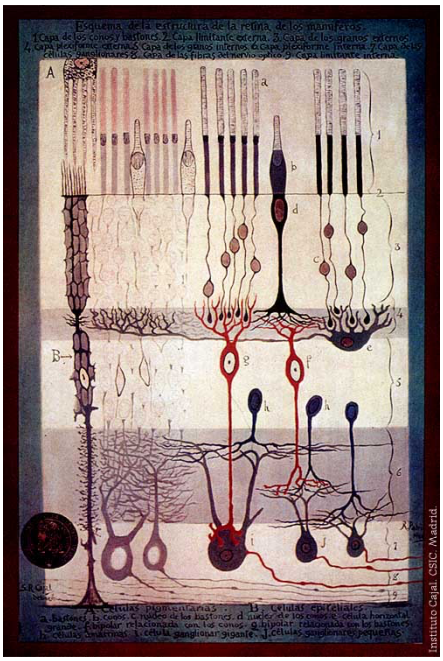
Neuromorphic Integrated Bioelectronics - Fall 2024

BENG 216, UC San Diego

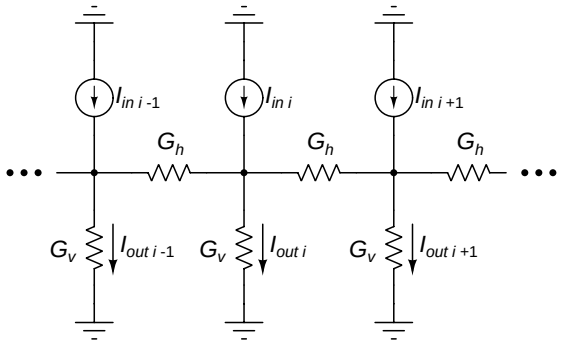
Homework 2: Due October 18

In the second homework we study a simple model of retinal function with an array of local contrast sensitive photoreceptors in a silicon integrated circuit, and use the electronic design automation (EDA) tools using the open-source SKY130 PDK for the 130nm CMOS Skywater process to implement the model with the layout of a small array of pixels. The model focuses on the outer plexiform layer of the mammalian retina, with rod and cone photoreceptors and horizontal and bipolar cells as shown in layers 1 through 4 in Ramón y Cajal's drawing on the right. We continue to assume subthreshold operation for all MOS transistors as in the first homework, and extend the use of Cadence® Virtuoso® tools for schematic capture with layout and LVS (layout vs schematic) for functional verification of the consistency between layout and schematic. In addition to analyzing the silicon retina circuit's properties and getting familiarized with the remainder of the full-custom EDA tools for its implementation, we aim for scalable and structured layout of the silicon retina pixel cell that readily tiles to implement larger arrays of cell instances, striving for high pixel density while also maximizing fill factor for high resolution and sensitivity of photoreception and visual processing.



1. *Diffusive kernel of spatial lowpass filtering of photocurrent in the retina outer plexiform layer* [40 points].

   (a) Consider the infinite linear network shown on the right modeling a 1-D version of the outer plexiform layer in the mammalian retina, with horizontal conductances G_h corresponding to horizontal cells shunting the photocurrent $I_{in\,i}$ received from rod and cone photoreceptors, and vertical conductances G_v corresponding to bipolar cells passing the residue as output current $I_{out\,i}$ to the retinal ganglion cells in the inner plexiform layer.

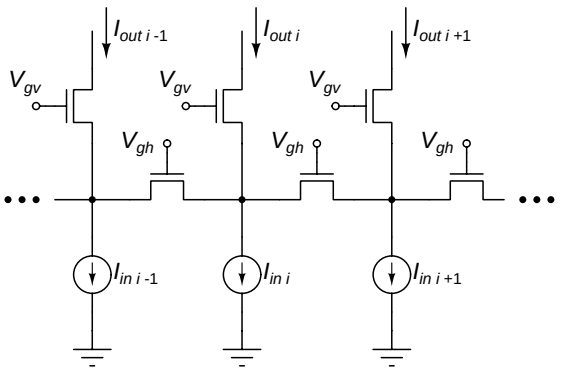


   Express the output $I_{out\,i}$ in terms of its neighbors $I_{out\,i-1}$ and $I_{out\,i+1}$, and the input $I_{in\,i}$. This output produces a spatial lowpass filtered version of the photocurrent input, smearing the linear profile of the photocurrent along the lateral axis x. Show the output response $I_{out\,i}$ to an impulse

at the input ($I_{in\,0} = 1$, and $I_{in\,i} = 0$ for $i \neq 0$) under two conditions of the network: *i*) $G_h = G_v$; and *ii*) $G_h = 0.1G_v$. Observe that the impulse response is a decaying exponential in the lateral distance x from the impulse center: $I_{out}(x) = \exp(-|x|/\lambda)$. Find an approximate expression for the space constant λ in terms of the conductances G_h and G_v, and the linear spacing L between nodes (*i.e.,* $x = i\,L$).

The space constant λ quantifies the spatial spread of smearing in the lowpass spatial filter response, and is the spatial equivalent of the time constant τ in a temporal lowpass filter. In the spatial frequency domain the lowpass filtering at the output is characterized by a transfer function $H(k) = I_{out}(k)\,/\,I_{in}(k) = 1\,/\,(1 + \lambda^2 k^2)$, where k is spatial frequency in radians per unit distance, with a cut-off frequency $k_c = 1/\lambda$.

(b) Now consider an nMOS implementation of the network of horizontal and vertical conductances, replacing each linear conductance with a corresponding translinear subthreshold MOS conductor as shown on the right. Photocurrents come in at the bottom, and outputs exit at the top; the vertical flipping of the topology amounts to consistent inversion of the polarities of all currents and voltages which is invariant to the conductances.



Express the translinear conductance $I_{ds} = G^*\,(V_d^* - V_s^*)$ for the subthreshold nMOS transistor under the transformations $G^* = I_n \frac{W}{L} \exp(\kappa_n V_{gb}/V_T)$, $V_d^* = -\exp(-V_{db}/V_T)$, and $V_s^* = -\exp(-V_{sb}/V_T)$ to show equivalence of the outputs $I_{out\,i}$ of this translinear MOS conductor network and its above linear conductance form under the transformations $G_h = I_n \frac{W}{L} \exp(\kappa_n V_{gh}/V_T)$ and $G_v = I_n \frac{W}{L} \exp(\kappa_n V_{gv}/V_T)$. Rewrite the space constant λ in terms of the bias gate voltages V_{gh} and V_{gv}, and the node spacing L.

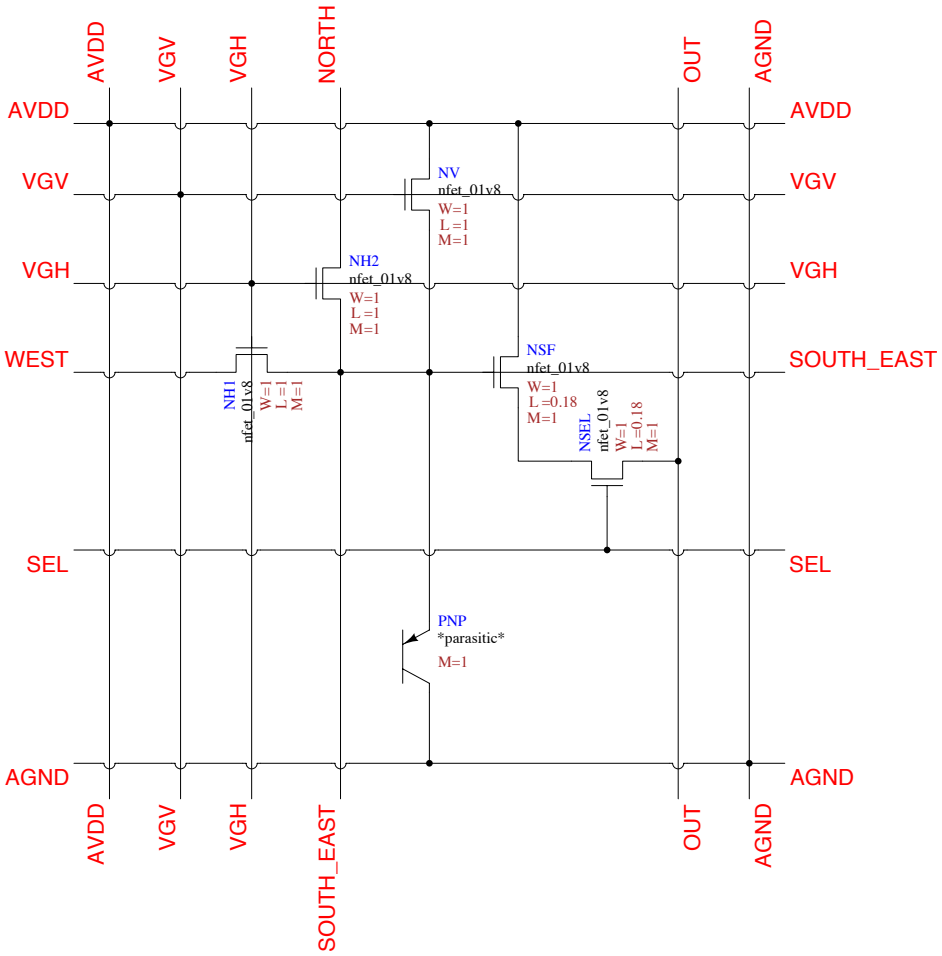
To what extent does it matter what voltage is applied at the $I_{out\,i}$ output nodes? What range of voltages would you recommend? Explain your reasoning.

(c) Finally, we consider 2-D array extensions of the linear and translinear network with horizontal conductances extending across nodes on a 2-D grid. Show (sketch) how to extend the MOS translinear network for a 2-D geometry of cells on a cartesian grid, and on a hexagonal grid.

The concept of spatial frequency k extends to the wave vector $\mathbf{k}$ in two dimensions. As in the 1-D network, a larger space constant λ across the 2-D grid corresponds to greater spatial extent of blurring in the output image. The 2-D lowpass filtering at the output is characterized by a transfer function $H(\mathbf{k}) = I_{out}(\mathbf{k})\,/\,I_{in}(\mathbf{k}) = 1\,/\,(1 + \lambda^2 |\mathbf{k}|^2)$, again with cut-off frequency $k_c = 1/\lambda$.

2. *Towards a local contrast-sensitive silicon retina: Pixel layout, array assembly, and schematic verification* [60 points].

We will implement, in the open 130nm Skywater process, the layout of a small-size (4×4) array of pixel cells each including a PNP bipolar junction phototransistor, vertical and horizontal translinear MOS conductances with neighboring cell connnectivity on a 2-D cartesian grid, and row-select multiplexed column readout of the outputs. A diagram with the pixel schematic, suggestive of the interconnect topology of the cell with its neighbors on the cartesian grid, is shown below using devices available in the 130nm PDK, with parameters for the MOS devices included. The corresponding netlist for the MOS transistors is also shown.




```
// Cell name: hw3_aps
// View name: schematic
NV (AVDD VGV SOUTH_EAST AGND) nfet_01v8 w=1u l=1u m=1 region=subth
NH1 (SOUTH_EAST VGH WEST AGND) nfet_01v8 w=1u l=1u m=1 region=subth
NH2 (NORTH VGH SOUTH_EAST AGND) nfet_01v8 w=1u l=1u m=1 region=subth
NSF (AVDD SOUTH_EAST OUT_BUF AGND) nfet_01v8 w=1u l=180n m=1 region=subth
NSEL (OUT_BUF SEL OUT AGND) nfet_01v8 w=1u l=180n m=1 region=subth
```

The parasitic PNP bipolar junction phototransistor (the element labeled PNP in the schematic) is intrinsic to the process layers and implied by the layout geometry, realized as a vertically nested stack of p-diffusion for the emitter, in n-well for the base, in the p-substrate for the collector. This structure is called a "parasitic" PNP bipolar junction transistor because every source or drain of a pMOS transistor has this vertical stack of junctions from p-diffusion to n-well to p-substrate. Under normal conditions where the n-well bulk of pMOS transistors are plugged to Vdd, each of these parasitic BJTs are reverse biased so they carry no active current and only reverse diode junction leakage current. For this reason, a parasitic BJT structure created in the layout is usually ignored and not included in the extracted netlist. Leaving the n-well floating, photocurrent generated by generation of electron-hole pairs due to photon absorption in the base-emitter depletion region forward biases the junction to amplify the photocurrent by a large gain (around 50) at the collector. Hence the parasitic PNP phototransistor structure is often preferred over a simpler n-diffusion to p-substrate photodiode structure for greater current sensitivity in photoreception down to very low light levels. Feel free to adjust the device sizing and physical layout to aid in the optimization for high pixel density and photosensor fill factor.

All nMOS transistors are implemented as `nfet_01v8` in the core library of the SKY130 PDK. Two horizontal translinear conductances (NH1 and NH2) connect to west and north neighbors, respectively; the cell further extends its central node to east and south neighbors to receive their inputs. The vertical translinear conductance (NV) is tied to AVDD in order to maximize voltage swing in saturation. Readout uses a standard active pixel sensor (APS) two-transistor circuit for time-multiplexing row selection (NSEL) and for sensing the bipolar cell source voltage for the cell output (NSF). The readout transistors NSEL and NSF are sized with minimum length, but the translinear conductance elements NV, NH1 and NH2 are sized with longer lengths for lower mismatch and less channel length modulation to minimize variability and drain conductance effects impacting the subthreshold characteristics. Although these transistors are shown at different orientations in the schematic, for greatest matching they should be oriented consistently in the layout, including their source and drain polarity, with all sources of NV, NH1 and NH2 either at the bottom of identical vertical geometry, or on the left of identical horizontal geometry.

As a benchmark, the cell will operate in the standard integrating APS mode by disabling horizontal shunting connectivity by driving V_{gh} to AGND and by pulsing V_{gv} for periodic pixel reset. By tying V_{gv} to a tunable constant voltage bias, the readout implements logarithmic intensity encoding owing to the exponential dependence of translinear conductance on source voltage of the vertical conductance transistor (NV). This logarithmic encoding, squashing 4 orders of magnitudes in photocurrent in just 300 mV of voltage on the transistor source provides for high dynamic range in photoreception bypassing the need for automatic gain control to dynamically adjust to environmental lighting conditions. Finally, adjusting the horizontal shunting conductance through V_{gh}, relative to V_{gv}, within the same readout cycle allows to collect two samples of the instantaneous intensity scene at two different space constants. Differencing these two measurements within the same readout produces a net output that directly codes local contrast at high sensitivity: the difference returns zero in those parts of the scene at near-zero contrast, whereas high-contrast regions generate highly positive (on center, off surround) or negative (off center, on surround) differences. Furthermore, the difference of log intensities, equivalently the log ratio of intensities, is insensitive to global illumination level ensuring that the same contrasts at drastically different ambient lighting conditions generate consistent difference measurements.

(a) Enter the schematic of the silicon retina pixel circuit, as shown, and complete and verify your layout implementing it in the Skywater 130nm process using Cadence® Virtuoso® for schematic capture, layout, and layout vs. schematic (LVS) design verification. As before you may use these tools from your class accounts on the `ieng6.ucsd.edu` linux cluster.

Show a printout of your schematic, layout, netlist, and LVS results. In addition to consistency between layout and schematic, your layout should be clear from design rule errors; this step may require explicit DRC runs. Make sure that all nodes in the circuit are accounted for, with no duplicate or missing nodes in the layout and schematic. This includes the substrate which by default is at ground potential. In the provided schematic, the substrate is represented by the `agnd` analog ground node variable. The bulk of the nMOS transistors, and the collector of the PNP phototransistor are directly connected to the substrate.

(b) Now extend your single pixel to an array of pixels. Instantiate your pixel cell as a 4×4 array, consistently in the schematic and in the layout. The cells should interface properly with their neighbors in the 4×4 array geometry in order to share common horizontal and vertical signal lines, and implement the horizontal conductance shunting between neighbors. The cell schematic as shown above purposely connects signals across with matching interfaces at the cell periphery to realize north-south and west-east interconnects, and to provide common grounding and supplies across the entire array.

Your top-level layout and schematic should have nothing in them other than array instances of the pixel cell. There should be no routing of signals by metal lines or other physical interconnects other than what is already provided within the cells being instanced. If you find any discrepancies, go back to your pixel design and reiterate. It is good practice to use the same metal layer for all vertical signal lines, and another layer for all horizontal signal lines, so as to avoid the need for via bridges jumping between metal levels in the signal crossing the cell.

It is often advantageous to share ground and supply lines between neighboring cell instances for increased cell density. Doing this correctly requires mirroring the cell instance in horizontal and/or vertical directions for every other column and/or row in the array. You may attempt to do this, although this is not necessary in this homework.

(c) Quantify the following metrics of your layout:

i. *Pixel density*: The number of pixels per unit area, or equivalently, the reciprocal of the pixel area. On a cartesian grid we expect your pixel to be square, with equal length and width L, so the area is L^2. L may be smaller or larger than the boundaries of the actual layout, and corresponds to the *pitch* between cells in the 2-D pixel array. Higher pixel densities are better for higher spatial resolution, although lower densities collect more photocurrent for the same illumination intensity, and may be preferable for applications operating under dim lighting conditions. Our design includes amplification of photocurrent by the current gain of the PNP bipolar junction transistor and hence we strive for maximum available density, constrained by fill factor.

ii. *Photoreceptor fill factor*: The ratio of the total fully exposed photosensitive junction area, over the pixel area as just defined. Higher fill factor is better in order to capture most of the incident light intensity as photocurrent and avoid spatial aliasing under the presence of illumulation patterns at spatial frequencies k greater than the $1/L$ linear density.

(d) **BONUS** [Extra 15 points]: Instead of the cartesian geometry, implement the silicon retina on a hexagonal grid. Each cell interfaces to 6 rather than 4 neighbors, but the cell geometry in the 4×4 array is still cartesian with row selection and column readout. Devise the dimensions of your pixel, and its sequence of orientations in the array instance, so as to realize a properly scaled hexagonal grid. The phototransistor needs to be positioned in the cell such that it is equidistant to its six neighbors.

Submission Guidelines: You are encouraged to work on teams and exchange solution strategies and share configuration of the EDA tools, but you must complete the homework yourself, and are not allowed to copy other's work. In particular you cannot share schematics and layout for homework submission and need to complete these entirely by yourself. It is anticipated that no two independent schematic and layout designs are identical.

Turn in your homework as a single PDF over canvas by the due date. Scanned handwritten notes are fine, including the printout of the design materials.