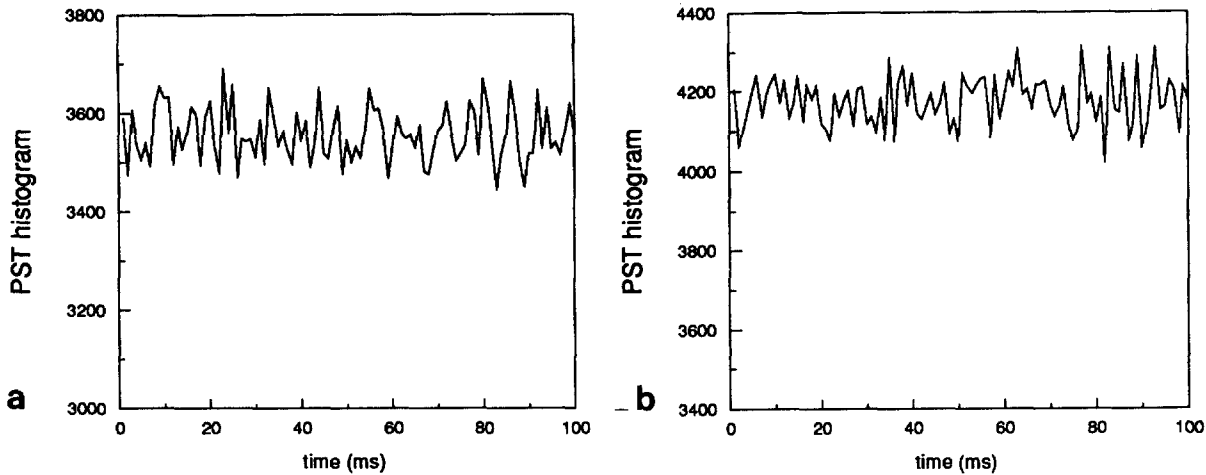$$= 1, \quad \text{otherwise}. \qquad (31b)$$

Fig. 6a, b. A portion of the PST histograms used to estimate the recovery functions and the stimulus. They were generated using $s(t) = 100$, $\Delta = 1$ ms, $N = 50000$, and recovery functions of a an absolute refractory period of 4 ms; b a linear rise from 0 to 1 over 4 ms. The recovery functions are shown in Fig. 7

The product $\lambda\Delta = 0.1$ so (22) can be used as the AR model of the histograms, and the Levinson–Durbin algorithm (Kay 1988) can be used to estimate the AR parameters $\lambda\Delta q(i\Delta)$.

From (22), it can be easily shown that the steady-state response of this system is

$$p = \lambda\Delta \Big/ \left(1 + \lambda\Delta \sum_{i=1}^{n} q(i)\right).\qquad(32)$$

Let the estimates $\hat{a}(i)$ of the AR parameters $q(i)\lambda\Delta$ be estimated from the Levinson–Durbin algorithm. The steady-state response, p, can be estimated by averaging $B(k)/N$ across the histogram, i.e.,

$$\hat{p} = \sum_{k=1}^{M} B(k)/(NM),\qquad(33)$$

where N is the number of presentations and M is the number of bins in the histograms. From (32), the estimate for λ is

$$\hat{\lambda} = \hat{p}\left(1 + \sum_{i=1}^{n} \hat{a}(i)\right)/\Delta.\qquad(34)$$

Using this estimate for λ, the recovery function can be solved to yield:

$$\hat{r}(i) = 1 - \hat{q}(i)$$
$$= 1 - \hat{a}(i)/(\hat{\lambda}\Delta).\qquad(35)$$

The length of the histogram used in case (a) was 300 points and the length of the histogram used in case (b) was 700 points. A larger number of data points were needed for case (b) in order to achieve the same estimator efficiency for the parameters as in case (a). The resulting estimates from these histograms are

$$\hat{p}_a = .07115\qquad(36a)\qquad\hat{p}_b = .08344\qquad(36b)$$

$$\underline{\hat{a}}_a(i) = \begin{bmatrix} .0984 \\ .0861 \\ .1325 \\ .0933 \end{bmatrix}\qquad(37a)\qquad\underline{\hat{a}}_b(i) = \begin{bmatrix} .0811 \\ .0686 \\ .0465 \\ .0296 \end{bmatrix}\qquad(37b)$$

$$\hat{\lambda}_a = 100.3\qquad(38a)\qquad\hat{\lambda}_b = 102.3\qquad(38b)$$

Equations (38a) and (38b) demonstrate the accuracy of estimating the constant stimulus without a priori knowledge of the recovery function or its estimate from an interval histogram. The recovery functions estimated using (35) as well as the actual recovery functions are shown in Fig. 7.
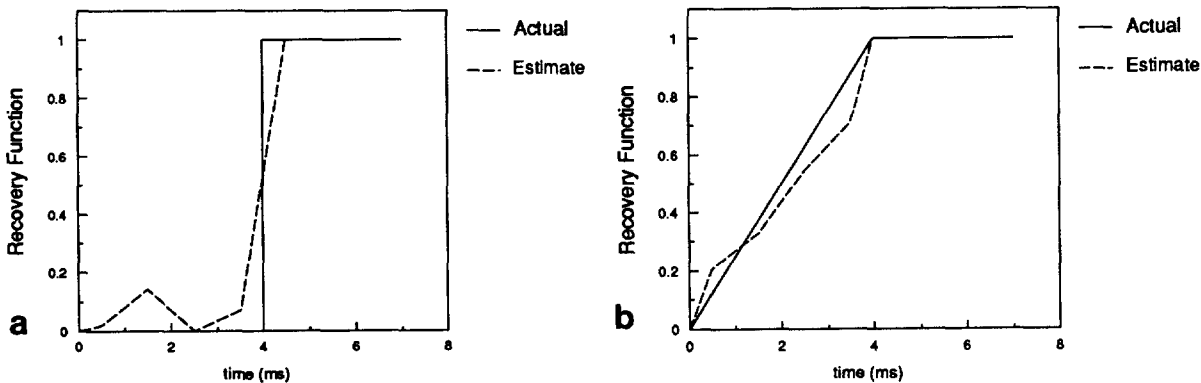

Fig. 7a, b. The actual recovery functions and their estimations using (35) and the PST histograms partially shown in Fig. 6

4 Conclusion

Point processes are often discretized in time for the simplicity of analysis. With the time-line divided into bins, a single probability of firing is assigned to the bin which produces a binomial instead of a Poisson process. The validity of this simplification using the Bernoulli approximation was examined by comparing some basic statistics using the actual Poisson distribution and the approximate Bernoulli distribution. The accuracy of the assumption was seen to be dependent upon the size of the product $\lambda\Delta$, where λ is the intensity within a time bin and Δ is the size of the binwidth.

Under the Bernoulli approximation, it is assumed that the process can fire only once within a bin. The probability of this occurring is close to the actual probability only when $\lambda\Delta$ is small, say less than 0.1. For $\lambda\Delta$ larger than this, the assumption gives a much larger probability of firing than exists in actuality. However, for any $\lambda\Delta$, the *expected* number of firings within a bin is exactly the same for both the Poisson and the Bernoulli process. Although the probability of firing for the Bernoulli process is greater, the actual Poisson process has non-zero probabilities for the process firing any number of times within the bin. This increases the mean number of firings to be equal to that of the Bernoulli process. Thus, if one is only concerned with first order moments, the Bernoulli approximation is exact. Nerve processes and PST histograms can then be modeled using the discrete approximation with no regard to bin size and parameters from the first order statistics can be estimated using the binomial representation. The Bernoulli approximation fails for higher order statistics when $\lambda\Delta$ is large, as seen for the variance plotted in Fig. 2.

The PST histogram was shown to have an AR structure given in (22). This AR representation provides a general method for modeling a PST histogram given the stimulus and recovery function, which leads to a well-defined estimator of these functions from the observed histogram data. When estimating the stimulus from the PST histogram, it is frequently assumed that the time since the most recent firing is known for each occurrence in a bin (Johnson and Swami 1983; Miller 1985). Such estimators require this additional side information concerning firing intervals which is not present in the PST histogram. Thus the exponential in (21) is introduced to estimate the probability of the time since the most recent occurrence. Estimators based on this latter representation can be considered as estimators using incomplete information.

The AR representation also provides the unique ability to estimate the recovery function from the PST histogram alone. The recovery function is typically estimated from the interval histogram, but the relation of the recovery function to the AR parameters allows an alternative estimate of the function using standard techniques such as the Levinson–Durbin algorithm. This second moment information has been typically ignored in the use of PST histograms and we have shown that it can be used to estimate both the recovery

function and the driving stimulus. This information could also be incorporated with the interval histogram to increase the accuracy of signal estimators.

Appendix A

Here we derive the unconditional probability of the nerve firing within bin B_k given that

$$s(t) = \lambda. \tag{A.1a}$$

$$r(t) = 0, \quad t \leqslant n\Delta,$$

$$= 1, \quad \text{else}. \tag{A.1b}$$

Three cases must be considered, corresponding to the three possible states of the recovery function:

I No refractory effect at the beginning of the bin. The probability of this case occurring is

$$1 - \sum_{i=1}^{n} \Pr\{B_{k-i}\}. \tag{A.2}$$

Given this case,

$$\Pr\{B_k \mid \text{case I}\} = 1 - e^{-\lambda\Delta}. \tag{A.3}$$

II A firing occurred in one of $\{B_{k-n+1}, \ldots, B_{k-1}\}$. The probability of this case occurring is

$$\sum_{i=1}^{n-1} \Pr\{B_{k-i}\}. \tag{A.4}$$

Then,

$$\Pr\{B_k \mid \text{case II}\} = 0 \tag{A.5}$$

since the refractory period will not have finished during bin B_k.

III A firing occurred in bin B_{k-n}. The probability of this case occurring is

$$\Pr\{B_{k-n}\}. \tag{A.6}$$

Then, the refractory period will end sometime during bin B_k. Assume that the firing distribution is uniform over the binwidth B_{k-n} so that the refractory effect ends with a uniform distribution over B_k, i.e.,

$$\Pr\{\text{refractory period ends at time } t\}$$

$$= 1/\Delta, \quad t \in [(k-1)\Delta, k\Delta). \tag{A.7}$$

This produces the result

$$\Pr\{B_k \mid \text{case III}\} = \int_0^\Delta 1/\Delta \int_0^{\Delta-s} \lambda e^{-\lambda w} \, dw \, ds$$

$$= 1 - (1 - e^{-\lambda\Delta})/(\lambda\Delta). \tag{A.8}$$

When the probability of each case is multiplied by its corresponding conditional probability and then summed, the result is the probability of a firing in bin B_k:

$$\Pr\{B_k\} = \left(1 - \sum_{i=1}^{n} \Pr\{B_{k-i}\}\right) \cdot (1 - e^{-\lambda\Delta})$$
$$+ \Pr\{B_{k-n}\} \cdot [1 - (1 - e^{-\lambda\Delta})/(\lambda\Delta)]$$
$$= (1 - e^{-\lambda\Delta}) - (1 - e^{-\lambda\Delta}) \sum_{i=1}^{n-1} \Pr\{B_{k-i}\}$$
$$+ [e^{-\lambda\Delta} - (1 - e^{-\lambda\Delta})/(\lambda\Delta)] \Pr\{B_{k-n}\} . \quad \text{(A.9)}$$

Appendix B

Here we derive the unconditional probability of the nerve firing in bin $B(k)$ for arbitrary $s(t)$ and $r(t)$. Two cases must be considered, corresponding to whether the nerve has recovered at the beginning of bin B_k:

I $q(t - w_{N_t})$ is zero at the beginning of the bin. The probability of this case occurring is

$$1 - \Pr\{\text{previous firing was within the last } n \text{ bins}\}$$

$$= 1 - \sum_{i=1}^{n} \Pr\{\text{doesn't fire before } B_k \mid \text{fires in } B_{k-i}\}$$
$$\times \Pr\{B_{k-i}\} . \quad \text{(B.1)}$$

Since the exact occurrence time of a firing in bin B_{k-i} is unknown, a uniform distribution is assumed over the bin for the firing time. Then, the first probability in the equation above is

$$\Pr\{\text{doesn't fire before } B_k \mid \text{fires in } B_{k-i}\}$$

$$= \int_{(i-1)\Delta}^{\Delta} \frac{1}{\Delta} \exp\left(-\int_0^w s([k-1]\Delta - y)r(y)\, dy\right) dw$$

$$\approx \exp\left(-\int_0^{(i-1)\Delta} s([k-1]\Delta - y)r(y)\, dy\right) \quad \text{(B.2)}$$

where to get to the second line we have assumed that $r(t)$ is constant over any binwidth and $s(t)\Delta$ is small. Assuming also that $s(t)$ is constant over any binwidth, we obtain

$$\Pr\{B_k \mid \text{case I}\} = s(k\Delta)\Delta \left[1 - \sum_{i=1}^{n} \Pr\{B_{k-i}\}\right]$$

$$6 \quad \times \exp\left(-\int_0^{(i-1)\Delta} s([k-1]\Delta - y)r(y)\, dy\right). \quad \text{(B.3)}$$

II $q(t)$ is non-zero at the beginning of bin B_k. The probability of this occurring is

$$\sum_{i=1}^{n} \Pr\{\text{doesn't fire before } B_k \mid \text{fires in } B_{k-i}\} \times \Pr\{B_{k-i}\} .$$
$$\text{(B.4)}$$

Again, assuming $r(t)$ is constant over a binwidth and $s(t)\Delta$ is small,

$$\Pr\{B_k \mid \text{case II}\} = s(k\Delta)\Delta \sum_{i=1}^{n} \Pr\{B_{k-i}\}$$

$$\times r(i\Delta) \exp\left(-\int_0^{(i-1)\Delta} s([k-1]\Delta - y)r(y)\, dy\right). \quad \text{(B.5)}$$

When the probabilities of cases I and II are added, the total bin probability is obtained:

$$\Pr\{B_k\} = s(k\Delta)\Delta - s(k\Delta)\Delta \sum_{i=1}^{n} \Pr\{B_{k-i}\}$$

$$\times \exp\left(-\int_0^{(i-1)\Delta} s([k-1]\Delta - y)r(y)\, dy\right)[1 - r(i\Delta)]$$

$$= s(k\Delta)\Delta - s(k\Delta)\Delta \sum_{i=1}^{n} \Pr\{B_{k-i}\}q(i\Delta)$$

$$\times \exp\left(-\int_0^{(i-1)\Delta} s([k-1]\Delta - y)r(y)\, dy\right). \quad \text{(B.6)}$$

References

Bi Q (1989) A closed-form solution for removing the dead time effects from the poststimulus time histograms. J Acoust Soc Am 85:2504–2513

Brillinger DR (1988) Maximum likelihood analysis of spike trains of interacting nerve cells. Biol Cybern 59:189–200

Chornoboy ES, Schramm LP, Karr AF (1988) Maximum likelihood identification of neural point process systems. Biol Cybern 59:265–275

Geisler CD, Goldberg JM (1966) A stochastic model of the repetitive activity of neurons. Biophys J 6:53–69

Gerstein GL, Kiang NYS (1960) An approach to quantitative analysis of electrophysiological data from single neurons. Biophys J 1:15–28

Gernstein GL, Maldelbrot B (1964) Random walk models for the spike activity of a single neuron. Biophys J 4:41–68

Johnson DH (1978) The relationship of post-stimulus time and interval histograms to the timing characteristics of spike trains. Biophys J 22:413–430

Johnson DH, Swami A (1983) The transmission of signals by auditory-nerve fiber discharge patterns. J Acoust Soc Am 74:493–501

Kay SM (1988) Modern spectral estimation. Prentice Hall, Englewood Cliffs, NJ

Maragos PA, Schafer RW, Mersereau RM (1984) Two-dimensional linear prediction and its application to adaptive predictive coding of images. IEEE Trans Acoust Speech Signal Process 32:1213–1229

Miller MI (1985) Algorithms for removing recovery-related distortion from auditory-nerve discharge patterns. J Acoust Soc Am 77:1452–1464

Palm G, Aertsen AMHJ, Gerstein GL (1988) On the significance of correlations among neuronal spike trains. Biol Cybern 59:1–11

Perkel DH, Gerstein GL, Moore GP (1967) Neuronal spike trains and stochastic point processes: I, the single spike train. Biophys J 7:391–418

Snyder DL (1975) Random point processes. Wiley, New York

Van Stokkum IHM, Johannesma PIM, Eggermont JJ (1986) Representation of time-dependent correlation and recurrence time functions. Biol Cybern 55:17–24

Brent W. Edwards
Department of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, MI 48109-1102
USA