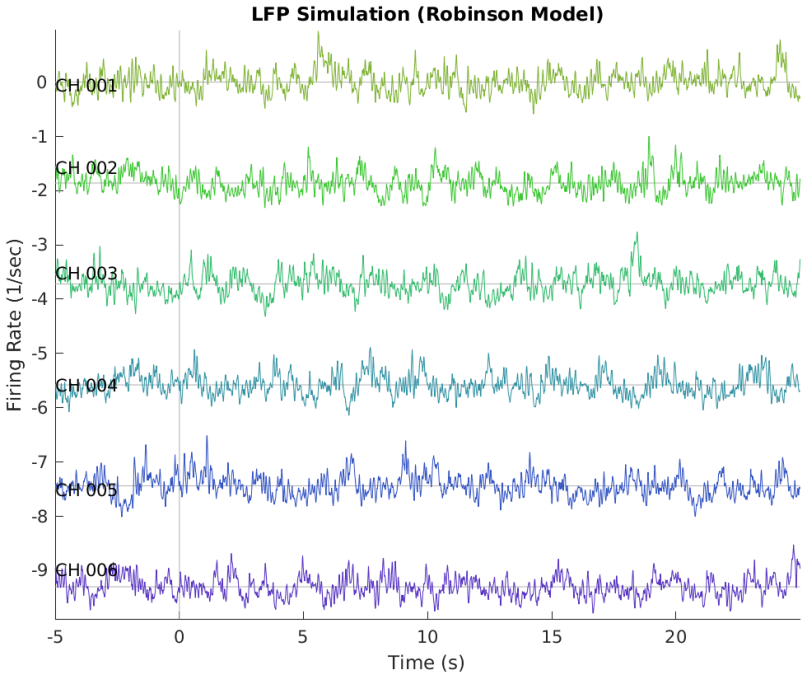
Data Synthesis Using the Extended Robinson Model

Written by Christopher Thomas – January 25, 2024.



Contents

Chapter 1

Overview

This document describes a model used to simulate interacting neural populations in the cortex and thalamus. The output of the model is mean firing rates of these neural populations, with noise-excited oscillations in the alpha band.

The original model was described and rigorously analyzed in Robinson 2002. Freyer 2011 describes an extension to the model where the noise input is modulated by the cortex population's output, providing a closer match with the oscillation spectra in biological data. Hindriks 2023 describes an extension to the model that adds multiple independent copies of the Robinson and Freyer model with coupling between instances. This is used to model co-oscillation of different brain regions.

The version of the model described in Hindriks 2023 is referred to in this document as the "extended Robinson model".

Chapter 2 describes the details of the model. Chapter 3 gives a summary of how the model is used; for a detailed treatment, consult the reference manual and sample code. Chapter 4 provides a detailed mathematical analysis of the model.

Reference details:

- P. A. Robinson, C. J. Rennie, and D. L. Rowe, *Dynamics of Large-Scale Brain Activity in Normal Arousal States and Epileptic Seizures*, Physical Review E, 65, 041924, April 2002
- F. Freyer, J. A. Roberts, R. Becker, P. A. Robinson, P. Ritter, and M. Breakspear, *Biophysical Mechanisms of Multistability in Resting-State Cortical Rhythms*, Journal of Neuroscience, 31, pp 6353–6361, April 2011
- R. Hindriks and P. K. B. Tewarie, *Dissociation Between Phase and Power Correlation Networks in the Human Brain is Driven by Co-Occurrent Bursts*, Communications Biology, 6, 286, March 2023

Chapter 2

Model Details

A diagram of the Robinson model with the extensions from Freyer 2011 and Hindriks 2023 is shown in Figure 2.1. To distinguish this from the Robinson 2002 model, this will be referred to as "the extended Robinson model".

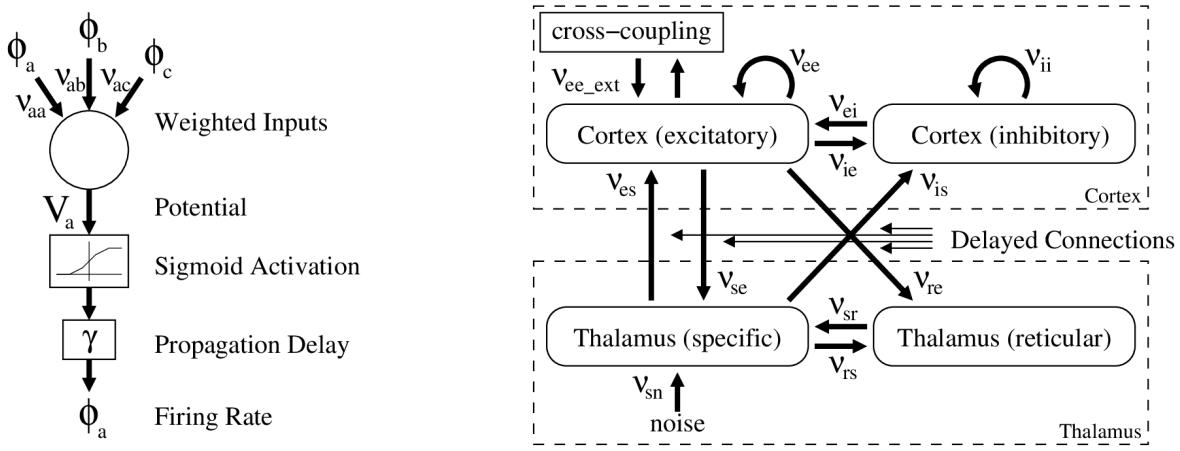


Figure 2.1: Extended Robinson model diagram, showing the neuron population model (left) and the population interactions (right).

The neural population model is described in Equations 2.1, 2.2 and 2.4. Per Equation 2.1, a weighted sum of input firing rates $\phi_b(t)$ is used to generate the cell body potential $V_a(t)$. A single dot indicates the first time derivative, and a double dot indicates the second time derivative. The parameters α and β are the inverse of the membrane potential fall time and rise time, respectively. The coupling parameter ν_{ab} is the strength of the connection from region b to region a (zero if no connection, negative if inhibitory).

NOTE: A delay is manually applied to some of the ϕ_b signals before this summation to reflect signal propagation time between the cortex and thalamus (per Robinson 2002), and to reflect cross-coupling delays within the cortex populations (per Hindriks 2023).

$$\left(\frac{1}{\alpha\beta}\right)\ddot{V}_a(t) + \left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\dot{V}_a(t) + V_a(t) = \sum_b \nu_{ab}\phi_b(t) \tag{2.1}$$

Per Equation 2.2, the cell body potential V is fed into a sigmoid activation function, to represent the collective action of many neurons with varying firing thresholds. The mean firing threshold is V_{th} and the standard deviation of the threshold is σ_{th}. This follows the convention of Freyer 2011; Robinson 2002 defined a related parameter $\sigma'_{th} = \frac{\sqrt{3}}{\pi}\sigma_{th}$ to simplify the activation equation (Equation 2.3).

$$Q(V) = \frac{Q_{max}}{1 + e^{-\left(\frac{\pi}{\sqrt{3}}\right)\left(\frac{V - V_{th}}{\sigma_{th}}\right)}} \tag{2.2}$$

$$Q(V) = \frac{Q_{max}}{1 + e^{-\left(\frac{V - V_{th}}{\sigma'_{th}}\right)}} \tag{2.3}$$

Per Equation 2.4, the local firing rate Q is propagated with damping and finite delay to give the non-local firing rate ϕ. A single dot indicates the first time derivative, and a double dot indicates the second time derivative. Per Robinson 2002, local propagation delay is assumed to only be relevant within the cortex, and $\gamma = \infty$ is assumed elsewhere. Per Freyer 2011, this further only applies to the excitatory neuron populations in the cortex.

$$\begin{cases} \frac{1}{\gamma^2}\ddot{\phi}(t) + \frac{2}{\gamma}\dot{\phi}(t) + \phi(t) = Q(t) & \text{Cortex excitatory neurons.} \\ \phi(t) = Q(t) & \text{All other populations.} \end{cases} \tag{2.4}$$

Noise is injected into the model as ϕ_n, and is described by Equation 2.5. Per Freyer 2011, there are three components: constant, additive, and multiplicative. Multiplicative noise (noise modulated by ϕ_e) is important for broadening the distribution of peak power levels of transient oscillations at frequencies above the fundamental oscillation mode of the cortex/thalamus loop.

In Equation 2.5, ϕ_n is the noise coupled to the thalamus via ν_{sn}, μ_n is the constant noise component (background firing rate), σ_n is the standard deviation of the independent component of the noise, and χ is a scaling parameter (per Freyer 2011) such that $\sigma_n\chi$ is the standard deviation of the component of the noise that is modulated by ϕ_e. Since the ϕ_e signal has to propagate from the cortex to the thalamus before modulating this noise component, it is manually delayed. The signals $g_1(t)$ and $g_2(t)$ denote two independent Gaussian noise sources with zero mean and with standard deviations of 1.

$$\phi_n(t) = \mu_n + \sigma_n g_1(t) + \sigma_n \chi g_2(t)\phi_e(t - t_{halfloop}) \tag{2.5}$$

As shown in Figure 2.1, cross-coupling between excitatory cortical populations is implemented per Hindriks 2023. This is described by Equation 2.6. Weight values w_{ab} represent the strength of connections between populations, and delay values $t_{coupling_{ab}}$ represent the propagation delays of these connections. While arbitrary weight values may be chosen, the recommended implementation is to use positive weights (purely excitatory), with the constraint that the sum of all weights contributing to a given ϕ_{ext_k} should sum to approximately unity. Cross-coupling propagation delay is typically no more than $\frac{1}{\gamma}$.

$$\phi_{ext_a}(t) = \sum_b w_{ab}\phi_{e_b}(t - t_{coupling_{ab}}) \tag{2.6}$$

$$\forall a, \sum_b w_{ab} \approx 1 \tag{2.7}$$

Typical parameter values for the extended Robinson model are shown in Table 2.1. Typical coupling coefficients are shown in Table 2.2. These are very similar to the parameter values and coupling coefficient values used in Hindriks 2023.

Parameter	**Value**	**Units**	**Notes**
Q_{max}	250	sec^{-1}	maximum firing rate
V_{th}	15	mV	potential threshold for firing
σ_{th}	6	mV	standard deviation of firing threshold
α	50	sec^{-1}	membrane potential inverse fall time
β	200	sec^{-1}	membrane potential inverse rise time
γ	100	sec^{-1}	cortex inverse propagation delay
$t_{halfloop}$	40	ms	one-way cortex/thalamus delay
μ_n	0	sec^{-1}	constant noise firing rate
σ_n	0.1	sec^{-1}	additive noise standard deviation
χ	0.3	dimensionless	multiplicative noise deviation coefficient

Table 2.1: Typical parameters for the extended Robinson model.

ν_{ee}	1.2
ν_{ei}	-1.8
ν_{es}	1.2
ν_{ie}	1.2
ν_{ii}	-1.8
ν_{is}	1.2
ν_{se}	1.2
ν_{re}	0.4
ν_{sr}	-0.8
ν_{rs}	0.2
ν_{sn}	0.5
$\nu_{ee_{ext}}$	0.07

Table 2.2: Typical coupling coefficient values for the extended Robinson model.

Robinson 2002 describes several oscillating modes: slow-wave/delta-wave oscillation, theta/fast delta oscillation with stable frequency (3 Hz) but varying shape, spindle oscillations at about 10 Hz driven by intra-thalamic resonance, and alpha oscillations at about 10 Hz. These are parameterized in terms of the amplification provided by each population of neurons (Equations 13-15 and Figure 3 in that reference).

The parameter values used in Freyer 2011 and Hindriks 2023 were chosen to support biologically relevant oscillations when driven by external noise (ϕ_n).

Chapter 3

Using the Model

FIXME: NYI

FIXME: Sample code for using the model to simulate neural populations.

FIXME: Sample code for finding operating points and analyzing gain and sensitivity to coupling parameters.

FIXME: Sample code for choosing parameters to support desired behavior.

Chapter 4

Mathematical Analysis

4.1 Low-Pass Filter Delays

Applying the Laplace transform to Equation 2.1 shows that the effect of α and β is to apply a low-pass filter to the weighted sum of input firing rates (a second-order exponential smoothing filter with poles at $-\alpha$ and $-\beta$).

$$\left(\frac{1}{\alpha\beta}\right)\ddot{V}_a(t) + \left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\dot{V}_a(t) + V_a(t) = \sum_b \nu_{ab}\phi_b(t) \tag{4.1}$$

$$\left(\frac{1}{\alpha\beta}\right)s^2V_a(s) + \left(\frac{1}{\alpha}+\frac{1}{\beta}\right)sV_a(s) + V_a(s) = \sum_b \nu_{ab}\Phi_b(s) \tag{4.2}$$

$$s^2V_a(s) + (\alpha+\beta)sV_a(s) + \alpha\beta V_a(s) = \alpha\beta\sum_b \nu_{ab}\Phi_b(s) \tag{4.3}$$

$$(s+\alpha)(s+\beta)V_a(s) = \alpha\beta\sum_b \nu_{ab}\Phi_b(s) \tag{4.4}$$

$$\frac{V_a(s)}{\sum_b \nu_{ab}\Phi_b(s)} = \frac{\alpha\beta}{(s+\alpha)(s+\beta)} \tag{4.5}$$

Applying the Laplace transform to Equation 2.4 shows that the effect of γ is to apply a low-pass filter to the firing rate (a second-order exponential smoothing filter with both poles at $-\gamma$).

$$\frac{1}{\gamma^2}\ddot{\phi}(t) + \frac{2}{\gamma}\dot{\phi}(t) + \phi(t) = Q(t) \tag{4.6}$$

$$\left(\frac{1}{\gamma^2}\right)s^2\Phi(s) + \left(\frac{2}{\gamma}\right)s\Phi(s) + \Phi(s) = Q(s) \tag{4.7}$$

$$s^2\Phi(s) + 2\gamma s\Phi(s) + \gamma^2\Phi(s) = \gamma^2 Q(s) \tag{4.8}$$

$$(s+\gamma)^2\Phi(s) = \gamma^2 Q(s) \tag{4.9}$$

$$\frac{\Phi(s)}{Q(s)} = \frac{\gamma^2}{(s+\gamma)^2} \tag{4.10}$$

The effect of both of these filters is to suppress high-frequency oscillations (those above the filter corner frequencies) and to delay low-frequency oscillations by an amount approximately equal to the filters' time constants. These delays and corner frequencies are listed in Table 4.1 (using the parameter values from Table 2.1).

Parameter	Value	Corner	Delay
α	50 sec^{-1}	8 Hz	20 ms
β	200 sec^{-1}	32 Hz	5 ms
γ	100 sec^{-1}	16 Hz	20 ms*

*Each pole at $-\gamma$ introduces a 10 ms delay; there are two such poles.

Table 4.1: Robinson model low-pass filter corners and low-frequency delays.

4.2 Oscillation Modes

A simplified diagram of the extended Robinson model is shown in Figure 4.1. This is intended to make it easy to identify the feedback loops that may support oscillations. Blue arcs indicate positive coefficients, and red arcs indicate negative (inhibitory) coefficients. As an approximation, the activity of different excitatory neuron populations within the cortex is assumed to be the same, combining ν_{ee} and ν_{ee_ext}. Additionally, multiplicative portion of the noise is treated as a contribution to the ν_{se} arc. The average contribution of the $\chi\sigma_n\nu_n$ term is zero, but the magnitude of that term compared to the magnitude of ν_{se} indicates whether or not multiplicative noise significantly contributes to that arc.

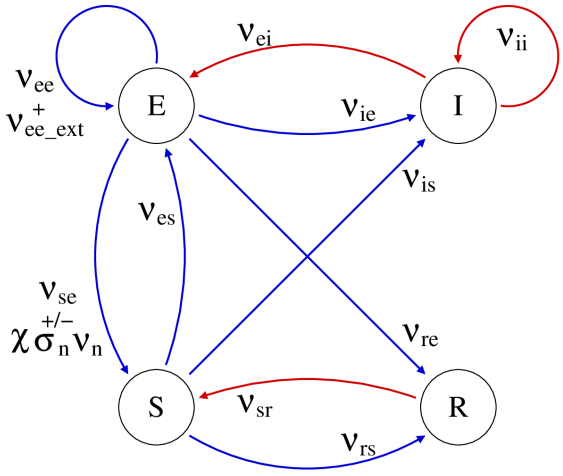


Figure 4.1: Simplified diagram of the extended Robinson model, showing feedback loops.

A list of potential oscillation loops and their oscillation frequencies is given in Table 4.2. For loops consisting entirely of positive coefficients, or with two negative coefficients, the oscillation period is the time needed to complete a single circuit. For loops with one negative coefficient, the oscillation period is the time needed to complete two circuits around the loop (in the same manner as a ring oscillator). Harmonics of these oscillation frequencies are also supported.

The number of arc traversals needed for one oscillation period is noted. Per above, this may reflect either one or two cycles around the loop. Each arc traversal involves gain from arc coefficients (noted in

the table), small-signal gain from the $Q(V)$ transfer function (omitted from the table), and delay from the α and β filter components. Delay contributions from the cortex excitatory population γ filter component and from the cortex-thalamus loop are noted in the table where applicable.

Label	Arcs	Gamma	C-T Loop	Period	Frequency	Coupling Gain
EE	1	Y	–	45 ms	22 Hz	$(\nu_{ee} + \nu_{ee\_ext})$
ES	2	Y	Y	150 ms	6.7 Hz	$\nu_{se} \cdot \nu_{es}$
ERSI	4	Y	Y	200 ms	5.0 Hz	$\nu_{re} \cdot \nu_{sr} \cdot \nu_{is} \cdot \nu ei$
II	2	–	–	25 ms	20 Hz	ν_{ii}
EI	4	Y	–	140 ms	7 Hz	$2 \cdot \nu_{ie} \cdot \nu_{ei}$
SR	4	–	–	100 ms	10 Hz	$2 \cdot \nu_{rs} \cdot \nu_{sr}$
ESI	6	Y	Y	350 ms	2.9 Hz	$2 \cdot \nu_{se} \cdot \nu_{is} \cdot \nu_{ei}$
ERS	6	Y	Y	350 ms	2.9 Hz	$2 \cdot \nu_{re} \cdot \nu_{sr} \cdot \nu_{es}$

Table 4.2: Extended Robinson model oscillation modes. Top modes: single-cycle (non-inverting). Bottom modes: two-cycle (inverting). Activation function gain is not shown.

Resonant loops explicitly described in Section IV of Robinson 2002 are the ones labelled ES, ERS, and SR in Table 4.2. Considering only those loops, the oscillation periods described in Section V of Robinson 2002 are consistent with the estimated periods of the ERS and SR loops.

The authors of Robinson 2002 were primarily concerned with noise-excited oscillations, and so only evaluated oscillation loops that included the specific nucleus. Evaluation was expressed in terms of the transfer function from the noise signal (input) to the firing rate of cortex excitatory neurons (output). Resonant oscillations were presumed to occur at frequencies for which this transfer function diverged (producing arbitrarily large output for finite input). This work instead considers gain within a loop, with resonant oscillations corresponding to a loop gain exceeding unity. As this does not explicitly consider noise excitation, all of the loops described in Table 4.2 may be analyzed.

4.3 Operating Points

As was described in Robinson 2002, the dynamics of the extended Robinson model can be analyzed by considering a state with unchanging (DC) firing rates, and evaluating the small-signal gain as firing rates are perturbed around this operating point. While this was used to find the transfer function $\frac{\phi_e(\omega)}{\phi_n(\omega)}$ with respect to noise in Robinson 2002, here it is used to find the small-signal loop gain (to determine which oscillating modes are dominant for given parameter values).

At a given operating point, the sensitivity of the gain of the dominant loops to each of the ν_{ab} coefficients provides insight into which network connections are most relevant for influencing dynamcis at that operating point. ν_{ee_ext} is particularly of interest, as this may be used as a proxy for the sensitivity of network dynamics to changes in the connectivity matrix between different excitatory cortex neuron populations. Since ν_{ee_ext} and ν_{ee} contribute to the same feedback arc in Figure 4.1, sensitivity of a single model instance to ν_{ee} may be used as a proxy for sensitivity of interacting model instances to ν_{ee_ext} and to changes in the connectivity matrix.

For time-independent operating point analysis, Equation 2.1 reduces to:

$$V_a = \sum_b \nu_{ab}\phi_b \tag{4.11}$$

Equation 2.2 (defining $Q(V)$) is unchanged, and Equation 2.4 reduces to:

$$\phi_a = Q(V_a) \tag{4.12}$$

Combining Equations 4.11 and 4.12 gives:

$$V_a = \sum_b \nu_{ab} Q(V_b) \tag{4.13}$$

$$\vec{V} = \mathbf{N}Q(\vec{V}) \tag{4.14}$$

In Equation 4.14, vector $\vec{V}$ contains as its elements all V_a, matrix $\mathbf{N}$ contains as its elements all ν_{ab}, and the activation function Q acts separately on each element of $\vec{V}$.

The activation function given in 2.3 can be rewritten as:

$$Q(V) = \frac{Q_{max}}{1 + e^{\left(\frac{V_{th}}{\sigma'_{th}}\right)} e^{-\left(\frac{V}{\sigma'_{th}}\right)}} \tag{4.15}$$

For firing rates that are much less than Q_{max} (such as those plotted in Robinson 2002), the exponential term dominates, and the activation function can be approximated as:

$$Q(V) \approx Q_{max}\, e^{-\left(\frac{V_{th}}{\sigma'_{th}}\right)} e^{\left(\frac{V}{\sigma'_{th}}\right)} \tag{4.16}$$

$$Q(V) \approx Q_0\, e^{\left(\frac{V}{\sigma'_{th}}\right)} \tag{4.17}$$

$$Q_0 = Q_{max}\, e^{-\left(\frac{V_{th}}{\sigma'_{th}}\right)} \tag{4.18}$$

The operating point equation then becomes Equation 4.19, where exponentiation is performed separately for each element of $\vec{V}$:

$$\vec{V} \approx \mathbf{N}\, Q_0\, e^{\left(\frac{\vec{V}}{\sigma'_{th}}\right)} \tag{4.19}$$

This system of exponential equations may be solved numerically.

Alternatively, a system of linear equations may be obtained by assuming that potentials are small compared to σ'_{th}. **This is not a robust assumption;** it is used as an estimate only. With this caveat

kept in mind, a linear approximation of the exponential function may be used (the first-order Taylor expansion), giving the following (where $\vec{1}$ denotes a vector where all elements are unity):

$$\vec{V} \approx \mathbf{N}\, Q_0 \left[\vec{1} + \left(\frac{1}{\sigma'_{th}}\right)\vec{V}\right] \tag{4.20}$$

$$\left(\frac{1}{Q_0}\right)\vec{V} \approx \mathbf{N}\,\vec{1} + \left(\frac{1}{\sigma'_{th}}\right)\mathbf{N}\,\vec{V} \tag{4.21}$$

$$\left[\left(\frac{1}{Q_0}\right)\mathbf{I} - \left(\frac{1}{\sigma'_{th}}\right)\mathbf{N}\right]\vec{V} \approx \mathbf{N}\,\vec{1} \tag{4.22}$$

Equation 4.22 has the form $\mathbf{A}\vec{V} = \vec{b}$, and so may be solved as a set of linear equations. The operating point estimated using Equation 4.22 **must** be examined to confirm that $|V_a| \ll \sigma'_{th}$ for all V_a. If this constraint does not hold, the estimate is not correct.

As a representative example, Table 4.3 shows the operating point corresponding to the model parameters given in Table 2.1 and the coupling coefficients given in Table 2.2. A single population was considered (no ν_{ee_ext} coupling). All of the operating point values satisfy the constraint for exponential approximation validity ($\phi \ll 250$), and while the constraint for linear approximation is not satisfied ($V \ll 3.3$), the potentials V_a are sufficiently close to the desired domain that the linear approximation result can be used as a seed value for obtaining the exponential approximation result via gradient descent.

	Measured	**Exponential**	**Linear**	
ϕ_e	4.2	4.1	4.8	/sec
ϕ_i	4.2	4.1	4.8	/sec
ϕ_s	3.3	3.2	4.0	/sec
ϕ_r	5.3	5.3	5.6	/sec
V_e	1.51	1.44	2.01	mV
V_i	1.51	1.44	2.01	mV
V_s	0.75	0.66	1.41	mV
V_r	2.34	2.31	2.49	mV

Table 4.3: Measured and estimated operating points for a representative model. "Measured" values were averaged over a 30-second simulation window. "Exponential" and "linear" values were estimated using Equations 4.19 (using gradient descent) and 4.22 (as a system of linear equations), respectively.

4.4 Small-Signal Gain and Dominant Oscillations

The small-signal gain G_{ab} of any given arc is the derivative of its output firing rate with respect to its input firing rate. For oscillations with frequencies much lower than the α, β, and γ filter corner frequencies, this may be estimated by taking the derivative of the DC operating point equations (rather than requiring an analysis of the full system dynamics):

$$G_{ab} = \frac{d\phi_a}{d\phi_b} \approx \frac{d}{d\phi_b}\left[Q(V_a)\right] \tag{4.23}$$

$$G_{ab} \approx Q'(V_a)\frac{dV_a}{d\phi_b} \tag{4.24}$$

$$G_{ab} \approx Q'(V_a)\frac{d}{d\phi_b}\left[\sum_c \nu_{ac}\phi_c\right] \tag{4.25}$$

$$G_{ab} \approx Q'(V_a)\nu_{ab} \tag{4.26}$$

The sigmoid response function is defined in terms of the logistic function:

$$\begin{cases} Q(V) = Q_{max}L\left(\frac{V-V_{th}}{\sigma'_{th}}\right) \\ L(x) = \frac{1}{1+e^{-x}} = \frac{e^x}{1+e^x} \\ \sigma'_{th} = \frac{\sqrt{3}}{\pi}\sigma_{th} \end{cases} \tag{4.27}$$

The derivative of the logistic function is:

$$L'(x) = L(x)\left(1 - L(x)\right) = \frac{e^x}{(1+e^x)^2} \tag{4.28}$$

This gives the derivative of the sigmoid response function:

$$Q'(V) = Q_{max}L'\left(\frac{V-V_{th}}{\sigma'_{th}}\right)\frac{d}{dV}\left[\frac{V-V_{th}}{\sigma'_{th}}\right] \tag{4.29}$$

$$Q'(V) = \frac{Q_{max}}{\sigma'_{th}}L'\left(\frac{V-V_{th}}{\sigma'_{th}}\right) \tag{4.30}$$

$$Q'(V) = \frac{Q_{max}}{\sigma'_{th}}L\left(\frac{V-V_{th}}{\sigma'_{th}}\right)\left(1 - L\left(\frac{V-V_{th}}{\sigma'_{th}}\right)\right) \tag{4.31}$$

$$Q'(V) = \frac{Q_{max}}{\sigma'_{th}}L\left(\frac{V-V_{th}}{\sigma'_{th}}\right)\left(1 - \frac{Q_{max}}{Q_{max}}L\left(\frac{V-V_{th}}{\sigma'_{th}}\right)\right) \tag{4.32}$$

$$Q'(V) = \frac{1}{\sigma'_{th}}Q(V)\left(1 - \frac{Q(V)}{Q_{max}}\right) \tag{4.33}$$

Combining this with Equation 4.26 gives:

$$G_{ab} \approx \frac{\nu_{ab}}{\sigma'_{th}}Q(V_a)\left(1 - \frac{Q(V_a)}{Q_{max}}\right) \tag{4.34}$$

$$G_{ab} \approx \frac{\nu_{ab}}{\sigma'_{th}}\phi_a\left(1 - \frac{\phi_a}{Q_{max}}\right) \tag{4.35}$$

This is Equation 10 from Robinson 2002.

The small-signal gain of a loop at low frequencies is the product of G_{ab} for each arc in the loop (as with the S_d, S_i, and S_r values in section IV of Robinson 2002). At frequencies that approach the low-pass filter cutoffs described in Section 4.1, Equation 4.5 must be applied to attenuate each edge, and Equation 4.10 must be applied to attenuate any edge leaving the excitatory cortex population node. The small-signal gains for the loops in Table 4.2 are shown in Table 4.4. These were evaluated at the exponential approximation operating point from Table 4.3, with low-pass filter attenuation applied.

Label	**Frequency**	**Cycle Time**	**Cycle Gain**	**Envelope Tau**
EE	22 Hz	45 ms	0.14	-23 ms
II	20 Hz	25 ms	-0.68	-66 ms
EI	7.1 Hz	70 ms	-1.39	**210 ms**
ES	6.7 Hz	150 ms	0.81	-690 ms
SR	10.0 Hz	50 ms	-0.09	-20 ms
ESI	2.9 Hz	175 ms	-2.93	**163 ms**
ERS	2.9 Hz	175 ms	-0.56	-300 ms
ERSI	5.0 Hz	200 ms	0.68	-520 ms

Table 4.4: Small-signal gains of extended Robinson model oscillation modes. Positive gains correspond to single-cycle loops; negative gains to two-cycle loops. Gains with absolute values smaller than 1 correspond to damped loops; gains with absolute values greater than 1 correspond to oscillation modes that grow over time with time constant τ_{loop}.

The single-cycle delay and single-cycle gain in a loop together define a time constant over which that loop's oscillation's envelope grows or decays, per Equations 4.36 and 4.37. An envelope time constant that is positive indicates that an oscillation mode grows with time; the smaller the time constant, the faster the growth:

$$|G_{loop}| = e^{\left(\frac{t_{loop}}{\tau_{env}}\right)} \tag{4.36}$$

$$\tau_{env} = \frac{t_{loop}}{\ln |G_{loop}|} \tag{4.37}$$

The dominant oscillation modes are expected to be those with the shortest positive envelope time constants. Only two such loops are present with the parameters from Tables 2.1 and 2.2: The EI loop in the cortex (at 7 Hz) and the ESI loop through the thalamus (at 3 Hz). The gain in the ESI loop may fluctuate due to the $\chi\sigma_n\nu_n$ contribution to ν_{se} (shown in Figure 4.1). A power spectrum from a simulation using these parameters is shown in Figure 4.2, showing a clear peak at 8 Hz (near the predicted EI loop resonance).

4.5 Adjusting Gain by Tuning Coupling Coefficients

Tuning the behavior of the simulated system is done by adjusting the internal coupling coefficients ν_{ab} (aggregated as matrix $\mathbf{N}$). The goal is typically to cause specific oscillation modes to become dominant. To facilitate this, the sensitivity of loop gain with respect to the coupling coefficients may be evaluated by computing the gradient of gain with respect to $\mathbf{N}$.

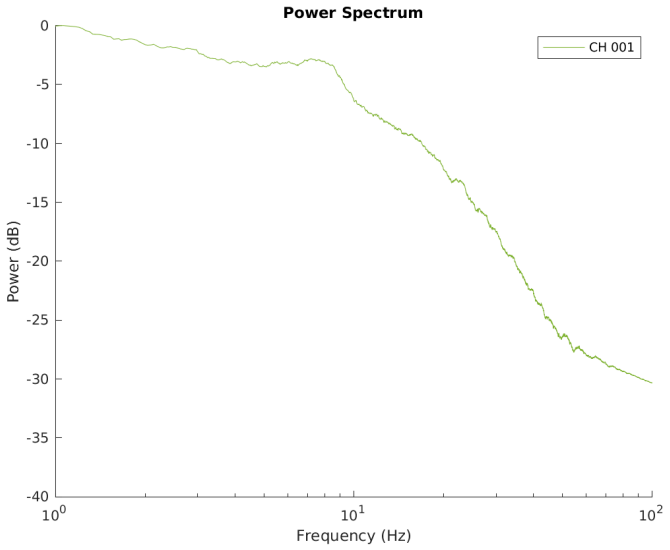


Figure 4.2: Power spectrum from a Robinson model simulation.

The gain across a given network edge may be expressed in terms of the edge's coupling coefficient and the derivative of the activation function, per Equations 4.26 and 4.33:

$$G_{ab} \approx \nu_{ab}\, Q'(V_a) \tag{4.38}$$

$$Q'(V_a) = \frac{1}{\sigma'_{th}} \phi_a \left(1 - \frac{\phi_a}{Q_{max}}\right) \tag{4.39}$$

Taking the gradient with respect to $\mathbf{N}$ gives:

$$\nabla_{\mathbf{N}} G_{ab} \approx Q'(V_a) \nabla_{\mathbf{N}} \nu_{ab} + \nu_{ab} \nabla_{\mathbf{N}} Q'(V_a) \tag{4.40}$$

$\nabla_{\mathbf{N}} \nu_{ab}$ is a matrix whose elements are zero except for a single element at row a column b, which is unity. The gradient of $Q'(V_a)$ is:

$$\nabla_{\mathbf{N}} Q'(V_a) = \nabla_{\mathbf{N}} \left[\frac{1}{\sigma'_{th}} \phi_a \left(1 - \frac{\phi_a}{Q_{max}}\right) \right] \tag{4.41}$$

$$\nabla_{\mathbf{N}} Q'(V_a) = \frac{1}{\sigma'_{th}} \left(\left(1 - \frac{\phi_a}{Q_{max}}\right) \nabla_{\mathbf{N}} \phi_a + \phi_a \nabla_{\mathbf{N}} \left[1 - \frac{\phi_a}{Q_{max}}\right] \right) \tag{4.42}$$

$$\nabla_{\mathbf{N}} Q'(V_a) = \frac{1}{\sigma'_{th}} \left(1 - 2\frac{\phi_a}{Q_{max}}\right) \nabla_{\mathbf{N}} \phi_a \tag{4.43}$$

In principle, the gradient of the firing rate $\nabla_{\mathbf{N}} \phi_a$ may be evaluated by taking the gradient with respect to $\mathbf{N}$ of Equations 4.17 and 4.19. In practice, the firing rate gradients are evaluated numerically by finding an operating point and making small perturbations to each ν_{ab}.

The gradient of the gain along a path consisting of multiple edges can be computed using the product rule. For two-, three-, and four-edge loops, this gives:

$$\nabla_{\mathbf{N}}[G_{ab}G_{ba}] = (\nabla_{\mathbf{N}}[G_{ab}])G_{ba} + G_{ab}(\nabla_{\mathbf{N}}[G_{ba}]) \tag{4.44}$$

$$\nabla_{\mathbf{N}}[G_{ab}G_{bc}G_{ca}] = (\nabla_{\mathbf{N}}[G_{ab}])G_{bc}G_{ca} + G_{ab}(\nabla_{\mathbf{N}}[G_{bc}])G_{ca} + G_{ab}G_{bc}(\nabla_{\mathbf{N}}[G_{ca}]) \tag{4.45}$$

$$\begin{aligned}\nabla_{\mathbf{N}}[G_{ab}G_{bc}G_{cd}G_{da}] = (\nabla_{\mathbf{N}}[G_{ab}])G_{bc}G_{cd}G_{da} + G_{ab}(\nabla_{\mathbf{N}}[G_{bc}])G_{cd}G_{da} \\ +G_{ab}G_{bc}(\nabla_{\mathbf{N}}[G_{cd}])G_{da} + G_{ab}G_{bc}G_{cd}(\nabla_{\mathbf{N}}[G_{da}])\end{aligned} \tag{4.46}$$

While Equations 4.44 through 4.46 can be expanded and written in terms of ν_{ab}, ϕ_a, and $\nabla_{\mathbf{N}}\phi_a$, it is not particularly useful to do so (since with or without such expansion the gradient will have to be evaluated numerically).

Plots of the gradients of the loop gains of the EI and ESI loops (from Table 4.4) are shown in Figure 4.3. The "raw" gradients are the change in loop gain with respect to each of the ν_{ab} coupling weights (i.e. $\nabla_{\mathbf{N}}G_{loop}$). The "normalized" gradients are divided by G_{loop}, to show the *relative* change in loop gain resulting from changes to ν_{ab}. In particular, this flips the sign of the gradient for loop gains that are negative (with a positive normalized gradient indicating that the loop gain is getting *larger*, not more positive).

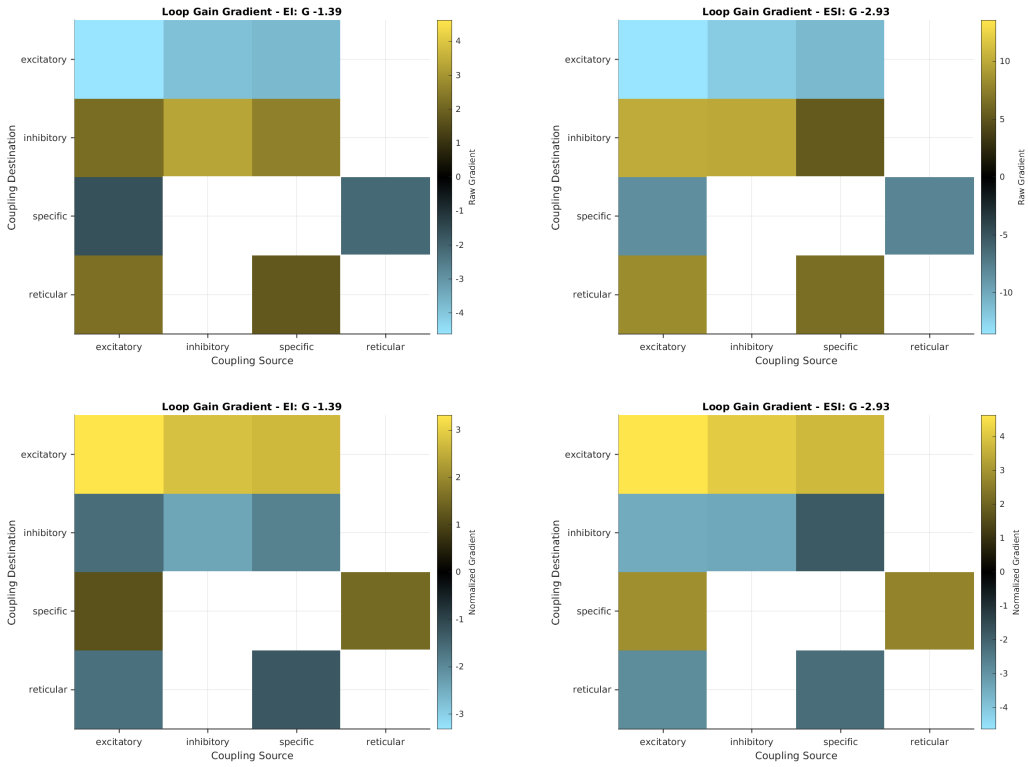


Figure 4.3: Gradient of loop gain with respect to ν_{ab}, for the EI loop (left) and the ESI loop (right). Top row: raw gradient. Bottom row: normalized gradient (gradient divided by G_{loop}).

The gradients may be visually inspected to gain a qualitative understanding of which ν_{ab} coupling weights have substantial impact on which loops. In particular, sensitivity to the strength of the ν_{ee}

coupling in this analysis indicates sensitivity to the strength of the ν_{ee_ext} coupling between populations, and to the weights in the population mixing matrix.

Inspection of the gradient plots may also guide hand-tuning of coupling weights. The main qualitative take-away is that while there are typically a small number of ν_{ab} coupling weights that most strongly affect a given loop, any given weight affects many loops, so weights must be jointly optimized against all loop gains rather than optimized individually against a single loop gain.

A function that jointly optimizes coupling weights to satisfy loop gain constraints is provided in `synthRFH_optimizeCouplings.m`. This is best used when making small adjustments, as for large adjustments the network may not end up in a biologically plausible operating state.