

## Parametric Modeling of Reaction Time Experiment Data

W. John Braun,<sup>1,4,\*</sup> Valentin Rousson,<sup>2,4</sup> William A. Simpson,<sup>3</sup> and Jennifer Prokop<sup>1</sup>

<sup>1</sup>Department of Statistical and Actuarial Sciences, University of Western Ontario,  
London, Ontario, Canada, N6A 5B7

<sup>2</sup>Department of Biostatistics, University of Zürich, Sumatrastrasse 30, 8006 Zürich, Switzerland

<sup>3</sup>SMART Section, DRDC Toronto, 1133 Sheppard Ave. West, P.O. Box 2000, Toronto,  
Ontario, Canada, M3M 3B9

<sup>4</sup>Centre for Mathematics and Its Applications, Australian National University,  
Canberra, ACT 0200, Australia

\**email*: braun@stats.uwo.ca

**SUMMARY.** A simple parametric model is proposed for data from a point-process version of a reaction time experiment. It is used to statistically check for the presence and nature of nonlinear inhibition in the eye-brain-hand system, as well as to study the nature of the reaction time delay distribution. The model tells us that, in principle, the second-order intensity estimate can be used to determine whether the experimental subject is systematically observing the first or the second of two flashes transmitted in short succession. Nonparametric estimates of second-order intensity functions are used in conjunction with this model. In particular, the model allows for the computation of good bandwidths for intensity curve estimation. A parametric bootstrap can also be implemented. Our methods are illustrated with 12 runs of data from a real reaction time experiment. It is found that nonlinear inhibition is present in the eye-brain-hand system. However, there are insufficient data to distinguish between log-normality and normality in the reaction time distribution, due partly to confounding with the particular kind of nonlinear inhibition present in the system.

**KEY WORDS:** Intensity functions; Parametric bootstrap; Point processes.

### 1. Introduction

In visual psychophysics, researchers study the brain mechanisms underlying vision by presenting visual stimuli and obtaining behavioral responses from an observer (Cornsweet, 1970). In the simple reaction time (RT) experiment, the experimenter presents a stimulus and measures the time it takes for the observer to hit a button. Simple RT is of interest to psychophysicists for two main reasons. First, RT is easy to measure and is thus a useful tool in studying various visual mechanisms. Second, the time taken to perform a visual-motor task is inherently interesting—it indicates the complexity of the operations taking place in the brain, for example. For these reasons, simple RT has been extensively studied. References can be found in Simpson et al. (2000).

The traditional simple RT experiment consists of a warning, a uniformly distributed delay, then the briefly flashed stimulus. The observer hits a button as quickly as possible, and the time from the stimulus onset to the response is the RT. This experiment has three problems. First, because of the uniform delay distribution, the observer can anticipate the stimulus. Second, the experiment, with its discrete trials and warning, is artificial. Third, since stimuli are presented in isolation, interaction effects cannot be studied.

Following Brillinger (1975a), an alternative approach consists of presenting the flashes according to a Poisson process and recording the time of each button press. The resulting data can then be analyzed using point-process techniques. For example, an estimate of the cross-intensity function (see Section 3) between the stimuli point process and the response point process yields information about the average impulse response of the eye-brain-hand system. The time at which the peak of the average impulse response occurs can be used as a measure of RT. The irreducible minimum RT is defined as the fastest possible RT that is approached as the stimulus strength (e.g., brightness of the flash) increases.

This modified simple RT experiment resolves the issues listed above: the stimulus cannot be anticipated, the experiment is quite similar to real-world tasks such as driving a car, and interactions among the stimuli can be studied.

This approach was adopted by Simpson et al. (2000), where it was used to more accurately estimate the irreducible minimum RT. That study suggested that the previously reported values of this quantity have been underestimated, because of stimulus predictability that confounds the traditional experiment. That article also reported evidence of nonlinear inhibition and facilitation between pairs of flashes. The

methodology was based on the point-process analogue of the Wiener-Volterra expansion (Brillinger 1975b, 1976, 1992). The construction of such an expansion requires the estimation of several point-process parameters. In particular, product moment densities or intensity functions must be estimated nonparametrically. Formal hypothesis testing was not undertaken by Simpson et al. (2000); thus, some of the conclusions of that work must be viewed as tentative.

Our present purpose is to make inferences, where possible, about the mechanisms in the eye-brain-hand system, using data from a reaction time experiment. The data only provide indirect information about such mechanisms. Thus, it is safest to make minimal statistical assumptions; nonparametric estimation is a natural methodology in such a setting. This was the program followed by Simpson et al. (2000), and our present work is a continuation along this line. However, there are a number of difficulties associated with nonparametric estimates that a reasonable model would assist in resolving. First, nonparametric curve estimates are sensitive to bandwidth choice; we propose the use of a parametric model to help choose a bandwidth according to a direct-plug-in procedure advocated strongly, for example, by Wand and Jones (1995). Second, confidence bands for curve estimates are desirable. The nonparametric point-process bootstrap described in Braun and Kulperger (1998) is not practical because very large samples are required to achieve reasonable accuracy. Parametric bootstrap methods are more accurate, provided the underlying model gives a good approximation. Finally, a parametric model can provide a basis for understanding the behavior of nonparametric estimates. We can look for features in the nonparametric estimates that are predicted by the parametric model, or we may detect features that are not so predicted. This last point is an example of what Vere-Jones (2002) has termed “an interrogative use of a model.”

In this article, we propose a parametric model for reaction time to be used in conjunction with nonparametric intensity function estimates. The model will be used to determine good bandwidths and it will be used as a basis for a parametric bootstrap. It will also play an interrogative role as described above.

The outline of the rest of the article is as follows. We introduce our parametric model in Section 2. A description of point-process intensity functions is given in Section 3, together with their functional forms under our parametric model. Nonparametric estimates are also described there. Section 4 describes our estimation and testing methodology, while Section 5 illustrates the use of our procedures on 12 runs of data from a real RT experiment.

## 2. A Parametric Model for the RT Experiment

In a specific reaction time experimental run, flashes were presented to a subject according to a homogeneous Poisson process with rate  $p_A$  per second. That is, a light was flashed on momentarily to be observed by a subject at each of the endpoints of a sequence of intervals having independent and exponentially distributed lengths (with mean  $1/p_A$ ). The output process consisted of the points in time at which the subject indicated observation of the stimulus by depressing (and releasing) a button.

The flash and response process can be thought to have been observed by the experimenter on an interval  $[0, T]$ .

We denote the number of flashes in the time interval  $X$  by  $A(X)$ , and we denote the set of times of the flashes by the sequence  $= \{A_1, A_2, \dots\}$ . Analogous notation is used for the subject's response times; that is,  $B = \{B_1, B_2, \dots\}$  is the set of response times. Note that  $A$  is taken to be a stationary Poisson process, and  $B$  is assumed to be a simple stationary point process (that is, its statistical properties are time invariant). More information on point processes can be found in the articles by Brillinger (1975a,b, 1976, 1992, 1994) as well as the monographs by Karr (1986), Daley and Vere-Jones (1988), Reiss (1993), and Stoyan, Kendall, and Mecke (1995).

We model the relation between the flash process and response process in the RT experiment using standard point-process operations. These operations are described by Reiss (1993), for example. Our model can be outlined as follows:

1. *Superposition.* Let  $p_S$  denote the rate of an additional independent Poisson process  $S$  which is superposed onto the process of flashes. Denote the resulting superposed process by  $A'$ . That is,  $A' = A \cup S$ .
2. *Thinning.* For  $j = 1, 2, 3, \dots$ , the  $j$ th point of  $A'$  is deleted with probability  $\pi_j$ , independently of all other flashes and perceived flashes. Consider two models for  $\pi_j$ . Either

$$\pi_j = D(A'_j - A'_{j-1}) \quad (\text{T.1})$$

or

$$\pi_j = D(A'_{j+1} - A'_j) \quad (\text{T.2})$$

$D(\cdot)$  is a nonincreasing function. In both cases, denote the thinned process by  $A''$ .

3. *Random Translation.* Let  $B_i = A''_i + V_i$ , for  $i = 1, 2, \dots$ , where  $\{V_i\}$  is a sequence of independent and identically distributed random variables with density  $f_V(v)$ .

A psychophysical interpretation for the superposition operation is as follows. Noise internal to the nervous system is added to the sequence of signals induced by the actual flashes. The observer makes a sequence of decisions as to whether a flash has occurred, based on noisy data. When a datum exceeds some threshold, the observer decides there is a flash. A false alarm occurs when the noise alone exceeds the criterion. It should be noted that the flash itself has fixed luminance with negligible noise.

The thinning operation corresponds to a failure to respond to the flash at  $A'_j$ , possibly because of an interaction (specifically, an inhibition) effect with either the flash at  $A'_{j-1}$  or at  $A'_{j+1}$ . It could also be caused by “temporal summation”; the idea here is that the human visual system is a lowpass temporal filter and so two stimuli in quick succession are summed. The two flashes are combined by the visual system and the observer sees only one bright flash (and so hits the button only once). One of the flashes incorporates the other; both are visible, but they are not seen as distinct.

We study the case where the deletion probability  $D(x)$  is a piecewise constant function:

$$D(x) = \begin{cases} 1, & x < d \\ p, & x \geq d \end{cases}. \quad (1)$$

Here,  $d \geq 0$ , and  $p \in [0, 1]$ . If  $d = 0$ , deletions occur with probability  $p$ , completely independently of the flash process. This is a linear model (see, e.g., Brillinger, 1976); individual flashes stimulate individual responses and there is no interaction effect due to consecutive flashes. If  $d > 0$ , some deletions occur independently of the process, but flashes or apparent flashes within  $d$  time units of their predecessor (alternatively, successor) are deleted with probability 1; there is nonlinear inhibition due to consecutive flashes occurring too close together in time.

The random variables  $V_i$  introduced by the translation operation correspond to the amounts of time required for the brain and hand to respond to the flash; these are the particular RT for each event. At various points in this article, we will consider normal and log-normal distributions. There is some empirical evidence from the classical form of the RT experiment that, in the absence of time pressure, the RT are often log-normally distributed (e.g., Green and Smith, 1982). Ulrich and Miller (1993) have provided some theoretical justification for this by considering the RT random variables as products of independent and identically distributed random variables. We will later see that it is difficult to discern a difference between the normal and log-normal distributions at parameter values that are realistic for the RT experiment.

### 3. Point-Process Intensity Functions and the RT Experiment

The response rate, or first-order intensity of the  $B$  process,  $p_B$ , satisfies

$$E\{B(Y)\} = p_B |Y|_L. \quad (2)$$

Here,  $|Y|_L$  refers to the length of the interval  $Y$ . Roughly speaking, the probability that a response occurs in an arbitrary time interval of length  $h$  is approximately  $p_B h$ . The second-order intensities  $p_{AB}(u)$  and  $p_{BB}(u)$  satisfy

$$E\{A(X)B(Y)\} = \int_X \int_Y p_{AB}(y-x) dx dy, \quad (3)$$

$$E\{B(Y_1)B(Y_2)\} = \int_{Y_1} \int_{Y_2} p_{BB}(y_2 - y_1) dy_1 dy_2 + p_B |Y_1 \cap Y_2|_L. \quad (4)$$

These functions and associated estimates have been studied by Brillinger (1975b), for example. Of course,  $p_{AA}(u)$  is also defined, but is not of interest here.

The quantity  $p_{AB}(u)h^2$  can be interpreted as the approximate probability that a flash occurs in an arbitrary interval of length  $h$ , and that a response occurs in a similar interval but  $u$  time units later. The quantity  $p_{AB}(u) - p_A p_B$  is proportional to the cross-intensity function (Brillinger, 1976). If flashes and responses occur independently of each other, we have  $p_{AB}(u) - p_A p_B = 0$ ; more realistically, there will be dependence between flashes and responses that will likely decrease as  $u$  becomes large, so that  $p_{AB}(u) - p_A p_B \doteq 0$ , for

large  $u$ . The function  $p_{BB}(u)$  is related to the autointensity function  $p_{BB}(u)/p_B - p_B$ . If responses occur independently of each other, which happens, for example, when the  $B$  process depends linearly on the  $A$  process, we have  $p_{BB}(u) - p_B^2 = 0$ . Again, short-range dependence among the responses might result in a nonzero autointensity function for small  $u$ . Brillinger (1975b) gives mixing conditions that generalize this notion of asymptotic independence to the higher-order intensities, or more precisely, to the cumulant intensities; these conditions are tacitly assumed in the rest of the article.

In the Appendix, we show that the response rate is given by

$$p_B = (1-p)(p_A + p_S)e^{-(p_A + p_S)d}. \quad (5)$$

We also show that

$$p_{AB}(u) = p_A p_B \left[ P\{V \notin (u-d, u)\} + \frac{f_V(u)}{p_A + p_S} \right] \quad (6)$$

under thinning assumption T.1, and

$$p_{AB}(u) = p_A p_B \left[ P\{V \notin (u, u+d)\} + \frac{f_V(u)}{p_A + p_S} \right] \quad (7)$$

under T.2. The function  $p_{BB}(u)$  is the same under both T.1 and T.2:

$$p_{BB}(u) = p_B^2 P(|V_2 - V_1 - u| > d). \quad (8)$$

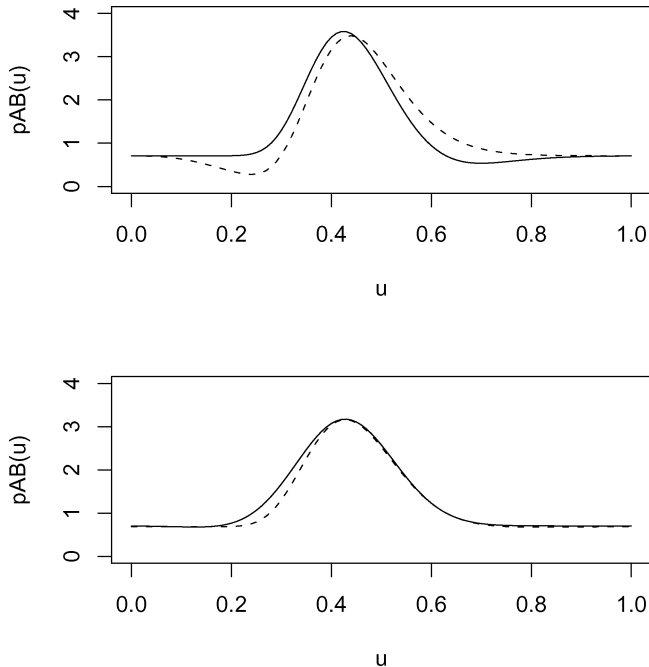
$V$ ,  $V_1$ , and  $V_2$  are independent random variables having common density  $f_V(v)$ .

Graphs of  $p_{AB}(u)$  for the log-normal density are plotted in the top panel of Figure 1 using a realistic parameter setting, and under the two different thinning assumptions. The effect of the  $d$  parameter is to depress the region to the right of the peak when thinning assumption T.1 is operative, and to depress the region to the left of the peak when T.2 is operative. Thus, in principle, this second-order intensity function can be used to determine whether someone is systematically observing the first or the second of two flashes observed in short succession.

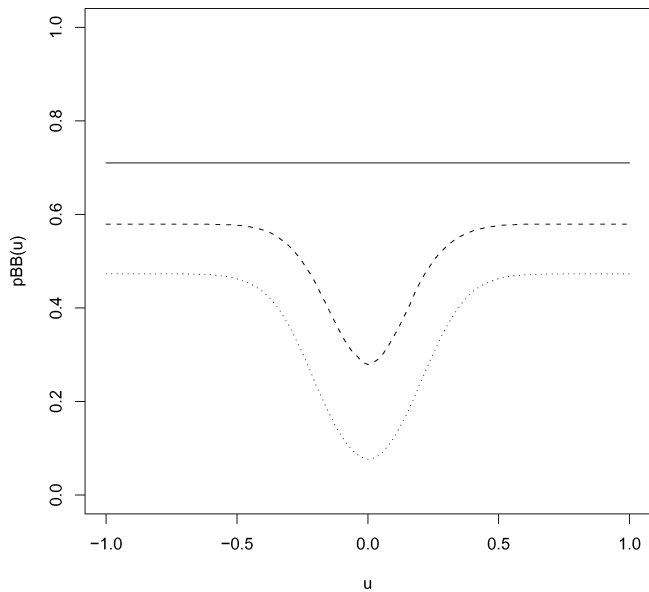
The presence of the term involving  $f_V(u)$  in (6) and (7) suggests that a plot of an estimate of  $p_{AB}(u)$  will be of some use in identifying the density of the  $V_i$ . However, the lower panel in Figure 1 indicates that it might be difficult to distinguish between normal and log-normal cases. The log-normal density exhibits some asymmetry, which would usually help to distinguish it from the normal density. However, the thinning mechanism also imposes a slight asymmetry on  $p_{AB}(u)$ , even in the normal case. Thus, the model tells us that distinguishing between normal and log-normal RT densities may be difficult. On the other hand, if the truth is really log-normal, the normal will provide a very good approximation.

More information about the nonlinear inhibition parameter  $d$  is contained in the function  $p_{BB}(u)$ . If  $d = 0$ , then  $p_{BB}(u)$  is constant. Otherwise, there is a trough centered at the origin. The width of this trough is governed by the magnitude of the parameter  $d$ , as exhibited in Figure 2, which contains plots of  $p_{BB}(u)$  for three values of  $d$  while the other parameters are held constant.

We have seen that the parametric model predicts certain behavior in the functions  $p_{AB}(u)$  and  $p_{BB}(u)$ . Nonparametric



**Figure 1.** The top panel shows the function  $p_{AB}(u)$  for log-normal RT under thinning assumptions T.1 (solid curve) and T.2 (dotted curve), with  $p_S = 0.015$ ,  $p = 0.1$ ,  $d = 0.25$ ,  $\mu = -0.8$ , and  $\sigma = 0.2$ . The lower panel shows the function  $p_{AB}(u)$  under the normal RT model with thinning assumption T.2 (solid curve) using  $\mu = 0.42$ ,  $\sigma = 0.1$ ,  $d = 0.16$ ,  $p = 0.182$ , and  $p_S = 0.015$ . The dotted curve represents  $p_{AB}(u)$  under the log-normal model with thinning assumption T.1 using  $\mu = -0.78$ ,  $\sigma = 0.22$ ,  $d = 0.2$ ,  $p = 0.17$ , and  $p_S = 0.015$ .



**Figure 2.** The function  $p_{BB}(u)$  in the log-normal case with  $p_S = 0.015$ ,  $p = 0.17$ ,  $\mu = -0.78$ ,  $\sigma = 0.22$ , and  $d = 0$  (solid curve),  $d = 0.1$  (dashed curve), and  $d = 0.2$  (dotted curve).

methods must be used to estimate these functions (e.g., Brillinger, 1975b, or Brillinger, 1994). An estimate for the first-order intensity  $p_B$  is given by

$$\hat{p}_B = B([0, T])/T. \tag{9}$$

For the second-order intensity estimates, let  $K(x)$  denote a symmetric second-order smoothing kernel with compact support, and let  $h$  be an appropriately chosen bandwidth. In this article, we employ the biweight kernel,  $K(x) = (15/16) \times (1 - x^2)^2 I_{(|x| < 1)}$ . Set

$$\hat{p}_{AB,h}(u) = (Th)^{-1} \sum_i \sum_j K \left\{ \frac{u - (B_j - A_i)}{h} \right\} \tag{10}$$

and

$$\hat{p}_{BB,h}(u) = (Th)^{-1} \sum_i \sum_{j \neq i} K \left\{ \frac{u - (B_j - B_i)}{h} \right\}. \tag{11}$$

Brillinger (1975b) has shown that under the cumulant mixing conditions alluded to above, the intensity function estimates are asymptotically normal as  $h \rightarrow 0$  and  $Th \rightarrow \infty$ . The respective means are

$$E\{\hat{p}_{AB,h}(u)\} = \int_{-1}^1 K(u_1) p_{AB}(u - u_1 h) du_1 \tag{12}$$

and

$$E\{\hat{p}_{BB,h}(u)\} = \int_{-1}^1 K(u_1) p_{BB}(u - u_1 h) du_1. \tag{13}$$

The bias in each case is of order  $O(h^2)$ . The asymptotic variances are given by

$$\frac{p_{AB}(u)}{Th} \quad \text{and} \quad \frac{p_{BB}(u)}{Th}.$$

#### 4. Estimation and Testing Methodology

We now give an outline of the calculations required to render accurate plots of and confidence bands for second-order intensity functions. Such accuracy requires a good bandwidth, and this, in turn, is facilitated by the estimation of the parameters in our model.

##### 4.1 Parameter Estimation

Maximum likelihood estimation is very difficult, if not impossible, for our parametric model, because of the lack of a closed-form expression for the likelihood function. The complete intensity function cannot be expressed in a useful form; thus, the likelihood function (e.g., Daley and Vere-Jones, 1988) cannot be obtained. One might try an EM-style algorithm where, for example, the  $V_j$ 's are considered as latent variables. Such an algorithm would be highly computationally intensive and likely unstable. Instead, we have adopted a quick approximation method.

Our algorithm requires several passes through the response data. During the first pass, each response time  $B_j$  is matched to the flash time immediately preceding it (denote this flash time by  $\hat{A}_j''$ ). The differences  $B_j - \hat{A}_j''$  are taken as initial estimates of the reaction times  $V_j$ . The mean and standard deviation of the  $V_j$  estimates are then computed (on the log

scale, if the log-normal model is used). On succeeding passes through the responses, each  $B_j$  is matched with a preceding flash now chosen so that the resulting difference  $B_j - \hat{A}_j''$  is closest to the mean of the current set of  $V_j$  estimates. The  $V_j$  estimates are thus updated, and after each pass through the responses, the mean and standard deviation are also updated accordingly. After a few such iterations, the  $V_j$  estimates and their corresponding mean and variance stabilize. We used six iterations to obtain our estimates.

Occasionally, a single flash  $\hat{A}_j''$  corresponds to more than one response. The number of such flashes divided by the duration of the experimental run is used as an estimate of  $p_S$ .

The parameters  $p$  and  $d$  can be estimated using a least-squares technique. First, note that the nonparametric estimator  $\hat{p}_{BB,h}(u)$  is, by definition, unbiased for  $E\{\hat{p}_{BB,h}(u)\}$ , for any bandwidth  $h$ . Thus, using an arbitrary bandwidth, a set of gridpoints  $u_i$  in an appropriate interval, and the previously obtained estimates for  $\mu$ ,  $\sigma$ , and  $p_S$ , the quantity

$$\sum_i [\hat{p}_{BB,h}(u_i) - E\{\hat{p}_{BB,h}(u_i)\}]^2$$

can be minimized with respect to  $d$  and  $p$ . The expectation can be computed numerically; the trapezoid method is satisfactory. A further simplification results from noting that the nonparametric estimator for  $p_B$  given at (9), together with the equation (5), allows the parameter  $p$  to be expressed as an explicit function of  $d$ . Thus,  $E\{\hat{p}_{BB,h}(u)\}$  can be expressed explicitly in terms of  $d$  only. Thus, the minimization is univariate. We used 401 evenly spaced gridpoints in  $(0, 1)$ , and bandwidth  $h = 0.4$  in our computations. Although any bandwidth can be used, the efficiency of the estimator for  $d$  will vary.

We should also mention an alternative method for estimating  $\mu$  and  $\sigma$ , in which one minimizes

$$\sum_i [\hat{p}_{AB,h}(u_i) - E\{\hat{p}_{AB,h}(u_i)\}]^2$$

with respect to  $\mu$  and  $\sigma$ . To do the minimization,  $E\{\hat{p}_{AB,h}(u)\}$  is evaluated at the estimates of  $d$ ,  $p$ , and  $p_S$ , obtained using the above algorithm, and under one of the two thinning assumptions.

Weak consistency of the least-squares estimators for  $\mu$ ,  $\sigma$ , and  $d$  can be established using Taylor's theorem, together with the fact that the nonparametric estimators for  $p_{AB}(u)$  and  $p_{BB}(u)$  are consistent. Consistency of the estimators for  $p$  and  $p_S$  may be more difficult to establish.

#### 4.2 Bandwidth Selection

The bandwidth minimizing the asymptotic mean squared error (AMSE) in  $\hat{p}_{BB,h}(u)$  is given by

$$h_{\text{AMSE}} = \left[ \frac{35p_{BB}(u)}{\{p_{BB}''(u)\}^2 T} \right]^{1/5} \quad (14)$$

when the biweight kernel is used. This bandwidth will render a very good approximation to  $p_{BB}(u)$ . Often, one tries to obtain a reasonable approximation over an interval. Then, one minimizes the asymptotic mean integrated squared error (AMISE) (e.g., Wand and Jones, 1995) over the given interval

with the bandwidth

$$h_{\text{AMISE}} = \left[ \frac{35 \int p_{BB}(u)}{\{\int p_{BB}''(u)\}^2 T} \right]^{1/5}. \quad (15)$$

Good estimates of both  $h_{\text{AMSE}}$  and  $h_{\text{AMISE}}$  are obtainable using our fitted model. Because of the similarity between the log-normal and normal distributions at the relevant parameter settings, it is sufficient to develop these expressions for the normal case. These expressions take a simple form; the log-normal case is feasible, but the slight increase in accuracy is not worth the accompanying increase in computational complexity.

Using the formula for  $p_{BB}(u)$  given at (8), and the fact that  $V_1$  and  $V_2$  are hypothesized to be independent normal random variables with mean  $\mu$  and variance  $\sigma^2$ , we can write

$$p_{BB}(u) = p_B^2 \left\{ 1 + \Phi\left(\frac{x-d}{\sqrt{2}\sigma}\right) - \Phi\left(\frac{x+d}{\sqrt{2}\sigma}\right) \right\}$$

where  $\Phi(\cdot)$  denotes the standard normal c.d.f. The second derivative of this is

$$p_{BB}''(u) = \frac{p_B^2}{2\sigma^2} \left\{ (x-d)\phi\left(\frac{x-d}{\sqrt{2}\sigma}\right) - (x+d)\phi\left(\frac{x+d}{\sqrt{2}\sigma}\right) \right\}.$$

Here,  $\phi(\cdot)$  denotes the standard normal p.d.f.

Replacing  $p_{BB}(u)$  by  $p_{AB}(u)$  everywhere in (14) and (15) yields the corresponding optimal bandwidths for the estimate of  $p_{AB}(u)$ . Again, the normal and lognormal cases give very similar numerical results. Furthermore, the results do not depend strongly on the thinning assumptions. In the normal case, under T.1, we use

$$p_{AB}(u) = p_A p_B \left\{ 1 - \Phi\left(\frac{u-\mu}{\sigma}\right) + \Phi\left(\frac{u-d-\mu}{\sigma}\right) + \frac{1}{(p_A + p_S)\sigma} \phi\left(\frac{u-\mu}{\sigma}\right) \right\}$$

and

$$p_{AB}''(u) = \frac{p_A p_B}{\sigma^2} \left[ (u-d-\mu)\phi\left(\frac{u-d-\mu}{\sigma}\right) + \left\{ \frac{(u-\mu)^2}{\sigma^2} - 1 - u + \mu \right\} \phi\left(\frac{u-\mu}{\sigma}\right) \right].$$

#### 4.3 Bootstrapped Confidence Bands

To obtain accurate estimates of the intensities, the bandwidth should be chosen as described in the previous section, but when producing confidence bands, a smaller bandwidth is usually recommended (e.g., Hall, 1992) because of bias. We sidestep this issue by studying the expected value of the intensity estimates instead of the intensities themselves.

The following technique can be used to obtain accurate confidence bands for  $E\{\hat{p}_{BB}(u)\}$  (see, e.g., Davison and Hinkley (1997)).

1. Estimate the model parameters for the given data.
2. Generate  $R$  samples using our model and the estimated parameters.
3. For each of these samples, calculate a nonparametric estimate  $\hat{p}_{BB,h}(u)$  and determine its minimum and

maximum deviation from  $E\{\hat{p}_{BB,h}(u)\}$

$$m = \min_{u \in I} \sqrt{\hat{p}_{BB,h}(u)} - E\{\sqrt{\hat{p}_{BB,h}(u)}\} \quad \text{and}$$

$$M = \max_{u \in I} \sqrt{\hat{p}_{BB,h}(u)} - E\{\sqrt{\hat{p}_{BB,h}(u)}\}$$

on the domain  $I$  of interest. The delta method is used to compute an approximation to  $E\{\hat{p}_{BB,h}(u)\}(1/2)$ .

4. Let  $D_m(h)$  and  $D_M(h)$  be the  $(\alpha/2) \cdot B$ th and the  $(1 - \alpha/2) \cdot B$ th quantiles  $m$  and  $M$ , respectively.

5. Define  $1 - \alpha$  confidence bands by

$$\{\sqrt{\hat{p}_{BB,h}(u)} + D_m(h)\}^2 \quad \text{and} \quad \{\sqrt{\hat{p}_{BB,h}(u)} + D_M(h)\}^2,$$

for  $u \in I$ , where  $\hat{p}_{BB,h}(u)$  is the nonparametric estimate computed for the original data. The bandwidth derived in the preceding section can be used here.

The validity of this parametric bootstrap procedure does not depend on the bandwidth  $h$ . However, the width of the confidence bands depends on  $h$ . A reasonable choice is to use the optimal AMISE bandwidth, as described in Section 4.2. Note the use of the variance-stabilizing square-root transformation (e.g., Brillinger, 1992).

These confidence bands can be used to test for nonlinear inhibition: in the absence of nonlinear inhibition,  $p_{BB}(u)$  and  $E\{\hat{p}_{BB,h}(u)\}$  should be constant; thus, we can conclude that there is nonlinear inhibition if a horizontal line cannot be contained within the confidence bands.

Confidence bands for  $E\{\hat{p}_{AB,h}(u)\}$  can be obtained by replacing  $\hat{p}_{BB}(u)$  by  $\hat{p}_{AB}(u)$  and  $\hat{p}_{BB,h}(u)$  by  $\hat{p}_{AB,h}(u)$  everywhere in the above algorithm. Using the confidence bands for  $E\{\hat{p}_{AB,h}(u)\}$ , one may test the appropriateness of the respective thinning assumptions in the following way. If the null hypothesis is T.1, then one obtains confidence bands for  $E\{\hat{p}_{AB,h}(u)\}$  and overlays a graph of  $E\{\hat{p}_{AB,h}(u)\}$  computed under the T.1 assumption and evaluated at the parameter estimates. (The alternative method of estimating  $\mu$  and  $\sigma$  works better here, since the thinning assumption can be explicitly used.) If the overlaid curve falls outside the confidence bands, we have evidence against the assumption T.1 as postulated by the model. Similarly, we can test the null hypothesis of T.2.

We can also test normality versus log-normality by overlaying the intensity function evaluated at the estimated parameter values, and identifying regions where the fitted intensity falls outside the confidence bands.

The power of the tests of the thinning assumptions increases with  $d$  and  $T$ , while the power of the tests of normality versus log-normality can be low unless  $T$  is very large. When  $d = 0.25$  and  $T = 300$ , simulations indicate that if the true model is log-normal with thinning assumption T.2, then the power at the normal T.2 alternative is around 10% for a size 5% test. By contrast, the power at a normal or log-normal T.1 alternative exceeds 70%. Thus, we have a hope of distinguishing between the thinning assumptions, but we will have lower power for distinguishing between normality and log-normality; this is not surprising, since we demonstrated the potential for such difficulties in Figure 1.

### 5. Results and Discussion

We applied our methodology to 12 runs of data from a real RT experiment. The runs lasted about five minutes (i.e.,  $T =$

**Table 1**

Parameter estimates for each of the twelve reaction experiment runs

Experiment no.	$\hat{\mu}$	$\hat{\sigma}$	$\hat{d}$	$\hat{p}$	$\hat{p}_S$
1	0.435	0.090	0.279	0.166	0.003
2	0.476	0.101	0.215	0.225	0.013
3	0.441	0.103	0.229	0.048	0.015
4	0.431	0.088	0.243	0.110	0.007
5	0.443	0.088	0.207	0.121	0.036
6	0.438	0.082	0.171	0.196	0.019
7	0.439	0.102	0.193	0.163	0.026
8	0.467	0.096	0.185	0.314	0.009
9	0.449	0.102	0.301	0.105	0.010
10	0.445	0.087	0.271	0.087	0.015
11	0.441	0.091	0.154	0.139	0.028
12	0.421	0.105	0.221	0.149	0.014

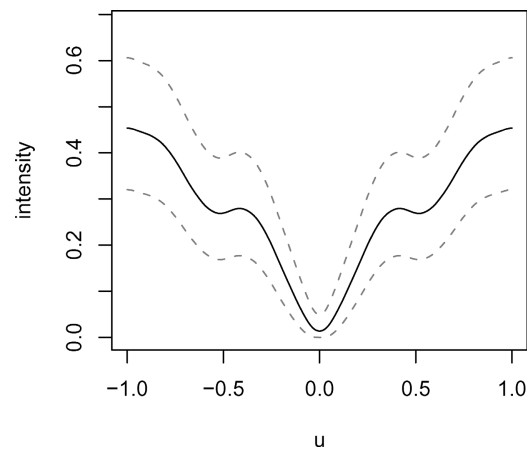
300 seconds), during which a single observer was presented with flashes of a small spot having diameter 0.5-minute arc, which was viewed at a distance of 3.4 m. The rate  $p_A = 1$  had been chosen by Simpson et al. (2000) to be roughly comparable with classical RT experiments. Details of the experiment are given in Simpson et al. (2000), and the data are available at <http://fhis.gcal.ac.uk/VS/wsi/software/>.

#### 5.1 The Fitted Model

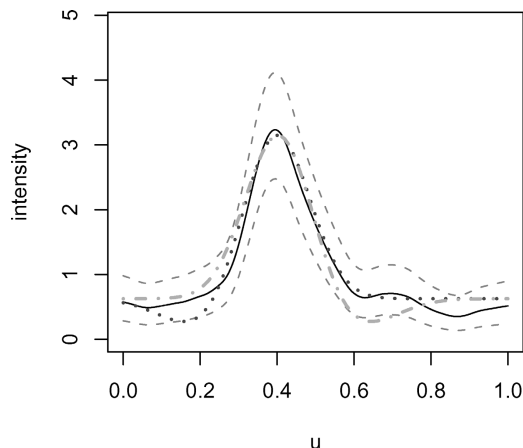
Table 1 displays the estimates of the model parameters, assuming normality. As might be expected, the estimates of  $\mu$  and  $\sigma$  are fairly stable, near 0.44 and 0.1, respectively. The estimates of  $d$  are considerably more variable, ranging from 0.15 to 0.3. Estimates of  $p$  are more variable still, while the estimates of  $p_S$  tend to be very small, suggesting that there is not a lot of superposed noise in the system.

#### 5.2 Is There Nonlinear Inhibition?

Figure 3 displays bootstrapped 95% confidence bands for the function  $p_{BB}(u)$  based on the first run of experimental data, using the AMISE-optimal bandwidth described in Section 4.2.



**Figure 3.** Bootstrapped 95% confidence bands (dashed curves) for the function  $p_{BB}(u)$  (estimated, using  $h = 0.35$ , is the solid curve) based on the first experimental run of reaction time data.



**Figure 4.** Bootstrapped 95% confidence bands (dashed curves) for the function  $p_{AB}(u)$  (solid curve is the estimate, using  $h = 0.1$ ) based on the first experimental run of reaction time data. Overlaying the plot are graphs of  $E\{\hat{p}_{AB,h}(u)\}$  under thinning assumptions T.1 (dashed-dotted curve) and T.2 (dotted curve).

The hypothesis of no nonlinear inhibition is rejected since no horizontal line can be contained within the confidence bands.

The other 11 data sets yield the same result. We have clear evidence, in each case, that the flashes are not being thinned in a completely random manner.

### 5.3 Testing the Thinning Assumptions Using $p_{AB}(u)$

Figure 4 displays 95% confidence bands for the function  $p_{AB}(u)$  based on the first run of experimental data, using the AMISE-optimal bandwidth. A slight asymmetry is evident.

Graphs of  $E\{\hat{p}_{AB,h}(u)\}$  under thinning assumptions T.1 and T.2 have been added to the plot. In both cases, we have used the normal model with parameters estimated by continuous least squares, as described in Section 4.3. It is clear that the data are not consistent with assumption T.1, while they may, at best, be marginally consistent with T.2. In any event, the inhibition mechanism suggested by the model does not sufficiently explain these data. One might speculate that a combination of the two types of thinning is operative here.

In studying the other experimental runs, different conclusions are reached. For example, in the second experimental run, we still see inconsistency with T.1, but the curve corresponding to T.2 lies within the confidence bands, except in the region near  $u = 0.8$ , where there is a statistically significant depression. This extra dip appears in several of the data sets and is not explained by the parametric model at all. One possibility is that this extra oscillation is induced by the mechanics of the subject's finger.

It was found that the hypothesis T.1 was rejected in all but 1 of the other 10 data sets. The hypothesis T.2 was rejected in 4 of these data sets. Thus, the thinning assumption T.1, as postulated under the parametric model with a normally distributed reaction time, is not consonant with the data. Under normality, thinning assumption T.2 often gives a better fit to the data, but it is not completely adequate. We also tested the two thinning assumptions under the log-normal assumption, whence T.1 was rejected only 5 out of 12 times, while T.2 was rejected in 6 cases.

We conclude that the eye-brain-hand system is deleting flashes in a nonlinear way, but the mechanism suggested by our model does not give a complete description as to how this is happening.

### 5.4 Conclusions

We have demonstrated how the interplay between a parametric model and nonparametric estimation can lead to improved analysis and interpretation of point-process data. The model tells us that we can detect the presence of a nonlinear inhibition by examining a plot of the the second-order intensity function estimate for the responses. If the subject is systematically observing the first or the second of two flashes observed in short succession, this kind of thinning behavior can be deduced from the estimate of the second-order intensity function relating the responses to the flashes.

The model facilitated the selection of good bandwidths for second-order intensity estimates as well as a parametric bootstrap procedure. It has allowed us to reach some firm conclusions about the given data: a nonlinear thinning mechanism is likely operative in the eye-brain-hand system of the subject. The nonlinear inhibition mechanism given by assumption T.1 appears to be incompatible with a normal reaction time distribution.

The model has revealed limitations in the nonparametric methodology: distinguishing between normal and lognormal reaction time densities would require substantially more data, because of a partial confounding with the type of thinning mechanism.

In turn, the nonparametric estimates allow us to detect some weaknesses in the proposed model. For example, the exact nature of the thinning operation is not clear; we can say with certainty that it is more complex than what the model postulates.

### ACKNOWLEDGEMENTS

The authors are grateful to two anonymous referees for constructive comments that have led to substantial improvements in the article. Part of this work was carried out when WJB and VR were visiting the Centre for Mathematics and Its Applications at the Australian National University. WJB, JP, and VR gratefully acknowledge the Centre and the additional support of Peter Hall. WJB and WAS were supported by grants from the Natural Sciences and Engineering Research Council of Canada. VR was partly supported by a grant from the Swiss National Science Foundation (81NE-54413). The authors also thank Rolf Turner for his careful reading of an earlier version of the manuscript, and Harry Joe for some helpful suggestions.

### RÉSUMÉ

Nous proposons un modèle paramétrique simple pour des données de processus ponctuel issues d'une expérience sur des temps de réaction. Il est utilisé pour vérifier statistiquement la présence et la nature d'une inhibition non-linéaire dans le système œil—cerveau—main, mais aussi pour étudier la nature de la distribution du retard de temps de réaction. Le modèle indique, en principe, que l'estimation de l'intensité de second ordre peut être utilisée pour déterminer si le sujet expérimental observe systématiquement le premier ou le second des deux flashes transmis pendant une période courte. En particulier, le modèle tient compte dans le calcul des

bonnes distances pour l'évaluation de la courbe d'intensité. Un bootstrap paramétrique peut aussi être mis en œuvre. Nos méthodes sont illustrées avec douze jeux de données d'une expérience réelle sur des temps de réaction. Nous trouvons que l'inhibition non-linéaire est présente dans le système œil—cerveau—main. Cependant, il n'y a pas assez de données pour distinguer entre la lognormalité et la normalité de la distribution du temps de réaction, cela est du, en partie, à une confusion avec la non-linéarité de l'inhibition présente dans le système.

REFERENCES

Braun, W. J. and Kulperger, R. J. (1998). A bootstrap for point processes. *Journal of Statistical Computation and Simulation* **60**, 129–155.

Brillinger, D. (1975a). The identification of point process systems. *Annals of Probability* **3**, 909–929.

Brillinger, D. (1975b). Statistical inference for stationary point processes. In *Stochastic Processes and Related Topics*, Volume 1, M. Puri (ed), 55–99. New York: Academic Press.

Brillinger, D. (1976). Identification of synaptic interactions. *Biological Cybernetics* **22**, 213–228.

Brillinger, D. (1992). Nerve cell spike train data analysis: A progression of technique. *Journal of the American Statistical Association* **87**, 261–271.

Brillinger, D. (1994). Time series, point processes, and hybrids. *Canadian Journal of Statistics* **22**, 177–206.

Cornsweet, T. (1970). *Visual Perception*. New York: Academic.

Daley, D. J. and Vere-Jones, D. (1988). *An Introduction to the Theory of Point Processes*. New York: Springer.

Davison, A. and Hinkley, D. V. (1997). *Bootstrap Methods and Their Application*. Cambridge, U.K.: Cambridge University Press.

Green, D. and Smith, A. (1982). Detection of auditory signals occurring at random times: Intensity and duration. *Perception and Psychophysics* **31**, 117–127.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. New York: Springer.

Karr, A. (1986). *Point Processes and Their Statistical Inference*. New York: Marcel-Dekker.

Reiss, R.-D. (1993). *A Course on Point Processes*. New York: Springer.

Simpson, W. A., Braun, W. J., Bargaen, C., and Newman, A. J. (2000). Identification of the eye-brain-hand system with point processes: A new approach to simple reaction time. *Journal of Experimental Psychology: Human Perception and Performance* **26**, 1675–1690.

Stoyan, D., Kendall, W. S., and Mecke, J. (1995). *Stochastic Geometry and its Applications*, 2nd edition. New York: Wiley.

Ulrich, R. and Miller, J. (1993). Information processing models generating lognormally distributed reaction times. *Journal of Mathematical Psychology* **37**, 513–525.

Vere-Jones, D. (2002). Tutorial on point process modelling. Presented at the IMA Workshop on Point Processes and Seismology, Minneapolis, Minnesota.

Wand, M. and Jones, M. (1995). *Kernel Smoothing*. London: Chapman and Hall.

Received July 2000. Revised January 2003.  
Accepted January 2003.

APPENDIX

Derivations of the Intensity Functions

We first derive the expression for the first-order intensity of the  $B$  process, noting first that for small  $h > 0$ ,

$$E\{B(x, x + h)\} = \int_x^{x+h} p_B(x_1) dx_1. \tag{A.1}$$

It is also true that

$$E\{dB(x, x + h)\} = \int_x^{x+h} E\{dB(x_1)\}. \tag{A.2}$$

Because of the translation mechanism in the model, we have

$$E\{dB(x_1)\} = E\{dA''(x_1 - V_{x_1})\} \tag{A.3}$$

where the  $V_x$  represent independent random variables from a population having density  $f(v)$ . Conditioning on the  $A''$  process, we have

$$E\{dB(x_1)\} = \int E\{dA''(x_1 - v)\}f(v) dv. \tag{A.4}$$

Under thinning assumption T.1, we have

$$\begin{aligned} E\{dA''(x_1 - v)\} &= E[dA'(x_1 - v)I\{A'(x_1 - v - d, x_1 - v) = 0\}I(U_{x_1-v} > p)] \\ &\tag{A.5} \end{aligned}$$

and under T.2, we have

$$\begin{aligned} E\{dA''(x_1 - v)\} &= E[dA'(x_1 - v)I\{A'(x_1 - v, x_1 - v + d) = 0\}I(U_{x_1-v} > p)]. \\ &\tag{A.6} \end{aligned}$$

Here, the  $U_x$  are independent uniform random variables on  $[0, 1]$ . Conditioning on the  $A'$  process and recognizing that  $A' = A + S$ , where  $S$  is an independent Poisson process representing the superposed “flashes,” we have

$$\begin{aligned} E\{dA''(x_1 - v)\} &= (1 - p)E\{\{dA(x_1 - v) + dS(x_1 - v)\} \\ &\quad \times I\{A(x_1 - v - d, x_1 - v) = 0\} \\ &\quad \times I\{S(x_1 - v - d, x_1 - v) = 0\}\}. \\ &\tag{A.7} \end{aligned}$$

or

$$\begin{aligned} E\{dA''(x_1 - v)\} &= (1 - p)E\{\{dA(x_1 - v) + dS(x_1 - v)\} \\ &\quad \times I\{A(x_1 - v, x_1 - v + d) = 0\} \\ &\quad \times I\{S(x_1 - v, x_1 - v + d) = 0\}\}. \\ &\tag{A.8} \end{aligned}$$

under T.1 and T.2, respectively.

Stationarity of  $A$  and  $S$  gives (in both cases)

$$\begin{aligned} E\{dA''(x_1)\} &= (1 - p)E\{\{dA(x_1) + dS(x_1)\} \\ &\quad \times I\{A(x_1 - d, x_1) = 0\} \\ &\quad \times I\{S(x_1 - d, x_1) = 0\}\}, \\ &\tag{A.9} \end{aligned}$$

so that

$$E\{dA''(x_1 - v)\} = (1 - p)(p_A + p_S)e^{-d(p_A + p_S)} dx_1. \quad (\text{A.10})$$

Inserting this expression into (A.4) and using (A.2) we obtain

$$E\{B(x, x + h)\} = h(1 - p)(p_A + p_S)e^{-d(p_A + p_S)} \int f(v) dv. \quad (\text{A.11})$$

Because  $f(v)$  is a probability density function and because of (A.1), we conclude that

$$\int_x^{x+h} p_B(x_1) dx_1 = h(1 - p)(p_A + p_S)e^{-d(p_A + p_S)}. \quad (\text{A.12})$$

Dividing by  $h$ , and letting  $h \rightarrow 0$ , we see that the first-order intensity function for the  $B$  process is given by

$$p_B(x) = (1 - p)(p_A + p_S)e^{-d(p_A + p_S)}. \quad (\text{A.13})$$

We now derive the expression for  $p_{AB}(u)$ , under the thinning assumption T.1. The argument needs only minor modifications (analogous to those described above in deriving  $p_B$ ) under T.2.

Let  $x, y$ , and  $h$  be real variables with  $h > 0$ . Then

$$E\{A(x, x + h)B(y, y + h)\} = E\left\{\int_x^{x+h} \int_y^{y+h} dA(x_1)dB(y_1)\right\} \quad (\text{A.14})$$

and

$$\begin{aligned} E\{dA(x_1)dB(y_1)\} &= E\{dA(x_1)dA''(y_1 - V_{y_1})\} \\ &= \int E\{dA(x_1)dA''(y_1 - v)\}f(v) dv. \end{aligned} \quad (\text{A.15})$$

The thinning mechanism leads to

$$\begin{aligned} E\{dA(x_1)dA''(y_1 - v)\} \\ = (1 - p)E[dA(x_1)dA'(y_1 - v)I\{A'(y_1 - v - d, y_1 - v) = 0\}]. \end{aligned} \quad (\text{A.16})$$

The independent increments property of the Poisson process allows us to write

$$\begin{aligned} E[dA(x_1)dA'(y_1 - v)I\{A'(y_1 - v - d, y_1 - v) = 0\}] \\ = p_A(p_A + p_S)e^{-d(p_A + p_S)}I\{x_1 \notin (y_1 - v - d, y_1 - v)\} \\ + p_Ae^{-d(p_A + p_S)}\delta(y_1 - x_1 - v) dx_1 d(y_1 - v), \end{aligned} \quad (\text{A.17})$$

where  $\delta(x)$  refers to the Dirac delta distribution. Upon substitution into (A.16) and then into (A.15) and (A.14), we have

$$\begin{aligned} \int_x^{x+h} \int_y^{y+h} E\{dA(x_1)dB(y_1)\} \\ = (1 - p) \left[ \int_x^{x+h} \int_y^{y+h} \int p_A(p_A + p_S)e^{-d(p_A + p_S)} \right. \\ \left. \times I\{x_1 \notin (y_1 - v - d, y_1 - v)\}f(v) dv dx_1 dy_1 \right. \end{aligned}$$

$$\begin{aligned} &+ \left. \int_x^{x+h} \int_y^{y+h} \int p_Ae^{-d(p_A + p_S)} \right. \\ &\left. \times \delta(y_1 - x_1 - v)f(v) dv dx_1 dy_1 \right] \\ &= (1 - p)p_Ae^{-d(p_A + p_S)}h^2 \\ &\times [(p_A + p_S)P\{x \notin (y - V - d, y - V)\} + f(y - x)] \\ &+ O(h^3). \end{aligned} \quad (\text{A.18})$$

The last equality follows upon application of a first-order Taylor expansion of  $f(v)$  about  $y - x$ . Dividing by  $h^2$  and letting  $h \rightarrow 0$ , we have

$$\begin{aligned} p_{AB}(y - x) &= (1 - p)p_Ae^{-d(p_A + p_S)}(p_A + p_S) \\ &\times \left[ P\{V \notin (y - x - d, y - x)\} + \frac{f(y - x)}{p_A + p_S} \right]. \end{aligned} \quad (\text{A.19})$$

Finally,  $y - x$  may be replaced everywhere by the variable  $u$ , and we may use (A.13) to obtain

$$p_{AB}(u) = p_Ap_B \left[ P\{V \notin (u - d, u)\} + \frac{f(u)}{p_A + p_S} \right]. \quad (\text{A.20})$$

Turning to  $p_{BB}(u)$ , we note that under the same thinning assumption as above, similar arguments to those given above lead to

$$\begin{aligned} E\{B(x, x + h)B(y, y + h)\} \\ = (1 - p)^2 \iint E \left[ \int_x^{x+h} \int_{y: y_1 \neq x_1}^{y+h} dA'(x_1 - v_1)dA'(y_1 - v_2) \right. \\ \times I\{A'(x_1 - v_1 - d, x_1 - v_1) = 0\} \\ \times I\{A'(y_1 - v_2 - d_1, y_1 - v_2) = 0\} \\ \left. \times f(v_1)f(v_2) dx_1 dy_1 \right] dv_1 dv_2 \\ + p_B|(x, x + h) \cap (y, y + h)|_L. \end{aligned} \quad (\text{A.21})$$

In view of (4), we focus on the first term on the right-hand side of the above equation. Independence of the Poisson increments allows us to rewrite this term as

$$\begin{aligned} \int \int h^2 p_A^2 e^{-2dp_A} f(v_1)f(v_2) \\ \times I\{x - v_1 \notin (y - v_2 - d, y - v_2)\} \\ \times I\{y - v_2 \notin (x - v_1 - d, x - v_1)\} dv_1 dv_2. \end{aligned} \quad (\text{A.22})$$

Another application of (A.13), division by  $h^2$ , and taking the limit as  $h \rightarrow 0$  gives the result

$$p_{BB}(y - x) = p_B^2 P(|V_2 - V_1 - y + x| > d). \quad (\text{A.23})$$

The other thinning assumption leads to this result as well. Again,  $y - x$  can be replaced by the variable  $u$ .