A 79μW 0.24mm^2 8-channel Neural Signal Recording Front-End Integrated Circuit

Sami Ur Rehman[1], Awais M.Kamboh[1], Yuning Yang[2]
[1]Department of Electrical Engineering, NUST School of Electrical Engineering and Computer Sciences, Islamabad, Pakistan.
{sami.rehman, awais.kamboh}@seecs.edu.pk
[2] Department of Electrical Engineering and Computer Engineering, Michigan State University, East Lansing, MI, USA

Abstract— **This brief presents a new architecture for an ultra-low power and area-efficient 8-channel prototype of a neural signal recording front-end circuitry. For implantable neural recording circuits, low power and low area are two of the most critical requirements. In contrast to architectures existing in the literature, the presented recording path is centered on a single high-performance programmable gain-bandwidth amplifier, instead of employing a separate stand-alone amplifier for each electrode. The resulting circuitry requires smaller area and less power compared to all previously published designs. Implemented in 0.5μm CMOS and a supply voltage of 1.8V, the 8-channel recording path is measured to consume a total of 79μW of power and a net area of 0.24mm^2. Therefore, allowing suitability of our design to be used in high channel count environments.**

***Keywords*—low power, neural recording, prosthesis, SAR ADC, neural amplifier, SNR.**

I. INTRODUCTION

Miniaturization at the level of integrated circuits and micro-electro-mechanical systems has enabled neuroscientists and prosthesis engineers to measure neural electric signals simultaneously at multiple locations on the surface of the brain. Traditional methods of recording neural activity using tethered electrodes present several challenges. Among them, the most notable are skin infection, and constricted patient mobility. These limitations restrict the use of such tethered techniques only to clinical and investigational experiments. To alleviate these problems wireless data telemetry and power has recently gained popularity as a viable alternate solution.

Electrical signal from a single brain neuron, called an action potential or a neural spike, is typically10μV to 50 μV in amplitude and occupy 300Hz to 7KHz spectrum. Signal from multiple neurons superimposed over each other, called Local Field Potential (LFP), can be as large as 1mV to 10mV and occupy a frequency band from 25mHz to 100Hz [1]. Both these signals are very weak, but contain important information about individual as well as collective neural activity. Recording them accurately is critical for any neural decoding algorithm to work. On the other hand, prosthetic devices that enable external stimulation of neural cells can help initiate inter-neuron communication, or inhibit certain spiking activity, can be used to help patients with certain neurological or psychological disorders. Wireless neural systems require design of silicon chips for neural recording and stimulation, which have a very small area and very low power consumption, making them safe to be implanted on the cortex. Design of such implants presents many challenges, in not only the back-end wireless telemetry, but also the front-end signal recording and stimulating circuitry. Noise, area, and power of the neural front-end are some of the most constricting requirements placed on such biomedical implants. A low input referred noise in the front-end circuitry is required because the neural signals are only a few tens of μVs in amplitude and can easily be distorted or completely drown in the physical noise of the front-end amplifiers. Such implants are placed on the brain surface beneath the skull and require a burr hole to be made in the skull. This severely restricts the area footprint of such modules due to medical reasons. Additionally, in-vivo neural recording is extremely safety critical because a mere 1ºC rise in temperature of such implants with respect to the surrounding tissue can severely damage the neuron. This requirement restricts the net power consumption density of 0.58mW/mm2 by the each implant [2]. However, existing high channel count neural recording architectures suffer from excessive power and area overhead, and, therefore, are not yet suitable for large scale commercial production. In this paper we present an 8-channel prototype of a new neural recording circuit architecture which being power-area efficient can support high channel count.

II. STATE OF THE ART AND THE PROPOSED SOLUTION

Design of low area-power neural front-ends with high channel count has gained significant popularity over the last decade and several low noise and low area-power designs have been proposed in literature. Harrison's architecture [1] has served as a benchmark and starting point for neural signal amplification and filtering. A 128-channel neural stimulation and recording interface has been discussed in [3] consuming moderate area and power, and a more robust 32-channel digitally programmed front-end is presented in [4]. A more recent work reports a channel count of 1024, however, the total power dissipation in recording path stands at astounding 75mW [5]. Looking more closely, we realize that in all multi-channel neural systems, like

ISBN 978-80-261-0642-5, © University of West Bohemia, 2017

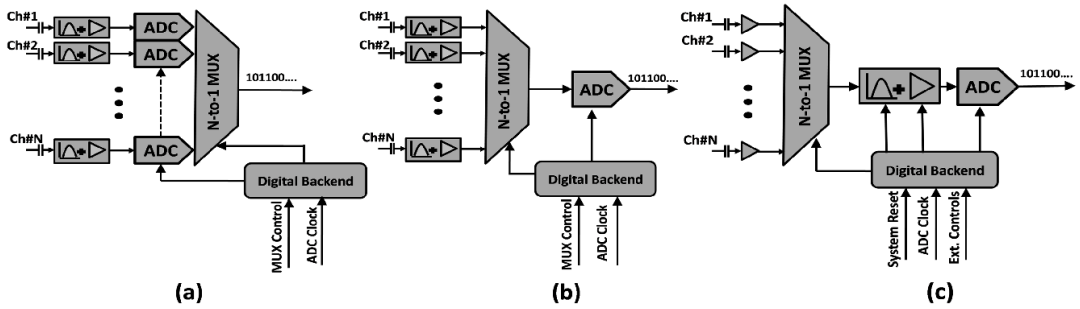

Fig. 1 a)Traditional neural recording architecture b) Improved neural recording architecture as reported in [4] and c) Proposed neural recording front-end

the ones referred above, each channel has its own high performance amplifier, its own filter and in some cases its own ADC in the recording path as shown in Fig 1a. Such designs become impractical in terms of power and area consumption when scaled to high channel count. To cater for this problem a single ADC combined with an analog mux was used in [4], to convert several input analog signals into a single digital output stream as shown in Fig 1b. Still, this design requires an exclusive filter-amplifier for each channel. Existing neural recording designs, therefore, report dedicated high performance neural amplifiers for each electrode. Such designs are not suitable for high channel count environments because the collective power consumption of amplifiers exceed the allotted power budget of around 0.58mW/mm2 as explained before. We propose, and demonstrate, that if somehow one super-performing amplifier could be designed that can serially record data from multiple channels replacing these amplifiers, the overall power and area consumption for a multi-channel system could be significantly curtailed. This work, therefore, proposes a design which employs power-area efficient low-key pre-amplifiers with each electrode, multiplexing the data from pre-amplifiers and incorporating one super-performing, gain-bandwidth programmable filter-amplifier to record both spikes and LFPs. The proposed architecture is shown in Fig 1c.One may argue that in the proposed design, with multiplexing shifted close to the recording electrodes, the acquired signal integrity and its resolution can be severely compromised since the entire recording path is dedicated to one channel only for τ/N amount of time (with τ being the total scan time of all channels and N being the total number of channels). We established the hypothesis of our proposed design on the fact that the neural information is embedded in the timing of the neural signals and not their amplitudes [6]. Therefore, our idea is not to capture a fine grained digital version of the neural signal but its frequency information or its spiking activity. We argue that if we know the maximum spike rate of a neural signal and if we can reliably multiplex the data from multiple channels close to the electrodes then employing dedicated amplifier for each electrode or pushing the multiplexing away from electrodes is a waste of power/area. In this context the channel scan time must be properly maintained so as not to miss a neural occurrence. This scan time consequently places settling time constraints on the filter-amplifier and the ADC. Since the highest neural signal frequency lies around 8KHz, we clocked the 8-to-MUX with 20KHz, scanning each electrode for 6.25µs after every 50µs.This restricts the settling time of filter amplifier and SAR estimation routine time of ADC to be less than 6.25µs. This paper develops on the theoretical concept proposed in [7, 8], and presents the measurement results on signal recorded from a chip fabricated in 500nm CMOS process.

III. Circuit Level Implementation

The complete recording path at circuit level can be seen in Fig 2. The low-key differential amplifiers with active loads were designed to achieve a gain of 10-15 dBs. Special care was required to minimize the thermal and flicker noise of this stage as it directly interacts with the neural signals. From preamplifiers the signal is input to a transmission gate based multiplexer and then fed into an amplification stage with controllable gain and bandwidth. This stage uses pseudo-resistors in feedback, discussed first in [1], to achieve very low cut-off frequencies. The signal is then fed into a filter stage which provides one degree of freedom to make the low-pass cut-off frequency adaptable to the input signal. The filter-amplifier adjusts its bandwidth and gain to measure spikes or LFPs as determined by the control signals. The control signals were generated off-chip. The digital backend-not shown in the figure-requires external controls, a system reset and a system clock. We dedicated one of the eight channels to measure AP while others for LFP signals. Therefore, during runtime, whenever 8:1 mux connects the filter amplifier to the preamp connected to channel dedicated for measuring AP, the digital backend generates signals to adjust gain and bandwidth (Gain is increased for AP, while the lower and higher cutoff frequencies are pushed towards the band in which AP signals lie). Digital backend provides digital control signals to tune a) gain b) low-pass cutoff frequency and c) high-pass cutoff frequency. The filter-amplifier block in Fig 2 consists of a gain stage followed by band tunable filter stage. The capacitive feedback sets the mid-band closed loop gain ($A_{CL}=C_{IN}/C_F$), the low cut-off frequency can be tuned by V_{TUNE}, and the high-pass cut-off frequency, given by $f_H=g_m/2\pi A_{CL}C_C$, depends on Cc, apart from OTA trans-conductance and its gain. Fig 3 contains the

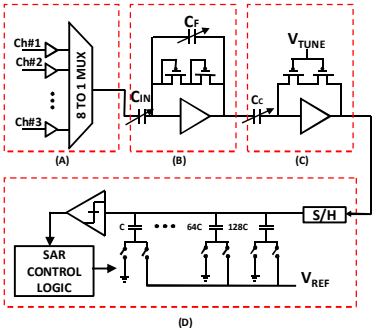

Fig 2 Single ended version of complete neural recording path. A: Low key Pre-amplifiers with 8to1 Mux. B: Amplifier Stage. C: Filter Stage. D: 8-bit SAR ADC. E) Telescopic differential amplifier used in amplifier and filter stage (common mode feedback circuit not shown)

circuit level design of telescopic OTA used in the design. PMOS differential pair, comprising of M1 and M2 and operating in sub-threshold region, was chosen because of their low immunity to flicker noise compared to NMOS devices. M3-M4 act as active load and serve to enhance the output impedance and hence the open loop gain of the OTA which is approximated by $A_{OL}=g_{m1,2}\times r_{o5}||r_{o3}+r_{o1}(1+r_{o3}g_{m3})$. This A_{OL}, however, becomes desensitized in closed loop configuration. Transistors M5-M12 implement current steering network. Successive Approximation Register (SAR) ADCs have demonstrated best power efficiency among all ADC architectures and is considered an ideal candidate for sensor and biomedical applications [9]. Therefore, we opted for a fully differential monotonically switched successive approximation scheme. As discussed above, the maximum channel scan rate of 20 kHz leaves a

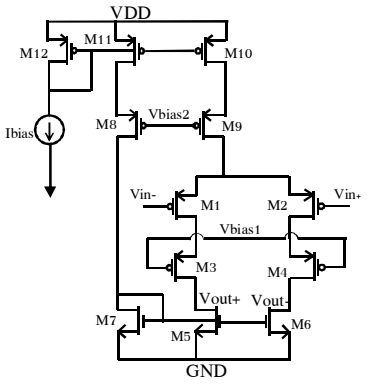

Fig 3. Telescopic differential amplifier used in amplifier and filter stage (common mode feedback circuit not shown)

window of only 6.25µs for amplification, filtering and digitization. We have chosen the bias currents and device sizes to achieve a settling time of around 2 µs for filter amplifier stage while the SAR ADC generates 8-bit digital code at a rate of 500 KHz.

IV. Measurement Results and Comparison

Micrograph of the designed microchip and the measurement setup can be seen in Fig 4. Measurement setup consists of DUT board, and a National Instruments (NI) IO card to serve clock generation and data acquisition interface. Very fine grained sinusoidal and ramp signals were generated from 16-bit NI card to acquire frequency response of filter amplifier and linearity plots of the ADC.

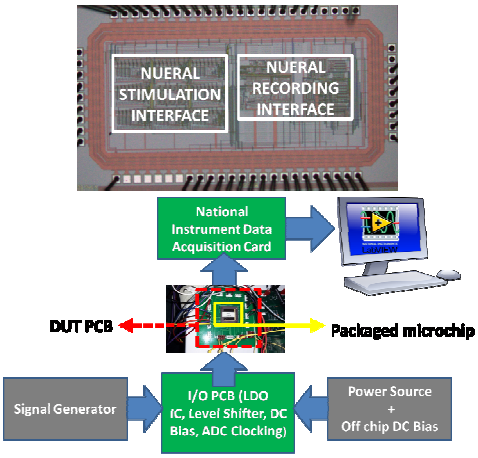

Fig. 4 Chip micrograph of the proposed architecture (top) and measurement setup (bottom)

Fig 5 contains frequency response of the filter amplifier measured independently from the ADC. Both the gain and bandwidth of the amplifier are tunable but in the plot shown the gain for neural spikes and LFP is kept at around 54 dB and 34 dB respectively. The dotted plot shows amplification setup for ultra low frequency LFPs. Its high pass 3dB cut-off was set at around 2.2 KHz. Similarly, the bold plot is set to measure neural spikes with its high pass cut-off at around 8 KHz and low pass cut-off around at around 1KHz. The coupling capacitor following the amplification stage, C_C, and the tunable voltage, V_{TUNE}, in filter stage were used to tweak high pass and low pass cut-off respectively. The unsmooth gain plot at maximum gain is likely a measurement artifact.

Fig 6 contains INL and DNL plots of the 8-bit SAR ADC. These plots were obtained after repeatedly applying a full scale ramp signal at the input of ADC and measuring its stair case response at the output. 100,000 data points were sampled and finally histogram method was used to construct the linearity plots. A DNL of greater than -1 and lesser than +1 manifests that the ADC is monotonic and has no wide codes, respectively. Right side of Fig 6 contains the frequency response of the ADC after applying a 750Hz of input signal. We observe the measured SNR of 45dB which translates to an Effective Number of Bits (ENOB) of around 7.2. The measured gain-bandwidth plots of the filter-amplifier and the linearity and dynamic plots of ADC clearly suggest that the recording path can very accurately capture the frequency information of the neural signals, which is the key argument our proposed concept is based upon.

Full measurement results of the recording path are reported in Fig 7. This measurement was obtained by disabling the functionality of the MUX and shorting pre-amplifier of channel 1 with the filter-amplifier and the ADC. Left hand side of Fig 7 shows

a time domain reconstructed output of the recording path when applied with 1KHz of input signal at the pre-amplifier and with ADC clocked at 200KHz.

The reconstruction of the analog signal was performed using an ideal DAC as a part of software estimation. The right side plot of Fig 7 contains the fft of the reconstructed signal.

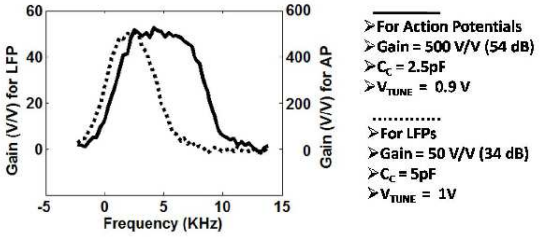

Fig. 5. Measured frequency response of the filter amplifier

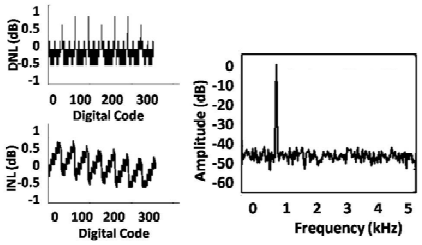

Fig. 6. Linearity plots (left) and the SNR plot (right) of the designed 8-bit ADC

Comparing the proposed architecture with the state of the art [3, 4] which also employ SAR ADC as digital interface, Table I shows our design is more power and area efficient. This power-area efficiency enables the proposed recording front end to be used in high channel count biomedical implants.

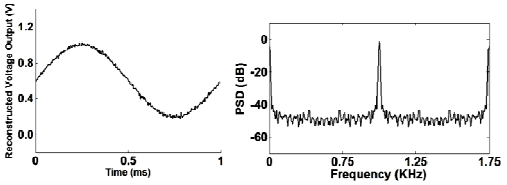

Fig. 7. Reconstructed analog signal at the output of ADC (left) and the FFT plot of the reconstructed signal (right).

TABLE 1. COMPARISON OF DESIGNED NEURAL RECORDING PATH WITH STATE OF THE ART

	[3]	[4]	This Work
Technology (µm)	0.35	0.18	0.5
Area (mm^2)/8 Channels	0.32	0.32	0.24
Power (µW)/8 Channel	102	82	79
Gain (dB)	73	49-66	45
Bandwidth (Hz)	10-5K	350-11K	100-7K
ADC Type	SAR	SAR	SAR
ADC ENOB	6.2	7.65	7.2
Input Referred Noise (µV)	6.08	5.4-11.2	5.8*

*simulated

CONCLUSIONS

A low power and low area multichannel neural recording circuit was presented in 0.5µm CMOS technology. Measurements show that this architecture saves power and area, has good scalability, and provides high fidelity signal at good sampling rate. We anticipate that by employing the reported neural recording architecture, high channel neural implants with medically tolerable power and form factor can be designed.

ACKNOWLEDGEMENT

The authors would like to thank Michigan State University for facilitating the fabrication of the chip through their MOSIS Education Program.

REFRENCES

1. R. R. Harrison *et al.*, "A Low-Power Integrated Circuit for a Wireless 100-Electrode Neural Recording System," in *IEEE Journal of Solid-State Circuits*, vol. 42, no. 1, pp. 123-133, Jan. 2007.
2. S. Kim, P. Tathireddy, R. A. Normann and F. Solzbacher, "Thermal Impact of an Active 3-D Microelectrode Array Implanted in the Brain," in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 4, pp. 493-501, Dec. 2007.
3. F. Shahrokhi, K. Abdelhalim, D. Serletis, P. L. Carlen and R. Genov, "The 128-Channel Fully Differential Digital Integrated Neural Recording and Stimulation Interface," in *IEEE Transactions on Biomedical Circuits and Systems*, vol. 4, no. 3, pp. 149-161, June 2010.
4. W. Wattanapanitch and R. Sarpeshkar, "A Low-Power 32-Channel Digitally Programmable Neural Recording Integrated Circuit," in *IEEE Transactions on Biomedical Circuits and Systems*, vol. 5, no. 6, pp. 592-602, Dec. 2011.
5. M. Ballini *et al.*, "A 1024-Channel CMOS Microelectrode Array With 26,400 Electrodes for Recording and Stimulation of Electrogenic Cells In Vitro," in *IEEE Journal of Solid-State Circuits*, vol. 49, no. 11, pp. 2705-2719, Nov. 2014.
6. Eric Kandel, James Schwartz, Thomas Jessell 'Principles of Neural Science' (McGraw-Hill 2000, fifth edition).
7. S.U.Rehman and A. M. Kamboh, "A new architecture for neural signal amplification in implantable brain machine interfaces," *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Osaka, 2013, pp. 2744-2747.
8. S.U. Rehman, A.M. Kamboh 'A CMOS Micro-power and Area Efficient Neural Recording and Stimulation Front-End for Biomedical Applications' *Circuits Systems and Signal Processing, Springer,* vol. 34, Iss. 6, pp. 1725-1746, 2015.
9. S. U. Rehman and A. Shabra, "Reference-less SAR ADC for on-chip thermal monitoring in CMOS," *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*, Lisbon, 2015, pp. 2441-2444.