CMOS micromachined probes by die-level fabrication for extracellular neural recording

This content has been downloaded from IOPscience. Please scroll down to see the full text.
2007 J. Micromech. Microeng. 17 283
(http://iopscience.iop.org/0960-1317/17/2/014)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 140.114.234.112
This content was downloaded on 26/02/2014 at 02:35

Please note that terms and conditions apply.
CMOS micromachined probes by die-level fabrication for extracellular neural recording

Meng-Han Ho¹, Hsin Chen¹,², Fouriers Tseng³, Shih-Rung Yeh³ and Michael S-C Lu¹,²,⁴

¹ Institute of Electronics Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan, Republic of China
² Department of Electrical Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan, Republic of China
³ Department of Life Science, National Tsing Hua University, Hsinchu 30013, Taiwan, Republic of China
⁴ Institute of Microelectromechanical Systems, National Tsing Hua University, Hsinchu 30013, Taiwan, Republic of China

E-mail: sclu@ee.nthu.edu.tw

Received 19 July 2006, in final form 5 December 2006
Published 9 January 2007
Online at stacks.iop.org/JMM/17/283

Abstract
In this paper, we present the design, fabrication and characterization of CMOS micromachined probes for extracellular neural recording. A convenient fabrication process is proposed for making integrated recording probes at the die level, providing a low-cost solution for academic research as compared to the more expensive wafer-level approach adopted in prior work. The devices are fabricated in a standard 0.35 \( \mu \)m CMOS process, followed by post-CMOS micromachining steps to form the probes. The on-chip circuit, used for recording action potential signals of neural activities, provides a stable dc bias when operating in electrolyte. The subthreshold transistor at the circuit input provides a tunable resistance value between 10 M\( \Omega \) up to G\( \Omega \). The circuit consumes a total power of 790 \( \mu \)W and has an output noise of 19.3 \( \mu \)V \( \text{Hz}^{-1/2} \) at 100 Hz. The recorded action potential from the stimulated ventral nerve cord of a crayfish is about 0.6 mV with a pulse width of about 1.2 ms.

(Some figures in this article are in colour only in the electronic version)

1. Introduction
Electrical signals represent a major physiological signal source to the understanding of nervous systems. The electrical signal of a neuron comes from the flow of ions across the cell membrane, whose overall effect is observable as the potential changes across the membrane. This characteristic facilitates the monitoring of neural activities with electronic devices, as well as interfacing electronic prostheses with impaired nervous systems. It is believed that simultaneous recording of multiple neural activities is important for understanding functions of nervous systems, as well as for developing improved neural prostheses [1]. To meet this demand, silicon-based recording chips with multiple thin-film electrodes fabricated by lithographic microelectronic technology have emerged [2–6]. Schemes for recording neural signals with these probes are categorized as intracellular and extracellular recording. Intracellular recording relies on penetration of a cell membrane, with a glass micropipette for example, to measure the potential inside a cell membrane. The advantage of intracellular recording is the measured high signal magnitudes up to tens of mV, yet the invasive recording scheme can reduce the lifetime of neurons and even alter the properties of cell membranes [7]. Low-invasive recording is therefore more favorable for neural prosthetic devices. Extracellular recording scheme minimizes the invasiveness
by measuring potential changes induced by ion flows outside cell membranes. As long as an electrode is placed close enough to neurons, voltage signals ranging from several tens of $\mu$V to a few mV can be recorded. Depending on the number of neurons attached to an electrode, the recorded signals correspond to either a series of spikes, so-called action potential, of a single neuron, or field potential induced by the activation of multiple neurons. Recording the spiking pattern of a neural tissue is important for understanding information coding and processing in a particular tissue [8], so as to restore impaired functions of nervous systems [1, 9–11]. Recording of spatial and temporal distributions of evoked potential using high-density microelectrode arrays, on the other hand, is important for inferring complex functions of nervous systems at the behavioral level, providing great potential for restoring impaired cognitive functions [12] and advanced neural computation based on hybrid systems [13, 14].

Silicon-based probes are attractive because the rapid progress of silicon technology facilitates the miniaturization of probe size, the maximization of electrode density, and the integration with signal conditioning circuits. The probes reported by Kewley [15] were fabricated by a CMOS-compatible process with plasma etch performed on wafer level, and the associated sensing circuitry was implemented on a separate chip. Yoon [16] used a combination of plasma and wet etch to form the probes, with the thickness being controlled by plasma etch using a low-temperature oxide (LTO) mask. Similarly the sensing circuitry was realized off-chip. The technique used in earlier work was based on using dopant-selective etch-stop [19, 20], but for more consistent thicknesses, silicon-on-insulator (SOI) wafers can be used with the top device layer defining the probe thickness [17, 18].

The development of neural recording probes with monolithically integrated circuits [18–22] is of great interest as the desired number of recording sites continues to grow. Routing and multiplexing of signals and reduction of channel crosstalk and noise can all take advantages of the integrated CMOS MEMS approach. The difficulty for integration, however, is the associated high investment cost in gaining access to a CMOS process, since certain micromachining steps performed either before or in the intermediate stage of CMOS fabrication are not tolerable by most foundries. The process integration between the IC and MEMS processes can be hindered when considering the issue of equipment contamination. It is known that most academic research efforts in IC design tend to adopt the multiple-project wafer (MPW) approach for reducing the shared cost per design. Designers receive dies, not a wafer, after fabrication is completed. Post-CMOS micromachining process is adopted to form released microstructures [23]. We provide in this paper a convenient micromachining process that can be performed at the die level with no lithographic steps needed. The materials that make up the neural recording probes can directly utilize those already available in a conventional CMOS process.

The rest of the paper is organized as follows. The CMOS micromachining process is introduced in section 2, followed by design of the neural recording probe in section 3. The experimental results on circuit characterization and recording of the action potential signal are presented in section 4, followed by discussions and conclusions in section 5.

2. Fabrication

The TSMC (Taiwan Semiconductor Manufacturing Company) 0.35 $\mu$m two-poly-silicon four-metal (2P4M) CMOS process is used for device fabrication. Dry-etching steps similar to those described in [24] are performed after CMOS dies are back from the foundry. As shown in figure 1, first the silicon beneath the probe region is thinned down to about 20 $\mu$m by a backside deep reactive etch using $\text{SF}_6$ and $\text{C}_4\text{F}_8$ (700 sccm: 4 sccm). The pressure and the RF power in the chamber are 5 Pa and 2500 W, respectively, and the silicon etch rate is about 20 $\mu$m min$^{-1}$. A hand-painted photoresist coating is applied on the backside as the etch-resistant mask in this step. Next a front-side reactive ion etch is performed to remove the exposed dielectric layers with the top metal as the etch-resistant mask. The etch rate of SiO$_2$ by using CHF$_3$ and O$_2$ gases (100 sccm: 5 sccm) is about 60 nm min$^{-1}$ for a chamber pressure of 13.3 Pa at a RF power of 250 W. Finally for structural release, a front-side deep reactive ion etch removes the silicon between the probe shanks with the top metal as the etch-resistant mask.

![Fabrication of CMOS micromachined neural probes](Image)

Figure 1. Fabrication of CMOS micromachined neural probes:

(a) after completion of CMOS, (b) backside deep reactive ion etch of silicon, (c) reactive ion etch of dielectric layers, (d) anisotropic silicon etch for structural release and (e) encapsulation of probes using parylene coating.
CMOS micromachined probes by die-level fabrication for extracellular neural recording

Silicon
Polymer

(a)

(b)

Figure 2. Micrographs of the CMOS die after the backside silicon etch: (a) bottom view of the die (backside) and (b) cross-sectional view.

The top metal layer is removed by about one-third of the total thickness due to the front-side plasma etch. The released probe structure contains multiple metal and dielectric layers and the silicon substrate. Figure 2 shows the top and the cross-sectional views of the CMOS die after the backside silicon etch, and the fabricated probe array after dry etch is shown in figure 3. The probe shank is 1.5 mm long, 50 µm wide, and about 20 µm thick. The separation between probes is 200 µm. The recording sites as shown are made of aluminum and have an area of 10 × 10 µm². The routing of electrodes to the sensing circuit is conveniently realized using either metal below the top metal layer. The metal-2 layer is used in our design to give an appropriate routing capacitance.

Bondwire connections from I/O pads on the CMOS die to a small printed circuit board are established after plasma etch. Then the whole chip, except the area with micromachined probes, is immersed in epoxy resin which is meant to provide electrical isolation when operating in electrolyte. This step ensures that the thin bondwires are well encapsulated and stabilized during operation in liquid. Figure 4 shows the micrograph of the packaged die on the board. Since the probe’s top metal and silicon are still exposed in air, we perform the final encapsulation of the whole board by deposition of a 60 nm parylene layer, which is known to be bio-compatible and have good resistances against solvents, acids and bases. Parylene coatings form from a gaseous monomer without experiencing an intermediate liquid stage. Component configurations with sharp edges are coated uniformly without voids.

3. Neural recording probe design

The neural probes carrying recording electrodes are designed in the shape of shafts in the aim to approach target neurons with minimized tissue trauma and disturbance in a biological system. It is desired that a recording electrode is placed closer to a neuron such that the recorded signal is better coupled to a neuron such that the recorded signal is better coupled into the sensing circuit. As shown in figure 1, the recording electrode is made of the top metal-4 layer which subsequently connects to the sensing circuit. The other overlaying metal-4 layer on top of the probe is electrically grounded as a shield.
The low-profile shafts need to be mechanically robust enough to survive the operation in liquid and contact with neurons. The main materials that make up a probe are the 0.9 μm metal and the 6 μm dielectric layers, and the remaining silicon of 20 μm. The force-displacement relationships in the axial and out-of-plane directions are obtained by finite-element simulations as shown in figure 5(a) and (b), respectively. The values of Young’s modulus used in simulation are 70 GPa, 75 GPa and 170 GPa for aluminum, silicon dioxide and silicon, respectively. For the axial-stress case, the nonlinear buckling phenomenon appears when the applied force is more than about 50 mN, as the displacement increases drastically with respect to the applied force beyond the point. The curve in figure 5(b) shows the linear relationship between the uniformly applied pressure on top of a probe and the produced displacement at the tip. The maximum displacement is about 30 μm under a pressure of 0.1 μN μm⁻².

The recording electrodes are coated with a 60 nm parylene layer for electrical insulation. The equivalent circuit of the recording interface is simplified as a capacitor and the action potential signal is capacitively coupled into a recording electrode. The thickness of the dielectric coating is preferred to be thin in order to produce a large coupling capacitance and therefore an enhanced signal-to-noise ratio (SNR). By using a relative permittivity of 3 for parylene, the calculated capacitance value for the interface of 10 × 10 μm² is 44 fF. It is then desired that the input capacitance of the sensing amplifier to be small as well for increase of the recorded signals.

The existence of high impedance at the sensing node, namely, the pre-amp input, could induce interferences from adjacent channels if signal lines are not properly shielded. The multiple metal layers provided by the commercial CMOS process can be conveniently utilized as ground shields in the design. On-chip sensing buffer can improve the SNR of the recorded signal by reducing signal attenuation and noise coupled through the output leads. The action potential signal of interest ranges from 100 Hz to 10 kHz with an amplitude between 50 μV up to mV. A low-noise signal processing circuit over the bandwidth of interest is desired for extracellular neural recording. To promote the measured SNR, it is desired to reduce the circuit noise contributed by the flicker (1/f) noise and the thermal noise of transistors. The former plays a more dominant role as it occupies the low-frequency range where the signals of interest appear. Reduction of the 1/f noise often resorts to the use of large-size transistors, which in turn produces an increased capacitance at circuit input. The use of PMOS rather than NMOS transistors is beneficial in further reducing the 1/f noise.

As shown schematically in figure 6, the sensing circuit is mainly a buffer amplifier consisting of the PMOS source follower M₁ and the biasing current mirror formed by the transistors M₂ and M₃. The p-type source follower can have a voltage gain close to 1 without the body effect. When operating a probe in a solution of unknown field potential, the dc bias at input of the buffer amplifier needs to be provided because the floating node is located in the middle of a capacitive divider formed by the electrode interface and the pre-amp input capacitance. This dc path is commonly established through the use of a large resistor, with the value in the order of GΩ to avoid signal attenuation due to the large impedance of the recording interface at the frequencies of interest. Direct use of the diffusion or the high-resistive polysilicon resistor provided by the CMOS process would result in very large silicon area for realizing a resistance of this order. Reset transistors have been used to set dc bias for micromachined capacitive sensors; however this approach is preferably used in a fully differential sensing circuit because the effect of charge injection due to transistor switching is more conveniently canceled than the case of a single-ended circuit. A feasible dc-bias scheme can use a transistor operated in subthreshold region (M₄ in figure 6) to provide the desired GΩ resistance. The drain-to-source current of a NMOS transistor operated in subthreshold region (M₄ in figure 6) is given by [26]:

\[
I_{ds} = I_{0} \left[ 1 - \exp \left( -\frac{V_{ds}}{V_{T}} \right) \right] \cdot \exp \left( \frac{V_{gs} - V_{th}}{nV_{T}} \right),
\]

where \(I_{ds}\) is the drain-to-source voltage, \(V_{gs}\) is the gate-to-source voltage, \(n\) is the subthreshold swing factor and \(V_{th}\) is the threshold voltage. The term \(V_{T}\) is the thermal voltage defined as \(k_{B}T/q\), where \(k_{B}\) is the Boltzmann’s constant, \(T\) is the absolute temperature and \(q\) is the electron charge. The term \(I_{0}\) is given by

\[
I_{0} = \mu_{0} \frac{W}{L} V_{T}^{2} \sqrt{\frac{q\varepsilon_{Si} N_{ch}}{4\phi_{F}}},
\]

where \(\mu_{0}\) is the carrier mobility, \(W\) and \(L\) are the transistor width and length, \(N_{ch}\) is the channel doping concentration and \(\varepsilon_{Si}\) is the permittivity of silicon. The term \(\phi_{F}\) is expressed by:

\[
\phi_{F} = -\frac{k_{B}T}{q} \ln \left( \frac{N_{sub}}{n_{i}} \right),
\]
Figure 6. Schematic of the sensing circuit. The unit of labeled transistor length and width is \( \mu m \).

\[
R_{ds} = \left( \frac{\partial I_{ds}}{\partial V_{ds}} \right)^{-1} = \left[ \frac{I_{th}}{V_T} \exp \left( \frac{V_{ds}}{V_T} \right) \cdot \exp \left( \frac{V_{gs} - V_{th}}{nV_T} \right) \right]^{-1}.
\]

(4)

The dc bias is provided by the NMOS transistor \( M_4 \) in figure 6. The width of the subthreshold transistor is less than its length to provide a small drain-to-source current. The current flown through \( M_4 \) is scaled down tremendously after conversions from the two preceding current mirrors [18]. The equivalent channel resistance of the transistor is adjusted through the tuning voltage \( V_{tune} \) as shown. The decrease of \( V_{tune} \) induces more current in the current mirrors and thus decreases the channel resistance of the subthreshold transistor. A simpler dc-bias scheme could use only the \( M_4 \) transistor with an independently controlled source voltage. The pitfall to be aware of is the potential damage of the polysilicon gate due to injected charges during plasma etch. To remedy this issue a reverse-biased diode connecting between the gate and the substrate is used to provide a discharging path. The sensing circuit is shared by four recording electrodes on the same probes. Activation of an individual channel is controlled by the respective switch realized by a transmission gate. A high-pass characteristic is formed by the subthreshold transistor and the total capacitance seen at the circuit input. The dc bias at circuit output can be stabilized without being affected by the field potential in solution; in addition, the flicker noise at low frequencies can be suppressed due to the high-pass characteristic.

4. Experiment

The buffer amplifier was first characterized after completion of the micromachining process. The measured power consumption is 790 \( \mu W \) by using a 3.3 V supply, and the measured output resistance is 3.96 k\( \Omega \). The frequency response in figure 7 was measured from a test circuit with a fixed on-chip capacitance of 7.5 pF in the front by using an Agilent 4295A network/spectrum analyzer. The voltage gain is 0.84 in the passband, somewhat less than 1, due to the capacitive divider in the front. The frequency response also shows a low-frequency pole associated with the dc-biasing transistor and the capacitance seen at the circuit input. The variation of this pole frequency was measured by adjusting the voltage \( V_{tune} \) for the same test circuit, and the equivalent resistance of the subthreshold transistor can be converted from the measured pole frequency. The curve in figure 8 shows that the resistance changes from 20 M\( \Omega \) to 500 M\( \Omega \) for \( V_{tune} = -3 \) to \(-1.5 \) V. The resistance can be tuned close to G\( \Omega \).
when we further increase the value of $V_{\text{tune}}$. The circuit output noise shown in figure 9 was measured using an Agilent 4295A network/spectrum analyzer with a grounded circuit input. The 1/f noise is dominant over the low-frequency range as shown, and the root spectral density is about $19.3 \, \mu V \, Hz^{-1/2}$ at 100 Hz. Table 1 summarizes the circuit performances obtained by experiment and simulation, respectively.

To make sure the high-pass characteristic of the sensing circuit can provide a stabilized dc bias and pass signals located in the band of interest, the whole chip was immersed in a saline solution and then an ac voltage was applied to the solution. Figure 10(a) shows the measured circuit output for a square-wave input of 400 mV at 400 Hz. The dc operating point remains at 0.8 V during the ac excitation. Figure 10(b) shows the measured signal when the frequency of the square wave increases to 10 kHz.

A Juvenile crayfish (*Procambarus clarkii*) of about 9 cm long was the subject for the physiological test of the neural recording probe. The ventral nerve cord of the crayfish extends from the thoracic to the abdominal part of the body. Surgery was performed to quickly remove the musculature of the crayfish such that a portion of the abdominal cord can be accessed by external stimulus. The crayfish was then immersed in a saline solution (210 mM NaCl, 15 mM CaCl$_2$, 5.4 mM KCl, 2.6 mM MgCl$_2$ and 5 mM HEPES). The experiment was performed using the setup as shown in figure 11. The packaged CMOS micromachined probes and a commercial stimulating probe were mounted on separate micromanipulators for positioning close to the crayfish under the microscope. The recorded signal was amplified by an off-chip ac amplifier (Model 1700 from A-M SYSTEMS) with a gain of 1000 and a passband between 300 Hz to 5 kHz. The stimulating pulses were provided by an isolated pulse stimulator (Model 2100 from A-M SYSTEMS).

The baseline noise waveform after off-chip amplification was measured with no contact to the neuron. The measured waveform has a maximum value of about 0.25 to 0.3 mV by referring back to the output of the on-chip circuit. Then the stimulating probe and the CMOS micromachined probe were positioned close to the third and fourth ganglions of the nerve cord. A 5 V biphasic square wave of 5 kHz was applied every 40 ms for stimulation and the corresponding action potential signal was recorded on an oscilloscope. Figure 12 depicts the recorded waveform by referring back to the output of the sensing circuit. The amplitude of the recorded action potential,
5. Discussion and conclusion

This work presents a convenient and low-cost CMOS micromachining approach for making integrated neural recording probes at the chip level, as compared to the conventional wafer-level fabrication process that often demands a high investment cost for achieving monolithic integration. The micromachining process is performed after completion of CMOS fabrication by successive steps of dry etch and conformal encapsulation using parylene. The thickness of the probe shank is controlled by the backside deep reactive etch of silicon. There is no stringent requirement regarding the polymer mask painted by hand for protection of the backside etch, and the front-side plasma etch can directly use the top metal layer for etch protection. The sensing circuits are well preserved after the backside etch since the remaining substrate is still much thicker than the doping depth of transistors. The cost for the backside etch can be greatly reduced by first grinding and polishing the substrate prior to the etch when a CMOS wafer is available. The probe thickness can thus be better controlled by using a recipe with a slower etch rate than that used in this work. The encapsulation using parylene provides the desired benefits of bio-compatibility, CMOS-compatibility and conformal coating. A thicker coating could give better protection for long-term operation, yet we should also be aware of the effect of more signal attenuation due to the reduced coupling capacitances on electrode surfaces.

It is conceivable that more degrees of freedom are allowed for integrated process development when wafer-level fabrication capabilities in both CMOS and MEMS are abundant. On the other hand, the chip-level microfabrication using a commercial CMOS process provides the benefit of yield enhancement, at least for the circuits. Besides, a commercial sub-µm CMOS process is better than a relatively low-end CMOS process as used in prior work in providing more metallization layers for routing and faster circuit speed for signal multiplexing. Both are beneficial in terms of realizing denser electrodes for recording. One concern regarding the use of a commercial CMOS process, as the technology continues to scale down, is the increase of flicker noise appearing at the frequencies of interest. It becomes inevitable to use large-size transistors for reduction of flicker noise, and the resulting large input capacitance could negatively impact the recorded signal amplitude.

The corner frequency of the high-pass characteristic at the pre-amp input has to be low enough to pass the signals of interest, which implies that the resistance value provided by the transistor operated in subthreshold should be close to the order of GΩ to realize this pole frequency. It is shown by the circuit characterization that this large resistance can be implemented successfully. Complete operation of the fabricated neural recording probe in this work has been validated by measurement of the action potential signal from the nerve cord of a crayfish.

Acknowledgments

This research effort is sponsored by the National Science Council, Taiwan, ROC, under grant number NSC 95-2221-E-007-240. The authors would like to thank the National Chip...
Implementation Center for support with CMOS fabrication and the National Center for High-Performance Computing for support with the simulation tool.

References