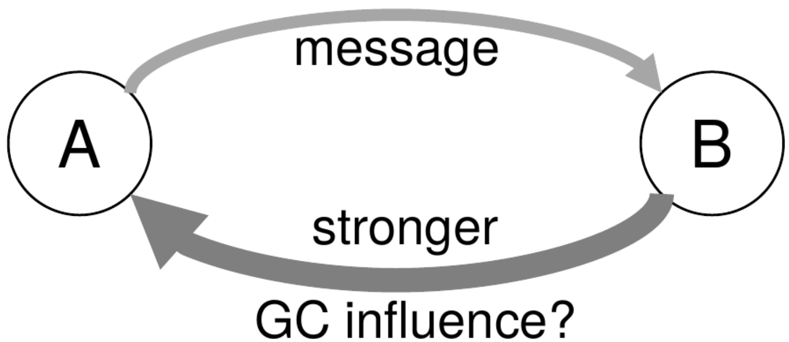
Is the direction of greater Granger causal influence the same as the direction of Information flow?



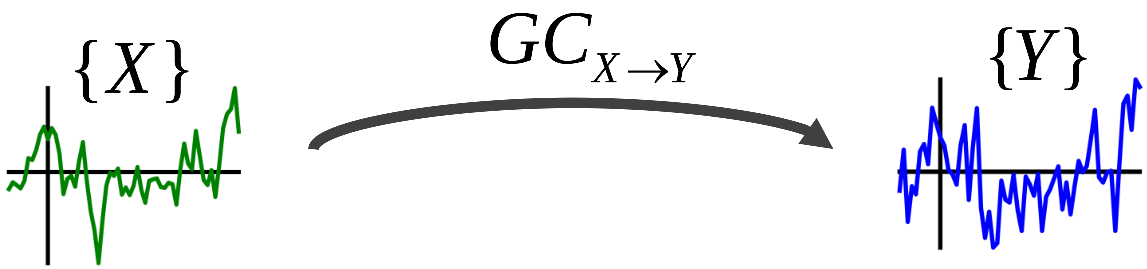
Praveen Venkatesh
(vpraveen@cmu.edu)

Pulkit Grover
(pulkit@cmu.edu)

Electrical & Computer ENGINEERING

Carnegie Mellon

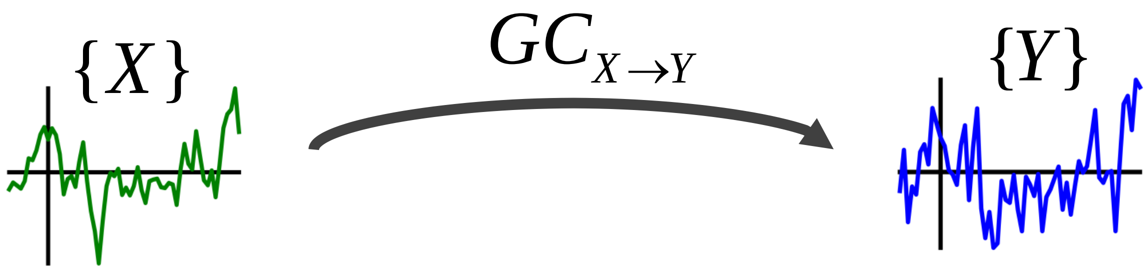
What is Granger causality?



C. W. Granger, *Econometrica*, 1969.

- Measure of *causal influence* between stochastic processes
- How much one process *causally predicts* another
- Directed information applied to Gaussian random variables (C. J. Quinn et. al., *J. Comp. Neuro*, 2011)

The Granger Causality (GC) metric

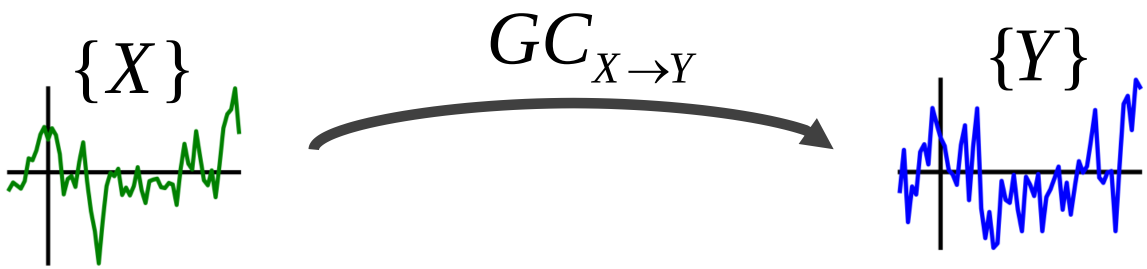


$$Y_i = \text{Past of Y} + \varepsilon_i$$

$$Y_i = \text{Past of Y} + \text{Past of X} + \tilde{\varepsilon}_i$$

$$GC_{X \to Y} = \frac{\mathrm{Var}(\epsilon)}{\mathrm{Var}(\tilde{\epsilon})}$$

The Granger Causality (GC) metric



$$Y_i = \sum_{j=1}^{p} \alpha_j Y_{i-j} + \varepsilon_i$$

$$Y_i = \sum_{j=1}^{p} \alpha_j Y_{i-j} + \sum_{j=1}^{p} \beta_j X_{i-j} + \tilde{\varepsilon}_i$$

$$GC_{X \to Y} = \frac{\mathrm{Var}(\epsilon)}{\mathrm{Var}(\tilde{\epsilon})}$$

Familiar objections to Granger causality

- Hidden (unmeasured) nodes
  can produce spurious influences
  J. Pearl, "Causality", *Cambridge Univ. press*, 2009

- Measurement noise
  can cause incorrect estimates
  H. Nalatore et. al., *Physical Review E*, 2007

- Subsampling
  can produce misleading GC relations
  M. Gong et. al., *Proc. Intl. Conf. on Machine Learning*, 2015

Familiar objections to Granger causality

- Hidden (unmeasured) nodes
  can produce spurious influences
- Measurement noise
  can cause incorrect estimates
- Subsampling
  can produce misleading GC relations

Deficiencies in measurement

Familiar objections to Granger causality

- Hidden (unmeasured) nodes
  can produce spurious influences
- Measurement noise
  can cause incorrect estimates
- Subsampling
  can produce misleading GC relations

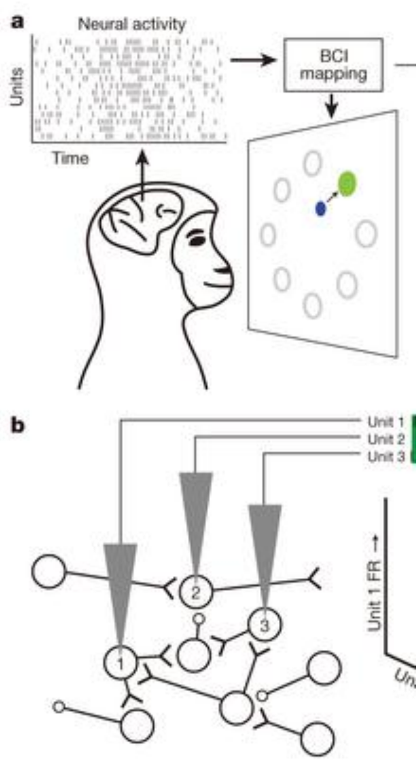
Deficiencies in measurement

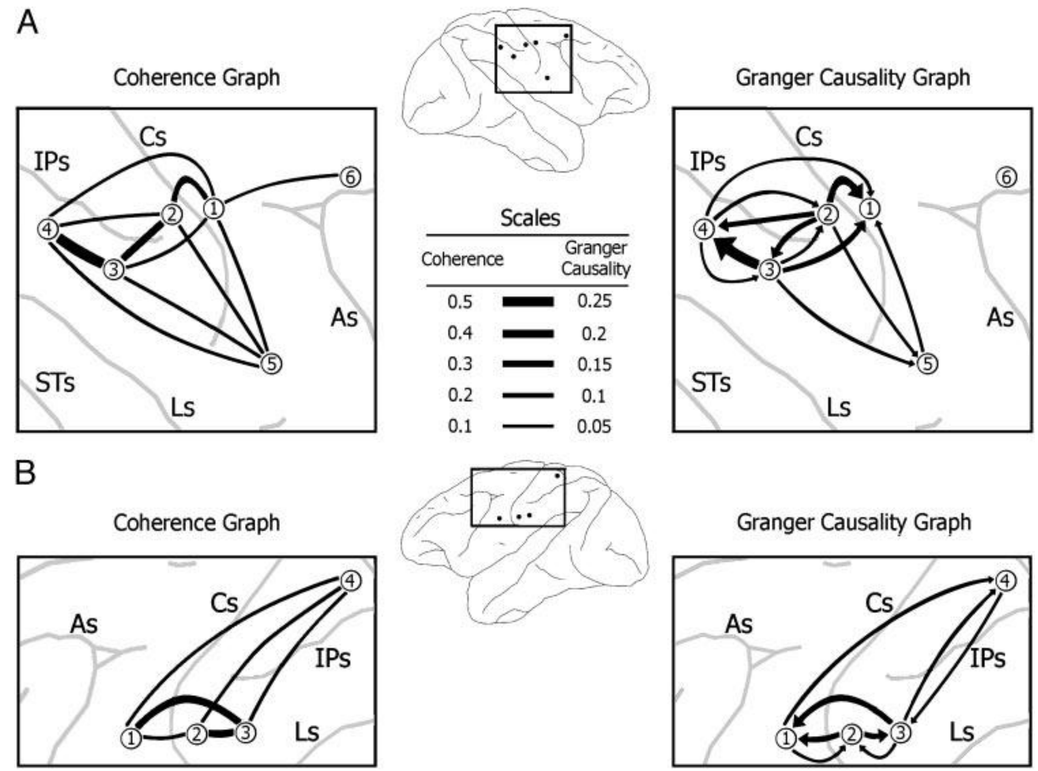
We assume perfect measurements
i.e. Granger causality estimated perfectly

How is GC used in neuroscience?

A neuroscientific question

- How does the brain compute?
- How does information flow in the brain?





P. T. Sadtler et. al., *Nature*, 2014

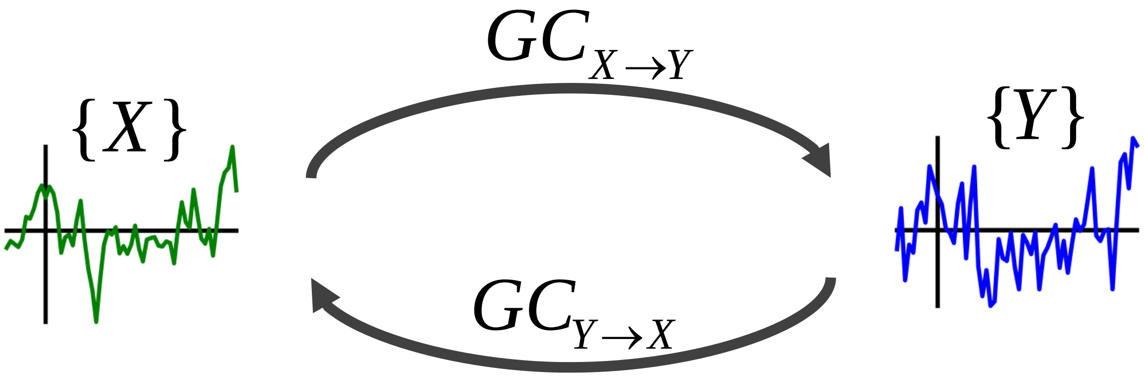
A. Brovelli et. al., *PNAS*, 2004

How is GC used in neuroscience?

- Measured brain signals modeled as stochastic processes
- GC used to find causal influence between processes = causal influence between brain regions
- Causal influence interpreted as information flow!

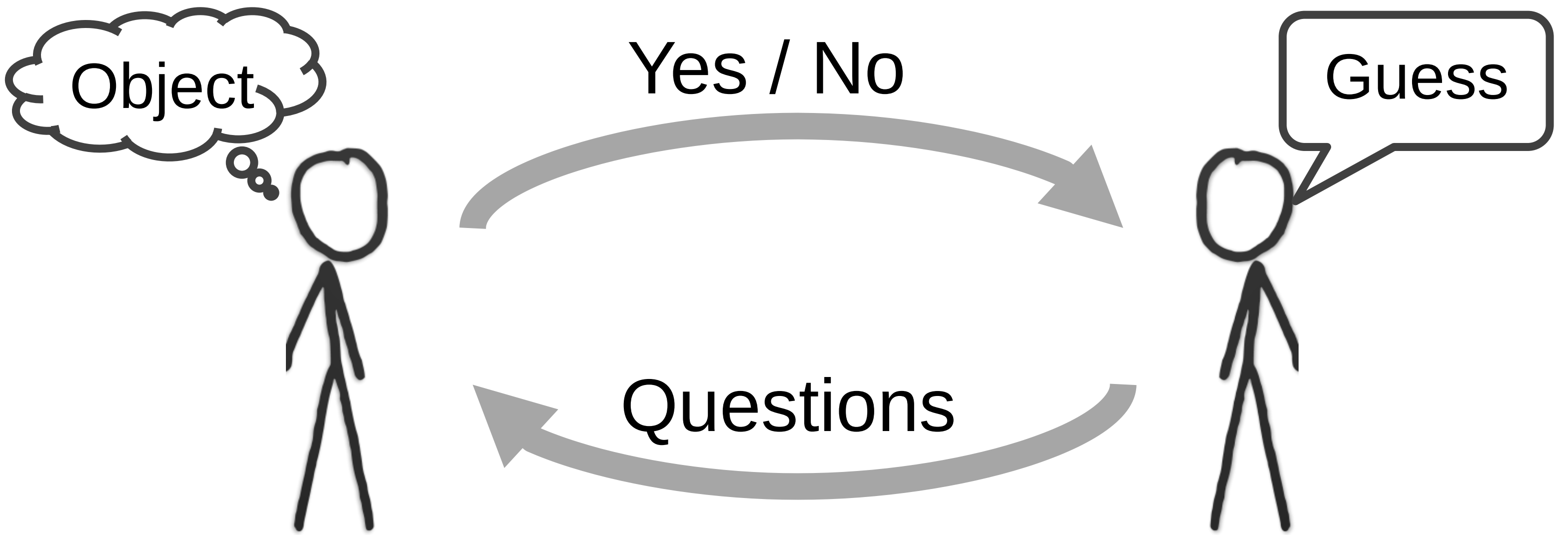
GC in feedback systems

- Comparison of GC metrics in the forward and reverse direction!



$$GC_{X \to Y} \overset{?}{>} GC_{Y \to X}$$

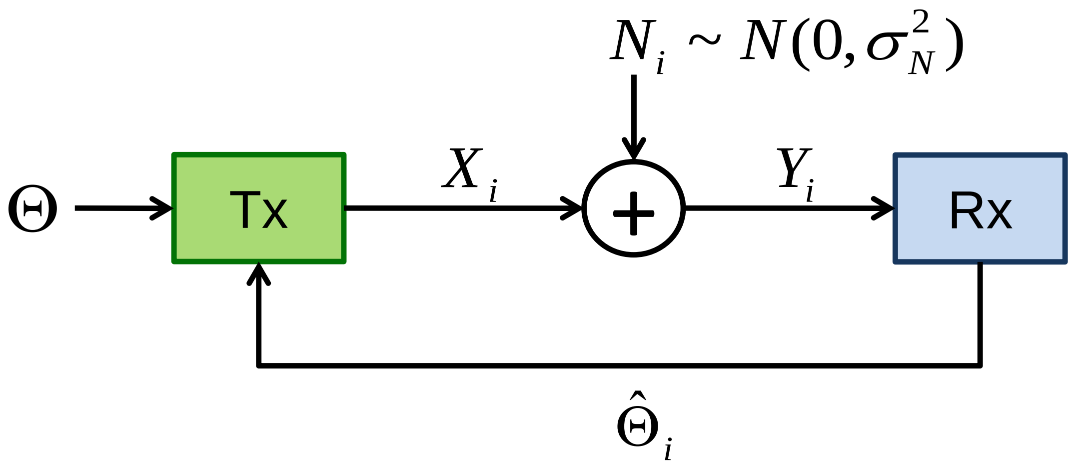
Is the direction of greater GC influence always the direction in which the message flows?


Object
Yes / No
Guess
Questions

A simple feedback communication strategy

The Schalkwijk-Kailath scheme

J. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback–i: No bandwidth constraint," *Information Theory, IEEE Transactions on, vol. 12, no. 2, pp. 172–182, 1966.*



- Tx sends message to Rx
- Rx communicates estimates back to Tx (noiseless)
- Tx sends errors in estimates to Rx

The Schalkwijk-Kailath scheme

- Tx sends

$$X_i = \Theta - \widehat{\Theta}_{i-1}$$

Error in estimate

- Rx receives

$$Y_i = X_i + N_i$$

Error + Noise

- Rx estimates and re-transmits

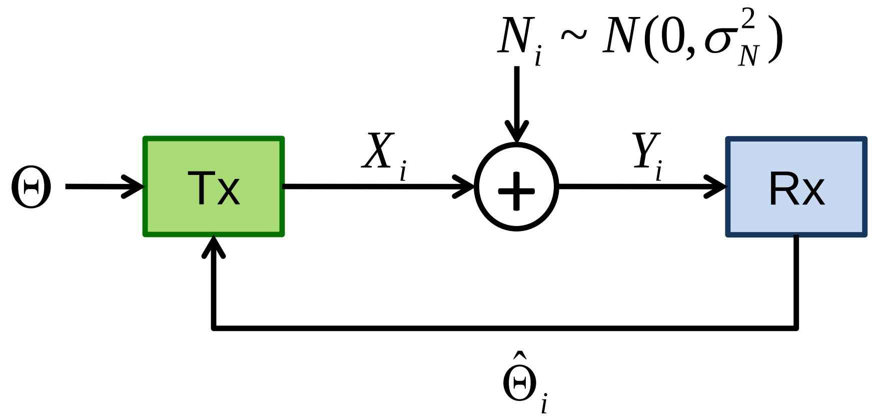
$$\widehat{\Theta}_i = \widehat{\Theta}_{i-1} + \frac{Y_i}{i}$$

New estimate

$$= \Theta + \frac{1}{i}\sum_{j=1}^{i} N_j$$

GC index computation

In the reverse direction (from Rx to Tx)



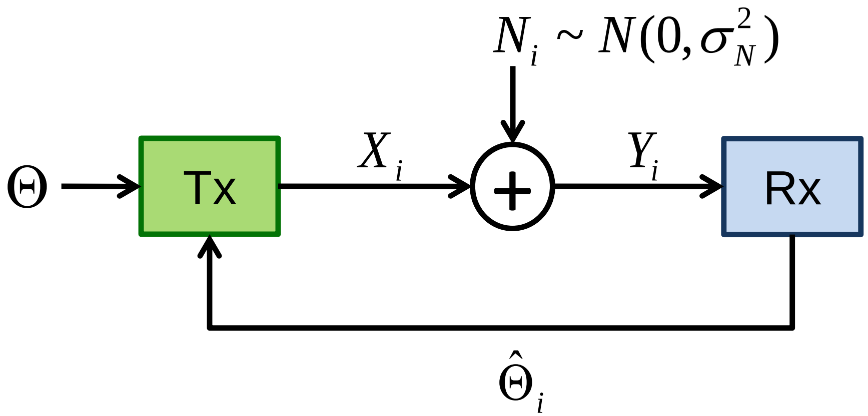
$$\boxed{\mathrm{X_i}} = \boxed{\text{Past of X}} + \varepsilon_i$$

$$\boxed{\mathrm{X_i}} = \boxed{\text{Past of X}} + \boxed{\text{Past of } \widehat{\Theta}} + \tilde{\varepsilon}_i$$

$$GC_{Rx \to Tx} = \frac{\mathrm{Var}(\epsilon)}{\mathrm{Var}(\tilde{\epsilon})}$$

GC index computation

In the reverse direction (from Rx to Tx)



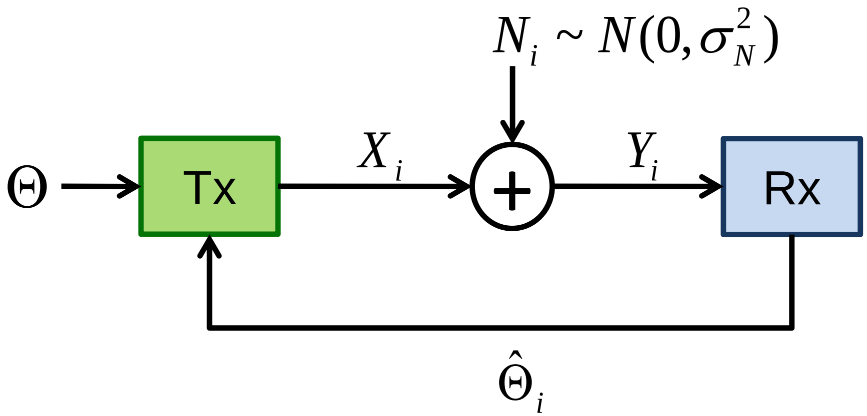
$$X_i = \frac{i-2}{i-1} X_{i-1} + \frac{N_{i-1}}{i-1}$$

$$X_i = \text{Past of } X + \text{Past of } \widehat{\Theta} + \tilde{\varepsilon}_i$$

$$GC_{Rx \to Tx} = \frac{\mathrm{Var}(\epsilon)}{\mathrm{Var}(\tilde{\epsilon})}$$

GC index computation

In the reverse direction (from Rx to Tx)



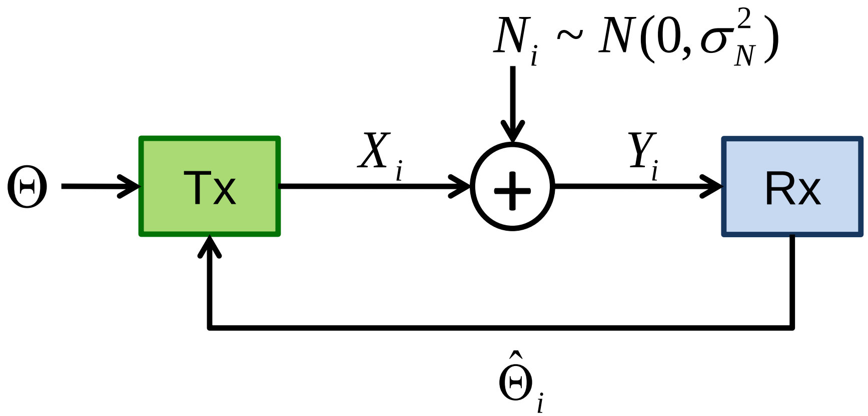
$$\boxed{X_i} = \frac{i-2}{i-1}X_{i-1} + \frac{N_{i-1}}{i-1}$$

$$\boxed{X_i} = X_{i-1} + \widehat{\Theta}_{i-2} - \widehat{\Theta}_{i-1}$$

$$GC_{Rx \to Tx} = \frac{\mathrm{Var}(\epsilon)}{\mathrm{Var}(\tilde{\epsilon})}$$

GC index computation

In the reverse direction (from Rx to Tx)

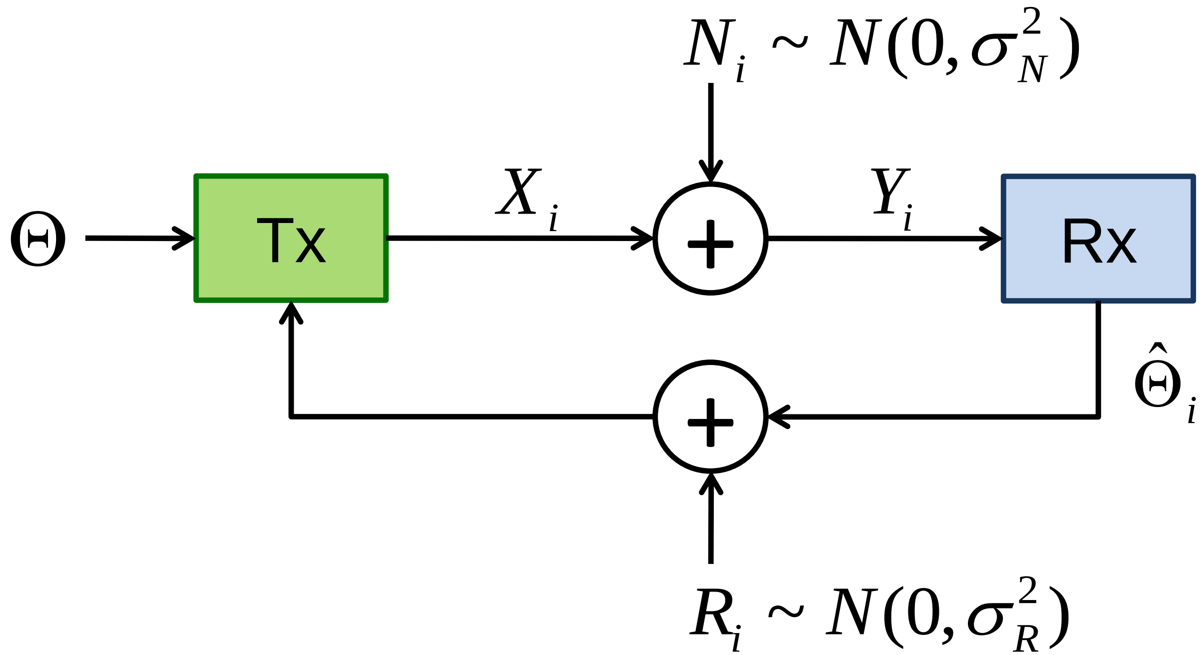


$$X_i = \frac{i-2}{i-1} X_{i-1} + \frac{N_{i-1}}{i-1} \quad \epsilon = finite$$

$$X_i = X_{i-1} + \widehat{\Theta}_{i-2} - \widehat{\Theta}_{i-1} \quad \tilde{\epsilon} = 0!$$

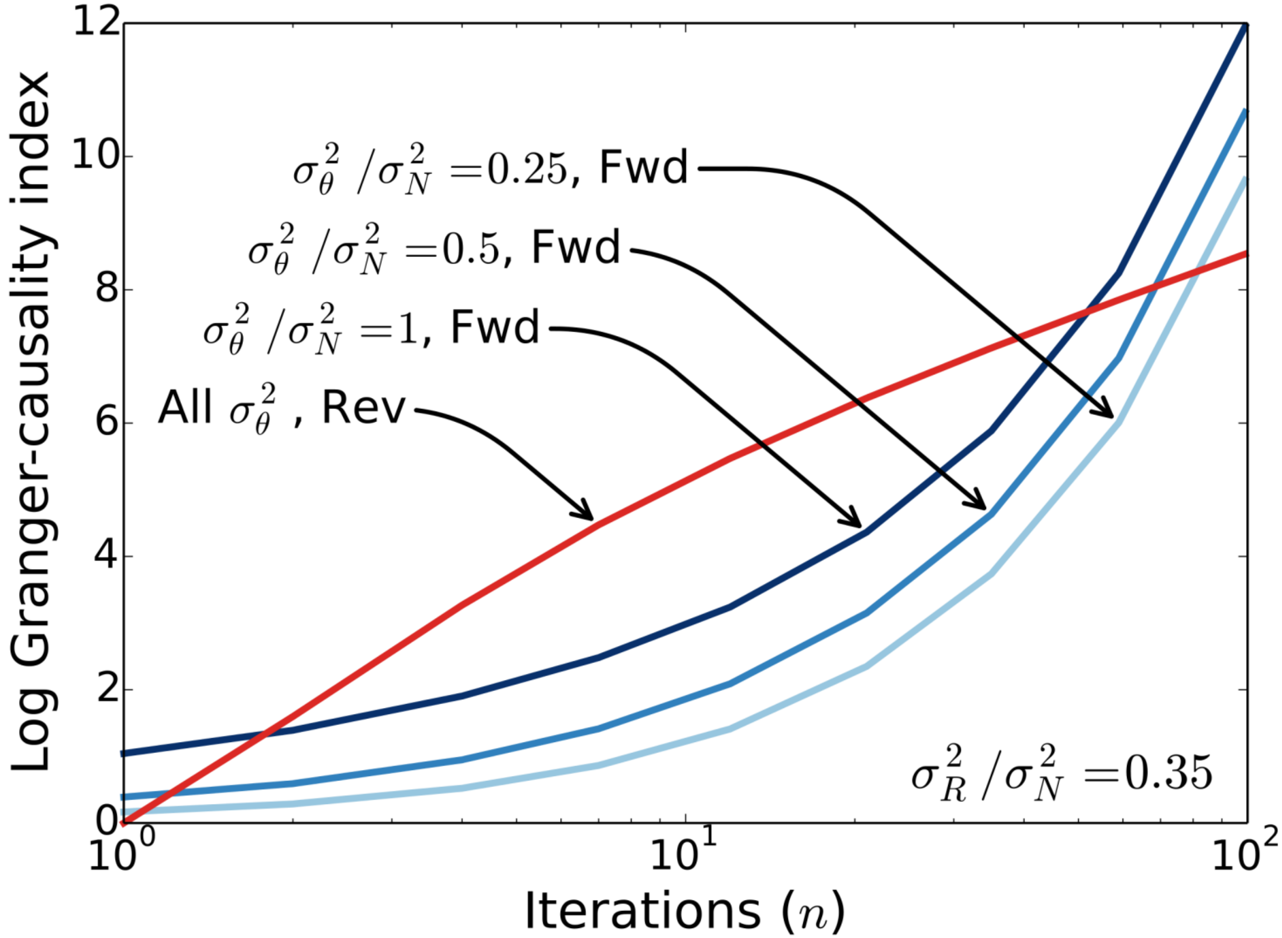
$$GC_{Rx \to Tx} = \frac{\mathrm{Var}(\epsilon)}{\mathrm{Var}(\tilde{\epsilon})} = \infty!$$

Noisy feedback

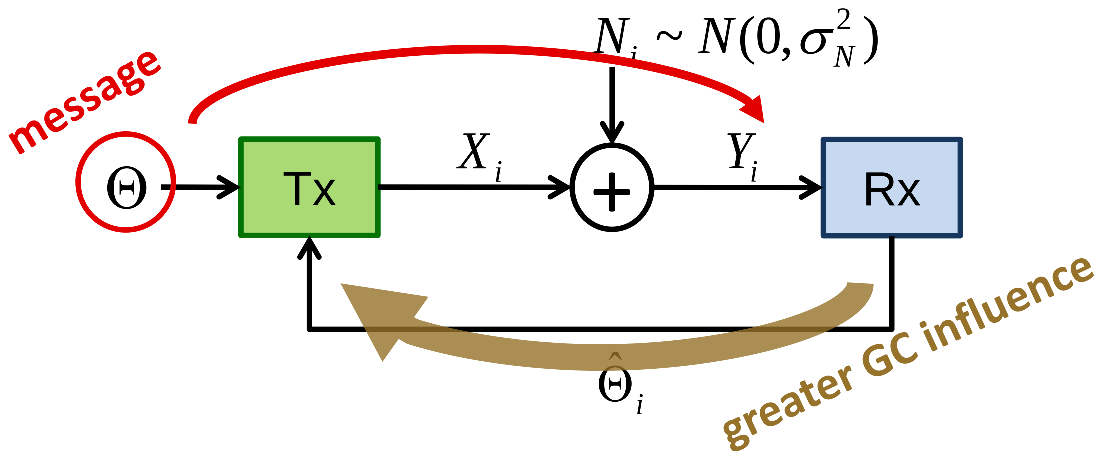


Optimal strategies not known, but S&K gives a consistent, unbiased estimate

Greater GC influence is opposite info flow



Summary



Counter-example shows greater Granger-Causal influence can be opposite to the direction of information flow!

Counter-example holds even if measurements are made perfectly

Future work:

- Better strategies for noisy feedback
- Evolving source: Θ changes with time

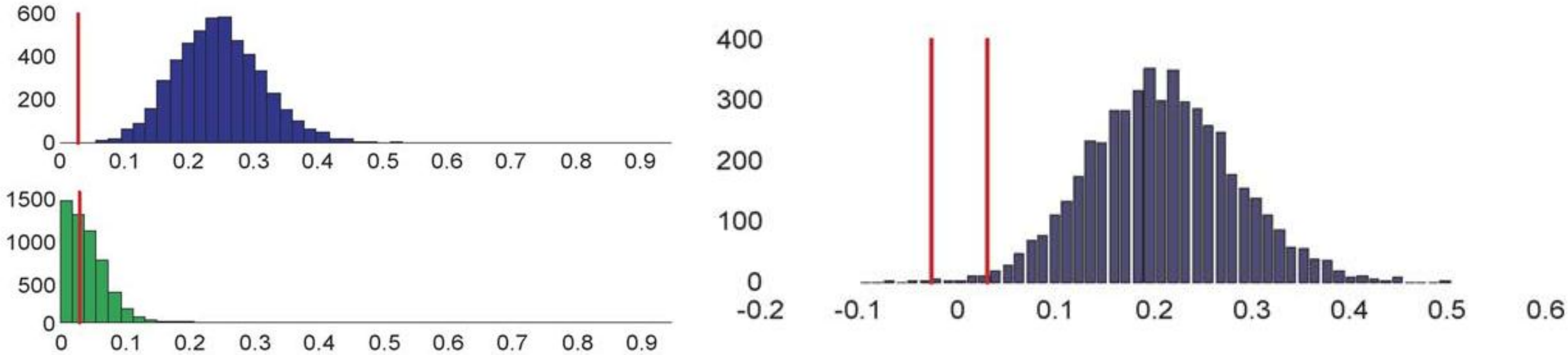
Strengthen the counter-example

- Searching for alternative ways to estimate direction of info flow

Look for new solutions

Why is comparison done?

- Might need comparison to get statistically significant results



A. Roebroeck et. al., *Neuroimage,* 2005

Noisy feedback

- Tx sends

$$X_i = \Theta - (\widehat{\Theta}_{i-1} + R_{i-1})$$

- Rx receives

$$Y_i = X_i + N_i$$

- Rx estimates and re-transmits

$$\begin{aligned}\widehat{\Theta}_i &= \widehat{\Theta}_{i-1} + \frac{Y_i}{i} \\ &= \Theta + \frac{1}{i}\sum_{k=1}^{i} N_k - \frac{1}{i}\sum_{k=1}^{i-1} R_k\end{aligned}$$

GC computation

In the reverse direction (from Rx to Tx)

$$X_i = \sum_{j=1}^{p} \alpha_j X_{i-j} + \epsilon_i \ = \frac{i-2}{i-1} X_{i-1} + \frac{N_{i-1}}{i-1}$$

$$\begin{aligned} X_i &= \sum_{j=1}^{p} \alpha_j X_{i-j} + \sum_{j=1}^{p} \beta_j \widehat{\Theta}_{i-j} + \tilde{\epsilon}_i \\ &= X_{i-1} + \widehat{\Theta}_{i-2} - \widehat{\Theta}_{i-1} \end{aligned}$$

$$\mathrm{Var}(\epsilon_i)/\mathrm{Var}(\tilde{\epsilon}_i) = \infty$$