Final Exam

ELEC 859: Unconventional Computing, W'2023
Instructor: Alex Tait

I. Instructions

You have **three hours** to complete the exam. Upload your responses to OnQ under Assessments>Assignments>Final-Exam.

I.A. Rules

This exam is open note. You are allowed to use anything that is not prohibited, including

- Your notes
- ELEC 859 lecture slides (OnQ:Content)
- The Internet
- Scripts you write to help with calculations
- Third party software
- Diagramming programs like Powerpoint and Inkscape

The following are **prohibited**

- Communication with other ELEC 859 classmates. This includes all communication related *or not related* to the exam.
- Artificial intelligence of any kind (e.g. GPT, QuillBot, Stable Diffusion, etc.). Exceptions: necessities like smartphone face recognition, enhancement filters for picture uploads, clippy.

The prohibitions are enforced through an honor code and the possibility that I have AI detection software.

I.B. Format

Allowable upload formats are PDF, markdown, Word, or pictures of hand written pages. Note, your responses will include **equations** and **diagrams**, so plan accordingly. Make sure to clearly label responses with the prompt number. Show your work for partial credit. Some guidelines:

- If uploading LaTeX, markdown, or Word: if you prefer not to typeset intermediate equations, include pictures of any hand-written intermediate work.
- If using scripts for calulation: you may submit these scripts for partial credit. Put final answers in the response document.
- If submitting pictures of pages, make sure they are legible. I recommend using a PDF scanner app, such as Genius Scan.

Any questions about format, ask me on the zoom room.

I.C. Announcements

If one of you finds an egregious error or ambiguity, I will post a clarification in OnQ under Communications>Announcements. Check for announcements periodically.

II. BROWNIAN MEMORY

II.A. Setup

There is a metal sphere. It has capacitance and voltage with respect to its environment. We can use the voltage on this capacitor to represent and remember one value. Call its voltage vs. time $v(t)$. At the start, we set its voltage to V_1 so that $v(0) = V_1$. Suppose this capacitor has infinite leakage resistance so that the voltage does not decay to zero over time.

There is thermal noise. It causes random ionizations of gas around the ball and random motions of charge carriers. All of these act like a current source giving $i_n(t)$, which is a Gaussian random variable with mean of zero and root-mean squared (RMS) amplitude of σ_n Amps. Math notation for this is $i_n(t) = N[0, \sigma_n]$. N stands for normal distribution.

A Gaussian random variable of $N[a,b]$ means that the probability of measuring a particular value x is a Gaussian function:

$$P\{N = x\} = \frac{1}{\sqrt{2\pi b^2}} e^{\frac{(x-a)^2}{2b^2}}$$

Over time, the voltage on the capacitor will vary randomly according to the integral of thermal current. The integral of a Gaussian random variable is called a Wiener process, a.k.a. Brownian motion, a.k.a. diffusion. At any given time, the Wiener process is also a Gaussian random variable; however, its std. dev. depends on how much time has elapsed since the process started at $t = 0$. In this case, the probability of measuring a certain voltage is a Gaussian distribution centered at V_1 with standard deviation σ_1. In math notation, $v(t) = N[V_1, \sigma_1(t)]$.

While σ_n is a constant parameter, σ_1 is time dependent: $\sigma_1(t) = \sqrt{Dt}$. That is the defining feature of Brownian motion. The diffusion parameter, D, is related to the current noise amplitude by $D = \sigma_n^2/C^2$ in volts-squared per second.

tl;dr voltage std. dev. depends on noise std. dev. and time like this

$$\sigma_1(t) = \frac{\sigma_n}{C}\sqrt{t}$$

Voltage depends on current like this

$$v(t) = C^{-1} \int_0^t i_n(t')dt'$$

and a Gaussian random variable $N[a,b]$ means that the probability of measuring a particular value of x is

$$P\{N = x\} = \frac{1}{\sqrt{2\pi b^2}} e^{\frac{(x-a)^2}{2b^2}}$$

II.B. Parameters

For each question below, give an analytical expression and a numeric value. To evaluate the numeric one, use these parameters

C	σ_n	$V_{\max}$
5×10^{-15} Farads	2×10^{-9} Amps	8 Volts

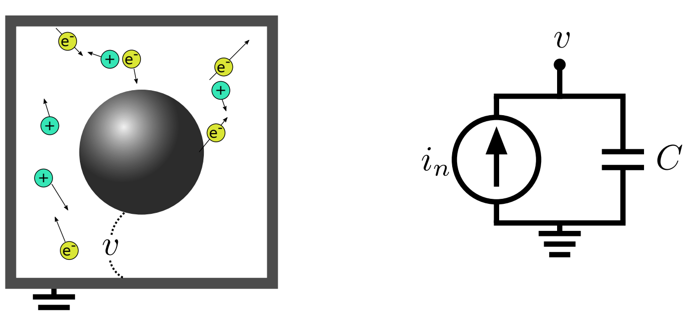

FIG. 1. Metal sphere with thermal noise. Circuit model.

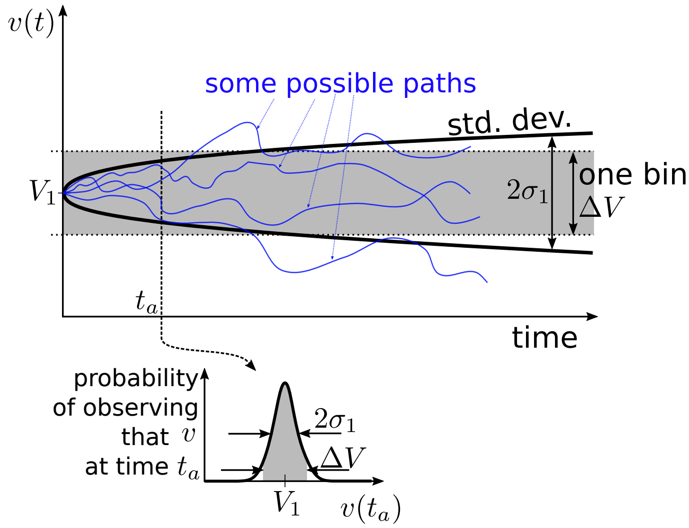

FIG. 2. Stochastic behavior of voltage over time.

II.C. Analog memory

Using Shannon's equation, we are going to calculate how fast information degrades over time. V_1 can range between 0 V and $V_{\max}$. Note, $v(t)$ is allowed to go outside of this range without bound. At $t = 0$, we know $v = V_1$ exactly.

> **Prompt 1.** How much information is conveyed in this variable at $t = 0$? What does σ_1 need to be such that we can get 8 effective bits of information about what V_1 was? What about 4 bits?

> **Prompt 2.** How long can it possibly store an analog value with an effective 8-bit precision? 4-bit precision? Remember, give both an analytical expression and numeric values for these.

II.D. Hybrid memory

To remember this variable longer, we can periodically regenerate it to discrete values. That means dividing voltages into bins with width ΔV. Bin 0 is the range $[0, \Delta V)$; Bin 1 is $[\Delta V, 2\Delta V)$; and so on. V_1 starts in the middle of one of the bins, in other words, a half-integer multiple of ΔV. After some time, T_{regen}, we measure $v\left(T_{\text{regen}}\right)$, determine which bin it is in, and then set it back to the center of that bin. An error happens if it meanders into a different bin. For example, if we measure $43.883 \times \Delta V$, then we will count that as Bin 43 and reset the voltage to $43.500 \times \Delta V$. If it started in bin 43, that's good; if it started in bin 44, that's an error.

> **Prompt 3.** What is the probability of an error, P_e? In other words, what proportion of the area under the Gaussian probability distribution lands outside of $V_1 \pm \frac{\Delta V}{2}$? Give a closed form analytical expression that contains ΔV and T_{regen}; Hint, there are plenty of tools for solving integrals analytically without using AI.

> **Prompt 4.** Suppose we need 8 bits of resolution and a P_e less than 10^{-12}. What should we use for ΔV and T_{regen} in order to barely meet this spec?

II.E. Quantum-ish memory

Instead of continually regenerating, we can use the fact that charge carriers are fundamentally quantized. First, let's eliminate thermal noise by cooling down to 0.01 Kelvin and sticking it in a vacuum so that $i_n = 0$. Easy stuff.

Prompt 5. When using individual electrons to define voltage levels, what is ΔV? Using this quantum-ish technique, and keeping the same $V_{\max}$ from above, how many bits of information can be stored on this metal sphere?

Electrons are wavelike. Although there is no thermal noise, there is still a small chance that electrons can tunnel across the capacitor. As above, we start from a known voltage state, wait some time, then measure voltage, and finally do something to put the sphere back in a known state. You don't have to calculate error rates again.

Prompt 6. In words, describe 1) the possible voltages we might measure and 2) after making a measurement, the steps we need to take to regenerate voltage back to a known state. How are these two aspects similar or different from the hybrid memory case?

II.F. Quantum memory

Let's limit ourselves to one bit. The metal ball can be completely charge neutral, or it can have one extra electron. Call these states $|0\rangle$ and $|1\rangle$ respectively. The states have different energies, which means that their quantum phases progress at different frequencies.

Prompt 7. What is the voltage difference between $|0\rangle$ and $|1\rangle$? What is the energy difference along the $|0\rangle$, $|1\rangle$ basis? What is the associated frequency difference?

It is possible for this system to be in a quantum superposition. We can refer to its state as $|\Psi\rangle = a|0\rangle + b|1\rangle$. A metal sphere is a lot simpler than a real qubit though, so there is not an obvious way to create the superposition. Let's just say that we can magically start it in any state $|\Psi\rangle$.

Prompt 8. If state starts at $t = 0$ with $a = \frac{1}{\sqrt{2}}$, what is one possible $b(t = 0)$? What will $a(t)$ and $b(t)$ be as functions of time? Include a Bloch sphere in the reponse.

III. Hardware oscillators

III.A. Setup

This is a thought experiment about a massively distributed computer for solving the ODE problems from Assignment 2. There are three kinds of problem considered: where each node is uncoupled, where all are coupled to a mean field, and where neighbors are coupled in a ring.

The circuit for a single "node" is pictured in Fig. 3. Once again, we are using capacitors to store physical state, but there is no stochastic behavior or noise in this section. There are two states per node. The rate of change of voltage is proportional to the current applied to these capacitors. We also have some kind of processing device that gives a voltage-dependent current

$$\frac{dv_1}{dt} = \frac{1}{C}\,\left[g f_1(\vec{v}) + i_{ext,1}\right] \tag{1}$$

$$\frac{dv_2}{dt} = \frac{1}{C}\,\left[g f_2(\vec{v}) + i_{ext,2}\right] \tag{2}$$

$$\tag{3}$$

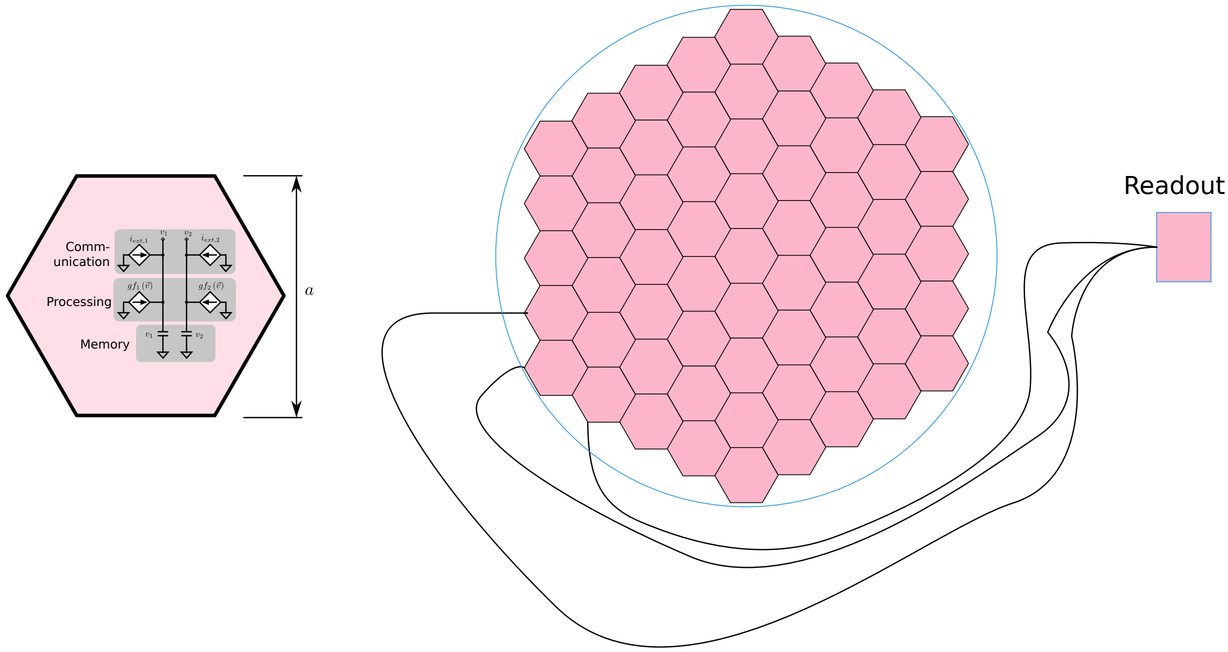

FIG. 3. Physical simulator layout.

where the terms in braces are units of electrical current. g is a conductance needed to make proper units. It is in Amps per Volt, a.k.a. Siemens or Mhos (℧). i_{ext} comes from neighboring nodes or mean field, depending on which of the three problems we are solving. For specificity, we make the function for f correspond to an oscillator. It is linear.

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} 0 & \omega \\ -\omega & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{4}$$

Let's impose some finite limits on voltage and current. All currents are limited to $I_{\text{min,max}}$. All voltages will fall between $V_{\text{min,max}}$.

Each node occupies some physical footprint on a chip. They can be tiled hexagonally with cell size a, resulting in a system layout with diameter D. The maximum of D is limited by room on a wafer with diameter D_w. The number of nodes is referred to as N.

We will benchmark against a CPU. Assume it is single-core with no caching behavior. It runs the Euler integration algorithm, calculating time derivative based on the state at the present timestep, then incrementing the state for the next timestep. To be clear, `dt` refers to a simulation timestep as a concept. Δt_{CPU} (in seconds) refers to the real, physical time that the CPU needs to complete one timestep.

III.B. Parameters

For each question below, give an analytical expression and a numeric value. To evaluate the numeric one, use these parameters

C	$I_{\text{min,max}}$	$V_{\text{min,max}}$	g	a	D_w
100 fF	$\pm$10 mA	$\pm$1 V	10 m℧	200 nm	300 mm

III.C. Physical bandwidth

Due to constraints on current and voltage, there is a minimum amount of time needed for voltage to swing all the way from V_{min} to V_{max}. Call this limit Δt_{ckt}. Δt_{ckt} is an analog circuit concept, but it is related to the step operation (SOP) that we defined as happening every `dt` in the discrete-time simulation.

Prompt 9. Calculate the minimum Δt_{ckt}.

III.D. Physical footprint

The tiled arrangement is shown in Fig. 3. We want a formula to determine system footprint based on a and N. This will take a bit of geometry, or you can approximate the system as a circle and take an area ratio.

> **Prompt 10.** Give an analytical relationship between a, N, and D. What is the maximum N that we can fit on a wafer with diameter D_w: analytical and numeric.

III.E. I/O limit

These circuits can go fast on their own, but we typically want to know what is going on within the simulator. We can read out from the edges of the mesh, meaning the maximum I/O bandwidth to the mesh scales with its perimeter (πD). This leads to another speed limit: Δt_{io}. Every step, every cell generates 1 byte of information about its state. The edge bandwidth coefficient, b, in bytes per second per meter, describes how much. Memory bandwidth is $B = \pi D b$. Assume b = 1.0 TB/s/m.

> **Prompt 11.** As functions of N, how many bytes about state are generated per step by the entire system? What is the maximum readout bandwidth in bytes/sec? What is the limit on Δt_{io}?

> **Prompt 12.** What is the arithmetic intensity of the computer assuming that we read out every value at every time step? Include units.

> **Prompt 13.** Suppose instead that we want the computer to go at full speed but only need to read out a subset of cells. What is the maximum number of cells we can read out? What is arithmetic intensity in that case?

III.F. Compare to CPU

Consider a single-threaded CPU implementation of the same problem. Every time step, the CPU loads a value, calculates the `dt` update, then writes the new value back. In addition, the new value is written to an array in memory, which will eventually contain all values over all time.

Suppose CPU can perform 150 MFLOPs/s, and the memory bandwidth is 3 GB/s. Neglect caching and scheduling effects: all reads and writes happen at the same memory bandwidth.

> **Prompt 14.** What is the arithmetic intensity of the CPU implementation as a function of N? What is the effective FLOPs/s as a function of N, and is it limited by processing or communication?

> **Prompt 15.** What is the real-time Δt_{CPU} (in seconds) as a function of N?

Compare the real-time Δt's between CPU and physical. They have different relationships to N. If these curves cross, we call the intersection N_{x}, i.e.

$$\Delta t_{\mathrm{CPU}}(N_{\mathrm{x}}) = \Delta t_{\mathrm{ckt}}(N_{\mathrm{x}})$$

Prompt 16. Case 1: suppose we don't probe anything, and call the possible intersection N_{x1}. What is N_{x1}?

Prompt 17. Case 2: suppose we probe all values at every time step, and call that possible intersection N_{x2}. What is N_{x2}?