35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

A New Architecture for Neural Signal Amplification in Implantable Brain Machine Interfaces

Sami ur Rehman, *Student Member, IEEE* and Awais M. Kamboh, *Member, IEEE*

Abstract— **This paper reports a new architecture for variable gain-bandwidth amplification of neural signals to be used in implantable multi-channel recording systems. The two most critical requirements in such a front-end circuit are low power consumption and chip area, especially as number of channels increases. The presented architecture employs a single super-performing amplifier, with tunable gain and bandwidth, combined with several low-key preamplifiers and multiplexors for multi-channel recordings. This is in contrast to using copies of high performing amplifier for each channel as is typically reported in earlier literature. The resulting circuits consume lower power and require smaller area as compared to existing designs. Designed in 0.5µmCMOS, the 8-channel prototype can simultaneously record Local Field Potentials and neural spikes, with an effective power consumption of 3.5 µW per channel and net core area of 0.407mm^2.**

I. Introduction

Neuroscientists and clinicians have grown increasingly interested in neural data recording in the past few years. The recorded neural data not only provides insights into the functioning of the brain; observing neural activity can also help monitor and diagnose neural disorders and can enable scientists to build smart neuro-prosthetic devices for amputated persons [1]. Non-invasive recordings of neural activity do not provide enough spatial and temporal resolution as required by the aforementioned applications. With the recent advances in wireless power and data transfer integrated circuit designs, it is now increasingly possible to use highly-invasive implantable modules, connected to high-channel count electrode arrays for brain machine interface (BMI) systems.

Neural signals comprise of two components, the action potentials, also known as neural spikes, and the local field potentials (LFP). Neural spikes have amplitudes ranging from 5µVpp to 50µVpp and lie in a band of 300Hz to 7.5KHz. LFPs have amplitudes typically ranging from 1mVpp to 10mVpp and lie in a band of 25mHz to 100Hz. Both these components provide important information about individual and collective neuron activity, respectively. Analog front-ends for neural signal acquisition typically consist of amplifiers, band-pass filters and analog to digital converters (ADCs). Wirelessly powered BMI systems only have limited power available. For a circuit to be implantable, it must have a small area footprint. Both these limitations become especially constricting when number of channels in the system is increased. Since the input signals are very weak, equally important is the requirement that the analog signal path, from the electrode to the ADC, must have a very low input referred noise. Input referred noise and power consumption have an inverse relationship and this trade-off is usually expressed as the noise efficiency factor (NEF) [2]. Similarly, reducing flicker noise demands enhanced device sizes and hence the overall area of the chip.

Neural signal amplifiers are considered the most critical components of the system, and thus there has been a lot of work in this domain. Harrison's [3] landmark architecture can record multiple signals, but is relatively power hungry. A gain and bandwidth tunable neural amplifier with improved NEF is described in [4]. A 128 channel neural stimulation and recording interface is discussed in [5] consuming moderate area and power. Sarpeshkar's [6] 32-channel neural interface is power and area efficient but slightly compromised in terms of NEF. In [7] authors have proposed a novel operational trans-conductance amplifier (OTA) sharing architecture to reduce the NEF. A technique to reduce the flicker noise and to enhance linearity of ultra low power neural amplifier is described in [8] and [9] respectively. In most multi-channel systems, each channel has its own high performance amplifier, its own filter and its own ADC[10]. In addition, many designs require a separate amplifier and filter for LFPs. Such designs become very large, in size and in power consumption, when number of channels is increased. To conserve area, some architectures use a single ADC combined with an analog mux [5, 6,7], as shown in Fig. 1, to convert several input analog signals into a single digital output stream.

This paper reports a new architecture for multi-channel neural recordings, efficient in terms of area, power, and noise, and hence best suited for systems with large number of electrodes. The new architecture differs from the architecture of Fig. 1, and thus a redesign of the tunable gain-bandwidth neural amplifiers and filters is presented, with additional constraints on settling time and signal propagation delay. The post-layout results of the 8-channel prototype are presented. Finally, a comparison to existing designs illustrates the area-power gains made by the new design.

II. Theoretical Framework and Design

Our hypothesis is based on the observation that in systems as in Fig. 1, each amplifier is a fully fledged high performance circuit which is the most power consuming component of the system. Multiple copies of this amplifier, required for multichannel recording, drastically increase the overall power consumption. If somehow, one super-

Sami ur Rehman and Awais M. Kamboh are with the School of Electrical Engineering and Computer Science at the National University of Sciences and Technology, Islamabad, Pakistan (phone: +92-52-90852119; fax: +92-51-8317363; e-mail: sami.rehman@seecs.edu.pk, awais.kamboh@seecs.edu.pk).

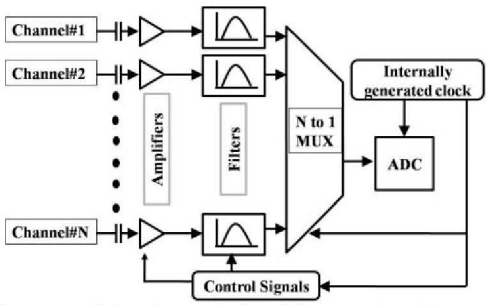

Fig. 1. Traditional neural signal front end architecture

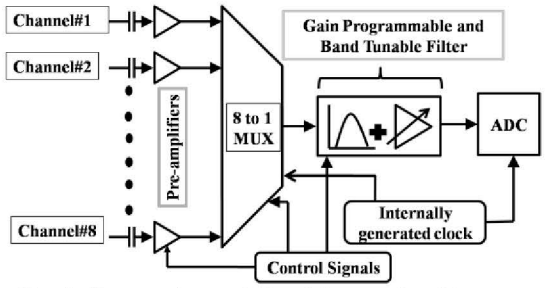

Fig. 2. Proposed neural signal front end architecture

performing amplifier could be designed that can handle multiple channels simultaneously and thus replace these amplifiers, the overall power consumption for a multi-channel system could be reduced.

We propose to divide the amplification into two stages. The fixed-gain small-sized preamplifiers are followed by an analog mux, and a super-performing gain-bandwidth programmable filter-amplifier to record both spikes and LFPs. In essence we have attempted to reduce the area and power by pushing the analog mux closer to electrodes and sequentially recording neural signals within a time window which ensures no signal on any channel is missed. The downside of this approach is the stringent requirement placed on the settling time of the super-performing amplifier (which in Fig. 1 was only faced by the Mux) or the speed of each component in general. However, by properly designing the amplifier for the required settling time, we can save a considerable amount of power.

The preamplifier in the proposed architecture, shown in Fig. 2, is a high input impedance common-source amplifier with the gain of a few decibels. It is used to pick the weak input signal immersed in noise. The signal obtained from all the pre-amps is input to a transmission gate based multiplexer and then fed in to a filter with controllable gain and bandwidth. The filter-amplifier adjusts its bandwidth and gain to measure spikes or LFPs as determined by the control signals. The amplifier is followed by an 8-bit successive approximation register (SAR) ADC, best suited for biomedical applications because of its balance in speed, resolution, power and area.

This paper presents the prototype of an 8-channel recording system as a proof of concept. Considering the highest neural signal frequency of 7kHz the sampling frequency is set at 20ksps per channel. Since the output signals from 8 channels have to be time-multiplexed, this translates to a maximum allowed scan time of 6.25µs for each channel. This implies that every channel will receive the signal, filter/amplify it, and digitize it within 6.25µs after every 50µs. Therefore, although the proposed architecture utilizes fewer circuit blocks to conserve power and area, the settling time of the filter-amplifier poses a serious challenge. The gain-bandwidth programmability means that any of the channels can be used to record LFPs or spikes as needed.

The filter-amplifier block in Fig. 2 consist of a gain stage followed by band tunable filter stage as shown in Fig. 3. The capacitive feedback sets the mid-band gain (C_{in}/C_{feed}), the low cut-off frequency is adjusted by $V_{tune,}$ and the high-pass cut-off frequency depends on C_c, apart from OTA trans-conductance and its gain.

The design uses a fully differential telescopic OTA in the gain stage as shown in Fig. 4. Input differential pair, M1 and M2, is the major source of device-size-dependent flicker noise in the designed OTA. Similarly, a high input trans-conductance is required to suppress thermal noise as given in (1) which describes the input referred noise of the OTA. Therefore large PMOS devices operating in sub-threshold regime/weak inversion regime instead of field effect regime are used not only to constrict noise but also to conserve power.

$$V_n^2, in = \frac{4kT}{2\kappa g_{m1}} + \frac{4kTg_{m5}}{2\kappa g_{m1}^2} + \frac{K}{C_{ox}(WL)_1 f} \qquad (1)$$

where K is a process dependent parameter on the order of $10^{-25}V^2F$, k is Boltzmann constant and κ is sub-threshold gate coupling coefficient. Transistors M3 and M4, also operating in sub-threshold are cascaded with input differential pair to enhance the output impedance. M5 to M12 comprise the current steering circuit while M13 to M20 together make the common mode feedback circuit (CMFB). Inversion

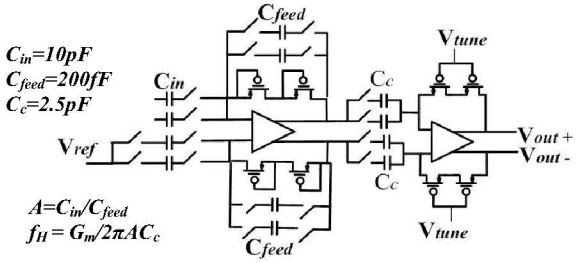

Fig. 3 Band-tunable and gain-programmable filter

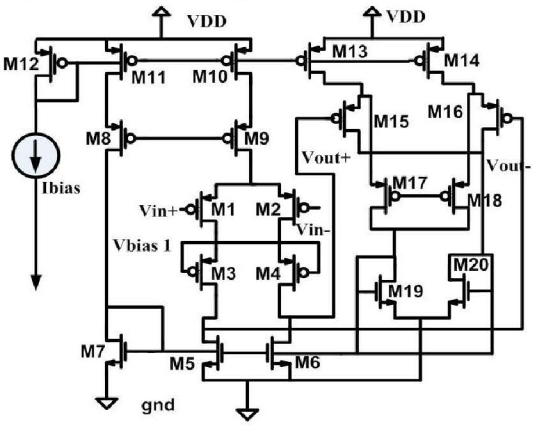

Fig. 4 Telescopic OTA used in neural amplifier

Coefficient (*IC*), given as drain current over saturation current, is the primary parameter that describes the region of operation of a MOSFET and must be less than 0.1 for a device in sub-threshold regime. With *IC* known, trans-conductance and device size in sub-threshold regime can easily be calculated using a procedure nicely elaborated in Harrison's amplifier architecture [11].

III. Results

Designed in 0.5µm CMOS process Fig. 5 shows the layout of 8 channel neural front end amplifier and 8 bit SAR ADC. Table I provides post-layout area and power consumption of each block in the design.

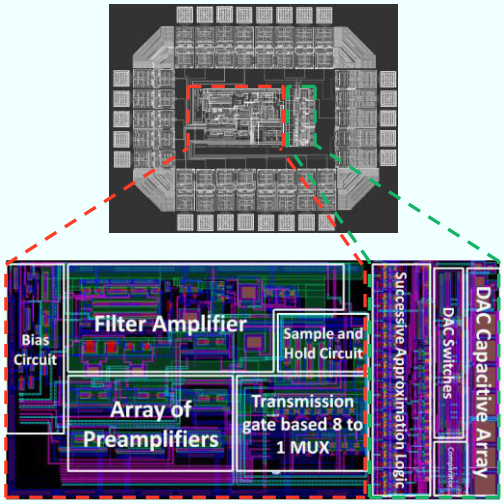

Fig. 5. Chip layout of the overall System displaying the amplification circuit on the left and the SAR ADC on the right

TABLE I. Power And Area Consumption Of Each Element In Our Design

	Power (µW)	Area (mm^2)
Filter Amplifier	12.32	0.11
Preamplifier (×8)	11.2	0.072
Analog MUX	4.54	0.057
SAR ADC	14	0.168
TOTAL	42.06	0.407

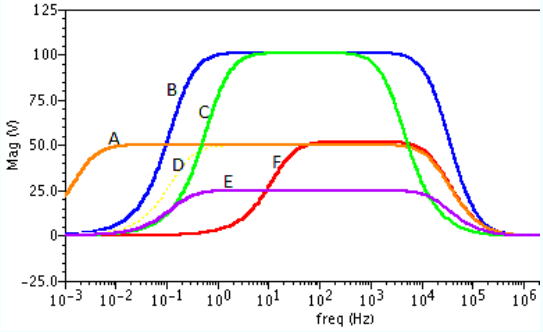

Fig. 6. Variable-gain & band-tuneable filter output plots

TABLE II Typical Values Of Factors Affecting Variable Gain And Bandwidth In Proposed Design

	Plot Curves					
	A	**B**	**C**	**D**	**E**	**F**
C_{in}/C_{feed}	50	100	100	50	25	50
V_{tune} (V)	1.2	0.7	0.7	0.7	0.7	0.5
C_c (pF)	5	2.5	2.5	5	5	5

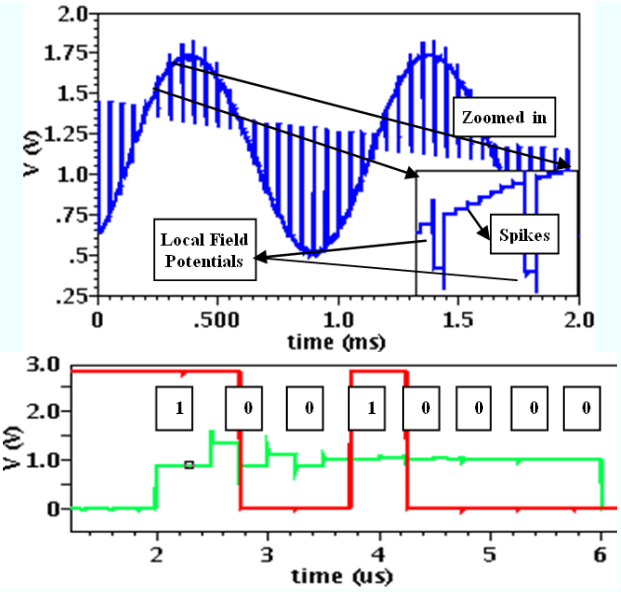

Fig. 7. (Top) Neural signal amplifier output. (Bottom)SAR ADC output with an input voltage to ADC held at 1V.

A. Gain Bandwidth Variability

The OTA has a DC gain of around 76dB and common mode rejection ratio of 72dB. Fig. 6 shows plots for different gains, lower and higher cut off frequencies obtained by setting appropriate values of feedback switching capacitors, 'V_{tune}', and coupling capacitor respectively in Fig. 3 using control circuit (not shown). Table II gives bandwidth parameters corresponding to plots in Fig. 6.

B. Transient Response

Fig. 7 shows the time-multiplexed output of neural signal amplifier after being sampled and held when the inputs to the electrodes were sinusoids of relevant frequencies and amplitudes. The SAR ADC output for a 1V input value along with the corresponding digital code and the voltage variation at DAC capacitive array for 8 clock cycles is also shown in the bottom part of Fig. 7.

C. Noise Performance

Noise efficiency factor is used to determine the noise characteristics of an analog front-end. The theoretical floor of NEF, when the two input differential PMOS are the only source of thermal noise, (ignoring flicker noise), is 2.9 as calculated in [12]. At 2.2µA the input referred noise is 6.7µVrms. This corresponds to a NEF of 4.6 which is quite acceptable to process the neural signals under consideration. Fig. 8 shows a plot of input referred noise with the bias current of the designed filter amplifier.

D. Settling Time

As discussed in section II, with a channel rate of 6.25µs/channel the settling time requirement of the neural filter amplifier becomes approximately 1µs, while 4.5µs are allotted to SAR ADC to produce the corresponding digital code. Settling time of the amplifier can be decreased or increased with a trade-off on power consumption. Smaller settling time requires more power. This becomes important if

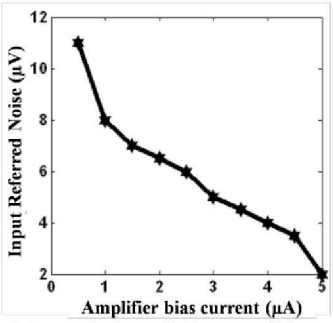

Fig. 8. Noise variation of filter amplifier with current.

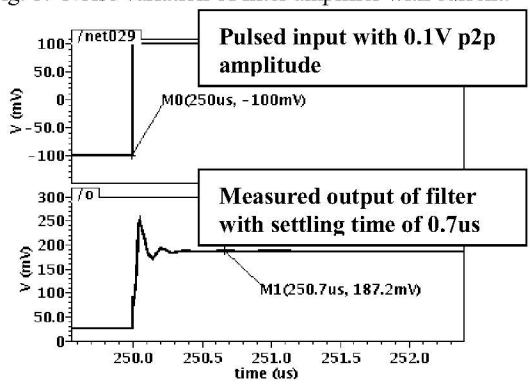

Fig. 9. Settling time of filter amplifier in Fig. 2

the number of channels in the system needs to be different from the 8-channel prototype. Fig. 9 shows that the settling time of the filter amplifier lies well within this specifications when a step input is applied.

Table III provides a comparison of the presented amplification-filtration architecture with existing designs. This includes the 8-channel preamplifiers, the mux and the programmable filter-amplifier. For fair comparison of multi-channel system performance, all designs have been scaled to 8 channels. It can be seen that for comparable NEF and bandwidth, the proposed architecture results in a smaller area and smaller power consumption. In addition, it provides good gain and uses the comparably inexpensive CMOS technology.

TABLE III. QUANTITATIVE COMPARISON OF NEURAL AMPLIFIERS

Ref/ Year	Tech (μm)	Area (mm^2)	Gain (dB)	Pwr. (uW)	NEF	BW
[12] 2009	0.18	0.3	36	90	3.6	up to 10K
[13] 2009	0.35	0.32	34	33.6	4.6	NA
[5] 2010	0.35	0.496	33	102	5.55	10-5K
[6] 2011	0.18	0.24	49-49	51.75	4.4-5.9	350-12K
[7] 2011	0.18	0.524	39.4	63.36	3.35	up to 7.2K
[8] 2012	0.5	NA	62	32	3.3	100-7K
This work	0.5	**0.239**	60	**28.06**	4.6	200-9.6K

NOTE: Area and power consumption figures for all designs are scaled for 8 channels for a fair comparison of a multi-channel system.

IV. CONCLUSION

A new architecture for neural signal amplification has been presented for simultaneous multi-channel neural recordings. Post layout simulations in 0.5um CMOS process prove that the proposed architecture consumes less area and smaller power than all the multi-channel architectures presented in recent literature. Configured for 8 channels, it consumes only 28.06μW of power, (3.5μW per channel) with an area of $0.239mm^2$ ($0.029mm^2$ per channel) excluding the ADC. The architecture is able to handle neural spikes as well as LFPs simultaneously. This is enabled by the gain-bandwidth programmability introduced in the design. Keeping high gain-bandwidth product, the fully differential architecture delivers a good noise efficiency factor and high common mode rejection ratio. Comparison to recently published works proves the advantages discussed above. These performance metrics make this architecture extremely suitable for implantable multi-channel neural signal acquisition systems.

REFERENCES

[1] W. Liu, M. Sivaprakasam, G. Wang, M. Zhou, J. Granacki, J. Lacoss, and J. Wills, "Implantable biomimetic microelectronic systems design," *IEEE Eng- Med. Biol. Mag*., vol. 24, no. 5, pp. 66–74, 2005.

[2] M. Steyaert and W. Sansen, "A micropower low-noise monolithic instrumentation amplifier for medical purposes," *IEEE J. Solid-State Cir*, vol. 22, no. 6, pp. 1163–1168, 1987.

[3] R.R. Harrison, "A versatile integrated circuit for the acquisition of biopotentials," in Proc. *IEEE Custom Integrated Circuits Conference* (CICC), pp. 115–122 , 2007

[4] M. Yin and M. Ghovanloo, "A low-noise preamplifier with adjustable gain and bandwidth for biopotential recording applications," *IEEE Int. Symp. Cir and Syst*., pp. 321-324, 2007

[5] F. Shahrokhi, K. Abdelhalim, D. Serletis, P. Carlen, and R. Genov,"128-channel fully differential digital integrated neural recording and stimulation interface," *IEEE Trans. on Biomed. Cir. and Syst*., vol. 4, no. 3, pp. 149–161, 2010.

[6] W. Wattanapanitch and R. Sarpeshkar, "A low-power 32-Channel digitally programmable neural recording integrated circuit," *IEEE Trans. Biomed. Cir. and Syst.*, vol. 5, pp. 592–602, 2011.

[7] V. Majidzadeh, A. Schmid, and Y. Leblebici, "Energy efficient low noise neural recording amplifier with enhanced noise efficiency factor," *IEEE Trans. Biomed. Cir. and Syst.*, vol. 5, no. 3, pp. 262-271, 2011.

[8] S. S. Saberhosseini, A. Zabihian, and A. M. Sodagar, "Low-noise OTA for neural amplifying applications," *Int. Caribbean Conf. on Dev., Cir. and Syst.*, pp.1-4, 2012.

[9] Y. –T. Liu, D. Y. C. Lie, W. Hu, and T. Nguyen, "An ultra- low power CMOS transconductor design with wide input linear range for biomedical applications," *IEEE Int. Symp. Cir. and Syst.*, pp.2211-2214, 2012

[10] M. S. Chae, W. Liu, and M. Sivaprakasam, "Design optimization for integrated neural recording systems," *IEEE J. Solid State Cir*, vol.43, pp. 1931–1939, 2008.

[11] R.R. Harrison, 'A low-power, low-noise CMOS amplifier for neural recording', *IEEE Int. Symp. Circuits and Syst.,* vol. 5, pp. 197-200, 2002.

[12] T. Yoshida, Y. Masui, R. Eki, A. Iwata, M. Yoshida, and K. Uematsu, "A neural signal detection amplifier with low-frequency noise suppression," *IEEE Int. Symp. Cir. and Syst.,* pp.661-664, 2009.

[13] J. N. Y. Aziz, K. Abdelhalim, R. Shulyzki, R. Genov, B. L. Bardakjian, M. Derchansky, D. Serletis, and P. L. Carlen, "256-channel neuralrecording and delta compression microsystem with 3D electrodes," *IEEE J. Solid State Cir*, vol. 44, pp. 995–1005, 2009