Development of a Signaling Network

Eleanor Brush

August 29, 2012

1 Model as a Stochastic Differential Equation

1.1 Derivation of the SDE from microscopic events There are three processes that can cause an animal's estimate of its dominance with respect to another to change:

1. an error can be made so that the animal's estimate randomly changes from one point in time to the next without an external stimulus

2. external observations provide evidence about each animal's dominance, e.g. by fighting each other they gather evidence about their relative strength

3. receiving a signal provides evidence that the receiver is dominant to the signaler

Assumption: For now, we will assume that the two animals have access to the same external observations, which would be the case, for instance, if they only gather evidence from fights that they both are engaged in. We can relax this assumption in the future.

Let $X_t^{(i)}$ denote animal i's estimate of its dominance at time t. We can write down equations describing how these estimates change over time:

$$X_{t+\tau}^{(1)} = X_t^{(1)} + \sum_{i=0}^{N_{e_1}(\tau)} E_i^{(1)} + \sum_{i=0}^{N_f(\tau)} F_i + b_{s_2} S_2(X_t^{(2)}, \tau)$$

$$X_{t+\tau}^{(2)} = X_t^{(2)} + \sum_{i=0}^{N_{e_2}(\tau)} E_i^{(2)} - \sum_{i=0}^{N_f(\tau)} F_i + b_{s_1} S_1(X_t^{(1)}, \tau)$$

where

- $E_i^{(j)}$ describes the magnitude of an error in estimate for animal j, which are identically distributed over time; $N_{e_j}(\tau)$ describes the number of error events in an interval of length τ so that $\sum_{i=0}^{N_{e_j}(\tau)} E_i^{(j)}$ gives the total magnitude of errors in an interval of length τ

- F_i describes the magnitude of a piece of external evidence, which are identically distributed over time; $N_f(\tau)$ describes the number of external observations in an interval of length τ so that $\sum_{i=0}^{N_f(\tau)} F_i$ gives the total magnitude of external evidence in an interval of length τ (and the evidence gathered changes the animal's estimate in opposite directions, i.e. if a fight is won, one animal's estimate is increased while the other's is decreased by the same amount)

- $S_j(X_t^{(j)}, \tau)$ gives the number of signals emitted by individual j in an interval of length τ and b_{s_j} describes the boost that each of those signals would give to animal $j + 1$'s estimate of its dominance

The fact that the two animals are privy to the same external observations is captured by the identity of the second sums in the above equations. **Assumption:** The above equations assume that τ is small enough so that X_t does not change significantly enough in the interval $[t, t + \tau)$ that signaling rates would change. (Error and fighting rates do not depend on the estimates– another assumption.) We could write the evidence from signals as a sum over several events in the interval $[t, t+\tau)$ as with errors and observations, but this is equivalent to the formulation above since we assume the size of a signal and the strength of a boost from receiving a signal is constant over time.

We can now describe the distributions from which these events are drawn:

- $N_{e_j}(\tau) \sim \mathcal{P}(r_{e_j}, \tau)$

- $E_i^{(j)} \sim \mathcal{N}(\mu_{e_j}, \sigma_{e_j}^2)$, i.i.d.

- $N_f(\tau) \sim \mathcal{P}(r_f, \tau)$

- $F_i \sim \mathcal{N}(\mu_f, \sigma_f^2)$, i.i.d.

- $S_j(X_t^{(j)}, \tau) \sim \mathcal{P}(f(X_t^{(j)}), \tau)$

where $\mathcal{P}(\lambda, \tau)$ denotes a Poisson process with rate λ and $\mathcal{N}(\mu, \sigma^2)$ denotes a Normal distribution with mean μ and σ^2. $f(x)$ gives the rate (probability) of emitting a signal as a function of the estimate x, which is decreasing in x.

We now make two approximations:

1. If Y_i, i.i.d., are drawn from some distribution and $N(\tau) \sim \mathcal{P}(\lambda, \tau)$, then $Z(\tau) = \sum_{i=0}^{N(\tau)} Y_i$ is a compound Poisson process with

$$\mathbb{E}[Z(\tau)] = \mathbb{E}[N(\tau)]\mathbb{E}[Y] = \lambda\tau\mathbb{E}[Y]$$
$$\text{and } Var(Z(\tau)) = \mathbb{E}[N(\tau)]\mathbb{E}[Y^2] = \lambda\tau\mathbb{E}[Y^2]$$

For computational convenience, we will approximate the distribution of $Z(\tau)$ by $\mathcal{N}(\lambda\tau\mathbb{E}[Y], \lambda\tau\mathbb{E}[Y^2])$. (HOW VALID IS THIS?)

2. If τ is big enough that enough events happen, then we can approximate the Poisson process $\mathcal{P}(\lambda, \tau)$ with $\mathcal{N}(\lambda\tau, \lambda\tau)$.

This allows us to rewrite our equations for X_t as:

$$X_{t+\tau}^{(1)} = X_t^{(1)} + Y_{e_1}(\tau) + Y_f(\tau) + b_{s_2} Y_{s_2}(X_t^{(2)}\tau)$$
$$X_{t+\tau}^{(2)} = X_t^{(2)} + Y_{e_2}(\tau) - Y_f(\tau) + b_{s_1} Y_{s_1}(X_t^{(1)}\tau)$$

where

- $Y_{e_j}(\tau) \sim \mathcal{N}(r_{e_j}\tau\mu_{e_j}, r_{e_j}\tau(\sigma_{e_j}^2 + \mu_{e_j}^2))$

- $Y_f(\tau) \sim N(r_f\tau\mu_f, r_f\tau(\sigma_f^2 + \mu_f^2))$

- $Y_{s_j}(\tau) \sim \mathcal{N}(f(X_t^{(j+1)})\tau, f(X_t^{(j+1)})\tau)$

Finally, if we define the following constants:

$$m_{e_j} = r_{e_j}\mu_{e_j},$$
$$n_{e_j}^2 = r_{e_j}(\sigma_{e_j}^2 + \mu_{e_j}^2),$$
$$m_f = r_f\mu_f,$$
$$n_f^2 = r_f(\sigma_f^2 + \mu_f^2),$$

then we get the following equations:

$$X_{t+\tau}^{(1)} = X_t^{(1)} + \left[m_{e_1} + m_f + b_{s_2}f(X_t^{(2)})\right]\tau + n_f\sqrt{\tau}Z^{(1)} + n_{e_1}\sqrt{\tau}Z^{(2)} + b_{s_2}\sqrt{f(X_t^{(2)})}\sqrt{\tau}Z^{(3)}$$
$$X_{t+\tau}^{(2)} = X_t^{(2)} + \left[m_{e_2} - m_f + b_{s_1}f(X_t^{(1)})\right]\tau - n_f\sqrt{\tau}Z^{(1)} + n_{e_2}\sqrt{\tau}Z^{(4)} + b_{s_1}\sqrt{f(X_t^{(1)})}\sqrt{\tau}Z^{(5)}$$

where $Z^{(1)}, \ldots, Z^{(5)} \sim \mathcal{N}(0,1)$, i.i.d., which are equivalent to the stochastic differential equations:

$$dX_t^{(1)} = \left[m_{e_1} + m_f + b_{s_2} f(X_t^{(2)}) \right] dt + n_f dW_t^{(1)} + n_{e_1} dW_t^{(2)} + b_{s_2} \sqrt{f(X_t^{(2)})} dW_t^{(3)}$$

$$dX_t^{(2)} = \left[m_{e_2} - m_f + b_{s_1} f(X_t^{(1)}) \right] dt - n_f dW_t^{(1)} + n_{e_2} dW_t^{(4)} + b_{s_1} \sqrt{f(X_t^{(1)})} dW_t^{(5)}$$

where $W_t^{(1)}, \ldots, dW_t^{(5)}$ are independent Brownian motions.

We could make a few simplifying assumptions to reduce the number of parameters. **Assumptions:**

- $r_{e_1} = r_{e_2}$

- $\mu_{e_1} = \mu_{e_2}$, and even simpler would be to have both equal 0 (expected error is 0)

- $\sigma_{e_1}^2 = \sigma_{e_2}^2$

- $b_{s_1} = b_{s_2}$

- and combining the above would give $m_{e_1} = m_{e_2}$ and $n_{e_1} = n_{e_2}$

1.2 Evolution of probability distribution over time More generally, consider a system

$$dX_t = a(X_t)dt + \sum_{i=1}^{5} \sigma_i(X_t)dW_t^{(i)}$$

with $W_t^{(1)}, \ldots, W_t^{(5)}$ independent Brownian motions. If $a, \sigma_1, \ldots, \sigma_5 : \mathbb{R}^2 \to \mathbb{R}^2$, are "nice" enough functions, and the random variable X_t has a "nice" enough density function (i.e. $\mathbb{E}[g(X_t)] = \int g(x)p(t,x)dx$ for every bounded measurable function g), then according to the Fokker-Planck / Kolmogorov forward equation

$$\frac{\partial p}{\partial t} = -\sum_{i=1}^{2} \frac{\partial}{\partial x_i}(a^{(i)}(x)p(t,x)) + \frac{1}{2} \sum_{i,j=1}^{2} \sum_{k=1}^{5} \frac{\partial^2}{\partial x_i \partial x_j}(\sigma_k^{(i)} \sigma_k^{(j)} p(t,x))$$

$$\Rightarrow \frac{\partial p}{\partial t} = -(m_{e_1} + m_f + b_{s_2} f(x_2))\frac{\partial p}{\partial x_1} - (m_{e_2} - m_f + b_{s_1} f(x_1))\frac{\partial p}{\partial x_2}$$

$$+ \frac{1}{2}(n_{e_1}^2 + n_f^2 + b_{s_2}^2 f(x_2))\frac{\partial^2 p}{\partial x_1^2} - n_f^2 \frac{\partial p}{\partial x_1 \partial x_2} + \frac{1}{2}(n_{e_2}^2 + n_f^2 + b_{s_1}^2 f(x_1))\frac{\partial^2 p}{\partial x_2^2}$$

using the functions a and σ_i as for our particular model. Note that this requires the function $f : \mathbb{R} \to \mathbb{R}$ be C^2.

1.3 Statistics of the Process WITHOUT Signaling Feedback For simplicity, we start with the case where there is no feedback from signaling to the estimates so that the model becomes

$$dX_t^{(1)} = [m_{e_1} + m_f] dt + n_f dW_t^{(1)} + n_{e_1} dW_t^{(2)}$$

$$dX_t^{(2)} = [m_{e_2} - m_f] dt - n_f dW_t^{(1)} + n_{e_2} dW_t^{(4)}$$

This is a simple drift-diffusion model. We can consider the dynamics of each estimate separately. We will show results for $X_t^{(1)}$, which follow similarly for $X_t^{(2)}$. As derived in [?], the expected time until $X_t^{(1)}$ reaches either $-T$ or T is

$$\frac{T}{m_e + m_f} \tanh\left(\frac{T(m_e + m_f)}{n_f^2 + n_e^2}\right)$$

and the probability that $X_t^{(1)}$ reaches $-T$ before it reaches T is

$$\frac{1}{1 + \exp\left(\frac{2T(m_e + m_f)}{n_f^2 + n_e^2}\right)}$$

(assuming $X_t^{(1)} = 0$).

The stationary distribution of the estimate will be (I THINK)

$$X^{(1)} \sim N(\exp((m_{e_1} + m_f)t), \frac{n_f^2 + n_e^2}{2}t) \tag{1}$$

1.4 Ornstein-Uhlenbeck Variant As shown in (1), both the mean and variance of the distribution of estimates increase with time for a pure drift-diffusion model. The dynamics are more interesting if the estimates don't just keep increasing or decreasing more or less deterministically. To add negative feedback and prevent this from happening, we can make the size of an error proportional to the current estimate:

$$dX_t^{(1)} = \left[m_{e_1} X_t^{(1)} + m_f + b_{s_2} f(X_t^{(2)})\right] dt + n_f dW_t^{(1)} + n_{e_1} dW_t^{(2)} + b_{s_2}\sqrt{f(X_t^{(2)})}dW_t^{(3)}$$

$$dX_t^{(2)} = \left[m_{e_2} X_t^{(2)} - m_f + b_{s_1} f(X_t^{(1)})\right] dt - n_f dW_t^{(1)} + n_{e_2} dW_t^{(4)} + b_{s_1}\sqrt{f(X_t^{(1)})}dW_t^{(5)}$$

The Fokker-Planck equation now becomes

$$\frac{\partial p}{\partial t} = -(m_{e_1} x_1 + m_f + b_{s_2} f(x_2))\frac{\partial p}{\partial x_1} - m_{e_1} p - (m_{e_2} x_2 - m_f + b_{s_1} f(x_1))\frac{\partial p}{\partial x_2} - m_{e_2} p$$

$$+ \frac{1}{2}(n_{e_1}^2 + n_f^2 + b_{s_2}^2 f(x_2))\frac{\partial^2 p}{\partial x_1^2} - n_f^2 \frac{\partial p}{\partial x_1 \partial x_2} + \frac{1}{2}(n_{e_2}^2 + n_f^2 + b_{s_1}^2 f(x_1))\frac{\partial^2 p}{\partial x_2^2}$$

If we ignore signaling feedback for the moment and look for a stationary distribution, we want to solve the equation

$$0 = (m_{e_1} x_1 + m_f)\frac{\partial p}{\partial x_1} - m_{e_1} p - (m_{e_2} x_2 - m_f)\frac{\partial p}{\partial x_2} - m_{e_2} p$$

$$+ \frac{1}{2}(n_{e_1}^2 + n_f^2)\frac{\partial^2 p}{\partial x_1^2} - n_f^2 \frac{\partial p}{\partial x_1 \partial x_2} + \frac{1}{2}(n_{e_2}^2 + n_f^2)\frac{\partial^2 p}{\partial x_2^2}$$

$$= (m_e x + m_f)\frac{\partial p}{\partial x} - (m_e y - m_f)\frac{\partial p}{\partial y} - 2m_e p + \frac{1}{2}(n_e^2 + n_f^2)\frac{\partial^2 p}{\partial x^2} - n_f^2 \frac{\partial p}{\partial x \partial y} + \frac{1}{2}(n_e^2 + n_f^2)\frac{\partial^2 p}{\partial y^2}$$

if we assume $m_{e_1} = m_{e_2}$

We can find a stationary distribution for each estimate individually:

$$X^{(1)} \sim N\left(-\frac{m_f}{m_e}, \frac{n_f^2 + n_e^2}{-2m_e}\right) \tag{2}$$

$$\text{and } X^{(2)} \sim N\left(\frac{m_f}{m_e}, \frac{n_f^2 + n_e^2}{-2m_e}\right) \tag{3}$$

and the covariance between the two estimates at steady state is

$$Cov(X^{(1)}, X^{(2)}) = \frac{-n_f^2}{-2m_e}.$$

2 Analyzing the Model without Stochasticity

As a first pass, we can ignore the Brownian motion / the noise and treat the model as a dynamical system:

$$\frac{dx}{dt} = m_e x + m_f + b f(y)$$
$$\frac{dy}{dt} = m_e y - m_f + b f(x)$$

If $b = 0$, there is an equilibrium at $(x, y) = \left(-\frac{m_f}{m_e}, \frac{m_f}{m_e}\right)$. If $b \neq 0$, we can find equilibrium by solving $\frac{dx}{dt} = \frac{dy}{dt} = 0$. We assume that $f(x) = 1$ for $x < -T$ and $f(x) = 0$ for $x > -T$.

1. $\left(\frac{-m_f}{m_e}, \frac{m_f}{m_e}\right)$ if $\frac{-m_f}{m_e} > -T$ and $\frac{m_f}{m_e} > -T$

2. $\left(\frac{-m_f}{m_e}, \frac{m_f - b}{m_e}\right)$ if $\frac{-m_f}{m_e} < -T$ and $\frac{m_f - b}{m_e} > -T$

3. $\left(\frac{-m_f - b}{m_e}, \frac{m_f}{m_e}\right)$ if $\frac{-m_f - b}{m_e} > -T$ and $\frac{m_f}{m_e} < -T$

4. $\left(\frac{-m_f - b}{m_e}, \frac{m_f - b}{m_e}\right)$ if $\frac{-m_f - b}{m_e} < -T$ and $\frac{m_f - b}{m_e} < -T$

Errors only provide negative feedback and prevent escalation when $m_e \leq 0$. In this case, equilibria 2. and 4. are impossible since $-m_f, -m_f - b < 0 \Rightarrow \frac{-m_f}{m_e}, \frac{-m_f - b}{m_e} > 0$ while $-T < 0$. Therefore, no equilibria where Animal 1 is signaling is possible. So we are left with

1. $\left(\frac{-m_f}{m_e}, \frac{m_f}{m_e}\right)$ if $\frac{m_f}{m_e} > -T$

2. $\left(\frac{-m_f - b}{m_e}, \frac{m_f}{m_e}\right)$ if $\frac{m_f}{m_e} < -T$

Note that if $b = 0$, the equilibria are identical. Equilibrium 1. is only possible if $\frac{m_f}{m_e} > -T$ while 2. is only possible if $\frac{m_f}{m_e} < -T$ so that the relative strength of the effects from fights and errors determines which equilibria is possible. If m_f is small, an equilibrium with neither animal signaling is possible, whereas is m_f is sufficiently large at equilibrium Animal 2 will signal Animal 1 and Animal 1's estimate is boosted by $\frac{b}{-m_e}$. In either case, Animal 2's equilibrium estimate is the same.

The Jacobian of this system at an equilibrium is

$$J = \begin{bmatrix} m_e & b f'(y^*) \\ b f'(x^*) & m_e \end{bmatrix}$$

with eigenvalues

$$\lambda = m_e \pm \frac{b}{2}\sqrt{f'(y^*)f'(x^*)}.$$

Since we have assumed that f is non-increasing everywhere, $f'(x) < 0$ everywhere and both eigenvalues are real. If $m_e < 0$, the eigenvalue with greater absolute value will be

$$\lambda = m_e - \frac{b}{2}\sqrt{f'\left(\frac{m_f}{m_e}\right)f'(x^*)}$$

since for both possible equilibria $y^* = \frac{m_f}{m_e}$. In addition, $\lambda < 0$ so that both equilibria are stable.

Let $\lambda_s = m_e - \frac{b}{2}\sqrt{f'\left(\frac{m_f}{m_e}\right)f'\left(\frac{-m_f}{m_e}\right)}$ and $\lambda_{ns} = m_e - \frac{b}{2}\sqrt{f'\left(\frac{m_f}{m_e}\right)f'\left(\frac{-m_f - b}{m_e}\right)}$ so that λ_s is the eigenvalue associated with the equilibrium where Animal 2 signals and λ_{ns} is the eigenvalue associated with the equilibrium where Animal 2 is not signaling. There are two effects of signaling feedbacks on

these eigenvalues. First, when $b \gtrless 0$, both λ_s and λ_{ns} decrease so that the decay to the equilibria is faster when the estimates are influenced by signaling behaviors.

If we further assume that f is a decreasing sigmoid, then f' is decreasing for $x < -T$ and increasing for $x > -T$ so that $0 < \left| f'\left(\frac{-m_f - b}{m_e}\right) \right| < \left| f'\left(\frac{-m_f}{m_e}\right) \right|$. Therefore, $0 > \lambda_s > \lambda_{ns}$. In general, this suggests that (SMALL!) perturbations around the non-signaling equilibrium, $\left(\frac{-m_f}{m_e}, \frac{m_f}{m_e}\right)$, will shrink faster than (SMALL!) perturbations around the signaling equilibrium, $\left(\frac{-m_f - b}{m_e}, \frac{m_f}{m_e}\right)$. This effect is greater when the transition from non-signaling to signaling behavior as a function of x is shallower: when the transition is steep, $f'(x)$ is essentially zero for $x > -T$ so that there is only a negligible difference between $f'\left(\frac{-m_f}{m_e}, \frac{m_f}{m_e}\right)$ and $f'\left(\frac{-m_f - b}{m_e}, \frac{m_f}{m_e}\right)$

3 Simulating the Stochastic Process

A simple updating algorithm allows us to simulate the process as described by the model above (ORNSTEIN-UHLENBECK VARIANT):

Fix $\Delta t > 0$. Set $t = 0$ and initialize $X_0^{(1)} = X_0^{(2)} = 0$.

1. Draw $Z_1, \ldots, Z_5$ i.i.d. from $N(0, 1)$.

2.

$$X_{t+\Delta t}^{(1)} = X_t^{(1)} + \left(m_{e_1} X_t^{(1)} + m_f + b_{s_2} f(X_t^{(2)})\right)\Delta t + n_f\sqrt{\Delta t}Z_1 + n_{e_1}\sqrt{\Delta t}Z_2 + b_{s_2}\sqrt{f(X_t^{(2)})}\sqrt{\Delta t}Z_3$$

$$X_{t+\Delta t}^{(2)} = X_t^{(2)} + \left(m_{e_2} X_t^{(2)} + m_f + b_{s_1} f(X_t^{(1)})\right)\Delta t + n_f\sqrt{\Delta t}Z_1 + n_{e_2}\sqrt{\Delta t}Z_4 + b_{s_1}\sqrt{f(X_t^{(1)})}\sqrt{\Delta t}Z_5$$

3. $t = t + \Delta t$

To confirm that our simulation is performing reasonably, we let the model run without signaling feedback until the estimates reach stationary distributions and compare these to the analytical prediction, as given by (2) and (3). This comparison is shown in Figure 1. As can be seen in the figure, the observed probability distribution of estimates is statistically indistinguishable from the prediction (for the first, K-S test $D = 0.0243$, $p = 0.5969$ and for the second, $D = 0.0211$, $p = 0.766$).

We are interested in the effect of signaling feedback, i.e. the b_s parameter, on the dynamics of the model. We run the algorithm 1000 times so that we obtain, at each point in time, a distribution of each animal's estimates. We then compare these distributions when there is and isn't signaling. In both regimes, each animal's estimates reach a relatively stationary distribution fairly quickly so that we can pick a typical time point at which the distributions are not changing any more. (A Kolmogorov-Smirnov test for the differences between distributions at different points in time does not allow us to reject the null hypothesis that the distributions are the same, at a significance level of $p = 0.05$.) If we allow the simulation to reach a stationary state, we find that the stationary distributions with and without signaling feedback ($b_{s_1} = b_{s_2} = 0$ and $b_{s_1} = b_{s_2} \gtrless 0$) are quite different for Animal 1 (K-S test statistic, $D = 0.995, p < 2.2 \times 10^{-16}$), whereas they are statistically indistinguishable for Animal 2 ($D = 0.048, p = 0.1995$) (Figure (2)).

By comparing distributions of the estimates at a given point in time, we compare the behavior of the stochastic process over many realizations. We can also compare the behavior of individual trajectories over time. We hypothesize that by increasing the strength of signaling feedback, we should increase the stability of signaling behavior and it should be less likely that the dominant animal's estimate sink to such an extent that it becomes subordinate. To test this hypothesis, we count, for each run of the simulation, the number of switches between dominance regimes. A dominance regime is defined by one animal's estimate being below the signaling threshold $-T$ and the other animal's estimate being above the positive threshold T. If a dominance regime is established with Animal i's estimate above the threshold (dominant) and Animal $i+1$'s estimate below the threshold (subordinate) and subsequently a dominance regime is established with Animal i subordinate and Animal $i+1$ dominant, we count
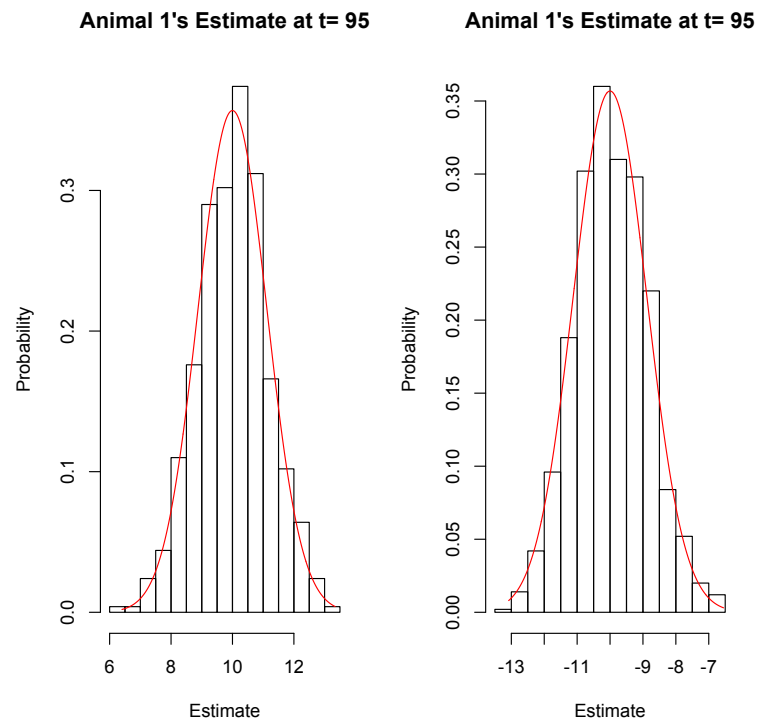
Figure 1: A comparison between the stationary distribution from a simulation of the Ornstein-Uhlenbeck model and the analytically predicted stationary distribution.

Animal 1's Estimate at t= 50   **Animal 2's Estimate at t= 50**
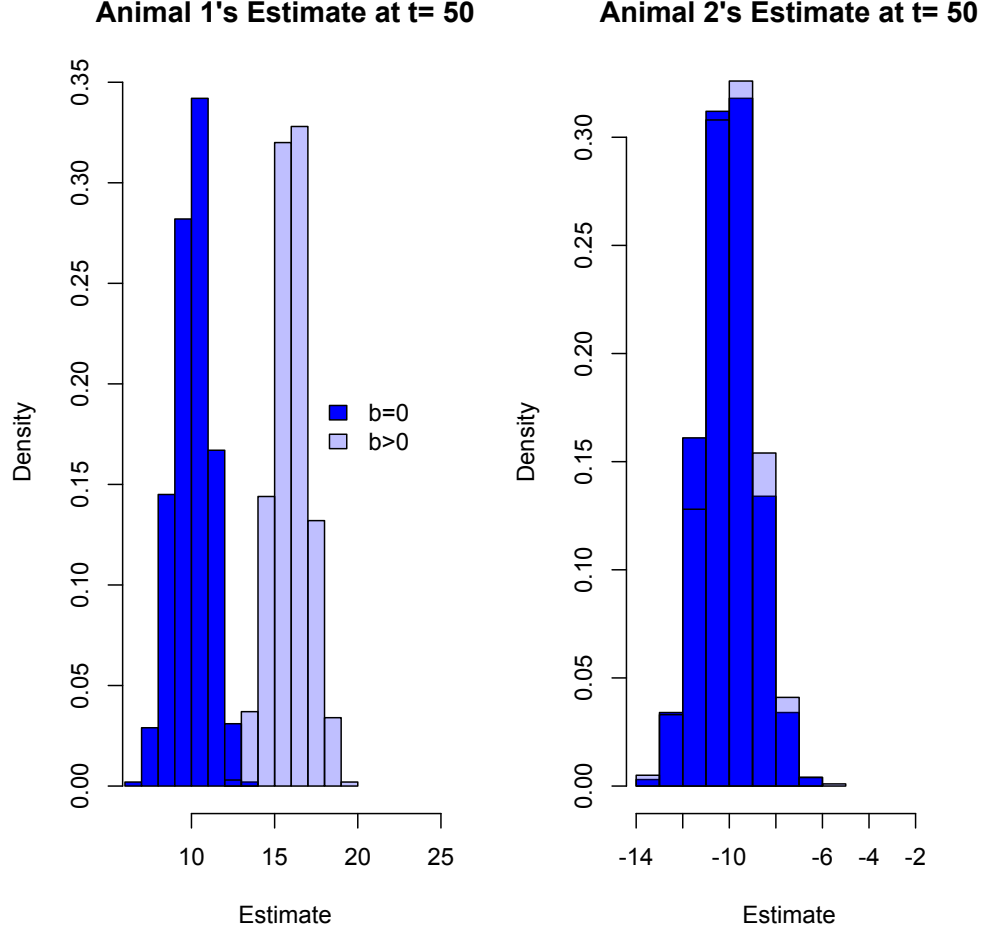
Figure 2: The distribution of estimates for each animal at time 200, for no signaling feedback and for signaling feedback $b_{s_1} = b_{s_2} = 3$. Other parameters are set as follows: threshold $T = 5$, $m_{e_1} = m_{e_2} = -.5$, $n_{e_1} = n_{e_2} = 1$, $m_f = 5$, $n_f = .5$.

this as one switch. Fixing the other parameters of the model, for each difference in fighting ability m_f and level of signaling feedback $b_{s_1} = b_{s_2}$, we can find the average number of switches that occur in a run of the model. As the difference in fighting ability increases fewer switches occur since Animal 1 is dominant the vast majority of the time. Additionally, the average number of switches between dominance regimes decreases as b_s increases, and the effect is greater for when the difference in fighting ability is smaller (Figure 3).

4   Linear Learning Formulation

So far, we have considered how the processes of engaging in fights and sending and receiving signals might affect the estimates two animals make of each other's dominance. We are interested in understanding these processes on the population level and therefore would like to embed these dynamics in a model of a social network. In order to make this more analytically tractable, we modify our model into a linear learning model and, for the moment, ignore feedback from signaling into the learning process. Our main question is how the structure of a social network can facilitate individuals' learning about
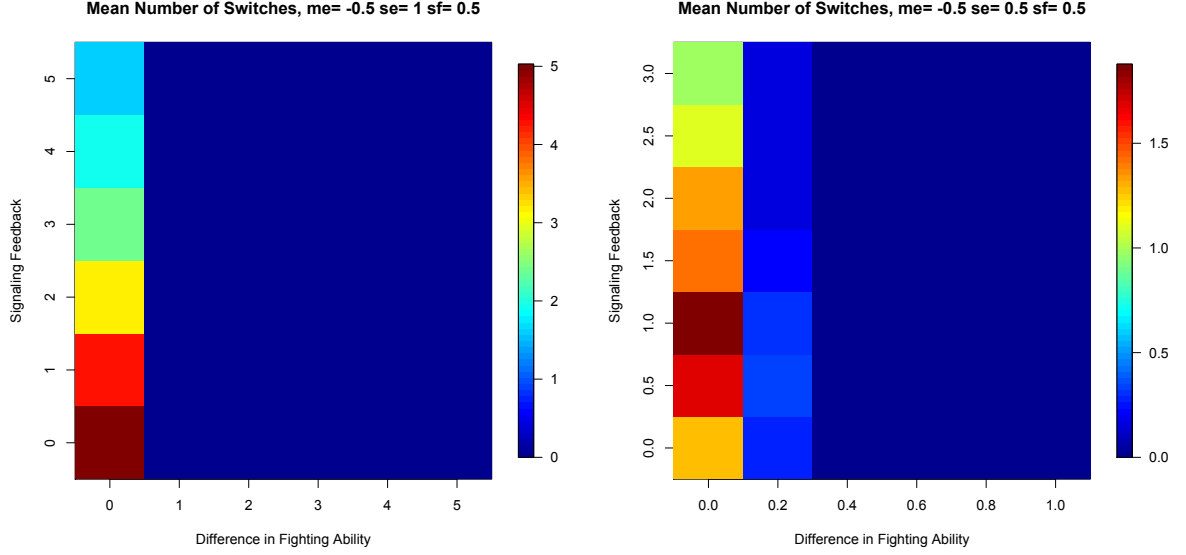
(a)
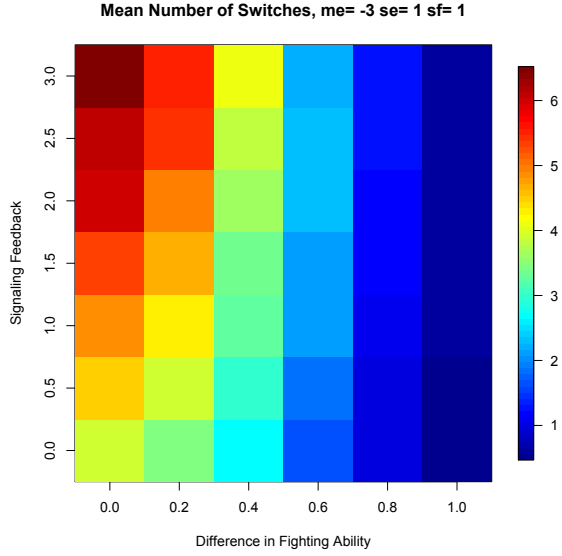
(b)

Figure 3: The average number of switches between dominance regimes as a function of the difference in fighting ability, m_f, between the animals and the degree to which receiving a signal affects an animal's estimate of its dominance, $b_{s_1} = b_{s_2}$. In general, even for small differences in fighting ability there are very few switches in dominance regimes. However, fighting ability can be made small enough to allow switches to occur, in which case the strength of signaling feedbacks affects how frequently they happen. When the leak rate is relatively low, increases in signaling feedback decrease the number of switches (3a). When the error rate is high / estimates leak back to 0 very quickly, increasing the signaling feedback actually increases the number of switches between dominance regimes / decreases the stability of a dominance regime (3b).

the states of other nodes in the network.

Suppose each individual has a true fighting ability $a_1, \ldots, a_N$ and let $f_{jk} = a_j - a_k$. Consider an individual in a social network observing its own interactions with other members of the group and the interactions between other members of the group and trying to learn each individual's true fighting ability. Let $x_j(t)$ be the individual's estimate of individual j's fighting ability at time t. The focal individual will be able to learn about other individual's fighting ability by engaging in fights with individuals it interacts with and observing fights between pairs of individuals both of whom it interacts with. Let

$$I_{jk} = \begin{cases} 1 & , \quad \text{if the focal individual interacts with both } j \text{ and } k \\ 0 & , \qquad\qquad\qquad \text{otherwise} \end{cases}.$$

Now, x_j will change according to

$$\dot{x}_j(t) = \sum_k I_{jk} \left[f_{jk} - (x_j(t) - x_k(t)) \right] \text{ for } j \lessgtr N$$

$$\text{and } x_N(t) = T - \sum_{j \lessgtr N} x_j(t),$$

where we impose a total on the estimates so that they do not diverge. Therefore,

$$\dot{x}_j(t) = \sum_{k \neq N} I_{jk} \left[f_{jk} - (x_j(t) - x_k(t)) \right] + I_{jN} \left[f_{jN} - (x_j(t) - x_N(t)) \right]$$

$$= \sum_{k \neq N} I_{jk} \left[f_{jk} - (x_j(t) - x_k(t)) \right] + I_{jN} \left[f_{jN} - (x_j(t) - (T - \sum_{k \neq N} x_k(t))) \right]$$

$$= \sum_k I_{jk} f_{jk} - x_j(t) \sum_k I_{jk} + \sum_{k \neq N} I_{jk} x_k(t) + T I_{jN} - \sum_{k \neq N} I_{jN} x_k(t)$$

$$= \sum_k I_{jk} f_{jk} + T I_{jN} - x_j(t) \sum_k I_{jk} + \sum_{k \neq N} (I_{jk} - I_{jN}) x_k(t)$$

$$= C_j + \sum_k L_{jk} x_k(t) \text{ where } C_j = \sum_{k \neq N} I_{jk} f_{jk} + T I_{jN} \text{ and}$$

L_{jk} is an $n - 1 \times n - 1$ matrix given by

$$L_{jk} = \begin{cases} -\sum_l I_{jl} + I_{jj} - I_{jN} & , \quad \text{if } j = k \\ I_{jk} - I_{jN} & , \quad\quad \text{if } j \neq k \end{cases}.$$

If $x_k(t) = a_k + \frac{T - \sum_k a_k}{N}$ for all k, $\dot{x}(t) = 0$. (If $T = \sum_k a_k$ then $x_k(t) = a_k$ is an equilibrium.)

Stability of the equilibrium is determined by the eigenvalues of L.

5   To Do

- More mathematical analysis.

  - We know that if signaling is constantly present or absent, the stationary distribution of the Ornstein-Uhlenbeck process is just shifted and the variance does not change. However, we do not know what the likelihood of switching from non-signaling to signaling behavior (or vice versa) and whether the estimates would stay in a non-signaling / signaling regime long enough for a stationary distribution to be established.

- Embed the signaling dynamics in a network.

  - What's the relationship between the distribution of fighting abilities and the distribution of signaling behaviors? Between fighting abilities and the distribution of power? And how does that depend on the metric used to calculate power?

– What's the timescale separation between fighting abilities / fight outcomes / signaling behaviors and the power distribution?

  * when formulated as a Langevin equation: measure timescales through the expected escape time using martingales
  * when formulated as a dynamical system: measure timescales through time constants of potential wells.

– What's the / is there a tradeoff between mutual information between fighting abilities and power distribution and stability of the power distribution?

- Compare to data.

  – What are the relative rates of fighting and signals observed / predicted?

  – Can we explain the data with our model?

  – Do the monkeys weight fights won and lost equally? Are they trying to maximize reward, minimize punishment, or are the two weighted equally?

  – Are the monkeys trying to minimize error tae or decision time? Can we tell if they're being forced to signal?

  – Is there a sense in which the monkeys are performing optimally?

  – (ALSO THEORY QUESTION) How efficiently are the monkeys representing /using the information they're observing about the world?

  – If we can estimate parameters like the thresholds for signaling, how are they related to other variables in the system?