Comparing Semiconductor and Superconducting Technologies for Ultra Large Scale Optoelectronic Neural Systems

bryce.primavera1

August 2020

1 Energy Consumed in Communication

Advantage: Wash

For semiconductor systems, the minimum energy/bit is likely around 1 fJ.[1] This is assuming a photodetecor with a very low capacitance. If the capacitance of the detector is low enough (femtofarads), then no amplifier is necessary. Even if a transimpedance amplifier is used, it seems unlikely to reach values of less than 1 fJ/bit, since that is the energy required to switch a transistor - something that is necessary if we seek to use logic-level voltage pulses to define spikes.

For SOENS, the signal could be as low as a handful of photons. Photons are of order 10^{-19} J, so let's assume that spikes contain roughly 10^{-18} J of energy. While this is three orders of magnitude less than the semiconducting case, it is only feasible at 4K. The cooling power adds a factor of 1000. Therefore, all other things being equal, the energy dedicated to communicating spikes in SOENS will likely be near to that of a room-temperature semiconductor approach.

2 Fan-in

Advantage: Wash

Any practical neural platform must be capable of implementing neurons with thousands of synapses. For SOENS, this is achieved with mutual inductors coupling into a current integration loop. In the semiconductor case this will be achieved through the direct addition of currents using Kirchoff's Law. This may be aided by an op amp establishing a virtual ground for all of the currents to combine at. More work needs to be done here, but there doesn't seem to be anything prohibitive. It also seems to have played a role in Mead's original thinking, which suggests that it's not a ridiculous idea.

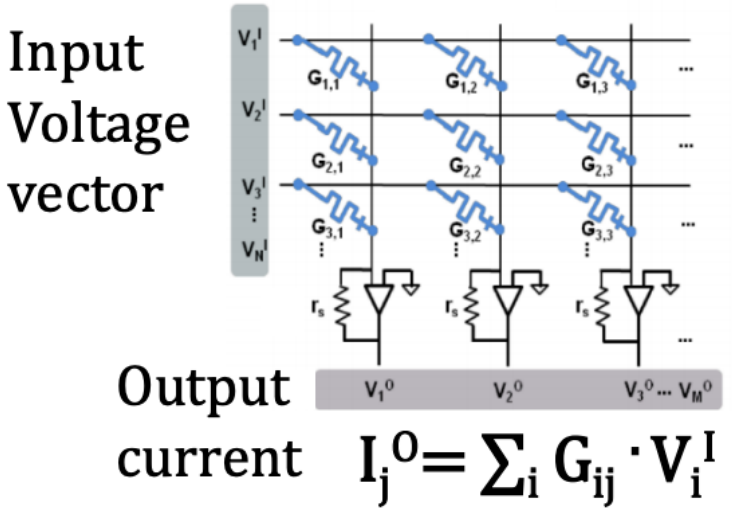

Figure 1: Typical Memristor crossbar array

In effect, I see no difference between memristor crossbar arrays and an array of optoelectronic synapses. Both generate a current that should be capable of summing with Kirchoff's Law. In Figure 1, you could replace the input voltages with optical pulses. The memristors would be replaced by full-on synaptic circuits with optical receivers. The output current would be the output of the synapse. The transimpedance amplifiers at each neuron may or may not be necessary. The horizontal lines would be waveguides while the vertical lines would just be wires.

3 Plasticity (Weight Adjustment)

Advantage: TBD

In SOENS the synaptic weights can be adjusted by adding current to flux storage loops associated with the plasticity circuits. In the semiconductor case, floating gate transistors seem to be the most natural fit. There is an advantage here in that floating gate memories are non-volatile. There are also various volatile bias-generating circuits that need to be looked at. We need to know how many bits are possible in each case. More work here...

4 Physical Size

Advantage: TBD

The size of the individual neurons will ultimately determine the size of a possible neuronal pool since both platforms will be communicating at the speed of light. For semiconductors, we have an estimate of 1360 μm^2 for a DPI synapse with significant functionality. [2] At least on first impression this seems comparable to SOENS?

5 3D Integration

Advantage: SOENS?

3D integration is a way to effectively shrink the area consumed by a circuit and increase the neuronal pool. For SOENS, it seems likely that SNSPDs can be deposited on dielectric layers, spreading synapses across many layers (up to 20?). 3D integration in CMOS has been an elusive goal for many years now. It may be more viable in this context than in previous digital attempts since the transistors are mostly subthreshold and will be spread out - easing the demands on cooling.

The fewer device layers are capable of being stacked, the more wafers will need to be linked together with off-wafer interconnects. 3D integration may save energy by reducing the number of these off-wafer connections.

6 Device Variability

Advantage: TBD

Fair to say that both JJ's and subthreshold transistors have significant questions about device variability? Subthreshold variability can be ameliorated by using larger device nodes. Maybe something similar will be the case for JJs. The essential question is which platform is most impacted by the amount of variability present in its respective devices?

7 Timeline for Feasibility

Advantage: Semiconductors?

There are significant technological hurdles for both of these platforms. Foremost for both is an on-chip optical source.

For SOENS, it is still an open question of how to interface JJ's with an LED, although hopefully this is not far off. The fabrication process is arguably more complicated for SOENS - potentially incorporating SNSPDs, JJ's, transistors, and LEDs. Additionally, for one million neurons per 300 mm wafer, you need 100,000 300 mm wafers to reach the scale of the human brain. How much of an investment would it take to build (and run) a cryostat that can cool 100,000 wafers to 4K?

There are also some questions endemic to the semiconductor platform. The idea of integrating femtofarad detectors with transistors in the "receiver-less" configuration is still somewhat speculative.[1] If III/V integration is not possible, then this platform will also require cryogenics. However, 77K becomes a possibility. How does a 77K cryostat capable of cooling 100 billion semiconductor neurons compare to SOENS? If 3D integration is not possible, then it might not be obvious that this is easier than the 4K SOENS version since you'll need more wafers.

Finally, I think it is worth emphasizing that a lot of the neuronal circuits in the semiconductor case have already been built and tested. While SOENS is in its infancy, a lot of groundwork has already been laid for the semiconductor case.

References

[1] DAB Miller 2009 pg 1177

[2] Bartolozzi 2007 pg 2592