Radiation Induced Transient Effects in HgCdTe IR Focal Plane Arrays

J.C. Pickel[a], R.A. Reed[*b], P.W. Marshall[c], A. Wacynski[d], E. Polidan[d], S. Johnson[d], R. McMurray[e], M. McKelvey[e], K. Ennico[e], R. Johnson[e], and G. Gee[f]
[a] PR&T, Inc., Fallbrook, CA 92028
[b] NASA Goddard Space Flight Center, Code 561, Greenbelt, MD, USA 20771
[c] NASA Consultant, Brookneal, VA 24528
[d] Global Science and Technology, NASA Goddard Space Flight Center, Greenbelt, MD 20771
[e] NASA Ames Research Center, Moffett Field, CA 94035-1000
[f] SGT, Inc, NASA/GSFC, Code561, Greenbelt, MD 20771

ABSTRACT

The operability requirements of NASA's James Webb Space Telescope (JWST) impose specific challenges on radiation effects mitigation and analysis. For example, the NIRSpec Instrument has the following requirements:

- The percentage of pixels defined as operable for target acquisition shall not be less than 97% (TBR) (goal 99%) of the total number of pixels… An inoperable pixel is:
  - A dead pixel: a pixel with no radiometric response
  - A noisy pixel: a pixel with a total noise greater than 21 e-, per Fowler 8 exposure
- The percentage of pixels defined as operable for science observations shall not be less than 92% (TBR) (goal 98%) of the total number of pixels… An inoperable pixel is:
  - A dead/low-DQE pixel: a pixel deviating by > 30% from the DQE mean value
  - A noisy pixel: a pixel with a total noise greater than 12 e- (goal 9e-)

With these performance requirements and operation in space, the radiation environment from galactic cosmic rays (GCR), energetic solar particles, and activation of spacecraft materials can contribute significantly to the number of inoperable pixels. The two most important issues to date are radiation-induced transient effects and hot pixels. This paper focuses on the methods used to assess the impact of ionizing radiation induced transients on the HgCdTe SCA selected by JWST. Hot pixel effects in these detectors has been previously presented [1]. Both effects are currently under investigation.

Keywords: Infrared focal plane arrays, radiation effects, IR, imager

1. INTRODUCTION

The James Webb Space Telescope (JWST) will use mega-pixel class infrared (IR) detector arrays in the near infrared (NIR) instruments, including the Near-Infrared Camera (NIRCAM), Near-Infrared Spectrometer (NIRSpec) and Fine Guidance Sensor (FGS). The focal plane array will consist of multiple sensor chip assemblies (SCA). The JWST SCAs must be able to survive the ionizing particle radiation environment (background cosmic rays and solar particle events) for its Lagrange-point (L2) orbit and ten year mission lifetime goal,, These test and analysis results for the tested SCAs were published in [1]. In addition, the SCAs must have adequate performance in the presence of the continuous ionizing particle environment from cosmic ray background levels to allow the science mission to be accomplished, and that is the focus of this paper.

To help determine their suitability for JWST and understand the performance limitations imposed by the space radiation environment, a comprehensive radiation test and analysis program is underway to assess candidate SCA technologies.

[*] Currently at Vanderbilt University, robert.a.reed@vanderbilt.edu XXX-XXX-XXXX

The NASA Goddard Radiation Effects Analysis Group (REAG) has overall responsibility for assessing radiation effects on JWST and providing guidance to the project for radiation effects management and mitigation. The radiation test and analysis described in this paper was performed in collaboration with the NASA Ames Research Center (ARC) Detector Test Laboratory. These testing, analysis and modeling activities will provide risk reduction to the project and will facilitate an optimized system design for the final technologies that are selected.

The tested devices were produced through NASA NIR focal plane development contract at Rockwell Scientific (5-μm cutoff HgCdTe) which ran from 1998 - 2003. Rockwell Scientific's H1RG is a mega-pixel class building block designed to be representative of the larger structures proposed for JWST. The devices are hybrid structures consisting of a silicon readout circuit mated to a detector array with indium bump bonds. The devices are in a 1024x1024 format and are optimized for JWST NIR requirements. The H1RG (and the H2RG, its 2048 X 2048 version) draws on Rockwell Scientific's 'HAWAII' technology and is comprised of a software configurable silicon readout circuit mated to a mercury-cadmium-telluride (HgCdTe) detector array at a pixel pitch of 18 μm and optimized for the JWST NIR spectral range. The designed operating temperature is near 37 K. The devices had the detector substrate removed after hybridization to allow detection across the wavelength range 0.6 to 5 μm.

Rockwell Scientific worked to demonstrate that JWST performance requirements could be met, and that the more ambitious performance goals could be approached. Among the performance challenges were achieving extremely low total noise (<10 electrons) over extended (~1000 second) integration times, near-vanishing dark current levels, and high-performance anti-reflection coatings to cover the 0.6 to 5 μm range. They developed both 1024 x 1024 and 2048 x 2048 element detector arrays and the associated two-dimensional multiplexers. Because in each case the architecture, processing steps, and detector geometries of the 1024^2 and 2048^2 devices were identical, the smaller-format devices were selected for testing convenience. Radiation testing included characterization of proton-induced transient effects and proton fluence degradation effects ─ total-ionizing-dose (TID) and displacement damage (DD). Measurements were made on the full 1024^2 hybrid arrays in five separate tests at the University of California Davis (UCD) cyclotron in Davis, CA.

This paper is organized as follows. First we provide a brief discussion of mechanisms and the expected radiation effects. We then describe the overall modeling approach and show how the test program integrates with the modeling and overall assessment. We then describe specific goals and rationale for the testing approaches, followed by detailed discussion of the test approach and test results for the transient testing. The total fluence (ionizing dose and displacement damage) testing was reported in [1]. We conclude with a summary.

2. TRANSIENT EFFECTS MECHANISMS

The key concerns for transient effects are illustrated in figure 1. The primary particles, mostly energetic protons, pass through the SCA and cause transients that are large with respect to the inherent noise level. When the primary protons pass through the surrounding material, there is a cascade effect with secondary electrons (deltas) generated, and occasionally, a nuclear reaction that releases additional particles and/or activates an atom that subsequently decays and releases particles. In addition, there is the possibility of particle emission from natural radioactivity from radio-nuclides that may be embedded as contaminants in the surrounding material. Large transients can be managed with established software algorithms for the data stream, however small transients that are near the inherent noise level are problematic. The delta electrons are of particular concern if they are not spatially correlated with the primary particle, since multiple pixels could be corrupted from a single particle hit.

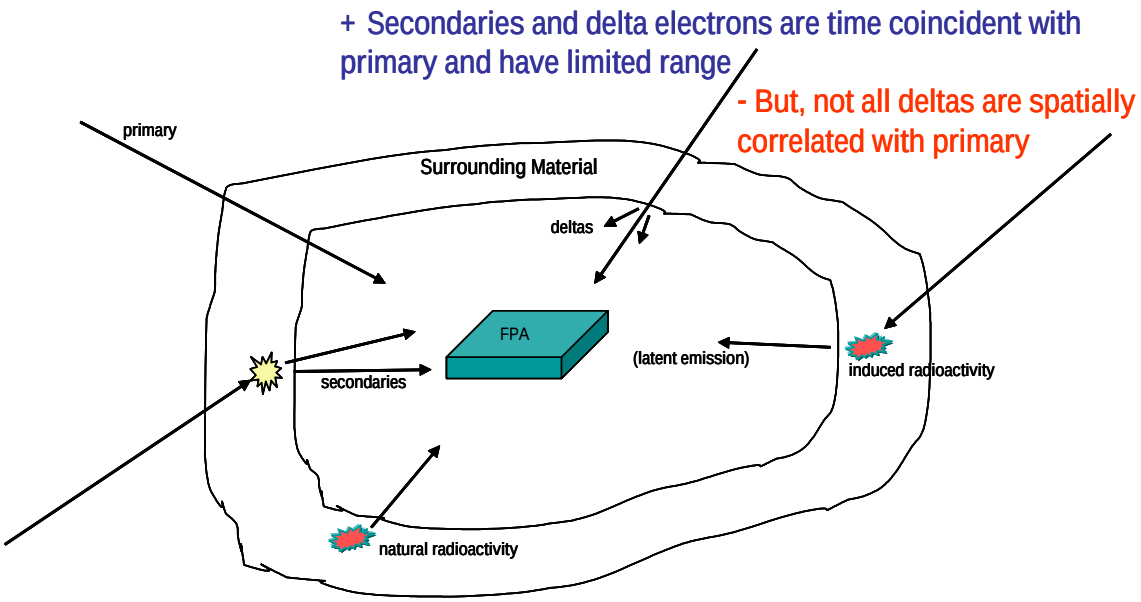

Figure 1: Ionizing particle impacts to FPA.

Because both the detector array and the ROIC contribute to charge collection, it is possible for more than one pixel to be corrupted by a single particle track. The SCA is a hybridized structure consisting of a detector array connected to a readout integrated circuit, as illustrated in figure 2.

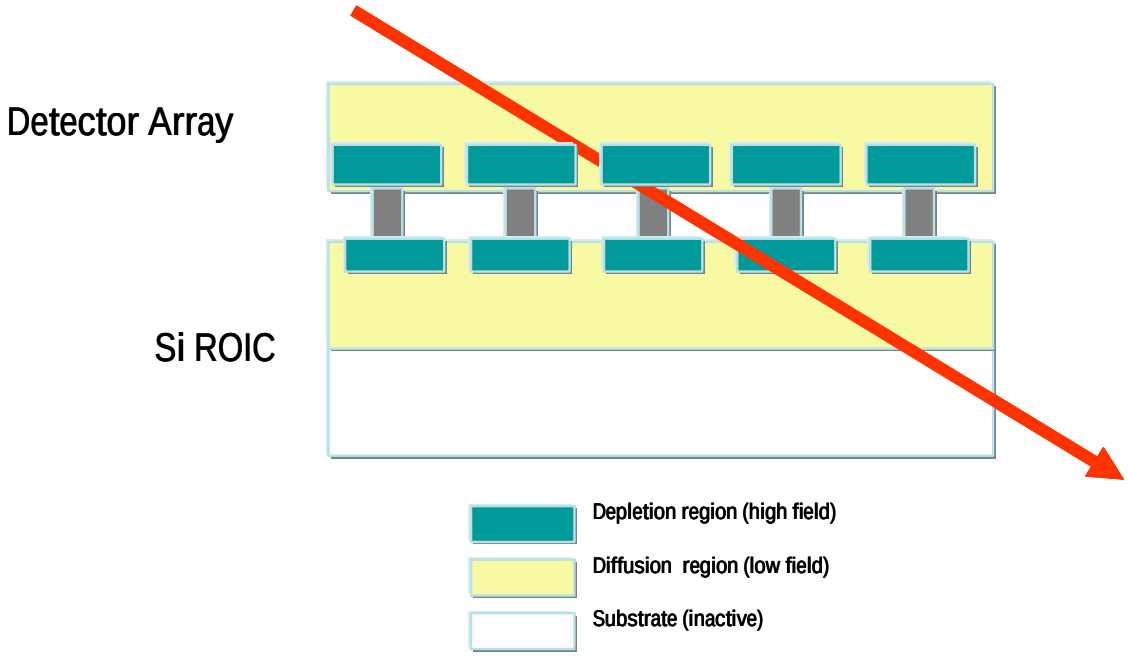

Figure 2: Hybrid FPA structure showing charge collection from both detector array and ROIC.

IR detectors efficiently detect ionizing particles – the detectors don't distinguish between ionization from an IR photon or from an energetic particle. Charge generated in depletion regions of p-n junctions is collected by drift in the high field region. Charge generated in quasi-neutral substrate regions can diffuse to junctions and be collected. Charge that diffuses to neighboring pixels results in crosstalk. ROICs can also detect particles (see figure 3). The charge diffusion

length in the Si substrate is comparable to the pixel pitch and all charge that diffuses to the sensitive junction during the integration time will be counted. Note that the particle is not required to hit the sensitive junction in the ROIC because of the charge diffusion and the long integration time.

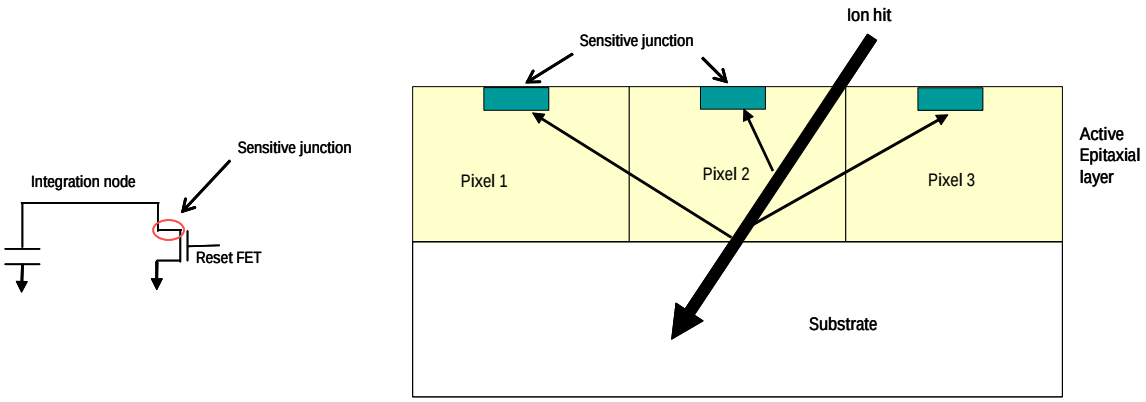

Figure 3: ROIC charge collection mechanism.

Because of the very low inherent noise level in these devices (~ 10 electrons), essentially every primary particle and every secondary particle that reaches the sensitive charge collection volume of the detector or readout integrated circuit (ROIC) will cause a transient that exceeds the noise level. Cosmic ray rejection algorithms can tolerate a limited number of hits within the integration time. However the problem is exacerbated by:

- Radiation Crosstalk (charge spreading to neighboring pixels)
- Multi-pixel hits (e.g., hit detector and ROIC in different pixels)
- Secondary particles that are not spatially correlated to primary

An important concern for the radiation-induced transients is the time that the disturbance remains above the noise floor. Cosmic ray rejection algorithms depend on multiple samples of the same pixel during the signal integration period to identify and remove corrupted pixel data. A particle hit can usually be identified because of the large charge in the hit pixel compared to the background noise level. Ideally a reset will restore the corrupted pixel. However, if the radiation-induced transient only slightly exceeds the 10 electron noise floor after reset, the corrupted pixel may not be detected by the rejection software and this will add to the noise burden of the measurement.

There are two possibilities for a transient that exceeds the noise floor for longer than a frame time. One is that the charge collection time for the radiation-induced charge is longer than the frame time. This effect is simply due to the time required for the radiation-induced charge along the particle track to be collected. Another is the phenomenon of persistence, where the diode dark current for the hit pixel is anomalously high for a period of time after the hit. Persistence is a well known effect for optical exposures and is thought to be related to temporary charge storage in surface passivation interfaces that affects the diode dark current. Persistence is likely to be present for particle-induced transients, especially at the low levels of concern.

3. TRANSIENT EFFECTS MODELING

The goal of the transient modeling is to predict the SCA response to incident energetic particles (protons, heavy ions and electrons). We determine the final charge contamination for each pixel in the array after a specified set of particle hits have occurred. The set of particle hits is determined by the external environment flux, transport of the external environment through the spacecraft and surrounding material, and the integration time. The model addresses radiation

crosstalk (charge spread from a hit pixel to neighboring pixels) and multi-pixel hits from a single particle due to a) delta electron spatial spread and b) non-pixel-correlated detector and ROIC hits.

The overall modeling approach is shown in figure 4. The external space environment is transported through the spacecraft structure and the material immediately surrounding the SCA using transport codes such as GEANT4, MCNPX and NOVICE. Full 3-D CAD models of the material structure are used. The local particle environment is then defined at the SCA. This environment consists of both primary and secondary particles, including protons, electrons, neutrons and photons. The interaction of the local particle environment with the SCA is analyzed with a charge transport code developed for analysis of charge collection in large arrays (REACT, see below). Other codes are used to study activation of surrounding material and predict the contribution to the local particle environment. The ultimate output of the modeling is a FITS-format file that shows the charge in the pixels of the array due to particle hits, overlaid on inherent random noise. This modeling can be exercised with variations to provide guidance on material selections and system operational choices that affect transient rates. In addition the model output can be used to test cosmic ray rejection algorithms.

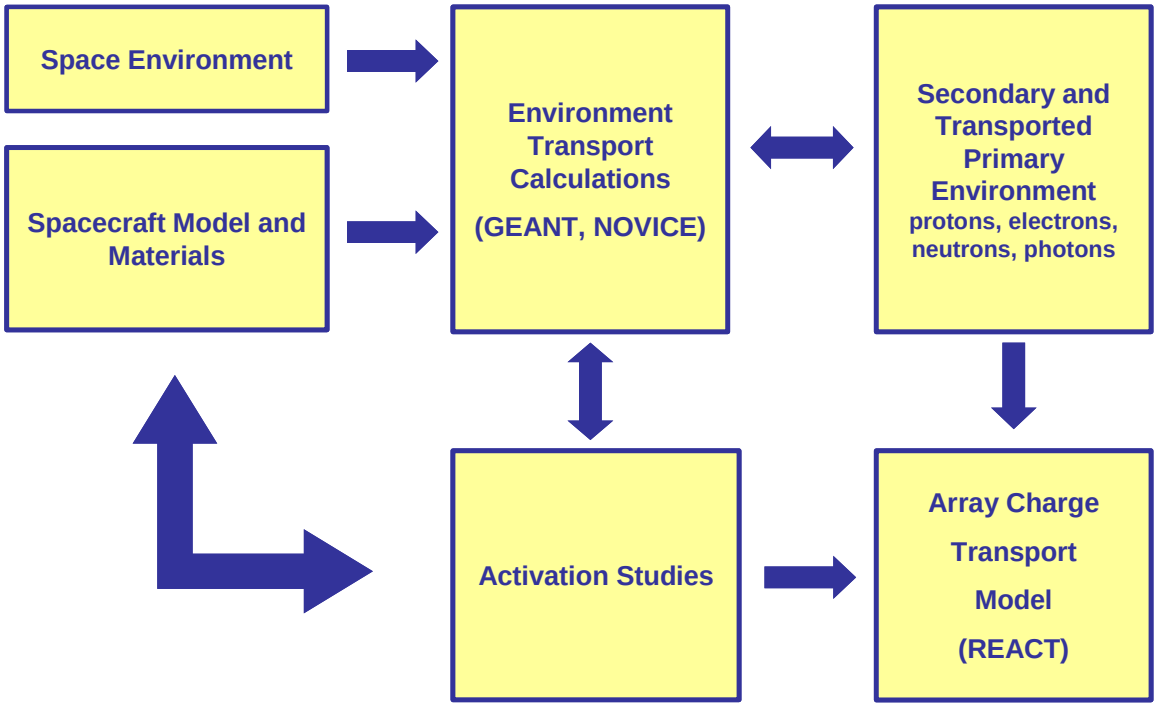

Figure 4: Overall modeling approach.

A key element of the modeling approach has been development of the array charge collection model, where the code is designated REACT (Radiation Environment Array Charge Transport) [2]. Figure 5 illustrates the REACT model. An initial line source of minority carrier distribution is based on particle linear energy transfer (LET) and trajectory and defined through the 3-D structure of the SCA. Appropriate drift and diffusion charge collection physics models are applied depending on particle location within the SCA. Charge carrier history ends when the carrier is either collected on a pixel or recombined. Thereby, charge is distributed to pixels across the array in accordance with drift and diffusion. The code output is pulse height distributions, crosstalk characterization, FITS files, etc.

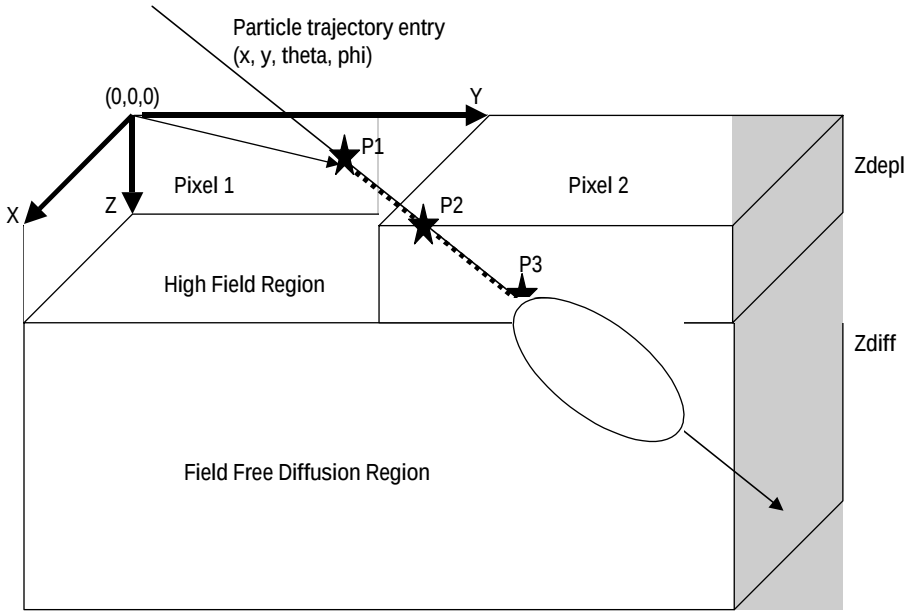

Figure 5: Radiation environment array charge collection model (REACT).

4. TEST PHILOSOPHY AND SCOPE

The radiation testing has three primary objectives:

- To validate and calibrate the transient effect models,
- To characterize the performance as a function of total proton fluence (reported in [1]), and
- To search for unexpected radiation effects.

Ideally we would like to test with the actual radiation environment that will be seen in the mission; however that is not possible with ground-based testing as accelerator beams are typically monoenergetic and monodirectional.

For transient effects, the environment of concern is the background cosmic rays. These are GeV-energy range particles and cannot be easily duplicated in ground-based accelerators. Thus we simulate parts of the environment and extrapolate the results to the mission case based on modeling. The primary job of the models for the transient effects analysis is a) to predict the secondary environment (particle type, energy and spatial distribution) at the sensitive volume of the SCA, and b) to predict the distribution of charge to pixels across the array after a particle hit. The transport codes GEANT4 and NOVICE are used for predicting the secondary particle environment. The REACT code is used for predicting charge distribution to the pixels across the array. A key reason for doing the extensive testing across a wide range of test variables is to calibrate the single event transient effects models.

The scope of the radiation testing to date has been limited to proton testing at 30 and 63 MeV on bare ROICs and SCAs (detector array mated to ROIC). Five test periods at the University of California Davis (UCD) cyclotron have been performed over the period May, 2002 to May, 2003. Transient effects testing was done at both energies, while all damage testing was limited to 63 MeV. UCD facility provides proton flux that can be varied to low levels (~1e3 p/cm^2) as required for single event testing. Beam dosimetry was performed using the well-established beam dosimetry system at UCD. Pre-irradiation and post-irradiation testing was performed at the low-background test facility at the NASA Ames Research Center Laboratory in Moffett Field, CA.

The primary objectives for the transient testing were:

- Characterize proton single events as a function of proton energy and angle of incidence
- Measure the charge spread (radiation crosstalk) to adjacent pixels

- Assess the transient recovery time
- Characterize the secondary particle emission from a material sample in the vicinity of the detector and from the interior of the dewar [3]

The strategy for the transient testing was to capture single event transients in the array as a function of proton energy and angle of incidence. In addition to acquiring detailed information about single event transients, we also had the goal of determining the recovery time for a pixel that had been hit by a proton, or assuring that the recovery time was sufficiently fast so as to not be a factor in JWST performance. Data acquisition speed and data storage make high-speed repeated reads of the entire array impractical. To accomplish the goal, we elected to measure transients in a small subarray (20x100) of the 1024x1024 array. The approach was to repeatedly read the subarray (patch) and save into the data acquisition system. The data patches were stored in two 2-dimensional data buffers containing an array of patches built up to represent a frame. Using the ARC Lab View software, the single event patches were organized into data frames, converted to FITS file format, organized into chronological order and saved as data cubes that were typically 20x100x800. This allowed the data to be analyzed with Interactive Data Language (IDL) software.

Data were taken at 30 MeV and at 63 MeV. At each energy, data were taken for angles of incidence of 0, 45 and 67 degrees. A low flux, typically in the range of 1e3 to 1e5 p/cm^2-s was used to assure that we were studying isolated single events. To capture the transient recovery, we elected to read at the maximum sampling rate of the data acquisition system (10 Hz), giving 100 ms between read time. We then varied the time between resets. That is, we used Fowler mode integrations with variable integration times (reset, read$_1$, read$_2$,…read$_i$, reset). Table 1 shows the various times between resets that were used for the testing.

Table 1: Test Modes for Transient Testing.

Mode	#Reads	Time Between Reset (ms)
F1	2	200
F2	4	400
F5	10	1000
F10	20	2000

5. TEST RESULTS AND COMPARISONS TO FIRST-ORDER MODELING

The data for the transient tests fall into five general areas:

- Single event morphology
- Pulse height distributions
- Charge spreading
- Transient recovery time
- Secondary particle energy deposits

The next three sections detail the results for the first three areas, the last area (secondary particles) is discussed in detail in [3]. The fourth area (transient recovery time) will be discussed in another paper to be published at a later date.

5.1 Single Event Morphology and Data Reduction

Figure 5 shows the average charge in individual pixels after a proton hit for the H1RG SCA. The data show the average charge (electrons) in a 5 x 5 array of pixels centered around a pixel that has been hit by a 30 MeV proton incident at 0 degrees. The average is formed by stacking 2000 single event hits (see Section 5.3). The inherent noise level in the measurements can be seen in the non-hit pixels. For the pixels that are hit, we see charge spreading to the neighboring pixels.

-16	-4	12	-5	-13
-7	339	1477	305	-9
7	1415	15985	1330	8
-7	270	1493	240	-7
-13	-1	32	-3	-12

Figure 5: Average proton-induced pulses (number of electrons) in 5x5 subarray of H1RG SCA from 30 MeV protons.

The proton generates charge by ionization, and to first order the charge generated in the active volume of the array is simply pathlength * LET / ionization energy, where LET is linear energy transfer (e/μm). Pathlength depends on the thickness of the active region and the angle of incidence. LET depends on the proton energy. We assume an ionization energy of 1.5 eV / carrier pair for the MWIR HgCdTe material.

However, the charge is not all collected to the pixel that is hit by the ion. There is charge spread to adjacent pixels by diffusion and the degree of charge spread is dependent on the hit location within the pixel (e.g., more spread for hits near edge than for center hits) and on the angle of incidence. We collected independent data sets at 3 angles (0, 45 and 67 degrees) and at 2 energies (30 and 63 MeV) so there are a variety of pulse distributions in the datasets. Angle of incidence, proton energy and the assumed thickness of the active volume determine the total charge for the expected pulse. However, since the charge is shared among multiple pixels in most cases, we should actually expect a distribution of charge deposits that should have a peak somewhat less than the maximum total charge.

The first step in reducing the data to pulse height distributions is to remove hot pixels and fixed-pattern noise from the data set and establish a baseline level. Hot pixels are identified by their constant presence in the clear data. Fixed pattern noise - non changing spatial variation in output levels across the array - is more difficult to deal with, however we have devised algorithms that smooth out the noise and establish a baseline. The baseline output charge represents the inherent noise level associated with the read process in the clear environment at the cyclotron. For the initial data processing (results reported below), events were identified as those pixels with charge that is at least 5 sigma above the noise floor. Note that this noise level represents the variation of the pixel charge for a single frame and does not represent the best noise levels that are achievable by Fowler sampling and averaging of multiple frames. Single frame data is required to capture single events that are completed within a single frame.

The distributions are formed after the dataset has been "cleaned" by removing hot pixels and fixed pattern noise (if needed). The charge in each pixel in the array is determined as the difference between the first read after reset and the final read. The data are then binned into 100 electron bins. There are three basic types of distributions of interest.

Peak Charge - the charge in each hit pixel is noted and the data set is histogrammed. In this case the charge that leaks to neighboring pixels is not counted. It is necessary to identify which pixels have been hit. Our initial analysis used a 5 sigma selection criterion. Those pixels that are at least 5 sigma larger than the mean noise are selected.

Total Charge - the charge in each event is summed and the data set is histogrammed. In this case we want to determine the total charge associated with each single event, including the hit pixel and its neighbors. A single event is when a single proton passes through one or more pixels of the array. The primary proton will induce charge and there may be secondary electrons (deltas) that are generated in the material near the sensitive region that accompany the primary proton and also induce charge. If the deltas hit other pixels, then these become indistinguishable from primary proton events except that they may deposit a different charge compared to primary protons. Developing the total charge histograms is a two step process:

1. Identify the primary hits. This can be done by looking across the array for pixels that have a charge above a threshold. Assuming that the fluence is sufficiently low that all hits are single events, the largest charges should be LET * pathlength as discussed above. However, if deltas from adjacent material along the proton path also hit the pixel, the charge could be greater. This is analogous to a dose enhancement effect.

2. Determine which neighboring pixels need to be summed as part of the event. For each primary hit location, look at the neighboring pixels and add to the sum if the charge is sufficiently enough (e.g., 5 sigma) above the noise.

Cluster Size - the number of pixels that are affected by each hit is noted. First isolated single event hits must be identified. Then selection criteria must be used (e.g., 5 sigma above the mean) to tag neighboring pixels as being affected by the hit. A histogram is then formed for the number of occurrences for each cluster size.

5.2 Pulse Height Distributions

Figure 6 shows the peak charge distribution for 0 and 67 degrees with 30 MeV (left) and 63 MeV (right) protons. The distributions scale with angle, as expected. The larger angles give longer pathlength and thus more charge generation. We also note that the pulses are smaller than for the 30 MeV case, as expected. The LET at 63 MeV is smaller than the LET at 30 MeV by a factor of ~0.58 for HgCdTe. Figure 7 shows total charge and peak charge distributions for the H1RG SCA for 30 (left) and 63 (right) MeV protons at 0 degrees, and figure 8 shows similar data for 30 (left) and 63 (right) MeV protons at 67 degrees.

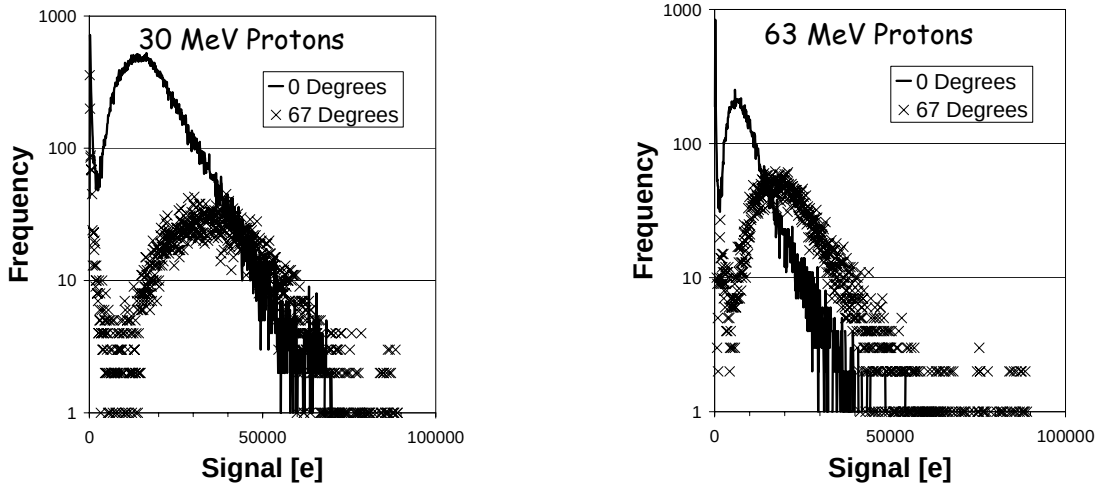

Figure 6: Peak charge pulse height distributions for SCA from 30 and 63 MeV protons at various angles.

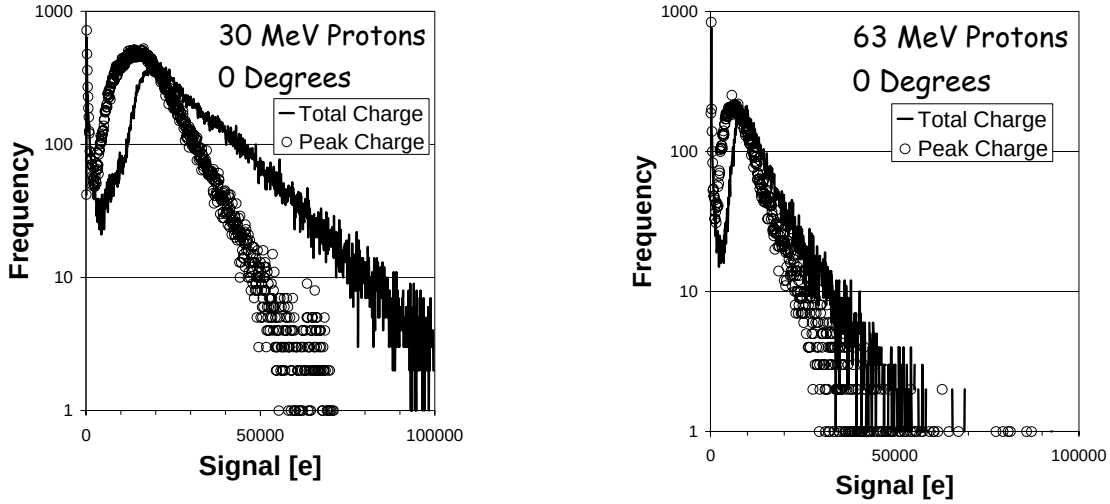

Figure 7: Peak charge and total charge pulse height distributions for SCA from 30 and 63 MeV protons at 0 degrees.

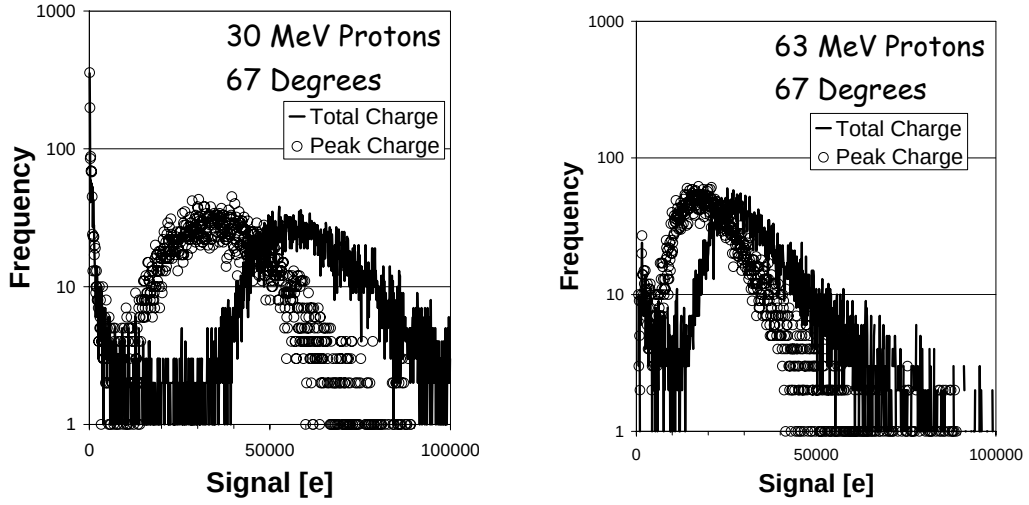

Figure 8: Peak charge and total charge pulse height distributions for SCA from 30 and 63 MeV protons at 67 degrees.

Table 2 summarizes the pulse size scaling (based on maxima of total charge distributions) with angle of incidence and with proton energy.. The location of the distribution maxima are summarized in the peak location table for the various cases. The expected scaling (normalized to 0 degrees and 30 MeV) is based on pathlength and LET and is shown in the left column for the angle scaling and energy scaling tables. We note that scaling with angle and energy generally follow expectations.

Table 2. Scaling of total charge distribution peaks with angle and energy.

Peak Locations			**Angle Scaling**		**Energy Scaling**	
Energy (MeV)	Angle (Deg)	Peak (e)	Predicted	Actual	Predicted	Actual
30	0	21,000	1.00	1.00	1.00	1.00
30	67	58,000	2.56	2.76	1.00	1.00
63	0	9,000	1.00	1.00	0.58	0.43
63	67	27,000	2.56	3.00	0.58	0.47

5.3 Charge Spreading

One measure of charge spreading from the hit pixel to adjacent pixels is the distribution of cluster sizes (number of pixels affected). Using a selection criterion of 5 sigma above the mean noise to identify corrupted pixels, we formed histograms of the number of affected pixels. Figure 9 shows the cluster size distribution data for the H1RG SCA for 30 MeV protons at 0 and 67 degrees; it also shows the same for 63 MeV protons at 0 degrees. Note that the general trend of that data in energy and angle are predictable from variation in proton LET and in path-length, respectively.

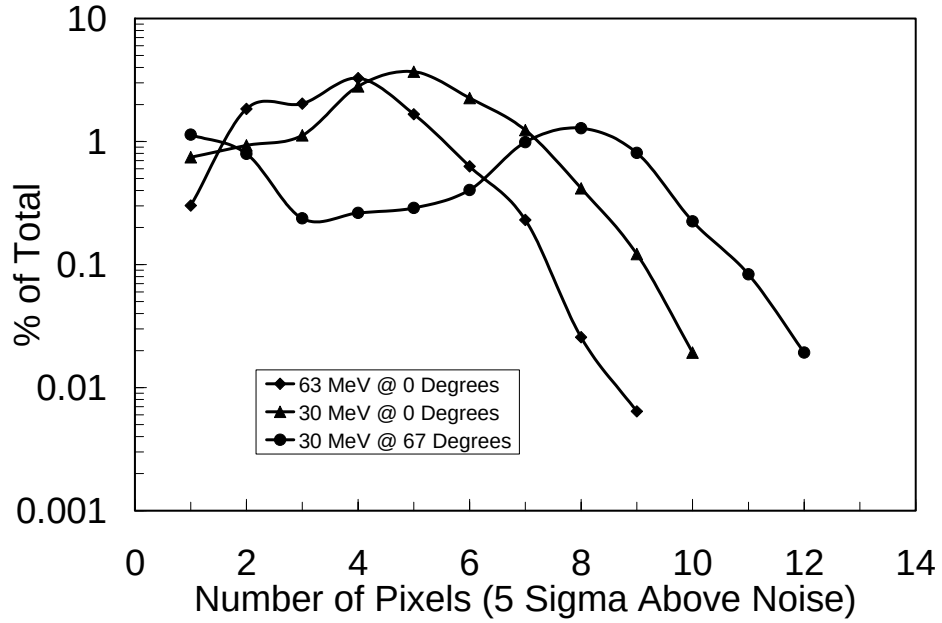

Figure 9: Number of pixels affected by a single proton event.

Another way to characterize charge spread is to perform a stacking analysis where multiple hits are co-added and then divided by the number of hits to give an average charge in adjacent pixels. For the stacking analysis, the hits are registered with respect to the hit pixel. This means that random hit locations across the pixel are averaged into the data, without regard to hit location within the pixel. This method of analysis differs from stacking analysis that has been performed by other groups where only hits to the central region of the pixel are picked for the averaging. However, hits will be randomly located across the pixel in the actual space environment case so this stacking method is more representative of the real problem.

Figure 5 is a stack average of nearly 2000 events in the H1RG SCA for 30 MeV protons at normal incidence. The results are shown as charge in a 5x5 pixel array centered around the hit pixel and are expressed as average number electrons in the hit pixel.

The REACT simulation tool (see Section 3) can be used to predict proton-induced charge spreading in an SCA. Figure 10 shows stack average results for the simulation code described in Section 3. These results are very similar to the experimental results.

Both the stack plots and the cluster size distribution plots are analyzed for noise floor in the 100's of electrons. Note that the JWST requirement is about an order of magnitude below these noise levels. We would expect that if the analysis could be done at something closer to the instrument noise level, the charge sharing would be greater and could possibly affect pixels further away than the nearest neighbor of the struck pixel.

0	1	2	1	0
1	109	928	99	0
2	987	16591	912	2
1	127	1098	109	1
0	1	2	0	0

Figure 10: REACT simulations for average proton-induced pulses (number of electrons) for normal incident 30 MeV protons.

6. SUMMARY

On-orbit assessment of transient effects in IR SCAs required detailed environment predictions, ground based testing, and a validated prediction method. Proton-induced SET testing on H1RG SCA show that transients induced by protons scale as expected with energy and angle of incidence of the proton. Multiple pixels are affected by a single particle event and there is an increased number of inoperable pixels. A modeling technique has been developed to provide instrument teams with reliable predictions of the FPA response to ionizing radiation. We also note that the reference [1] demonstrates that there is a significant increase in hot pixel count due to exposure to the space radiation environment, and this effect is a key issue that must be addressed.

REFERENCES

[1] M.E.McKelvey, K.A.Ennico, R.R.Johnson, P.W.Marshall, R.E.McMurray, Jr., C.R.McCreight, J.C.Pickel and R.A.Reed, "Radiation Environment Performance of Prototype JWST FPAs," SPIE Proceedings of Vol. #5167

[2] J.C.Pickel, R.A.Reed, R.Ladbury, B.Rauscher, P.W.Marshall, T.M.Jordan, B.Fodness and G.Gee, "Radiation-Induced Charge Collection in Infrared Detector Arrays," IEEE Trans. Nucl. Sci., Vol. 49, No. 6, December 2002.

[3] J.C.Pickel, R.A.Reed, P.W.Marshall, T.M.Jordan, G.Gee, B.Fodness, M.McKelvey, R.E.McMurray, K.A.Ennico, R.R.Johnson and C.McCreight, "Proton-Induced Secondary Particle Environments for Infrared Infrared Sensor Applications," IEEE Trans. Nucl. Sci., Vol. 50, No. 6, December 2003.