Abstract—A systematic design guideline is presented for the noise performance of preamplifier for semiconductor neural probe which contains on-chip electronic circuitry. The overall signal-to-noise ratio (SNR) is calculated considering the spectral characteristics of the measured extracellular action potential and the low-frequency noise spectrum of the CMOS device from typical fabrication processes. An analytical expression of the output noise power is derived, and utilized to tailor the frequency response and device parameters which are controllable by the circuit designer. An analysis of the output SNR of a two-stage CMOS differential amplifier is given and the major factors which have significant effects on the SNR are determined. We showed that a little deviation of the input device sizes and transconductance ratio from the optimal values can significantly deteriorate the SNR. Quantitative information of the preamplifier circuit parameters for satisfactory noise performance is provided.

Index Terms—Active neural probe, CMOS, differential amplifier, extracellular recording, signal-to-noise ratio.

I. INTRODUCTION

THE RECORDING of an extracellular action potential by microelectrodes represents a major neurophysiological tool for the investigation of the nervous system at the cellular level. Metal microelectrodes are typically used for this purpose [1], [2], but they possess limitations, in that it is difficult to manufacture such electrodes with reproducible tip sizes and shapes. As a result, experiments involving their use cannot be well-controlled. Moreover, simultaneous recording of signals from neurons placed very close to one another is impossible, because of limitations in the resolution of the electrode site. Advances in semiconductor-based microelectronics technology have greatly changed this situation. Using semiconductor micromachining technology, it is possible to precisely define the location, the area and, thus, the impedance level of the electrode site. The fabrication of one- or two-dimensional array of electrodes having minutely controlled inter-electrode spacings and sizes on a semiconductor substrate is also possible. This enables the simultaneous monitoring of electrical activities from several adjacent neurons and their networks in both the central nervous system and the peripheral nervous system [3]–[7].

The semiconductor microelectrode provides another significant and important benefit. Because the electrode sites are made on a semiconductor substrate, on which active circuits can be integrated, it is possible to perform basic signal conditioning such as amplification and multiplexing by monolithically integrated active circuitry. The advantages of this capability are obvious. First, this eliminates much of the off-chip signal processing circuits so that the overall instrumentation channel is greatly simplified, possibly to a single chip. Second, because the distance between the front-end amplifier and the electrode site is very short, the output signal is less contaminated by the electromagnetic interference (EMI) which results from the connecting wires. Third, the number of connecting leads can be reduced by an on-chip multiplexer [5]. In addition, further signal conditioning circuitry can also be integrated. Even wireless transmission is feasible when the radio-frequency (RF) telemetry circuits are built on-chip [8].

The most important feature of the monolithic integration of active circuits is on-chip amplification. Extracellular action potentials show an amplitude in the range of tens to hundreds of millivolts and, thus, amplification is necessary for further signal processing such as multiplexing, filtering, and analog-to-digital conversion. However, this small amplitude of the neural signal and the high impedance level of the electrode site make the design of the front-end amplifier somewhat difficult. In addition, semiconductor devices inevitably add unwanted noise to the signal. Extracellular neural signals have frequency components in the range of about 50–5 kHz, where MOSFET’s show an undesirably high level of noise spectral density (NSD) [11]–[13]. This poor noise characteristic is mainly due to 1/f noise. Although no precise explanation for the mechanism of 1/f noise is available, it is widely accepted and has been experimentally verified that the NSD of 1/f noise is inversely proportional to the gate area of MOSFET, as well as the characteristic capacitance of the gate oxide [11]–[13]. Thus, a large gate area is necessary, in order to reduce 1/f noise.

However, the silicon area is a valuable resource that must be saved in the procedure of integrated circuit design, especially when simultaneous multichannel recording from up to hundreds of channels is being considered. It is desirable to minimize the layout area of devices while still satisfying the necessary requirements for the noise performance. In addition, although the
II. DESIGN OF THE LOW-NOISE PREAMPLIFIER FOR EXTRACELLULAR NEURAL SIGNAL ACQUISITION

A. Considerations of the Noise Characteristics in Front-End Amplifier Design

In the design of communication or instrumentation systems, the noise characteristic of the front-end amplifier is an important design issue, since it determines the required minimum level of the input signal amplitude. It is generally considered that any unwanted signal which contaminates the desired signal is noise. Major noise sources can be grouped into two categories. One is EMI from external sources such as 60-Hz power line, coupled through various parasitic capacitances. Proper shielding, grounding and isolation are necessary, in order to reduce the coupling of EMI. In the case of the design of a semiconductor neural probe, careful layout is required.

Another factor is intrinsic noise sources which result from the random fluctuations of physical processes within active and passive electronic devices. Our main concern in this paper is to minimize the influence of these intrinsic noise sources by appropriate circuit design. If the system is constructed using discrete components on a printed circuit board (PCB), devices which have low-noise characteristics can be chosen to meet the requirement, but in case of a monolithically integrated circuit, the available device is limited by the given fabrication technology. Most of today’s integrated circuit (IC) and microelectromechanical systems (MEMS) are fabricated using CMOS technology whose constituent devices are MOSFET’s which show poor noise characteristics over a bipolar transistor or JFET, especially for the low-frequency range where 1/f noise is dominant. Active semiconductor neural probe is no exception to this [3]–[5], [28]. Hence, it is evident that poor noise performance of MOSFET poses a potentially serious problem.

Considerable effort has been expended in the implementation of the low-noise CMOS amplifier at the low-frequency range. Methods such as correlated double sampling and auto-zeroing [24] have been proposed and utilized for the offset reduction and low-frequency noise elimination. These are attractive solutions when the amplifiers are part of the sampled-data system, but it is impossible to use them for true continuous-time applications, such as neural signal recording.

The output noise power is determined from equivalent input noise sources and transfer function of the system by calculating the following integral,

$$
\sigma_{\text{out}}^2 = \int_0^\infty |H(f)|^2 N_{\text{in}}(f) \, df
$$

(1)

where $H(f)$ is transfer function of the system and $N_{\text{in}}(f)$ is power spectral density of the total equivalent input noise. In our problem, $N_{\text{in}}(f)$ is determined from the equivalent impedance of the electrode and the topology and constituent devices of the preamplifier. Noise sources which contribute to $N_{\text{in}}(f)$ are: thermal noise resulting from the equivalent impedance of the electrode, noise from the active semiconductor devices which constitute the amplifier. In the case...
of semiconductor devices, three fundamental mechanisms of thermal, flicker (1/f) and shot noise contribute to the total noise power [11]–[13]. Thermal noise is the result of the thermal excitation of carriers in a conductor and shows a white spectral density whose level is proportional to the absolute temperature. Shot noise also shows a white spectral density which is dependent on the dc bias current. In the case of a MOSFET, where dc bias current at the gate is almost negligible, shot noise is ignored [13]. Flicker noise, or 1/f noise, is the least well understood of the three mechanisms, but it is generally accepted that its noise spectral density $N_{1/f}(f)$ can be expressed by (2) in the case of a MOSFET [9], [13], [18]

$$N_{1/f}(f) = \frac{K}{C_{ox}WL} \frac{1}{f^\alpha}$$  \hspace{1cm} (2)

where

- $L$ is the gate length;
- $W$ is the gate width;
- $C_{ox}$ is the gate-oxide capacitance per unit area;
- $K$ are parameters which are dependent on specific fabrication and a cation process.

Because its spectral density is inversely proportional to the frequency, the contribution is significant in the low-frequency range. Equation (2) clearly demonstrates that it is possible to reduce noise level by increasing the area of the devices, but only at the expense of the increased circuit areas and input capacitance. These result in a serious problem in our application because the input of the system has a large impedance level (i.e. the equivalent impedance of electrode–electrolyte interface [15]), and because the active neural probe usually contains multiple recording sites so that one preamplifier is required for each recording site. In this regard, it is very important to minimize the areas of devices at the input stage while meeting the specification of the noise characteristics required for the following system.


Fig. 2 shows a typical two-stage CMOS differential amplifier. The large gain and phase margin which is required for op-amp application can be satisfied using this structure [9], [16], [18]. For our extracellular neural signal preamplifier application, the required voltage gain is several hundreds and the cutoff frequency is several kilohertz, respectively. These requirements are easy to achieve with this structure. Several active neural probes having a preamplifier which is similar to this structure have been reported [3], [5], [28]. As previously mentioned, the output noise power can be calculated from the input equivalent noise sources and the transfer function. From circuit analysis and by assuming that the devices $M_1$ and $M_2$, $M_3$, and $M_4$, respectively, are identical, it can be shown that the equivalent noise spectral density referred to the input of the preamplifier, $N_{in}(f)$, can be expressed as follows [17]:

$$N_{in}(f) = N_{electrode} + N_{MOS, thermal}$$

$$N_{electrode} = \left( \frac{4kT}{g_{m1}} \right) \left( \frac{R_{seal} + R_e}{1 + \left( \frac{g_{m3}}{g_{m1}} \right)^2} \right)$$

$$N_{MOS, thermal} = \left( \frac{4kT}{3g_{m1}} \right) \left( 1 + \left( \frac{g_{m3}}{g_{m1}} \right)^2 \right) \frac{1}{R_t}$$

$$N_{MOS, 1/f}(f) = 2 \frac{K}{C_{ox}} \left( \frac{K_p}{L_1W_1} + \frac{K_n}{L_3W_3} \left( \frac{g_{m3}}{g_{m1}} \right)^2 \right) \frac{1}{f^\alpha}$$  \hspace{1cm} (6)

Here, $g_{m1}$ is the transconductance of the device $M_1$, $R_{seal}$ and $R_e$ are the constituent elements of the equivalent impedance of the electrode–electrolyte interface (Fig. 3) whose physical meanings can be found in [2] and [15]. $K_p$ and $K_n$ are $K$ parameters defined for PMOS and NMOS, respectively. $K_p$, $K_n$, and $a$ can be determined from the measured noise spectra of the devices, $q$ is in the range of 0.7–1.2. $N_{electrode}$, $N_{MOS, thermal}$ and
$N_{\text{MOS},1/f}(f)$ denote the input referred noise spectral densities due to the thermal noise of the electrod–electrolyte equivalent impedance, thermal noise sources of the MOS devices, and $1/f$ noise sources of the MOS devices, respectively. When deriving (3)–(6), we considered only the contributions from the electrode and the active devices at the first stage, since these determine the overall noise characteristics of the amplifier.

The transfer function of the amplifier shown in Fig. 2 can be expressed by the following second-order system function [9]:

$$H(f) = \frac{A(1 + j2\pi f/z)}{(1 + j2\pi f/p_1)(1 + j2\pi f/p_2)}.$$  \hspace{1cm} (7)

Here, $z$, $p_1$ and $p_2$ are the angular frequencies of the zero, the dominant pole and the second pole, respectively, and $A$ is the dc gain. These values can be obtained from small-signal analysis of the equivalent circuit of the amplifier. The dc gain has no influence on SNR but the location of the poles and zeros alters the transfer function, and as a result, the output signal and noise power. Regardless of this fact, in previous studies the preamplifiers for the neural signal have been designed so that they have a cutoff frequency of just “several” kilohertz (6–10 kHz) [3], [5], [6]. Here, we attempted to determine the major factors which affect the output noise power, and to optimize these so that the required SNR specification is satisfied.

From (3)–(7), the total output noise power $\sigma_{\text{out}}^2$ can be expressed as follows:

$$\sigma_{\text{out}}^2 = \sigma_{\text{electrode}}^2 + \sigma_{\text{MOS,thermal}}^2 + \sigma_{\text{MOS},1/f}^2.$$  \hspace{1cm} (8)

$$\sigma_{\text{electrode}}^2 = 4kT(R_{\text{seal}} + R_e) A^2 \int_0^\infty \frac{(1 + j2\pi f/z)^2}{(1 + j2\pi f/p_1)(1 + j2\pi f/p_2)} df.$$  \hspace{1cm} (9)

$$\sigma_{\text{MOS,thermal}}^2 = \frac{16kT}{3g_{m1}} \left(1 + \left(\frac{g_{m3}}{g_{m1}}\right)^2\right) A^2 \int_0^\infty \frac{(1 + j2\pi f/z)^2}{(1 + j2\pi f/p_1)(1 + j2\pi f/p_2)} df.$$  \hspace{1cm} (10)

$$\sigma_{\text{MOS},1/f}^2 = \frac{2A^2}{C_{ox}} \left(\frac{K_P}{L_1W_1} + \frac{K_n}{L_3W_3} \left(\frac{g_{m3}}{g_{m1}}\right)^2\right) \int_0^\infty \frac{(1 + j2\pi f/z)^2}{(1 + j2\pi f/p_1)(1 + j2\pi f/p_2)} \frac{1}{f_0} df.$$  \hspace{1cm} (11)

As is done in this paper, in many cases it is possible to obtain a satisfactory fit to the measured noise spectrum with only $K_P$ and $K_n$ by setting a equal to one [11], [18]. To evaluate the integrals in the above equations, we obtained the following explicit expressions after a rather lengthy calculus by setting $a$ equal to one and by integrating from $f_L$ to $f_H$

$$\sigma_{\text{out}}^2 = 4kT(R_{\text{seal}} + R_e) A^2 I_1 + \frac{16kT}{3g_{m1}} \left(1 + \left(\frac{g_{m3}}{g_{m1}}\right)^2\right) A^2 I_1 + \frac{2}{C_{ox}} \left(\frac{K_P}{L_1W_1} + \frac{K_n}{L_3W_3} \left(\frac{g_{m3}}{g_{m1}}\right)^2\right) A^2 I_2$$  \hspace{1cm} (12)

where $I_1$ and $I_2$ are scaling factors of the output thermal noise power and the output $1/f$ noise power, respectively, and which are dependent on the spectral characteristics of the transfer function and noise sources. These can be evaluated analytically as given in (13) and (14), shown at the bottom of the next page.

In order to estimate the output noise more accurately, the numerically calculated value of the frequency response using a circuit simulator like HSPICE might be used [10]. However, the above explicit expressions for $I_1$ and $I_2$ give directly the dependency of output noise power on the locations of poles and zero. Moreover, as shown in Fig. 4, we found that the second-order system function (7) is able to approximate the actual frequency response obtained from HSPICE simulation precisely, and as a result, no substantial errors were found when the analytical expressions were used instead for the calculation of the output noise power.

C. Results

In order to determine the major factors which influence noise performance and to estimate their numerical values, we performed a set of calculations of the output noise power based on (12)–(14) while varying a number of parameters which can be changed by the circuit designer.
We determined the locations of poles and zero and the geometry of the devices by means of a quantitative analysis of the output noise. 1/f noise parameters of the MOSFET used for the preamplifier design were obtained from curve fitting to the measured noise spectrum from the literature which contains detailed measurements of 1/f noise parameters for several CMOS IC fabricators [14]. \( K_p \) and \( K_n \) values used to obtain the results in this section were \( 3.01 \times 10^{-24} \) and \( 1.78 \times 10^{-22} \), respectively.

Careful considerations of (12)–(14) provides insights into the factors determining the noise power at the preamplifier output. These are as follows.

- The frequency of the dominant pole, the second pole and the zero. However, experiences of analytical calculation and HSPICE simulation revealed that the locations of the second pole and zero do not have significant effects on the overall noise characteristics in most cases.
- The transconductance ratio of the devices at the input stage, \( g_{m3}/g_{m1} \).
- The gate areas of the devices at the input stage, \( L_1W_1 \) and \( L_3W_3 \).
- \( K_p, K_n, \) and \( C_{ox} \). However, these parameters are not controllable by the circuit designer, since they are determined by the given fabrication process.
- The equivalent impedance of electrod–electrolyte interface. But in our case, we found that the total output noise is mainly determined by the 1/f noise as shown in Fig. 5(a), so that a rough estimation of \( R_{\text{leak}} \) and \( R_{\text{seal}} \) is sufficient for the analysis.

Therefore, we can summarize the factors that have great effects on the noise performance of the preamplifier, and can be adjusted by circuit designer as follows. These factors include the frequency of dominant pole, the transconductance ratio of the devices at the input stage, \( g_{m3}/g_{m1} \), and the gate area of the devices at the input stage, \( L_1W_1 \) and \( L_3W_3 \). With this preliminary information, we performed an analysis, in order to estimate quantitative value of SNR.

Fig. 5 shows the effect of the dominant pole frequency variation on the output noise power. The 1/f noise power of the preamplifier is much larger than the thermal noise power of the preamplifier and the thermal noise power of the electrod–electrolyte interface, unless the areas of input devices are tremendously large. The 1/f and thermal noise powers both increase with increasing dominant pole frequency, but the rate of increase becomes slower. While the dominant pole frequency should be high enough to pass all the frequency components of the input

\[
I_1 = \int_0^\infty |H(f)|^2 df = \int_0^\infty \left| \frac{(1 + j2\pi f/z)}{(1 + j2\pi f/p_1)(1 + j2\pi f/p_2)} \right|^2 df \\
= \left[ \frac{1}{2\pi(p_1^2 - p_2^2)x^2} \arctan(2\pi f/p_1) + \frac{1}{2\pi(p_1^2 - p_2^2)x^2} \arctan(2\pi f/p_2) \right]_{f_L} \tag{13}
\]

\[
I_2 = \int_0^\infty |H(f)|^2 \frac{1}{f\alpha} df = \int_0^\infty \left| \frac{(1 + j2\pi f/z)}{(1 + j2\pi f/p_1)(1 + j2\pi f/p_2)} \right|^2 \frac{1}{f\alpha} df \\
= \left[ \log f + \frac{(p_1^2 + p_2^2)z^2}{2(p_1^2 - p_2^2)x^2} \log(4\pi^2 f^2 + p_1^2) + \frac{(2p_1^2 - p_2^2)z^2}{2(p_1^2 - p_2^2)x^2} \log(4\pi^2 f^2 + p_2^2) \right]_{f_L} \tag{14}
\]
signal, it cannot be increased unrestrictedly high manner because of the inherent limitation of the MOSFET [13]. Furthermore higher dominant pole frequency increases the output noise power as shown in Fig. 5, thus, this might lead to a deterioration of SNR. Therefore, it is desirable to determine the minimum value of the dominant pole frequency while still meeting the required SNR criteria. This necessitates, not only the calculation of the output noise power, but also the signal power.

Fig. 6(a) shows a typical extracellular action potential whose amplitude is approximately 60 μV. This was recorded from a metal microelectrode inserted in the abdominal ganglia of Aplysia. Details of the recording experiment are given elsewhere [19]. The power spectrum of the signal shown in Fig. 6(b) is calculated using Welch’s method [20]. The output signal power is calculated from this signal power spectrum and the transfer function of the preamplifier. Fig. 7 shows the effect of the dominant pole frequency on the output signal power. As the dominant pole frequency increases, the output signal power increases up to 3 kHz, but the signal power saturates at around 5 kHz. A further increase in the dominant pole frequency is not beneficial in increasing the output signal power and the SNR.

The variation of SNR with respect to each parameter can now be quantitatively obtained from the calculation of signal power and noise power at the preamplifier output. Fig. 8 shows SNR variation as a function of the dominant pole frequency. SNR increases as the dominant pole frequency increases, but no further enhancement of SNR was obtained when it is higher than about 5 kHz. Each trace in Fig. 8 is for various $g_{m3}/g_{m1}$ ratios. A larger $g_{m3}/g_{m1}$ ratio decreases SNR. In order to investigate this more carefully, we observed SNR variation with respect to the $g_{m3}/g_{m1}$ ratio while fixing the dominant pole frequency. Fig. 9 depicts the variation of SNR with respect to the transconductance ratio at the input stage when the dominant pole frequency is 5 kHz. SNR rapidly decreases as the $g_{m3}/g_{m1}$ ratio increases to 0.4-0.5.

Now we investigate the effect of device size on SNR. Fig. 10 shows the result of slightly misjudged input device sizes on
the noise performance. When the area of $M1$ and $M3$ are 4000 $\mu m^2$ and 900 $\mu m^2$, respectively [Fig. 10(a)], and as the $g_{m3}/g_{m1}$ ratio is 0.25, the output noise power is significantly larger than the output signal power although the areas of input devices appear to be sufficiently large. Fig. 11 shows the SNR variation with respect to the two input device areas. The $g_{m3}/g_{m1}$ ratio is 0.1 and the dominant pole frequency is 5 kHz. As we can easily estimate qualitatively from (2), an increase in the device areas led to an improvement in SNR, but Fig. 11 gives the numerical value of SNR with respect to the device areas. The shaded region in contour plot Fig. 11(b) indicates the domain where the values of SNR exceeds five. The SNR value of five is an arbitrarily chosen lower limit of SNR, but in recent studies of neural spike sorting, correct classification of up to 80%–90% was possible provided the SNR is higher than five [25], [26]. Fig. 11(a) and (b) provide the specific numerical values of device sizes necessary to achieve the required SNR once the dominant pole frequency and transconductance ratio is determined.

III. DISCUSSION AND CONCLUSION

In this paper, we presented an organized design scheme for a preamplifier, the most important building block of the signal-conditioning circuitry of an active neural probe with on-chip electronics monolithically integrated with the recording electrode sites. The major issue is to provide the quantitative information necessary to design the noise performance, using conventional CMOS fabrication processes. We demonstrated that the output noise power can be much higher than the output signal power, provided the devices at the input stage of the preamplifier were made as small as present fabrication technology permits. Thus, if the typical CMOS process is used for the fabrication of the active neural probe, the design method proposed in this paper is indispensable. Otherwise special low-noise process will be necessary to facilitate noise performance design. Although there is general knowledge available for the low-noise design method (for example, it is well known that large device area reduces $1/f$ noise), in the real design of integrated circuits, quantitative estimation of parameters meeting the given specification can be a great help to the circuit designer.

One thing to note is that the $1/f$ noise parameters are directly dependent on the Si–SiO$_2$ interface and, thus, it might vary slightly even among the wafers from the same process run. We suggest that one should use the conservative value of $1/f$ noise parameter, i.e. largest $K_p$ and $K_n$ values obtained by measuring the noise spectrum of the devices from the given fabrication process run.

The methods in this paper can be applied to any CMOS-integrated sensor device used for acquisition of low-frequency, continuous-time signal. Possible examples are active EEG [27], [28] or an ECG recording device [29] and a cultured neuronal network device having on-chip CMOS circuitry. The methods can also be used for other preamplifier types, such as simple
source-follower buffer [21], after identifying the equivalent input noise sources.

REFERENCES


Kyung Hwan Kim was born in Seoul, Korea, in 1973. He received the B.S. degree from Department of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Taejon, Korea, in 1995. In 1997, he received the M.S. degree from School of Electrical Engineering, Seoul National University, Seoul, Korea, where he is currently working toward the Ph.D. degree.

His research interests include biomedical signal processing and instrumentation with emphasis on their application to neural signals, for neuroscience research and neural prostheses.
Sung June Kim (S’79–M’84) received B.S. degree in electronics engineering from Seoul National University, Seoul, Korea, in 1978. He received M.S. and Ph. D. degrees in electrical engineering from Cornell University, Ithaca, NY, in 1981, and 1983, respectively. His Ph.D. thesis was on the photoinduced fabrication of transsubstrate microelectrode arrays based on silicon substrate and their neurophysiological applications.

From 1983 to 1989, he worked as an MTS in Bell Laboratories, Allentown, PA and Murray Hill, NJ. At Bell Labs, he studied design and processing of silicon VLSI, and the process and device development of optoelectronic integrated circuits (OEIC’s) for long-wavelength optical communications. In 1989, he returned to Korea to join the Department of Electronics Engineering and the Inter-University Semiconductor Research Center (ISRC) of Seoul National University, where he is now a full Professor in the School of Electrical Engineering. His research interests are in areas of bioelectronics, bioinstrumentation, neural prosthesis, and optoelectronic semiconductor devices. He has published some 60 papers, and has managed many research projects in the areas mentioned.