

Nonlinearity Corrections and Statistical Uncertainties Associated with Near-Infrared Arrays

WILLIAM D. VACCA

SOFIA-USRA, NASA Ames Research Center, MS 144-2, Moffett Field, CA 94035, and Department of Astronomy, 601 Campbell Hall,
University of California, Berkeley, CA 94720; wvacca@mail.arc.nasa.gov

AND

MICHAEL C. CUSHING AND JOHN T. RAYNER

Institute for Astronomy, University of Hawai'i, 2680 Woodlawn Drive, Honolulu, HI 96822; cushing@ifa.hawaii.edu, rayner@ifa.hawaii.edu

Received 2003 August 18; accepted 2004 January 15; published 2004 April 5

ABSTRACT. We derive general equations for nonlinearity corrections and statistical uncertainty (variance) estimates for data acquired with near-infrared detectors employing correlated double sampling, multiple correlated double sampling (Fowler sampling), and uniformly spaced continuous readout techniques. We compare our equation for the variance on each pixel associated with Fowler sampling with data taken with the array installed in the near-infrared cross-dispersed spectrograph (SpeX) at the NASA Infrared Telescope Facility and find that it provides an accurate representation of the empirical results. This comparison also reveals that the read noise associated with a single readout of the SpeX array increases with the number of nondestructive reads, n_r , as $n_r^{0.16}$. This implies that the *effective* read noise of a stored image decreases as $n_r^{-0.34}$, shallower than the expected rate of $n_r^{-0.5}$. The cause of this read noise behavior is uncertain but it may be due to heating of the array as a result of the multiple readouts. Such behavior may be generic to arrays that employ correlated or multiple correlated double sampling readouts.

1. INTRODUCTION

Correlated double sampling (CDS) and multiple correlated double sampling (MCDS, also known as Fowler sampling; Fowler & Gatley 1990) are common techniques used to read out near-infrared (NIR) arrays. These techniques are used, for example, for recording data obtained with the arrays in SpeX at the NASA Infrared Telescope Facility (Rayner et al. 2003), in NIRSPEC at the Keck telescope (McLean et al. 1998), in the IRCS at the Subaru telescope (Kobayashi et al. 2000), and in ISAAC at the Very Large Telescope (Moorwood 1997). An alternative readout technique, known as continuous, or “up the ramp,” sampling, was used for recording data obtained with the CGS2 spectrograph on the United Kingdom Infrared Telescope (Chapman et al. 1990). A variant of this technique is the recommended readout scheme for the detectors in the Near-Infrared Camera and Multi-Object Spectrometer (NICMOS) aboard the *Hubble Space Telescope* (Royer et al. 2003). The relative advantages and disadvantages of these techniques in various noise-limited operating regimes have been summarized by Garnett & Forrest (1993).

As part of the effort to develop a set of semiautomated routines to reduce SpeX data (Spextool; see Cushing, Vacca, & Rayner 2003), we have derived equations for the nonlinearity corrections and the uncertainties on each pixel value recorded by an array that employs CDS, MCDS, and uniformly spaced

continuous sampling readout techniques. In this paper, we present a simple model for the CDS and MCDS readout procedures of an NIR array, define the relevant parameters, and derive equations for the nonlinearity corrections and variances on pixel values associated with these techniques (§ 2). The corresponding model and equations for the uniformly spaced continuous sampling technique are presented in § 3. Verification of the MCDS equations and their application to the SpeX array for the determination of the gain and read noise are presented in § 4. A summary of our results is given in § 5.

2. CORRELATED AND MULTIPLE CORRELATED DOUBLE SAMPLING

2.1. Ideal Detector

Figure 1a shows a schematic of the signal recorded on an integrating node as a function of time for an ideal detector exposed to a constant source flux and read out using the MCDS technique. The ideal detector is perfectly linear over its entire well depth, is thermally stable, and is not subject to any time-dependent effects that vary during integrations or readouts (such as bias drift, image persistence that decays over time, or amplifier glow associated with readouts). Excellent descriptions of such systematic effects, which are assumed to be either negligible or correctable (and therefore ignored) in our analysis,

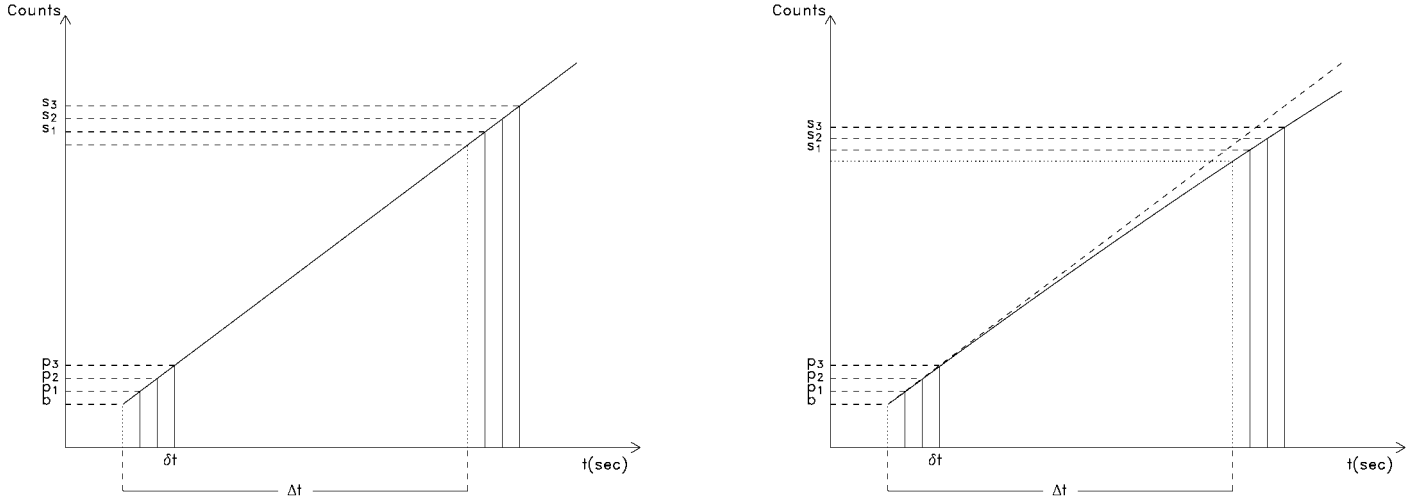


FIG. 1.—Schematic of the signal recorded by an integrating node as a function of time on a perfectly linear detector (*left*) and a real (i.e., nonlinear) detector (*right*) employing the MCDS readout technique with $n_r = 3$. The two dotted vertical lines represent the starting time of the first read of the pedestal and the start of the first read of the signal. Each read is assumed to last δt s. Each solid vertical line represents the end of a read. The exposure time is Δt s.

are given by Skinner & Bergeron (1997) and Roye et al. (2003) for the NICMOS array.

At time $t = 0$, the array is reset to the bias level \mathbf{b} , which is assumed to be constant over time during the integrations and readouts.¹ In NIR arrays, the bias is a negative voltage set by the electronics and therefore is “noiseless.” Because it is assumed to be simply an additive constant that defines a zero point for count levels at each pixel, it can be removed from all equations without any loss of generality (as long as it does not vary). We assume it takes a time t_{reset} to reset an individual pixel. The reset procedure may be global (all pixels in the array reset at once, as in SpeX) or sequential (array reset occurring pixel by pixel, as in NIRSPEC). After reset, the array is immediately read out nondestructively n_r times. (In the CDS technique, $n_r = 1$.) The time needed to read an individual pixel is t_{read} , and the pixels are read sequentially during an array readout. Reading the entire array, therefore, requires a total time δt , given by the time needed to read a single pixel times the number of pixels read, n_p , plus any overhead or waiting time between readouts:

$$\delta t = n_p t_{\text{read}} + t_{\text{wait}}. \quad (1)$$

In SpeX, for example, the 1024×1024 array is divided into quadrants that are read simultaneously, each of which has eight

¹ Variables written in roman boldface denote two-dimensional quantities (i.e., images, which may have different values at each pixel in the image); mathematical operations involving such variables should be understood to represent image manipulation in which the operations are performed on each pixel in the image. Boldface variables with subscripts (e.g., \mathbf{s}_i) denote individual images in a sequence. Variables written in italics with subscripts, e.g., $s_{k,l}$, denote individual values at a given pixel k, l .

readout outputs. For this array, $t_{\text{reset}} \approx 30 \mu\text{s}$, $t_{\text{read}} \approx 10 \mu\text{s}$, and $n_p = 2^{15}$; therefore, δt is on the order of 300 ms.

This first set of n_r readouts provides an estimate of the mean pedestal level $\bar{\mathbf{p}}$. After an exposure time Δt , the array is again read out n_r times, which provides an estimate of the mean signal level $\bar{\mathbf{s}}$. Therefore, Δt is given by the time difference between the start of the first read of the pedestal \mathbf{p}_1 and the start of the first read of the signal \mathbf{s}_1 . In most implementations of these techniques, the values of neither the individual readouts nor the mean pedestal and signal levels are stored; instead, usually only the net source counts (in data numbers, DN) for the n_r nondestructive readouts after an integration time of Δt , given by

$$\mathbf{S}_{\text{net}} = \sum_{i=1}^{n_r} \mathbf{s}_i - \sum_{i=1}^{n_r} \mathbf{p}_i \quad (2)$$

$$= n_r \bar{\mathbf{s}} - n_r \bar{\mathbf{p}}, \quad (3)$$

is recorded. The estimated source flux \mathbf{I} (which includes contributions from the dark current, background, sky, and object), in units of DN s^{-1} , is then given by

$$\mathbf{I} = \frac{1}{n_r \Delta t} \mathbf{S}_{\text{net}} = \frac{1}{n_r \Delta t} \left(\sum_{i=1}^{n_r} \mathbf{s}_i - \sum_{i=1}^{n_r} \mathbf{p}_i \right). \quad (4)$$

Hence, \mathbf{I} is the slope of the solid diagonal line in Figure 1a.

Our goal is to determine the uncertainty, or variance, in the estimated source flux \mathbf{I} . As can be seen from equation (4) and in § 2.4 below, expressions for the individual pedestal and signal readout values are needed for the computation of the variances in \mathbf{I} . From Figure 1a, it can be seen that for a linear

detector, after a reset the counts (in DN) recorded at the end of the i th read of the pedestal \mathbf{p}_i and the signal \mathbf{s}_i can be expressed in terms of the recorded net source counts \mathbf{S}_{net} or the estimated source flux \mathbf{I} as

$$\mathbf{p}_i = (i - \mathbf{f}) \frac{\mathbf{S}_{\text{net}} \delta t}{n_r \Delta t} \quad (5)$$

$$= (i - \mathbf{f}) \mathbf{I} \delta t \quad (6)$$

and

$$\mathbf{s}_i = \frac{\mathbf{S}_{\text{net}}}{n_r} + (i - \mathbf{f}) \frac{\mathbf{S}_{\text{net}} \delta t}{n_r \Delta t} \quad (7)$$

$$= \frac{\mathbf{S}_{\text{net}}}{n_r} + \mathbf{p}_i \quad (8)$$

$$= \mathbf{I} \Delta t + \mathbf{p}_i, \quad (9)$$

respectively. Here \mathbf{f} is a fraction, $\mathbf{f} < 1$, whose value at a given pixel depends on the position of the pixel in the readout sequence. The quantity $(i - \mathbf{f}_{k,l})\delta t$ is the time after the reset when the pixel k, l is read out for the i th time. If a total number of n_p pixels are read out, f for the m th pixel in the readout sequence is given by

$$f = 1 - m \frac{t_{\text{read}}}{\delta t} - \frac{t_{\text{reset}}}{\delta t} \quad (10)$$

for a global reset, and

$$f = 1 - m \frac{t_{\text{read}}}{\delta t} - (n_p - m + 1) \frac{t_{\text{reset}}}{\delta t} \quad (11)$$

for a sequential reset.

A common observing technique employed with NIR arrays is to immediately repeat an integration n_c times and “co-add” the results. In this case, the entire array addressing procedure (reset, n_r pedestal reads, n_r signal reads) is repeated n_c times. Usually, the net source counts resulting from the individual integrations are not recorded. Rather, the total source counts given by

$$\mathbf{S}_{\text{tot}} = \sum_{j=1}^{n_c} \mathbf{S}_{\text{net},j} = \sum_{j=1}^{n_c} \sum_{i=1}^{n_r} \mathbf{s}_{i,j} - \sum_{j=1}^{n_c} \sum_{i=1}^{n_r} \mathbf{p}_{i,j} \quad (12)$$

$$= n_c n_r \bar{\mathbf{s}} - n_c n_r \bar{\mathbf{p}} \quad (13)$$

is stored. The estimated source flux \mathbf{I} in units of DN s^{-1} is

given by

$$\mathbf{I} = \frac{1}{n_c n_r \Delta t} \mathbf{S}_{\text{tot}}. \quad (14)$$

Therefore,

$$\mathbf{p}_i = (i - \mathbf{f}) \frac{\mathbf{S}_{\text{tot}} \delta t}{n_c n_r \Delta t} \quad (15)$$

and

$$\mathbf{s}_i = \frac{\mathbf{S}_{\text{tot}}}{n_c n_r} + (i - \mathbf{f}) \frac{\mathbf{S}_{\text{tot}} \delta t}{n_c n_r \Delta t} \quad (16)$$

$$= \frac{\mathbf{S}_{\text{tot}}}{n_c n_r} + \mathbf{p}_i. \quad (17)$$

2.2. Nonlinear Detector

Unfortunately, few astronomical detectors are perfectly linear over their entire well depth. NIR arrays are inherently nonlinear devices because the detector capacitance varies during an integration as photons are detected (see, e.g., McCaughrean 1988). Figure 1b shows a schematic of the signal recorded on an integrating node as a function of time for a nonlinear NIR array read out using the MCDS technique. The recorded signal (*solid line*) falls below that expected in the ideal case (*diagonal dashed line*). In the absence of other systematic and time-variable effects (such as image persistence, bias drift, or thermal instability of the array), the ratio of the theoretically expected signal to the recorded signal yields the nonlinearity correction curve of the array, which increases steadily from unity as a function of the recorded counts:

$$\mathbf{C}_{\text{nl}}(\mathbf{s}) = \frac{\mathbf{s}^{\text{lin}}}{\mathbf{s}}, \quad (18)$$

where \mathbf{s}^{lin} is the signal recorded by an ideal (linear) detector.

Therefore, the values of the individual readouts of the pedestal and signal are given by

$$\mathbf{s}_i = \frac{\mathbf{s}_i^{\text{lin}}}{\mathbf{C}_{\text{nl}}(\mathbf{s}_i)} \quad \text{and} \quad \mathbf{p}_i = \frac{\mathbf{p}_i^{\text{lin}}}{\mathbf{C}_{\text{nl}}(\mathbf{p}_i)}. \quad (19)$$

If the nonlinearity of the detector is not accounted for, the estimated source flux \mathbf{I} can be substantially less than the true source flux. Furthermore, as can be seen from Figure 1b, the values of the individual pedestals and signal readouts derived from equations (5)–(9) will be underestimated.

The corrections \mathbf{C}_{nl} for the nonlinearity of optical and NIR arrays are often derived from flat-field data acquired expressly

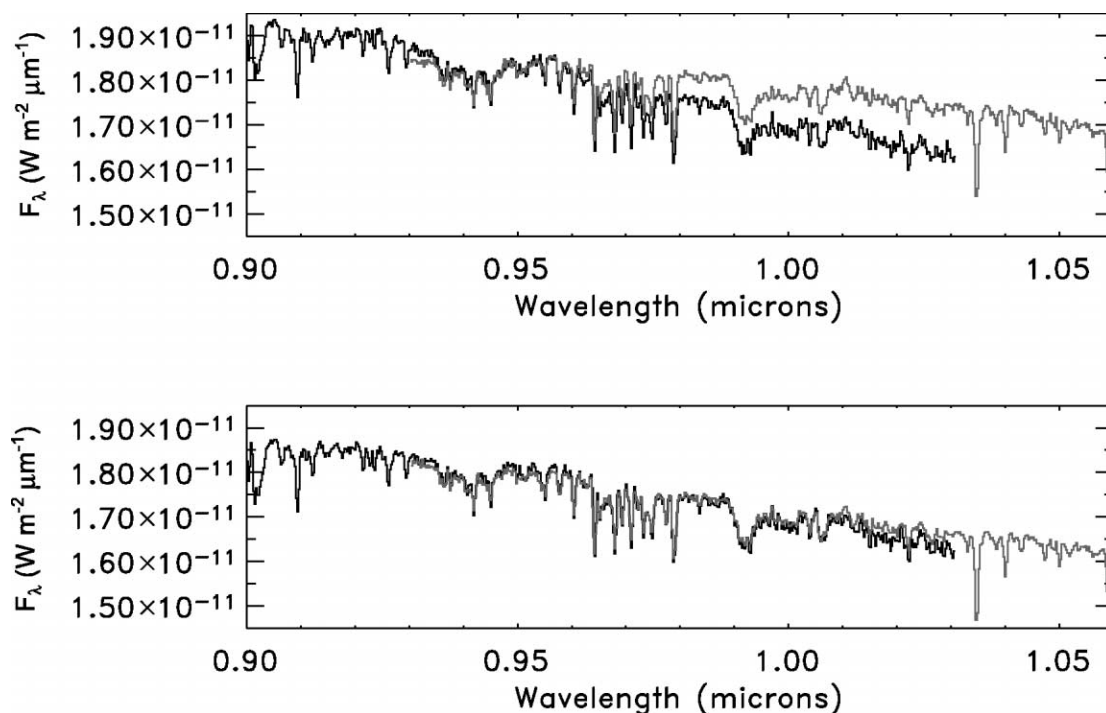


FIG. 2.—Spectra of Gl 846 from orders 6 and 7 of SpeX in the SXD mode (see Rayner et al. 2003). The spectra were reduced with the *Spextool* reduction package (Cushing et al. 2003). The spectra shown in the upper panel were reduced without applying the nonlinearity corrections discussed in the text; the corrections were incorporated in the reduction of the spectra shown in the lower panel. Incorporation of the nonlinearity corrections removes, to a large degree, the mismatch in both the flux levels and the slopes of the spectra in the overlap regions.

for this purpose. A series of flat-field images taken with gradually increasing exposure times is used to generate a curve of the recorded signal at each pixel as a function of the exposure time, or equivalently, the estimated count rate at each pixel as a function of the recorded counts. (Of course all systematic effects, such as those mentioned above, must be corrected for and removed before the nonlinearity curve can be determined from the flat-field data.) The curve can then be fitted with a continuous function (e.g., a polynomial), and the deviation of the function (at large count values for long exposures) from a straight line (fitted to the count values recorded during short exposures) provides a measure of the nonlinearity correction $C_{nl}(s_{k,l})$ as a function of the detected counts $s_{k,l}$ in pixel k, l . For the SpeX array, for example, it was found that the nonlinearity correction curve can be approximated extremely well over the range 0 to ~ 8500 DN (which represents $\sim 90\%$ of the well depth and is close to saturation) by the rational function

$$C_{nl}(s) \approx \frac{1}{1 + \mathbf{a}_1 \mathbf{x} + \mathbf{a}_2 \mathbf{x}^2 + \mathbf{a}_3 \mathbf{x}^3}, \quad (20)$$

where \mathbf{a}_1 , \mathbf{a}_2 , and \mathbf{a}_3 are constants, $\mathbf{x} = s - \mathbf{p}_{\text{flat}}$, and \mathbf{p}_{flat} is the pedestal level estimated for the flat-field images (using

eq. [23] below). This curve can then be used to correct the signals recorded by the detector (after removal of the bias, dark current, and any systematic effects) to the “true” values that would be recorded by a perfectly linear detector.

2.3. Nonlinearity Corrections

When applying a nonlinearity correction to data acquired with CDS or MCDS techniques, it is important to keep in mind that the corrections must be made to *both* the signal and pedestal values separately, not to the net or total source counts \mathbf{S}_{net} or \mathbf{S}_{tot} . However, because these readout procedures do not record separate signal and pedestal values (see § 2.1), we must estimate these a posteriori from \mathbf{S}_{tot} . Fortunately, this can be done relatively easily with the definitions given in § 2.1, although in general an iterative procedure is required.

We assume that the nonlinearity corrections apply to the mean signal and pedestal values above the bias level (that is, the relative values above the bias rather than the absolute values). Furthermore, we assume the bias level remains constant during the integration and readouts, the dark current has been subtracted, and any other systematic effects have been corrected for before the nonlinearity corrections are applied. In this case, first estimates of the mean pedestal and signal values can be

calculated from

$$\bar{\mathbf{p}}^{(1)} = \frac{1}{n_c n_r} \sum_{j=1}^{n_c} \sum_{i=1}^{n_r} \mathbf{p}_{i,j} \quad (21)$$

$$= \frac{1}{n_c n_r} \sum_{j=1}^{n_c} \sum_{i=1}^{n_r} (i - \mathbf{f}) \frac{\mathbf{S}_{\text{net},j}}{n_r} \frac{\delta t}{\Delta t} \quad (22)$$

$$= \frac{\mathbf{S}_{\text{tot}}}{n_c n_r} \frac{\delta t}{\Delta t} \left[\frac{n_r + 1}{2} - \mathbf{f} \right] \quad (23)$$

and

$$\bar{\mathbf{s}}^{(1)} = \frac{1}{n_c n_r} \mathbf{S}_{\text{tot}} + \bar{\mathbf{p}}^{(1)}, \quad (24)$$

respectively.

These values are then corrected for nonlinearity using the \mathbf{C}_{nl} curve

$$\bar{\mathbf{s}}^{(2)} = \bar{\mathbf{s}}^{(1)} \mathbf{C}_{\text{nl}}(\bar{\mathbf{s}}^{(1)}), \quad \bar{\mathbf{p}}^{(2)} = \bar{\mathbf{p}}^{(1)} \mathbf{C}_{\text{nl}}(\bar{\mathbf{p}}^{(1)}) \quad (25)$$

and a new estimate of the net source counts is computed from

$$\mathbf{S}_{\text{tot}}^{(2)} = n_c n_r \bar{\mathbf{s}}^{(2)} - n_c n_r \bar{\mathbf{p}}^{(2)}. \quad (26)$$

It should be noted that the corrected values of $\bar{\mathbf{s}}$, $\bar{\mathbf{p}}$, and \mathbf{S}_{tot} depend on the number of nondestructive reads n_r (see eq. 23).

Using the corrected estimate of the total source counts $\mathbf{S}_{\text{tot}}^{(2)}$, a new estimate of the mean pedestal level $\bar{\mathbf{p}}^{(3)}$ can be computed from equation (23). Adding this to the *original* values of \mathbf{S}_{tot} using equation (24) yields a new estimate of the signal, $\bar{\mathbf{s}}^{(3)}$. The new estimates of $\bar{\mathbf{s}}$ and $\bar{\mathbf{p}}$ can then be corrected for nonlinearity using the \mathbf{C}_{nl} curve and then used to generate another estimate of true source counts $\mathbf{S}_{\text{tot}}^{(3)}$. The procedure can be repeated as many times as necessary to achieve a desired convergence criterion.

For SpeX data we have found that convergence is rapidly achieved and that unless the mean detected signal $\bar{\mathbf{s}}$ is close to the top of the well (>8500 DN, highly nonlinear), the correction to the estimate of the true source counts is less than 1% after three iterations; therefore, $\mathbf{S}_{\text{tot}}^{(4)} \approx \mathbf{S}_{\text{tot}}^{\text{lin}}$.

An example of both the necessity of making corrections for nonlinearity and the effectiveness of the procedure outlined above can be seen in Figure 2, taken from the paper by Cushing et al. (2003). This figure shows that spectra in adjacent spectral orders, reduced without incorporating the nonlinearity corrections, do not match one another in either intensity level or spectral slope in the overlapping wavelength region. After the nonlinearity corrections described above are made, however, the spectra agree in both intensity and slope to better than a few percent over the entire overlap region.

2.4. Variance Estimates

Because the array is read out nondestructively in the CDS and MCDS (as well as the ‘‘up the ramp’’) techniques, the values recorded at a given pixel for each readout are correlated. Therefore, the variance \mathbf{V}_i of the estimated source flux \mathbf{I} cannot be estimated simply from the photon noise and the read noise for the individual readouts added in quadrature. The covariance between any two readouts must be included in the variance estimate (Garnett & Forrest 1993).

We start by considering the case in which $n_c = 1$. The variance of \mathbf{I} , \mathbf{V}_i , is given by the standard error propagation formula (Bevington 1969; Garnett & Forrest 1993)

$$\begin{aligned} \mathbf{V}_i = & \sum_{i=1}^{n_r} \left[\sigma_{s_i}^2 \left(\frac{\partial \mathbf{I}}{\partial s_i} \right)^2 + \sigma_{p_i}^2 \left(\frac{\partial \mathbf{I}}{\partial p_i} \right)^2 \right] \\ & + 2 \sum_{j=2}^{n_r} \sum_{i < j} \left[\text{Cov}(s_i, s_j) \left(\frac{\partial \mathbf{I}}{\partial s_i} \right) \left(\frac{\partial \mathbf{I}}{\partial s_j} \right) \right. \\ & \quad \left. + \text{Cov}(p_i, p_j) \left(\frac{\partial \mathbf{I}}{\partial p_i} \right) \left(\frac{\partial \mathbf{I}}{\partial p_j} \right) \right] \\ & + 2 \sum_{i=1}^{n_r} \sum_{j=1}^{n_r} \left[\text{Cov}(p_i, s_j) \left(\frac{\partial \mathbf{I}}{\partial p_i} \right) \left(\frac{\partial \mathbf{I}}{\partial s_j} \right) \right], \quad (27) \end{aligned}$$

where $\sigma_{x_i}^2$ is the variance of the i th readout value and $\text{Cov}(\mathbf{x}_i, \mathbf{x}_j)$ is the covariance between the i th and j th readouts.

The variance of the i th readout value is given by the sum of the square of the photon noise and the square of the read noise. The read noise σ_{read} is usually specified in terms of electrons. If the array has a gain \mathbf{g} (with units of electrons per data number, $e^- \text{DN}^{-1}$), conversion to DN gives a variance associated with the read noise of $\sigma_{\text{read}}^2 / \mathbf{g}^2$. The variance associated with the photon noise of each readout value, expressed in DN^2 , is given by

$$\sigma_{s_i}^2 = \frac{s_i}{\mathbf{g}} \quad \text{and} \quad \sigma_{p_i}^2 = \frac{p_i}{\mathbf{g}}. \quad (28)$$

The covariance between the i th and j th readout is given by the square of the photon noise of the i th readout for $j > i$ (Garnett & Forrest 1993). This can easily be shown as follows. For any two reads i, j , with $j > i$, the associated readout values are \mathbf{x}_i and \mathbf{x}_j , which are related by

$$\mathbf{x}_j = \mathbf{x}_i + \Delta_{j-i}, \quad (29)$$

where Δ_{j-i} is the difference in counts between the two reads.

Then

$$\begin{aligned}
 \text{Cov}(\mathbf{x}_i, \mathbf{x}_j) &= \langle (\mathbf{x}_i - \langle \mathbf{x}_i \rangle)(\mathbf{x}_j - \langle \mathbf{x}_j \rangle) \rangle \\
 &= \langle \mathbf{x}_i^2 \rangle - \langle \mathbf{x}_i \rangle^2 + \langle \mathbf{x}_i \Delta_{j-i} \rangle - \langle \mathbf{x}_i \rangle \langle \Delta_{j-i} \rangle \\
 &= \text{Cov}(\mathbf{x}_i, \mathbf{x}_i) + \text{Cov}(\mathbf{x}_i, \Delta_{j-i}) \\
 &= \sigma_{\mathbf{x}_i}^2,
 \end{aligned} \tag{30}$$

where the angle brackets $\langle \dots \rangle$ denote the average. Therefore, we have

$$\text{Cov}(\mathbf{s}_i, \mathbf{s}_j) = \sigma_{\mathbf{s}_i} \text{ and } \text{Cov}(\mathbf{p}_i, \mathbf{p}_j) = \sigma_{\mathbf{p}_i} \text{ for } j > i \tag{32}$$

and

$$\text{Cov}(\mathbf{p}_i, \mathbf{s}_j) = \sigma_{\mathbf{p}_i}. \tag{33}$$

Upon substitution, equation (27) becomes

$$\begin{aligned}
 \mathbf{V}_i &= \frac{1}{n_r^2 \Delta t^2} \left\{ \sum_{i=1}^{n_r} \left(\frac{\mathbf{s}_i}{\mathbf{g}} + \frac{\sigma_{\text{read}}^2}{\mathbf{g}^2} + \frac{\mathbf{p}_i}{\mathbf{g}} + \frac{\sigma_{\text{read}}^2}{\mathbf{g}^2} \right) \right. \\
 &\quad \left. + 2 \sum_{j=2}^{n_r} \sum_{i=1}^{j-1} \left(\frac{\mathbf{s}_i}{\mathbf{g}} + \frac{\mathbf{p}_i}{\mathbf{g}} \right) - 2 \sum_{i=1}^{n_r} \sum_{j=1}^{n_r} \frac{\mathbf{p}_i}{\mathbf{g}} \right\} \tag{34}
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{\mathbf{g} n_r^2 \Delta t^2} \left\{ \sum_{i=1}^{n_r} (\mathbf{s}_i + \mathbf{p}_i) + 2 \sum_{i=1}^{n_r} (n_r - i) (\mathbf{s}_i + \mathbf{p}_i) \right. \\
 &\quad \left. - 2 n_r \sum_{i=1}^{n_r} \mathbf{p}_i + \frac{2 n_r}{\mathbf{g}} \sigma_{\text{read}}^2 \right\} \tag{35}
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{\mathbf{g} n_r^2 \Delta t^2} \left\{ \sum_{i=1}^{n_r} (\mathbf{s}_i + \mathbf{p}_i) + 2 n_r \sum_{i=1}^{n_r} \mathbf{s}_i \right. \\
 &\quad \left. - 2 \sum_{i=1}^{n_r} i (\mathbf{s}_i + \mathbf{p}_i) + \frac{2 n_r}{\mathbf{g}} \sigma_{\text{read}}^2 \right\}. \tag{36}
 \end{aligned}$$

We now assume that the array has a perfectly linear response, such that $\mathbf{s}_i = \mathbf{s}_i^{\text{lin}}$, $\mathbf{p}_i = \mathbf{p}_i^{\text{lin}}$, and $\mathbf{S}_{\text{net}} = \mathbf{S}_{\text{net}}^{\text{lin}}$. If the array is not linear, the signal and pedestal images must be first corrected for nonlinearity as outlined above. Explicitly accounting for nonlinearity would introduce factors of $\mathbf{C}_{\text{nl}}(\mathbf{s}_i)$ and $\mathbf{C}_{\text{nl}}(\mathbf{p}_i)$ in the equations given below, which would result in an increase in the variance of any given pixel value beyond that expected for an intrinsically linear array and would also prevent us from deriving completely general equations for the variance in terms of recorded quantities. However, as we demonstrate below (§ 4), this

increase can easily be accounted for by modifying the values of the gain and read noise from their ‘‘intrinsic’’ values.

Substituting the definitions of \mathbf{s}_i and \mathbf{p}_i for a linear, stable (constant \mathbf{b}) detector from equations (5) and (7) into equation (36), we have

$$\begin{aligned}
 \mathbf{V}_i &= \frac{1}{\mathbf{g} n_r^2 \Delta t^2} \left\{ \sum_{i=1}^{n_r} \left(\frac{\mathbf{S}_{\text{net}}}{n_r} + (i - \mathbf{f}) \frac{2 \mathbf{S}_{\text{net}}}{n_r} \frac{\delta t}{\Delta t} \right) \right. \\
 &\quad \left. + 2 n_r \sum_{i=1}^{n_r} \left(\frac{\mathbf{S}_{\text{net}}}{n_r} + (i - \mathbf{f}) \frac{\mathbf{S}_{\text{net}}}{n_r} \frac{\delta t}{\Delta t} \right) \right. \\
 &\quad \left. - 2 \sum_{i=1}^{n_r} i \left(\frac{\mathbf{S}_{\text{net}}}{n_r} + (i - \mathbf{f}) \frac{2 \mathbf{S}_{\text{net}}}{n_r} \frac{\delta t}{\Delta t} \right) + \frac{2 n_r}{\mathbf{g}} \sigma_{\text{read}}^2 \right\}. \tag{37}
 \end{aligned}$$

Using the relations

$$\sum_{i=1}^n i = \frac{1}{2} n(n+1), \quad \sum_{i=1}^n i^2 = \frac{1}{6} n(n+1)(2n+1), \tag{38}$$

we can simplify equation (37):

$$\mathbf{V}_i = \frac{\mathbf{S}_{\text{net}}}{\mathbf{g} n_r \Delta t^2} \left[1 - \frac{1}{3} \frac{\delta t}{\Delta t} \frac{(n_r^2 - 1)}{n_r} \right] + \frac{2 \sigma_{\text{read}}^2}{\mathbf{g}^2 n_r \Delta t^2} \tag{39}$$

$$= \frac{\mathbf{I}}{\mathbf{g} \Delta t} \left[1 - \frac{1}{3} \frac{\delta t}{\Delta t} \frac{(n_r^2 - 1)}{n_r} \right] + \frac{2 \sigma_{\text{read}}^2}{\mathbf{g}^2 n_r \Delta t^2}. \tag{40}$$

Note that the factor \mathbf{f} drops out of the variance estimate.

We now consider the case of n_c co-additions. The variance then becomes

$$\mathbf{V}_i = \frac{1}{n_c^2} \sum_{i=1}^{n_c} \mathbf{V}_i \tag{41}$$

$$= \frac{1}{n_c^2} \sum_{i=1}^{n_c} \left(\frac{\mathbf{S}_{\text{net},i}}{\mathbf{g} n_r \Delta t^2} \left[1 - \frac{1}{3} \frac{\delta t}{\Delta t} \frac{(n_r^2 - 1)}{n_r} \right] + \frac{2 \sigma_{\text{read}}^2}{\mathbf{g}^2 n_r \Delta t^2} \right) \tag{42}$$

$$= \frac{\mathbf{S}_{\text{tot}}}{\mathbf{g} n_r n_c^2 \Delta t^2} \left[1 - \frac{1}{3} \frac{\delta t}{\Delta t} \frac{(n_r^2 - 1)}{n_r} \right] + \frac{2 \sigma_{\text{read}}^2}{\mathbf{g}^2 n_r n_c \Delta t^2} \tag{43}$$

$$= \frac{\mathbf{I}}{\mathbf{g} n_c \Delta t} \left[1 - \frac{1}{3} \frac{\delta t}{\Delta t} \frac{(n_r^2 - 1)}{n_r} \right] + \frac{2 \sigma_{\text{read}}^2}{\mathbf{g}^2 n_r n_c \Delta t^2}. \tag{44}$$

Finally, let us define the *effective read noise* $\sigma_{\text{read,eff}}$ as

$$\sigma_{\text{read,eff}} = \frac{\sqrt{2} \sigma_{\text{read}}}{\sqrt{n_r n_c}}, \tag{45}$$

which has the units of electrons. The variance on \mathbf{I} is then

given by

$$\mathbf{V}_I = \frac{\mathbf{I}}{gn_r \Delta t} \left[1 - \frac{1}{3} \frac{\delta t}{\Delta t} \frac{(n_r^2 - 1)}{n_r} \right] + \frac{\sigma_{\text{read,eff}}^2}{\mathbf{g}^2 \Delta t^2}. \quad (46)$$

Systematic effects produce additional terms in the equations for the variances. If these effects are present but not accounted for (as in our analysis), they will manifest themselves as an increased (i.e., larger than intrinsic) or perhaps variable “read noise” (see § 4).

3. CONTINUOUS SAMPLING TECHNIQUE

3.1. Ideal Detector

The notation given in § 2 can also be used to estimate the pixel variances of an ideal detector read out with the “continuous sampling” (also known as the line-fitting or “up the ramp”) technique (Chapman et al. 1990; Finger et al. 2000). In this technique, the array is read out repeatedly during the entire exposure. The count values from the readouts are then fitted with a straight line, and the slope gives the estimated source flux \mathbf{I} (in DN s⁻¹). For n_r reads, each of duration δt during an exposure of Δt , the slope can be written as follows (Bevington 1969):

$$\mathbf{I} = \frac{n_r \sum_{i=1}^{n_r} \mathbf{t}_i \mathbf{s}_i - \sum_{i=1}^{n_r} \mathbf{t}_i \sum_{i=1}^{n_r} \mathbf{s}_i}{n_r \sum_{i=1}^{n_r} \mathbf{t}_i^2 - (\sum_{i=1}^{n_r} \mathbf{t}_i)^2}, \quad (47)$$

where \mathbf{t}_i is the time of the i th readout, and \mathbf{s}_i is the count value recorded during this readout. This equation assumes that all readouts are given equal weight in the fitting process. If the readouts are equally spaced in time, $\mathbf{t}_i = (i - \mathbf{f})\delta t$, $\Delta t = (n_r - 1)\delta t$, and the above equation can be simplified to yield

$$\mathbf{I} = \frac{\sum_{i=1}^{n_r} \mathbf{s}_i [i - (n_r + 1)/2]}{\alpha}, \quad (48)$$

where $\alpha = n_r(n_r + 1)\Delta t/12$. (We refer the reader to the report by Sparks [1998] for the case in which the readouts are not equally spaced in time.) For a perfectly linear detector exposed to a source with constant flux, the values recorded at the end of each individual read are given by

$$\mathbf{s}_i = \mathbf{I} \mathbf{t}_i = (i - \mathbf{f})\mathbf{I}\delta t. \quad (49)$$

3.2. Nonlinearity Corrections

Nonlinearity corrections are fairly straightforward to implement in this readout technique. Again, we assume that the bias level is constant, the bias and dark current have been subtracted, and systematic effects have been accounted for before the nonlinearity corrections are applied. Equation (49) yields the first

estimate of the signal values $\mathbf{s}_i^{(1)}$ for each read. The nonlinearity-corrected signal values can then be estimated in a manner similar to that for Fowler sampling:

$$\mathbf{s}_i^{(2)} = \mathbf{s}_i^{(1)} \mathbf{C}_{\text{nl}}(\mathbf{s}_i^{(1)}). \quad (50)$$

Once the signal values for all i readouts have been corrected, the corrected slope can be redetermined from equation (48). The procedure can then be repeated to yield more accurate values of \mathbf{s}_i and \mathbf{I} , until a suitable convergence criterion is reached.

3.3. Variances

From equation (48) the variance in the slope recorded by a linear array can be calculated in a manner similar to that used to determine the uncertainties in the MCDS estimate of \mathbf{I} :

$$\mathbf{V}_I = \sum_{i=1}^{n_r} \sigma_{\mathbf{s}_i}^2 \left(\frac{\partial \mathbf{I}}{\partial \mathbf{s}_i} \right)^2 + 2 \sum_{j=2}^{n_r} \sum_{i < j} \text{Cov}(\mathbf{s}_i, \mathbf{s}_j) \left(\frac{\partial \mathbf{I}}{\partial \mathbf{s}_i} \right) \left(\frac{\partial \mathbf{I}}{\partial \mathbf{s}_j} \right) \quad (51)$$

$$= \sum_{i=1}^{n_r} \left(\frac{\mathbf{s}_i}{\mathbf{g}} + \frac{\sigma_{\text{read}}^2}{\mathbf{g}^2} \right) \left[\frac{i - \frac{1}{2}(n_r + 1)}{\alpha} \right]^2 + 2 \sum_{j=2}^{n_r} \sum_{i=1}^{j-1} \frac{\mathbf{s}_i}{\mathbf{g}} \left[\frac{i - \frac{1}{2}(n_r + 1)}{\alpha} \right] \left[\frac{j - \frac{1}{2}(n_r + 1)}{\alpha} \right]. \quad (52)$$

Using the definition of \mathbf{s}_i above (eq. [49]), along with the relations given in equation (38) and the additional definitions

$$\sum_{i=1}^n i^3 = \frac{n^2(n+1)^2}{4}, \quad \sum_{i=1}^n i^4 = \frac{n}{30} (n+1)(2n+1)(3n^2+3n-1), \quad (53)$$

we find

$$\mathbf{V}_I = \frac{\mathbf{I} \delta t n_r (n_r^4 - 1)}{\mathbf{g} 120 \alpha^2} + \frac{\sigma_{\text{read}}^2 n_r (n_r^2 - 1)}{\mathbf{g}^2 12 \sigma^2} \quad (54)$$

$$= \frac{6}{5} \frac{\mathbf{I}}{gn_r \Delta t} \left(\frac{n_r^2 + 1}{n_r + 1} \right) + \frac{12 \sigma_{\text{read}}^2}{\mathbf{g}^2 n_r \Delta t^2} \left(\frac{n_r - 1}{n_r + 1} \right) \quad (55)$$

$$= \frac{6}{5} \frac{\mathbf{I}}{gn_r \Delta t} \left(\frac{n_r^2 + 1}{n_r + 1} \right) + \frac{6 \sigma_{\text{read,eff}}^2}{\mathbf{g}^2 \Delta t^2} \left(\frac{n_r - 1}{n_r + 1} \right), \quad (56)$$

where we have used the definition of the effective read noise given earlier (eq. [45]). These equations are identical to that given by Garnett & Forrest (1993) once the different definitions of the exposure time are taken into account. However, they are substantially different from that derived by Chapman et al.

(1990), who did not account for the correlation of the individual readouts in their analysis.

4. VERIFICATION AND APPLICATION

To verify that the relation between observed count rate and variance predicted by equation (44) for the MCDS technique yields realistic estimates of the variances, we obtained a series of flat-field exposures with the SpeX instrument (whose array is read out using Fowler sampling) in the SXD mode (Rayner et al. 2003). (Unfortunately, because we lacked easy access to an array that is read out with the continuous readout technique, we were unable to obtain similar data to verify the procedures and equations given in § 3.) Details about the readout procedures and electronics for SpeX can be found in Rayner et al. (2003) and references therein. Because SpeX is a cross-dispersed spectrograph, the flat fields exhibit a large range in observed count values, which makes them ideal for our purposes. Two series of exposures were taken, with $\Delta t = 17$ s, $\delta t = 0.51$ s, and successive n_r values of 1, 4, 8, 16, 32, 1 and 32, 16, 8, 4, 1, 32. In all cases, $n_c = 1$. For each n_r , 19 flat-field exposures were obtained, from which a mean count rate and a standard deviation were computed at each pixel. Non-linearity corrections were then applied to the individual frames, and the means and standard deviations were recomputed at each pixel.

For plotting purposes only, we then computed for each n_r the array-averaged mean values and standard deviations of the variances V_r within bins of the observed count values S_{tot}/n_r that were 50 DN wide. In Figure 3a, the mean variances are plotted versus the mean count values for one of the $n_r = 1$ sets of flats. The results from the original data are shown, as well as those obtained after applying linearity corrections. The same plot for one of the $n_r = 32$ sets of flats is shown in Figure 3b. As can be seen from these figures, a linear trend of the variances of the nonlinearity-corrected data with the observed counts, like that predicted by equation (44), provides a very good representation of the results.

Equation (44) was then fitted to the actual (i.e., unbinned) flat-field data with a least-squares procedure in which all data points were equally weighted, to determine the array-averaged gain g and read noise σ_{read} for each n_r value. We included only those pixels with values between $0 < S_{\text{tot}}/n_r < 4000$ DN in our analysis. Given the exposure parameters for the flat fields, this range of values represents most of the usable well for the SpeX array. SpeX documentation strongly advises users to keep values of S_{tot}/n_r below 4000 for all observations. Because the nonlinearity corrections depend on n_r , the corrections for the $n_r = 1$ data sets are less than $\sim 5\%$, while the nonlinearity corrections for the $n_r = 32$ data sets are less than about 15% over this range. (These estimates have been derived from the SpeX nonlinearity curve and simulations we have performed with SpeX flat-field images.) The best fits are shown as solid

lines in Figures 3a and 3b. Since most of the data points have values below 1500 DN, the fits are heavily biased to these lower values. (The distribution of S_{tot}/n_r values in the flat fields are shown at the top of the two plots.)

The gain and read noise values resulting from the fits are shown in Figure 4, plotted as a function of the number of reads. The values for the gain (Fig. 4a) were found to be reasonably consistent with one another, with a mean of $12.1 e^- \text{DN}^{-1}$, although there is some indication of a slight decrease in g as n_r increases. We note that the smallest gain values were derived from the first images in both series of exposures, which were taken after the instrument and detector had been sitting idle for a long period. As seen in Figure 4b, however, there is a clear correlation between σ_{read} and n_r . The variation in σ_{read} with n_r is well described by a power law with an index of 0.16:

$$\sigma_{\text{read}} \approx 36n_r^{0.16}. \quad (57)$$

This relation implies that the effective read noise ($\sigma_{\text{read, eff}}$, eq. [45]) does not decrease as the square root of the number of reads, as might be expected for a constant σ_{read} , but rather decreases at a much slower rate, as $n_r^{-0.34}$. This is demonstrated in Figure 5, where we plot the effective read noise determined from the flat-field data and the best-fitting power law.

An alternative description of the variation of the effective read noise with the number of reads is the following:

$$\sigma_{\text{read, eff}} = An_r^{-0.5} + \text{constant}, \quad (58)$$

where A is a constant proportional to the read noise σ_{read} . In this formulation, the read noise σ_{read} does have a constant value, and the effective read noise does decrease as $n_r^{-0.5}$, but there is an additional constant noise source that results in an additive offset. A fit of this form to the SpeX data is also shown in Figure 5, where it can be seen that it also provides an acceptable representation of the data.

The variation of the effective read noise with n_r determined from the SpeX data appears remarkably similar to that presented in Fig. 1 of Garnett & Forrest (1993). Although these authors claim that the effective read noise decreases as $n_r^{-0.5}$, close inspection of their figure reveals that a power law with $n_r^{-0.5}$ cannot reproduce the data points without an additive constant offset (eq. [58]). A power law with a smaller index of about -0.4 would probably also fit their data well.

As mentioned in § 2, variation in the read noise with n_r is a sign that systematic effects in the behavior of the array or the readout procedure are present but not accounted for in our analysis. Although there are several possible candidates for such an overlooked systematic effect (see § 2), the exact cause for the observed variation of the read noise with the number of reads is uncertain at present. The variation is probably not a result of the nonlinearity corrections, as the fits are biased to a range of data values for which the nonlinearity corrections

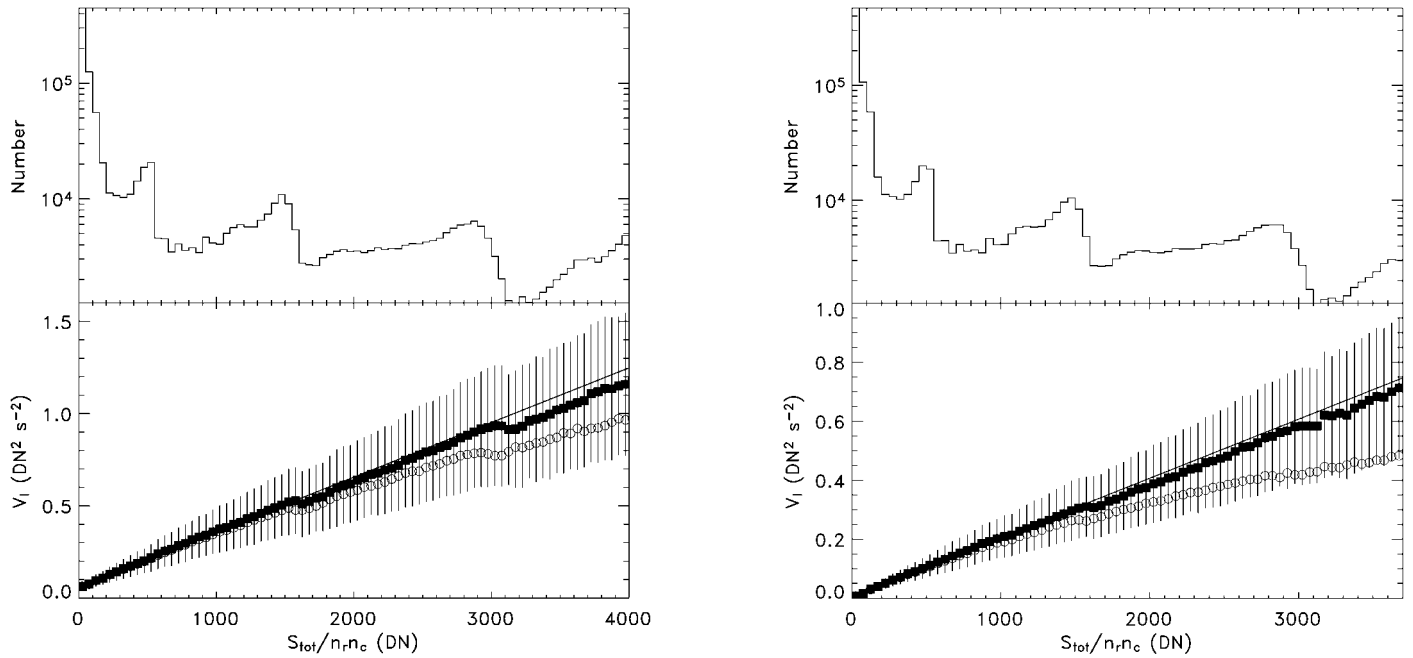


FIG. 3.—Measured mean variances and standard deviations as a function of count value for two sets of flat-field data $n_r = 1$ (left) and $n_r = 32$ (right). The values have been computed in bins 50 DN wide. The open points are the original values, while the solid points represent the values measured after making nonlinearity corrections to the individual frames (see text). The solid line represents a fit of eq. (47) to the nonlinearity-corrected data points. Since most of the data have count values below 1500 DN, the fit is heavily biased to match the variances in this range. The histogram of data values is shown at the top of both plots.

are negligible. Furthermore, variations in the read noise with the number of reads have been observed before by others. For example, Finger et al. (2000) have suggested that the deviation from $n_r^{-0.5}$ in the effective read noise measurements of the Aladdin array installed in the ISAAC instrument at the VLT is due to additional Poisson noise from multiplexer glow and dark current. As the dark current is implicitly included in our formulation for the variance, increased dark current in the SpeX array is probably not the source of the additional noise. Furthermore, it is straightforward to include an additional source term in equations (5)–(9) to represent amplifier glow, which is present only during the readouts. This term leads to an additional noise term in the expression for the variance that increases as n_r , which is significantly larger than the variation we observe.

However, additional noise might result from the slight heating of the array due to the numerous reads. In order to reduce residual image effects, the SpeX array is read out at 1 Hz when not integrating (Rayner et al. 2003). The act of reading out the array dissipates 20 mW. An exposure consists of an initial set of reads of the pedestal (heating), followed by a period of almost no power dissipation during the actual integration (cooling), terminated by another set of reads of the signal (heating). Although the array mount temperature is controlled to ± 0.01 K, power is dissipated locally in the array, so heating effects due to the readout might possibly cause much larger temperature varia-

tions across the array. Changes in temperature (or the resulting instabilities in the readout process) as a function of the number of reads would lead to a variation in the dark current rate and/or the bias level during the readouts and therefore might explain the observed variation in read noise (see, e.g., Böker et al. 2001). This may well be a common phenomenon in many readout techniques for NIR arrays.

Because we have not explicitly included these systematic effects, or the corrections for nonlinearity, in our equations, we should point out that the gain and read noise derived from our fits are not necessarily the “true,” or intrinsic, values for the array. The nonlinearity corrections of the individual images, for example, increase the variance of any given pixel value above that expected for the same pixel value recorded on an intrinsically linear array. Therefore, the estimated gain and read noise values determined from the fits will generally be smaller and larger, respectively, than the intrinsic values for the array. As far as data reduction is concerned, however, this does not pose a problem, as long as all images subject to the same systematic effects are reduced with the same parameter values. For example, the *Spextool* software first corrects each individual image for nonlinearity before generating a variance image or proceeding with any reduction steps (see Cushing et al. 2003). The increased variances due to the nonlinearity corrections are then correctly accounted for by adopting the gain and read noise values derived from our linear fits. Hence, these

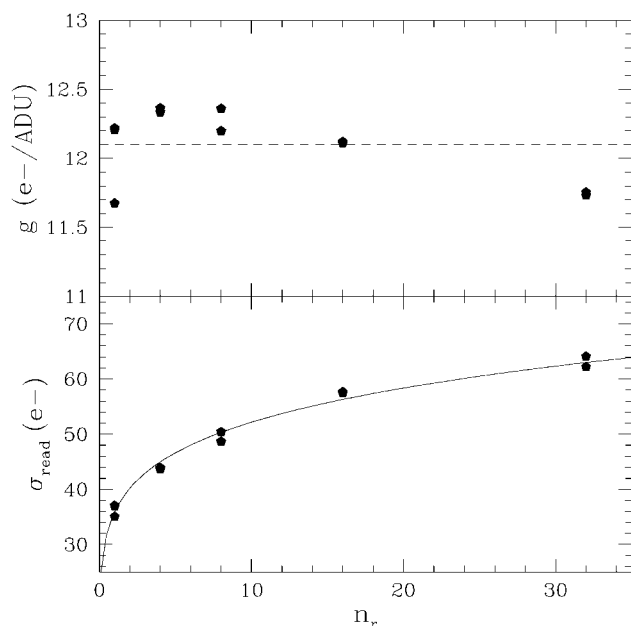


FIG. 4.—Gain (in electrons/DN; *top*), and read noise (in electrons, *bottom*) measured for the SpeX array as a function of the number of reads. The dashed line is the estimated mean value of the gain (12.1); the solid line is a fit of a power law to the read noise as a function of the number of reads. The power law index is 0.16.

gain and read noise values are necessary for consistency with the data reduction routines.

5. SUMMARY

We have derived fairly simple equations for the estimated variance as a function of the observed count rate at each pixel in NIR arrays employing correlated double sampling, multiple correlated double sampling, and equally spaced continuous sampling readout techniques. We also give a general prescription for implementing corrections for nonlinearity at each pixel. We have applied these equations to sets of flats obtained with the SpeX instrument at the NASA Infrared Telescope Facility and find that they provide a good representation to the nonli-

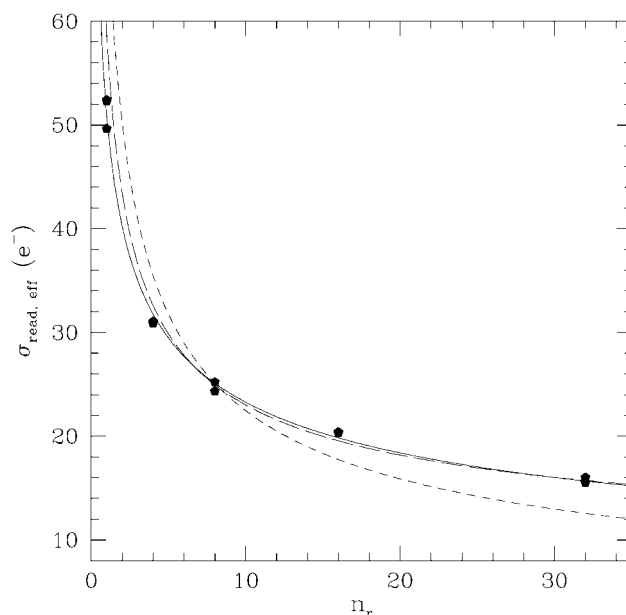


FIG. 5.—Effective read noise as a function of the number of reads. The solid line is the best-fit power law, which decreases as $n_r^{-0.34}$, significantly shallower than expected; the short-dashed line is the expected decrease as $n_r^{-0.5}$, normalized to the solid fitted curve at $n_r = 8$. The long-dashed line is the alternative fitting form in which the effective read noise decreases as $n_r^{-0.5}$, plus an additive constant.

nearity-corrected data. We also find that the read noise of the SpeX array varies with the number of reads as $n_r^{0.16}$, perhaps due to heating of the array as a result of multiple reads. Such heating may be generic to NIR arrays that employ (M)CDS as the readout technique.

W. Vacca thanks Alan Tokunaga and James Graham for support and for reading an earlier draft of this paper. M. Cushing acknowledges financial support from the NASA Infrared Telescope Facility. We thank J. Leong and P. Onaka for acquiring some of the SpeX flat fields used in our analysis. We also thank P. Onaka for discussions regarding the operation of NIR array electronics and readout techniques.

REFERENCES

- Bevington, P. R. 1969, *Data Reduction and Error Analysis for the Physical Sciences* (New York: McGraw-Hill)
- Böker, T., et al. 2001, *PASP*, 113, 859
- Chapman, R., Beard, S., Mountain, M., Pettie, D., & Pickup, A. 1990, *Proc. SPIE*, 1235, 34
- Cushing, M. C., Vacca, W. D., & Rayner, J. T. 2003, *PASP*, 116, 362
- Finger, G., Mehrgan, H., Meyer, M., Moorwood, A. F. M., Nicolini, G., & Stegmeier, J. 2000, *Proc. SPIE*, 4008, 1280
- Fowler, A. M., & Gatley, I. 1990, *ApJ*, 353, L33
- Garnett, J. D., & Forrest, W. J. 1993, *Proc. SPIE*, 1946, 395
- Kobayashi, N., et al. 2000, *Proc. SPIE* 4008, 1056
- McCaughrean, M. J. 1988, Ph.D. thesis, Univ. Edinburgh
- McLean, I. S., et al. 1998, *Proc. SPIE*, 3354, 566
- Moorwood, A. F., 1997, *Proc. SPIE*, 2871, 1146
- Rayner, J. T., Toomey, D. W., Onaka, P. M., Denault, A. J., Stahlberger W. E., Vacca, W. D., Cushing, M. C., & Wang, S. 2003, *PASP*, 115, 362
- Roye, E., et al. 2003, *NICMOS Instrument Handbook, Version 6.0* (Baltimore: STScI)
- Skinner, C. J., & Bergeron, L. 1997, *Characteristics of NICMOS Detector Dark Observations* (Instrum. Sci. Rep. 97-026; Baltimore: STScI)
- Sparks, W. B. 1998, *Statistics of Accumulating Signal* (Instrum. Sci. Rep. 98-008; Baltimore: STScI)