The effect of channel correlation on the accuracy of photon counting digital autocorrelators (photon counting spectroscopy)

1973 J. Phys. A: Math. Nucl. Gen. 6 1897

(http://iopscience.iop.org/0301-0015/6/12/014)

J. Phys. A: Math., Nucl. Gen., Vol. 6, December 1973. Printed in Great Britain. © 1973

The effect of channel correlation on the accuracy of photon counting digital autocorrelators

B E A Saleh and M F Cardoso

Electrical Engineering Department, University of Santa Catarina, Florianopolis, Santa Catarina, Brazil

Received 5 March 1973, in final form 17 July 1973

Abstract. We investigate the estimation of spectral parameters of optical fields by digital autocorrelation of photon counting fluctuations, taking into account the correlations that exist between readings of the autocorrelator channels. A general expression for the covariance of the normalized intensity autocorrelation estimators is derived for a gaussian field with general lineshape. Two estimation schemes are outlined, the maximum likelihood and the least-squares. Dependence of the estimation accuracy on the channel correlations and the system parameters is discussed.

1. Introduction

The methods of photon correlation spectroscopy have recently been a subject of great interest (for a review, see Pike and Jakeman 1973). Several authors (Jakeman and Pike 1968, Jakeman *et al* 1971, Degiorgio and Lastovka 1971, Kelly 1971) have investigated the problem of the statistical accuracy of these methods when applied to estimate the spectral linewidth of gaussian–lorentzian light. Jakeman *et al* (1971) in an exhaustive theoretical study have determined an appropriate weighting for a least-squares fitting procedure under the assumption that the readings of the various channels are uncorrelated. This assumption, they have noted, is reasonable when the average counting rate is sufficiently small or when the sample time interval (the resolution time of the instrument) is sufficiently longer than the spectral coherence time. In practice, these conditions are not always satisfied. The results of Jakeman *et al* (1971) show that the optimum sample time should be much shorter than the coherence time especially when the counting rates are high. This has also been confirmed by computer simulated experiments (Swinney 1973, private communication). Moreover, in many applications, such as those involving laser sources, the counting rates are not sufficiently small to justify the assumption that the channels are uncorrelated. The new technique of local oscillator scattering, introduced by light scattering experimentalists to circumvent the trouble with parasitic scattering, implies that new experiments will be performed at very high counting rates.

In this paper, the overall problem of statistical errors in photon counting spectroscopy is formulated in a more general form than is to be found in earlier papers on the subject. We calculate the channel correlations in the general case of a field that is not necessarily lorentzian, and study the optimum data processing schemes for estimating one or several parameters in the situations when the channels are not necessarily uncorrelated. The interest in a general light spectrum stems from the fact that the light scattered from binary liquid mixtures or from biological cells or macromolecules (two subjects of recent great interest) is not lorentzian.

In the present work, several assumptions are used. (i) The sample time interval is much shorter than the field's coherence time. (ii) The total time of the experiment is much longer than the coherence time. This is valid in most practical applications (if the total time of the experiment is 50–500 s and the coherence time is 0·5 μs–5 ms, then their ratio is of the order of 10^4–10^9). (iii) The optical field is stationary, cross-spectrally pure and gaussian.

The effects of clipping, scaling, dead time and of the finite aperture are not considered here. However, the effect of biasing is included.

Our estimation procedure is based on the calculation of the normalized (biased) estimators

$$\hat{g}_l = \frac{(1/N)\,\Sigma_{m=1}^{N}\, n(m)n(m+l)}{[(1/N)\,\Sigma_{r=1}^{N}\, n(r)]^2}, \qquad l = 1, 2, \ldots, L, \tag{1}$$

where the total available time T is divided into N intervals each of width T_d, $n(m)$ is the number of counts in the mth interval, and L is the number of channels. If the coherence time is τ_c, then the assumptions (i) and (ii) above mean that $T \gg \tau_c \gg T_d$.

In § 2, we study the statistical properties of the set $\{\hat{g}_l\}$ defined above. In § 3, two data processing schemes are described, the maximum likelihood (ML) estimation and the least-squares (LS) estimation. This involves demonstrating the estimation procedure and calculating the estimation error (or accuracy). In § 4, examples are given and the results are compared with those of Jakeman *et al* (1971). The final section is devoted to discussions of the general conclusions of this work.

2. Statistical properties of the normalized intensity autocorrelation estimators $\{\hat{g}_l\}$

In order to simplify the expressions of the statistical moments of $\{\hat{g}_l\}$, we follow the approach of Jakeman *et al* (1971) and define the denominator of (1) as

$$\hat{n} = (1/N) \sum_{m=1}^{N} n(m). \tag{2}$$

We also define

$$\hat{g}_{0l} = \frac{1}{\bar{n}^2}\frac{1}{N}\sum_{m=1}^{N} n(m)n(m+l). \tag{3}$$

The statistic $\hat{n}$, which is measurable, is an unbiased estimator of $\bar{n}$, which is unknown, ie

$$E\{\hat{n}\} = \bar{n}, \tag{4}$$

where E is the expectation value. Also, the statistic $\hat{g}_{0l}$ is an unbiased estimator for the normalized intensity coherence function g_{0l} which, from now on, we shall call g_l,

$$E\{\hat{g}_{0l}\} = g_l, \qquad l \neq 0. \tag{5}$$

However, $\hat{g}_{0l}$ is not measurable because $\bar{n}$ is not precisely known. Fortunately, because N is very large, $\hat{n}$ has a value very close to that of $\bar{n}$ and also $\hat{g}_{0l}$ is approximately equal to g_l. This is used in (1) by expanding $\hat{n}$ around $\bar{n}$ and $\hat{g}_{0l}$ around g_l and thus getting

$$\hat{g}_l \simeq \hat{g}_{0l} + g_l\left[-2\left(\frac{\hat{n}-\bar{n}}{\bar{n}}\right) + 3\left(\frac{\hat{n}-\bar{n}}{\bar{n}}\right)^2 - 2\left(\frac{\hat{n}-\bar{n}}{\bar{n}}\right)\left(\frac{\hat{g}_{0l}-g_l}{g_l}\right)\right]. \tag{6}$$

We are now in a position to study the statistical moments of $\hat{g}_l$ by using (6) instead of (1). By taking the expectation value of $\hat{g}_l$ as given by (6) we find that $\hat{g}_l$ is a biased estimator of g_l. However, the bias terms can be shown to be proportional to $(TW)^{-1}$ where TW is the number of coherence lengths in an observation time ($W = \tau_c^{-1}$ is the bandwidth). On the other hand, we will see that the standard deviation of $\hat{g}_l$ (the square root of the variance) is proportional to $(TW)^{-1/2}$. As we assume that $TW \gg 1$, it is reasonable to assume that $\hat{g}_l$ is an unbiased estimator of g_l, ie

$$E\{\hat{g}_l\} \simeq g_l. \tag{7}$$

The covariance between $\hat{g}_l$ and $\hat{g}_k$ is defined as

$$\Lambda_{lk} = \mathrm{cov}\{\hat{g}_l, \hat{g}_k\} = E\{\hat{g}_l\hat{g}_k\} - E\{\hat{g}_l\}E\{\hat{g}_k\}. \tag{8}$$

By using (6) and neglecting terms of higher orders in $[(\hat{n}-\bar{n})/\bar{n}]$ and $[(\hat{g}_{0l}-g_l)/g_l]$, we get

$$\Lambda_{lk} = \Lambda_{lk}^{(0)} + \Lambda_{lk}^{(1)}, \tag{9}$$

where

$$\Lambda_{lk}^{(0)} = \mathrm{cov}\{\hat{g}_{0l}, \hat{g}_{0k}\}, \tag{10}$$

and

$$\Lambda_{lk}^{(1)} = 4g_lg_k\left(\frac{1}{\bar{n}^2}\right)\mathrm{var}(\hat{n}) - 2g_l\frac{1}{\bar{n}}E[\hat{g}_{0k}(\hat{n}-\bar{n})] - 2g_k\frac{1}{\bar{n}}E[\hat{g}_{0l}(\hat{n}-\bar{n})]. \tag{11}$$

It is shown in the appendix that these two terms can be written as

$$\Lambda_{lk}^{(1)} = (TW)^{-1}[d_{l,k} - (4/\bar{n}_c)c_lc_k] \tag{12}$$

and

$$\Lambda_{lk}^{(0)} = (TW)^{-1}[a_{l,k} + (1/\bar{n}_c)b_{l,k} + (1/\bar{n}_c\bar{n})c_l\delta_{l,k}] \tag{13}$$

where $\bar{n}_c$ is the average number of counts in a coherence time,

$$a_{l,k} = 4 + z_{l-k} + z_{l+k} + 2\,\mathrm{Re}(\chi_l^*\chi_k y_{l-k}) + 4\,\mathrm{Re}(\chi_l^* y_l + \chi_k^* y_k) + 2\,\mathrm{Re}(\chi_l\chi_k y_{l+k}^*) + u_{l,k}, \tag{14}$$

$$b_{l,k} = 2[1 + \chi_l^2 + \chi_k^2 + \chi_{l-k}^2 + 2\,\mathrm{Re}(\chi_l^*\chi_k\chi_{l-k})], \tag{15}$$

$$c_l = 1 + \chi_l^2, \tag{16}$$

and

$$d_{l,k} = -4[1 + \chi_l^2 + \chi_k^2 + \chi_l^2\chi_k^2 - \mathrm{Re}(\chi_l^* y_l + \chi_k^* y_k) - \chi_l^2\,\mathrm{Re}(\chi_k^* y_k) - \chi_k^2\,\mathrm{Re}(\chi_l^* y_l)]. \tag{17}$$

In (14–17) χ, y, z and u are related to $\chi(\tau)$ by

$$\chi_l = \chi(\tau_l), \tag{18}$$

$$y_l = W\int_{-\infty}^{+\infty} \chi(t)\chi^*(t-\tau_l)\,\mathrm{d}t, \tag{19}$$

$$z_l = W\int_{-\infty}^{+\infty} |\chi(t+\tau_l)|^2|\chi(t)|^2\,\mathrm{d}t, \tag{20}$$

$$u_{l,k} = 2W\,\mathrm{Re}\int_{-\infty}^{+\infty} \chi^*(t)\chi(t-\tau_l)\chi(t+\tau_k)\chi^*(t+\tau_k-\tau_l)\,\mathrm{d}t. \tag{21}$$

Thus, we have found an expression for the covariance of the set of estimators $\{\hat{g}_l\}$ as a function of $\chi(\tau_l)$ of the optical field. It is to be remembered that this expression is not valid when either τ_l or τ_k or both are zero. In such cases, additional terms have to be included, but these cases are of no interest to us here.

A very important special type of optical field is that which is quasimonochromatic and has a lorentzian spectrum. In this case,

$$\chi_l = \exp(-|x_l|)\exp(-\mathrm{i}\omega_0\tau_l), \tag{22}$$

where $x_l = \Gamma\tau_l, \Gamma = W$ is the bandwidth and ω_0 is the central frequency. By substituting in the above general expression we get

$$\begin{aligned}\Lambda_{lk} = (TW)^{-1}\Big((\tfrac{1}{2}+|x_l-x_k|)\exp(-2|x_l-x_k|)-(\tfrac{5}{2}+x_l+x_k)\exp[-2(x_l+x_k)]\\ +4(1+|x_l-x_k|)\exp[-(x_l+x_k)-|x_l-x_k|]+\frac{2}{\bar{n}_\mathrm{c}}\{\exp(-2|x_l-x_k|)\\ +2\exp[-(x_l+x_k)-|x_l-x_k|]+\exp[-2(x_l+x_k)]\}\\ +\delta_{l,k}\frac{1}{\bar{n}_\mathrm{c}\bar{n}}[1+\exp(-2x_l)]\Big),\end{aligned} \tag{23}$$

from which the variance is

$$\begin{aligned}\Lambda_{ll} = (TW)^{-1}\Big(\tfrac{1}{2}-(\tfrac{5}{2}+2x_l)\exp(-4x_l)+4\exp(-2x_l)\\ +\frac{2}{\bar{n}_\mathrm{c}}[1+\exp(-2x_l)]^2+\frac{1}{\bar{n}_\mathrm{c}\bar{n}}[1+\exp(-2x_l)]\Big),\end{aligned} \tag{24}$$

which is the same as the expression derived by Jakeman *et al* (1971). It is to be noted that if the count rate $\bar{n}$ is very small, the last term in (23) dominates and $\{\hat{g}_l\}$ are uncorrelated. On the other hand, if $\bar{n}$ is large, the matrix Λ can be very correlated.

3. Optimum processing of data

In this section we are concerned with the problem of finding the best estimates of parameters of the light field spectrum given the set of observations $\{\hat{g}_l\}$, $l = 1, \ldots, L$. We examine two methods, the maximum likelihood (ML) estimation and the method of least-squares (LS).

3.1. Maximum likelihood estimator

In order to find the ML estimator, the joint probability distribution of the set of statistics $\{\hat{g}_l\}$ has to be found. This is, in general, an extremely difficult task. However, fortunately, in most practical applications TW is a very large number (10^4–10^9), and each of the statistics $\{\hat{g}_l\}$ is effectively the sum of a very large number of independent components and therefore it seems reasonable to use the central limit theorem and conclude that $\{\hat{g}_l\}$ has, approximately, a multivariate gaussian distribution. Adopting this assumption, we now have an expression for the joint probability distribution $P(\{\hat{g}_l\})$, because we have already found expressions for the means and the covariance (equations (7), (9), (12), (13)).

Let the parameters to be estimated be $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_S)$, and the ML estimates be $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \ldots, \hat{\theta}_S)$, then $\hat{\boldsymbol{\theta}}$ is the value of $\boldsymbol{\theta}$ which makes $P(\{\hat{g}_l\})$ maximum. Using the multivariate gaussian expression, taking the logarithm and using the identity $\ln \det \boldsymbol{\Lambda} = \mathrm{Tr} \ln \boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda}$ is the covariance matrix whose elements are $\{\Lambda_{lk}\}$, we are led to the minimization problem

$$\min_{\hat{\boldsymbol{\theta}}}\{\tfrac{1}{2}\,\mathrm{Tr}\ln \boldsymbol{\Lambda}(\boldsymbol{\theta}) + \tfrac{1}{2}[\hat{\boldsymbol{g}} - \boldsymbol{g}(\boldsymbol{\theta})]^T \boldsymbol{\Lambda}^{-1}[\hat{\boldsymbol{g}} - \boldsymbol{g}(\boldsymbol{\theta})]\} \tag{25}$$

where $\hat{\boldsymbol{g}} - \boldsymbol{g}$ is the vector whose components are $\{\hat{g}_l - g_l\}$. This minimization problem can be solved by finding the value of $\boldsymbol{\theta}$ which makes the first derivatives with respect to each of the components of $\boldsymbol{\theta}$ vanish, ie

$$\frac{1}{2}\mathrm{Tr}\left(\boldsymbol{\Lambda}^{-1}\frac{\partial \boldsymbol{\Lambda}}{\partial \theta_s}\right) + \frac{1}{2}[\hat{\boldsymbol{g}} - \boldsymbol{g}(\hat{\boldsymbol{\theta}})]^T \frac{\partial \boldsymbol{\Lambda}^{-1}}{\partial \theta_s}[\hat{\boldsymbol{g}} - \boldsymbol{g}(\hat{\boldsymbol{\theta}})] + [\hat{\boldsymbol{g}} - \boldsymbol{g}(\hat{\boldsymbol{\theta}})]^T \boldsymbol{\Lambda}^{-1}\frac{\partial \boldsymbol{g}(\hat{\boldsymbol{\theta}})}{\partial \theta_s} = 0, \quad s = 1, \ldots, S. \tag{26}$$

This set of S nonlinear equations in S unknowns $\hat{\theta}_1, \ldots, \hat{\theta}_S$, can, in general be solved numerically (eg using the Newton–Raphson method) and $\hat{\boldsymbol{\theta}}$ can be determined.

In order to calculate the resulting estimation errors we find the variances of the estimates $\{\theta_s\}$. This can be obtained by making use of the fact that TW is very large and therefore the variances of its ML estimators are given by (see Mood and Graybill 1963, theorem 10.8)

$$\mathrm{var}\,\hat{\theta}_j = [\boldsymbol{R}^{-1}]_{jj}, \tag{27}$$

where $\boldsymbol{R}$ is the matrix whose elements are

$$R_{ij} = -E\left(\frac{\partial^2}{\partial \theta_i \partial \theta_j}\ln(P)\right), \tag{28}$$

and P is the multivariate normal distribution which when substituted in (28) gives approximately

$$R_{ij} \simeq \frac{\partial \boldsymbol{g}^T}{\partial \theta_i}\boldsymbol{\Lambda}^{-1}\frac{\partial \boldsymbol{g}}{\partial \theta_j} \tag{29}$$

(when terms of order $(TW)^{-1}$ are neglected). The case $S = 1$ gives simply

$$\mathrm{var}(\theta) = \left(\frac{\partial \boldsymbol{g}^T}{\partial \theta}\boldsymbol{\Lambda}^{-1}\frac{\partial \boldsymbol{g}}{\partial \theta}\right)^{-1}. \tag{30}$$

Thus, the ML estimators $\hat{\boldsymbol{\theta}}$ are given by solving the set of nonlinear equations (26) and the error is given by (27) and (29). This holds good provided that our assumption of a gaussian distribution is justifiable. It may be argued that the central limit theorem is least accurate in the tails of the distribution which may contribute significantly in (28) when the logarithm is taken. Hence, the expression of the error may not be very accurate. However, it is not the error that we are interested in estimating with maximum accuracy; it is rather the parameters $\boldsymbol{\theta}$ themselves which (26) can supply with good accuracy. Yet (27) and (29) give a reasonable approximation for the accuracy of ML estimation.

3.2. Least-squares method

The square error is written as

$$e = \sum_{l=1}^{L} \zeta_l [\hat{g}_l - g_l(\boldsymbol{\theta})]^2,$$

where ζ_l are weighting numbers to be chosen later. The LS estimate $\hat{\boldsymbol{\theta}}$ is the value of $\boldsymbol{\theta}$ which makes e a minimum. By equating to zero the first derivatives with respect to each of the parameters θ_s, we get S equations

$$\sum_{l=1}^{L} \zeta_l [\hat{g}_l - g_l(\hat{\boldsymbol{\theta}})] \frac{\partial g_l(\hat{\boldsymbol{\theta}})}{\partial \theta_s} = 0, \qquad s = 1, \ldots, S. \tag{31}$$

This, in principle, determines $\hat{\theta}$ except that the weights are not yet known. We next find expressions for the variances of $\hat{\boldsymbol{\theta}}$. This can be done by linearization. Expanding $g_l(\hat{\boldsymbol{\theta}})$ around $\boldsymbol{\theta}$, we obtain S linear equations in $\hat{\theta}_s - \theta_s$, which can be solved and from which

$$\operatorname{var}(\hat{\theta}_s) = \sum_{l,k} \zeta_l \zeta_k f_l^s f_k^s \Lambda_{lk}, \tag{32}$$

where

$$f_l^s = \sum_{r=1}^{S} [F^{-1}]_{sr} \frac{\partial g_l}{\partial \theta_r}, \tag{33}$$

and $\boldsymbol{F}^{-1}$ is the inverse of the matrix $\boldsymbol{F}$ whose elements are

$$F_{sr} = \sum_{l=1}^{L} \zeta_l \frac{\partial g_l}{\partial \theta_r} \frac{\partial g_l}{\partial \theta_s}. \tag{34}$$

Given that it is of equal importance to estimate correctly each of θ_s, the optimum weights ζ_l are those which minimize the sum of the variances of the estimates, ie

$$\sum_{l,k} \zeta_l \zeta_k \left(\sum_s f_l^s f_k^s \right) \Lambda_{l,k} = \text{minimum}. \tag{35}$$

Of course, another set of weights may be included to account for the importance of estimating the various parameters.

Thus two steps have to be followed in order to obtain the LS estimates; the values of $\{\zeta_l\}$ that minimize (35) have to be determined, and the set of nonlinear equations (31) have to be solved for $\hat{\boldsymbol{\theta}}$. The estimation accuracies are given by (32).

If only one parameter is to be estimated, $S = 1$, the estimation accuracy is given by

$$\operatorname{var}(\hat{\theta}) = \frac{\sum_{l,k} \zeta_l \zeta_k (\partial g_l / \partial \theta)(\partial g_k / \partial \theta) \Lambda_{l,k}}{[\sum_l \zeta_l (\partial g_l / \partial \theta)^2]^2}. \tag{36}$$

3.3. Comparison between the ML and LS estimates

The ML and LS estimation methods are equivalent when the observations are uncorrelated. In this case the set of equations (27) and (29) and the set (32), (33), (34) and (35) both give the same result. When the observations are correlated, the two methods are different. The LS method is more conservative and it makes no use of the knowledge that the observations are approximately normal, thus it gives larger errors. It also involves longer computations (to solve the optimization problem). Hence, it seems that, for very large TW it is better to use the ML method.

4. Examples

In this section, we use the results of the previous sections to find the accuracy of estimating parameters of some spectral distributions of practical importance, and also demonstrate numerically the effect of taking into account the channel correlations.

4.1. Lineshape with one parameter

We consider the problem of estimating the spectral width of a light field with lorentzian lineshape, ie

$$g_l = 1+\exp(-2\Gamma|\tau_l|), \tag{37}$$

and Γ is to be estimated. In this example $S = 1$. The covariance matrix Λ_{lk} has already been calculated and it is given by (23).

4.1.1. ML *estimator.* Taking $\theta = \Gamma$ and substituting (37) in (30), we get

$$\epsilon_{\hat{\Gamma}}^2\bigg|_{\mathrm{ML}} = \frac{\operatorname{var}\hat{\Gamma}}{\Gamma^2} = \frac{1}{4}\left(\sum_{l,k} d_l d_k (\Lambda^{-1})_{lk}\right)^{-1}, \tag{38}$$

where

$$d_l = x_l \exp(-2x_l), \qquad x_l = \Gamma|\tau_l|. \tag{39}$$

4.1.2. LS *estimator.* Similarly by substituting (37) in (35), we get

$$\epsilon_{\hat{\Gamma}}^2\bigg|_{\mathrm{LS}} = \frac{\operatorname{var}\hat{\Gamma}}{\Gamma^2} = \min_{\{\zeta_l\}} \frac{1}{4}\,\frac{\Sigma_{lk}\,\zeta_l\zeta_k d_l d_k \Lambda_{lk}}{(\Sigma_l \zeta_l d_l^2)^2}. \tag{40}$$

The values of $\{\zeta_l\}$ which minimize (40) are given by

$$\zeta_l = (\boldsymbol{\beta}^{-1})_{l,1}, \tag{41}$$

where $\boldsymbol{\beta}$ is the matrix whose elements are

$$\beta_{l,k} = d_k(d_k\delta_{k,1} + d_1\Lambda_{l,k} - d_l\Lambda_{1,k}). \tag{42}$$

It is easy to see that in the limit when Λ_{lk} is uncorrelated both the ML and LS variances given by (38) and (40) are equal and have a value

$$\epsilon_{\hat{\Gamma}}^2\bigg|_{\mathrm{LS,ML,uncorrelated}} = \frac{1}{4}\left(\sum_l (d_l^2/\Lambda_{ll})\right)^{-1}. \tag{43}$$

We have calculated the errors $\epsilon_{\hat{\Gamma}}$ given by (38), (40) and (43) for several values of the counting rate per coherence time, $\bar{n}_c$, and at a fixed $\gamma = \bar{n}/\bar{n}_c = 0{\cdot}1$ and $\Gamma T = 10^4$. Results are plotted in figure 1. It is not very surprising that, in this example the ML and LS methods have almost the same error even when the channels are correlated. After all, the two criteria of estimation are very similar. The set of curves (calculated for 20 channels) shows that the effect of correlation increases as $\bar{n}_c$ increases. At $\bar{n}_c = 10$ and $\bar{n} = 1$ the estimation error calculated assuming uncorrelated channels is smaller than the correct estimation error by approximately a factor of 3. Jakeman *et al* (1971) have previously pointed out that the variance calculated assuming no correlation serves as a lower bound to the real variance. Our present results give exact values for the errors and thus help for a proper design of experiments.

The errors plotted in figure 1 are calculated for 20 channels with time delays $\{x_l\}$ equally spaced between 0·1 and 2. We have found that the error is sensitive to the choice

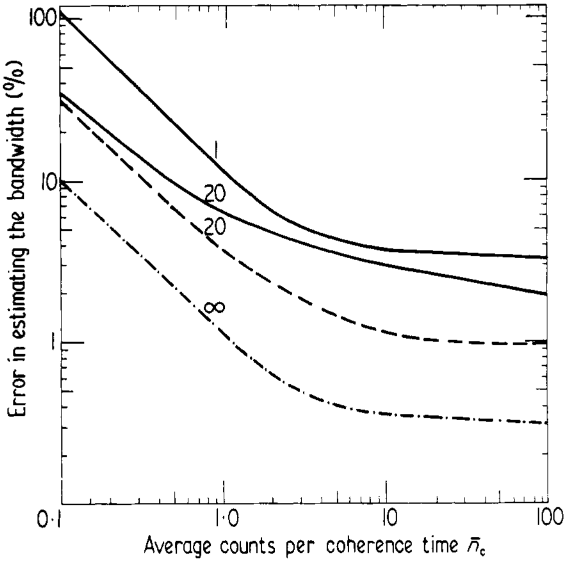

Figure 1. Error in estimating the bandwidth Γ by fitting to $g_l = 1+\exp(-2\Gamma|\tau_l|)$ (§ 4.1) as a function of the counting rate per coherence time $\bar{n}_c$, for $\Gamma T = 10^4$, $\gamma = 0{\cdot}1$, and the indicated number of channels L: full curves, correlation taken into consideration; broken curve, channels assumed uncorrelated; chain curve, infinite uncorrelated channels. Single channel estimates are calculated at a time delay $x_l = \Gamma|\tau_l| = 0{\cdot}6$. The 20 channel estimates are calculated for time delays $\{x_l\} = \{0{\cdot}1, 0{\cdot}2, \ldots, 2\}$.

of $\{x_l\}$. For example, if instead, the time delays are equally spaced between 0·1 and 3·9, the error at $\bar{n}_c = 0{\cdot}1$ is 47·6 % instead of 35·5 %. In order to demonstrate the importance of choosing the right time delays we plot (figure 2) the error in estimating Γ using a single channel against its delay. In this case, (38), (40) and (43) all give the same result,

$$\epsilon^2 = \Lambda_{ll}/4d_l^2. \tag{44}$$

It is seen from figure 2 that the error is minimum when $x = 0{\cdot}6$. This is similar to results we have previously found (Saleh 1973) when using digital autocorrelators to estimate the spatial coherence length. The variation of the optimum error based on a single channel with $\bar{n}_c$ is also plotted in figure 1. By comparing this to the set of error curves at $L = 20$, we can see how much gain in accuracy we can get by increasing the number of channels from 1 to 20. Also for comparison, the estimation error when infinite channels are used, is plotted. This is calculated under the assumption of uncorrelated channels.

4.2. Lineshape with two parameters

In this important example, we assume

$$g_l = 1+C\exp(-2\Gamma|\tau_l|), \tag{45}$$

where Γ is the parameter to be estimated, but C is just an unknown parameter which appears because of the finite sampling time, the finite detector area, as well as clipping effects. These effects have been neglected when we derived our expression for the covariance matrix of $\hat{g}_l$. It would be extremely difficult to include such effects and it is likely that more than a single factor C would appear. But it seems reasonable (when

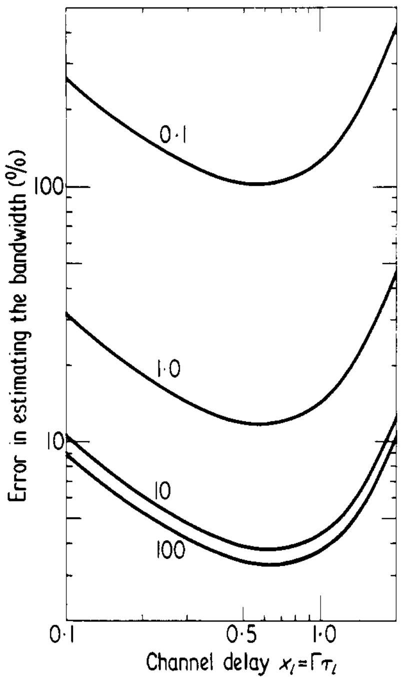

Figure 2. Percentage error in estimating the bandwidth Γ using a single channel (§ 4.1) as a function of the time delay $x_l = \Gamma|\tau_l|$ for several counting rates $\bar{n}_c$, $\Gamma T = 10^4$ and $\gamma = 0{\cdot}1$.

$\Gamma T \gg 1$) to include these effects in the function to be fitted g_l, and approximately neglect them when calculating the covariance of its estimators. Taking $\theta_1 = \Gamma$ and $\theta_2 = C$ we can use the results of the previous section directly to find the estimation error.

4.2.1. ML *estimator.* Using (45) in (27) and (29) we get

$$\epsilon^2_{\hat{\Gamma}}\Big|_{\mathrm{ML}} = \frac{\mathrm{var}(\hat{\Gamma})}{\hat{\Gamma}^2} = \frac{1}{4}\frac{R_{22}}{R_{11}R_{22}-R_{12}R_{12}}, \tag{46}$$

where

$$\begin{aligned} R_{11} &= \sum_{l,k} d_l d_k \Lambda_{lk}^{-1}, \\ R_{12} &= \sum_{l,k} d_l h_k \Lambda_{lk}^{-1}, \\ R_{22} &= \sum_{l,k} h_l h_k \Lambda_{lk}^{-1}, \\ h_l &= \exp(-2x_l), \\ d_l &= x_l \exp(-2x_l). \end{aligned} \tag{47}$$

4.2.2. LS *estimator.* Similarly, we use (45) in (32), (33) and (34), and get

$$\epsilon^2_{\hat{\Gamma}}\Big|_{\mathrm{LS}} = \frac{\mathrm{var}(\hat{\Gamma})}{\Gamma^2} = \min_{\{\zeta_l\}} \frac{1}{4C^2}\frac{\Sigma_{l,k}\,\zeta_l\zeta_k\bar{d}_l\bar{d}_k\Lambda_{l,k}}{(\Sigma_l\,\zeta_l d_l\bar{d}_l)^2}, \tag{48}$$

where

$$\tilde{d}_l = (x_l - \bar{x}) \exp(-2x_l), \qquad \bar{x} = \frac{\Sigma_j \zeta_j x_j \exp(-4x_j)}{\Sigma_i \zeta_i \exp(-4x_i)}. \tag{49}$$

Equation (48) is the same as that derived by Jakeman *et al* (1971) who considered only the case when $\Lambda_{l,k}$ is diagonal. In this case, the optimum weights ζ_l are proportional to $1/\Lambda_{ll}$, and the ML and LS estimation errors given by (47) and (48) are equal. Now that we have found an expression for the covariance Λ, we are in a position to find more precise values for the estimation accuracies. However, we face the difficult problem of finding the optimum weights. We resort to numerical methods. By using the method of steepest descent, we have calculated the optimum weights and the accuracy of the LS estimators.

Results are presented in figures 3 and 4, which show, for different values of $\bar{n}_c$, the ML error, the LS error, and the error had the channels been uncorrelated. In figure 3, $\gamma = 0{\cdot}01$ and the errors are plotted for 20 channels taken such that $\{x_l\} = \{0{\cdot}1, 0{\cdot}15, \ldots, 1{\cdot}05\}$. Here also the error is sensitive to the choice of $\{x_l\}$. For example, if instead, $\{x_l\}$ is chosen to be $\{0{\cdot}5, 0{\cdot}51, \ldots, 0{\cdot}7\}$ the error is 731% instead of 149% for $\bar{n}_c = 0{\cdot}1$ ($\Gamma T = 10^4$). That explains why our errors when assuming 20 uncorrelated channels are lower than the errors calculated by Jakeman *et al* (1971) under the same conditions but for a different set of $\{x_l\}$. Therefore, for a certain counting rate, our expressions should be used to find out the optimum set of time delays. In figure 3, we also plot the error (calculated by Jakeman *et al* 1971) when infinite uncorrelated channels are used and $\gamma = 0{\cdot}01$. In figure 4, $\gamma = 0{\cdot}1$ and the errors are plotted for 20 channels with

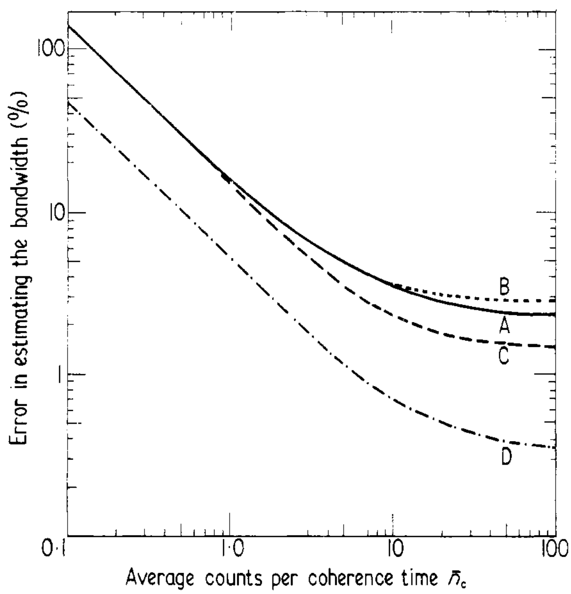

Figure 3. Percentage error in estimating the bandwidth Γ by fitting to $g = 1 + C\exp(-2\Gamma|\tau_l|)$ (§ 4.2) as a function of the counting rate per coherence time $\bar{n}_c$. The vertical axis is scaled by a factor $(1/C)$, $\Gamma T = 10^4$ and $\gamma = 0{\cdot}01$: curve A, ML estimate based on 20 channels; curve B, LS estimate based on 20 channels; curve C, ML and LS estimates assuming 20 uncorrelated channels; curve D, infinite uncorrelated channels. The 20 channel estimates are calculated for time delays such that $\{x_l\} = \{0{\cdot}1, 0{\cdot}15, \ldots, 1{\cdot}05\}$.

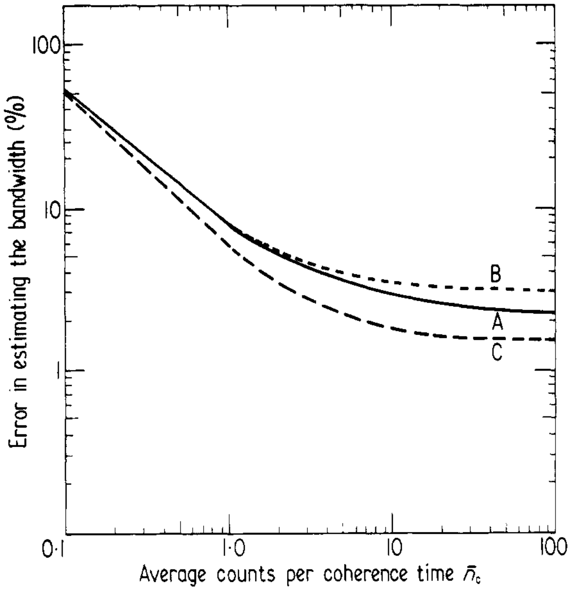

Figure 4. As in figure 3, but with $\gamma = 0{\cdot}1$ and the 20 channel estimates are calculated for time delays such that $\{x_l\} = \{0{\cdot}1, 0{\cdot}2, \ldots, 2\}$.

$\{x_l\} = \{0{\cdot}1, 0{\cdot}2, \ldots, 2\}$. If the channels are assumed uncorrelated we obtain the same errors as those calculated by Jakeman *et al* (1971) (slight differences may be attributed to our assumption that $\gamma \ll 1$).

The effect of correlation is clear in the graphs of figures 3 and 4. This effect increases as $\bar{n}_c$ increases. We also show in these graphs that the difference in the ML and LS errors is almost insignificant.

It is of interest to see that the errors in estimating with an unknown C (§ 4.2) are larger than those when C is known and equals 1 (§ 4.1), under the same conditions. Thus, ignoring our theoretical knowledge of the value of C leads to higher estimation errors.

5. Conclusions

We have found a general expression for the correlation between the readings of various channels of a photon counting digital autocorrelator. This expression is valid for any spectral distribution and is not limited to the lorentzian case. The knowledge of this correlation is necessary in estimating spectral parameters. Two data processing schemes which take into consideration this correlation have been analysed and general expressions for the estimation accuracies found. Following the 'recipes' we have given, any number of parameters that describe the field's spectrum can be estimated and the estimation error determined. We have shown that for $\bar{n}_c > 1{\cdot}0$, it is important to take into account the channel correlations, and that the ML and LS estimates have almost the same error. Although we have only given two examples of simple spectra, there are many practical spectra with one or several parameters to which our general formalism can be directly applied.

Acknowledgments

We wish to acknowledge the support of the Organization of American States and the Banco Nacional de Desenvolvimento Econômico. We are also grateful to the light scattering group of the New York University for valuable comments on the manuscript. One of us (BEAS) appreciates stimulating correspondence with Dr Harry Swinney.

Appendix

In this appendix, we derive the general expression of the covariance of two channel readings.

Let the optical field at the location of the point detector be described by its complex degree of coherence γ_{12} at two times t_1 and t_2, and also by its higher order normalized intensity coherence functions $g_{12\ldots M}$ at points in time $t_1, t_2, \ldots, t_M$ (see eg Mandel and Wolf 1965). Because we assume that the field is complex circular gaussian, the intensity coherence functions can be expanded in terms of γ_{12} by the general rule

$$g_{12\ldots M} = \sum_{(i \neq j \neq \ldots \neq k) = (1,2,\ldots,M)} \gamma_{1i}\gamma_{2j} \cdots \gamma_{Mk} \tag{A.1}$$

where the summation is taken over the indicated permutations. We also assume that the field is stationary which means that

$$\gamma_{12} = \chi(t_1 - t_2). \tag{A.2}$$

It is customary to choose the normalization $\chi(0) = 1$, and to define the bandwidth W such that

$$W^{-1} = \int_{-\infty}^{\infty} |\chi(t)|^2 \, dt, \tag{A.3}$$

and the effective coherence length $\tau_c = W^{-1}$.

Now we consider the photodetector. Let $n(m)$ be the number of photons counted in a time interval T_d centred around t_m. We need the statistical moments of $n(m)$. Expressions of these moments are, in general, difficult (see, eg, Jakeman 1970), but with our assumption that $T_d/\tau_c = \gamma \ll 1$, the moments are related to the intensity coherence functions of the detected field by the following equations derived from the properties of the Poisson distribution:

$$E\{n(1)n(2)\} = \bar{n}^2 g_{12} + \bar{n}\delta_{1,2}, \tag{A.4}$$

$$E\{n(1)n(2)n(3)\} = \bar{n}^3 g_{123} + \bar{n}^2 g_{13}\delta_{1,2} + (2\bar{n}^2 + \bar{n})\delta_{1,2}\delta_{2,3}, \qquad t_3 \geqslant t_2 \geqslant t_1, \tag{A.5}$$

$$\begin{aligned} E\{n(1)n(2)n(3)n(4)\} &= \bar{n}^4 g_{1234} + \bar{n}^3 g_{134}\delta_{1,2} + (2\bar{n}^3 g_{114} + \bar{n}^2 g_{14})\delta_{1,2}\delta_{2,3} + (3\bar{n}^3 + 6\bar{n}^2 + \bar{n}) \\ &\quad \times \delta_{1,2}\delta_{2,3}\delta_{3,4}, \qquad t_4 \geqslant t_3 \geqslant t_2 \geqslant t_1, \end{aligned} \tag{A.6}$$

where $E\{n(t)\} = \bar{n}$ is the average counting rate.

Now we are in a position to find $\Lambda_{lk}^{(0)}$ and $\Lambda_{lk}^{(1)}$. We use the definitions of $\{\hat{g}_{0l}\}$ and $\hat{n}$ given by (4) and (2). We relate the moments of $n(m)$ to the intensity coherence functions using (A.4), (A.5) and (A.6), and relate these to the complex degree of coherence $\gamma(\tau)$ using (A.1) and (A.2). We change all our summations over m into integrations over t making use of the assumption that $T_d \ll \tau_c$. By using stationarity and the definition of W given by (A.3) we finally get the required general expressions of the covariance (12) and (13) as a function of $\chi(\tau)$, T, T_d and $\bar{n}$.

References

Degiorgio V and Lastovka L B 1971 *Phys. Rev.* A **4** 2033–50
Jakeman E 1970 *J. Phys.* A: *Gen. Phys.* **3** 201–15
Jakeman E and Pike E R 1968 *J. Phys.* A: *Gen. Phys.* **1** 690–3
Jakeman E, Pike E R and Swain S 1971 *J. Phys.* A: *Gen. Phys.* **4** 517–34
Kelly H C 1971 *IEEE J. quantum Electron.* **7** 541–50
Mandel L and Wolf E 1965 *Rev. mod. Phys.* **37** 231–87
Mood A M and Graybill F A 1963 *Introduction to the Theory of Statistics* (New York: McGraw-Hill)
Pike E R and Jakeman E 1973 *Adv. quantum Electron.* **2** in the press
Saleh B A 1973 *J. Phys.* A: *Math., Nucl. Gen.* **6** 980–6