

Higher Order Cumulants and Inference for a Class of Filtered Poisson Processes¹

R.J. Kulperger

Department of Statistical and Actuarial Sciences
The University of Western Ontario, London, Canada

Abstract: Spectral methods are useful in the analysis of time series and point process data in \mathbf{Z}^d or \mathbf{R}^d . Parameter estimates based on these often have a limiting Gaussian distribution, whose limiting variance depends upon integrals of the second, third and fourth order spectral densities. The effective evaluation of these spectral density integrals is needed. In image analysis ($d = 2$), these integrals are very time consuming to evaluate. This paper considers a particular class of integrated or filtered Poisson processes. Using higher order cumulants, one can identify parameters up to a certain order. In some parametric cases these identify all the parameters of the process. In such cases it would then be possible to construct and hence simulate a filtered Poisson process for the estimator functional has the same asymptotics. Simulation or bootstrapping this process is a more efficient way of estimating or approximating the distribution of the parameter estimate. This paper demonstrates the validity for such a method. Specific examples of scanning fluorescence correlation spectroscopy (S-FCS) and image correlation spectroscopy (ICS) are discussed and used as motivating examples. In these cases one can obtain estimates of an object that previously could only be inferred indirectly.

1 Introduction

Filtered Poisson processes have many applications; see Daley and Vere-Jones (1988) and Snyder and Miller (1991) for definition and examples. Usually one does not observe complete information, that is they are incomplete data problems. The EM algorithm can be applied with success in some problems (Snyder and Miller, 1991), but it is not usually a tractable method. Another family of filtered Poisson processes has been studied in Benn and Kulperger (1996, 1997) in a 1D and a 2D (random field) problem. The random field problem example is called image correlation spectroscopy (ICS), and will play the role of a prototype example. The model for the ICS data is an integrated marked Poisson process, and is described in Section 2. Any point event viewed through a small aperture will give rise to such an integrated point process with an appropriate smoothing kernel (see Freeman, 1990), giving rise to a fairly general class.

The method used in Benn and Kulperger (1997) is based on second order statistics, specifically periodograms for a random field, while Benn and Kulperger (1996) used a

¹This work has been supported by the Natural Sciences and Engineering Research Council of Canada. This paper is part of a presentation made at the 1997 Environmetrics Meeting in Innsbruck. The author wishes to thank Professor Ian McLeod for the invitation to the meeting and to thank the referee for some helpful comments.

real time domain analysis in a 1 dimensional problem. These gave rise to the information sought by the biochemist (Professor Nils Petersen, Department of Chemistry, University of Western Ontario) who posed the question to us. This information is then used to further infer some properties about the protein aggregate distributions that are of primary interest.

It is possible to obtain information about the marking distribution, and hence in the ICS problem information on the protein aggregates. This can be done by making some further parametric assumptions on the marking or protein aggregate distribution and by the use of higher cumulants or equivalently higher moments, using essentially a method of moments, and then identifying all the parameters of the filtered Poisson process. The use of higher order cumulant methods has also been suggested in the chemistry literature; see Palmer and Thompson (1987) and several references therein. The term used there is higher order autocorrelation. As a by product one can estimate the intensity of the Poisson field.

This paper discusses how higher order moments or cumulants can be used in this class of filtered Poisson processes. There are some partial goodness of fit methods available to check these extra assumptions. The estimation methods are described in Section 3. A discussion of the partial goodness of fit methods is given in Section 4.

Aside from the additional information that be extracted from the experimental data, there is another possible use. As seen in Benn and Kulperger (1997), estimation of the variances of the estimators is not always easy. The ICS data is observed as a T by T lattice, where $T = 512$. Fourth order spectral integrals appear in the limit variance. Using an estimator based on fourth order periodograms (see Brillinger, 1981) yields a consistent estimate (Keenan, 1987, and Taniguchi, 1982), but one of order T^6 calculations. These computations are estimated to take about 364 days on a 2K MASPAC computer (Benn and Kulperger, 1997). A MASPAC is a Single Instruction Multiple Data (SIMD) massively parallel machine with 2048 subprocessors. Such a method is of course not numerically feasible. Thus an alternative was used in Benn and Kulperger (1997). Another important by product of the method discussed in this paper is that all parameters for the filtered Poisson process are estimable. This means that one can study the sampling distribution of a parameter estimate by a parametric bootstrap. This can completely bypass the estimation of the limiting variance. It may not be as efficient as the method for the essentially second order statistics used in Benn and Kulperger (1997), but of course a complete structure of the process can be estimated. The method is numerically feasible. This paper only shows this is possible in theory, and does not carry out the numerical implementation.

2 A Filtered Poisson Process and The Observable Data

The model discussed in Benn and Kulperger (1996, 1997) is a filtered marked Poisson process to model scanning fluorescence correlation spectroscopy (S-FCS) and image scanning spectroscopy (ICS). S-FCS is an experimentally slow method yielding a 1D scan, while ICS is an experimentally faster 2D image scan. S-FCS is a higher resolution than ICS. From a mathematical point of view, they are similar processes, both being an integrated marked Poisson process. Some of the cell biology uses and background can be found in Magde and Elson (1974), Petersen (1986, 1993), St-Pierre and Petersen (1990,

1992); see Benn (1996) for a relevant summary to statisticians. S-FCS and ICS are experimental techniques to obtain information about cell protein aggregate sizes. ICS is a state of the art method.

The class of processes is now defined. The following ingredients are used in the model.

1. N is a stationary, mean λ Poisson process in R^2 .
2. $Y_{(x,y)}$ is an iid array of marks, with finite moments, and independent of N . The marked Poisson process has a mark $Y_{(x,y)}$ at each point (x, y) in the Poisson field. Let $\mu_{Y,k} = E(Y^k)$ be the k -th moment of the marking distribution.
3. $f(x, y)$ is a weight function or kernel in R^2 . In the S-FCS and ICS applications it represents the image light intensity kernel of a confocal scanning microscope lens (CSML), and is a Gaussian shaped kernel with parameter w

$$f(x, y) = \exp \left\{ -\frac{1}{w^2} (x^2 + y^2) \right\} . \quad (1)$$

4. The scanning microscope is shifted by an amount δ on the cell surface to produce observations. The observations are indexed by k in dimension 1 (S-FCS experiment) and by $k = (k_1, k_2)$ in dimension 2 (ICS experiment). Let S_k be a disc centred at $(k\delta, 0)$ or $(k_1\delta, k_2\delta)$ in dimensions 1 and 2 respectively. When the CSML is positioned at k , the illuminated region is S_k and the kernel at position k is

$$\begin{aligned} f_k(x, y) &= f(x - k\delta, y), \text{ dimension} = 1 \\ f_k(x, y) &= f(x - k_1\delta, y - k_2\delta), \text{ dimension} = 2 \end{aligned}$$

5. The observable data is of the form

$$X_k = \int_{R^2} f_k(x, y) Y_{(x,y)} N(dx, dy) + \epsilon_k \equiv Z_k + \epsilon_k \quad (2)$$

where k is interpreted in either \mathbf{Z} or \mathbf{Z}^2 , and ϵ are noise terms. If ϵ has zero variance, then $X_k = Z_k$ is a noiseless filtered Poisson process. Let $\text{Cov}_X(k)$ and $\text{Cov}_Z(k)$ be the lag k covariances of the X and Z processes.

6. The noise process ϵ is independent of the noiseless filtered Poisson process, and is an iid mean zero Gaussian sequence (array). Let $\text{Var}(\epsilon) = \beta \text{Cov}_Z(0)$.

In terms of the S-FCS and ICS processes, the Poisson points (x_i, y_i) represents the i -th monomer and the mark $Y_{(x_i, y_i)}$ represents the protein aggregate size of the monomer. The protein aggregate process is not observed directly. Instead it is observed through a confocal scanning device. Astronomical data is often observed through a small aperture, so this type of model may also be useful in that setting. The kernel is generally described by (1); see Freeman (1990).

For the S-FCS process, a cell is moved under an electron scanning microscope. The microscope observes the process centred at some position on the cell. The observation is

then a weighted sum of the protein process centred at this position and under the effective region of illumination of the microscope; that is the integrated protein process with kernel f . The kernel is of type (1). Observations are made by then moving the cell an amount δ under the microscope, yielding a one dimensional scan of the cell. The observed data is of the form X_k , $k = 1, \dots, T$. This experimental technique is laborious (St-Pierre and Petersen, 1990). There is evidence that the noise process ϵ is Gaussian and independent of Z_k , justifying the model.

For the ICS process, a 2 dimensional scan similar in description to the S-FCS is made. The scanning device in this case is confocal scanning microscope lens (CSML). The cell is moved under the CSML, shifting an amount δ in each of two dimensions. At a given position the observation is a weighted sum of the protein aggregates under the effective area of illumination of the lens. The weight function or kernel is again of the type (1), with w describing the lens properties. This data set is then a random field indexed by Z^2 . The ICS method is much easier to use experimentally than S-FCS, but is thought to be not quite as precise (Petersen, 1993). However it has T^2 observations as opposed to T in S-FCS.

The observable data is given by (2), and hence the underlying marked Poisson process is not directly observable. Thus the statistical inference problem is an incomplete data problem. The observable data X_k forms an observation sequence or array from a stationary process in Z^d .

Moments and cumulants of the stationary process X of (2) can be calculated. They are laborious to obtain. The mean and lag covariances functions, or equivalently the first and second cumulants, for Z and X are given here. The third and fourth cumulants are given in the appendix. Parameters of the underlying filtered Poisson process can only be estimated through the process X . Also notice that (2) can easily be simulated since the integral with respect to the Poisson process N is actually a sum, and the support of the kernel function f_k can be truncated to a finite disc, say S_k or radius $3w/\sqrt{2}$.

$$\begin{aligned} \mathbf{E}(Z) &= \mathbf{E}(X) = \mathbf{E} \left(\int_{R^2} f_k(x) Y_x N(dx) \right) \\ &= \lambda \mu_{Y,1} \int_{R^2} f(x) dx = \lambda \mu_{Y,1} \pi w^2. \end{aligned} \quad (3)$$

The lag k covariance for Z is

$$\begin{aligned} \text{Cov}_Z(k) &= \lambda \mu_{Y,2} \int_{R^d} f_0(x) f_k(x) dx \\ &= \lambda \mu_{Y,2} w^2 \frac{\pi}{2} \exp\left\{-\frac{\delta^2 |k|^2}{2w^2}\right\} \end{aligned} \quad (4)$$

where the index $k \in Z$ for S-FCS, and hence $|k|^2 = k^2$, and for ICS, $k \in Z^2$ and hence $|k|^2 = k_1^2 + k_2^2$. Alternatively one can use the index $(k, 0)$ for S-FCS and $k = (k_1, k_2)$ for ICS. Then

$$\text{Cov}_X(k) = \begin{cases} \text{Cov}_Z(0) + \gamma^2 & \text{if } k = 0 \\ \text{Cov}_Z(k) & \text{if } k \neq 0 \end{cases} \quad (5)$$

where $\gamma^2 = \text{Var}(\epsilon)$.

3 Parameter Identification and Estimation

As indicated in Section 2, moments of the observable process X can be calculated in terms of the underlying process, specifically in terms of the moments of Z , the noise free filtered Poisson process, plus the Gaussian noise terms. Since X is a stationary process, sample moments or sample spectra can be used to estimate the corresponding quantities for the process X .

Parameters for the underlying process can be identified only in terms of the moments of the observable process can then be estimated. This may only include functions of the parameters, and not all parameters. This is what generally happens with an incomplete data problem. This section examines the used of essentially a method of moments for the filtered Poisson processes of section 2. In special cases one can identify and hence estimate all the parameters of the process. This can then be used to do a parametric bootstrap of the filtered Poisson process.

The parameters of the process X are

1. λ the Poisson parameter
2. w the kernel parameter
3. the parameters of the marking distribution. Let $\mu_{Y,k} = E(Y^k)$ be the k -th moment of the marking distribution.
4. β the noise parameter, where $\text{Var}(\epsilon) = \beta \text{Cov}_Z(0)$.

Using moments or cumulants up to order $\ell \geq 2$, one can identify $\theta_0 = w$, and $\theta_k = \lambda \mu_{Y,k}$, $k \leq \ell$, and β , the noise variance term. This is due to the fact that given X moments up to order k , the $(k+1)$ -th moment or cumulant involves the same parameters, plus $\lambda \mu_{Y,k+1}$. Using the data wisely, one can obtain \sqrt{n} -consistent estimators of these parameters, where n is the sample size.

For example if the marking random variables are $Y \sim \text{Poisson}$, ν , then $\mu_{Y,1} = \nu$, and $\mu_{Y,2} = \nu(\nu+1)$. Then

$$\frac{\theta_2}{\theta_1} = \frac{\nu(\nu+1)}{\nu} = \nu + 1.$$

Thus ν and hence λ can be obtained. From these and w , $\text{Cov}_Z(0)$ can be obtained. Finally

$$\text{Cov}_X(0) = \text{Cov}_Z(0) (1 + \beta)$$

and hence β can be obtained.

An example requiring third order moments is where the marking distribution is $Y \sim \text{Gamma}(\alpha, \gamma)$. Then $\mu_{Y,1} = \alpha\gamma$ and $\mu_{Y,2} = \alpha(\alpha+1)\gamma^2$. and $\mu_{Y,3} = \alpha(\alpha+1)(\alpha+2)\gamma^3$. Then

$$\frac{\theta_2}{\theta_1} = \frac{\lambda \mu_{Y,2}}{\lambda \mu_{Y,1}} = \gamma(\alpha+1)$$

and

$$\frac{\theta_3}{\theta_2} = \frac{\lambda \mu_{Y,3}}{\lambda \mu_{Y,2}} = \gamma(\alpha+2).$$

Thus α and γ , and hence λ can be identified. Next the Gaussian noise variance can be estimated.

An example of using fourth order moments is the case where the marking distribution is a mixture of two Poissons. This case is interesting in the ICS and S-FCS settings, as there is some speculation that the protein aggregates are made up of two general sizes or types (Petersen, 1986, p. 810). Specifically the marking random variables are

$$Y = UV_1 + (1 - U)V_2$$

where $U \sim \text{Bernoulli}(\alpha)$, $V_1 \sim \text{Poisson}, \gamma_1$ and $V_2 \sim \text{Poisson}, \gamma_2$, and U, V_1, V_2 are independent. Notice that $U(1 - U) = 0$. Thus

$$E(Y^k) = \alpha E(V_1^k) + (1 - \alpha)E(V_2^k) .$$

For $V \sim \text{Poisson}, \gamma$,

$$E(V^k) = \begin{cases} \gamma & \text{if } k = 1 \\ \gamma + \gamma^2 & \text{if } k = 2 \\ \gamma + 3\gamma^2 + \gamma^3 & \text{if } k = 3 \\ \gamma + 7\gamma^2 + 6\gamma^3 + \gamma^4 & \text{if } k = 4 . \end{cases}$$

This example is of a mixture type distribution. These are generally difficult to estimate, and it is no different here.

In these cases one can use a higher order moment as a model check of the assumed distribution of Y . The sampling distribution of the appropriate partial goodness of fit statistic can be approximated by a parametric Monte Carlo simulation. The use of third sample moments for the Poisson marking case, and fourth sample moments for the Gamma marking case would be appropriate.

Aside from these, the parametric Monte Carlo method allows one to estimate sampling distributions of the ratio parameter estimator (6) of interest in Petersen (1986, 1993).

$$\hat{R} = \frac{\text{Cov}_Z(0; \hat{\theta})}{\bar{X}^2} . \quad (6)$$

3.1 Estimation

The illustration will be for the S-FCS and ICS processes. The second order estimating equation is taken to be a least squares equation, minimizing the sum of squares between the periodograms and spectral densities. If one can estimate all the parameters of the underlying process, one can do a parametric bootstrap, that is simulate at the estimated parameters. This will then yield an estimate of the standardized or centred estimator \hat{R} from (6). If the parameterization is reasonable, this will give an alternative method to the least squares methods in Benn and Kulperger (1996, 1997). It will also yield direct information on λ , the Poisson parameter. This is one of the interesting objects in the chemistry problem.

Two natural methods can be used to obtain the second order estimates, that is estimators of θ_k , $k = 0, 1, 2$. There are many variations of these. The methods are unweighted

least squares of either the empirical sample covariances, or of the periodograms. To exploit computational properties of fast Fourier transforms, the periodogram method is perhaps better for the ICS problem, that is the random field filtered process. Given estimates of $w = \theta_0$ and θ_2 , one can then estimate θ_1 from \bar{X} .

The least squares function is given by

$$H_T(\theta_0, \theta_2) = \sum_{k=1}^L (\hat{c}(k) - \text{Cov}_X(k; \theta_0, \theta_2))^2 \quad (7)$$

where $\hat{c}(k) = \frac{1}{n+1} \sum_{i=0}^{n-k} (X_i - \bar{X})(X_{i+k} - \bar{X})$ is the empirical covariance, and Cov_X is given by (5). With the parameters θ_0, θ_2 explicitly shown this becomes

$$\text{Cov}_X(k; \theta_0, \theta_2) = \text{Cov}(Z_0, Z_k) = (\lambda \mu_{Y,2}) (w^2 \frac{\pi}{2}) \equiv \theta_2 \theta_0^2 \frac{\pi}{2}.$$

The parameter L refers to the number of lags used in the least squares equation. The main goal in the chemistry problem is the determination of $R \equiv c(0; \theta_1, \theta_2) / \mu^2$ (St-Pierre and Petersen, 1990). This can be estimated by (6).

If one uses a periodogram least squares method, the least squares function is then

$$L_T(\theta) = \sum_s \left(I^{(T)}(\lambda_s) - \mathbb{E}(I^{(T)}(\lambda_s); \theta_0, \theta_2) \right)^2, \quad (8)$$

where the periodogram least squares is over a set of frequencies $\lambda_s = \frac{2\pi s}{T}$ for a 1D transect, such as S-FCS, and is $\lambda_s = \frac{2\pi s}{T^2}$, $s = (s_1, s_2)$ for a 2D image such as ICS. The sum is over a subset of $s_1, s_2 \in (0, T)^2$, thus excluding frequency $\mathbf{0}$. The periodograms are centred at $\mathbb{E}(I^{(T)}(\lambda_s); \theta_0, \theta_2)$. In a 1D transect, such as S-FCS or ordinary time series, this can be replaced by the second order spectral density. In the case of a 2D transect, such as ICS, or a random field, this centering matters. See Benn and Kulperger (1998) for further discussions about this. The short reason is that $\mathbb{E}(I^{(T)}(\lambda)) = f(\lambda) + O(T^{-1})$, where f is the appropriate spectral density. The normalizing rate of convergence is $\sqrt{T^D}$ in a D dimensional problem. The normalizing rate times the error term from the expectation of the periodogram is $O(T^{1-D/2})$, which tends to zero if $D = 1$, but is $O(1)$ if $D = 2$.

One of the major difficulties in applying these results to experimental data is the evaluation of the limit variances. Theorem 1 is a typical result for the distribution of the estimators in connection with (7), involving up to fourth moments of the underlying process. A typical limit result for the periodogram case (8) involves integrals of second, third and fourth order spectral densities. These are computationally intensive in a random field, involving $O(T^6)$ computations. Some interesting methods of getting around this difficulty are discussed in Benn (1996) and Benn and Kulperger (1997). When using higher order moments, the limit variances become even more complicated, and hence more difficult to estimate. A nonparametric bootstrap or simulation method is perhaps the best that one can do computationally.

Theorem 1

$$\sqrt{n} \begin{bmatrix} \bar{X}_n - \mu \\ \hat{c}(1) - c(1; \theta_0) \\ \vdots \\ \hat{c}(L) - c(L; \theta_0) \end{bmatrix} \Rightarrow N_{L+1}(\mathbf{0}, \Sigma)$$

where

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix},$$

and the component matrices are given below.

$$\Sigma_{11} = \gamma^2 + \sum_{j=-\infty}^{\infty} c(j; \boldsymbol{\theta}_0),$$

$$\Sigma_{12} = \Sigma'_{21} = \left[\sum_{j=-\infty}^{\infty} \text{cov}(X_0, (X_j - \mu)(X_{j+k} - \mu)) \right]_{1 \times L}$$

and

$$\Sigma_{22} = \left[\sum_{j=-\infty}^{\infty} \text{cov}((X_i - \mu)(X_{i+k} - \mu), (X_0 - \mu)(X_l - \mu)) \right]_{L \times L}.$$

In order to estimate for example the next order θ , one could use a weighted moment least squares

$$H_{T,3}(\theta_0, \theta_2, \theta_3) = \alpha \sum_{k=1}^L (\hat{c}(k) - \text{Cov}_X(k; \theta_0, \theta_2))^2 + (1 - \alpha) \sum_{k_1, k_2} (\hat{\mu}_3(k_1, k_2) - \mu_3(k_1, k_2))^2 \quad (9)$$

where $0 < \alpha < 1$ is a weight,

$$\mu_3 = \mathbf{E}((X_0 - \mu_X)(X_{k_1} - \mu_X)(X_{k_2} - \mu_X))$$

and $\hat{\mu}_3$ is the sample centred third moment. An analogous estimation method can be written for the periodogram case.

$$L_{T,3}(\theta_0, \theta_2, \theta_3) = \alpha \sum_s \left(I^{(T)}(\lambda_s) - \mathbf{E}(I^{(T)}(\lambda_s); \theta_0, \theta_2) \right)^2 + (1 - \alpha) \sum_{s_1, s_2} \left(I_3^{(T)}(\lambda_{s_1}, \lambda_{s_2}) - \mathbf{E}(I_3^{(T)}(\lambda_{s_1}, \lambda_{s_2}) | \theta) \right)^2. \quad (10)$$

The second term in the sum is over $s_1, s_2 \in \mathbf{Z}^2$, where $1 \leq s_j \leq t_1 T$, where t_1 is a design parameter giving the number of terms in the sum. In practice one will take $t_1 > 0$, but not too large for computational reasons.

One can obtain a central limit theorem for the resulting estimators from (9). The limit variances will involve up to 6–th moments in this case, and 6–th order spectral densities in the periodogram estimation method (10). Estimating these limiting variances becomes even less tractable than in the ICS where these can take hundreds of days on a massively parallel computer (Benn and Kulperger, 1997).

With a parametric bootstrap, these of course are as amenable as a simulation study of the relevant estimators. That is one, in theory, simply estimates the parameters, and then simulate at these values. Even with the ICS problem, with a sample of size T^2 , with $T = 512$, this can be done in 5 to 10 minutes per simulation run and estimate. Thus within a day, a simulation sample of size a few hundred can be produced. This should be compared with a direct evaluation in the order of a year for the second order parameters. With the third order parameters, this increases by a factor of another T^2 , putting the calculation into many years, even on a state of the art data parallel computer.

4 Discussion and Summary

This paper discusses the use of higher moments to extract additional information about the marking distribution for a class of integrated marked Poisson processes. The processes includes a large class of processes that can model point events observed through a small viewing aperture, such as a scanning electron or laser microscope or a telescope.

Section 3.1 discusses a least squares method of parameter estimation. It is based on having the correct model. The simplest idea is to use the next order moment. If one uses second moments, then third moments can be used as a goodness of fit check. That is suppose that the proposed process is identified by second order moments. Then given estimates of the parameters, one can simulate the second order identified process, and simulate in particular functions of third order moments. If the underlying mechanism for the process observed is different from the postulated second order model, the third order statistic may have different sampling properties than that under the second order model. One can construct various such statistics thereby generating partial goodness of fit statistics, and hence have some power to detect model departures.

Only some initial numerical studies of the feasibility of this type of proposed model and partial goodness of fit have been considered. Additional work on this question needs to be undertaken. However it appears that one needs some type of estimation method other than the least squares mixture in the Poisson mixtures example. It is an example of an unstable least squares, if one is estimating both the mixture parameter and the two protein aggregate size mean parameters. One would need some other information in a mixtures case, or a more efficient estimating equation. However in theory the method should work. In another case with some limited simulation results, a protein aggregate distribution that is identified by second order statistics does work reasonably well.

There are some design parameters in the estimation method, such as the mixture parameter α on the sum of least squares terms. Is there a good way of choosing this? Is there something better than least squares as a basis for the estimating equation? One needs to also consider computational tractability in these methods. As discussed in Benn and Kulperger (1997), the naive method of estimating a limiting variance takes in the order of 1 year to numerically evaluate, which is practically useless. While these ideas look promising, there are important implementation questions to be considered. However if these can be overcome, there is much further useful information to be extracted from spectroscopy experiments that has not been accessible at present. The question of using higher order cumulant information has been raised in the chemistry literature as early as Palmer and Thompson (1987).

5 Appendix

In this appendix the fourth order cumulant is calculated. It will be readily seen that in general if one calculates a $(k + 1)$ -st order cumulant, all parameters are used in a k -th order cumulant except the additional parameter $\mu E(Y^{k+1})$.

The underlying process is Z , an integrated marked Poisson process. The observable data is $X = Z + \epsilon$, where ϵ is an independent white noise process. A k -th cumulant will be denoted $\text{cum}_{U,k}$ for a process U . When there is no confusion, drop the process

subscript U . Since $X = Z + \epsilon$ is a sum of independent components, one obtains $\text{cum}_{X,k} = \text{cum}_{Z,k} + \text{cum}_{\epsilon,k}$. Since ϵ is Gaussian, $\text{cum}_{\epsilon,k} = 0$ for $k \geq 3$.

The third-order cumulant spectrum of X is

$$f_{X,3}(\lambda^{(1)}, \lambda^{(2)}) = \frac{1}{(2\pi)^4} \sum_{a=-\infty}^{\infty} \sum_{b=-\infty}^{\infty} \text{cum}_{X,3}(\mathbf{a}, \mathbf{b}) e^{-i(\langle \lambda^{(1)}, \mathbf{a} \rangle + \langle \lambda^{(2)}, \mathbf{b} \rangle)}$$

where

$$\text{cum}_{X,3}(\mathbf{a}, \mathbf{b}) = \text{cum}_3(X(\mathbf{0}), X(\mathbf{a}), X(\mathbf{b})) = \text{cum}_{Z,3}(\mathbf{a}, \mathbf{b}) = \text{cum}_3(Z(\mathbf{0}), Z(\mathbf{a}), Z(\mathbf{b}))$$

by stationarity. Moments of Z can be computed using properties of point processes (Benn, 1996). Using cumulant and moment relations, Brillinger (1981), yields

$$\begin{aligned} & \text{cum}_{Z,3}\{\mathbf{a}, \mathbf{b}, \mathbf{c}\} \\ &= (-1)^2 \cdot 2! \text{E}(Z(\mathbf{a}))\text{E}(Z(\mathbf{b}))\text{E}(Z(\mathbf{c})) \\ & \quad + (-1)^1 \cdot 1! \{ \text{E}(Z(\mathbf{a})Z(\mathbf{b}))\text{E}(Z(\mathbf{c})) + \text{E}(Z(\mathbf{a})Z(\mathbf{c}))\text{E}(Z(\mathbf{b})) \\ & \quad \quad + \text{E}(Z(\mathbf{b})Z(\mathbf{c}))\text{E}(Z(\mathbf{a})) \} \\ & \quad + \text{E}(Z(\mathbf{a})Z(\mathbf{b})Z(\mathbf{c})) \\ &= 2\mu_Z^3 - \mu_Z \{ \text{E}(Z(\mathbf{a})Z(\mathbf{b})) + \text{E}(Z(\mathbf{a})Z(\mathbf{c})) + \text{E}(Z(\mathbf{b})Z(\mathbf{c})) \} \\ & \quad + \text{E}(Z(\mathbf{a})Z(\mathbf{b})Z(\mathbf{c})) \\ &= -\mu_Z^3 - \mu_Z \{ \text{Cov}_Z(\mathbf{a} - \mathbf{b}) + \text{Cov}_Z(\mathbf{a} - \mathbf{c}) + \text{Cov}_Z(\mathbf{b} - \mathbf{c}) \} \\ & \quad + \text{E}(Z(\mathbf{a})Z(\mathbf{b})Z(\mathbf{c})) \\ &= \lambda\mu_{Y,3} \frac{\pi w^2}{3} \int_{R^2} f_{\mathbf{a}}(x) f_{\mathbf{b}}(x) f_{\mathbf{c}}(x) dx \\ &= \lambda\mu_{Y,3} \frac{\pi w^2}{3} \exp \left[-\frac{\delta^2}{w^2} (\langle \mathbf{a}, \mathbf{a} \rangle + \langle \mathbf{b}, \mathbf{b} \rangle + \langle \mathbf{c}, \mathbf{c} \rangle \right. \\ & \quad \left. - \frac{1}{3} \langle \mathbf{a} + \mathbf{b} + \mathbf{c}, \mathbf{a} + \mathbf{b} + \mathbf{c} \rangle \right]. \end{aligned}$$

This result uses the first two moments of Z calculated in Section 2, $\mu_Z = \text{E}(Z)$ and

$$\begin{aligned} & \text{E}(Z(\mathbf{a})Z(\mathbf{b})Z(\mathbf{c})) \\ &= \lambda\mu_{Y,3} \frac{\pi w^2}{3} \int_{R^2} f_{\mathbf{a}}(x) f_{\mathbf{b}}(x) f_{\mathbf{c}}(x) dx \\ & \quad \mu_Z \lambda\mu_{Y,2} \left\{ \int_{R^2} f_{\mathbf{a}}(x) f_{\mathbf{b}}(x) dx + \int_{R^2} f_{\mathbf{a}}(x) f_{\mathbf{c}}(x) dx + \int_{R^2} f_{\mathbf{b}}(x) f_{\mathbf{c}}(x) dx \right\} \\ & \quad + \mu_Z^3 \\ &= \frac{\pi \lambda \mu_{3Y} w^2}{3} \exp \left[-\frac{\delta^2}{w^2} (\langle \mathbf{a}, \mathbf{a} \rangle + \langle \mathbf{b}, \mathbf{b} \rangle + \langle \mathbf{c}, \mathbf{c} \rangle \right. \\ & \quad \left. - \frac{1}{3} \langle \mathbf{a} + \mathbf{b} + \mathbf{c}, \mathbf{a} + \mathbf{b}, \mathbf{c} \rangle \right] \\ & \quad + \mu_Z (\text{Cov}_Z(\mathbf{a} - \mathbf{c}) + \text{Cov}_Z(\mathbf{b} - \mathbf{c}) + \text{Cov}_Z(\mathbf{a} - \mathbf{b})) + \mu_Z^3. \end{aligned}$$

The corresponding expression for the fourth-order product moment of Z is required to obtain cum_Z .

$$\begin{aligned}
& \mathbf{E}(Z(\mathbf{a})Z(\mathbf{b})Z(\mathbf{c})Z(\mathbf{d})) \\
&= \lambda\mu_{Y,4} \int_{R^2} f_{\mathbf{a}}(x)f_{\mathbf{b}}(x)f_{\mathbf{c}}(x)f_{\mathbf{d}}(x)dx + \mu_Z\lambda\mu_{Y,3} \left\{ \int_{R^2} f_{\mathbf{a}}(x)f_{\mathbf{b}}(x)f_{\mathbf{c}}(x)dx \right. \\
&\quad \left. + \int_{R^2} f_{\mathbf{a}}(x)f_{\mathbf{b}}(x)f_{\mathbf{d}}(x)dx + \int_{R^2} f_{\mathbf{a}}(x)f_{\mathbf{c}}(x)f_{\mathbf{d}}(x)dx + \int_{R^2} f_{\mathbf{b}}(x)f_{\mathbf{c}}(x)f_{\mathbf{d}}(x)dx \right\} \\
&\quad + \left\{ (\lambda\mu_{Y,2})^2 + m\mu_Z^2\lambda\mu_{Y,2} \right\} \left\{ \int_{R^2} f_{\mathbf{a}}(x)f_{\mathbf{b}}(x)dx + \int_{R^2} f_{\mathbf{a}}(x)f_{\mathbf{c}}(x)dx \right. \\
&\quad \left. + \int_{R^2} f_{\mathbf{a}}(x)f_{\mathbf{d}}(x)dx + \int_{R^2} f_{\mathbf{b}}(x)f_{\mathbf{c}}(x)dx + \int_{R^2} f_{\mathbf{b}}(x)f_{\mathbf{d}}(x)dx \right. \\
&\quad \left. + \int_{R^2} f_{\mathbf{c}}(x)f_{\mathbf{d}}(x)dx \right\} + \mu_Z^4 \\
&= \frac{\pi\lambda\mu_{4Y}w^2}{4} \exp \left[-\frac{\delta^2}{w^2}(\langle \mathbf{a}, \mathbf{a} \rangle + \langle \mathbf{b}, \mathbf{b} \rangle + \langle \mathbf{c}, \mathbf{c} \rangle + \langle \mathbf{d}, \mathbf{d} \rangle \right. \\
&\quad \left. - \frac{1}{4} \langle \mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d}, \mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d} \rangle \right] \\
&\quad + \mu_Z \{ \text{cum}_Z(\mathbf{a}, \mathbf{b}, \mathbf{c}) + \text{cum}_Z(\mathbf{b}, \mathbf{c}, \mathbf{d}) + \text{cum}_Z(\mathbf{a}, \mathbf{b}, \mathbf{d}) \} \\
&\quad + \{ \text{cum}_Z(\mathbf{a} - \mathbf{b})\text{cum}_Z(\mathbf{c} - \mathbf{d}) + \text{cum}_Z(\mathbf{a} - \mathbf{c})\text{cum}_Z(\mathbf{b} - \mathbf{d}) \\
&\quad + \text{cum}_Z(\mathbf{a} - \mathbf{d})\text{cum}_Z(\mathbf{b} - \mathbf{c}) \} \\
&\quad + \mu_Z^2 \{ \text{cum}_Z(\mathbf{a} - \mathbf{b}) + \text{cum}_Z(\mathbf{a} - \mathbf{c}) + \text{cum}_Z(\mathbf{a} - \mathbf{d}) + \text{cum}_Z(\mathbf{b} - \mathbf{c}) \\
&\quad \quad \text{cum}_Z(\mathbf{b} - \mathbf{d}) + \text{cum}_Z(\mathbf{c} - \mathbf{d}) \} + \mu_Z^4,
\end{aligned}$$

Thus,

$$\begin{aligned}
& f_4(\lambda^{(1)}, \lambda^{(2)}, \lambda^{(3)}) \\
&= \frac{1}{(2\pi)^6} \sum_{a=-\infty}^{\infty} \sum_{b=-\infty}^{\infty} \sum_{d=-\infty}^{\infty} \text{cum}_{X,4}(\mathbf{a}, \mathbf{b}, \mathbf{d}) e^{-i(\langle \lambda^{(1)}, \mathbf{a} \rangle + \langle \lambda^{(2)}, \mathbf{b} \rangle + \langle \lambda^{(3)}, \mathbf{d} \rangle)}
\end{aligned}$$

where

$$\text{cum}_{X,4}(\mathbf{a}, \mathbf{b}, \mathbf{d}) = \text{cum}_4\{X(\mathbf{0}), X(\mathbf{a}), X(\mathbf{b}), X(\mathbf{d})\}.$$

After some algebraic cancellation, and since the kernel is Gaussian,

$$\begin{aligned}
& \text{cum}_4\{Z(\mathbf{a}), Z(\mathbf{b}), Z(\mathbf{c}), Z(\mathbf{d})\} \\
&= \frac{\pi\lambda\mu_{4Y}w^2}{4} \exp \left[-\frac{\delta^2}{w^2}(\langle \mathbf{a}, \mathbf{a} \rangle + \langle \mathbf{b}, \mathbf{b} \rangle + \langle \mathbf{c}, \mathbf{c} \rangle + \langle \mathbf{d}, \mathbf{d} \rangle \right. \\
&\quad \left. - \frac{1}{4} \langle \mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d}, \mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d} \rangle \right].
\end{aligned}$$

See Benn (1996) for the algebraic details.

References

A.G. Benn. Ph.D. thesis, Department of Statistical and Actuarial Sciences, University of Western Ontario, 1996.

- A.G. Benn and R.J. Kulperger. Statistical Inference for a Poisson Process Model of Scanning Fluorescence Correlation Spectroscopy. *Environmetrics*, 7, 167–83, 1996.
- A.G. Benn and R.J. Kulperger. Integrated Marked Poisson Processes with Application to Image Correlation Spectroscopy. *Canadian Journal of Statistics*, 25, 215–31, 1997.
- A.G. Benn and R.J. Kulperger. On the use of FFT properties for sums of periodograms. *Annals of the Institute of Statistical Mathematics*, 50, 187–202, 1998.
- D.R. Brillinger. *Time Series: Data Analysis and Theory*. Holden-Day, San Francisco, 1981.
- D.J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes*. Springer-Verlag, New York, 1988.
- M.H. Freeman. *Optics*. Tenth edition, Butterworths, London, 1990.
- D.M. Keenan. Limiting behaviour of functionals of higher-order sample cumulant spectra. *The Annals of Statistics*, 15, 134–151, 1987.
- D. Magde and E.L. Elson. Fluorescence correlation spectroscopy. 1. Conceptual basis and theory. *Biopolymers*, 13, 1–27, 1974.
- A. Palmer and N. Thompson. Molecular aggregation characterized by high order autocorrelation in fluorescence correlation spectroscopy. *Biophys. J.*, 52, 257–70, 1987.
- N.O. Petersen. Scanning fluorescence correlation spectroscopy. 1. Theory and simulation of aggregation measurements. *Biophys. J.*, 49, 809–815, 1986.
- N.O. Petersen. Quantitation of membrane receptor distributions by image correlation spectroscopy: concept and application. *Biophys. J.*, 65, 1135–1146, 1993.
- P. St-Pierre. Ph.D. thesis, Department of Chemistry, University of Western Ontario.
- P.R. St-Pierre and N.O. Petersen. Relative ligand binding to small or large aggregates measured by scanning correlation spectroscopy. *Biophys. J.*, 58, 503–511, 1990.
- P.R. St-Pierre and N.O. Petersen. Average density and size of microclusters of epidermal growth factor receptors on A431 cells. *Biochemistry*, 31, 2459–63, 1992.
- D.L. Snyder and M.I. Miller. *Random Point Processes in Time and Space*. 2nd ed., Springer-Verlag, New York, 1991.
- M. Taniguchi. On estimation of the integrals of the fourth-order cumulant spectral density. *Biometrika*, 69, 117–122, 1982.

Author's address:

R.J. Kulperger

Department of Statistical and Actuarial Sciences

The University of Western Ontario

London, Canada, N6A 5B7

E-mail: rjk@fisher.stats.uwo.ca