Dynamic-range limitations of intensified CCD photon-counting detectors

J. L. A. Fordham,¹ C. F. Moorhead¹ and R. F. Galbraith²

¹Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT ²Department of Statistical Science, University College London, Gower Street, London WC1E 6BT

Accepted 1999 September 7. Received 1999 June 21

ABSTRACT

The bright limit to the dynamic range of intensified CCD photon-counting detectors is governed by coincidence losses. In this paper a theoretical analysis of the loss mechanisms is carried out and verified using laboratory data. For applications where the input source is stable, such as star field imaging from space, the theoretical dynamic-range curve can then be used for accurate quantification.

Key words: instrumentation: detectors – instrumentation: miscellaneous – techniques: photometric.

1 INTRODUCTION

The dynamic range of any detector can be defined as the range of input light intensities over which, within the constraints of any particular programme, scientifically analysable data can be obtained.

The faint limit is governed by signal-to-noise ratio requirements and is ultimately limited by shot noise on the input.

The bright limit is governed by saturation effects within the detector. For the direct, cryogenically cooled, CCDs that are used in most optical astronomy applications the bright limit can, in principle, be assumed to be unlimited, input flux intensity defining only the integration time to prevent saturation in a charge-coupled device (CCD) pixel. For photon-counting detectors, though, the limit is governed by a number of factors. Here, we analyse the limitations imposed on such detectors, particular emphasis being attached to the maximum brightness that can be observed, this being placed in the context of the MIC photon-counting detector developed at University College London for ground- and spacebased astronomical applications. A version of this detector has been selected for incorporation as the Blue Detector (Horner et al. 1994.) in the Optical Monitor experiment on the ESA XMM Space Observatory due for launch in 1999 December and the results given here have direct relevance to the scientific performance of that experiment.

2 THE MIC DETECTOR

A schematic diagram of a typical CCD capture photon-counting system, the MIC detector (Fordham et al. 1992), is shown in Fig. 1. Initially photons are converted to electrons by the photocathode of an image intensifier (Norton et al. 1991). Electron amplification is then provided by a stack of microchannel plates (MCPs) providing an electron gain of $\sim 5 \times 10^5$. The charge cloud output is proximity focused on to a phosphor screen creating a scintillation with an intensity of $\sim 10^7$ photons. Each scintillation is then

captured by a fast-scanning, frame-transfer CCD (this being optically coupled to the phosphor screen by a fibre-optic taper) and centroided to minimize resolution losses in the intensifier prior to accumulation in computer memory.

3 FAINT LIMIT

The faint limit is governed by the noise contribution from the detector. This could potentially come from two sources; the image intensifier and the CCD camera.

3.1 CCD camera noise

Photon-counting detectors employ a threshold to discriminate against electronic noise (see Section 4.2) and hence this can be excluded. Another source of noise can exist, though, if a fast phosphor is used on the intensifier output.

The frame period with a frame transfer CCD can be split into two sections: (1) integration of a new image in the image zone whilst reading out the previous image from the storage zone and (2) transfer of that new image from the image to storage zones. Typically, a clock speed of $0.5 \,\mu s$ to $2 \,\mu s$ is used for shifting the data between the two zones. In the MIC detector, which has 288 rows on the CCD, the clock period is $0.8 \,\mu s$ giving a transfer period of $230.4 \,\mu s$. Events that arrive during this period (as no shutter is employed) result in either:

(i) Detection in an incorrect vertical position if the phosphor decay is fast enough to allow sufficient charge to be accumulated in $0.8 \,\mu s$ for the events to be recognized above the noise discrimination threshold, as is the case with a P46 phosphor for example.

(ii) Smearing over many CCD rows if a slow phosphor such as a P20 is employed. Here no incorrectly positioned event detection takes place owing to the smeared charge being below the detection threshold. However, correct detection may take place as, owing to



Figure 1. Schematic diagram of the MIC detector. Individual photoelectrons are initally amplified by a high-gain image intensifier producing a scintillation on the phosphor screen anode. Each scintillation is then captured by a CCD camera, centroided within the event recognition electronics to recover as far as possible the resolution loss in the intensifier, and stored in memory.

the long decay, sufficient charge can be accumulated in the imaging period following the transfer.

With a fast phosphor, incorrectly positioned events can be treated as noise. The level will be dependent upon the input light level and the frame period of the CCD. As an example, for *XMM* a subset of 256 of the 288 CCD rows are used for data accumulation with a frame period of 10.8 ms. Here 0.2304/10.8 = 2.1 per cent of events detected would be treated as noise.

3.2 The image intensifier

The image intensifier has a number of noise sources:

(i) Thermal emission from the photocathode. Thermally generated electrons will look identical to photoelectrons and cannot be recognized as a separate component in an integrated image.

(ii) Fluorescence within the input window. With quartz windows this has not been measurable. However, alternative windows such as fibre optic or sapphire can show appreciable fluorescence owing to radioactive decay of impurities.

(iii) Cosmic rays. These can impact with either the window creating a track of photons that are detected by the cathode, or MCPs creating a track of electrons. For the MCPs, only MCP1 is of concern as noise produced in later stages will not have sufficient gain to be recognized as events. Additionally, cosmic rays can impact directly on the CCD but the accumulated charge is too low for events to be recognized.

(iv) Radioactive decay in the lead-glass MCP substrate. This leads to a very low noise contribution from MCP1 that can be measured with the cathode switched off. Noise generated in later stages will not have sufficient gain to be recognized as events.

The dominant mechanism in the highest performance intensifiers is the thermal noise from the cathode which is dependent upon the red sensitivity – a higher red response being associated with a lower band gap and hence higher probability of thermal release. For an S20 photocathode the typical dark count at room temperature is ~50 count cm⁻² s⁻¹. This is equivalent to ~50×10⁻⁶ count pixel⁻¹ s⁻¹ with a 10 µm pixel size. With a bi-alkali cathode, this background reduces to ~10 count cm⁻² s⁻¹ owing to the lower red response. In both cases the background can be reduced by cooling if required, following the classical cooling curve for semiconductors that gives rise to a factor of 2 reduction for each ~7°C drop in temperature.



Figure 2. Schematic of the functioning of the Data Analysis Array. The left group of figures represent a section of a CCD frame with a single event captured. Overlaid is the DAA which, in synchronism with the read out of the CCD, shifts pixel by pixel (top row) and row by row (top to bottom row) through each frame. The right group of figures represent typical digitized CCD data in the DAA for each position. The central data set in the bottom row is recognized as a valid event.

4 COINCIDENCE BETWEEN EVENTS

The bright limit on dynamic range in photon-counting detectors is governed by spatial coincidence losses where two or more events detected by the intensifier photocathode within the temporal resolution of the detector lead to a loss in detective quantum efficiency (DQE). For CCD capture detectors a number of factors are associated:

(i) The frame rate of the camera. The frame period is typically in the range 1-12 ms, this being the limiting temporal resolution.

(ii) The size of an event as captured by the CCD. This is dependent upon the scintillation size on the intensifier output and the demagnification to the CCD.

(iii) Decay time of the phosphor on the intensifier output.

(iv) Pore paralysis in the MCPs. The electron-amplification process strips charge from the associated MCP pores and a finite time, dependent upon the gain used, is required for replenishment. Laboratory measurements have shown that the recharge time when operating at the MIC electron gain of $\sim 5 \times 10^5$ is typically in the range 0.3–1 ms.

Coincidence between events can then be split into two categories: global and point-source coincidence.

4.1 Global coincidences

Global coincidences are related to diffused input sources such as a flat field. As CCDs are pixellated, in principle any two events that are spatially separated within a CCD frame, i.e. are captured in different CCD pixels, can be recognized. However, there is a localized limitation to this which is governed by the eventrecognition electronics within the detector.

With the MIC detector a 3×3 Data Analysis Array (DAA) is, in effect, stepped pixel by pixel and row by row through each frame of CCD data to allow recognition of events. All events in each frame will, at some point, be then centred in this Array. A typical situation is shown in Fig. 2.

Two variants on the electronic design are then available for event recognition:

(i) An event is recognized as being central in the DAA when the conditions B2 > B1, B2 > A2, $B2 \ge B3$, $B2 \ge C2$, B2 > threshold are met. The first four of these define B2 as having the highest value. The threshold is then included to reject noise and is typically placed in the valley of the intensifier PHD (see Fig. 3). When all the conditions are met it is taken that a valid



Figure 3. Typical pulse height and energy distributions after capture by the CCD.



Figure 4. (a) Unresolved coincidence. Events are centred in adjacent pixels and are counted as a single event. (b) Resolved coincidence. There are two peaks in the signal output from the CCD camera enabling recognition of the individual events.

event is centred in the DAA and is subsequently counted by the detector system.

(ii) An event is recognized as being central in the DAA when the conditions (A2 + B2 + C2) > (A1 + B1 + C1) (B1 + B2 + B3) > (A1 + A2 + A3) $(A2 + B2 + C2) \ge (A3 + B3 + C3)$ $B2 \ge (C1 + C2 + C3)$ B2 > threshold are met. This is similar in form to Option 1 but utilizes the complete data set associated with each event instead of a cross-hair through the data set.

Option 2 provides the highest centroiding accuracy (Michel, Fordham & Kawakami 1997) whilst Option 1 maximizes the bright limit on the dynamic-range curve.

The event recognition circuit then places constraints on the individual recognition of events in close proximity to each other. For example, when events are centred in both A2 and B2 they will just be recognized as a single event. Thus, in general, for global coincidence with a CCD capture detector there is a 'coincidence area' surrounding each detected event, this being dependent upon the size of an event as captured by the CCD and on the criteria used for event detection. For an event size that is equal at FWHM to a CCD pixel, as is typically the case with the MIC detector, examples of the effect of the coincidence area are shown in Fig. 4.

In principle, the possibility exists for using event energy to recognize that two events are present in the unresolved coincidence case, summing together the data in all CCD pixels in the coincidence area and using a threshold to discriminate between a single event and coincident events. In practice, though, this is not possible. Three factors, in particular, provide limitations:

(i) Although the intensifier is quoted as operating in a saturated mode, this is something of a misnomer. Fig. 3 shows typical pulse energy (PED) and pulse height (PHD) distributions for the scintillations on the output phosphor. (ii) Pore paralysis in the MCPs. The second event could (dependent upon its arrival time with respect to the first event) have lower gain owing to charge depletion in the MCP pores.

(iii) Typically P20 phosphors are used in image intensifiers as they have high electron to photon conversion efficiency and are well matched to the quantum efficiency characteristics of a CCD. These phosphors, though, have a relatively long decay curve after excitation by the arrival of an event. For each detected event the charge accumulated in a single CCD frame will be dependent upon the arrival time of the event within that frame leading to broadening of the PED and infilling of the valley in the distribution curve. This effect can be overcome by using a fast phosphor, such as a P46, but higher electron gain in the MCPs will be required as the conversion efficiency is lower and, as detailed in Section 3.1, higher detector noise is associated.

When two events occur within the coincidence area in the same CCD frame, the combined event energies will depend upon the original event energies, which could be any value within the PED (Fig. 3) and on the arrival time of each event within the CCD frame, governing the integrated signal on the individual phosphor decay curves. This results in a broadening of the PED but no defined threshold being available for discrimination.

4.2 Point-source coincidence

Coincidences of this type are related to input features within an image that can be considered as point sources. Examples are individual stars or, in spectroscopy, narrow emission lines. Here each photon from that source is captured at the same position within the CCD array. If two photons are detected within the frame period of the CCD then, on the video signal, they will appear as a 'single' event, the extreme case of an unresolved coincidence. The intensity of that 'event' may be greater than for a single event owing to the summation of the two electron clouds on the output phosphor of the CCD but, as detailed in Section 4.1, no discrimination between the cases can be made.

5 THEORETICAL ANALYSIS OF COINCIDENCE LOSS

In CCD capture photon-counting detectors, there are four DQE loss mechanisms:

 (i) The responsive quantum efficiency (RQE) of the photocathode.

(ii) The counting efficiency of the image intensifier. Single photoelectrons are released from the photocathode and are accelerated towards the MCPs. A probability exists that these will hit the webbing between the pores of the first MCP and not initiate an electron cascade process, this probability being dependent upon the open area ratio – the area of the pores compared with the area of the webbing – of the MCP.

(iii) The arrival time of the event within the CCD frame period.(iv) Coincidence losses.

The first two of these are a constant for any detector. Loss due to the arrival time of an event is related to events occurring in the transfer period from the image to storage area of a frame transfer CCD. The frame period for the CCD is of the form shown in equation (1).

$$FP = (N_{\rm r} \times \tau_{\rm v}) + (N_{\rm r} \times \tau_{\rm v}) + (N_{\rm p} \times N_{\rm rw} \times \tau_{\rm h}), \tag{1}$$

where $(N_r \times \tau_v)$ is the transfer period from image zone to storage zone which equals that from storage zone to readout register, with N_r being the number of CCD rows and τ_v the vertical clock period. $(N_p \times N_{rw} \times \tau_h)$ is the time to readout all rows with N_p being the number of pixels in a row, N_{rw} being the number of rows selected for readout using an available windowing facility, and τ_h being the horizontal clock period.

When using a slow phosphor, some events that occur during the frame transfer period from the image to storage zones of the CCD are acquired as sufficient event energy is accumulated in the CCD frame following arrival. This will be a fixed percentage for any selected CCD frame format, the loss *TP* being governed by equation (2).

$$TP = A \times \lambda_{\rm INT} \times \frac{N_{\rm r} \times \tau_{\rm v}}{FP},\tag{2}$$

where A = a constant, $\lambda_{INT} =$ number of events per frame on intensifier output, $N_r \times \tau_v =$ frame transfer period.

The constant A will be dependent upon the decay characteristics of the output phosphor of the intensifier and will, hence, vary between detector systems. For the MIC detector it has been found that, when operating with the full CCD area, ~65 per cent of events occurring in the transfer period are lost (i.e. A = 0.65) when a P20 output phosphor is incorporated in the intensifier. For fast phosphors A = 0, all events owing to incorrect positioning contributing to the background noise.

Hence, for a fixed CCD frame period, the only variable that affects DQE is coincidence loss and this will be dependent upon the input flux rate. Two types of coincidence, global and point source, have previously been defined. The theoretical expected count per pixel, allowing for coincidence loss of either type, is derived in the Appendix.

5.1 Global coincidence

For a flat field or diffuse source input, the theoretical expected number of events counted in any pixel is governed by equation (3).

$$C_{\rm FF} = \frac{1 - e^{-n\lambda_{\rm F}/N}}{n},\tag{3}$$

where

N = area in CCD pixels that is illuminated,

n = coincidence area in CCD pixels over which an unresolved coincidence can take place,



Figure 5. Theoretical coincidence loss on a flat field for a CCD capture photon-counting detector with a CCD format of 384×288 pixels and a

frame period of 12.12 ms.



Captured by CCD (Photons/point source/sec)

Figure 6. Theoretical fit to point-source data acquired with the MIC detector for 3 CCD frame rates.

 $\lambda_{\rm P}$ = no. of events/illuminated area/CCD frame, which is dependent upon *TP*: $\lambda_{\rm P} = (\lambda_{\rm INT} - TP)$.

Theoretical dynamic-range curves with varying coincidence area (*n*) for the MIC detector with a CCD format of 384×288 pixels and a frame period of 12.12 ms are shown in Fig. 5 based upon a uniform flat-field illumination being applied. The accumulated signal per second is ($N \times C_{FF} \div FP$). Here, the importance of the event size can be seen. With n = 15, 10 per cent coincidence loss occurs at 120 000 events per second over the area of the detector. At n = 11 this increases to 200 000 and at n = 7 increases to 300 000.

5.2 Point-source coincidence

For CCD-based detectors, all point sources within an image (for example, a star field) can be treated independently, as long as they are separated by a distance greater than the coincidence area. For each point source, it can be assumed that all coincidences are directly on top of each other with minimal blurring of a single event profile and equation (3) simplifies to equation (4), with n and N both being equal.

$$C_{\rm PS} = 1 - e^{-\lambda_{\rm P}}.\tag{4}$$

Theoretical dynamic-range curves, where the accumulated signal is $FP \times C_{PS}$, are overlaid on acquired data from the MIC detector in Fig. 6.



Captured by CCD (Photons/pixel/sec)

Figure 7. Flat-field data acquired with the MIC detector for 4 CCD frame rates. The best theoretical fit to the 288 CCD row data is n = 13.1. This value was then used as the coincidence area for the other formats.

6 COMPARISON BETWEEN THEORETICAL AND MEASURED COINCIDENCE LOSS

To compare the theoretical curves with acquired data, a series of dynamic-range experiments were carried out with the MIC detector. For these, a beta light was used for illumination, providing a constant intensity source, the input to the detector being varied by using ND filters. Both flat-field and point-source dynamic-range curves were obtained for different CCD frame periods. The period was varied by controlling the number of rows read out from the CCD. The 12.12-ms period is associated with reading out the full 288 rows, the 5.69-ms, 3.11-ms and 1.18-ms periods are associated with reading out 129, 64 and 16 rows, respectively. The point-source data was acquired by illuminating the detector through a 20-µm diameter pinhole.

The results are shown in Figs 6 and 7. Overlaid are theoretical curves where, for the flat-field cases, the coincidence area is taken as 13.1 pixel, this being the best fitting to the 12.12-ms data.

7 DISCUSSION

The dynamic-range curves for photon-counting detectors are inherently non-linear owing to coincidence losses. When comparing detectors, the standard figure of merit employed is the 10 per cent coincidence loss point as shown on Fig. 5. Of importance to note is that detectors of this type have been developed for lowlight-level applications where coincidences will be minimal and errors will be dominated by shot and/or thermal noise. However, by prior characterization of the detector, a viable dynamic-range curve can be produced which, for applications where the input conditions are stable, will allow accurate calculation of the input light level from the detected signal at relatively high flux levels. This is particularly important for space applications such as XMM. For ground-based observations, seeing variations owing to atmospheric turbulance will lead to variations in the size of features within an input image resulting in distortions to the dynamicrange curve that could lead to significant errors in deduction of the input signal.

Here, it has been shown that from experimental data a value for the effective coincidence area can be derived for a detector that

Intensified CCD photon-counting detectors 87

can then be used for providing a theoretical dynamic-range curve for both point and diffused sources. It is important to note that for diffused sources, equation (3) is an approximation, errors arising at the highest input light levels. These result from multiple (>2 events) coincidences. As an example, a third event in a CCD frame can create a bridge between two resolved events leading to coincidence loss but with an effective coincidence area greater than that theoretically derived.

From equation (3) it can be seen that it is essential that the event size as captured by the CCD be as small as possible to minimize the effective coincidence area, this being highlighted by the results shown in Fig. 5. However, the centroiding employed in the detectors for enhancing spatial resolution requires a reasonable event width to maximize centroiding accuracy. Thus a compromise between maximizing flat-field dynamic range and minimizing centroiding errors has to be found. An analysis of centroiding errors (Michel et al. 1997) has shown that an event width of 1-CCD-pixel FWHM is a minimum acceptable limit. This is equivalent to 'n' in equation (3) being equal to ~ 8 , giving a 10 per cent coincidence loss of 250 000 events detector area⁻¹ s⁻¹ using the full area of the MIC detector, a significant increase on the 140 000 figure related to the detector measurements given in Fig. 7.

For the point-source dynamic range, the bright limit can be improved by replacing the CCD by a detector such as a Charge Integration Device (CID) where each pixel is individually addressable. This allows read out of only the pixels associated with the point source(s) in an image and hence maximizes the frame rate (Morrissey et al. 1998). However, a fundamental limit is reached that is related to the recharge time required for depleted MCP pores. If this is assumed to be 0.3 ms then the limit, applying equation (4), for 10 per cent coincidence loss will be ~600 event s⁻². Low resistance 'hot' MCPs are available that will provide an improvement but these do have a limited lifetime.

Noise in most detectors of this type is minimal, being dominated by thermally emitted electrons from the photocathode. However, as discussed in Section 3.1, if a fast phosphor is employed in the image intensifier this can create an appreciable noise source that can contribute to coincidence? One event instead of two is added to the data accumulated. Hence, in effect, the input event is acquired and the noise event lost, leading to an effective lowering of the noise. During data reduction, the noise must be accounted for in the quantification of features within an image and the effect of coincidence on this noise level known. The expected count per pixel with input rates of λ_s and λ_n signal and noise events per pixel, respectively, is given by equation (5).

$$C_{\rm SN} = \frac{1 - e^{-(\lambda_{\rm s} + \lambda_{\rm n})}}{n}.$$
(5)

The localized background noise must then be quantified using this equation prior to data reduction.

ACKNOWLEDGMENTS

We are very grateful to Hajime Kawakami at the Mullard Space Science Laboratory and Peter Sandford in the Department of Physics & Astronomy at UCL for their suggestions and comments. CFM was in receipt of a PPARC studentship during the period of this work.

REFERENCES

- Fordham J. L. A., Bone D. A., Oldfield M. K., Bellis J. G., Norton T. J., 1992, in Guyenne T. D., Hunt J. J., eds, ESA SP-356, Photon Detectors for Space Instrumentation. ESA Publications, Noordwijk, p. 103
- Horner S. D. et al., 1994, in Crawford D. L., Craine E. R., eds, Proc. SPIE Vol. 2198, Instrumentation in Astronomy VIII, p. 1238
- Michel R., Fordham J., Kawakami H., 1997, MNRAS, 292, 611
- Morrissey P. F., Norton T. J., Kimble R. A., 1998, in Bely P. Y., Breckenridge J. B., eds, Proc. SPIE Vol. 3356, Space Telescopes and Instruments V, p. 1036
- Norton T. J., Airey R. W., Morgan B. L., Fordham J. L. A., Bone D. A., 1991, in Morgan B. L., ed., Proc. 10th McGee Symp., Photoelectronic Imaging Devices, IoP Conf. Ser. No. 121. IoP Publishing, Bristol, p. 97

APPENDIX A: DERIVATION OF EQUATION (3)

Consider the locations of events within an illuminated area of N pixels in a given time frame. A natural (and reasonable) probability model is that these locations form a spatial Poisson process with intensity λ per pixel area. The number of events occurring at any given pixel then has a Poisson distribution with mean λ , independently of events at other pixels. In particular, if θ is the probability that one or more events occur at a given pixel, then

 $\theta = 1 - \mathrm{e}^{-\lambda}$.

Equation (3) gives a formula for the probability that 1 count will be accumulated by the computer at any given pixel. We will derive this by two different arguments.

For any given pixel, consider a coincidence area of n pixels centred on it, such that an event occurring outside this area cannot result in an unresolved coincidence with one at the central pixel (e.g. as in Fig. 2, where n = 9). 1 count will be accumulated if one or more events occur at the central pixel *and* one or more events

occur in each of r of the other pixels, but the data signal is highest in the central one, for r = 0, 1, ..., n - 1.

Note that if one or more events occur at each of r + 1 pixels, then each of these pixels has an equal chance 1/(r + 1) of recording the highest data signal. Hence the probability that the CCD counter will count 1 ($P_{ct=1}$) is:

$$P_{ct=1} = \theta \sum_{r=0}^{n-1} \frac{1}{r+1} \frac{(n-1)!}{r!(n-1-r)!} \theta^r (1-\theta)^{n-1-r}$$

= $\frac{1}{n} \sum_{r+1=1}^n \frac{n!}{(r+1)!(n-r-1)!} \theta^{r+1} (1-\theta)^{n-r-1}$
= $\frac{1}{n} [1-(1-\theta)^n]$
= $(1-e^{-n\lambda})/n.$

This is equivalent to equation (3) with $\lambda = \lambda_P/N$. Furthermore, this applies to any pixel, so the expected count for a frame of area N pixels is $N(1 - e^{-n\lambda})/n$.

In the above argument, the coincidence area *n* is an integer number of pixels, although this may in principle be relaxed by redefining a pixel area. The same expected count may be obtained by the following cruder argument that does not attempt to mimic the counting process (and that requires no restriction on *n*). Consider an arbitrary region of area *n* and suppose that 1 count will be accumulated by the computer if one or more events occur in this region. The probability of this is $1 - e^{-\lambda n}$, which is thus the expected count in an area *n*. By symmetry, the expected count per pixel is therefore $(1 - e^{-n\lambda})/n$, and the expected count for the whole area *N* is $N(1 - e^{-n\lambda})/n$, as before.

In each of these arguments, the coincidence area n is assumed not to depend on how many events coincide, which may not strictly be true in practice. However, by treating n as an adjustable parameter, a suitable 'average' value is used and the resulting loss formula agrees closely with observation.

This paper has been typeset from a TEX/LATEX file prepared by the author.