A Synchronous Neural Recording Platform for Multiple High-Resolution CMOS Probes and Passive Electrode Arrays

Gian Nicola Angotzi, Mario Malerba, Fabio Boi, Ermanno Miele, Alessandro Maccione, Hayder Amin, and Luca Berdondini

Abstract—Electrophysiological signals in the brain are distributed over broad spatial and temporal scales. Monitoring these signals at multiple scales is fundamental in order to decipher how brain circuits operate and might dysfunction in disease. A possible strategy to enlarge the experimentally accessible spatial and temporal scales consists in combining the use of multiple probes with different resolutions and sensing areas. Here, we propose a neural recording system capable of simultaneous and synchronous acquisitions from a new generation of high-resolution CMOS probes (512 microelectrodes, 25 kHz/electrode whole-array sampling frequency) as well as from a custom-designed CMOS-based headstage. While CMOS probes can provide recordings from a large number of closely spaced electrodes on single-shaft devices, the CMOS-based headstage can be used to interface the wide range of available intra- or epi-cortical passive electrode array devices. The current platform was designed to simultaneously manage high-resolution recordings from up to four differently located CMOS probes and from a single 36-channels low-resolution passive electrode array device. The design, implementation, and performances for both ICs and for the FPGA-based interface are presented. Experiments on retina and neuronal culture preparations demonstrate the recording of neural spiking activity for both CMOS devices and the functionality of the system.

Index Terms—BMI, CMOS neural probe, neural recording devices.

I. INTRODUCTION

DURING the last decades, a strong research effort was allocated to the development of technologies for neuroscience research, with the aim of improving the experimental capabilities to precisely and selectively perturb neurons and to record neuronal signals. In addition, very recently, the topic aroused also the interest of private companies who are investing in neurotechnologies for biomedical applications [1]. However, while the capabilities to perturb brain activity were remarkably improved, for instance with the advent of optogenetics [2], advances in recording neurotechnologies are still needed.

Electrophysiological signals in the brain are distributed over broad spatial and temporal scales, ranging from single neuron spikes to slow signal oscillations emerging from collective neuronal activations in different brain regions [3]. Furthermore, a fundamental interplay existing among spiking activity in local brain circuits and the state imposed by low-frequency field potential signals has been recently revealed [4]. This gives rise to a massively distributed cellular processing that characterizes the implementation and execution of brain functions. However, studying such orchestrated communications among cells and brain circuits involving a wide spectrum of temporally and spatially distributed electrophysiological signals, remains challenging.

A possible approach investigated in this work, consists in developing an electrode-based neural recording system capable of combining recordings of neural activity from an extensive number of single neurons in brain circuits as well as bioelectrical signals from large brain areas. Indeed, the simultaneous recording of the spiking activity of many single neurons in brain circuits with high spatial and temporal resolution is fundamental to study how complex neural systems process information from the micro to the mesoscopic spatial scales [5]. Different approaches were proposed to this aim and, nowadays, it is possible to record neural spiking activity in vivo from up to hundreds of single neurons. Among the different technologies proposed for further upsampling over several thousands the number of simultaneously recorded neurons, the approach based on the use of monolithic CMOS circuits is one among the most promising [6] and a first CMOS-probe capable of recordings from 384 over 960 electrodes was recently demonstrated [7]. CMOS technology permits the implementation of compact monolithic devices integrating dense arrays of microelectrodes (micro electrode arrays or MEAs) with on-chip/in-pixel circuits for signal conditioning and multiplexing and capable of whole-array sub-millisecond temporal resolution for spike recordings.
at high signal to noise ratios. On the other hand, the resolution of conventional passive MEA devices, such as epi-cortical probes, might be sufficient for large-area recordings at higher spatial scales. Further, a wide range of different MEA-probes already exists and could be exploited.

In order to combine the performances of both generations of active and passive devices, in this work we propose a system that exploits two custom designed CMOS ICs, respectively for high-resolution neural recordings with active monolithic CMOS-probes and for recordings from passive probes integrating a small number of electrodes. In particular, we designed and engineered the system for simultaneous and synchronous acquisition of signals from multiple probes. As recently introduced in [8], the system exploits high-resolution CMOS neural probes to resolve a large number of spiking neurons in local circuits with sub-millisecond resolution [9], [10]. Concomitant low-density recordings for mapping of neuronal signals across brain regions can be achieved with a custom designed digital output ASIC that integrates 36 low-noise amplifiers and that can be connected to commercially available or custom made passive electrodes. Here, we present the design, implementation and validation of the sensing performance of the two recording sub-systems and of the entire platform as obtained prior to in vivo through bench-tests and recordings from ex vivo retinal mounts and in vitro neuronal networks.

The paper is organised as follows: the architecture of the whole recording system will be introduced in Section II while the detailed circuit description of the two recording frontends will be given in Sections III and IV respectively for the CMOS neural probe and for the low density CMOS headstage. The measured electrical, electrochemical and mechanical properties will be discussed in Section VI, right before preliminary results obtained from ex vivo and in vitro recordings (see Section VII). Conclusions and future perspective will be finally discussed in Section VIII.

II. SYSTEM ARCHITECTURE

The proposed system for multiscale bioelectrical recordings of brain activity consists of a base unit that can mount one headstage for low-density recordings from up to 36 passive channels and up to four distinct headstages for high-density recordings from implantable CMOS neural probes (see Fig. 1).

For the case of the low-density CMOS headstage, both area and power constraints can be relaxed. As shown on the right side of Fig. 1, signal amplification and analog to digital conversion are both performed on-chip while the resulting digital data is streamed out through a serial peripheral interface toward the acquisition board.

The design of implantable CMOS neural probes capable of simultaneous high-density recordings in the brain from several hundreds to thousands of electrodes, on the contrary, is challenging because minimum area and ultra low power are mandatory constraints to limit brain tissue damages during recordings. For this reason, in our solution, neural signal amplification of the CMOS-probe is performed on-chip (see Section III for details, first generation with 512 electrodes) while digital conversion is performed off-chip using commercial off-the-shelf ADCs. More in detail, neural signals are amplified by the on-probe circuitry, multiplexed into 16 different analog traces and routed off-probe through I/O pads. Each analog output is buffered with an operational amplifier (OPA365, Texas Instruments) before processing it with a 12-bit, 3Msample/s analog to digital converter (MAX11105, Maxim Integrated). The resulting stream of digital data is transmitted to the acquisition board via 16 separated serial interfaces.

The base unit finally comprises an FPGA-based acquisition module, an acquisition board and a graphical user interface running on a PC (modified BrainWave X, 3Brain GmbH-Switzerland [11]). The acquisition module is designed on an Opal Kelly ZEM4310 integration module based on an Altera Cyclone IV FPGA. The unit provides control signals required for synchronous operation of the headstages, performs programmable band pass filters on the acquired digital samples and implements a high data rate camera link interface with a PC for a data acquisition software for visualization and saving. Furthermore, the acquisition board provides a low noise power supply to the headstages, mandatory requirement for neural recordings.

III. CMOS NEURAL PROBE

Standard CMOS technologies can be used to design a novel class of neuroelectronic devices for monitoring the electrical activity of large neuronal networks by means of thousands of densely packed recording channels integrated on single-shaft silicon probes. Prior passive [12]–[14] neural probes used dedicated metal wires to route the sensing electrodes toward low noise amplifiers. Such one-to-one mapping results in a limited number and/or density of recording sites. Prior active solutions, conversely, integrated a large number of recording sites along the shaft but, still, they aimed at routing out only a subset of electrodes from such large array of electrodes [7],[15], [16]. By doing so, circuits required for amplification are placed outside of the sensing electrode area, where area and power design constraints are relaxed, thus permitting to reduce the electronic contribution to the overall noise on the recorded signals. A time multiplexed active neural probe was recently presented [17]. However, also such device places signal conditioning circuits outside of the active area and, as a result, the proposed architecture is hardly scalable in terms of number of recording sites. Our CMOS neural probe, conversely, exploits the area and power efficient modular solution previously introduced in [10] to implement a probe with whole array readout capabilities.

As shown in Fig. 2, in our first prototype realized in a standard 180 nm CMOS technology, multiple instances of the same analog frontend are repeated along the probe shaft for a total number of 512 electrode-pixels while a digital unit is integrated on-probe for control of the device operation and for bidirectional communication with the base unit. Each analog frontend module comprises 32 electrode-pixels with in-situ amplification, a column buffer and a programmable gain amplifier for further signal amplification. An active feedback loop for the auto-zeroing circuit, shared among multiple electrode-pixels, permits the cancella-
Fig. 1. Picture (left side) and block diagram (right side) of the proposed acquisition system designed for synchronous recordings from multiple CMOS neural probes and passive multielectrode array devices connected to a custom-designed CMOS headstage.

Fig. 2. Schematic circuit architecture of the CMOS neural probe (top side) and example of a timing diagram in which the readout sequence does not include pixel number two (bottom side). As shown in the layout floor plan illustrated on the left side, the sensing area along the elongated shaft integrates multiple instances of the same low-area and low-power analog frontend comprising 32 electrode-pixels that are read out in a time division multiplexed fashion. A digital unit is integrated to provide timing signal required for correct operation and for bidirectional communication with the base unit.

A. Pixel-Electrode

The in-pixel circuit (see panel A in Fig. 2) was derived and modified from the solution described in [18]. The sensing electrode is implemented through a 14 μm × 14 μm opening on the top insulation layer (SiO₂), exposing the metal plate realized with the 180 nm CMOS technology top metal layer. The input stage is a PMOS source follower (transistor M₁) while the feed-forward amplifier A_FF is a PMOS differential pair that provides 40 dB of signal amplification in the 0–4 kHz signal band. During normal operation the main amplifier A_FF works in open
loop mode while the current $I_{pxl} + I_{AZ}$ sets the bias point at the inverting input of $A_{EF}$. This permits to compensate the DC offset arising from the electrode-electrolyte interface. Due to the drifting of such DC offset, however, the feedback loop has to be periodically closed to refresh the capacitor $C$ with the correct voltage that controls the current $I_{AZ}$. Differently from the solution proposed in [18], to save area and power consumption, the feedback amplifier $A_{FBAZ}$ (panel B in Fig. 2) is shared among a group of 32 electrode-pixels. The pixel occupies an area of 25 $\mu$m $\times$ 25 $\mu$m while maximum power consumption is 3 $\mu$W. A further reduction of power consumption is achieved by turning on the $A_{FBAZ}$ feedback amplifier only when the autozeroing loop is active (signal $AZ[i]$ high).

### B. Column Buffer

The circuit of the column buffer is implemented by transistors $M_2$ to $M_6$ (see panel C in Fig. 2). Transistors $M_2$ and $M_3$ are integrated into each recording site while the rest of the circuit is shared among the 32 electrode-pixels of the analog frontend. By enabling the readout of the $i$-th pixel through $M_3$ (signal $read[i]$) data locally amplified at the electrode-pixel site are read trough the circuit in a time division multiplexing fashion. Similarly to the architecture described in [19], the local biasing circuitry (transistors $M_6$ to $M_1$) improves the precision by forcing the current on the $M_4-M_5$ branch to be equal to $I_{CB}/2$ while the feedback circuitry through amplifier $A_{FBCB}$ allows to mitigate short channel effects by matching the drain-source drop voltages of transistors $M_2$ and $M_4$. Finally, a group of switches ($S1$ to $S4$) permits to reduce the settling time between consecutive readouts by resetting the large parasitic capacitor at the $out_{line}$ node. Similarly to the autozeroing loop, a dynamic bias control is implemented also for this module. More in detail, during the reset phase (signal $rst$ high-switches $S1$, $S4$ closed and $S2$, $S3$ open) only transistor $M_6$ is turned ON while the rest of the circuit is disabled (included the feedback amplifier $A_{FBCB}$) to save power consumption (compare timing diagram at the bottom of Fig. 2).

### C. Programmable Gain Amplifier

The amplification of the signal chain is completed by a programmable gain amplifier whose schematic is depicted in panel D of Fig. 2. It implements the well-known switched capacitor architecture described in [20]. Two phases are needed to produce the amplified output (see timing diagram at the bottom of Fig. 2). During the first phase, $\Phi_{SAMPLE}$ high, the input voltage is sampled on capacitor $C$ while the amplifier is connected in a unitary gain configuration. During the subsequent phase, $\Phi_{HOLD}$ high, the amplified output is produced, with the gain given by the ratio $C/C_B$. In our implementation $C$ is designed as a 5-bit programmable metal-insulator-metal (MIM) capacitor array to obtain a programmable gain from 0 to 30 dB.

### D. Readout Strategy

The readout and the autozeroing time division multiplexing are managed using two different circular shift registers that are integrated into each 32 electrode-pixels analog frontend and implemented by cascading multiple D-type flip-flops. For both shift registers, a global reset (not shown in panel E and F of Fig. 2) sets the output of the first flip-flop in the chain ($Q[1]$) to logic high and the remaining $Q[i]$ ($i = 2;32$) to logic low. After such reset phase, the access of any sequence of electrode-pixels in the module is managed by two control signals, namely $clk_{read}$ and $clk_{AZ}$, that respectively permit to scan the autozeroing and the readout shift registers.

### E. Digital Control Unit

This unit is used for bidirectional serial communication with the base unit and for generation of the timing signals required for correct operation of the device. Three distinct operating phases are available: configuration, validation, and recording. During the configuration phase, a bank of six system registers is loaded with user-defined parameters that are used to set the operating conditions of the device (i.e. gain factor, sampling frequency, the desired readout sequence within each 32 electrode-pixels module and the autozeroing frequency). During the subsequent validation phase, the content of the registers is sent toward the FPGA-based acquisition module for verification purpose only. Finally, the recording phase enables data acquisition with the desired sampling frequency and readout sequence. Two finite state machine (FSM) modules were designed for the purpose. The first manages the readout sequence while synchronization with the back-end acquisition module is kept through the use of signal $ADC_{en}$ which is provided by the FPGA to control the off-chip analog to digital converters (c.f. timing diagram of Fig. 2). The second periodically performs the autozeroing procedure to guarantee correct amplification of neural data. As from post layout circuit simulations, the pixel readout and further amplification can be performed within about 1.2 $\mu$s, which leads to a maximum sample frequency of about 25 kHz per electrode-pixel in full readout mode. The sampling rate can be proportionally increased if only a randomly selected subset of the 32 electrode-pixels module is accessed.

### F. CMOS Post-Processing

The CMOS devices received from the foundry (dimensions $Length \times Width \times Thickness = 7$ mm $\times 1$ mm $\times 250$ $\mu$m) are...
post-processed at the IIT clean-room using photolithography and dry micro-machining techniques to permit implantation. As depicted in Fig. 3, the workflow consists of the following steps. To preserve the aluminium (Al) electrodes, a thin (50 nm) layer of SiO$_2$ is first deposited on the top side by Atomic Layer Deposition (ALD), while 350 nm of Chromium (Cr) is sputtered on the top of this layer to obtain an integrated protective mask for the successive etching steps used for shaping the probe (step 1). After photolithography (step 2), the Cr-unmasked regions are etched with an ICP-RIE system (Sentech), first with a SiO$_2$/Si$_3$N$_4$ etching recipe for approximately one hour (step 3), thus removing completely the SiO$_2$ (step 4), and then with a standard Si etching Bosch process (step 5). Afterwards, the dies are flipped topside down and Si is uniformly etched from the backside for approximately 200 $\mu$m, until reaching a final thickness of the shaft of 50 $\mu$m (step 6). Finally, the devices are cleaned by removing the Cr and SiO$_2$ protective layers (step 7). The resulting structured CMOS neural probes are mounted using epoxy glue and wire-bonded on printed circuit boards (PCB). After chip-mounting, the native CMOS-metal of the 512 electrode pads (flat Al) are modified using electrodeposition techniques. The aim of this post-processing step is to lower the electrodes electrochemical impedance, improve their chemical stability, and to achieve a protruding electrode morphology. This is performed by electroplating a first layer of Platinum (Pt), followed by the deposition of PEDOT:PSS. The Pt layer ensures a good adhesion with the native Al and provides a rough texture at the microscale. The PEDOT:PSS was used to obtain a protruding semi-spherical shape for each electrode, to achieve a nano-scale roughness and to ensure a good wettability of the electrodes. To do so, the native Al-oxide was first removed using an Al$_2$O$_3$ wet etchant and successively Pt and than PEDOT:PSS were electrodeposited [21]. Both Pt and PEDOT depositions were carried out using Cyclic Voltammetry (CV) electrochemical processes, using a potentiostat/galvanostat (Parstat 2273, Princeton Applied Research, USA).

IV. THE 36-CHANNELS CMOS HEADSTAGE

The 36-channels headstage exploits a CMOS integrated circuit that was designed and fabricated in a standard 0.35 $\mu$m CMOS technology. The schematic block diagram of the realized circuit is depicted in panel A of Fig. 4 while a view of the chip is shown in panel B. The circuit integrates 36 low-noise amplifiers that are read in a time division multiplexed fashion through a programmable gain amplifier before on-chip analog to digital conversion. The timing signals required for correct operation of the device as well as for bidirectional communication with the base unit are provided by the integrated Digital Control Unit.

In the following sections, the circuit implementation for only the low-noise amplifier and analog to digital converter (ADC) are detailed. The programmable gain amplifier and the DCU were designed similarly to the previously described circuits integrated on the CMOS-probe. Specifically, for the DCU, a bank of integrated system registers is available to permit configuration of the CMOS frontend with user defined parameters with respect to both low- and high-pass cut-off frequencies of the integrated amplifiers (see Section IV-A), programmable gain (in the range 1:32 $\times$ ) and sampling frequency (see Section IV-B).

A. Low-Noise Amplifier

Differently from the CMOS neural probe, and due to the lack of constraints on the circuit size, an AC-coupled amplification stage was preferred for the 36-channels CMOS headstage. A more classical circuit architecture was implemented (similar to the one presented in [22]) permitting to optimize the design with respect to the noise. The circuit schematic, which was derived from the solutions previously presented in [23], [24], is shown in Fig. 5. The first stage is an AC-coupled high-pass filter whose cut-off frequency (below 1 Hz) permits to reject the DC offset and the voltage drift of the electrodes. A high-pass filter is inserted after the first stage to remove the residual offset while a second gain stage consisting of a single-ended capacitive-coupled voltage amplifier provides low-pass filtering and further amplification for an overall gain of 60 dB. Both the high- and low-pass cut-off frequencies are programmable. More in detail, the high-pass cut-off frequency can be moved up to 300 Hz by tuning the analog bias voltage, and thus the $g_{m}$, of the $g_{m}$-C stage, while the low-pass cut-off frequency (eight different values in the range 2–10 kHz) is tuned by changing the...
bias current of the last amplification stage. Each LNA occupies an area of 0.07 mm² and consumes 30 μW of power.

B. Integrated Analog to Digital Converter

A two-step recycling technique is applied to implement a 10-bit, 1Msample/s CMOS analog-to-digital converter (ADC). The ADC occupies an area of 0.2 mm² and consumes 5 nJ per conversion. For the classical architecture depicted in panel A of Fig. 6, M = N/2 bits are produced each time a new cycle is launched, starting from the MSB until the LSB, with N = 10 being the resolution of the converter. As discussed in [25], the sensitivity against circuit non ideals, which may reduce the actual resolution, can be mitigated through the redundancy resulting from the introduction of extra decision levels. The double threshold concept underlying the RSD (Redundant Sign Digit) conversion algorithm was first introduced in [26] for a cyclic converter where a single bit is produced per conversion-cycle. The algorithm, widely used in the design of pipeline ADCs (see for instance [27]–[29]) was extended in this work to design the 10-bit, two-step converter integrated in the 36-channels CMOS headstage. Such algorithm is based on the Sweny-Robertson-Tocher division principle [30], in which the analog input is compared simultaneously with two distinct thresholds (Tₚ and Tₒ) and a single RSD digit is produced according to the scheme summarized in Table I.

1) Parallel ADC: The circuit architecture of the parallel A/D converter implementing such approach is depicted in panel B of Fig. 6, where FSR defines the full scale range of the converter which, in the current implementation, is between +1.65 V and −1.65 V. Note that, since the single-bit solution of the reference can be implemented with a \(\{1.5R - R - 1.5R\}\) divider, in the M-bit implementation the idea is generalized using the \(\{1.5R - R - \cdots - R - 1.5R\}\) divider to produce the references. As far as the comparators are concerned, they are typically implemented by means of the Track&Latch stage [31], whose off-set is fixed by the ratio W/L of the input differential pair of the preamplifier stage: the larger the aspect ratio the smaller the off-set. However, when used in flash converters, the large input capacitance associated with the high number of parallel comparators required to simultaneously compare the input voltage with the references significantly impacts on the power consumption of the amplifier stage used to achieve the double gain factor. To avoid the use of large devices, and hence all the related issues, AC coupling was preferred to DC coupling (see inset in panel B of Fig. 6). Two phases are needed to perform comparisons: during the first phase (Fₚ and Fₚ high) the off-set of the input differential pair is sampled and held to be cancelled during the comparison that takes place in the second phase (F₂ high). The insensitivity against the off-set of the preamplifier stage allows to use smaller devices without affecting the output response of the converter, with a significant saving of area and power of the amplifier stage.

2) MDAC: For the general case of multistep converters, the D to A conversion, the computation, and amplification of the residues \(R_i\) resulting from subsequent conversion cycles can be obtained by means of a single module named MDAC. In panel C of Fig. 6 is depicted the switched capacitor circuit implementation for this module adopted in our circuit. Three cycles are needed to produce the whole conversion: the sample phase (F₁), the hold phase (F₂) and the amplification phase (F₃). The word produced by the flash during the hold phase (the so called coarse conversion) is used to drive the MDAC during Fig. 6. Panel A: standard architecture of a two-step analog to digital converter; panel B: circuit implementation of the parallel A/D converter and, inset, AC-coupled comparator used to implement the RSD conversion algorithm; panel C: Switched capacitor circuit solution for the MDAC implementing the parallel D/A conversion and the gain amplification (factor \(2^q\)) in the two-step architecture.

---

**Table I**

<table>
<thead>
<tr>
<th>RSD Conversion Algorithm</th>
<th>(b_p)</th>
<th>(b_q)</th>
<th>(b_{RSD})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(T_p &lt; V_{in} &lt; T_p)</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>(T_p \geq V_{in})</td>
<td>0</td>
<td>0</td>
<td>−1</td>
</tr>
<tr>
<td>(V_{in} \geq T_p)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>not allowed</td>
</tr>
</tbody>
</table>
the amplification phase. More in detail, during such phase the top plate of the capacitors is connected to 0, FSR/2 or -FSR/2 voltage rails depending on the word produce during the coarse conversion cycle, according to the scheme shown in panel C of Fig. 6 (inset).

Observe that the duration of each one of the previously described phases is defined by the user during the configuration phase and all the timing signals required for the correct operation of the integrated ADC are generated and provided by the integrated DCU.

V. FPGA-BASED ACQUISITION MODULE

The FPGA-based acquisition module permits complete control over the system for synchronous recordings from both our high-density and 36-channels CMOS devices. For the purpose, the module implements a bidirectional serial communication with the CMOS ICs, provides the clock signals required for their operation (20 MHz and 100 MHz respectively for the CMOS neural probe and for the 36-channels CMOS headstage) and performs a high rate data transfer to a PC through a high-speed cameralink interface, as adopted in [18].

Before each experimental session, the recording frontends are configured with user defined parameters (c.f. configuration and validation phases described in Section III-E for the case of the CMOS neural probe, which also applies to the 36-channels CMOS headstage). If the validation phase is concluded correctly (i.e., data sent to the configuration registers match with data read back) the CMOS ICs move to the subsequent recording phase. In such phase, data from the CMOS neural probes and from the 36-channels CMOS headstage is read by the FPGA module and organized in frames of respectively 32 (the number of pixels in each analog frontend) and 36 channels before transmission to a PC. Fig. 7 shows how data readout is managed for the case of a synchronous acquisition from 4 distinct CMOS neural probes and one low-density CMOS headstage. More in detail, as previously described in Section III, the 512 recording sites integrated into each CMOS neural probe are organized in 16 analog frontends. These are used for amplification and time division multiplexed readout of 32 electrode-pixels on dedicated analog outputs. According to the timing diagram in Fig. 7), the digital conversion of the i-th electrode-pixels from each of such analog frontends is performed in parallel by a bank of 16 ADCs and enabled by the falling edge of signal $ADC_{enPi}$ (which is produced by the FPGA module) upon detection of the data ready signal $DreadyPi$ (which is produce by the CMOS neural probe). For each probe, after 32 conversion cycles, a flag $EOF_{Pi}$ (End Of Frame) is asserted. Similarly, the FPGA module also acquires data from the 36-channels CMOS headstage, but this time the A/D conversion in performed on-chip and the signal $EOF IC36$ is produced once that the FPGA module has acquired 36 data samples. The so collected data are temporarily saved and sent to a PC via a cameralink protocol as soon as all the ICs have completed the acquisition of one frame. At that point the FPGA module enables the acquisition of a new frame (signal $New Frame$) by releasing all the $EOF$ flags and the procedure is repeated, making data acquisition synchronous and robust against time slacks among the operation of the different recording devices.

| TABLE II | SUMMARY OF THE METRIC PERFORMANCES FOR THE CMOS NEURAL PROBE AND THE 36-CHANNEL CMOS HEADSTAGE |
|-----------------|-----------------|-----------------|
| Recording headstage | CMOS neural probe | 36-channel |
| Number of channels | 512 | 36 |
| high-pass filter [Hz] | 0 | <1–300 |
| low-pass filter [kHz] | 4 | 2–10 |
| Gain [dB] | 40 | 60 |
| Area/channel [mm$^2$] | $625 \times 10^{-6}$ | $70 \times 10^{-3}$ |
| Power/channel [$\mu$W] | 2.7 | 30 |
| Noise (AP band) [$\mu$V RMS] | 20.6 | 6.3 |
VI. BENCH TEST CHARACTERIZATION

The gain and noise performances of the 36-channels CMOS headstage and of the CMOS-probes, are summarized in Table II together with other metric performances. These were separately characterized in a phosphate buffered saline (PBS) solution (see panel A and B of Fig. 8). For the CMOS neural probe, electrical tests were performed by dipping the devices in beaker containing a PBS solution as illustrated in the top side of Panel B in Fig. 8. For testing the 36-channels CMOS headstage, the device was connected to a custom passive planar multi electrode array (MEA) realized at the IIT clean room following a micro-fabrication process similar to the one described in [32]. Successively the planar MEA was wire bonded to a printed circuit board (PCB) and a glass ring acting as a reservoir was placed around the active area before sealing it with epoxy glue, as depicted in the top side of panel B in Fig. 8.

For both test conditions the gain frequency response was evaluated by injecting pure sine waves of different frequencies and with constant amplitudes (corresponding to configuration 1 in the diagrams of Fig. 8). Results are reported on the bottom side of panel A and B of Fig. 8. For the CMOS neural probe, in-line with our design specifications, a gain of 40 dB below the 4 kHz cut off frequency was measured for the in-pixel circuit (panel A). It can be noted that, although the maximum sampling frequency for the CMOS neural probe is of about 25 kHz in full readout mode (see Section III-E for details), in our experiments data sampling was limited to 16 kHz, due to limitations imposed by the data acquisition software (Brainwave, 3Brain AG, [11]). For the 36-channels CMOS headstage a gain of about 74 dB (obtained from a programmable gain of about 14 dB) was measured for different programmed low-pass circuit configurations (panel B). To evaluate the input referred noise of the recording systems, the PBS solution was finally biased to the reference potential and the data recorded in such configuration was divided by the overall gain provided by the signal amplification chain (corresponding to configuration 2 in Fig. 8). The spectral power density was computed and used to measure the input referred noise of the recording systems. For the CMOS neural probe, it accounts for about 8.4 $\mu V_{RMS}$ and 20.6 $\mu V_{RMS}$ when integrated respectively over the 300 Hz-5 kHz and 1 Hz-5 kHz frequency bands. Finally, an input sine wave (1 kHz frequency, 400 $\mu V$ amplitude) was recorded from a CMOS neural probe (dark line, 46 dB gain) and from the 36 channels CMOS headstage (gray line, 66 dB gain) (panel C).

For the CMOS neural probes, the electrochemical impedance of the on-chip electrodes and the mechanical performances were also evaluated. For mechanical tests, an experimental setup consisting in an adapted Dektak Stylus profiler was used. By means of the 5 $\mu m$ radius stylus, forces spanning from 1 mg to 5 mg were applied to the backside of structured probes and the corresponding deflection was measured for each point while scanning on the backside surface of the probe along its 7 mm length (resolution of 0.33 $\mu m$/sample). Such deflection curve (i.e. vertical displacement in the direction parallel to the applied force and orthogonal to probe axis) was finally used to calculate a Young modulus of 50 GPa.

The electrochemical impedance of the electrodes was measured in a PBS solution for the native Al, and post-processed Pt and PEDOT:PSS electrodes, at ten frequencies per decade over
5–10⁶ Hz. The measured value of the impedance modules at 1.1 kHz of frequency account respectively for about 40 MΩ (Al), 500 kΩ (Pt) and 200 kΩ (PEDOT:PSS) (see panel C of Fig. 9).

VII. NEURAL RECORDINGS

The recording capability of the system was finally assessed through ex vivo and in vitro experiments separately for the two frontends. All work was done in accordance with the Italian guidelines and regulations and all animal procedures carried out in this work were approved by the institutional IIT Ethic Committee and by the Italian Ministry of Health and Animal Care (Authorization No. 110/2014-PR of the 19th of December 2014). For the purpose, the 36-channels CMOS headstage was validated by connecting it to a planar MEA using the same experimental setup described in Section VI). For the high-density headstage, a polydimethylsiloxane (PDMS) culture chamber acting as a reservoir to contain the culture media was created around a CMOS neural probe that was previously mounted on a PCB. The culture chambers were first sterilized by filling them with 70% EtOH and then rinsed after 20 minutes 4 × with sterile Double Distilled Water (DDW). Finally the devices were dried under a sterile laminar-flow-hood.

As shown in Fig. 10, both systems could record spiking activity from both ex vivo retinas as well as from neuronal cultures grown on the devices for 16 days-in-vitro. For ex vivo recordings, retinas from adult mice were dissected and isolated after sacrificing the anaesthetized animal by cervical dislocation. By removing and placing retinas with the Ganglion Cell Layer (GCL) facing down onto the sensing area of the devices we recorded stable spontaneous as well as light-induced activity (see panel A of Fig. 10). For neuronal cell cultures, devices and primary neurons were prepared following protocols that we have previously described in [33]. Briefly, the CMOS neural probe and the planar MEA were pre-conditioned, i.e. incubated overnight at 37 °C and 5% CO₂. The day after, the devices were rinsed 4 × with sterile DDW and were left drying under the hood before cell seeding. Primary hippocampal neurons were obtained from brain tissues of mouse embryos at day 18 (E18). Cells were seeded at the final density of 1000 cell/mm² on the coated substrates and then incubated at 37 °C with 5% CO₂ and 95% humidity. After 1.5 hour, 3 ml of Complete Neurobasal Medium (CNM), containing 2% B-27, 1% penicillin/streptomycin and 1% GlutaMax supplements were added to the culture chamber and incubated at the same conditions. For maintenance and cell culture growth, one-third of the medium was routinely replaced with a new CNM every four days. All reagents were obtained, unless indicated differently, from Life Technologies. At 16 DIVs, the spontaneous spiking activity was recorded for 5 minutes (see a sample in panel B of Fig. 10).

It can be noted that for neural recordings, both the electrodes impedance and noise values are highly affected by the neuron-electrode interface. In our experiments, tissue and culture were both prepared following protocols that were previously optimized for CMOS MEAs. For this reason data in Fig. 10 show a better measuring performances for the CMOS neural probe despite their larger input referred noise (see data presented in Section 8).

VIII. CONCLUSION

Here we presented a platform designed for simultaneous and synchronous sub-millisecond resolution recordings of neural signals by combining the use of active and passive multielectrode arrays (MEAs) with custom designed CMOS ICs. This opens the opportunity to perform recordings of bioelectrical activity in the brain at multiple spatial scales by exploiting the capabilities of different electrode-based probes. Large-scale neuronal interfacing throughout the depths of the brain can be performed by means of our implantable monolithic CMOS neural probe while concomitant and synchronous recordings exploiting the wide range of available conventional passive devices can be achieved with the custom designed digital
output CMOS headstage integrating 36 low-noise amplifiers. It is worth to notice that nowadays complete neural acquisition systems integrating few tens of bioamplifiers are commercially available (see for instance [34]). However, since full control of their operation is not permitted, such devices are not suitable for synchronous operation with our CMOS neural probes.

The FPGA based acquisition board can manage synchronous recordings (sampling rate up to 25 kHz) from up to 4 different CMOS neural probes and one 36-channels CMOS headstage. Each of such CMOS neural probes integrates multiple 32 electrode-pixels modules to achieve (in the current generation) 512 recording channels in a 180 nm CMOS technology. These devices, when delivered from the foundry, occupy a volume of 7 mm × 1 mm × 250 μm and post-processing is required both for shaping and thinning as well as for modifying the electrodes material. This allowed to achieve final dimensions comparable with those of classic silicon probes: thickness and width were reduced along the shaft respectively down to ≤ 50 μm and 100 μm. To guarantee electrochemical stability of the integrated electrodes, the native aluminium metal was electroplated with platinum, followed by PEDOT:PSS. The process also permitted to reduce the electrochemical impedance of the small (14 μm × 14 μm) integrated electrodes of about 2 orders of magnitude.

To assess the recording capabilities of this platform prior to in vivo experiments, here we reported results obtained by electrical testing of the two systems in PBS as well as recordings of neural spiking activity from ex vivo retinas and in vitro neuronal networks. For the case of the CMOS neural probe, the stable recording for several hours from the retina indicates the absence of thermal effects. Further, recordings from 16-days old neural cultures are indicative of the adequate level of bio-compatibility of our devices. Also, synchronous acquisition from one CMOS neural probe and from the 36-channels CMOS headstage was demonstrated in bench test conditions.

ACKNOWLEDGMENT

The authors would like to thank Giacomo Pruzzo, senior technician at the Neuroscience and Brain Technology department, for his help in preparing the bench test setup.

REFERENCES


Hayder Amin received the M.Eng. degree in biomedical engineering from Martin-Luther University, Halle, Germany, in 2011, and the Ph.D. degree in microtechnology for neuromorphic and neuroscience from the Italian Institute of Technology, Genova, Italy, in 2015. He employs competencies in Engineering, Biophysics, and Biology to bridge technology and science. He has been designing, developing, and implementing bio-assays based on the development of microelectronics, micro-/nano-structuring, and biotechnologies. His research interests include enabling high information content on-CMOS-microelectrode arrays and novel optical imaging to improve multiscale monitoring in large-scale neuronal networks and brain circuits, hence serving to further develop effective tools for neurodegenerative disease-specific therapies and drug development.

Gian Nicola Angotzi received the M.Sc. degree in electronic engineering and the Ph.D. degree in electronic and computer science from the University of Cagliari, Cagliari, Italy, in 2003 and 2007, respectively. He is currently a Research Technologist with the Istituto Italiano di Tecnologia, Genova, Italy. He has a strong background on the conception, design, realization, and validation of very large-scale integration circuits. His research interest focuses on the design of electronic front ends for neural recording systems.

Hayder Amin is a Physicist by education. He is currently working with the Institute for Fundamental Electronics, University of Paris-Sud, Orsay, France. His research interests include the development of novel nanofabrication methods in the fields both of nano-optics and of technology for biology and neuroscience.

Alessandro Maccione received the Master’s degree in robotics from the University of Genova, Genova, Italy, in 2005, and then a Researcher with the National University of Singapore, Singapore, in 2007. He has a strong background on the conception, design, realization, and validation of very large-scale integration circuits. His research interest focuses on the design of electronic front ends for neural recording systems.

Alessandro Maccione received the Master’s degree in biomedical engineering and the Ph.D. degree in biophotonics from the University of Genova, Genova, Italy, in 2015. He employs competencies in Engineering, Biophysics, and Biology to bridge technology and science. He has been designing, developing, and implementing bio-assays based on the development of microelectronics, micro-/nano-structuring, and biotechnologies. His research interests include enabling high information content on-CMOS-microelectrode arrays and novel optical imaging to improve multiscale monitoring in large-scale neuronal networks and brain circuits, hence serving to further develop effective tools for neurodegenerative disease-specific therapies and drug development.

Mario Malerba is a Physicist by education. He is currently working with the Institute for Fundamental Electronics, University of Paris-Sud, Orsay, France. His research interests include the development of novel nanofabrication methods in the fields both of nano-optics and of technology for biology and neuroscience.

Hayder Amin is a Physicist by education. He is currently working with the Institute for Fundamental Electronics, University of Paris-Sud, Orsay, France. His research interests include the development of novel nanofabrication methods in the fields both of nano-optics and of technology for biology and neuroscience.

Fabio Boi received the M.Sc. degree in robotics engineering and the Ph.D. degree in robotics, cognition and interaction technologies from the University of Genoa, Genoa, Italy, in 2012 and 2016, respectively. He is currently a Post-Doc with the Istituto Italiano di Tecnologia, Genova, Italy. He has a strong background on programming languages, electrophysiological signals, and surgical procedures. His research interests include neural recordings and data analysis of signals recorded from the central nervous system.

Marco Crepaldi received the Engineering degree (summa cum laude) and the Ph.D. degree in electronic engineering from the Politecnico di Torino (Polito), Turin, Italy, in 2005 and 2009, respectively. During 2008, he was a Visiting Scholar with the Electrical Engineering Department, Columbia University, New York, NY, USA. After the Ph.D. degree, he worked as a Post-Doc with the VLSI-Lab, Electrical Engineering Department, Polito. He is currently the Coordinator of the Electronics Design Lab, Istituto Italiano di Tecnologia, Genova, Italy. His scientific activity is related to the analysis, simulation, and development of integrated event-driven and all-digital impulse-radio ultrawide band systems.

Luca Berdondini received the M.Sc. degree in microengineering from the Swiss Federal Institute of Technology of Lausanne (EPFL), Lausanne, Switzerland, in 1999, with a Master thesis at Caltech, Pasadena, CA, USA, and the Ph.D. degree in nano-/micro-fabricated interfaces and CMOS devices for electrophysiology from Samlab, EPFL, in 2003. He is currently a Senior Researcher leading the IIT-NetS3 Laboratory, Department of Neuroscience and Brain Technologies, Istituto Italiano di Tecnologia, Genova, Italy. He is among the pioneers of CMOS-based multielectrode arrays for electrophysiology, and a Co-Founder of 3Brain GmbH, Landquart, Switzerland. His research team focuses on the development and application of innovative neuroelectronic solutions to measure and analyze electrical activity at cellular resolution in large neuronal assemblies and brain circuits.

Ermanno Miele received the M.Sc. degree in biomedical engineering from Rome Tor Vergata University, Rome, Italy, and the Ph.D. degree from the University of Genoa, Genoa, Italy, in 2014. He has been enrolled as a Research Fellow with the Istituto Italiano di Tecnologia, Genova, Italy (2014 and 2017), a Visiting Scientist with Case Western Reserve University, Cleveland, OH, USA (2015), and a Research Fellow with the National University of Singapore, Singapore (2016). His research interests include advanced fabrication techniques, surface chemistry and functionalization, and templated assembly at the nanoscale.