Bio-signal Data Conditioning Optimization Method for Hardware Applications

Lijun Chen
Electrical and Electronics Engineering Section
EPFL
Lausanne, Switzerland
Email: lijun.chen@epfl.ch

Linghan Zhao
Electrical and Electronics Engineering Section
EPFL
Lausanne, Switzerland
Email: linghan.zhao@epfl.ch

***Abstract*—The development of neural recording systems for high-density, implantable applications has advanced through innovations in feature extraction, spike detection, and data management techniques. Kamboh and Mason (2010) introduced on-chip feature extraction for spike sorting to simplify processing. Rodriguez-Perez et al. (2012) designed a low-power spike detection channel with embedded calibration and compression. Gosselin et al. (2009) developed a mixed-signal multichip interface to address bandwidth constraints. These innovations enhance the performance of implantable neural interfaces.**

***Index Terms*—Neural recording, spike detection, feature extraction, low power, data compression**

I. Introduction

In recent years, implantable neural recording systems have emerged as a significant focus in biomedical engineering, driven by advances in neuroscience, electronics, and computational technologies. These systems aim to capture neural signals to help researchers decode the functioning of neural networks, diagnose neurological disorders, or develop neural prosthetics. Such systems must support high-density, multi-channel recording to capture the brain's complex neural activity while balancing constraints such as power consumption, size, and signal quality to ensure long-term feasibility and stability for implantation.

However, although with the development of micro-electronics, this field has made many progresses, it also faces several technical challenges.

A. Spike Sorting

Spike sorting is a critical step in classifying neural signals into distinct neuron spikes, which is crucial for accurate neural monitoring and interaction with prosthetic devices. But implementing spike sorting on implantable devices is challenging due to their limited processing power and the need for real-time analysis. The primary challenge is how to perform efficient on-chip feature extraction for spike sorting while minimizing computational complexity and resource requirements.

B. Power Consumption

Implantable neural interfaces are often constrained by power consumption, as long-term operation within the human body requires energy-efficient designs. Additionally, the large amount of data generated by neural recordings requires effective data management, including compression and calibration, to ensure that signals are accurately detected and transmitted with minimal energy consumption. The challenge here is designing a low-power, programmable neural spike detection channel that can operate efficiently while integrated with embedded calibration and data compression.

C. Band-width Limitation

Neural recording devices require handling multiple channels, resulting in large data bandwidth, especially those with high-density electrode arrays, the amount of data generated can be overwhelming. For implantable systems, this data must be processed and transmitted with limited bandwidth. The main difficulty is that reducing the data rate while still preserving the quality and accuracy of the recorded neural signals.

II. Spike Sorting System

A. Spike Detection

For spike detection, the normal methods to use are data transformations, derivatives and template matching. But for these methods often required intense computation. A more simple way to detect spike is to set two threshold values, one for plus magnitude and the other for minus magnitude. It's called double threshold (DT) to detect spike signals, they are shown in Fig. 1 with red lines.

B. Feature Selection and Extraction

Traditionally, PCA and Wavelet Transforms (WT) are comment way to sort spikes. However for on-chip multi-channel feature extraction, they seem to be unsuitable. Here we just simply use the accumulation of amplitude between the zero crossing points, as:

$$ZC_1 = \sum_{n=0}^{Z-1} x(n); \qquad ZC_2 = \sum_{n=Z}^{K-1} x(n)$$

where ZC_1 and ZC_2 are the two features, K is the number of samples in a spike and Z is the index of first zero crossing after the spike has been detected. The value of Z is thus the first zero crossing after a significant amount of energy in the spike has already been recorded.

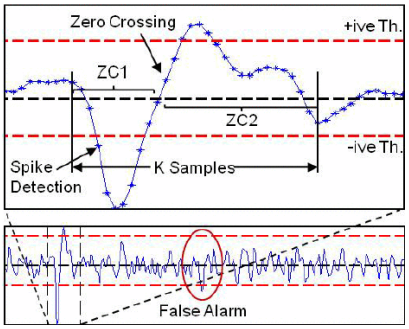


Fig. 1. Graphical representation of ZCF features for a spike detected using dual thresholds. [1]

Fig. 1 shows that the spike is detected when the electrode potential crosses the negative threshold. It should be mentioned that there are still some samples before the detection that belong to the spike and may contain useful information and thus need to be stored in a buffer. The hardware structure could be found in Fig. 2, where light colored blocks represent memory elements.

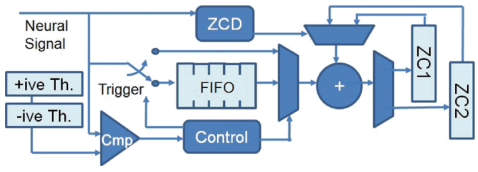


Fig. 2. Hardware architecture of proposed signal detection and feature extraction stages. [1]

Obviously, there are some strengths of this system. Firstly, it can realize low power consumption while have a acceptable accuracy of sorting. The feature extraction and classification algorithms are optimized for implantable systems, minimizing floating-point computations and hardware resource usage to save power, making it ideal for long-term neural implants. Besides, the circuit design reduces dependence on complex memory and computational resources, saving chip area, which is critical for implantable devices.

However, those strengths highly depend on the signal quality. Fixed threshold-based detection is susceptible to background noise and artifacts, leading to missed spikes or false positives in noisy environments. And while effective, features like peak amplitude and PCA may be insufficient for distinguishing complex or highly similar spike waveforms, limiting classification accuracy.

III. Low Power Spike Process System

A. System Architecture

In [2], a sophisticated designed system was proposed for neural signal processing. As the Fig. 3 shows, the raw signal process rough flow is:

1) Signal Acquisition: First, neural signals are captured by microelectrode arrays (MEAs) placed near neurons. A bandpass low-noise amplifier (BP-LNA) boosts weak neural signals while minimizing noise, at the same time isolates the neural spike frequency range (commonly 300 Hz to 3 kHz, could be adjusted).
2) Spike Detection: In detection, the system compares the signal amplitude against a pre-defined or dynamically adjusted threshold. If the signal exceeds the threshold, it is classified as a spike.
3) Spike Feature Extraction: Extracts features of each detected spike. Rather than record the complete signal, the system only collect some key features of the raw signal shown in Fig. 4.
4) Transmission or Storage: Processed spike data is sent to an external device (e.g., for storage, analysis, or real-time use in closed-loop systems).

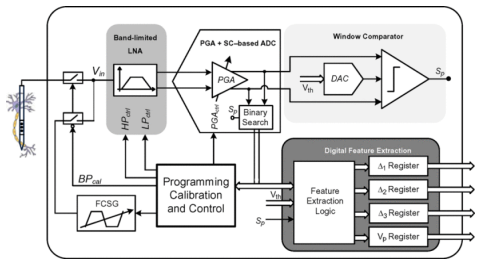
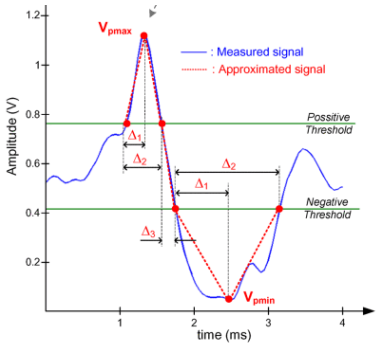


Fig. 3. Block diagram of the neural channel architecture. [2]



	V_{peak}	Δ_1	Δ_2
Possitive peak	"11101110" (1.12 V)	"00010010" (180 us)	"00111110" (620 us)
Negative peak	"00001010" (49 mV)	"01010010" (820 us)	"10000011" (1310 us)

Fig. 4. Piece-Wise Linear spike approximation. [2]

B. Low Power Design Consideration

With the limited power supply and heat generation, low power design should be considered in each part of the system. Overall, the low-power design in the system is achieved through modular design, application-specific integrated circuit (ASIC) implementation and spike-specified processing. Specifically, the design consideration are:

- Modular Working
  For the task of spike processing, it is not necessary to utilize all components of the system. For practice, the digital core can point out which blocks remain idle at each operation mode and powered them off, either totally or

partially, by means of power gating techniques. It can also incorporate clock division and clock gating strategies, so that non-idle blocks work at the minimum clock frequency able to comply with the selected operation mode.

- Spike-specified Processing
  Rather than always standby to analyze the spikes, the system will start spike processing mode and analyze when the input signal reaches a specific threshold value. Only detected spikes are processed further to extract key features (e.g., amplitude, width, and timestamp), avoiding computations on irrelevant data.
- Spike-optimized Circuit
  The BP-LNA and DAC parts are specified for the voltage swing and frequency of spike biosignal. For example, the employed DAC part utilizes a 6-bit R-2R resistive current divider structure [3], which achieves requirements by very low area occupation and a minimum power consumption.

C. System Evaluation

Overall, this novel system offers a way to collect and process spike biosignal. There are significant advantages of the system:

TABLE I
STATE-OF-ART PERFORMANCE COMPARISON. [2]

	[4]	[5]	[6]	This Work
Supply Voltage (V)	1.5	1.8	3	1.2
Sampling Rate (kS/s)	62.5	30	-	90
Midband Gain (dB)	59.5	70	67.8-78	47.5-65.5
Power (uW)	24.6	42.5	75	3.1

- Low power consumption
  The complete channel consumes 2.8 uW at 1.2 V voltage supply when operated in the signal tracking mode, and 3.1 uW when the feature extraction mode is enabled. This power consumption is standout among other designs while retaining other performances as table I.
- Programmability
  The system features programmable elements such as gain control in the BP-LNA and dynamic threshold adjustment in the spike detection module. This flexibility makes the design adaptable to a wide range of neural recording conditions and electrode arrays.

However, the simplification of spike signal and data storage could also bring some limitations to the system:

- Limited Feature Extraction
  The system focuses only on basic spike features such as amplitude and timestamp, which might limit its utility for applications requiring more detailed spike shape information or advanced classification algorithms.
- Simplistic Spike Detection Algorithm
  The threshold-based spike detection mechanism, while power-efficient, may struggle with overlapping spikes or spikes with low signal-to-noise ratios (SNR).

IV. BANDWIDTH REDUCTION

A. Mixed-Signal Architecture

The prototype is composed of a mixed-signal front-end IC and a digital processing and control IC.

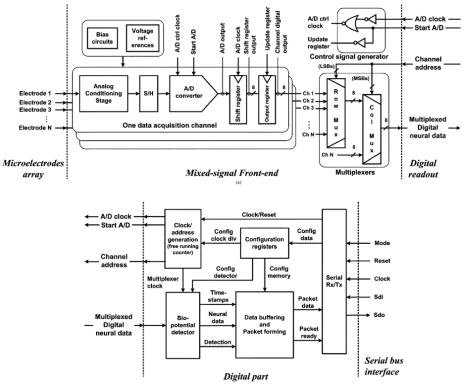


Fig. 5. Block diagram of the multichip neural interface.[5]

In Fig.5, the mixed-signal front end is composed of 16 fully parallel data-acquisition channels including conditioning circuits, data converters, and output registers. Each channel is supplied with its own DAQ chain, which is scalable, and provides signal conditioning, sampling, digitization, and each channel output is multiplexed toward a digital readout with channel decoding.

The parallel acquisition approach enables scalability for various topologies, supports true simultaneous sampling for precise neural code interpretation, and eliminates analog multiplexing, reducing noise, crosstalk, and power consumption.

The digitized samples undergo time-division multiplexing (TDM) toward a digital readout which interfaces the digital ASIC. The digital part provides control and enables communications with a remote host interface through a four-wire serial bus. It also provides bandwidth reduction to avoid overwhelming data rates that result when recording from a large amount of channels.

B. Action Potential Detection

AP detection identifies neural spikes using an absolute value threshold detector and adaptive data buffering to capture complete biopotential waveforms. Figure 6 illustrates SRAM usage during neuronal AP detection. Initially, the first 16 samples are buffered in a FIFO and copied into the DPRAM buffer upon detection. Subsequently, the remaining 48 samples are transferred to memory, forming a 64-sample data packet that is then delivered on the bus. The FIFO continuously buffers incoming samples to ensure that waveform segments preceding the detection point are not missed, preserving the complete detected AP for storage in the DPRAM.

When bandwidth reduction is enabled, the detector compares each sample's amplitude against a programmable threshold (3-5 standard deviations of input voltage). Upon threshold crossing, a finite state machine (FSM in Fig.7) triggers waveform capture, storing detected events, timestamps, and

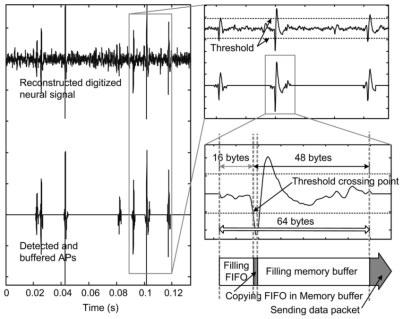


Fig. 6. Detection/buffering of digitized AP waveforms.[5]

channel addresses in on-chip memory. The memory buffers 64-sample windows (2 ms), ensuring full waveform preservation and efficient data reduction while maintaining signal integrity. Only detected APs are transmitted, eliminating the need to send raw data.

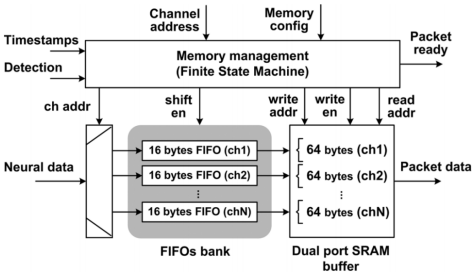


Fig. 7. Schematic of the data buffering block.[5]

This AP detection presents an excellent tradeoff between effectiveness and computational complexity, and uses bilateral threshold detectors. The neural interface combines absolute thresholding with data buffering to achieve a bandwidth reduction factor of up to 50 within typical recordings, and decreases the raw neural data rate by 98%. This strategy enhances energy efficiency and reduces data transmission demands without compromising the quality of neural information.

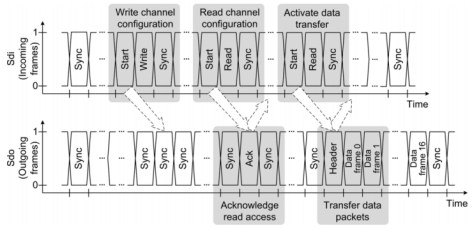


Fig. 8. Communication protocol employed between the multichip interface and the remote host. The five types of transactions allowed are shown.[5]

Then the digital ASIC optimizes neural data processing and transmission. It integrates a configurable communication protocol(as in Fig.8), enabling dynamic adjustment of thresholds and parameters to adapt to various electrode configurations and applications. Data is structured into frames and packets, containing metadata (channel ID, timestamp) and compressed neural data, ensuring efficient transmission and seamless decoding. To maintain real-time operation, the system minimizes processing delays and uses synchronized clocks across all channels, preserving the temporal accuracy of action potentials.

C. Vertical Integration

The multichip neural interface employs a vertical integration approach, implemented on the back of the epoxy base of a metallic array to mount the ASICs and passive components.

1) Design and Fabrication of the Microelectrode Array: The microelectrode array is fabricated from medical-grade stainless steel using precision micromachining techniques. Electrodes are arranged in a 4x4 matrix, with each pin shaped and polished for smooth surfaces and biocompatibility. The electrodes are electrically isolated using biocompatible epoxy, ensuring accurate and independent recordings.

2) Postprocessing of the Array Base: Conductive metal traces (chromium and gold) are applied to the array base using photolithography and metal deposition techniques. These traces serve as electrical connections between the electrodes and the stacked integrated circuits (ICs).

3) Vertical Stacking of ASICs: Two application-specific integrated circuits (ASICs), a Mixed-Signal ASIC and a Digital ASIC, are stacked vertically. The chips are connected and protected using wire bonding and epoxy layers.

4) Compact and Biocompatible Design: The entire stacked system occupies a 2.3 mm^2 footprint for a 16-channel system and is designed to fit within the subarachnoid space above the cortex. A Parylene C coating is applied for electrical insulation and biocompatibility, leaving the electrode tips exposed for neural recording.

V. Conclusion

All references address critical challenges in neural spike processing for implantable systems, focusing on low-power, efficient designs suitable for high-density neural recording. Kamboh and Mason's work in [1] emphasizes real-time on-chip feature extraction and classification with simplicity and scalability, making it ideal for resource-constrained environments. Meanwhile, Rodriguez-Perez et al. in [2] introduce advanced programmability with embedded calibration and data compression, enhancing adaptability and data handling efficiency.

The proposed mixed-signal neural recording interface integrates efficient bandwidth reduction techniques, and compact chip design. However, it faces limitations in noise resilience and real-world performance, as most testing relied on synthetic signals, necessitating more in vivo validation. The design complexity of mixed-signal integration and vertical stacking increases manufacturing challenges, and high firing rates or neural seizures could overwhelm memory buffers, risking data loss.

These works represent a significant advancement in neural recording technologies, offering promising solutions for future brain-machine interfaces and neuroscience research.

REFERENCES

[1] A. M. Kamboh and A. J. Mason, "On-chip feature extraction for spike sorting in high density implantable neural recording systems," 2010 Biomedical Circuits and Systems Conference (BioCAS), Paphos, Cyprus, 2010, pp. 13-16, doi: 10.1109/BIOCAS.2010.5709559.

[2] A. Rodriguez-Perez, J. Ruiz-Amaya, M. Delgado-Restituto and Á. Rodriguez-Vazquez, "A Low-Power Programmable Neural Spike Detection Channel With Embedded Calibration and Data Compression," in IEEE Transactions on Biomedical Circuits and Systems, vol. 6, no. 2, pp. 87-100, April 2012, doi: 10.1109/TBCAS.2012.2187352.

[3] R. J. Kier, R. R. Harrison and R. D. Beer, "An MDAC synapse for analog neural networks," 2004 IEEE International Symposium on Circuits and Systems (IEEE Cat. No.04CH37512), Vancouver, BC, Canada, 2004, pp. V-V, doi: 10.1109/ISCAS.2004.1329917.

[4] A. M. Sodagar, G. E. Perlin, Y. Yao, K. Najafi and K. D. Wise, "An Implantable 64-Channel Wireless Microsystem for Single-Unit Neural Recording," in IEEE Journal of Solid-State Circuits, vol. 44, no. 9, pp. 2591-2604, Sept. 2009, doi: 10.1109/JSSC.2009.2023159.

[5] B. Gosselin et al., "A Mixed-Signal Multichip Neural Recording Interface With Bandwidth Reduction," in IEEE Transactions on Biomedical Circuits and Systems, vol. 3, no. 3, pp. 129-141, June 2009, doi: 10.1109/TBCAS.2009.2013718.

[6] S. B. Lee, H. -M. Lee, M. Kiani, U. -M. Jow and M. Ghovanloo, "An inductively powered scalable 32-channel wireless neural recording system-on-a-chip for neuroscience applications," 2010 IEEE International Solid-State Circuits Conference - (ISSCC), San Francisco, CA, USA, 2010, pp. 120-121, doi: 10.1109/ISSCC.2010.5434028.