Fundamental performance differences of CMOS and CCD imagers: Part VI

James Janesick[*a] Tom Elliott[a],
James Andrews[b], John Tower[b]
[a] SRI-Sarnoff, 4952 Warner Av., Suite 300, Huntington Beach, CA. 92649
[b] SRI-Sarnoff, CN5300, 201 Washington Road, Princeton, NJ 08543-5300

ABSTRACT

Past papers demonstrated advancements made on developing scientific PMOS/NMOS CMOS imagers that match or exceed CCD performance. New data and discussions presented in this paper present further progress on subject matters that include: 1). subcarrier read noise performance with understandings for how the noise floor can be reduced further, 2). comprehensive correlated double sampling (CDS) signal processing noise fundamentals in response to random telegraph and flicker noise sources, 3). high energy radiation damage test data from NASA's BSI SoloHi/WISPR CMOS imager and 4). update on a new scientific BSI PMOS/NMOS stitched Mk x Nk x 10 um 5TPPD pixel imager being fabricated for Lawrence Livermore National Lab (LLNL) and NASA's Europa Clipper mission.

Keywords: CMOS and CCD scientific imagers.

1. OPTIMUM CDS PROCESSING OF 1/f AND RTS NOISE

1.1 Introduction

CCD and CMOS imagers have achieved a remarkable low read noise level of ~ 1 carrier rms. The limiting read noise sources are flicker noise (1/f) and random telegram noise (RTN) generated by the pixel's output source follower (SF) amplifier. This section updates our physical understandings about these two noise sources and how correlated double sampling (CDS) responds to them in hopes to reduce noise further (i.e., sub carrier rms). We will approach this study analytically and simulation style. Experimental data will then be presented to confirm that the detailed analysis performed is on track. Discussions will then focus on where the two noise sources are located within the pixel's output amplifier and present design and process remedies to hopefully lower them. First, a few CDS equations that predict read noise performance given that 1/f and RTN are present.

For upcoming discussions a correlated double sampling (CDS) circuit used by our CMOS imagers is presented in Fig. 1. The general read noise equation employed to determine CDS output noise for this circuit is given by [1],

$$R(e-/h+) = \frac{1}{S_V\ A_{SF}\ (1 - e^{-t_s/\tau_D})}\ [\int_0^\infty |S_{DET}(f) H_{CDS}(f)|^2\ df]^{1/2} \quad (1)$$

where R(e-/h+) is the read noise (rms e- or rms h+), S_V is the sense node sensitivity of the detector (V/e- or V/h+), A_{SF} is the pixel's SF gain (V/V), t_s is the sample-to-sample time or often called clamp-to-sample time if a clamp is used for the first sample (sec), τ_D is the dominate time constant of the CDS processor and f is frequency (Hz). $H_{CDS}(f)$ is the CDS transfer function given by,

$$H_{CDS}(f) = \left(\frac{1}{1 + (2\ \pi\ f\ \tau_D)^2}\right)(2\ -\ 2\ \cos(2\ \pi\ f\ t_s)) \quad (2)$$

The second term in Eq. 2 is the CDS transfer function and the first term sets the CDS equivalent noise bandwidth by adjusting τ_D (i.e., B=1/4τ_D Hz).

The pixel's output noise power spectrum, $S_{DET}(f)$, input to the CDS processor is described by,

[*]e-mail: CMOSCCD@AOL.COM, Paper Number: 9591-1 , San Diego, 9-13 August 2015 .

$$S_{DET}(f) = \frac{a}{f^k} + \sum_1^n \frac{2\tau_{RTN\,n}}{4 + (2\pi\tau_{RTNn})^2} + W \qquad (3)$$

where the first term is 1/f noise power of magnitude 'a' with slope 'k' and the second term is collective sum of RTN noise sources within the CDS passband each with a different characteristic time constant τ_{RTN}. The last term is the white noise power. In general, CCD read noise performance is limited by 1/f noise assuming white noise is lowered by sufficiently reducing the bandwidth through τ_D until 1/f noise dominates. Discussions below will assume that white noise content is negligible. In contrast CMOS read noise is limited by 1/f and RTN without significant white noise present. Note also, unlike the CCD which typically has one output amplifier, each CMOS pixel has its own output amplifier and different $S_{DET}(f)$ spectrum. The physical differences between 1/f and RTN will be discussed below.

Figure 2 plot the CDS frequency transfer curve using Eq. 2 on log scales for three different t_s and τ_D. The curves assume $t_s = 2\tau_D$ for this will be a special timing relationship in achieving optimum signal-to-noise (S/N) performance as examined below. Notice that the CDS is a band-pass (BP) filter centered at a frequency of BP~$1/2\pi\tau_D$ Hz with it's gain dropping by 20 dB/octave on either side. Figure 3a plots the CDS transfer function for different τ_D (10, 0.5, 0.1, 0.01 us) assuming t_s is fixed at 1 us. The signal gain suffers if τ_D is set too high (i.e., BP is too low) whereas excessive noise comes through the CDS if τ_D is set too low (i.e., BP is too high). Setting $t_s = 2\tau_D$ appears to be a good compromise in adjusting the bandwidth for a good S/N ratio (the signal gain is only down by 14 %). Figure 3b keeps τ_D fixed at 1 us and varies t_s (100, 10, 2, 1 us). Here we find a similar outcome i.e., $t_s > 2$ us yields excess noise and $t_s < 2$ us yields insufficient signal gain. Optimum S/N appears to also follow $t_s = 2\tau_D$.

1.2 Flicker noise

We now introduce 1/f noise to the CDS processor and adjust t_s and τ_D to find the timing combination that realizes the best S/N result. Figure 4a plots simulated 1/f noise as a function of time displaying high frequency noise that rides on low frequency excursions. The top plot of Fig. 4b magnifies the time scale of Fig. 4a by 1000x. Within this smaller time window the peak-to-peak noise is significantly less than what Fig. 4a shows. In general 1/f noise excursions decrease as the time window shrinks. This unique 1/f characteristic and is unlike thermal white noise where equal amounts of noise frequencies exist for any time window size assuming the bandwidth is infinite. The lower plot of Fig. 4b is the output noise of the CDS filter (τ_D =5 us). It can be noted that noise decreases from 1 V to 0.41 V rms. The corresponding power frequency spectrum of the filtered noise is shown in Fig. 5a. The flicker noise slope drops at a 10 db/octave rate (1/f) until encountering the filter's break frequency of $1/2\pi\tau_D$ (31.8 kHz) at which point the noise drops at 30 dB/octave ($1/f^3$). The upper trace of Fig. 5b shows filtered 1/f noise (τ_D = 5 us) while the lower trace is the CDS output time response assuming $t_s = 2\tau_D$. The CDS rejects lower 1/f frequencies lowering the noise level from 0.41 V rms to 0.197 V rms. Figure 6a presents CDS frequency responses (τ_D =5 us and τ_D =0.5 us) along with filtered noise spectrums for comparison. The area under a CDS curve is proportional to read noise (carriers rms). It is difficult to say from these plots which response produces the lowest output noise. Figure 6b plots the same data on a linear vertical scale to better show that noise powers are approximately the same suggesting there is no difference in noise performance given that $t_s = 2\tau_D$.

Figures 7 plots CDS output noise as a function of t_s for various τ_D. For a specified τ_D the output noise increases with t_s because noise correlation is being lost between the two CDS samples. For very long t_s the CDS output noise will eventually level off to $2^{.5}$ times that of the input noise from the CDS filter output because the samples are uncorrelated and differenced. It is important to note from this plot that output noise is constant with t_s given that $t_s = 2\tau_D$ (i.e., the dotted horizontal line shown on the plot). Under this boundary condition we find that the noise decreases with T_D at the same rate as the noise increases with t_s. This is the central reason why it is difficult to change read noise in CCD and CMOS by CDS timing. The noise is also constant for $t_s = N\tau_D$ (N=3, 4, 5. .). However, additional CDS process pixel time is required which is usually not desirable for high speed camera systems.

If flicker noise does not exhibit a characteristic unity slope a different optimum timing adjustment other than $t_s = 2\tau_D$ may be necessary. For instance, Fig. 8a plots noise as a function of t_s for different τ_D for a 1/f noise slope of k= 1.2. A slope of > 1 physically implies that there are more low frequencies than higher ones. In this case the read noise decreases as the BW is decreased. Figure 8b shows the opposite situation for a 1/f noise slope of k=0.8. Here there are more high frequencies than lower ones which results in lower noise as the BW is reduced.

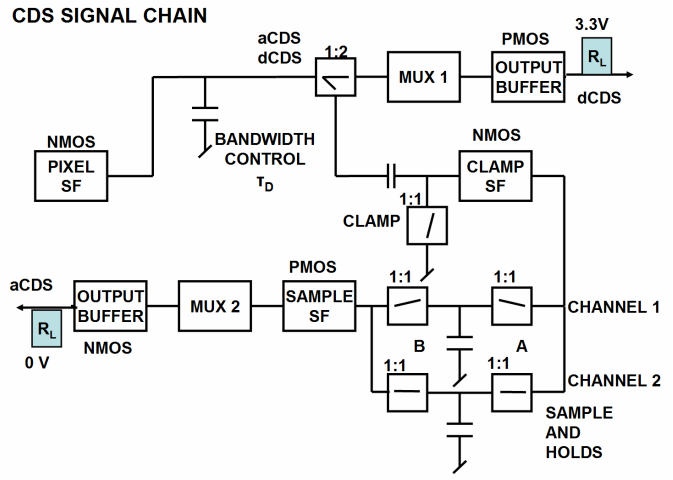

Fig. 1. Pixel CDS signal processor.

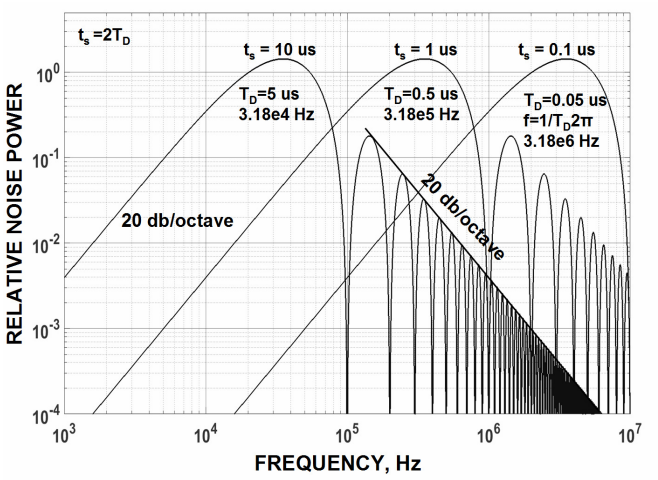

Fig. 2. CDS transfer function for three t_s.

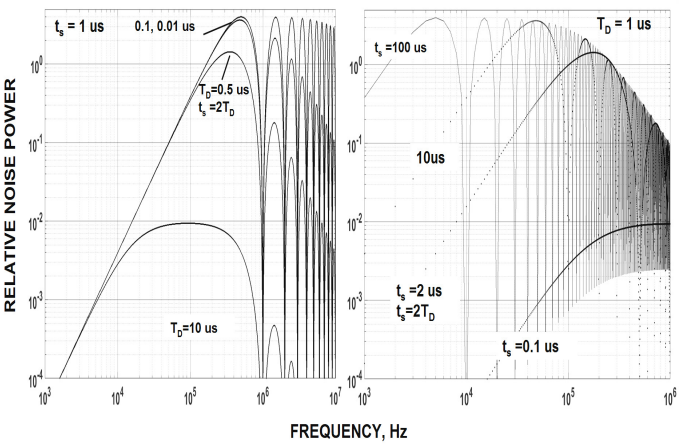

Fig. 3a,b. CDS transfer functions for different τ_D and t_s.

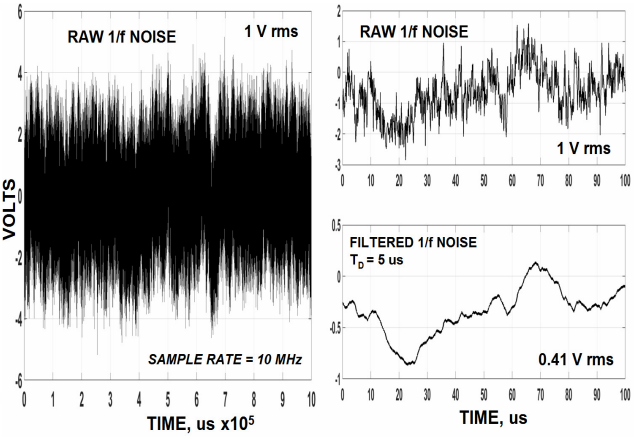

Fig. 4a,b. Raw 1/f noise and filtered 1/f noise.

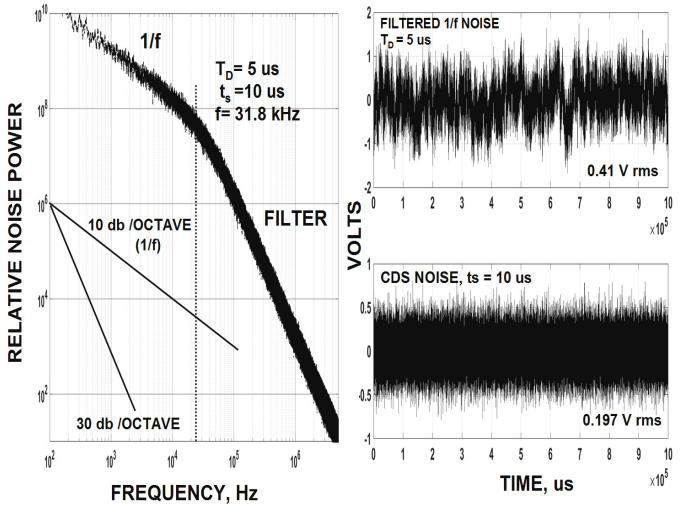

Fig. 5a,b. Filtered 1/f frequency response and CDS output noise.

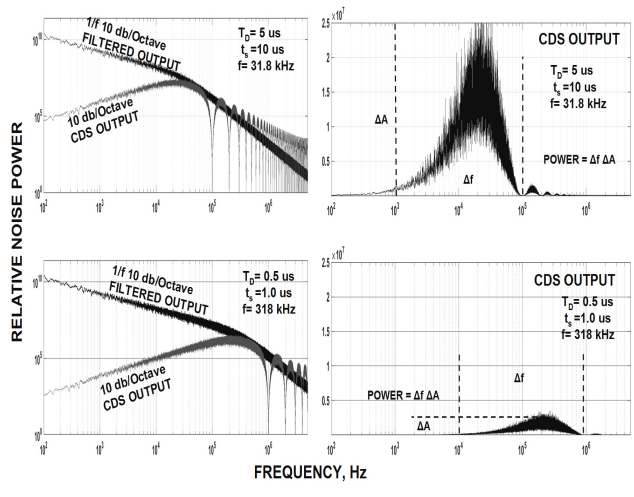

Fig. 6a/b. CDS frequency responses for τ_D= 5 us and 0.5 us.

1.3 Random telegraph noise

Let's now analyze RTN in a similar fashion as we did for 1/f noise above. The top plot of Fig. 9 shows simulated RTN with a time constant of τ_{RTN} = 1 us. The lower plot is the filtered response with τ_D =0.5 us. The filtering action reduces the 1 V rms input RTN down to 0.338 V rms output. Figure 10a presents three RTN sources summed together (τ_{RTN}=1, 10 and 100 us). Their corresponding individual frequency spectrums are displayed in Fig. 10b. Each response is initially level with frequency and fall at a 20 db octave rate (i.e., $1/f^2$) at a break frequency of $1/2\pi T_{RTN}$. Figure 11a shows the spectrum of two RTN sources when added together (τ_{RTN}=1 and 100 us). Figure 11b shows net noise when all three sources are summed producing a 1/f response. This plot provides an explanation to why flicker noise is fundamentally associated to multiple RTN sites working together in producing it's characteristic slope of unity.

The CDS response to RTN is presented in Fig. 12 as a function of t_s and τ_D for three different τ_{RTN} (τ_{RTN}=1, 10, 100 us). Note unlike 1/f noise the noise level for a given τ_{RTN} is not constant with t_s assuming $t_s = 2\tau_D$. Instead the noise peaks at $t_s = \tau_{RTN}$ where the RTN frequency is approximately centered within the CDS pass band. It is important to note that the noise change is very small as t_s varies about the peak. In practice we will see some RTN pixels will remain nearly constant with t_s where others will slightly increase and decrease depending on their τ_{RTN} characteristics. Figure 13 plots the CDS response to three RTN sources that produced the 1/f spectral response in Fig. 11b. The output noise is nearly constant with t_s just as 1/f noise exhibited above. Lastly Fig. 14a applies Eq. 1 to plot CDS noise with t_s for various τ_{RTN} showing individual noise peaks (at $t_s = \tau_{RTN}$). Figure 14b adds the RTN sources listed in Fig. 14a one at a time to demonstrate how the noise becomes insensitive to t_s timing as the input noise looks more and more 1/f like. In general, the average noise spectrum for CMOS imagers is very 1/f like and therefore not significantly sensitive to t_s timing as demonstrated in the next section.

1.4 Experimental data

Figure 15a presents photon transfer (PT) curves[2] generated by a row of 512 PMOS CMOS 5T pinned photo diode (PPD) pixels assuming τ_D =3.5us and t_s =48 us). A read noise of 1.7 h+ rms is measured based on a conversion constant of 0.15 h+/DN directly derived from the DN plot. Figure 15b presents a read noise histogram along with a Guassian curve fit displaying that some RTN is present (pure 1/f noise exhibits a Guassian distribution). It can be noted that the sample time of 35 us used is considerably longer than the desired timing of $t_s = 2\tau_D$ (7 us). This is because additional transfer gate (TG) overhead time is necessary after the clamp is released (i.e., charge transfer time and TG clock feedthrough settling time).

Stacked raw clamped dark video plots for a quantity of pixels used by the PT curve in Fig. 16a. The multiple row stacking response shown is how the output noise would appear on a high persistence oscilloscope before it enters a analog digital converter (ADC) with an internal sample and hold. The magnified plot presented in Fig. 16b shows timing starts by resetting and clamping pixel (the TG clock is absent in this measurement). The white noise seen in the clamp period is associated with the clamp switch (~ 0.1 h+ rms). Following clamp release and τ_D settling time the noise slowly increases with t_s characteristic of 1/f behavior (white noise would immediately stabilize to some level). Figure 17a plots read noise (rms h+) buildup for the first sixteen pixels shown in Fig. 16a with time. Figure 18 magnifies the time scale for the first two pixels (data points are taken every 1 us). It is important to note that the video should be sampled as soon as the clamp is released after the required video τ_D settling time is satisfied (i.e., $t_s = 2\tau_D$ = 7 us). This special timing condition is illustrated by the 'squares' on the plot. An average noise level of ~ 0.9 h+ rms is measured for the entire row of pixels. The data points with 'stars' include TG overhead time where a higher noise is determined (1.7 h+ rms). This is the same noise level measured in the PT curve in Fig. 15. It is interesting to point out that a 3T pixel can achieve lower noise than 4T pixels because the 3T does not require TG overhead time. This claim however highly depends on how 3T pixels are readout (most 3T applications can not benefit from the advantage).

The lower traces presented in Figure 19 reduce the bandwidth (τ_D =23 us). The upper traces assume τ_D =3.5 us as used above. The noise voltage decreases for all pixels since more filtering action takes place with the bandwidth reduction. Fig. 20 plots read noise as a function t_s with τ_D =23 us for 15 pixels. The data points indicated by 'squares' sample the noise for the ideal timing condition (i.e., $t_s = 2\tau_D$ = 46 us). Comparing Fig. 17 to Fig. 20 shows that noise levels for the two τ_D settings are virtually the same (i.e., 0.9 h+ rms in contrast to 0. 82 h+rms). In other words, increasing τ_D does not significantly lower CDS output noise level (maintaining $t_s = 2\tau_D$). Unlike 1/f noise, thermal white noise would decrease by a factor 2.5 (i.e., [23 /3.5]0.5). The analysis performed in Secs. 1.2 and 1.3 predict that τ_D would have little effect on read noise if ideal 1/f noise and RTN were present. However as already noted, TG overhead

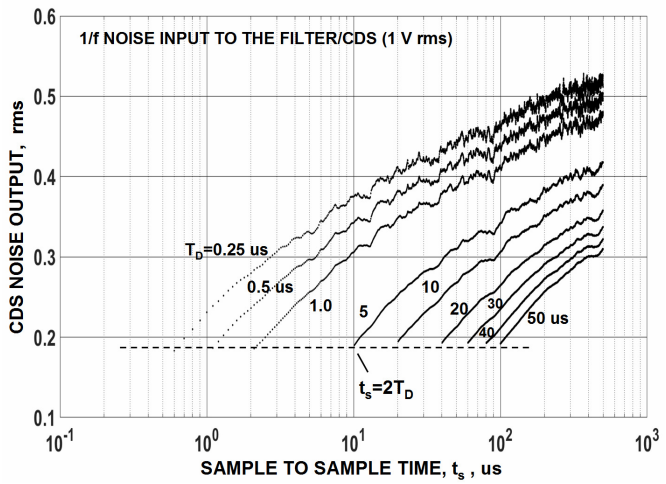

Fig. 7. CDS 1/f noise response as a function of t_s and T_D.

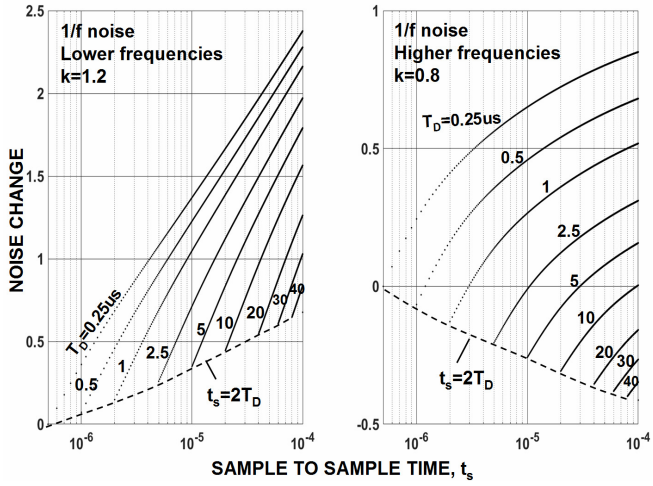

Fig. 8a/b. CDS noise change for different sloped 1/f noise.

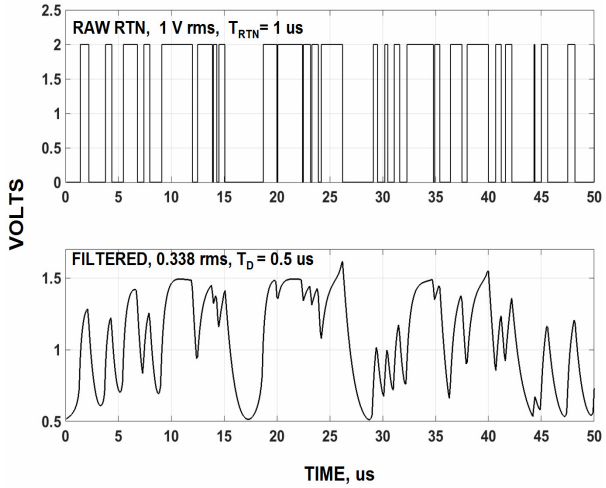

Fig. 9. RTN and filtered RTN.

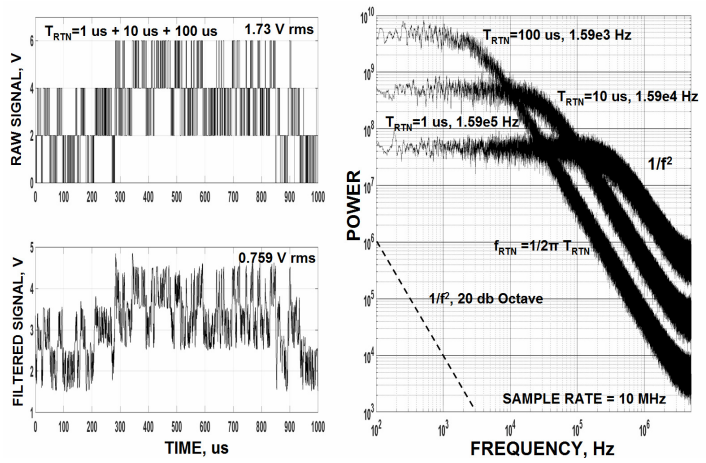

Fig. 10a/b. RTN time and frequency responses for three different τ_{RTN}.

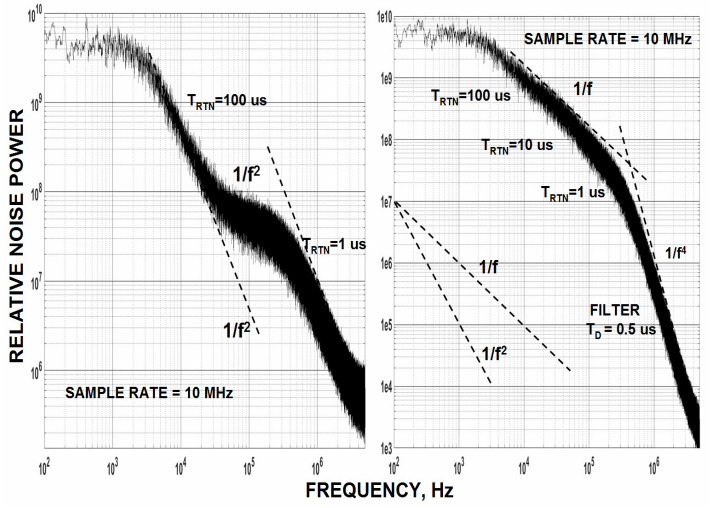

Fig. 11a/b. Spectral plots for two and three τ_{RTN}.

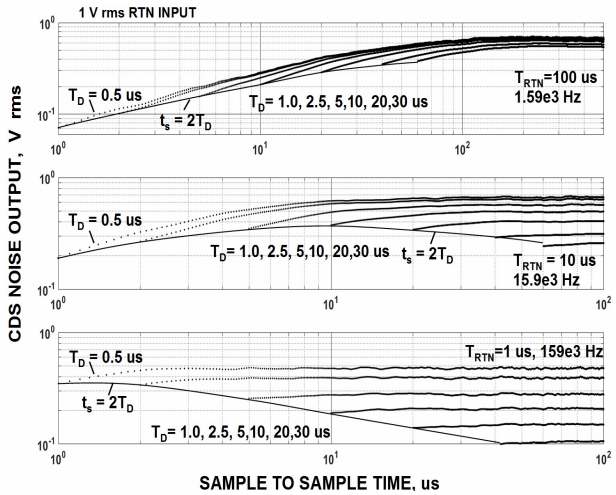

Fig 12. CDS RTN noise response as a function of ts, τ_D and τ_{RTN}.

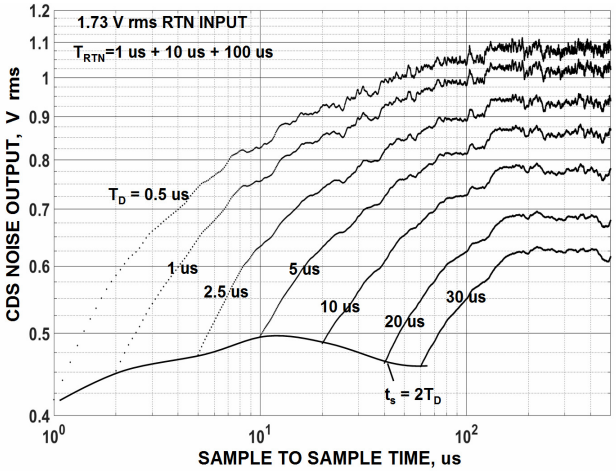

Fig. 13. CDS noise as function t_s and τ_D for three τ_{RTN}.

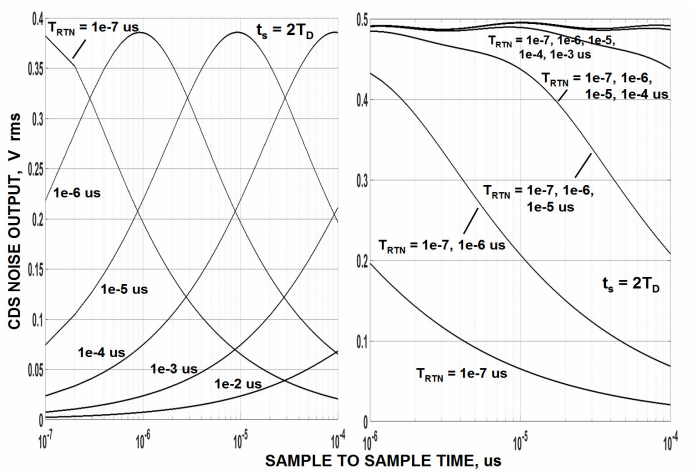

Fig. 14a/b. CDS noise as function t_s and τ_D for many different τ_{RTN}.

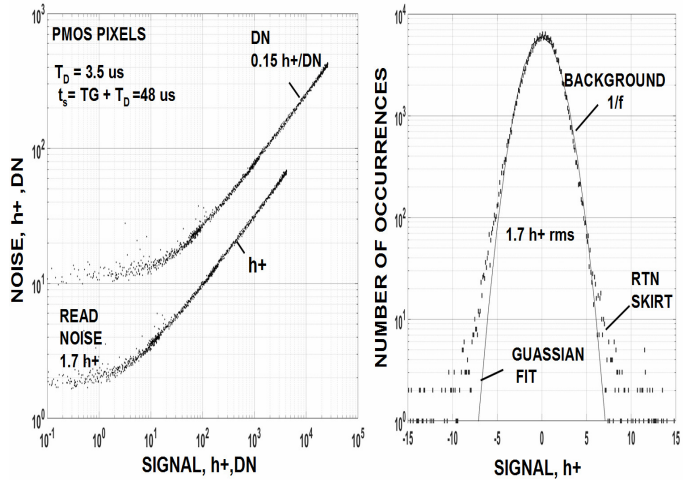

Fig. 15 a/b. PTC and read noise histogram.

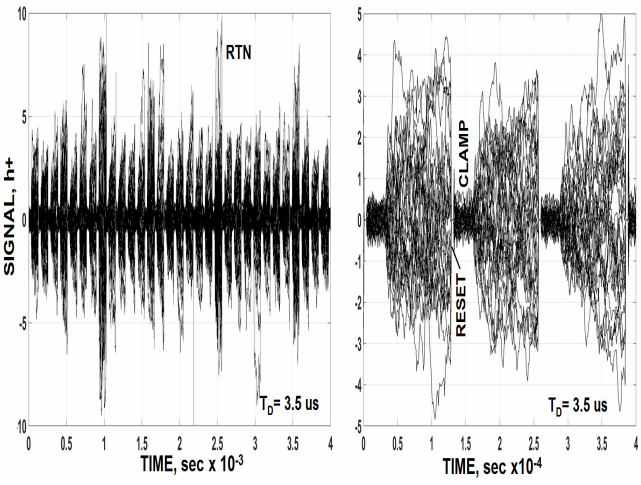

Fig. 16a/b. Clamped video for several pixels.

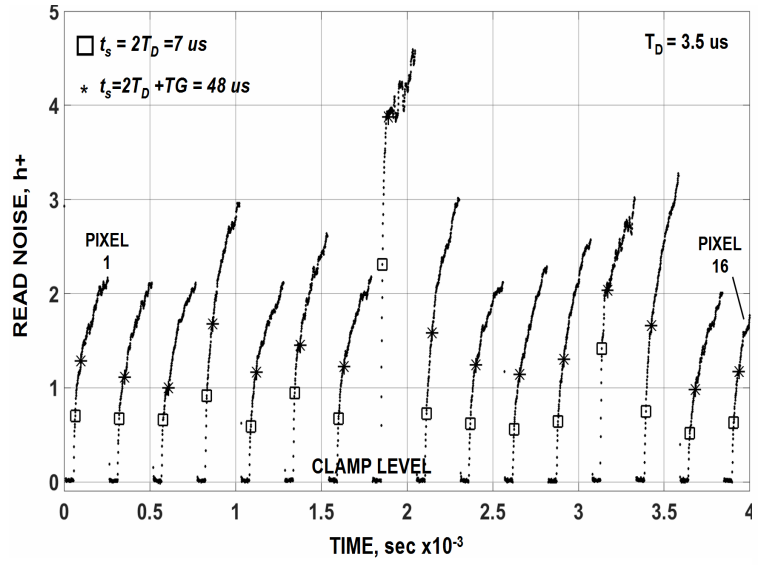

Fig. 17. Clamped video showing noise increase with time.

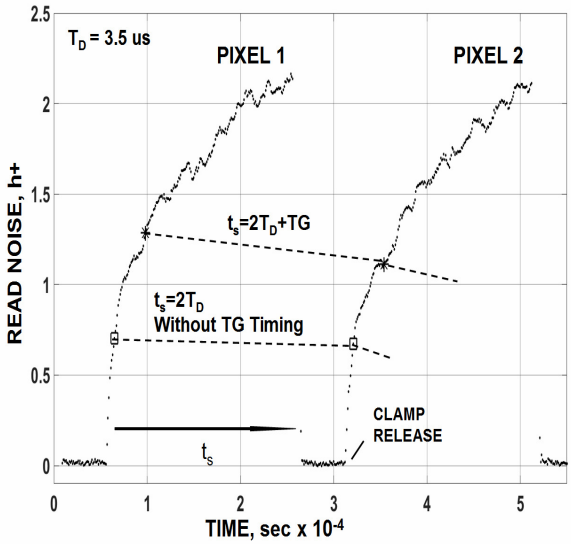

Fig. 18. Magnified view of Fig. 17.

time is required which will increase the noise from the ideal. Although TG clock width can remain fixed with τ_D, the TG settling time must increase proportionally to τ_D. Hence, for this reason it is desirable to keep τ_D small as long as $t_s = 2\tau_D$ is assumed. For CMOS imagers the lower τ_D limit is usually governed by the time constant associated with the SF output impedance and the video column bus capacitance.

Figure 21a are raw dark video traces from a single pixel showing RTN riding on 60 Hz background noise. The top plot of Fig. 21b displays a magnified view of the clamped response. The bottom plot of Fig. 21b shows a single noise trace and the samples taken 7 us after clamp is released (i.e., $t_s = 2\tau_D$). Notice that the samples hit and miss the RTN signal. For example, the first sample catches the RTN at the zero signal level whereas the second sample catches the RTN when it is in the high state. Figure 22a shows the raw video, clamped video and the sampled video signal plotted together. Respective noise levels are 6.68, 2.8 and 1.5 h+ rms. The clamping action eliminates 60 Hz noise since it's frequency is very low. The clamp also suppresses the RTN to 1.5 h+ rms through hit and miss action given the $t_s = 2\tau_D$ timing. Figure 22b presents a non clamped filtered frequency plot for this RTN pixel along with the CDS transfer function for reference. The RTN frequency response initially drops at a $1/f^2$ slope until the CDS filter is encountered at 45 kHz at which point the noise decreases with a $1/f^4$ slope. The RTN frequency (~16 kHz) is lower than the CDS bandpass frequency (~45kHz). Decreasing t_s will lower the noise for this input noise. Increasing t_s will at first increase the noise until RTN and CDS frequencies are equal whereupon the noise reaches a peak. Increasing t_s further will cause it to decrease again.

We can separate RTN pixels from 1/f noise pixels by examining their frequency characteristics. Figure 23 presents a collection of RTN pixels where both time and frequency responses are presented. For example RTN #1 is low frequency one (~1 kHz) that starts it's $1/f^2$ fall at 1 kHz until reaching a CDS cutoff frequency of 45 kHz at which point the response drops by $1/f^4$. RTN#2 frequency is the same as the CDS cutoff frequency such that the $1/f^2$ slope is not apparent. The response drops with $1/f^4$. RTN #3 is composed of at least two traps (~ 15kHz and 1 kHz). RTN #4 reveals two traps where a portion of spectrum follows 1/f slope but quickly goes into $1/f^2$ slope as frequency decreases. Figure 24a magnifies the time axis for the RTN #1 signal. Note that the RTN source is accompanied with background noise. The RTN signature would have only two distinct levels if it weren't for the fact that this background noise shakes it around. Figure 24b inspects this background noise in a time region where RTN is not active. For the most part this noise exhibits 1/f characteristics although there is still a small amount of RTN behavior present. This 1/f background noise is observed in every pixel. It sets the lower read noise limit for the sensor and camera system. Figure 25 presents a collection of pure flicker noise limited pixels. Note that the time response in Fig. 25a has the same appearance as the simulated flicker noise displayed in Fig. 4. The multiple stacked pixel frequency responses of Fig. 25b are tightly grouped together indicating there is little RTN present. The majority of pixels for the imager exhibit the same 1/f response with little rms noise variation from pixel to pixel. The read noise floor for these pixels will be essentially independent of clamp-to-sample timing (given $t_s = 2\tau_D$).

1.5 Discussion

It is generally believed that RTN is associated with traps within the gate oxide. Free carriers from the SF channel tunnel in and out of these traps at different rates depending on the tunneling distance to the trap and the barrier height encountered. The local 'on' and 'off' potential in the region of a trapped carrier modulates the channel current in a digital manner and is seen as RTN voltage frequency at the SF output. NMOS imagers have a great deal more RTN pixels than PMOS devices. For example, Fig. 26 compares the read noise for NMOS and PMOS pixels. Very few RTN pixels are seen for the PMOS device. The reason for the NMOS/PMOS RTN population difference is holes are much less likely to tunnel into oxide traps and generate RTN than electrons because the barrier height is greater.[3]

There are two theories that describe flicker noise generation: Hooge's mobility fluctuation and McWhorter's carrier density (or number fluctuation) models.[4,5] The latter mechanism is based on the random trapping and release of conduction band carriers located at the $Si\text{-}SiO_2$ interface of the SF MOSFET (i.e., generation recombination (GR) noise). The former is related to conductance or resistance fluctuations within the channel of the SF MOSFET. There are several possible causes for channel mobility noise. For example, fluctuations may be associated with a lower surface mobility compared to bulk silicon because carriers scatter off the $Si\text{-}SiO_2$ interface. Another scattering source could be linked to the channels implanted impurities that have not been fully thermally activated into the silicon lattice properly.

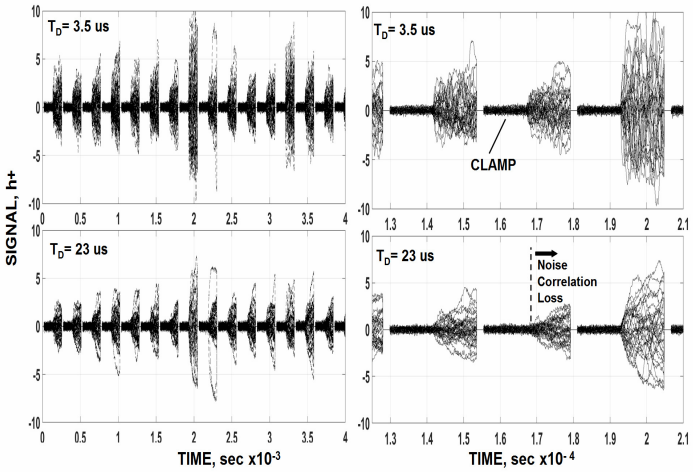

Fig. 19. Clamped video comparison for two τ_D (3.5 and 23 us).

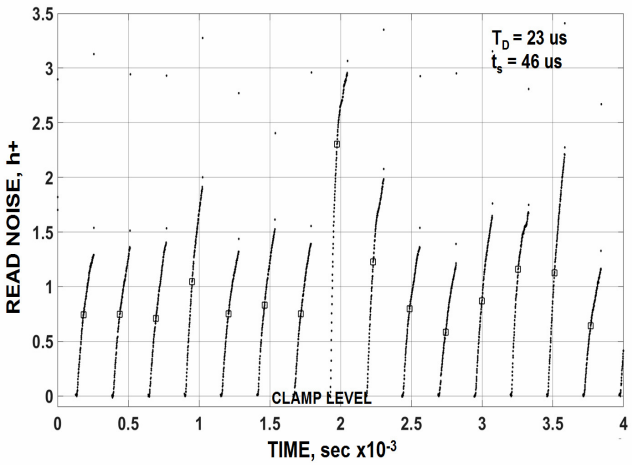

Fig. 20. Clamped video showing noise increase with time for τ_D =23us.

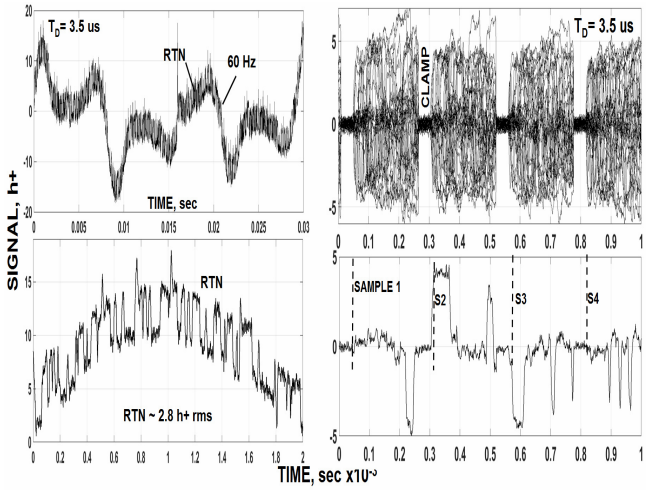

Fig. 21a/b. Raw 60 Hz and RTN and clamped video. response.

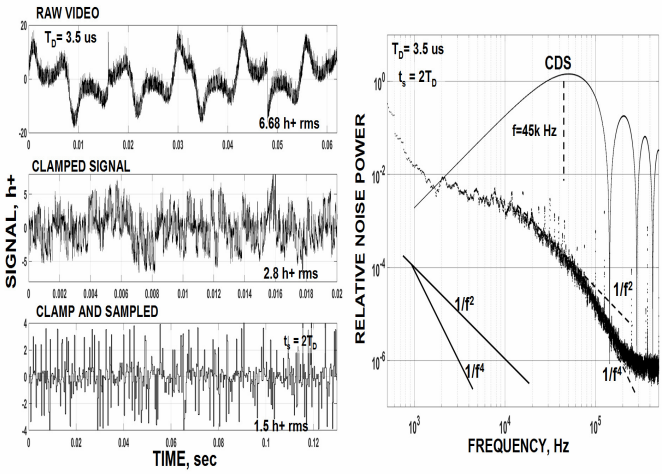

Fig. 22a/b. Raw, clamped, and sampled video and spectral

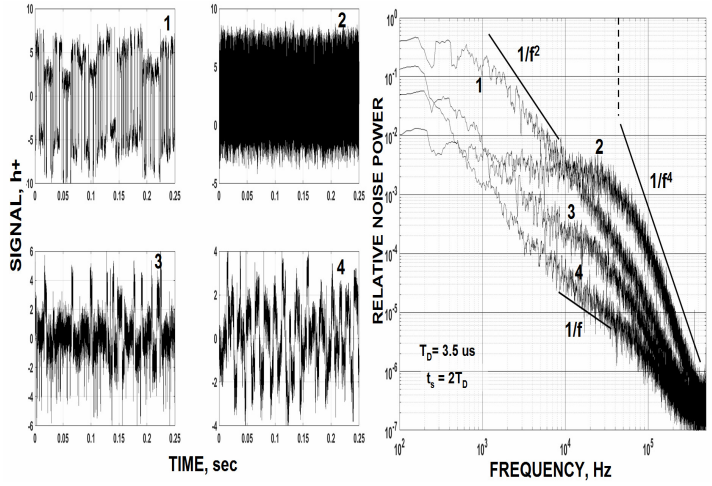

Fig. 23a/b. Various RTN time and frequency responses.

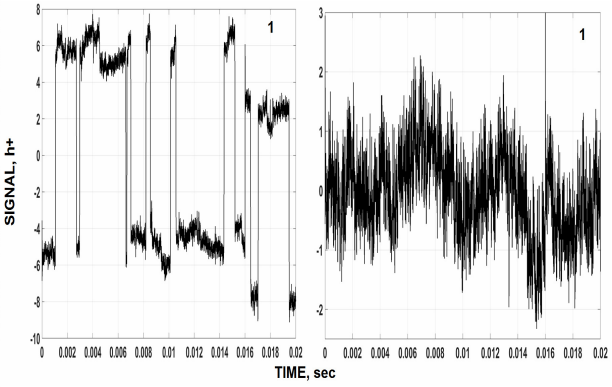

Fig. 24a/b. Magnified views of RTN #1 from Fig. 23a.

To help understand where 1/f noise sources are located, pixels have been irradiated with high energy radiation sources to deliberately increase interfacial surface states. In general when this experiment is performed we find that read noise does not significantly change. For example, Fig. 27a shows read noise that remains nearly constant after irradiating an unbiased NMOS imager to different dose levels of high energy electrons. The read noise for 10 Mrd exposure is higher because of thermally generated dark current. Cooling the pixels slightly as displayed in Fig. 27b brings the noise level down to the same floor level as other irradiated pixels. Figure 28 also shows a unchanged read noise when irradiated by 50 MeV protons at a dose of 1 Mrd. All the imagers irradiated did exhibit CTE degradation related to the TG. This test demonstrated that surface states were in fact being generated.[3] Some flat-band shift (~ 0.3V max) was also noticed indicating the surface was damaged. Other research groups experienced the same outcome.[5] The finding steers one to believe that that mobility noise is generating the 1/f noise measured above.

1.6 Flicker noise reduction

Applying the lessons learned above, this section opens by presenting the PT curve in Fig. 29. Notice that the read noise floor achieved is just shy of 'one' h+ rms. This level of noise is representative of the limiting noise performance for today's CCD and CMOS imagers. It is extraordinary to end up (after 43 years of development by this author) with an absolute level of 'one' carrier rms considering numerous entangled solid state phenomena at fabrication foundries all over the world are at work in producing it. Why 'one' ? Is this consequence simply coincidental? One also ponders further why this 'one' carrier of noise along with 'one' carrier of signal forces the minimum detection limit (MDL) of the detector to be 'one' (i.e., S/N='one'). And the conundrum that today's imagers are very close (but apparently yet so far) from regularly seeing 'one' single visible photo generated carrier because of the 'one' carrier of 1/f noise wall. For example, Fig. 30a is a simulated histogram of different read noise levels for an average signal level of 1 photon/pixel. Notice that the read noise would need to be ~ 0.4 e- rms before a histogram starts to appear 'quantized' and show the number of photon interacting counts within the pixels. Occasionally we do run into an individual low noise pixel that is capable (but barely) of doing this allowing us to see the Poisson distribution profile that governs photon counting statistics. The top histogram of Fig. 30b is generated by an exceptionally low noise cherry picked 3T pixel with a 0.35 e- noise floor taken under cooled conditions (~ - 80C). An average signal level of 0.8 e- is assumed (i.e., 0.8 interacting photons/pixel). The histogram below this plot is a computer simulation showing that data and theory match quite well. Figure 31a is a different 3T pixel that generates a PT curve with a 0.78 e- noise level when τ_D is set at 6.6 us. These low noise pixels were fabricated with a past 0.25um process which is known to have lower 1/f noise than the current 0.18 um process utilized (SF implants are annealed longer for the 0.25 um process). That said, the 'one' carrier noise puzzle may not be fundamental but simply related to it being 'good enough'. This is because shot noise dominates read noise as the signal falls below 'one' carrier. For extended images the S/N is always less than 'one' regardless what the read noise is. Hence, imager manufactures do not feel obligated to aggressively lower 1/f noise (especially monetarily). This stance gives us the illusion that there is a physical barrier of some kind. Although not mainstream, there are some applications where read noise less than 'one' is advantageous. Photon counting is an example as demonstrated in Fig. 30a/b. X-ray imaging spectrometers is another. Night vision applications would also welcome slightly lower read noise.

Is it impossible to lower 1/f noise significantly to the 0.1 carrier level ? Future CMOS development work will attempt to reduce SF 1/f noise without knowing exactly where it is coming from. Three approaches will be employed to help reduce SF MOSFET carrier interaction with the surface (be it for mobility or GR reasons). The first experiment involves further use of buried channel MOSFETs. We've had some past success with buried channel NMOS MOSFETs.[6] For example Fig. 32 shows a noise difference between surface and buried channel devices. We plan to fabricate deeper buried channels while providing the same SF threshold voltage adjust to maintain high dynamic range for the pixel. Special attention will be given to punch-through issues which can emerge with buried channel use. The second tactic in lowering 1/f noise will entail increasing the thermal anneal temperature to better activate the SF implants which could lower mobility noise which we suspect to be our dominant 1/f noise source. The third path will be finding the optimum SF geometry for the lowest noise performance. We typically find that a larger SF MOSFET relative to using minimum design rules achieves lower read noise (carrier rms) even though it's sensitivity (V/carrier) may be lower. For example, Fig. 33a/b plots read noise, h+/DN and sensitivity (V/h+) as a function of SF gate width for two different gate lengths. According to theory both 1/f and RTN (rms V) decrease with increasing gate length and width along with sensitivity (V/h+). This fact has been verified experimentally for test CMOS imagers.[7] Which parameter (1/f and V/h+) changes more rapidly with SF geometry determines if read noise (carrier rms) will increase or decrease. It is a tedious trial and error data taking process to determine which one wins out. This optimum geometry may not work for future pixel designs and fabrication processes. Experiments would need to be repeated to find the best size.

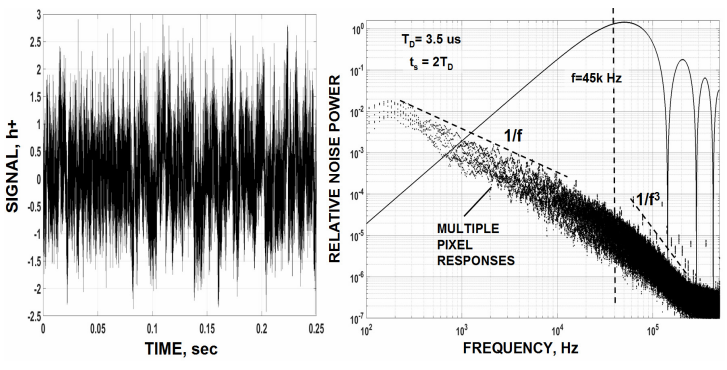

Fig. 25a/b. Pixel flicker noise and spectral noise for different pixels.

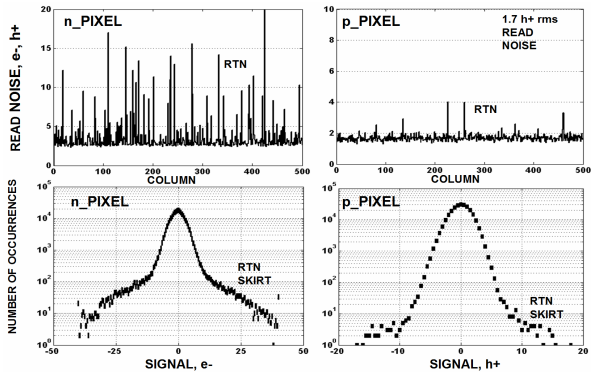

Fig. 26a/b. NMOS and PMOS RTN comparison.

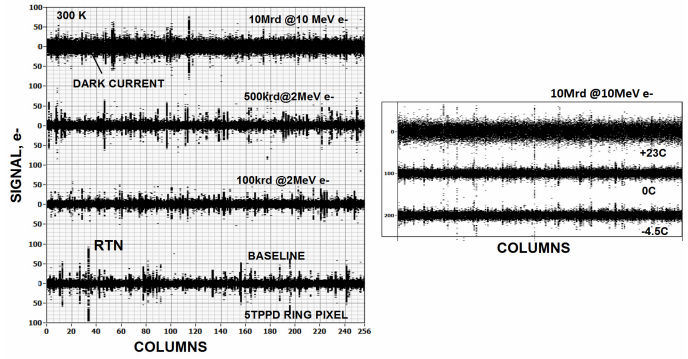

Fig. 27. Read noise at different irradiated dose levels.

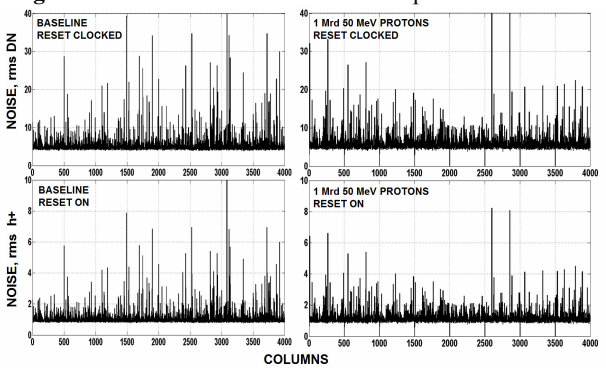

Fig. 28. Read noise before and after proton irradiation.

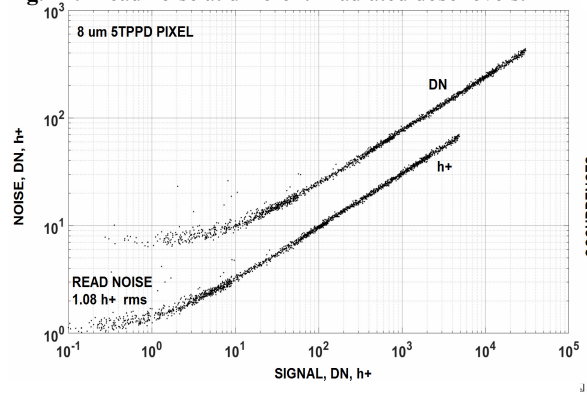

Fig. 29. PT curve demonstrating 1 h+ rms read noise.

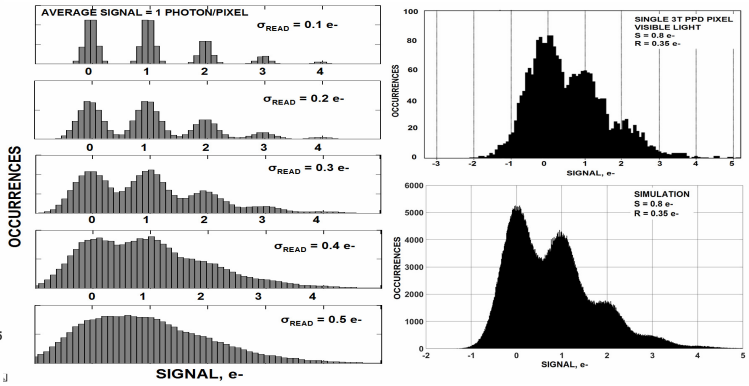

Fig. 30. Single photon counting quantized histograms.

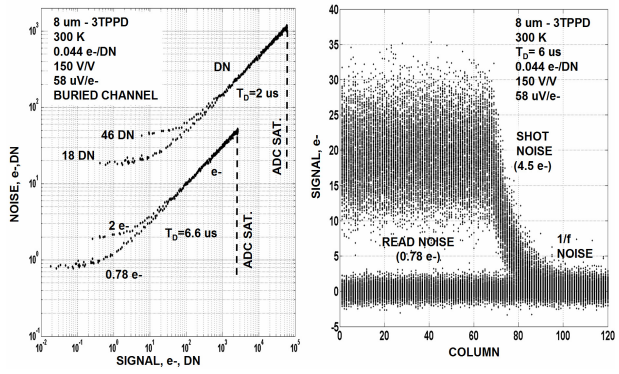

Fig. 31a/b. PTC Sub electron noise PT curve.

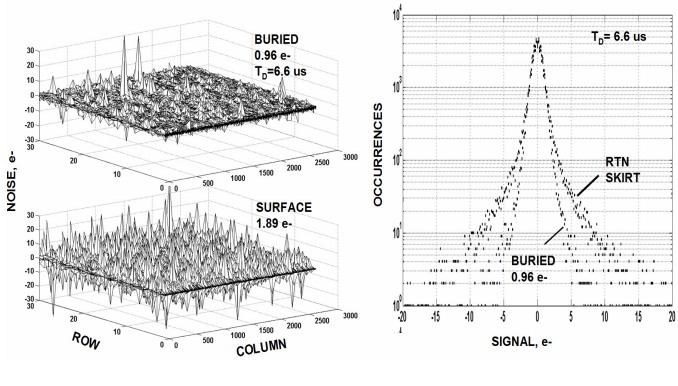

Fig. 32a/b. Read noise for surface and buried channel SF MOSFETs.

2. SoloHi /WISPR and Mk x Nk CMOS Imagers

2.1 SoloHi/WISPR imagers

The SoloHi imager has been previously discussed[3,7,8,9]. Recently (June 2015) 3840 x 4096 x 10 um 5TPPD pixel flight units have been delivered to the NRL/NASA satisfying their scheduled mission launch date in Oct. 2018 (refer to Fig. 44). Work is now proceeding to supply SoloHi imagers to the Wide-Field Imager Camera used on Solar Probe Plus (WISPR) mission scheduled to be launched shortly after SoloHi. Recent SoloHi improvements have involved thinning and subjecting the device to high radiation sources. This work was initiated by the Europa Clipper mission in that SoloHi's imager technology matches their need for upcoming CMOS imagers (refer to Sec. 2.2). The primary objective for this project was to determine if a backside illuminated (BSI) imager tolerance to radiation was different than frontside illuminated (FSI) experienced by NRL. Also a concern was brewing if the device could be read out at high rates (10 Mpixel/sec) for Europa given that SoloHi only reads the imager at 4M pixels/sec. This concern was shorted lived for it was demonstrated that the device could be read at 10 Mpixels/sec by generating decent PT curves. Also, an updated imager design now being fabricated would read 4x faster than SoloHi in anticipation of Europa's readout need (refer to Sec. 2.2).

Previous FSI SoloHi tests showed that dark current was the central vulnerable performance parameter that reacted to radiation. Hence BSI radiation tests for Europa focused on the same parameter. Figure 34a presents SoloHi pre-irradiated baseline dark current data when the TG is accumulated (AC) and non accumulated (NAC). The top histogram was generated after the imager thermally stabilized as on-chip power dissipation causes the temperature to increase. In contrast the lower plot was taken immediately after the sensor was powered for an ambient temperature measurement (~300K). Figure 34b are log-linear dark current histograms under AC conditions that produce an average dark level of ~ 200 h+/sec at 300K. Figure 35a presents a BSI flat field SoloHi image taken with narrow-band 940 nm light. Interference fringing is clearly seen at this wavelength because of thinning nonuniformities (each fringe represents a device thickness change of 1/7 um). Figure 35b is a corresponding image taken by the sensor using a broadband light source.

Figure 36a/b are side-by-side FSI and BSI responses in reaction to 100 krd of 63 MeV protons (a dose 2x higher than SoloHi is expecting at end of life). The semi-non-irradiated shielded portion of the detector serves as a baseline reference to measure performance degradation for the pixels and on-chip analog CDS processing circuitry. CDS circuitry is located at the top of the image. Figure 37a/b are FSI and BSI row traces showing that dark current generation rates are essentially identical for the imagers. Note that some protons do make their way into the semi-shielded region and generate individual dark spikes (also referred to as hot pixels).[1] This data demonstrated that FSI and BSI SoloHi devices are equally sensitive to proton irradiation in terms of dark current. Figure 38a presents a dark current histogram for the FSI imager showing irradiated and non-irradiated regions. The average dark current increase is ~ 8x higher after the proton exposure compared to baseline. Figure 38b is a dark current plot for the worst case dark spike pixel found in the histogram as a function of temperature (1e5 e-/sec at 300K). Such data was used to determine the operating temperature for the SoloHi imager. A temperature of -67 C was finally selected for flight operation. At this temperature the highest dark spike generates ~ 1 e-/sec of dark current which meets SoloHi's end-of-life noise specification.

A proton test energy of 177 MeV with a fluence of 5e10 protons/cm^2 (100krd) was also performed. This energy produces nuclear displacement damage to silicon lattice similar to what high energy neutrons would create. Neutrons have played havoc on past flight virtual-phase pinned CCDs by producing very large dark spikes.[1] Silicon cluster damage in conjunction with high internal electric field conditions creates these spikes. This proton test sought to find out if CMOS detectors were more sensitive than CCDs because both use pixel pinning technology where internal fields are typically higher. Figure 39a/b are row traces taken through the 177 MeV irradiated and semi-shielded regions under TG NAC and AC bias conditions. Individual large spikes are seen riding on the pre-irradiated dark floors. Figure 40a differences the NAC and AC responses leaving for the most part pre-irradiated NAC TG dark current. Proton induced dark spikes associated with the PPD region subtract out. Lastly, Fig. 38a is a histogram of the semi-shield region when TG is NAC. Comparing it to Fig. 34a shows the pre irradiated TG dark current (~2000 h+/sec) along with the proton induced dark spikes. It should be mentioned that proton interaction at 177 MeV is considerably less than 63 MeV. Hence 177 MeV protons create single spiked pixels whereas 63 MeV protons produce multiple spikes within each pixel (compare Figs. 39 and 37). However, its hard to tell if the amplitude of the 177 MeV spike is significantly greater than

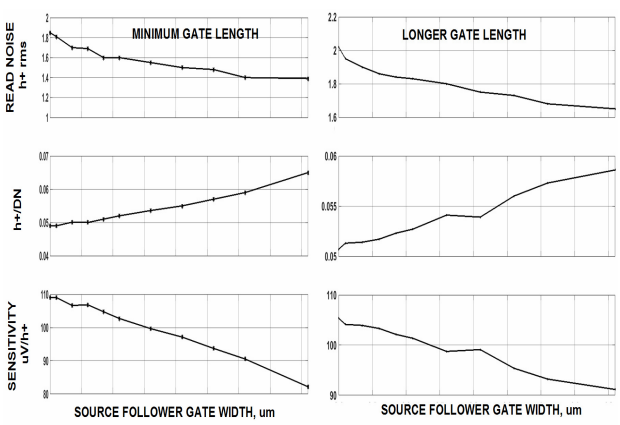

Fig. 33a/b. Read noise, h+/DN, uV/h+ as a function of SF gate width.

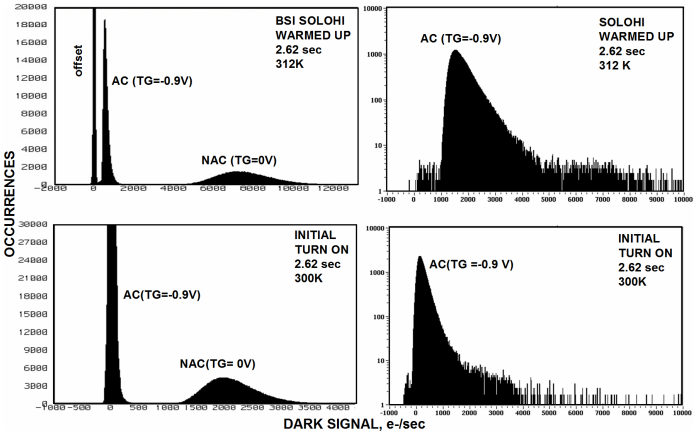

Fig. 34a/b. SoloHi baseline dark current histograms.

Fig. 35a/b. 940 nm response and image from a BSI SoloHi imager.

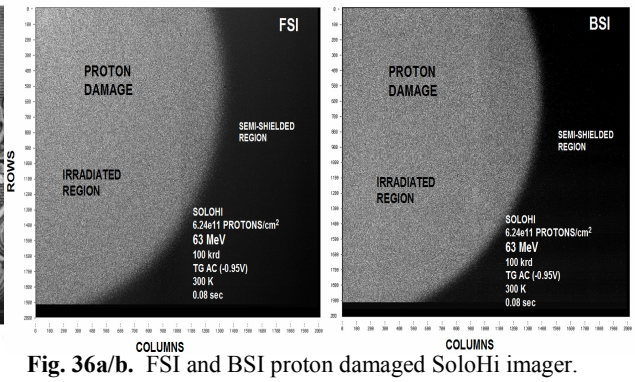

Fig. 36a/b. FSI and BSI proton damaged SoloHi imager.

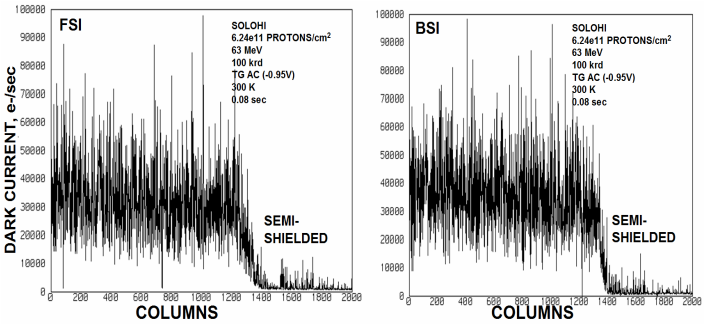

Fig. 37a/b. Dark video traces for FSI and BSI SoloHi imager.

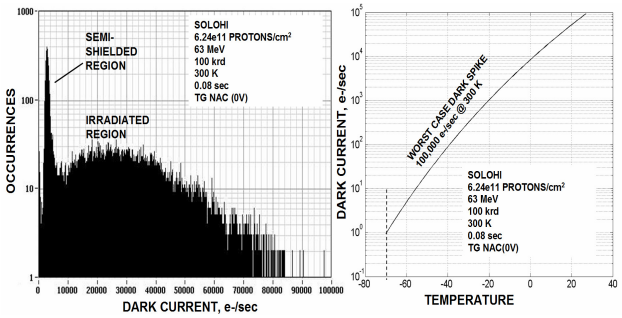

Fig. 38a/b. Dark histogram and generation rate with temperature.

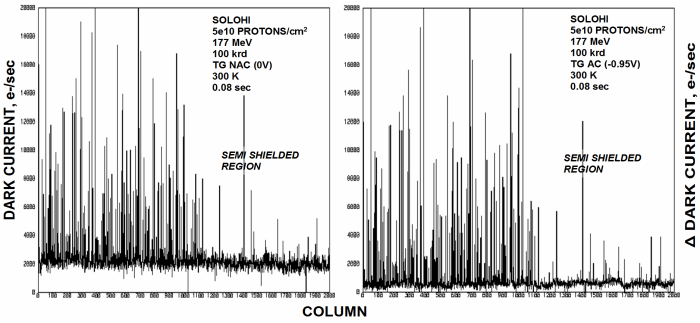

Fig. 39a/b. 177 MeV proton damaged responses.

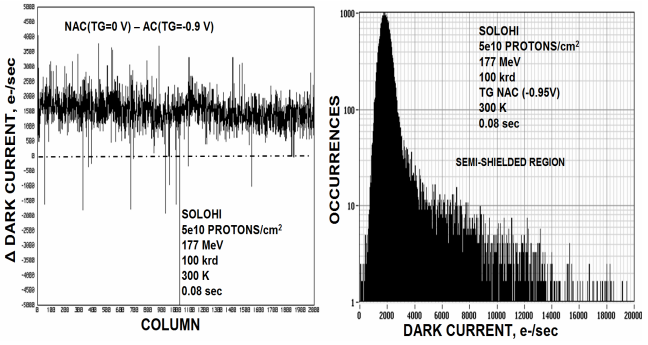

Fig. 40a/b. Difference NAC and AC response and dark histogram.

the 63 MeV spike. Spikes for both energies can be 100 x greater than the pre irradiated AC dark current floor and must be accounted for in selecting the operating temperature for the imager.

Changes for other performance parameters after proton irradiation were shown to be insignificant for the application. For example, read noise, flat-band shift and QE performance essentially remained the same for the imager. Also there was no noticeable shift in offset between the irradiated and non irradiated sides of the imager. Figure 41a shows raw video before CDS processing. The four different offset levels seen are related to the four parallel output sections of the imager (Sections 1 - 4) which inherently have different global offsets. There is no apparent offset shift at the irradiated and shielded boundary within Section 3 (indicated by the vertical dotted line). The CDS output response shown in Fig. 41b also shows no offset difference. The results are rather surprising in view of the fact that offset is the most unstable operating voltage for the imager (in general this is usually the case for most CCD or CMOS camera systems). For this reason SoloHi elected to have the offset voltage commandable in flight even though radiation test performed didn't appear to need it.

Some deferred charge (image lag) was detected for the imager as revealed in Fig. 42 for a low light level signal (1600 e-). The sensor is stimulated by a light source (LED) that is switched on and off as each pixel in a single row of pixels is CDS processed (referred to as the square wave CTE test).[2] Deferred charge (image lag) is seen when the LED turns off. The CTE problem stems from proton induced phosphorus-vacancy (P-V) traps[1] generated in the PPD region of the pixel. The measurement was performed at a worst case operating temperature of 300 K where the trap's emission time constant is near equal to the TG clock width. This timing and temperature combination produces the most measurable amount of deferred charge. At SoloHi's operating temperature of -67 C deferred charge is immeasurable as most P-V traps freeze-out at this low temperature. Figure 43 shows a similar response at an elevated signal (~20,000 e-). At this signal level CTE is limited by pre-irradiated TG surface state traps which dominate the P-V traps. It was shown previously that there is little difference in TG deferred charge for a proton fluence of only 100krd.[3]

2.2 Mk x Nk imager

The Mk x Nk imager layout and operation was discussed in considerable detail last year.[10] The imager was finally released for fabrication at Jazz Semiconductor (July 2015). The first lot run of Mk x Nk imagers will stitch 1k x 1k, 2k x 4k and 4k x 4k x 10 um 5TPPD pixel NMOS imagers. Lawrence Livermore National Lab (LLNL) will be the first user of the 4k x 4k imager at their National Ignition Facility (NIF).[10] Also NASA's Europa Clipper mission has recently chosen to use the imager for the Europa Imaging System (EIS) to conduct detailed reconnaissance of Jupiter's moon Europa in search of environmental conditions there that could potentially support life. This project will be testing the 2k x 4k device.

The Mk x Nk design can either support NMOS or PMOS processing depending on the needs of the customer. PMOS exhibits lower read noise than NMOS as explained in Sec.1.5. Also PMOS is more radiation resistant since proton induced P-V traps are not created in the boron doped PPD region.[3] However, PMOS carrier mobility is ~3x lower which can hinder device performance when working at very high speeds (specifically related to SF response and TG transfer time). NMOS technology at this time also carries a high TRL rating because of the SoloHi/WISPR missions which is important to new flight NASA projects such as Europa.

If desired multiple Mk x Nk imagers can be mosaiced together depending on how the wafers are cut. Although a monolithic Mk x Nk imager can be stitched much larger than a 4k x 4k, process yield issues could force the user to mosaicing instead (refer to substrate shorts discussed below). For example, four 4k x 4k's can be assembled together in producing a 8k x 8k. Mosaicing was the scheme that SoloHi employed to avoid a stitched design in producing their 3840(V) x 4096(H) x 10 um 5TPPD imager (refer to Fig. 44). Figure 45 shows where a 1k x 1k device is cut for either mosaicing or monolithic imagers and the associated bond pad sets used. Cutting for mosaicing leaves the top and right hand sides of the device without circuitry or bond pads. When devices are mosacied the gap between devices can be kept to < 100 um depending on how accurately the wafer is sawed. Figure 45 also labels the four patterns contained in a reticule that are stitched together to produce the whole imager. Pattern 1 is for the pixels. Pattern 2 (bottom side) is used for the CDS signal processor, X address decoders, column mux and output video buffers. Pattern 3 (left side) is for the pixel drivers and Y address decoders. And lastly Pattern 4 (corner) is used to readout the entire chip through a single output amplifier by only clocking the bottom left 1k x 1k block for any Mk x Nk imager fabricated. Figure 46a is an assembled 4k x 4k imager using the four patterns.

The substrate short has been shown to be the most important yield issue for CMOS and CCD imagers. The problem is usually related to a single pixel where a poly silicon gate (e.g., TG) and the substrate (i.e., epi layer) are shorted. Often near IR (1 um) luminescence accompanies this type of short. The short can also be responsible for producing dark rows and columns for CMOS imagers. The blemish is typically 'catastrophic' to a device's operation and reliability.[1] The chip is therefore usually discarded. Figure 46b is a dark response taken from a BSI SoloHi imager that contains five glowing 'catastrophic blemishes' (Cat Blems) located near the bottom of the chip. The quantity of light generated by each Cat Blem is different (Cat Blem 3 generates the most charge). Blem luminescence may decrease or increase in strength as a function of device use and vary as operating voltages to the imager are changed. When wafers are probed and screened one makes an effort to use a long exposure to better find the blems. Figure 47 illustrates how Cat Blem 1 luminescence erupts when the polarity of a pixel clock is switched during charge integration. The response located the substrate short to the pixel's reset MOSFET.

3.0 TRANSFER GATE RECOMBINATION LOSS

Accumulating the TG with carriers from the substrate is critically important to achieving lowest dark current generation. However, the user may encounter signal loss during charge transfer when TG is biased in this manner. The problem is technically referred to as recombination loss, a subject discussed in previous SPIE papers.[7] This section will furnish additional insights to the solid state physics behind the phenomenon and ways to reduce it (a PMOS pixel is assumed in the discussion). Figure 48 demonstrates the TG recombination problem by using the LED square-wave testing method while slowly switching the TG in and out of AC. When the TG is AC electrons from the substrate diffuse towards the TG and saturate it's surface state traps. For PMOS pixels the voltage necessary to have this come about is ~ 4.3 V assuming a substrate potential of 3.3 V. When the TG is clocked low to transfer charge (~ 2.0 V for this device) most of electrons return to the substrate. However some electrons remain trapped and recombine with the signal holes that are being transferred to the SN. The probability that a trapped electron meets up with a signal hole decreases as charge transfer time decreases. Ideally this transfer time is < 100 ns for 10 um pixel.[10] Unfortunately there usually exists a front edge barrier of the TG that slows the transfer time making the pixel vulnerable to the recombination effect. The barrier height encountered is a complex function of PPD and TG implants (dose, energy and alignment), specific details with the layout of the pixel and fabrication process used.

Figure 49 magnifies the top and bottom portions of the response in Fig. 48. Note from the upper plot that signal exponentially builds up to ~5700 h+ immediately after the LED turns on at the start of the TG NAC regime. During this time period is when signal holes recombine with trapped electrons placed there by the previous TG AC switching condition. At the same time signal holes begin to be trapped in empty TG surface states where electrons had just occupied. Given sufficient time the majority of trapped electrons will depart the TG region and leave only trapped signal holes which will later be observed as deferred charge when the LED turns off. This is seen in the lower plot of Fig. 49. It is noticed that within the TG AC region deferred change is not seen because electrons immediately recombine with the trapped holes. The detrapping discharge time for the holes and electrons is fairly long because there are long lived traps. Data in Fig. 49 indicates that ~100 us will allow most trapped electrons to find their way back to the substrate and the trapped holes back to the PPD as image lag. Figure 50a measures recombination loss in percent as a function of signal level. Unfortunately the loss increases with decreasing signal. For applications where every signal carrier counts to S/N performance any kind of loss can't be tolerated (e.g., night vision imagery). The loss experienced also increases with TG surface area simply because there are more traps involved with the recombination process. For example, Fig. 50b shows the loss that occurs when the TG length is varied in steps from 0.5 um to 2.5 um.

The observations made above opens up a remedy to reduce recombination loss by not transferring charge until the electrons detrap themselves and safely return to the substrate. One means to accomplishing this is to use 'tri-level' clocking. That is, we can reduce the TG voltage slightly from the AC to NAC (4.3 to 3.3 V) without charge transfer taking place. Electrons are then allowed to detrap. After a sufficient discharge time passes (~100 us) the TG is then clocked low to transfer charge. A different method that applies to progressive readout is by having the row of pixels that is immediately behind the row being readout assume the NAC condition before it is read out. All other rows will remain in the AC state for low dark current generation. Assuming row time is ~100 us or greater the majority of the charge is transferred without significant recombination.

A buried channel TG is a direct solution to prevent recombination and CTE loss by keeping signal charge away from the

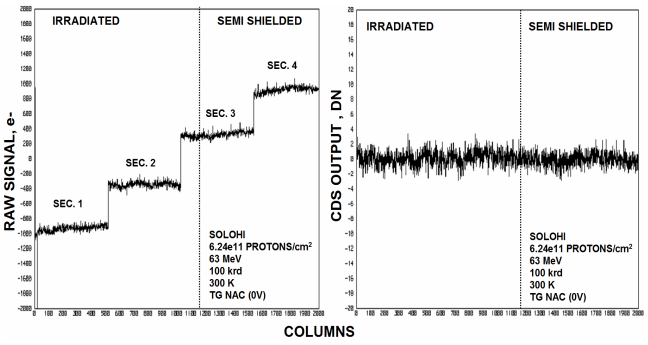

Fig. 41a,b. Raw and CDS lack of offset shift after irradiation.

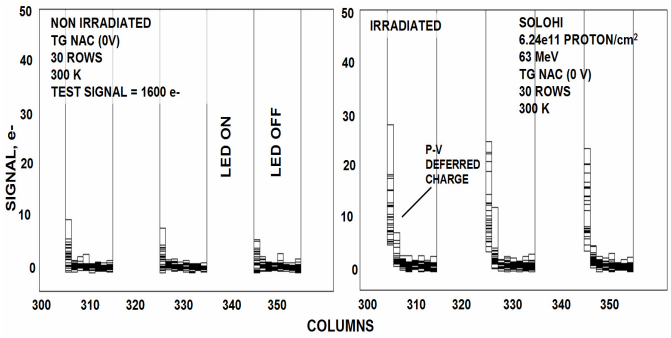

Fig. 42a,b. Deferred charge increase due to proton induced P-V traps.

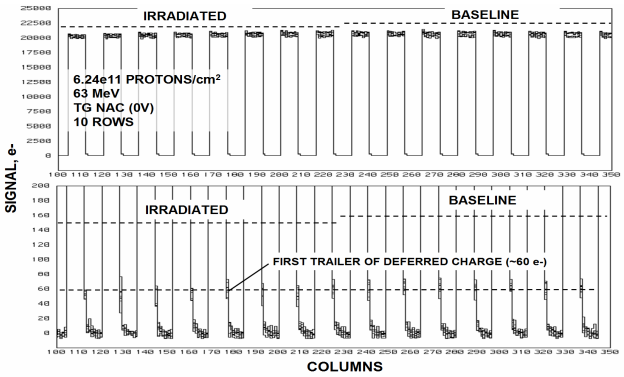

Fig. 43a,b. Deferred charge due to TG traps.

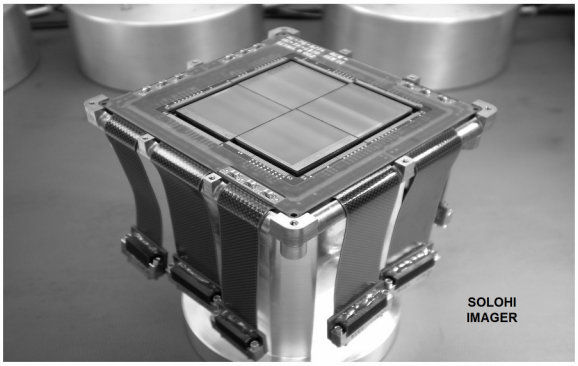

Fig. 44. SoloHi imager.

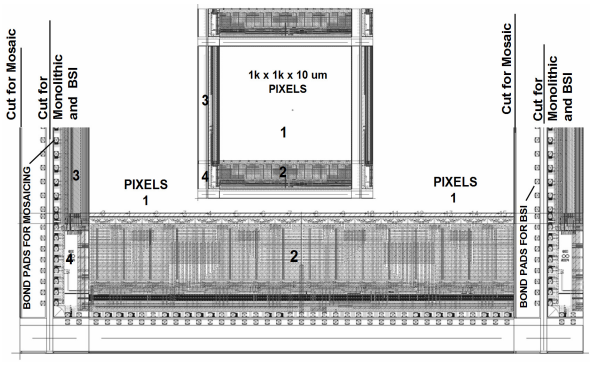

Fig. 45. Mk x Nk dicing for monolithic and mosaicing.

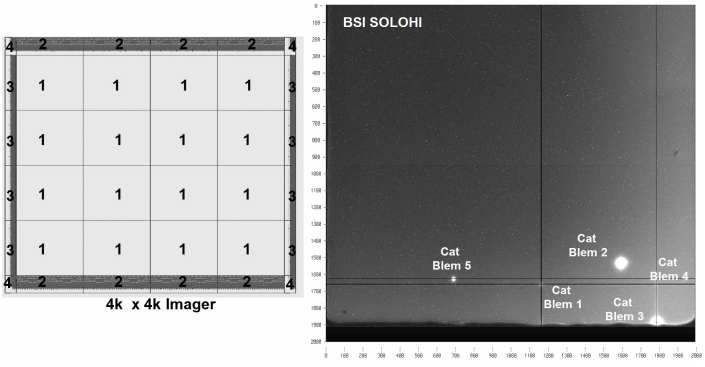

Fig. 46a,b. 4k x 4k imager and SoloHi Cat Blems.

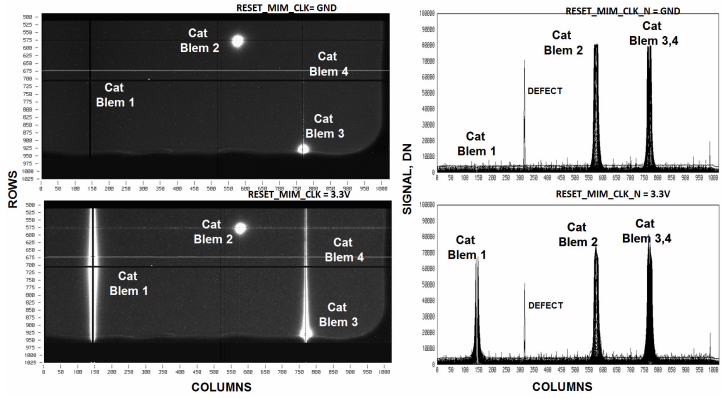

Fig. 47a,b. An erupting Cat Blem caused by change in pixel voltage.

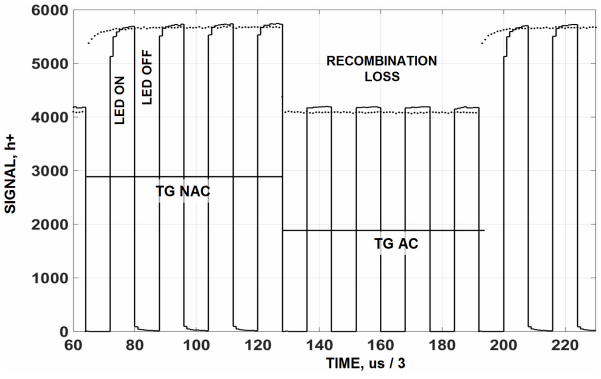

Fig. 48a,b. TG AC recombination loss.

surface. Some work along these lines has been performed but further development is required. It has been shown that keeping a surface channel design and processing can produce a pixel with excellent CTE and without recombination loss. For example, Fig. 51a presents two PT curves that are stacked upon one another under AC and NAC bias conditions. The responses are identical. Figure 52a shows two stacked square-wave responses as a function of signal with different light settings. The plots perfectly line up to one another indicating no recombination is taking place. Figure 52b greatly magnifies the vertical axis of Fig.52a. Deferred charge is immeasurable.

ACKNOWLEDGMENTS

The authors are appreciative of Richard Mann for the fabrication of our imagers at Jazz/Tower Semiconductor. We express our gratitude to Mike Lesser of the University of Arizona for his collaboration on thinning our sensors. We are grateful to Joe Kimbrough and Perry Bell at LLNL for their continued support in developing the Mk x Nk imager. We are indebted to Elizabeth Turtle, Kim Strohbehn, and John Bolt of John Hopkins University Applied Physics Laboratory and Alfred McEwen of the University of Arizona Lunar and Planetary Laboratory for their support on thinning and testing the BSI SoloHi imager and the lead up to Mk x Nk imager for Europa Clipper use. Thanks goes to Clarence Korendyke of Navel Research Laboratory in supplying SoloHi devices for thinning.

REFERENCES

1). J. Janesick, *Scientific Charge-Coupled Devices*, SPIE Press, PM 83, Bellingham, WA, (2001).

2). J. Janesick, *Photon Transfer* $\lambda \rightarrow DN$, SPIE Press, PM 170, Bellingham, WA, (2007).

3). James Janesick, Tom Elliott, James Andrews, John Tower and Jeff Pinter, Fundamental performance differences of CMOS and CCD imagers: Part V, SPIE Photonics West, San Francisco, Ca., Feb. 2013, paper #8659-1.

4). Lode K. J. Vandamme and F. N. Hooge, What Do We Certainly Know About 1*/f* Noise in MOSTs?, IEEE Transactions on Electron Devices, Vol. 55, No. 11, Nov 2008.

5). l/f Noise in MOS Devices, Mobility or Number Fluctuations?, L. K. J. Vandamme, Xiaosong Li, and Dominique Rigaud, IEEE Transactions on Electron Devices, Vol. 41, No. 11, Nov. 1994.

6). James Janesick, Jeff Pinter, Robert Potter, Tom Elliott, James Andrews, John Tower, Mark Grygon, Dave Keller, Fundamental performance differences of CMOS and CCD imagers: Part IV, San Diego June 2010, paper # 7742-11.

7). J. Janesick, J. Pinter, R. Potter, J.Andrews, J. Tower, T. Elliott, J. Cheng and J. Bishop, Fundamental performance differences between CMOS and CCD imagers; Part III, "Focal Plane Arrays for Space Telescopes IV," August 2009 in San Diego, paper # 7439A-6.

8). Russell Howard, et.al, The solar orbiter imager (SoloHI) instrument for the solar orbiter mission, Proc. of SPIE Vol. 8862 88620H-2.

9). Clarence M Korendyke, et.al, Development and test of an active pixel sensor detector for heliospheric imager on solar orbiter and solar probe plus, SPIE Vol. 8862 88620J-1.

10). James Janesick, Tom Elliott, James Andrews, John Tower, Perry Bell, Alan Teruya, Joe Kimbrough and Jeanne Bishop, Mk x Nk gated CMOS imager, SPIE Optics and Photonics, San Diego, August 18, 2014, paper # 9211-5 .

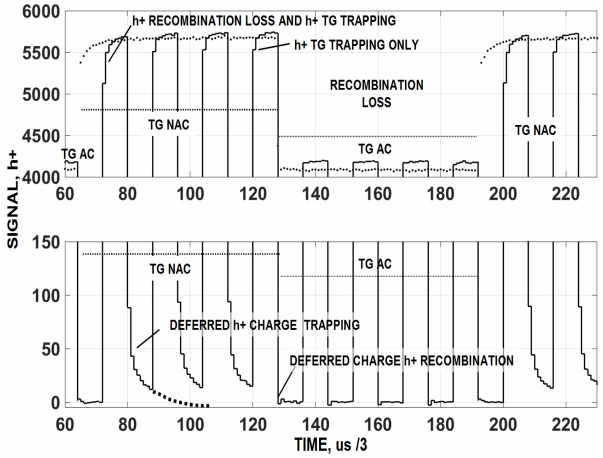

Fig. 49. TG recombination loss and deferred charge.

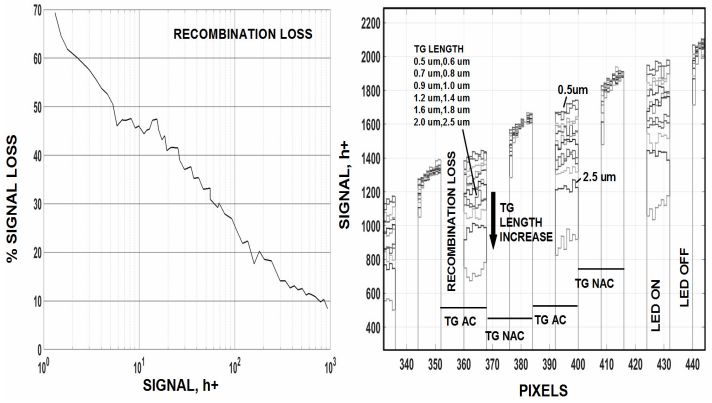

Fig. 50a,b. TG recombination loss versus signal and TG gate width.

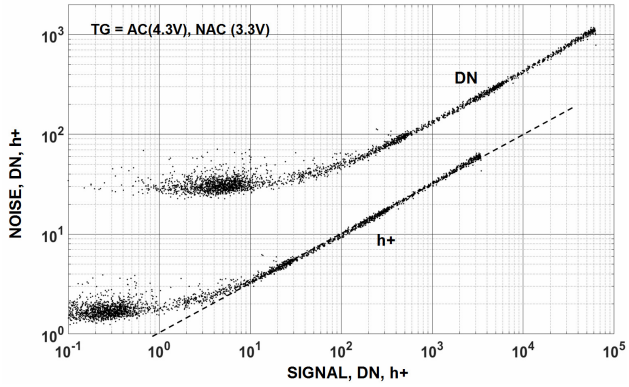

Fig. 51a,b. PT curve for an imager without TG loss. charge.

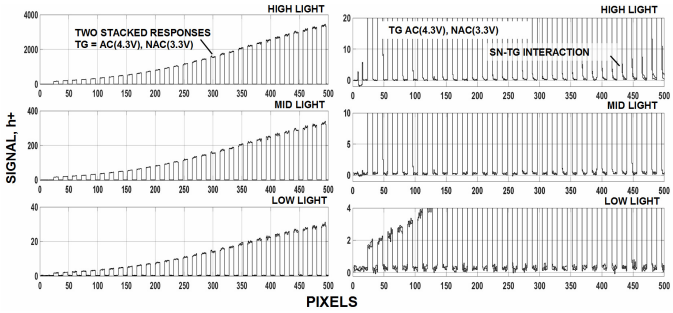

Fig 52a,b. Square-wave test without recombination loss or deferred