A 1 V, Compact, Current-Mode Neural Spike Detector with Detection Probability Estimator in 65 nm CMOS

Enyi Yao* and Arindam Basu†

School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore
Email: *eyao1@e.ntu.edu.sg;†arindam.basu@ntu.edu.sg

***Abstract*— In this paper, we describe a novel low power, compact, current-mode spike detector circuit for real-time neural recording systems where neural spikes or action potentials (AP) are of interest. Such a circuit can enable massive compression of data facilitating wireless transmission. This design operates by approximating the popularly used nonlinear energy operator (NEO) through standard current mode analog blocks that can operate at low voltages. To reduce sensitivity of threshold setting, this work uses a current-mode oscillator based detection probability estimator (DPE) to reject false positives caused by the background noise. The circuit is implemented in a 65 nm CMOS process and occupies 200 μm x 150 μm of chip area. Operating from a 1 V power supply, it consumes about 88 nW of static power and 10 nJ of dynamic energy per input spike.**

I. Introduction

With the advance of today's semiconductor technology, hundreds of channels have been integrated in single-chip neural recording system to monitor multiple neurons' activities in the brain simultaneously. One emerging application is the brain machine interfaces (BMIs) which is a direct real-time interface between brain and computers or other devices. It could be utilized to treat conditions like Parkinson's disease, restore motor impairment or even for neuroscience research purpose. Usually, the sensory part should be implantable under the skull, record the neurons' action potentials, which are also called spike signals and wirelessly transmit the raw neural data to avoid tissue infections. After that a signal processing unit outside our body will derive the information from these spike signals and generate the corresponding control parameters for some robotic prosthetic device or a stimulator in the brain, forming a close-loop feedback. Since the power supply of the implanted neural recording system is typically obtained from wireless transfer, the implanted circuits have to operate within very limited power budgets

Fig. 1 illustrates a general architecture of the neural recording system composed of a pre-amplifier (A), a variable gain amplifier (VGA), an analog-to-digital converter (ADC) and a wireless radio transmitter. Consider the case where we want to transmit over 100 channels of neural data sampled at 20 kHz with 10 bit resolution. The resultant data rate of 20 Mbps that is beyond the capacity of current low power transmitters [1]. However, since spikes are sparsely distributed in the neural signal, one possible approach of compression

The authors would like to thank MediaTek 65-nm IC Shuttle Project for chip fabrication facility.

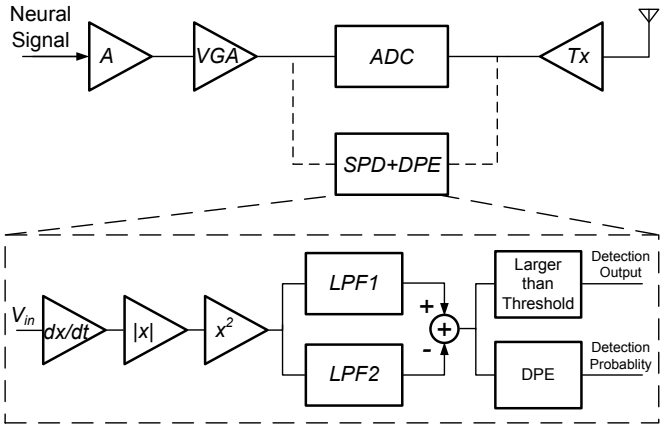

Fig. 1: Neural recording system with real-time spike detection circuit.

in applications requiring only spike information is just to transmit the occurrence of a spike instead of the raw neural data. This alternative, shown in a dashed line in Fig. 1, can lead to compression factors of $50 - 5000$ depending on input spike rate and exact compression scheme used. Of course, this scheme is useful only if the circuit to detect spikes consumes considerably less power compared to the transmitter.

The original and simplest method for spike detection is amplitude thresholding. The threshold is set to be several times of the background noise level. However, the performance of this method is highly dependent on the chosen threshold value and has a poor performance in high noise scenario. Much better performance can be obtained by using nonlinear energy operator (NEO) to perform spike detection. It is defined as

$$NEO(x) = \left(\frac{dx}{dt}\right)^2 - \frac{d^2x}{dt^2} \cdot x. \quad (1)$$

Many new hardware designs based on NEO have been proposed in recent publications [2] [3] [4]. However, all of these work have either had poor detection accuracy or required the setting of a well calibrated threshold for high performance.

In this paper, a low-power, compact analog spike detector based on an approximation to NEO (similar to [4]) is presented. Current mode design is used for lowering supply voltage. To reduce the dependence of performance on threshold, an estimate of detection probability is also generated which is used to reduce false positives thus increasing the robustness of the proposed method. Measurement results from a prototype fabricated in 65nm CMOS are presented to validate

978-1-4799-8391-9/15/$31.00 ©2015 IEEE

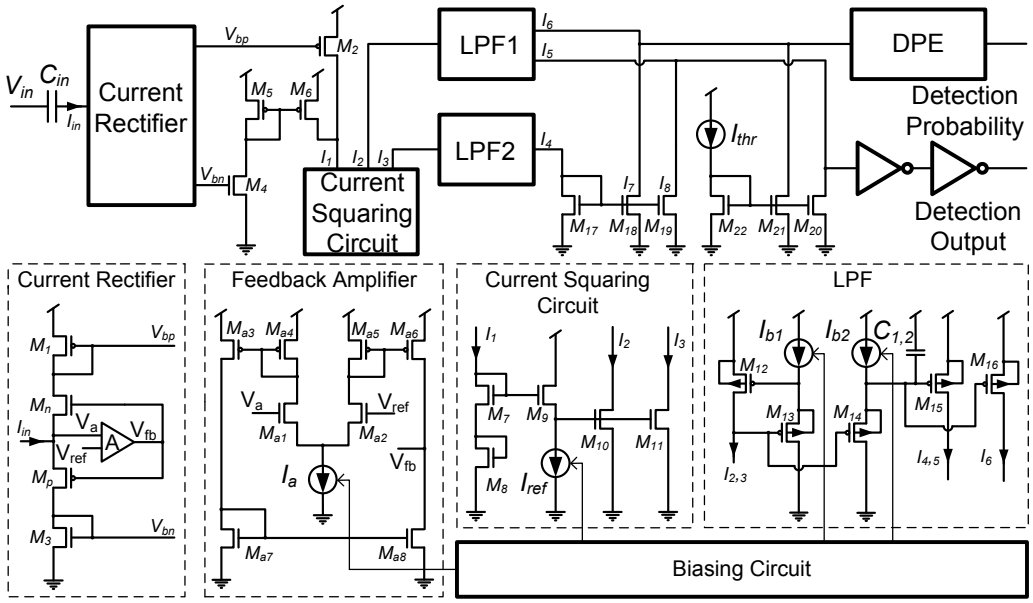

Fig. 2: Schematic of the spike detector with transistor level circuit of each sub-block.

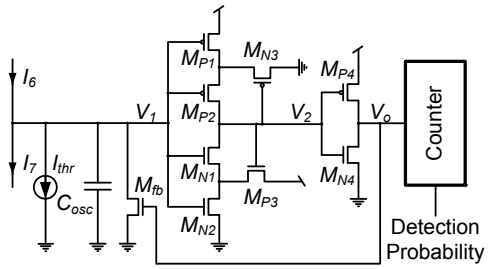

Fig. 3: Schematic of the oscillator based DPE circuit.

the concept.

II. System Architecture

The architecture of the spike detector shown in Fig. 1 consists of a differentiator, a current rectifier, a current squaring circuit, two low pass filter with different cut-off frequencies, a comparator and a oscillator based detection probability estimator (DPE). Each of these blocks are described next and shown in detail in Fig. 2.

A. Spike Detector

The architecture of our detector is based on an approximation to the NEO similar to the ED method in [4]. NEO output for a sine wave input $Asin(\omega t)$ is a constant of value $A^2\omega^2$. This can also be obtained by low pass filtering the square of derivative of the input with the advantage of much easier implementation. In contrast to [4], we have explicitly included a low pass filter that helped reduce spurious detections.

As shown in Fig. 2, the capacitor C_{in} acts as a differentiator and converts V_{in} to a current proportional to (dV_{in}/dt). Since the current squaring circuit can only process uni-directional current, a class-B current rectifier is used before it. An OTA in negative feedback is added in the class-B current mirror to reduce the input impedance. M_p and M_n are thin oxide devices with low threshold to reduce minimum required VDD. Also, the OTA uses long channel length devices instead of cascodes to get high gain at low voltages. The current squarer is a log domain translinear loop circuit using transistors biased in sub-threshold region with the following characteristic:

$$I_{2,3} = \frac{I_1^2}{I_{ref}}, \tag{2}$$

where I_{ref} is tunable with 8 digital control bits.

After the current squaring block, high frequency components need to be filtered by LPF1 (cut-off frequency $\omega 1 = 300$ Hz) to produce the desired estimate of NEO. Another low pass filter, LPF2 (cut-off frequency $\omega 2 = 64$ Hz) is also used to estimate the background noise level including the low frequency noise from LFP and power line interference. A detection pulse is generated if the difference of the two LPF currents is larger than a threshold. A copy of this current is also fed to a DPE that relaxes the accuracy needed in setting threshold. Fig. 2 shows the current-mode design of the low pass filter with the following transfer function:

$$\frac{I_{out}(s)}{I_{in}(s)} = \frac{I_{b1}/I_{b2}}{1 + s/\omega_0} = \frac{I_{b1}/I_{b2}}{1 + s\tau_0}, \tag{3}$$

where $\tau_0 = \frac{CU_t}{I_{b2}\kappa}$ and U_t is the thermal voltage. The cut-off frequency of the low pass filters can be adjusted by tuning either the bias current or the capacitance $C_{1,2}$. Thick oxide transistors are used to remove gate leakage induced error in cut-off frequency.

B. Detection Probability Estimator

Fig. 3 illustrates details of the current-controlled oscillator (CCO) based DPE. The capacitor C_{osc} is charged by the current $I_6 - I_7 - I_{thr}$. When V_1 crosses the higher threshold of the Schmitt Trigger inverter [5], it would pull V_2 down

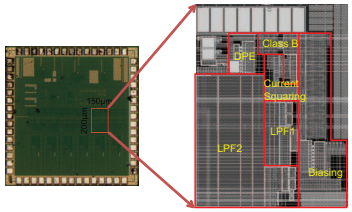

Fig. 4: The micrograph of the spike detector chip.

to the ground and V_o to VDD immediately. Then the feedback transistor M_{fb} turns on, discharging the capacitor C_{osc}. When V_1 reaches the lower threshold of the Schmitt Trigger inverter, the inverters will switch and C_{osc} will be charged by the currents again. Denoting the hysteresis window of the Schmitt Trigger inverter by ΔV, the frequency of this current-controlled oscillator can be obtained as:

$$f_{osc} = \frac{1}{\frac{\Delta V C_{osc}}{I_6 - I_7 - I_{thr}} + \frac{\Delta V C_{osc}}{I_{dis} - (I_6 - I_7 - I_{thr})}} \approx \frac{I_6 - I_7 - I_{thr}}{\Delta V C_{osc}} \tag{4}$$

since I_{dis}, the discharge current through M_{fb}, is very large. The CCO pulses clock a counter to create a digital representation of $I_6 - I_7 - I_{thr}$ which in turn is an estimate of the probability of detection. The counter is enabled for 1 ms every time there is a spike detection and the 6bit count, C_i for every consecutive 100μs is stored (total of 10 such values for 1 ms). We expect spike signals to have large amplitude and temporal duration exceeding 1 ms–noise signals that falsely trigger the detector will probably not satisfy at least one of the above criteria. To capture these two qualities, we use the following two features (f_1 and f_2):

$$f_1 = \Sigma_i C_i, f_2 = max(C_i) \tag{5}$$

We have found that there is a lot of variability in C_1. So the above features are computed excluding it. In the current chip, feature calculation is not included. Without it, the data rate for one channel for 50 detections/sec is 3kbps–a compression of $\approx 7X$. With feature extraction included on-chip, compression ratio would increase to about $25X$.

III. Power Dissipation

The power dissipation for the proposed spike detector can be separated into two parts: static power and dynamic power. The static power is mainly consumed by the biasing blocks, of which 540 nW is consumed by the main biasing block (comprising current reference and splitter) shared by all channels. For each channel, there are also some local current splitters with total power dissipation of 88 nW. The dynamic power is decided by the input signal characteristics. For a 100 mV amplitude spike signal as the input, the dynamic energy (dominated by DPE) simulated is about 10.3 nJ per spike. Thus we have as power per channel

$$P_t\,(nW) = 88 + 10.3f, \tag{6}$$

where f is the firing rate of the spike signal. For $f = 65$ Hz, dynamic power dissipation is about 670 nW leading to total power dissipation of 738 nW (1.3μW including global bias).

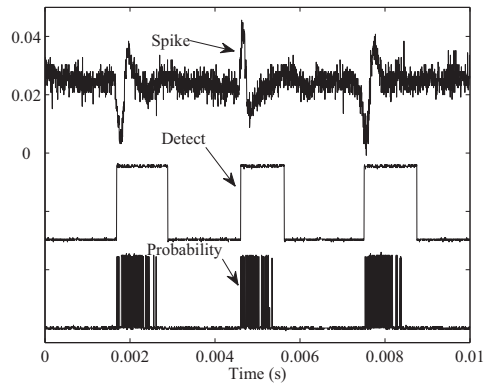

Fig. 5: Input neural signal, detection and oscillation outputs.

IV. Measurement Results

To validate the function of the proposed design, we implemented the system in a 65 nm CMOS process with a die photo illustrated in fig. 4. The input signals are constructed by selecting three distinct spike shapes from a database with 594 different spikes from intra-cortical neural recordings. The spike signals are randomly distributed on top of a noise generated by superimposing spikes selected from the database with random times and amplitudes to mimic a realistic scenario [6]. To evaluate the performance of our spike detector, we use false positive rate (FPR) and false negative rate (FNR) for comparison, which are defined as:

$$\text{FPR} = \frac{\text{\# of false alarms}}{\text{\# of total detections}}, \text{FNR} = \frac{\text{\# of missed spikes}}{\text{\# of actual spikes}}. \tag{7}$$

Figure 5 shows an example of one piece of input neural signal and the corresponding detection output and the CCO output before counting. The big advantage of the DPE is the weak dependence of performance the input amplitude or the threshold setting. To validate this, we test the DPE performance by varying the input signal level for a fixed value of threshold. Fig. 6 shows the measurement results of the DPE for 2 different noise levels: 0.1 and 0.25. These numbers indicate the noise standard deviation relative to the amplitude of the spike classes–corresponding SNR values are 7 dB and 1.5 dB respectively. As the noise level increases, input spike signal is corrupted by the background noise more significantly, making it more difficult to separate the noise from spike signal. However, for the DPE we expect the values of f_1 and f_2 to increase equally for both signal and noise classes–thus separation between the two should not be affected.

Figure 6(a) and (d) show the distribution of features f_1 and f_2 for the two noise cases demonstrating clear separability of clusters of signal and noise. As expected, valid spike signals have large values for both f_1 and f_2. Figure 6(b) and (c) demonstrate FPR and FNR when input singal strength is varied from 60 mV to 100 mV for same threshold level in the low noise scenario. The X axis plots the ratio of $I_{in,max}/I_{thr}$ normalized to the smallest value in each case. larger values of this ratio result in increased FPR for the conventional detection

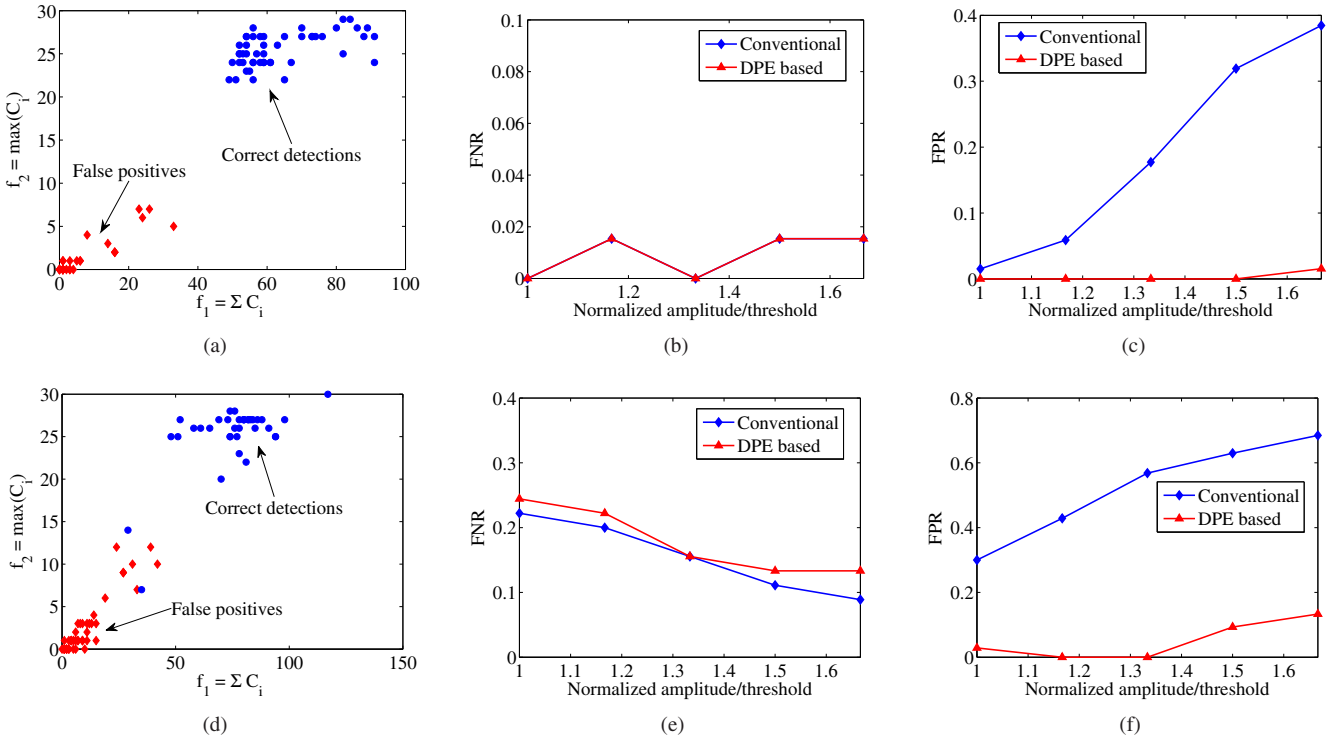

Fig. 6: The DPE output feature distributions and FNR FPR plots for low noise level: fig 6(a), fig 6(b), 6(c) and high noise level: fig 6(d), fig 6(e), 6(f). .

TABLE I: Comparison Table

	CICC 2008 [2]	ISCAS 2013 [3]	TBCAS 2012 [7]	EMBC 2013 [8]	This work
Process (nm)	130	180	130	130	65
Area (mm^2)	0.044	0.03	0.16	6.17	0.03
VDD (V)	1.0	1.8	1.2	-	1.0
Power (μW)	0.95	1.5	2.8	85	1.3
Result	Test	Simulation	Test	Test	Test
Feature	NEO Analog	NEO Analog	Threshold Analog	EC-PC Digital	NEO+DPE Analog

as can be expected if the threshold is too low compared to noise. However, for the DPE based detection, the clear separation between spike and noise features enable a drastic reduction of FPR using simple k-means clustering. A similar conclusion also holds for the high noise case in Fig. 6(e) and (f), though in this case FNR increases slightly for DPE.

V. Conclusions

We have reported a 1-V, low-power neural spike detector for real-time neural recording system. The circuit performs detection by approximating NEO using current mode techniques. Further, it uses a detection probability estimate to make performance relatively insensitive to threshold setting. The chip, fabricated in a 65 nm CMOS process with a die area of 200 μm x 150 μm consumes less than 2μW including global biasing.

References

[1] M. Shoaib, N. Jha, and N. Verma, "Enabling Advanced Inference on Sensor Nodes through direct use of Compressively Sensed Signals," in *Proc. of DATE*, 2012, pp. 437–442.

[2] J. Holleman, A. Mishra, C. Diorio, and B. Otis, "A micro-power neural spike detector and feature extractor in .13 μm cmos," in *Custom Integrated Circuits Conference*, Sept 2008, pp. 333–336.

[3] E. Koutsos, S.E. Paraskevopoulou, and T.G. Constandinou, "A 1.5 uW NEO-based spike detector with adaptive-threshold for calibration-free multichannel neural interfaces," in *IEEE ISCAS*, May 2013, pp. 1922–1925.

[4] Yang-Guo Li, Qingyun Ma, M.R. Haider, and Y. Massoud, "Ultra-low-power high sensitivity spike detectors based on modified nonlinear energy operator," in *Circuits and Systems (ISCAS), IEEE International Symposium on*, May 2013, pp. 137–140.

[5] N. Lotze and Y. Manoli, "A 62mv 0.13 μm cmos standard-cell-based design technique using schmitt-trigger logic," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), IEEE International*, Feb 2011, pp. 340–342.

[6] R Quian Quiroga, Zoltan Nadasdy, and Yoram Ben-Shaul, "Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering," *Neural computation*, vol. 16, no. 8, pp. 1661–1687, 2004.

[7] 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," *Biomedical Circuits and Systems, IEEE Transactions on*, vol. 6, no. 2, pp. 87–100, April 2012.

[8] Tong Wu and Zhi Yang, "A multichannel integrated circuit for neural spike detection based on ec-pc threshold estimation," in *Engineering in Medicine and Biology Society (EMBC), 35th Annual International Conference of the IEEE*, July 2013, pp. 779–782.