Design Challenges for Sense Amplifier and Wireless Link in High-Density Neural Recording Implants

Mohamed N. Elzeftawi, C. Patrick Yue, and Luke Theogarajan
Department of Electrical and Computer Engineering
University of California, Santa Barbara
Santa Barbara, CA 93106, USA

ABSTRACT

In this paper we discuss the challenges in designing high-density biomedical neural implants. We discuss in more detail the constraints of the sense amplifier and the wireless link. Different techniques to design the front-end low-noise sense amplifier are discussed. We compare between different wireless designs and we show that trading off the transmitter power consumption and complexity with that of the external receiver will be mandatory to minimize the power dissipation and area of the implant.

INTRODUCTION

The concept that neural signal recording can be used for several useful applications, such as controlling artificial arms, started at least in the late 1960s [1-2]. Other applications are paralysis prosthetics, stroke, Parkinson's disease, prosthetics for blindness, and experimental neuroscience systems. Several attempts were conducted in recent years to implement large-scale multi-electrode neural recording [3-5]. Some experiments were conducted on, rats [6], monkeys [7], and humans [8-9].

It has been shown that the ability to guide a robotic arm to a limited number of targets in 2 and 3 dimensions can be implemented with tens of neurons [10-11]. Useful linear control in 2 and 3 dimensions with higher accuracy has been shown using several hundred neurons [7]. It is extremely likely that more sophisticated systems in the future will utilize hundreds to thousands of neural inputs. Thus, high density neural recording implants are required for more accurate recordings.

Figure 1 shows an artistic picture of a brain implant system. The implant is wirelessly powered and the neural signal recordings data go through a first hop (few cm) to an external repeater unit, then through a second hop to a monitoring station. Control signals are also wirelessly sent to the implant. It is common in the literature to refer to the control signals transmission as the "downlink", while the neural recordings transmission is referred to as the "uplink".

When the recording electrodes are implanted in the brain, faradic reactions at the electrode-tissue interface leads to DC offsets than can be as high as few volts. To protect the amplifier from saturation, input signal is usually ac-coupled by using a DC blocking input capacitor. The closed loop gain is thus given by the ratio of the input capacitor to the feedback capacitor.

There are strict constraints on the amount of implant's power dissipation due to the close proximity of the electrodes to living tissues. If the cells are exposed to elevated temperatures for extended periods of time, they will die. A temperature increase more than 3 ^{o}C above normal body temperature has been reported to lead to physiological abnormalities such as angiogenesis or necrosis [12]. In guinea pig olfactory cortical slices, aberrant activity began at 2 ^{o}C over normal [13]. Temperature increases greater than 1 ^{o}C above normal can have long-term effects on the brain tissue in an anesthetized rat [14]. Therefore implants should be designed to limit the chronic heating of surrounding tissue to less than 1 ^{o}C.

The temperature increase coefficient due to an implant's power dissipation in the brain of anesthetized cat was measured to be 0.0673 ^{o}C/mW for *in vitro*, and 0.05 ^{o}C/mW for *in vivo* conditions. The experimental results have also shown that the implanted IC, measuring roughly 6 x 6 x 2 mm^3, can safely dissipate approximately 10 mW of power [15].

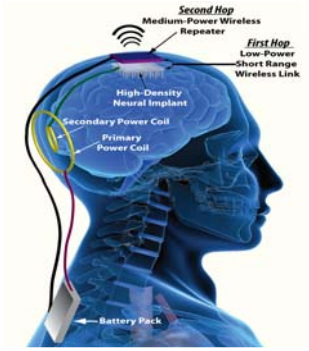

FIGURE 1. POWER AND TWO-HOP WIRELESS LINKS FOR AN IMPLANT

Another constraint on the implant system is area. The area has to be small enough to be minimally invasive and the system should preferably use few or no off-chip components to minimize the size.

Neural signals have very small magnitudes and occupy low frequencies. There are two types of neural signals: *(a) Neural action potentials (spikes):* they have magnitudes ~10's 500 µV and roughly occupy the 100 Hz 6 Khz frequency band [16]; *(b)Local field potentials (LFPs):* they have magnitudes ~10's µV 5 mV and roughly occupy less than 100 Hz frequency band. Unfortunately, flicker noise is cumbersome at such low frequencies and techniques for reducing/avoiding it should be used.

SENSE AMPLIFIER FLICKER NOISE AND ITS REDUCTION TECHNIQUES

MOSFET transistors suffer from low frequency flicker noise due to carrier trapping/detrapping at the channel/insulator interface traps. These traps have long occupation time constants which causes the flicker noise to appear only at low frequencies. MOSFET input-referred flicker noise PSD is inversely proportional to the transistor area and the operating frequency.

There are mainly two widely used techniques in the literature to deal with flicker noise, namely, auto-zeroing (sometimes referred to as correlated double sampling) and chopper stabilization techniques.

A- Auto-zeroing (AZ) / Correlated double sampling (CDS)

This technique is usually used to eliminate the DC offset voltage at the input of an amplifier. The block diagram of auto-zeroing/correlated double sampling technique is shown in figure 2. In the offset sampling phase (Φ_1), only the offset and noise are sampled and stored by using S&H circuit. In the input sampling phase (Φ_2), the signal is sampled, but with the previous value subtracted from it. This leaves the signal pure from offset and low frequency noise. Notice that this is only valid if the previous noise sample correlate well with the present sample. This is true for DC offsets and low-frequency flicker noise. However, thermal/shot noise doesn't

correlate well and their samples add together because they are under-sampled. This is commonly referred to as "noise folding".

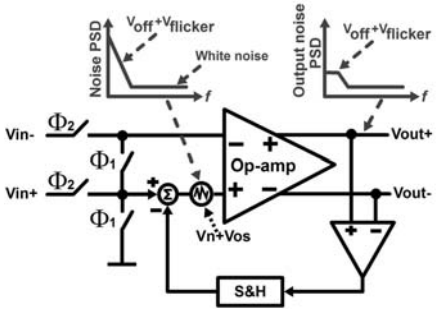

FIGURE 2. AZ / CDS TECHNIQUE

B- Chopper stabilization

Chopper stabilization is another technique that avoids flicker noise effect instead of just reducing it. The block diagram of the chopping technique is shown in figure 3. The key idea is to up-modulate the input signal to a high frequency once it has been received before much flicker noise has been added to it by minimizing the number of MOS transistors in the signal path before the signal is up-modulated or "chopped". The output, after demodulation, will consist of the even harmonics of the signal and the odd harmonics of the flicker noise and thus must be low-pass-filtered. The following question needs to be addressed:

What is the best node to chop at?

Usually chopping is done at the input node right before the operational amplifier. Since the neural recording are high impedance nodes (>100 KΩ), chopping at the input nodes will cause charge injection right into the incoming signal. Although that technique was used in [17], the authors had to keep their chopping frequency low enough (4 KHz) so that the input differential impedance of the operational amplifier is much higher (>8 MΩ) than the electrode impedance to avoid loading it. This approach has two limitations: the sum of the signal bandwidth and the flicker noise corner frequency have to be lower than the used chopper frequency (4 KHz in case of [17]) which forces the used devices to be really large and sets a lower limit on their area. The other limitation is that harmonics at the multiples of the chopping frequency will appear within the band of interest, even after the LPF. The authors in [17] may not have faced that second problem because they were recording LFPs which have frequency components much lower than neural spikes.

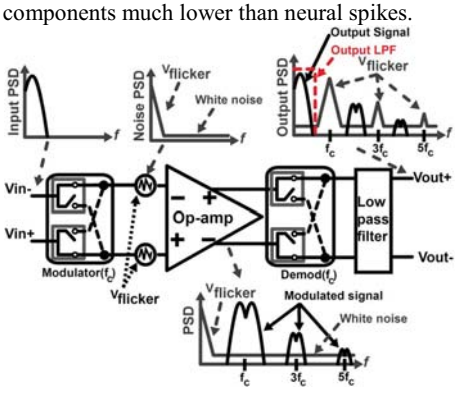

FIGURE 3. CHOPPING TECHNIQUE

Another approach is to chop at low impedance nodes instead of chopping at the high input impedance nodes. This can be done using cascode devices and chop just before them as shown in figure 4. The downside is that the input devices' flicker noise will be added to the signal. However, this noise can be minimized by increasing their layout area (but there is no lower limit as the previous case).

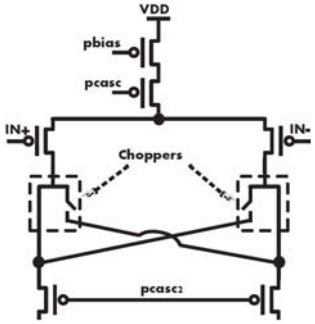

FIGURE 4. CHOPPING AT LOW IMPEDANCE NODES

Comparing chopping to auto-zeroing, it is obvious that chopping is better in terms of dealing with flicker noise because one can almost completely avoid flicker noise instead of just decreasing it.

It is worth mentioning that a commonly used figure of merit to compare different sense amplifiers is the noise efficiency factor (NEF) originally proposed in [18] which is a very important figure of merit. However this figure of merit doesn't take into account the total amplifier area. Using large amplifier area, by using large devices that contribute the most to the input referred noise, improves the NEF but limits the scalability of the system to achieve high-density neural implants that uses several hundred or few thousand electrodes. The authors in [17,19] additionally used a resistive degeneration technique for the amplifier first stage's active load. They had to size the active load area to be 6.25 larger than the input transistor and used 200 KΩ resistors in [17] and 2.5 larger than the input transistors and 240 KΩ resistors in [19]. This of course improves the NEF by minimizing the thermal/shot noise contribution from the active loads but leads to a large penalty in area. In addition, high resistance values may not be available in all the processes and also have large mismatches. Considering these factors we believe that other topologies should be used to minimize the sense amplifier area that avoids the use of such large resistances and active load devices.

Furthermore, we believe an alternative figure of merit should be used to provide a fair comparison between different sense amplifiers performances. This alternative figure of merit should factor in the total amplifier area. We propose the high density noise efficiency factor (HDNEF) as a figure of merit which calculates the NEF normalized to the number of amplifiers, N, that can fit a 1 mm^2 area, given an amplifier with a total area A (mm^2).

$$HDNEF = NEF / N,$$

$$N = 1 / A,$$

Table 1 shows a comparison between different neural amplifiers in the literature. The table's 2nd entry has a better NEF than the 3rd one, but a worse HDNEF. Similarly, the 1st and 3rd entries have the same NEF, but the HDNEF in the 3rd entry is much better. It is apparent that using the NEF as the only figure of merit can be quite misleading as a fair comparison for HDNI.

Table 1. Comparison between different neural amplifiers

Ref.	BW (Hz)	P_d / LNA (µW)	Input Device W/L(µm)	Noise floor (nV/rtHz)	NEF	Area (mm^2)	HDNEF
[17]	0.05–180	2	80/8	100	4.6	>1.7	7.82
[20]	25m–7.2K	80	800/4	25.9	4.0	0.16	0.64
[5]	1–5K	15	75/2	52.7	4.6	0.04	0.184

We were able to design a chopper sense amplifier that occupies less than 100 x 100 μm^2 and consumes total current less than 1.2 µA from a 1.5 V supply. More details and measurement results will be published in a subsequent publication.

A typical block diagram of a chopper-based front end system using, for instance, 1024 electrodes is shown in figure 5. Such system will have an adaptive compression block, to take advantage of the correlation between nearby neural signals, low power ADC array, a digital compression block, and an RF transceiver.

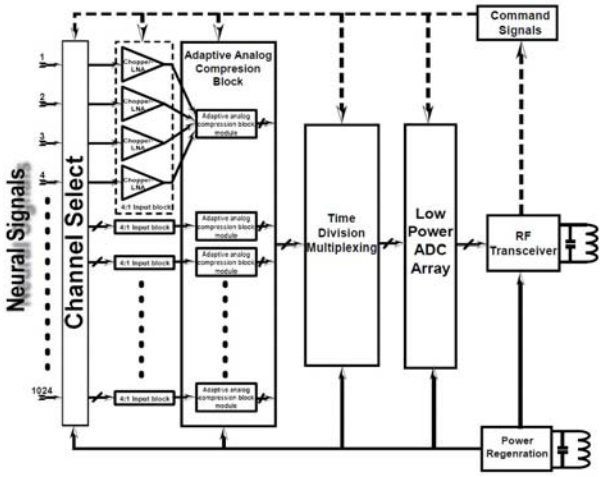

FIGURE 5. CHOPPING AT LOW IMPEDANCE NODES

OPEN LOOP VS. CLOSED LOOP CHOPPING

Simulink simulations were done to show the difference between using open loop and closed loop chopping. Figure 6A shows the model used for a 2-stage low-noise sense amplifier where non-idealities such as flicker noise, thermal noise, non-linearities, saturation, slew rate, etc are taken into account. Frequencies f_{sig} =10 KHz, and f_{chop}=100 KHz were used. We assumed a thermal noise floor of 85 nV/rtHz and a flicker noise corner frequency of 1 KHz based on Cadence simulations of our low-noise sense amplifier. The overall unity gain frequency of the amplifier is 1MHz. Figure 6B shows a closed loop system that has a mid-band gain of 40 dB and a high pass corner frequency at 1Hz. Notice that the up-modulated signal is an error signal and that the LPF is after the feedback signal. Figure 7 shows the output spectrum just before the LPF. The closed loop system is more robust and reduces the harmonics as can be seen from the figure. This is because the system keeps on trying to minimize the error signal at the input over time. That's a key advantage of closed loop systems.

HIGH DENSITY IMPLANT WIRELESS LINK

As we mentioned earlier that the total power consumption should be less than roughly 10 mW to avoid damaging the cells. There have been several attempts to implement the wireless link for a neural implant [21-25]. For a high density neural implant, the wireless link should consume no more than ~ 2 mW. The downlink, from external unit to the implant, carries only control signals and thus have low data rates. The uplink, from implant to the external unit, carries the neural activity and therefore has huge data rates. For 1024 channels the uplink data rate is very high. Best case scenario will yield a data rate of 12.8 Mbps if we assumed 25% neural activity, 50% compression rate, and 100 Kbps ADC. This means that the transmitter efficiency has to be better than 156 pJ/bit. Worst case scenario if we have full uncorrelated neural activity, in which case the data rate is 102.4 Mbps which imposes transmitter efficiency better than 20 pJ/bit. These are very tight constraints and digital compression has to be done on chip for such large electrode count.

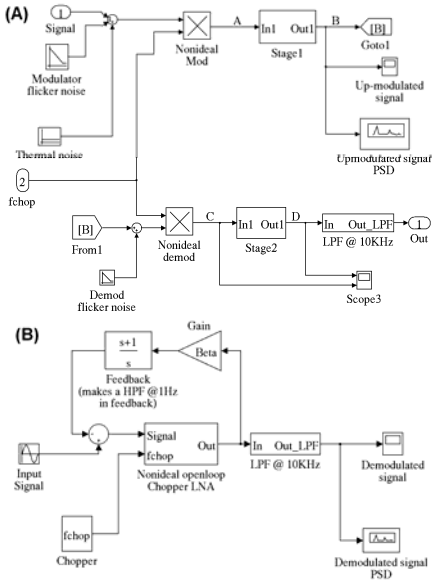

FIGURE 6. (A) OPEN-LOOP LNA. (B) CLOSED-LOOP SYSTEM

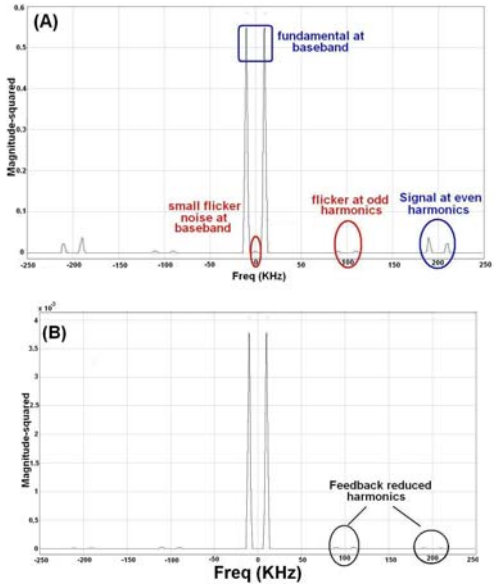

FIGURE 7. (A) OPEN-LOOP (B) CLOSED-LOOP, SPECTRUM BEFORE LPF

In [21] the authors were able to achieve 330 Kbps up to a separation of 13cm with a BER of $10^{-2.5}$. They implemented a fully integrated FSK transmitter with power consumption of 6.75 mW and energy efficiency of 20.5 nJ/bit. This is, however, very inefficient. In [22] the authors used Load Shift Keying (LSK) modulation and were able to achieve a much lower power consumption of ~1 mW, a BER of 10^{-5} for 4 cm separation but on the expense of a much worse data rate of 32 Kbps and a degraded energy efficiency of ~30 nJ/bit. In [23] the authors designed a 350 µW FSK/MSK direct modulation transmitter with a data rate of 120 Kbps whose efficiency is only 2.9 nJ/bit. The receiver was a 400 µW on-off keying (OOK) super-regenerative receiver (SRR) with a sensitivity of -93 dBm and data rate of 120 Kbps consuming 3.3 nJ/bit.

In [24] they used different modulation schemes in the implant and the external unit. For the downlink, they used OOK which encode the data with RZ pulse-width-modulation (PWM) technique. This eliminates the need of a PLL in the implant and the clock can be recovered from the data itself, taking advantage of the low data rate

in the downlink. This is a common approach that is used to minimize power consumption [26]. For the uplink, they used NRZ impedance-modulation to push as much of the power and complexity to the external unit. In the uplink mode the implant power was 100 µW with a data rate of 4 Mbps and BER of 10^{-3} (or 2.8 Mbps for a BER of 10^{-6}) over a 2 cm separation. In the downlink mode the implant power was 140 µW with a data rate of 300 Kbps and a BER < 10^{-6}. Their best uplink energy efficiency was ~25 pJ/bit. In [25] the authors used IR-UWB (3-5 GHz) transmitter with a data rate of 90Mbps and power consumption of 1.6mW and efficiency of 18 pJ/bit. They also traded off the transmitter power consumption and complexity with that of the external receiver. The two main advantages of the implementation in [25] are the transmitter efficiency and the potential of integrating the antenna on-chip. We can see that such trade off will be mandatory to minimize power consumption and area of the implant.

Figure 8 shows the implant transmit power vs. data rate for a BER of 10^{-5}. For a high density neural implant it is better to be at the lower right corner of the graph. Therefore [25] provides the best performance with respect to other papers in the same figure.

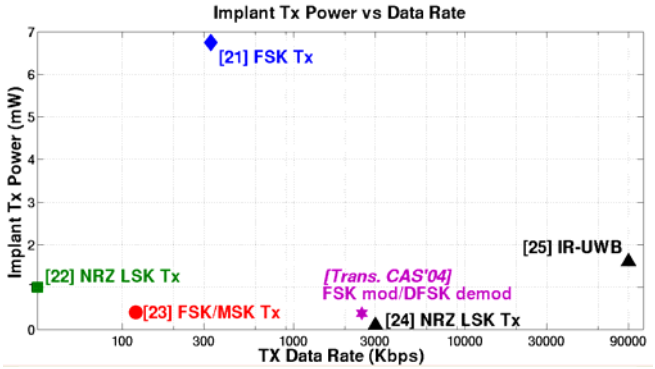

FIGURE 8. IMPLANT TX POWER VS. DATA RATE FOR BER≤10^{-5}

CONCLUSION

High density neural recording implants are required for more accurate recordings. The power and area constraints are very strict in such systems. To reduce sense amplifier area, techniques (like chopping) to avoid flicker noise are mandatory to avoid large input devices. Simulink simulations were presented to show that using chopping prevents the neural signal and the flicker noise from being in the same frequency. HDNEF was proposed to provide a fair comparison by taking the total amplifier area into consideration. Wireless link for the implanted system has to minimize the power dissipation in the implant by pushing most of it to the external receiver.

REFERENCES

[1] E. E. Fetz, *Operant conditioning of cortical unit activity*. Science, 1969, vol. 163, pp. 955-957.

[2] D. R. Humphrey, E. M. Schmidt, and W. D. Thompson, *Predicting measures of motor performance from multiple cortical spike trains*. Science, 1970, vol. 170, pp. 758-762.

[3] R. R. Harrison et al., *Wireless Neural Recording With Single Low-Power Integrated Circuit*. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2009, vol. 17.

[4] L. S. Theogarajan, *A Low Power Fully Implantable 15-Channel Retinal Stimulator Chip*. IEEE J. Solid-State Circuits, 2008, vol. 43, no. 10, pp. 2322-2337.

[5] J. Aziz et al., *256-Channel Neural Recording and Delta Compression Microsystem with 3D Electrodes*, IEEE J. Solid State Circuits, 2009, vol. 44, no. 3, pp. 995-1005.

[6] J. K. Chapin et al., *Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex,* Nature Neurosci., 1999, vol. 2, pp. 664–670.

[7] J. Wessberg et al., *Real-time prediction of hand trajectory by ensembles of cortical neurons in primates,* Nature, 2000, vol. 408, no. 6810, pp. 361–365.

[8] P. R. Kennedy and R. A. E. Bakay, *Restoration of neural output from a paralyzed patient by a direct brain connection*, NeuroReport, 1998, vol. 9, pp. 1707–1711.

[9] P. R. Kennedy et al., *Direct control of a computer from the human central nervous system*, IEEE Trans. Neural Systems and Rehabilitation, 2000, vol. 8, no. 2., pp. 198-202.

[10] D. M. Taylor, S. I. Tillery, and A. B. Schwartz, *Direct cortical control of 3D neuroprosthetic devices.* Science, 2002, vol. 296, no. 5574, pp. 1829–1832.

[11] M. D. Serruya et al., *Instant neural control of a movement signal*. Nature, 2002, vol. 416, no. 6877, pp. 141–142.

[12] T. M. Seese et al., *Characterization of tissue morphology, angiogenesis, and temperature in the adaptive response of muscle tissue in chronic heating*, Lab. Investigation, vol. 78, pp. 1553–1562, 1998.

[13] T. Fujii and Y. Ibata, *Effects of heating on electrical activities of guinea pig olfactory cortical slices*, Eur. J. Physiol., vol. 392, pp. 257–260, 1982.

[14] J. C. LaManna et al., *Stimulus-activated changes in brain tissue temperature in the anesthetized rat*, Metabolic Brain Disease, vol. 4, pp. 225–237, 1989.

[15] S. Kim et al., *In Vitro and in vivo study of temperature increases in the brain due to a neural implant*, Int. IEEE/EMBS Conf. Neural Engineering, 2007, pp. 163-166.

[16] K. Najafi, and K. D. Wise, *An implantable multielectrode array with on-chip signal processing*, IEEE J. Solid State Circuits, 1986, vol. 21, no. 6, pp. 1035-1044.

[17] T. Denison et al., *A 2 µW 100 nV/rtHz Chopper-Stabilized Instrumentation Amplifier for Chronic Measurement of Neural Field Potentials*. IEEE J. Solid State Circuits, 2007, vol. 42, no. 12, pp. 2934-2945.

[18] M. S. Steyaert, W. M Sansen, and C. Zhongyaun, *A micropower low-noise monolithic instrumentation amplifier for medical purposes*, IEEE J. Solid State Circuits, 1987, vol. 22, no. 6, pp. 1163-1168.

[19] W. Wattanapanitch, M. Fee, and R. Sarpeshkar, *An Energy-Efficient Micropower Neural Recording Amplifier*, IEEE Trans. on BIOCAS, 2007, vol. 1, no. 2, pp. 136-147.

[20] R. R. Harrison, *The Design of Integrated Circuits to Observe Brain Activity,* Proc. of IEEE, 2008, vol. 96, pp. 1203-1216.

[21] R. R. Harrison et al., *A Low-Power Integrated Circuit for a Wireless 100-Electrode Neural Recording System*, IEEE J. Solid State Circuits, 2007, vol. 41, no. 1, pp. 123-133.

[22] M. Mollazadeh et al., *Wireless Multichannel Acquisition of Neuropotentials*, IEEE BIOCAS 2008, pp. 53-56.

[23] J. L. Bohorquez et al., *A 350 µW CMOS MSK Transmitter and 400 µW OOK Super-Regenerative Receiver for Medical Implant Communications*, IEEE J. Solid-State Circuits, 2009, vol. 44, no. 4, pp. 1248–1259.

[24] S. Mandal, and R. Sarpeshkar, *Power-Efficient Impedance-Modulation Wireless Data Links for Biomedical Implants*, IEEE Trans. on BIOCAS, 2008, vol. 2, no. 4, pp. 301-315.

[25] M. Sung Chae et al., *A 128-Channel 6 mW Wireless Neural Recording IC With Spike Feature Extraction and UWB Transmitter*, IEEE Trans. TNSRE, vol. 17, no. 4, pp. 312-321.

[26] W. Liu, et al., *Neuro-stimulus chip with telemetry unit for retinal prosthetic device*, IEEE J. Solid-State Circuits, 2000, vol. 35, no. 10, pp. 1487–1497.