Design Considerations for Minimizing Noise in Micropower CMOS Integrated Circuits

by

Bruce C. Larson

B.S., Electrical Engineering
University of Houston, 1992

Submitted to the Department of Electrical Engineering and Computer Science In Partial Fulfillment of the Requirements for the Degree of

Master of Science

at the

Massachusetts Institute of Technology

June 1996

Signature of Author _____
Department of Electrical Engineering and Computer Science
May 16, 1996

Certified by ___
David J. Edell
Thesis Supervisor

Accepted by ___
rederic R. Morgenthaler
Chairman, Department Committee on Graduate Students

Design Considerations for Minimizing Noise in Micropower CMOS Integrated Circuits

by
Bruce C. Larson

Submitted to the Department of Electrical Engineering and Computer Science on May 16, 1996 in partial fulfillment of the requirements for the Degree of Master of Science in Electrical Engineering

ABSTRACT

A theoretical and experimental study of the relationship between noise and bias in the metal-oxide-semiconductor field effect transistor (MOSFET) was performed in order to facilitate the design of circuits optimized for low noise and low power. The immediate motivation for this research was the on-going development of a clinically applicable intra-cortical neural activity sensor that is free to move with the brain and that is not tethered to the skull by any wires. Because neural information sensors must be small, low mass, untethered, and able to relay many channels of signal information, they require electronics optimized for minimum power consumption and acceptably low noise levels. Since the literature concerning noise in electronic devices and circuits did not clearly indicate the relationship between noise and power consumption, a study of noise in MOS transistors was carried out in order to develop a low noise, low power circuit design strategy. The study focused on MOS field-effect transistors because they permit the construction of amplifiers with the high impedance, low leakage inputs that are necessary for recording from the high impedance microelectrodes that will be used to detect single-unit neural activity.

While the noise of the MOSFET has been described separately for the conventional high-current "strong inversion" mode, and (less extensively) in the more recently explored low-current "weak inversion" region, it was not clear which would ultimately yield better signal-to-noise performance, and at what cost in terms of power consumption. Thus a model of the white noise component in these operating regions was derived from the perspective of velocity fluctuation of charge carriers in the transistor channel. The resulting model showed that the white noise spectral intensity of the drain current was directly proportional to the number of carriers in the channel, which could be related to drain current and transconductance. In fact, the drain current spectral intensity was shown to be directly proportional to transconductance in both weak and strong inversion, with a constant of proportionality that was similar in magnitude in the two cases. The gate-referred voltage spectral intensity, obtained by dividing the drain current spectral intensity by the square of the transconductance, was thus inversely proportional to the transconductance.

Our investigation of the transistor operating characteristics revealed that the ratio of transconductance to drain current was maximum in weak inversion,

where it was constant, and decreased as the level of inversion increased. Hence, for a set amount of drain current, the minimum gate-referred white noise was obtained when the transistor was operated in weak inversion. Similarly, for an acceptable level of gate-referred noise, the bias current needed to achieve this was minimized by weak inversion operation. Weak inversion operation was achieved for a particular drain current by making the width/length ratio of the transistor sufficiently large, since the saturation drain current at which the transition from weak to strong inversion took place was proportional to the width/length ratio of the transistor (for a particular fabrication process).

Because the low frequency "1/f" component of MOSFET noise was not well understood in either weak or strong inversion, and appears to be highly dependent on the specifics of device fabrication, a computer-interfaced noise measurement system capable of measuring the minute fluctuations in drain current of the transistor at low bias levels was constructed to empirically quantify this form of noise. A theoretical derivation of 1/f noise in MOSFETs based on carrier trapping at the semiconductor-oxide interface was presented which suggested that the gate-referred voltage spectral intensity was inversely proportional to the gate area and oxide capacitance of the transistor. Noise measurements of transistors fabricated on a commercial 2 micron "low noise analog CMOS" process supported this theory for strong inversion operation, but the noise increased at low bias levels. In the context of the model presented, this could be explained by a higher "effective trap density" at lower energy levels. The noise measurements were also used to verify the predictions of the velocity-fluctuation white noise model. Finally, a design strategy for optimizing low noise, low power electronics was developed from the results of this research.

Thesis Supervisor: Dr. David J. Edell

Title: Principal Research Scientist

Acknowledgments

I wish to thank David Edell for guiding the research that went into this thesis, and for providing a stimulating and enjoyable work environment. I thank the Department of Veteran's Affairs for the funding which made this research possible. In addition, I wish to thank my wife, Lisa, for her encouragement and patience, and my parents for 27 years of complete and unwavering support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Increasingly, as portable electronics, medical implants, and other applications necessitate decreased power consumption, the circuit designer is faced with the challenge of maintaining acceptably low noise levels. As a result, a clear understanding of the relationship between noise and power consumption in electronic devices is becoming a vital tool for the modern circuit designer. Unfortunately, the literature concerning noise in electronic devices and circuits does not clearly indicate the relationship between noise and power consumption in metal-oxide-semiconductor (MOS) integrated circuits. MOS circuits have many valuable traits that make the technology appropriate for a wide variety of digital and analog electronics applications. The technology is readily available commercially, even for prototyping and small-scale production work, allows high density of integration, and permits integrated fabrication of good-quality passive components as well as transistors. The high input resistance of MOS field-effect transistors (MOSFETs) makes them particularly suited to the transduction of signals from high-impedance sources. This prompted a study of the current-voltage relationships and noise characteristics of MOSFETs, including empirical measurements of various transistors fabricated with a commercial "low noise analog N-well CMOS" process. The purpose of the study was to develop a method for optimizing the power consumption and noise performance trade-off for circuit applications in which small signals are involved and the supply of power is limited. The long-term goal of this study was the development of a direct neuroelectric link for an improved control interface for functional electrical stimulation, prosthetics, and assistive devices for the spinal cord injured.

A system that could chronically sense many independent channels of neural information could provide an abundant source of motor control information if placed in appropriate areas of the nervous system. The link could be made by a neuroelectric interface consisting of an array of microelectrodes and amplifiers that would detect individual cell activity for a large number of neurons [Edell '80; Edell '86; Edell, McNeil, Clark, and Van; Clark; Agnew and McCreery; Hambrecht and Reswick]. Since a clinical neural information sensor based on a microelectrode array implanted in the human brain would likely be wireless (fixed only to the brain) for mechanical stability, it would require its own amplifiers, multiplexers, signal transmitter, and power supply. Because power supplies of sufficiently small size and mass will be limited in their supply capabilities, the implant's electronics must be of low power design. With limited power, it was difficult to design circuits capable of detecting the small extracellular neural potentials that are anticipated.

One possible approach to low power was to design subthreshold circuits, in which MOSFETs are operated in weak inversion. While the noise of the MOSFET has been described separately for the conventional high-current "strong inversion" mode, and (less extensively) in the more recently explored low-current "weak inversion" region, it was not clear which will ultimately yield better signal-to-noise performance, and at what cost in terms of power consumption. The noise in strong inversion operation was generally modeled by a white noise component from the thermal noise of the resistive channel, and by a "1/f" component that is not well understood but often attributed to charge traps at the semiconductor-oxide interface [Van der Ziel '86; Buckingham; Das and Moore; Christensson, Lundstrom, and Svensson; Christensson and Lundstrom; Berz; Hsu; Sah]. The limited amount of literature concerning MOSFET noise characteristics for weak inversion operation modeled the drain current noise by a white noise component described as shot noise on the drain current, and a "1/f" component that may depend on bias quite differently than in strong inversion [Fellrath; Reimbold; Kornfeld; Duh and Van der Ziel; Van der Ziel '86; Sarpeshkar, Delbrück, and Mead; Schutte and Rademeyer]. Unfortunately, since the operating characteristics of the MOSFET are also considerably different in weak and strong inversion operation, it was not clear how a transistor should be biased in order to optimize the noise-power relationship for any given application [Tsividis; Sze; Streetman; Deen and Yan; Godfrey; Gray and Meyer]. Additionally, most of the literature concerning MOSFET 1/f noise has focused on the mechanisms that might be responsible for generating this type of noise, and is of greater interest to the device physicist than the integrated circuit designer because many of the models presented contain process-dependent parameters which make them difficult to apply quantitatively [Das and Moore; Berz; Buckingham; Christensson, Lundstrom, and Svensson; Christensson and Lundstrom; Duh and Van der Ziel; Hsu; Kornfeld; Reimbold; Sah; Van der Ziel '86]. From the designer's perspective, such models can provide some insight into general directions to take, but in most cases they don't permit critical optimization of analog circuitry. Since the designer generally cannot change the process with which circuits are fabricated, and often does not even have extensive information about it, the device-physics-based models were of little practical use. What would be of greater value in the design of low-power, low-noise electronics would be a clear understanding of the performance trade-offs that can be made with the parameters that the designer can manipulate, and a means to empirically quantify them for the devices available in a given process. In MOS integrated circuits, the choices that are usually available to the designer are physical, such as transistor type (n-channel or p-channel) and geometry (size, width/length ratio), and operational (bias levels).

To determine the effects of these parameters, a noise measurement system was constructed to bias transistors to desired operating conditions and measure the spectral intensity of the drain current. By also measuring the transconductance, the noise was referred to the gate as an equivalent input voltage noise. Input-referred noise determines the smallest signals that can be

detected. For applications where high-speed electronics are not required (such as the neural information sensor which prompted this study), subthreshold transistor operation may be advantageous, so it was investigated. Since there was not extensive literature regarding noise in subthreshold operation, the noise measurement system was designed to bias transistors in the weak inversion region (as well as in strong inversion) while measuring the small noise signals. By characterizing transistors of various sizes and geometries under multiple bias conditions, transistor and circuit design guidelines were developed in terms of the parameters that a circuit designer can specify.

Chapter 2

Theory of MOSFET Operation

A brief review of the characteristics of the MOSFET is presented to aid in the development of low-noise, low-power circuit design guidelines by providing the terminology and models of operation through which a discussion of noise may be presented.

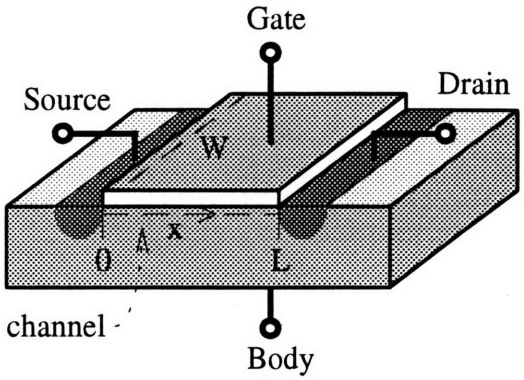

Figure 1: The Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET)

In its most basic form, which is sufficient for most aspects of our discussion, the MOSFET consists of an MOS capacitor, with plates formed by a conductive layer (the "Metal") and the doped semiconductor substrate (the "Semiconductor") separated by a thin dielectric layer (the "Oxide"). The MOS capacitor is bounded in the substrate at two ends by regions of complementary doping to form the MOSFET. This structure is depicted in Figure 1. The "metal" layer is referred to as the *gate* electrode, and the two doped regions at the ends are called the *source* and *drain* (since they are physically indistinguishable, the applied electrical bias will distinguish the two according to convention). The semiconductor under the gate is referred to as the *substrate* or *body*. All four elements have electrical contacts of the same name.

The source and drain regions form back-to-back diodes in the substrate, so current cannot normally flow between them. The MOSFET operates by the creation of a layer of concentrated minority charge carriers in the region of the substrate under the gate (called the *channel*), which allows current to flow between the source and drain. The conductive layer of charge is capacitively induced by biasing the gate electrode relative to the substrate such that it attracts minority charge carriers to the channel surface. As the gate potential first begins to change, majority charges in the substrate are driven away from the channel surface, leaving ionized dopant atoms behind. Thus a depletion region of fixed charge forms below the gate, and extends downward as the gate

potential is changed further. In equilibrium (when the source and drain are left at the substrate potential or unconnected, so that there is no flow of current in steady state), the concentration of minority carriers is an exponential function of potential in the substrate. Thus at low gate bias levels, very few mobile charges are present at the channel surface. But as the depletion layer grows, the gate and depletion layer appear as two capacitors in series: one from the gate electrode across the gate insulator to the channel surface, and the other from the channel surface across the insulating depletion layer to the substrate. So the channel surface potential begins to change in proportion to the gate potential changes, and the minority carrier concentration at the channel surface (where the potential difference relative to the substrate is greatest) begins to increase exponentially. The charge on the gate electrode "plate" of the MOS capacitor is therefore balanced in part by fixed charges in the depletion layer, which we will refer to as *depletion charge*, and in part by a thin layer of mobile minority carriers, called *inversion charge*. If the minority carrier concentration at the channel surface is less than the surface concentration of majority carriers, the MOS capacitor is said to be in *depletion*. When the concentration of minority carriers at the surface exceeds that of majority carriers, we say that the MOS capacitor is in *inversion*. The onset of inversion, when the total inversion charge is small compared to the depletion charge, is called *weak inversion*. Weak inversion is characterized by a linear relationship between gate potential and surface potential due to the capacitive divider formed by the gate oxide and the now well-developed depletion region. Since the inversion charge concentration has an exponential dependence on the surface potential, it also has an exponential dependence on the gate potential in weak inversion. As the potential difference between the gate and substrate is increased further, the total inversion charge increases rapidly, until it eventually exceeds the depletion charge. As this happens, the surface potential ceases to increase linearly with gate potential, because most of the charge added below the gate oxide is now provided by minority carriers at the channel surface rather than by the increasing depth of the depletion layer. Under this condition, the capacitive divider paradigm no longer holds. Since the inversion charge is concentrated very close to the channel surface, the addition of inversion charge causes very little change in the surface potential. Thus for further increases in potential difference between the gate electrode and the substrate, the surface potential remains almost constant. This condition is called *strong inversion*, and from the above discussion, is characterized by a linear relationship between changes in gate potential and inversion charge concentration, set by the gate oxide capacitance. The transition range of operation between weak and strong inversion is called *moderate inversion*.

Changing the potential of the source and/or drain regions relative to the substrate to make them more attractive to minority carriers in the channel will cause the inversion charge to move out of the channel and into the source and drain regions. If the source and drain are at different potentials, a current will flow between them. Because the gate potential relative to the substrate

determines the amount of inversion charge in the channel, it also determines the conductivity of the connection between the source and drain regions. Hence a flow of current between the source and drain terminals may be modulated with the gate potential. This property allows the MOSFET to provide signal amplification.

Considering first the case that the source and drain potentials are changed equally relative to the substrate potential, the flow of inversion charge out of the channel may be prevented by changing the surface potential in the channel by an equal amount (via the gate electrode) such that the source and drain are no longer *more* attractive to minority charges than the channel surface. Thus the source/drain potential has the ability to set the surface potential limits for weak, moderate, and strong inversion by drawing inversion charge out of the channel. Since only the mobile inversion charge can be moved out of the channel by the source/drain potential, the influence that the source/drain potential can have on the channel surface potential (when the gate potential is held constant) depends on how much inversion charge there is relative to depletion charge. In weak inversion, the inversion charge comprises such a small fraction of the total stored charge that it has very little effect on the surface potential. Thus while the inversion charge concentration may be dramatically decreased by the change in source/drain potential, the surface potential will not be significantly affected -- it will still be set primarily by the gate potential through the capacitive divider discussed earlier. In strong inversion, the inversion charge constitutes a significant portion of the total charge stored under the gate, and removing it will change the surface potential dramatically. In fact, if strong inversion is maintained (not enough inversion charge is removed to put the device into weak inversion or depletion), the surface potential will change by the same amount that the source/drain potential changes.

If a potential difference is created between the source and drain regions, then there will be a net flow of current between them, as they compete to draw inversion charges out of the channel. The nature of the current flow, and its dependence on source, drain, and gate potentials, will depend on the state of inversion of the channel at the source and drain boundaries. By convention, the mode of operation of the MOSFET is described by the state of the most heavily inverted end of the channel. This end of the channel is traditionally defined to be the source since it is the "source" of mobile carriers flowing into the channel, and the other end is called the drain because it is the place where the mobile carriers "drain" out of the channel. If the source end of the channel (which we shall assume to be the most heavily inverted end for the remainder of our discussion) is in weak inversion, then the surface potential there will be essentially the same as at the drain end of the channel, so there will not be any significant drift current along the channel. However, the minority carrier concentrations at the two ends may differ greatly, giving rise to diffusion current along the channel, which characterizes weak inversion, or *subthreshold* operation of the MOSFET. On the other hand, if the source end of the channel is in strong inversion, then the

surface potential there may be quite different from the surface potential at the drain end, and the electric field along the channel will have a longitudinal component causing drift current to flow. Drift dominates the channel current in strong inversion, and diffusion current can be neglected to good approximation. The level of inversion at the drain end of the channel determines how much influence the drain potential has on the channel currents, in both weak and strong inversion operation. For weak inversion operation, if the minority carrier concentration at the drain end of the channel is negligible compared to the concentration at the source end, then increasing the drain-source potential difference will have negligible effect on the channel current (since diffusion current is set by the difference in concentrations at the two ends). This condition is referred to as *saturation*, since the potential at the drain no longer affects the current that flows through it. As will be seen, the ratio of the concentrations is an exponential function of drain-source potential, so saturation is achieved for small potential differences in weak inversion. When the transistor is in strong inversion, the channel current will be proportional to the drain voltage as long as the drain end of the channel remains strongly inverted as well. This is the case because there is a continuous layer of inversion charge connecting the source and drain regions which behaves like a resistor. The range of drain voltages for which this holds, called the *linear region*, extends from the source voltage to the potential at which the drain end of the channel is depleted of minority carriers (the *pinch-off* condition), and the drain potential therefore has diminishing affect on the channel surface potential. At that point, the current through the drain is almost independent of the drain potential, and the transistor is in saturation.

2.1 General Model of MOSFET Operation

A description of MOSFET operation has been presented in a particularly clear and intuitive manner by Tsividis, which will be followed here [Tsividis 1-164]. Since the fundamental principles of operation of the MOSFET are the same whether current in the channel is carried by electrons or holes, only the equations for the *n-channel* case (NMOS, where the substrate is doped p-type with n-type source and drain regions, so that the induced channel is n-type), will be presented here. The extension to the *p-channel* case is made simply by reversing the polarities of all potentials and currents, and considering hole concentrations in the channel rather than electron concentrations. Some basic assumptions are made in Tsividis' development, but the resulting theory shows good agreement with empirical data for most cases. It is assumed that the substrate doping is uniform and light, and that the source and drain regions are heavily doped. The assumption of a lightly doped substrate allows us to treat the depletion layer as fixed in strong inversion, and minimizes the depletion layer capacitance in weak inversion. The assumption of heavily doped source and drain regions permits us to neglect any component of current flow due to channel majority carriers. In addition, it is assumed that the channel is long and wide, so that edge effects may be neglected, and that the horizontal component of the

electric field under the gate is always much smaller than its vertical component (the *gradual channel approximation*) [Tsividis 103-105].

Tsividis' general model is based on the assumption that the induced channel charge is concentrated at the channel surface (under the gate oxide) in an infinitesimally thin layer. It achieves its generality by accounting for the channel current due to both drift *and* diffusion, so that it is able to describe drain current continuously throughout weak, moderate, and strong inversion operation. By considering the inversion charge concentration per unit area Q_I' and channel surface potential ψ_s (relative to the body, or bulk of the semiconductor) as functions of position x from source to drain, equations for the drift and diffusion components of drain current may be obtained [Tsividis 108-110]:

$$I_{D,drift} = \mu W(-Q_I')\frac{d\psi_s}{dx} = \frac{W}{L}\int_{\psi_{s0}}^{\psi_{sL}} \mu(-Q_I')d\psi_s$$

$$I_{D,diffusion} = \mu W \phi_t \frac{dQ_I'}{dx} = \frac{W}{L}\phi_t \int_{Q_{I,source}'}^{Q_{I,drain}'} \mu dQ_I'$$

where

$$Q_I' = -C_{ox}'\left(V_{GB} - V_{FB} - \psi_s - \gamma\sqrt{\psi_s}\right)$$

In these equations, μ is the *surface mobility* of the mobile electrons in the channel, W and L are the channel width and length, and ϕ_t is the thermal voltage (equal to kT/q, where k is the Boltzmann constant, T is the absolute temperature, and q is the magnitude of the electron charge). The channel surface potential at the source ($x = 0$) and drain ($x = L$) ends of the channel are ψ_{s0} and ψ_{sL}, and the inversion charge concentrations there are $Q_{I,source}'$ and $Q_{I,drain}'$. The *flat-band voltage*, V_{FB}, accounts for the constant offset in surface potential due to the difference in the semiconductor and gate metal work functions (their *contact potential*) and fixed charges in the oxide [Tsividis 39-41]. Gate oxide capacitance per unit area is denoted by C_{ox}', and V_{GB} denotes the gate-to-body potential. Finally, the *body effect* or *substrate effect*, which describes how charges that are induced below the gate are divided between depletion charge and inversion charge, is taken into account by the *body effect coefficient*, γ, which depends on the substrate doping concentration N_A and oxide capacitance as follows:

$$\gamma \equiv \frac{\sqrt{2q\varepsilon_s N_A}}{C_{ox}'}$$

where ε_s is the permittivity of the semiconductor [Tsividis 54]. As a result of the body effect, if the source and drain potentials are raised by a given amount such that inversion charge is moved out of the channel, the gate potential will have to be increased by a *greater* amount in order to return the channel to the original level of inversion. This is because the surface potential must be increased by the same amount as the source and drain regions, but the gate potential has only an indirect influence. As the surface potential is increased, the depletion

layer expands, requiring a portion of the total induced charge. But the additional charge required on the gate electrode to balance the added depletion charge means that the gate potential will have to be higher. Thus transistors with more highly doped substrates and thicker oxides will exhibit a more pronounced body effect, since the ratio of the capacitance of the depletion layer to that of the oxide will be greater [Tsividis 86-88].

Performing the integrations in the drift and diffusion current equations and summing the two contributions gives the total drain current [Tsividis 110-111],

$$I_D = \underbrace{\frac{W}{L}\mu\int_{\psi_{s0}}^{\psi_{sL}}(-Q_I')d\psi_s}_{drift} + \underbrace{\frac{W}{L}\mu\phi_t\int_{Q_{I,source}'}^{Q_{I,drain}'}dQ_I'}_{diffusion}$$

$$= \underbrace{\frac{W}{L}\mu C_{ox}'\left[(V_{GB}-V_{FB})(\psi_{sL}-\psi_{s0})-\tfrac{1}{2}(\psi_{sL}^{\,2}-\psi_{s0}^{\,2})-\tfrac{2}{3}\gamma(\psi_{sL}^{\,3/2}-\psi_{s0}^{\,3/2})\right]}_{drift}$$

$$+\underbrace{\frac{W}{L}\mu C_{ox}'\left[\phi_t(\psi_{sL}-\psi_{s0})-\phi_t\gamma(\psi_{sL}^{\,1/2}-\psi_{s0}^{\,1/2})\right]}_{diffusion}$$

The surface potentials at the source and drain ends of the channel are related to the applied gate, source and drain potentials (V_{GB}, V_{SB}, and V_{DB}, relative to the body) by:

$$\psi_{s0} = V_{GB} - V_{FB} - \gamma\sqrt{\psi_{s0} + \phi_t e^{(\psi_{s0}-2\phi_F-V_{SB})/\phi_t}}$$

$$\psi_{sL} = V_{GB} - V_{FB} - \gamma\sqrt{\psi_{sL} + \phi_t e^{(\psi_{sL}-2\phi_F-V_{DB})/\phi_t}}$$

in which ϕ_F, the *Fermi potential*, is the surface potential at which the surface concentration of electrons and holes are equal (the transition between depletion and weak inversion conditions) [Tsividis 47-50]. Except at very high or low temperatures, the Fermi potential is a function of the doping concentration and the intrinsic carrier concentration of the semiconductor, n_i:

$$\phi_F \approx -\phi_t \ln\left(\frac{N_A}{n_i}\right)$$

While the charge sheet model above has the advantage of being quite general, accurately describing the MOSFET from weak inversion through strong inversion, the equations must be solved numerically, rendering them unsuitable for hand calculations and thus of limited value for first-order circuit design. However, by separately considering the case that the drain current is dominated by diffusion (weak inversion operation) and the case that the drain current is dominated by drift (strong inversion operation), these equations may be reduced to expressions which may be solved analytically, and which provide excellent accuracy in their respective regions of operation.

2.2 Weak Inversion Approximation of MOSFET Operation

In weak inversion operation, we assume that the surface potential depends only on the gate potential and is constant along the channel, so that drift current may be neglected. Taking only the diffusion component of the general drain current equation given above,

$$I_{D,weak\ inversion} \approx \frac{W}{L}\mu\phi_t \int_{Q'_{I,source}}^{Q'_{I,drain}} dQ'_I = -\frac{W}{L}\mu\phi_t \left(Q'_{I,source} - Q'_{I,drain}\right)$$

The inversion charge concentrations are given by [Tsividis 137]:

$$Q'_{I,source} = -\frac{\gamma C'_{ox}}{2\sqrt{\psi_{sa}(V_{GB})}}\phi_t e^{\left[\psi_{sa}(V_{GB})-2\phi_F\right]/\phi_t} e^{-V_{SB}/\phi_t}$$

$$Q'_{I,drain} = -\frac{\gamma C'_{ox}}{2\sqrt{\psi_{sa}(V_{GB})}}\phi_t e^{\left[\psi_{sa}(V_{GB})-2\phi_F\right]/\phi_t} e^{-V_{DB}/\phi_t}$$

in which the surface potential has been denoted by $\psi_{sa}(V_{GB})$ to indicate that it is dependent only on the gate potential, as follows [Tsividis 137]:

$$\psi_{sa}(V_{GB}) = \left(-\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + V_{GB} - V_{FB}}\right)^2$$

Noting the similarity between the expressions for $Q'_{I,source}$ and $Q'_{I,drain}$, the drain current equation may be re-written as [Tsividis 138]:

$$I_{D,weak\ inversion} = \frac{W}{L}\mu\phi_t\left(-Q'_{I,source}\right)\left(1 - e^{-V_{DS}/\phi_t}\right)$$

where V_{DS} is the drain-to-source voltage. Since the surface potential only ranges from approximately ϕ_F to $2\phi_F$ and the slope of ψ_s versus V_{GB} is almost constant in weak inversion, a good approximation can be made for fixed V_{SB} by assuming that ψ_s varies linearly with V_{GB} about the center of this region ($1.5\phi_F$) [Tsividis 96-97,139]:

$$I_D = \frac{W}{L} I'_X e^{(V_{GS}-V_X)/(n\phi_t)}\left(1 - e^{-V_{DS}/\phi_t}\right)$$

Here V_{GS} is the gate-to-source voltage, and I'_X and V_X are the drain current and gate-to-source voltage in the middle of the weak inversion operating region, when the surface potential is $1.5\phi_F$. The slope of V_{GB} versus ψ_s at this point is called the *ideality factor*, n. These quantities are given by [Tsividis 139]:

$$V_X = V_{FB} + 1.5\phi_F + \gamma\sqrt{1.5\phi_F + V_{SB}}$$

$$I'_X = \mu C'_{ox}\phi_t^{\,2}\frac{\gamma}{2\sqrt{1.5\phi_F + V_{SB}}} e^{-0.5\phi_F/\phi_t}$$

$$n \equiv \left.\frac{dV_{GB}}{d\psi_s}\right|_{\psi_s = 1.5\phi_F} = 1 + \frac{\gamma}{2\sqrt{1.5\phi_F + V_{SB}}}$$

Another, more intuitive description of the ideality factor is obtained by considering how the gate voltage sets the surface potential by the "capacitive

divider" effect of the gate oxide and depletion layer [Tsividis 62-69]. Because almost all of the charge deposited incrementally under the oxide layer in weak inversion is in the form of depletion charge, the incremental capacitance associated with the forming inversion layer may be neglected, and the ideality factor is simply the inverse of the capacitive divider that the oxide and depletion layer incremental capacitances (C'_{ox} and C'_d) form:

$$n = \left. \frac{C'_{ox} + C'_d}{C'_{ox}} \right|_{\psi_s = 1.5\phi_F}$$

The relationship between incremental changes in the gate-source potential and the drain current is the *gate transconductance*, g_m, which is the fundamental transfer characteristic of the MOSFET. It is defined as the derivative of drain current with respect to gate-source voltage, which for the case of weak inversion operation is:

$$\begin{aligned} g_m &\equiv \left. \frac{\partial I_D}{\partial V_{GS}} \right|_{V_{BS}, V_{DS} \text{ fixed}} \\ &= \frac{1}{n\phi_t} \frac{W}{L} I'_X e^{(V_{GS} - V_X)/(n\phi_t)} \left(1 - e^{-V_{DS}/\phi_t}\right) \\ &= \frac{I_D}{n\phi_t} \end{aligned}$$

This has the important properties of being independent of device geometry (channel length and width), and increasing in direct proportion to drain current. The significance of the ideality factor becomes evident in this expression, since the maximum possible transconductance is achieved as n approaches its theoretical minimum limit of unity. Saturation of the drain current occurs at small drain-source voltages, requiring a potential difference of only a few ϕ_t for the exponential term containing V_{DS} to rapidly approach zero.

2.3 Strong Inversion Approximation of MOSFET Operation

In strong inversion operation, a substantial inversion charge layer exists in the channel which allows the source and drain regions to strongly influence the surface potential. If the source and drain potentials differ, their influence will cause the surface potential to vary along the length of the channel, resulting in drift flow of inversion charge. This drift current dominates the current due to diffusion in the strong inversion case, so that the diffusion term in the general drain current equation may be neglected. By using a Taylor series approximation to eliminate the square root term in Q'_I (and hence the 3/2 power in the current equation), a useful expression for the drain current may be obtained [Tsividis 118-130]:

$$I_D = \begin{cases} \frac{W}{L}\mu C'_{ox}\left[(V_{GS}-V_T)V_{DS} - \frac{1}{2}(1+\delta)V_{DS}^2\right], & V_{DS} \le \frac{V_{GS}-V_T}{1+\delta} \\ \frac{W}{L}\mu C'_{ox}\left[\frac{(V_{GS}-V_T)^2}{2(1+\delta)}\right], & V_{DS} > \frac{V_{GS}-V_T}{1+\delta} \end{cases}$$

In this expression, V_T is the *extrapolated gate-source threshold voltage*, which is so named because it may be extrapolated from the curve of $\sqrt{I_D}$ versus V_{GS}. It depends on the source potential through the body effect according to:

$$V_T = V_{T0} + \gamma\left(\sqrt{\phi_B + V_{SB}} - \sqrt{\phi_B}\right)$$

$$V_{T0} = V_{FB} + \phi_B + \gamma\sqrt{\phi_B}$$

where ϕ_B is the *equilibrium strong inversion surface potential*,

$$\phi_B \approx 2\phi_F + 6\phi_t$$

and V_{T0} is the threshold voltage when the source and body terminals are at the same voltage. The term δ accounts for the change in depletion layer thickness along the length of the channel corresponding to the surface potential variation, and various values have been suggested ranging from $\delta = 0$ to $\delta = \frac{\gamma}{2\sqrt{\phi_B + V_{SB}}}$ [Tsividis 128-130]. For the case of transistors with lightly doped substrates and thin gate oxides, and thus small γ, it is often assumed that the depletion layer is relatively constant along the channel, and a value of $\delta = 0$ is chosen. Drain current saturation occurs in strong inversion when the drain-source voltage is slightly greater than the gate-source voltage minus the threshold voltage, "pinching off" the inversion layer at the drain end of the channel. Thus the farther into strong inversion the transistor is operated, the higher the drain-source voltage will have to be in order to operate in saturation. It is convenient to combine the linear and saturation region equations by defining the term α,

$$\alpha = \begin{cases} 1 - \frac{V_{DS}}{V'_{DS}}, & V_{DS} \le V'_{DS} \\ 0, & V_{DS} > V'_{DS} \end{cases}$$

where V'_{DS} denotes the *drain-source pinch-off voltage*,

$$V'_{DS} = \frac{V_{GS}-V_T}{1+\delta},$$

so that the drain current is described by the single expression:

$$I_D = \frac{W}{L}\mu C'_{ox}\frac{(V_{GS}-V_T)^2}{2(1+\delta)}\left(1-\alpha^2\right)$$

As in the weak inversion case, the gate transconductance is defined as the derivative of drain current with respect to gate-source voltage:

$$g_m = \begin{cases} \frac{W}{L}\mu C'_{ox}(V_{DS}), & V_{DS} \le \frac{V_{GS}-V_T}{1+\delta} \\ \frac{W}{L}\mu C'_{ox}(V_{GS}-V_T), & V_{DS} > \frac{V_{GS}-V_T}{1+\delta} \end{cases}$$

Alternatively,

$$g_m = \frac{W}{L}\mu C'_{ox} V'_{DS}(1-\alpha)$$
$$= \sqrt{2\frac{W}{L}\frac{\mu C'_{ox}}{1+\delta}\left(\frac{1-\alpha}{1+\alpha}\right)I_D}$$

which is clearly dependent on the width/length ratio, and is proportional to the square root of drain current.

2.4 Comparison of Weak and Strong Inversion Performance: The Transconductance-to-Current Ratio

In order to compare weak and strong inversion performance of the MOSFET, a useful parameter to consider is the transconductance-to-current ratio, which gives us a measure of the signal amplification available for a given expenditure of current (or power). In the weak inversion case, the ratio is constant:

$$\frac{g_m}{I_D} = \frac{1}{n\phi_t}$$

whereas in strong inversion, it is inversely proportional to the square root of drain current:

$$\frac{g_m}{I_D} = \sqrt{2\frac{W}{L}\frac{\mu C'_{ox}}{1+\delta}\left(\frac{1-\alpha}{1+\alpha}\right)\left(\frac{1}{I_D}\right)}$$
$$= \left(\frac{1}{1+\alpha}\right)\frac{2}{V_{GS}-V_T}$$

Since the transition from weak to strong inversion is continuous, the transconductance-to-current ratio is maximum in weak inversion and decreases monotonically as current increases through moderate and strong inversion, as sketched in Figure 2.

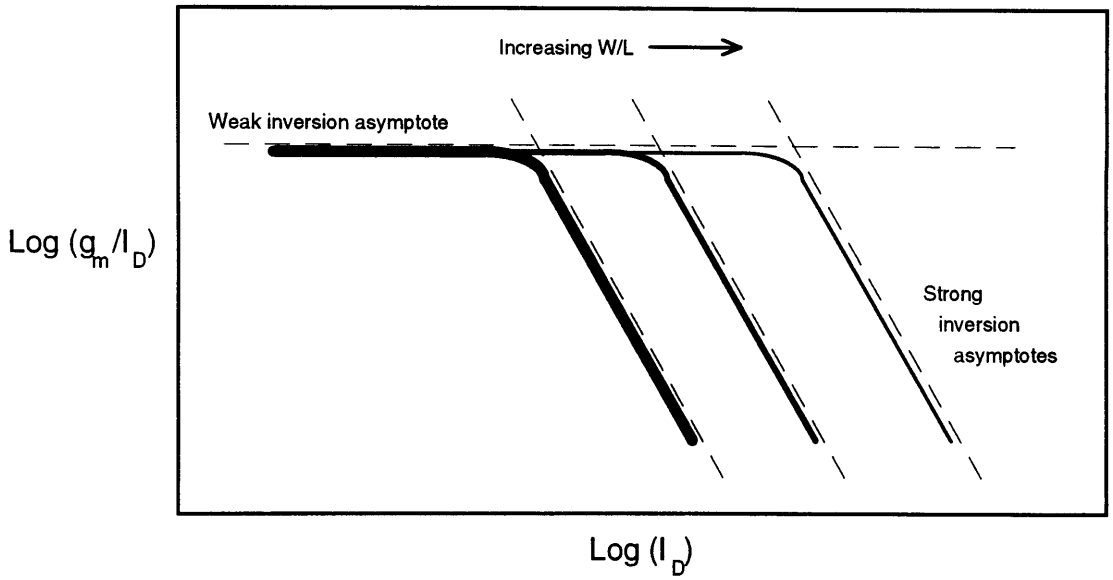

Figure 2: Transconductance-to-current ratio in weak and strong inversion

This suggests that the maximum transconductance (and thus signal amplification) for a given flow of drain current will be achieved in weak inversion. The design parameter that allows the transistor to be operated in weak or strong inversion at a specified current is the width/length ratio, as evident in this expression for the $2\phi_F$ current threshold between weak and strong inversion operation adapted from Deen and Yan [505]:

$$I_T = \frac{W}{L} I'_X e^{0.5\phi_F/\phi_t}\left(1 - e^{-V_{DS}/\phi_t}\right)$$

$$= \frac{W}{L}\mu C'_{ox}\phi_t^{\,2}(n-1)\left(1 - e^{-V_{DS}/\phi_t}\right)$$

Clearly, for a given fabrication process, the width/length ratio scales the threshold current, which can be determined by extracting the various parameters of the weak and strong inversion drain current equations through curve-fits of empirical data. Thus a transistor's width/length ratio may chosen such that the device will operate in weak inversion at a specified current level, maximizing its transconductance.

Chapter 3

Noise Analysis Techniques for the MOSFET

Noise in the MOSFET has traditionally been modeled by two primary sources, "$1/f$" and "white" noise. $1/f$ noise has a frequency spectrum that exhibits an approximately $1/f$ frequency dependence. White noise has a "flat" frequency spectrum, with power distributed evenly from very low frequencies to very high frequencies. The white noise component has generally been described in weak inversion as shot noise on the drain current, and in strong inversion as thermal (Johnson) noise of the channel resistance. Because of the disparate equations used to describe the transistor's operation and noise in weak and strong inversion, it was awkward to compare noise performance in the two cases. A more consistent description of noise in the two cases would be valuable in making circuit design decisions. With this in mind, the derivation of white noise in the MOSFET was approached from the perspective of *velocity fluctuation noise* of the charge carriers that make up the inversion layer. By relating the random motions of all the charge carriers in the inversion layer to electrical currents, the total drain current fluctuation was obtained.

The process of calculating noise observed at the terminals of a device can be a daunting task considering the various non-uniformities that may be present, such as charge distribution in the MOSFET channel. Fortunately, Van der Ziel has presented a relation that allows one to determine the terminal current noise of a device from an oftentimes simpler calculation of the AC short-circuit current noise for a small section of the device. His derivation is worth repeating, in order to show its generality and to lend insight into its utility. With only minor notation changes to apply his derivation to the general case of a resistive medium, his theory is presented below [Van der Ziel '86 74-75,290-291].

Van der Ziel begins by relating the terminal current of a resistive device (the MOSFET channel, in his presentation) to the electrical potential and resistivity along its length. In order to do so, it is convenient to deal with the somewhat unfamiliar quantity of conductance *for* unit length, g. This is simply the inverse of resistance *per* unit length, and it may be a function of position or the DC potential along the device. If a DC potential $V_0(x)$ exists along the device relative to one end, a current will flow, described by:

$$I = g(V_0)\frac{dV_0}{dx}$$

where the conductance for unit length has been expressed as a function of potential along the length in order to include any dependence. Now, if a current noise source $h(x,t)$ exists in the device at x, the changes in the potential and conductance at x as well as the direct effect of the noise source must be

considered in determining the terminal current fluctuation. The terminal current will now be described by:

$$I + \Delta I(t) = g(V)\frac{dV}{dx} + h(x,t)$$

where the DC potential at x has been replaced by the total voltage $V = V_0 + \Delta V$ which includes the potential fluctuation $\Delta V(x,t)$ caused by $h(x,t)$. Expanding in ΔV and neglecting the second-order term in order to obtain a linear approximation,

$$I + \Delta I(t) = \left(g(V_0) + \frac{dg(V_0)}{dV_0}\Delta V \right)\left(\frac{dV_0}{dx} + \frac{d\Delta V}{dx} \right) + h(x,t)$$

$$= \underbrace{g(V_0)\frac{dV_0}{dx}}_{I} + \frac{dg(V_0)}{dx}\Delta V + g(V_0)\frac{d\Delta V}{dx} + \underbrace{\frac{dg(V_0)}{dV_0}\frac{d\Delta V}{dx}\Delta V}_{neglect} + h(x,t)$$

gives the terminal current fluctuation,

$$\Delta I(t) \approx \frac{dg(V_0)}{dx}\Delta V + g(V_0)\frac{d\Delta V}{dx} + h(x,t)$$

$$\approx \frac{d}{dx}\left[g(V_0)\Delta V(x,t) \right] + h(x,t)$$

To find the "short-circuit" current noise, this is multiplied by dx and integrated over the length of the device,

$$\int_0^l \Delta I(t)dx = \int_0^l d\left[g(V_0)\Delta V(x,t) \right] + \int_0^l h(x,t)dx$$

noting that the terminals will be connected for small signals, so that the potential fluctuation at both ends (defined relative to one end) will be zero ($\Delta V(0,t) = \Delta V(l,t) = 0$), giving the result:

$$\Delta I(t)l = \underbrace{\int_0^l d\left[g(V_0)\Delta V(x,t) \right]}_{0} + \int_0^l h(x,t)dx$$

$$\Delta I(t) = \frac{1}{l}\int_0^l h(x,t)dx$$

According to the Wiener-Khintchine theorem [Van der Ziel '86 10-14, 283-285], the spectral intensity is found by taking the Fourier transform of the autocorrelation function (and multiplying by two):

$$S_I(f) = 2\Im\left\{ \overline{\Delta I(t)\Delta I(t+\tau)} \right\}$$

$$= 2\Im\left\{ \frac{1}{l^2}\int_0^l \int_0^l \overline{h(x,t)h(x',t+\tau)}dxdx' \right\}$$

$$= \frac{1}{l^2}\int_0^l \int_0^l S_h(x,x',f)dxdx'$$

where $S_I(f)$ denotes the spectral intensity of the terminal current fluctuation, $\Im\{\cdot\}$ is the Fourier transform operator, and $S_h(x,x',f)$ is the spatial cross-spectral

intensity of the noise source. Since $h(x,t)$ was the localized noise source at position x, $S_h(x,x',f)$ must be a delta function in $(x'-x)$:

$$S_h(x,x',f) = F(x,f)\delta(x'-x)$$

so that the double integral in $S_I(f)$ reduces to a single integral in x:

$$S_I(f) = \frac{1}{l^2}\int_0^l F(x,f)dx$$

Now the challenge lies in determining $F(x,f)$. Van der Ziel cleverly notes that the equation above holds for sections of any length l, so if it is applied to a small section of the device (of length $l = \Delta x$) at position x that is hypothetically short-circuited, the spectral intensity of its current will be:

$$\begin{aligned} S_{\Delta I}(x,f) &= \frac{1}{\Delta x^2}\int_x^{x+\Delta x} F(x,f)dx \\ &\approx \frac{1}{\Delta x^2}\left[F(x,f)\Delta x\right] \\ &\approx \frac{F(x,f)}{\Delta x} \end{aligned}$$

Hence $F(x,f)$ is obtained directly from the short-circuit current noise of a sufficiently small section of the device at x:

$$F(x,f) = S_{\Delta I}(x,f)\Delta x$$

which is integrated over the length of the device to find the terminal current spectral intensity. This "sectioning" technique transforms the calculation of total noise into a determination of the noise of small, homogeneous segments of an otherwise potentially complicated device. Since Van der Ziel's derivation (presented above) does not require a DC potential gradient along the device, it applies to the MOSFET channel in weak inversion as well.

Chapter 4

Theory of White Noise in the MOSFET

4.1 Drain Current Noise from Velocity Fluctuation of Carriers

The derivation of white noise in the MOSFET from the perspective of velocity fluctuation noise begins with the application of Ramo's theorem, which is presented here for convenience [Van der Ziel '86 68-69]. To understand Ramo's theorem, first consider the current flow resulting from a charge carrier moving between two parallel planar electrodes separated by distance d with some component of velocity perpendicular to the plates v_x, as illustrated in Figure 3.

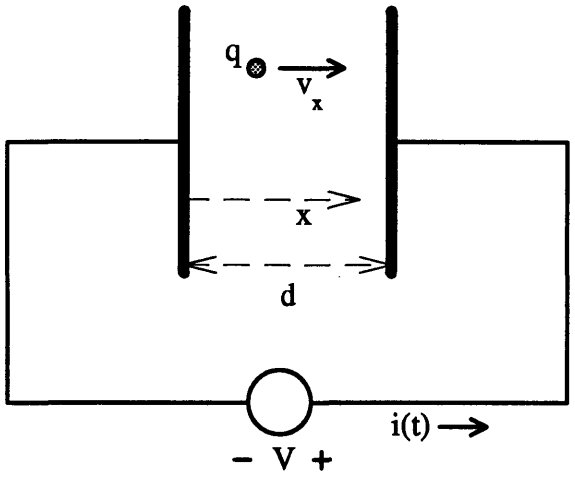

Figure 3: Current associated with the motion of a single charge

If the carrier has charge q and there is a voltage source V connecting the two plates, the energy gained (or lost, if V is negative) by the charge during time dt will be:

$$\Delta E_q = q\varepsilon_x dx = q\frac{V}{d}dx$$

where dx is the distance traversed in time dt. This must be equal to the energy lost/gained by the external voltage source:

$$\Delta E_V = V \cdot i(t)dt$$

so that the current flow resulting from the velocity of the charge carrier is

$$i(t) = q\frac{1}{d}\frac{dx}{dt} = \frac{qv_x}{d}$$

which is known as Ramo's theorem [Van der Ziel '86 23]. Clearly this relation holds in the absence of the external voltage source, i.e. when $V = 0$.

Ramo's theorem can be applied to mobile charges in a transistor channel by considering a small section of length Δx at distance x from the source, in order to take advantage of Van der Ziel's "sectioning" technique for noise calculation described previously. As an inversion charge moves through the channel, it will repeatedly collide with the semiconductor lattice, changing velocity in both direction and magnitude. But if the ends of the channel segment were (hypothetically) shorted together for AC signals, the component of the carrier's velocity between the ends $v_x(t)$ would give rise to a current, which is described by Ramo's theorem. The average value of this velocity $\bar{v}_x$, such as would result from an applied electric field, corresponds to the drift current. The fluctuation in velocity (due to the random nature of the carrier-lattice collisions),

$$\Delta v_x(t) = v_x(t) - \bar{v}_x$$

is thus responsible for the noise current associated with this carrier.

Taking the autocorrelation of the velocity fluctuation and applying the Wiener-Khintchine theorem gives the spectral intensity of $v_x(t)$, the "velocity-fluctuation noise" [Van der Ziel '86 68]:

$$S_{\Delta v_x}(x,f) = 2 \cdot \Im\left\{\overline{\Delta v_x(t)\Delta v_x(t+\tau)}\right\}$$

$$= 4\int_0^\infty \overline{\Delta v_x(t)\Delta v_x(t+\tau)}\cos(\omega\tau)d\tau$$

Kubo has defined the high-frequency diffusion coefficient D as the Fourier transform of $\overline{\Delta v_x(t)\Delta v_x(t+\tau)}$ [Van der Ziel '86 68)]:

$$D(f) = \int_0^\infty \overline{\Delta v_x(t)\Delta v_x(t+\tau)}e^{-j\omega\tau}d\tau$$

so that

$$S_{\Delta v_x}(x,f) = 4\,\mathrm{Re}\{D(f)\}$$

and Van der Ziel has shown that this relationship holds for the reasonably low-frequency case where the diffusion coefficient $D_0 \equiv D(f \to 0)$ becomes almost entirely real ['86 68-69]:

$$S_{\Delta v_x}(x, f_{low}) = 4D_0$$

Now that the spectral intensity of the velocity fluctuation is known, Ramo's theorem may be applied for the individual carrier:

$$\Delta i_1(t) = \frac{q\Delta v_x(t)}{\Delta x}$$

so that the spectral intensity of the short-circuit current through the segment is

$$S_{\Delta i_1}(x,f) = \frac{q^2 S_{\Delta v_x}(f)}{\Delta x^2} = \frac{4q^2\,\mathrm{Re}\{D(f)\}}{\Delta x^2}$$

$$= \frac{4q^2 D_0}{\Delta x^2}, \text{ modest frequencies}$$

for the single charge carrier. As long as the random motions of the charge carriers in the section Δx are independent, their individual current spectral

intensities will sum to give the total short-circuit current spectral intensity of the segment, $S_{\Delta i_{\Delta N}}(f)$. Thus if $N'(x)$ is the number of inversion charges per unit area in the channel at x, there are $\Delta N = N'(x)W\Delta x$ total mobile charge carriers in the segment, so the noise for the small section is:

$$S_{\Delta i_{\Delta N}}(x,f) = \frac{4q^2 D_0}{\Delta x^2} N'(x) W \Delta x$$

$$= \frac{4q^2 W D_0}{\Delta x} N'(x)$$

To determine the total drain current noise of the transistor by Van der Ziel's approach, the function $F(x,f)$ associated with this segment noise is needed:

$$F(x,f) = S_{\Delta i_{\Delta N}}(x,f)\Delta x$$

$$= \left(\frac{4q^2 W D_0}{\Delta x} N'(x) \right) \Delta x$$

$$= 4q^2 W D_0 N'(x)$$

Integrating this over the length of the transistor channel yields the total drain current spectral intensity,

$$S_{I_D}(f) = \frac{1}{L^2} \int_0^L F(x,f) dx$$

$$= \frac{1}{L^2} \int_0^L \left(4q^2 D_0 W N'(x) \right) dx$$

$$= \frac{4q^2 W D_0}{L^2} \int_0^L N'(x) dx$$

assuming the diffusion coefficient is constant throughout the channel, which simply reduces to:

$$S_{I_D}(f) = \frac{4q^2 D_0}{L^2} N$$

where N is the total number of inversion charges in the channel. Applying Einstein's relationship between diffusion and mobility:

$$\frac{D_0}{\mu} = \frac{kT}{q} = \phi_t$$

yields an equivalent expression:

$$S_{I_D}(f) = \frac{4kTq\mu}{L^2} N, \quad f \to 0$$

Interestingly, applying Ramo's theorem to a charge in the inversion layer to determine its contribution to the drain current noise directly (instead of considering small segments of the channel) and then multiplying by N to sum the independent contributions of all the carriers yields the same result for the drain current spectral intensity. This underscores how the white noise in the MOSFET is the direct result of random thermal motion of the inversion layer charges (neglecting other minor sources of white noise in the device).

Thus in order to know the white noise current observed at the drain and source terminals of the MOSFET, only the number of charges in the inversion layer must be determined (assuming that the channel length and some physical constants are known). The number of carriers in the channel may be determined by integrating the surface density of the inversion charge over the channel area:

$$N = -\frac{1}{q}Q_{total} = -\frac{1}{q}\iint_{WL} Q_I' dxdy$$

Because it was assumed that the channel is uniform across its width, that integration can be performed first, leaving only the integral along the channel length:

$$N = -\frac{1}{q}W\int_0^L Q_I'(x)dx$$

4.2 White Noise in Weak Inversion Operation

For the weak inversion case, the inversion charge density varies linearly from source to drain because the current is constant along the channel and due only to diffusion. Integration of the charge reduces to finding the area of a trapezoid, so that

$$N = -\frac{WL}{q}\frac{\left(Q_{I,source}' + Q_{I,drain}'\right)}{2} = -\frac{WL}{q}\frac{Q_{I,source}'}{2}\left(1+e^{-V_{DS}/\phi_t}\right)$$

which is conveniently expressed in terms of drain current and transconductance as follows:

$$N = \frac{L^2}{2q\mu\phi_t}\left(\frac{1+e^{-V_{DS}/\phi_t}}{1-e^{-V_{DS}/\phi_t}}\right)I_D = \frac{L^2}{2qD_0}\left(\frac{1+e^{-V_{DS}/\phi_t}}{1-e^{-V_{DS}/\phi_t}}\right)I_D$$

$$= \frac{nL^2}{2q\mu}\left(\frac{1+e^{-V_{DS}/\phi_t}}{1-e^{-V_{DS}/\phi_t}}\right)g_m$$

Thus the drain current spectral intensity in weak inversion is

$$S_{I_D}(f) = 2q\left(\frac{1+e^{-V_{DS}/\phi_t}}{1-e^{-V_{DS}/\phi_t}}\right)I_D$$

in terms of drain current, which limits to the traditional result from the "shot noise" derivation when the drain current saturates:

$$S_{I_D}(f) = 2qI_{D,saturation}$$

The noise may also be expressed as a function of the transconductance:

$$S_{I_D}(f) = 4kT\left(\frac{n}{2}\right)\left(\frac{1+e^{-V_{DS}/\phi_t}}{1-e^{-V_{DS}/\phi_t}}\right)g_m$$

$$\Rightarrow 4kT\left(\frac{n}{2}\right)g_m, \text{ saturation}$$

4.3 White Noise in Strong Inversion Operation

The inversion charge density equation associated with the strong inversion drain current approximation is [Tsividis 124-125]

$$Q_I'(V_{CB}) = -C_{ox}'\left[V_{GS} - V_T - (1+\delta)(V_{CB} - V_{SB})\right]$$

which is expressed in terms of a useful construct, V_{CB}, the *effective reverse bias* between the induced n-type inversion layer and the substrate. The effective reverse bias is the deviation in surface potential from the equilibrium strong inversion surface potential,

$$V_{CB}(x) = \psi_s(x) - \phi_B$$

and it varies from $V_{CB}(0) = V_{SB}$ at the source end to $V_{CB}(L) = V_{DB}$ at the drain end of the channel. To find the number of inversion charge carriers, Q_I' must be integrated over the area of the channel as before,

$$N = -\frac{1}{q} W \int_0^L Q_I'(x)dx$$

but since Q_I' is known as a function of V_{CB} rather than x, it will be expedient to convert the integration over distance to an integral over V_{CB}. Recognizing that

$$\frac{dV_{CB}}{dx} = \frac{d}{dx}\left[\psi_s(x) - \phi_B\right] = \frac{d\psi_s(x)}{dx} = I_D \frac{dR(x)}{dx}$$

where the *incremental channel resistance* per unit length in strong inversion is:

$$\frac{dR(x)}{dx} = -\frac{1}{\mu W Q_I'(x)}$$

allows dx and $Q_I'(x)$ to be replaced:

$$dx = -\frac{\mu W Q_I'(x)}{I_D} dV_{CB}$$

$$Q_I'(x) = Q_I'\left(V_{CB}(x)\right)$$

so that the integral becomes

$$N = \frac{1}{q} W \int_{V_{CB}(0)}^{V_{CB}(L)} Q_I'(V_{CB}) \frac{\mu W Q_I'(V_{CB})}{I_D} dV_{CB}$$

$$= \frac{\mu W^2}{q I_D} \int_{V_{SB}}^{V_{DB}} \left[Q_I'(V_{CB})\right]^2 dV_{CB}$$

Performing the integration yields the number of inversion charge carriers:

$$N = \frac{\mu W^2}{qI_D}\int_{V_{SB}}^{V_{DB}}\left[-C'_{ox}\left[V_{GS}-V_T-(1+\delta)(V_{CB}-V_{SB})\right]\right]^2 dV_{CB}$$

$$= \frac{\mu W^2 C'^{\,2}_{ox}}{qI_D}\frac{-1}{3(1+\delta)}\left[V_{GS}-V_T-(1+\delta)(V_{CB}-V_{SB})\right]^3\Big|_{V_{CB}=V_{SB}}^{V_{CB}=V_{DB}}$$

$$= \frac{\mu W^2 C'^{\,2}_{ox}}{3qI_D(1+\delta)}\left[(V_{GS}-V_T)^3-(V_{GS}-V_T-(1+\delta)(V_{DB}-V_{SB}))^3\right]$$

$$= \frac{\mu W^2 C'^{\,2}_{ox}}{3qI_D(1+\delta)}\left[((1+\delta)V'_{DS})^3-(\alpha(1+\delta)V'_{DS})^3\right]$$

$$= \frac{\mu W^2 C'^{\,2}_{ox}(1+\delta)^2}{3qI_D}V'^{\,3}_{DS}(1-\alpha^3)$$

$$= \frac{2WLC'_{ox}}{3q}\left(\frac{1-\alpha^3}{1-\alpha^2}\right)(V_{GS}-V_T)$$

In terms of transconductance,

$$N = \frac{2L^2(1+\delta)}{3q\mu}\left(\frac{1+\alpha+\alpha^2}{1-\alpha^2}\right)g_m$$

Substituting N into the white noise drain current spectral intensity formula yields

$$S_{I_D}(f) = 4kT\mu C'_{ox}\left(\frac{W}{L}\right)\left(\frac{2}{3}\right)\left(\frac{1-\alpha^3}{1-\alpha^2}\right)(V_{GS}-V_T)$$

for the strong inversion case. In terms of the transconductance,

$$S_{I_D}(f) = 4kT\left(\frac{2}{3}(1+\delta)\right)\left(\frac{1+\alpha+\alpha^2}{1-\alpha^2}\right)g_m$$

$$\Rightarrow 4kT\left(\frac{2}{3}(1+\delta)\right)g_m, \text{ saturation}$$

which is the familiar thermal noise expression with the addition of the depletion layer correction factor δ. As stated earlier, δ is often assumed to be zero, giving the typical equation $4kT(\frac{2}{3})g_m$ for the noise.

4.4 Connection Between Velocity-Fluctuation, Thermal, and Shot Noise in Resistors and MOSFETs

It is worth noting that the regular Johnson (thermal) noise expression is obtained when the velocity fluctuation noise derivation is applied to a resistor, as expected. For example, consider a uniform resistor R of length L and cross-sectional area A. The resistance is related to the carrier volume density $\hat{N}$ by:

$$G = \frac{1}{R} = q\mu\hat{N}\frac{A}{L}$$

so that

$$\hat{N} = \frac{L}{q\mu RA}$$

and

$$N = AL\hat{N} = \frac{L^2}{q\mu R},$$

and hence the spectral intensity of the short-circuit current is

$$S_{I_D}(f) = \frac{4kTq\mu}{L^2} N$$
$$= \frac{4kT}{R}$$

as it should be.

Since random motion of charge carriers is the fundamental basis of both "shot" noise and "thermal" noise, it is reasonable that both approaches may be used to describe current fluctuations in resistive media [Kingston 69,55-71; Sarpeshkar, Delbrück and Mead 23-29], as suggested by the weak inversion noise equation found above. The key to connecting the two approaches in MOSFETs and resistors is to note that diffusion currents are a direct result of random thermal motion of carriers, and are thus inherently linked with noise. The random motions of many individual charge carriers usually combine to cancel each other, such that no net current is produced. But in the special case that there is a concentration gradient, the random motions of the carriers may combine (in the absence of other influences) to produce a net diffusion current. If the carrier concentration in a region with cross-sectional area A transitions linearly from $\hat{N}(0)$ at one boundary to zero at the other, over distance L, as depicted in Figure 4, the diffusion current due to this gradient will be

$$I = qAD\frac{\hat{N}(0)}{L}$$

and the total number of carriers will be

$$N = \tfrac{1}{2}AL\hat{N}(0)$$

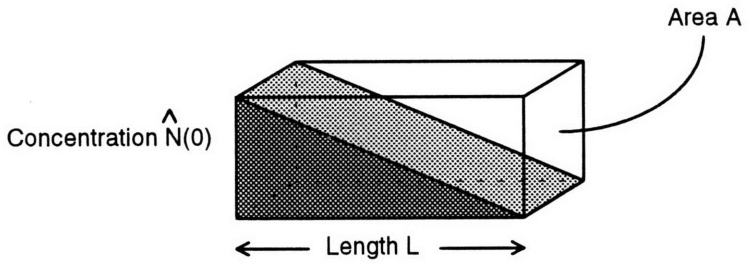

Figure 4: Volume with carrier concentration linearly decreasing to zero

By applying the velocity-fluctuation-based noise formula,

$$S_I(f) = \frac{4q^2 D}{L^2}\left(\tfrac{1}{2} AL\hat{N}(0)\right)$$
$$= 2q\left(qAD\frac{\hat{N}(0)}{L}\right)$$
$$= 2qI$$

which is equivalent to shot noise *on the diffusion current*. This diffusion current *is* the current through the device terminals in this case, so the shot noise formula may be applied to the terminal current as well. Shot noise specifically describes the fluctuation in the occurrence of Poisson events, such as the random arrival of charge carriers at the terminal of this device. In the case that the concentration of carriers is held to zero at one boundary, the rate of arrival of charges at that terminal is directly proportional to the terminal current, since all the carriers that arrive there will contribute to the terminal current. However, if the carrier concentration does not go to zero at one boundary, then the net diffusion current is not directly proportional to the rate of arrival of carriers at the terminal, and accordingly the shot noise formula applied to the terminal current underestimates the actual noise. But if the carriers that originate from each end of the device could be labeled, their random thermal motions could be tracked by watching them diffuse across the device to the opposite ends, where their concentration is (by definition) zero. This is illustrated in Figure 5.

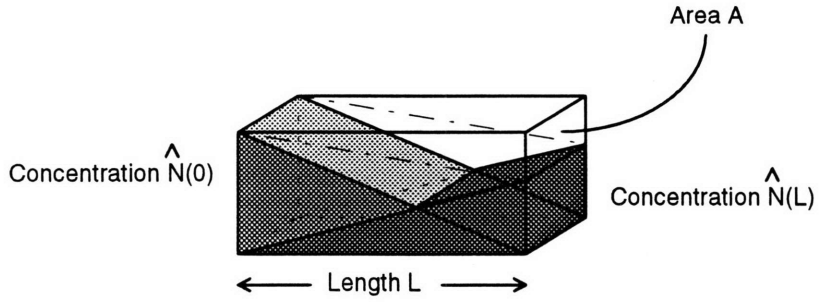

Figure 5: Volume with linear concentration profile but non-zero boundary conditions, modeled as sum of two linear profiles that do go to zero

Thus there would be two independent currents moving in opposite directions,

$$I_f = qAD\frac{\hat{N}(0)}{L}$$

and

$$I_r = qAD\frac{\hat{N}(L)}{L}$$

where I_f and I_r are the diffusion currents originating from the boundaries at the left and right, respectively. The net current through the terminals would be the *difference* between these forward and reverse currents, $I_{net} = I_f - I_r$, whereas the

current fluctuation observed at the terminals would be the *sum* of the shot noise contributions of each component:

$$S_{I_{total}}(f) = 2qI_f + 2qI_r$$
$$= 2q\left(qAD\frac{\hat{N}(0)}{L}\right) + 2q\left(qAD\frac{\hat{N}(L)}{L}\right)$$

Now the random thermal motion of all the charge carriers is accounted for, and it follows that the spectral intensity should be the same as that obtained when the velocity fluctuation of the total number of carriers $N = AL\left(\frac{\hat{N}(0)+\hat{N}(L)}{2}\right)$ is considered:

$$S_I(f) = \frac{4q^2D}{L^2}\left[AL\left(\frac{\hat{N}(0)+\hat{N}(L)}{2}\right)\right]$$
$$= 2q\left[qAD\frac{\hat{N}(0)+\hat{N}(L)}{L}\right]$$
$$= 2q\left(qAD\frac{\hat{N}(0)}{L}\right) + 2q\left(qAD\frac{\hat{N}(L)}{L}\right)$$
$$= 2qI_f + 2qI_r$$

This derivation can be applied to a resistor, where the carrier concentration at either end is, in terms of the resistance,

$$\hat{N}(0) = \hat{N}(L) = \frac{L}{q\mu RA}$$

as noted earlier. The mock diffusion currents are given by I_f and I_r:

$$I_f = I_r = qAD\frac{\frac{L}{q\mu RA}}{L} = \frac{D}{\mu R} = \frac{kT}{qR}$$

where the Einstein relation has been employed, so that the current spectral intensity given by shot noise theory is

$$S_I(f) = 2qI_f + 2qI_r$$
$$= 2q\left(\frac{kT}{qR}\right) + 2q\left(\frac{kT}{qR}\right)$$
$$= \frac{4kT}{R}$$

which simplifies to the thermal noise formula. It is important to remember that drift current, due to an applied electric field, has no effect on the noise of the resistor (excluding nonlinearities and heating of the resistor) because the drift current corresponds to the average velocity of charge carriers. The noise is associated with the deviations from this average carrier velocity.

Chapter 5

Theory of 1/f Noise in the MOSFET

$1/f$ noise is so named because of the frequency dependence of its spectral intensity, which varies as $1/|f|^{\alpha}$, where α is close to unity, typically between 0.8 and 1.4. This type of noise is also referred to as "flicker noise". While $1/f$ fluctuations have been observed in many different settings, including biological, geological, mechanical, and electronic systems, no entirely satisfactory physical explanation has been developed, and even in electronic systems it appears that the origins of the noise may be quite different in various types of devices [Buckingham 143-145]. In homogeneous resistors, $1/f$ noise is usually seen in the presence of a DC current or voltage (as a fluctuation in the voltage or current, respectively), and varies as the square of this DC bias. This suggests that the noise is attributable to a fluctuation in the resistance, since

$$S_V(f) = I^2 S_R(f) = \frac{V^2}{R^2} S_R(f)$$

or

$$S_I(f) = V^2 S_G(f) = \frac{I^2}{G^2} S_G(f)$$

where $S_V(f)$ and $S_I(f)$ are the voltage and current spectral intensity in the presence of DC current I or voltage V, due to fluctuations in the average resistance R or conductance G (spectral intensities $S_R(f)$ and $S_G(f)$). According to this view, the DC bias is not a cause of $1/f$ noise phenomena. Instead, the DC bias merely provides a means to observe the effects of some $1/f$-type fluctuation (i.e. in the resistance). In fact, experiments have shown that when sinusoidal signals (instead of DC signals) are applied to resistors, $1/f$-type sidebands will appear in the power spectrum on either side of the applied signal frequency. This is consistent with the theory that $1/f$ noise is due to resistance fluctuation, since the amplitude of the current or voltage sinusoid resulting from the applied sinusoid would be modulated by the resistance fluctuation. When both DC and AC signals are applied, the $1/f$ fluctuations centered at $f = 0$ and at the sinusoid frequency have been shown to be highly correlated, as predicted. In addition, the thermal noise level of a resistor (which depends directly on the resistance) has been reported to vary with a $1/f$ dependence. [Buckingham 159-160]

Although there is much evidence to support resistance fluctuations as the source of $1/f$ noise in homogeneous specimens, the cause of these resistance fluctuations in thermal equilibrium is not clear. Since resistance depends on the

number and mobility of charge carriers, fluctuation in either of these quantities could be responsible for resistance fluctuations. There is evidence supporting and contradicting both mechanisms as the cause of resistance fluctuations [Buckingham 162-163]. In the case of the MOS transistor, the flicker noise level is greatly affected by surface treatments, suggesting that charge traps at the semiconductor-oxide interface are responsible by causing significant fluctuations in the number of carriers. The fact that JFETs often have very little $1/f$ noise suggests that in the MOSFET case $1/f$ noise may be primarily a surface effect (in many other cases $1/f$ noise appears to be a bulk effect). Although there is some evidence to support mobility fluctuation as at least a partial cause of resistance fluctuations in the MOSFET channel, the number fluctuation due to surface traps is likely to be the main cause since the number of trapping centers may be significant compared to the number of charge carriers.

$1/f$ noise due to carrier number fluctuations is a special case of generation-recombination noise [Van der Ziel '86 120-130]. Generation-recombination noise is the fluctuation in the number of carriers resulting from their random generation and recombination, as the name implies. Charge traps trap and release (recombine and generate) carriers with a characteristic time constant which determines the frequency content of the fluctuation. If the population of charge traps is such that their time constants τ are distributed continuously over some range $\tau_0 \le \tau \le \tau_1$ with probability proportional to $1/\tau$, the spectral intensity of the number fluctuation will have a $1/f$ spectrum over the range $1/\tau_1 << \omega << 1/\tau_0$ [Van der Ziel '86 126]. For the MOSFET, it has been hypothesized that a trapping mechanism capable of having time constants distributed in this manner is the tunneling of charges from the semiconductor to traps at various distances into the oxide [Christensson, Lundstrom, and Svensson 797-812; Van der Ziel '86 127-128]. If the traps are distributed uniformly over some distance from the oxide surface, the distribution of time constants will be as $1/\tau$. For completeness, the "microscopic model of $1/f$ noise" developed by Van der Ziel will be presented here.

According to Van der Ziel ['86 296-297], if the number of traps with energy between E and $E+\Delta E$ in a small section of the gate oxide with volume $\Delta x \Delta y \Delta z$ is $\Delta N_T = N_T(E)\Delta E \Delta x \Delta y \Delta z$, and ΔN_t of them are occupied and the trapped charges obey Binomial statistics, then

$$\overline{\Delta N_t} = \Delta N_T f_t$$

and

$$\overline{\delta \Delta N_t^2} = \Delta N_T f_t (1 - f_t)$$

where the probability of occupancy is the Fermi function:

$$f_t = \frac{1}{\left(1 + e^{\frac{E - E_f}{kT}}\right)}$$

If the characteristic time constant of the traps in this volume element is τ, the spectrum of the fluctuation in number of occupied traps $\delta\Delta N_t$ is

$$S_{\Delta N_t}(f) = 4\Delta N_T f_t(1-f_t)\left(\frac{\tau}{1+\omega^2\tau^2}\right)$$

$$= 4N_T(E)\Delta E\Delta x\Delta y\Delta z f_t(1-f_t)\left(\frac{\tau}{1+\omega^2\tau^2}\right)$$

Assuming that trapping occurs by tunneling of charges to the traps in the oxide, the time constant of the traps will be an exponential function of distance z from the semiconductor surface:

$$\tau = \tau_0 e^{\gamma_t z}$$

where $\gamma_t \approx 10^8 cm^{-1}$ is the tunneling parameter. If it is further assumed that the distribution of traps in the oxide is uniform from the semiconductor surface at $z=0$ to distance $z=z_1$ into the oxide, and zero outside of this region, the time constants will be distributed as $1/\tau$ between τ_0 and $\tau_1 = \tau_0 e^{\gamma_t z_1}$. Replacing dz by $d\tau/\gamma_t\tau$ and integrating $S_{\Delta N_t}(f)$ with respect to τ yields

$$S_{\Delta N_t}(f) = \frac{N_T(E)\Delta E\Delta x\Delta y}{\gamma_t f} f_t(1-f_t), \text{ for } \frac{1}{2\pi\tau_1} < f < \frac{1}{2\pi\tau_0}$$

While the trap density may be a complicated function of energy (electrical potential in the oxide), only those traps whose probability of occupancy is not close to zero or unity will contribute significantly to the fluctuations. These "active" traps are the ones with energy close to the Fermi level. This dependence is reflected in the sharp peak of $f_t(1-f_t)$ at $E=E_f$. Because of this sharp peak, the integration of $S_{\Delta N_t}(f)$ with respect to energy can be simplified by using the "effective trap density":

$$N_T(E_f)_{eff} = \frac{1}{kT}\int_{-\infty}^{\infty} N_T(E) f_t(1-f_t) dE$$

For the case that $N_T(E)$ varies slowly with E, this has the value of $N_T(E_f)$, suggesting an "effective width" of the energy distribution of active traps equal to $4kT$. The integration of $S_{\Delta N_t}(f)$ using the effective trap density results in

$$S_{\Delta N_t}(f) = \frac{N_T(E_f)_{eff} kT\Delta x\Delta y}{\gamma_t f}$$

and integrating over the uniform width of the channel (from $y=0$ to W) yields

$$S_{\Delta N_t}(x,f) = \frac{N_T(E_f)_{eff} kT}{\gamma_t f} W\Delta x$$

Reimbold has suggested that a change in the quantity of trapped charge will be compensated by appropriate changes in the local inversion and depletion charge to maintain charge conservation [Reimbold, 1191]. The fraction of the compensation charge that is provided by (or added to) the inversion charge is a

function of the surface potential, since this determines the incremental distribution of charges at the channel surface. Small-signal capacitances are associated with each component of charge stored at the gate for incremental changes in surface potential $\delta\psi_s$, as follows:

Gate electrode charge fluctuation: $\delta Q'_g = -C'_{ox}\delta\psi_s$

Fast surface state charge fluctuation: $\delta Q'_{ss} = -C'_{ss}\delta\psi_s$

Depletion region charge fluctuation: $\delta Q'_d = -C'_d\delta\psi_s$

Inversion layer charge fluctuation: $\delta Q'_I = -C'_i\delta\psi_s$

Trapped charge fluctuations $\delta Q'_t$ must be balanced by these fluctuations according to the conservation of charge, so that

$$\delta Q'_t = -\left(\delta Q'_g + \delta Q'_{ss} + \delta Q'_d + \delta Q'_I\right)$$

Thus the ratio of inversion charge fluctuation to trapped charge fluctuation in the small section of channel Δx is

$$\frac{\delta\Delta N}{\delta\Delta N_t} = \frac{\delta Q'_I}{\delta Q'_t} = \frac{\delta Q'_I}{-\left(\delta Q'_g + \delta Q'_{ss} + \delta Q'_d + \delta Q'_I\right)}$$

$$= \frac{-C'_i\delta\psi_s}{C'_{ox}\delta\psi_s + C'_{ss}\delta\psi_s + C'_d\delta\psi_s + C'_i\delta\psi_s}$$

$$= \frac{-C'_i}{C'_{ox} + C'_{ss} + C'_d + C'_i}$$

so the spectral intensity of the inversion charge fluctuation in Δx is

$$S_{\Delta N}(x,f) = \left(\frac{\delta\Delta N}{\delta\Delta N_t}\right)^2 S_{\Delta N_t}(x,f)$$

$$= \left(\frac{\delta\Delta N}{\delta\Delta N_t}\right)^2 \frac{N_T\left(E_f\right)_{eff} kT}{\gamma_t f} W\Delta x$$

The modulation of inversion charge along the channel will affect the drain current through different mechanisms in weak and strong inversion. Characteristic of $1/f$ noise, the effect will be seen in the presence of a DC current through the device in either case. In strong inversion, the change in the number of mobile carriers in the channel will cause local resistance fluctuations. These resistance fluctuations will cause voltage fluctuations along the length of the channel (in the presence of a DC current), which will in turn result in drain current fluctuations. In weak inversion, the channel potential is assumed to be constant along its length, and the "resistance" variations due to inversion charge fluctuation will not create voltage fluctuations. Rather, the inversion charge fluctuations along the channel will change the carrier concentration profile, and hence the diffusion current. The effect of charge trapping in weak inversion is perhaps more appropriately described as changing the local "solubility" of the channel for carriers diffusing from source to drain.

In spite of the different mechanisms of drain current modulation in weak and strong inversion, Van der Ziel's sectioning technique can be applied to calculate the drain current noise based on the local resistance fluctuation paradigm in both cases. Because Van der Ziel's technique calculates drain current fluctuations from the *short-circuit* current fluctuations of small segments of the channel (one at a time), it imposes the condition that there is no voltage fluctuation along the channel. With zero voltage fluctuation there will be no perturbation of the rest of the channel to consider, so the task of determining the affect of local fluctuations on the whole channel is avoided. This is illustrated by the circuit model of the channel shown in Figure 6, in which a large capacitor illustrates the hypothetical short-circuiting of transient noise currents in the section Δx. Thus the fact that the DC current flowing through the channel in weak inversion is due to diffusion rather than drift does not affect our calculation of the change in current flowing through a section whose carrier concentration (and thus resistance to current flow) has changed.

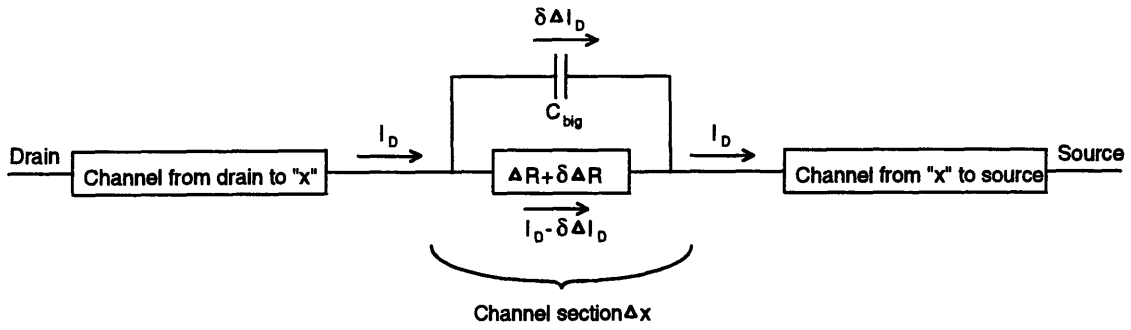

Figure 6: Circuit representation of Van der Ziel's sectioning method

The resistance of a small section of channel between x and $x+\Delta x$ is:

$$\Delta R(x) = -\frac{\Delta x}{\mu W Q_I'(x)}$$

$$= \frac{\Delta x^2}{q\mu \Delta N(x)}$$

since $Q_I'(x) = -\frac{q\Delta N(x)}{W\Delta x}$. A fluctuation of $\delta\Delta N$ in the inversion charge at x will result in a fluctuation of the resistance of this segment according to the derivative:

$$\frac{\partial \Delta R(x)}{\partial \Delta N} = -\frac{\Delta x^2}{q\mu(\Delta N(x))^2}$$

so that

$$\delta\Delta R(x) = -\frac{\Delta x^2}{q\mu(\Delta N(x))^2}\delta\Delta N$$

If the fluctuation in the resistance of Δx is small relative to its average value ($\delta\Delta R << \Delta R$ or, equivalently, $\delta\Delta N << \Delta N$), then the short-circuit current fluctuation $\delta\Delta I_D$ through the segment resulting from resistance fluctuation may be found by linearizing Ohm's law for small deviations from the average:

$$I_D + \delta\Delta I_D = \frac{\Delta V}{\Delta R + \delta\Delta R}$$

$$(I_D + \delta\Delta I_D)(\Delta R + \delta\Delta R) = \Delta V$$

$$I_D\Delta R + \delta\Delta I_D\Delta R + I_D\delta\Delta R + \underbrace{\delta\Delta I_D\delta\Delta R}_{neglect} = \Delta V$$

Subtracting $I_D\Delta R = \Delta V$, neglecting the second-order term to make the linear approximation, and re-arranging, yields

$$\delta\Delta I_D = -I_D\frac{\delta\Delta R}{\Delta R}$$

Substituting for the resistance terms gives the short-circuit current fluctuation of the section in terms of the change in the number of mobile carriers:

$$\delta\Delta I_D = -I_D\left(\frac{-\frac{\Delta x^2}{q\mu(\Delta N(x))^2}\delta\Delta N}{\frac{\Delta x^2}{q\mu\Delta N(x)}}\right) = I_D\frac{\delta\Delta N}{\Delta N(x)}$$

This can be used to relate the spectral intensity of the current fluctuations through Δx to the spectral intensity of its carrier fluctuations:

$$S_{\Delta I_D}(x,f) = \frac{I_D{}^2}{(\Delta N(x))^2}S_{\Delta N}(x,f)$$

According to Van der Ziel's sectioning method, the noise function in the channel is

$$F(x,f) = \Delta x S_{\Delta I_D}(x,f)$$

$$= \Delta x\frac{I_D{}^2}{(\Delta N(x))^2}S_{\Delta N}(x,f)$$

Replacing $S_{\Delta N}(x,f)$ with the previously derived expression, and $\Delta N(x)$ by its inversion charge representation gives the function that will be integrated over the channel length:

$$F(x,f) = \Delta x\frac{I_D{}^2}{\left(-\frac{W\Delta x Q_I'(x)}{q}\right)^2}\left[\left(\frac{\delta\Delta N}{\delta\Delta N_t}\right)^2\frac{N_T(E_f)_{eff}kT}{\gamma_t f}W\Delta x\right]$$

$$= \frac{q^2 I_D{}^2 N_T(E_f)_{eff}kT}{W\gamma_t f(Q_I'(x))^2}\left(\frac{\delta\Delta N}{\delta\Delta N_t}\right)^2$$

Finally, integration yields the total drain current spectral intensity:

$$S_{I_D}(f)=\frac{1}{L^2}\int_0^L\left[\frac{q^2 I_D{}^2 N_T(E_f)_{eff}kT}{W\gamma_t f(Q_I'(x))^2}\left(\frac{\delta\Delta N}{\delta\Delta N_t}\right)^2\right]dx$$

The ratio of inversion charge to trapped charge fluctuation was previously found to be

$$\frac{\delta\Delta N}{\delta\Delta N_t}=\frac{-C_i'}{C_{ox}'+C_{ss}'+C_d'+C_i'}$$

and the small-signal capacitance of the inversion layer is

$$C_i'=-\frac{Q_I'(x)}{\left(\frac{kT}{q}\right)}$$

so the integral becomes

$$S_{I_D}(f)=\frac{1}{L^2}\int_0^L\left[\frac{q^2 I_D{}^2 N_T(E_f)_{eff}kT}{W\gamma_t f(Q_I'(x))^2}\left(\frac{Q_I'(x)/\left(\frac{kT}{q}\right)}{C_{ox}'+C_{ss}'+C_d'-Q_I'(x)/\left(\frac{kT}{q}\right)}\right)^2\right]dx$$

$$=\frac{1}{L^2}\int_0^L\left[\frac{q^2 I_D{}^2 N_T(E_f)_{eff}kT}{W\gamma_t f\left(\left(\frac{kT}{q}\right)(C_{ox}'+C_{ss}'+C_d')-Q_I'(x)\right)^2}\right]dx$$

$$=\frac{q^2 I_D{}^2 kT}{WL^2\gamma_t f}\int_0^L\left[\frac{N_T(E_f)_{eff}}{\left(\left(\frac{kT}{q}\right)(C_{ox}'+C_{ss}'+C_d')-Q_I'(x)\right)^2}\right]dx$$

At this point, the functional dependence of the effective trap density $N_T(E_f)_{eff}$ on position (or at least surface potential) in the channel must be determined, since it is generally regarded to be a variable quantity within the semiconductor bandgap [Das and Moore 248]. If the channel is assumed to be uniform so that $N_T(E_f)_{eff}$ depends only on surface potential, it can be determined empirically by measuring the current noise when the drain-source bias is small compared to the gate bias. In this case the inversion charge density is simply a weak linear function of x and the effective trap density is almost constant in the channel, so the integration above becomes quite simple and $N_T(E_f)_{eff}$ can be computed from the measured drain current noise. Once $N_T(E_f)_{eff}$ is known as a function of surface potential, numerical integration may be used to solve for the drain current spectral intensity for arbitrary drain-source bias. Van der Ziel suggests a procedure for accomplishing this [Van der Ziel '86 170]. However, if $N_T(E_f)_{eff}$ is treated as a constant, as done by Reimbold [1191], the term may be taken out of the integral:

$$S_{I_D}(f)=\frac{q^2 I_D{}^2 N_T(E_f)_{eff}kT}{WL^2\gamma_t f}\int_0^L\left[\frac{1}{\left(\left(\frac{kT}{q}\right)(C_{ox}'+C_{ss}'+C_d')-Q_I'(x)\right)^2}\right]dx$$

For many devices this seems to be a reasonably accurate approximation, and it permits the integral to be solved in closed form for the weak and strong inversion cases, as shown below.

5.1 1/f Noise in Weak Inversion

For the case of weak inversion operation, the small-signal capacitance of the inversion layer is negligible compared to the oxide and depletion layer capacitances,

$$Q_I'(x)/\left(\tfrac{kT}{q}\right) << \left(C_{ox}' + C_{ss}' + C_d'\right)$$

so that

$$Q_I'(x) << \left(\tfrac{kT}{q}\right)\left(C_{ox}' + C_{ss}' + C_d'\right)$$

The drain current spectral intensity of $1/f$ noise in weak inversion becomes

$$S_{I_D}(f) \approx \frac{q^2 I_D{}^2 N_T\left(E_f\right)_{eff} kT}{WL^2 \gamma_t f} \int_0^L \left[\frac{1}{\left(\left(\tfrac{kT}{q}\right)\left(C_{ox}' + C_{ss}' + C_d'\right)\right)^2}\right] dx$$

$$= \frac{q^2 I_D{}^2 N_T\left(E_f\right)_{eff} kT}{WL^2 \gamma_t f} \left[\frac{L}{\left(\left(\tfrac{kT}{q}\right)\left(C_{ox}' + C_{ss}' + C_d'\right)\right)^2}\right]$$

$$= \frac{q^4 I_D{}^2 N_T\left(E_f\right)_{eff}}{kTWL\left(C_{ox}' + C_{ss}' + C_d'\right)^2 \gamma_t f}$$

which is proportional to the square of drain current and inversely proportional to gate area and the square of the dominant capacitances. The gate-referred voltage spectral intensity is valuable for design, so $S_{I_D}(f)$ is divided by the square of transconductance to obtain its value:

$$S_{V_{GS}}(f) = \frac{S_{I_D}(f)}{g_m{}^2} \approx \frac{\dfrac{q^4 I_D{}^2 N_T\left(E_f\right)_{eff}}{kTWL\left(C_{ox}' + C_{ss}' + C_d'\right)^2 \gamma_t f}}{\left(\dfrac{I_D}{n\phi_t}\right)^2}$$

$$\approx \frac{n^2 q^2 N_T\left(E_f\right)_{eff} kT}{WL\left(C_{ox}' + C_{ss}' + C_d'\right)^2 \gamma_t f}$$

Recalling that the ideality factor n is

$$n = \left.\frac{C_{ox}' + C_{ss}' + C_d'}{C_{ox}'}\right|_{\psi_s = 1.5\phi_F}$$

(where C_{ss}' was neglected in the derivation of subthreshold operation), the noise formula reduces to:

$$S_{V_{GS}}(f) \approx \frac{q^2 N_T(E_f)_{eff} kT}{WLC_{ox}'^2 \gamma_t f}$$

which is inversely proportional to the channel area and the square of oxide capacitance.

5.2 1/f Noise in Strong Inversion

The strong inversion case is more complicated. In the strongly inverted channel, the inversion layer capacitance dominates,

$$Q_I'(x)/\left(\tfrac{kT}{q}\right) >> (C_{ox}' + C_{ss}' + C_d')$$

which means that the change in inversion charge is almost equal to the change in trapped charge:

$$\frac{\delta\Delta N}{\delta\Delta N_t} = \frac{-C_i'}{C_{ox}' + C_{ss}' + C_d' + C_i'} \approx -1$$

Accordingly, the drain current spectral intensity integral:

$$S_{I_D}(f) = \frac{q^2 I_D^2 N_T(E_f)_{eff} kT}{WL^2 \gamma_t f} \int_0^L \left[\frac{1}{\left(\left(\tfrac{kT}{q}\right)(C_{ox}' + C_{ss}' + C_d') - Q_I'(x)\right)^2} \right] dx$$

can be approximated by

$$S_{I_D}(f) \approx \frac{q^2 I_D^2 N_T(E_f)_{eff} kT}{WL^2 \gamma_t f} \int_0^L \left[\frac{1}{(-Q_I'(x))^2} \right] dx$$

but this is *only valid for linear operation*, where the whole channel is strongly inverted. When the transistor is saturated, this approximation fails and the integral becomes infinite. This is because the drain end of the channel is weakly inverted when the transistor is saturated ("pinched off"), so the approximation that $Q_I'(x)/\left(\tfrac{kT}{q}\right) >> (C_{ox}' + C_{ss}' + C_d')$ is not true there.

A general expression for the drain current spectral intensity, valid for the linear and saturation regions of strong inversion operation, requires that the integration be performed without this approximation. Replacing dx and $Q_I'(x)$ by their equivalent functions of the effective reverse bias $V_{CB}(x)$ in the channel,

$$dx = -\frac{\mu W Q_I'(x)}{I_D} dV_{CB}$$

$$Q_I'(x) = Q_I'(V_{CB}(x))$$

as done previously for the calculation of white noise drain current spectral intensity in strong inversion, allows the integral for the $1/f$ component of drain current noise to be expressed as a function of the effective reverse bias:

$$S_{I_D}(f)=\frac{q^2 {I_D}^2 N_T(E_f)_{eff} kT}{WL^2\gamma_t f}\int_{V_{CB}(0)}^{V_{CB}(L)}\left[\frac{1}{\left(\left(\frac{kT}{q}\right)\left(C'_{ox}+C'_{ss}+C'_d\right)-Q'_I(V_{CB})\right)^2}\left(-\frac{\mu W Q'_I(V_{CB})}{I_D}\right)\right]dV_{CB}$$

$$=-\frac{q^2\mu I_D N_T(E_f)_{eff} kT}{L^2\gamma_t f}\int_{V_{SB}}^{V_{DB}}\left[\frac{Q'_I(V_{CB})}{\left(\left(\frac{kT}{q}\right)\left(C'_{ox}+C'_{ss}+C'_d\right)-Q'_I(V_{CB})\right)^2}\right]dV_{CB}$$

where the mobility has been assumed to be constant along the channel. This can be simplified further by converting to an integral over Q'_I. Recalling that in strong inversion

$$Q'_I(V_{CB})=-C'_{ox}\left[V_{GS}-V_T-(1+\delta)(V_{CB}-V_{SB})\right]$$

the derivative can be taken to find

$$dV_{CB}=\frac{dQ'_I(V_{CB})}{C'_{ox}(1+\delta)}=\frac{1}{C'_{ox}(1+\delta)}dQ'_I$$

Upon substitution the integral becomes

$$S_{I_D}(f)=-\frac{q^2\mu I_D N_T(E_f)_{eff} kT}{L^2\gamma_t f}\int_{Q'_I(V_{SB})}^{Q'_I(V_{DB})}\left[\frac{Q'_I}{\left(\left(\frac{kT}{q}\right)\left(C'_{ox}+C'_{ss}+C'_d\right)-Q'_I\right)^2}\left(\frac{1}{C'_{ox}(1+\delta)}\right)\right]dQ'$$

$$=-\frac{q^2\mu I_D N_T(E_f)_{eff} kT}{L^2 C'_{ox}(1+\delta)\gamma_t f}\int_{Q'_I(V_{SB})}^{Q'_I(V_{DB})}\left[\frac{Q'_I}{\left(\left(\frac{kT}{q}\right)\left(C'_{ox}+C'_{ss}+C'_d\right)-Q'_I\right)^2}\right]dQ'$$

Integration by parts gives the result:

$$S_{I_D}(f)=-\frac{q^2\mu I_D N_T(E_f)_{eff} kT}{L^2 C'_{ox}(1+\delta)\gamma_t f}\left[\frac{Q'_I}{\left(\left(\frac{kT}{q}\right)\left(C'_{ox}+C'_{ss}+C'_d\right)-Q'_I\right)}+\ln\left(\left(\frac{kT}{q}\right)\left(C'_{ox}+C'_{ss}+C'_d\right)-Q'_I\right)\right]_{Q'_I(V_{SB})}^{Q'_I(V_{DB})}$$

where

$$Q'_I(V_{DB})=-C'_{ox}\left[V_{GS}-V_T-(1+\delta)(V_{DB}-V_{SB})\right]$$

and

$$Q'_I(V_{SB})=-C'_{ox}\left[V_{GS}-V_T\right]$$

Using the drain-source pinch-off voltage V'_{DS} and the parameter α to describe the linear and saturation regions as before,

$$V'_{DS}=\frac{V_{GS}-V_T}{1+\delta}$$

$$\alpha=\begin{cases}1-\dfrac{V_{DS}}{V'_{DS}}, & V_{DS}\le V'_{DS}\\ 0, & V_{DS}>V'_{DS}\end{cases}$$

the limits of integration may be written as

$$Q_I'(V_{DB}) = -C_{ox}'\left[V_{GS} - V_T - (1+\delta)((1-\alpha)V_{DS}')\right]$$
$$= -C_{ox}'(1+\delta)\left[V_{DS}' - (1-\alpha)V_{DS}'\right]$$
$$= -C_{ox}'(1+\delta)\alpha V_{DS}'$$

and

$$Q_I'(V_{SB}) = -C_{ox}'(1+\delta)\left[\frac{V_{GS} - V_T}{(1+\delta)}\right]$$
$$= -C_{ox}'(1+\delta)V_{DS}'$$

Solving with these limits of integration gives the final result:

$$S_{I_D}(f) = -\frac{q^2\mu I_D N_T(E_f)_{eff} kT}{L^2 C_{ox}'(1+\delta)\gamma_t f}\left[\begin{array}{l} \dfrac{-C_{ox}'(1+\delta)\alpha V_{DS}'}{\left(\left(\frac{kT}{q}\right)\left(C_{ox}' + C_{ss}' + C_d'\right) + C_{ox}'(1+\delta)\alpha V_{DS}'\right)} \\ -\dfrac{-C_{ox}'(1+\delta)V_{DSI}'}{\left(\left(\frac{kT}{q}\right)\left(C_{ox}' + C_{ss}' + C_d'\right) + C_{ox}'(1+\delta)V_{DS}'\right)} \\ +\ln\left(\left(\frac{kT}{q}\right)\left(C_{ox}' + C_{ss}' + C_d'\right) + C_{ox}'(1+\delta)\alpha V_{DS}'\right) \\ -\ln\left(\left(\frac{kT}{q}\right)\left(C_{ox}' + C_{ss}' + C_d'\right) + C_{ox}'(1+\delta)V_{DS}'\right) \end{array}\right]$$

$$= \frac{q^2\mu I_D N_T(E_f)_{eff} kT}{L^2 C_{ox}'(1+\delta)\gamma_t f}\left[\begin{array}{l} \dfrac{\alpha V_{DS}'}{\left(\left(\frac{kT}{q}\right)\left(\frac{C_{ox}'+C_{ss}'+C_d'}{C_{ox}'(1+\delta)}\right) + \alpha V_{DS}'\right)} - \dfrac{V_{DS}'}{\left(\left(\frac{kT}{q}\right)\left(\frac{C_{ox}'+C_{ss}'+C_d'}{C_{ox}'(1+\delta)}\right) + V_{DS}'\right)} \\ +\ln\left(\dfrac{\left(\frac{kT}{q}\right)\left(\frac{C_{ox}'+C_{ss}'+C_d'}{C_{ox}'(1+\delta)}\right) + V_{DS}'}{\left(\frac{kT}{q}\right)\left(\frac{C_{ox}'+C_{ss}'+C_d'}{C_{ox}'(1+\delta)}\right) + \alpha V_{DS}'}\right) \end{array}\right]$$

This expression may be simplified for the separate cases of linear and saturation operation. In the linear region, far away from saturation, $0 << \alpha < 1$ and

$$V_{DS}' > \alpha V_{DS}' >> \left(\frac{kT}{q}\right)\left(\frac{C_{ox}'+C_{ss}'+C_d'}{C_{ox}'(1+\delta)}\right)$$

so that the spectral intensity of drain current is approximately

$$S_{I_D}(f) \approx \frac{q^2\mu I_D N_T(E_f)_{eff} kT}{L^2 C_{ox}'(1+\delta)\gamma_t f}\ln\left(\frac{1}{\alpha}\right)$$

Dividing by the transconductance squared gives the gate-referred voltage spectral intensity:

$$S_{V_{GS}}(f)=\frac{S_{I_D}(f)}{g_m{}^2}\approx\frac{\dfrac{q^2\mu I_D N_T(E_f)_{eff}kT}{L^2C'_{ox}(1+\delta)\gamma_t f}\ln\left(\dfrac{1}{\alpha}\right)}{2\dfrac{W}{L}\dfrac{\mu C'_{ox}}{1+\delta}\left(\dfrac{1-\alpha}{1+\alpha}\right)I_D}$$

$$\approx\frac{q^2N_T(E_f)_{eff}kT}{2WLC'^{2}_{ox}\gamma_t f}\left(\frac{1+\alpha}{1-\alpha}\right)\ln\left(\frac{1}{\alpha}\right)$$

In saturation, $\alpha=0$ and

$$V'_{DS}>>\left(\tfrac{kT}{q}\right)\left(\tfrac{C'_{ox}+C'_{ss}+C'_d}{C'_{ox}(1+\delta)}\right)$$

so the spectral intensity becomes

$$S_{I_D}(f)\approx\frac{q^2\mu I_D N_T(E_f)_{eff}kT}{L^2C'_{ox}(1+\delta)\gamma_t f}\left[0-1+\ln\left(\frac{V'_{DS}}{\left(\tfrac{kT}{q}\right)\left(\tfrac{C'_{ox}+C'_{ss}+C'_d}{C'_{ox}(1+\delta)}\right)+0}\right)\right]$$

$$=\frac{q^2\mu I_D N_T(E_f)_{eff}kT}{L^2C'_{ox}(1+\delta)\gamma_t f}\left[-\ln(e)+\ln\left(\frac{V'_{DS}}{\left(\tfrac{kT}{q}\right)\left(\tfrac{C'_{ox}+C'_{ss}+C'_d}{C'_{ox}(1+\delta)}\right)}\right)\right]$$

$$=\frac{q^2\mu I_D N_T(E_f)_{eff}kT}{L^2C'_{ox}(1+\delta)\gamma_t f}\ln\left(\frac{V'_{DS}}{\left(\tfrac{kT}{q}\right)\left(\tfrac{C'_{ox}+C'_{ss}+C'_d}{C'_{ox}(1+\delta)}\right)e}\right)$$

$$=\frac{q^2\mu I_D N_T(E_f)_{eff}kT}{L^2C'_{ox}(1+\delta)\gamma_t f}\left[\ln\left(\left(\frac{C'_{ox}}{C'_{ox}+C'_{ss}+C'_d}\right)\frac{(V_{GS}-V_T)}{\left(\tfrac{kT}{q}\right)}\right)-1\right]$$

Again, dividing by the square of transconductance gives the gate-referred voltage spectral intensity:

$$S_{V_{GS}}(f)\approx\frac{\dfrac{q^2\mu I_D N_T(E_f)_{eff}kT}{L^2C'_{ox}(1+\delta)\gamma_t f}\left[\ln\left(\left(\dfrac{C'_{ox}}{C'_{ox}+C'_{ss}+C'_d}\right)\dfrac{(V_{GS}-V_T)}{\left(\tfrac{kT}{q}\right)}\right)-1\right]}{2\dfrac{W}{L}\dfrac{\mu C'_{ox}}{1+\delta}I_D}$$

$$=\frac{q^2N_T(E_f)_{eff}kT}{2WLC'^{2}_{ox}\gamma_t f}\left[\ln\left(\left(\frac{C'_{ox}}{C'_{ox}+C'_{ss}+C'_d}\right)\frac{(V_{GS}-V_T)}{\left(\tfrac{kT}{q}\right)}\right)-1\right]$$

Thus in strong inversion the $1/f$ component of the drain current spectral intensity is directly proportional to drain current and inversely proportional to the oxide capacitance per unit area and the square of channel length. It also has a relatively weak (logarithmic) dependence on α (which is a function of V_{DS} and V_{GS}) in the linear region, and on $V_{GS}-V_T$ in saturation. Referred to the gate, the noise is inversely proportional to the gate area and the square of the oxide capacitance per unit area, and has the same log terms as the current noise in

the linear and saturation regions. Comparing the gate-referred noise equations for weak and strong inversion operation, striking similarities are evident. The basic dependencies are the same for all cases, with only a slowly changing or constant multiplier distinguishing them. The important thing to remember for design purposes is the inverse dependence on gate area and oxide capacitance per unit area.

The preceding derivation of $1/f$ noise, based on Van Der Ziel's and Reimbold's work, is appealing in its physical basis, especially in light of the derivation presented earlier to describe the operation of the MOSFET. However, there is no conclusive evidence that this model of flicker noise is correct, particularly with regards to the assumption of constant effective trap density, and several alternative theories and their variations have been presented [Sah; Christensson, Lundstrom, and Svensson; Hsu; Berz; Klaassen; Duh and Van der Ziel; Kornfeld]. Das and Moore have compared many of these, and show how the basic assumptions behind any theory affects the interpretation of experimental results. In particular, assumptions made regarding the physical distribution and bias dependence of the trap density will lead to very different geometry and bias dependence predictions from the various noise theories. By the same token, these assumptions can be used to support quite disparate theories for mechanisms of drain current or surface potential modulation, given any set of experimental results. For this reason it is probably worthwhile to empirically quantify the $1/f$ noise, including its geometry and bias dependencies, for the particular fabrication process that is to be used for low noise circuit design.

Chapter 6

Experimental Methods

Measurements of the $I_D - V_{GS}$ characteristics of the test transistors were made using a Keithley semiconductor characterization instrument. Since MOSFETs are typically operated in saturation for analog circuits, this region was the focus of the empirical study. The drain-source bias was five volts for all of the $I_D - V_{GS}$ sweeps, which was sufficient to saturate the transistors for the gate bias sweep used. Drain current was recorded as the gate-source voltage was stepped from zero to four volts in 10mV increments. Transconductance was determined as a function of drain current by smoothing the $I_D - V_{GS}$ data and finding its slope.

To empirically quantify $1/f$ noise and to test the basic predictions of the presented noise theory, a computer-based noise measurement system was built. It allowed a transistor to be biased to a desired operating condition so that the spectral intensity of its drain current could be determined at this bias by the technique of periodogram averaging (which is described further on). The key elements of the bias and noise amplifier circuitry are illustrated in the block diagram of Figure 7. A data acquisition board provided digital-to-analog outputs for the control signals, and analog-to-digital conversion for the various system outputs.

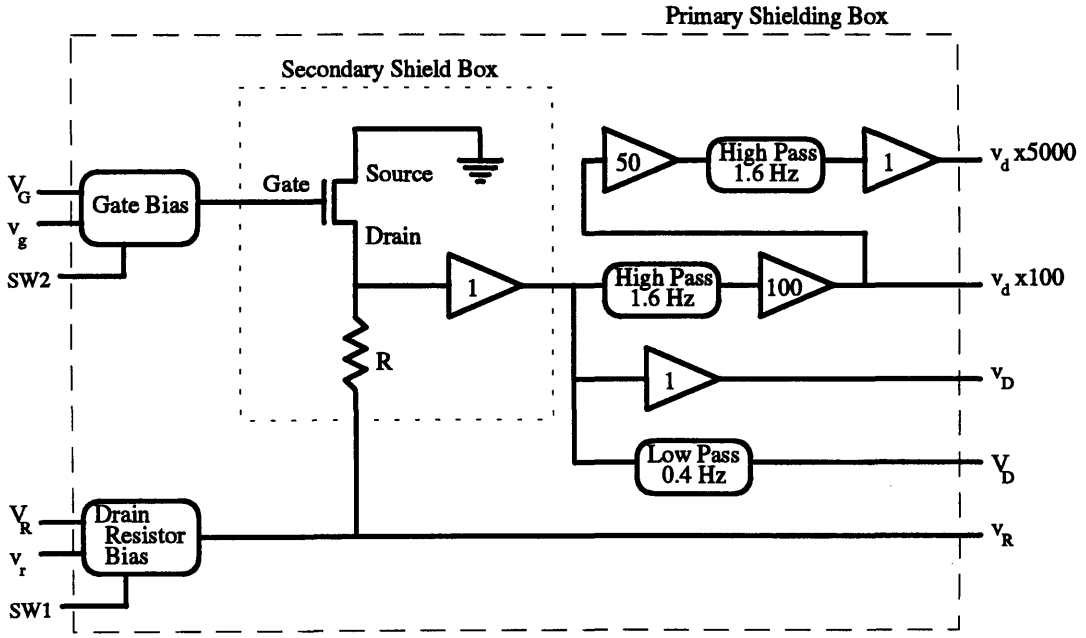

Figure 7: Computer-based noise measurement system

An operational amplifier with ultra-low leakage inputs, the Analog Devices AD549, was used for the drain voltage buffer because of its correspondingly low input noise current. All other biasing and amplifying circuits were implemented using low voltage-noise Analog Devices OP-27 op-amps. System inputs V_G and V_R were used to set the DC gate and drain bias voltages, so they were low-pass filtered to minimize their noise contributions. AC signals could be applied to the gate and drain (via the load resistor) for transconductance and impedance measurements by the use of inputs v_g and v_r, which are physically connected to the transistor by switch controls SW2 and SW1 when in use. Initially, a transimpedance topology was used to convert the drain current to a voltage for amplification and analog-to-digital conversion, but stability problems made this approach unsuitable. Since the stability of the operational amplifier current-to-voltage converter circuit depended on the source impedance, it had significant gain peaking when the source impedance was mostly capacitive. This was the case for the MOSFET drain when the transistor was saturated, and especially for weak inversion operation where the real part of the impedance was very high. While this could have been compensated, the uncertainty in the transistor output impedance made such an approach difficult without severely limiting the circuit's bandwidth. Unpredictable gain peaking occurred when different transistors were tested and when the bias conditions changed the drain-to-source impedance. A passive, purely resistive load made the system dynamics much easier to take into account.

Each noise measurement began with the setting of bias conditions. The control program allowed either the gate and load resistor voltages to be set directly, or the drain voltage and current to be selected. In the latter case, the load resistor voltage was set so that the drain voltage would be correct at the desired current, and a bisection algorithm was used to iteratively set the gate voltage until the desired drain current was achieved. Once the transistor bias had settled to its steady-state value, the impedance of the drain node was determined in order to correct the measured noise data for the frequency response of the passive current-to-voltage converter. The signal source (everything connected to the drain) was modeled as a resistor and capacitor in parallel, as shown in Figure 8.

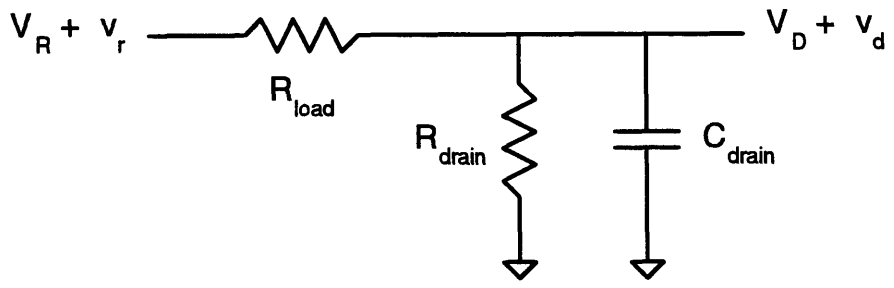

Figure 8: Model of noise amplifier input

The impedance at this node was determined each time a noise measurement was made by AC-coupling a small voltage sinewave to the load resistor (via v_r) and sampling the voltage there and at the drain. Through Fourier analysis, the equivalent resistance and capacitance of the drain node was obtained. The procedure was as follows:

1. Apply a low amplitude 10 kHz sinewave, v_r, at the maximum rate of D/A output. This consists of 34 samples per cycle at 3 μs each.
2. Sample the input (v_r) and output (v_d) sinusoids synchronously with the D/A output, and at the same rate (3 μs).
3. Average 10,000 cycles of each sampled waveform point-by-point.
4. Get the Fourier series coefficient for the input and output sinusoids as follows, where $N_{\sin}$ is the period (34 points, in this case) and $x_n = \{x_0, x_1, \ldots, x_{N-1}\}$ are the points in the averaged cycle:

$$\text{Real Part} = \sum_{n=0}^{N_{\sin}-1} x_n \cos\left(\frac{2\pi n}{N_{\sin}}\right)$$

$$\text{Imaginary Part} = -\sum_{n=0}^{N_{\sin}-1} x_n \sin\left(\frac{2\pi n}{N_{\sin}}\right)$$

5. Compute the real and imaginary parts of V_r/V_d:

$$\text{Re}\left\{\frac{V_r}{V_d}\right\} = \frac{\text{Re}\{V_r\}\text{Re}\{V_d\} + \text{Im}\{V_r\}\text{Im}\{V_d\}}{\text{Re}\{V_d\}^2 + \text{Im}\{V_d\}^2}$$

$$\text{Im}\left\{\frac{V_r}{V_d}\right\} = \frac{\text{Im}\{V_r\}\text{Re}\{V_d\} + \text{Re}\{V_r\}\text{Im}\{V_d\}}{\text{Re}\{V_d\}^2 + \text{Im}\{V_d\}^2}$$

6. Determine R_{drain} and C_{drain}, where T is the sampling period (3 μs in this case):

$$R_{drain} = \frac{R}{\text{Re}\left\{\frac{V_r}{V_d}\right\} - 1}$$

$$C_{drain} = \frac{\text{Im}\left\{\frac{V_r}{V_d}\right\}}{\omega_{test} R}, \qquad \omega_{test} = \frac{2\pi}{NT}$$

7. Calculate the admittance of the signal source at the input of the amplifier (everything at the drain node), $Y_{tot}(\omega)$, which will be necessary to determine the drain current from the measured voltage data:

$$Y_{tot}(\omega) = \frac{1}{R_{tot}} + j\omega C_{drain}, \qquad R_{tot} = R_{drain} \| R = \frac{R}{\text{Re}\left\{\frac{V_r}{V_d}\right\}}$$

$$\|Y_{tot}(\omega)\| = \sqrt{\frac{1}{R_{tot}^{\,2}} + \omega^2 C_{drain}^{\,2}}$$

The next step in the measurement of the drain current noise was to sample the amplified drain voltage, and to use a *Fast Fourier Transform* (FFT) algorithm and periodogram averaging to estimate its spectral intensity. The steps in this procedure were:

1. Sampling at a rate 1/T that satisfies the Nyquist criteria, collect a block of data whose length N is an integer power of 2 so that the FFT may be used most efficiently.
2. Subtract the mean of the block of data from each of its values in order to minimize "leakage" from the zero frequency component of the spectrum when the FFT is performed.
3. Apply a windowing function to the data. This will decrease the influence that neighboring frequency bins in the FFT spectrum have on each other (the "leakage" between them).
4. Perform the FFT on the data block to determine the power spectrum, or *periodogram*, of the drain voltage noise.
5. Divide each bin in the resulting frequency spectrum by the width of the frequency bin, which is 1/(4NT) in units of Hertz. This gives the spectral intensity.
6. Repeat steps 1-4 many times, averaging the periodograms from all of the trials on a point-by-point basis to get an estimate of the spectral intensity of the drain voltage. Averaging is required in order for the estimate to converge to the correct value (the mean value of the power in each frequency bin).
7. Divide the averaged power spectrum by the mean-square value of the windowing function that was applied earlier, in order to correct for the amplitude scaling it introduced.
8. Find the amplifier gain A(f) at the frequency of each bin in the frequency spectrum (using a logarithmic interpolation between points in a table of measured gain versus frequency data), and divide the bin by the square of this gain. This will correct the power spectrum for the frequency response of the amplifier.

At this point the spectral intensity of the total equivalent voltage noise at the input to the amplifier has been estimated. This includes contributions from the amplifier itself, the load resistor, and the drain current fluctuations of interest. To extract the spectral intensity of the drain current, the noise contributions of the extraneous sources were subtracted from the spectrum, and then the drain voltage spectral intensity was converted to a drain current spectral intensity. The circuit of Figure 9 was used to model the lumped noise sources at the amplifier input.

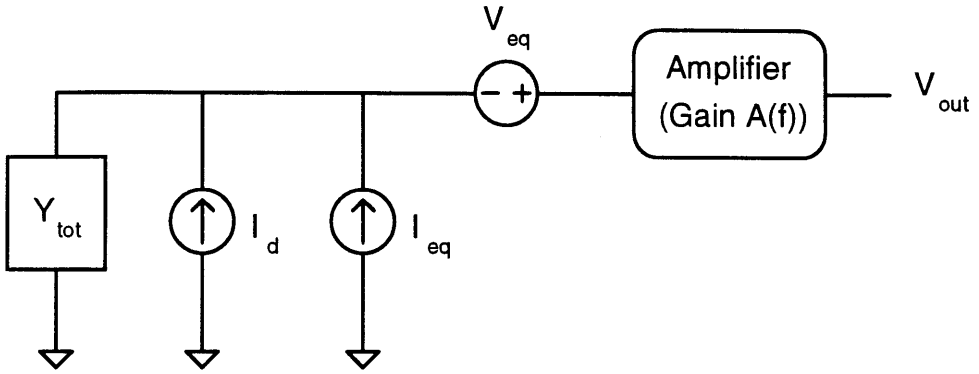

Figure 9: Noise modeled by sources at input of noiseless amplifier:
I_{eq} = equivalent input noise current of amplifier and load resistor
V_{eq} = equivalent input noise voltage of amplifier
I_d = drain current noise

Computing each noise source's contribution to the output voltage and adding their spectral intensities gave the total output voltage spectral intensity of this model:

$$S_{V_{out}}(f) = \left(\frac{S_{I_d}(f) + S_{I_{eq}}(f)}{\|Y_{tot}\|^2} + S_{V_{eq}}(f) \right) (A(f))^2$$

Rearranging terms allowed the transistor's drain current spectral intensity to be determined:

$$S_{I_d}(f) = \left(\frac{S_{V_{out}}(f)}{(A(f))^2} - S_{V_{eq}}(f) \right) \|Y_{tot}\|^2 - S_{I_{eq}}(f)$$

Thus, prior to making transistor noise measurements, it was first necessary to determine $S_{V_{eq}}(f)$, the equivalent input voltage noise of the amplifier, and $S_{I_{eq}}(f)$, the equivalent input current noise of the amplifier and load resistor. $S_{V_{eq}}(f)$ was found by shorting the input of the amplifier to ground and measuring the noise as outlined above. To calculate $S_{I_{eq}}(f)$, the input to the amplifier was opened so that the only thing connected was the load resistor. The impedance measurement technique described earlier was used to find the source resistance (which should simply be the load resistor) and the stray capacitance at the input. The noise in this configuration was measured, and $S_{V_{eq}}(f)$ was subtracted from its spectral intensity point-by-point. Multiplying each point in the resulting frequency spectrum by the square of the admittance of the source at that frequency yielded $S_{I_{eq}}(f)$. Once these spectra were found, they were saved for use in subsequent transistor noise measurements. Finally, the process used to extract the drain current spectral intensity $S_{I_d}(f)$ was:

1. Measure the noise with the transistor biased as desired.
2. Subtract the spectral intensity of V_{eq} from the total noise (which has already been corrected for the gain of the amplifier, A(f)).
3. Multiply each frequency point in the result by the square of the admittance $|Y_{tot}(2\pi f)|^2$ that was determined earlier by the impedance measurement made with the transistor in place (and biased for the noise measurement).
4. Finally, subtract $S_{I_{eq}}(f)$ to obtain $S_{I_d}(f)$.

Chapter 7

Experimental Results and Discussion

7.1 Transistor Operating Characteristics

A set of n-channel and p-channel transistors fabricated with the "Orbit 2μm Analog" process available through the MOS Implementation Service (MOSIS) was chosen for characterization. The geometries of these transistors are given in Table 1.

Table 1: Transistors used for operating characteristic measurements

W [μm]	L [μm]	W/L	W•L [μm^2]
4	2	2	8
6	2	3	12
8	2	4	16
16	2	8	32
32	2	16	64
100	2	50	200
6	3	2	18
100	3	33.3	300
100	5	20	500
4	8	0.5	32
6	8	0.75	48
8	8	1	64
16	8	2	128
32	8	4	256
100	8	12.5	800
4	16	0.25	64
6	16	0.375	96
8	16	0.5	128
16	16	1	256
32	16	2	512
100	16	6.25	1600
4	32	0.125	128
6	32	0.1875	192
8	32	0.25	256
16	32	0.5	512
32	32	1	1024
100	32	3.125	3200

The transistor $I_D - V_{GS}$ measurements were mostly in agreement with the theoretical predictions regarding weak and strong inversion operation. All measurements were taken with the transistors in saturation, since this was of primary interest for analog circuit design. Figure 10 shows the $I_D - V_{GS}$ relationships for typical NMOS and PMOS transistors on the test chips.

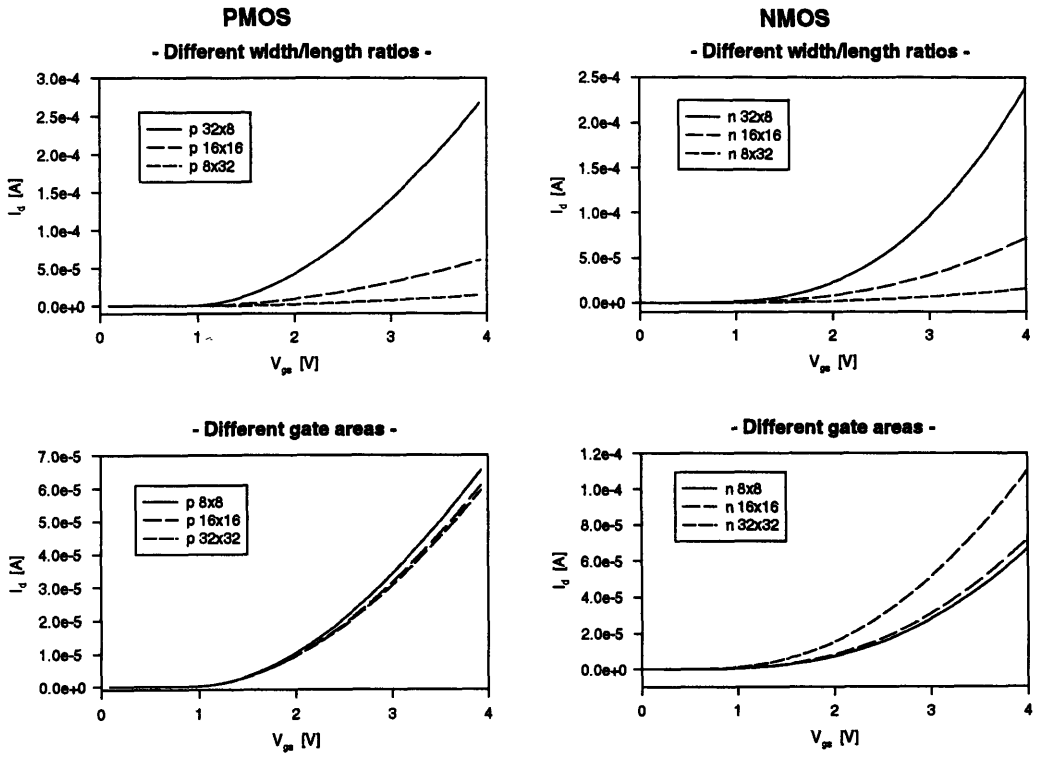

Figure 10: $I_D - V_{GS}$ curves for test transistors

These plots show the affect of transistor geometry on $I_D - V_{GS}$ characteristics. As channel width/length ratio is varied for constant gate area, the curves shift dramatically (above threshold) as seen on the top two plots. As gate area is varied with constant width/length ratio, there should be no change in the curves, as verified by the lower-left plot of the PMOS devices. However, the NMOS transistors used for the lower-right plot showed some unexpected variation. The value of empirical studies of any specific process to be used for fabricating circuits is that discrepancies such as this may be found and possibly corrected for.

The important features of the $I_D - V_{GS}$ relationships, the weak and strong inversion characteristics, are more readily apparent when the drain current axis is plotted on a log scale (for subthreshold behavior), and after taking the square root of the current values (to see strong inversion behavior). These plots are given in the upper and lower half of Figure 11, respectively.

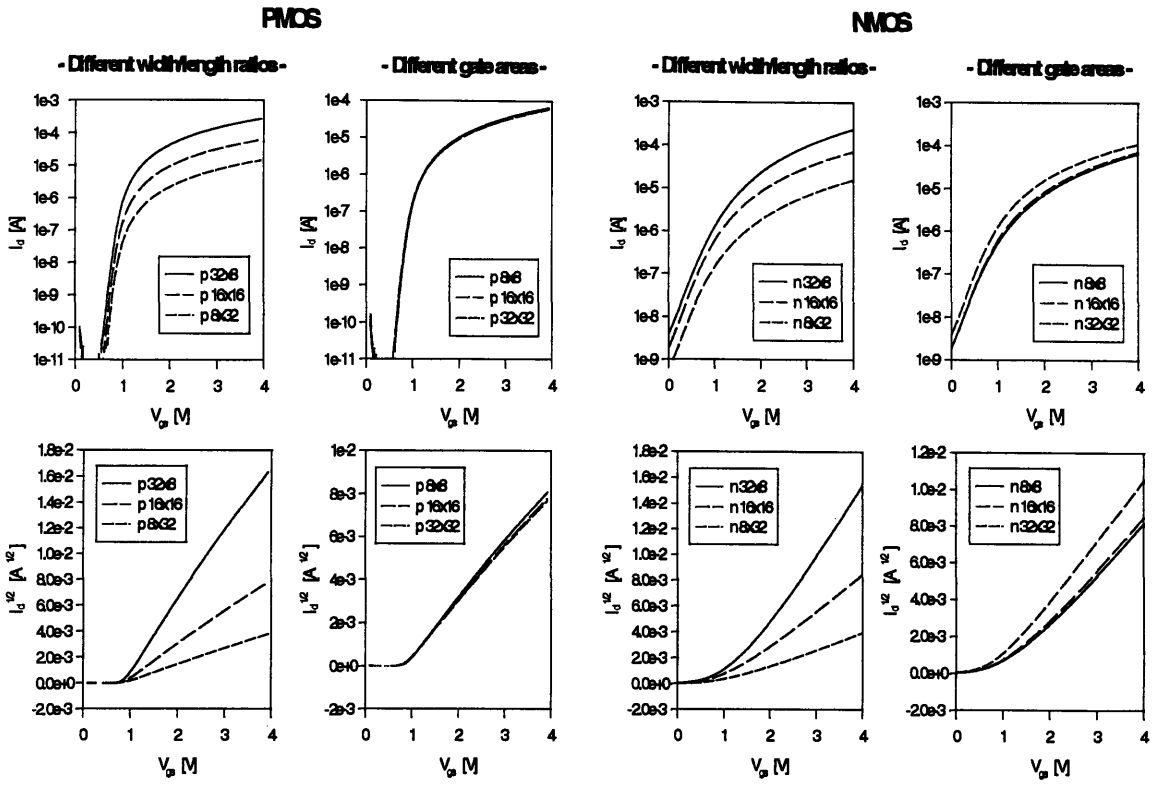

Figure 11: Weak and strong inversion $I_D - V_{GS}$ characteristics

The subthreshold region lies to the left of the "knee" in the curves on the semilog plots. Note that the slopes of the $I_D - V_{GS}$ curves are constant in subthreshold, and equal for all devices of the same type (n or p). Since the slope of the $\ln(I_D)$ vs V_{GS} curve equals the transconductance-to-current ratio, these plots show that this ratio is constant and maximum in subthreshold, and independent of device geometry. Dividing q/kT by this slope gave the weak inversion ideality factor, n. It was found to be approximately 1.5 for the p-channel transistors and about 6 for the n-channel transistors. The strong inversion behavior is seen in the plots of $\sqrt{I_D}$ vs V_{GS} as the curves slope upward above the threshold voltage. The square-law relationship between drain current and gate-source voltage is evident in the fairly linear curves above threshold, as is the effect of width/length ratio on the transconductance. The non-ideality of the n-channel transistors is probably an artifact of the manufacturing process rather than a general feature of n-type devices. The fact that in this process the NMOS transistors are fabricated in the substrate while the PMOS transistors are fabricated in an n-type well may account for the difference in quality.

To more clearly show how the transconductance-to-current ratio decreases when the drain current is increased beyond threshold, the ratio was calculated from measured data for three transistors, and is shown in Figure 12.

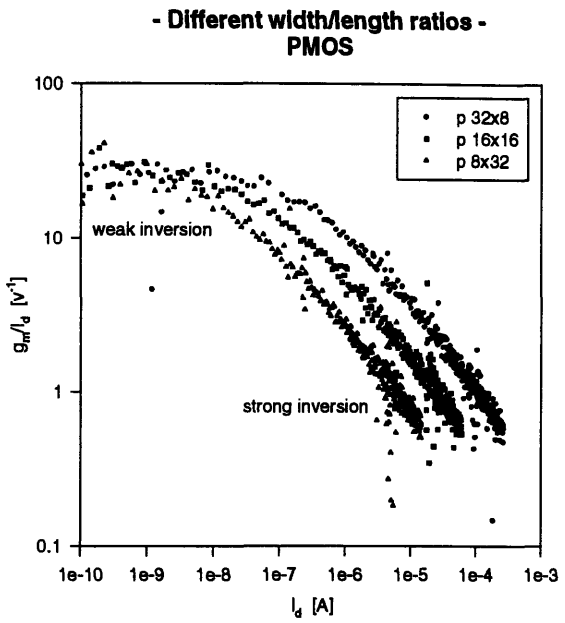

Figure 12: Transconductance-to-current ratio vs drain current

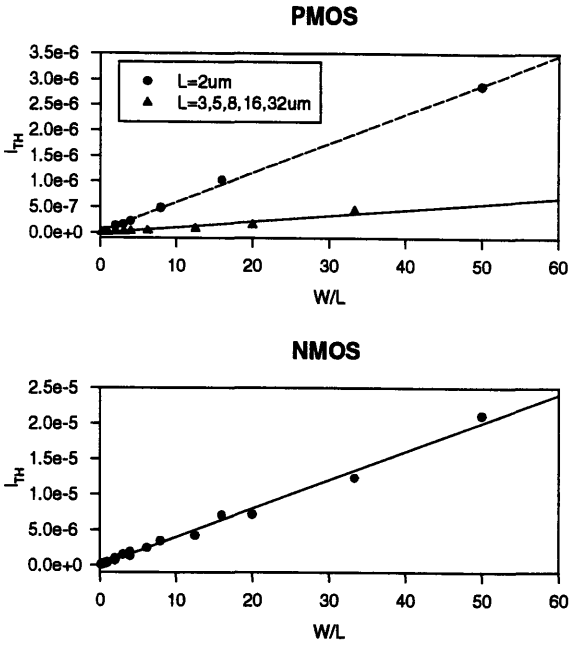

Figure 13: Threshold current vs width-to-length ratio

The proportionality between threshold current and the width/length ratio is shown in Figure 13 for several transistors. The plotted $2\phi_F$ thresholds were calculated from the measured $I_D - V_{GS}$ data using the method presented by Deen and Yan [503-511]. The p-channel transistor thresholds follow two different lines -- in this process, the minimum length transistors (L = 2 μm) evidently have a different constant of proportionality than the other devices, but each group is fairly linear.

7.2 Noise Measurement System

The performance of the noise measurement system was partially assessed by baseline noise characteristics. The equivalent input current noise component was a function of the load resistor, so this must be kept in mind when interpreting the baseline noise of the system. In order to keep the bandwidth of the system large, a 1 MΩ load resistor was used for all of the measurements made thus far. A sampling period of 3 μs was used for all of the noise measurements, and a block size of 2048 was used. This produced a 1024 point power spectrum that extended up to 167 kHz. The fast sampling rate was chosen to observe high frequency noise behavior and to minimize aliasing. Averaging of 1000 periodograms was performed for each measurement. Spectral intensities of the equivalent input voltage and current noise of the system, $S_{V_{eq}}(f)$ and $S_{I_{eq}}(f)$, are plotted in Figure 14.

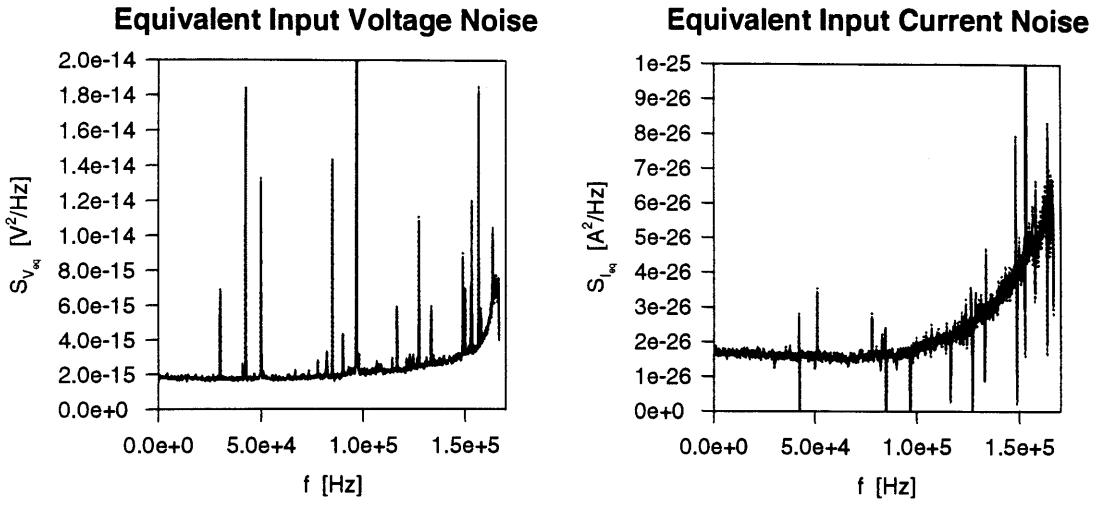

Figure 14: Baseline noise components of measurement system

The narrow spikes in the spectra are from "environmental" noise sources, such as the computer and monitor, switching power supplies, etc. While careful shielding attenuated these signals greatly, they were still present as seen in these plots. Baseline noise subtraction also tends to remove them from the transistor noise measurements, but the cancellation is far from perfect (this is the reason for the downward-pointing spikes in $S_{I_{eq}}(f)$, since its calculation includes a correction for $S_{V_{eq}}(f)$). The increasing noise for frequencies above 100kHz seen in these plots can be explained by aliasing. While the actual detected amplitude of signals was very small at these high frequencies, they were non-zero, and extended beyond 167 kHz. Since the spectrum has been corrected for all of the dynamics of the measurement system, the inevitable aliasing that must occur in a physically realizable system is seen. The slope of the rise corresponds to the fall-off of the frequency response of the system after 167 kHz, which could be made steeper by the use of an additional anti-aliasing filter. The increased "width" of the data at high frequencies reflects the greater uncertainty in this data, since it was calculated from smaller amplitude signals.

The equivalent input voltage noise was about 45 $\mathrm{nV}/\sqrt{\mathrm{Hz}}$, which was principally due to the AD549 input buffer op-amp. Equivalent input current noise was about 130 $\mathrm{fA}/\sqrt{\mathrm{Hz}}$, the current noise of the 1 MΩ load resistor. As a final check of the noise measurement electronics and the data processing algorithms, resistors were connected to the input. Since the thermal noise of resistors was well known, they provided a good test signal source. The spectral intensity of the current noise of a 1 MΩ metal-film resistor is shown in Figure 15. Dotted lines (barely distinguishable) showing the standard error of the mean are plotted

along with the periodogram average, and the theoretical noise of the resistor is the horizontal line at 1.66×10^{-26} A^2/Hz.

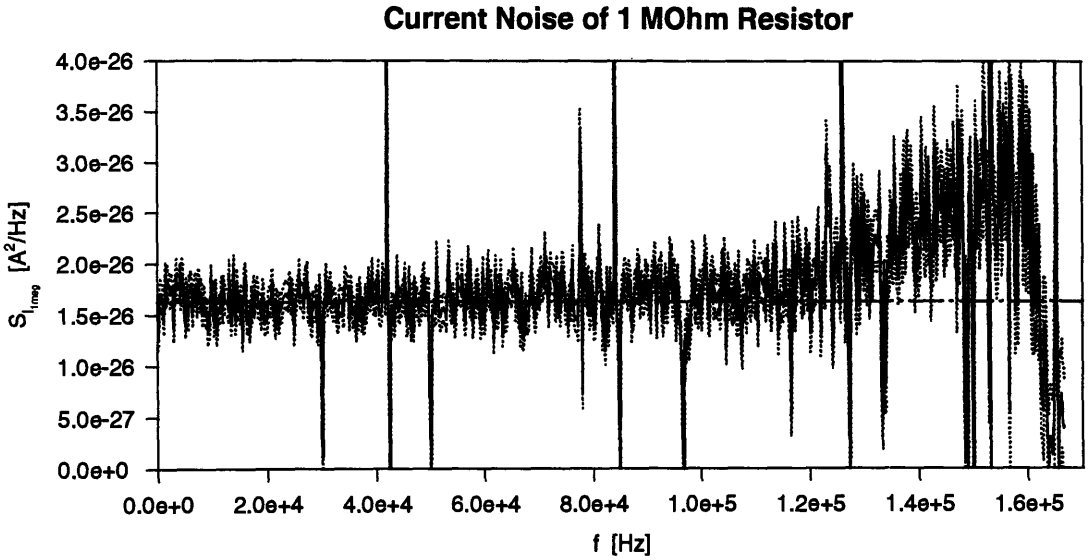

Figure 15: Measured spectral intensity of a 1 MΩ resistor

The measured noise shows excellent agreement with theory, except past 120 kHz where the aliasing described earlier introduced error in the noise measurement.

7.3 Transistor Noise Measurements

Noise measurements were made for transistors with the geometries listed in Table 2. For each transistor, noise measurements were taken with drain bias currents of 1/8, 1/4, 1/2, 1, 2, and 4 μA, and a few transistors were measured at 8 μA. The drain-source bias was 5 V.

Table 2: Transistors used for noise measurements

W [μm]	L [μm]	W/L	W•L [μm^2]
16	8	2	128
32	8	4	256
16	16	1	256
32	16	2	512
4	32	0.125	128
8	32	0.25	256
8	16	0.5	128

Current spectral intensities typical of the test transistor measurements are plotted in Figure 16.

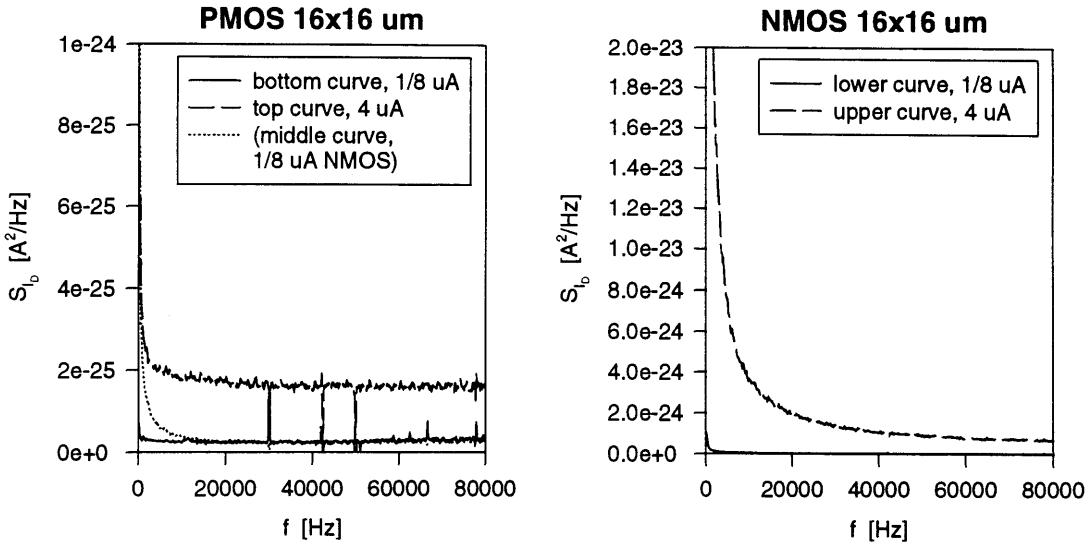

Figure 16: Typical drain current spectral intensity measurements for n- and p-channel transistors at 1/8 μA and 4 μA bias

As evident in the two plots, the $1/f$ noise of the n-channel transistors is significantly greater than for p-channel transistors, especially at higher current levels. This is consistent with other reports of MOSFET $1/f$ noise. Curve-fitting was used to determine the thermal (white) noise and $1/f$ noise components from the measurements. At high frequencies where the spectra flattened out the noise was assumed to be dominated by thermal noise. The average value of the noise over the "flat" spectral range was used as the measure of thermal noise. $1/f$ noise was found by fitting a straight line to the low-frequency end of the spectrum, plotted on a log-log scale, where the slope was approximately -1. The slope of this line is the actual exponent of the $1/f$ dependence, and the constant term is the $1/f$ multiplier which is the important measure of this noise component. Application of these curve-fit techniques to the noise spectra obtained for several transistors at the drain currents specified above produced the data presented in Figures 17 through 21.

7.4 Transistor Noise Measurements -- White Noise

The graph of Figure 17 shows the thermal noise component of the drain current spectral intensity measured for the test transistors at the selected drain currents.

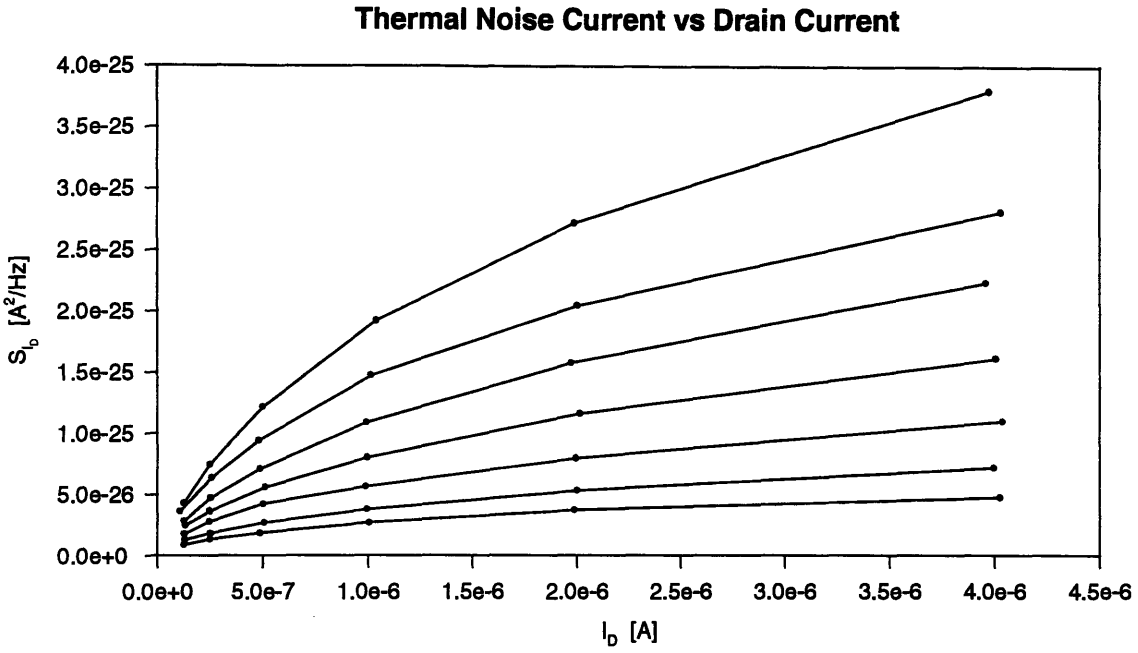

Figure 17: Thermal noise component of drain current spectral intensity measured for multiple transistor geometries and bias currents

Figure 18 presents the same noise data versus transconductance.

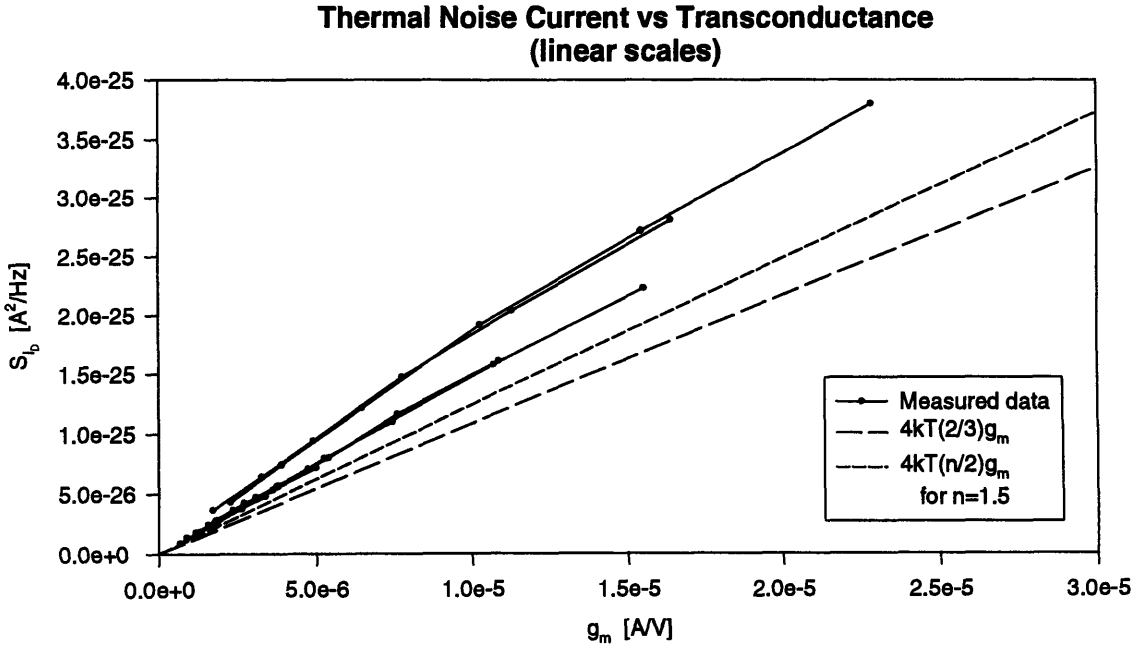

Figure 18: Linear dependence of thermal noise drain current spectral intensity on transconductance, with lines predicted by weak and strong inversion theory

This plot shows the linear relationship between $S_{I_D,\text{thermal}}(f)$ and transconductance. Recalling the theoretical expressions for white noise in weak inversion when the current is saturated:

$$S_{I_D}(f) = 4kT\left(\frac{n}{2}\right)g_m$$

$$S_{V_{GS}}(f) = 4kT\left(\frac{n}{2}\right)\frac{1}{g_m}$$

and in saturated strong inversion:

$$S_{I_D}(f) \approx 4kT\left(\frac{2}{3}\right)g_m$$

$$S_{V_{GS}}(f) \approx 4kT\left(\frac{2}{3}\right)\frac{1}{g_m}$$

it is evident from the plot that the actual slope of $S_{I_D,\text{thermal}}(f)$ versus g_m was slightly greater than predicted. This discrepancy is most likely due to the fact that the derivation of thermal noise assumed a lightly doped substrate, and these test transistors have a more heavily doped channel. The derived relationship between transconductance and inversion charge underestimates the number of carriers in the channel when the substrate is not lightly doped, and since the unaccounted carriers contribute to the noise (by their velocity fluctuations) the drain current noise is underestimated by the resulting theory.

The error in the relationship between drain current spectral intensity and transconductance is the deviation from the predicted slope of 4kT(n/2) in weak inversion and 4kT(2/3) in strong inversion. Figure 19 illustrates this error by plotting the same noise measurements versus transconductance on a log-log scale, and by plotting the factor $S_{I_D}/(4kTg_m)$ against drain current. On the log-log plot on the top, the slope error is seen as a constant offset from the predicted weak and strong inversion thermal noise. The weak inversion ideality factor measured for the p-channel transistors (n=1.5) is used for the predicted subthreshold noise. The graph on the bottom shows $S_{I_D}/(4kTg_m)$, which is the actual multiple of $4kTg_m$ that gives the measured noise. This factor should be n/2 in weak inversion, and 2/3 in strong inversion. The true $4kTg_m$ multiplier for the devices had a slight dependence on the drain current and was higher for the devices with the larger width/length ratios.

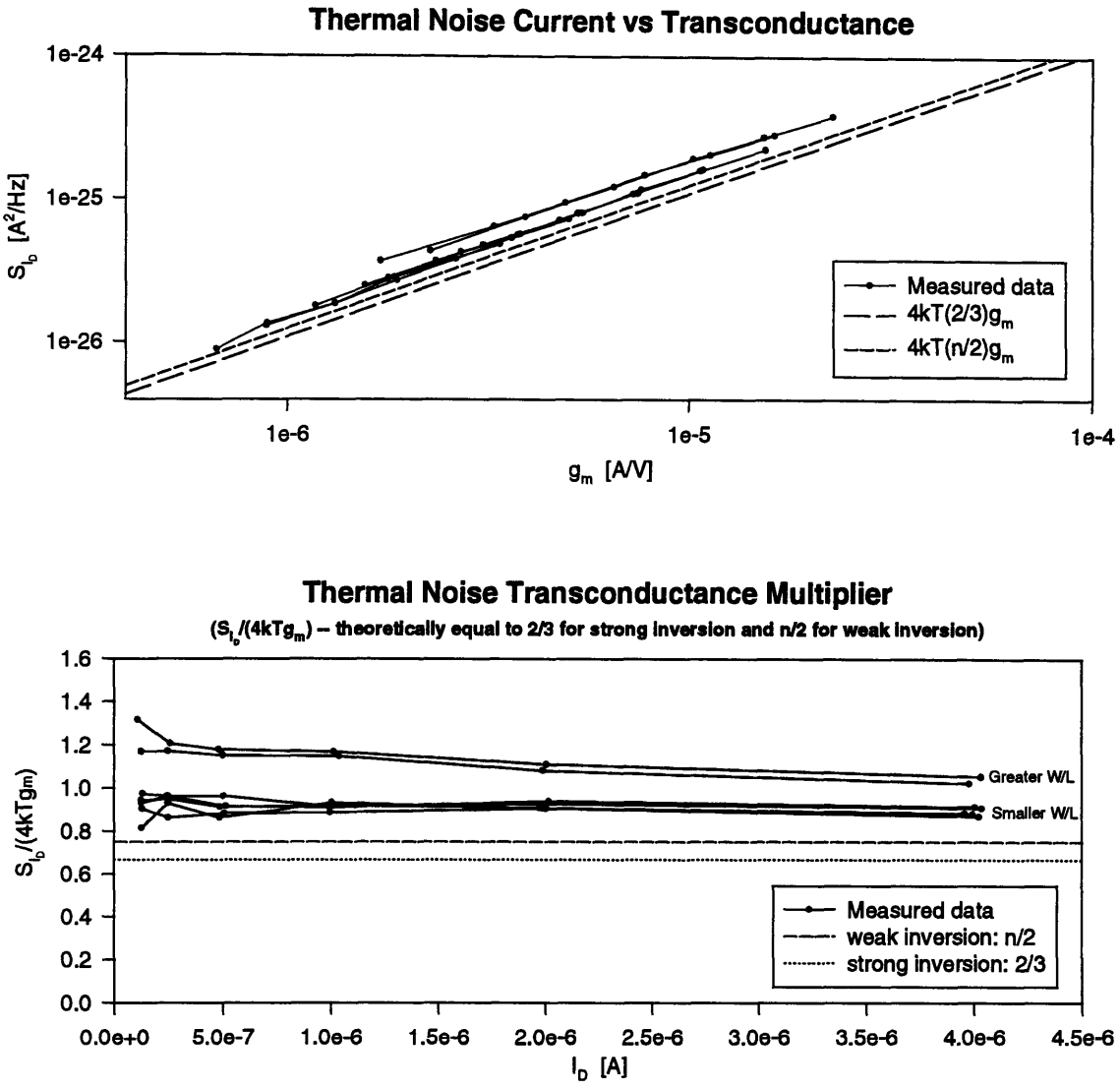

Figure 19: Thermal noise dependence on transconductance

7.5 Transistor Noise Measurements -- 1/f Noise

$1/f$ noise was most easily interpreted by normalizing the spectral intensity for the factors that were predicted to influence it. To review, the equations that were derived for $1/f$ noise are

$$S_{I_D}(f) = \frac{q^4 I_D{}^2 N_T(E_f)_{eff}}{kTWL(C'_{ox} + C'_{ss} + C'_d)^2 \gamma_t f}$$

$$S_{V_{GS}}(f) = \frac{q^2 N_T(E_f)_{eff} kT}{WLC'_{ox}{}^2 \gamma_t f}$$

for weak inversion operation, and

$$S_{I_D}(f) \approx \frac{q^2 \mu I_D N_T (E_f)_{eff} kT}{L^2 C'_{ox}(1+\delta)\gamma_t f}\left[\ln\left(\left(\frac{C'_{ox}}{C'_{ox}+C'_{ss}+C'_d}\right)\frac{(V_{GS}-V_T)}{\left(\frac{kT}{q}\right)}\right)-1\right]$$

$$S_{V_{GS}}(f) \approx \frac{q^2 N_T (E_f)_{eff} kT}{2WLC'^{2}_{ox}\gamma_t f}\left[\ln\left(\left(\frac{C'_{ox}}{C'_{ox}+C'_{ss}+C'_d}\right)\frac{(V_{GS}-V_T)}{\left(\frac{kT}{q}\right)}\right)-1\right]$$

for saturated strong inversion operation.

Since the spectral intensity of the flicker noise component in strong inversion was expected to be proportional to the drain current and inversely proportional to the channel length squared, *dividing* by current and *multiplying* by length squared should give a constant for all of the measurements. A logarithmic curve fit was used to find the constant $1/f$ multiplier from the measured data, so the measured spectral intensities were already normalized for frequency (as $S_{I_D} \cdot f$). The graph on the left in Figure 20 shows the linear dependence of $1/f$ noise on drain current by plotting the measured flicker noise which has only been normalized for the length squared and frequency. The graph on the right shows the measured noise data that has been normalized for drain current, length squared, and frequency. This plot shows that the predicted dependencies of $S_{I_D,\text{flicker}}(f)$ in strong inversion were fairly accurate, since the normalized values of the measured noise did indeed cluster around a constant value at high currents, where the transistors were strongly inverted.

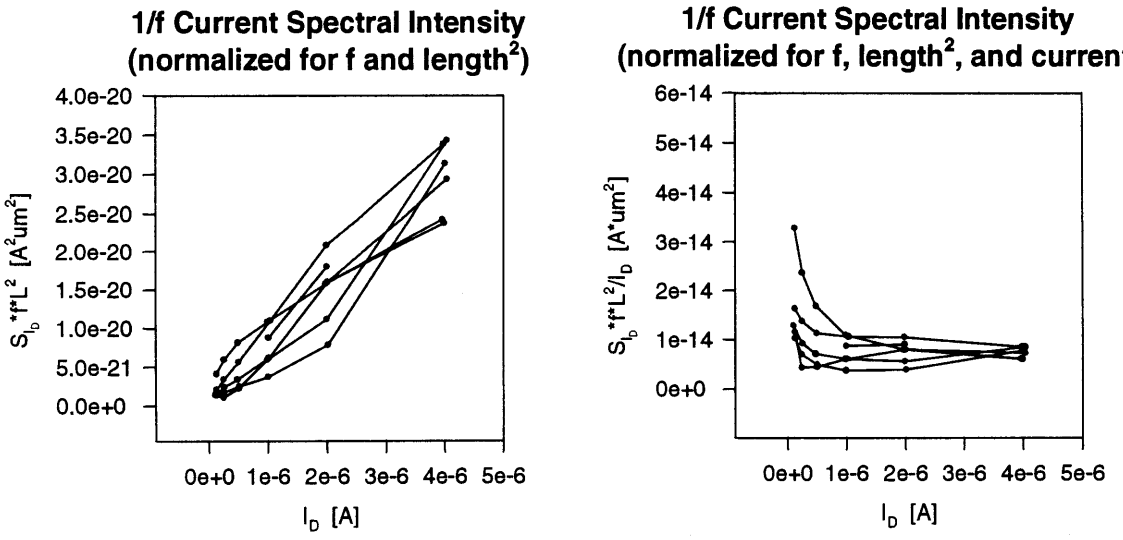

Figure 20: Normalized $1/f$ drain current spectral intensity measurements

Similarly, the gate-referred voltage spectral intensity of the $1/f$ noise (obtained by dividing $S_{I_D,\text{flicker}}(f)$ by g_m^2) could be normalized for gate area, which was the main dependence predicted by the theory. This is plotted in Figure 21.

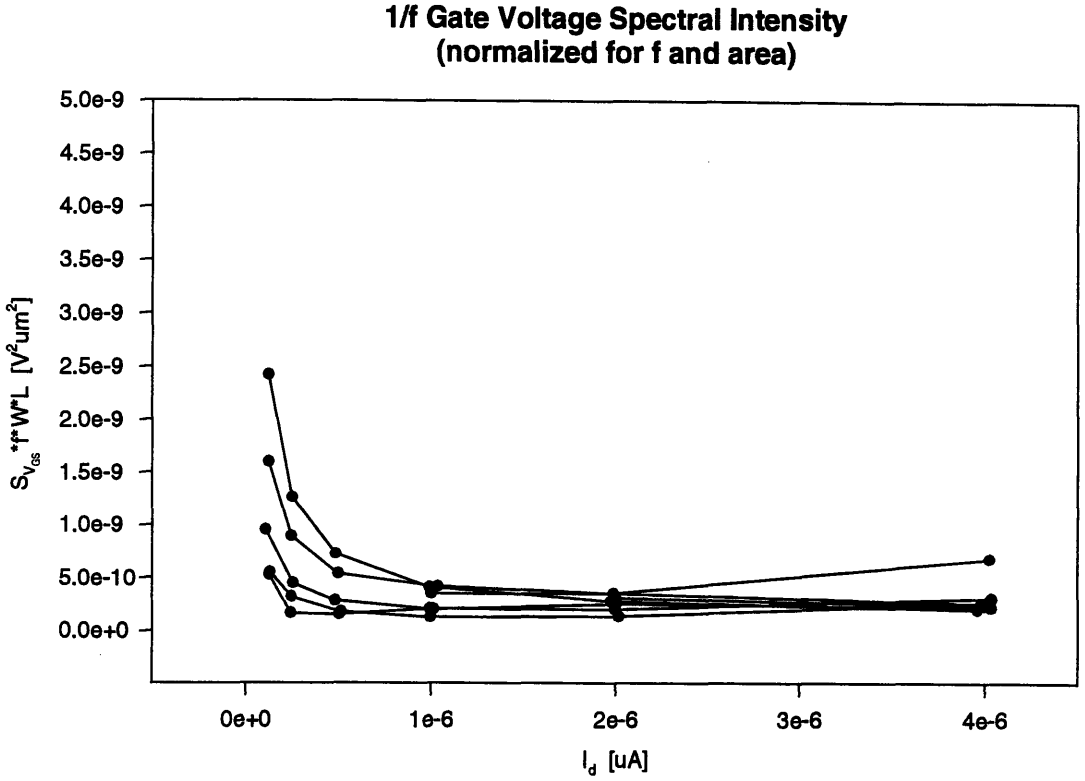

Figure 21: Normalized $1/f$ gate voltage spectral intensity measurements

At high current there is good agreement, but at drain currents below 1μA the data appears to be inversely proportional to current. Many factors may account for this. If the basic trapping mechanism upon which the noise derivation is based is not in error, then perhaps the assumption that the effective trap density is constant is flawed. Higher effective trap density at lower surface potentials would then explain the increase in gate-referred noise at low current levels. In addition, the $1/f$ noise becomes harder to determine accurately for low bias currents (because the white noise begins to dominate), and this could lead to experimental errors. More accurate assessment of the $1/f$ noise could be made by increasing the load resistor for higher sensitivity and sampling at a lower rate to focus on the low frequency end of the spectrum.

Chapter 8

Conclusions

8.1 Summary of MOSFET Theory

Low noise, low power circuit design requirements motivated this study of MOSFET performance, so the results were combined to form a design strategy. To briefly summarize the theory, the MOSFET has two distinct regions of operation: weak and strong inversion. In weak inversion there are very few inversion charges in the channel, so the surface potential is essentially constant in the channel and independent of the drain-source bias. Current in this region is dominated by diffusion of carriers, and the transconductance-to-current ratio is maximum and constant. Strong inversion operation is characterized by a large concentration of mobile carriers in the channel that form a resistive connection between the source and drain. Drift is the dominant mechanism of conduction, and the transconductance-to-current ratio decreases in proportion to the square root of drain current. The drain current threshold between weak and strong inversion operation is proportional to the width/length ratio of the gate.

The noise of a MOSFET is dominated by two sources: a white noise component due to the random thermal motion of carriers comprising the inversion layer, and a $1/f$ component that may be due to trapping of carriers at the semiconductor-oxide interface. The equations describing these noise components in weak and strong inversion operation are repeated here for convenience:

- Thermal noise in weak inversion:

$$S_{I_D}(f) = 2q\left(\frac{1+e^{-V_{DS}/\phi_t}}{1-e^{-V_{DS}/\phi_t}}\right)I_D \qquad S_{V_{GS}}(f) = 2q\left(\frac{1+e^{-V_{DS}/\phi_t}}{1-e^{-V_{DS}/\phi_t}}\right)\frac{(n\phi_t)^2}{I_D}$$

$$S_{I_D}(f) = 4kT\left(\frac{n}{2}\right)\left(\frac{1+e^{-V_{DS}/\phi_t}}{1-e^{-V_{DS}/\phi_t}}\right)g_m \qquad S_{V_{GS}}(f) = 4kT\left(\frac{n}{2}\right)\left(\frac{1+e^{-V_{DS}/\phi_t}}{1-e^{-V_{DS}/\phi_t}}\right)\frac{1}{g_m}$$

- Thermal noise in strong inversion:

$$S_{I_D}(f) = 4kT\left(\frac{2}{3}(1+\delta)\right)\left(\frac{1+\alpha+\alpha^2}{1-\alpha^2}\right)g_m$$

$$S_{V_{GS}}(f) = 4kT\left(\frac{2}{3}(1+\delta)\right)\left(\frac{1+\alpha+\alpha^2}{1-\alpha^2}\right)\frac{1}{g_m}$$

- $1/f$ noise in weak inversion:

$$S_{I_D}(f)=\frac{q^4 I_D{}^2 N_T(E_f)_{eff}}{kTWL(C'_{ox}+C'_{ss}+C'_d)^2\gamma_t f} \qquad S_{V_{GS}}(f)=\frac{q^2 N_T(E_f)_{eff}kT}{WLC'^2_{ox}\gamma_t f}$$

- $1/f$ noise in strong inversion:

$$S_{I_D}(f)=\frac{q^2\mu I_D N_T(E_f)_{eff}kT}{L^2C'_{ox}(1+\delta)\gamma_t f}\left[\frac{\alpha V'_{DS}}{\left(\left(\frac{kT}{q}\right)\left(\frac{C'_{ox}+C'_{ss}+C'_d}{C'_{ox}(1+\delta)}\right)+\alpha V'_{DS}\right)}-\frac{V'_{DS}}{\left(\left(\frac{kT}{q}\right)\left(\frac{C'_{ox}+C'_{ss}+C'_d}{C'_{ox}(1+\delta)}\right)+V'_{DS}\right)}+\ln\left(\frac{\left(\frac{kT}{q}\right)\left(\frac{C'_{ox}+C'_{ss}+C'_d}{C'_{ox}(1+\delta)}\right)+V'_{DS}}{\left(\frac{kT}{q}\right)\left(\frac{C'_{ox}+C'_{ss}+C'_d}{C'_{ox}(1+\delta)}\right)+\alpha V'_{DS}}\right)\right]$$

linear region:

$$\begin{cases} S_{I_D}(f)\approx\dfrac{q^2\mu I_D N_T(E_f)_{eff}kT}{L^2C'_{ox}(1+\delta)\gamma_t f}\ln\left(\dfrac{1}{\alpha}\right) \\ S_{V_{GS}}(f)\approx\dfrac{q^2 N_T(E_f)_{eff}kT}{2WLC'^2_{ox}\gamma_t f}\left(\dfrac{1+\alpha}{1-\alpha}\right)\ln\left(\dfrac{1}{\alpha}\right) \end{cases}$$

saturation:

$$\begin{cases} S_{I_D}(f)\approx\dfrac{q^2\mu I_D N_T(E_f)_{eff}kT}{L^2C'_{ox}(1+\delta)\gamma_t f}\left[\ln\left(\left(\dfrac{C'_{ox}}{C'_{ox}+C'_{ss}+C'_d}\right)\dfrac{(V_{GS}-V_T)}{\left(\frac{kT}{q}\right)}\right)-1\right] \\ S_{V_{GS}}(f)\approx\dfrac{q^2 N_T(E_f)_{eff}kT}{2WLC'^2_{ox}\gamma_t f}\left[\ln\left(\left(\dfrac{C'_{ox}}{C'_{ox}+C'_{ss}+C'_d}\right)\dfrac{(V_{GS}-V_T)}{\left(\frac{kT}{q}\right)}\right)-1\right] \end{cases}$$

8.2 Low Noise, Low Power Circuit Design -- Example and Guidelines

To demonstrate how these noise models could be applied to circuit design, the equivalent input-referred noise of the simple differential amplifier of Figure 22 was considered.

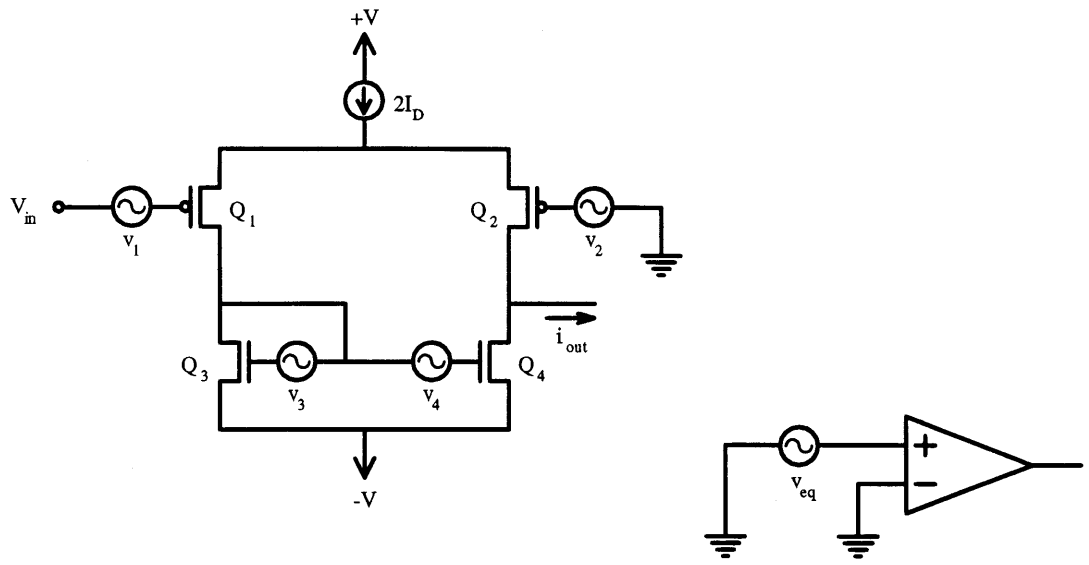

Figure 22: Differential amplifier circuit with transistor noise sources, modeled by noiseless amplifier with equivalent input noise source

Either the drain current noise source or the gate voltage noise source may be used to model each transistor, but it is often more convenient to use the latter. By referring each transistor's noise sources back to the input of the amplifier, the spectral intensities of all of the components can be added to get the equivalent input noise of the amplifier. This is the quantity that sets the minimum detectable signal limit of the amplifier. Since a MOSFET gate serves as the input to the amplifier, the equivalent input current noise is negligible, and the noise of the amplifier may be modeled by a single input-referred voltage noise source. If transistors Q_1 and Q_2 are matched and have transconductance $g_{m_{1,2}}$, and transistors Q_3 and Q_4 are matched with transconductance $g_{m_{3,4}}$, the equivalent voltage noise spectral intensity at the input to the amplifier will be:

$$S_{v_{eq}} = S_{v_1} + S_{v_2} + \left(\frac{g_{m_{3,4}}}{g_{m_{1,2}}}\right)^2 \left(S_{v_3} + S_{v_4}\right)$$

where for each transistor i,

$$S_{v_i} \equiv S_{V_{GSi},\text{thermal}} + S_{V_{GSi},\text{flicker}}$$

In general, the only parameters that can be specified by the designer are the width, length, and bias current of the transistors. In terms of these parameters, the gate-referred noise sources of the transistors may be approximated by:

$$S_{V_{GS},\text{thermal}} \approx \frac{K_t}{g_m}$$

$$S_{V_{GS},\text{flicker}} \approx \frac{K_f}{WLf}$$

Substituting these approximations into the equation for $S_{v_{eq}}$ reveals the trends in the total noise of the amplifier:

$$S_{v_{eq}} \approx \underbrace{2\frac{K_{t_{1,2}}}{g_{m_{1,2}}} + 2\left(\frac{g_{m_{3,4}}}{g_{m_{1,2}}}\right)^2 \frac{K_{t_{3,4}}}{g_{m_{3,4}}}}_{\text{thermal}} + \underbrace{2\frac{K_{f_{1,2}}}{W_{1,2}L_{1,2}f} + 2\left(\frac{g_{m_{3,4}}}{g_{m_{1,2}}}\right)^2 \frac{K_{f_{3,4}}}{W_{3,4}L_{3,4}f}}_{\text{flicker}}$$

$$\approx \underbrace{\frac{2}{g_{m_{1,2}}}\left(K_{t_{1,2}} + \left(\frac{g_{m_{3,4}}}{g_{m_{1,2}}}\right)K_{t_{3,4}}\right)}_{\text{thermal}} + \underbrace{\frac{2}{f}\left(\frac{K_{f_{1,2}}}{W_{1,2}L_{1,2}} + \left(\frac{g_{m_{3,4}}}{g_{m_{1,2}}}\right)^2 \frac{K_{f_{3,4}}}{W_{3,4}L_{3,4}}\right)}_{\text{flicker}}$$

where the transistors in each matched pair are assumed to have the same noise characteristics, and the $1/f$ and thermal components have been grouped. With the transistors in the amplifier circuit operated in saturation, the K_t coefficients should be almost constant through weak and strong inversion. Thus the thermal noise of the amplifier may be minimized by making $g_{m_{1,2}}$ large, and $g_{m_{3,4}}$ small. And since g_m is maximum in weak inversion for a particular amount of drain bias current, transistors Q_1 and Q_2 should be operated in weak inversion. By the same reasoning, transistors Q_3 and Q_4 should be operated in strong inversion in order to minimize their transconductance at this current level. These objectives may be accomplished by making the width/length ratio of Q_1 and Q_2 large enough that they are in weak inversion at the selected bias current, and by making the width/length ratio of Q_3 and Q_4 much smaller. Flicker noise of the amplifier will be minimized simply by making the transistors large. The ratio of the transconductances that was used to minimize thermal noise will reduce the total $1/f$ noise as well. It is also important to use p-channel transistors for Q_1 and Q_2 because of the lower level of $1/f$ noise typically exhibited by these devices.

Combining these design considerations, a strategy can be formed to minimize power (by minimizing drain current) given a maximum acceptable noise level:

1. Use p-channel transistors for the critical input devices Q_1 and Q_2.
2. Set bias current large enough that the thermal noise of Q_1 and Q_2 is a reasonably small fraction of the total noise budget. Because these transistors will be operated in weak inversion, the thermal noise will be described by the shot noise formula, so that:

$$S_{V_{GS1,2},\text{thermal}} = \frac{K_{t_{1,2}}}{g_{m_{1,2}}} = \frac{2(nkT)^2}{qI_D}$$

This can be used to set the bias current I_D by re-arranging terms:

$$I_D = \frac{2(nkT)^2}{qS_{V_{GS1,2},\text{thermal}}}$$

3. Make the width/length ratio of Q_1 and Q_2 just large enough that they will operate in weak inversion at this bias current. This is achieved by measuring the threshold current of a particular transistor from the same fabrication process, and multiplying its width/length ratio by the ratio of the desired threshold current to the measured value.
4. Make the width/length ratio of Q_3 and Q_4 much smaller than Q_1 and Q_2 so that their thermal and flicker noise contributions will be significantly attenuated.
5. Based on noise measurements (of transistors from the same process), scale the area of all four transistors so that their combined $1/f$ noise is within the noise budget.

Following a similar approach, circuits of greater complexity may be designed for minimum noise and power.

8.3 Future Work

Some suggestions for future work that would further increase the applicability and general value of this study are:

1. Noise measurements at lower bias levels, focusing on the lower end of the frequency spectrum where $1/f$ noise is dominant. A simple change in the load resistor of the measurement circuit and in sampling rate would facilitate this. Most of the noise measurements taken for this study happened to be at the upper end of weak inversion and in strong inversion. While the upper limit of weak inversion is a desirable operating range from a low noise, low power design perspective (because the apparent benefits of subthreshold operation are obtained without requiring excessively wide (i.e. large) transistors), a more thorough characterization of weak inversion $1/f$ noise would be valuable. This could lead to a better understanding, or at least a better ability to predict, this form of noise. If the $1/f$ noise is found to be significantly higher or lower deep into weak inversion, then this could lead to a revised design strategy.
2. Evaluation of the transistor threshold and noise characteristics of multiple processes, and of particular processes over time. The ability to predict the threshold current (between weak and strong inversion) and noise behavior of transistors is essential to the design of optimal integrated circuits.
3. Tests of the bipolar junction transistors (BJTs) available in BiCMOS processes, for possible replacement of noisy NMOS transistors. The $1/f$ noise of BJTs should be very low, and for components of a circuit that do not require the high input impedance of a MOSFET gate, the noise performance may be superior. As done for the MOSFET, the bipolar transistor's noise sources must be included in a circuit model and interpreted within this context in order to determine the relative merit of the BJT.

References

Agnew, William F., and Douglas B. McCreery, eds. Neural Prostheses: Fundamental Studies. Englewood Cliffs, N.J.: Prentice-Hall, 1990.

Berz, F. "Theory of Low Frequency Noise in Si MOST's." Solid State Electronics 13 (1970): 631-647.

Buckingham, M.J. Noise in Electronic Devices and Systems. Chichester: Ellis Horwood Limited, 1983.

Christensson, S., I. Lundstrom, and C. Svensson. "Low Frequency Noise in MOS Transistors -- I: Theory." Solid State Electronics 11 (1968): 797-812.

Christensson, S., and I. Lundstrom. "Low Frequency Noise in MOS Transistors -- II: Experiments." Solid State Electronics 11 (1968): 813-820.

Clark, Lloyd Dan, Jr. "A System for Chronic Neural Signal Transduction, Processing, and Control." Diss. Massachusetts Institute of Technology, 1990.

Das, Makunda B., and James M. Moore. "Measurements and Interpretation of Low-Frequency Noise in FET's." IEEE Transactions on Electron Devices ED-21, No. 4, April 1974: 247-257.

Deen, M.J., and Z.X. Yan. "A New Method for Measuring the Threshold Voltage of Small-Geometry MOSFETs from Subthreshold Conduction." Solid State Electronics 33 (1990): 503-511.

Duh, K.H., and A. van der Ziel. "Thermal and 1/f Noise at Weak Inversion and Limiting Noise in MOSFETs." Proceedings of the 7th International Conference on Noise in Physical Systems and 3rd International Conference on 1/f Noise: 291-293. Montpellier, France, May 1983. Elsevier Science Publishers B.V., 1983.

Edell, David J. "Development of a Chronic Neural Interface." Diss. University of California at Davis, 1980.

Edell, David J. "A Peripheral Nerve Information Transducer for Amputees: Long-Term Multichannel Recordings from Rabbit Peripheral Nerves." IEEE Transactions on Biomedical Engineering BME-33, No. 2, February 1986: 203-214.

Edell, D.J., V.M. McNeil, L.D. Clark, and T.V. Van. "Insertable Neural Information Transducer Technology." In 39th ACEMB, 1986.

Fellrath, J. "Shot Noise Behaviour of MOS Transistors." Revue De Physique Appliquée Dec. 1978: 719-723.

Godfrey, Michael D. "CMOS Device Modeling for Subthreshold Circuits." IEEE Transactions on Circuits and Systems 39 (1992): 532-539.

Gray, Paul R., and Robert G. Meyer. Analysis and Design of Analog Integrated Circuits. 3rd ed. New York: John Wiley & Sons, 1993.

Hambrecht, F. Terry, and James B. Reswick, eds. Functional Electrical Stimulation: Applications in Neural Prostheses. New York: Marcel Dekker, 1977.

Hsu, S.T. "Surface State Related 1/f Noise in MOS Transistors." Solid State Electronics 13 (1970): 1451-1459.

Kingston, Robert H. Detection of Optical and Infrared Radiation. New York: Springer-Verlag, 1978.

Kornfeld, A. "Subthreshold Conduction and 1/f Noise in Si-MOS Transistors at Cryogenic Temperatures." 14th Convention of Electrical and Electronic Engineers in Israel Proceedings: 1.4.2/1-4. Tel Aviv, 1985. New York: IEEE, 1985.

Lineback, J. Robert. "Subthreshold FETs Keep Pacemaker Batteries Alive." Electronics 22 July 1985: 16.

McNeil, Vincent Maurice. "Design and Fabrication of an Optimized Preamplifier for a Chronic Neural Transducer." Master's Thesis, Massachusetts Institute of Technology, 1988.

Reimbold, G. "Modified 1/f Trapping Noise Theory and Experiments in MOS Transistors Biased from Weak to Strong Inversion -- Influence of Interface States." IEEE Transactions on Electron Devices ED-31, No. 9, Sept. 1984: 1190-1198.

Sah, C.T. "Theory of Low-Frequency Generation Noise in Junction-Gate Field-Effect Transistors." Proceedings of the IEEE 52 (1964): 795-814.

Sarpeshkar, Rahul, Tobias Delbrück, and Carver A. Mead. "White Noise in MOS Transistors and Resistors." IEEE Circuits and Devices Magazine Nov. 1993: 23-29.

Schutte, C., and P. Rademeyer. "Subthreshold 1/f Noise Measurements in MOS Transistors Aimed at Optimizing Focal Plane Array Signal Processing." Analog Integrated Circuits and Signal Processing 2 (1992): 171-177.

Streetman, Ben G. Solid State Electronic Devices. 3rd ed. Englewood Cliffs, N.J.: Prentice-Hall, 1990.

Sze, S.M. Physics of Semiconductor Devices. 2nd ed. New York: John Wiley & Sons, 1981.

Tsividis, Yannis P. Operation and Modeling of the MOS Transistor. New York: McGraw-Hill, 1987.

Van der Ziel, Aldert. Noise. Englewood Cliffs, N.J.: Prentice-Hall, 1956.

Van der Ziel, Aldert. Noise in Solid State Devices and Circuits. New York: John Wiley & Sons, 1986.

Yariv, Amnon. Introduction to Optical Electronics. 2nd ed. New York: Holt, Rinehart and Winston, 1976.

About the Author

Bruce Larson grew up in Omaha, Nebraska. After graduating from Creighton Preparatory School in 1987, he entered the University of Houston as a National Merit Scholar. There he was named the Outstanding Senior Engineering Student in the both the Department of Electrical Engineering and the Cullen College of Engineering in 1992. He graduated Summa Cum Laude with a Bachelor of Science degree in Electrical Engineering in May of 1992, and began attendance at the Massachusetts Institute of Technology in September of that year. He is presently a Ph.D. candidate in the Department of Electrical Engineering and Computer Science. His research interests include low noise, low power analog circuit and system design, optical power and communication techniques, and chronic neural interface technology.