

## Chapter 6 Linear Networks: By Diffusion in MOS Transistors

In this chapter, I extend the device-level charge-based formulation of the MOS transistor to the circuit level by introducing the concepts of terminal and node charges, and the equivalence principle. With this formalism, we can exploit the linear current–charge relationship of the MOS transistor at the circuit level, enabling us to simulate the diffusion of ions in cell syncytia, or the spread of current in resistive networks, extremely efficiently.

Ions spread from cell to cell in a syncytium through ion channels that are part of the gap-junction synaptic complex formed between these cells. At gap junctions, the membranes of two cells are in close juxtaposition, and pores in the two membranes are lined up. Hence, a channel is formed, and ions cross from the intracellular fluid of one cell directly into the intracellular fluid of another.

When the ion channel sees a small concentration gradient—as it does when a gap junction is formed between cells of the same type—transport of ions is primarily by drift. As we saw in Chapter 5, in the constant-field approximation, drift produces a linear current–voltage relationship. It is difficult to reproduce this linear current–voltage characteristic with the transistor. The transistor’s behavior is close to linear over a range of only a few thermal voltages. For higher voltages, its current either increases exponentially or saturates, depending on the polarity of the voltage.

I show here how we can obtain linear behavior by exploiting the inherently linear current–charge relationship of the transistor in the subthreshold regime.

## 6.1 Symmetric MOS Transistor Model

Using the charge-based formulation for the channel current in a MOS transistor, we can develop an intuitive, physically accurate, circuit-level abstraction of this device. The parallel-plate-capacitor approximation yields a simple and symmetric relationship between the channel current and the charge at the ends of the channel. As it turns out, this intuitively appealing model does indeed provide a good fit to the experimental measurements [118, 119]. Thus, we can confidently use it to develop circuit models of the transistor.

In addition to preserving our intuition about the symmetrical construction of the MOS transistor, a symmetric model is easier to use than an asymmetric one. When you use asymmetric models, you have to figure out which terminal is the source, and then reference the voltages of all the other terminals to that source terminal before you can apply the model. With a symmetric model, all voltages are referenced to the bulk, and it is not necessary to know a priori which terminal is the drain and which is the source. Because the roles of the channel terminals are determined by the direction in which the charge carriers flow, source and drain can be determined by the circuit design, by the bias conditions—and even by the input signals. Using a symmetric model makes it easier to understand the behavior of circuits with such flexibility, and enhances our ability to design circuits that use unidirectional currents as well as bidirectional currents.

I adopt the conventions for voltages and currents in pMOS and nMOS transistors shown in Figure 6.1a, to preserve symmetry between the channel terminals. For example, in an n-well process, the local reference for the nMOS transistors is the p-substrate  $V_{\text{NBB}}$  (usually  $V_{\text{SS}}$ ); for the pMOS transistors, it is the n-well  $V_{\text{PBB}}$  (usually  $V_{\text{DD}}$ ). This notation is consistent with that used by Mead [119], and also with that commonly used in the subthreshold MOS literature [121]. Because the drain and source terminals are treated symmetrically by these conventions, and because the circuit model for the device is itself symmetric, we can assign labels of source and drain to the channel terminals arbitrarily, without regard to the actual direction of

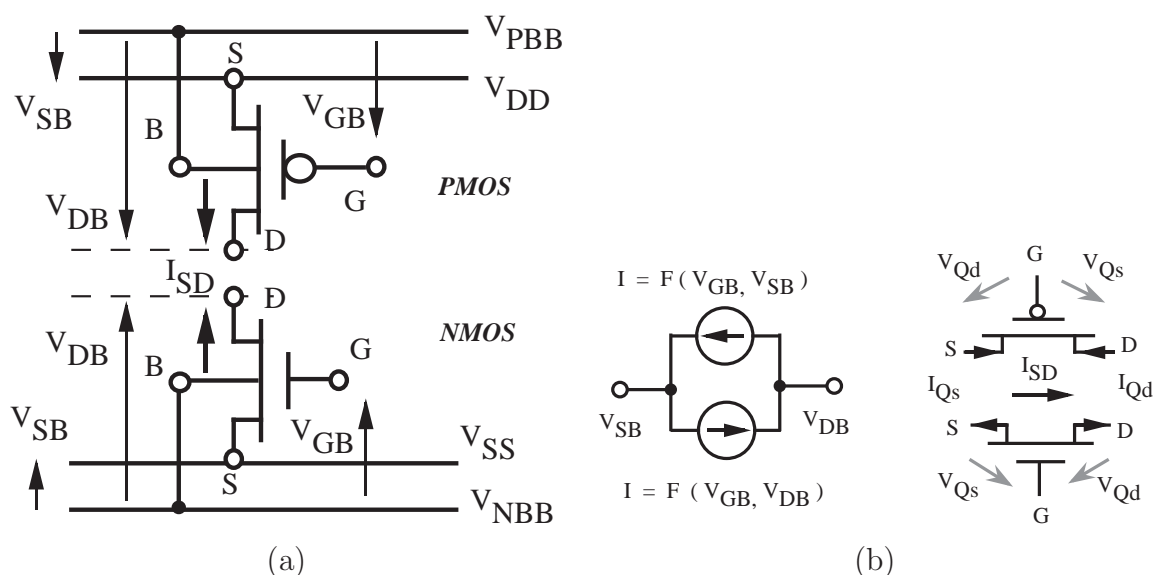


Figure 6.1: CIRCUIT CONVENTIONS FOR THE MOS TRANSISTOR

(a) Adopted conventions for voltages polarities and currents directions. (b) Symmetric decomposition of channel current into source and drain components.

the channel current.

## 6.2 Source- and Drain-Current Components

As shown in Section 5.2.1, for the parallel-plate capacitor approximation, both diffusion and drift are proportional to the charge-concentration gradient. This linear relation yields a quadratic expression for the current that consists of two symmetric, independent, opposing components (see Equation 5.10), as has been shown by other device physicists [122, 123, 120]. Therefore, the channel current can be decomposed into a **source component** and a **drain component**, as illustrated in Figure 6.1b.

One component is a function of the mobile charge at the source end of the channel; the other component is a function of the mobile charge at the drain end. Because the charge concentrations are related to voltages applied to the source terminal and to the drain terminal, respectively, as well as to the gate voltage, we can express the

channel current, per unit width, of an nMOS transistor in the form

$$I_{SD} = \mathcal{P}(L)(\mathcal{Q}_n(V_{GB}, V_{SB}) - \mathcal{Q}_n(V_{GB}, V_{DB})). \quad (6.1)$$

We can write a similar equation for the pMOS transistor in terms of  $\mathcal{Q}_p(V_{GB}, V_{SB})$ .

As we saw in Section 5.2.2, the source component is related to the source voltage and the gate voltage by *exactly* the same function that relates the drain component to the drain voltage and the gate voltage. Therefore, which component we call the drain and which we call the source has nothing to do with the device—it is purely a question of which direction we prefer for the current in the device.

As we saw in Section 5.2.2, the charge–voltage relationship,  $\mathcal{Q}_n(V_1, V_2)$ , has a complicated form, due to the exponential dependence of charge concentration on potential, and the compressive relationship between surface potential and gate voltage. These highly nonlinear dependencies arise because three different charged species are involved in the electrostatics. The complexity of the electrostatics obscures the simplicity of the drift–diffusion transport mechanisms that determine the charge–current relationship (Equation 6.1).

As shown in Section 5.2.1, in the subthreshold regime, the mobile-charge concentration grows exponentially as the gate voltage increases. This behavior is due to the linear relationship between gate voltage and surface potential when the mobile-charge concentration is much smaller than the depletion-layer charge. When the relationship is exponential, changing the gate voltage multiplies the mobile-charge concentrations at the source and the drain ends of the channel by the same amount. This multiplicative effect, together with the strictly linear relationship between current and charge in the diffusion-dominated subthreshold regime, allows us to factor out the dependence on the gate voltage. Hence, we can express the channel current per unit width in the form

$$I_{SD} = \frac{\mathcal{K}_n(V_{GB})\mathcal{D}}{L}(\mathcal{Q}_n(V_{SB}) - \mathcal{Q}_n(V_{DB})), \quad (6.2)$$

for an nMOS transistor; and similarly for a pMOS transistor in terms of  $\mathcal{K}_p(V_{GB})$  and

$\mathcal{Q}_p(V_{CB})$ , where

$$\mathcal{D} = \mu V_T, \quad (6.3)$$

$$\mathcal{K}_n(V_{GB}) = n_0 e^{\kappa_n V_{GB}/V_T}, \quad (6.4)$$

$$\mathcal{K}_p(V_{GB}) = p_0 e^{-\kappa_p V_{GB}/V_T}, \quad (6.5)$$

$$\mathcal{Q}_n(V_{CB}) = -C_{\text{dep}} V_T e^{-V_{CB}/V_T}, \quad (6.6)$$

$$\mathcal{Q}_p(V_{CB}) = C_{\text{dep}} V_T e^{V_{CB}/V_T}. \quad (6.7)$$

### 6.3 Conditions for Symmetric Current Decomposition

In general, current decomposition is difficult to achieve for devices whose charge transport is governed by the drift–diffusion process. The difficulty arises because decomposition precludes any dependence of  $\mathcal{P}(L)$  on the terminal voltages. Integrating the diffusionlike formulation of the Nernst–Planck equation (Equation 5.5), with the flux held constant, tells us that

$$\frac{1}{\mathcal{P}(L)} \propto \int_0^L e^{\psi(x)/V_{Tn}} dx.$$

Hence,  $\mathcal{P}(L)$  is constant when the integral is constant. The preceding integral is independent of the voltages applied at the ends of the channel only if the dependence of the potential profile on these voltages is constrained appropriately. This constraint is not satisfied in general; in particular, the cell membrane does not satisfy it. The constant-field assumption makes the potential change linearly along the channel, with its values at the ends equal to the voltages applied there. Hence, the equivalent permeability of the membrane is not independent of the voltages on either side of the membrane, as is evident in the solution for the membrane current (Equation 5.8).

Any voltage dependence of the nominally constant physical and geometrical properties of the device also violates decomposition. For the transistor, the susceptible

physical constants that are factored into  $\mathcal{P}(L)$  and  $\mathcal{D}$  are the mobility and the channel length  $L$ . The mobility degrades at high electric fields due to velocity saturation, and the effective channel length is reduced by the depletion layers at the source–bulk and drain–bulk junctions. The widths of these depletion layers depend on the voltages across those junctions. The channel width is also prone to modulation due to field fringing. Both velocity saturation and channel-length and channel-width modulation are negligible for devices with dimensions well over  $1\mu\text{m}$ , but they become critically important for submicron devices.

The susceptible constant that is factored into  $\mathcal{Q}_n(V_{\text{GB}}, V_{\text{SB}})$  and  $\mathcal{Q}_p(V_{\text{GB}}, V_{\text{DB}})$  is the depletion-layer capacitance. This voltage-dependent capacitance is fairly constant below threshold, because the surface potential is virtually independent of the source and drain voltages below threshold—it is determined primarily by the gate voltage. Above threshold, the mobile-charge concentration is limited by the gate-oxide capacitance, which is constant. Therefore, the voltage dependence of the depletion capacitance does not limit decomposition.

Symmetry is violated by any changes in doping profile along the channel, or by differences in doping between the source–drain regions, such as in the lightly doped drain (LDD) structure employed in submicron devices. Symmetry is also violated by differences in area between the source and drain regions. When such differences are present, we can no longer use the same function,  $\mathcal{Q}_n(V_{\text{GB}_n}, V_i)$ , to compute the terminal charge at both ends of the channel. Instead, we must use one function  $\mathcal{Q}_{\text{L}_n}(V_{\text{GB}_n}, V_i)$ , for the left end of the channel, and a different function,  $\mathcal{Q}_{\text{R}_n}(V_{\text{GB}_n}, V_i)$ , for the right end.

## 6.4 The Terminal Charge

I introduce the concept of **terminal charge** to exploit the inherent linearity of the MOS transistor. Each drain–source terminal is assigned a fictitious charge that is given by  $\mathcal{Q}_n(V_{\text{GB}_n}, V_i)$ , where  $i$  is the label of the node to which the terminal is connected, and  $n$  is the label of the transistor to which the terminal belongs. Terminal charge

is negative for an nMOS transistor because electrons serve as charge carriers, and is positive for a pMOS transistor because holes serve as charge carriers.

In terms of these terminal charges, the current that flows from node  $i$  to node  $j$ , via transistor  $n$ , is in general given by

$$I_{i,j} = W_n \mathcal{P}(L_n) (\mathcal{Q}_n(V_{GB_n}, V_i) - \mathcal{Q}_n(V_{GB_n}, V_j)), \quad (6.8)$$

from Equation 6.1. There is a perfectly *linear* relationship between the difference in terminal charges and the current, as though terminal-charge transport occurs by diffusion across a device with **permeability**  $\mathcal{P}(L_n)$  per unit width. The effective permeability  $W_n \mathcal{P}(L_n)$  is fixed because it depends on only physical constants, such as mobility, thermal voltage, and the width and the length of the channel—the designer specifies the channel width and the channel length.

In the special case, where we restrict operation to the subthreshold regime, the current is given by

$$I_{i,j} = W_n \frac{\mathcal{K}(V_{GB_n}) \mathcal{D}}{L_n} (\mathcal{Q}_n(V_i) - \mathcal{Q}_n(V_j)), \quad (6.9)$$

from Equation 6.2. Therefore, in subthreshold, the permeability can be factored into a **diffusivity**  $\mathcal{D}$ , which is constant, and a **partition coefficient**  $\mathcal{K}(V_{GB})$ , which is a function of the gate voltage. Thus, we can use the gate voltage to control how the charge concentration partitions between the source–drain regions and the channel, and thereby we can control the permeability of the device electronically. In contrast, the only way to change the permeability in the general case is to change the channel length or the channel width, and we cannot do that after the device has been fabricated.

I have shown how, in theory, we can transform the nonlinear MOS transistor into a linear element by performing a mapping  $\mathcal{Q}(V_1, V_2)$  on the voltages applied to its terminals. We may achieve this linearity in practice if we adhere to the design and operation constraints under which symmetric current decomposition is valid, as

discussed in Section 6.2.

## 6.5 Diffusors, Pseudoconductances, and Ohm's Law

I draw analogies between the MOS transistor operating below threshold and diffusion across a permeable membrane because Equation 6.9 arises from diffusion-dominated charge transport in the transistor. This mode of transport gives rise to an inherently linear charge–current relationship—unlike for drift-dominated transport. Therefore, when I use transistors that exploit this linear relationship in a circuit, I call them **diffusors** [124]. The analogy with the physical process of diffusion serves our intuition well, and allows us to make comparisons with neurobiology; I got the idea of using a transistor in this inherently linear fashion by making an analogy between the transistor and a gap junction [125].

The analogy between charge flow in a transistor and diffusion of uncharged particles across a porous membrane is perfect in the subthreshold regime, where transport is due primarily to diffusion, and the charge on the mobile carriers is negligible compared to the charge on the gate and the immobile charge in the depletion layer. Above threshold, however, transport is primarily due to drift, and the mobile charge is the dominant charge species. To the extent that the parallel-plate–capacitor model is valid, the derivative of charge concentration with respect to potential is constant, and therefore the charge-concentration profile along the channel is simply a scaled version of the potential profile. Therefore, at a microscopic level, we can model the charge transport as a diffusion process with a diffusion coefficient proportional to the local charge concentration, as shown by Equation 5.9, the diffusionlike formulation for the transistor. Therefore, the analogy is not disingenuous, as long as we bear in mind the proportionality between permeability and charge concentration above threshold. This dependence explains the macroscopic quadratic relationship between current and charge above threshold.

Other researchers have proposed viewing the transistor as a pseudoconductance, with a linear relationship between current and pseudovoltage [126, 127]. However, this

view disconnects us from the physics of the transistor. Conductance, or resistance, is a property of devices that obey Ohm's law, which states that the current density is proportional to the potential gradient. Transistors do not obey Ohm's law because the charge-carrier-concentration gradient along the channel is not zero. Linearity between potential gradient and current density holds only when the concentration is constant, and hence transport is due entirely to drift, and drift is linear at the macroscopic level when the charge-carrier-concentration profile is flat. However, in the above-threshold regime, where carriers drift, the concentration profile along the channel is not flat, and, in the subthreshold regime, carriers do not drift.

Actually, it is physically impossible to achieve linearity at the macroscopic level by satisfying Ohm's law at the microscopic level. Constant flux and constant concentration along the length of a conductor imply constant electric field. But this uniformity in both the charge density and the electric field is inconsistent with Gauss' law, which states that the electric field is the integral of the charge. Introducing an oppositely charged species to neutralize the charge does not produce linear behavior either, as we saw for the membrane. Carbon-film resistors, as well as other varieties of resistors, use thousands of elements connected in series to achieve linearity, by ensuring that the voltage drop across each element is smaller than the thermal voltage. Such small-signal operation makes the nonlinear behavior of these element irrelevant, effectively linearizing the element.

We can satisfy Ohm's law at the microscopic level by introducing another pair of terminals, in addition to the pair that conduct the current, as Tsividis and his colleagues have shown [128]. We achieve the correct microscopic behavior by placing voltage gradients on the gate and on the bulk that match the voltage gradient at the channel surface, such that  $C_{\text{ox}}(V_{\text{G}}(x) - \psi_{\text{s}}(x)) + C_{\text{dep}}(\psi_{\text{s}}(x) - V_{\text{B}}(x))$  does not change with  $x$ . This arrangement makes the mobile-charge concentration, as well as the depletion charge and the hole charge, constant everywhere, thereby satisfying Ohm's law for the entire range of operation of the device. The concentration of mobile charge can be controlled linearly above threshold by the potential difference between the gate and surface, or exponentially below threshold by the potential difference

between the gate and the source–drains. Thus, the conductivity of the material can be controlled electronically. We need additional circuitry, however, to make the pair of voltages applied to the ends of the gate, and the pair of voltages applied to the ends of the bulk, track the voltages on the source and the drain, and we must use a highly-resistive gate layer to limit power dissipation.

By thinking in terms of terminal charges and viewing the transistor as a diffuser, I created a simple circuit-level abstraction for the device. This viewpoint also gives us a powerful analogy between transistors and the porous membranes found in nerve cells. In particular, when operation is restricted to the subthreshold regime, we can control the permeability of the device by changing the gate voltage, which modulates the partition coefficient. Electronic control gives us the ability to model active properties of ion channels in the cell membrane. However, we must extend this abstraction to the circuit level, if we wish to use our abstract diffusers to build a real linear network.

## 6.6 The Node Charge

I introduce the concept of **node charge** to extend the device-level terminal-charge concept to the circuit level. It is possible to extend the latter concept to the circuit level if the **equivalence property** holds:

All terminals connected to the same node have the same terminal charge.

Equivalence between node voltage and terminal charge allows devices to communicate their terminal charges using the voltage on the common node to which they are connected. When equivalence holds, we can replace device-level terminal charges with circuit-level node charges. Thus, equivalence allows us to exploit the linearity between terminal charge and current at the circuit level.

In the general case (Equation 6.1), equivalence holds if

$$V_i = V_j \Rightarrow \mathcal{Q}_n(V_{GB_m}, V_i) = \mathcal{Q}_n(V_{GB_n}, V_j) \Rightarrow V_{GB_m} = V_{GB_n}.$$

Therefore, equivalence limits the ways in which transistors can be connected: Transistors connected to the same node cannot have different gate voltages. When two source–drains are connected together, the gates of the corresponding devices must be connected together as well—as must their bulks. Thus, for transistors that are part of the same circuit, all the gates must be connected together, and all the bulks must be connected together. For the special case of subthreshold operation (Equation 6.2), there is no such restriction, because  $\mathcal{Q}_n$  does not depend on the gate voltage in this region. In both regions of operation, however, equivalence precludes us from connecting together the drain–source terminals of devices of different type, and from connecting channel terminals to gate terminals.

Equivalence implies a one-to-one relationship between node charge and node voltage. Such an invertible relationship requires that terminal charge,  $\mathcal{Q}_n(V_{GB_n}, V_i)$ , be a strictly monotonic function of terminal voltage,  $V_i$ . For the nMOS transistor, the mobile electrons decrease with increasing source or drain voltages. Hence,  $-\mathcal{Q}_n(V_{GB_n}, V_i)$  is a monotonically decreasing function. And, for a pMOS transistor, the mobile holes increase with increasing source voltage. Hence,  $\mathcal{Q}_p(V_{GB_n}, V_i)$  is a monotonically increasing function. These functions change either exponentially or quadratically with the voltage on the channel terminals.

Due to the expansive nature of  $\mathcal{Q}_n(V_{GB_n}, V_i)$ , we have

$$\begin{aligned} V_i \gg V_j &\Rightarrow |\mathcal{Q}_n(V_{GB_m}, V_i)| \ll |\mathcal{Q}_n(V_{GB_n}, V_j)| \\ &\Rightarrow \left| \frac{d\mathcal{Q}_n}{dV_i} \right| \ll \left| \frac{d\mathcal{Q}_n}{dV_j} \right|. \end{aligned}$$

The same relationships apply to the pMOS transistor when the signs of the latter's voltages are reversed. Hence, we have

$$\begin{aligned} V_i \gg V_j &\Rightarrow I_{i,j} \approx I_j \equiv -W_n \mathcal{P}(L_n) \mathcal{Q}_n(V_{GB_n}, V_j) \\ &\Rightarrow \left| \frac{dI_{i,j}}{dV_i} \right| \ll \left| \frac{dI_{i,j}}{dV_j} \right|. \end{aligned}$$

Therefore, when the voltage on node  $i$  is much larger than that on node  $j$ , the node charge at  $i$  becomes negligible, and the current asymptotically approaches  $I_j$ , the value of the component driven by node  $j$ .

In the dichotomous **ohmic–saturation** voltage-mode viewpoint, the device is said to enter a different regime of operation when the current becomes independent of the node voltage, called the **saturation region**. When the current is decomposed into source and drain components, however, there is no such dichotomy. For an nMOS transistor, we have

$$\begin{aligned} V_i > V_j + V_{\text{sat}} &\Rightarrow |I_i| \ll |I_j| \Rightarrow I_{i,j} \approx I_j, \\ V_j > V_i + V_{\text{sat}} &\Rightarrow |I_j| \ll |I_i| \Rightarrow I_{i,j} \approx I_i, \\ |V_i - V_j| < V_{\text{sat}} &\Rightarrow I_j \simeq I_i \Rightarrow I_{i,j} = I_i - I_j, \end{aligned}$$

where  $V_{\text{sat}} = 5V_{T_n}$  below threshold, and  $V_{\text{sat}} = V_{\text{GB}_n} - V_{\text{thr}}$  above threshold. The functional dependence of the current components,  $I_i$  and  $I_j$ , on the terminal voltage,  $V_i$  and  $V_j$ , is fixed, and remains the same throughout the ohmic and saturation regions. There is no dichotomy between these two regions from the current-component perspective: We split the current into two components, rather than split the voltage range into two regions. I prefer to split the current because this choice preserves the symmetry of the device, whereas splitting the voltage does not.

The linear dependence of the current on only one of the terminal charges in the saturation region gives us the capability to measure our fictitious terminal charges and node charges. Gaining access to the node charge is extremely important if we want to apply external inputs to the circuit, process them using the inherent linearity of the transistor, and read out the results from the circuit. Therefore, it is most convenient to use currents as our inputs and outputs, if we want to exploit the inherent linearity of the MOS transistor.

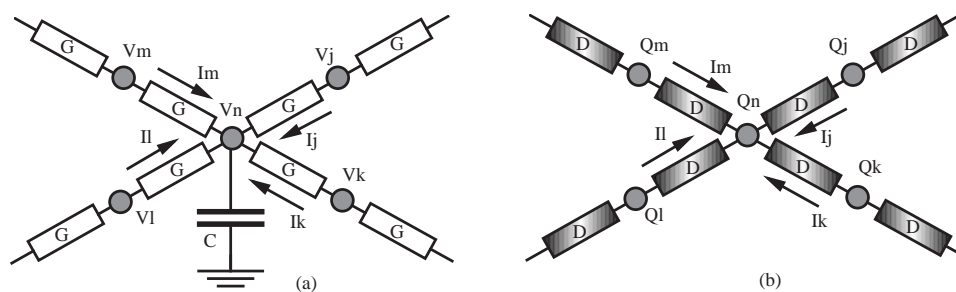


Figure 6.2: LOCAL AGGREGATION

(a) Aggregation using voltage–current linearity and conductances. (b) Aggregation using charge–current linearity and diffusors.

## 6.7 Diffusive Networks

**Local aggregation**—the linear summation of signals over a confined region of space—is a computation that occurs throughout the nervous system. A voltage-mode circuit that performs this extremely useful computation is described by Mead (Chapter 6 of [119]); he uses this computation in several examples of neuromorphic systems presented in the book. In this section, I present a more efficient current-mode technique for performing local aggregation. My technique exploits diffusion in subthreshold MOS devices, much as cell syncytia in the nervous system use diffusion to distribute and sum signals over a local neighborhood.

The diffusion of particles through a continuous medium—or of heat in a solid—is described by the following partial differential equation:

$$\frac{dc}{dt} = D\nabla^2 c(x, y), \quad (6.10)$$

where  $c(x, y)$  is the concentration profile over space—it is assumed to be uniform in the third ( $z$ ) dimension. Here,  $\nabla^2 \equiv \partial^2/\partial x^2 + \partial^2/\partial y^2$  is the Laplacian operator, and  $D$  is the diffusion coefficient. This equation is an application of Fick’s law, which governs diffusion, and of the continuity equation, which guarantees conservation, as we discussed in Sections 5.1.1 and 5.1.2.

The discrete networks shown in Figure 6.2 both simulate diffusion in a continuous

medium. The first network (Figure 6.2a) uses voltages, currents, and conductances; its node equation is

$$\frac{dV_n}{dt} = \frac{4G}{C} \left( \frac{V_j + V_k + V_l + V_m}{4} - V_n \right), \quad (6.11)$$

which is homologous with Equation 6.10. The term in parentheses is the second-difference approximation to the Laplacian, when distance is measured in units of the internode spacing, as we discussed in Section 4.1. This solution, however, is not amenable to VLSI integration, because we must expend large amounts of area and power to make the nonlinear conductances of transistors appear linear over a voltage range larger than a few thermal voltages.

The second network (Figure 6.2b) uses charges, currents, and diffusors; its node equation is

$$\frac{dQ_n}{dt} = 4D \left( \frac{Q_j + Q_k + Q_l + Q_m}{4} - Q_n \right). \quad (6.12)$$

Note that  $dQ_n/dt$  is the same as the current supplied to node  $n$  by the network. This solution is amenable to VLSI implementation. We can realize diffusion with diffusors—transistors operating below threshold, as described in Section 6.6. The diffusion coefficient  $D$  is related to the diffusivity,  $\mathcal{D}$ , and to the partition coefficient,  $\mathcal{K}(V_{\text{GB}})$ , of the diffusors by

$$D = W\mathcal{K}(V_{\text{GB}})\mathcal{D}.$$

In both of these networks, we can set up the boundary conditions by injecting current into the appropriate nodes. In the voltage-mode network, the solution is the node voltages, and we can read these voltages without disturbing the network. In the current-mode network, however, the solution is the node charges, and these fictitious charges are not directly accessible. We can infer the node charge from the node voltages  $V_i$  if we have an accurate description of  $\mathcal{Q}(V_i)$ . In practice, we can use a transistor to compute  $\mathcal{Q}(V_i)$ , by tying it to the node in question, and operating it in saturation so that it passes a current proportional to the node charge, as we discussed in Section 6.6. Unfortunately, this approach draws current from the node, and the

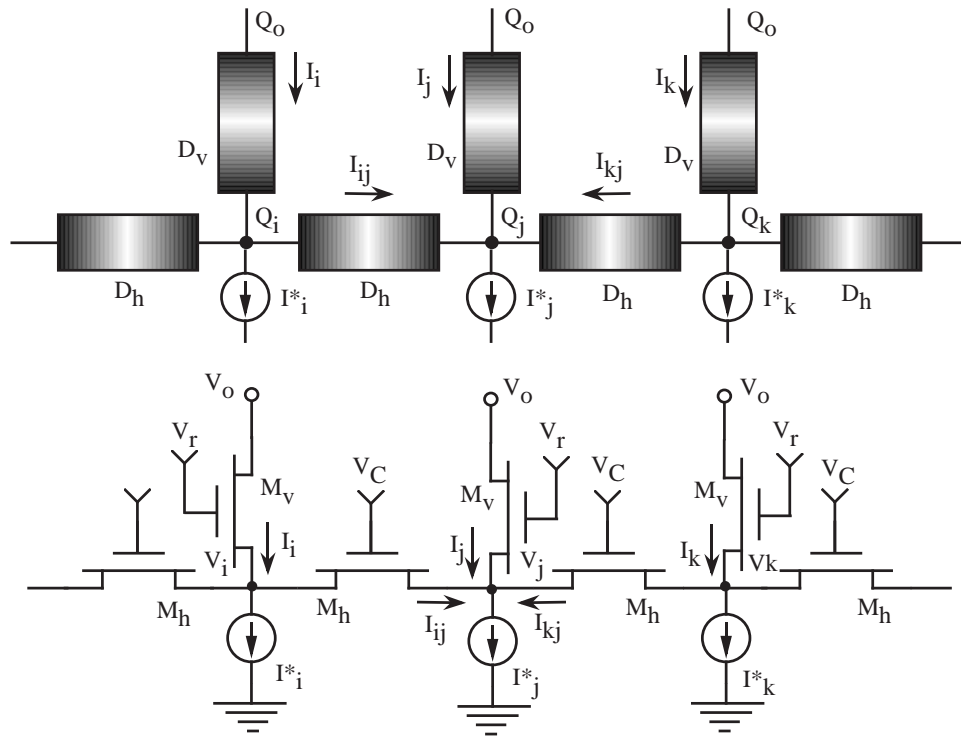


Figure 6.3: CELL-SYNCYTIA CIRCUIT MODEL

The lateral diffusers model gap junctions between cells; the vertical diffusers model the membrane leakage. (a) Schematic representation. (b) MOS transistor implementation.

permeability of the added device must be extremely small so that the disturbance is negligible.

Biological diffusive media, such as cell syncytia, are hardly ever isotropic (i.e.,  $D$  varies from place to place). Nerve cells make gap junctions of varying area, and neuromodulators such as dopamine can vary the pore permeability. Thus, nerve cells can control actively the permeability of membranes between them and neighboring cells or the extracellular fluid. The dependence of the diffuser's partition coefficient on its gate voltage (see Equation 6.2) gives us the ability to control permeability locally in our circuit model of the diffusion network.

We can add a loss term to the diffusion equation to model the sequestering of

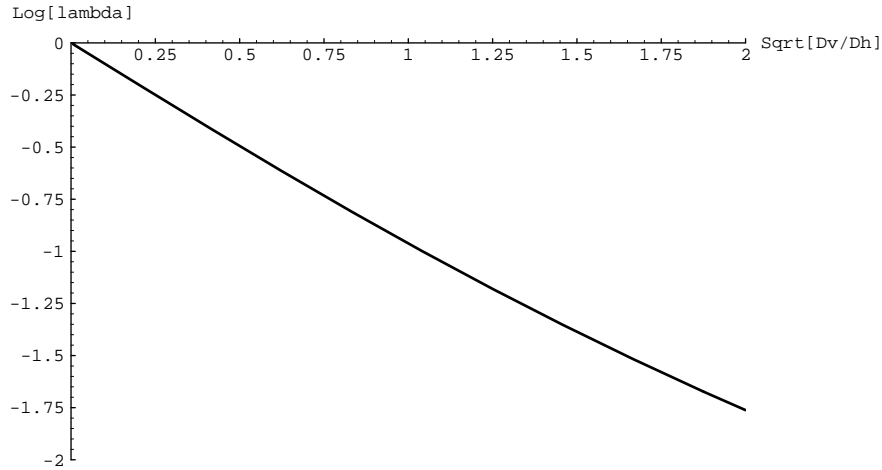


Figure 6.4: CONCENTRATION DECAY RATES FOR DIFFUSION

The exponential decay rate in the discrete network,  $\log(\lambda)$ , is plotted versus that in a continuous medium,  $1/L = \sqrt{D_v/D_h}$ , with the same values of vertical and horizontal diffusivities,  $D_v$  and  $D_h$ . The decay rates are nearly equal for low decay rates—the error is 12.5 percent when the charge density decays by a factor of  $e^2$  each time that  $x$  changes by  $L$ .

particles by buffers or the leakage of particles out of the diffusive medium:

$$\frac{dc}{dt} = D_h \nabla^2 c(x, y) - D_v c(x, y). \quad (6.13)$$

This version of the diffusion equation is realized by the diffusor network shown in Figure 6.3, for the one-dimensional case. Its node equation is

$$I_j^* = I_{ij} + I_{kj} + I_j = \mathcal{K}(V_c) \mathcal{D}(Q_i + Q_k - 2Q_j) - \mathcal{K}(V_r) \mathcal{D}(Q_j - Q_0). \quad (6.14)$$

When  $V_o > \max_{vi}(V_i + V_{\text{sat}})$ ,  $Q_0$  is negligible and the node equation becomes the discrete analog of Equation 6.13. In this case, we can read out the node charges simply by monitoring the currents at the drains of the vertical elements as these currents flow into the voltage source applying  $V_o$ .

The discrete version of the lossy-diffusion equation admits solutions of the form

$Q_i = \lambda^i$ , where

$$\lambda = 1 + \frac{D_v}{2D_h} \left( 1 - \sqrt{1 + \frac{4D_h}{D_v}} \right),$$

and  $D_v \equiv \mathcal{DK}(V_r)$ , and  $D_h \equiv \mathcal{DK}(V_c)$ . The solution to the continuous lossy-diffusion differential equation in one dimension has the form  $Q(x) = \exp(-x/L)$ , where

$$L = \sqrt{\frac{D_h}{D_v}}.$$

Thus, the discrete simulation reproduces the exponential form of the decay; its decay rate matches that of the continuous medium when  $\log(\lambda) \approx 1/L$ . These quantities are plotted against each other in Figure 6.4; the slope is close to unity for low decay rates.

Using the exponential dependence of the partition coefficients on the gate voltages, we can relate the space constant to the biases voltages  $V_r$  and  $V_c$  applied to the diffusers:

$$L \approx \exp\left(\frac{\kappa_n(V_c - V_r)}{2V_T}\right).$$

It becomes obvious why the ratio has an exponential dependence on the voltage difference between  $V_c$  and  $V_r$  if you observe that  $M_h$  and  $M_v$  constitute a differential pair operating in subthreshold. These devices act as a current-divider for current driven by the charge at their common node. The divider ratio is set by their effective widths, which depend on the geometrical width as well as on the surface potential. Here, we have used the  $\kappa$  approximation to relate the surface potential to the gate-bulk voltage. The surface potential is constant as long as the gate and bulk voltages are fixed—assuming that the mobile charge is negligible. Therefore, the divider ratio is constant, and linear division occurs. However, as we enter the transition and above-threshold regions, this assumption fails, and the surface potential starts to follow the source voltage. Consequently, the divider ratio is no longer independent of the current level. This variation of the divider ratio limits the dynamic range of the current divider, and hence the linear operating range of the diffuser network.

The diffuser network is a particularly attractive circuit for implementing local

aggregation because of the area efficiency that we realize by using a single device to model a linear element, the power efficiency that we obtain by operating with subthreshold currents, and the enhanced functionality available with electronically adjustable coupling strength.

The diffusive network in Figure 6.3 has been described in terms of pseudoconductances [127]. I prefer the charge-based formulation using diffusors, originally proposed in [124] and elaborated in [129, 130], because of the physically accurate intuition that it provides. The essence of this approach is the representation of variables and parameters by charge, current, and diffusivity—voltages and conductances are not used explicitly.

Bult and Geelen proposed an identical network for linear current division above threshold, and used it in a digitally controlled attenuator [126]; they also analyzed its subthreshold behaviour. However, they stipulated that all gate voltages must be identical, and controlled the division by manipulating the geometrical factor  $W/L$  of the devices. I showed here, and previously in [124], that this constraint can be relaxed in subthreshold without disrupting linear operation. This flexibility is a real bonus, because it allows us to modify the divider ratio or space constant of the network after the chip is fabricated by varying  $V_c - V_r$ . Tartagni and colleagues have demonstrated a current-mode centroid network [131] using subthreshold MOS devices whose operation is described by the diffusors discussed here.

## 6.8 Test Results

In this section, I present results from experiments designed to demonstrate the linearity of diffusor circuits, and to measure the dependence of diffusivity on the gate voltage, the dynamic range of operation, and the spread of currents in diffusor networks.

The circuit designed to test the functionality of the current divider and the diffusor is shown in Figure 6.5. The measurements obtained from this test circuit are shown in Figures 6.6 and 6.7. These measurements demonstrate the linearity of the current

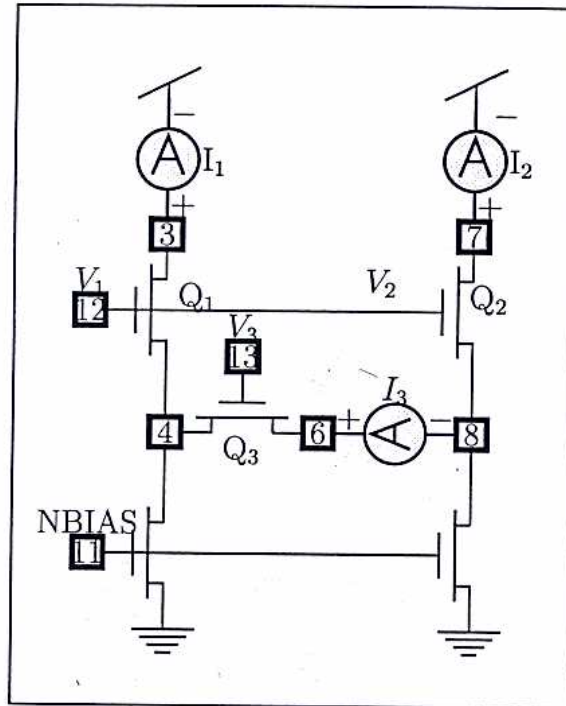


Figure 6.5: CURRENT DIVIDER AND DIFFUSOR TEST CIRCUIT

Using this test circuit, I measured current division in the differential pair formed by transistors  $Q_1$  and  $Q_2$  when  $Q_3$  is shorted, and the linear dependence of the current,  $I_3$ , in transistor  $Q_3$  on the current differential,  $I_1 - I_2$ , between transistors  $Q_1$  and  $Q_2$ . I have plotted the current-divider measurements, taken for various voltage differentials ( $V_1 - V_2$ ), in Figure 6.6, and the diffusor measurements, taken for various voltage differentials ( $V_R - V_G$ , where  $V_R = V_3$ ,  $V_G = V_1 = V_2$ ), in Figure 6.7.

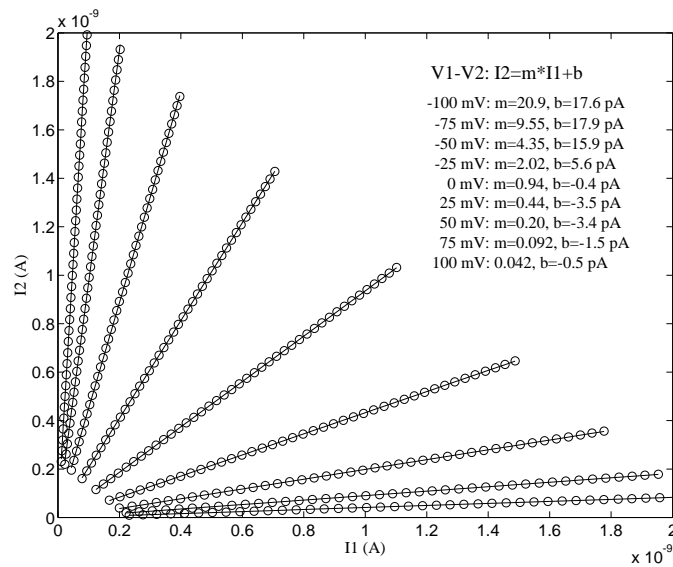


Figure 6.6: CURRENT-DIVIDER CURRENTS

These measurements, obtained from the test circuit shown in Figure 6.5, demonstrate that the differential pair splits its tail current between its two arms according to a fixed ratio, and this ratio depends on the voltage differential. The ratios  $I_1/I_2$  obtained from the slopes of these curves are plotted against the voltage differentials  $V_1 - V_2$  in Figure 6.8.

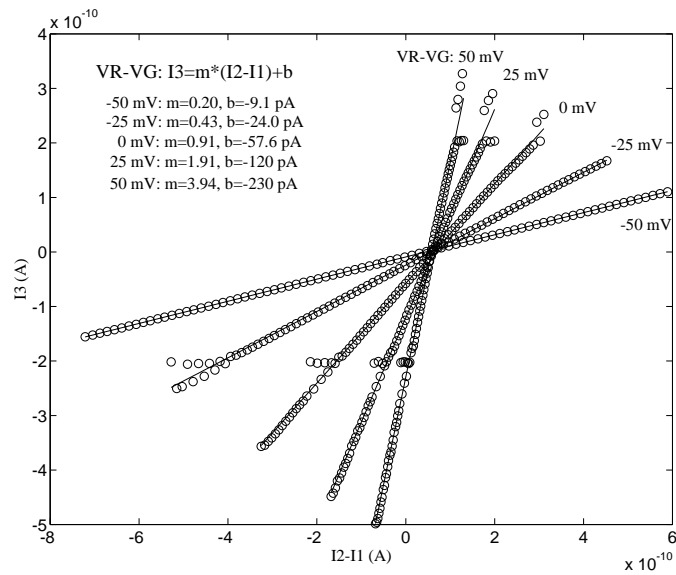


Figure 6.7: DIFFUSOR CURRENT VERSUS CURRENT-DIFFERENCE

These measurements, obtained from the test circuit shown in Figure 6.5, demonstrate that the current in the horizontal diffuser is directly proportional to the current differential in the vertical diffusers, and the ratio  $I_3/(I_1 - I_2)$  depends on the voltage differential  $VR - VG$  between the horizontal and vertical diffusers. The ratios obtained from the slopes of these curves are plotted against the voltage differential in Figure 6.9.

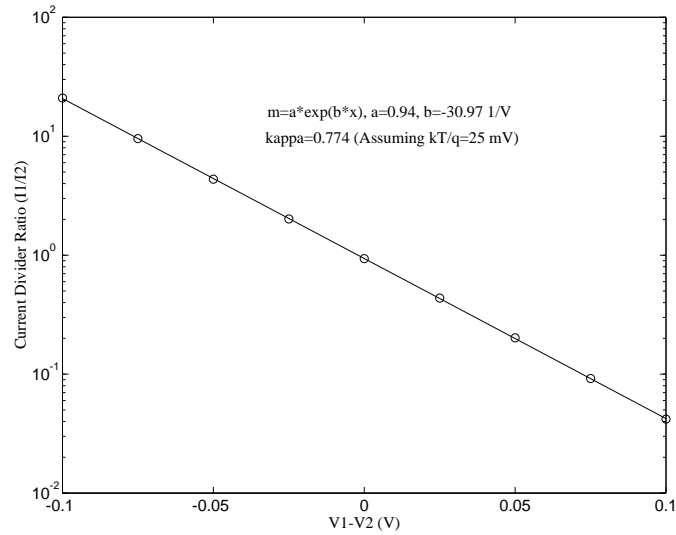


Figure 6.8: CURRENT-DIVIDER RATIO VERSUS VOLTAGE DIFFERENTIAL

These measurements demonstrate an exponential dependence of the current-divider ratio on the voltage differential, just as predicted by the theory.

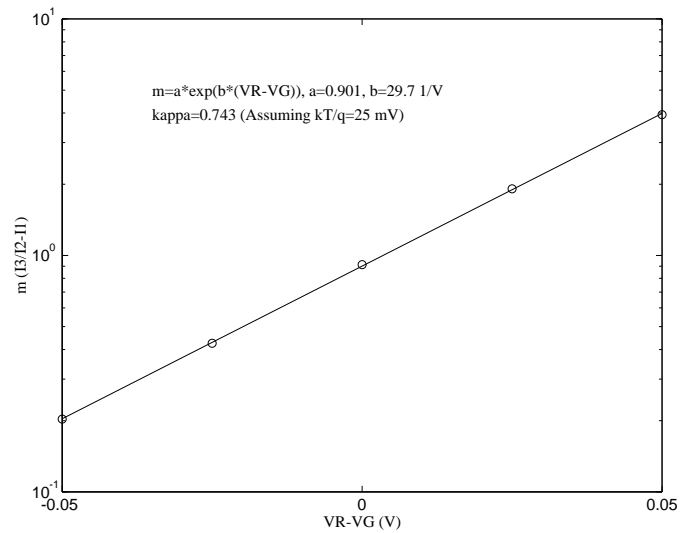


Figure 6.9: DIFFUSOR PERMEABILITY VERSUS GATE VOLTAGE

These measurements demonstrate an exponential dependence of permeability on gate voltage, just as predicted by the theory.

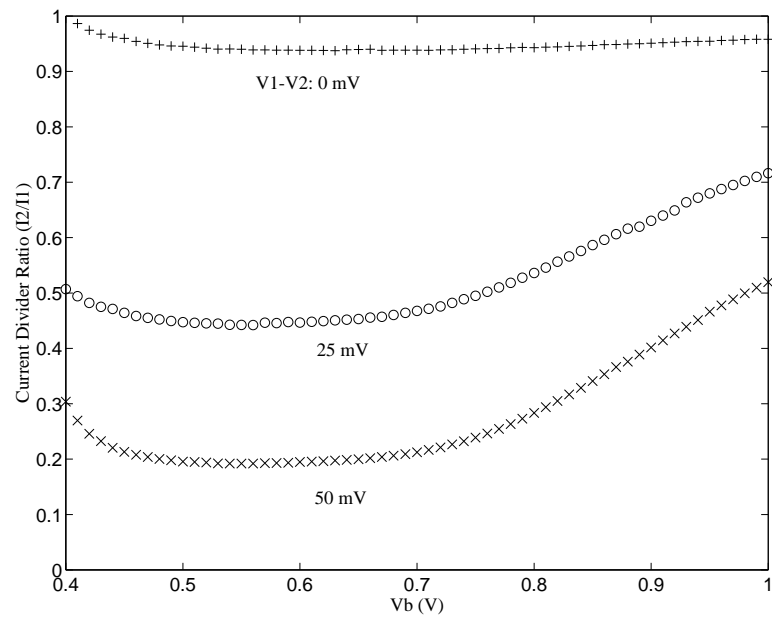


Figure 6.10: DYNAMIC RANGE OF CURRENT DIVIDER CIRCUIT

These measurements, which I took by sweeping the voltage applied to pin NBIAS of the test circuit shown in Figure 6.5, and taking the ratio between the currents in  $Q_1$  and  $Q_2$ , with  $Q_3$  shorted, for various voltage differentials  $V_1 - V_2$ , show the limited range of operation of current divider, and diffusor, circuits.