Gatsby Computational Neuroscience Unit
Combined Qualifying Examination

30 January 2018

There are 10 questions. Answer all of them, to the best of your ability. Each is worth 10 marks. No reference materials are allowed.

It should take no more than 6 hours

Good luck!

1. The just noticeable difference (JND) in the frequency of a pure tone may be assessed using (at least) two different behavioural experimental designs:

   - A 2-interval 2-alternative forced choice (2I-2AFC) task where two pure tones are heard one after the other and the subject must pick out the higher one.
   - A 3-interval 2-alternative forced choice (3I2AFC) task in which the first and last of three tones differ and the middle one matches one of them. Subjects must pick the odd tone out.

   Assume that a neural code for a sound with spectrum $a(f)$ is defined by mean activity in the ith neuron given by

   $$r_i(a) = \int df\, \phi_i(f) a(f)\,,$$

   where $\phi_i(f)$ is an effective tuning curve, but that this activity is corrupted by noise that is independent from neuron to neuron and sound to sound.

   (a) Assume first that the noise is Gaussian with constant variance, and that subjects are able to use the knowledge that each sound in the experiment is a pure tone. Derive the optimal decoding strategy for each experiment (assuming perfect knowledge of the tuning curves and perfect memory). Which design do you think will yield the smaller JND?

   (b) How would your answers change if subjects were unaware that the sounds were pure tones (and interpreted the 2I-2AFC question as being about the spectral centre of mass)?

   (c) Would the conclusions differ for Poisson noise?

2. Many decision-making experiments are based on a stimulus sequence (say a display of moving dots) which may be reduced to a scalar time-series $s_1 \ldots s_T$ (say the net horizontal velocity of the dot field). A simple model of the stimulus may be that it is randomly generated with

   $$s_t \sim \mathcal{N}(\mu, \sigma^2)\,.$$

   The correct decision is based on the sign of μ.

   Suppose that the s_t are iid, with known variance. Two possible neural integration strategies might be

   - neural firing at t represents (a monotonic function of) an optimal running estimate of μ given $s_1 \ldots s_t$.
   - neural firing at t represents (a monotonic function of) the probability or log-odds of $\mu > 0$ given $s_1 \ldots s_t$.

   (a) Assuming that the neural activity defined above is the only integration signal available in the brain, which of these two strategies will be the more accurate?

   (b) How would the recorded neural activity differ under the two strategies? (Remember to account for the possible monotonic nonlinear mappings).

   (c) Would either answer above change if σ^2 were unknown? How?

   (d) What would an optimal neural representation need to encode if the variance were only piecewise constant – i.e., it sometimes jumped to a new value mid-trial? Describe and justify a possible neural code for this case, making reference to experimental data and established coding hypotheses.

3. Sketch, and discuss the evidence for and against, the two dominant theories of the circuit origin of orientation tuning curves in V1 simple cells:

   - The Hubel and Wiesel model, in which they are determined by the pattern of feedforward connections from the LGN.
   - The lateral-interaction model in which they are shaped primarily by the fixed points of a "bump attractor" model with mexican-hat connectivity.

4. Consider a simple neuron model which receives input $x_i(t)$ on synapses $i = 1 \ldots n$, and has an instantaneous linear output $y(t) = \sum_i w_i x_i(t) + x_0(t)$, where $x_0(t)$ is an obligate input with a fixed weight of 1. Think of $x_i(t)$ and $y(t)$ as continuous rates.

   (a) a time-dependent Hebbian learning rule can be written as

   $$\Delta w_i \propto \int d\tau\ x_i(t) H(\tau) y(t + \tau)$$

   where $H(\tau)$ is the learning kernel. Justify this form, and describe a kernel associated with conventional spike-time dependent plasticity (STDP).

(b) the suggestion has been made that a rule of this form (for a more complicated neuron model) could implement a form of temporal difference learning. Specifically, if $x_0(t)$ is 0 until just before the end of a learning episode; and then 1, say. Is there a form of $H(\tau)$ for which this is (perhaps approximately) true? If so, what form?

(c) if the standard STDP kernel doesn't have this form, describe qualitatively, with justification the consequence of using it instead. Can we expect the resulting learning rule to converge stably? Explain your answer.

5. Consider the case of an agent faced with the problem of planning to make a single choice. It has two possible actions $a_0; a_1$, but is uncertain about their values. This uncertainty starts off describable by Gaussian distributions $(Q(a_0), Q(a_1)) \sim \mathcal{N}((\mu_0, \mu_1), \Sigma)$ where Σ is a covariance matrix.

   (a) Which action should the agent take if it is risk neutral and has to act without collecting further information?

   (b) Consider the case that the agent is allowed, for a cost, to reduce its uncertainty about the action value of one of the actions to 0. Write formulas (in terms of the cumulative distribution function Φ of a Gaussian distribution) for the amount it should be willing to pay to do this for action a_0. Make sure you have considered all the relevant cases.

   (c) Would it ever be worth reducing the uncertainty to 0 for *both* actions? If so, would the order of doing so matter? A qualitative argument with sketches would be fine.

6. Consider a forager in a collection of statistically identical patches, with mean travel time $\bar{\tau} > 0$ between patches. One standard foraging policy (called the giving up time policy) is to continue foraging in a patch until one has waited for time $\hat{t}$ without catching a prey, at which point one leaves immediately.

   (a) Let $F(x)$ be the probability that the intercapture interval is less than or equal to x, and let $G(x) = 1 - F(x)$. Show that:

      i. the expected number of prey $E_{\hat{t}}(Y)$ caught in a single patch is $E_{\hat{t}}(Y) = F(\hat{t})/G(\hat{t})$

      ii. the expected time in the patch $E_{\hat{t}}(R)$ is

$$E_{\hat{t}}(R) = \frac{\int_0^{\hat{t}} dx\ G(x)}{G(\hat{t})}$$

   (b) write an expression for the long run rate of prey capture given this policy

   (c) if $F(x) = 1 - e^{-\lambda x}$, show that it is not optimal to leave a patch. Suggest and justify a condition under which leaving would be optimal.

   (d) if there are n patch types, of probability $p_1 \ldots p_n$ of being encountered, and $\hat{t}_i$ is the giving up time for patch i, then if the mean travel time remains $\bar{\tau}$, write an expression for the long run rate ξ of prey capture

   (e) show that if ξ has a maximum, then this occurs where

$$\xi = \frac{dE_{\hat{t}_i}(Y_i)/d\hat{t}_i}{dE_{\hat{t}_i}(R_i)/d\hat{t}_i}$$

   for all i, where Y_i is the (random) number of prey caught in a single instance of being in a patch of type i, and R_i is the (random) time which the forager spends in a patch of type i.

   (f) if patch i has $F_i(x) = 1 - e^{-\lambda_i x}$, describe in qualitative terms the optimal policy for the forager.

7. Consider a linear integrate and fire neuron augmented by the so-called H-current,

$$\tau \frac{dV}{dt} = -(V - \mathcal{E}_L) - x(V - \mathcal{E}_{Na}) \tag{1a}$$

$$\tau_x \frac{dx}{dt} = x_\infty(V) - x \tag{1b}$$

where

$$x_\infty(V) = \frac{1}{1 + \exp\big((V - V_0)/\Delta V\big)}.$$

When the voltage reaches V_{th}, a spike is emitted and the voltage is instantaneously rest to $\mathcal{E}_L$. The numerical values of the parameters are

$$\begin{aligned} \mathcal{E}_L &= -65 \text{ mV} \\ V_{th} &= -50 \text{ mV} \\ \mathcal{E}_{Na} &= 0 \text{ mV} \\ V_0 &= -70 \text{ mV} \\ \Delta V &= 10 \text{ mV} \\ \tau_x &\gg \tau. \end{aligned}$$

(a) Draw the nullclines for $V < V_{th}$. Verify that for the above parameters, there's one intersection between the two nullclines.

(b) You "voltage clamp" the neuron at V_{clamp} (less than V_{th}) for a time much longer than τ_x, and then turn off the voltage clamp and let the dynamics evolve according to Eq. (1). (This is a model, not an experiment, so "voltage clamp" means set the voltage to V_{clamp} and let the system evolve for a long time.) What is the maximum value of V_{clamp} that will ensure a spike after $t = 0$? You may assume that τ_x is sufficiently large that x remains constant for $t > 0$, at least on the scale of spike generation.

(c) Suppose you want a spike to occur with a delay after you turn off the voltage clamp, and your only constraint is that the delay should be much greater than τ. What current could you add to Eq. (1a)? Provide qualitative activation/inactivation functions and approximate time constants, and any other parameters you think are important.

8. Consider a leaky integrate and fire neuron receiving orientation-tuned excitatory input and untuned inhibitory input from N other neurons,

$$\tau \frac{dV}{dt} = -(V - \mathcal{E}_L) - \frac{1}{N}\sum_{i=1}^{N} W_E(\theta - \theta_i) g(\theta_i - \theta_0)(V - \mathcal{E}_E) - \frac{1}{N}\sum_{i=1}^{N} W_I g(\theta_i - \theta_0)(V - \mathcal{E}_I)$$

where $\mathcal{E}_E$ and $\mathcal{E}_I$ are the excitatory and inhibitory reversal potentials, respectively, $g(\theta_i - \theta_0)$ is the angle dependent conductance due to spikes on neuron i, and $W_E(\theta - \theta_i)$ and W_I are the excitatory and inhibitory weights, respectively, from neuron i to our neuron. This model looks like it violates Dale's law, but think of the inhibitory input as coming from fast inhibitory neurons that receive input from the excitatory cells. Those inhibitory neurons are nonspecific, which is why W_I is independent of orientation. You will not need the last couple of facts to answer the questions.

Assume both the excitatory weights and the excitatory conductances have simple cosine tuning,

$$\begin{aligned} W(\phi) &\equiv W_E(1 + \cos\phi) \\ g(\phi) &\equiv g_0(1 + \cos\phi). \end{aligned}$$

(a) Given the above definitions of $W(\phi)$ and $g(\phi)$, how should you interpret θ and θ_0?

(b) Assume the θ_i are evenly distributed between 0 and 2π. In the large N limit, compute the total excitatory drive, $h(\theta, \theta_0)$, defined by

$$h(\theta, \theta_0) \equiv \frac{1}{N}\sum_{i=1}^{N} W(\theta - \theta_i) g(\theta_i - \theta_0).$$

(c) Assume the firing rate is a threshold linear function of $V - \mathcal{E}_L$. Show that if the inhibitory and leak reversal potentials are equal, $\mathcal{E}_I = \mathcal{E}_L$, and the inhibitory weights are large compared to the excitatory ones which in turn are "strong," $W_I g_0 \gg W_E g_0 \gg 1$, then the firing rate tuning curves of the above neuron first increase linearly (with respect to threshold) as the total conductance, g_0, increases, but saturate at high conductance, when g_0 is large.

9. Consider a network of N neurons that obeys the set of equations

$$\tau \frac{dx_i}{dt} = \Theta\left(\sum_{j=0}^{N} W_{ij} x_j - \theta\right) - x_i$$

where Θ is the Heaviside step function ($\Theta(x) = \max[x, 0]$), θ is a threshold, and W_{ij} is a random matrix,

$$W_{ij} = \frac{w}{N} + \frac{\sigma}{N^{1/2}} \xi_{ij}$$

with ξ_{ij} drawn *i.i.d.* from a zero mean, unit variance, Gaussian distribution,

$$\xi_{ij} \sim \mathcal{N}(0, 1).$$

For all questions, assume $N \gg 1$.

(a) Show that the population averaged firing rate

$$\overline{x} \equiv \frac{1}{N} \sum_{i=1}^{N} x_i,$$

obeys the time evolution equation

$$\tau \frac{d\overline{x}}{dt} = \Phi\left(\frac{w - \theta/\overline{x}}{\sigma \langle x^2 \rangle^{1/2}}\right) - \overline{x} \tag{1}$$

where Φ is the cumulative normal function and $\langle x^2 \rangle$ is the population averaged second moment of the firing rate.

(b) Show that if σ is sufficiently large and $\theta > 0$, there are two stable solution to Eq. (1).

(c) Show that if $w < \theta$ and $\theta > 0$, then if σ is sufficiently small there is one solution to Eq. (1).

10. Consider a set of N neurons that evolve according to

$$\tau \frac{dx_i}{dt} = -x_i + u_i \frac{1}{N} \sum_j v_j \tanh(x_j) + I_i. \tag{1}$$

Assume that u_i, v_i, and I_i are drawn from a zero mean multivariate Gaussian distribution with a covariance matrix specified by

$$\begin{aligned} \langle v_i u_i \rangle &= \rho_{vu} \\ \langle v_i I_i \rangle &= \rho_{vI} \\ \langle u_i I_i \rangle &= \rho_{uI} \\ \langle v_i^2 \rangle = \langle u_i^2 \rangle = \langle I_i^2 \rangle &= 1 \\ \langle v_i v_j \rangle = \langle u_i u_j \rangle = \langle I_i I_j \rangle = \langle v_i u_j \rangle = \langle v_i I_j \rangle = \langle u_i I_j \rangle &= 0 \text{ for } i \neq j. \end{aligned}$$

Assume the parameters are such that Eq. (1) admits a stable equilibrium. Also assume, as usual, that $N \gg 1$.

(a) Show that at equilibrium,

$$x_i = u_i \kappa + I_i$$

where

$$\kappa = \frac{1}{N} \sum_j v_j \tanh(u_j \kappa + I_j).$$

(b) Write down an equation for κ involving integrals over the distribution of u_i, v_i and I_i.

(c) Show that if $\rho_{vu} = \rho_{vI} = 0$, then $\kappa = 0$.

(d) Show that if $\rho_{vI} = \rho_{uI} = 0$ and $\rho_{vu} \neq 0$, then $\kappa \neq 0$.

(e) Show that if $\rho_{vu} = \rho_{uI} = 0$ and $\rho_{vI} \neq 0$, then $\kappa \neq 0$.