Assessing Directional Connectivity via Network Localized Granger Causality Extracted from MEG Data

Behrad Soleimani[*1,2], *I.M. Dushyanthi Karunathilake*[1,2], *Proloy Das*[3], *Stefanie E. Kuchinsky*[4], *Jonathan Z. Simon*[1,2,5], *Behtash Babadi*[1,2]

[1]Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, [2]Institute for Systems Research, University of Maryland, College Park, MD, [3]Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, Boston, MA, USA, [4]Audiology and Speech Pathology Center, Walter Reed National Military Medical Center, Bethesda, MD, USA, [5]Department of Biology, University of Maryland College Park, MD, USA

*behrad@umd.edu

Introduction

- Identifying causal relationships between different cortical areas for understanding mechanisms behind sensory processing
- Connectivity characterized by the temporal predictability of activity across brain regions via Granger causality (GC)
- Challenges with Magnetoencephalography (MEG): the data are low-dimensional, noisy, and linearly mixed versions of underlying source activities
- Conventional methods (two-stage procedure):

MEG Data → Source Localization → GC Inference

- Drawbacks: bias propagation, spatial leakage

- Goal: *directly* localize GC influences without an intermediate source localization step
- Method: Network Localized Granger Causality (NLGC)

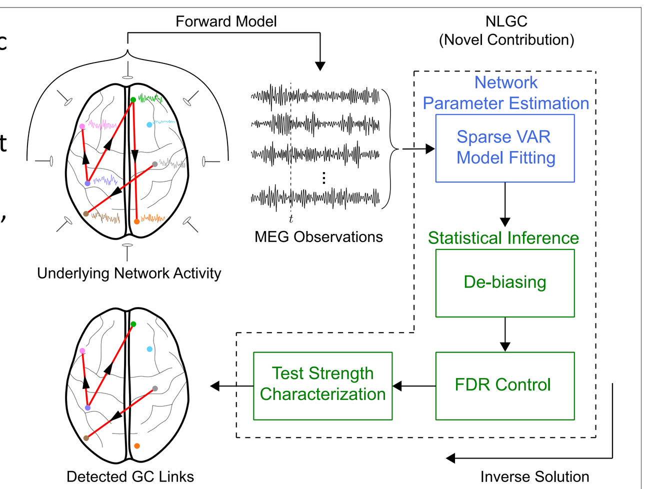


Fig. 1. Schematic depiction of the proposed NLGC inference. Without an intermediate source localization, the cortical connectivity is obtained directly from MEG observations.

Model

- Observation model:

$$\mathbf{y}_t = \mathbf{C}\mathbf{x}_t + \mathbf{n}_t, \quad t = 1, 2, \cdots, T$$

$\mathbf{y}_t \in \mathbb{R}^M$ MEG observation, $\mathbf{C} \in \mathbb{R}^{M\times N}$ lead field matrix
$\mathbf{x}_t \in \mathbb{R}^N$ source activity, $\mathbf{n}_t \in \mathbb{R}^M$ measurement noise

- Source dynamic model (auto-regressive):

$$\mathbf{x}_t = \sum_{k=1}^{q} \mathbf{A}_k \mathbf{x}_{t-k} + \mathbf{w}_t, \quad t = 1, 2, \cdots, T$$

$\mathbf{A}_k \in \mathbb{R}^{N\times N}$ coefficient matrix, $\mathbf{w}_t \in \mathbb{R}^N$ noise process

- Distributional assumptions:

$\mathbf{n}_t \sim$ zero-mean Gaussian (known covariance)
$\mathbf{w}_t \sim$ zero-mean Gaussian, independent sources (unknown diagonal covariance $\mathbf{Q}$)

Granger Causality

- Consider link $(\tilde{i} \to i)$ with following models:

$$\text{Full: } \mathbf{x}_t^{(i)} = \sum_j \sum_k a_{i,j,k} \mathbf{x}_{t-k}^{(j)} + \mathbf{w}_t^{(i)}, \quad \mathbf{w}_t^{(i)} \sim \mathcal{N}(0, \sigma_i^2)$$

$$\text{Reduced: } \mathbf{x}_t^{(i)} = \sum_{j\neq\tilde{i}} \sum_k a'_{i,j,k} \mathbf{x}_{t-k}^{(j)} + \mathbf{w'}_t^{(i)}, \quad \mathbf{w'}_t^{(i)} \sim \mathcal{N}(0, \sigma_{i\setminus\tilde{i}}^2)$$

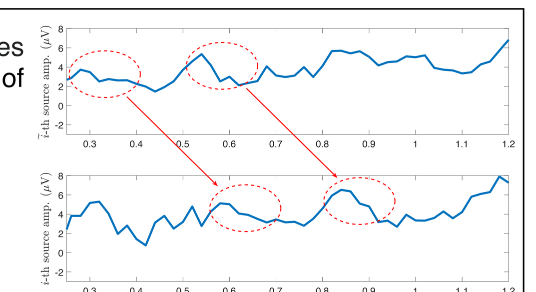
- Granger Causality (GC) measure:

$$\mathcal{F}_{(\tilde{i}\to i)} = \log\left(\frac{\sigma_{i\setminus\tilde{i}}^2}{\sigma_i^2}\right)$$

relative predictive variance explained

- $\mathcal{F}_{(\tilde{i}\to i)} \gg 0$: GC link exists.

Fig. 2. GC link $(\tilde{i} \to i)$ implies temporal predictability of source i by $\tilde{i}$.



Parameter Estimation†

- Challenge: source activities are unknown
- Solution: Expectation Maximization (EM)
- At the l-th iteration:

$$\text{E-step: } Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(l)}) = \mathrm{E}\left[\log p(\mathbf{x}_{1:T}, \mathbf{y}_{1:T}; \boldsymbol{\theta}) \middle| \mathbf{y}_{1:T}; \widehat{\boldsymbol{\theta}}^{(l)}\right]$$

$$\text{M-step: } \widehat{\boldsymbol{\theta}}^{(l+1)} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \left\{Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(l)}) + R_{\ell_1}(\boldsymbol{\lambda}, \boldsymbol{\theta})\right\}$$

- ℓ_1-norm regularization is utilized at the M-step to mitigate the ill-posedness resulting from the low-dimensional measurements

Statistical Inference†

- Two hypotheses for link $(\tilde{i} \to i)$:

$H_{(\tilde{i}\mapsto i),0}$: there is no GC influence
$H_{(\tilde{i}\mapsto i),1}$: there is a GC influence

- Asymptotic distributions:

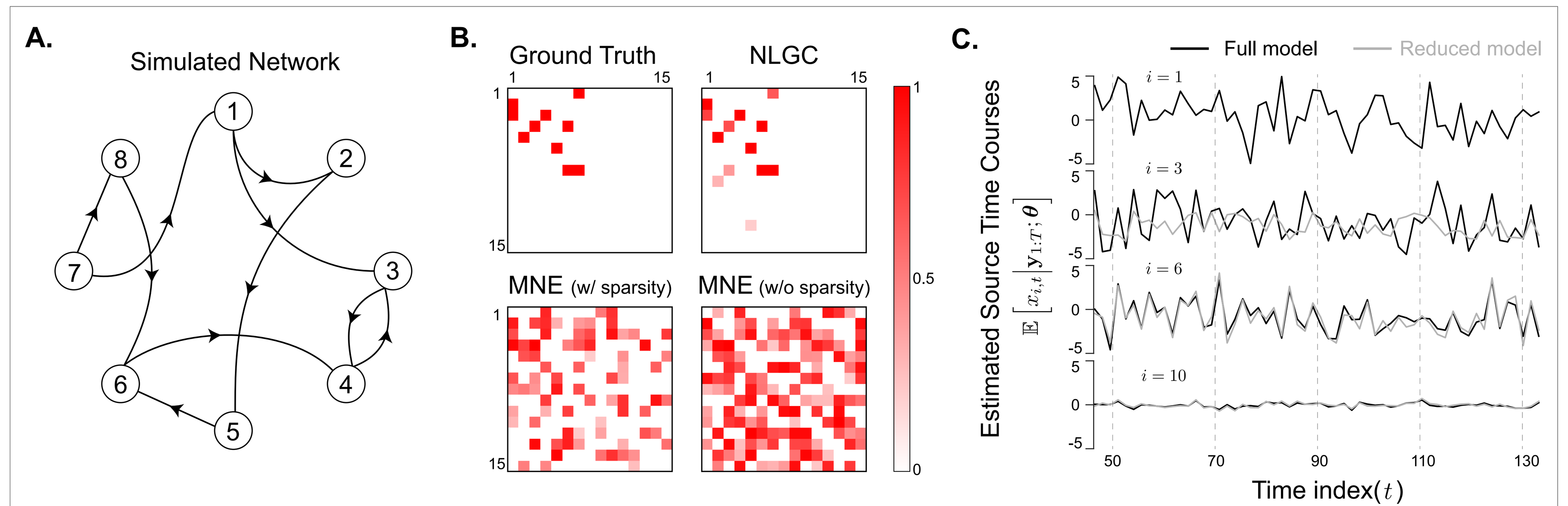
$$[\mathcal{D}_{(\tilde{i}\to i)} | H_{(\tilde{i}\mapsto i),0}] \xrightarrow{d} \chi^2(q)$$

$$[\mathcal{D}_{(\tilde{i}\to i)} | H_{(\tilde{i}\mapsto i),1}] \xrightarrow{d} \chi^2(q, \nu_{(\tilde{i}\to i)})$$

- False discovery rate (FDR) control:
  - Reject null hypothesis at a confidence level α
  - Control FDR via BY procedure

†For details and more explanations, please check the paper.

Results: Simulation†

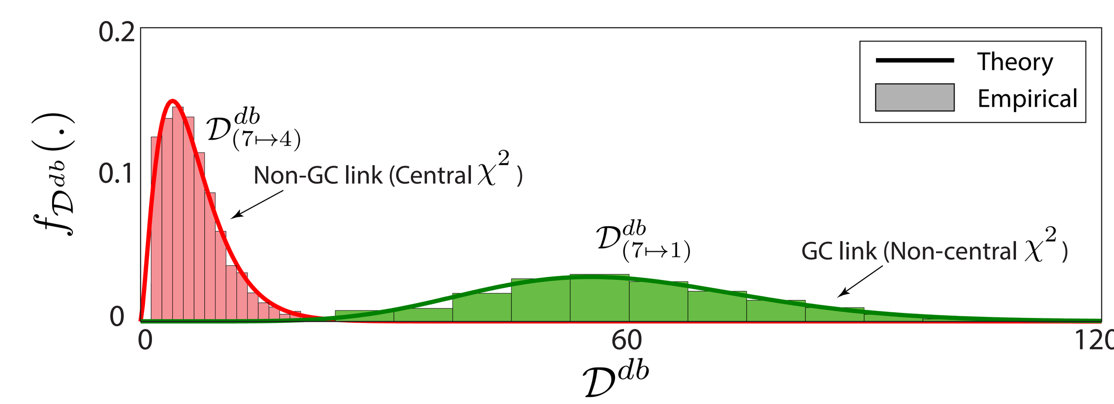


NLGC outperforms the two-stage procedure both in terms of hit rate and false alarm rate.

Fig. 3. A. The underlying true GC network between the active sources indexed by 1, 2, … , 8 (explaining 90% of the power of the 84 sources). The remaining 76 sources are silent and are modeled as independent white noise accounting for the remaining 10% of the source power. **B.** The ground truth and estimated GC maps using NLGC and MNE (with and without accounting for sparsity). Only a subset of sources indexed by 1, 2, … , 15 are shown for visual convenience. NLGC fully captures the true links with only a few false detections; on the other hand, the two-stage approaches using MNE, capture around half of the true links, but also detect numerous spurious links. While enforcing sparsity mildly mitigates the false alarm performance of the two-stage approach, it is unable to resolve it. **C.** Estimated activity time-courses of the patches with index 1, 3, 6, and 10 based on full models and the reduced models corresponding to the GC link (1 ↦ 3) and non-GC links (1 ↦ 6) and (1 ↦ 10) as examples. As expected, since the GC link (1 ↦ 3) exists, removing the 1st patch contribution from the VAR model of the 3rd patch dramatically changes the predicted activity of patch 3 (second line). However, this is not the case for the other two examples, since the links (1 ↦ 6) and (1 ↦ 10) do not exist (third and fourth lines).

Empirical distribution of the deviance closely matches the theoretical results.

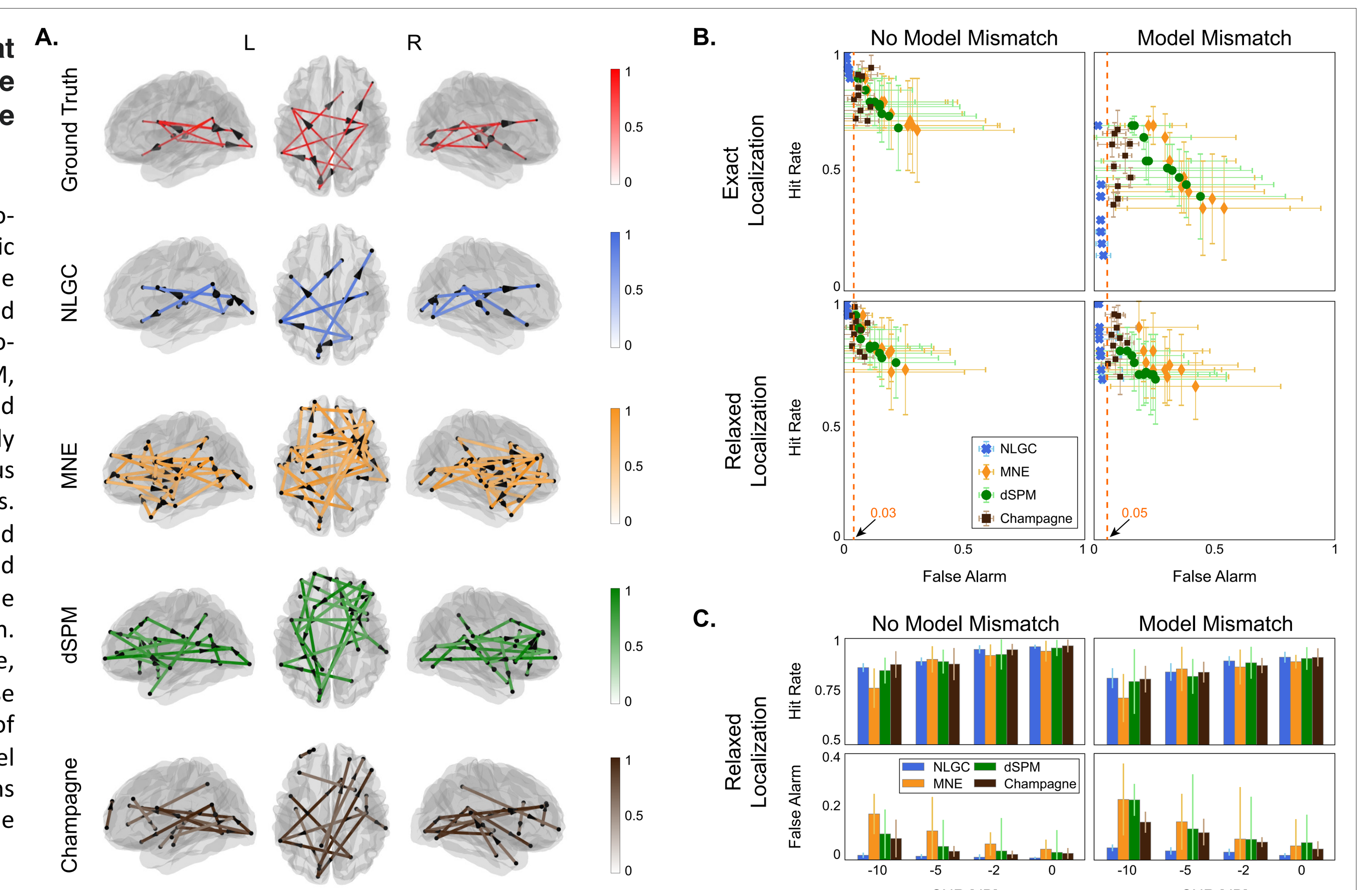
Fig. 4. Theoretical and empirical distributions of the de-biased deviance differences corresponding to the GC link (7 ↦ 1) and non-GC link (7 ↦ 4) from the setting of Fig. 3., for 200 different realizations of the noise processes.



Results: Synthetic Data†

Simulation results suggest that NLGC is more reliable compared to the two-stage procedures.

Fig. 5. Comparison of NLGC with two-stage procedures using a realistic simulation setting. **A.** Example of the ground truth GC network, and estimates obtained by NLGC and two-stage approaches based on MNE, dSPM, and Champagne overlaid on dorsal and lateral brain plots. NLGC captures nearly all the existing GC links with no spurious detection. **B.** ROC curves (hit rate vs. false alarm) corresponding to NLGC, and two-stage approaches for exact/relaxed link localization and in the presence/absence of model mismatch. NLGC provides equal or better hit rate, while consistently maintaining low false alarm rate. **C.** Evaluating the effect of SNR in presence/absence of model mismatch. NLGC consistently maintains low false alarm rates across a wide range of SNR settings.

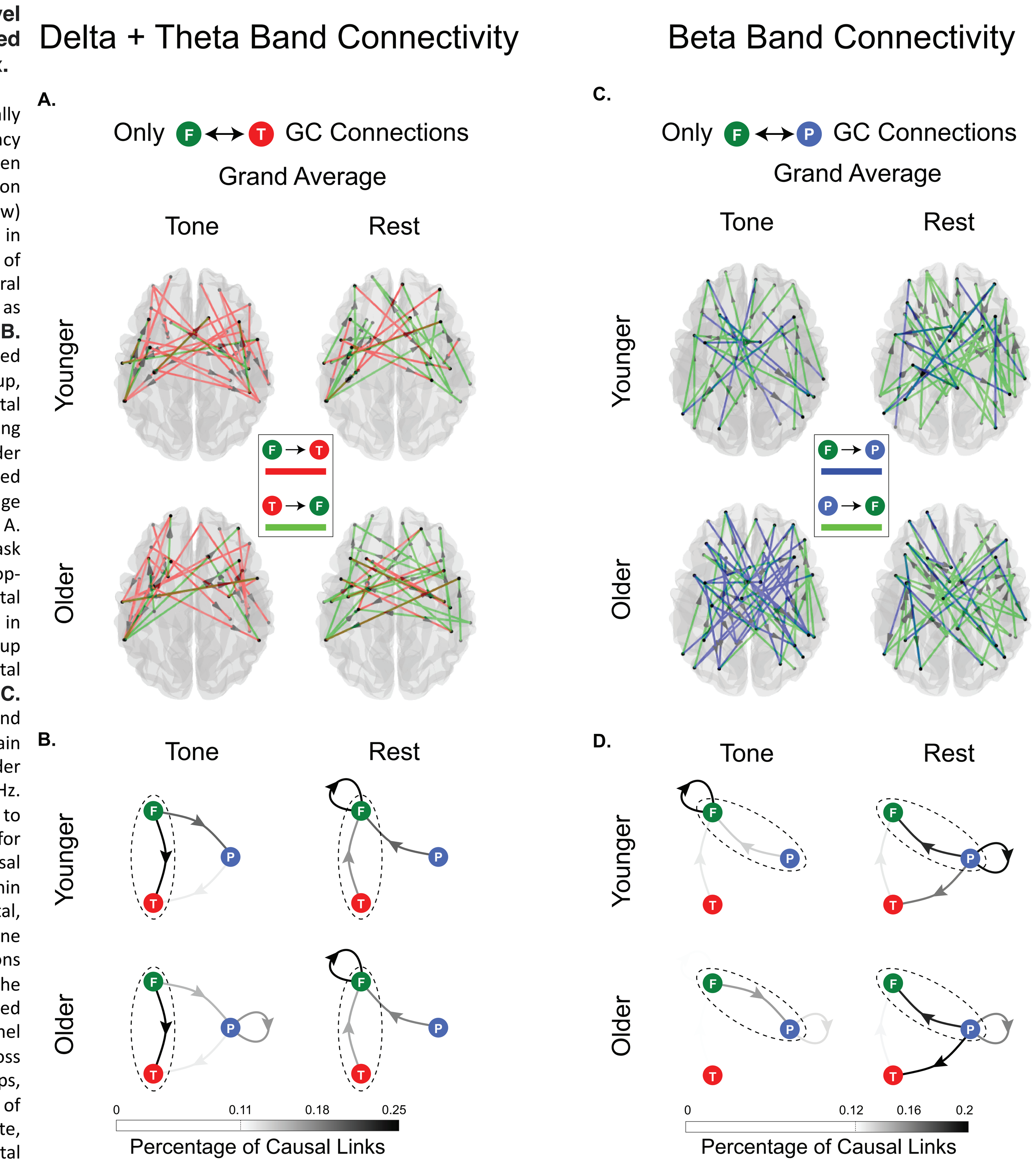


Results: Tone Processing vs. Resting State†

- 13 younger and 9 older adults
- 100 repetitions of tone pips presented at the end of resting state recordings
- Two 40-second trials per subject/condition
- Connectivity in auditory cortex is investigated

NLGC identifies network-level age- and condition-related changes in the auditory cortex.

Fig. 6. NLGC analysis of experimentally recorded MEG data in two frequency bands. **A.** Extracted GC links between frontal and temporal areas overlaid on dorsal brain plots for younger (top row) and older (bottom row) participants in 0.1-8 Hz. There is a notable increase of top-down links from frontal to temporal areas during tone processing as compared to the resting state. **B.** Percentage of causal links, averaged over subjects within each age group, between frontal, temporal, and parietal areas for tone processing vs. resting state conditions and younger vs. older participants in 0.1-8 Hz. The dashed ovals indicate the normalized average number of links shown in panel A. There are notable changes across task conditions, including dominantly top-down frontal to temporal/parietal connections during tone processing, in contrast to dominantly bottom-up temporal/parietal to frontal connections during resting state. **C.** Extracted GC links between frontal and parietal areas overlaid on dorsal brain plots for younger (top row) and older (bottom row) participants in 13-25 Hz. There is a notable increase of frontal to parietal links under tone processing for older adults. **D.** Percentage of causal links, averaged over subjects within each age group, between frontal, temporal, and parietal areas for tone processing vs. resting state conditions and younger vs. older participants. The dashed ovals indicate the normalized average number of links shown in panel C. There are notable changes across both task conditions and age groups, including the higher involvement of parietal areas during resting state, increase of frontal to frontal connections for younger participants and top-down links from frontal to parietal areas for older participants, during tone processing.



Reference

Paper:
B. Soleimani, P. Das, I.M. D. Karunathilake, S. E. Kuchinsky, J. Z. Simon, and B. Babadi, NLGC: Network Localized Granger Causality with Application to MEG Directional Functional Connectivity Analysis. *NeuroImage, Vol. 260, 119496, 2022.* DOI: https://doi.org/10.1016/j.neuroimage.2022.119496

Python Package:
Soleimani B, Das P. Network Localized Granger Causality. (2022) GitHub Repository at https://github.com/BabadiLab/NLGC

Supported by NSF (OISE2020624, SMA1734892 and CCF1552946) and NIH (R01-DC019394, R01- 970 DC014085, P01-AG055365, and R21-AG068802).