





Neuro-Current Response Functions: A Unified Approach to MEG Source Analysis Under Continuous Stimuli Paradigm

Proloy Das, Christian Brodbeck, Jonathan Z Simon, Behtash Babadi University of Maryland

proloy@umd.edu

Presentation available at



Encoding Model of Speech Processing



Linear filter model:

predict the M/EEG response from continuous stimulus

Lalor et. al., 2006, 2009, 2010

Encoding Model of Speech Processing



Linear filter model:

predict the M/EEG response from continuous stimulus

Functional Role:

- Different peaks are associated with different modes of processing
- M50: encoding of acoustic-level features
- M100: encoding of attentional state



Encoding Model of Speech Processing



Linear filter model:

predict the M/EEG response from continuous stimulus

Functional Role:

- Different peaks are associated with different modes of processing
- M50: encoding of acoustic-level features
- M100: encoding of attentional state

Beyond acoustic processing:

- Phoneme level processing
- Lexical processing
- semantic processing



Di Liberto et. al., 2015 Brodbeck et. al., 2018 Boderick et. al., 2018

Insight from intracranial recordings

- Electrophysiology, electrocorticography
- Limited spatial range

Existing M/EEG approaches works in two-stage:

- estimate the TRFs for each sensor \rightarrow localize them on cortical mantle Lalor et. al., 2009
- Decompose MEG signals to source time courses \rightarrow estimate TRF for each source

Challenges:

Bias propagation, spatio-temporal leakage, sensitivity to forward model



Mesgarani et. al., 2008, 2014

Insight from intracranial recordings

- Electrophysiology, electrocorticography
- Limited spatial range

Existing M/EEG approaches works in two-stage:

- estimate the TRFs for each sensor \rightarrow localize them on cortical mantle Lalor et. al., 2009
- Decompose MEG signals to source time courses \rightarrow estimate TRF for each source

Challenges:

Bias propagation, spatio-temporal leakage, sensitivity to forward model





Mesgarani et. al., 2008, 2014

Brodbeck et. al., 2018

N: sensors

Insight from intracranial recordings

- Electrophysiology, electrocorticography
- Limited spatial range

Existing M/EEG approaches works in two-stage:

- estimate the TRFs for each sensor \rightarrow localize them on cortical mantle Lalor et. al., 2009
- Decompose MEG signals to source time courses \rightarrow estimate TRF for each source

Challenges:

Bias propagation, spatio-temporal leakage, sensitivity to forward model



N: sensors



Mesgarani et. al., 2008, 2014

Insight from intracranial recordings

- Electrophysiology, electrocorticography
- Limited spatial range

Existing M/EEG approaches works in two-stage:

- estimate the TRFs for each sensor \rightarrow localize them on cortical mantle Lalor et. al., 2009
- ► Decompose MEG signals to source time courses → estimate TRF for each source

Challenges:

Q

Bias propagation, spatio-temporal leakage, sensitivity to forward model

MEG observationsLead-field matrixSource time courses
$$\mathbf{Y}_{N \times T} = \mathbf{L}_{N \times 3M} \mathbf{J}_{3M \times T} + \mathbf{W}_{N \times T}$$
uasi-static solution to
Maxwell's equations $\mathbf{j}_{m,t} := [j_{m,t,R}, j_{m,t,A}, j_{m,t,S}]^\top \in \mathbb{R}^3$ $j_{m,t,i} = (\boldsymbol{\tau}_{m,i})^\top \mathbf{e}_t + v_{m,t,i}$ $i \in \{R, A, S\}$ $\boldsymbol{\Phi} := [\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \cdots, \boldsymbol{\tau}_M]^\top$ $\mathbf{S} := [\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_T]$ N : sensors T : time points M : sources L : lags



Mesgarani et. al., 2008, 2014

Insight from intracranial recordings

- Electrophysiology, electrocorticography
- Limited spatial range

Existing M/EEG approaches works in two-stage:

- estimate the TRFs for each sensor \rightarrow localize them on cortical mantle Lalor et. al., 2009
- Decompose MEG signals to source time courses \rightarrow estimate TRF for each source

Challenges:

Bias propagation, spatio-temporal leakage, sensitivity to forward model





Mesgarani et. al., 2008, 2014

Insight from intracranial recordings

- Electrophysiology, electrocorticography
- Limited spatial range

Existing M/EEG approaches works in two-stage:

- estimate the TRFs for each sensor \rightarrow localize them on cortical mantle Lalor et. al., 2009
- Decompose MEG signals to source time courses \rightarrow estimate TRF for each source

Challenges:

Bias propagation, spatio-temporal leakage, sensitivity to forward model





Mesgarani et. al., 2008, 2014

NCRF Learning: Bayesian Estimation

Measurement noise: Gaussian

Gaussian
$$\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_w)$$

 $p(\mathbf{Y}|\mathbf{J}) = |(2\pi)\mathbf{\Sigma}_w|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{J}\|_{\mathbf{\Sigma}_w^{-1}}^2\right)$

Background activity: Gaussian $\mathcal{N}(\mathbf{0}, \mathbf{\Gamma})$

$$p(\mathbf{V}|\mathbf{\Gamma}) = \left(\prod_{m=1}^{M} |(2\pi)\mathbf{\Gamma}_{m}|^{-T/2}\right) \exp\left(-\frac{1}{2}\sum_{m=1}^{M} \|\mathbf{V}_{m}\|_{\mathbf{\Gamma}_{m}^{-1}}^{2}\right)$$
$$p(\mathbf{J}|\mathbf{\Phi},\mathbf{\Gamma}) = |(2\pi)\mathbf{\Gamma}|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{J}-\mathbf{\Phi}\mathbf{S}\|_{\mathbf{\Gamma}^{-1}}^{2}\right)$$

Joint density of measurements and current dipoles:

$$p\left(\mathbf{Y}, \mathbf{J} | \boldsymbol{\Phi}, \boldsymbol{\Gamma}\right) = |(2\pi) \boldsymbol{\Sigma}_w|^{-T/2} |(2\pi) \boldsymbol{\Gamma}|^{-T/2} \exp\left(-\frac{1}{2} \|\mathbf{Y} - \mathbf{L}\mathbf{J}\|_{\boldsymbol{\Sigma}_w^{-1}}^2 - \frac{1}{2} \|\mathbf{J} - \boldsymbol{\Phi}\mathbf{S}\|_{\boldsymbol{\Gamma}^{-1}}^2\right)$$





NCRF Learning: Bayesian Estimation

Measurement noise: Gaussian

$$p(\mathbf{Y}|\mathbf{J}) = |(2\pi)\boldsymbol{\Sigma}_w|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{J}\|_{\boldsymbol{\Sigma}_w^{-1}}^2\right)$$

 $\mathcal{N}(\mathbf{0} \mathbf{\Sigma}_{\mathrm{out}})$

Background activity: Gaussian $\mathcal{N}(\mathbf{0}, \mathbf{\Gamma})$

$$p(\mathbf{V}|\mathbf{\Gamma}) = \left(\prod_{m=1}^{M} |(2\pi)\mathbf{\Gamma}_{m}|^{-T/2}\right) \exp\left(-\frac{1}{2}\sum_{m=1}^{M} \|\mathbf{V}_{m}\|_{\mathbf{\Gamma}_{m}^{-1}}^{2}\right)$$
$$p(\mathbf{J}|\mathbf{\Phi},\mathbf{\Gamma}) = |(2\pi)\mathbf{\Gamma}|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{J}-\mathbf{\Phi}\mathbf{S}\|_{\mathbf{\Gamma}^{-1}}^{2}\right)$$

Joint density of measurements and <u>current dipoles</u>:

$$p(\mathbf{Y}, \mathbf{J} | \mathbf{\Phi}, \mathbf{\Gamma}) = |(2\pi) \mathbf{\Sigma}_w|^{-T/2} |(2\pi) \mathbf{\Gamma}|^{-T/2} \exp\left(-\frac{1}{2} \|\mathbf{Y} - \mathbf{L}\mathbf{J}\|_{\mathbf{\Sigma}_w^{-1}}^2 - \frac{1}{2} \|\mathbf{J} - \mathbf{\Phi}\mathbf{S}\|_{\mathbf{\Gamma}^{-1}}^2\right)$$

Marginal density of the measurements given NCRFs:

$$p\left(\mathbf{Y}|\boldsymbol{\Phi},\boldsymbol{\Gamma}\right) = \left|\left(2\pi\right)\left(\boldsymbol{\Sigma}_{w} + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^{\top}\right)\right|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{Y} - \mathbf{L}\boldsymbol{\Phi}\mathbf{S}\|^{2}_{\left(\boldsymbol{\Sigma}_{w} + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^{\top}\right)^{-1}}\right)$$



not obsorved

NCRF Learning: Bayesian Estimation

Measurement noise: Gaussian

$$p(\mathbf{Y}|\mathbf{J}) = |(2\pi)\boldsymbol{\Sigma}_w|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{J}\|_{\boldsymbol{\Sigma}_w^{-1}}^2\right)$$

 $\mathcal{N}(\mathbf{0} \mathbf{\Sigma}_{\mathrm{out}})$

Background activity: Gaussian $\mathcal{N}(\mathbf{0}, \mathbf{\Gamma})$

$$p(\mathbf{V}|\mathbf{\Gamma}) = \left(\prod_{m=1}^{M} |(2\pi)\mathbf{\Gamma}_{m}|^{-T/2}\right) \exp\left(-\frac{1}{2}\sum_{m=1}^{M} \|\mathbf{V}_{m}\|_{\mathbf{\Gamma}_{m}^{-1}}^{2}\right)$$
$$p(\mathbf{J}|\mathbf{\Phi},\mathbf{\Gamma}) = |(2\pi)\mathbf{\Gamma}|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{J}-\mathbf{\Phi}\mathbf{S}\|_{\mathbf{\Gamma}^{-1}}^{2}\right)$$

Joint density of measurements and <u>current dipoles</u>:

$$p(\mathbf{Y}, \mathbf{J} | \mathbf{\Phi}, \mathbf{\Gamma}) = |(2\pi) \mathbf{\Sigma}_w|^{-T/2} |(2\pi) \mathbf{\Gamma}|^{-T/2} \exp\left(-\frac{1}{2} \|\mathbf{Y} - \mathbf{L}\mathbf{J}\|_{\mathbf{\Sigma}_w^{-1}}^2 - \frac{1}{2} \|\mathbf{J} - \mathbf{\Phi}\mathbf{S}\|_{\mathbf{\Gamma}^{-1}}^2\right)$$

Marginal density of the measurements given NCRFs:

$$p\left(\mathbf{Y}|\boldsymbol{\Phi},\boldsymbol{\Gamma}\right) = \left|\left(2\pi\right)\left(\boldsymbol{\Sigma}_{w} + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^{\top}\right)\right|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{Y} - \mathbf{L}\boldsymbol{\Phi}\mathbf{S}\|_{\left(\boldsymbol{\Sigma}_{w} + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^{\top}\right)^{-1}}^{2}\right)$$

Maximum likelihood estimate:

$$\min_{\mathbf{\Phi}} \quad \frac{1}{2} \|\mathbf{Y} - \mathbf{L} \mathbf{\Phi} \mathbf{S}\|_{(\mathbf{\Sigma}_w + \mathbf{L} \mathbf{\Gamma} \mathbf{L}^\top)^{-1}}^2$$

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

not obcorved



Maximum likelihood estimate:

$$\min_{\mathbf{\Phi}} \quad \frac{1}{2} \|\mathbf{Y} - \mathbf{L} \mathbf{\Phi} \mathbf{S}\|_{(\mathbf{\Sigma}_w + \mathbf{L} \mathbf{\Gamma} \mathbf{L}^\top)^{-1}}^2$$

But

Maximum likelihood estimate:

$$\min_{\mathbf{\Phi}} \quad \frac{1}{2} \|\mathbf{Y} - \mathbf{L} \mathbf{\Phi} \mathbf{S}\|_{(\mathbf{\Sigma}_w + \mathbf{L} \mathbf{\Gamma} \mathbf{L}^\top)^{-1}}^2$$

But

NCRFs may not be smooth



Maximum likelihood estimate:

$$\min_{\mathbf{\Phi}} \quad \frac{1}{2} \|\mathbf{Y} - \mathbf{L} \mathbf{\Phi} \mathbf{S}\|_{(\mathbf{\Sigma}_w + \mathbf{L} \mathbf{\Gamma} \mathbf{L}^\top)^{-1}}^2$$

But

NCRFs may not be smooth

Sol: Represent Φ using a Gabor dictionary, ${f G}$

 $egin{array}{ll} \Phi = \Theta \mathbf{G}^{ op} \ \widetilde{\mathbf{S}} := \mathbf{G}^{ op} \mathbf{S} \end{array}$

 $\min_{\Theta} \qquad \frac{1}{2} \|\mathbf{Y} - \mathbf{L}\boldsymbol{\Theta}\widetilde{\mathbf{S}}\|_{\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^{\top - 1}}$



Maximum likelihood estimate:

$$\min_{\mathbf{\Phi}} \quad \frac{1}{2} \|\mathbf{Y} - \mathbf{L} \mathbf{\Phi} \mathbf{S}\|_{(\mathbf{\Sigma}_w + \mathbf{L} \mathbf{\Gamma} \mathbf{L}^\top)^{-1}}^2$$

But

NCRFs may not be smooth

Sol: Represent $oldsymbol{\Phi}$ using a Gabor dictionary, $oldsymbol{G}$

$$egin{array}{lll} m{\Phi} = m{\Theta} \mathbf{G}^ op\\ \widetilde{\mathbf{S}} := \mathbf{G}^ op \mathbf{S} \end{array}$$

 $\min_{\Theta} \qquad \frac{1}{2} \|\mathbf{Y} - \mathbf{L}\Theta\widetilde{\mathbf{S}}\|_{\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^{\top - 1}}$



Maximum likelihood estimate:

$$\min_{\boldsymbol{\Theta}} \qquad \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top - 1}}$$



Maximum likelihood estimate:

Θ

$$\min_{\boldsymbol{\Theta}} \qquad \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top - 1}}$$

But

NCRFs may not be smooth ►

→ Consider Gabor dictionary representation

It is an ill-posed problem ►

 $N \sim 10^2 \ll M \sim 10^4$

Maximum likelihood estimate:

Θ

$$\min_{\Theta} \qquad \frac{1}{2} \|\mathbf{Y} - \mathbf{L}\boldsymbol{\Theta}\widetilde{\mathbf{S}}\|_{\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^{\top - 1}}$$

But

NCRFs may not be smooth

Consider Gabor dictionary representation

It is an ill-posed problem ►

 $N \sim 10^2 \ll M \sim 10^4$

Incorporate prior knowledge:

- Not all regions are processing ► same stimulus
- TRFs exhibits peaks and troughs

Maximum likelihood estimate:

Θ

$$\min_{\Theta} \qquad \frac{1}{2} \|\mathbf{Y} - \mathbf{L}\boldsymbol{\Theta}\widetilde{\mathbf{S}}\|_{\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^{\top - 1}}$$

But

NCRFs may not be smooth

Consider Gabor dictionary representation

It is an ill-posed problem ►

 $N \sim 10^2 \ll M \sim 10^4$

Incorporate prior knowledge:

- Not all regions are processing ► same stimulus
- TRFs exhibits peaks and troughs

Enforce spatial and lag-domain sparsity

Maximum likelihood estimate:

 \min_{Θ}

$$\frac{1}{2} \|\mathbf{Y} - \mathbf{L} \mathbf{\Theta} \widetilde{\mathbf{S}}\|_{\mathbf{\Sigma}_w + \mathbf{L} \mathbf{\Gamma} \mathbf{L}^{\top - 1}}$$

 \rightarrow

But

- NCRFs may not be smooth
- It is an ill-posed problem

 $N\sim 10^2 \ll M\sim 10^4$

Incorporate prior knowledge:

- Not all regions are processing same stimulus
- TRFs exhibits peaks and troughs

Enforce spatial and lag-domain sparsity



Consider Gabor dictionary representation



Maximum likelihood estimate:

 \min_{Θ}

$$\frac{1}{2} \|\mathbf{Y} - \mathbf{L} \mathbf{\Theta} \widetilde{\mathbf{S}}\|_{\mathbf{\Sigma}_w + \mathbf{L} \mathbf{\Gamma} \mathbf{L}^{\top - 1}}$$

 \rightarrow

But

- NCRFs may not be smooth
- It is an ill-posed problem

 $N\sim 10^2 \ll M\sim 10^4$

Incorporate prior knowledge:

- Not all regions are processing same stimulus
- TRFs exhibits peaks and troughs

Enforce spatial and lag-domain sparsity

Rotational Invariance



Consider Gabor dictionary representation



Maximum likelihood estimate:

$$\min_{\boldsymbol{\Theta}} \ \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{(\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^\top)^{-1}}^2 + \eta \mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

But

- NCRFs may not be smooth
- It is an ill-posed problem

 $N\sim 10^2 \ll M\sim 10^4$

Incorporate prior knowledge:

- Not all regions are processing same stimulus
- TRFs exhibits peaks and troughs

Enforce spatial and lag-domain sparsity

Rotational Invariance



Consider Gabor dictionary representation

 \rightarrow Enforce spatial and lag-domain sparsity



Ø

/ Regularization parameter

$$\min_{\boldsymbol{\Theta}} \ \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{(\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^\top)^{-1}}^2 + \eta \mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

But

- NCRFs may not be smooth
- It is an ill-posed problem

Consider Gabor dictionary representation
 Enforce spatial and lag-domain sparsity

 $N\sim 10^2 \ll M\sim 10^4$

Incorporate prior knowledge:

- Not all regions are processing same stimulus
- TRFs exhibits peaks and troughs

Enforce spatial and lag-domain sparsity

Rotational Invariance



Maximum likelihood estimate:

$$\min_{\boldsymbol{\Theta}} \ \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{(\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^\top)^{-1}}^2 + \eta \mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

But

- NCRFs may not be smooth ——
- It is an ill-posed problem ——
- Nuisance covariance, Γ is unknown

Consider Gabor dictionary representation
 Enforce spatial and lag-domain sparsity

Maximum likelihood estimate:

$$\min_{\boldsymbol{\Theta}} \ \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{(\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^\top)^{-1}}^2 + \eta \mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

But

NCRFs may not be smooth — Consider Gabor dictionary representation
 It is an ill-posed problem — Enforce spatial and lag-domain sparsity
 Nuisance covariance, Γ is unknown — Estimate Γ jointly with Θ

Estimate Γ jointly with Θ :

$$\min_{\boldsymbol{\Theta},\boldsymbol{\Gamma}} \qquad \frac{T}{2} \log \left| \boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top} \right| + \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{(\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top})^{-1}}^2 + \eta \mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

Maximum likelihood estimate:

$$\min_{\boldsymbol{\Theta},\boldsymbol{\Gamma}} \qquad \frac{T}{2} \log \left| \boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^\top \right| + \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{(\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^\top)^{-1}}^2 + \eta \mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

But

NCRFs may not be smooth — Consider Gabor dictionary representation
 It is an ill-posed problem — Enforce spatial and lag-domain sparsity
 Nuisance covariance, Γ is unknown — Estimate Γ jointly with Θ

Estimate Γ jointly with Θ :

$$\min_{\boldsymbol{\Theta},\boldsymbol{\Gamma}} \qquad \frac{T}{2} \log \left| \boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top} \right| + \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{(\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top})^{-1}}^2 + \eta \mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

Maximum likelihood estimate:

$$\min_{\boldsymbol{\Theta},\boldsymbol{\Gamma}} \qquad \frac{T}{2} \log \left| \boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top} \right| + \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{(\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top})^{-1}}^2 + \eta \mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

But

- NCRFs may not be smooth —
- It is an ill-posed problem ——
- Nuisance covariance, Γ is unknown –
- Hard to solve the non-convex problem
- Consider Gabor dictionary representation
 Enforce spatial and lag-domain sparsity
 Estimate Γ jointly with Θ

Maximum likelihood estimate:

$$\min_{\boldsymbol{\Theta},\boldsymbol{\Gamma}} \qquad \frac{T}{2} \log \left| \boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top} \right| + \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{(\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top})^{-1}}^2 + \eta \mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

But

- NCRFs may not be smooth ——
- It is an ill-posed problem ————
- Nuisance covariance, Γ is unknown –
- Hard to solve the non-convex problem

Idea: coordinate descent approach

Maximum likelihood estimate:

$$\min_{\boldsymbol{\Theta},\boldsymbol{\Gamma}} \qquad \frac{T}{2} \log \left| \boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top} \right| + \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{(\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top})^{-1}}^2 + \eta \mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

But

- NCRFs may not be smooth ——
- It is an ill-posed problem ————
- Nuisance covariance, Γ is unknown
- Hard to solve the non-convex problem

Idea: coordinate descent approach

Initialize $\Theta^{(0)} = \mathbf{0}$

 $\begin{array}{r} & \longrightarrow \\ & \bigcirc \\ & \bigcirc \\ & \frown \\ & \bullet \\$

Maximum likelihood estimate:

$$\min_{\boldsymbol{\Theta},\boldsymbol{\Gamma}} \quad \frac{T}{2} \log \left| \boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top} \right| + \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{(\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top})^{-1}}^2 + \eta \mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

But

- NCRFs may not be smooth —
- It is an ill-posed problem —
- Nuisance covariance, Γ is unknown –
- Hard to solve the non-convex problem

Idea: coordinate descent approach

Initialize $\mathbf{\Theta}^{(0)} = \mathbf{0}$

$$\mathbf{C}_{v}^{(r)} = \frac{1}{T} (\mathbf{Y} - \mathbf{L} \boldsymbol{\Theta}^{(r)} \widetilde{\mathbf{S}}) (\mathbf{Y} - \mathbf{L} \boldsymbol{\Theta}^{(r)} \widetilde{\mathbf{S}})^{\top}$$

Update Γ

$$\begin{split} \mathbf{\Gamma}^{(r+1)} &= \underset{\mathbf{\Gamma}}{\operatorname{argmin}} \quad \operatorname{tr} \left(\mathbf{\Sigma}_{v}^{-1} \mathbf{C}_{v}^{(r)} \right) + \log |\mathbf{\Sigma}_{v}| \\ & \text{s.t.} \ \mathbf{\Sigma}_{v} = \mathbf{\Sigma}_{w} + \mathbf{L} \mathbf{\Gamma} \mathbf{L}^{\top} \end{split}$$

Champagne Wipf et. al, 2010 Consider Gabor dictionary representation
 Enforce spatial and lag-domain sparsity
 Estimate Γ jointly with Θ

Maximum likelihood estimate:

$$\min_{\boldsymbol{\Theta},\boldsymbol{\Gamma}} \qquad \frac{T}{2} \log \left| \boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top} \right| + \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{(\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top})^{-1}}^2 + \eta \mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

But

- ►
- Nuisance covariance, Γ is unknown ►
- Hard to solve the non-convex problem ►

Idea: coordinate descent approach

Initialize
$$oldsymbol{\Theta}^{(0)} = oldsymbol{0}$$

$$\mathbf{C}_{v}^{(r)} = \frac{1}{T} (\mathbf{Y} - \mathbf{L} \boldsymbol{\Theta}^{(r)} \widetilde{\mathbf{S}}) (\mathbf{Y} - \mathbf{L} \boldsymbol{\Theta}^{(r)} \widetilde{\mathbf{S}})^{\top}$$

Update Γ ►

$$\begin{split} \mathbf{\Gamma}^{(r+1)} &= \mathop{\mathrm{argmin}}_{\mathbf{\Gamma}} & \operatorname{tr}\left(\mathbf{\Sigma}_{v}^{-1}\mathbf{C}_{v}^{(r)}\right) + \log|\mathbf{\Sigma}_{v}| \\ & \text{s.t. } \mathbf{\Sigma}_{v} &= \mathbf{\Sigma}_{w} + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^{\top} \\ \end{split} \\ \begin{array}{l} \mathbf{\Theta}^{(r+1)} &= \mathop{\mathrm{argmin}}_{\mathbf{\Theta}} & \frac{1}{2}\|\mathbf{L}\mathbf{\Theta}\widetilde{\mathbf{S}} - \mathbf{Y}\|_{\mathbf{\Sigma}_{v}^{(r+1)-1}}^{2} \\ & +\eta\mathcal{P}_{2,1,1}(\mathbf{\Theta}) \\ \end{array} \\ \begin{array}{l} \mathbf{C} \text{hampagne} \\ & \text{Wipf et. al, 2010} \\ \end{array} \\ \begin{array}{l} \mathbf{\Sigma}_{v}^{(r+1)} &= \mathbf{\Sigma}_{w} + \mathbf{L}\mathbf{\Gamma}^{(r+1)}\mathbf{L}^{\top} \\ \end{array} \\ \begin{array}{l} \mathbf{\Theta}^{(r+1)} &= \mathop{\mathrm{argmin}}_{\mathbf{\Theta}} & \frac{1}{2}\|\mathbf{L}\mathbf{\Theta}\widetilde{\mathbf{S}} - \mathbf{Y}\|_{\mathbf{\Sigma}_{v}^{(r+1)-1}}^{2} \\ & +\eta\mathcal{P}_{2,1,1}(\mathbf{\Theta}) \\ \end{array} \\ \begin{array}{l} \mathbf{E}^{(r+1)} &= \mathbf{E}^{(r+1)}\mathbf{E}^{(r+1)}\mathbf{E}^{(r+1)}\mathbf{E}^{(r+1)}\mathbf{E}^{(r+1)}\mathbf{E}^{(r+1)} \\ \end{array} \\ \begin{array}{l} \mathbf{E}^{(r+1)} &= \mathbf{E}^{(r+1)}\mathbf{E}^{(r+1$$

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

NCRFs may not be smooth \rightarrow Consider Gabor dictionary representation It is an ill-posed problem — > Enforce spatial and lag-domain sparsity \rightarrow Estimate Γ jointly with Θ

Maximum likelihood estimate:

$$\min_{\boldsymbol{\Theta},\boldsymbol{\Gamma}} \quad \frac{T}{2} \log \left| \boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top} \right| + \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{(\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top})^{-1}}^2 + \eta \mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

But

- NCRFs may not be smooth Consider Gabor dictionary representation
- It is an ill-posed problem > Enforce spatial and lag-domain sparsity
- Nuisance covariance, Γ is unknown —
- Hard to solve the non-convex problem

Idea: coordinate descent approach

Initialize
$$\Theta^{(0)} = \mathbf{0}$$

 $\mathbf{C}_{v}^{(r)} = \frac{1}{T} (\mathbf{Y} - \mathbf{L}\Theta^{(r)}\widetilde{\mathbf{S}}) (\mathbf{Y} - \mathbf{L}\Theta^{(r)}\widetilde{\mathbf{S}})^{\top}$ Update $\mathbf{\Gamma}$
 $\mathbf{\Gamma}^{(r+1)} = \operatorname{argmin}_{\mathbf{\Gamma}} \operatorname{tr} \left(\mathbf{\Sigma}_{v}^{-1} \mathbf{C}_{v}^{(r)} \right) + \log |\mathbf{\Sigma}_{v}|$
 $\operatorname{s.t.} \mathbf{\Sigma}_{v} = \mathbf{\Sigma}_{w} + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^{\top}$
 $\mathbf{Champagne}_{\text{Wipf et. al, 2010}}$
 $\mathbf{\Sigma}_{w}^{(r+1)} = \mathbf{\Sigma}_{w} + \mathbf{L}\mathbf{\Gamma}^{(r+1)}\mathbf{L}^{\top}$
 $\mathbf{Update } \Theta$
 $\mathbf{P}^{(r+1)} = \operatorname{argmin}_{\Theta} \frac{1}{2} \|\mathbf{L}\Theta\widetilde{\mathbf{S}} - \mathbf{Y}\|_{\mathbf{\Sigma}_{v}^{(r+1)-1}}^{2}$
 $\mathbf{P}^{(r+1)} = \mathbf{P}^{(r+1)}\mathbf{L}^{\top}$

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

 \rightarrow Estimate Γ jointly with Θ

Maximum likelihood estimate:

$$\min_{\boldsymbol{\Theta},\boldsymbol{\Gamma}} \qquad \frac{T}{2} \log \left| \boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top} \right| + \frac{1}{2} \| \mathbf{Y} - \mathbf{L} \boldsymbol{\Theta} \widetilde{\mathbf{S}} \|_{(\boldsymbol{\Sigma}_w + \mathbf{L} \boldsymbol{\Gamma} \mathbf{L}^{\top})^{-1}}^2 + \eta \mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

But

NCRFs may not be smooth \rightarrow Consider Gabor dictionary representation

- ►
- Nuisance covariance, Γ is unknown \longrightarrow Estimate Γ jointly with Θ
- Hard to solve the non-convex problem \longrightarrow Champ Lasso algorithm ►

It is an ill-posed problem — > Enforce spatial and lag-domain sparsity



MEG Data:

Brodbeck & Simon, 2018



- 17 young-adult participants.
- ► Two 60-second segments from 'The Legend of Sleepy Hollow' by W. Irving.
- 3 repetitions for each segment.

MEG Data:

Brodbeck & Simon, 2018



- 17 young-adult participants.
- ► Two 60-second segments from 'The Legend of Sleepy Hollow' by W. Irving.
- 3 repetitions for each segment.



MEG Data:

Brodbeck & Simon, 2018



- 17 young-adult participants.
- ► Two 60-second segments from 'The Legend of Sleepy Hollow' by W. Irving.
- 3 repetitions for each segment.



- No MRI was available:
- 'fsaverage' morphed in individual head shape

Semantic composition Westerlund et. al., 2015 Brysbaert et. al., 2009 Word frequency

Acoustic envelope Yang et. al., 1992 Ø

Semantic composition Westerlund et. al., 2015 Brysbaert et. al., 2009 Word frequency





P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

Acoustic envelope Yang et. al., 1992

Westerlund et. al., 2015 Semantic composition Brysbaert et. al., 2009

Word frequency





P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

Acoustic envelope Yang et. al., 1992



Ø

Acoustic envelope:



- M50 at ~ 30-35 ms, M100 at ~ 110 ms (reversed polarity, stronger in r.h.)
- Bi-lateral motor activity at ~ 50 ms

Ø

Word frequency:



- Strong auditory component in the l.h. at ~ 150 ms.
- Weak frontal and inferior temporal components in l.h.

Semantic composition:



- Bilateral auditory component at ~ 155 ms, late auditory component at ~ 475 ms.
- Auditory-frontal dynamics at ~ 175-210 ms ($rMT_{sc} \rightarrow rF2_{sc} \rightarrow rF1_{sc} \rightarrow rF2_{sc}$).







- A tool for directly extracting the cortical dynamics that underlie continuous stimuli processing from MEG.
- Novel spatiotemporal prior that not only combats overfitting and spatio-temporal dispersion but also handles lack of MR scans.
- The NCRFs are readily interpretable in a meaningful fashion without any recourse to post-hoc processing.

NCRFs as a powerful source localization tool for continuous stimuli experiments.

Acknowledgements

Advisor:

Behtash Babadi

Collaborators:

- Christian Brodbeck
- Jonathan Z. Simon

Funding:

- National Science Foundation (Award No: 1552946 and 1734892)
- Defense Advanced Research Projects Agency (Award No:N6600118240224)
- National Institutes of Health (Award No: R01-DC014085

Github Repo Link



Thank you!

Presentation available at



