Impact on the Signal-to-Noise Ratio due to Stochastic Release of Multiple Electrons in the NIRSpec Detectors at Short Wavelengths

Peter Jakobsen, v1.0, 11 February 2022

Abstract: A simple statistical model is constructed for the additional noise added to the extracted detector electron signal due to the quantum yield effect - i.e. the NIRSpec detectors generating multiple electron-hole pairs per sufficiently energetic incoming photons in a stochastic manner. The model is used to demonstrate that while the degradation of the signal-to-noise ratio due to the effect is modest ($\leq 6\%$) over the wavelength range of NIRSpec, not taking the effect properly into account in signal-to-noise calculations or error propagation estimation, can lead to the signal-to-noise ratio being over-estimated by as much as $\simeq$26%. The model is also used to generalize the conventional expression for the variance of the signal extracted through up-the-ramp sampling to also include the quantum yield effect.

Introduction and Background

The $\simeq$ 5 μm long wavelength cut-off Teledyne HgCdTe detectors employed in NIRSpec (as well as NIRISS/FGS and the long wavelength channel of NIRCam) are know to occasionally generate more than a single electron-hole pair for incident photons having energies sufficiently above the band gap of the device. The additional statistical noise added to the Poisson noise of the incoming photon signal by this random process is referred to as Fano noise. The purpose of this note is to model and quantify this additional noise source in order that it can be properly included in NIRSpec signal-to-noise calculations at the shorter wavelengths, and eventually be incorporated in the error propagation algorithms of the NIRSpec data reduction pipeline.

The multiple electron generation effect is customarily quantified by the so-called quantum yield, $\Phi(E)$, denoting the average number of electrons released per detected photon as a function of energy. In this picture one distinguishes between the photonic quantum efficiency of the detector, $QE(E)$ and its responsive quantum efficiency $RQE(E)$, such that $RQE(E) = \Phi(E)\, QE(E)$ with $\Phi(E) \geq 1$ at all photon energies.

The quantum yield of the two NIRSpec flight detectors has been characterized by the NASA GSFC Detector Characterization Laboratory. As shown in Figure 1 taken from reference [1], $\Phi(E)$ is customarily described as equalling unity for photon energies below some threshold energy E_t, after which $\Phi(E)$ increases linearly with increasing photon energy.

If instead expressed in terms of the photon wavelength, $\Phi(\lambda)$ can be parameterized as

$$\Phi(\lambda) = \max\left[1, 1 + hc\beta(\frac{1}{\lambda} - \frac{\alpha}{\lambda_c})\right] \qquad (1)$$

where $\alpha = 2.6562$ and $\beta = 0.4708$ eV^{-1} are empirically determined constants, and λ_c is the red cutoff wavelength of the detector response [1,2]. The NIRSpec SCA491 array has a measured cutoff at $\lambda_c \simeq 5.45$ μm and the SCA492 array at $\lambda_c \simeq 5.37$ μm. In the following we for simplicity assume $\lambda_c = 5.40$ μm for both devices.

The shortest wavelength covered by NIRSpec is $\lambda \simeq 0.6$ μm, for which Eq. (1) predicts $\Phi(0.6\mu\text{m}) = 1.69$. That this number lies well below 2 suggests that no more than one additional electron is released over the photon energy range relevant to NIRSpec. With this assumption the quantum yield process can be readily modeled statistically.

The simple statistical model

Consider a single detector pixel illuminated by photons of a fixed wavelength λ. Let n_γ denote the number of photons detected during an integration, and let n_e denote the corresponding number of electrons generated and read out from the pixel.

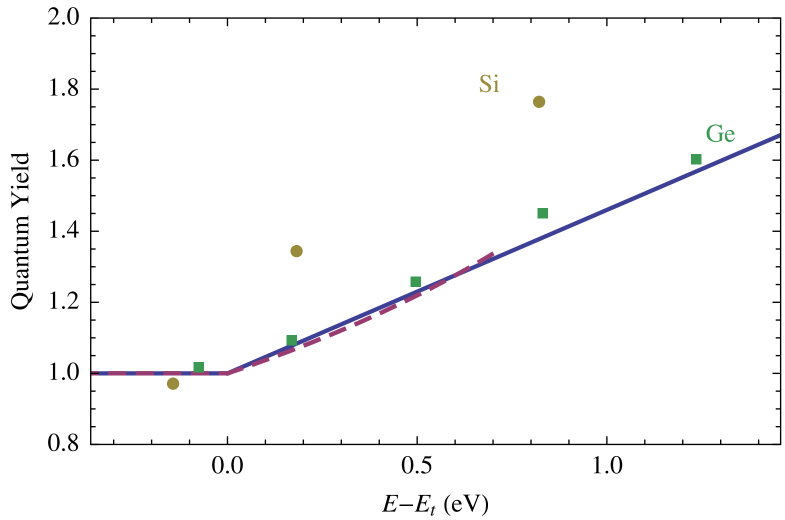

Figure 1: Measured quantum yield $\Phi(E)$ as a function of photon energy for the NIRSpec detectors according to [1]

Let m denote the number of electrons generated for a given incoming photon and p denote the probability that $m = 2$ electrons are released. The probability that $m = 1$ electron is released is then $1 - p$.

The mean of m is

$$E(m) = 2 \times p + 1 \times (1 - p) = 1 + p \tag{2}$$

and its second moment

$$E(m^2) = 4 \times p + 1 \times (1 - p) = 1 + 3p \tag{3}$$

The total number of incoming photons n_γ is Poisson distributed. Therefore

$$E(n_\gamma) = Var(n_\gamma) \tag{4}$$

In this notation the total number of electrons in the accumulated signal n_e results as a random sum of n_γ trials of m

$$n_e = \sum_{i=1}^{n_\gamma} m_i \tag{5}$$

where both n_γ and m_i are stochastic variables. The mean and variance of such a compounded distribution are given by the general expressions

$$E(n_e) = E(n_\gamma)E(m) = E(n_\gamma)(1 + p) \tag{6}$$

$$Var(n_e) = E(n_\gamma)Var(m) + E(m)^2 Var(n_\gamma) = E(n_\gamma)E(m^2) = E(n_\gamma)(1 + 3p) \tag{7}$$

where the second-to-last simplification in eq. (7) exploits the defining property (4) valid for a compound Poisson distribution. It follows from eqs. (6) and (7) that when $p > 0$, the additional randomly generated extra electrons lead to n_e no longer being Poisson distributed (i.e. $Var(n_e) \neq E(n_e)$).

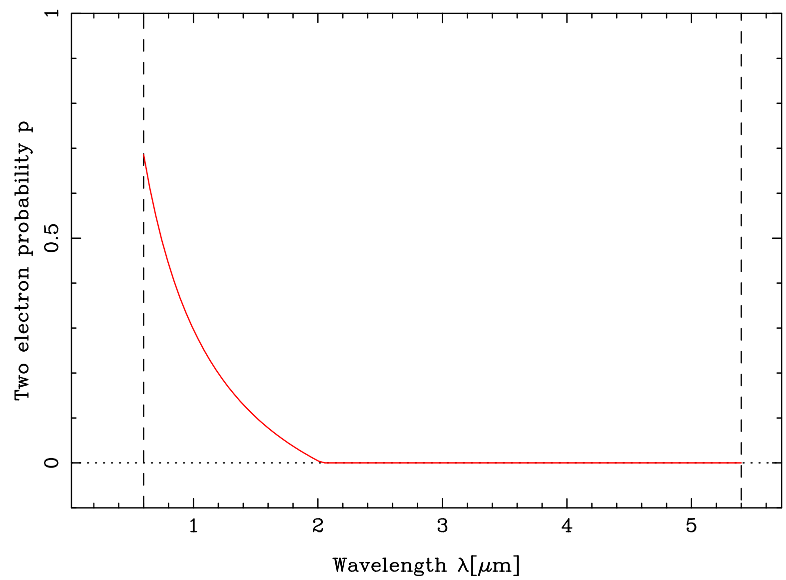

Figure 2: Probability that an incoming photon will generate two electron-hole pairs as a function of wavelength inferred from references [1] and [2].

Comparing eq. (2) to the conventional notation of the previous section, it is evident that we can equate $\Phi(\lambda)$ with $1 + p(\lambda)$. The empirical parameterization of eq. (1) can therefore in our model be re-formulated as

$$p(\lambda) = \begin{cases} 0.5838(\frac{1}{\lambda[\mu\mathrm{m}]} - \frac{1}{2.033}) & \text{for } \ 0.60\,\mu\mathrm{m} \le \lambda < 2.033\,\mu\mathrm{m} \\ 0 & \text{for } \ 2.033\,\mu\mathrm{m} \le \lambda \le 5.40\,\mu\mathrm{m} \end{cases} \tag{8}$$

valid over the $0.60\,\mu\mathrm{m} \le \lambda < 5.40\,\mu\mathrm{m}$ operational range of NIRSpec.

As shown in Figure 2, the maximum value of $p = 0.69$ is reached at the blue-most edge of the NIRSpec coverage at $\lambda = 0.60\,\mu\mathrm{m}$. Note that since the above model by construction does not allow more than two electrons to be released per incoming photon, the quantum yield process becomes non-stochastic with $p(\lambda) \equiv 1$ for wavelengths $\lambda \le 0.454\,\mu\mathrm{m}$. In this extreme limit the multiple electron generation is modeled as a simple scaling of n_γ by a constant factor of 2. This is clearly not realistic since the stochastic creation of triple electron-hole pairs will presumably begin to appear at sufficiently energetic wavelengths. This unrealistic asymptotic behavior notwithstanding, the adopted model can still be assumed to be a reasonable description at the wavelengths relevant for NIRSpec at which $p \le 0.69$.

The NIRSpec radiometric reduction scheme has the extra $(1 + p(\lambda))$ factor in eq. (6) implicitly included in the photon conversion efficiency (PCE) curve. The extra electrons generated at the shorter wavelengths will therefore automatically be taken into account when converting the detected electron signal $\hat{n}_e$ into an incoming physical light flux. However, in this approach the extra noise from the additional generated electrons carried by the extracted signal needs to be explicitly considered in the error propagation calculation used to infer the statistical noise on $\hat{n}_e$.

In the case of an ideal noiseless and background-free detector that always releases a single electron for each incoming photon, the statistical fluctuations in the observed electron signal $\hat{n}_e$ in a raw unresampled pixel are solely caused by the the photon noise. The variance of $\hat{n}_e$ is therefore in this case estimated per eq. (4) in the conventional manner as

$$Var_C(\hat{n}_e) \simeq \hat{n}_e \tag{9}$$

that is, by the value of the measured electron signal itself.

With the additional electron releases, this estimator will in the mean display a variance per eq. (6)

$$E(Var_C(\hat{n}_e)) = E(n_e) = E(n_\gamma)(1 + p) \tag{10}$$

But from eq. (7) we know that the true variance of n_e is

$$Var_T(\hat{n}_e) = E(n_\gamma)(1+3p) \tag{11}$$

Combining eqs. (10) and (11), it follows that the relationship between the true statistical noise on n_e and that calculated in the conventional manner ignoring the quantum yield effect is

$$Var_T(n_e) \simeq Var_C(\hat{n}_e)(\frac{1+3p}{1+p}) = \hat{n}_e(\frac{1+3p}{1+p}) \tag{12}$$

The correction factor $(1+3p)/(1+p)$ takes on the value of unity at wavelengths $\lambda \geq 2.033\,\mu$m and steadily rises with decreasing wavelength to 1.81 at $\lambda = 0.60\,\mu$m. In other words, ignoring the additional noise from the extra released electrons and incorrectly calculating the variance of the signal as if it stemmed from a pure Poisson process, leads to the actual error on the extracted signal being systematically underestimated by a factor of up to $\sqrt{1.81} = 1.35$, and the signal-to-noise ratio therefore being overestimated by up to $\simeq 26\%$, at the shortest wavelengths accessible with NIRSpec.

Eqs. (6) and (7) give for the true signal-to-noise ratio of the net signal n_e including the quantum yield

$$(\frac{S}{N})_T = \sqrt{n_\gamma}\frac{1+p}{\sqrt{1+3p}} \tag{13}$$

For constant n_γ (i.e. constant assumed QE) the maximum degradation factor w.r.t. the underlying photon statistics occurs for $p(\lambda) = \frac{1}{3}$ where it reaches $2\sqrt{2}/3 = 0.9428$. Per the parameterization of eq. (8), this occurs at a wavelength of $\lambda = 0.941\,\mu$m. It follows that for a given fixed number of photons, the additional electrons released by the finite quantum yield will in the present model degrade the total signal-to-noise ratio relative to the pure photon ($p = 0$) case by $\leq 6\%$.

However, this is not the situation we are normally faced with. In practice we are given the measured electron signal $\hat{n}_e$, and need to appeal to external knowledge of the wavelength of the incoming photons λ and the corresponding value of $p(\lambda)$ to partition $\hat{n}_e$ between the incident photons and the extraneous generated electrons to correctly assess the true statistical noise on $\hat{n}_e$. This corresponds to the situation where $\hat{n}_e$ and RQE are known, and the inferred value of n_γ implicitly decreases with the assumed value of p as $n_\gamma = \hat{n}_e/(1+p)$. In this case eq. (13) becomes

$$(\frac{S}{N})_T \simeq \sqrt{\hat{n}_e}\sqrt{\frac{1+p}{1+3p}} \tag{14}$$

In this situation the signal-to-noise reduction factor w.r.t. the pure photon ($p = 0$) case decreases with increasing p, dropping to a factor of 0.74 for $p(0.60\,\mu\text{m}) = 0.69$ in accord with eq. (12). Note that the dominant error made when ignoring the quantum yield effect in estimating the statistical noise on $\hat{n}_e$ is not so much ignoring the additional source of noise caused by the extra electrons themselves, but over-estimating the number of detected photons as $n_\gamma \simeq \hat{n}_e$ when in reality only $n_\gamma \simeq \hat{n}_e/(1+p)$ photons are detected.

This simple analysis serves to show that while the wavelength-dependent quantum yield effect in itself only leads to a modest ($\leq 6\%$) degradation in the net signal-to-noise, the mistake of not taking the effect properly into account by reducing the data as if the detected electron signal were Poisson distributed, results in the inferred signal-to-noise ratio being overestimated by as much as $\simeq 26\%$ at the shortest NIRSpec wavelengths.

Application to up-the-ramp sampling

The key finding of our simple model is that whenever the quantum yield effect is in play, the photon-generated electron signal n_e ceases to be Poisson distributed, but instead displays a variance given by eq. (12).

The real-life NIRSpec detectors, however, also display significant read-out noise which further adds to the variance of n_e. Moreover, the accumulated electron signal is extracted by means of a complicated up-the-ramp slope-fitting processing algorithm designed to average down the read noise. The canonical expression for the variance of the accumulated electron signal extracted from up-the-ramp processing is given in [3] and [4]. The observation that the lengthy derivation of this expression nowhere makes use of the fact that the variance of the photon signal is Poissonian, suggests that it can be readily generalized to include the quantum yield effect simply by substituting the variance (12) for that of the photon noise. This results in

$$Var(n_e) = \frac{12(n-1)}{mn(n+1)}\sigma^2_{\text{read}} + (\frac{6(n^2+1)}{5n(n+1)} - \frac{2(m^2-1)}{n(n+1)m^2})(\frac{1+3p}{1+p})n_e \tag{15}$$

with

$$n_e = (n-1)\, m\, t_f\, f \tag{16}$$

Here n is the number of groups of m averaged reads that the exposure extends over, t_f is the array frame time, f the signal electron rate (slope - including the quantum yield), and σ^2_{read} is the read noise per read.

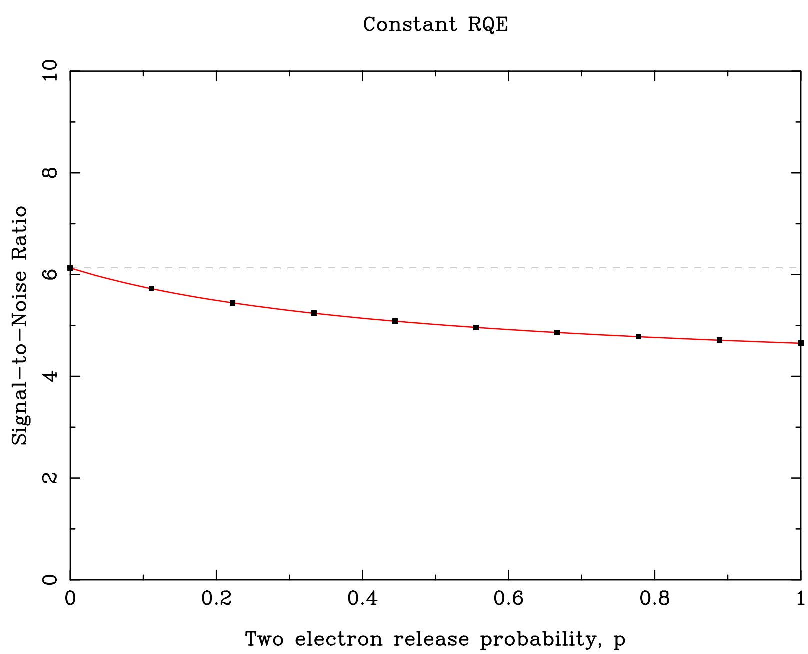

Figure 3: Detailed pixel-level Monte Carlo simulation of a ($n = 14$, $m = 5$) sub-exposure of the reference continuum source reference signal-to-noise calculation described in [5], expanded to also include the quantum yield. The squares show the outcome of the simulations for each assumed value of p. The continuous curve shows the signal-to-noise ratio calculated per eqs. (15) and (16). The excellent agreement holds up for any assumed value of the input source brightness.

Eqs.(15) and (16) were verified numerically by means of detailed pixel-level Monte Carlo simulations of the up-the-ramp signal accumulation and extraction process. Figure 3 shows the result of expanding the already existing detailed simulation of one sub-exposure of the $3\,\mu$m continuum source reference signal-to-noise calculation presented in [5] to also simulate the quantum yield for different assumed values of p, while holding RQE constant. Although multiple electrons are strictly speaking not generated at a wavelength of $3\,\mu$m, this is immaterial for the purpose of verifying eq. (15). This particular simulation was merely chosen for convenience as a representative example involving significant sky background and detector noise. It is seen that the agreement between the simulations (squares) and the predicted signal-to-noise ratio calculated from eqs.(15) and (16) (full curves) is excellent for all assumed values of p. This also holds if the source input flux is increased or decreased by a factor of ten from its nominal value.

As expected from eq. (15), the inclusion of the read-noise dilutes the net reduction in signal-to-noise w.r.t to the $p = 0$ case caused by the quantum yield slightly for a given value of p compared to the ideal detector case of eq. (14). At the largest value of $p = 0.69$ relevant for NIRSpec, the signal-to-noise reduction factor inferred from the simulation is 0.79 compared to 0.74 calculated from eq. (14). The latter value is approached if the source brightness in the simulation in increased, thereby making the photon-induced signal more dominant w.r.t. to the detector noise. Conversely, decreasing the assumed source brightness causes the reduction factor to slowly increase.

It follows that eq. (14) captures the worst-case error made in the calculated signal-to-noise made when overlooking the quantum yield effect, also in the presence of significant read-noise and when up-the-ramp sampling is employed.

Conclusions

Armed with the parameterization of $p(\lambda)$ given by eq. (8) (or some other preferred alternate version thereof) and the generalized expression eq. (15) for the variance of the up-the-ramp extracted accumulated electron signal including the quantum yield, it is straightforward to refine any NIRSpec signal-to-noise calculation, following the recipe described in [5], to correctly account for the multiple electron events occurring at the affected shorter wavelengths.

Taking the effect into account in the error propagation calculation during the reduction of a given NIRSpec exposure is slightly more tricky. It should ideally be done at the individual raw (i.e. unresampled) pixel level, employing the NIRSpec instrument model to first work out the wavelength of the light illuminating each pixel (assuming no spectral overlap), and then using eq. (15) with the appropriate value of $p(\lambda)$ obtained from eq. (8), and the measured accumulated electron signal $\hat{n}_e$ (including the quantum yield) substituted for n_e and p in eq. (15) to estimate the noise on $\hat{n}_e$.

Lastly, we note that although the parameterization of $p(\lambda)$ given by eq. (8) has $p(\lambda) > 0$ setting in already at $\lambda \leq 2.033\,\mu$m (i.e. at a wavelength notably longer than the $\lambda \leq 1.4\,\mu$m value usually quoted for NIRSpec), the value of p at the blue-most end of Band II at $\lambda = 1.70\,\mu$m is only $p = 0.056$, which, if ignored, according to eq. (14) leads to a worst case $\simeq 5\%$ error in the signal-to-noise estimate. Therefore in practice, the considerations of this note only need apply to NIRSpec observations taken with the PRISM or either of the two Band I gratings.

References

[1] Rauscher, B. et al. 2014, PASP, 126, 739
[2] Ferruit, P. 2014 *Computation of the quantum yield correction*, ESA NTN-2014-002
[3] Ferruit, P. 2006 *Computation of noise levels for different readout patterns*, NIRS-CRAL-TN-0024
[4] Rauscher, B. et al. 2007, PASP, 119, 768 - Erratum: 2010, PASP, 122, 1254
[5] Jakobsen, P. et al. 2022, A&A, in press, arXiv:2202.03305