Development of a Signaling Network

Eleanor Brush
October 03, 2012

Contents

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.

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

- $\mu_{e_1} = \mu_{e_2}$, 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:

$$\frac{dx}{dt} = -ex + f + br(y + T)$$

$$\frac{dy}{dt} = -ey - f + br(x + T)$$

The isoclines of this system are given by

$$\dot{x} = 0 \Rightarrow x = \frac{f}{e} + \frac{b}{e}r(y + T)$$

$$\dot{y} = 0 \Rightarrow y = -\frac{f}{e} + \frac{b}{e}r(x + T)$$

The Jacobian of this system is given by

$$J(x, y) = \begin{pmatrix} -e & br'(y + T) \\ br'(x + T) & -e \end{pmatrix}$$

Letting $\tau = Tr(J(x, y))$ and $\Delta = \det(J(x, y))$, we see that $\tau = -2e$ and $\Delta = e^2 - b^2 r'(y + T)r'(x + T)$. Since $e > 0$, $\tau < 0$ and any fixed point will be a saddle point, a stable sink, or a stable spiral, depending on the sign of Δ.

$$\Delta < 0 \Leftrightarrow e^2 - b^2 r'(y + T)r'(x + T) < 0$$

$$\Leftrightarrow e < b\sqrt{r'(y + T)r'(x + T)} \text{ since } e, b, \text{ and } r'(y + T)r'(x + T) \geq 0$$

$$\Leftrightarrow \begin{cases} b > 0 & \text{if } e = 0 \\ b > 2\varepsilon e & \text{if } e > 0 \end{cases}$$

Therefore, if there is a fixed point with $r'(y + T) \neq 0$ and $r'(x + T) \neq 0$ (i.e. the isoclines cross where their slopes are not zero), the fixed point is a saddle and there are two other stable fixed points only if $b > 2\varepsilon e$. Other conditions are required, as well. In the following, we let $c = \frac{f}{e}$. Then we also need

$$-T + \varepsilon < -c + \frac{b}{e}r(c + T) \tag{4}$$

$$T \quad \quad b \quad \quad b \tag{5}$$

These lead to the following conditions on $\frac{f}{e} = c$:

$$(4) \Rightarrow -T + \varepsilon < -c - \frac{b}{2\varepsilon e}c - \frac{bT}{2\varepsilon e} + \frac{b}{2e}$$

$$\Rightarrow c < \left(\frac{2\varepsilon e - b}{2\varepsilon e + b}\right)T - \frac{2\varepsilon^2 e}{2\varepsilon e + b} + \frac{\varepsilon b}{2\varepsilon e + b} =: M \tag{6}$$

$$(5) \Rightarrow -T - \varepsilon > -c - \frac{b}{2\varepsilon e}c - \frac{b^2}{2\varepsilon e^2} - \frac{bT}{2\varepsilon e} + \frac{b}{2e}$$

$$\Rightarrow c > \left(\frac{2\varepsilon e - b}{2\varepsilon e + b}\right)T + \frac{2\varepsilon^2 e - b^2/e}{2\varepsilon e + b} + \frac{\varepsilon b}{2\varepsilon e + b} =: m \tag{7}$$

Note that $m < M$ if and only if $b > 2\varepsilon e$, so that m and M are valid lower and upper bounds on x only if b is sufficiently large for a saddle point to exist. Equations (6) and (7) can be rewritten to provide bounds on T that allow a bifurcation to occur:

$$(6) \Rightarrow (2\varepsilon e - b)T > (2\varepsilon e + b)c + 2\varepsilon^2 e - \varepsilon b$$

$$\Rightarrow T < \frac{(2\varepsilon e + b)c + 2\varepsilon^2 e - \varepsilon b}{2\varepsilon e - b}$$

$$\Rightarrow T < \frac{2\varepsilon e(c + \varepsilon) + b(c - \varepsilon)}{2\varepsilon e - b} \tag{8}$$

$$(7) \Rightarrow (2\varepsilon e - b)T < (2\varepsilon e + b)c - 2\varepsilon^2 e + b^2/e - \varepsilon b$$

$$\Rightarrow T > \frac{(2\varepsilon e + b)c - 2\varepsilon^2 e + b^2/e - \varepsilon b}{2\varepsilon e - b}$$

$$\Rightarrow T > \frac{2\varepsilon e(c - \varepsilon) + b(c - \varepsilon + b/e)}{2\varepsilon e - b} \tag{9}$$

If the equivalent conditions (4),(6),(8) are violated, there can be no stable fixed point with animal 1 signaling at a nonzero rate and animal 2 not signaling because the difference in fighting ability is too great or the threshold below which signaling occurs is too low.

Another interesting phenomenon is the possibility of each animal signaling the other at a non-zero rate. In order for this to occur, we need $\dot{x}(x, -\frac{f}{e}) < 0$ for al $x > x^*$ and in particular for $x = -T + \varepsilon$ so that

$$-e(-T + \varepsilon) + f + br(-\frac{f}{e} + T) < 0$$

$$\Rightarrow -T + \varepsilon > c + \frac{b}{e}r(-\frac{f}{e} + T)$$

$$\Rightarrow -T + \varepsilon > c + \frac{b}{e}\left(\frac{c}{2\varepsilon} - \frac{T}{2\varepsilon} + \frac{1}{2}\right)$$

$$\Rightarrow \frac{(b - 2\varepsilon e)}{2\varepsilon e}T > c - \varepsilon + \frac{bc}{2\varepsilon e} + \frac{b}{2e}$$

$$\Rightarrow (b - 2\varepsilon e)T > 2\varepsilon e(c - \varepsilon) + bc + b\varepsilon$$

$$\Rightarrow T < \frac{2\varepsilon e(\varepsilon - c) - b(c + \varepsilon)}{2\varepsilon e - b}$$

where the last inequality follows because this can only happen if $b < 2\varepsilon e$. (The condition $b < 2\varepsilon e$ is equivalent to the slope of the x isocline being steeper than the slop of the y isocline where they have non-zero slopes. Otherwise, if x and y are both signaling at non-zero rates, a bifurcation will occur.)

Note that if $f > 0$ so that $\frac{f}{e} > -T - \varepsilon$, there can be no steady state with animal 1 signaling at rate $= 1$. The possible steady states are as follows:

1. Neither animal signaling:

$$\Rightarrow -T + \varepsilon < \frac{f}{e} \text{ and } -T + \varepsilon < -\frac{f}{e}$$

$$\Rightarrow T > \varepsilon - \frac{f}{e} \text{ and } T > \varepsilon + \frac{f}{e}$$

$$\Rightarrow T > \varepsilon + \frac{f}{e}$$

2. Animal 1 not signaling, animal 2 signaling at a rate $\in (0, 1)$:

   (a)

$$-T - \varepsilon < -\frac{f}{e}$$

$$\Rightarrow T > \frac{f}{e}$$

(b)

$$\frac{f}{e} + \frac{b}{e}r\left(-\frac{f}{e}+T\right) > -T + \varepsilon > \max\left\{-\frac{f}{e} + \frac{b}{e}r\left(\frac{f}{e}+T\right), -\frac{f}{e}\right\}$$

$$\Rightarrow \varepsilon - \frac{f}{e} - \frac{b}{e}r\left(-\frac{f}{e}+T\right) < T < \varepsilon - \max\left\{-\frac{f}{e} + \frac{b}{e}r\left(\frac{f}{e}+T\right), -\frac{f}{e}\right\}$$

$$\Rightarrow \varepsilon - \frac{f}{e} - \frac{b}{e}r\left(-\frac{f}{e}+T\right) < T < \min\left\{\varepsilon + \frac{f}{e} - \frac{b}{e}r\left(\frac{f}{e}+T\right), \varepsilon + \frac{f}{e}\right\}$$

$$\Rightarrow \begin{cases} \frac{2\varepsilon e(\varepsilon - \frac{f}{e}) + b(-\frac{f}{e} - \varepsilon)}{2\varepsilon e - b} < T < \varepsilon + \frac{f}{e} & \text{if } b < 2\varepsilon e \\ \frac{2\varepsilon e(\varepsilon + \frac{f}{e}) + b(\frac{f}{e} - \varepsilon)}{2\varepsilon e - b} < T < \varepsilon + \frac{f}{e} & \text{if } b > 2\varepsilon e \end{cases}$$

3. Animal 1 not signaling, animal 2 signaling at rate = 1:

   (a)

   $$-T - \varepsilon > -\frac{f}{e}$$

   $$\Rightarrow T < \frac{f}{e} - \varepsilon$$

   Therefore, for the following condition we assume that $\frac{f}{e} - \varepsilon > 0$ so that $\varepsilon - \frac{f}{e} < 0$.

   (b)

   $$\frac{f}{e} + \frac{b}{e} > -T + \varepsilon > -\frac{f}{e} + \frac{b}{e}r\left(\frac{f}{e}+T\right)$$

   $$\Rightarrow \varepsilon - \frac{f}{e} - \frac{b}{e} < T < \varepsilon + \frac{f}{e} - \frac{b}{e}r\left(\frac{f}{e}+T\right)$$

   $$\Rightarrow T \begin{cases} < \frac{2\varepsilon e(\varepsilon + \frac{f}{e}) + b(\frac{f}{e} - \varepsilon)}{2\varepsilon e - b} & \text{if } b < 2\varepsilon e \\ > \frac{2\varepsilon e(\varepsilon + \frac{f}{e}) + b(\frac{f}{e} - \varepsilon)}{2\varepsilon e - b} & \text{if } b > 2\varepsilon e \end{cases}$$

   $$\Rightarrow 0 < T < \frac{f}{e} - \varepsilon$$

4. Both animals signaling at a non-zero rate, but only one steady state:

   (a) $b < 2\varepsilon e$ ensures that there is no unstable equilibrium, i.e. there is only one equilibrium

   (b)

   $$-T + \varepsilon > -\frac{f}{e} + \frac{b}{e}r(-\frac{f}{e}+T)$$

   $$\Rightarrow T < \frac{2\varepsilon e(\varepsilon - \frac{f}{e}) - b(\frac{f}{e} + \varepsilon)}{2\varepsilon e - b}$$

5. Two stable equilibria, one with animal 1 signaling at a non-zero rate and animal 2 not signaling and one with animal 1 not signaling and animal 2 signaling at a non-zero rate:

   (a) $b > 2\varepsilon e$

   (b)

   $$-T + \varepsilon < -\frac{f}{e} + \frac{b}{e}r\left(\frac{f}{e}+T\right)$$

   $$\Rightarrow T > \varepsilon + \frac{f}{e} - \frac{b}{e}r\left(\frac{f}{e}+T\right)$$

   $$\Rightarrow T < \frac{2\varepsilon e(\varepsilon + \frac{f}{e}) + b(\frac{f}{e} - \varepsilon)}{2\varepsilon e - b}$$

   (c)

   $$-T - \varepsilon > -\frac{f}{e} + \frac{b}{e}r\left(\frac{f}{e} + \frac{b}{e} + T\right)$$

   $$\Rightarrow T > \frac{2\varepsilon e(\frac{f}{e} - \varepsilon) + b(\frac{f}{e} + \frac{b}{e} - \varepsilon)}{2\varepsilon e - b}$$

   $$\rightarrow T > 0 \text{ if } b > 2\varepsilon e \text{ and } \frac{f}{e} <$$

(Two stable equilibria, one with animal 1 signaling at a non zero rate and animal 2 not signaling and the other with animal 1 not signaling and animal 2 signaling at rate $= 1$ would require $T > \varepsilon - \frac{f}{e}$ and $T < \varepsilon - \frac{f}{e}$, which is clearly not possible.)

6. A single steady state with animal 1 signaling at a non-zero rate is not possible.

Let

$$\psi(b) = \frac{2\varepsilon e(\varepsilon - \frac{f}{e}) - b(\frac{f}{e} + \varepsilon)}{2\varepsilon e - b} \text{ and } \varphi(b) = \frac{2\varepsilon e(\varepsilon + \frac{f}{e}) + b(\frac{f}{e} - \varepsilon)}{2\varepsilon e - b}$$

Then, there are three general cases, depending on the magnitude of $\frac{f}{e}$ (Figure 1):

1. $\frac{f}{e} < \varepsilon$:

   (a) neither animal signaling if $T > \varepsilon + \frac{f}{e}$

   (b) animal 1 not signaling and animal 2 signaling at a non-zero rate if $T < \varepsilon + \frac{f}{e}$ and

   $$T > \begin{cases} \varphi(b) & \text{if } b > 2\varepsilon e \\ \psi(b) & \text{if } b < 2\varepsilon e \end{cases}$$

   (c) both animals signal at a non-zero rate if $b < 2\varepsilon e$ and $T < \psi(b)$

   (d) two stable equilibria if $b > 2\varepsilon e$ and $T < \varphi(b)$

2. $\frac{f}{e} > \varepsilon$:

   (a) neither animal signaling if $T > \varepsilon + \frac{f}{e}$

   (b) animal 1 not signaling and animal 2 signaling at a non-zero rate if $T < \varepsilon + \frac{f}{e}$ and $T > \frac{f}{e} - \varepsilon$

   (c) animal 1 not signaling and animal 2 signaling at a rate $= 1$ if $T < \frac{f}{e} - \varepsilon$

In summary, when the difference in fighting ability is low, there is more uncertainty about which animal is dominant and there can even be two stable equilibria. Once the difference in fighting ability is great enough, there can never be a stable equilibrium where the stronger animal thinks he is subordinate, but the equilibrium estimate of the weaker animal depends on the threshold below which signaling occurs.

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 2. 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 $b_{s_1} = b_{s_2} \geq 0$) are quite different for Animal 1(K-S test statistic, $D = 0.995, p < 2.2 \times 10^{-16}$), whereas they are statistically
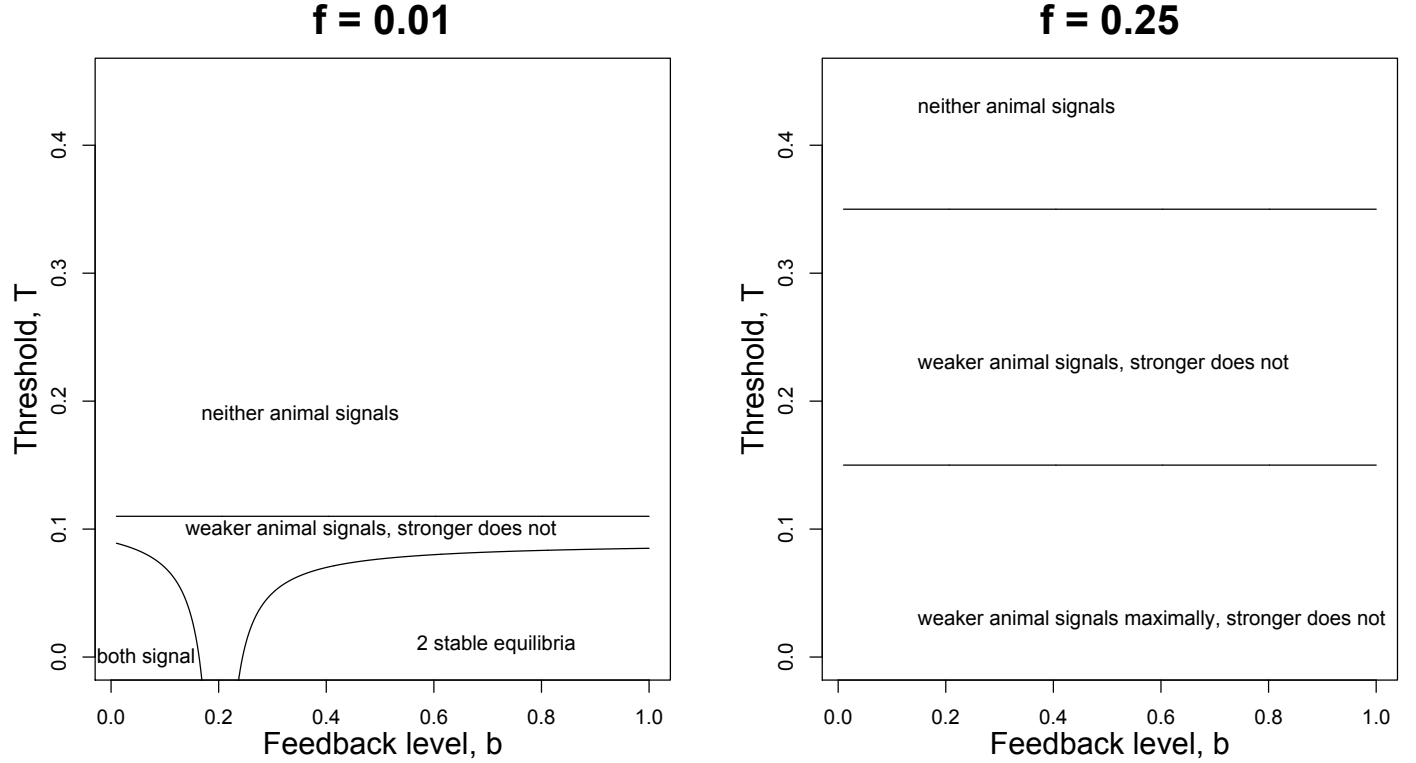
Figure 1: **The equilibria of the signaling model as a function of feedback level, b, and signaling threshold, T.** For both plots, $e = .5$, $\varepsilon = .1$. As f increases from less than εe to greater than εe, the possible equilibria change.

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 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 4).

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 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

$$I_{jk} = \begin{cases} 1 & , \quad \text{if the focal individual interacts with both } j \text{ and } k \\ 0 & \quad \text{otherwise} \end{cases}$$
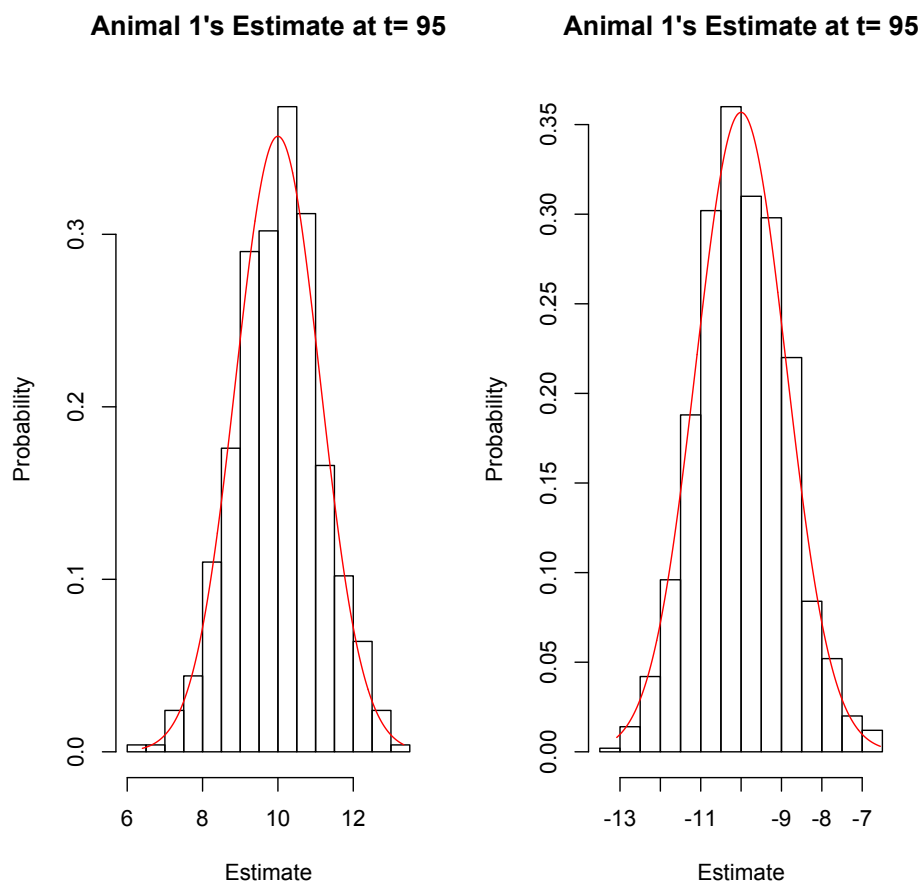
Figure 2: A comparison between the stationary distribution from a simulation of the Ornstein-Uhlenbeck model and the analytically predicted stationary distribution.
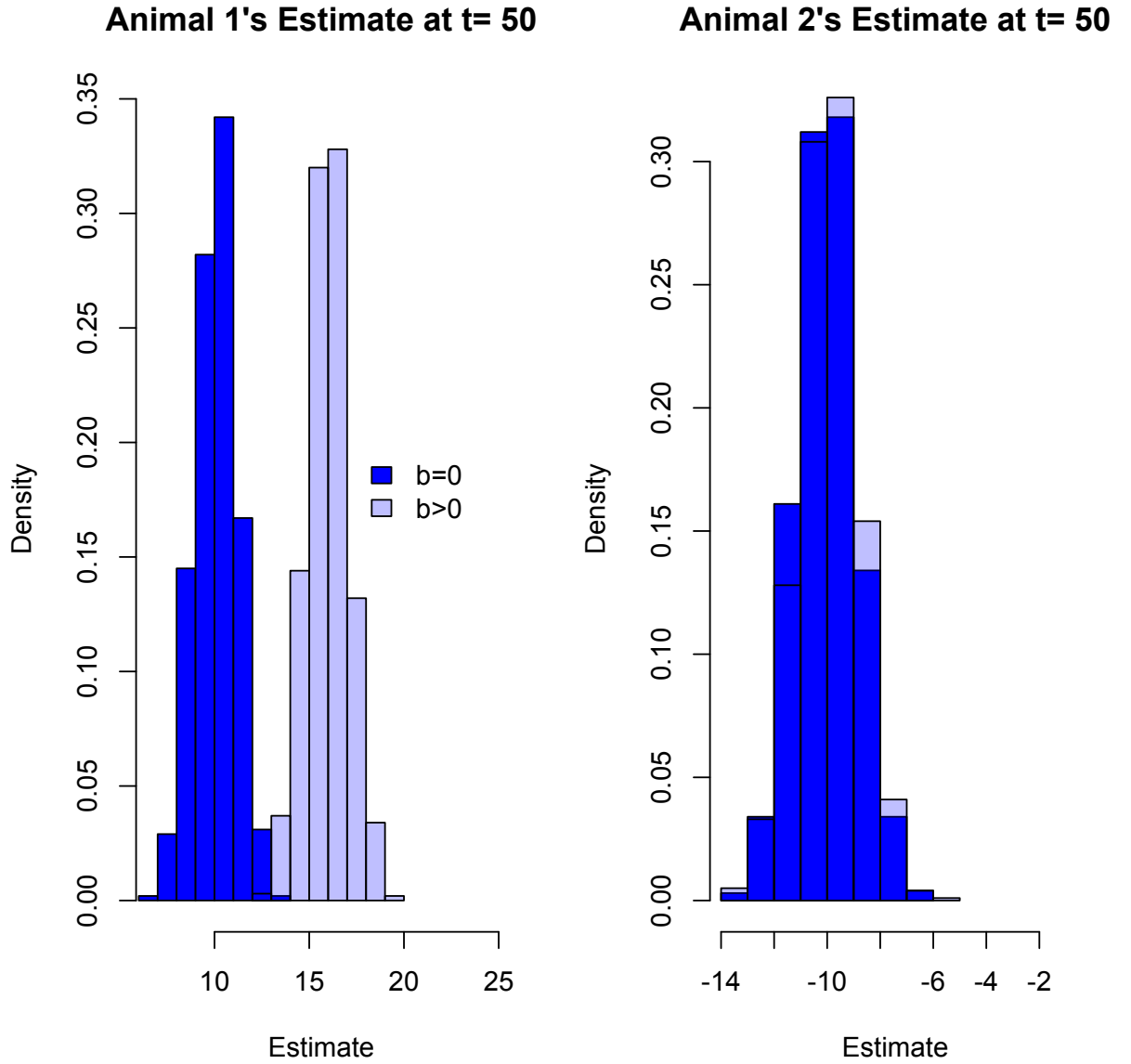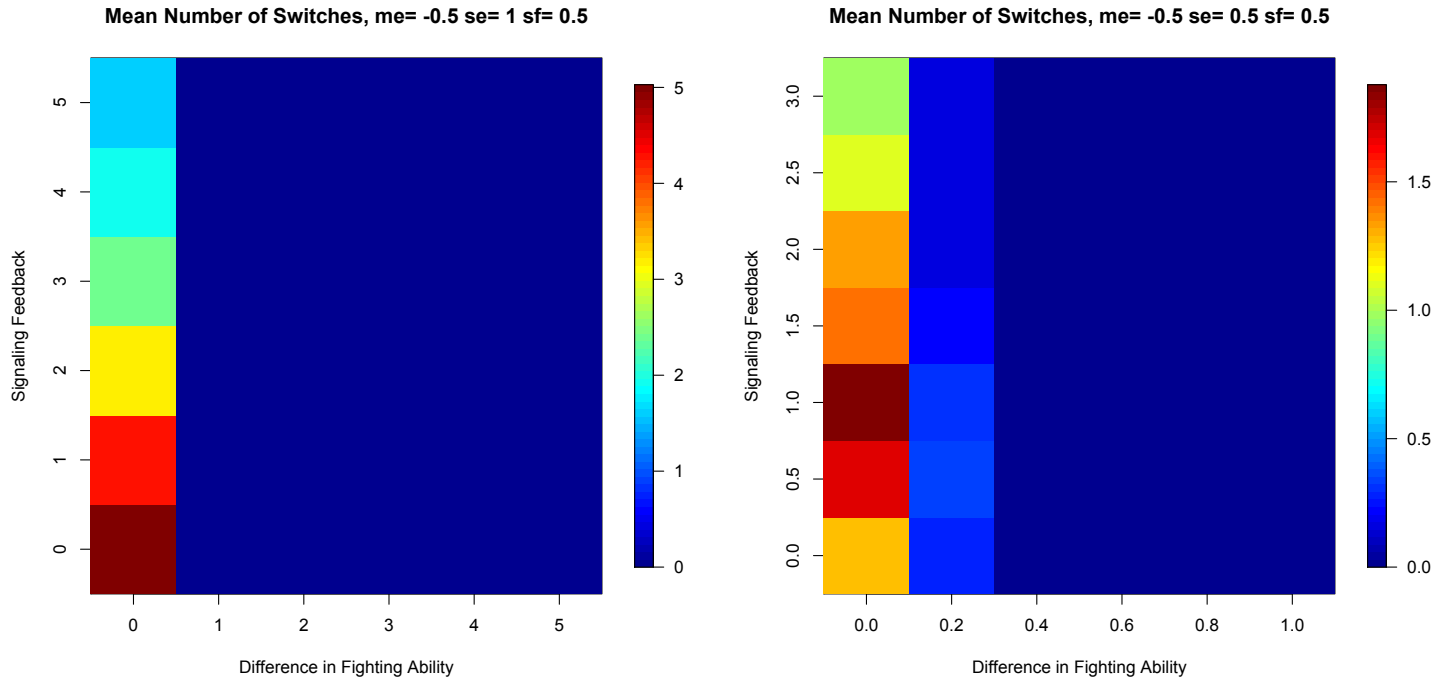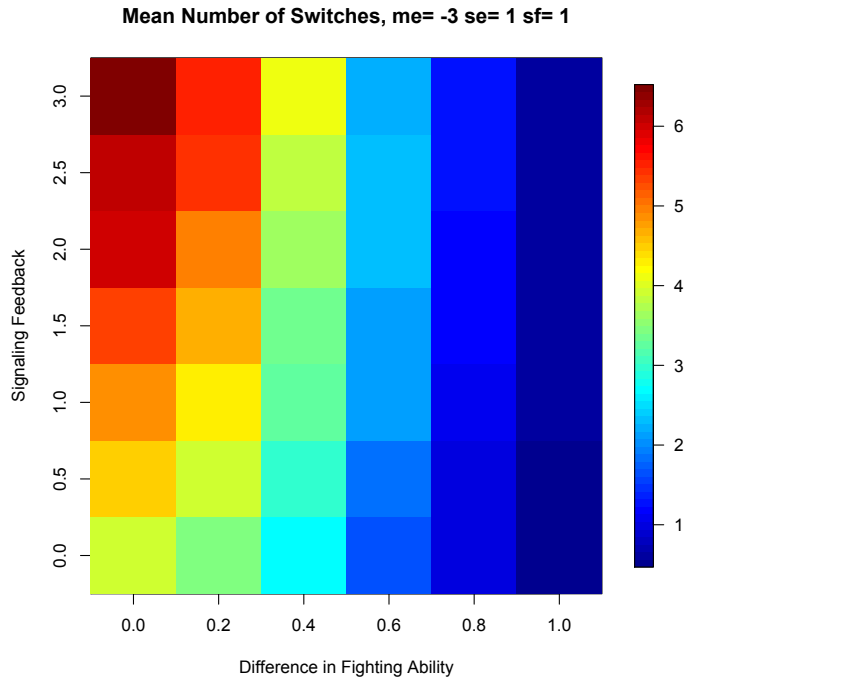
Figure 3: The distribution of estimates for each animal at time 200, for no signaling feedback and for signaling feedback $b_{s_1} = b_{s_2} = 3$. Other parameters are set as follows: threshold $T = 5$, $m_{e_1} = m_{e_2} = -.5$, $n_{e_1} = n_{e_2} = 1$, $m_f = 5$, $n_f = .5$.

Mean Number of Switches, me= -0.5 se= 1 sf= 0.5

Mean Number of Switches, me= -0.5 se= 0.5 sf= 0.5

(a)

Mean Number of Switches, me= -3 se= 1 sf= 1

(b)

Figure 4: 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 (4a). 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 (4b).

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

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

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?