A Power Efficient Neural Spike Recording Channel with Data Bandwidth Reduction

Alberto Rodríguez-Pérez, Jesús Ruiz-Amaya, José Antonio Rodríguez-Rodríguez, Manuel Delgado-Restituto and Ángel Rodríguez-Vázquez

Institute of Microelectronics of Seville (IMSE-CNM-CSIC) and University of Seville
e-mail: alberto@imse-cnm.csic.es Tlf.: +34 954 466 667 Fax: +34 954 466 600

***Abstract* - This paper presents a mixed-signal neural spike recording channel which features, as an added value, a simple and low-power data compression mechanism. The channel uses a band-limited differential low noise amplifier and a binary search data converter, together with other digital and analog blocks for control, programming and spike characterization. The channel offers a self-calibration operation mode and it can be configured both for signal tracking (to raw digitize the acquired neural waveform) and feature extraction (to build a first-order PWL approximation of the spikes). The prototype has been fabricated in a standard CMOS 0.13μm and occupies 400μmx400μm. The overall power consumption of the channel during signal tracking is 2.8μW and increases to 3.0μW average when the feature extraction operation mode is programmed.**

I. Introduction

The design of implantable neural recording microsystems for the wireless transmission of brain activity has focused the attention of many research groups during the last years [1-4]. These microsystems can be very useful both in clinical (as part of therapeutic procedures in patients with neurological diseases) and neuroscience applications. Compared to wired systems, these solutions provide untethered communications with the external world, allowing for more freedom of movement and avoiding the risk of infections associated with percutaneous plugs.

Wireless connectivity implies two major challenges on the design of implantable biosensors. First, as microsystems are powered by small batteries or even rely on transcutaneous inductive links, the need for ultra-low power design strategies is mandatory. Second, the data rate that can be transmitted by wireless telemetry circuits is limited and, therefore, some data reduction mechanisms must be performed in the implanted device [1-4]. This is particularly true in the case of microsystems mounted on Multi-Electrode Arrays (MEAs) in which, the transmission bandwidth requirement linearly increases with the number of recording channels.

In this paper we present a low power neural spike recording channel suitable for raw brain activity acquisition at 22.5kS/s and on-chip detection/characterization of firing action potentials. This channel is intended to be part of a large multi-channel array and uses first-order Piecewise-Linear (PWL) approximations for spike feature extraction. This data compression mechanism could be useful in cortically-controlled neuroprosthetics for spinal cord injured patients. The proposed solution employs mixed-signal techniques, and uses a foreground digital calibration process to counteract technology process variations. The complete channel has been integrated in a CMOS standard 130nm technology and it occupies 400μmx400μm, thus fitting in the pitch of commercially available MEAs [1].

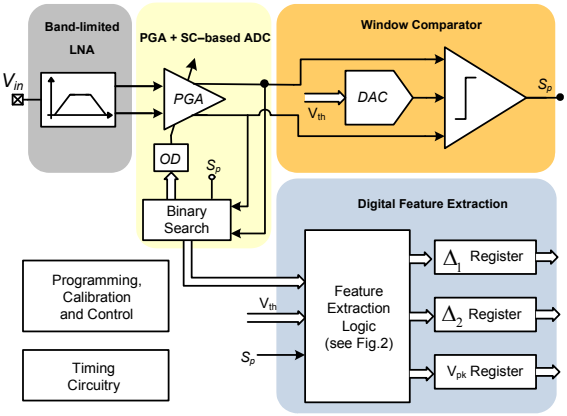

Fig. 1: Simplified neural recording channel architecture.

The paper is structured as follows. Section II describes the architecture of the channel and its operation modes, while Section III and Section IV describe, respectively, two of its basic building blocks, the band-limited low noise amplifier and the data converter. Section V shows post-layout simulations of the recording channel and summarizes the achieved performance. Finally, Section VI gives some conclusions.

II. Recording Channel Architecture

Fig. 1 shows the block diagram of the proposed neural spike recording channel. It consists of a band-limited fully-differential Low-Noise Amplifier (LNA), a reconfigurable Analogue-to-Digital Converter (ADC) which embeds a Programmable Gain Amplifier (PGA), a Switched-Capacitor (SC) window comparator and some digital circuitry for control and feature extraction purposes, which is described in [5].

Three different operation modes can be programmed in the recording channel. They are the *calibration*, *signal tracking* and *feature extraction* modes. Depending on the operation mode, idle blocks are total or partially powered off for power saving. Active blocks are operated by dedicated clock signals (derived from a master 1.6MHz reference) at frequencies that depend on the particular system configuration.

In the *calibration* mode, the bandpass characteristic of the

LNA is automatically tuned using the master clock as reference frequency and, afterwards, the gain of the PGA and the threshold voltage magnitude for spike detection are digitally adjusted. In this mode, only the LNA, the ADC and some calibration specific circuitry (not shown in Fig.1) are active. The ADC operates at a 200kHz clock frequency to provide a throughput rate of 22.5kS/s and 8-bit resolution.

In the *tracking* mode, the signal captured by the neural electrode is conditioned by the LNA and 8-bits digitized by the ADC at a 22.5kS/s throughput rate. The output of the ADC is serially read out at a rate of 1.6MHz to allow for real-time external reconstruction of the neural data.

The *feature extraction* mode provides a data compressing mechanism to reduce the bandwidth of the information transmitted from the channel. Two phases can be distinguished in this operation mode which are referred to as *spike detection* and *spike processing*.

During *spike detection*, the output of the PGA is directly processed by the window comparator at a clock frequency of 100kHz. The SC window comparator detects if the voltage magnitude of the PGA output exceeds the threshold value calculated during calibration. Once a positive or negative spike is detected, the window comparator is disabled, and the *spike processing* phase starts. The objective of this phase is to create a PWL approximation of the peak by means of two time segments (one for the peak position and the other for the overall duration of the spike) and a measurement of the peak amplitude. Biphasic spikes are regarded as a concatenation of two monophasic spikes and they are characterized by two sets of the aforementioned variables (one per phase), together with an additional measurement of the separation between the phases. All the variables are codified into 8-bit digital vectors.

Spike processing is implemented in digital domain. An 8-bit counter, working at 100kHz provides the temporal information required by the spike characterization. The ADC, operated at a clock frequency of 800kHz (90kS/s throughput rate) is used to keep track of the signal amplitude along the spike. Finally, some registers and digital comparators are used to determine the peak position and the end of the spike. It is worth observing that the ADC and the window comparator do not operate simultaneously during spike processing, what favors a more homogenous power consumption of the recording channel.

III. Low Noise Amplifier

The LNA must be able to boost the weak signals detected by the microelectrodes and filter out the undesired frequency components [6-8]. Our proposed LNA follows the capacitive feedback structure in [7] but it introduces three main modifications for improved performance. First, it uses a fully-differential architecture, shown in Fig. 2(a), to increase the dynamic range of the amplifier and improve PSRR. Second, the Operational Transconductance Amplifier (OTA) uses a high output swing two-stage topology with feedforward capacitive compensation, shown in Fig. 2(b), to obtain a higher low-pass filter rejection of –40dB/dec . Third, the first stage of the OTA

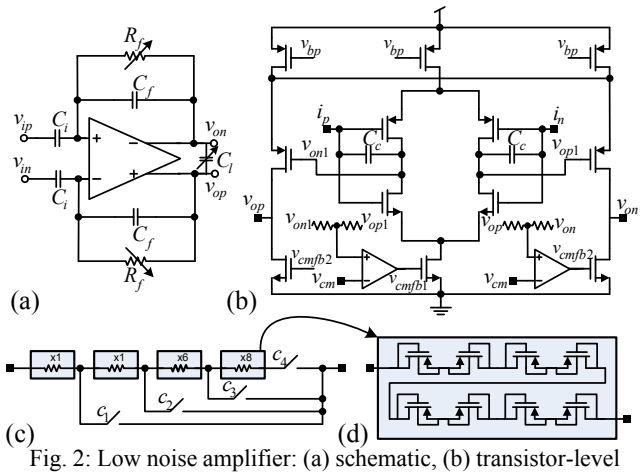

Fig. 2: Low noise amplifier: (a) schematic, (b) transistor-level OTA implementation, (c) programmable feedback resistor and (d) pseudo-resistor realization.

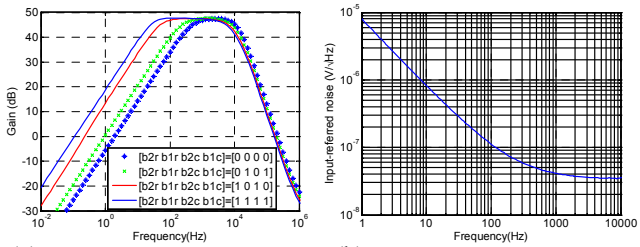

Fig. 3: (a) Transfer characteristic of the low noise amplifier obtained from post-layout simulations. (b) Input-referred voltage noise spectrum.

uses a complementary input differential pair to reuse the tail current and nearly double the achieved transconductance [9].

The high-pass and low-pass frequency corners of the characteristic are 2-bits adjustable, by programming the value of the feedback resistors (control word [b2r,b1r]) and the load capacitor (control word [b2c,b1c]), respectively. Fig. 2(c) shows the structure of the feedback resistors. They are formed by as many pseudo-resistors in series as determined by the 1-out-4 code $\mathbf{c} = [c_1, ..., c_4]$ defined, in turn, by the control word [b2r,b1r]. The schematic of the pseudo-resistors is shown in Fig. 2(d) [9]. They are formed by eight MOS-bipolar devices in series to reduce distortion for large output signals. Mismatch effects and leakage currents have been carefully simulated to ensure feasibility.

Fig. 3(a) shows the transfer characteristics of the LNA for different programming configurations. A mid-band gain of about 47.5dB is achieved by means of the capacitors ratio C_i/C_f. Fig. 3(b) depicts the input-referred voltage noise spectrum for a nominal passband of 217Hz - 7.16kHz. Integration under this curve yields an rms noise voltage of 2.84μVrms. Considering that the power consumption of the amplifier is 1.92μW including CMFB circuits, the Noise Efficiency Factor (NEF) [7] of the LNA is 1.62 (1.75 when the biasing circuitry is also considered).

IV. PGA and ADC

The proposed ADC follows a binary search algorithm for

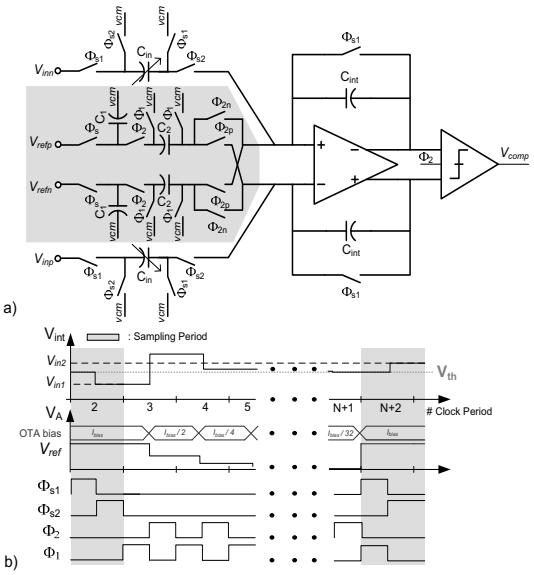

Fig. 4: SC-based ADC schematic.

data conversion but, instead of using a SAR architecture [1-3], it employs the SC structure of Fig. 4(a) [10]. The circuit contains two active blocks, a S&H programmable gain amplifier/integrator and a comparator, as well as a conventional SAR register (not shown in the figure). Circuit operation is as follows. The signal is first sampled and stored in the integrator by transferring charge from capacitor C_{in} to C_{int}. Afterwards, the conversion phase starts. It is simply realized by successively comparing the integrated voltage with $V_{ref}/2^n$, where V_{ref} is the full-scale reference voltage of the converter, and index n =1, ..., N, indicates the conversion step (N represents the converter resolution). The weighted voltages $V_{ref}/2^n$ are generated by a passive SC arrangement (shaded area in Fig. 4(a)). If the integrated voltage is larger (alt., smaller) than $V_{ref}/2^n$ the comparator sets to '1' (alt., '0') the n-th bit of the conversion, and the integrator is updated by substracting (adding) such increment. As illustrated in the timing diagram of Fig. 4(b), conversion takes N+1 clock cycles. Depending on the operation mode, the circuit is clocked at 100, 200 or 800kHz.

During spike detection, the voltage divider and the comparator are disabled and the block operates as a PGA. The gain of the PGA can be adjusted from 0 to 18dB at discrete steps of 3dB (this gives a total gain for the channel front-end of 48-66dB, including the LNA contribution) by 3-bits programming the input capacitance.

Fig. 5(a) shows the schematic of the fully differential OTA in the ADC+PGA. It is a folded-cascode topology with a SC-based common-mode feedback circuit (not shown). The OTA uses a dynamic biasing scheme in which the tail current of the block is adjusted according to the slew-rate demand (it decreases along the conversion process, as shown in Fig. 4(b)) and the selected operation mode. This is accomplished by controlling the widths of transistors M3-M7. Fig. 5(b) shows the schematic of the comparator. It is a current-controlled dynamic-latch buffered with class-A amplifiers. As in the OTA, the bias current is adapted according to the selected

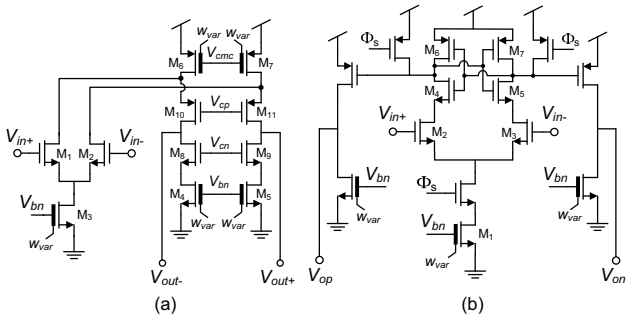

Fig. 5: (a) Schematic of the FD-OTA used in the SC-based ADC. (b) Schematic of the comparator

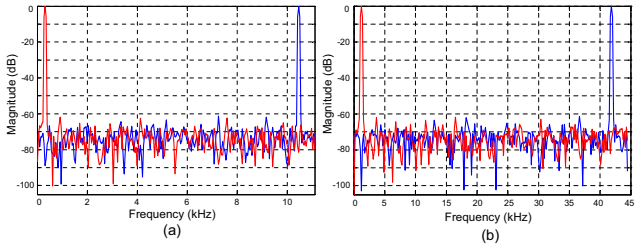

Fig. 6: Spectra of the ADC output response for low and near-Nyquist input frequencies sampled at: (a) 22.5kS/s, (b) 90kS/s

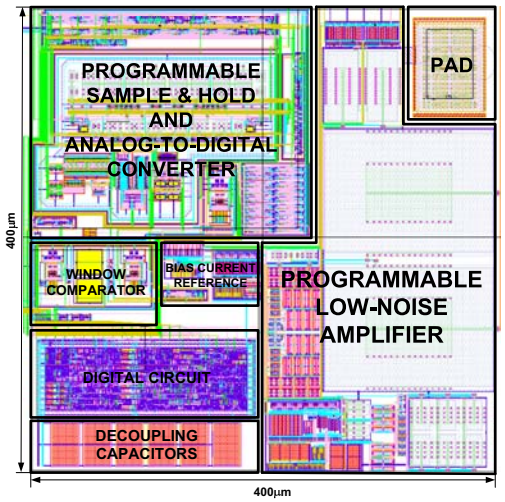

Fig. 7: Layout of the neural spike recording channel.

sampling frequency. The same comparator architecture is also used to implement the window comparator of Fig. 1.

Fig. 6 shows two FFT analyses of the ADC at sampling rates of 90kS/s (spike processing) and 22.5kS/s (signal tracking). The *SFDR* and *SNDR* of the converter are approximately 62dB and 50dB along the whole Nyquist band, respectively. This corresponds to an *ENOB* of almost 8 bits. These features remain unaltered regardless of the gain setting of the embedded PGA.

V. Post-Layout Simulation Results

Fig. 7 shows the layout of the complete neural spike recording cell. It has been implemented in a 130nm CMOS technology and occupies 400µm x 400µm. Fig. 8 illustrates the operation of the proposed neural recording channel in signal tracking mode and feature extraction modes. The plot

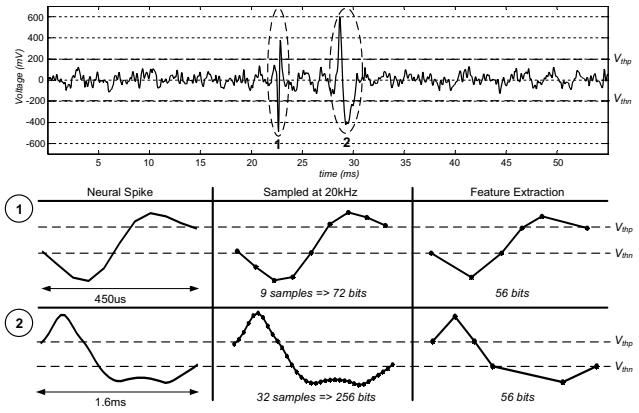

Fig. 8: Illustration of the neural recording channel operation.

Table 1. Performance of the neural spike recording channel

Operation Mode	Signal Tracking	Spike Feature Extraction	
		Processing	Detection
Band-Pass Low-Noise Amplifier			
Input-referred noise	2.84μVrms		
Nominal passband	217Hz - 7.16kHz		
Low freq. corner	20-400Hz (2-bits tunable)		
High. freq. corner	6.25-12.6kHz (2-bits tunable)		
THD ($2mV_{pp}$ input)	0.04%		
Power Consumption	2.17μW (including biasing circuitry)		
NEF	1.75 (including biasing circuitry)		
PGA and Analog-to-Digital Converter			
Enabled Blocks	PGA+ADC	PGA+ADC	PGA
Sampling Rate	22.5kS/s	90kS/s	100kS/s
ENOB	7.98-bits	7.96-bits	7.98-bits
SNDR	49.8dB	49.68dB	49.8dB
Power Consumption	500nW	1.8μW	360nW
Window Comparator			
Time response	5μs		
Power Consumption	OFF	OFF	200nW
Neural Recording Channel			
Process	Standard CMOS 0.13μm		
Supply Voltage	1.2 V		
Overall Gain	47.5-65.5dB (3dB step)		
Power Consumption	2.8μW	4.1μW	2.9μW
Die Area	$0.160mm^2$ (400μm x 400μm)		

at the top is an exemplary synthesized spike sequence from Waveclus (http://www.vis.caltech.edu/~rodri/data.htm). The two sequences at the bottom of the figure illustrate the circuit operation at the two detected spikes. The first column shows the outputs of the band-limited LNA for both spikes, while the second column represents the outputs of the ADC operated (signal tracking mode). The last column shows the PWL reconstructions of the spikes when the channel is configured for feature extraction. Observe that the spike information to be transmitted is substantially reduced, the more as the duration of the spike increases. Moreover, taking into account that cortical neurons exhibit firing rates around 10 Hz and that no information is transmitted during unless a spike is detected, the feature extraction mode offers a substantial data reduction as compared to the signal tracking process.

Table I summarizes the performance of the presented neural spike recording channel. The power consumption depends on the selected operation mode. During signal tracking, only the band-limited LNA and the ADC with a throughput rate of 22.5kS/s are enabled, and the overall power consumption of the channel is 2.8μW, also including the dissipation of the biasing circuitry. During the feature extraction mode, the average power consumption is of abut 3μW, assuming that the spike detection phase takes the 95% of the total evaluation time, while the remaining 5% is employed for spike processing.

VI. Conclusions

A neural spike recording channel with feature extraction capabilities has been presented. The complete channel has been integrated in a CMOS standard 130nm technology and it occupies 400μmx400μm. The circuit is reconfigurable and offers different operation modes, including foreground self-calibration, signal tracking and feature extraction using first-order PWL approximations. Power-down strategies and frequency adaptation techniques are adopted in the chip for power saving. The channel consumes 2.8μW during signal tracking at 22.5kS/s, and 3.0μW average when performing feature extraction operation.

Acknowledgments

This work has been supported by the Spanish Ministry of Science & Innovation under grant TEC2009-08447, the Junta de Andalucía under grant TIC-02818 and the 2007-2013 FEDER Programme.

References

[1] 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(1), pp. 123-133, Jan. 2007.

[2] A. M. Sodagar, *et al*., "An Implantable 64-Channel Wireless Microsystem for Single-Unit Neural Recording," *IEEE J. Solid-State Circ.*, vol. 44(9), pp. 2591-2604, Sept. 2009.

[3] M.S. Chae, *et al*., "A 128-Channel 6mW Wireless Neural Recording IC With Spike Feature Extraction and UWB Transmitter," *IEEE T. on Neural Systems and Rehabilitation Eng.,* vol. 17(4), pp. 312-321, Aug. 2009.

[4] B. Gosselin, *et al*., "A Mixed-Signal Multichip Neural Recording Interface with Bandwidth Reduction," *IEEE T. Biomed. Circ. Sys.,* vol. 3(3), pp. 129-141. June 2009.

[5] A. Rodríguez-Pérez *et al*., "A Reconfigurable Neural Spike Recording Channel with Feature Extraction Capabilities," in *IEEE Biomedical Circ. Sys. Conf.*, Nov. 2010.

[6] B. Gosselin, M. Sawan and C. A. Chapman, "A low-power integrated bioamplifier with active low-frequency suppression", *IEEE T. Biomed. Circ. & Syst.*, vol. 1, pp. 184-192, Sep. 2007.

[7] R. R. Harrison and C. Charles, "A low-power low-noise CMOS amplifier for neural recording applications", *IEEE J. Solid-State Circ.*, vol. 38, no. 3, pp. 958-965, June 2003.

[8] W. Zhao, H. Li, Y. Zhang: "A Low-Noise Integrated Bioamplifier with Active DC Offset Suppression", *IEEE T. Biomed. Circ. & Syst.,* pp. 5-8, Nov. 2009.

[9] S. Rai *et al*., "A 500uW neural tag with 2uVrms AFE and frequency-multiplying MICS/ISM FSK transmitter," in *IEEE Int. Solid-State Circuits Conference* (ISSCC 2009), pp. 212-214, 2009.

[10] A. Rodríguez-Pérez, *et al.*, "A low-power reconfigurable ADC for biomedical sensor interfaces," in *IEEE Biomedical Circ. Sys. Conf. 2009,* pp. 253-256, Nov. 2009.