A 128×128 Current-Mode Ultra-High Frame Rate ISFET Array with In-Pixel Calibration for Real-Time Ion Imaging

Junming Zeng, Student Member, IEEE, Lei Kuang, Student Member, IEEE, Nicholas Miscourides, and Pantelis Georgiou, Senior Member, IEEE

Abstract—An ultra-high frame rate and high spatial resolution ion-sensing Lab-on-Chip platform using a 128 × 128 CMOS ISFET array is presented. Current mode operation is employed to facilitate high-speed operation, with the ISFET sensors biased in the triode region to provide a linear response. Sensing pixels include a reset switch to allow in-pixel calibration for non-idealities such as offset, trapped charge and drift by periodically resetting the floating gate of the ISFET sensor. Current mode row-parallel signal processing is applied throughout the readout pipeline including auto-zoning circuits for the removal of fixed pattern noise. The 128 readout signals are multiplexed to eight high-sample-rate on-chip current mode ADCs followed by an off-chip PCIe-based readout system on a FPGA with a latency of 0.15 s. Designed in a 0.35 μm CMOS process, the complete system-on-chip occupies an area of 2.6 × 2.2 mm² with a pixel size of 18 × 12.5 μm² and the whole system achieves a frame rate of 3000 fps which is the highest reported in the literature for ISFET arrays. The platform is demonstrated in the application of real-time ion-imaging through the high-speed visualization of sodium hydroxide (NaOH) diffusion in water at 60 fps on screen in addition to slow-motion playback of ion-dynamics recorded at 3000 fps.

Index Terms—ISFET, ultra-high frame rate, linearity conversion, current mode, in-pixel calibration.

I. INTRODUCTION

CMOS based Lab-on-Chip (LoC) technology has shown great promise over the last decade as it has enabled the realization of micro-systems with integrated sensors and instrumentation capable of biochemical monitoring and analysis [1]. Integrated sensors have benefited from their fabrication in CMOS technology for Lab-on-Chip applications, leveraging sub-micron scale devices to allow miniaturization of sensors, which has enabled creation of large scale arrays, mass production, as well as implementation of high sample rate systems with fast data transmission [2]. Such platforms have enabled breakthroughs in various applications ranging from next generation semiconductor-based DNA sequencing [3] to point-of-care diagnostics [4].

A major enabler for CMOS-based Lab-on-Chip platforms has been the invention of the Ion-Sensitive Field Effect Transistor in the 1970s [5], allowing the ability to introduce a chemical input to electronic circuits and systems. Modern ISFETs are formed using a MOSFET fabricated in an unmodified CMOS process with its gate extended to the top metal layer. Sensing takes place through the CMOS passivation layer (SiO₂,Nₓ) in contact with a solution. Due to the nature of the SiO₂,Nₓ being inherently sensitive to hydrogen ions (H⁺), variations in [H⁺] (concentration) will change the overdrive voltage of the ISFET device and thus will be sensed and transformed into a corresponding voltage or current signal. Additionally, selectivity to other ions can be realized by depositing ion-selective membranes on the passivation surface and ISFET-based solutions have been reported which can sense [K⁺], [Na⁺], [Ca²⁺] and [Mg²⁺] ions [6]–[10], making them suitable for biomedical applications. ISFET sensors are also generally small in size and easily scalable in CMOS, therefore a large array of sensors can be integrated with peripheral signal processing to form a complete System-on-Chip (SoC) as shown in Fig. 1.

![Fig. 1. On-chip ion imaging system using an array of ISFET pixels and on-chip parallel signal processing peripherals.](image-url)

Despite their benefits, ISFETs suffer from sensor non-idealities such as trapped charge and drift [11]. Trapped charge directly adds non-uniformly distributed DC offsets on individual ISFETs across an array which has been reported to be in the range of several volts [12] [13]. On the other hand, drift produces a temporal variation in the signal that could either have chemical or electrical origin and shifts the ISFET’s gate voltage even at chemical equilibrium [14] [15]. Consequently, several compensation methods have already been proposed which can be broadly classified into two main categories - chip post processing and on-chip circuit compensation [16]. In terms of post processing, the chip can be exposed to UV light irradiation to remove trapped charge, but this additional processing steps increase the cost and time associated with fabrication [17]. Alternatively, on-chip circuit compensation techniques can be used to eliminate non-idealities, which are further sub-divided into two main categories - using Programmable Gate ISFETs (PG-ISFET) [18] or using a reset switch on the floating gate [19]. PG-ISFETs are realized by adding a capacitor to the floating gate of the ISFET to compensate the effect of trapped charge at the expense of reducing the pH sensitivity. Conversely, direct feedback to the floating gate with a reset switch can effectively set the gate voltage and thus the operating point of the ISFET while eliminating the effect.
of trapped charge. Furthermore, the effect of chemical drift can also be minimized by periodic resetting. Adding a reset switch however, introduces electrical drift to the floating gate through leakage which can affect long term measurements. Nevertheless, it is suitable for high speed applications where the frequency of resetting is sufficiently high such that the effect of leakage at the floating gate is minimized.

Currently, there is strong focus on implementing robust sensing arrays with sufficient compensation for the aforementioned non-idealities. ISFET arrays are typically used for low frequency biomedical applications such as DNA amplification detection [4] with low requirements on speed and resolution. Other applications reported in the literature include, but are not limited to, monitoring of enzyme kinetics [21], food safety [22] and recording of neural activity [23]. However, we anticipate that there is significant unexplored potential in the area of chemical sensing for a high speed and high resolution ion imager that is able to visualize fast chemical reactions in real-time with integrated on-chip compensation techniques. Specifically, such an ion imager will allow characterizing sub-second diffusion of reactions across the array which will deliver new insights as to how ions diffuse in space and time. This will be beneficial in areas such as visualization of proton generation to allow characterizing reaction dynamics, demonstrated in this work using diffusion of NaOH, but also in future to allow the mapping of proton generation during DNA amplification with a high spatial-temporal accuracy, delivering new insights to molecular methods which could allow the optimization of the amplification technique [24].

In order to achieve a high speed readout, current mode signal processing methods are explored and applied here which, when compared to the traditional voltage mode methods, eliminate the need for op-amps as well as charging of voltage nodes that reduce bandwidth [25]. Instead, analogue signal processing is carried out using the current conveyor, with a similar role as an op-amp in voltage mode, while other signal operations such as amplification, addition and subtraction are greatly simplified typically by joining current branches together or through the use of current mirrors [26]. Multiple CMOS arrays have been reported to employ current mode operation especially in motion imagers [27], [28]. Among those, several demonstrate high resolution and high speed operation such as in [29], a 40 × 40 array with 25000 fps in a 0.35 μm process and in [30], a 192 × 124 array with 48000 fps in 0.5 μm CMOS technology. As a result, similar techniques can be applied in the case of high speed ion imaging as well.

Independent of the mode of operation, parallel processing is the main enabler for increasing the frame rate of a sensing array, with the drawbacks of mismatch and power consumption. Previously reported systems utilize single global readout circuits sharing the entire pixel array [19], [31], [32] whereas the temporal resolution of the sampled information is greatly reduced. Readout channels are on a per row/column basis and increasing the number of on-chip ADCs can decrease the number of pixels shared by one ADC and thus accelerate the scan rate of the ISFET array. Nevertheless, as the data throughput of the chip increases, the data bottleneck moves off-chip, since an external high-speed readout system is required for real-time acquisition and frame recovery. Traditional micro-controller-based systems with SPI-UART transmission links are limited on both speed and storage space. Instead, for ultra-high frame acquisition, more advanced solutions are preferred such as PCIe-based FPGA systems with on board SDRAM chips to ensure very high time resolution without being limited by the storage space of the off-chip memory.

In this context, we present a complete platform for ultra high-speed ion imaging comprising: (1) A Lab-on-Chip platform including a 128 × 128 ISFET array that uses a reset switch in pixel for compensating trapped charge and drift. The array operates in current-mode and employs high speed parallel signal processing on a per-row basis to facilitate ultra-high readout speeds. (2) A high speed digital readout system utilizing DDR3 memory and PCIe 2.0 × 4 transmission link achieves real-time frame acquisition at 3000 fps.

The rest of the paper is organized as follows: Section II introduces the ISFET and how linear input-output characteristics are achieved in current-mode. Section III describes the design considerations for each block of the signal pipeline covering the ISFET pixel to the current mode ADC. Section IV discusses the chip architecture and the system operation. Section V introduces the experimental setup together with the digital readout system and Section VI shows the system characterization as well as demonstration for real-time ion imaging. Lastly, Section VII concludes the paper. This paper is an extension of the work in [20], where Section II, III, IV are adopted from the original work.

II. THE ION SENSITIVE FIELD EFFECT TRANSISTOR

In this section, the ISFET sensor fundamentals are discussed as well as the biasing regime to achieve linear pH conversion.

A. pH Sensitivity

ISFET sensors fabricated in unmodified CMOS processes are designed as floating gate MOSFETs with the gate extended to the top metal layer which resides below the passivation layer [33]. As shown in Fig. 2, hydrogen ions presenting in an
electrolytic solution bind to the top passivation layer and are capacitively coupled to the floating gate. As a result, the ion concentration modifies the operating point of the device and can therefore be detected. The electro-chemical effects at the interface are modeled as in [18] and modify the gate overdrive voltage

\[
V_{OV(ISFET)} = V_{GS} - V_{th} 
\]

\[
V_{OV(ISFET)} = V_{chem} + V_{OV(MOSFET)} 
\]

\[
V_{chem} = \gamma + \alpha S N pH
\]

where \( \gamma \) captures any non-pH dependent terms and \( \alpha \) represents the deviation from the (maximum) Nernstian sensitivity at the electrical-chemical interface: \( S_N = 59 \text{mV} / \text{pH} \). For our fabricated chip the pH sensitivity is 16.7 mV and \( \alpha \) corresponds to 16.7 mV/59 mV= 0.28.

**B. ISFET Operation in the Triode Region**

Since the chemically-dependent \( V_{chem} \) is essentially sensed at the floating gate, the ion concentration can be conveyed to an output current in different forms, such as quadratic, logarithmic or linear. The quadratic form needs the highest requirements. However, this form is obtained with an ISFET operating in the weak inversion region where mismatch is critical. Furthermore, translinear circuits which are commonly applied to process signals from ISFETs operating in the weak inversion region are in need of critical matching therefore pixels using translinear circuits generally occupy a large area [34]. Obtaining linear \( V_{GS} \) to \( I_D \) characteristics is achieved by biasing in either the triode or the velocity saturation region of operation. However, operation in velocity saturation introduces large amounts of ficker noise and is normally associated with short channel MOSFETs, as long channel devices would require high electric fields and power to operate and would not be suitable for mass-integration [35].

Therefore, the ISFET is designed to operate in the triode region which linearly conveys \( V_{GS} \) to \( I_D \) in order to achieve a linear \( pH = I_D \) with less intrinsic device noise [4]. The output current in this regime is given by:

\[
I_D = \mu C_{ox} \frac{W}{L} (V_{OV(ISFET)} - V_{chem}) V_{DS} - \frac{V_{DS}^2}{2} \tag{4}
\]

where \( \mu \) is the charge-carrier effective mobility, \( C_{ox} \) is the oxide capacitance and \( W \) and \( L \) represent the device dimensions. As a result, the underlying pH sensitivity in this region (assuming any \( V_{DS}^2 \) terms are negligible) is given by:

\[
\frac{\partial I_D}{\partial pH} = -\mu C_{ox} \frac{W}{L} \alpha S_N V_{DS} \tag{5}
\]

which evaluates to a constant for a given \( V_{DS} \), transistor size and sensing material properties.

**III. CIRCUIT DESIGN**

In this section, the design considerations of the system are discussed. This includes the complete signal pipeline from the front-end sensing pixel to the ADC. The pixel architecture is discussed with the reset switch for in-pixel calibration while the biasing of the current mode pixel will be provided by the current conveyor in the following stage. The buffered signal will then be amplified and passed through an Auto Zering stage where the Fixed Pattern Noise (FPN) is suppressed, and it will subsequently be sampled using a current mode ADC. With this configuration, current mode signal processing is applied throughout the entire pipeline for high-speed and low-noise readout.

**A. Pixel Architecture**

In order to achieve in-pixel calibration, high spatial resolution, and high-speed readout, the pixel comprises of three transistors, as shown in Fig. 4. Switch S1 is used to reset the gate voltage of the ISFET and switch S2 acts as the readout selection switch. Resetting takes place by defining \( V_{GS} \) and \( V_{DS} \) respectively as discussed in Section II-B, where
S1 defines the $V_G$ to 1 V directly through a biasing voltage and S2 defines the $V_D$ to 2 V by the current conveyor shared by each row, so that $V_{GS}$ is biased at 2.3 V, and $V_{DS}$ is biased at 1.3 V with a PMOS operating under 3.3 V power supply. In this case, the non-uniformly distributed dc offset introduced by trapped charge that exists on the floating gate will no longer be a problem since a direct path to the floating gate exists during resetting. Resetting the gate is asynchronous and externally triggered since it is only used to compensate for non-idealities of floating gate devices. Specifically, the array is designed to reset once at the start of operation to remove trapped charge and set the same biasing point of the devices. In order to overcome the leakage problem introduced by the switch, periodic resetting ensures that the contribution of leakage at the floating gate is minimized between any two resets. This resetting interval is externally controlled via the on-chip digital controller serving as a globally-controlled compensation method to the switch leakage. Since both sides of the switch are roughly at a fixed bias voltage, the potential difference across them is negligible. Subsequently, periodic reset can also be used to compensate for long term drift as long as the amplitude of the chemical signal within the reset period can be detected by the system. When resetting happens, S1 and S2 are switched on at the same time for biasing but S2 is designed to have a large W/L ratio to minimize the capacitance which reduces the attenuation at the floating gate. S1 is designed to have a small W/L ratio to reduce its own parasitic capacitance and degrades the coupling effect of the pixel. As a result, for a compact pixel with limited top metal area, the parasitic capacitance added by the reset switch to the floating gate is an important factor in this capacitive network and therefore needs to be addressed. S1 is designed to have a small W/L ratio rather than using the minimum length transistor to reduce flicker noise. The total pixel area is 18 $\mu$m $\times$ 12.5 $\mu$m with a corresponding top metal area of 17.2 $\mu$m $\times$ 11.7 $\mu$m (due to the design rule of the CMOS process node) serving as the ISFET sensing membrane. We designed the pixel layout so that the pixel height is sufficient for the readout circuit which will be shared per row of pixels and 16 of such pixels in an array share the same height as the ADC in the later stage that enables easy expansion of the array in future designs.

### B. Current Conveyor & Auto-zeroing

In current mode signal processing, the current conveyor acts similarly to a voltage buffer in voltage mode. The current conveyor designed here is used to set a fixed $V_{DS}$ across each ISFET pixel and mirror the pixel current to subsequent stages as shown in Fig. 5. The front end is based on the topology in [23] an additional current mirror added at the output for further current amplification. The differential pair serves to copy the voltage from the Y side towards the X side while sinking no current into the X node. As a result, all the input current of the current conveyor will flow through the output transistor feeding the auto zeroing stage with an amplification gain of 4. The offset defined by the loop gain of the differential pair is simulated with input current varying within the full pH range.
as shown in Fig. 6, with a maximum absolute error of 0.22%, and a maximum relative error of 0.28% obtained from a post layout simulation.

The auto-zeroing stage is used for offset cancellation from the previous stages. Additionally, it reduces fixed pattern noise due to mismatch between pixels even when they are all reset at the same $V_{GS}$. S3 stores the offset by the intrinsic overlap capacitance at the gate node of the storage transistor (mainly $C_{GS}$), which is subsequently subtracted from the signal after double sampling. The charge injection introduced by S3 is a constant dc offset and thus can be ignored in a current subtraction operation. In addition, the auto-zeroed offset current during the first sample is several times larger than the second sample (actual signal) to guarantee its operation, and this is controlled by the size of the transistors.

C. Sample and Hold

The improved signal after the auto-zeroing stage (subtracted current) is then sampled and held towards the 10-bit current mode ADC. As shown in Fig. 5, the sampling switch is isolated by a series of cascoded current mirrors to eliminate high frequency charge injection noise. A regulated cascode current mirror at the output stage serves to boost the output impedance and ensure better driving ability. In addition, the option for an external ADC is considered by bypassing the sampled signal towards the output pad. An off-chip high speed ADC designed with smaller CMOS process node can be used to achieve an even higher frame rate in the future.

D. Current-mode ADC

The ADC introduces the most critical delay in the signal pipeline which is set by the time needed per conversion and this delay is higher than the pixel settling time which is in the order of nano-seconds. Therefore, a high speed ADC is essential in order to boost the frame rate. A current-mode ADC is preferred in this case to eliminate the need for op-amps performing I-V conversion which introduces additional delays even in cases of wide bandwidth op-amps. The Successive

![10-bit Current Steering DAC](image)

Fig. 7. The designed 10-bit current mode SAR ADC schematic diagram.

![10-bit Current Mode ADC](image)

![SAR Logic](image)

![10-bit Current Mode ADC](image)

![SAR Logic](image)

![10-bit Current Mode ADC](image)

![SAR Logic](image)

![10-bit Current Mode ADC](image)

![SAR Logic](image)

![10-bit Current Mode ADC](image)

![SAR Logic](image)

![10-bit Current Mode ADC](image)

![SAR Logic](image)

![10-bit Current Mode ADC](image)

![SAR Logic](image)

(a) Current comparator switching delay

(b) Current comparator switching delay point

Fig. 8. The maximum delay time of the current comparator is 1.55 ns with a switching point of 33.8 nA for both directions extracted from a post-layout simulation of the circuit block. In is the input current of the current comparator where $V_{comp}$ is the comparison result in voltage.

![Current Comparator](image)

(a) INL

(b) DNL

Fig. 9. The linearity plot of the designed ADC operating in its maximum speed (6 MHz) from post layout simulation. The dotted line states the effective INL and DNL defined by the full pH range verified in the experiment stage. An effective INL of 3 LSB and DNL of 2.6 LSB is observed.

Approximation Register (SAR) ADC designed here comprises three parts, a current comparator, a current steering Digital to Analogue Converter (DAC), and a digital controller logic. The block diagram and the relative schematic diagram are shown in Fig. 7.

1) Current Comparator: The current comparator’s output is determined by the direction of the input current. Therefore, it includes a two node circuit block with one input node and one output node. The current subtraction happens by connecting the two current branches $I_{in}$ and $I_{dac}$ together while sending the subtraction result to the input branch of the comparator. The current comparator designed here is adapted from [37], which achieves faster switching time than the traditional Trâff’s current comparator [38]. An additional inverter pair is added
at the end of the circuit to increase the switching strength even further. From post layout simulation as shown in Fig. 8, the maximum switching delay is 1.55 ns which is less than the delay caused by the digital block and the switching point is 33.8 nA, which is approximately 1 LSB of the designed ADC.

2) Current Steering DAC: The DAC designed here is based on the architecture from [39], with additional dummy switches to reduce instantaneous current transients during each approximation. In addition, all the outputs from each branch of the DAC are collected by a self-biased regulated cascode current mirror for higher driving ability. This architecture has the benefit of providing accurate current reference without using large size transistors. The maximum transistor size applied here is 32 µm/2 µm. The full scale output current is set to 40 µA with a maximum setup time of 15 ns for every iteration.

3) SAR Logic: The SAR logic is synthesized from a hardware description language (Verilog) to perform the transition stages as well as data communication. In order to reduce the setup delay for every computation, firstly the next sample-and-hold takes place one clock cycle before the actual computation, and secondly, the reset register always returns to half of the ADC’s full range which is exactly the same as the first approximation value.

The ADC linearity extracted from post-layout simulation at its maximum speed of 6 MS/s is shown in Fig. 9. The effective INL and DNL are extracted based on experiment result of the full pH range, corresponding to 3 LSB and 2.6 LSB respectively. At its maximum sampling rate, a column and the output MUX is located close to the bond pads. Scaling to 128 rows results in a frame rate of 2929 fps and a data throughput of 479.88 Mbps.

IV. SYSTEM-ON-CHIP

Fig. 10 shows the system architecture. The array comprises 128 × 128 ISFET pixels with each row sharing the same readout bus and each column sharing the same readout selection switch (S2). To eliminate charge redistribution when resetting and autozeroing the noise within the pixels, switch S2<6:4> stays closed throughout the reset phase together with the globally controlled S1. In this case, we are assuming the offset is within a certain range for one row of pixels.

Overall, this configuration is based on row parallel processing in order to increase the readout speed with a dedicated readout pipeline per row designed. Whenever a column is selected, all pixel outputs are amplified to the auto-zeroing block. Subsequently, eight 16-to-1 multiplexers cycle 16 rows to a dedicated sample and hold circuit and current-mode ADC. Pipelining is also applied at the input of the ADC by sampling the next pixel when the current one is still in its ADC conversion period, allowing extra clock cycles for the signal to settle.

A digital controller was designed and synthesized to receive incoming commands from a SPI master which are encoded as a 16-b SPI command. The command includes instructions for the chip such as generating the control signals for switches. A digital output multiplexer used in standard optical imagers is also included which effectively reduces the pin number of the chip without sacrificing the operating speed. The output MUX cycles between the eight 10-bit outputs from the ADCs during the time needed for the subsequent conversion.

The chip micro-photograph is shown on the right hand side of Fig. 11 with the chip fabricated in AMS 0.35 µm standard CMOS technology occupying a die area of 2.6 mm × 2.2 mm. The sensing membrane is the unmodified Si3N4 passivation layer used in the CMOS process. For compactness, each row-parallel readout channel is designed with the same height as the pixels. Further scaling this structure could be easily realized by stacking more numbers of pixels and its peripheral readout channel. ADCs are designed to have the same height as 16 channels so that 8 current mode ADCs match exactly as the height of the array. The digital controller is located by the side of the array for global control with minimum delay and the output MUX is located close to the bond pads.

V. EXPERIMENT LAB-ON-CHIP PLATFORM AND READOUT

This section describes the complete Lab-on-Chip platform as well as the digital readout system used to achieve low latency and high throughput data acquisition.

A. Platform setup

An outline of the data acquisition platform is shown in Fig. 12, where the left hand side of Fig. 11 shows the exposed sensing chip. The chip is glued on a disposable cartridge PCB with the bond wires and sides of the chip encapsulated using glob top epoxy. A plastic tube is glued around the encapsulation to create a well on top of the chip in order to
host a solution and a Ag/AgCl reference electrode is inserted in the solution provides biasing. The reference electrode is made by a 1 mm-thick silver wire chloridized using a redox reaction in KCL 3M biased at 1.3 V. The cartridge PCB is mounted on top of a main PCB which provides biasing circuits and buffering of the digital data output from the chip. The main PCB is subsequently connected to a FPGA via GPIO with the FPGA plugged into a PC motherboard using a PCIe connection (described below).

The chip clock is provided by the FPGA which is clocked up to 66.7 MHz in order for the ADC to achieve its maximum sampling rate of 6 MS/sec. This clock is also programmable for data acquisition at lower rates. The total power consumption for the platform is 2.726 W comprising 376 mW for the chip, 132 mW for the digital buffer on the main PCB, 218 mW for the biasing peripheral circuits, and 2 W for the FPGA. At lower sampling operations, the chip power reduces down to 320 mW.

B. PCIe-based Digital Readout System

In order to process data at high throughput from the chip, a digital readout system is designed based on a FPGA development board. The purpose of this system is to ensure that there is no bandwidth restriction introduced by the peripheral systems so that the chip can be operated at its maximum speed. Previous high-speed data acquisition implementations use on-board memory to buffer the data in high speed but read out slowly [40]. This way, the stream of ion images can only be captured for a limited time period before the buffer overflows. In this case, we present a system that can read and write data communicated from the chip at its full speed and display the ion image in real time. The system can also remove the constraint on when the reaction happens due to the limited buffer size, since the high speed PCIe interface supports much faster data rate than that between the chip-FPGA interface.

The system block diagram is shown in Fig. 13, which consists of three main sub-blocks: (1) A chip interface block for sending instructions and receiving data. (2) A DDR3 memory controller that buffers data in an array-based cluster at 100 MHz. (3) A PCIe controller that fetches data from the DDR3 memory and sends to the PC via Direct Memory Access (DMA). The timing diagram describing how the system operates is shown in Fig. 14, which includes the chip calibration stage for trapped charge removal, the continuous readout stage with ultra high speed, periodic calibration during readout, and the data exchange from the chip to the FPGA. After the chip enters the readout mode, array data are continuously sent out from the chip to the FPGA, synchronized by the update signal, which triggers the FPGA for further data processing.

Every 25 samples of the 10-bit ADC quantization result are concatenated into the array SPIQ_out_buf[255:0] while the stored data in SPIQ is ready to be sent to the FIFO after one frame of the array has been fetched. Periodic asynchronous calibration is applied with a flexible interval for compensation.

The system realizes read and write simultaneously to achieve real-time data transmission. During data read, the readout controller instructs the chip to start data transmission through the chip MUX, buffered to a FIFO in between chip clock domain and DDR3 controller clock domain. In the meantime, the PCIe controller keeps cycling the DDR3 flag and monitors if a cluster of data is received and stored. If the data has been correctly stored in the DDR3 chip, the PCIe controller then fetches the data towards the PC.

On the PC side, a user interface is programmed in Python which utilizes multi-threading to achieve minimum latency. The main thread controls the PCIe driver that buffers the received data from the DDR on PC, with other threads handling frame recovery, frame display (at 60 fps for a standard display), and frame storage on an SSD disk (full data packet). Overall, the system achieves a latency of 0.15 s which is suitable for real-time ion imaging display.

The digital readout system is implemented based on a FPGA development board namely ARTIX A7103 including
a Xilinx Artix-7 (xc7a100t-2fgg484i) FPGA. 2 sets of DDR3 chips (Micron mt41j256m16ha-125) with 8 Gb storage, and a physical PCIe interface that supports PCIe 2.0 protocol with a 4 channel configuration. This way, a maximum throughput of 180 MB/sec can be achieved. For comparison, the chip’s maximum data throughput is 61.4 MB/sec calculated using:

\[
3000 \text{fps} \times 128 \times 128 \times 10 \text{bit} = 61.4 \text{ MB/sec}
\]

VI. CHARACTERIZATION AND DEMONSTRATION OF ION IMAGING

This section is divided into two parts. In the first part, we characterize the sensor performance at an array level. In the second part, real-time ion imaging is demonstrated using the array, with the chip operating at different clock speeds.

A. ISFET Array Characterization

1) pH characterization: The linear pH response of the ISFETs operating in the triode region is verified by inserting...
Various pH buffers ranging from 4 to 10 to a starting reference
solution of pH 7. Fig. 15 shows the array output and pixel
spread after offset cancellation due to row and ADC mismatch,
using the same method as stated in [22]. There exists a
variation in the ADC output image due to the IR drop across
the chip since the chip drains high current. However, for the
measurement of pH change, this is not a problem because the
difference in pH will be converted into a difference in ADC
quantization result. Additionally, the designed ADC

\[ R^2 = 0.999 \]

is over-engineered to have a wide dynamic range and is
therefore still able to sample the array output including any
DC variation. Utilizing the Digital Double Sampling [41] in
the FPGA also results in the output at pH 7 being relatively
flat, which also indicates that trapped charge is compensated
when using the pixel reset switch. During the experiments the

\[ \text{Mean} = 513 \]
\[ \text{Sigma} = 5.9 \]
\[ \Delta \text{pH} = 0.012 \]
reference electrode is grounded and the clock speed is set at 22 MHz corresponding to a 1 MS/sec ADC and a frame rate of ~1000 fps.

The resulting pH after the reaction is recorded across five designed chips and also validated using a commercially available pH meter. The resulting output at different pH follows a linear relationship as shown in Fig.16 whereas the increase in pH is reflected as an increase in the ADC output since the ISFET device is designed as a PMOS. A gain variation in each chip is observed due to mismatch across the channels and the growth of error bar as pH changes reflects the process mismatch of the fabricated chips. Furthermore, the pH resolution can be improved in future designs by reducing the dynamic range of the ADC which in this case is too large even for the whole pH range. The pH sensitivity is recorded as 50 LSBs/pH with a linearity of $R^2 = 99.9\%$. The quantization noise in this case is 0.5 LSB = 0.01 pH. Later in this section we show that the noise of the circuit is the main noise factor which downgrades the minimum detectable pH performance.

2) Surface Membrane Sensitivity: Fabricated in AMS 0.35 μm CMOS technology, this chip has an intrinsic passivation layer $Si_3N_4$ serving as the hydrogen ion sensing membrane. Its sensitivity is calculated as

$$S_{ref} = \Delta V_{ref} / \Delta D_{out}$$

(7)

$$S_{pH} = \Delta pH / \Delta D_{out}$$

(8)

$$S = S_{ref} / S_{pH} = mV / pH$$

(9)

where $S_{ref}$ is the sensitivity acquired from the reference electrode ramp, $S_{pH}$ is the sensitivity acquired from the pH buffer injection, $D_{out}$ is the ADC output digital code and $S$ is the membrane sensitivity.

A reference electrode ramp test was applied from 0 V to 250 mV at constant chemical conditions using a pH 7 buffer. Fig.17 shows the resulting output from a 80-pixel cluster. Using a linear approximation, 250 mV/820 LSB can be extracted corresponding to a 16.7 mV/pH membrane sensitivity. This is comparable to other recent works in the same process using $Si_3N_4$ as the passivation layer [36] even though the introduction of the reset switch could impact the sensitivity. Leakage through the switch sinks a portion of the useful pH signal rather than coupling all the change to the floating gate. A very low leakage switch or high speed sampling and recovery could however improve this.

Taking into account the capacitive network formed at the floating gate, a 0.45-fold attenuation is observed from the surface membrane to the floating gate verified by post layout simulations. Consequently, the effective sensitivity at the floating gate reduces to 7.5 mV/pH.

3) Noise Characterization: For noise characterization, the power spectral density of the raw output is calculated while recording at conditions of chemical equilibrium (buffer pH 7, $V_{ref} = 0 V$) for 30 minutes. The input referred PSD is then derived by dividing the output PSD by $S_{ref}^2$ which was previously reported in [36], [42] and the result is shown in Fig.18. In order to calculate the minimum detectable pH, the noise amplitude is obtained by integrating the input-referred PSD across the frequencies of interest - cross array diffusion in milli-seconds, (10 Hz to 10000 Hz was selected for high speed imaging) [43]. As a result, a minimum detectable pH of 0.24 pH is obtained at 1000 fps. This number downgrades to 0.45 pH at 3000 fps mainly due to the ADC bandwidth. The figure also suggests that the dominant noise source is the flicker noise which could be reduced in future designs using larger transistors for lower noise performance. However, we have seen in typical reactions generating protons such as DNA amplification, a 1 pH change is expected with a noise margin of 0.5 pH [24], therefore in this context a 0.45pH resolution is acceptable.

4) Minimum and Maximum Reset Frequency: The introduction of the reset switch at the floating gate of the ISFET compensates trapped charge by providing direct access to the floating gate. Additionally, it acts as a high pass filter that only allows high frequency chemical changes to be monitored while introducing a leakage path which is similar to chemical drift in behaviour. The drawback, however, is that only a certain rate of pH change can be monitored provided the switch resets periodically.

To investigate this, an experiment using a pH ramp from 7 to 9 is shown in Fig.19 which is then left to leak. The result suggests a leakage rate of three pH per eight seconds. This number does not indicate the actual leakage rate of
the switch since during pH ramp, leakage is taking effect as well. However, the experiment suggests a typical reset rate for applications such as monitoring ion diffusion which happens extremely quickly. The maximum reset frequency is limited by the chemical to noise ratio, provided that a minimum detectable pH is observed within each reset period. This number is flexible since different diffusion rates can have different reset periods. Experimentally, based on a pH injection, we observe a 7 pH change in around 0.1 sec, suggesting a 6.4 ms/0.45 pH, where 0.45 is the minimum detectable pH. As a result, the maximum reset frequency is around 150 Hz. The minimum reset frequency is limited by the leakage rate of the switch, provided that the floating gate of the ISFET is brought back to the maximum linear biasing region while leakage is causing drift of the floating gate potential away from it. Fig. 19 suggests a minimum reset frequency of around 10 Hz for good data recovery. Both the maximum and the minimum reset frequency is defined by estimation and they only suggest a range for the reset period. As we test across different chips this value changes accordingly with the experiment setup and the mismatch in leakage rate.

### Table I

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Technology</strong></td>
<td>0.35 µm AMS 2P4M CMOS</td>
</tr>
<tr>
<td><strong>Supply voltage</strong></td>
<td>3.3 V (regulated from 12 V)</td>
</tr>
<tr>
<td><strong>Membrane (Sensitivity)</strong></td>
<td>Si$_3$N$_4$ (16.7 mV / pH)</td>
</tr>
<tr>
<td><strong>Array size</strong></td>
<td>128 × 128</td>
</tr>
<tr>
<td><strong>Chip size</strong></td>
<td>2.6 mm × 2.2 mm</td>
</tr>
<tr>
<td><strong>Pixel size</strong></td>
<td>18 µm × 12.5 µm</td>
</tr>
<tr>
<td><strong>Power consumption</strong></td>
<td>376 mW$^1$</td>
</tr>
<tr>
<td><strong>Reference Electrode</strong></td>
<td>Ag/AgCl</td>
</tr>
<tr>
<td><strong>pH resolution</strong></td>
<td>0.24 pH (0/1000 fps)</td>
</tr>
<tr>
<td><strong>ADC pH resolution</strong></td>
<td>50 LSB / pH</td>
</tr>
<tr>
<td><strong>ADC dynamic range</strong></td>
<td>0-100 µA</td>
</tr>
<tr>
<td><strong>ADC clock speed</strong></td>
<td>up to 66.7 MHz</td>
</tr>
<tr>
<td><strong>Chip throughput</strong></td>
<td>up to 61.4 MB / sec</td>
</tr>
<tr>
<td><strong>System latency</strong></td>
<td>0.15 s</td>
</tr>
<tr>
<td><strong>Frame rate</strong></td>
<td>3000 fps</td>
</tr>
<tr>
<td><strong>Calibration</strong></td>
<td>Yes (with reset switch)</td>
</tr>
</tbody>
</table>

$^1$ 2.726W total power consumption for the entire system.

### B. Ion Imaging

Following the pH characterisation, we then demonstrate the complete platform for ultra-fast ion imaging visualizing...
ph diffusion in real time. At constant conditions of pH 7 and $V_{ref}$ biased at 0 V, sodium hydroxide (NaOH) powder is inserted at different locations on top of the array. This way, a significant pH spike can be induced as NaOH is a strong base and the overall solution pH increases sharply from 7 to 14. Fig. 20 shows the diffusion across the array recorded at 1000 fps, starting from the top left corner until the whole array is covered. Furthermore, Fig. 21 shows recording in 3000 fps by placing a single NaOH tablet on the top right corner of the array. Real-time observation at 60 fps can also be visualized reliably on screen using the platform. Comparing the two measurements, the ultra-high speed at 3000 fps contains more noise and fewer active pixels but it recovers additional diffusion dynamics. Still, both measurements show promising performance in revealing diffusion at a very high speed.

VII. CONCLUSION

The work presented here achieves the highest frame rate reported in literature. The complete system specification is summarised in Table I. Current-mode operation has been shown to be advantageous in two ways, by ensuring linear pH-current conversion and by facilitating high-speed readout with low settling time and high-speed peripherals. Auto-zeroing and pixel reset have been included to reduce fixed-pattern noise and compensate for ISFET non-idealities whereas subsequent parallelization and simultaneous sampling of 8 rows serve to increase the readout speed. A high-speed readout and data acquisition system has been designed to ensure that the chip can operate at its full speed. The complete system has been demonstrated in ultra-high-speed ion imaging, showing real-time imaging of NaOH diffusion in water with resolution of the order of milliseconds.

A comparison of this work with similar works in the literature has been summarized in Table II. A Figure of Merit has been derived for comparison which represents the array speed and is defined as FoM = Number of Pixels × Frame Rate. The proposed system achieves the highest frame rate and the highest FoM in the literature considering prior works. We anticipate that this ion imager will be used in the next-generation chemical sensing platforms, offering novel insight into real-time ion dynamics.

REFERENCES


# Table II

<table>
<thead>
<tr>
<th>Specification</th>
<th>[3]</th>
<th>[44]</th>
<th>[45]</th>
<th>[21]</th>
<th>[46]</th>
<th>[22]</th>
<th>[23]</th>
<th>[40]</th>
<th>This Work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process [(\mu \text{m})]</td>
<td>0.35</td>
<td>0.3</td>
<td>0.18</td>
<td>0.35</td>
<td>0.18</td>
<td>0.065</td>
<td>0.15</td>
<td>0.5</td>
<td>0.35</td>
</tr>
<tr>
<td>Array size</td>
<td>13 M</td>
<td>128 × 128</td>
<td>64 × 64</td>
<td>256 × 256</td>
<td>32 × 32</td>
<td>512 × 576</td>
<td>128 × 128</td>
<td>128 × 128</td>
<td>128 × 128</td>
</tr>
<tr>
<td>Pixel size [(\mu \text{m}^2)]</td>
<td>3.8 × 3.8</td>
<td>23.55 × 23.55</td>
<td>10 × 10</td>
<td>10.2 × 10.2</td>
<td>26 × 26</td>
<td>4.4 × 4.4</td>
<td>2 × 2</td>
<td>7.8 × 7.8</td>
<td>18 × 12.5</td>
</tr>
<tr>
<td>Frame Rate [fps]</td>
<td>N/A</td>
<td>58</td>
<td>1200</td>
<td>500</td>
<td>1000</td>
<td>375</td>
<td>1933</td>
<td>2000</td>
<td>3000</td>
</tr>
<tr>
<td>FoM [×10^3]</td>
<td>N/A</td>
<td>950.2</td>
<td>4915</td>
<td>32,768</td>
<td>1024</td>
<td>110,592</td>
<td>31,670</td>
<td>32768</td>
<td>49,152</td>
</tr>
<tr>
<td>Power [mW]</td>
<td>N/A</td>
<td>N/A</td>
<td>32</td>
<td>N/A</td>
<td>11.286</td>
<td>195.5</td>
<td>N/A</td>
<td>N/A</td>
<td>376</td>
</tr>
<tr>
<td>Real-time</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Yes</td>
<td>Yes</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Yes</td>
</tr>
<tr>
<td>Compensation</td>
<td>N/A</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

1 The ADC is off-chip.
2 The recorded frame rate is based on simulation.


Nicholas Miscourides (S'15) received the M.Eng. degree in Electrical and Electronic Engineering in 2014, the M.Res. degree in Advanced Computing in 2015 and the Ph.D degree in Electrical Engineering in 2019 all from Imperial College London, UK. He is currently a Research Associate at the Centre for Bio-inspired Technology, Department of Electrical and Electronic Engineering, Imperial College London. His research interests include mixed-signal IC design for integrated chemical sensing in CMOS, large-scale ion imaging and optimization of chemical sensing for biomedical applications. He was the recipient of the Sir Bruce White Prize 2014 for the best M.Eng. project with a focus on analogue IC design and co-author of the Best Paper Award at ISCAS 2018.

Pantelis Georgiou (AM'05-M'08-SM'13) received the M.Eng. degree in electrical and electronic engineering and the Ph.D. degree from Imperial College London (ICL), London, U.K., in 2004 and 2008, respectively. He is currently a Reader with the Department of Electrical and Electronic Engineering, ICL, where he is also the Head of the Bio-Inspired Metabolic Technology Laboratory, Centre for Bio-Inspired Technology. His research includes bio-inspired circuits and systems, CMOS based Lab-on-Chip technologies, and application of microelectronic technology to create novel medical devices. He has made significant contributions to integrated chemical-sensing systems in CMOS, conducting pioneering work on the development of ISFET sensors, which has enabled applications, such as point-of-care diagnostics and semiconductor genetic sequencing and has also developed the first bio-inspired artificial pancreas for treatment of Type I diabetes using the silicon-beta cell. He received the IET Mike Sergeant Medal of Outstanding Contribution to Engineering in 2013. In 2017, he was also awarded the IEEE Sensors Council Technical Achievement award. He is a member of the IET and serves on the BioCAS and Sensory Systems technical committees of the IEEE CAS Society. He is also the CAS representative on the IEEE Sensors council and the IEEE Distinguished Lecturer in Circuits and Systems.