Neuronal Signals, Granger Causality and Time Series Analysis

A. E. Brockwell

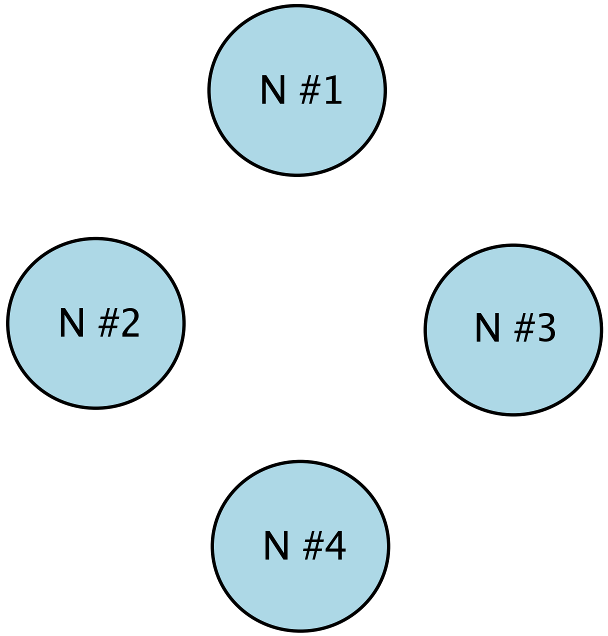
Dept. of Statistics

Carnegie Mellon University

abrock@stat.cmu.edu

A Problem

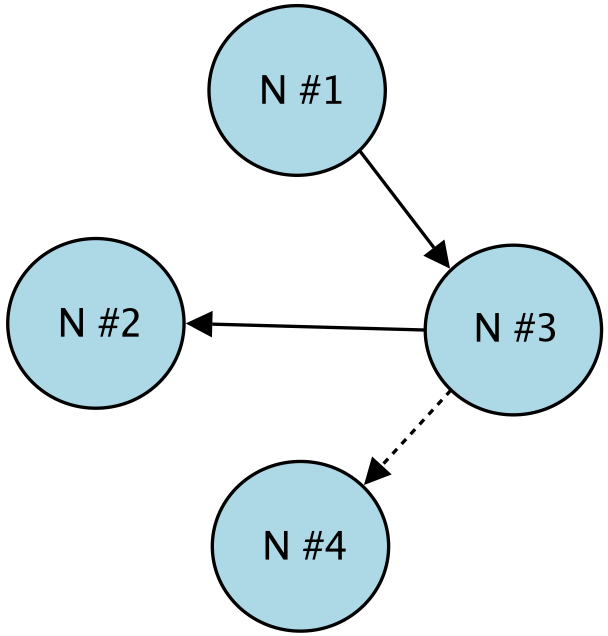
Problem: Given a set of time series associated with objects, determine which components are driving other components.



Neurons could be replaced with other objects.

A Problem

Problem: Given a set of time series associated with objects, determine which components are driving other components.



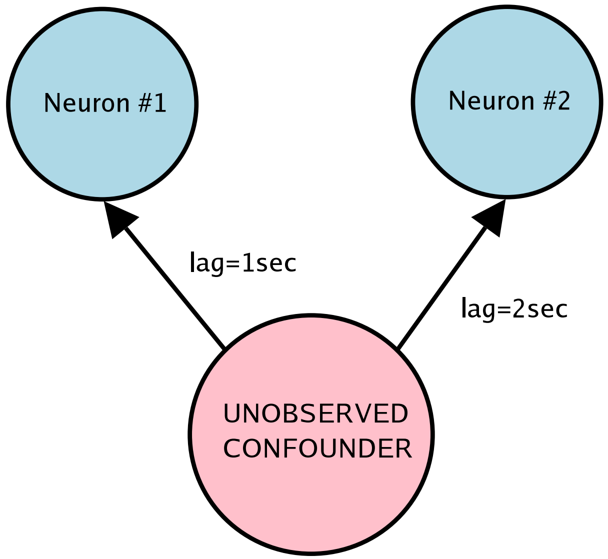
Neurons could be replaced with other objects.

Causality

It's tempting to say

"Neuron #1 has a causal effect on Neuron #2"

This is fundamentally problematic.



Granger Causality

"Granger causality" is not a new idea

- "Investigating Causal Relations by Econometric Models ...", C. Granger, Econometrica, 1969.
- "Analying Multiple Nonlinear Time Series with Extended Granger Causality", Chen, Rangarajan, Feng and Ding, Physics Letters A, 2004.

Granger Causality: Defn

Definitions:

$$
\begin{aligned}
P(Y_t|A) &= \text{unbiased min. var. predictor of } Y_t \text{ given} \\
&\quad\ \text{info in A at times } \leq t \\
\epsilon(Y_t|A) &= Y_t - P(Y_t|A) \\
\sigma^2(Y_t|A) &= \text{Var}(\epsilon(Y_t|A)) \\
U &= \text{all information available in the universe}
\end{aligned}
$$

We say $\{X_t\}$ "*Granger causes*" $\{Y_t\}$ if

$$\sigma^2(Y_t|U) < \sigma^2(Y_t|U \setminus X).$$

In practice, we can't take into account all the information in U.

So replace U by $\{$all measured processes$\}$.

An Index: We can define a *Granger Causality Index* by

$$GCI(X,Y) = 1 - \frac{\sigma^2(Y_t|U)}{\sigma^2(Y_t|U \setminus X)}$$

Granger Causality in Practice

Procedure:

1. Fit full multivariate time series model to all processes.
2. Fit sub-models, leaving out one process at a time.
3. Carry out model diagnostic tests.
4. Compute indices.

Example

2-Neuron Example: Suppose 2 neurons have firing rates M_t and N_t, satisfying the VAR(1) ("vector autoregression of 1st order") equation

$$\begin{bmatrix} M_t \\ N_t \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0 & 0.5 \end{bmatrix} \begin{bmatrix} M_{t-1} \\ N_{t-1} \end{bmatrix} + \begin{bmatrix} Z_t^{(1)} \\ Z_t^{(2)} \end{bmatrix},$$

where

$$\begin{bmatrix} Z_t^{(1)} \\ Z_t^{(2)} \end{bmatrix} \sim \mathsf{N}(0, I_{2\times 2}).$$

Note: This model doesn't allow for instantaneous Granger causal relationships.

Marginal Models

The VAR(1) has marginal models

$$M_t = 0.6M_{t-1} + V_t, \quad \{V_t\} \sim \mathsf{N}(0, 1.3)$$

$$N_t = 0.5N_{t-1} + W_t, \quad \{W_t\} \sim \mathsf{N}(0, 1)$$

(These could be calculated theoretically, or simply fit to observed data.)

GCI(M, N)

$$
\begin{aligned}
P(N_t|U) &= 0.5N_{t-1} \\
\epsilon(N_t|U) &= Z_t^{(2)} \\
\sigma^2(N_t|U) &= \text{Var}(Z_t^{(2)}) = 1
\end{aligned}
$$

$$
\begin{aligned}
P(N_t|U \setminus M) &= 0.5N_{t-1} \\
\epsilon(N_t|U \setminus M) &= W_t \\
\sigma^2(N_t|U \setminus M) &= \text{Var}(W_t) = 1
\end{aligned}
$$

So

$$\text{GCI}(M, N) = 1 - \frac{1}{1} = 0.$$

GCI(N, M)

$$\begin{aligned} P(M_t|U) &= 0.5M_{t-1} + 0.5N_{t-1} \\ \epsilon(M_t|U) &= Z_t^{(1)} \\ \sigma^2(M_t|U) &= \text{Var}(Z_t^{(1)}) = 1 \end{aligned}$$

$$\begin{aligned} P(M_t|U \setminus N) &= 0.6M_{t-1} \\ \epsilon(M_t|U \setminus N) &= V_t \\ \sigma^2(M_t|U \setminus N) &= \text{Var}(V_t) = 1.3 \end{aligned}$$

So

$$\text{GCI}(N, M) = 1 - \frac{1}{1.3} \simeq 0.23.$$

Fitting a VAR Model

Goal: Given K time series $\{N_t^{(j)}\},\ j = 1, 2, \ldots, K$, find a model of the form

$$N_t = \Phi_1 N_{t-1} + \Phi_2 N_{t-2} + \ldots + \Phi_p N_{t-p} + \epsilon_t,$$

where $N_t = (N_t^{(1)}, \ldots, N_t^{(K)})^T$ and Φ_j is a $K \times K$ matrix, and

$$\epsilon_t \sim \mathrm{N}(0, \Sigma).$$

Some VAR Fitting Methods

1. Compute sample cross-correlations and match with theoretical cross-correlations for model.
2. Compute cross-spectra and apply a spectral analog of the above procedure.
3. Cast the model as a *state-space model* and use the Kalman filter to compute likelihood as a function of parameters. Maximize over parameters.

A Useful Modification

Bivariate VAR(1):

$$N_t^{(1)} = \phi_{11} N_{t-1}^{(1)} + \phi_{12} N_{t-1}^{(2)} + \epsilon_t^{(1)}$$

$$N_t^{(2)} = \phi_{21} N_{t-1}^{(1)} + \phi_{22} N_{t-1}^{(2)} + \epsilon_t^{(2)}.$$

A Useful Modification

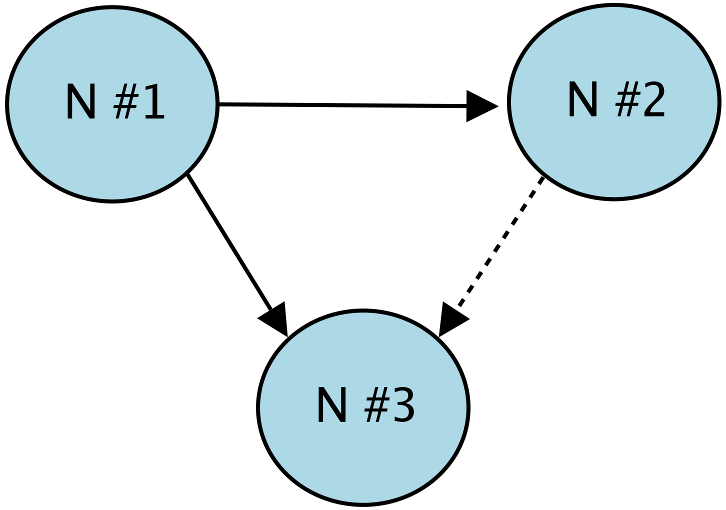
Bivariate VAR(1): Simultaneous Dep.

$$\begin{aligned} N_t^{(1)} &= \alpha_{12} N_t^{(2)} + \phi_{11} N_{t-1}^{(1)} + \phi_{12} N_{t-1}^{(2)} + \epsilon_t^{(1)} \\ N_t^{(2)} &= \alpha_{21} N_t^{(1)} + \phi_{21} N_{t-1}^{(1)} + \phi_{22} N_{t-1}^{(2)} + \epsilon_t^{(2)}. \end{aligned}$$

Crude method for fitting:

Use previously-mentioned methods, shifting time series by one time unit.

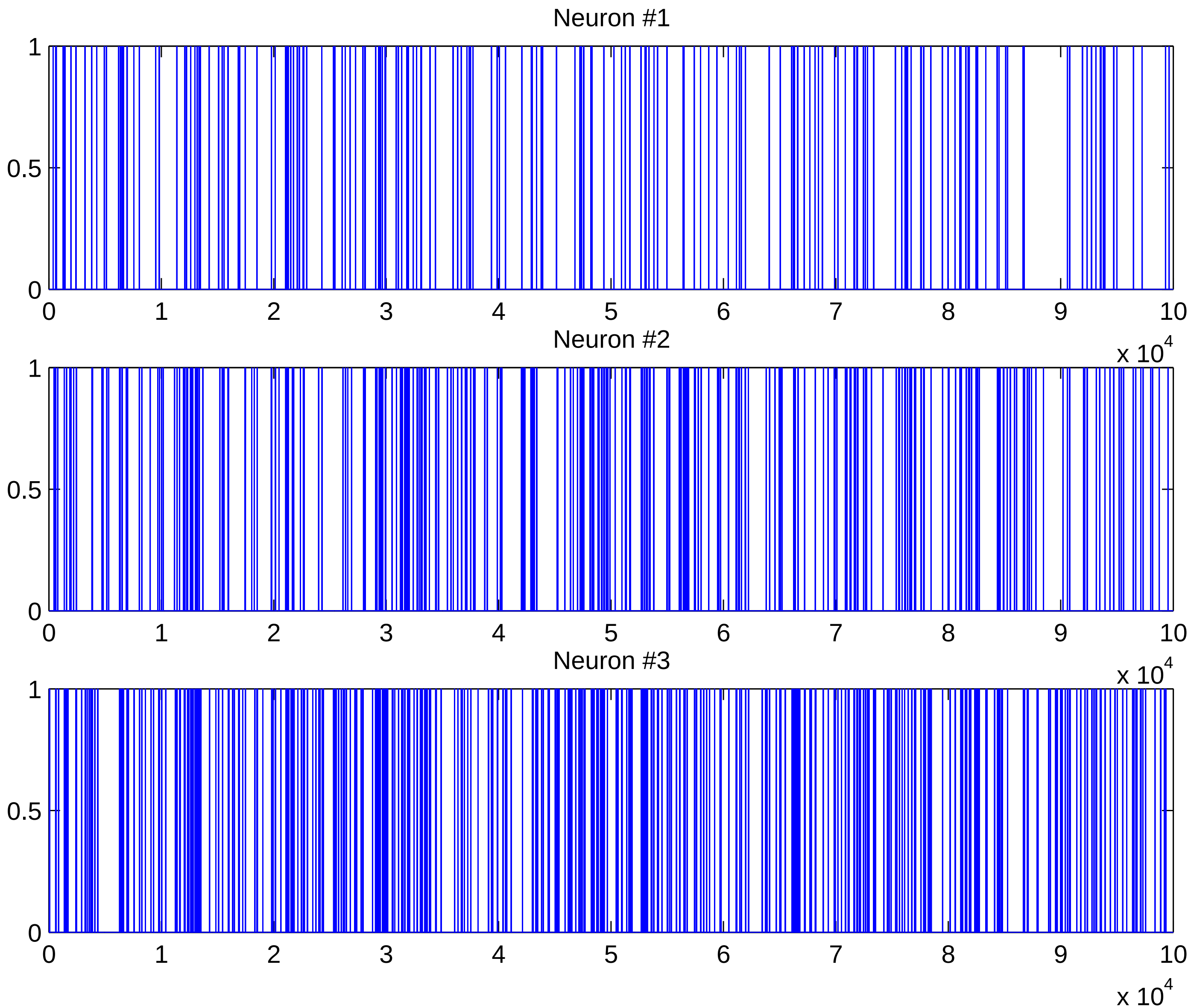
Simulation Study



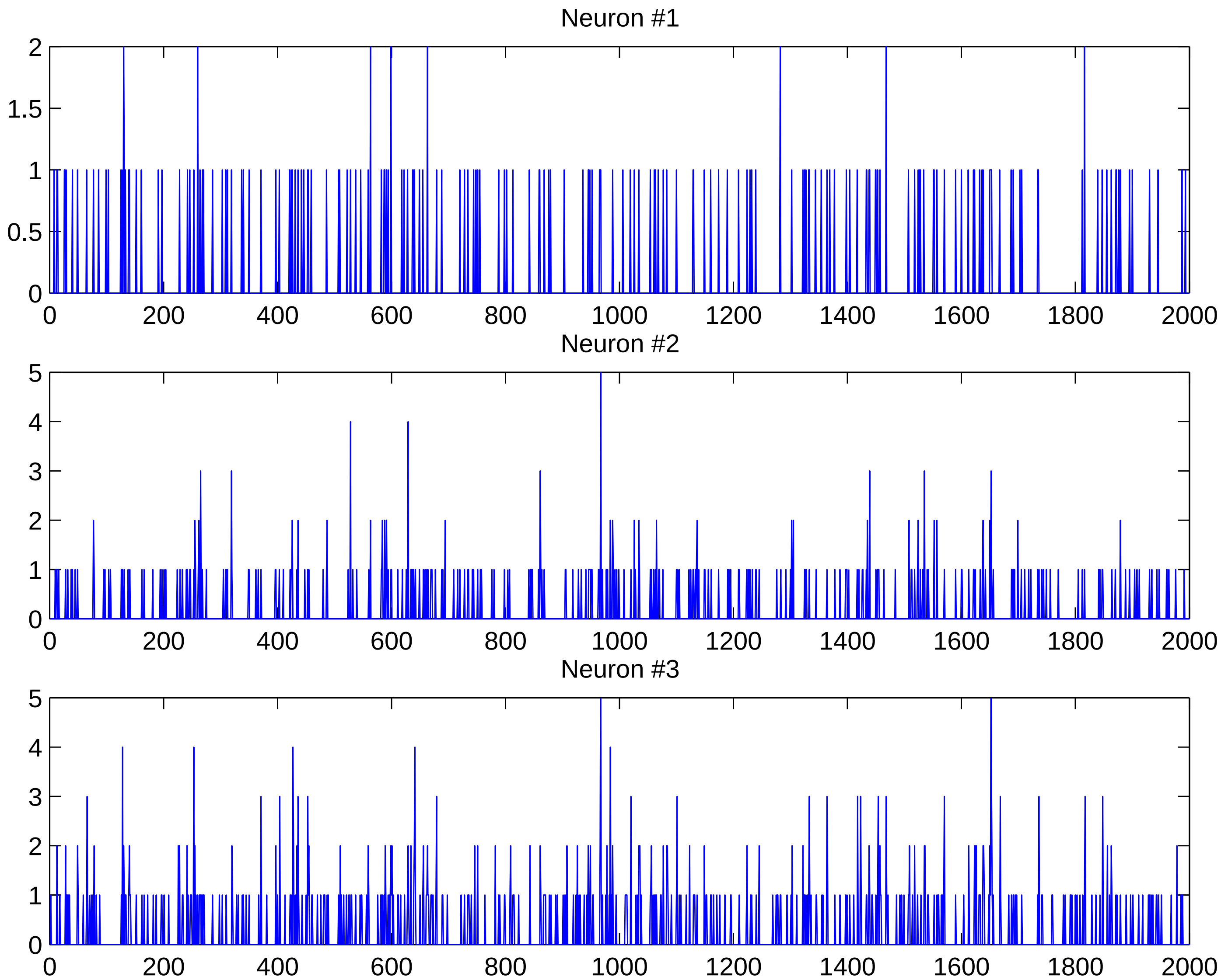
Neuron #1: Rate = $20Hz$
Neuron #2: Rate = $20Hz + 300\lambda_{12}$
Neuron #3: Rate = $20Hz + 300\lambda_{13} + 150\lambda_{23}$
Simulation Time Unit = 0.0001 seconds.
Total Time Simulated = 10 seconds.

Data is binned into 5ms bins.

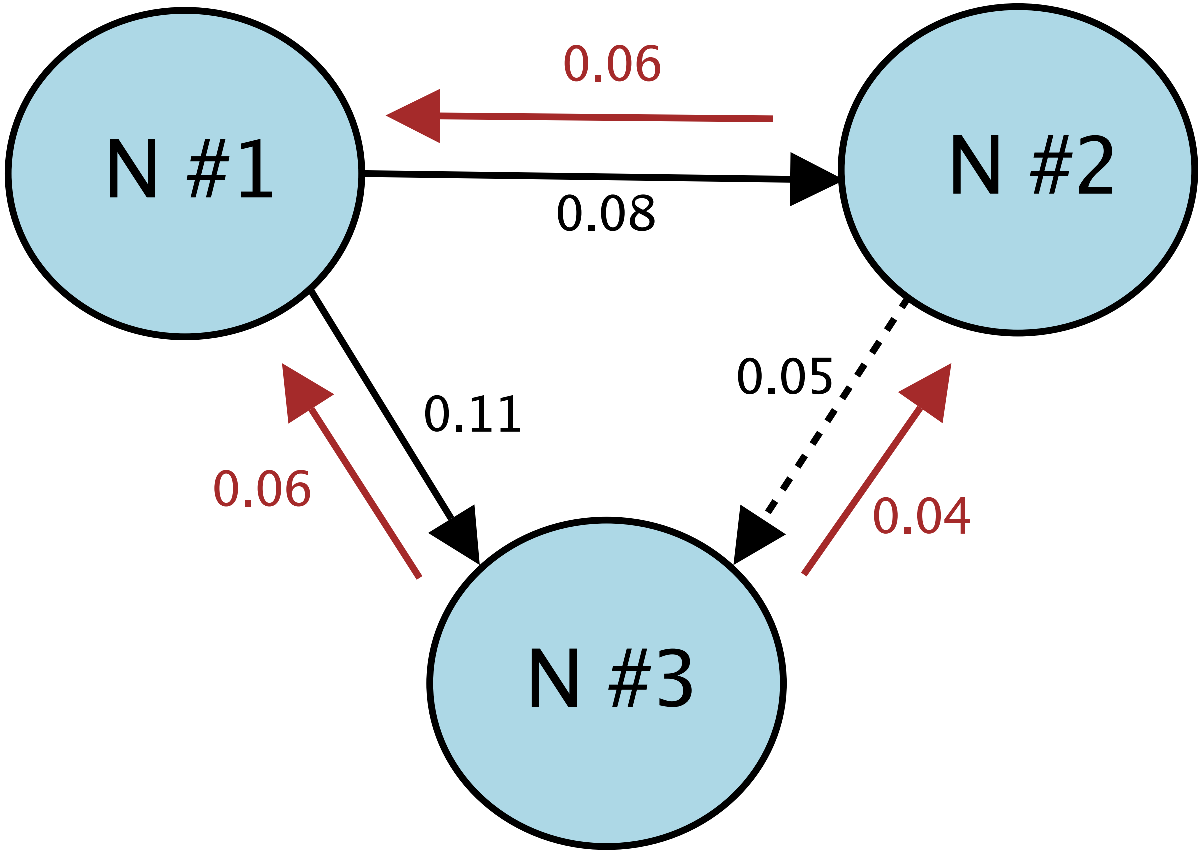
Simulated Spike Trains



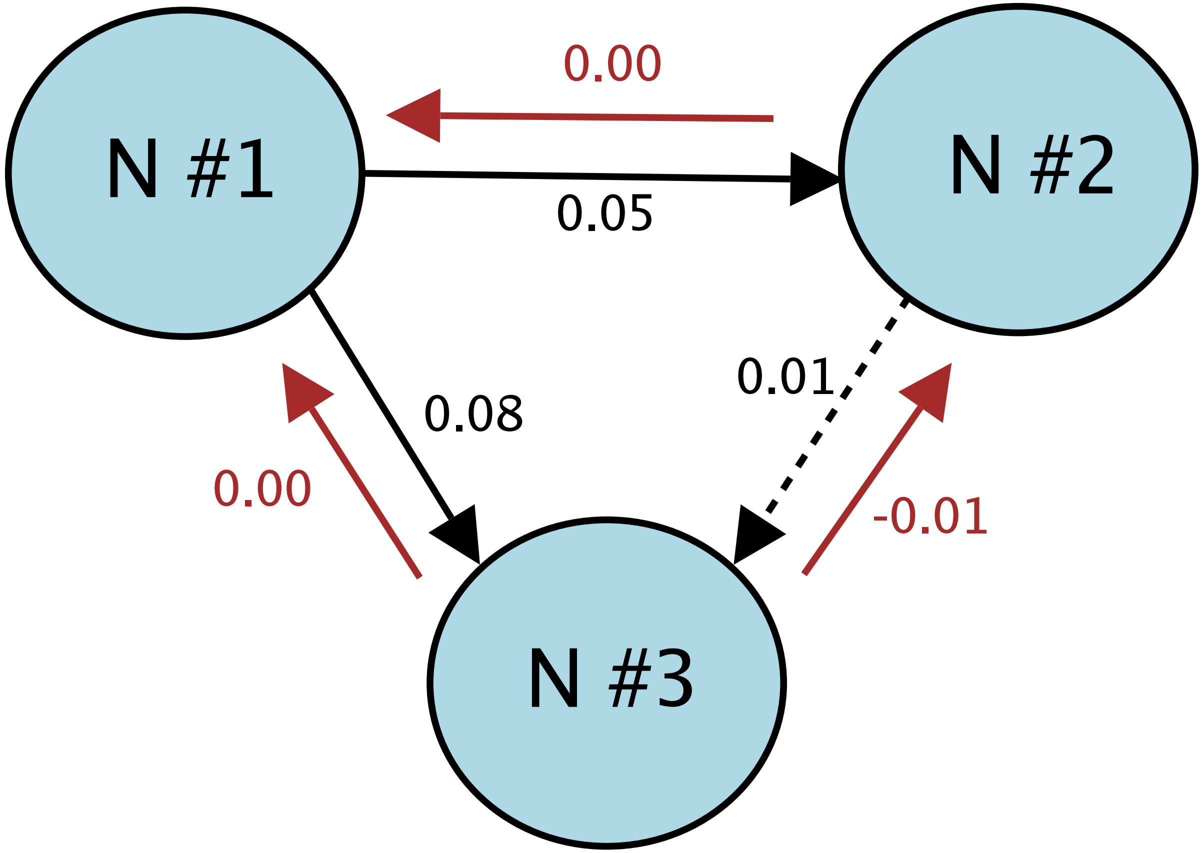
Binned Data


Neuron #1
Neuron #2
Neuron #3
0
200
400
600
800
1000
1200
1400
1600
1800
2000

Results - Allowing Simultaneous Dependence



Results - No Simultaneous Dependence



Comments on Time Series Analysis

- The VAR model is *linear*.
- Standard estimation procedures for VAR either implicitly or explicitly assume ϵ_t is *Gaussian*.

These assumptions are often unrealistic.

Something Better?

For modeling binned spike counts, perhaps

$$\begin{aligned} \{X_t^{(1)}, X_t^{(2)}, X_t^{(3)}\} &\sim \text{VAR} \\ N_t^{(j)} &\sim \text{Poisson}(\exp(X_t^{(j)})) \end{aligned}$$

would be more realistic. ($\{X_t\}$ is a hidden process.)

This is an example of a **generalized state-space model**.

Techniques are being developed for handling these kinds of models.

Additional Comments

- It's safer to use the term "Granger causality" than "causality".
- Results depend on your definition of the "universe". For optimal results, you should measure as much as possible.
- Results may also depend on sampling period.
- Model-fitting is *critical*. Hence diagnostics are important.