Journal of the Korean Statistical Society (2002), 31: 3, pp 343–357

Inhomogeneous Poisson Intensity Estimation via Information Projections onto Wavelet Subspaces

Woo-Chul Kim[1] and Ja-Yong Koo[2]

ABSTRACT

This paper proposes a method for producing smooth and positive estimates of the intensity function of an inhomogeneous Poisson process based on the shrinkage of wavelet coefficients of the observed counts. The information projection is used in conjunction with the level-dependent thresholds to yield smooth and positive estimates. This work is motivated by and demonstrated within the context of a problem involving gamma-ray burst data in astronomy. Simulation results are also presented in order to show the performance of the information projection estimators.

Keywords. Astronomy, gamma-ray bursts, Newton-Raphson algorithm, wavelet shrinkage.
AMS 2000 Subject Classifications. Primary 62G07.

1. Introduction

Consider an inhomogeneous Poisson process N_t on $I = [0, 1]$. Let $t_i = i/n$ for $i = 0, \ldots, n$ and define the differences

$$Y_i = N_{t_i} - N_{t_{i-1}}, \quad i = 1, \ldots, n.$$

The data may be expressed as n independent Poisson random variables $Y_1, \ldots, Y_n$. The problem is to estimate the *positive* intensity function of the inhomogeneous Poisson process based on $Y_1, \ldots, Y_n$.

This kind of problems arise in fields, such as astronomy and medical imaging, where an increasingly larger and larger amounts of data take the form of binned

Received November 2001; accepted May 2002.

[1]Department of Statistics, Seoul National University, Seoul 151-742, Korea (e-mail : wckim@stats.snu.ac.kr)

[2]Department of Statistics, Inha University, Inchon 402-751, Korea (e-mail : jykoo@stat.inha.ac.kr)

Poisson counts with either a temporal or spatial index. Such Poisson signals and images arise quite naturally, for example, when the data are acquired using a CCD (charged-coupled device) device, in which the arrival of individual photons of energy is recorded (Babu and Feigelson, 1996). In the field of medical imaging, similar devices play a central role in experiments involving emission tomography. In collecting Poisson signal and image data, there is typical interest on the part of the scientist in estimating the underlying intensity function. This estimation problem may be of primary interest in and of itself, or of secondary interest as a pre-processing step prior to further scientific analysis.

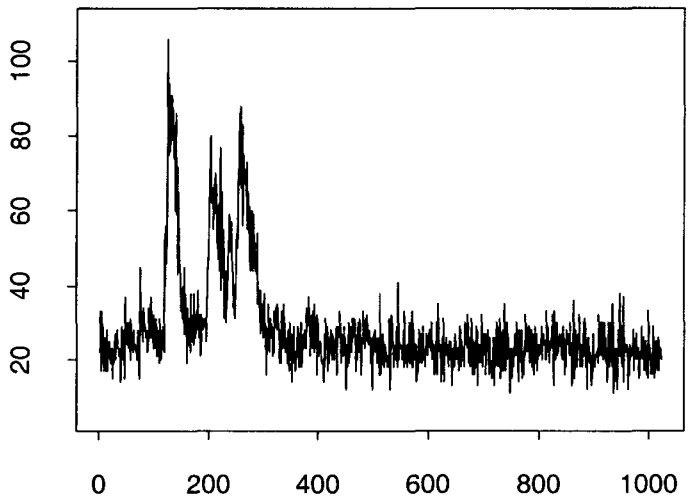

FIGURE 1.1 *Data from gamma-ray burst 551, recorded over 0.94 seconds. The number of bins is 1024 and the length of each bin is 0.918 milliseconds.*

In Figure 1.1 is shown a gamma-ray burst (GRB) signal recorded by the BATSE (Burst and Transient Source Experiment) instruments on board NASA's Compton Gamma Ray Observatory (Meegan *et al.*, 1992). Gamma-rays are at the high end of the electro-magnetic spectrum, and the original data are in the form of arrival times for individual high-energy photons (*i.e.*, 'pieces' of light). These arrival times have been collected into bins of approximately 0.918 milliseconds in length for the purposes of analysis and visual display. Nonparametric estimation of the intensity functions (called 'intensity profiles' in the astrophysical literature) underlying GRBs is potentially useful for studying these phenomena, both on an individual basis and for larger studies of entire catalogues of bursts.

Nonparametric methods for estimating a varying point process intensity have been developed by a number of authors, including Diggle (1985), Diggle and Marron (1988), and Ellis (1986). Kolaczyk (1997) has considered a wavelet method in the estimation of inhomogeneous Poisson processes following the translation-invariant de-noising approach of Donoho and Coifman (1995), and Kolaczyk (1999) has developed corrected versions of the usual Gaussian-based shrinkage thresholds.

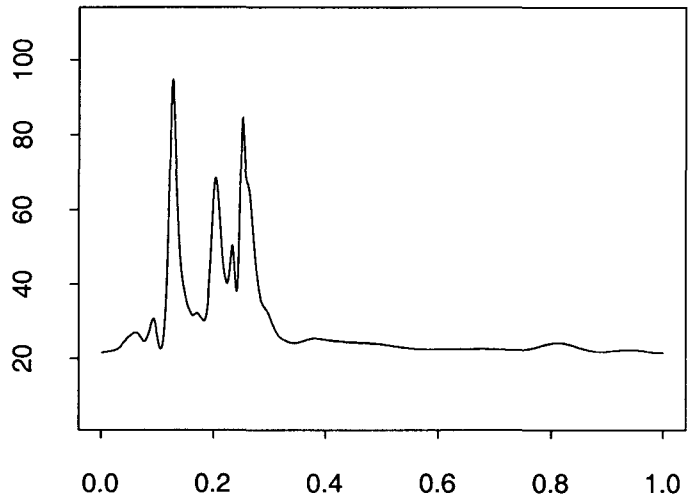

FIGURE 1.2 *Estimate of intensity function for gamma-ray burst 551*

In this paper, we introduce a procedure for the nonparametric estimation of inhomogeneous Poisson intensities, in which the log-intensity is modeled by an orthonormal wavelet basis, up to a given resolution. The implementation is carried out within the context of a standard Poisson regression framework. The model is fit using an implementation of the concept of an *information projection* (Barron and Sheu, 1991). The use of wavelets in our model allows for flexibility at multiple scales in modeling such highly inhomogeneous phenomena as that described above. Also, by incorporating the wavelet parameterization into the exponential family model for Poisson data (as opposed to modeling the intensity function itself with respect to a wavelet basis), we insure the positivity of our estimate. We provide empirical evidence, in the form of simulations and applications, that suggest our method can be expected to perform well in practical

contexts.

Figure 1.2 shows an estimate of the intensity profile underlying Burst 551, obtained using the method proposed herein. The estimate is visually smooth, capturing the three main peaks in the first third of the signal and the constant background throughout the remaining two thirds. It is interesting to note that the small spike between the second and the third peaks stands out clearly in the information projection fit.

The rest of this paper is organized as follows. In Section 2, we formally introduce the model and define our estimator. Section 3 is devoted to a summary of the simulation study. Section 4 is used to pull together a few concluding thoughts and comments.

2. Information Projection Estimator

Recall that we are given an inhomogeneous Poisson process on $I = [0, 1]$:

$$N_t \equiv N(0, t] \sim \mathcal{P}\Big(F_n(t)\Big),$$

where $F_n(t) = \int_0^t f_n(x)dx$ for $t \in I$, the intensity function f_n is positive on I, and $\mathcal{P}(\mu)$ denotes a Poisson distribution with mean μ. Assume that this process is observed at intervals $I_i = [t_{i-1}, t_i)$, $1 \leq i < n$, and $I_n = [t_{n-1}, t_n]$. Define the differences

$$Y_i = N(I_i) \equiv N_{t_i} - N_{t_{i-1}}.$$

The data may be expressed as n independent Poisson random variables $Y_1, \ldots, Y_n$:

$$Y_i \sim \mathcal{P}\Big(\int_{I_i} f_n\Big).$$

We cannot expect to estimate f_n well using only local information, unless the mean of each Y_i is bounded below. As in Cowling, Hall and Phillips (1996), we adopt the intensity model

$$f_n(t) = nf(t),$$

where the smooth function f, which will be referred to as the scaled intensity function, is held fixed. The problem is to estimate the scaled intensity function f based on $Y_1, \ldots, Y_n$ which are independent Poisson random variables having, respectively, the means $\int_{I_i} f_n$, $i = 1, \ldots, n$.

There are two ways of interpretation of Figure 1.2. At first, it displays an estimate of the intensity function f_n such that the mean of Y_i might be estimated

by the integration of the estimate over the interval I_i. Secondly, it shows an estimate of the scaled intensity function f, for which case the estimate of the mean of Y_i is given by the ordinate of the plot at t_i. This is due to the observation: if f is smooth and n is large, then the mean of Y_i, $\int_{I_i} f_n$ would be approximately equal to $f(t_i)$.

Assume we are given an orthonormal wavelet basis of $L_2(I)$. One may use the construction from Meyer (1992) or Cohen, Daubechies and Vial (1993). See also Donoho and Johnstone (1998). In this paper, periodized wavelets are used since the background noise of the data in Figure 1.1 appears to behave periodically. Given a multiresolution analysis with scaling function ϕ and mother wavelet ψ, we define periodized scaling function and mother wavelet by

$$\phi^\circ_{j,k}(t) = \sum_{l\in Z} \phi_{j,k}(t+l), \quad \psi^\circ_{j,k}(t) = \sum_{l\in Z} \psi_{j,k}(t+l)$$

where $\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k)$ and $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$. A periodic function f in $L_2(I)$ has a formal expansion

$$f(t) = \sum_{k=0}^{2^{j_0}-1} a_{j_0,k}\phi^\circ_{j_0,k}(t) + \sum_{j\geq j_0}\sum_{k=0}^{2^j-1} b_{j,k}\psi^\circ_{j,k}(t) \tag{2.1}$$

where $a_{j_0,k}$, $j_0 \geq 0$ and $b_{j,k}$, $j \geq j_0$ are, respectively, the wavelet coefficients of f given by

$$a_{j_0,k} = \int_I f(t)\phi^\circ_{j_0,k}(t)dt \quad \text{and} \quad b_{j,k} = \int_I f(t)\psi^\circ_{j,k}(t)dt.$$

For integers j_0 and j_1 with $j_1 \geq j_0$, let

$$A_{j_0} = \{(j_0,k) : k = 0,\ldots,2^{j_0}-1\}$$

and

$$B_j = \{(j,k) : k = 0,\ldots,2^j - 1\}$$

for $j_0 \leq j \leq j_1$. An element of A_{j_0} or B_j is denoted by $\lambda = (j,k)$. Set

$$w(t;\boldsymbol{a},\boldsymbol{b}) = \sum_{\lambda\in A_{j_0}} a_\lambda\phi^\circ_\lambda(t) + \sum_{j=j_0}^{j_1}\sum_{\lambda\in B_j} b_\lambda\psi^\circ_\lambda(t), \tag{2.2}$$

where $\boldsymbol{a} = (a_\lambda)$ and $\boldsymbol{b} = (b_\lambda)$. Let $\Lambda(j_0,j_1) = A_{j_0}\cup B_{j_0}\cup\cdots\cup B_{j_1}$, and let J denote the number of elements of $\Lambda(j_0,j_1)$. Let $\boldsymbol{\Theta}$ denote the collection of all J-dimensional vectors

$$\boldsymbol{\theta} = \{\theta_\lambda : \lambda\in\Lambda(j_0,j_1)\} = \{a_\lambda : \lambda\in A_{j_0}\}\cup\{b_\lambda : \lambda\in B_j, j_0\leq j\leq j_1\}.$$

It is convenient to use for ϕ°_λ, $\lambda \in A_{j_0}$, the notation ψ°_λ. Then (2.2) can be written as

$$w(t;\boldsymbol{\theta}) = \sum_{\lambda \in \Lambda(j_0,j_1)} \theta_\lambda \psi^\circ_\lambda(t). \tag{2.3}$$

Given $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, set

$$f(t;\boldsymbol{\theta}) = \exp\Big(w(t;\boldsymbol{\theta})\Big).$$

We will refer to the class of positive functions $f(\cdot;\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, as the exponential models. For notational convenience, let $w(\boldsymbol{\theta})$ and $f(\boldsymbol{\theta})$ denote the functions $w(\cdot;\boldsymbol{\theta})$ and $f(\cdot;\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, respectively.

To give an object function to be optimized for the definition of estimators, we note that the log-likelihood of $\{Y_1, \ldots, Y_n\}$ is given by

$$\sum_{i=1}^{n} \left[Y_i \log\Big(\int_{I_i} f_n\Big) - \int_{I_i} f_n \right]. \tag{2.4}$$

If f is smooth and n is sufficiently large, then $\int_{I_i} f_n \approx f(t_i)$ so that an approximation to the log-likelihood (2.4) is given by

$$\sum_{i=1}^{n} \Big[Y_i \log f(t_i) - f(t_i) \Big].$$

Now we define an object function

$$l(\boldsymbol{\theta}) = \sum_{i=1}^{n} Y_i w(t_i;\boldsymbol{\theta}) - C(\boldsymbol{\theta}), \tag{2.5}$$

where

$$C(\boldsymbol{\theta}) = \sum_{i=1}^{n} f(t_i;\boldsymbol{\theta}).$$

Then we can use the Poisson regression idea from the generalized linear models to maximize $l(\boldsymbol{\theta})$. However, note that the object function defined by (2.5) is not necessarily interpretable as a log-likelihood.

Let

$$\hat{\boldsymbol{\theta}} = \operatorname*{argmax}_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} l(\boldsymbol{\theta})$$

and define an estimator of f by $\hat{f} = f(\hat{\boldsymbol{\theta}})$. We will refer to $\hat{f}$ as the information projection estimator (IPE) of the scaled intensity function f. Observe that $\hat{f}$ satisfies the equation

$$\overline{\psi}^\circ_\lambda = \sum_{i=1}^{n} \Big[\psi^\circ_\lambda(t_i) \hat{f}(t_i) \Big],$$

where

$$\overline{\psi}_\lambda^\circ = \sum_{i=1}^{n} Y_i \psi_\lambda^\circ(t_i) \text{ for } \lambda \in \Lambda(j_0, j_1).$$

3. Numerical Results

3.1. Numerical implementation

For the figures in this paper, the Symmlet of Daubechies (1993), which has seven vanishing moments and support length $S = 15$, is used. One difficulty in implementing IPEs is computational. Namely, except for the Haar wavelet, all higher order Daubechies mother and father wavelet functions, for example, the Symmlet used in this paper, have no closed functional form. A simple solution is to have values of the mother and father wavelet given in a table. Evaluation of $\psi_{j,k}^\circ(x)$, for given x, then can be performed by interpolating table values. The package MATLAB is used to generate the table for the values of the mother and father wavelets at equally spaced points. Since the set $\{t_i : i = 1, \ldots, n\}$ is chosen as a subset of the set of points at which those functions are computed, no interpolation is required. Programs for implementing IPE as described in this article have been written in C. Pseudo-random numbers for Poisson variates are generated by `poidev` and linear equations are solved by `gaussj` in Press, *et al.* (1992).

To find the estimate $\hat{\boldsymbol{\theta}}$, we use a modified version of the Newton-Raphson method. Let $\boldsymbol{S}(\boldsymbol{\theta})$ denote the J-dimensional vector of elements

$$\boldsymbol{S}(\boldsymbol{\theta})_\lambda = \partial l(\boldsymbol{\theta})/\partial\theta_\lambda = \left(\overline{\psi}_\lambda^\circ - \sum_i \psi_\lambda^\circ(t_i) f(t_i; \boldsymbol{\theta})\right)$$

and $\boldsymbol{H}(\boldsymbol{\theta})$, the $J \times J$ Hessian matrix of $-l(\boldsymbol{\theta})$ whose (λ, λ')-th element is given by

$$\partial^2 C(\boldsymbol{\theta})/\partial\theta_\lambda \partial\theta_{\lambda'} = \sum_{i=1}^{n} \Big[f(t_i; \boldsymbol{\theta}) \psi_\lambda^\circ(t_i) \psi_{\lambda'}^\circ(t_i) \Big].$$

Our method of computing $\hat{\boldsymbol{\theta}}$ is to start with an initial guess $\boldsymbol{\theta}^0$ and iteratively determine $\boldsymbol{\theta}^{m+1}$ according to the formula

$$\boldsymbol{\theta}^{m+1} = \boldsymbol{\theta}^m + \boldsymbol{H}^{-1}(\boldsymbol{\theta}^m)\boldsymbol{S}(\boldsymbol{\theta}^m).$$

Now we employ the step-halving, in which $\boldsymbol{\theta}^{m+1}$ is determined from $\boldsymbol{\theta}^m$ according to the formula

$$\boldsymbol{\theta}^{m+1} = \boldsymbol{\theta}^m + 2^{-r}\boldsymbol{H}^{-1}(\boldsymbol{\theta}^m)\boldsymbol{S}(\boldsymbol{\theta}^m),$$

where r is the smallest nonnegative integer such that

$$l\Big(\boldsymbol{\theta}^m + 2^{-r}\boldsymbol{H}^{-1}(\boldsymbol{\theta}^m)\boldsymbol{S}(\boldsymbol{\theta}^m)\Big) > l(\boldsymbol{\theta}^m).$$

We stop the iteration when $l(\boldsymbol{\theta}^{m+1}) - l(\boldsymbol{\theta}^m) < 10^{-6}$.

During the Newton-Raphson iteration, we approximate several integrals by the simple rectangular rule. For example, the (λ, λ')-th element of the Hessian matrix can be approximated by $\sum_{i=1}^{n} \psi_\lambda^\circ(t_i)\psi_{\lambda'}^\circ(t_i)f(t_i;\boldsymbol{\theta})$. The property that the wavelets are of compact support can be used to reduce the computing time (for large levels, the number of t_i for which $\psi_\lambda^\circ\psi_{\lambda'}^\circ$ is not zero is quite small).

3.2. Selection of wavelets

Wavelet shrinkage techniques (Donoho *et al.*, 1995) have emerged recently as powerful methods for the nonparametric estimation objects which may be characterized as 'spatially variable'. To select a set of wavelet basis functions, we use the wavelet shrinkage technique following the level-dependent wavelet thresholding idea (Kolaczyk, 1997). Suppose that the level j of wavelets ranges from j_0 to j_1. Let $\hat{\sigma}$ be the Mean Absolute Deviation of the wavelet coefficients at the finest level j_1, divided by 0.6745 (Donoho and Johnstone, 1994). The set of wavelets at level j for an IPE consists of

$$\Big\{\psi_\lambda^\circ : |\overline{\psi}_\lambda^\circ| > \sqrt{\log_2(n_j)}\hat{\sigma}\Big\},$$

where n_j is the number of wavelets at level j. Here we do not apply this hard thresholding rule to the coefficients of the scaling function. The estimate presented in Figure 1.2 is obtained by this procedure: (i) choose the cut-off levels such that $j_0 = 4 < j_1$; (ii) apply the wavelet thresholding for levels $j_0 \le j \le j_1$ and run the Newton-Raphson algorithm to find the IPE. Here $j_0 = 4$ is chosen such that the wrap-arounded wavelets do not intersect on the interval I and $j_1 = \log_2(n/2)$ such that $2^{j_1} = n/2$.

3.3. Artificial intensity function

To assess the performance of IPE for the gamma-ray burst phenomenon, we carried out a simulation study. Burst phenomena have been modeled successfully by astrophysicists using fast rise-exponential (FRED) models of the form

$$f(t) = f_0 + \sum_{i=1}^{\kappa} f_i(t), \tag{3.1}$$

where f_0 models the relatively constant background level of gamma-ray photon arrivals, κ is the number of 'peaks' in the burst and

$$f_i(t) = \begin{cases} A_i \exp\left(-|t-u_i|/\sigma_{r,i}^{\nu_i}\right), & \text{if } t \leq u_i \\ A_i \exp\left(-|t-u_i|/\sigma_{d,i}^{\nu_i}\right), & \text{if } t > u_i. \end{cases}$$

Here u_i is the location of the i–th peak, $\sigma_{r,i}$ and $\sigma_{d,i}$ control the 'rise' and 'decay' of the peak, and ν_i determines the peakedness.

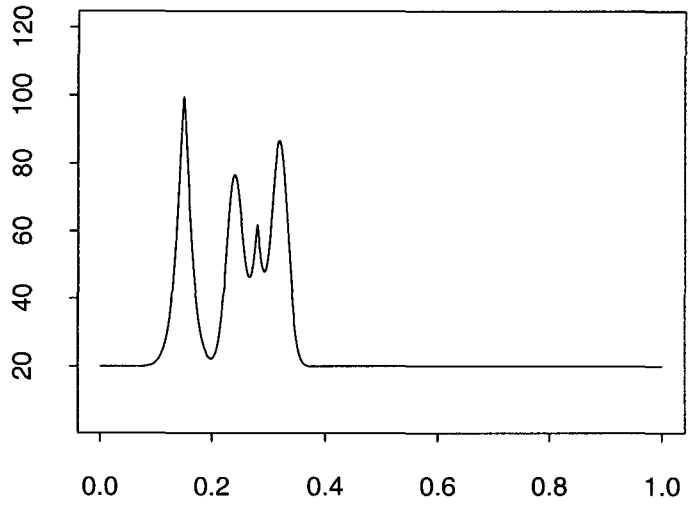

FIGURE 3.1 *Artificial intensity function. This function is used to model the burst phenomena using fast rise-exponential models.*

Figure 3.1 shows an intensity function constructed to mimic the structure which might model the gamma-ray burst signal in Figure 1.1. The three main peaks and a small peak between the second and third peaks of the burst signal have been modeled with FRED functions, with a constant intensity of 20 assigned to the background. Using this intensity function, Poisson counts are generated in each of $n = 1024$ equally spaced bins, where numerical integration is used to compute approximate value of the mean of Poisson count on each bin. Figure 3.2 shows a Poisson signal generated this way. IPE of the intensity function is computed from the data using the same method which provides us the fit in Figure 1.2 for the GRB signal.

Figure 3.2 shows a simulated signal that looks similar to the GRB signal presented in Figure 1.1. We can note that there is a small spike between the second

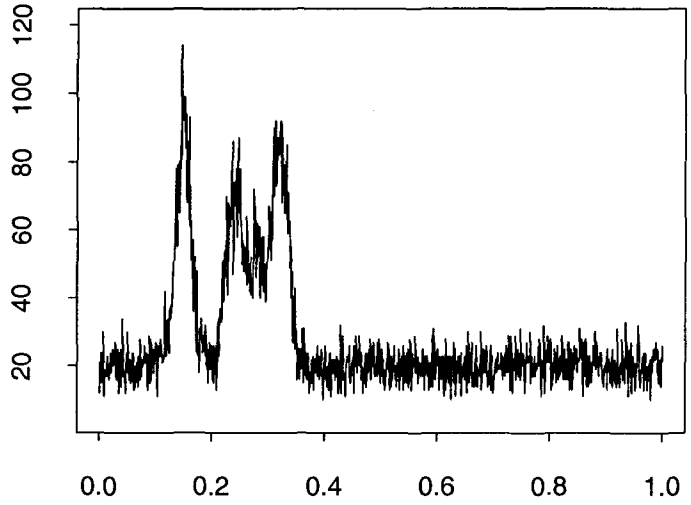

FIGURE 3.2 *Simulated data imitating gamma-ray burst 551*

and third peaks, which may confirm the performance of IPE for the gamma-ray burst signal around that point.

Figure 3.3 shows the IPE of the simulated signal in Figure 3.2. The number of parameters selected by the level-dependent wavelet thresholding is 47. The estimate is visually smooth, capturing the three main peaks in the first third of the signal and the constant background throughout the remaining two thirds. The small spike between the second and the third peaks is rather under-estimated, which might imply that the small spike of the IPE of the gamma-ray burst 551 is also under-estimated. However, the location of such small spike is identified quite accurately. Considering this simulation study, we may conclude that the IPE for the gamma-ray burst 551 signal is reliable.

3.4. Bin size

Figure 3.4 displays the effect of the bin size n on the performance of IPEs for $n = 1024, 512, 256, 128$. The bands are pointwise 5th and 95th percentiles of the 100 runs of IPEs.

Let $F_n(i/n)$ denote the Poisson mean of Y_i for $i = 1, \ldots, n$. For the simulation, $F_n(i/n)$ are computed by the following procedure.

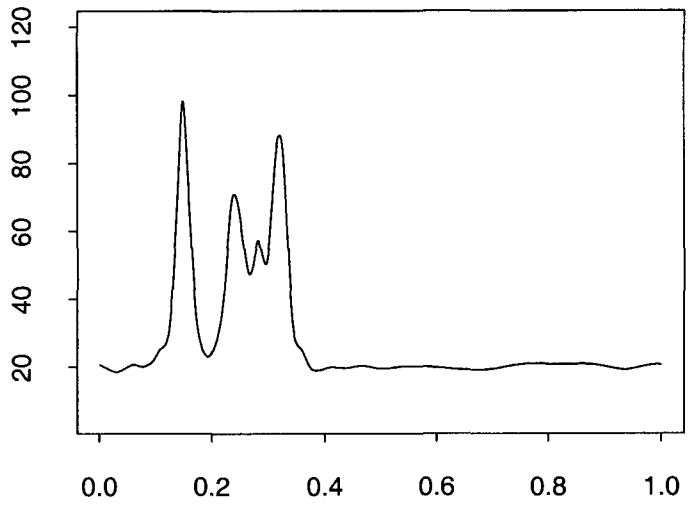

FIGURE 3.3 *Estimate of the intensity function presented in Figure 3.1*

(1) Compute $F_{1024}(i/1024), i = 1, \ldots, 1024$ from the approximate formula

$$F_{1024}(i/1024) \approx f(i/1024),$$

where f is the artificial intensity function (3.1).

(2) For $n = 512, 256, 128$, let $K_n = 1024/n$ and compute

$$F_n(i/n) = \frac{1}{K_n} \sum_{k=1}^{K_n} F_{1024}\left(\frac{(i-1)n + k}{1024}\right), \quad i = 1, \cdots, n.$$

According to Figure 3.4, the small spike between the second and third peaks disappears and the width of the band thicker as n decreases.

To summarize the behavior of IPEs for several n numerically, we consider two criteria defined as follows:

1. Root mean square error (RMSE). If $\hat{f}(t_i)$ is the estimated function value at t_i and n is the bin size, then

$$\text{MSE} = \frac{1}{n} \sum_i \{f(t_i) - \hat{f}(t_i)\}^2.$$

The MSE is computed for each run and averaged over the 100 runs. Then its square root is taken to give RMSE.

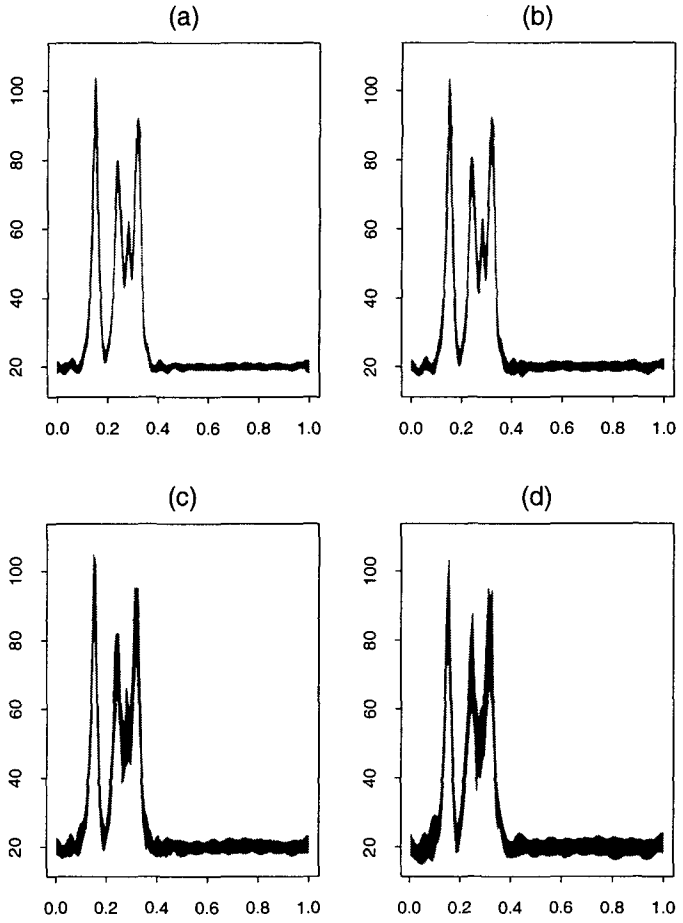

FIGURE 3.4 *The pointwise bands: (a)* $n = 1024$, *(b)* $n = 512$, *(c)* $n = 256$, *(d)* $n = 128$

2. Maximum deviation (MXDV). This is the average over the 100 runs of

$$\max_i |f(t_i) - \hat{f}(t_i)|.$$

Table 3.1 summarizes the simulation results for various sample sizes. Note that both RMSE and MXDV increase as n decreases as one can guess from Figure 3.4.

TABLE 3.1 *RMSE and MXDV for several* $n = 128, 256, 512, 1024$

n	*RMSE*	*MXDV*
1024	1.633423706	10.990727804
512	2.306339268	14.445330352
256	3.577758078	18.089362107
128	5.015181126	21.219912997

4. Concluding Remarks

The research for this article was motivated by the context of a problem involving gamma-ray burst data in astronomy. The simulations showed that the IPEs perform comparably with the TIPSH of Kolacyzk (1997). Though we use wavelets smoother than the Haar wavelet, the performance of IPEs looks reasonable.

There are some problems which are worth of investigating in the future. It is of interest to study the development of appropriate thresholds in the case where wavelets smoother than the Haar wavelet are applied to Poisson data. To give a theoretical justification for such an adaptive method, it would be interesting to study asymptotic optimality when the logarithm of the intensity function belongs to a Besov spaces. Work extending the IPEs algorithm to problems of two-dimensional image processing would be an important subject. The usual tensor–product of wavelets can be used in this context for practical implementation and theoretical study.

Acknowledgements

Both authors wish to acknowledge the financial support of the Korea Research Foundation (1998-015-D00047) made in the program year of 1998. The research

of the first author was also supported by the Brain Korea 21 project. The authors thank Eric D. Kolaczyk for helpful comments and providing the Burst data, and Cheol-Woo Park for help in programming.

REFERENCES

Babu, G. J. and Feigelson, E. (1996). "Spatial point processes in Astronomy", *Journal of Statistical Planning and Inference*, **50**, 311–326.

Barron, A. R. and Sheu, C.-H. (1996). "Approximations of density functions by sequences of exponential families", *The Annals of Statistics*, **19**, 1347–1369.

Cohen, A., Daubechies, I. and Vial, P. (1993). "Wavelets on the interval and fast wavelet transforms", *Applied Computational and Harmonic Analysis*, **1**, 54–81.

Cowling, A., Hall, P. and Phillips, J. P. (1996). "Bootstrap confidence regions for the intensity of a Poisson point process", *Journal of the American Statistical Association*, **91**, 1516–1524.

Daubechies, I. (1993). "Orthonormal bases of compactly supported wavelets II: variations on a theme", *SIAM Journal of Mathematical Analysis*, **24**, 499–519.

Diggle, P. (1985). "A kernel method for smoothing point process data", *Applied Statististics*, **34**, 138–147.

Diggle, P. and Marron, J. S. (1988). "Equivalence of smoothing parameter selectors in density and intensity estimation", *Journal of the American Statistical Association*, **83**, 793–800.

Donoho, D. L. and Coifman, R. R. (1995). "Translation-invariant de-noising", In *Wavelets and Statistics* (A. Antoniadis and G. Oppenheim, eds.), Springer-Verlag, New York.

Donoho, D. L. and Johnstone, I. M. (1994). "Ideal spatial adaptation via wavelet shrinkage", *Biometrika*, **81**, 425–455.

Donoho, D. L. and Johnstone, I. M. (1998). "Minimax estimation via wavelet shrinkage", *The Annals of Statistics*, **26**, 879–921.

Donoho, D. L., Johnstone, I. M., Kerkyacharian, G. and Picard, D. (1995). "Wavelet shrinkage: asymptopia? (with discussion)", *Journal of the Royal Statististical Society*, **B57**, 301–369.

Ellis, S. P. (1986). "A limit theorem for spatial point processes", *Advanced Applied Probability*, **18**, 649–659.

Kolaczyk, E. D. (1997). "Non-parametric estimation of gamma-ray burst intensities using Haar wavelets", *Astrophysical Journal*, **483**, 340–349.

Kolaczyk, E. D. (1999). "Wavelet shrinkage estimation of certain Poisson intensity signals using corrected thresholds", *Statistica Sinica*, **9**, 119–135.

Meegan, C. A., Fishman, G. J., Wilson, R. B., Paciesas, W. S., Pendleton, G. N., Horack, J. M., Brock, M. N. and Kouveliotou, G. (1992). "The spatial distribution of gamma ray bursts observed by BATSE", *Nature*, **355**, 143–145.

Meyer, Y. (1992). *Wavelets and Operators.* Cambridge University Press, Oxford.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1992). *Numerical Recipes in C, the Art of Scientific Computing, Vol. 2*, Cambridge University Press, Oxford.