A 64-Channel Inductively-Powered Neural Recording Sensor Array

Alberto Rodríguez-Pérez, Jens Masuch, José A. Rodríguez-Rodríguez, Manuel Delgado-Restituto and Ángel Rodríguez-Vázquez
Institute of Microelectronics of Sevilla and University of Sevilla
Avda. Americo Vespucio s/n, 41092-Seville, SPAIN
Email: {alberto, masuch, rodriguez, mandel, angel}@imse-cnm.csic.es

***Abstract*—This paper reports a 64-channel inductively powered neural recording sensor array. Neural signals are acquired, filtered, digitized and compressed in the channels. Additionally, each channel implements a local auto-calibration mechanism which configures the transfer characteristics of the recording site. The system has two operation 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 transmitted. Data streams coming from the channels are serialized by an embedded digital processor and transferred to the outside by means of the same inductive link used for powering the system. Simulation results show that the power consumption of the complete system is 377μW.**

I. Introduction

In the last years, there has been a growing interest on the design of multichannel neural recording interfaces with wireless transmission capabilities for the untethered measurement of brain activity [1]–[3]. These interfaces are expected to play a significant role both in clinical (as part of therapeutic procedures in patients with neurological diseases), brain-machine interfaces and neuroscience applications. As these recording interfaces are implanted below the skull, the use of ultra-low power consumption techniques is mandatory, not only to prevent from harmful effects in the brain, but also to avoid the need for batteries. Thus, by making the power dissipation low, it becomes feasible to use energy harvesting strategies for supplying the implant. This is illustrated in Fig. 1 in which the intracranial device is powered via a wireless inductive link from an external unit placed on the head. The same link or a dedicated one could be also employed for data transfer to such external unit from where information could be communicated to a specific hub for compiling and processing the recorded brain activity.

This paper aims to contribute to this scenario and presents a multichannel wireless neural sensor array designed in a standard 0.13μm CMOS process. It is composed of 64 channels in which neural signals are acquired, filtered, digitized and optionally compressed [4]. The system has two transmission modes; in one case the information captured from a selected set of channels is transmitted as uncompressed raw data, in the other, feature vectors are extracted from the detected neural spikes at every channel and transmitted to the external unit for further processing. A single wireless inductive link, inspired in RFID technologies, is used both for powering

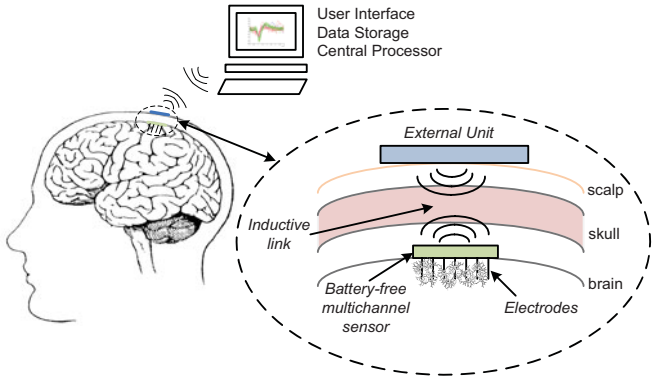

Figure 1. Implanted solution of the wireless neural array.

the implant and for data transfer to/from the external unit. This link uses a 40.68MHz carrier signal and employs On-Off Keying (OOK) modulation for data transfer from the external unit to the implant (forward link) and Load-Shift Keying (LSK) in the reverse direction (backward link). A 4MHz clock is used to send information through the backward link. This is enough for the implant operated in the feature extraction mode to characterize and serialize the detected spikes even in the unlikely case all the channels fire at the same instant. Post-layout simulations show that the total power consumption of the system, including the recording array and the communication protocol, is only 377μW, i.e., about one order of magnitude below prior art.

The paper is organized as follows. The architecture of the neural sensor is detailed in Section II. Section III presents the design of the RF front-end, while the simulation results are given in Section IV. Finally, Section V ends the paper with some conclusions.

II. Neural Sensor Architecture

Fig. 2 shows the architecture of the proposed system. It consists of a 8x8 neural recording array, each of them serially connected to an Event-Based Processor Unit (EBPU), which stores the information generated by the channel. The data stored in these EBPUs are read and classified by an embedded digital processor, which also handles the timing of the implant. A communication block implements the link to/from the external unit. Additionally, the system includes one

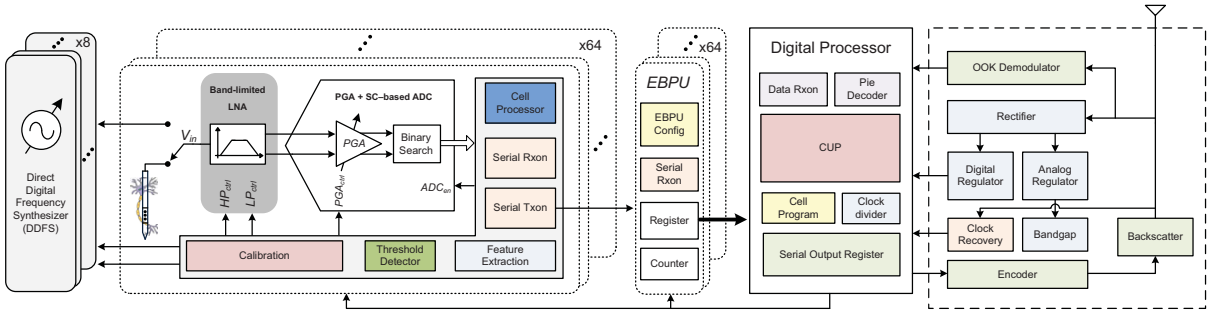

Figure 2. Architecture of the multichannel neural array.

tunable Direct Digital Frequency Synthesizer (DDFS) per row for calibration purposes [5].

Each channel embeds all the needed circuitry to acquire and digitize neural waveforms including a Low Noise Amplifier (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 channel architecture is similar to that in [4] but, in this version, spike detection is accomplished in digital domain and the decision threshold is adaptively updated according to the noise floor of the captured signal. Further, in order to increase the granularity of the calibration process, three control bits are used to adjust the high-pass pole of the bandpass filter.

A. Modes of operation

Together with the two already mentioned transmission modes, denoted as signal tracking and feature extraction modes, the system also offers a foreground calibration mode. They are briefly described next.

Calibration: In this mode, the transfer characteristic and gain of the recording channels are individually adjusted. This is done by sequentially adjusting the pass-band of the filters and the gain of the PGAs using the algorithm in [5]. First, the programming words for the high-pass (3-bit) and low-pass (2-bit) poles of the channel bandpass filter are tuned so that its passband ranges from about 200Hz to 7kHz, corresponding to the spike spectral range. This is done for every channel by using the output signals of the DDFSs as frequency references. As there is one DDFS per row, passband calibration is done in a column-wise manner. Afterward, every channel starts capturing neural signals at a rate of 27kS/s and the gain of each PGA is adjusted so that its output fits into the input dynamic range of the corresponding ADC. Digitized signals are transmitted out column by column so that an external observer validates the completion of the calibration process. This is done because neural spiking is random by nature and channels can be silent for long periods. After validation, the observer can change to a different column or finish the calibration process by applying corresponding commands.

Signal Tracking: In this mode, one column/row of the array is arbitrarily selected for neural signal monitoring while remaining channels are disabled for power saving. Neural signals are acquired at a sampling rate of 27kS/s, 8-bit per sample, to give an overall throughput rate of 1.92Mbps. No data compression is applied in this mode.

Feature Extraction: In this case, the system is employed for spike detection tasks. All the 64 channels are enabled during feature extraction. Every detected spike is locally compressed at channel level by means of a Piece-Wise Linear (PWL) approximation of its waveform. This approximation involves amplitude and time interval values, an results in a 47-bit representation per spike, enough for sorting and clustering purposes [4]. During the characterization of the spike the channel operates at a sampling rate of 90kS/s.

B. Event-Based Communication

EBPU units are the responsible for temporarily storing the information provided by the channels. In the calibration and signal tracking modes, channels serialize and transfer data to the EBPUs, where information is retained until it is read out by the system digital processor. In the feature extraction mode, EBPUs not only provide storing resources but also contribute on the calculation of the time intervals involved in the PWL representation of spikes. Peaking and threshold crossing events along spikes are transmitted to corresponding EBPUs. Such units keep track of the duration between the events by means of counters. When spikes end, channels send to the EBPUs the amplitude related information to complete the associated PWL feature vectors. Once vectors are gathered, they are stored in the EBPUs ready to read out. It is worth observing this approach reduces the information transfer from the channels to the EBPUs by about 50%, as single events instead of complete time interval measurements (coded in 8-bit words) are transmitted.

The main digital processor cyclically reads the enabled EBPUs. If it is found the stored information in the EBPU is complete, the digital processor retrieves data at a 4MHz rate, builds up the transmission frame and sends this stream to the telemetry unit for wireless transmission.

C. Communication Protocol

Similar to RFID technologies, the system uses Pulse Interval Encoding (PIE) of symbols in the forward link. Fig. 3(a) shows the symbol representations for data-0 and data-1, which essentially differ on the duration of the high-level state.

Fig. 3(b) illustrates the structure of data frames in the forward link, i.e. towards the sensor array. They are used to configure the neural recording sensor array. A forward frame

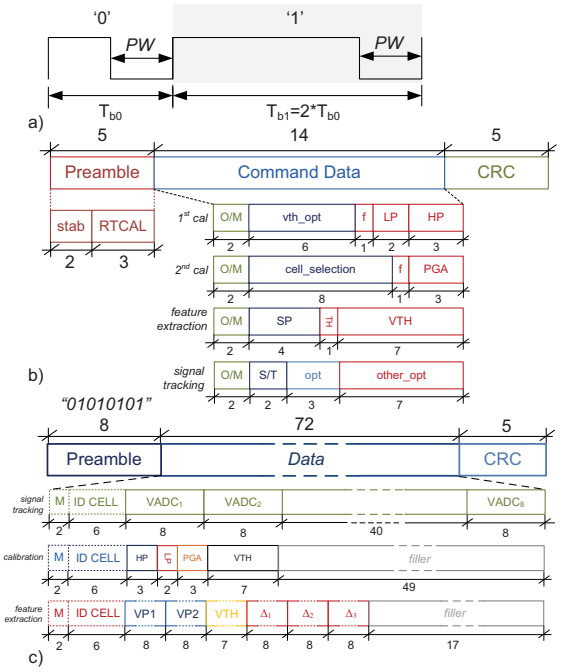

Figure 3. Communication protocol of the proposed system: a) PIE format, b) forward frame, c) backward frame.

consists of 24-bit, including preamble (5 bit), command (14 bit) and cyclic redundancy check (CRC) word (5 bit). As shown in Fig. 3(b), the structure and parameters included in the command word depends on the selected operation mode.

Fig. 3(c) shows the structure of data frames in the backward link, i.e. from the sensor array to the outside. The backward frame is 85-bit long and includes a fixed 8-bit preamble "01010101", followed by a 72-bit output data set, and completed by a 5-bit CRC word. The first 8-bit of the output data set inform about the operation mode (2-bit) and the channel identification (6-bit). In the signal tracking mode, the system collects the sampled data in groups of eight (by column or row, depending on the selected option), and only the first channel of the column/row has to be identified. In the feature extraction mode, the output data set is formed by three bytes of temporal information, two bytes of amplitude information and 7-bit representing the applied threshold voltage. In the calibration mode, the system generates 15-bit which inform on the settings for the bandpass filter, PGA and threshold voltage.

III. Telemetry Unit

Fig. 4 shows the schematics of the power and data telemetry unit. It is based on inductive link techniques and operates in the worldwide available ISM band centered at 40.68MHz. Data reception employs (OOK) modulation whereas data transmission is accomplished by modulating the amplitude of the carrier by means of a switchable antenna matching network driven by the digital processor. In this latter case, the modulation depth is less than 50% and the output data is encoded using a Manchester encoder.

Not shown in the figure, the telemetry unit also includes a timing recovery circuit which extracts the 4MHz clock of

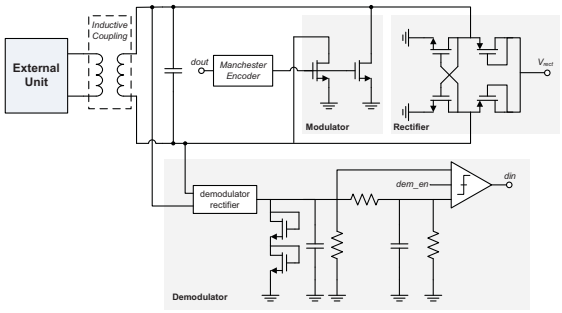

Figure 4. Schematic of the telemetry front-end.

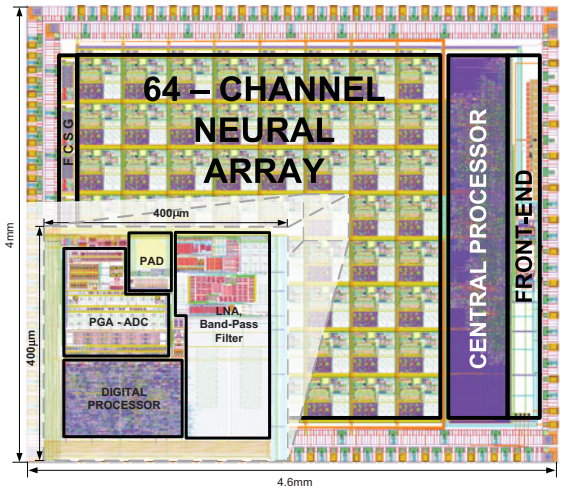

Figure 5. Layout of the multichannel neural sensor.

the system from the incoming RF signal, which is also used to modulate the backward link. This is accomplished by means of divide by 2 circuits based on single-transistor-clocked dynamic latches [6].

The telemetry unit also includes a power management circuitry which harvest energy from the inductive link using a rectifier. Analog and digital supply lines of 1.2V are obtained from corresponding regulators, while a bandgap circuit generates the analog voltage references. The efficiency of the rectifier is 60% at 1mW RF input power.

IV. Post-Layout Results

Fig. 5 shows the layout of the proposed system. It has been designed in a 6M2P 0.13μm standard CMOS technology. Each channel includes an internal pad for flip-chip connection to a microelectrode. For the sake of testability, the channel input nodes can be also accessed from an external padring. Clamp cells are placed along the chip periphery to protect the microelectrode nodes from ESD damages. The system occupies 18.4mm-sq.

Fig. 6 illustrates the operation of the adaptive threshold algorithm implemented in the local digital processor of the channels. The signal-to-noise ratio of the neural signal has been intentionally varied to better appreciate the evolution of the threshold detection level. As can be seen, the algorithm reacts in less than 0.5s to changes in the background noise.

Fig. 7 illustrates the system operation in the feature extraction mode. Dots represent the spikes detected by the neural

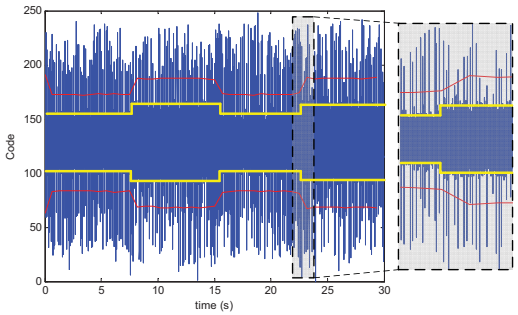

Figure 6. Adaptive threshold voltage algorithm: Neural signal (blue), voltage threshold (red), noise level (yellow).

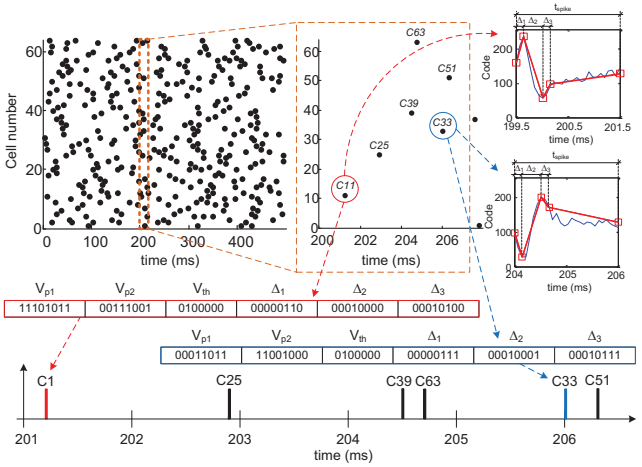

Figure 7. Data output stream under feature extraction mode.

array in a time slot of 500ms. Once a spike is detected in a channel and its PWL representation derived (47-bits, as figure 7 illustrates), the feature vector is stored in the associated EBPU. The main digital processor cyclically reads the EBPUs every 237μs. Considering the 85-bit length of the backward frame detailed in Section II, the system requires 21.25μs to transmit the information of one spike at 4MHz. Therefore, we can calculate the maximum possible delay by summing up the delay of the EBPU reading and the transmission delay, which results 258.25μs. This is much lower than a typical spike duration (around 2ms) and, of course, much lower than the time basis for firing occurrences. It means, that no information is lost not even in the unlikely case all the channels fire at the same instant (only a small delay no larger than about 10% the duration of a spike could be observed in some of the records).

The performance of most of the blocks comprised in the channels (LNA, filter and ADC) were measured and reported in [4]. The new channel implementation in this paper also includes an additional digital processor which, together with the needed buffers to communicate along the array, rise the power consumption per channel to 4.54μW. From the simulated power consumption it can be extrated that most of the power is consumed by the neural channels (290.56μW). The main digital processor and EBPUs, which make extensive use of clock gating and clock frequency division techniques, consumes 40μW (5μW of them dissipated by leakage currents). Bandgap references, regulators and current conveyors

Table I
PERFORMANCE SUMMARY AND COMPARISON

	[1]	[2]	[3]	This work
Technology (μm)	0.18	0.18	0.13	0.13
Supply voltage (V)	1.8	1.8 / 1	0.5	1.2
Number of channels	16	32	16	64
Total power (μW)	680	325	18	377
Power / channel (μW)	42.5	10.1	1.13	5.9
High pass freq. (Hz)	100	350	400	200
Low pass freq. (kHz)	9.2	12	7.5	6.9
Input ref. noise (μV_{rms})	5.4	5.4	5.32	3.8
NEF	4.9	4.4	3.09	2.16
ENOB (bits)	7	7.65	7.32	7.65
Sampling freq. (kS/s)	30	31.25	30	27 / 90
Data bitrate reduction	Yes	No	No	Yes

consume 32μW. The clock recovery block, the Manchester encoder and the demodulator require, respectively, 12.5μW, 1.5μW and 400nW. All in all, the total power consumption of the system sums 377μW.

Table I summarizes the performance of the neural recording system and compares it with some state-of-the-art works. Note that the presented work presents one of the lowest power dissipation per channel, even though it is the only one that includes a wireless communication circuitry.

V. CONCLUSIONS

A 64-channel neural array with embedded data reduction techniques, fabricated in a standard CMOS 130nm process, has been presented. Inspired by RFID systems, an inductive link is used for both powering the implant and transferring information to/from an external unit placed on the head. 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. The total power consumption of the system has been estimated in 377μW from a nominal voltage supply of 1.2V, i.e., about one order of magnitude below prior art.

ACKNOWLEDGMENTS

This work has been supported by the Spanish Ministry of Science & Innovation under grant TEC2009-08447 and the 2007-2013 FEDER Program.

REFERENCES

[1] B. Gosselin *et al.*, "A mixed-signal multichip neural recording interface with bandwidth reduction," *IEEE Trans. Biomed. Circuits*, vol. 3, no. 3, pp. 129–141, 2009.

[2] W. Wattanapanitch and R. Sarpeshkar, "A low-power 32-channel digitally programmable neural recording integrated circuit," *Biomedical Circuits and Systems, IEEE Transactions on*, vol. 5, no. 6, pp. 592–602, 2011.

[3] L. Wen-Sin, Z. Xiaodan, and L. Yong, "A 0.5-v 1.13uw/channel neural recording interface with digital multiplexing scheme," in *ESSCIRC (ESSCIRC), 2011 Proceedings of the*, pp. 219–222.

[4] A. Rodriguez-Perez, J. Ruiz-Amaya, M. Delgado-Restituto, and A. Rodriguez-Vazquez, "A low-power programmable neural spike detection channel with embedded calibration and data compression," *IEEE Trans. Biomed. Circuits*, vol. 6, no. 2, pp. 87 –100, april 2012.

[5] A. Rodriguez-Perez *et al.*, "A self-calibration circuit for a neural spike recording channel," in *Proc. IEEE Biomed. Circ. and Systems Conf.*, nov. 2011, pp. 464 –467.

[6] J. Yuan and C. Svensson, "New single-clock cmos latches and flipflops with improved speed and power savings," *IEEE J. Solid-State Circ.*, vol. 32, no. 1, pp. 62–69, 1997.