A 330μW, 64-Channel Neural Recording Sensor with Embedded Spike Feature Extraction and Auto-calibration

Alberto Rodríguez-Pérez[1], Manuel Delgado-Restituto[1], Angela Darie[1], Cristina Soto-Sánchez[2],
Eduardo Fernández-Jover[2] and Ángel Rodríguez-Vázquez[1]
[1]Institute of Microelectronics of Sevilla (IMSE-CNM) and University of Sevilla
[2]Bioengineering Institute, University Miguel Hernández and CIBER-BBN

Abstract—**This paper reports an integrated 64-channel neural recording sensor. Neural signals are acquired, filtered, digitized and compressed in the channels. Additionally, each channel implements an auto-calibration mechanism which configures the transfer characteristics of the recording site. The system has two transmission modes; in one case the information captured by the channels is sent as uncompressed raw data; in the other, feature vectors extracted from the detected neural spikes are released. Data streams coming from the channels are serialized by an embedded digital processor. Experimental results, including *in vivo* measurements, show that the power consumption of the complete system is lower than 330μW.**

I. Introduction

Besides fostering advances in neuroscience, wireless neural prostheses for the measurement of intracranial neural activity are expected to play a significant role in the development of novel treatments for some neurological diseases and in the implementation of untethered brain-machine interfaces [1]–[5]. As long as these prostheses are implanted, they have to achieve and maintain stable long-term recordings so that the need for re-surgery is essentially eliminated. This poses important challenges on the hardware implementation of the prostheses as they have to exhibit utra-low power consumption, not only to prevent from harmful effects in the brain but also to minimize energy requirements; low form factor; versatility, to prove useful in different scenarios as determined by neurologists; and adaptability to deal not only with the intrinsic statistical deviations of the fabrication process but also with the non-stationary nature of the electrode-tissue interface.

In this scenario, this paper presents an integrated 64-channel neural recording sensor suitable for acquiring Local Field Potentials (LFPs) and Action Potentials (APs). An on-chip dedicated processor defines the operation mode of the channels and implements a full-duplex communication protocol for data transmission through a wireless link. In one operation mode, the recording system can be configured to detect and compress neural spikes so that feature vectors instead of raw signal samples are transferred. In another mode, the system runs a self-calibration mechanism which automatically adapts the filter bandwidth and the gain setting of the channels. The sensor also offers different alternatives for raw data transmission in which the number of active channels and the effective sampling rate are traded off. In all cases, the total throughput rate of
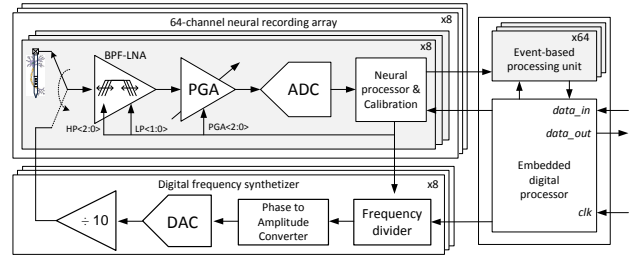

Figure 1.   Architecture of the 64-channel neural recording array

the sensor keeps below 4Mbps as imposed by the wireless link. The sensor has been fabricated in a 0.13μm standard CMOS process and consumes 330μW from a 1.2V voltage supply in the spike compression mode, the most demanding one. The architecture of the recording sensor as well as a description of the different operation modes are presented in Section 2. Afterward, Section 3 presents some experimental measurements and *in vivo* validation results. Finally, Section 4 concludes the paper and compares the proposed recording system to others in the current state-of-the-art.

II. Neural sensor architecture

Fig. 1 shows the architecture of the proposed system. It consists of an array of 8×8 neural recording channels, each of them serially connected to an Event-Based Processing Unit (EBPU). The data stored in these EBPUs are read and classified by an embedded digital processor, which also handles the timing of the implant. A wireless transceiver (not shown) provides the link to/from an external hub. Additionally, the system includes one tunable Digital Frequency Synthesizer (DFS) per row for calibration purposes.

A. Channel sensor interface

Each channel comprises all the needed circuitry to acquire and digitize neural waveforms including a Band-Pass Filter Low Noise Amplifier (BPF-LNA), a digitally tunable band-pass filter, a Programmable Gain Amplifier (PGA), an Analog-to-Digital Converter (ADC) and a local digital processor to detect neural spikes and extract their features. The spike detection is accomplished in digital domain and the decision threshold is adaptively updated according to the noise floor of the captured signal.
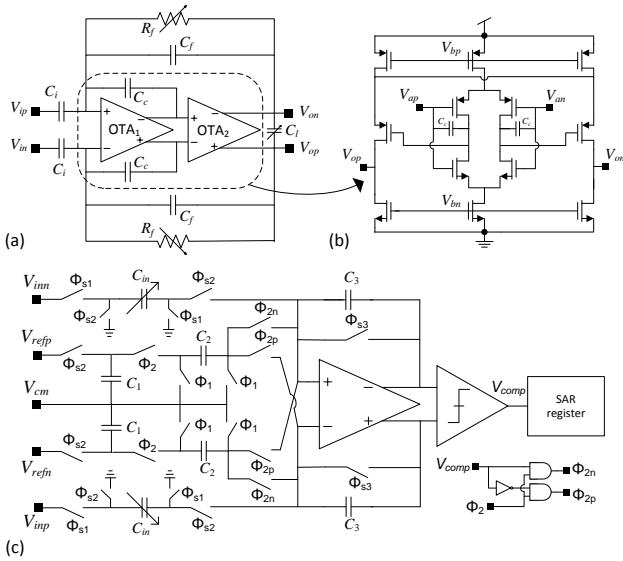
Figure 2.   (a) Schematic of the BPF-LNA and (b) the OTAs. (c) Schematic of the PGA-ADC

Fig. 2(a) shows the schematic of the BPF-LNA which uses a two-stage capacitive feedback network architecture. Its midband gain is given by the ratio between the input and the feedback capacitance, C_i/C_f. The position of the high-pass and low-pass corners of the bandpath characteristic are approximately given by $1/R_fC_f$ and $2g_{m2}C_f/C_cC_l$, respectively. The first amplifier stage (OTA_1) has been designed so that its output pole matches the low-pass corner of the bandpass, thus leading to a 40dB/dec roll-off, beneficial for suppressing high frequency noise components. Additionally, as Fig. 2(b) shows, a complementary input differential pair has been used for OTA_1 to nearly double its equivalent transconductance for the same biasing current. Each feedback resistor in Fig. 2(a) is implemented as a 3-b digitally-controlled tapped cascade of transistors biased in deep subthreshold region [6]. This makes the feedback resistors programmable thus, allowing to externally tune the position of the high-pass pole of the BPF-LNA. Similarly, the load capacitor can be digitally adjusted through 2 programming bits and, hence, the low-pass pole is tunable as well.

Fig. 2(c) shows the schematic of the implemented ADC [7]. It is built around a SC integrator whose gain can be controlled from 0 to 18dB by digitally programming the input capacitor bank C_{in}. Hence, besides conversion, the circuit also features PGA capabilities. Outputs bits are derived by successively detecting the sign of the voltage stored in the integrator. Depending on the output of the comparator, the integrated voltage is updated by adding or substracting binary scaled versions of a voltage reference V_{ref}. These voltages $V_j = V_{ref}/2^j$, $j = 1, \ldots, n - 1$, , where n is the output resolution of the converter, are obtained by capacitive voltage division at every step of the conversion process. Solved bit are stored in a SAR register. In the presented design, the bias current of the OTA is dynamically adapted for power saving by taking advantage that settling requirements are progressively relaxed along the conversion.

B. Modes of operation

Three operation modes are available in the proposed neural recording system. They are the calibration, signal tracking and feature extraction modes. These modes and their associated parameters are specified through a custom communication protocol implemented in the embedded digital processor.

Calibration: In this mode, the passband and gain of the recording channels are individually adjusted. In one case, the objective is to automatically tune the programming words for the high- and low-pass corners of the BPF-LNA in order to satisfy a given capture frequency range. In the other, the calibration mode automatically programs the gain of the PGA in order to maximize the voltage swing at the input of the ADC. Different to [8], reference tones to set the passband corners of the recording channels are provided by programmable DFSs. As there is one DFS per row, passband calibration is done in a column-wise manner. As shown in Fig. 1, each DFS consists of a programmable frequency divider, followed by a phase-to-amplitude converter. This block cyclically accesses the registers of a ROM memory which stores equally spaced samples of a sine function. The phase-to-amplitude converter is followed by a DAC and an analog attenuator which adapts the amplitude of the generated tone to the LNA's input swing. By controlling the clock rate provided by the programmable divider, the frequency of the output tone can be adjusted to the desired value.

Signal tracking: Under the signal tracking mode, the system transfers the uncompressed recorded data acquired from the selected channels. The system offers different tracking possibilities which trade-off the number of selected channels and the time interval between samples. In all cases, the overall throughput rate of the system remains below 4Mbps, as imposed by the wireless transceiver. Regardless of the configuration, a sampling rate of 30kS/s is used in all the selected channels.

Feature extraction: In this case, the system is configured for spike detection and data compression tasks. All the 64 channels are enabled during feature extraction. Spike compression is implemented by obtaining on-the-fly a Piece-Wise Linear (PWL) representation of its waveform. As shown in Fig. 3, this representation comprises two amplitude values (V_{p1} and V_{p2}) and three time slots (Δ_{1-3}), together with the magnitude of the threshold voltage used in the detection of the spike. All these parameters are coded in 8-b words, with the exception of the threshold amplitude which uses 7-b; hence, the whole PWL representation occupies 47-b. The feature extraction process is similar to the one presented in [6], however, in the proposed implementation both the spike detection and the spike feature extraction tasks are performed in the digital domain.

C. Event-Based Processing Units

EBPUs are used for temporarily storing the information provided by the channels. During the calibration and signal tracking modes, channels serialize and transfer data to the EBPUs, where information is retained until it is retrieved by the system digital processor. In the feature extraction mode,
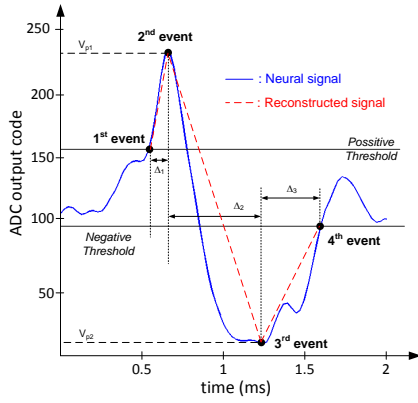
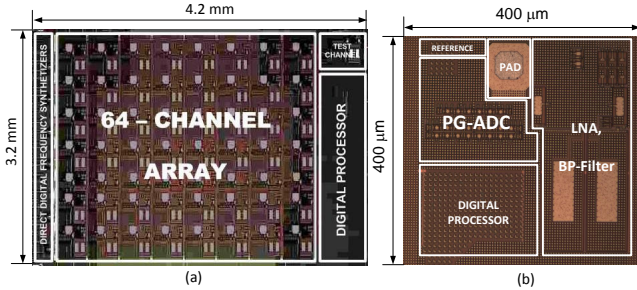Figure 3. Piecewise linear approximation of spikes.

Figure 4. Microphotography of the (a) 64-channel neural sensor, (b) channel

EBPUs also calculate the time intervals comprised in the PWL representation of spikes. Every channel informs to the associated EBPU on the instants in which waveform peaking and threshold crossing events are taking place, as Fig. 3 shows. Then, the EBPU uses these triggering signals for evaluating the time slots between events by means of counters. In order to obtain increased time resolution for the measurement of the interval durations, the sampling frequency of the PGA-ADC is changed from the nominal 30kS/s rate to 90kS/s. When a spike finishes, the channel sends to the EBPU the amplitude parameters needed to complete the PWL feature vector. Once the vectors of the representation are gathered, the EBPU asserts a flag to inform that the stored data is ready to be read out. The system digital processor cyclically checks the state of these flags. In case a flag is enabled, it retrieves the information from the EBPU at a rate of 4Mbps, builds up a transmission frame and sends this stream to the wireless transceiver for data transmission.

III. EXPERIMENTAL RESULTS

Fig. 4(a) shows the microphotography of the presented 64-channel neural array sensor fabricated in a standard $0.13\mu m$ CMOS process. The whole system occupies an area of $13.45mm^2$ and is organized in 8×8 channels, each of them laid out in a square of $400 \times 400\mu m^2$, as shown in Fig. 4(b). Note that each channel includes an internal pad for flip-chip connection to a microelectrode. DFSs and embedded digital processor are respectively placed at both sides of the channel array.

Fig. 5(a) shows the frequency response of the BPF-LNA under different configurations for the LP and the HP pole positions. The HP pole can be adjusted between 15 and 232Hz
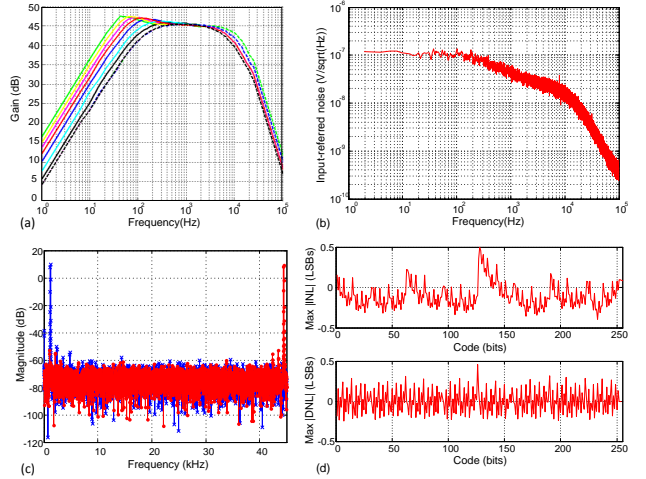

Figure 5. Measurement results of the channel sensor interface. BPF-LNA: (a) Bandpass transfer function, (b) Power spectral density of the input referred noise. PGA-ADC: (c) FFT spectrum of the output at 90kS/s for low and Nyquist input frequencies, (d) INL and DNL.
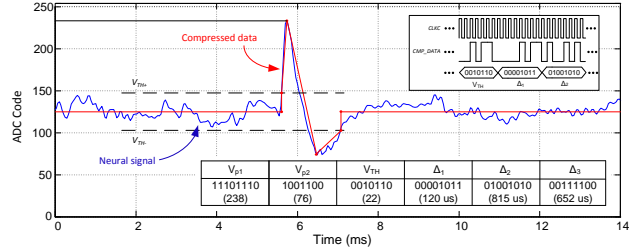

Figure 6. Measurement under feature extraction mode

along 8 different positions, while the LP pole can be tuned at four different frequencies between 5.2kHz and 10.15kHz. The midband gain is 45dB. Fig. 5(b) illustrates the power spectral density of the BPF-LNA's input referred noise. Note that the 1/f noise contribution is attenuated at low frequencies by the HP transfer pole, so it results in a flat noise level band. The total integrated noise power is 3.8µVrms when integrated between 1Hz and 100kHz, and reduces to 3.2µVrms within the spike recording band, between 200Hz and 7kHz. The BPF-LNA consumes 1.92µW, resulting in a Noise Efficiency Factor (NEF) of 2.16.

Fig. 5(c) shows the spectrum of the PGA-ADC operated at 90kS/s conversion rate (similar results were obtained for 30kS/s) for inputs tones at low frequency and close to Nyquist rate. The Signal-to-Noise and Distortion Ratio (SNDR) is above 47.0 dB and the Equivalent Number of Bits (ENOB) is about 7.65 bit. These results remains unaltered regardless of the selected amplification setting (0-18dB). As shown in Fig. 5(d), both the integral and differential nonlinearity are bounded between ±0.5dB. The power consumption of the PGA-ADC is 1.52µW and 515nW for the 90kS/s and the 30kS/s sampling modes, respectively.

Fig. 6 illustrates the operation of a neural channel under the feature extraction mode. If no spike is detected, the EBPU associated to the channel remains idle. When a spike is detected, both the channel and the EBPU work together to obtain the parameters of the PWL representation. Once the spike ends, the EBPU stores the parameters and enables the
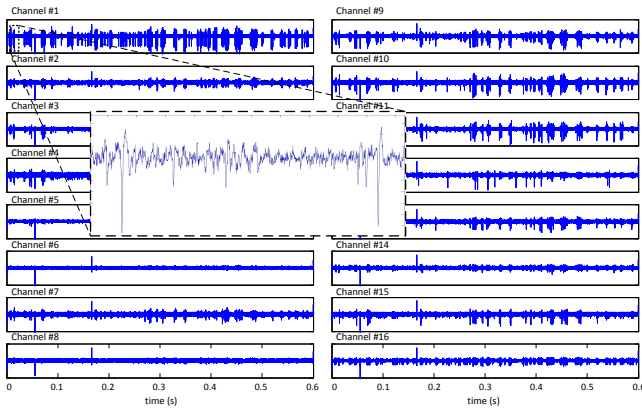
Figure 7. *In vivo* neural spike activity recorded from 16 channels using an intracranial microelectrode array. Vertical axis correspond to the ADC output code and range between 0 and 255.
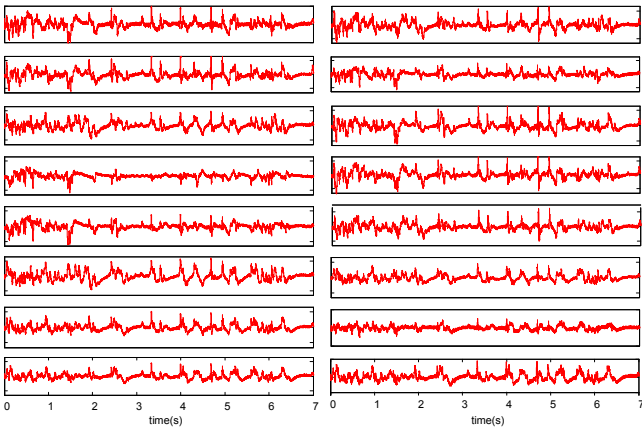

Figure 8. *In vivo* LFP activity recorded from 16 channels using a flexible non-penetrating sub-dural microelectrode array. Vertical axis correspond to the ADC output code and range between 0 and 255.

ready flag. Afterward, the digital processor reads and serializes the feature vector (see the inset of the figure). In this operation mode, all channels are active, each consuming 4.54µW. This together with the power consumption of the EBPUs and the digital processor gives a total dissipation of 330µW.

Fig. 7 shows 16 *in vivo* recordings captured by the system by using an intracranial microelectrode array by Blackrock Microsystems (adult male Long Evans rat model). The bandpath characteristics of the 16 selected channels were set between 200Hz and 7kHz in order to capture spike activity. As can be observed, isolated action potentials and bursts of spikes are clearly noticeable in the recordings. The power consumption of the system in this experiment was 98µW.

Another in vivo experiment with a different rat model was conducted using a flexible non-penetrating sub-dural microelectrode array (Multi Channel Systems MCS GmbH) with TiN electrodes separated by 300µm. In this case, all the 64 channels of the array were selected and their bandpass characteristics were set between 15Hz and 5.2kHz. The throughput rate per channel was reduced to 4KS/s. Fig. 8 shows that LFPs were successfully recorded (only 16 channels are presented for simplicity). The power consumption in this case was 241.5µW.

Table I
PERFORMANCE SUMMARY AND COMPARISON

	[1]	[2]	[3]	[4]	This work
Technology (nm)	130	65	180	180	130
Supply voltage (V)	1.2/3.3	0.5	1/1.8	0.45	1.2
Number of channels	64	64	100	100	64
Total power (μW)	640	147.2	1160	94	< 330
Power / channel (μW)	10	2.3	11.6	0.94	3.04/4.54
Max. system gain (dB)	60	N.A.	60.9	55.5	63
High pass freq. (Hz)	1	1-8	248	0.25	15-232
Low pass freq. (kHz)	5	0.4	5.1	10	5.2-10.15
Input ref. noise (μV_{rms})	5.1	1.27	4	3.2	3.8
NEF	4.4	4.76	1.9	1.57	2.16
ENOB (bit)	7.6	15	7.2	8.27	7.65
Sampling freq. (kS/s)	12.5	1	24.5	20	30 / 90
Data bitrate reduction	Yes	No	No	No	Yes

IV. CONCLUSIONS

A 64-channel neural array with embedded data reduction techniques, fabricated in a standard CMOS 130nm process, is presented. Each channel embeds all the circuitry to filter, amplify and digitizes the input data, as well as compress the detected neural spike activity, minimizing the amount of generated data. A distributed digital signal processing approach, with tasks at channel- and array levels, has been found an efficient solution for reducing the power consumption of the SoC and simplifying communications through the array. Table I compares the proposal to the state-of-the-art. Note that the most power efficient solutions don't include any embedded data compression technique.

ACKNOWLEDGMENTS AND ETHICAL APPROVAL

REFERENCES

[1] K. Abdelhalim *et al.*, "64-channel uwb wireless neural vector analyzer soc with a closed-loop phase synchrony-triggered neurostimulator," *IEEE J. of Solid-State Circ.*, vol. 48, no. 10, pp. 2494–2510, 2013.

[2] R. Muller *et al.*, "A miniaturized 64-channel 225uw wireless electrocorticographic neural sensor," in *Digest ISSCC*, 2014, pp. 412–413.

[3] X. Zou *et al.*, "A 100-channel 1-mW implantable neural recording IC," *IEEE Trans. on Circ. and Systems I*, vol. 60, no. 10, pp. 2584–2596, 2013.

[4] D. Han *et al.*, "A 0.45V 100-channel neural-recording IC with sub-uw/channel consumption in 0.18um CMOS," in *Digest of ISSCC*, 2013, pp. 290–291.

[5] R. R. Harrison *et al.*, "A low-power integrated circuit for a wireless 100-electrode neural recording system," *IEEE J. Solid-State Circ.*, vol. 42, no. 1, pp. 123–133, 2007.

[6] A. Rodriguez-Perez *et al.*, "A low-power programmable neural spike detection channel with embedded calibration and data compression," *IEEE Trans. on Biom. Circ. and Systems*, vol. 6, no. 2, pp. 87 –100, 2012.

[7] A. Rodríguez-Perez *et al.*, "A 515 nw, 0-18 db programmable gain analog-to-digital converter for in-channel neural recording interfaces," *IEEE Trans. on Biom. Circ. and Systems*, vol. 8, no. 3, pp. 358–370, 2014.

[8] A. Rodriguez-Perez *et al.*, "A self-calibration circuit for a neural spike recording channel," in *Proc. of IEEE BioCAS*, 2011, pp. 464–467.