Journal of
Low Power Electronics
Vol. 12, 1–9, 2016

Power-Efficiency of Signal Processing Circuits in Implantable Multichannel Brain-Machine Interface

Amir Zjajo

Circuits and Systems Group, Delft University of Technology, Delft, 2628 CD, The Netherlands

(Received: 28 June 2016; Accepted: 4 October 2016)

The nature of the neural signals, increasing density in multichannel arrays, information quality, and feasible data bandwidth pose significant challenges encountered in a power-efficient design of implantable brain-machine interface. In this paper, we propose a set of solutions to address this design problem at both circuit- and system abstraction level. In particular, we review circuits for real time read-out of neural signals and discuss the role of classification in hardware neural processing architectures; we review the challenges of realizing power-efficient circuits in physical systems and present examples of mixed-signal electronic circuits that implement them; we provide a broad view of optimization approaches, and their possible combination in effective complimentary techniques. We validate the approach with experimental results obtained from our own circuits and systems, and argue how the circuits and systems presented in this work represent a valid set of components for power-efficient design of implantable multichannel brain-machine interface.

Keywords: Brain-Machine Interface, Neural Recording Interface, Power Optimization.

1. INTRODUCTION

Brain machine interface (BMI) enable the interaction with neural cells, either by recording (to facilitate early diagnosis and predict intended behavior before undertaking any preventive or corrective actions), or by stimulation (to prevent the onset of detrimental neural activity). Monitoring the activity of a large population of neurons with high-density microelectrode arrays in multichannel BMI is a prerequisite for understanding the cortical structures and can lead to a better conception of stark brain disorders, such as Alzheimer's and Parkinson's diseases, epilepsy and autism,[1] or can aid to reestablish sensory (e.g., hearing and vision) and motor (e.g., movement and speech) functions.[1] One of the main goals of the current neural probe technologies[2] is to minimize the size of the implants while including as many recording sites as possible, with high spatial resolution. This enables the fabrication of devices that match the feature size and density of neural circuits,[3] and facilitates the spike classification process.[4] Since electrical recording from single neurons is invasive, monitoring large numbers of neurons using large implanted devices inevitably increases the tissue damage. Although existing neural probes can record from many neurons, the limitations in the interconnect technology constrains the number of recording sites that can be routed out of the probe.[6]

The study of highly localized neural activity requires, besides implantable microelectrodes, electronic circuitry for accurately amplifying and conditioning the signals detected at the recording sites. While neural probes have become more compact and denser in order to monitor large populations of neurons, the interfacing electronic circuits have also become smaller and more capable of handling large amounts of parallel recording channels. Some of the challenges in the design of analog front-end circuits for neural recording are associated with the nature of the neural signals; the recording circuits have to be designed with sufficiently low input-referred noise (i.e., to achieve a high signal-to-noise ratio (SNR)) and sufficient gain and dynamic range. The raw data rates that are generated by simultaneous monitoring of hundreds and even thousands of neurons are large.[7] Communicating large volumes of neuronal data over battery-powered wireless links, while maintaining reasonable battery life, is hardly possible with common methods of low-power wireless communications. Evidently, some form of data reduction or lossy data compression to reduce the raw waveform data capacity, e.g., wavelet transform,[8] must be applied. Alternatively, only significant features of the neuronal signal could be extracted and the transmitted data

Email: amir.zjajo@ieee.org

could be limited to those features only,[6] which may lead to an order of magnitude reduction in the required data rate.[9] Additionally, if the neuronal spikes are classified on the chip,[10] and mere notifications of spike events are transmitted to the host, another order of magnitude reduction can be achieved. Adapting power-efficient spike sorting algorithms can yet lead to significant power savings, with only a limited accuracy loss.[11]

In this paper, we address design challenges in respect to circuit miniaturization and power reduction of the neural recording system, along with circuit topologies, architecture trends, and (post-silicon) circuit optimization algorithms. Power-efficiency is the main focus of both, the circuit design (including circuits for signal conditioning, quantization and classification), as well as the system design: design variables include the number of obtainable channels, feasible data rate, and available (battery) power of the implantable system.

The paper is organized as follows: Section 2 focuses on the architectural overview of a multi-channel neural recording interface, circuit parameters formulation, and associated process variability and noise. Section 3 discusses power-efficiency of analog front-end signal conditioning and quantization circuits and its main limitations. In Section 4, characterization of power-efficiency of the digital back-end signal detection and classification circuits is presented. Finally, Section 5 provides a summary and the main conclusions.

2. IMPLANTABLE MULTICHANNEL BRAIN-MACHINE INTERFACE

2.1. Architectural Overview of the Brain-Machine Interface

The block diagram of a M-channel brain-machine interface is illustrated in Figure 1. With an increase in the range of applications and their functionalities, neuroprosthetic devices are evolving to a closed-loop control system composed of a front-end neural recording interface and a back-end neural-signal processing, containing features such as spike detection and classification circuits. The neural data acquired by the recording electrodes is conditioned using analog circuits. The low-noise amplification (LNA), band-pass filtering, and programmable gain amplification (PGA) (Fig. 2) of the neural signals is performed before the signals are quantized by an A/D converter (ADC) (Fig. 3). The A/D converter output containing the time-multiplexed neural signals is fed to a back-end digital signal processing unit, which provides additional filtering and executes a spike detection and classification. The relevant information is then utilized for K-channel brain stimulation in a closed-loop framework, or alternatively, transmitted to an outside receiver for off-line processing.

2.2. Neural Signals

The neural signals have amplitude in the order of a few μV to several mV and frequency spans from *dc* to a few kHz. Local-field potentials (LFP), representing averaged activity from small sets of neurons surrounding the recording sites, are in the low-frequency range (~0.5–300 Hz), while action potentials (AP) or spikes, indicating single-cell activity, are located in the higher-frequency range (~300–20 kHz). The test dataset in Figure 4 illustrate recordings from the human neocortex and basal ganglia. In Figure 5, we illustrate statistical voltage trace of a neuron signal composed of a spike burst and biological noise. Recording both local-field potentials and action potentials using implanted electrodes yields the most informative signals for studying neuronal communication and computation. Consequently, the recording circuits have to be designed with sufficiently low input-referred noise (i.e., to achieve a high signal-to-noise ratio (SNR)) and sufficient gain and dynamic range.

In the Hodgkin and Huxley neural cell model, a configuration of a neural channel is predisposed by the states of its constituent subunits, where each subunit can be either in an open or closed state. Adding a noise term to this formulation is consistent with the behavior of the Markov process for channel gating. Subsequently, the neural cell noise is modelled as Brownian motion, i.e., as a Gauss-distributed nonstationary stochastic process with independent increments and heuristically fixed constant variance.[12] In intra-cortical microelectrode recordings, neural cell noise mainly originates from the firing of several neurons in the tissue surrounding the recording microelectrode, while thermal noise levels are influenced by the electrode-tissue interface impedance in each individual recording site and the recording bandwidth, i.e., a 36 μm diameter probe (1000 μm^2) may have a capacitance of 200 pF, equivalent to ~40 kΩ impedance at

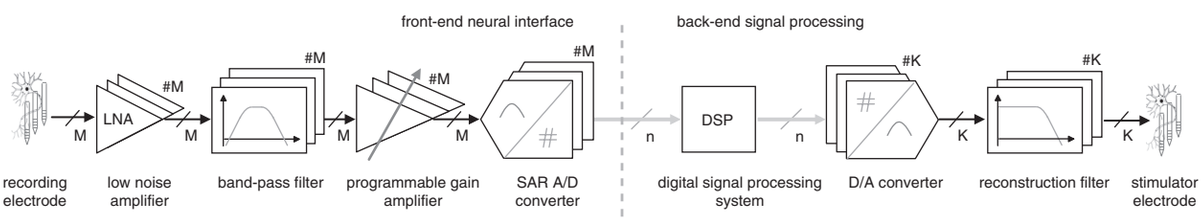

Fig. 1. Block diagram of a brain machine interface with M-channel front-end neural recording interface and K-channel back-end signal processing.

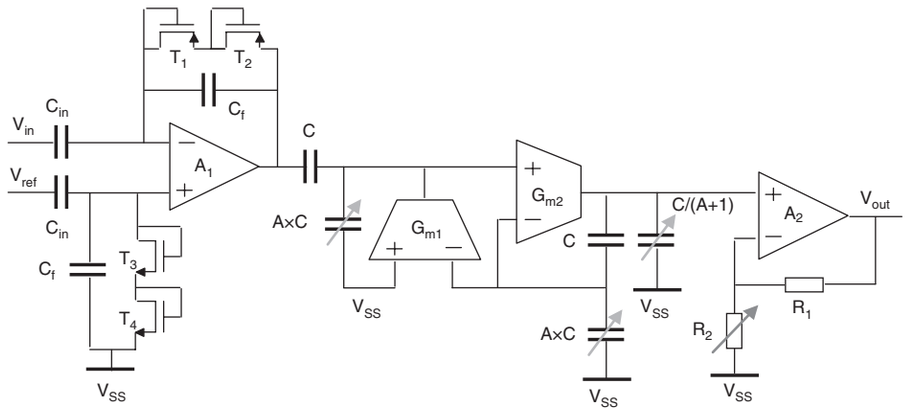

Fig. 2. Block level implementation of signal conditioning circuit.

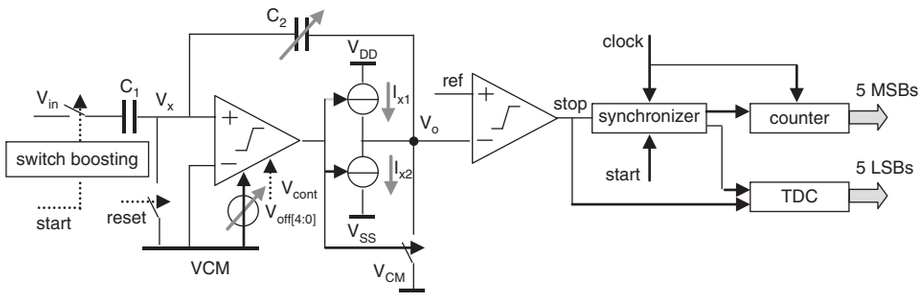

Fig. 3. Block level of an ADC with two-step TDC.

20 kHz, which determine the amount of noise added to the signal (3.5 μVrms at 37 °C). An effective way of lowering high electrode impedances is reducing the distance between the electrode sites and the readout circuits, e.g., by placing an amplifier very close to the signal source (the amplifier converts the high-impedance node at the electrode site to a low-impedance node)[13] and/or changing material composition of the electrode, e.g., gold electroplating.[14]

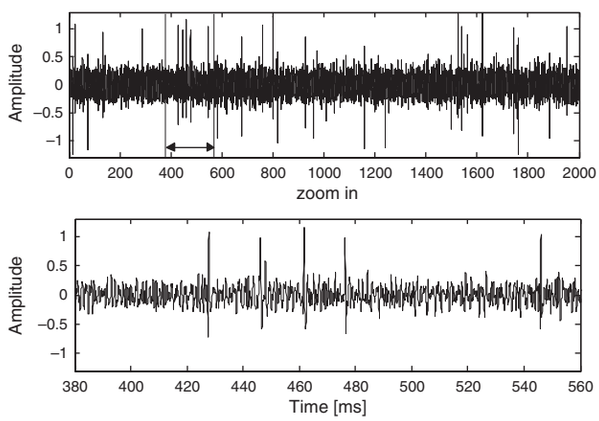

Fig. 4. The test data set, the *y*-axis is arbitrary—top: Raw signal after amplification, not corrected for gain, bottom: Zoom in of the raw signal.

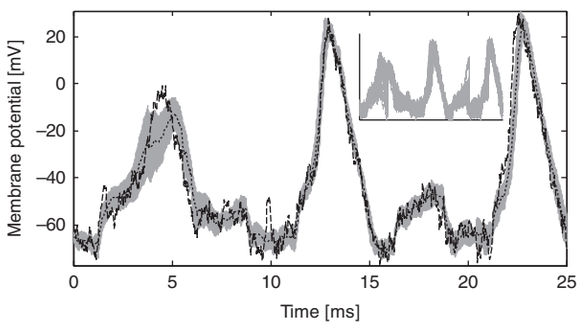

Fig. 5. Statistical voltage trace of a raw neural signal; dark gray—nominal voltage trace, black area—expected voltage trace; smaller figure: light grey area—voltage traces from 1000 randomly selected neural channel compartments.

3. POWER-EFFICIENCY OF ANALOG FRONT-END SIGNAL CONDITIONING AND QUANTIZATION CIRCUITS

3.1. Signal Conditioning Circuits

The design constraints of front-end neural amplifiers such as low-power, high-gain and low-noise operation, stable *dc* interface with the sensors (microprobes), and small silicon area are even more pronounced when the number of recording sites increases to several hundred for typical multi-electrode arrays. The power dissipation of front-end amplifiers (Fig. 2) is dictated

by the tolerable input-referred thermal noise (IRN). For minimum IRN, the transistors should operate in the subthreshold region; IRN can be than expressed as $V_{\mathrm{rms,ni}} = \sqrt{[(4kT\pi U_{\mathrm{T}}BW)/(2\kappa^2 I_{\mathrm{D}})]}$[15] (assuming first order frequency response), where k is Boltzmann constant, T is the absolute temperature, U_{T} is the thermal voltage, κ is the subthreshold gate coupling coefficient, I_{D} is total drain current, and BW is the -3 dB bandwidth of the amplifier. Consequently, for a given bandwidth the noise is inversely proportional to the square root of the supply current, hence, there exists the steep power cost of achieving low-noise performance in an amplifier. To adapt its noise per unit bandwidth, the bias current of the LNA can be made variable (consequently, to keep the overall bandwidth constant when the bias current of the gain stage is varied, a bandpass filter is added to the output of the LNA). The constant power, area and gain contours are illustrated in Figure 6. The total area is shown as the hyperbolic-shaped contour, while elliptic contours define the total drain current, I_{D}. In subthreshold region a transistor has a maximum $g_{\mathrm{m}}/I_{\mathrm{D}}$, such that $g_{\mathrm{m}} = \kappa I_{\mathrm{D}}/U_{\mathrm{T}}$. Typically, desired high g_{m} is obtained at the cost of an increased bias current (increased power) or area (wide transistors). However, for very short channel the carrier velocity quickly reaches the saturation limit at which the g_{m} also saturates, becoming independent of gate length or bias. The intrinsic gain degradation can be alleviated with open-loop residue amplifiers, comparator-based switched capacitor circuits,[16] and correlated level shifting.[17]

On a transistor level, the noise contribution of the amplifier input pair is reduced by increasing the transconductance g_{m}, increasing the current, by using a cascode resistive loading rather than current-source loads, or increasing the aspect ratio of the devices. The effect of the last method, however, is partially canceled by the increase in the noise excess factor. The noise voltages of the transistors used as current sources (or mirrors) are multiplied by the g_{m} of the device itself i.e., minimizing g_{m} minimizes noise. Since the current is usually set by other requirements, the only possibility is to decrease the aspect ratio of the device. This leads to an increase in the gate overdrive voltage, which, as a positive side effect, also decreases the thermal noise coefficient γ. IRN of the fully differential LNA in (Fig. 2)[18] is 3.1 μVrms over 0.1–20 kHz. The circuit achieves 55 dB closed loop gain, THD is below 1% for typical extracellular neural signals (smaller than 10 mV peak-to-peak). The common-mode rejection ratio (CMRR), and the power-supply rejection ratio (PSRR) exceeds 75 dB. The capacitive-attenuation band-pass filter with first-order slopes achieves 65 dB dynamic range, 210 mVrms at 2% THD, and 140 μVrms total integrated output noise. The total signal conditioning circuit, including LNA and band-pass filter, consumes 2.1 μV/channel, and occupies an area of 0.036 mm^2/channel. In Table I, we compare the state of the art signal conditioning circuits to the system in Ref. [18].

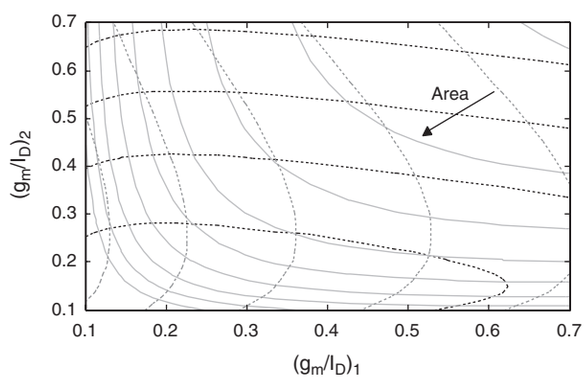

Fig. 6. Two stages $g_{\mathrm{m}}/I_{\mathrm{D}}$ versus constant gain (plain), constant area (plain hyperbolic), and constant current (dashed elliptic) contours.

3.2. Signal Quantization Circuits

Neural pattern classification and recognition require simultaneous recording from a large number of neurons (and recording the LFP and AP simultaneously). This however leads to the requirement of large dynamic range and signal bandwidth for the analog front-end. In the worst case, we assume that spikes with amplitude of tens of μV added on LFPs with amplitudes of about 2 mV appear at the input of a recording channel. If an input-referred noise of 2 μV is needed to meet the SNR requirement of the neuron signal, the dynamic range of the channel is around 60 dB, resulting in a 10-bit A/D conversion. Additionally, this sampling has to be done fast enough to capture the information in neuron signals, e.g., 40 kS/s sampling rate. For a neural recording device with 128 channels this results in a data rate of 51.2 Mbs^{-1}. To relax the in-channel integration density, part of the front-end electronics, usually an ADC, is moved out of the channel to the periphery of the recording area (e.g., for the sampling rate of 40 kS/s for one channel to avoid extensive interpolation of spike samples, 640 kS/s ADC can be shared among 16 channels). Consequently, a 128-channel front-end interface can be built in 16×8 configuration.

Several circuit topologies, such as current reuse,[23] time multiplexing,[24] sleep modes,[25] adaptive duty-cycling of the entire analog front-end,[26] and adaptive system bandwidth

Table I. Signal conditioning: Comparison with prior art.

	[19]	[20]	[21]	[22]	[18]*
Technology	0.18	0.13	0.18	0.065	0.065
V_{DD} [V]	0.45	1.2	1.8	1	1
Gain [dB]	52	54–60	30–72	52.1	65
INF [μVrms]	3.2	4.7	3.2	4.13	3.1
Bandw. [Hz]	10k	10–5k	300–6k	1–8.2k	20k
P/chann. [μW]	0.73	3.5	5.4	2.8	2.1
A/chann. [mm^2]	0.2	0.09	0.08	0.042	0.036

Note: *simulation data

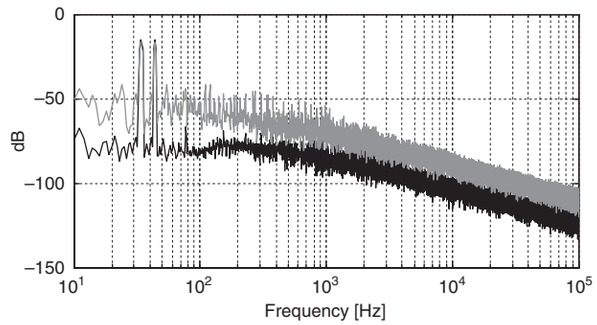

Fig. 7. Spectral signature of SAR A/D converter-two tone test; black line–spectral content with nominal gain, grey line–spectra with 20% gain reduction, equivalent to 4 LSB loss in the dynamic range.

or resolution,[27] can be used to improve power efficiency of signal quantization circuits by exploiting the fact that neurons spikes are irregular and low frequency. Signal quantization in BMI is usually performed with SAR A/D converter. The accuracy of the SAR A/D converter is determined either by the *kT/C* noise requirement, the required capacitor matching, the size of parasitic capacitance, or design rules of the technology. The matching properties of the capacitors as well as the parasitic capacitances affect the linearity characteristics of the converter. Additionally, the input parasitic capacitance of the comparator is dependent on the input voltage of the comparator, which itself varies during the conversion cycles, further degrading the linearity performance of the converter. The noise of the signal conditioning electronic circuits is mainly determined by the thermal and flicker noise, e.g., nonstationary stochastic process. The noise in signal quantization circuits is, however, designated as wide-sense cyclostationary.[28] The observed SNR of the system increases as the system is isomorphically scaled up (Fig. 7), which suggests a fundamental trade-off between SNR and speed of the system (note that over-dimensioning in a case of higher yield, leads to a larger area and higher power consumption). This lower bound on the speed in a SAR A/D converter loop is primarily a function of the technology's gate delay and *kT/C* noise multiplied by the number of SAR cycles necessary for one conversion.

Recently, several alternatives for the SAR A/D converter have emerged on a system-, e.g., combining signal amplification and data conversion into a single circuit,[29] or operational-level, e.g., current-mode converters,[30] time-mode converters (Fig. 3).[31] For example, time-based ADC (Fig. 3) in Ref. [31], achieves 9.4 ENOB at 640 kS/s, occupy an area of 0.022 mm^2 in 90 nm CMOS, and consumes less than 2.7 μW, corresponding to a FoM ($=P/(2f_{in} \times 2^{ENOB})$) of 6.2 fJ/conversion-step. In the circuit, a voltage signal is converted to a time-domain representation using a comparator-based switched-capacitor circuit and a continuous-time comparator. To improve the power efficiency, resulting time domain information is converted to the corresponding digital code with a two-step time-to-digital converter (TDC), where fine quantization of the resulting residue is obtained with folding Vernier converter. In Table II, we compare[31] to the state of the art circuit realizations of the signal quantizers.

4. POWER-EFFICIENCY OF DIGITAL BACK-END SIGNAL DETECTION AND CLASSIFICATION CIRCUITS

4.1. Signal Detection Circuits

The ability to distinguish spikes from noise, and to distinguish spikes from different sources from the superimposed waveform, depends on both the discrepancies between the noise-free spikes from each source and the SNR in the recording system. The time occurrences of the AP are detected by the absolute value thresholding (AT)[38] based on root-mean-square (AT-RMS), or based on median (AT-median), the nonlinear energy operator (NEO),[39] an exponential-polynomial-component Hilbert space detection (EC-PC)[40] or discrete- or continuous-wavelet transform.[41] The power-efficiency in a neuron signal detector is not only constrained by practical imperfections of a neuron signals, e.g., spike overlapping, waveform variation, low SNRs, unresolved artifacts, and interferences, but with a requirement for nonparametric and unsupervised detection to avoid frequent manual parameter tuning. The structure of NEO and AT-RMS are the most power-efficient, yet the EC-PC detector offers the most balanced trade-off among hardware complexity, functionality, and detection performance.

Table II. Signal quantization: comparison with prior art.

	[19]	[21]	[32]	[33]	[30]*	[34]*	[35]	[36]*	[37]*	[31]*	
Technology	0.18	0.18	0.12	0.09	0.18	0.09	0.18	0.35	0.09	0.09	
Type	SAR	SAR	Time	Time	Current	SAR	ΣΔ	SAR	SAR	Time	
V_{DD} [V]	0.45	1.8	1.2	1	1.2	1	1.8	3.3	0.5	1	
f_S [kS/s]	200	120	1000	1000	16	1000	50	16	1280	640	40
ENOB	8.3	9.2	10	7.9	8	9.34	10.2	8.9	9.95	9.4	9.5
FoM [fJ/conv-st]	21	382	175	188	132	2.87	0.22	93	2.36	6.2	21
Power [μ W]	1.35	27	180	14	0.45	1.79	13	3.06	3	2.7	1.6
Area [mm^2]	NR	NR	0.105	0.06	0.078†	NR	0.038	NR	0.048†	0.022†	

Notes: *simulation data, †estimated, nr—not reported.

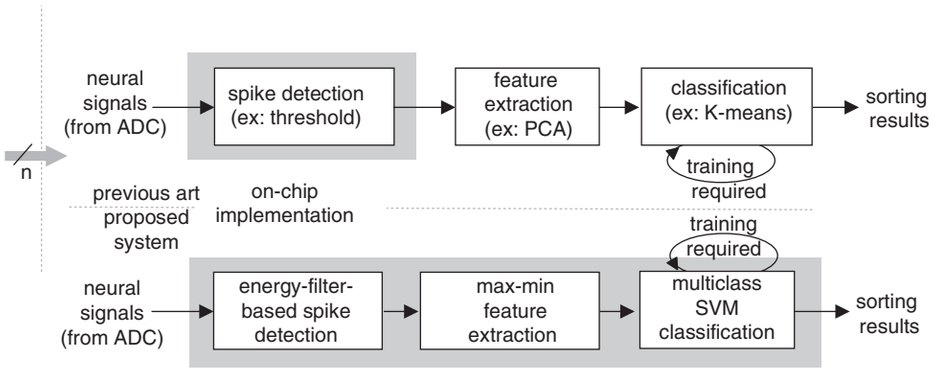

Fig. 8. Back-end signal processing: previous art versus proposed method.

4.2. Signal Classification Circuits

After the waveform alignment, to simplify the classification process, a feature extraction step, e.g., principal component analysis,[42] or wavelet decomposition,[38] characterizes detected APs and represent each detected AP in a reduced dimensional space. Based on these features the spikes are classified (e.g., by *k*-means,[43] expectation maximization,[44] template matching,[45] Bayesian clustering,[46] artificial neural network (ANN), support vector machine (SVM)[47–50]), in the clusters, where each cluster corresponds to the spiking activity of a single neuron. The feature extraction and spike classification significantly reduce the data requirements prior to data transmission (in multi-channel systems, the raw data rate is substantially higher than the limited bandwidth of the wireless telemetry). The accuracy of the neural spike classification in a digital back-end signal processing unit directly increase with A/D converter resolution, although it saturates beyond 5–6 bit resolution, ultimately limited by the SNR. However, since the amplitude of the observed neural signals can vary, typically, by one order of magnitude, additional resolution is needed (i.e., 2–3 bit), if the amplification gain is fixed. Additionally, increasing the sampling rate of the A/D converter improves neural classification accuracy, since this captures finer features further differentiating the signals.

The SVM has been introduced to bioinformatics and neural signal classification[47–50] because of its excellent generalization, sparse solution and concurrent utilization of quadratic programming. Like ANN classifiers, applications of SVMs to any classification problem require the determination of several user-defined parameters, e.g., choice of an appropriate kernel and related parameters, determination of regularization parameter and an appropriate optimization technique. Correspondingly, SVM applies the structure risk minimization instead of the empirical risk minimization and solves the non-linear, dimensionality curse problems efficiently. A programmable neural spike classifier based on multiclass kernel SVM for 128-channel spike sorting system in Figures 8 and 9[50] tracks the evolution of clusters in real-time. Threshold crossings of a local energy measurement[41] are used to detect spikes (Fig. 10). A frequency-shaping filter significantly attenuates the low frequency noise and helps to differentiate similar spikes from different neurons. The feature extraction based on maximum and minimum values of spike waveforms first derivatives[51] is employed due to its small computation and little memory requirement, while preserving high information score. To simplify the kernel classifier trained by the SVM, we extend iterative greedy optimization reduced set vectors approach with boosted cascade classifier (Fig. 11). Consequently, we assess the reduced expansion in a cascaded way, such that in most

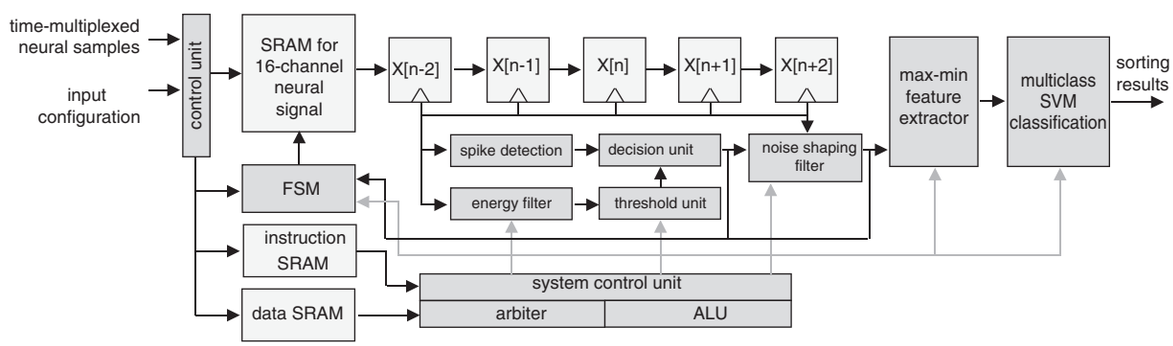

Fig. 9. Architecture of the proposed back-end signal processing.

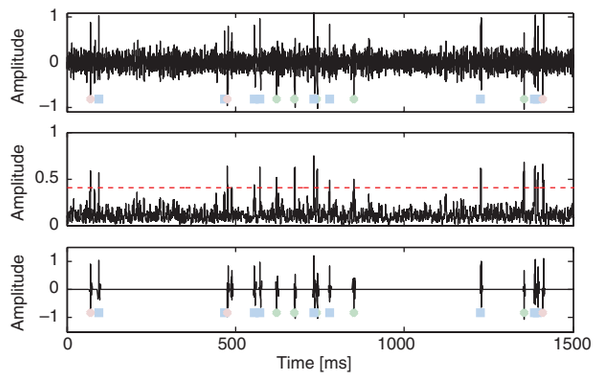

Fig. 10. Spike detection from continuously acquired data, the *y* axis is arbitrary-top: raw signal after amplification, not corrected for gain, middle: threshold (line) crossings of a local energy measurement with a running window of 1 ms, and bottom: detected spikes.

cases a very small number of support vectors are applied. The training data (*td*) is split into subsets, and each one is evaluated individually for support vectors in the first layer. Hence, eliminating non-support vectors early from the classification, significantly accelerates SVM procedure. The scheme requires only modest communication from one layer to the next, and a satisfactory accuracy is often obtained with a single pass through the cascade. The required arithmetic over feature vectors (the element-wise operands as well as SVM model parameters) is executed with two-stage pipeline (i.e., to reduce glitch propagation) processing unit (Fig. 12). Flip-flops are inserted in the pipeline to lessen the impact of active-glitching, and to reduce the leakage energy.

Figure 13 illustrate a multi-class classification, where the bold lines represent decision boundaries. The SVM spike sorting performance has been summarized and benchmarked (Fig. 14) versus four different, relatively computationally-efficient methods for spike sorting: template matching, principal component analysis (PCA), Mahalanobis, and Euclidean distance. The performance is quantified using the effective accuracy, i.e., total spikes classified versus spikes correctly classified (excluding spike detection). The power-efficient clustering is achieved[50] (Fig. 15) by a combination of the several algorithm and circuit techniques, namely, the Kesler's transformation, a boosted cascade reduced set vectors approach, a two-stage pipelined processing units, the power-scalable kernels, the register-bank memory, a high-V_T devices, and a near-threshold supply. The number of support vectors required is partly governed by the complexity of the classification task, i.e., as the SNR decreases more support vectors are needed in order to define a more complex decision boundary. For our dataset, the number of support vectors required is reduced within the range of 300–310. The kernels yield increasing levels of strength (e.g., false alarm for linear kernel of 18 per day decrease to 1.2 per day for RBF kernel). However, the required power for each kernel varies by orders of magnitude. The signal detection and classification circuit implementation includes 31 k logic gates resulting in a 2.64 mm^2 area, and consumes only 41 μW of power from a 0.4 V supply voltage.[50] The consumed power corresponds to a temperature increase of

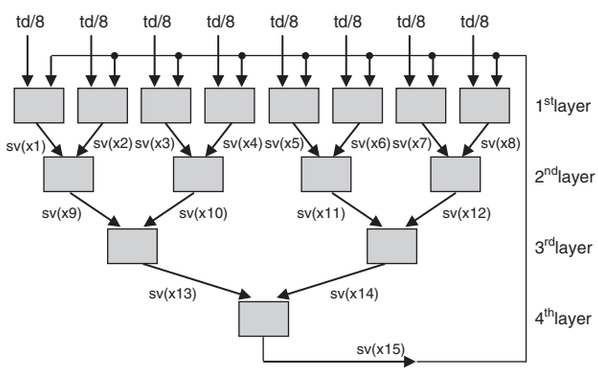

Fig. 11. Binary boosted cascade architecture.

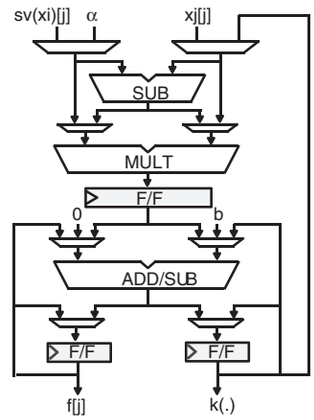

Fig. 12. Two-stage pipeline processing unit.

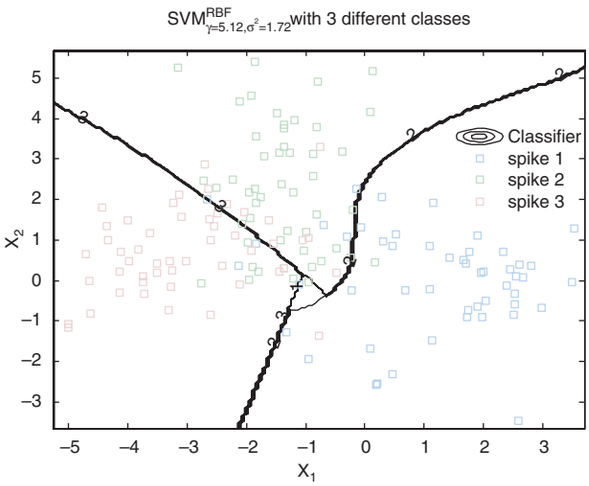

Fig. 13. The SVM separation hypersurface for the RBF kernel.

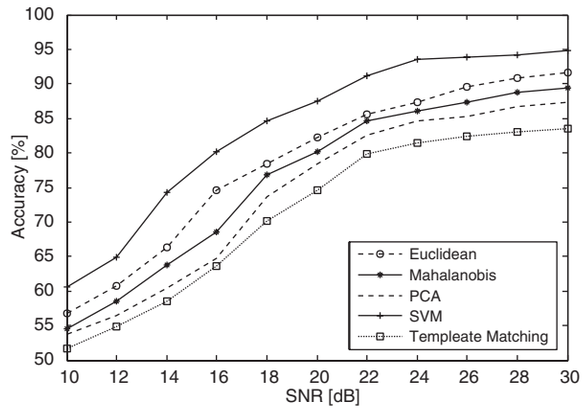

Fig. 14. Effect of SNR on overlapping spikes of three classes on sorting accuracy of the BMI system.

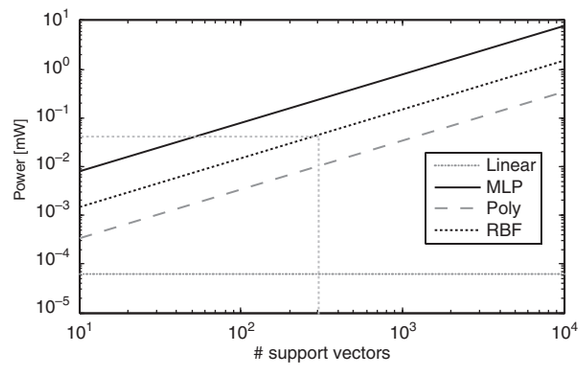

Fig. 15. Energy per cycle versus various SVM kernels.

Table III. Signal classification: comparison with prior art.

	[52]	[53]	[54]	[50]*
Technology [nm]	65	90	65	65
Programmability	no	yes	no	yes
V_{DD} [V]	0.27	1	0.3	0.4
No. of channels	16	128	1	128
Pow. Dens. [μW/mm^2]	60.9	9.8	43.4	15.5
Power [μW]	75	87	2.17	41
Area [mm^2]	1.23	8.9	0.05	2.64

Note: *simulated data.

0.11 °C (i.e., assuming the 0.029 °C/mW model), which is ∼9 times lower than the required consumed power in a neural implants safe range (<1 °C). In Table III, we compare the state of the art spike sorting systems to Ref. [50].

5. CONCLUSION

In this paper, we address the power efficiency of brain-machine interface at various abstraction levels, i.e., circuit- and system level. It, therefore, provides a broad view of the various solutions and their possible combinations in effective (complementary) techniques, without compromising on required power and area, even in a neural interfaces with a low SNR.

Acknowledgment: This research was supported in part by the European Union and the Dutch government, as part of the CATRENE program under Heterogeneous INCEPTION project.

References

1. G. Buzsaki, Large-scale recording of neuronal ensembles. *Nature Neurosc.* 7, 446 **(2004)**.
2. F. A. Mussa-Ivaldi and L. E. Miller, Brain-machine interfaces: Computational demands and clinical needs meet basic neuroscience. *Trends Neurosc.* 26, 329 **(2003)**.
3. K. Faligkas, L. B. Leene, and T. G. Constandinou, A novel neural recording system utilising continuous time energy based compression. *IEEE Int. Symp. Circ. Syst.* 3000 **(2015)**.
4. J. T. Robinson, M. Jorgolli, and H. Park, Nanowire electrodes for high-density stimulation and measurement of neural circuits. *Front. Neur. Circ.* 7, 1 **(2013)**.
5. K. D. Harris, D. A. Henze, J. Csicsvari, H. Hirase, and G. Buzsáki, Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *Jour. Neurophys.* 84, 401 **(2000)**.
6. G. Buzsaki, Large-scale recording of neuronal ensembles. *Nature Neurosc.* 7, 446 **(2004)**.
7. R. R. Harrison, A low-power integrated circuit for adaptive detection of action potentials in noisy signals, *Int. Conf. IEEE Eng. Med. Biol. Soc.* 3325 **(2003)**.
8. K. Oweiss, K. Thomson, and D. Anderson, A systems approach for real-time data compression in advanced brain-machine interfaces, *IEEE Int. Conf. Neur. Eng.* 62 **(2005)**.
9. Y. Perelman and R. Ginosar, Analog frontend for multichannel neuronal recording system with spike and lfp separation. *Jour. Neurosci. Meth.* 153, 2126 **(2006)**.
10. Z. S. Zumsteg, et al., Power feasibility of implantable digital spike-sorting circuits for neural prosthetic systems, *Int. Conf. IEEE Eng. Med. Biol. Soc.* 4237 **(2004)**.
11. A. Zviagintsev, Y. Perelman, and R. Ginosar, Low power architectures for spike sorting, *IEEE Int. Conf. Neur. Eng.* 162 **(2005)**.
12. A. Saarinen, et al., Stochastic differential equation model for cerebellar granule cell excitability. *PLoS Comput. Biol.* 4, 1 **(2008)**.
13. R. H. Olsson and K. D. Wise, A three-dimensional neural recording microsystem with implantable data compression circuitry. *IEEE J. Solid-State Circ.* 40, 2796 **(2005)**.
14. J. Du, et al., Multiplexed, high density electrophysiology with nanofabricated neural probes. *PLoS One* 6, e26204 **(2011)**.
15. W. Wattanapanitch, M. Fee, and R. Sarpeshkar, An energy-efficient micropower neural recording amplifier. *IEEE Trans. Biomed. Circuits Syst.* 1, 136 **(2007)**.
16. T. Sepke, et al., Comparator-based switched-capacitor circuits for scaled CMOS technologies, *IEEE Int. Solid-State Circ. Conf. Dig. Tech. Papers* 220 **(2006)**.
17. B. R. Gregoire and U.-K. Moon, An over-60db true rail-to-rail performance using correlated level shifting and an opamp with 30db loop gain, *IEEE Int. Solid-State Circ. Conf. Dig. Tech. Papers* 540 **(2008)**.
18. A. Zjajo and R. van Leuken, A 2.1 μW/channel current-mode integrated neural recording interface. *IEEE Biomed. Health Infor.* 505 **(2016)**.
19. D. Han, et al., A 0.45 V 100-channel neural-recording IC with sub-μW/channel consumption in 0.18 μm CMOS. *IEEE Trans. Biomed. Circ. Syst.* 7, 735 **(2013)**.

20. K. Abdelhalim, et al., 64-channel UWB wireless neural vector analyzer SoC with a closed-loop phase synchrony-triggered neurostimulator. *IEEE J. Solid-State Circ.* 48, 2494 **(2013)**.

21. C. M. Lopez, et al., An implantable 455-active-electrode 52-channel CMOS neural probe, *IEEE Int. Solid-State Circ. Conf.* 288 **(2013)**.

22. K. A. Ng and Y. P. Xu, A multi-channel neural-recording amplifier system with 90dB CMRR employing CMOS-inverter-based OTAs with CMFB through supply rails in 65 nm CMOS, *IEEE Int. Solid-State Circ. Conf.* 206 **(2015)**.

23. X. Zou, et al., A 100-channel 1-mW implantable neural recording IC. *IEEE Trans. Circuits Syst.-I: Reg. Papers* 60, 2584 **(2013)**.

24. C. Chae, et al., A 128-channel 6 mw wireless neural recording IC with spike feature extraction and UWB transmitter. *IEEE Trans. Neural Syst. Rehab. Eng.* 4, 312 **(2009)**.

25. R. F. Yazicioglu, et al., A 200 W eight-channel EEG acquisition ASIC for ambulatory EEG systems, *IEEE Int. Solid-State Circ. Conf. Dig. Tech. Papers* 164 **(2008)**.

26. J. Lee, et al., 64 channel programmable closed-loop neurostimulator with 8 channel neural amplifier and logarithmic ADC. *IEEE J. Solid-State Circ.* 45, 1935 **(2010)**.

27. X. D. Zou, et al., A 1-V 450-nW fully integrated programmable biomedical sensor interface chip. *IEEE J. Solid-State Circ.* 44, 1067 **(2009)**.

28. A. Zjajo, C. Galuzzi, and R. van Leuken, Stochastic noise analysis of neural interface front end, *IEEE Int. Symp. Circ. Syst.* 169 **(2015)**.

29. A. Rodríguez-Pérez, et al., A 64-channel inductively-powered neural recording sensor array, *IEEE Biomed. Circ. Syst. Conf.* 228 **(2012)**.

30. B. Haaheim and T. G. Constandinou, A sub-1 μW, 16 kHz Current-Mode SAR-ADC for Single-Neuron Spike Recording, *IEEE Biomed. Circ. Syst. Conf.* 2957 **(2012)**.

31. A. Zjajo and R. van Leuken, A 2.7 μW 10b 640kS/s time-based A/D converter for implantable neural recording interface, *IEEE Int. Symp. Circ. Syst.* 1074 **(2016)**.

32. E. Allier, et al., 120 nm low power asynchronous ADC, *IEEE Int. Symp. Low Pow. Electr. Des.* 60 **(2005)**.

33. S. Naraghi, M. Courcy, and M. P. Flynn, A 9-bit, 14 μW and 0.006 mm^2 pulse position modulation ADC in 90 nm digital CMOS. *IEEE J. Solid-State Circ.* 45, 1870 **(2010)**.

34. T. Rabuske, et al., A self-calibrated 10-bit 1MSps SAR ADC with reduced-voltage charge-sharing DAC, *IEEE Int. Symp. Circ. Syst.* 2452 **(2013)**.

35. C. Gao, et al., An ultra-low-power extended counting ADC for large scale sensor arrays. *IEEE Int. Symp. Circ. Syst.* 81 **(2014)**.

36. L. Zheng, et al., An adaptive 16/64 kHz, 9-bit SAR ADC with peak-aligned sampling for neural spike recording. *IEEE Int. Symp. Circ. Syst.* 2385 **(2014)**.

37. Y.-W. Cheng and K. T. Tang, A 0.5-V 1.28-MS/s 10-bit SAR ADC with switching detect logic. *IEEE Int. Symp. Circ. Syst.* 293 **(2015)**.

38. R. Q. Quiroga, Z. Nadasdy, and Y. Ben-Shaul, Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput.* 16, 1661 **(2004)**.

39. K. Kim and S. Kim, Neural spike sorting under nearly 0-db signal-to-noise ratio using nonlinear energy operator and artificial neural-network classifier. *IEEE Trans. Biomed. Eng.* 47, 1406 **(2000)**.

40. Z. Yang, et al., A new EC-PC threshold estimation method for *in-vivo* neural spike detection. *J. Neural. Eng.* 9, 1 **(2012)**.

41. K. H. Kim, et al., A wavelet-based method for action potential detection from extracellular neural signal recording with low signal-to-noise ratio. *IEEE Trans. Biomed. Eng.* 50, 999 **(2003)**.

42. D. A. Adamos, E. K. Kosmidis, and G. Theophilidis, Performance evaluation of pca-based spike sorting algorithms. *Comp. Meth. Progr. Biomed.* 91, 232 **(2008)**.

43. S. Takahashi, Y. Anzai, and Y. Sakurai, A new approach to spike sorting for multi-neuronal activities recorded with a tetrode-how ICA can be practical. *Neurosc. Res.* 46, 265 **(2002)**.

44. F. Wood, M. Fellows, J. Donoghue, and M. Black, Automatic spike sorting for neural decoding, *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.* 4009 **(2004)**.

45. C. Vargas-Irwin and J. P. Donoghue, Automated spike sorting using density grid contour clustering and subtractive waveform decomposition. *J. Neurosc. Meth.* 164, 1 **(2007)**.

46. J. Dai, et al., Experimental study on neuronal spike sorting methods, *IEEE Fut. Gen. Communic. Netw. Conf.* 230 **(2008)**.

47. R. J. Vogelstein, et al., Spike sorting with support vector machines, *Int. Conf. IEEE Eng. Med. Biol. Soc.* 546 **(2004)**.

48. K. H. Kim, S. S. Kim, and S. J. Kim, Advantage of support vector machine for neural spike train decoding under spike sorting errors, *Int. Conf. IEEE Eng. Med. Biol. Soc.* 5280 **(2005)**.

49. R. Boostani, et al., A comparison approach toward finding the best feature and classifier in cue-based BCI, *Med. Biolog. Eng. Comp.* 45, 403 **(2007)**.

50. A. Zjajo and R. van Leuken, A 41 μW real-time adaptive neural spike classifier. *IEEE Biomed. Health Infor.* 489 **(2016)**.

51. T. Chen, et al., NEUSORT2.0: A multiple-channel neural signal processor with systolic array buffer and channel-interleaving processing schedule, *Int. Conf. IEEE Eng. Med. Biol. Soc.* 5029 **(2008)**.

52. V. Karkare, S. Gibson, and D. Marković, A 75-μW, 16-channel neural spike-sorting processor with unsupervised clustering. *IEEE J. Solid-State Circ.* 48, 2230 **(2013)**.

53. T.-C. Ma, T.-C. Chen, and L.-G. Chen, Design and implementation of a low power spike detection processor for 128-channel spike sorting microsystem, *IEEE Int. Conf. Acous., Speech Sig. Proc.* 3889 **(2014)**.

54. Z. Jiang, Q. Wang, and M. Seok, A low power unsupervised spike sorting accelerator insensitive to clustering initialization in sub-optimal feature space, *IEEE Des. Autom. Conf.* 1 **(2015)**.

Amir Zjajo

Amir Zjajo *received the M.Sc. and DIC degrees from the Imperial College London, London, U.K., in 2000 and the Ph.D. degree from Eindhoven University of Technology, Eindhoven, The Netherlands in 2010, all in electrical engineering. In 2000, he joined Philips Research Laboratories as a member of the research staff in the Mixed-Signal Circuits and Systems Group. From 2006 until 2009, he was with Corporate Research of NXP Semiconductors as a senior research scientist. In 2009, he joined Delft University of Technology as a Faculty member in the Circuit and Systems Group. Dr. Zjajo has published more than 70 papers in referenced journals and conference proceedings, and holds more than 10 US patents or patents pending. He is the author of the books Brain-Machine Interface: Circuits and Systems (Springer, 2016), Low-Voltage High-Resolution A/D Converters: Design and Calibration (Springer, 2011, Chinese translation, China Machine Press, 2015) and Stochastic Process Variations in Deep-Submicron CMOS: Circuits and Algorithms (Springer, 2014). He serves as a member of Technical Program Committee of IEEE Design, Automation and Test in Europe Conference, IEEE International Symposium on Circuits and Systems, IEEE International Symposium on VLSI, IEEE International Symposium on Nanoelectronic and Information Systems, and IEEE International Conference on Embedded Computer Systems. His research interests include power-efficient mixed-signal circuit and system design for health and mobile applications, neuromorphic electronic circuits for autonomous cognitive systems, and molecular and genetic circuits and networks. Dr. Zjajo won the best paper award at BIODEVICES 2015 and DATE 2012.*