A Micropower Neural Recording Amplifier with Improved Noise Efficiency Factor

Vahid Majidzadeh, Alexandre Schmid, and Yusuf Leblebici
Microelectronic Systems Laboratory (LSM)
Ecole Polytechnique Fédérale de Lausanne (EPFL) 1015 Lausanne, Switzerland
e-mail: vahid.majidzadeh@epfl.ch

***Abstract*—This article presents a neural recording amplifier suitable for large-scale integration with multi-electrode arrays (MEAs) in very low-power microelectronic cortical implants. The proposed amplifier is the most energy-efficient structure reported to date, which achieves an effective noise efficiency factor (NEF) smaller than the theoretical limit that was claimed in literature for any existing amplifier (NEF=2.02). The proposed technique, which is referred to as partially OTA sharing technique, achieves a significant reduction of power dissipation as well as silicon area, in addition to the very low NEF. The effect of systematic mismatch on crosstalk between adjacent channels and the trade-off between noise and crosstalk are theoretically analyzed. For an array of four neural amplifiers, simulation results show a midband gain of 39.2 dB and a -3 dB bandwidth from 10 Hz to 10.6 kHz. The input referred noise is simulated to be 2.21 μV_{rms} and the power consumption is 7.92 µW from 1.8 V supply, which refers to NEF=1.8. The worst-case crosstalk within the desired bandwidth is -46.1 dB.**

Keywords— Cortical implants, Neural Amplifier, noise efficincy factor, Partially OTA sharing technique, crosstalk

I. INTRODUCTION

Minimally invasive monitoring of the electrical activity of specific brain areas using implantable microsystems offers the promise for diagnosing brain diseases, as well as detecting and identifying neural patterns which are specific to a behavioral phenomenon. Neural pattern classification and recognition require simultaneous recording from large number of neurons [1]. However, massive recording *in-vivo* requires complying with severe safety requirements. For example, the maximum temperature elevation in brain tissue due to the operation of the implant should be kept at less than 1°C [2]. This requirement constrains the maximum allowable power dissipation in the implant, which reaches at most 4 to 5mA drawn from a 1.8 V supply [3]. The limited total power budget impose severe limitations to circuit design, especially when the number of recording sites increases to a range of one to several hundred for typical MEAs.

Front-end neural amplifiers are the most important building blocks in developing implantable cortical microsystems. Low-power and low-noise operation, stable DC interface with the sensors (microprobes), and small silicon area are the main design specifications of these amplifiers. The power dissipation is dictated by the input referred thermal noise of the amplifier, where the trade-off is expressed in terms of NEF [4]. The contribution of flicker noise to the input referred noise of the amplifier can be reduced to a negligible level either by proper sizing of the input devices or by using circuit techniques such as chopper stabilization [5]. In practice, the total input referred noise of the amplifier should be kept smaller than the background noise of the electrode (~ 5μV_{rms}). Neural amplifiers should pass the action potential signal spanning over the frequency range of 100 Hz-10 kHz, while rejecting the large DC offset (up to several hundred of millivolts) generated at the electrode-tissue interface. Loading the recording site with a large value resistor [6], active low-frequency suppression [7], and capacitive feedback network [8],[9] are three main techniques applied to reject low-frequency components. The last scheme provides a robust suppression without using additional biasing circuitry [6] or any active circuitry [7] at the cost of increased silicon area, due to bulky on-chip capacitors. Only few amplifiers reported in recent literatures fulfill the noise, power, and area requirements explained above [8], [9]. Useful design techniques introduced in [8] results in a NEF=4, which is close to the theoretical limit (NEF=2.9) for that particular OTA structure. The authors in [9] show that the minimum NEF for any existing amplifier using a differential pair as input stage is equal to 2.02. Moreover, they measure a NEF=2.67 using the folded-cascode OTA structure, which is in very close agreement with theory.

In this paper, we demonstrate that NEF can be reduced below the theoretical limit stated in [9] by proposing the partially OTA sharing technique. The proposed technique not only improves the NEF figure of merit, but also reduces the silicon area, mostly dominated by on-chip feedback capacitors.

This paper is organized as follows. Section II describes the proposed technique and side effects. Section III presents simulation results of the amplifier configured for action potential recording. Finally, IV concludes the paper.

II. MICROPOWER NEURAL AMPLIFIER

Fig. 1(a) shows the conventional structure for an array of n neural amplifiers, which is adopted from [8] with slight modifications. Diode connected transistors $M_{3\text{-}8}$ act as a high value resistor and adjust the high-pass cut-off frequency of the amplifier. The midband gain A_d is set by C_1/C_2 and the low-pass cut-off frequency is approximately placed at $g_m/(A_d C_L)$, where g_m is the transconductance of the input differential pair

The authors gratefully acknowledge the support of the Swiss NSF, under projects number 200021-113883 and 200020-122082

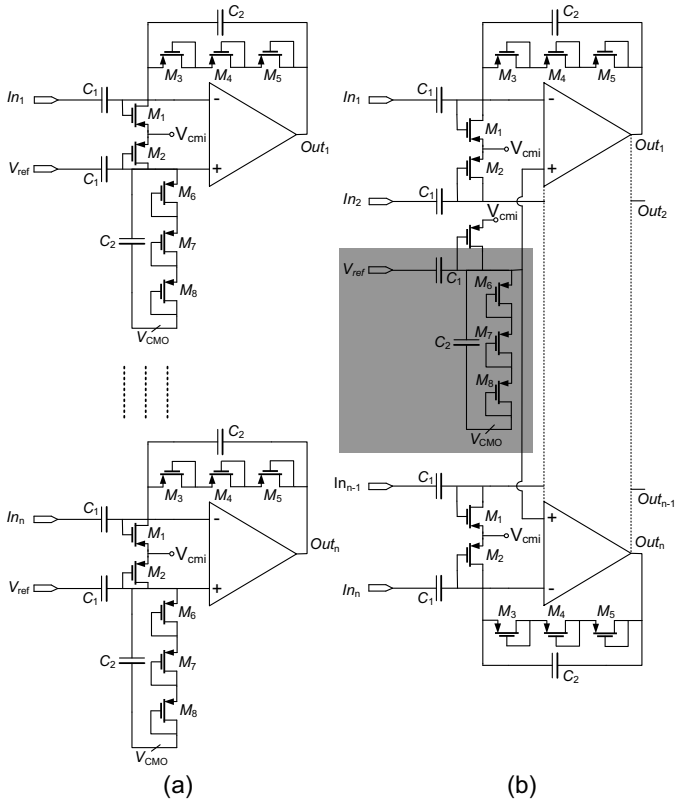

Fig 1. (a) Conventional structure: array of n neural amplifier, and (b) proposed partially OTA sharing structure: array of n neural amplifier.

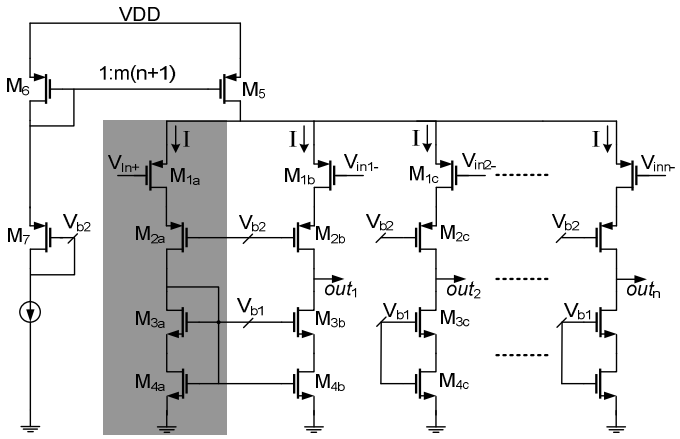

Fig 2. Circuit schematic of the partially OTA sharing structure.

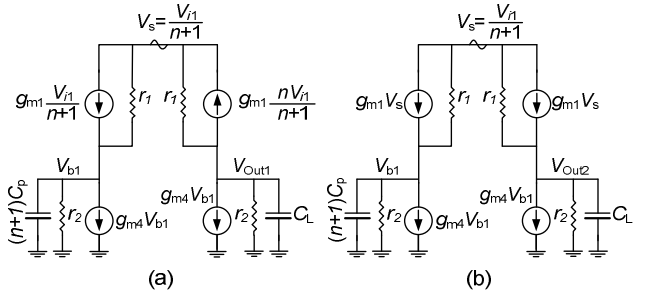

Fig 3. (a) Small-signal model of the partially OTA sharing structure for a single channel, (b) Small signal model for crosstalk analysis.

and C_L is the effective load capacitance of the amplifier. The OTA benefits from the telescopic-cascode structure, which offers the best noise-power trade-off. However, it requires different input and output common-mode voltages to be properly biased. $M_{1\text{-}2}$ in series with $M_{3\text{-}8}$ acts as voltage divider and sets the input common-mode voltage. The contribution of $M_{1\text{-}2}$ to the input referred noise of the amplifier can be made negligible by proper sizing of these devices. Since a large numbers of amplifiers share the same V_{cmi}, overheads in area, power, etc, related to the generation of this voltage are considered to be negligible. In a conventional structure, the power consumption of the array linearly increases with the number of amplifiers. The effective power consumption of each amplifier in the array is equal to that of any individual amplifier. Therefore, NEF is limited by the amplifier topology, which has been proven to be higher than 2.02 [9].

The total power consumption of the amplifier array can be reduced by applying the partially OTA sharing technique. Fig. 1(b) shows the proposed structure. Each n amplifier in the array shares the passive part corresponding to the reference electrode (V_{ref}), which is shown in shaded box. This helps reducing the silicon area, thanks to sharing the bulky capacitor C_1. The improvement factor depends on the number of shared amplifiers. Fig. 2 shows the circuit schematic of the amplifiers in the proposed technique. The non-inverting input of the amplifiers is shared, which is referred to as partially OTA sharing technique. The total current drawn from supply, excluding the bias circuitry, is calculated to be $(n+1)$I and $2n$I for the proposed and conventional architectures respectively, where I refers to the bias current of each inverting input. For example, the total power consumption and area occupied by feedback capacitors are reduced by 37.5% for n=4. The diode connected transistor M_{3a} mitigates the need for dedicated bias circuitry to generate V_{b1}. In order to approach the minimum limit of NEF, NMOS load devices $M_{4a,b,...}$ are biased in strong inversion, while all other devices are biased in weak inversion regime of operation.

A. Frequency Response

Fig. 3 shows the small-signal model of the partially OTA sharing structure for a single channel, when all other channels are connected to ground. g_{m1} is the transconductance of the input devices, r_1 and r_2 are the equivalent output resistance of the PMOS cascode devices (M_{1a}…-M_{2a}…) and NMOS cascode devices ($M_{3a...}$-$M_{4a...}$), C_p refers to the parasitic gate capacitance of each device $M_{4a...}$, g_{m4} is the transconductance of the tail devices $M_{4a...}$, and finally C_L is the effective load capacitance.

In contrary to the conventional structure, the source terminal of the input devices is not a virtual signal ground as shown in Fig. 3(a). In practice this node includes the superposition of attenuated input signals of all channels, $\Sigma V_i/(n+1)$. The signal transfer function from each input to the corresponding output is as follows:

$$H_d(S) = \frac{V_{out1}(S)}{V_{in1}(S)} = -g_m\left(r_1 \| r_2\right)\frac{1+\tau_z S}{\left(1+\tau_1 S\right)\left(1+\tau_2 S\right)} \qquad (1)$$

Where

$$\tau_z = nC_p / g_{m4},\ \tau_1 = (r_1 \| r_2) C_L,\ \tau_2 = (n+1) C_p / g_{m4} \quad (2)$$

The dominant pole occurs at the output as expected, and there is a pole-zero doublet effect due to parasitic pole τ_2 and zero τ_z. If the amplifier phase margin is designed to be large enough which guarantees stability for n=1, then the amplifier will be stable for larger values of n, since, τ_z/τ_p approaches unity with increasing n.

B. Channel Crosstalk

As mentioned before, the common source terminal of the input transistor gathers the superposition of attenuated input signals from all channels. Since there is systematic mismatch due to the presence of NMOS current mirror, a small fraction of each input signal leaks to the non-corresponding output, which is referred to as crosstalk between channels. Fig. 3(b) shows the small-signal model for crosstalk analysis, where the desired input/output is the first channel, the non-corresponding output is the second channel, and total number of shared amplifiers is n. The transfer function from the input of the first channel to the output of the second channel is as follow:

$$H_C(S) = \frac{V_{out2}(S)}{V_{in1}(S)} = \frac{g_{m1}}{(n+1) g_{m4}} \frac{1+\tau_{zc} S}{(1+\tau_1 S)(1+\tau_2 S)} \quad (3)$$

Where $\tau_{zc}=(n+1)C_p(r_1\|r_2)$ and τ_1 and τ_2 are as same as (2). Thus, crosstalk between these two channels can be characterized as:

$$\begin{aligned} Crosstalk_{1-2}(S) &= 20 \log \left| \frac{H_c(S)}{H_d(S)} \right| \\ &= 20 \log \left| \frac{1}{(n+1) g_{m4} (r_1 \| r_2)} \frac{1+\tau_{zc} S}{1+\tau_z S} \right| \end{aligned} \quad (4)$$

The statement of (4) demonstrates the trade-off between noise and crosstalk; low input referred noise necessitates a small value of g_{m4}, which can be realized by operating of $M_{4a,b,...}$ in strong inversion regime. However, crosstalk between two channels increases by decreasing g_{m4}.

C. Noise Efficiency factor

The input referred noise of the amplifier is composed of flicker and thermal noise. The flicker noise contribution can be made negligible by proper sizing of the $M_{1a,b,...}$ and $M_{4a,b,...}$, while thermal noise is usually limited by limiting power consumption. The theoretical circuit analysis reveals that the input referred noise power spectral density (excluding the contributions of the flicker noise and noise of the bias circuit) is as follows:

$$\overline{v_{ni}^2} = \frac{1}{g_{m1}} \left(\frac{4kT}{\kappa} + \frac{16kT}{3} \frac{g_{m4}}{g_{m1}} \right) \quad (5)$$

Where k is Boltzmann's constant and κ is the gate coupling factor, which is the reciprocal of the subthreshold slope factor n. The noise-power trade-off is characterized by the NEF [4]:

$$\mathrm{NEF} = v_{ni,rms} \sqrt{\frac{2I_{tot}}{\pi \cdot U_T \cdot 4kT \cdot BW}} \quad (6)$$

Where $v_{ni,rms}$ is the total input referred noise, BW is the bandwidth of the amplifier, U_T refers to the thermal voltage, and $I_{tot}=(n+1)I/n$ is the average current consumption for each amplifier in the proposed architecture, where I is the bias current of each input devices $M_{1a,b,...}$. Thus, the theoretical limit of the NEF of the proposed architecture, $g_{m4}<<g_{m1}$, is:

$$\mathrm{NEF} = \frac{\sqrt{2}}{\kappa} \cdot \sqrt{\frac{n+1}{2n}} \quad (7)$$

Where $\sqrt{2}/\kappa$ is the theoretical limit of the NEF for any existing amplifier [9]. The impressive result in (7) shows that for $n\geq 2$, the NEF reduces below the theoretical limit stated in [9]. This statement will also be verified by circuit simulation.

III. Simulation Results

In order to validate the theoretical achievement mentioned in the previous section, an array of neural amplifiers for different values of n is designed in a 0.18 µm CMOS technology. Fig. 4 shows the simulated frequency response of the amplifier for n=8, and considering different corner cases of the process at body temperature. The midband gain is 39.2 dB with the -3dB frequency bandwidth of 10Hz<BW<10.6 kHz for typical conditions. Fig. 5 depicts the input referred noise of the amplifier with two different setups; the solid line shows the output noise divided by H_d(s) and the dotted line shows the output noise divided by midband gain which is usually used to calculate the input referred noise. For our design the total input referred noise integrated from 100Hz to 100 kHz is 2.21 μV_{rms}. The input referred noise is same for any arbitrary value of n. However, the effective power of each amplifier is scaled down by increasing n. The resulting NEF is plotted in Fig. 6. For a single amplifier, the expected NEF is equal to 2.33 and decreases to 1.71 for n=8. Fig. 7 shows the crosstalk between two channels, which is simulated from the desired input electrode In_1 to the non-corresponding output, V_{out2}. The worst-case crosstalk in pre-layout simulations is equal to -46.1 dB, which is negligible considering the intrinsic spatial and temporal correlations between the channels. The crosstalk decreases to -44 dB in post-layout simulations, which indicates that systematic mismatch in (4) is more pronounced than layout induced mismatch.

A test prototype has been fabricated considering n=4, but measurement results are not yet available at the moment of

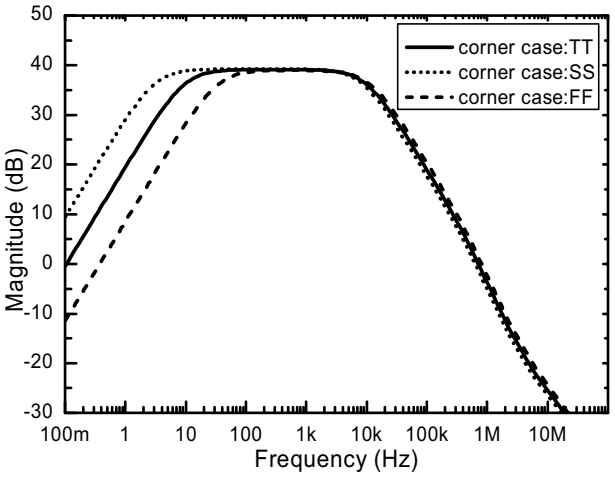

Fig 4. Frequency response of the amplifier for n=8.

publishing. Table I represents summary of the results of the fabricated prototype and comparison with recently published works. The main disadvantage of the proposed technique is complex layout required in order to achieve high CMRR. Pre-layout simulations show 85 dB of CMRR, while it drops to 63 dB in post-layout simulations. Moreover, Monte-Carlo simulations show that standard deviation of the CMRR is less than 1.3 dB for 200 runs, which indicates that layout induced mismatch is much important than process induced mismatch.

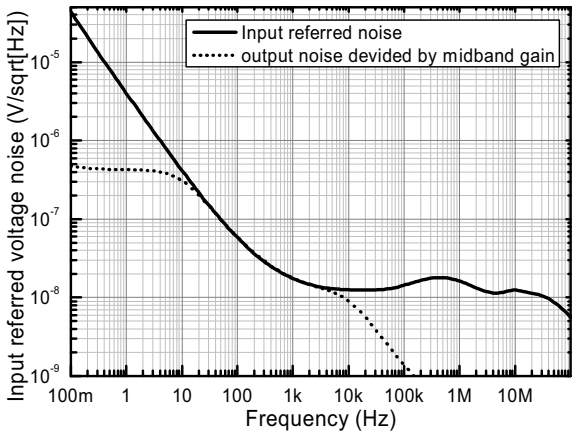

Fig 5. Input referred noise of the amplifier with two different setups.

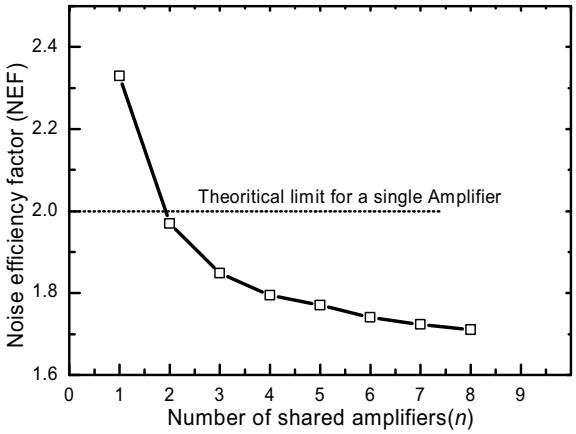

Fig 6. NEF versus number of shared amplifiers, *n*.

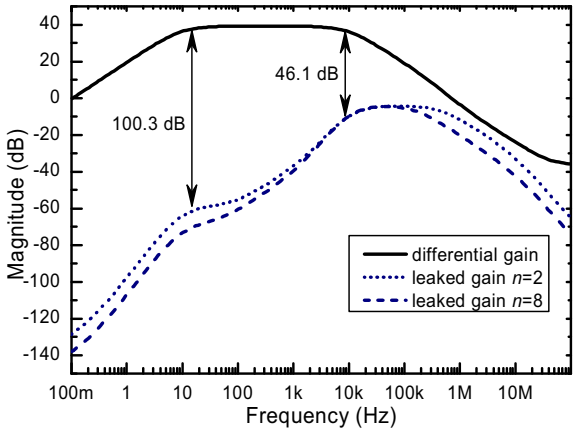

Fig 7. Crosstalk between two channels.

TABLE I. SUMMARY OF RESULTS AND COMPARISON WITH OTHER WORKS

Parameter	[8]	[9]	This Work (*n*=4)	
Technology (CMOS)	1.5 µm	0.5 µm	0.18 µm	
Supply Voltage	±2.5 V	2.8 V	1.8 V	
Supply Current	16 µA	2.7 µA	4.4 µA	
Gain	39.5 dB	40.8 dB	39.2 dB	
Bandwidth (Hz)	25 <f<7.2 k	45 <f<5.3 k	10 <f<10.6 k	
Input Referred Noise	2.2 μV_{rms}	3.06 μV_{rms}	2.21 μV_{rms}	
NEF	4.0	2.67	1.8	
CMRR (dB)	83	66	*Pr-L	85
			*Po-L	63
PSRR (dB)	85	75	Pr-L	64
			Po-L	63
Crosstalk (dB)	-64 dB	-	Pr-L	-46.1
			Po-L	-44

* Pr-L and Po-L refers to as pre-layout and post-layout simulation, respectively.

IV. CONCLUSION

A micropower neural recording amplifier is presented. It outperforms any existing amplifiers in terms NEF by implementing a novel technique called partially OTA sharing. A noise efficiency factor of 1.8 is achieved for an array of four amplifiers, *n*=4. Moreover, a 37.5% improvement of power consumption, and reduction of the occupied silicon area (feedback capacitors) are other advantages of the proposed technique. A trade-off between NEF and crosstalk has to be considered for target specifications, and full theoretical developments are provided.

REFERENCES

[1] J. Wessberg, C. Stambaugh, J. Kralik, P. Beck M. Laubach, J. Chapin, J. Kim, S.Biggs, M. Srinivasan, and M. Nicolelis, "Real-time rediction of hand trajectory by ensembles of cortical neurons in primates," *Nature*, vol. 408, pp. 361–365, Nov. 2000.

[2] "IEEE standard for safety levels with respect to human exposure to radio frequency electromagnetic fields, 3 kHz to 300 GHz, " *IEEE Std. C*95.1-2005, 2006.

[3] K. M. Silay, C. Dehollain, and M. Declercq, "Numerical analysis of temperature elevation in the head due to power dissipation in a cortical implant," in *Proc. IEEE* EMBC'08, pp. 951–956.

[4] M. Steyaert, W. Sansen, and C. Zhongyuan, "A Micropower Low-Noise Monolithic Instrumentation Amplifier for Medical Purposes," *IEEE J. Solid-State Circuits*, vol. sc-22, no. 6, pp. 1163-1168, Dec. 1987.

[5] C. C. Enz, G. C. temes, "Circuit techniques for reducing the effects of op-amp imperfections: autozeroing, correlated double sampling, and chopper stabilization," *Proceedings of the IEEE*, vol. 84, no. 11, pp. 1584-1614, Nov. 1996.

[6] P. Mohseni, K. Najafi, "A Fully Integrated Neural Recording Amploifier With DC Input Stablization," *IEEE Trans. Biomedical Engineering*, vol. 51, no. 5, pp. 832-837, May. 2004.

[7] B. Gosselin, M. Sawan, C. A. Chapman, "A Low-Power Integrated Bioamplifier With Active Low-Frequency Suppression," *IEEE Trans. Biomedical Circuits Syst*, vol. 1, no. 3, pp. 184-192, Sep. 2007.

[8] R. R. Harrison, C. Charles, "A Low-Power Low-Noise CMOS Amplifier for Neural Recording Applications," *IEEE J. Solid-State Circuits*, vol. 38, no. 6, pp. 958-965, June. 2003.

[9] W. Wattanapanitch, M. Fee, and R. Sarpeshkar, "An Energy-Efficient Micropower Neural Recording Amplifier," *IEEE Trans. Biomedical Circuits Syst*, vol. 1, no. 2, pp. 136-147, June. 2007.