Proceedings MicroNeuro '97, Dresden, pp. 163-173

Pseudo-Resistive Networks and Their Applications to Analog Collective Computation

E.A.Vittoz
CSEM Centre Suisse d'Electronique et de Microtechnique SA
Jaquet-Droz 1, CH-2007 Neuchâtel, Switzerland.
E-mail: eric.vittoz@csem.ch

Abstract

The basic concept of using a standard MOS transistor as a pseudo-conductance is explained. It offers the possibility to implement any network of linear resistors by means of transistors only, and to control the value of each of these pseudo-resistors by a voltage or a current. Applications to linear attenuation, moment computation, diffusion networks, fuzzy logic computation and emulation of physical media are described. Several examples are given, including D/A conversion, elementary shape recognition, path-finding, place coding and emulation of biological organs such as the retina and the cochlea.

1 Introduction

Analog VLSI is believed to be an ideal medium for implementing functions which require the collective processing of data in massively parallel systems [1]. These functions, of perceptive or evaluative nature, include those of biologically inspired neural networks and fuzzy computation [2]. The precision of the data and that of the operators is replaced by their large number, and therefore analog approaches are expected to be more efficient in power and in chip area [3]. Moreover, analog circuits allow a better opportunistic exploitation of all the properties of available components. The following sections focus on one particular property of the standard MOS transistor, which opens the way to completely novel architectures. This property, which was only clearly identified some 30 years after the advent of working MOS transistors, is only understandable if the device is adequately modelled. Section 2 will recall such a model and explain the general principle, whereas section 3 will focus on the particularly useful case of weak inversion (or subthreshold) operation. The subsequent sections will introduce several domains of application, and examples thereof.

2 General principle

As shown by the schematic representation of Fig. 1, the MOS transistor is essentially symmetrical. The source and drain ends are in principle not distinguishable and are here labelled as the two terminals A and B of the channel, with potentials V_A and V_B with respect to the local substrate (general substrate of the chip, or local well).

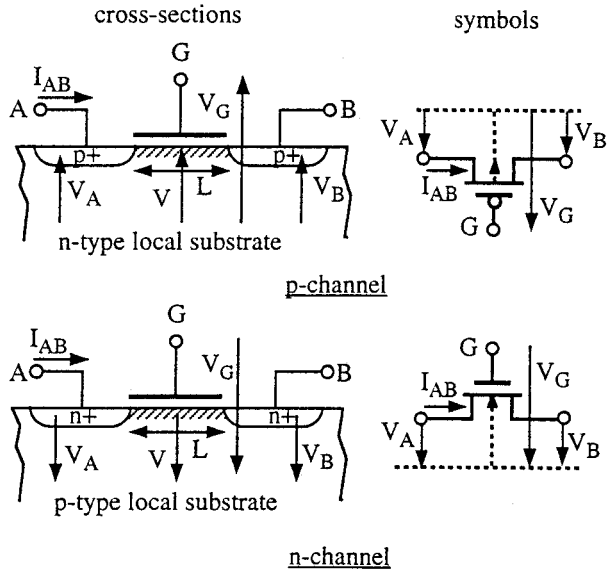


Fig. 1: Cross-section, symbol and definitions of n- and p-channel transistors.

V is the local value of channel "potential". It is not the electrostatic potential but a measure of the disequilibrium

in the distribution of electrons or holes in the channel that is produced by the application of voltages V_A and/or V_B (quasi-Fermi potential of electrons or holes). It has the value V_A and V_B at the two respective ends of the channel.

For a given value of gate voltage, the local sheet conductivity g_S of the channel is a decreasing function of the channel potential V as illustrated by Fig. 2 [4].

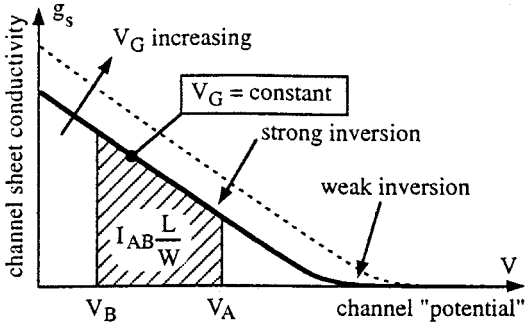


Fig. 2: Sheet conductivity of the channel and current through the transistor for constant gate voltage V_G.

This function decreases approximately linearly with V for large values of g_S (strong inversion of the channel) and exponentially for low values of g_S (weak inversion). It can be shown [4] that the current flowing through a n-channel transistor is simply given by

$$I_{AB} = \frac{W}{L}\int_{V_B}^{V_A} g_S(V_G, V)\, dV =$$

$$\frac{W}{L}\left[\int_{V_B}^{\infty} g_S(V_G, V)\, dV - \int_{V_A}^{\infty} g_S(V_G, V)\, dV\right] \quad (1)$$

Due to the particular definition of V, this equation includes the two possible mechanisms for transport of charge carriers: conduction (which dominates in strong inversion) and diffusion (which dominates in weak inversion). It is therefore valid for any positive value of V_B and V_A.

The decomposition into two terms is possible since g_S tends to 0 for V large. It provides a symmetrical expression with respect to V_A and V_B, which can be written

$$I_{AB} = \pm I_S\,[f(V_G, V_A) - f(V_G, V_B)] \quad (2)$$

(+ for p-channel, - for n-channel)

where I_S is a specific current proportional to the width-to-length ratio W/L of the transistor. The transistor is said to be saturated when the smaller of the two terms of (2) becomes negligible (the area under the curve of Fig. 2 becomes independent of the larger of V_A or V_B). Channel shortening degrades the precision of (2) by making I_S itself slightly dependent on V_A and/or V_B. The relation is no more applicable if the channel length is reduced below the short channel limit.

When needed, the function $f(V_G, V)$ can be approximated by [4]

$$f(V_G, V) = \ln^2\left(1 + \exp\frac{(V_G - V_{T0})/n - V}{2U_T}\right) \quad (3)$$

$$\text{with} \quad I_S = 2n\frac{W}{L}\mu C_{ox}U_T^2 \quad (4)$$

where V_{T0} is the gate threshold voltage for V=0 (process parameter), $U_T = kT/q$ is the thermal voltage of about 26mV at room temperature T, μ is the mobility of electrons or holes in the channel and C_{ox} is the gate oxide capacitance per unit area. The slope factor n ranges from 1.4 to 1.8 for V_G close to V_{T0}. It slowly tends to 1 when V_G tends to infinity.

By defining a *pseudo-voltage* V^* [5] given by

$$V^* = \pm V_0 f(V_G, V) \quad (+ \text{ for } p\text{-ch}, - \text{ for } n\text{-ch}) \quad (5)$$

where V_0 is an arbitrary positive scaling voltage, (2) can be rewritten as:

$$I_{AB} = G^*(V_A^* - V_B^*) \quad (6)$$

which corresponds to a linear *pseudo-Ohm's* law, with constant *pseudo-conductance* $G^* = I_S/V_0$, proportional to W/L according to (4). The pseudo-voltage V^* is always positive for a p-channel transistor (negative for a n-channel). Moreover, it tends to 0 for V large. Thus, the *pseudo-ground* 0* (0-reference for the pseudo-voltage V^*) is obtained by imposing $V = V_{pg}$ large enough to make $f(V_G, V_{pg})$ negligible.

As consequence of (6), a current I imposed through two transistors T_1, T_2 connected in parallel splits linearly [6] into two components I_1, I_2 respectively proportional to G_1^*, G_2^*. This is identical to the current splitting through two linear conductances G_1, G_2, as illustrated in Fig. 3.

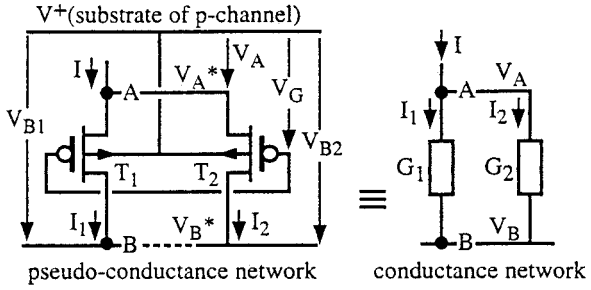


Fig. 3: Linear splitting of current I.

If V_B =0 (ground potential) in the conductance network, V_B* in the pseudo-conductance network should be at pseudo-ground 0*. Voltages V_{B1} and V_{B2} need not be fixed or equal anymore; they must just be sufficiently large to make $f(V_G, V_{Bi})$ negligible. This means that the transistors must be saturated.

This mapping of a resistive network into its transistor equivalent on the basis of (6) can be generalized: any arbitrary network of linear resistors can be implemented by replacing each resistor R_i by a transistor T_i, with all transistors in the *same substrate* [5]. If all gate voltages V_{Gi} are the same, the transistor network behaves linearly with respect to currents. The value of each pseudo-conductance G_i* can be made proportional to the corresponding conductance G_i=1/R_i by adjusting the ratio W/L of transistor T_i. As long as only currents (split by the network) are imposed, the absolute value of each G_i* (and thus of V_0) is irrelevant.

Since the pseudo-ground level is obtained for any value of V_{pg} large enough, any current flowing to the pseudo-ground can be easily extracted by means of a current mirror made of transistors complementary to those of the network. This is not possible in the corresponding conductance network.

3 Particular case of weak inversion

For a given value of V_G, a transistor is in weak inversion if both V_A and V_B are large enough to obtain f (V_G, V) « 1 at both ends of its channel [4]. This corresponds to ensuring a value of its saturation current much smaller than the specific current I_S. If V_G<V_{T0}, this is already the case when the smaller of V_A and V_B is equal to zero. Equation (3) then reduces to

$$f(V_G, V) = \exp\left(\frac{V_G - V_{T0}}{nU_T}\right)\exp\left(-\frac{V}{U_T}\right) \ll 1 \qquad (7)$$

and is *separable* in exponential functions of V_G and V. Pseudo-voltage and pseudo-conductance may then be redefined as

$$V^* = \pm V_0 \exp\left(-\frac{V}{U_T}\right) \quad (+ \text{ for } p\text{-ch}, - \text{ for } n\text{-ch}) \qquad (8)$$

$$G^* = \frac{I_S}{V_0}\exp\left(\frac{V_G - V_{T0}}{nU_T}\right) \qquad (9)$$

The linear pseudo-Ohm's law (6) is still valid, but the pseudo-conductance G^* of each transistor is now controllable independently by the value of its gate voltage V_G. The linearity of currents is available in the whole range of weak inversion, which may correspond to 3 to 6 orders of magnitude. The corresponding range of real voltage V is only 7 to 14U_T (175 to 350mV).

Each pseudo-conductance may alternatively be controlled by a current by means of a control transistor, as shown in Fig. 4.

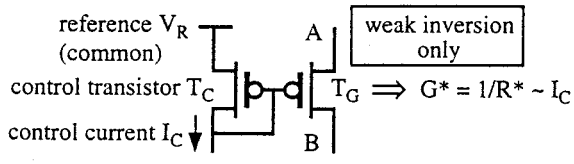


Fig. 4: Current-controlled pseudo-conductance.

All transistors are in the same substrate and the reference voltage V_R is common to all the control transistors of the network. Several pseudo-conductances can be controlled by the same current by repeating transistor T_G.

Except when specifically indicated, it will be assumed that the transistors of the same network are all of a size (thus have same value of I_S). Each value of $G_i^* = 1/R_i^*$ is then proportional to the corresponding control current I_{Ci}. However, *precision is limited* by the uncorrelated mismatches δV_T and δI_S of threshold voltages and specific currents of transistors T_C and T_G:

$$\frac{\delta G^*}{G^*} = \frac{\delta I_S}{I_S} - \frac{\delta V_T}{nU_T} \qquad (10)$$

The RMS value of $\delta I_S/I_S$ is usually no more than a few percents, but that of δV_T may be as high as 5mV, corresponding to a very large error of about 12% RMS.

As a first simple example of application, Fig. 5 shows again the implementation of two parallel controlled-conductances.

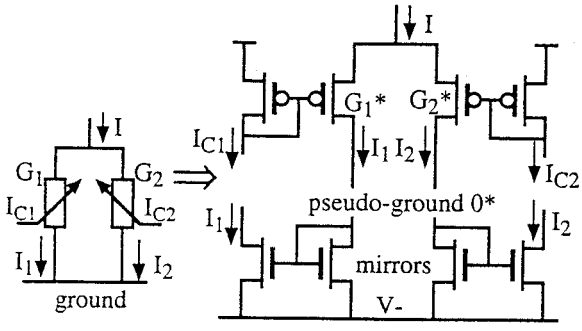
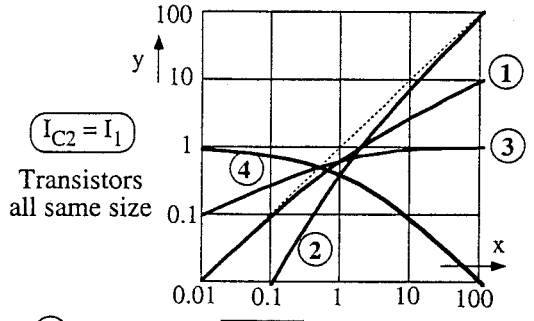


Fig. 5: Parallel controlled-conductances.

The input current is split proportionally to the control currents I_{Ci}, thus:

$$\frac{I_1}{I} = \frac{1}{1+G_2/G_1} = \frac{1}{1+G_2^*/G_1^*} = \frac{1}{1+I_{C2}/I_{C1}} \qquad (11)$$

As shown in the figure, the output currents I_1 and I_2 can be extracted at the pseudo-ground by current mirrors. If needed, they may then be used as control currents of controlled-conductances. As an example, Fig. 6 shows the various functions which can be obtained by imposing $I_{C2} = I_1$ (right-hand conductance proportional to left-hand current).



① $y = -0.5 + \sqrt{x + 0.25}$ for $x = I/I_{C1}$, $y = I_1/I_{C1}$

② $y = x + 0.5 - \sqrt{x + 0.25}$ for $x = I/I_{C1}$, $y = I_2/I_{C1}$

③ $y = -x/2 + \sqrt{x + x^2/4}$ for $x = I_{C1}/I$, $y = I_1/I$

④ $y = 1 + x/2 - \sqrt{x + x^2/4}$ for $x = I_{C1}/I$, $y = I_2/I$

Fig. 6: Parallel conductances with feedback.

This simple example already illustrates the potential of such networks of controllable pseudo-conductances for processing currents.

The linear properties of MOS transistors in weak inversion can be traced back to the linear behaviour of diffusing carriers [7]. They could be exploited with bipolar transistors as well, provided the base current in direct and reverse modes can be neglected. In principle, *pnp* (*npn*) bipolar transistors could be mixed with *p*-ch. (*n*-ch.) MOS transistors in such circuits, but adjustments would be needed to control the relative value of bipolar and MOS pseudo-conductances (no intrinsic matching).

The concept of pseudo-conductance shares some characteristics with the well-known principle of translinear circuits [8]. Indeed, some particular circuits can be analyzed by either of these approaches.

4 Linear attenuators

The parallel controlled-conductances of Fig. 5 can be used as a current attenuator according to (11). As an example, Fig. 7 shows a learning synapse intended for a continuous-operation Kohonen map [9].

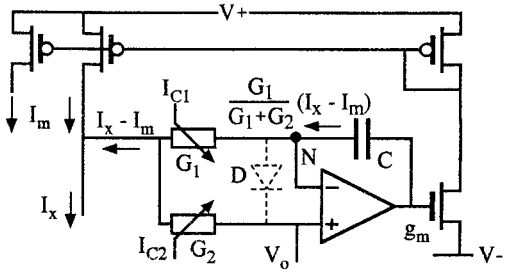


Fig. 7: Learning synapse of Kohonen map. *

The attenuator G_1-G_2 is used to reduce the current flowing into the integrating capacitor C. The synaptic component I_m is slowly tending to the input value I_x with time constant

$$\tau = \frac{G_1+G_2}{G_1}\frac{C}{g_m} = \frac{I_{C1}+I_{C2}}{I_{C1}}\frac{C}{g_m} \qquad (12)$$

where g_m is the transconductance of the *n*-ch. transistors. The time constant τ can be made very large if $I_{C2} \gg I_{C1}$. The well of the *p*-ch. transistors implementing G_1 and G_2 is connected to virtual-ground voltage V_0. Thus, the only junction leaking from critical node N (shown as diode D) has no voltage across it and the very small current delivered through G_1 is not affected. The input voltage can be kept close to V_0 by choosing $I_{C2} \gg I_x$ (small voltage across G_2, corresponding transistor in conduction).

* Error in loop sign in the original publication, corrected here

Connecting m identical conductances in parallel offers the possibility to avoid current mirrors to create m replicas of a current. Each of them is of course attenuated by the same factor m.

Current splitting can be used to produce binary weighted currents by means of the well-known R-$2R$ network. A digital-to-analog converter based on this principle is shown in Fig. 8 (for 3 bits as an example) [10]. The resistors and the switches of the original resistive version are replaced by transistors only in the pseudo-resistance version; in this version, each switch is even used to implement half of the pseudo-resistance $2R^*$.

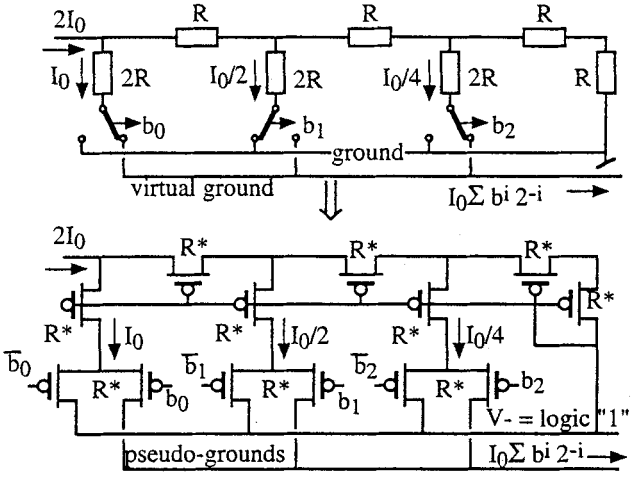


Fig. 8: R-2R digital-to-analog converter (3 bits).

For best precision, the transistors should be operated in strong inversion, which is possible since all gate voltages of non-blocked transistors have the same value V-. This dense circuit may be used to convert a digitally stored analog values (for example a synaptic weight) with a precision up to 6 bits.

5 Moment computation

The computation of moments in the 1-dimension space can be obtained by a resistive line of length 1 and resistivity $\rho(x)$ per unit length as shown in Fig. 9A.

The input signal is a current density $J(x)$ injected along this line and the result is carried by the currents I_a and I_b flowing out of the two grounded ends of the line. The value of right-hand current I_b is obtained by integrating the elementary contributions of current $J(x)dx$ injected and split at position x:

$$I_b = \frac{1}{R}\int_{x=0}^{1} R_x J(x)dx \tag{13}$$

$$\text{with} \quad R_x = \int_{x=0}^{x} \rho(x)dx \quad \text{and} \quad R = \int_{x=0}^{1} \rho(x)dx \tag{14}$$

$$\text{If} \quad \rho(x) = k_n x^{n-1} \tag{15}$$

Then (13) and (14) result in

$$I_b = \int_{x=0}^{1} x^n J(x)dx = M_x n \tag{16}$$

which is the n^{th} order moment of distribution $J(x)$ with respect to the origin. In particular, the center of gravity of the distribution is obtained with a constant value of $\rho(x)$:

$$x_0 = \frac{M_x}{M_0} = \int_{x=0}^{1} xJ(x)dx \;/\; \int_{x=0}^{1} J(x)dx = \frac{I_b}{I_a + I_b} \tag{17}$$

The continuous line may be approximated by a network of N series connected resistors ΔR_j as illustrated in Fig. 9B.

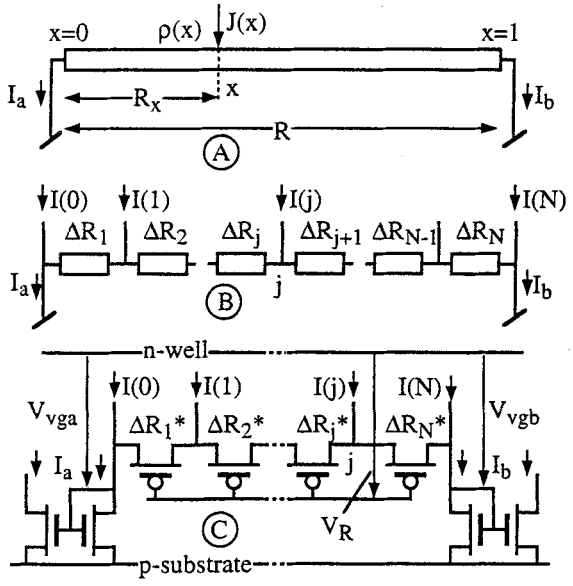


Fig. 9: 1-dimensional moment computation.

The injected current density $J(x)$ is then replaced by a series of current sources $I(j)$ at each node j. All ΔR_j are identical to compute the first order moment [11]. Their value increases with the $(n-1)^{th}$ power of their index for the n^{th} order moment. In particular, for the 2^{nd} order moment $M_x 2$, it increases linearly (ΔR, $2\Delta R$, $3\Delta R$, $4\Delta R$,...); (16) is then replaced by:

$$I_b = \sum_{j=0}^{N} \frac{j(1+j)I(j)}{N(1+N)} \cong \sum_{j=0}^{N} (j/N)^2 I(j) = M_x 2 \qquad (18)$$

Each resistor may in turn be replaced by a transistor used as a pseudo-resistor with a common value of gate voltage V_R, as shown in Fig. 9C. The value of each ΔR_j may be adjusted by the length-to-width ratio L/W of the corresponding transistor or, for better accuracy, by connecting several unit transistors in series. The output currents I_a and I_b are collected by two n-channel current mirrors. The pseudo-ground is maintained as long as the first and the last p-channel transistors are kept saturated, which, in strong inversion, requires

$$V_{vga} \text{ and } V_{vgb} > (V_G - V_{T0})/n \qquad (19)$$

In the 2-dimensional case, moments $M_x n$ and $M_y m$ with respect to the y and x co-ordinate axes can be computed separately by creating two copies of each current $I(j,k)$ and by summing separately the currents of column j and the currents of row k. These sums are then injected at node j, respectively k, of two separate 1-dimensional resistive (or pseudo-resistive) networks of Fig. 9 (of $(n\text{-}1)^{th}$ and $(m\text{-}1)^{th}$ power). The same approach is applicable to the computation of $M_r n$ and $M_\theta m$ in polar co-ordinates [12].

Separate calculation with respect to each axis is not possible for moments

$$M_x{}^n{}_y{}^m = \int_x \int_y x^n y^m J(x,y)\,dx\,dy \qquad (20)$$

which would require one computation per node. The particular and useful case of M_{xy} can be treated by dimensional reduction of data possible by means of the simple uniform resistive grid of Fig. 10 [13].

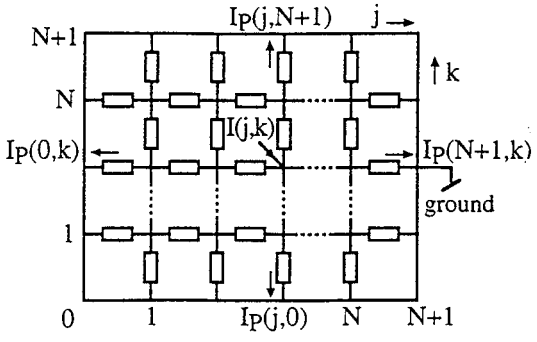


Fig. 10: Resistive grid for dimensional reduction of data.

Input currents $I(j,k)$ are injected at each internal node. The perimeter P is grounded and the currents flowing to ground are $I_P(j,k)$ with j or k equal to 0 or $N+1$. It can be shown [13] that

$$\sum_j \sum_k f(j,k) I(j,k) = \sum_P f(j,k) I_P(j,k) \qquad (21)$$

if the weighting function $f(j,k)$ is harmonic (its Laplacian is identically 0). This yields, for $f(j,k)=jk$:

$$M_{xy} = \sum_P jk I_P(j,k) =$$
$$(N+1)\left[\sum_j j I_P(j,N+1) + \sum_k k I_P(N+1,k)\right] \qquad (22)$$

which is the sum of first order moments.

This approach has been used to build a single-chip image sensor computing the orientation (and the position) of an object [14]. All resistors can be replaced by pseudo-resistors, and the perimeter can be connected to a pseudo-ground to facilitate the collection of currents.

6 Diffusion networks

A 2-dimensional diffusion network is shown in Fig. 11.

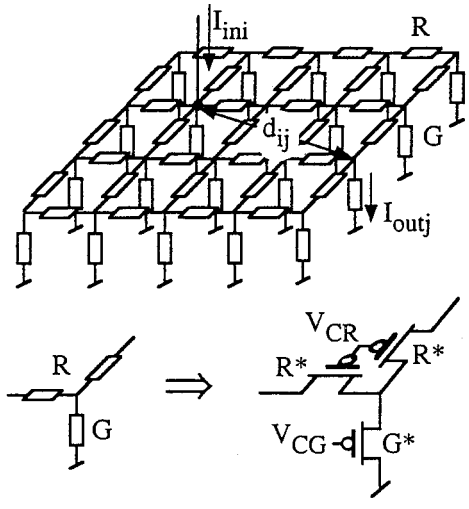


Fig. 11: 2-dimensional diffusion network.

It is a large array of lateral resistors R and vertical grounded conductances G. The current I_{outj} flowing to ground at node j as the result of a current I_{ini} injected into node i at distance d_{ij} (measured in pitch of the grid) decays with this distance according to the approximate law [15]:

$$\frac{I_{outj}}{I_{ini}} \sim \frac{\exp(-d_{ij}/L)}{\sqrt{d_{ij}/L}} \qquad \text{for } d_{ij} > L \tag{23}$$

with a diffusion length L (measured in pitch) given by:

$$L = 1/\sqrt{RG} \tag{24}$$

Since the network is linear, the effects of currents injected into all the nodes are superimposed and the network behaves essentially as a low-pass spatial filter.

Here again resistors and conductances can be replaced by transistors as illustrated in the same figure. If operation is maintained in weak inversion, the transistors implementing R^* and G^* can be controlled by different gate voltages V_{CR} and V_{CG}, and the diffusion length L can be electrically adjusted:

$$L = 1/\sqrt{R^*G^*} = \exp\left(\frac{V_{CR} - V_{CG}}{nU_T}\right) = \sqrt{I_{CR}/I_{CG}} \tag{25}$$

where I_{CR} and I_{CG} are the control currents of R^* and G^* according to Fig. 4.

The network can be hexagonal (or higher order polygonal) instead of rectangular for a better isotropy. The anisotropy may be electrically controlled by having different gate voltages for horizontal pseudo-resistors of different orientations. It can of course also be simply 1-dimensional.

A first classical example of application is a silicon retina providing edge enhancement by spatial differentiation in an isotropic hexagonal network [16]. Fig. 12 shows a possible implementation of one cell k of such a network using pseudo-conductances [5].

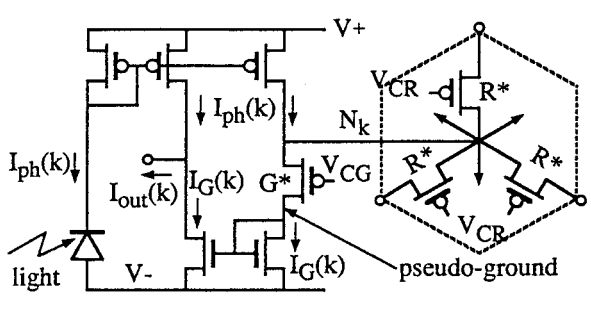


Fig. 12: Cell of an edge-enhancing silicon retina.

The light received on the light-sensing diode of the cell (pixel) produces a photo current $I_{ph}(k)$, a copy of which is injected into the local node N_k of a R^*G^* diffusion network. The current $I_G(k)$ flowing through pseudo-conductance G^* carries the low-pass filtered local contribution to the image, and is collected at a pseudo-ground. A mirrored copy of $I_G(k)$ is subtracted from a second copy of $I_{ph}(k)$ (non-filtered image) to produce the high-pass filtered local contribution $I_{out}(k)$.

The principle of another retina providing adaptation to local light intensity is illustrated in Fig. 13 [17].

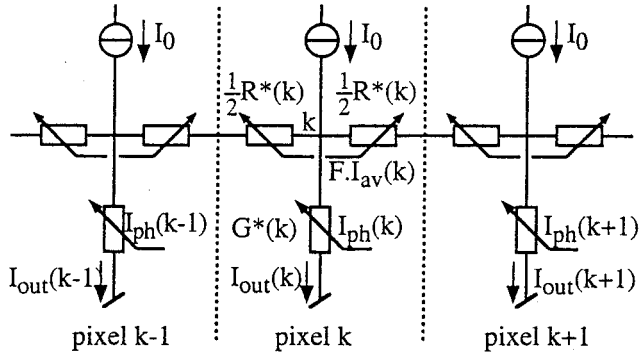


Fig. 13: Retina for local adaptation.

It is a 2-dimensional diffusion network (figure shown for 1-dimensional) in which each vertical pseudo-conductance $G^*(k)$ is controlled by the local photo-current $I_{ph}(k)$. A constant current I_0 is injected into each node and is distributed across an equivalent diffusion area A proportional to the square of the diffusion length L given by (25). The local output current $I_{out}(k)$ flowing through $G^*(k)$ is thus given by

$$I_{out}(k) = \frac{G^*(k)}{G_A{}^*(k)} I_0 = \frac{I_{ph}(k)}{I_{phA}(k)} I_0 \tag{26}$$

where $G_A{}^*(k)$ and $I_{phA}(k)$ are the average values of G^* and I_{ph} in the area A centered on pixel k. This retina thus adapts locally its sensitivity to the average light intensity in surrounding area A. In order to achieve a value of this area approximately independent of the average light intensity, a local average $I_{av}(k)$ of the photo current is computed by an auxiliary diffusion network according to Fig. 11 and weighted by a factor F to control the lateral resistor $R^*(k)$. The equivalent diffusion area A is then proportional to F.

A last example of application of the diffusion network is the realisation of the lateral weights w_{ij} of a Kohonen self-organized map, as illustrated in Fig. 14 [18].

The weight w_{ij} of the connection from neurone i to neurone j is obtained by the difference of the transfer functions of two linear diffusion networks D_1 and D_2

having different diffusion lengths L_1 and L_2. The weights are positive (excitatory) for small distances d_{ij} and negative (inhibitory) for d_{ij} large. If w_{ij} does not need to tend to 0 for very large d_{ij}, L_2 may be made infinite ($R = 0$). Two copies of the output current of each neurone are used to drive the two separate networks.

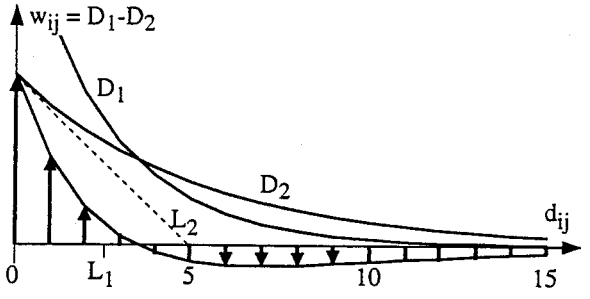


Fig. 14: Synaptic weights of lateral connection in a Kohonen self-organized map.

The diffusion network of Fig. 11 can be modified by replacing each grounded conductance G^* by a current source I_0. When a single input current I_{ini} is injected at node i, it is collected evenly by the N closest current sources I_0, so that $NI_0=I_{ini}$. Beyond this limit distance d_B, no more current is available and the current sources (saturated transistors) pull the node potential to 0. This creates a well defined "bubble of activity" centered at node i, of radius

$$d_B \cong \sqrt{I_{ini}/\pi I_0} \tag{27}$$

which can be centered at the single winner neurone in a Kohonen map [19].

This modified diffusion network is not linear anymore, and is therefore not applicable when more than one current is injected (no superposition anymore). An interesting manner to circumvent this limit is to use a communication scheme based on address-coding events [20], which permits to linearly superpose the effects of several signals by using this nonlinear network *sequentially* for each signal. An anisotropic version has been exploited to build an orientation enhancing image processor [21].

7 Fuzzy logic operations

The series connection of N conductances G_i results in a conductance G given by

$$G = 1 / \sum 1/G_i \tag{28}$$

which is the harmonic mean of the G_i divided by N. The harmonic mean I_{hm} of N currents I_i can therefore be obtained by the resistive circuit of Fig. 15a.

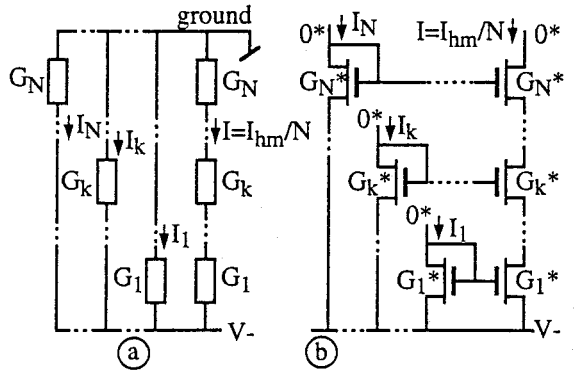


Fig. 15: Calculation of harmonic mean I_{hm} of N currents; can be used as a fuzzy AND gate. Pseudo-grounds are labelled 0*.

By exploiting the concepts of pseudo-conductance and pseudo-ground, this circuit can immediately be transformed into that of Fig. 15b. This last circuit can be used as the fuzzy logic AND gate needed to implement each rule of a fuzzy controller [22], in which each current I_i is a measure of the proximity of an input variable i with respect to some preferred value (value of the membership function). Several rules can be implemented by several such AND gates (branches) in parallel sharing the same bias current. The current flowing through each branch represents the relative weight of the corresponding rule. All branch currents are themselves weighted and summed to produce an output variable.

The membership function can be implemented by using the circuit of Fig. 15b with N=2 (two pseudo-conductances in series). The resulting output current

$$I = I_1 I_2/(I_1+I_2) \tag{29}$$

is maximum for $I_1 = I_2$ if I_1+I_2 is maintained constant. This can be approximately obtained by embedding the circuit inside a differential pair as shown in Fig. 16.

The output current of this "bump" or coincidence circuit [23] is given by

$$\frac{I}{I_0} = \frac{x}{x^2+3x+1} \quad \text{with} \quad x = \frac{I_a}{I_b} = \exp\left(\frac{V_a - V_b}{nU_T}\right) \tag{30}$$

and can be controlled by currents I_b/I_a or by voltages V_a -V_b (the input transistors can then be removed). This function, also shown in Fig. 16, can be used as a membership function with I_a (or V_a) as the input value and I_b (or V_b) the preferred value.

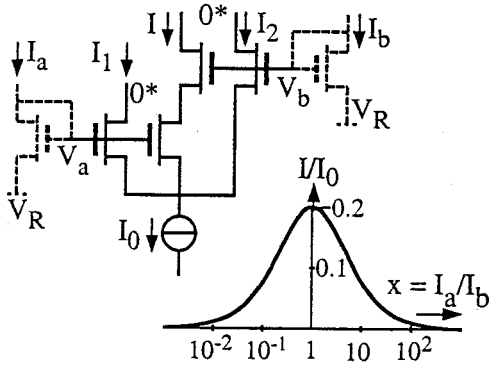


Fig. 16: Coincidence circuit implementing a membership function.

Another implementation of membership function based on the series connection of two pseudo-conductances is represented in Fig. 17 [22].

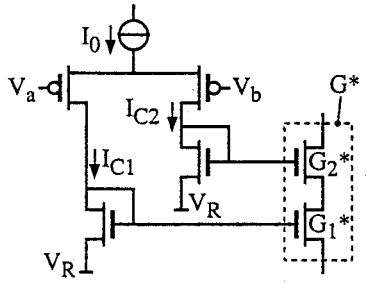


Fig. 17: Pseudo-conductance membership function.

The value of G_1*+G_2* is kept constant by a differential pair that imposes the sum I_0 of the control currents I_{C1} and I_{C2}. The membership function has about the same shape as that of Fig. 16 and is directly represented by the pseudo-conductance G* of the series connected G_1* and G_2*. It can therefore be directly used as one of the series-connected conductances G_i of the fuzzy AND gate implementing one rule [22].

The concept of membership function can be used to realize place coding in analog VLSI circuits [24]. This attractive alternative to purely digital or analog coding, inspired from the biology, can combine the precision of digital with the continuous amplitudes of analog.

8 Analog simulation of physical media

A n-dimensional resistive grid can be used to approximate the following equations:

$$\mathbf{grad}V = -\rho \boldsymbol{J} \quad \text{and} \quad \mathrm{div}\boldsymbol{J} = 0 \tag{31}$$

which relates the potential V and the current density $\boldsymbol{J}$ (unit: Am^{1-n}) in a n-dimensional continuous medium of resistivity ρ (unit: $\Omega\mathrm{m}^{n-2}$). A first and direct application of these laws is in the implementation of a path-finding circuit explained by Fig. 18 [25].

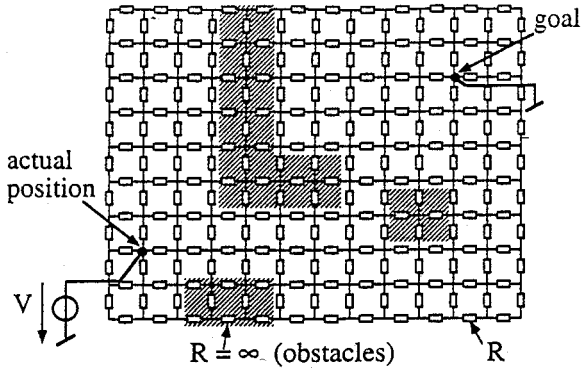


Fig. 18: Principle of a path-finding circuit.

A known landscape is represented by an homogeneous rectangular resistive grid which is interrupted ($R = \infty$) at the positions occupied by obstacles. The goal to be reached is grounded and a voltage source V is applied at the node corresponding to the present position. An auxiliary circuit identifies the direction of steepest gradient by comparing the potential of the four neighbouring nodes. The voltage source is then moved to the selected neighbour and the process is iterated until the goal is reached, which can be detected by an abrupt step of the current delivered by the source.

Each resistor can be replaced by a transistor operated as a pseudo-resistor. The whole circuit is then equivalent to a very large transistor of irregular shape. If this transistor is saturated by a large value of V, the gradient of potential is maximum at the actual position. This facilitates the identification of the steepest gradient and permits the implementation of larger grids than with simple linear resistors.

Such a system operates perfectly for any maze with a single possible path. It may be locked in cycles of

"hesitation" in situations with multiple similar paths, which can be solved by slightly complicating the biasing scheme [25, 26]. Only fixed (or infinite) pseudo-resistance values R^* are needed for the binary landscape of Fig. 18, thus the transistors may be operated at any current level. More elaborate analog landscapes could be implemented by exploiting the possibility offered by weak inversion to spatially modulate the value of R^*.

A n-dimensional resistive grid can also approximate equations

$$\mathbf{grad}P = -\gamma \boldsymbol{A} \qquad \text{and} \qquad \mathrm{div}\boldsymbol{A} = 0 \tag{32}$$

which relates the pressure P (unit: Nm^{1-n}) and the acceleration $\boldsymbol{A}$ in a n-dimensional volume of incompressible (and non viscous) liquid of density γ (unit: $\mathrm{Kg.m}^{-n}$). An interesting application is the electronic emulation of a cochlea, according to the principle explained in Fig. 19 [27].

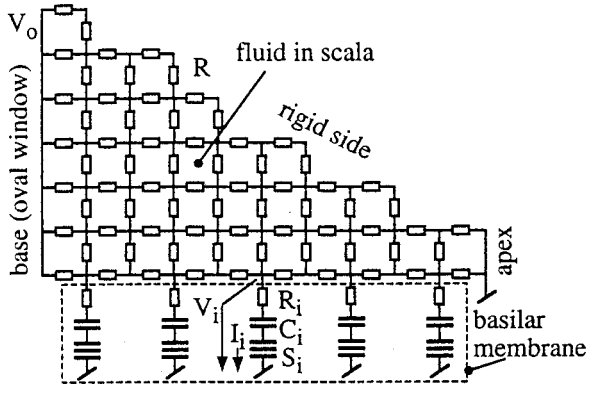


Fig. 19: Electronic emulation of cochlea hydrodynamics.

The shape of the longitudinal cross-section of the fluid filling the *scala* is replicated as a resistive network. The basal side (input) is driven by a voltage $V_o(t)$ corresponding to the pressure applied to the oval window. The rigid side is floating (no fluid motion) whereas the basilar membrane is implemented by a bank of grounded series resonators, each providing an impedance

$$\frac{V_i}{I_i} = R_i + \frac{1}{sC_i} + \frac{1}{s^2 S_i} \tag{33}$$

where R_i represents the local mass of the membrane, C_i the inverse of the local damping and S_i the inverse of the local stiffness. S_i is a "super-capacitor" realized by an active circuitry. The resonance frequency $\omega_{0i} = 1/\sqrt{R_i S_i}$ is reduced exponentially with the basal distance by increasing exponentially the value of S_i. The current I_i flowing through each resonator represents the local acceleration of the membrane and can be integrated (if it is not available inside the detailed resonator circuit) and rectified to emulate the output of the local inner hair cells.

The resistive network can be implemented by means of pseudo-conductances. Those may be varied in space to emulate the variation of mass due to the variation of size in the direction perpendicular to the cross-section. Voltage V_o must be transformed into a pseudo-voltage before driving the pseudo-conductance network. This is not needed for V_i if each resonators is built as a logarithmic filter [28] or if a current proportional to V_i is available inside the active circuit.

Another completely different possible application of pseudo-resistive networks is in the simulation of the non-linear behaviour of large transcontinental power distribution networks in which each separate line of the network is emulated by just two pseudo-resistors of value proportional to its length [29].

9 Conclusion

The concept of pseudo-conductance allows the implementation of any resistive circuit by means of MOS transistors only, while keeping its linearity properties with respect to currents. An additional advantage is the availability of pseudo-grounds, by which currents flowing to the ground can be collected by current mirrors without affecting the current distribution. Moreover, if the current levels are kept sufficiently low to remain in weak inversion, the value of each pseudo-resistor can be controlled by a voltage or by a current. These features open the way to very innovative architectures, well suited for the collective processing of data by means massively parallel analog VLSI. A wide range of applications are already known but the full potential of the approach is still being brought to light. So far, only time-independent processing has been described. Maintaining the linearity in time-dependent pseudo-resistive circuits requires pseudo-capacitors, which are active circuits closely related to the new field of logarithmic filters [28].

References

[1] E.Vittoz, "Analog VLSI signal processing: why, where, and how", Journal of VLSI Signal Processing, Vol. 8, pp.27-44, 1994.

[2] E.Vittoz, "Analog VLSI implementation of neural networks", to be published in the *Handbook of Neural Computation*, Institute of Physics Publishing and Oxford University Press, USA, 1996.

[3] E. Vittoz, "Future trends of analog in the VLSI environment", Proc. ISCAS'90, pp.1372-1375 New Orleans , May 2, 1990.

[4] E.Vittoz, "Micropower Techniques", in *Design of VLSI Circuits for Telecommunication and Signal Processing*, Editors J. Franca and Y. Tsividis, Prentice Hall, Englewood Cliffs, 1994.

[5] E.Vittoz and X.Arreguit, "Linear networks based on transistors", Electron. Lett., vol. 29, pp.297-299, 1993.

[6] K.Bult and G.Geelen, "A inherently linear and compact MOST-only current division technique", Dig. ISSCC Tech. Papers, February 1992, pp.198-199. Also in IEEE J. Solid-State Circuits, vol.27, pp.1730-1735, December 1992.

[7] K.W.Boahen and A.G.Andreou, "A contrast sensitive silicon retina with reciprocal synapses", *Advances in Neural Information Processing Systems*, Vol.4, pp.764-772, Morgan Kaufmann Publishers, San Mateo, 1992.

[8] B.Gilbert, "Translinear circuits: a proposed classification", Electron. Lett., vol.11, p.14-16, 1975.

[9] R.Maeder, "Réseau de Kohonen analogique à fonctionnement continu", Travail de Diplôme EPF-Lausanne, 1997 (in French).

[10] C.Enz and E.Vittoz, "CMOS low-power analog circuit design", in *Emerging Technologies*, edit. R.Cavin and W.Liu, The Institute of Electrical and Electronics Engineers, Piscataway, NJ, 1996, pp. pp.79-133.

[11] M.Tartagni and P.Perona, "Computing centroids in current-mode technique", Electron. Lett., vol.29, pp.1811-1813, 1993.

[12] P.Venier *et al.*,"Analog CMOS photosensitive array for solar illumination monitoring", ISSCC'96 Dig. Tech. Paper, pp.96-97, San Francisco, 1996.

[13] B.K.P.Horn, "Parallel networks for machine vision", The Artificial Intelligence Lab., M.I.T, Cambridge, MA, A.I. Memo No. 1071, Dec. 1988.

[14] D.L.Standley, "An object position and orientation IC with embedded imager", IEEE J. Solid-State Circuits, vol.26, pp.1853-1859, December 1991.

[15] C.A.Mead, *Analog VLSI and neural systems*, Addison-Wesley, Reading, 1989.

[16] C.A.Mead and M.A.Mahowald, "A silicon model of early visual processing", Neural Networks, vol.1, pp.91-97, 1988.

[17] P.Venier, "A contrast sensitive silicon retina based on conductance modulation in a diffusion network", MicroNeuro'97, Dresden.

[18] E.Vittoz *et al.*,"Analog VLSI implementation of a Kohonen map", Dig. of the Journées d'Electronique 1989 on Artificial Neural Networks, EPF-Lausanne, pp.291-301.

[19] P.Heim *et al.*, "Generation of Learning Neighborhood in Kohonen Feature Maps by means of Simple Nonlinear network", Electronics Letters, vol.27, No 3, pp.275-277, 1991.

[20] A.Mortara and E.Vittoz, "A communication architecture tailored for analog VLSI artificial neural networks: intrinsic performance and limitations", IEEE Trans. on Neural Networks, Vol. 5, pp.459-466, May 1994.

[21] P.Venier *et al.*, "An integrated cortical layer for orientation enhancement", IEEE J. Solid-State Circuits, vol. 32, pp.177-186, Febr. 1997.

[22] O.Landolt, "Low-power analog fuzzy rule implementation based on a linear transistor network", Proc. MicroNeuro'96, pp. 86-93, Lausanne, 1996.

[23] T.Delbrück,"Bump circuits for computing similarity and dissimilarity of analog voltages", Proc. of International Joint Conference on Neural Networks, vol.1, pp.I-475-479, 1991.

[24] O.Landolt, "Place coding in analog VLSI and its application to the control of a light deflection system", MicroNeuro'97, Dresden.

[25] L.Tarassenko *et al.*, "Real-time autonomous robot navigation using VLSI neural networks", in *Advances in Neural Information Processing Systems*, vol.3, R.P.Lippmann, J.E.Moody and D.S.Touretzky (editors), pp.422-428, Morgan Kaufmann, 1991.

[26] S.Scotzniovsky, "Détection de chemin dans un environnement avec obstacles", Travail de Diplôme EPF-Lausanne, 1995 (in French).

[27] L.Watts *et al.*,"A bidirectional analog VLSI cochlear model", in *Advanced Research in VLSI, Proc. of the 1991 Santa Cruz Conference*, M.I.T.Press, Cambridge, MA, pp.153-163.

[28] C.Enz *et al.*, Low-voltage log-domain signal processing in CMOS and BiCMOS", ISCAS'97, Hong Kong.

[29] R.Fried *et al.*,"On chip transient stability simulator 10^4 times faster than real time"Dig. 2nd Int. Conf. on Digital Power System Simulators-ICDS'97, Montréal, pp.79-84.