

Granger Mediation Analysis for Multiple Time Series

Xi (Rossi) LUO

The University of Texas
Health Science Center
School of Public Health
Dept of Biostatistics
and Data Science



**BigComplex
DATA.com**

JSM, Philadelphia
August 6, 2020

Funding: NIH R01EB022911, P01AA019072, P20GM103645, P30AI042853; NSF/DMS (BD2K) 1557467

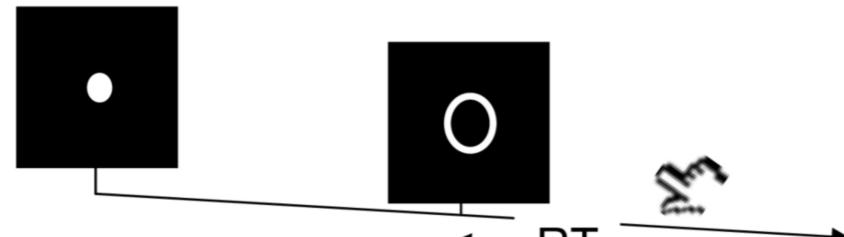
Co-Author



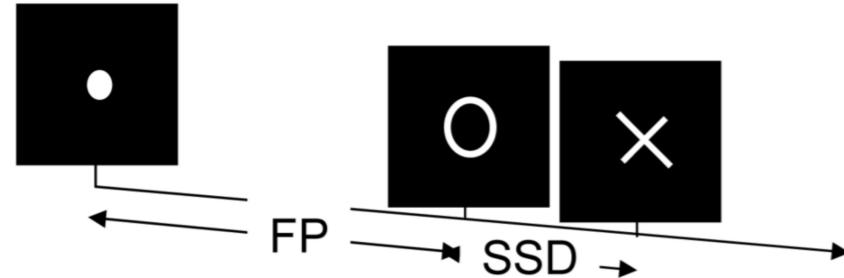
Yi Zhao
Assistant Prof
Indiana Univ

fMRI Experiments

(a) GO trials: Success **G**; Error **F**



STOP trials: Success **SS**; Error **SE**



- Task fMRI: performs tasks under brain scanning
- **Randomized** stop/go task:
 - press button if "go";
 - withhold pressing if "stop"
- Not resting-state: "do nothing" during scanning

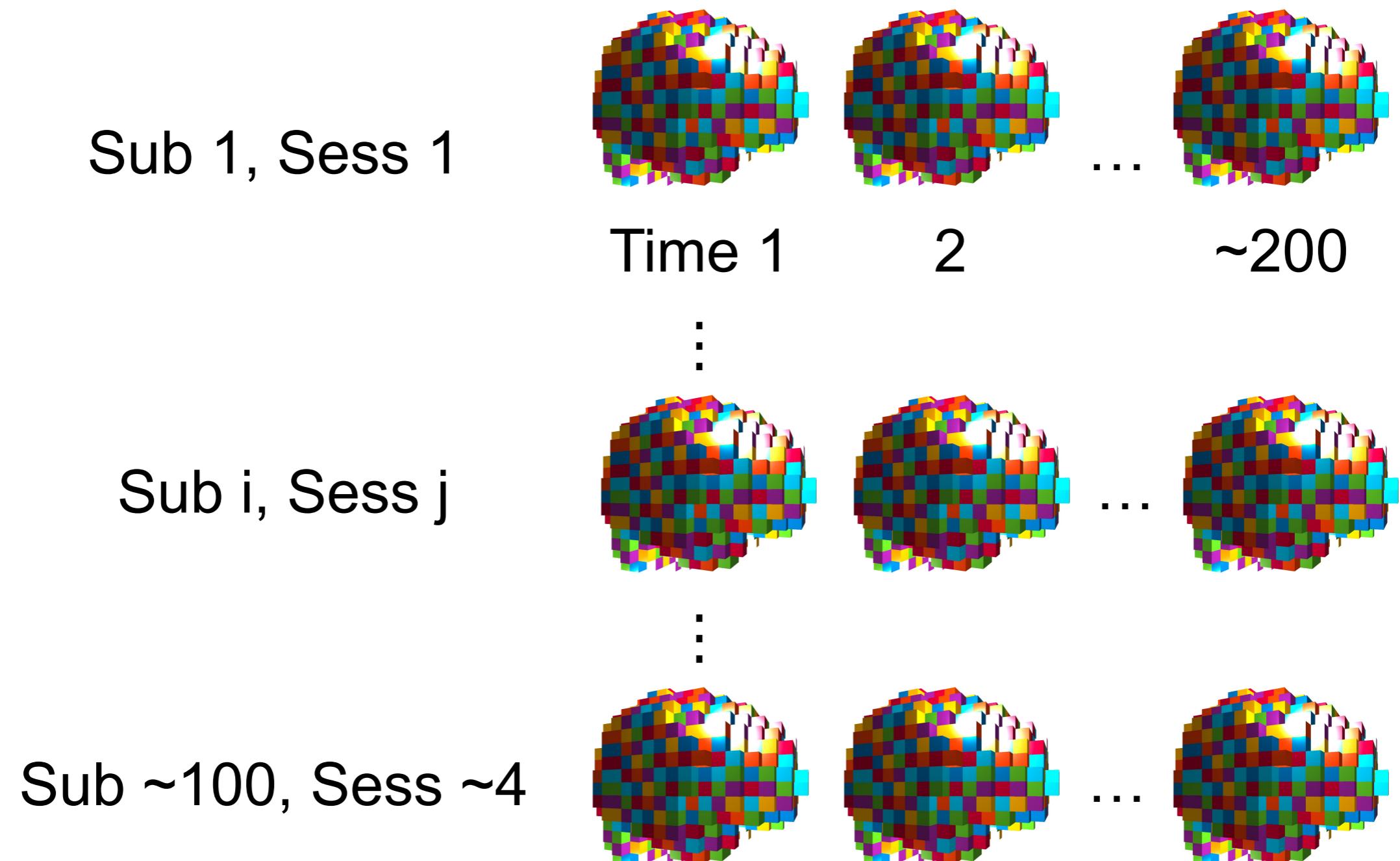
Goal: infer brain activation and connectivity





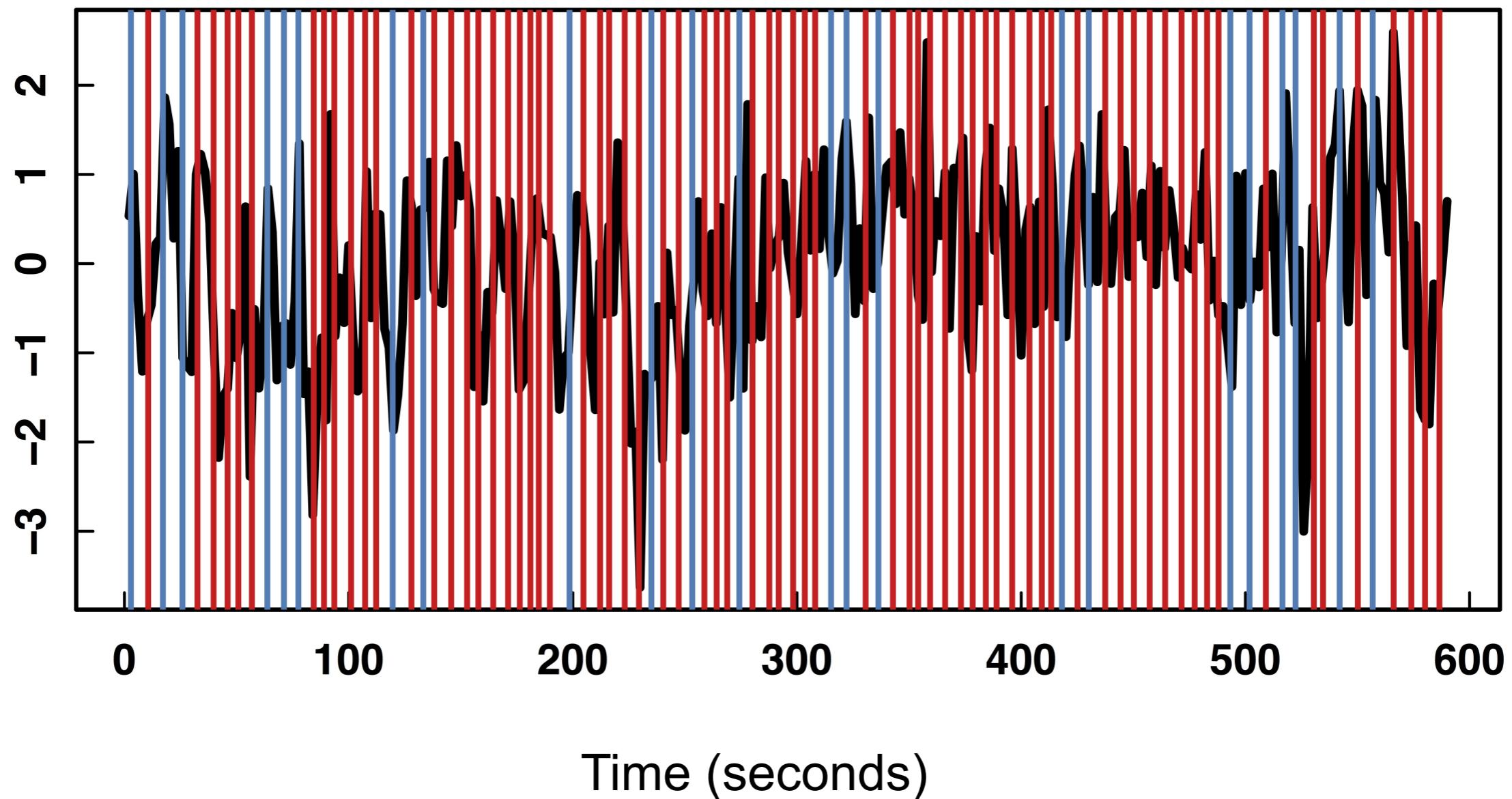
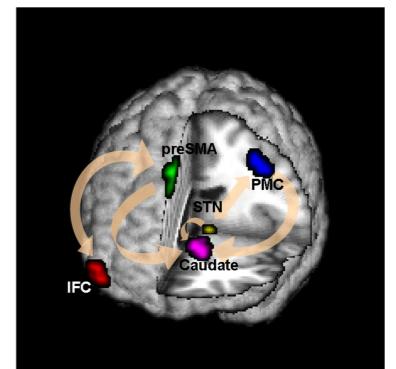
fMRI data: blood-oxygen-level dependent (BOLD) signals from
≡ each **cube**/voxel (~millimeters), $10^5 \sim 10^6$ voxels in total.

Multilevel fMRI Studies



Large, multilevel (subject, session, voxel) data
e.g. $1000 \times 4 \times 300 \times 10^6 \approx 1 \text{ trillion data points}$

Raw Data: Motor Region

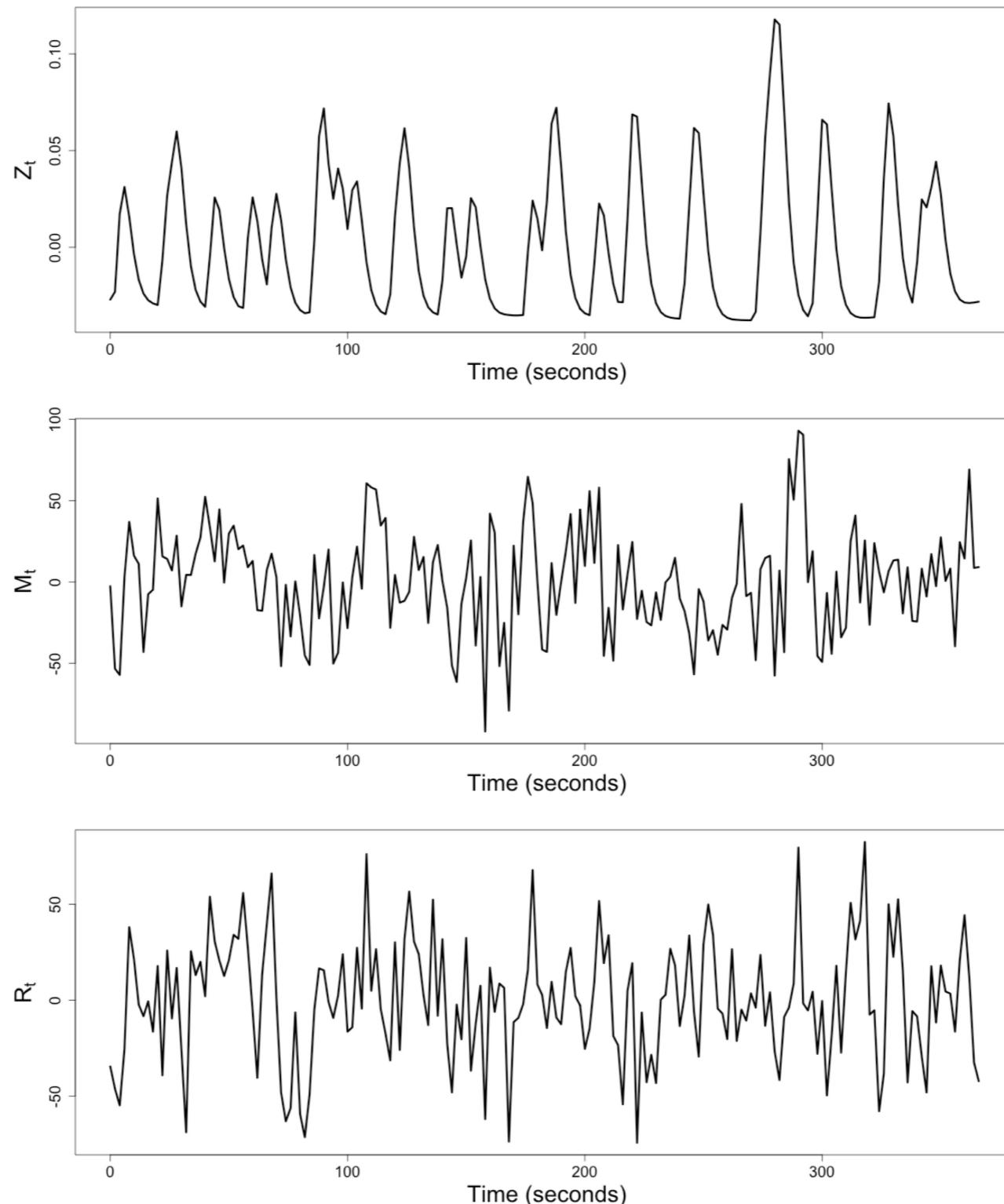


Time (seconds)

Black: fMRI BOLD activity

Blue: stop onset times; **Red:** go onset times





Z_t : Stimulus onsets convoluted with Canonical HRF

M_t, R_t : fMRI time series from two brain regions

Review: Granger Causality/VAR

- Given two (or more) time series x_t and y_t

$$x_t = \sum_{j=1}^p \psi_{1j} x_{t-j} + \sum_{j=1}^p \phi_{1j} y_{t-j} + \epsilon_{1t}$$

$$y_t = \sum_{j=1}^p \psi_{2j} y_{t-j} + \sum_{j=1}^p \phi_{2j} x_{t-j} + \epsilon_{2t}$$

- Also called vector autoregressive models
- y Granger causes x if $\phi_{1j} \neq 0$ [Granger, 69]
- Models **pair-wise** connections not **pathways**



Articles

About 8,690 results (0.04 sec)

[Any time](#)[Since 2018](#)[Since 2017](#)[Since 2014](#)[Custom range...](#)[Sort by relevance](#)[Sort by date](#) include patents include citations[Create alert](#)[Mapping directed influence over the brain using Granger causality and fMRI](#)[A Roebroeck, E Formisano, R Goebel - Neuroimage, 2005 - Elsevier](#)

We propose **Granger causality** mapping (GCM) as an approach to explore directed influences between neuronal populations (effective connectivity) in fMRI data. The method does not rely on a priori specification of a model that contains pre-selected regions and ...

99 Cited by 848 Related articles All 22 versions

[Wiener–Granger causality: a well established methodology](#)[SL Bressler, AK Seth - Neuroimage, 2011 - Elsevier](#)

For decades, the main ways to study the effect of one part of the nervous system upon another have been either to stimulate or lesion the first part and investigate the outcome in the second. This article describes a fundamentally different approach to identifying **causal** ...

99 Cited by 407 Related articles All 12 versions

[\[HTML\] Analysing connectivity with Granger causality and dynamic causal modelling](#)[K Friston, R Moran, AK Seth - Current opinion in neurobiology, 2013 - Elsevier](#)

... Such data are generated by a wide range of **neuroimaging** and neurophysiological methods ... This paper introduced the operationalization of **Granger causality** in the form of linear vector autoregressive models ... Dynamic **causal** modeling of evoked responses in EEG and MEG ...

99 Cited by 292 Related articles All 15 versions

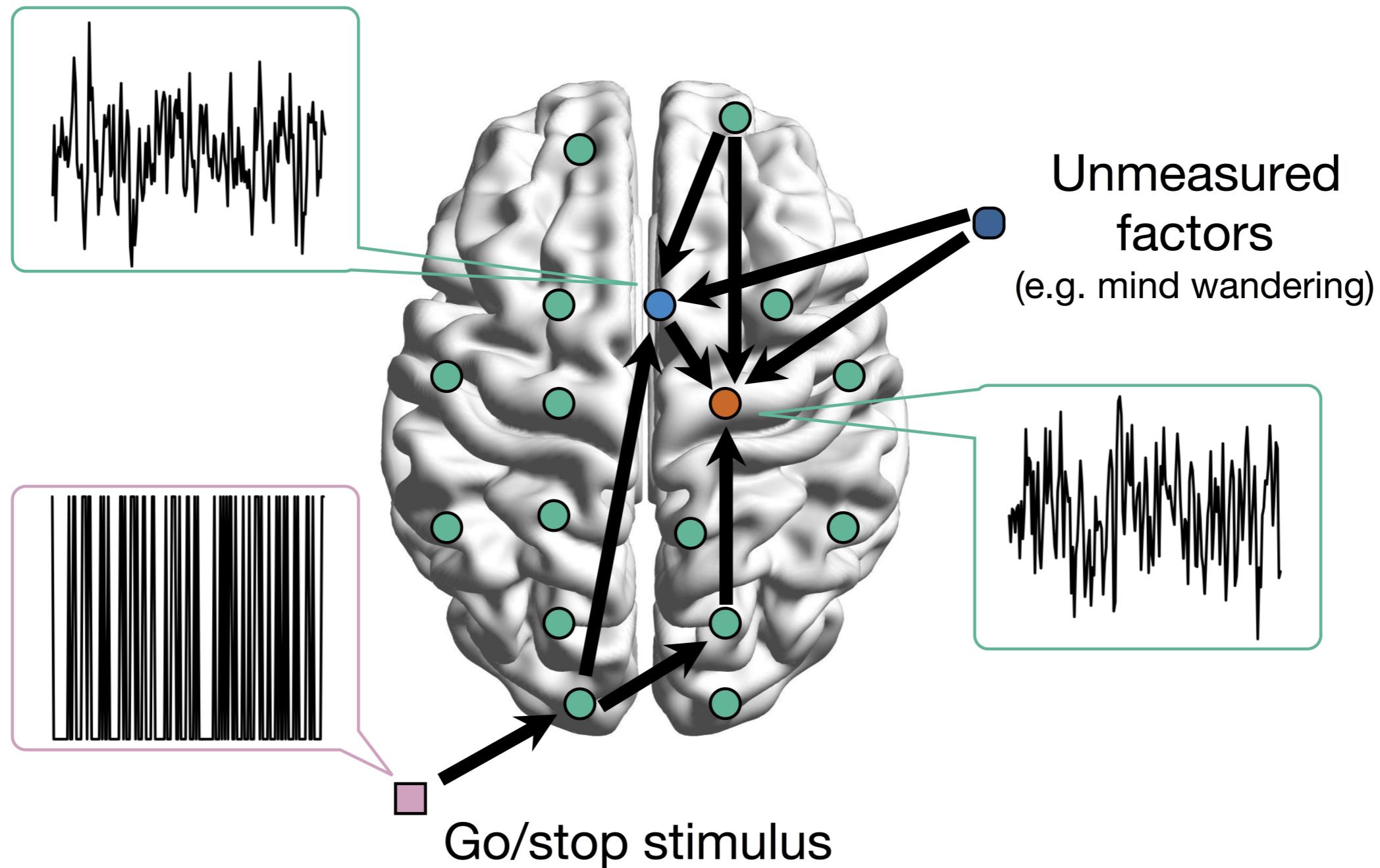


Granger Causality/VAR

- Granger Causality (VAR) popular for fMRI
 - Over **10,000** google scholar results on "granger causality neuroimaging", as of May 29, 2019
- Models multiple **stationary** time series
- AR(p) (small p) fits fMRI well [Lingdquist, 08]
- **Not** for **non-stationary**/task fMRI
- **Cannot** model stimulus effects

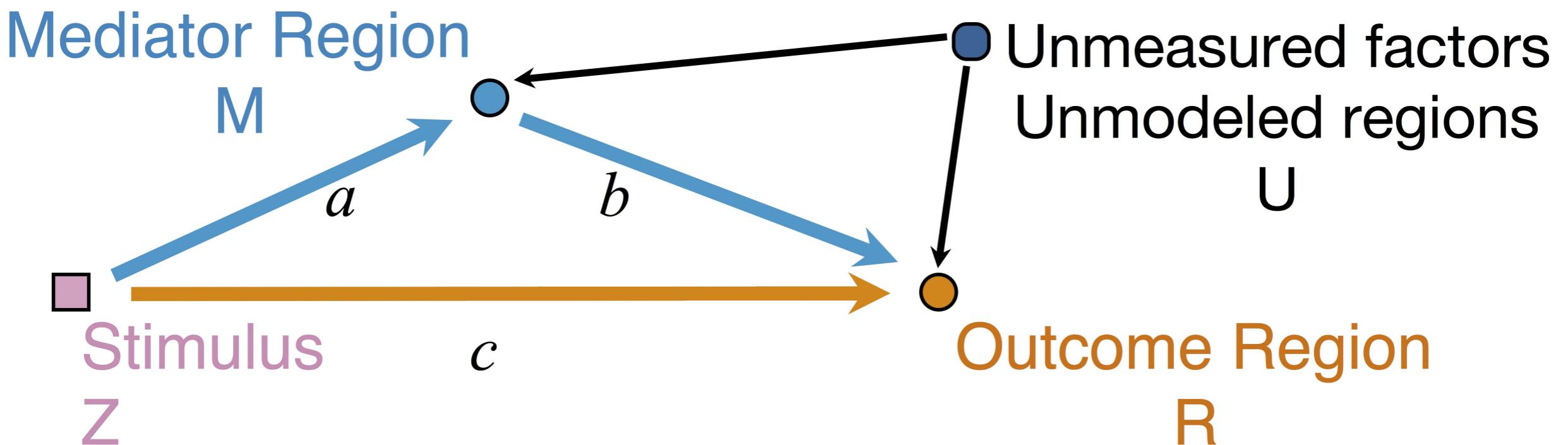


Conceptual Brain Model with Stimulus



Goal: quantify effects **stimuli** → **preSMA** → **PMC regions**

Model: Mediation Analysis and SEM



$$M = Za + \underbrace{U + \epsilon_1}_{E_1}$$
$$R = Zc + Mb + \underbrace{Ug + \epsilon_2}_{E_2}, \quad \epsilon_1 \perp \epsilon_2$$

- **Indirect** effect: $a \times b$; **Direct** effect: c
- Correlated errors: $\delta = \text{cor}(E_1, E_2) \neq 0$ if $U \neq 0$

Mediation Analysis in fMRI

- Mediation analysis (usually assuming $U = 0$)

[Baron&Kenny, 86; Sobel, 82; Holland 88; Preacher&Hayes 08; Imai et al, 10; VanderWeele, 15;...]

- Parametric [Wager et al, 09] and functional [Lindquist, 12] mediation, under (approx.) independent errors
 - Stimulus → brain → **user reported ratings**, one brain mediator
 - Assuming $U = 0$ between ratings and brain
- Multiple mediator and multiple pathways
 - Dimension reduction by arXiv1511.09354 [Chen, Crainiceanu, Ogburn, Caffo, Wager, Lindquist, 15]
 - Pathway Lasso penalization [Zhao, Luo, 16]
- **This talk:** integrating Granger causality and mediation analysis



Model & Method



Our Mediation Model

$$M_t = Z_t a + E_{1t}, \quad R_t = Z_t c + M_t b + E_{2t}$$

- **Temporal** VAR errors

$$E_{1t} = \sum_{j=1}^p (\omega_{11_j} E_{1,t-j} + \omega_{21_j} E_{2,t-j}) + \epsilon_{1t}$$

$$E_{2t} = \sum_{j=1}^p (\omega_{12_j} E_{1,t-j} + \omega_{22_j} E_{2,t-j}) + \epsilon_{2t}$$

- **Spatial** errors: $\epsilon_{1t}, \epsilon_{2t}$

$$\begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} \sim \mathcal{N}(0, \Sigma), \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \delta\sigma_1\sigma_2 \\ \delta\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$



Equivalent Form

$$M_t = Z_t A + \sum_{j=1}^p (\phi_{1j} Z_{t-j} + \psi_{11j} M_{t-j} + \psi_{21j} R_{t-j}) + \epsilon_{1t}$$

$$R_t = Z_t C + M_t B + \sum_{j=1}^p (\phi_{2j} Z_{t-j} + \psi_{12j} M_{t-j} + \psi_{22j} R_{t-j}) + \epsilon_{2t}$$

- Nonzero ϕ 's and ψ 's denote the temporal influence from stimulus to mediator/outcome and etc
- A, B, C are causal following a similar proof in

[Sobel, Lindquist, 04]



Causal Conditions

- The treatment randomization regime is the same across time and participants.
- Models are correctly specified, and no treatment-mediator interaction.
- At each time point t , the observed outcome is one realization of the potential outcome with observed treatment assignment $\bar{\mathbf{S}}_t$, where
$$\bar{\mathbf{S}}_t = (\mathbf{S}_1, \dots, \mathbf{S}_t).$$
- The treatment assignment is random across time.
- The causal effects are time-invariant.
- The time-invariant covariance matrix is not affected by the treatment assignments.
≡

Estimation: Conditional Likelihood

- The full likelihood for our model is too complex
- Given the initial p time points, the conditional likelihood is

$$\begin{aligned}\ell(\boldsymbol{\Theta}, \delta | \mathbf{Z}, \mathcal{I}_p) &= \sum_{t=p+1}^T \log f((M_t, R_t) | \mathbf{X}_t) \\ &= -\frac{T-p}{2} \log \sigma_1^2 \sigma_2^2 (1 - \delta^2) - \frac{1}{2\sigma_1^2} \|\mathbf{M} - \mathbf{X}\boldsymbol{\theta}_1\|_2^2 \\ &\quad - \frac{1}{2\sigma_2^2(1 - \delta^2)} \|(\mathbf{R} - \mathbf{MB} - \mathbf{X}\boldsymbol{\theta}_2) - \kappa(\mathbf{M} - \mathbf{X}\boldsymbol{\theta}_1)\|_2^2\end{aligned}$$



Multilevel Data: Two-level Likelihood

- Second level model, for each subject i

$$(A_i, B_i, C_i) = (A, B, C) + (\eta_i^A, \eta_i^B, \eta_i^C)$$

where errors η are normally distributed

- The two level likelihood is conditional convex
- Two-stage fitting: plug-in estimates from the first level
- Block coordinate fitting: jointly optimize first level likelihood + second level likelihood



Theorem: Assume assumptions (A1)-(A6) are satisfied.

Assume $E(Z_{i_t}^2) = q < \infty$, for $i = 1, \dots, N$. Let

$$T = \min_i T_i.$$

1. If Λ is known, then the two-stage estimator $\hat{\delta}$ maximizes the profile likelihood of model asymptotically, and $\hat{\delta}$ is \sqrt{NT} -consistent.
2. If Λ is unknown, then the profile likelihood of model has a unique maximizer $\hat{\delta}$ asymptotically, and $\hat{\delta}$ is \sqrt{NT} -consistent, provided that $1/\varpi = \bar{\kappa}^2/\varrho^2 = O_p(1/\sqrt{NT})$, $\kappa_i = \sigma_{i_2}/\sigma_{i_1}$, $\bar{\kappa} = (1/N) \sum \kappa_i$, and $\varrho^2 = (1/N) \sum (\kappa_i - \bar{\kappa})^2$.

Using the two-stage estimator $\hat{\delta}$, the CMLE of our model is consistent, as well as the estimator for $\mathbf{b} = (A, B, C)$.

Theory: Summary

- Under regularity conditions, N subs, T time points
- Our $\hat{\delta}$ is \sqrt{NT} -consistent
 - This relaxes the unmeasured confounding assumption in mediation analysis
- Our $(\hat{A}, \hat{B}, \hat{C})$ is also consistent



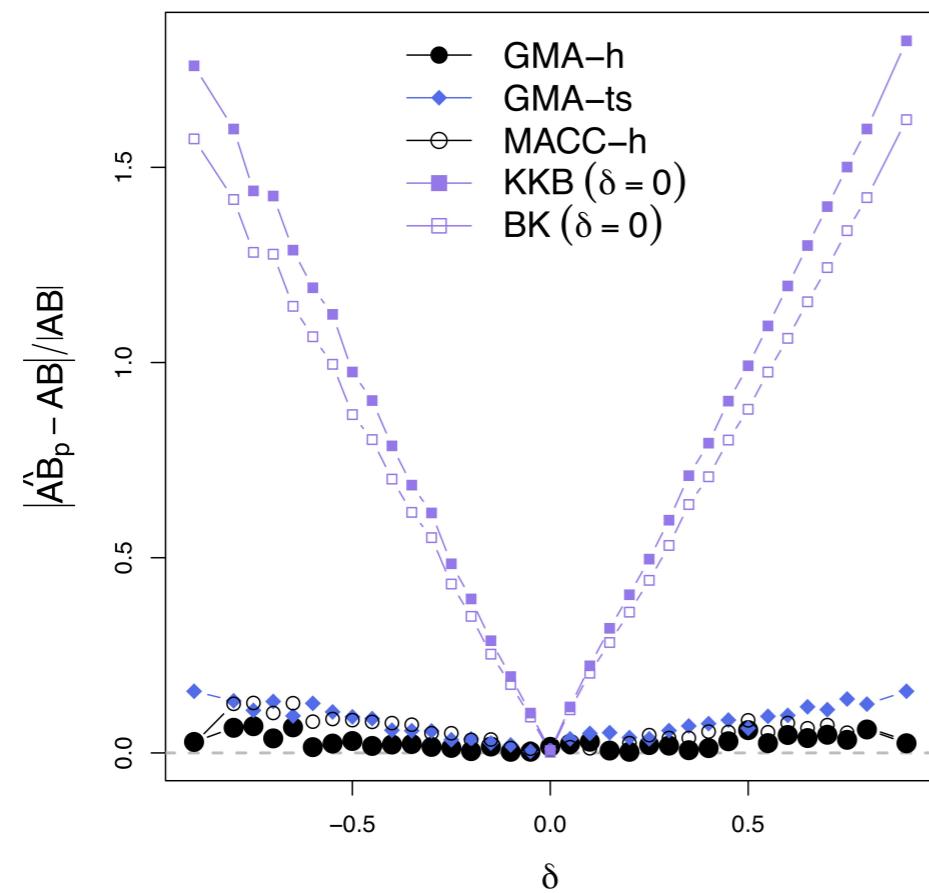
Simulations & Real Data

Comparison

- Our methods: **GMA-h** and **GMA-ts**
- Previous methods: BK^[Baron & Kenny], MACC^[Zhao and Luo], KKB^[Kenny et al]
- Other methods do not model the temporal correlations or time series like ours

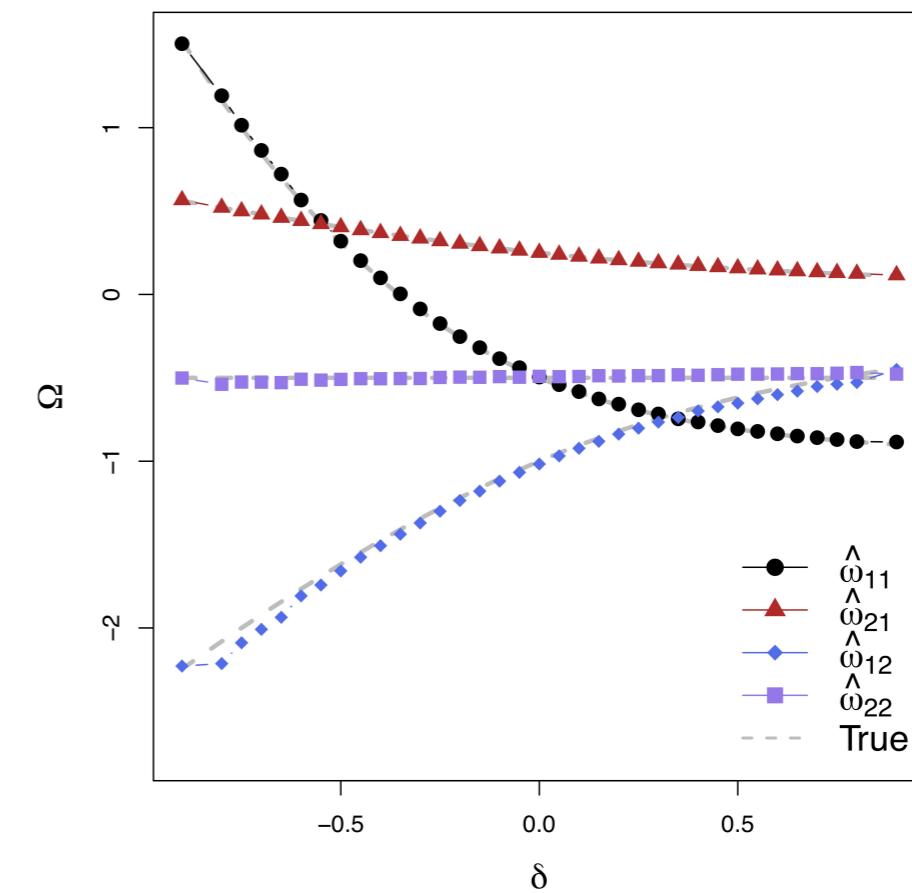


Simulations



Low bias for AB

Gray dash lines are the truth



Low bias for temporal cor

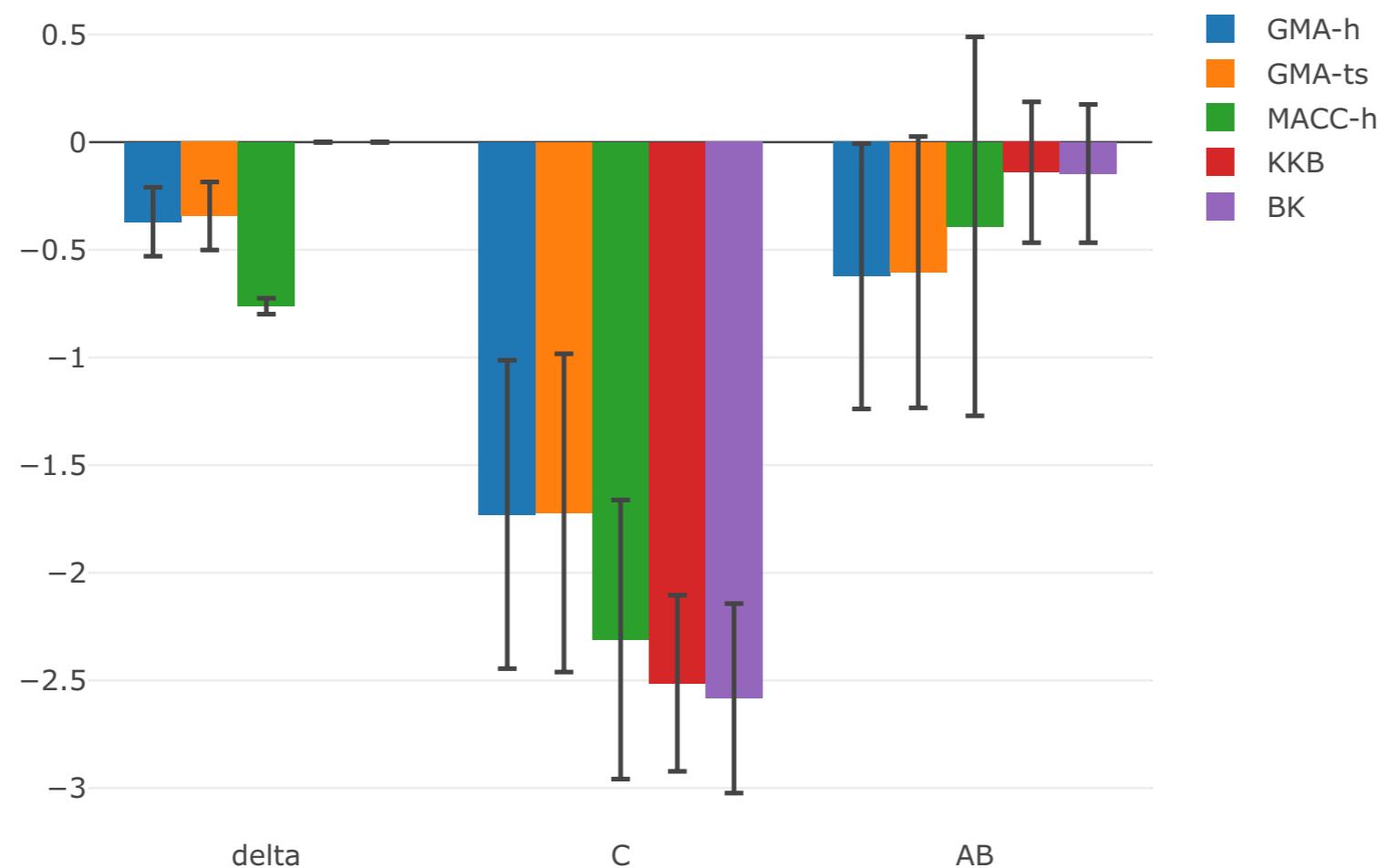
GMA performs the best, and recovers the temporal correlations

Real Data Experiment

- Public data: OpenFMRI ds30
- Stop-go experiment: withhold (STOP) from pressing buttons
- Expect "STOP" stimuli to deactivate brain region M1
- Goal: quantify the role of region preSMA



Result



Result

Method	δ	C	A	B	AB_p
GMA-h	-0.370 (-0.530, -0.156)	-1.729 (-2.445, -0.964)	-0.739 (-1.487, 0.035)	0.838 (0.644, 0.999)	-0.623 (-1.239, 0.033)
GMA-ts	-0.343 (-0.501, -0.163)	-1.722 (-2.461, -0.904)	-0.740 (-1.442, 0.080)	0.810 (0.656, 0.965)	-0.604 (-1.234, 0.055)
MACC-h	-0.762 (-0.799, -0.721)	-2.310 (-2.958, -1.641)	-0.259 (-0.810, 0.303)	1.511 (1.410, 1.619)	-0.391 (-1.271, 0.465)
KKB	-	-2.513 (-2.922, -2.073)	-0.225 (-0.772, 0.326)	0.617 (0.589, 0.647)	-0.140 (-0.467, 0.196)
BK	-	-2.583 (-3.023, -2.142)	-0.235 (-0.774, 0.352)	0.616 (0.588, 0.647)	-0.146 (-0.467, 0.211)

- STOP deactivates M1 directly (C) and indirectly (AB)
- preSMA mediates a good portion of the total effect
 - Help resolve the debates among neuroscientists
- Other methods under-estiamte the effects
- Novel feedback findings: M1 → preSMA after lag 1 and 2 (not shown)



Discussion

- Mediation analysis for multiple time series
- Method: Granger causality + mediation
 - Optimizing complex likelihood
- Theory: identifiability, consistency
- Result: low bias and improved accuracy
- Extension: functional mediation
- Paper in Biometrics
- CRAN pkg: **gma** and references within



Thank you!

Comments? Questions?

BigComplexData.com

or BrainDataScience.com

