Statistics of electron-multiplying charge-coupled devices

Brian M. Sutin*
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, United States

Abstract. Electron-multiplying charge-coupled devices are efficient imaging devices for low-surface-brightness ultraviolet astronomy from space. The large amplification allows photon counting (PC), the detection of events versus nonevents. This paper provides the statistics of the observation process, the photon-counting process, the amplification process, and the compression. The expression for the signal-to-noise of PC is written in terms of the polygamma function. The optimal exposure time is a function of the clock-induced charge. The exact distribution of amplification process is a simple-to-compute powered matrix. The optimal cutoff for comparing to the read noise is close to a strong function of the read noise and a weak function of the electron-multiplying gain and photon rate. A formula gives the expected compression rate.
 [DOI: 10.1117/1.JATIS.9.2.028001]

Keywords: electron multiplying charge-coupled devices; photon counting; data analysis and techniques; clock-induced charge; ultraviolet instrumentation; read noise.

Paper 23015G received Jan. 28, 2023; accepted for publication Mar. 10, 2023; published online Apr. 6, 2023.

1 Science Motivation

The Earth's atmosphere is essentially opaque to ultraviolet (UV) radiation at wavelengths shorter than about 300 nm. All sorts of interesting astrophysical phenomena exist to be imaged in the UV, such as emission lines from interstellar gas, redshifted Ly-α from galaxies, and zodiacal light caused by Rayleigh scatter of sunlight. Most of the UV sources mentioned above have very low-UV surface brightness, so observation requires an imaging telescope above the Earth's atmosphere that integrates for long periods, coupled with an imaging detector with very high efficiency. Typical photon rates are on the order of a photon per 1000 pixels per second. As space-based missions are limited in the returned data, compression of data consisting of non-detections is relevant.

Electron-multiplying charge-coupled devices (EMCCDs) are the enabling technology for space-based UV imaging, especially with enhanced UV quantum efficiency. The Teledyne e2v CCD201-20 EMCCD has flight heritage from the Faint Intergalactic Redshifted Emission Balloon.[1–3] Roman coronagraphic instrument has extensively tested and qualified the CCD201-20.[4] Other space applications using the CCD201-20 are SPARCS,[5–7] currently in production, and the Polarized Zodiacal Light Experiment concept.[8]

2 What is an EMCCD?

The EMCCD is a CCD modified to achieve high signal-to-noise ratio (SNR) by rendering the read noise effectively zero. Compared to conventional CCDs, EMCCDs have an additional serial register (604 extra charge-coupled "pixels" [see Fig. 1]), where one of the register clocks is replaced by a high-voltage clock (25 to 50 V). The higher voltage causes a multiplication process that stochastically turns one electron into many, resulting in thousands of electrons at the output amplifier. This allows detection of single-photon events by thresholding above amplifier read noise.[10]

In the amplification process, if the charge transfer from one register to the next is done with enough input energy, a signal electron will knock lose another electron. This is the avalanche

*Address all correspondence to Brian M. Sutin, brian.m.sutin@jpl.nasa.gov

photodiode (APD) effect. Repeating this process makes an "APD staircase,"[11] magnifying the signal by having some extra CCD-like charge transfers explicitly run using overdrive. Since the digitization takes place after magnification, the read noise is proportionally smaller when compared to the original, unamplified signal.

EMCCDs are generally run in "photon counting" (PC) mode, meaning that the detector is read out often enough that the expected signal counts in each pixel are $\ll 1$. After magnification, the final counts are generally either very large or very small, signifying detecting or not detecting signal photons. This is similar to a photomultiplier tube (PMT). The downside, as compared to a PMT or an intensified CCD, is that the dark current and clock-induced charge (CIC) may also add significant noise. The EMCCD data output is then a combination of science signal, dark current, CIC, and readout noise. Dark current decreases with temperature, but sufficient cooling is not always available to make the dark current negligible. CIC is fixed per exposure and is discussed in Sec. 6. Without PC, sufficient EM amplification will make the read noise comparatively small, but the discussion in Sec. 11 shows that in PC mode, the read noise is eliminated with the proper choice of parameters.

3 Probability Notation

The rest of this paper mainly consists of probability theory related to processing and understanding the signal from these detectors. As the intended target of this paper is engineers and scientists rather than mathematicians, neither mathematical rigor nor mathematically rigorous notation are prioritized. The notation used is as follows:

X, X_Y: a random variable; i.e., a variable that can take on randomly generated values

$P(X)$: a probability distribution

$E(X)$: the expected value of X, also known as the first moment or mean

$V(X)$: the variance of X, equal to the second moment of $X - E(X)$

The various symbols used are, in order of definition:

λ: the number of counts expected in a pixel's well for a given exposure [Eq. (1)]

$Q(n, \lambda)$: regularized gamma function [Eq. (3)]

ε: probability of one or more signal counts in a pixel's well [Eq. (6)]

N: number of exposures in an observation for a single pointing [Eq. (7)]

PC_N: number of exposures out of N that have one or more counts [Eq. (7)]

c: an observed value of PC_N [Eq. (7)]

$B(\alpha, \beta)$: Beta function with parameters α and β [Eq. (9)]

$M_n(\lambda)$: n'th moment of the probability distribution of λ [Eq. (11)]

ψ, ψ_1: digamma and trigamma functions [Eqs. (13) and (14)]

T: total exposure time in seconds for series of exposures [Eq. (15)]

λ_{CIC}: CIC count rate in counts per exposure [Eq. (15)]

η: photon count rate from the signal source in counts per second [Eq. (15)]

$\bar{\varepsilon}$: expected value of ε [Eq. (17)]

λ_{opt}: the counts per exposure that maximize SNR [Eq. (19)]

$\hat{\mu}$: read out noise when not using PC [Eq. (21)]

$P_m(n)$: the probability of n counts after m stages of amplification [Eq. (25)]

B_{nk}: a matrix containing the probability of k counts multiplying to n counts [Eq. (27)]

G: amplification gain of a set of multiplication elements [Eq. (30)]

μ: read out noise when using PC [Eq. (33)]

τ: cutoff criterion for amplified counts to assuming a detection [Eq. (36)]

$\hat{\tau}$: optimal value for photon-counting cutoff τ [Eq. (41)]

4 Poisson Distributions

For independent events arriving at a detector, the Poisson distribution is appropriate to use. For a Poisson-distributed random variable X, $E(X) = V(X) = \lambda$, and X is distributed as

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}. \tag{1}$$

Similar to the Gaussian assumption, the SNR goes as the square root of the counts as a result of $E(X) = V(X)$. For no events,

$$\begin{aligned} P(X = 0) &= e^{-\lambda}, \\ P(X > 0) &= 1 - e^{-\lambda}. \end{aligned} \tag{2}$$

Useful is the cumulative distribution function, which is

$$F_X(n) = P(X \le n) = \sum_{k=0}^{n} \frac{\lambda^k e^{-\lambda}}{k!} = Q(n+1, \lambda). \tag{3}$$

Here $Q(n, \lambda)$ is the regularized gamma function; e.g., evaluated in Excel as

$$\mathrm{Q}(n, \lambda) = 1 - \mathrm{GAMMADIST}(\lambda, n, 1, \mathrm{true}) \tag{4}$$

5 Photon Counting

Consider an ideal photon-counting detector with infinite gain. If any photons at all are detected, then the detector returns 1, otherwise 0. Let the random variable X_W be the number of electrons in the well of a pixel. Using Eq. (2), the expected value (and second and all higher moments) of this process is then

$$\begin{aligned} E(\mathrm{Ideal\,PC}) &= \sum_{w=1}^{\infty} (1) P(X_W = w) \\ &= P(X_W > 0) \\ &= 1 - e^{-\lambda}. \end{aligned} \tag{5}$$

In practice, a single exposure is not useful. Rather, λ is estimated by taking a sequence of N exposures, each in PC mode. Define the ε as the probability of detection

$$\varepsilon = 1 - e^{-\lambda}. \tag{6}$$

The number of exposures with a signal detection are then given by the binomial distribution

$$P(\mathrm{PC}_N = c|\varepsilon) = \binom{N}{c} \varepsilon^c (1-\varepsilon)^{N-c}. \tag{7}$$

Here c is the total number of counts for the N exposures. If ε is ½, then this is the distribution of the number of heads after N flips of a fair coin. What we want to know is ε, which is on the wrong side of the equation. Using Bayes' theorem,

$$P(\varepsilon|\mathrm{PC}_N = c) = \frac{P(\mathrm{PC}_N = c|\varepsilon) P(\varepsilon)}{P(\mathrm{PC}_N = c)}. \tag{8}$$

We need a prior for $P(\varepsilon)$ and select the most commonly used prior for the binomial distribution, the beta distribution, given by

$$P(\varepsilon) = \frac{\varepsilon^{\alpha-1}(1-\varepsilon)^{\beta-1}}{B(\alpha, \beta)} \tag{9}$$

The denominator is the beta function $B(\alpha,\beta)=\Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$, and the distribution is valid over $\{\alpha>0,\beta>0\}$. A common choice is the uniform prior $\{\alpha=1,\beta=1\}$. For a reader unfamiliar with Bayesian statistics, the prior can be used to bias the result using prior knowledge (thus the name). However, the actual choice of prior makes little difference with sufficient data. More useful is to reverse this last statement; if the result of an observation depends significantly on the choice of prior, then the data quantity is insufficiently robust. Using the above prior,

$$P(\varepsilon|\mathrm{PC}_N=c)=\frac{\varepsilon^{c+\alpha-1}(1-\varepsilon)^{N-c+\beta-1}}{B(c+\alpha,N-c+\beta)}. \tag{10}$$

We did not need to explicitly calculate the Bayesian denominator $P(\mathrm{PC}_N=c)$. The denominator is determined by noticing that the binomial distribution combined with a beta distribution is another beta distribution, so the new denominator is the normalization factor from the new beta distribution.

The expression for the moments of λ is found using the probability distribution for ε to find the powers of λ written as a function of ε

$$M_n(\lambda)=\int_{\varepsilon=0}^{1}(-\ln(1-\varepsilon))^n\frac{\varepsilon^{c+\alpha-1}(1-\varepsilon)^{N-c+\beta-1}}{B(c+\alpha,N-c+\beta)}\mathrm{d}\varepsilon \tag{11}$$

A change of variable gives, to within a sign, the formula for the geometric moments of the beta distribution

$$M_n(\lambda)=\int_{x=0}^{1}(-\ln x)^n\frac{x^{N-c+\beta-1}(1-x)^{c+\alpha-1}}{B(N-c+\beta,c+\alpha)}\mathrm{d}x. \tag{12}$$

So, the mean and variance of λ are (within a sign) the geometric mean and geometric variance of a beta distribution, respectively. The expected value can be written in terms of the digamma function ψ

$$E(\lambda)=-M_1(\lambda)=\psi(N+\alpha+\beta)-\psi(N-c+\beta), \tag{13}$$

and the variance in terms of the trigamma function ψ_1

$$V(\lambda)=M_2(\lambda)-M_2^2(\lambda)=\psi_1(N-c+\beta)-\psi_1(N+\alpha+\beta). \tag{14}$$

The polygama function ψ_n is a transcendental function available in language math libraries or evaluated with a series approximation.

6 Optimal Exposure Time

λ, the expected number of counts in an exposure, depends on the exposure time, while the relevant quantities for an observation are the rate of photons per unit time and the total exposure time available. The obvious choice is to have a very large number of exposures and make the exposure time arbitrarily small. However, in practice, an EMCCD has a noise source called CIC that is added to every exposure,[3,12] where an extra charge is added during clocking of the charge-coupled pixels. CIC has been measured as low as 7×10^{-4} e^-/pixel/readout,[13] while achieving is 5×10^{-3} e^-/pixel/readout is relatively easy.

CIC is not constant across the detector. For a 1k × 1k-pixel frame transfer device, the number of pixel-to-pixel transfers can vary from 1k to 3k (frame transfer plus row number plus column number) before arriving at the readout and amplification chain (see Fig. 1).

For this section, we define the following variables:

- T- the total exposure time in seconds for a series of exposures
- λ_{CIC}- the CIC count rate in counts per exposure
- η- the photon count rate from the science signal source in counts per second

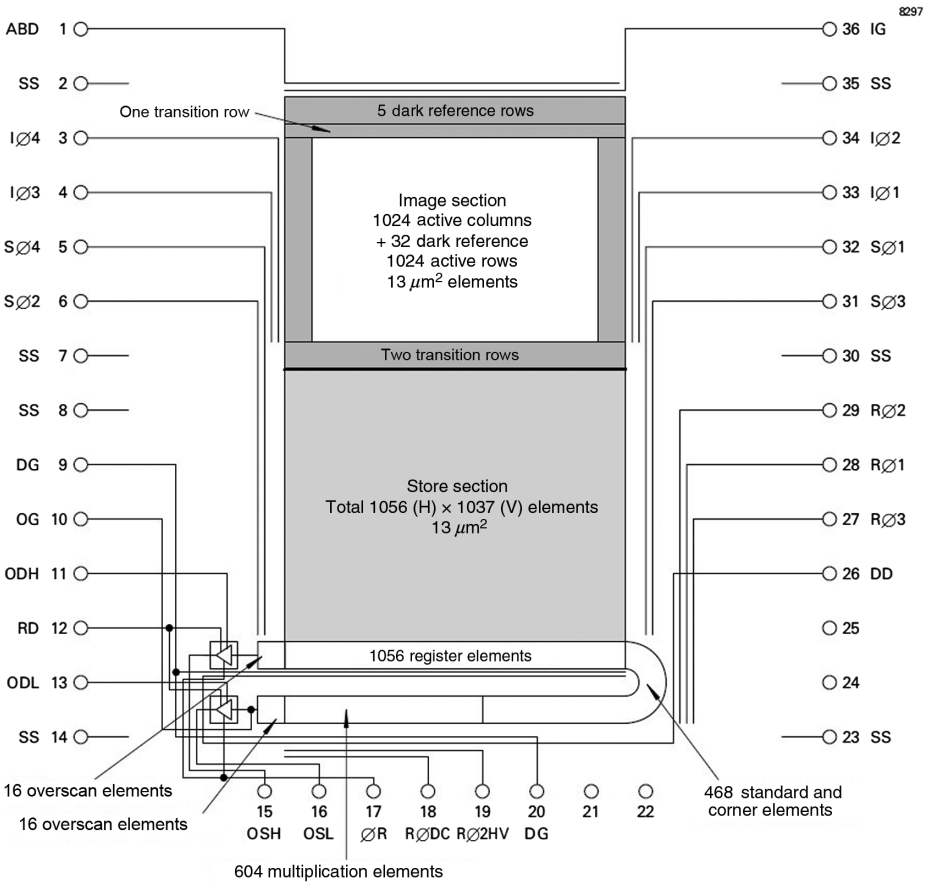

Fig. 1 Device architecture diagram for the e2v CCD201-20 EMCCD.[9] The device has two separate output amplifiers, one of which is amplified with the 604 multiplication elements.

The variable η is the fundamental astronomical parameter being measured, but includes dark current. Our expected counts per exposure λ is then

$$\lambda = \lambda_{\mathrm{CIC}} + \frac{\eta T}{N}. \tag{15}$$

The expected photon-counting event detection rate is

$$\overline{\varepsilon} = 1 - e^{-\lambda_{\mathrm{CIC}} - \eta T/N}. \tag{16}$$

To avoid the difficult polygama function, we will find the optimal number of exposures N by maximizing the SNR with respect to ε rather than λ. This should be acceptable, as CIC is only critical for small λ. Choose the number of exposures N to be $N \gg \alpha$ and $N \gg \beta$ (or choose $\alpha = \beta = 1$). Since ε is beta distributed, we know the mean and variance, and thus the SNR

$$\begin{aligned} E(\overline{\varepsilon}) &= c/N, \\ V(\overline{\varepsilon}) &= \frac{(c/N)(1 - c/N)}{N}, \\ \mathrm{SNR}(\overline{\varepsilon}) &= \left(\frac{c}{1 - c/N}\right)^{1/2}. \end{aligned} \tag{17}$$

The SNR evaluated at the expected $\overline{\varepsilon}$ is then

$$\mathrm{SNR}(\overline{\varepsilon}) = N^{1/2}(e^{\lambda_{\mathrm{CIC}} + \eta T/N} - 1)^{1/2}. \tag{18}$$

Setting the derivative with respect to N to zero gives the transcendental equation

$$(1 + \lambda_{\mathrm{CIC}} - \lambda_{\mathrm{opt}})e^{\lambda_{\mathrm{opt}}} = 1. \tag{19}$$

Over the typical region of $10^{-3} < \lambda_{\mathrm{CIC}} < 5 \times 10^{-3}$, λ_{opt} is well approximated by the fit

$$\lambda_{\mathrm{opt}} \approx \frac{3}{2}\lambda_{\mathrm{CIC}}^{1/2}. \tag{20}$$

So, for CIC ~ 0.001, the optimal number of well counts per exposure is $\lambda_{\mathrm{opt}} \sim 0.045$ counts for best SNR. Given that CIC for current state-of-the-art, EMCCD controllers is in the range of 0.001 to 0.005, λ for optimal SNR is restricted to (0.045 to 0.10). Since CIC varies across the detector, λ_{opt} can never be more than a compromise. Other considerations might lead to using shorter exposure times than for optimal SNR. For example, bright stars in the field can cause "blooming," where the wells of the amplification staircase elements overfill and spill charge into adjacent elements. The exposure time may also be limited by the spacecraft stability for applications where spatial resolution is important.

7 Comparison to Standard Mode

For bright sources, the detector frame rate may not be high enough to achieve the optimal count rate λ per exposure. In this case, observing with standard mode (SM) might provide better SNR. The SNR in SM is

$$\mathrm{SNR(SM)} = \frac{\eta T}{(\eta T + \hat{\mu})^{1/2}}. \tag{21}$$

Here $\hat{\mu}$ is the readout noise for standard mode, which may be different than the readout noise μ for PC mode if the EMCCD device has separate readout amplifiers (see Fig. 1). Given the total number of expected events ηT for an observation and a known λ_{CIC}, the number of PC-mode exposures is

$$N = \eta T/(\lambda - \lambda_{\mathrm{CIC}}). \tag{22}$$

The expected number of PC counts is

$$E(c) = N - Ne^{-\lambda}. \tag{23}$$

Assuming N is large enough that the prior is irrelevant and replacing c with the expected counts, the PC-mode SNR is

$$\mathrm{SNR(PC)} = \frac{\psi(N) - \psi(Ne^{-\lambda})}{(\psi_1(Ne^{-\lambda}) - \psi_1(N))^{1/2}}. \tag{24}$$

Equations (21) and (24) can be compared to choose which mode is best. A worked example is shown in Fig. 2. In practice, the standard-mode exposure time T is limited by cosmic rays, whereas for PC mode the exposure time T/N is limited by readout speed.

8 Amplification

The amplification process of EMCCD amplification is similar to the amplification process in an APD.[11,14–18] The amplification of a single electron from Matsuo 1984[11] (hereafter "Matsuo") is given by a recurrence relation for each stage m by

$$P_m(n) = (1 - Q)P_{m-1}(n) + Q\sum_{k=0}^{n} P_{m-1}(n-k)P_{m-1}(k), \quad P_0(n) = \delta_{1,n}. \tag{25}$$

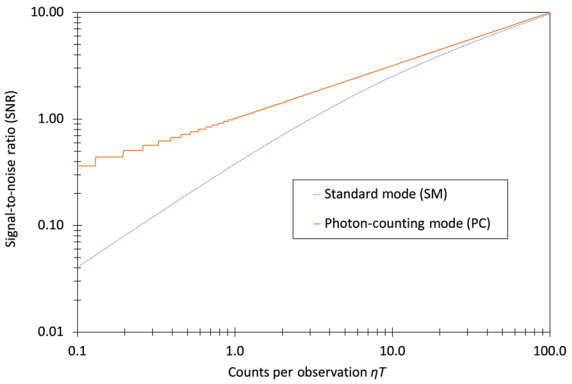

Fig. 2 Assuming $\lambda_{\mathrm{CIC}} = 0.002\mathrm{e}^-$ and $\mu = \hat{\mu} = 6\mathrm{e}^-$ for this plot, PC mode always has better SNR than standard mode. The staircase effect for lower counts in PC mode is real and due to rounding up to the nearest *N* integer number of exposures.

Q is the probability of an electron generating a secondary. The distribution of counts for each stage $P_m(n)$, where n is the number of counts, depends on the distribution of counts at the previous stage. What happens if the initial distribution of counts is not restricted to a single count, or $P_0(n) \neq \delta_{1,n}$? Then the recurrence relation spectacularly fails. However, the most important issue with this recurrence relation is that it physically makes no sense; the relation connects the opposite tails of the distribution to generate the next stage. So, although the relation may hold true, it is more of a mathematical curiosity than anything else.

The APD staircase effect can be described as a process where, at each step, there is a chance when transferring the charge that an electron generates a second electron. Now make two assumptions. First, the electrons are not aware of each other, equivalent to the process at each step being linear in the number of electrons generated on average. Second, assume that an original electron can never generate two or more additional electrons, equivalent to assuming that the energy of an electron, once spent on generating a second electron, is gone. With these two assumptions, we can write down a recurrence relation.

Consider, at some step m in the APD staircase, the distribution of electrons $P_m(n)$. How was this distribution generated if the previous step had k electrons? Clearly, $n-k$ electrons each generated a single electron (first assumption), while $k-(n-k) = 2k-n$ electrons did not. Since the electrons are independent (second assumption), the applicable distribution is the binomial distribution

$$P_m(n) = \sum_{k=n/2}^{n} \binom{k}{n-k} (1-Q)^{2k-n} Q^{n-k} P_{m-1}(k). \tag{26}$$

This expression looks unwieldy; however, the right-hand side is nothing more than a fixed matrix multiplication to get from one stage of the APD to the next. The matrix is

$$B_{nk} = \binom{k}{n-k} (1-Q)^{2k-n} Q^{n-k}, \quad k \in \left[\left\lceil \frac{n}{2} \right\rceil, n\right]. \tag{27}$$

Once the matrix B_{nk} is formed with a size corresponding to the largest electron counts of interest, the recurrence relation becomes

$$P_m(n) = B_{nk} P_{m-1}(k). \tag{28}$$

The final amplified distribution is then

$$P_m(n) = (B_{nk})^m P_0(k). \tag{29}$$

Computing $P_m(n)$ is only a few lines of MATLAB code,

```
% Create the transform matrix for a single stage of a staircase APD
Q = gain^(1/stages) - 1;
for n = (0:nmax);
  for k = (0:nmax);
    if((k <= n) && (n-k <= k))
      B(k+1,n+1) = exp(gammaln(k+1) - gammaln(n-k+1) -
        gammaln(2*k-n+1) + (2*k-n)*log(1-Q) + (n-k)*log(Q));
    end % k
  end % n
% Transform initial distribution to the final distribution
P = Po*BB^stages; % Po and P are row vectors here
```

This recurrence relation gives identical results to Matsuo for an initial distribution of a single electron $P_0(n) = \delta_{1,n}$, while also, unlike Matsuo, working for any initial distribution.

9 Comparison to Basden

Basden 2003[19] (hereafter "Basden") gives the formula for the distribution of final amplified counts X_A as [Basden, Eq. (1)]

$$\tilde{P}(X_A = a|X_W = w) = \frac{a^{w-1}e^{-a/G}}{G^w(w-1)!} \quad a \geq w,\ w > 0,$$
$$\tilde{P}(X_A = 0|X_W = 0) = 1. \tag{30}$$

G here is the gain of the entire APD staircase, so the gain at stage m would be $G_m = (1+Q)^m$.

Basden's derivation is based on assuming the first assumption above (independent electrons) and a few approximations. We write $\tilde{P}$ because the formula is not a probability distribution; the sum over all values is not equal to unity. The formula has the functional form of an Erlang distribution, but the Erlang distribution is a continuous distribution in a. The Erlang distribution is the distribution of the sum of w independent identically distributed (IID) exponentially distributed random variables. This is a hint that the Basden equation could be derived as a sum of IID geometrically distributed random variables. In fact, a plot of the Matsuo distribution for a single initial electron does look very similar to an exponential distribution.

The recurrence relation of Eq. (25) deviates from that of Eq. (26), mostly for larger counts (Fig. 3). Basden explicitly states that their approximation should not be used in this region.

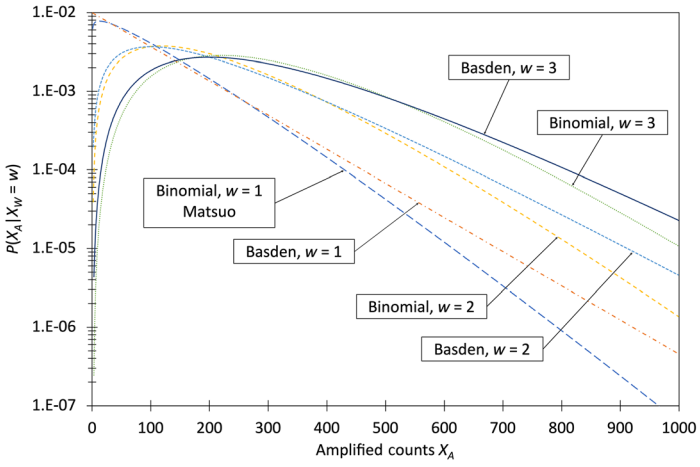

Fig. 3 The Basden approximation is close to the binomial model for moderate counts but fits poorly for the less likely counts. For this plot, the gain is 100, with 50 stages. "Matsuo" is Eq. (25), "Basden" is Eq. (30), and "Binomial" is Eq. (26).

In the limit as the number of stages m becomes large and the amplification Q becomes small for fixed G, a closed-form solution in terms of G based on Eq. (26) should exist. At a minimum, the recurrent relation can be written in terms of the eigenvectors and eigenvalues

$$B_{nk}^{m} = V\Lambda^{m}V^{-1}. \tag{31}$$

Since the matrix B_{nk} is triangular, the diagonal elements are the eigenvalues. Taking the limit as m goes to infinity for fixed G gives the eigenvalues as

$$\Lambda_i = G^{-i}. \tag{32}$$

In the limit, the eigenvectors V are independent of G.

10 Combining Signal with Read Noise

In practice, we add the amplified well counts to the read noise X_R. Read noise is also Poisson distributed, so

$$P(X_R = y) = \frac{\mu^{y}e^{-\mu}}{y!}. \tag{33}$$

The expected value of the read noise μ is trivially measured by taking a zero-exposure-time frame or alternatively running the analog-to-digital converter without advancing the CCD charge. Since the read noise and amplified counts are independent, the distribution of the sum read out by the detector is

$$P(X_A + X_R = c) = \sum_{a=0}^{c} P(X_A = a)P(X_R = c - a). \tag{34}$$

Inserting in the expressions from Eqs. (1), (29), and (33)

$$P(X_A + X_R = c) = \sum_{a=0}^{c} \frac{\mu^{c-a}e^{-\mu}}{(c-a)!} B_{ak}^{m} \frac{\lambda^{k}e^{-\lambda}}{k!}. \tag{35}$$

11 PC Cutoff Setting

The algorithm used for PC is that c is considered a single count if higher than some threshold τ. The expected value (and second and all higher moments) of this process is then

$$E_{\tau}(\mathrm{PC}) = \sum_{c=\tau+1}^{\infty} (1)P(X_A + X_R = c). \tag{36}$$

Since the argument of the summation is a probability distribution over c, this is equal to

$$E_{\tau}(\mathrm{PC}) = 1 - \sum_{c=0}^{\tau} P(X_A + X_R = c). \tag{37}$$

Substituting in Eq. (35) and swapping the summations gives

$$E_{\tau}(\mathrm{PC}) = 1 - e^{-\lambda} \sum_{a=1}^{\tau} \sum_{m=1}^{\tau-a} \frac{\mu^{m}e^{-\mu}}{m!} B_{ak}^{m} \frac{\lambda^{k}}{k!}. \tag{38}$$

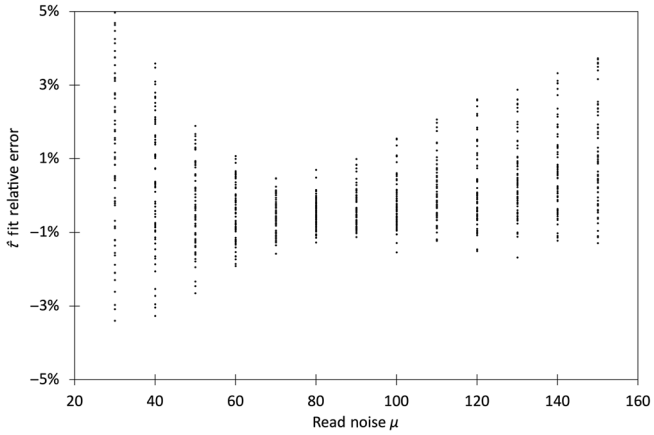

Fig. 4 The residuals of the Eq. (42) fit to the cutoff τ in Eq. (41) are within 2%. The scatter is due to variations in λ over 0.04 to 0.1 and G over 10 to 1000.

The sum over m is again the regularized gamma function from Eq. (3)

$$E_\tau(\mathrm{PC}) = 1 - e^{-\lambda} \sum_{a=0}^{\tau} \mathrm{Q}(\tau - a + 1, \mu) B_{ak}^{m} \frac{\lambda^k}{k!}. \tag{39}$$

The optimal choice for the cutoff parameter τ is such that the read noise does not bias the PC process. This is equivalent to setting τ such that false positives equal false negatives. With this choice, lost counts from low-amplification signal counts balances added counts from high noise. The PC method is then an "ideal" photon counter, as in Sec. 5. One can prove that setting false positives to false negatives is equivalent to the read noise having no effect, and so the expected value of ε with noise is equal to ε without noise, or

$$E_\tau(\mathrm{PC}) \cong \varepsilon. \tag{40}$$

Let $\hat{\tau}$ be the optimal τ that makes Eq. (39) satisfied. Then

$$\sum_{a=0}^{\hat{\tau}} \mathrm{Q}(\hat{\tau} - a + 1, \mu) B_{ak}^{m} \frac{\lambda^k}{k!} = 1. \tag{41}$$

Over the range of typical values for λ, μ, and G, the optimal $\hat{\tau}$ is well fit to within 5% by

$$\hat{\tau} \approx 1.72 \mu^{0.88} (G/\lambda)^{0.03}. \tag{42}$$

$\hat{\tau}$ only has a weak dependence on λ and G. The residuals of the fit are due to the μ dependence not being sufficiently well-fit with a power law. The residuals are shown in Fig. 4.

12 Compression of Photon-Counting Mode Data

For a typical photon-counting exposure, the expected image data are a list of binary values, the vast majority of which are zeros. Since CIC may be a significant contributor to the signal and is unstructured random noise, compression algorithms based on entropy (structure in the image) are likely not optimal. On the other hand, an algorithm that depends on most of the data being zero might be.

If the events in each pixel are independent, then the statistics are Poisson distributed with a "rate" or expected number of pixels between events, of λ. For a detector where the noise is CIC-dominated and CIC is ~0.001 e^-/pix/frame, the number of pixels between each photon detection is about $1/(0.001) = 1000$.

The simplest compression method for sparse binary data is to count zeros between events. This nominally gives a compression ratio of approximately the number of pixels divided by the number of bits needed to encode the number, or $\sim(1/\lambda)/\log_2(1/\lambda)$. A compression ratio roughly on the order of 100:1 should be possible for the above CIC alone.

This algorithm is a variation of run-length encoding (RLE), which is to count the number of bits before a change. The heritage of the RLE algorithm stretches back to at least the early 1960s.[20] For mostly empty data, the appropriate standard compression algorithm is the run-length limited (RLL) variation, which has been common for encoding hard disks and optical storage since the 1980s, and only counts zeros. We use $(0, b)$ RLL.

The encoding algorithm for RLL$(0, b)$ is as follows:

1. Count number of zeros until a nonzero.
2. For an b-bit encoding, do not count more than $2^b - 2$ zeros.
3. If no nonzero encountered, then return $2^b - 1$.
4. Else, return number of zeros.

For the data generated by the photon-counting process, we can compute the expected compression ratio. Equation (1) gives the probability for an event (1) and a nonevent (0) for an image pixel. Assuming independent events, the probability of n zeros followed by a nonzero is

$$P(X_{\leq n} = 0, X_{n+1} = 1) = e^{-n\lambda}(1 - e^{-\lambda}). \tag{43}$$

The expected number of inputs bits consumed by a b-bit encoding is then the probability of all zeros (consuming one encoding value), or some run of zeros followed by a nonzero

$$E(\text{bits consumed}) = 2^b P(X_{\leq 2^b} = 0) + \sum_{n=0}^{2^b-1}(n+1)P(X_{\leq n} = 0, X_{n+1} = 1). \tag{44}$$

Simplifying,

$$E(\text{bits consumed}) = 2^b e^{-2^b\lambda} + (1 - e^{-\lambda})\sum_{n=0}^{2^b-1}(n+1)e^{-n\lambda}. \tag{45}$$

Table 1 Optimal code length for given counts per exposure λ.

Code length	λ_{lo}	λ_{hi}
2	0.2272	0.4114
3	0.1636	0.2272
4	0.0955	0.1636
5	0.0531	0.0955
6	0.0288	0.0531
7	0.0154	0.0288
8	0.0082	0.0154
9	0.0043	0.0082
10	0.0023	0.0043
11	0.0012	0.0023
12	0.0006	0.0012

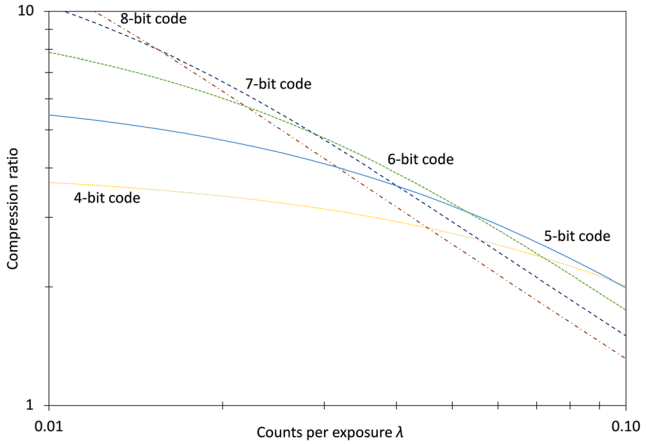

Fig. 5 An encoding with 5-bit codes has the highest compression ratio for $0.053 < \lambda < 0.095$, which encompasses most of the range of λ for optimal SNR from Sec. 6. The most likely lossless compression assuming optimal SNR is thus from 2:1 to 3:1. The optimal compression ratio for $\lambda < 0.01$ is well fit by the approximation $0.264 \times \lambda^{-0.825}$.

Using the derivative of the formula for the sum of a geometric series, this simplifies to

$$E(\text{bits consumed}) = 1 + 2^b e^{-2^b\lambda} - 2^b e^{-(2^b-1)\lambda} + \frac{e^{-\lambda} - e^{-2^b\lambda}}{(1 - e^{-\lambda})}. \tag{46}$$

The final compression ratio is then $E(\text{bits consumed})/b$. The boundaries in λ for the ranges of optimal compression are found by looking for the λ where compression (λ, b) = compression $(\lambda, b+1)$. The results are shown in Table 1.

13 Conclusions

The following steps specify the operating mode (PC or SM), expected SNR, optimal exposure time, and data volume. For a space-based astronomical observatory, knowing these parameters is critical to mission design. The steps to set up an EMCCD for an observation are

1. Measure the CIC λ_{CIC} by amplifying short dark frames.
2. Use Eqs. (19) or (20) to find the target expected counts per exposure λ.
3. Choose a target SNR from science requirements.
4. Use Eqs. (22) and (24) to find the number of exposures N required.
5. Estimate the photon rate η (from radiometry).
6. Check if dynamic range (e.g., bright stars, stability) requires adjusting N and thus λ.
7. Use Eq. (15) to find the total observation time T.
8. Use Eq. (21) to check if SM has superior SNR.
9. Adjust T depending on maximum frame rate (PC) and cosmic ray tolerance (SM).
10. Choose amplification gain G based on device operation considerations.
11. Measure the readout noise μ by reading out unamplified short dark frames.
12. Use Eq. (41) to find the optimal cutoff $\hat{\tau}$.
13. Use Table 1 to find the compression code length and Fig. 5 for the compression ratio.

For a given observation of time T, the number of exposures N and exposure time T/N must be chosen to ensure that the required SNR is met for all image pixels, taking into account variations in CIC, dark current, and signal brightness across the detector. The optimal photon-counting cutoff $\hat{\tau}$ can be chosen per pixel, either in postprocessing or adaptively as exposures are taken. Using these equations, even bright (but unsaturated) stars will have correctly computed expected values.

Acknowledgments

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (Grant No. 80NM0018D0004). Government sponsorship acknowledged.

References

1. A. D. Jewell et al., "Detector performance for the FIREBall-2 UV experiment," *Proc. SPIE* **9601**, 96010N (2015).
2. E. Hamden et al., "FIREBall-2: the faint intergalactic medium redshifted emission balloon telescope," *Astrophys. J.* **898**(2), 170 (2020).
3. G. Kyne et al., "Delta-doped electron-multiplying CCDs for FIREBall-2," *J. Astron. Telesc. Instrum. Syst.* **6**(1), 011007 (2020).
4. L. K. Harding et al., "Technology advancement of the CCD201-20 EMCCD for the WFIRST coronagraph instrument: sensor characterization and radiation damage," *J. Astron. Telesc. Instrum. Syst.* **2**(1), 011007 (2015).
5. A. D. Jewell et al., "Ultraviolet detectors for astrophysics missions: a case study with the star-planet activity research cubesat (SPARC)," *Proc. SPIE* **10709**, 107090C (2018).
6. D. R. Ardila et al., "The Star-Planet Activity Research CubeSat (SPARCS): a mission to understand the impact of stars in exoplanets," arXiv:1808.09954 (2018).
7. P. A. Scowen et al., "Monitoring the high-energy radiation environment of exoplanets around low-mass stars with SPARCS (Star-Planet Activity Research CubeSat)," *Proc. SPIE* **10699**, 106990F (2018).
8. N. Turner et al., "PoZoLE: a tiny space telescope to snap our solar system's ultraviolet selfie," in *Amer. Astron. Soc. Meeting Abstracts*, Vol. 54 (2022).
9. Teledyne e2v, "CCD201-20 datasheet electron multiplying CCD sensor," A1A-100013 Version 7 (2019).
10. P. Jerram et al., "The LLCCD: low-light imaging without the need for an intensifier," *Proc. SPIE* **4306**, 178–186 (2001).
11. K. Matsuo, M. Teich, and B. Saleh, "Noise properties and time response of the staircase avalanche photodiode," *J. Lightwave Technol.* **3**(6), 1223–1231 (1985).
12. O. Daigle et al., "CCCP: a CCD controller for counting photons," *Proc. SPIE* **7014**, 70146L (2008).
13. N. Bush et al., "Measurement and optimization of clock-induced charge in electron multiplying charge-coupled devices," *J. Astron. Telesc. Instrum. Syst.* **7**(1), 016002 (2021).
14. M. S. Robbins and B. J. Hadwen, "The noise performance of electron multiplying charge-coupled devices," *IEEE Trans. Electron. Devices* **50**(5), 1227–1232 (2003).
15. O. Daigle, C. Carignan, and S. Blais-Ouellette, "Faint flux performance of an EMCCD," *Proc. SPIE* **6276**, 62761F (2006).
16. O. Daigle et al., "Extreme faint flux imaging with an EMCCD," *Publ. Astron. Soc. Pacific* **121**(882), 866 (2009).
17. O. Daigle and S. Blais-Ouellette, "Photon counting with an EMCCD," *Proc. SPIE* **7536**, 753606 (2010).
18. O. Daigle et al., "Astronomical imaging with EMCCDs using long exposures," *Proc. SPIE* **9154**, 91540D (2014).
19. A. Basden, C. Haniff, and C. Mackay, "Photon counting strategies with low-light-level CCDs," *Mon. Not. R. Astron. Soc.* **345**(3), 985–991 (2003).
20. C. Cherry et al., "An experimental study of the possible bandwidth compression of visual image signals," *Proc. IEEE* **51**(11), 1507–1517 (1963).

Brian M. Sutin is an instrument systems engineer at Caltech/Jet Propulsion Laboratory. He specializes in the design and performance modeling of space-based sensing instruments, particularly the radiometry, end-to-end system engineering, and calibration plans required to demonstrate that an instrument concept will meet science requirements.