Characterization of the column-based priority logic readout of Topmetal-II\textsuperscript{−} CMOS pixel direct charge sensor

M. An,\textsuperscript{a} W. Zhang,\textsuperscript{a,1} L. Xiao,\textsuperscript{a} C. Gao,\textsuperscript{a} C. Chen,\textsuperscript{a} M. Han,\textsuperscript{b} G. Huang,\textsuperscript{a} R. Ji,\textsuperscript{a} X. Li,\textsuperscript{c} J. Liu,\textsuperscript{a} Y. Mei,\textsuperscript{b} H. Pei,\textsuperscript{a} Q. Sun,\textsuperscript{d} X. Sun,\textsuperscript{a} K. Wang,\textsuperscript{a} P. Yang\textsuperscript{d} and W. Zhou\textsuperscript{a}

\textsuperscript{a}PLAC, Key Laboratory of Quark and Lepton Physics (MOE), Central China Normal University, Wuhan, Hubei, 430079 China
\textsuperscript{b}Nuclear Science Division, Lawrence Berkeley National Laboratory, Berkeley, California, 94720 U.S.A.
\textsuperscript{c}Institute of High Energy Physics Chinese Academy of Sciences, Beijing, 100049 China
\textsuperscript{d}Department of Physics, Southern Methodist University, Dallas, TX, 75205 U.S.A.

E-mail: stone156324@mails.ccnu.edu.cn

ABSTRACT: We present the detailed study of the digital readout of Topmetal-II\textsuperscript{−} CMOS pixel direct charge sensor. Topmetal-II\textsuperscript{−} is an integrated sensor with an array of 72 \times 72 pixels each capable of directly collecting external charge through exposed metal electrodes in the topmost metal layer. In addition to the time-shared multiplexing readout of the analog output from Charge Sensitive Amplifiers in each pixel, hits are also generated through comparators in each pixel with individually adjustable thresholds. The hits are read out via a column-based priority logic structure, retaining both hit location and time information. The in-array column-based priority logic features with a full clock-less circuitry hence there is no continuously running clock distributed in the pixel and matrix logic. These characteristics enable its use as the charge readout device in future Time Projection Chambers without gaseous gain mechanism, which has unique advantages in low background and low rate-density experiments. We studied the detailed working behavior and performance of this readout, and demonstrated its functional validity and potential in imaging applications.

KEYWORDS: Electronic detector readout concepts (solid-state); Front-end electronics for detector readout; Pixelated detectors and associated VLSI electronics; VLSI circuits

\textsuperscript{1}Corresponding author.
1 Introduction

Highly pixelated sensors such as CMOS image sensors and Monolithic Active Pixels Sensors (MAPS), have been successfully deployed in various fields. In many nuclear/particle physics applications, the traditional readout is orders of magnitude slower than what is required for signal/data acquisition in order to achieve the physics goals. Therefore, many novel readout schemes, designed to provide faster access to the information collected in the pixels, have been developed over the years. These schemes often exploit the characteristics of signal distribution among pixels, such as the sparseness or clustering in space and time of pixel hits, and the similarity in amplitudes. Notable examples include column-based readout [1–3] and row-based compression [4, 5].

We implemented a column-based priority logic readout in a prototype pixel sensor called Topmetal-II− [6], aimed at improving the latency between a pixel hit and the availability of data off the chip. Topmetal-II− is implemented in a 0.35 µm CMOS process. It features a 72 × 72 pixel array with 83.2 µm pixel pitch. Pixels in Topmetal-II− are sensitive to external charges arriving at an exposed metal electrode in the topmost layer in each pixel. Pixel hits are generated by pixel-local comparators with tunable thresholds. We chose a scheme that is clock-less in the pixel array to minimize the potential interference from digital switching noise. Digital activity in the array only happens when the sensor receives hits. The in-array column-based priority logic (without clock) drives the address of the pixel that is hit to the edge of the array immediately upon a hit, which minimizes the latency. A sequential logic (with clock) is employed to sense the hit location and time at the edge of the array and then ship such information off the sensor. A detailed study of the analog characteristics of Topmetal-II− is reported in [6]. This paper focuses on the details of operation and functional validity of the digital readout.
2 Sensor structure and operation

A Topmetal-II sensor, as shown in figure 1, contains an array of $72 \times 72$ sensitive pixels occupying a $6 \times 6$ mm$^2$ area. Each pixel has an exposed metal patch in the topmost layer that can directly collect charge. The charge signal is amplified by a Charge Sensitive Amplifier (CSA) in each pixel (figure 2). The sensitivity of Topmetal-II to light is due to the parasitic p-n junction of n-type MOSFET Md and Mf (figure 2), between source/drain (n+ region, above GND) and p-well (body, connected to GND). Such p-n junctions are reverse biased and light illumination would induce leakage. Light-induced leakage currents in Md and Mf are injected into the input of the CSA and amplified. The amplified charge signal is accessible through two channels. The analog voltage signal is read out through a “rolling shutter” style time-shared multiplexer controlled by the array scan unit. A digital hit signal is generated by an in-pixel comparator with per-pixel adjustable threshold, then read out through the column-based priority logic.

Figure 1. Photograph of a Topmetal-II sensor (left) and its top-level block diagram (right). The chip is $8 \times 9$ mm$^2$ (blue box) in size, in which a $6 \times 6$ mm$^2$ charge sensitive area (red box) is located in the center of the sensor. Major functional units are shown in the top-level block diagram. The sensor is powered by analog supply AVDD and digital supply DVDD, which are individually regulated at nominal voltage of 3.3 V.

2.1 Column-based priority logic readout

The overall readout has two parts: an in-array clock-less logic and a sequential logic at the bottom edge of the array. The clock-less logic consists of a Priority Logic (PL) in each pixel and an Address Bus (AB) in each column. The sequential logic includes a Column Readout Unit (CRU) placed at the bottom edge of each column and a multiplexer (MUX) collecting the outputs of all CRUs. CRUs monitor the address changes on the ABs. CRUs and the MUX are synchronous to shared clock CLK and reset RST signals. The maximum allowable frequency of the shared clock is 80 MHz.

2.1.1 Pixel hits and Priority Logic (PL)

A schematic view of the circuit in a single pixel is shown in the red dashed box in figure 2. The exposed Topmetal electrode is directly connected to the input of the CSA. A ring electrode (Gring), which is in the same topmost metal layer as the Topmetal, surrounds the Topmetal while being
isolated from it. The stray capacitance between the Gring and the Topmetal, \( C_{\text{inj}} \approx 5.5 \, \text{fF} \), is a natural test capacitor that allows pulses applied on Gring to inject charge into the CSA. The CSA with \( C_f \approx 5 \, \text{fF} \) converts the injected charge to voltage signal (CSA\_OUT) and feeds it into the comparator. The threshold voltage applied to the \( i \)th pixel comparator \( \text{V}_{\text{th}} \) is the sum of a globally set voltage \( \text{V}_{\text{thg}} \) with a locally adjustable offset \( \text{V}_{\text{thp}} \). The step size of all the 4-bit DACs is globally adjustable as well. The pixel-local 4-bit DAC is intended for compensating the threshold dispersion of the comparator across the entire array. The CSA and the comparator are constantly active.

**Figure 2.** Digital readout pathway from a single pixel to the edge of array. Structures at pixel, column and edge-of-array levels are indicated in the red, green and blue dashed boxes, respectively.

Upon an event of CSA\_OUT surpassing the threshold \( \text{V}_{\text{th}} \), the comparator asserts \( \text{Flag} = 1 \), which propagates to an AND gate \( \text{G0} \) (figure 2). The other input of \( \text{G0} \), \( \text{Mask} \), is used for disabling pixels from responding digitally. This feature is exploited during the digital readout tests and imaging demonstration. The \( \text{Mask} \) together with the 4-bits for DAC in each pixel are set by a pixel-local 5-bit SRAM. Writes to SRAMs are synchronous to array scan. When \( \text{G0} \) outputs 1, a hit is generated (\( \text{Hit} = 1 \)) and the PL module is notified. Each PL is a full clock-less logic that controls the reset (CSA\_RST) of the CSA upon the readout of a hit and drives the hit information through the column structure. The internal structure of PL and its truth table are shown in figure 3 and table 1.

### 2.1.2 Column-wise priority chain and Address Bus (AB)

The priority logic signals propagate in columns. For the \( i \)th pixel, its PFI\(_i \) is connected to the previous \((i - 1)\)th pixel’s PFO\(_{i-1} \), and its PFO\(_i \) is fed into the next \((i + 1)\)th pixel’s PFI\(_{i+1} \). Pixels in the same column are daisy-chained in this fashion. Every pixel of a given column has a unique hard-coded 7-bit address in the form of pull-down switches. Encoded pull-down switches are connected to the column-shared Address Bus (AB) (green dashed box in figure 2). AB is weakly pulled up to all high by default. When AddrEN becomes active in a pixel, it pulls down the AB to its own unique address. The topmost pixel (0th) in a column has PFI\(_0 \) = 0. If there is no hit in any pixel (\( \text{Hit} = 0 \)), the PFO output is forced to 0 by G2, G5 & G7, which dictates that every pixel in the column has PFI = PFO = 0. When there is no active COL\_RST sent from the CRU module to every
pixel in the column simultaneously, the outputs of G3 & G4 are forced to 0. Once a pixel (e.g. \textit{i}th) gets a hit, due to the effects of G2, G5 & G7, PFO\textsubscript{i} = 1. Forced by G7, all pixels below the \textit{i}th pixel (denoted by \textit{j}th, \textit{j} > \textit{i}) will have PFI\textsubscript{j} = PFO\textsubscript{j} = 1. Forced by G8, any pixel with PFI = 1 won’t enable AddrEN even if it gets a hit. The above described logic forms a column-wise priority chain: only the pixel with a hit that has the lowest \textit{i} (highest priority) enables its AddrEN, and it disables all the pixels lower in the chain from asserting their individual addresses on the AB.

### 2.1.3 Column Readout Unit (CRU)

Each priority chain (column) is terminated by a Column Readout Unit (CRU) at the bottom of the column. CRU monitors the AB and validates the address change, then records the 7-bit address & 10-bit time stamp for the corresponding hit pixel. Upon the read of a hit, the CRU asserts COL\_RST = 1, which is fed back simultaneously to all the pixels in the column. Only the pixel that is pulling on the bus will respond to COL\_RST (see figure 3), which results in the analog reset of the CSA (CSA\_RST = 1), the removal of hit, and the release of the bus. When the bus is successfully released, the address seen by the CRU returns to all high. The CRU senses such condition and outputs Ready = 1. It indicates that a hit has been registered in the column and has not yet been read by the MUX. COL\_RST and Ready are kept high until this CRU is read by the MUX. R\_en is set to high by the MUX when it reads the associated CRU.

### 2.1.4 Multiplexer (MUX)

As shown in the blue dashed box of figure 2, a digital multiplexer (MUX) polls the status of each CRU sequentially, advancing at the falling edge of each clock cycle. It picks up the valid addresses and time stamps for the hit pixels from each CRU, then ships them off the sensor. A MARKER signal is asserted when the 0th column is polled to indicate the start of a frame. The index of the column being read can be calculated externally referencing to MARKER. A VALID signal is asserted when the column being read has a hit. The address and time stamp outputs are valid only when VALID = 1.

### 2.2 Readout operation and timing

A timing diagram of the readout process of a valid hit is shown in figure 4. It is assumed that there is only one hit pixel at Row 50, Column 0 and the system counter has an initial value of Sys\_Time[9 : 0] = 100. We also set Mask = 1 to enable the pixel response to hits.

| Table 1. Truth table of PL. X=do not care. |
|---------------------------------------------|
| \begin{tabular}{|c|c|c|c|c|c|}
| Input & AddrEN & \textbf{COL\_RST} & Initial & Final & \textbf{PFO} & \textbf{CSA\_RST} |
| \hline
| 0 & 1 & 1 & 0 & 0 & 0 & 0 |
| 0 & 1 & 1 & 1 & 1 & 1 |
| 0 & 1 & 0 & X & 1 & 1 & 0 |
| 0 & 0 & X & X & 0 & 0 & 0 |
| 1 & X & X & X & 0 & 1 & 0 |
| \end{tabular} |

\textbf{Figure 3.} Schematic of in-pixel Priority Logic (PL) circuitry.
Charges arrive at $t_1$, causing the CSA output to exceed the threshold of the comparator, resulting in $\text{Flag} = 1$. Since $\text{Mask} = 1$, a hit is generated ($\text{Hit} = 1$); hence, the single-pole-double-throw (SPDT) switch (figure 2) grounds the gate of $M_f$ from its original bias $\text{FB	extunderscore VREF}$ so the CSA maximally retains the charge signal. As $\text{PFI} = 0$, $\text{PFO}$ and $\text{AddrEN}$ become 1 accordingly. At this moment ($t_1$), the AB is pulled to the address of this pixel as well $\text{Addr} = 50$. At $t_2$ (rising edge of the clock in the CRU), the CRU senses the address change and outputs $\text{Addr}[6 : 0] = 50$, and waits for 4 clock cycles to confirm that the address change is not a transient phenomenon. At $t_3$, the CRU latches the address value and the time stamp from the system counter $\text{Time}[9 : 0] = \text{Sys	extunderscore Time}[9 : 0] = 105$. It also sends a reset signal $\text{COL	extunderscore RST} = 1$ back to the column. Although $\text{COL	extunderscore RST}$ is sent to every pixel in the column, forced by G3 in figure 3, only the pixel that is pulling the AB and is being read out will respond to the reset. The reset sets $\text{CSA	extunderscore RST} = 1$, which turns on the feedback transistor $M_f$, discharging $C_f$ so that the CSA output comes down towards the baseline. At $t_4$, the CSA output falls below the threshold, causing $\text{Hit} = 0$ hence $\text{AddrEN} = 0$ and $\text{PFO} = 0$. Once $\text{AddrEN} = 0$, CSA reset is done ($\text{CSA	extunderscore RST} = 0$) and $\text{Addr} = 50$ returns to all high. The CRU also sets $\text{Ready} = 1$ indicating there is a valid hit waiting to be read. Both the $\text{COL	extunderscore RST}$ and $\text{Ready}$ are removed when the CRU is polled at $t_7$. The time between $t_4$ and $t_7$ is non-deterministic and can be as high as 72 clock cycles. During $t_5 \sim t_7$, the MUX is Polling the CRU and shipping the data $\text{Addr}[6 : 0] = 50$.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Timing diagram of relevant signal activities during a hit and its readout. In-pixel, in-column, in-CRU and in-MUX signals are indicated in red, green, blue and purple dashed boxes, respectively.}
\end{figure}
and TIME[9 : 0] = 105 off the sensor. Since this pixel is in the 0th column, besides generating a VALID = 1, a synchronous MARKER is also simultaneously asserted. As the signal Polling (R_en) is driven by the falling edge of the clock, it has a half-clock-cycle delay behind the MARKER; therefore, it’s high from t6 to t8. When multiple pixels in the same column are hit simultaneously, the logic reads out and resets the hit pixels sequentially following their priorities in descending order. No hit is missed.

3 Measurements and experimental results

3.1 Threshold and noise

We applied a repetitive tail pulse with an amplitude $V_{TP} = 10\, \text{mV}$ on Gring (see the top-left insert in figure 2). An equivalent negative charge $Q_i = C_{inj} \times V_{TP}$ is injected at every falling edge of the pulse into the CSA in every pixel simultaneously. The response amplitude of the CSA is expected to be $V_{TP} \cdot \left( C_{inj}/C_f \right) \approx 11 \, \text{mV}$. The CSA responds to both positive and negative charges equally well; however, only the negative equivalent charge can bring the CSA_OUT above the threshold to generate hits. Also, we would like to avoid undershoots of the CSA output due to positive charge injections; therefore, tail pulses are chosen over a square wave. The repetition rate of tail pulses is chosen to be low enough such that all the hit pixels have sufficient time to be readout and reset before the next pulse arrives.

As shown in figure 5, an S-Curve for a single pixel is obtained by scanning the threshold, which is gradually lowered from well above the signal height where hit probability = 0. When the threshold is close to the injected signal height, a characteristic tapered transition from probability 0 to 1 due to noise appears. When the threshold is close to the baseline, the logic registers a hit every cycle regardless of the injected signal pulse; therefore, the computed probability (Apparent Probability) is bogusly well above 1. When the threshold is well below the baseline, the logic saturates and outputs no hit, although internally the discriminator constantly outputs 1. We determine the median and width of the baseline using the Apparent Probability $> 2$ part of the curve. We fit the transition part using the Cumulative Distribution Function (CDF) of Gaussian, $f(x) = \frac{1}{\sqrt{2\pi}} \left[ 1 - \text{erf} \left( \frac{x - \mu}{\sigma \sqrt{2}} \right) \right]$, to determine the mean ($\mu$) and width ($\sigma$) of the transition. This procedure is repeated for every pixel in the array. 4-bit DACs are set to 0 for all pixels while the global $V_{thg}$ is varied to achieve the threshold scan. Through the threshold scan procedure for the entire array, we extracted the baseline and transition’s location and width from recorded S-Curves of every pixel. The width ($\sigma$) of transition, which is an indicator of the noise of CSA output presented to the comparator, has a mean value of 1.2 mV (see the insert in figure 5). It is consistent with the analog noise measurement reported in [6]. The distribution of the baseline median over the array is shown in figure 6(a) and 6(c).

We write a set of values into the SRAM in each pixel to drive the 4-bit DAC to trim (reduce) the threshold differences between pixels in the array. The set of DAC values, $\{n_i\}$, are calculated from the parameters extracted from the threshold scans. All the 4-bit DACs share a globally adjustable step size $V_{\text{step}}$. $V_{\text{th}p} = V_{\text{step}} \times n_i$. Ideally, $V_{\text{th}p}$ should be as close to the baseline median while kept above the baseline noise width, to detect minimal signal amplitudes. This requirement points to a small $V_{\text{step}}$. However, at the same time, $V_{\text{th}p}$ should cover a maximal threshold dispersion of the
array in order to reduce the number of dysfunctional pixels due to insufficient trimming. Since \( n_t \) has only 16 values, it points to a large \( V_{\text{step}} \), contradicting the low threshold requirement. To find a balanced set of parameters, we minimize the quadratic sum, \( \sum_i (V_{\text{th}} - \text{baseline median})_i^2 \), by varying \( \{n_t\} \), \( V_{\text{step}} \) and \( V_{\text{thg}} \). We allow a small fraction of pixels with baselines that are far off to be excluded and subsequently disabled. We also disable defective and noisy pixels by setting \( \text{Mask} = 0 \). Disabled pixels are marked with black points in the 2D-figures. A representative set of parameters are \( V_{\text{step}} = 9 \text{ mV} \), \( V_{\text{thg}} = 532 \text{ mV} \), and 10% disabled pixels. After trimming with the optimized setting, we varied \( V_{\text{thg}} \) to perform the threshold scan again. The results show a greatly reduced width in baseline median distribution (figure 6).

### 3.2 Imaging with pulsed LED illumination

We placed a purple light LED \( \sim 2 \text{ cm} \) above the top surface of a Topmetal-II\(^{-} \) sensor. The sensor is covered by an opaque photo mask with a transparent T-shaped pattern. The T-shaped pattern is aligned with the center of the sensor (figure 7). The LED is driven by a train of narrow pulses with 10 \( \mu \)s width, 50 ms interval and 14 V Peak-to-Peak amplitude. A \( \sim 1 \text{ MHz} \) clock drives the CRUs and the MUX; therefore, the time it takes to read one frame (all 72 columns for once) is \( T_f \approx 72 \times 1 \mu \text{s} = 72 \mu \text{s} \). The sensor operates at the optimized threshold settings. We recorded many frames of hits induced by a large number of LED pulses. The photo mask was also rotated and displaced to cover different regions of the sensor. Hit location and time are reconstructed from data. A set of images showing the T-shape at four different orientations is in figure 8.

When a LED light pulse arrives at the sensor, multiple active pixels that receive the light generate a hit in each of them. Due to the column readout logic, hit pixels that are in the same column will have only one pixel that has the highest priority registered by the CRU. Since CRUs
Figure 7. Light pulse injection setup. Light from the LED is filtered by a photo mask with a T-shaped transparent opening before arriving at the sensor.

Figure 8. A set of images collected with the T-shaped photo mask placed at four different orientations. Black points mark the disabled pixels.

from each column work concurrently while reading off a single globally shared time, each CRU registers the hit time of the highest priority pixel in its column. Since the light pulse arrives at each pixel simultaneously, the initially registered time, which is from the highest priority pixel, is the same for all the CRUs (figure 9(a)). The MUX reads the registered time from each CRU in a round-robin fashion from one column to the next. When a CRU is read, the pixel of the highest priority in its column is reset, and the CRU subsequently registers the second-highest priority pixel. Since only the CRU has access to the global time, the hit time of the second-highest as well as all the lower priority pixels is determined by the readout rather than the actual arrival of the signal. Only the hit time of the highest priority pixel is physically meaningful. It is worth noting that starting from the second-highest priority pixel, the time difference between the \(i\)th-priority pixel and the \((i + 1)\)th-priority pixel in the same column equals the number of columns (72), which is the time interval between consecutive reads for a given CRU (readout time for one full frame). Figure 9(b) illustrates this phenomenon.

4 Summary and outlook

We successfully implemented a CMOS pixel sensor, Topmetal-II, for direct charge collection and imaging. The detailed design, behavior and performance of a column-based priority logic readout in the sensor are presented. The electrical measurements and imaging applications demonstrated the validity of such a readout scheme. In the current design, although the in-array clock-less logic could drive the hit pixel’s address to the edge of the array with minimal latency, the sequential logic nature of the CRU and the MUX limits the time it takes to discover the hit information to be beyond one clock cycle. To further reduce the readout latency, analog and clock-less logic could be designed at the edge of the array to detect the activities in the AB promptly. A polling style MUX could be replaced by a priority logic to read out the columns as well. We will investigate these options in future Topmetal sensor development in addition to improving the array uniformity.
Figure 9. Time stamping of hits. (a) Time stamps printed at the location of corresponding pixels resulting from one light pulse. The T-shaped light hit pattern is clearly visible. Insert illustrates the readout of column No. 38 in this fashion. (b) Time stamp read out as a function of clock cycle. Frame 0 reads out the initial time stamps from highest priority pixels in each column. Subsequent frames read out pixels hit by the same light pulse but with progressively lower priority.

Acknowledgments

This work is supported, in part, by the Thousand Talents Program at CCNU and by the National Natural Science Foundation of China under Grant No. 11375073. We also acknowledge the support through the Laboratory Directed Research and Development funding from Berkeley Lab, provided by the Director, Office of Science, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We would like to thank Christine Hu-Guo and Nu Xu for fruitful discussions.

References

[1] J. Millaud and D. Nygren, The column architecture-a novel architecture for event driven 2D pixel imagers, IEEE Nucl. Sci. Symp. Med. Imag. Conf. 1 (1995) 321.
[2] P. Yang et al., Low-power priority Address-Encoder and Reset-Decoder data-driven readout for Monolithic Active Pixel Sensors for tracker system, Nucl. Instrum. Meth. A 785 (2015) 61.
[3] M. Garcia-Sciveres et al., The FE-I4 pixel readout integrated circuit, Nucl. Instrum. Meth. A 636 (2011) S155.
[4] C. Hu-Guo et al., CMOS pixel sensor development: A fast read-out architecture with integrated zero suppression, 2009 JINST 4 P04012.
[5] C. Hu-Guo et al., First reticule size MAPS with digital output and integrated zero suppression for the EUDET-JRA1 beam telescope, Nucl. Instrum. Meth. A 623 (2010) 480.
[6] M. An et al., A Low-Noise CMOS Pixel Direct Charge Sensor, Topmetal-II , Nucl. Instrum. Meth. A 810 (2016) 144 [arXiv:1509.08611].