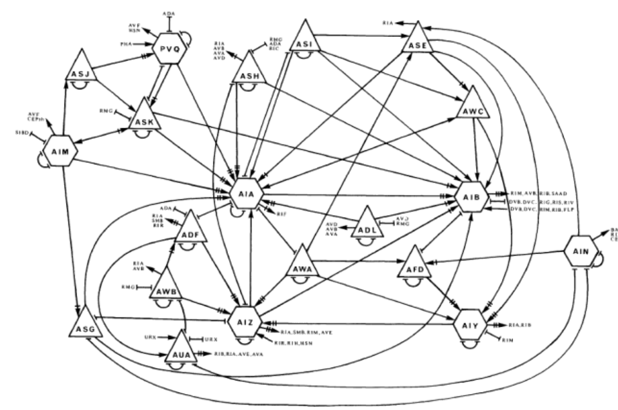
Non-parametric directionality analyses of electrophysiological and neuroimaging signals.

David Halliday

Department of Electronic Engineering
University of York



Neural circuitry in *C elegans* (Newman, 2010)

Overview of Talk

UNIVERSITY *of York*

- Motivation: Brain Networks & Network Theory.
- Statistical signal processing:
  Measures of association, Spectra & Coherence.
- Parametric approaches to directionality.
- Non-parametric directionality:
  Unconditional: decomposition of coherence.
  Conditional: decomposition of partial coherence.
- Results:
  Simulated data, Comparison with Granger causality,
  Experimental data.
- Conclusions.

Brain Networks:
Anatomical, Functional & Effective Connectivity

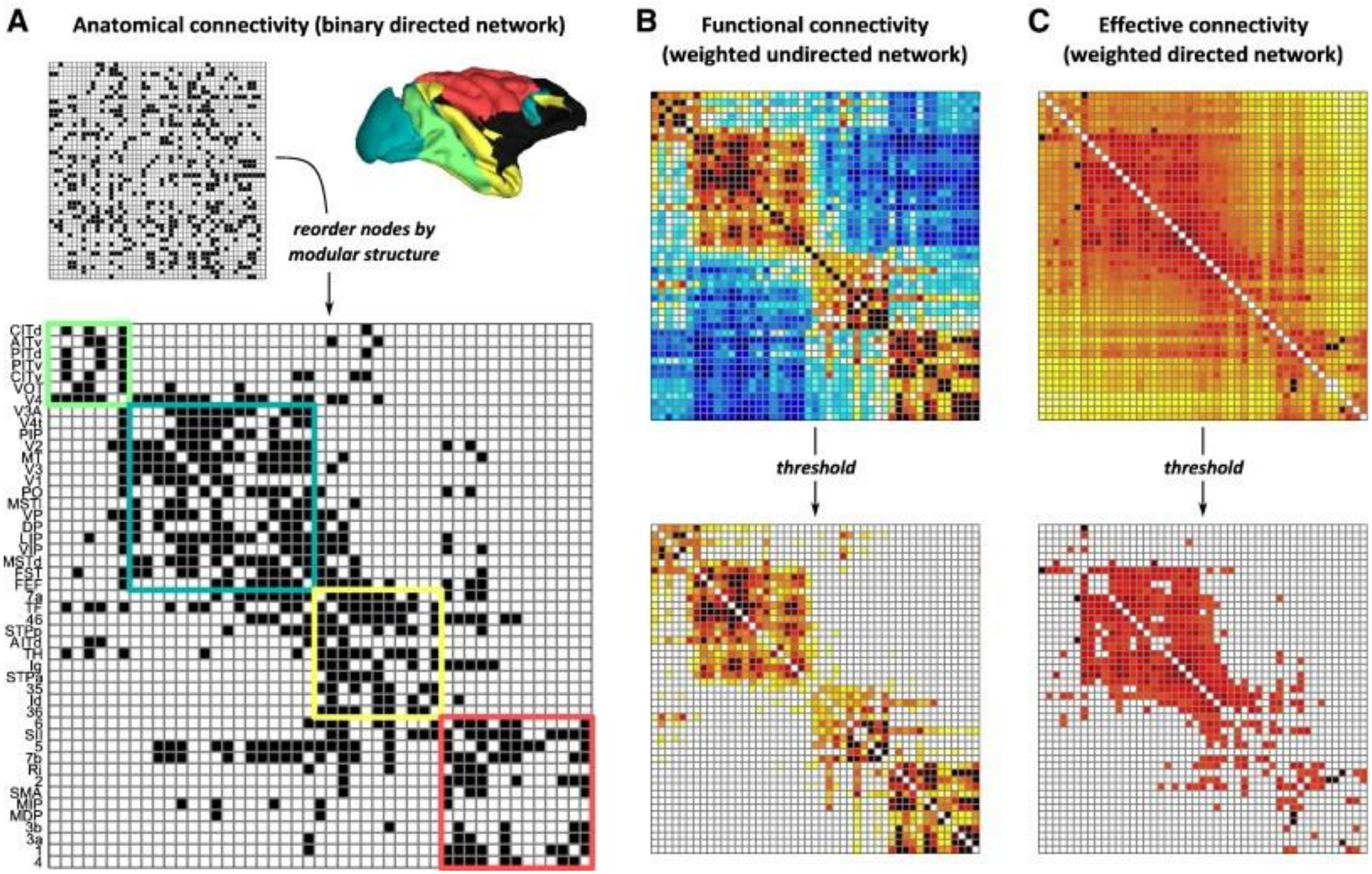


Figure: Rubinov & Sporns (2010). NeuroImage 52: 1059-1069.

Brain Networks

Directed Networks
v.s.
Undirected Networks

Non-parametric directionality: May provide unified approach.

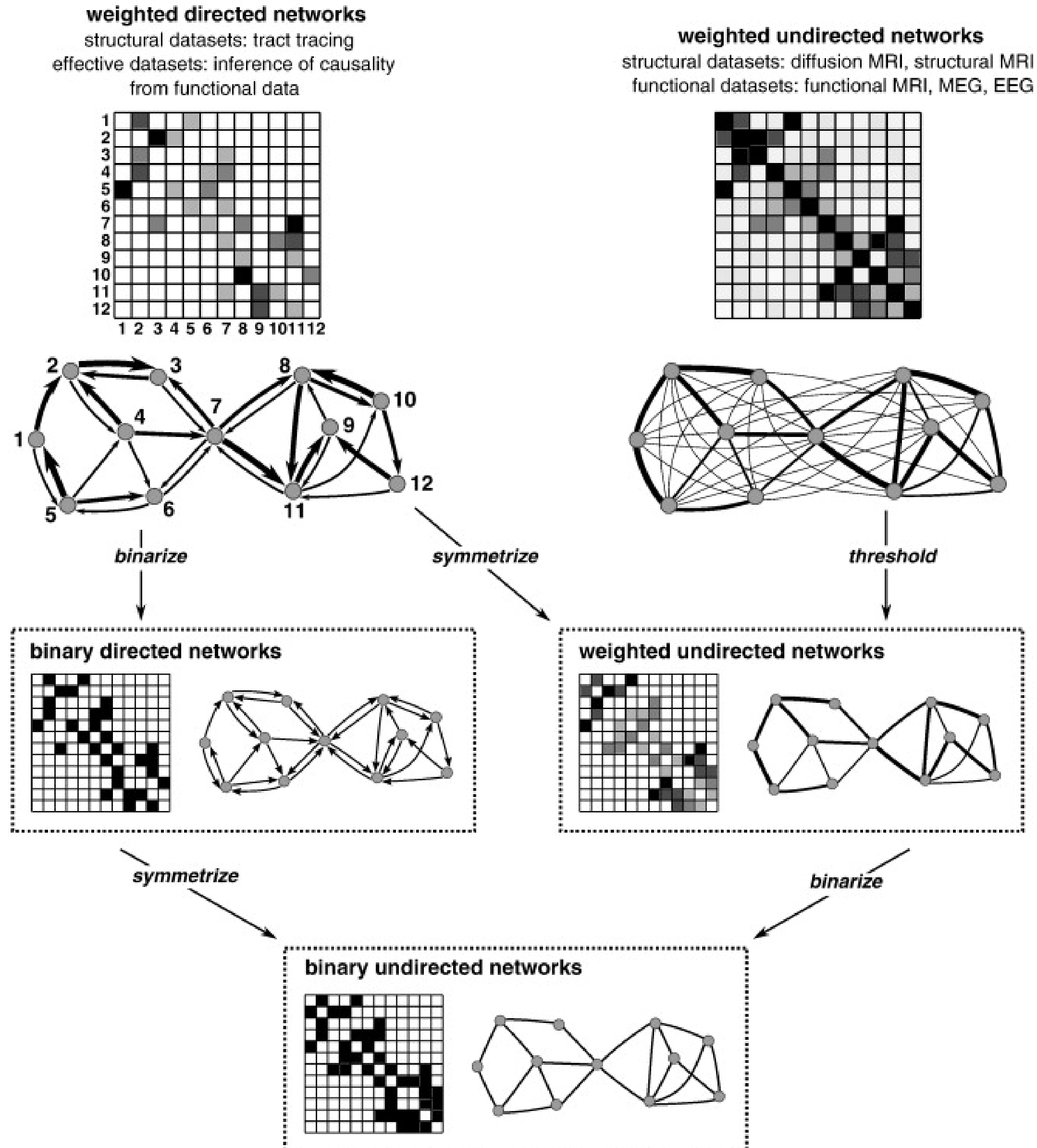
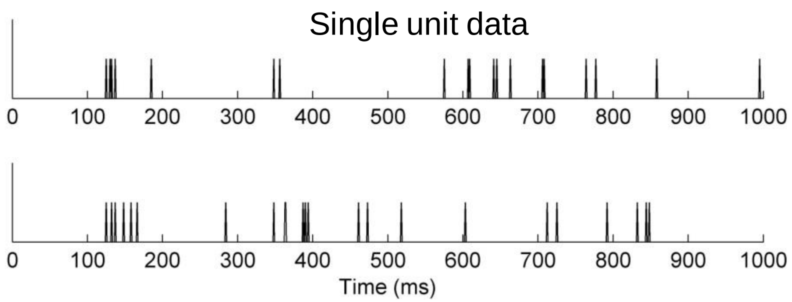
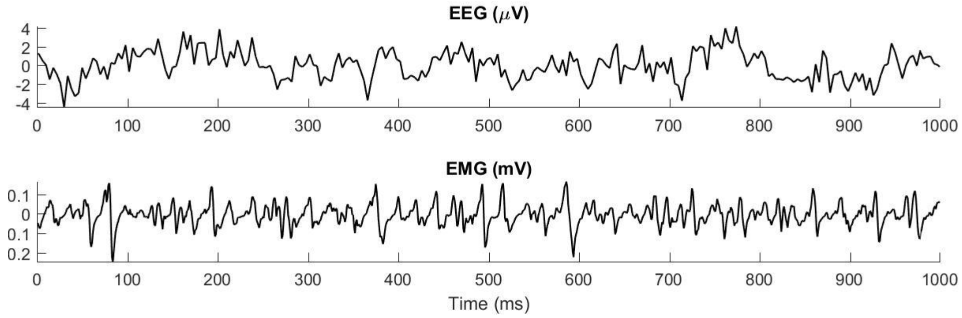


Figure: Rubinov & Sporns (2010). NeuroImage 52: 1059-1069.

Functional connectivity: Time domain.





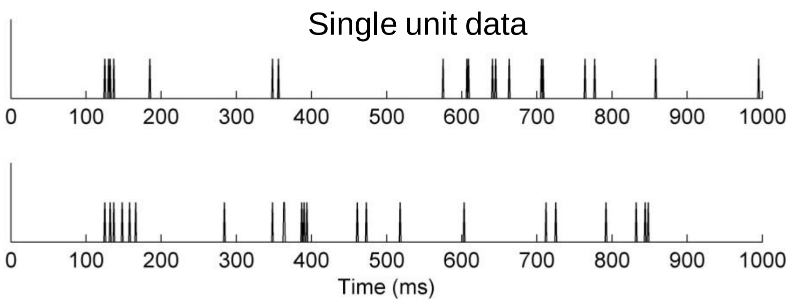
Point process differential increments: $dN_1(t), dN_2(t+\tau)$

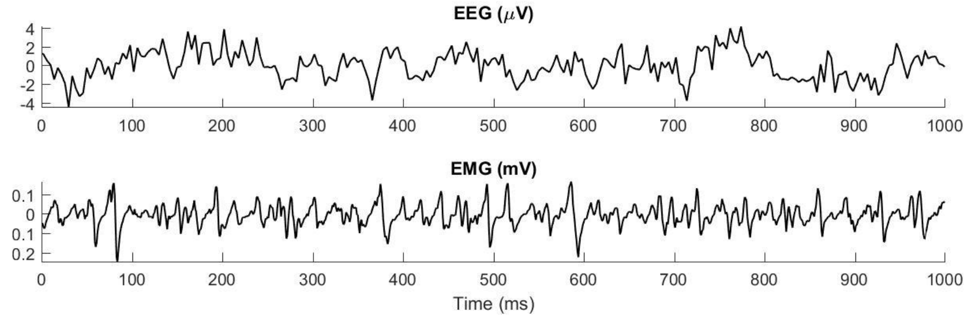
Cross covariance: $\operatorname{cov}\left\{dN_1(t), dN_2(t+\tau)\right\}$ or $\operatorname{cov}\left\{x(t), y(t+\tau)\right\}$

Point process, estimate using: $\#\left\{\left(r_j, s_k\right): \left(\tau - \tfrac{h}{2}\right) < \left(s_k - r_j\right) \leq \left(\tau + \tfrac{h}{2}\right)\right\}$

Times series, estimate using: $\frac{1}{T}\sum_{t=1}^{T-|\tau|} x_t \, y_{t+\tau}$

Functional connectivity: Frequency domain.





Consider correlation of function of $dN_1(t)$, $dN_2(t)$ and $x(t)$, $y(t)$.

Correlation of Fourier transforms.

Finite FT: $d_{N_1}^T(\lambda) = \sum_{r_j \in (0,T]} \exp(-i\lambda r_j)$ or $d_x^T(\lambda) = \sum_{t=0}^{T-1} x_t \exp(-i\lambda t)$

Magnitude squared correlation: $\lim_{T\to\infty} \left| corr\left\{ d_x^T(\lambda), d_y^T(\lambda) \right\} \right|^2$

Coherence: measure of functional connectivity.

Parametric approaches to directionality

Granger Causality:

Model x on past history of y.

Model y on past history of x.

Causality from residual variances.

$$x_t = \sum_{j=1}^{m} a_j \; x_{t-j} + \sum_{j=1}^{m} b_j \; y_{t-j} + \varepsilon_t$$

$$y_t = \sum_{j=1}^{m} c_j \; x_{t-j} + \sum_{j=1}^{m} d_j \; y_{t-j} + \eta_t$$

Granger, 1969 eq (5.1)

Other approaches:

Geweke: Similar approach to Granger,

Uses $\log_e$ of ratio of residual variances to estimate causality.

Application to electrophysiology and brain signals:

Directed coherence, Directed Transfer function, Partial directed coherence, Multivariate Granger causality, Variants…

Parametric approaches to directionality

Granger Causality:
  Model *x* on past history of *y*.
  Model *y* on past history of *x*.
Causality from residual variances.

$$x_t = \sum_{j=1}^{m} a_j \, x_{t-j} + \sum_{j=1}^{m} b_j \, y_{t-j} + \varepsilon_t$$

$$y_t = \sum_{j=1}^{m} c_j \, x_{t-j} + \sum_{j=1}^{m} d_j \, y_{t-j} + \eta_t$$

Granger, 1969 eq (5.1)

What are the issues with these approaches?

- Validity of VAR models.
- Selection of model order.

A non-parametric approach to directionality

Starting point is linear regression:

Measures dependence of *y* on past, present and future values of *x*.

$$y_t = \sum_{k=-\infty}^{\infty} a_k \, x_{t-k} + e_t$$

Contains information regarding directionality.

Measure association using residual variance: $R_{yx}^2 = \dfrac{\left(\sigma_y^2 - \sigma_e^2\right)}{\sigma_y^2}$

Can get this by integrating coherence: $R_{yx}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| R_{yx}(\lambda) \right|^2 d\lambda$

BUT – coherence is a ratio:
(Not easy to decompose)

$$\left| R_{yx}(\lambda) \right|^2 = \frac{\left| f_{yx}(\lambda) \right|^2}{f_{xx}(\lambda) \, f_{yy}(\lambda)}$$

What if the spectra were completely white?

Coherence would depend only on cross-spectrum

$$\left|R_{yx}(\lambda)\right|^2 = \frac{\left|f_{yx}(\lambda)\right|^2}{f_{xx}(\lambda)\, f_{yy}(\lambda)}$$

$$\downarrow$$

Cross spectrum between whitened processed: $\left|f_{yx}^{w}(\lambda)\right|$

$$\left|R_{yx}(\lambda)\right|^2 = \left|f_{yx}^{w}(\lambda)\right|^2$$

This can be achieved using MMSE Whitening approach:

1846 IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 49, NO. 7, JULY 2003

MMSE Whitening and Subspace Whitening

Yonina C. Eldar, *Member, IEEE,* and Alan V. Oppenheim, *Fellow, IEEE*

***Abstract*—This correspondence develops a linear whitening transformation that minimizes the mean-squared error (MSE) between the original and whitened data, i.e.,one that results in a white output that is as close as possible to the input, in an MSE sense. When the covariance matrix of the data is not invertible, the whitening transformation is designed to optimally whiten the data on a subspace in which it is contained. The optimal whitening transformation is developed both for the case of finite-length data vectors and infinite-length signals.**

***Index Terms*—Mean-squared error (MSE) whitening, subspace whitening, whitening.**

I. INTRODUCTION

Data whitening arises in a variety of contexts in which it is useful to either decorrelate a data sequence prior to subsequent processing, or to control the spectral shape after processing. Examples in which data whitening has been used to advantage include enhancing direction of arrival algorithms by prewhitening [1], [2], and improving probability of correct detection in multisignature systems [3], [4] and multiuser wireless communication systems [5] by prewhitening.

Whitening of a random sequence parallels closely the concept of orthogonalization of a set of vectors. Specifically, orthogonalizing a set in quantum mechanics [6], and later applied to the design of optimal frames [9], [10].

Paralleling the concept of LS orthogonalization, in this paper we develop an optimal linear whitening transformation. Our criterion for optimality is motivated by the fact that, in general, whitening a data vector or signal introduces distortion to the values of the data relative to the unwhitened data. In certain applications of whitening, it may be desirable to whiten the data while minimizing this distortion. Therefore, in this correspondence we propose choosing a linear whitening transformation that minimizes the mean-squared error (MSE) between the original and whitened data, i.e., that results in a white output that is as close as possible to the input in an MSE sense. We refer to such a whitening transformation as a minimum MSE (MMSE) whitening transformation. Extensions of this concept to other forms of covariance shaping are considered in [4], [11].

Applications of MMSE whitening and subspace whitening to matched-filter detection, multiuser detection, and LS estimation are considered in [3], [5], [12]–[14]. The essential idea in the detection applications is to improve the detection performance by optimally whitening the output of conventional receivers prior to detection using an MMSE or subspace MMSE whitening transformation. As we show by simulations in [3] and analytically in [5], in many cases this approach can, in fact, lead to improved detection performance.

To illustrate the use of MMSE whitening and subspace whitening in more detail, we consider here an application of these ideas to LS estimation. This application is developed and explored in more detail in

Reducing coherence to the magnitude squared cross spectrum using MMSE whitening

Optimal whitening filter for process *x*: $w_{xx}(\lambda) = f_{xx}(\lambda)^{-\frac{1}{2}}$

Spectrum of derived/whitened processes now 1 at all frequencies: $f_{xx}^{w}(\lambda) = 1, \quad f_{yy}^{w}(\lambda) = 1$

Coherence direct from cross-spectrum: $\left|R_{yx}(\lambda)\right|^2 = \left|f_{yx}^{w}(\lambda)\right|^2$

Correlation from cross spectrum: $R_{yx}^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|f_{yx}^{w}(\lambda)\right|^2 d\lambda$

Construction of directional metrics

Overall correlation measure:

$$R_{yx}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| f_{yx}^{w}(\lambda) \right|^2 d\lambda$$

Define time domain correlation:

$$\rho_{yx}(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{yx}^{w}(\lambda) e^{i\lambda\tau} \, d\lambda$$

Overall correlation is now:

$$R_{yx}^2 = \int_{-\infty}^{\infty} \left| \rho_{yx}(\tau) \right|^2 d\tau$$

Summative decomposition by direction, scalar coefficients:

$$R_{yx}^2 = \int_{\tau<0} \left| \rho_{yx}(\tau) \right|^2 d\tau + \rho_{yx}(0) + \int_{\tau>0} \left| \rho_{yx}(\tau) \right|^2 d\tau$$

$$R_{yx}^2 = R_{yx;-}^2 + R_{yx;0}^2 + R_{yx;+}^2$$

Reverse   Zero-Lag   Forward

Decomposition of coherence by direction

Use $\rho_{yx}(\tau)$ to decompose directional effects by frequency:

$$f'_{yx;-}(\lambda) = \int_{\tau<0} \rho_{yx}(\tau) e^{-i\lambda\tau} \, d\tau$$

$$f'_{yx;-}(\lambda) = \rho_{yx}(0)$$

$$f'_{yx;+}(\lambda) = \int_{\tau>0} \rho_{yx}(\tau) e^{-i\lambda\tau} \, d\tau$$

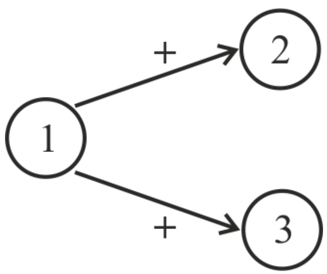
Can generate summative decomposition of coherence by direction:

$$\left|R_{yx}(\lambda)\right|^2 = \left|R'_{yx;-}(\lambda)\right|^2 + \left|R'_{yx;0}(\lambda)\right|^2 + \left|R'_{yx;+}(\lambda)\right|^2$$

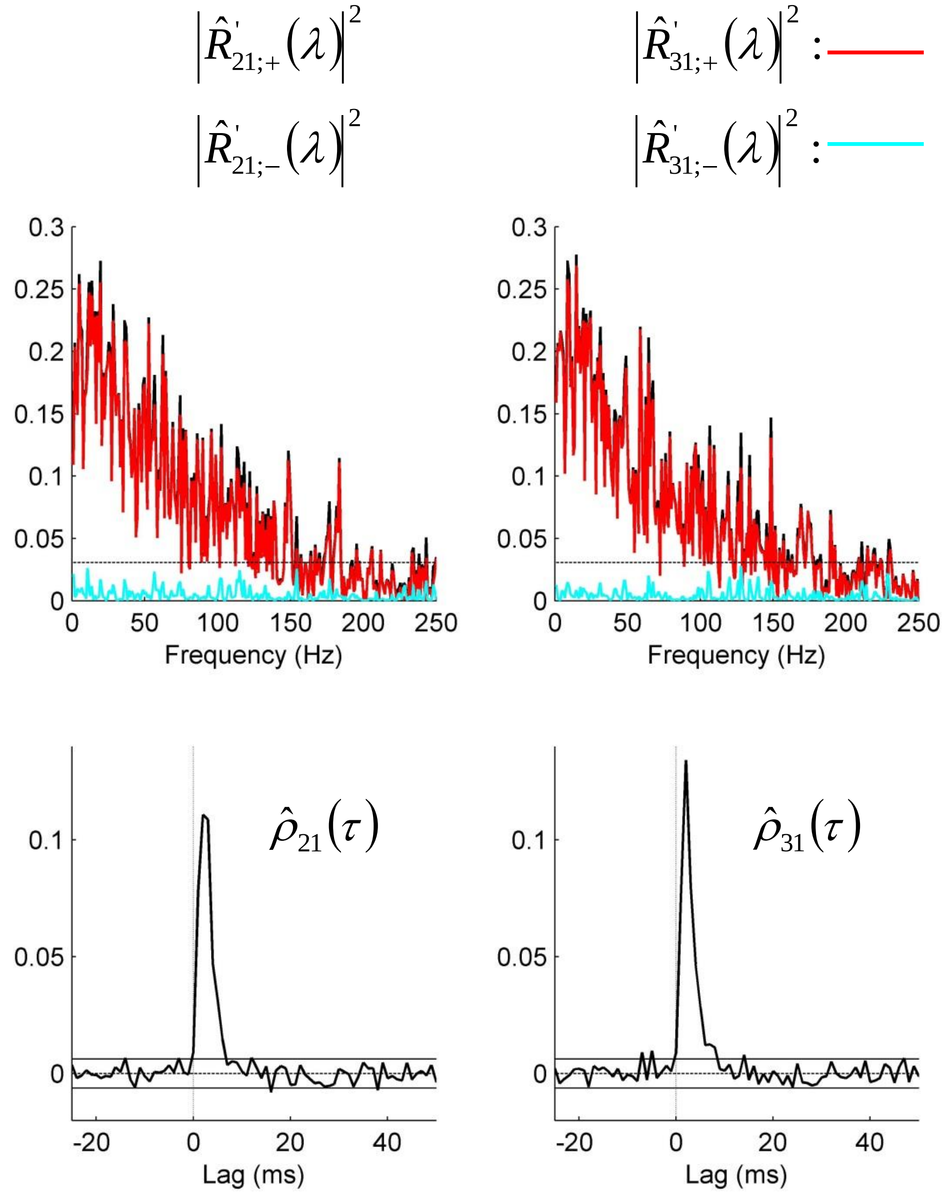
Reverse Zero-Lag Forward

Results: Simulation 1

Three neurone network, cortical neurone model *in-vivo* conditions.



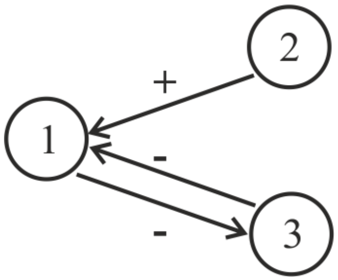
Network configuration (Divergent)



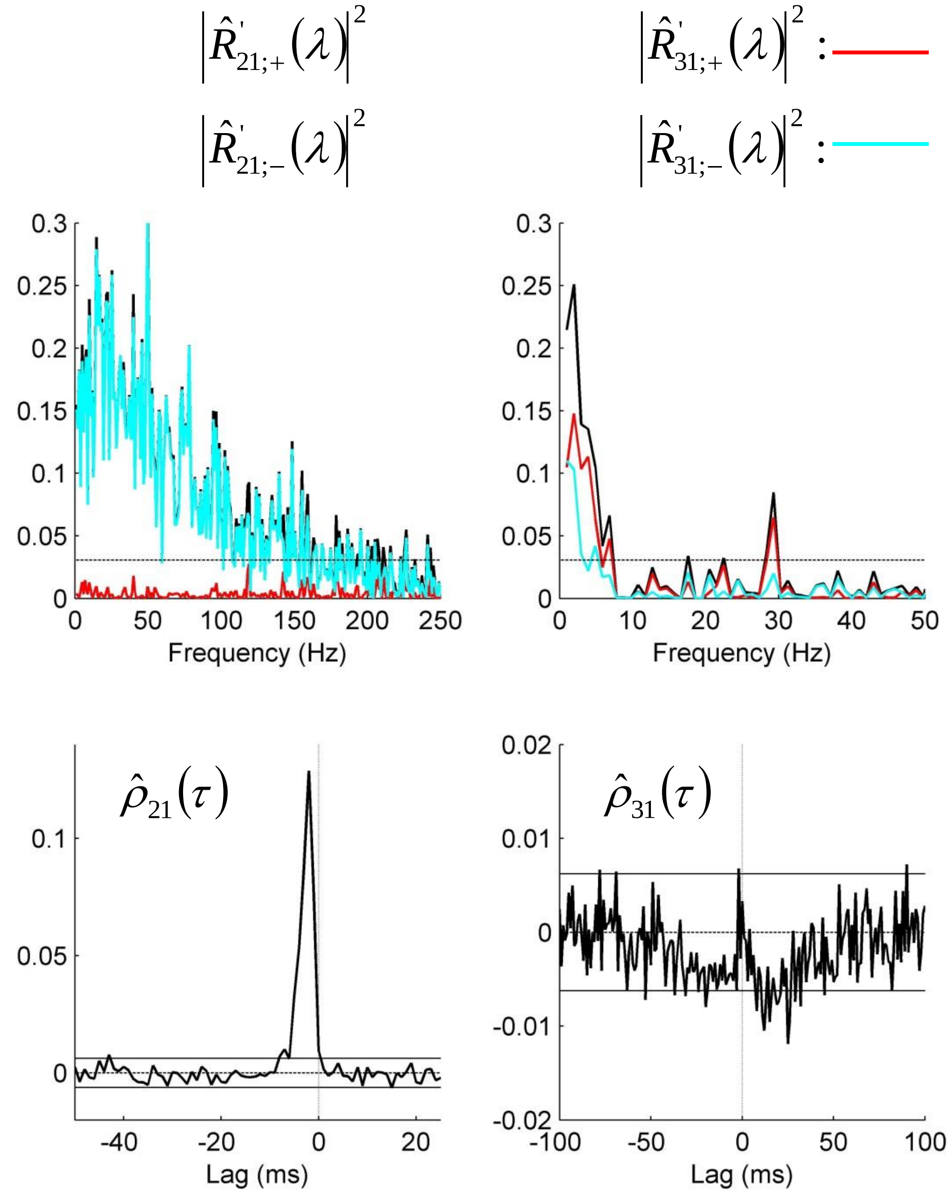
Figures: Halliday (2015). J Integrative Neuroscience 14: 253-277.

Results: Simulation 2

Three neurone network, cortical neurone model *in-vivo* conditions.



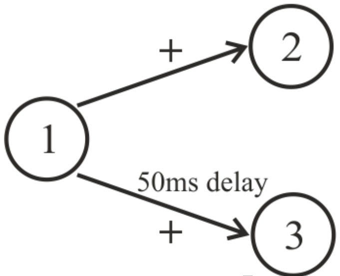
Network configuration (Reciprocal inhibitory)



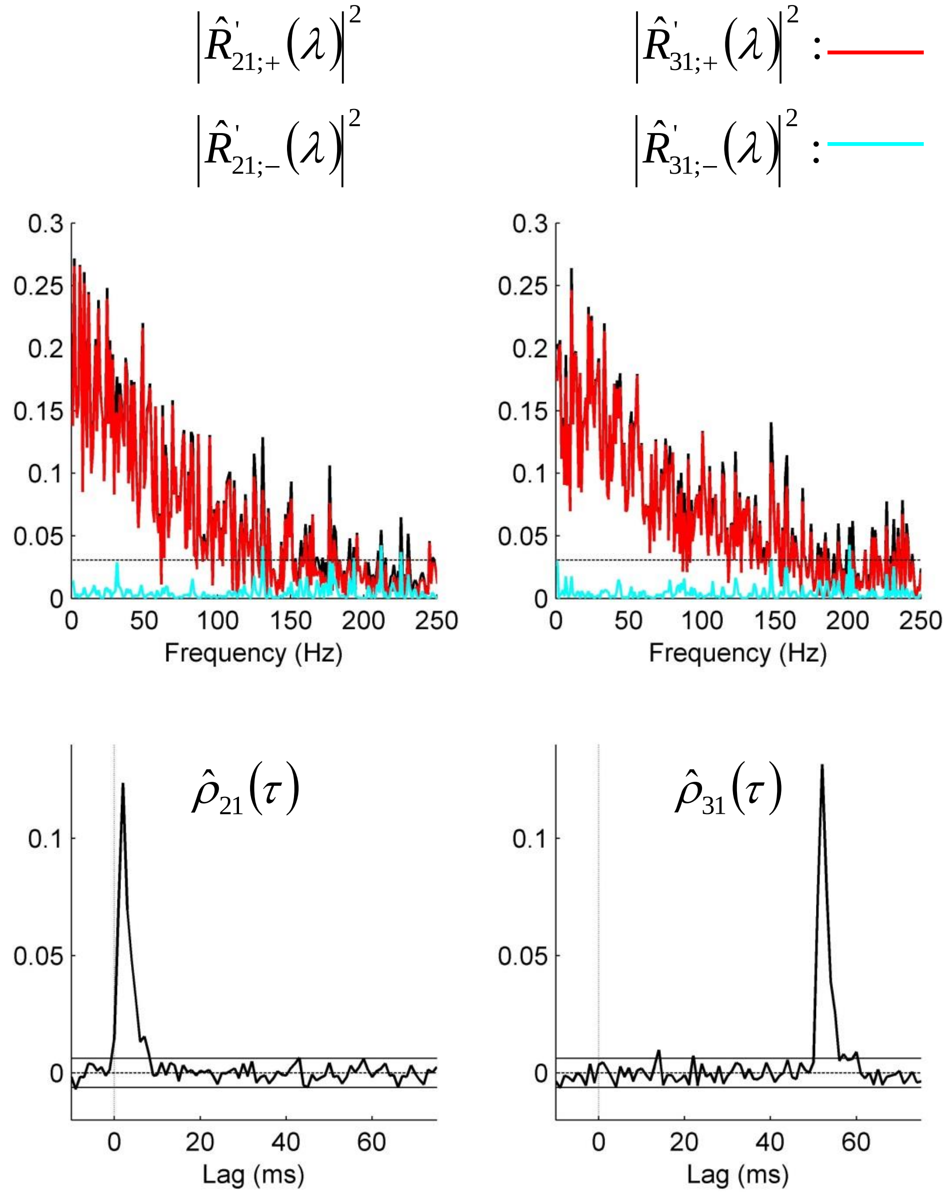
Figures: Halliday (2015). J Integrative Neuroscience 14: 253-277.

Results: Simulation 3

Three neurone network, cortical neurone model *in-vivo* conditions.



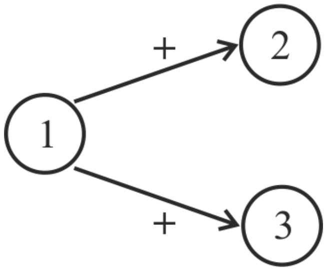
Network configuration (Delay)



Figures: Halliday (2015). J Integrative Neuroscience 14: 253-277.

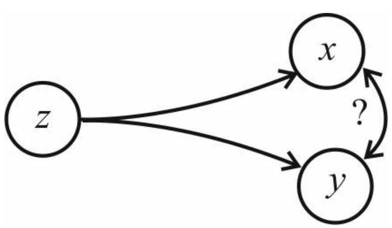
Extension to multivariate non-parametric directionality.

Removal of induced correlation due to common influences:



Correlation induced by neuron 1: $\hat{\rho}_{32}(\tau) \neq 0$

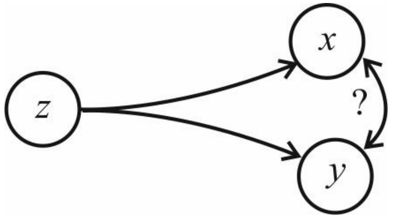
Move to framework based on multivariate partial coherence:



Coherence is significant: $|R_{yx}(\lambda)|^2 > 0$

Partial coherence is zero: $|R_{yx|z}(\lambda)|^2 \sim 0$

Conditional directionality analysis:



Decomposition from partial coherence: $R^2_{yx|z} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| R_{yx|z}(\lambda) \right|^2 d\lambda$

Conditioned Fourier transform of *x*: $d^T_{x|z}(\lambda) = d^T_x(\lambda) - \frac{f_{xz}(\lambda)}{f_{zz}(\lambda)} d^T_z(\lambda)$

MMSE filter for *x*: $w_{xx|z}(\lambda) = f_{xx|z}(\lambda)^{-\frac{1}{2}}$

Pre-whitened conditional FT for *x*: $dw^T_{x|z}(\lambda) = d^T_{x|z}(\lambda)\,\hat{w}_{xx|z}(\lambda)$

Whitened partial spectrum is 1: $f^w_{xx|z}(\lambda) = 1, \quad f^w_{yy|z}(\lambda) = 1$

Correlation from partial cross spectrum: $R^2_{yx|z} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| f^w_{yx|z}(\lambda) \right|^2 d\lambda$

Construction of conditional directional metrics

Overall partial correlation measure: $R^2_{yx|z} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| f^{w}_{yx|z}(\lambda) \right|^2 d\lambda$

Define time domain correlation: $\rho_{yx|z}(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f^{w}_{yx|z}(\lambda)\, e^{i\lambda\tau}\, d\lambda$

Overall partial correlation is: $R^2_{yx|z} = \int_{-\infty}^{\infty} \left| \rho_{yx|z}(\tau) \right|^2 d\tau$

Summative decomposition by direction with scalar coefficients:

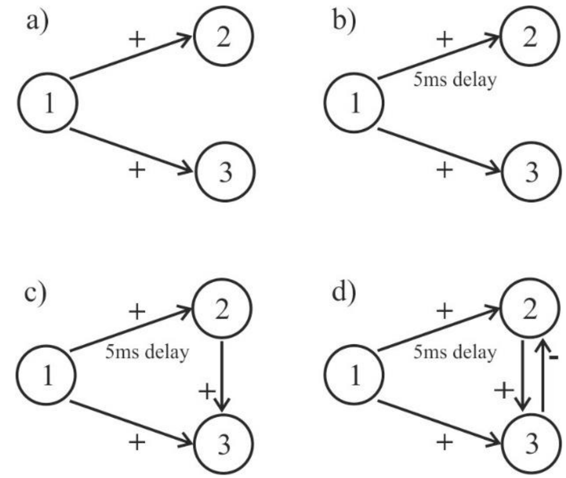
$$R^2_{yx|z} = \int_{\tau<0} \left| \rho_{yx|z}(\tau) \right|^2 d\tau + \rho_{yx|z}(0) + \int_{\tau>0} \left| \rho_{yx|z}(\tau) \right|^2 d\tau$$

$$R^2_{yx|z} = R^2_{yx|z;-} + R^2_{yx|z;0} + R^2_{yx|z;+}$$

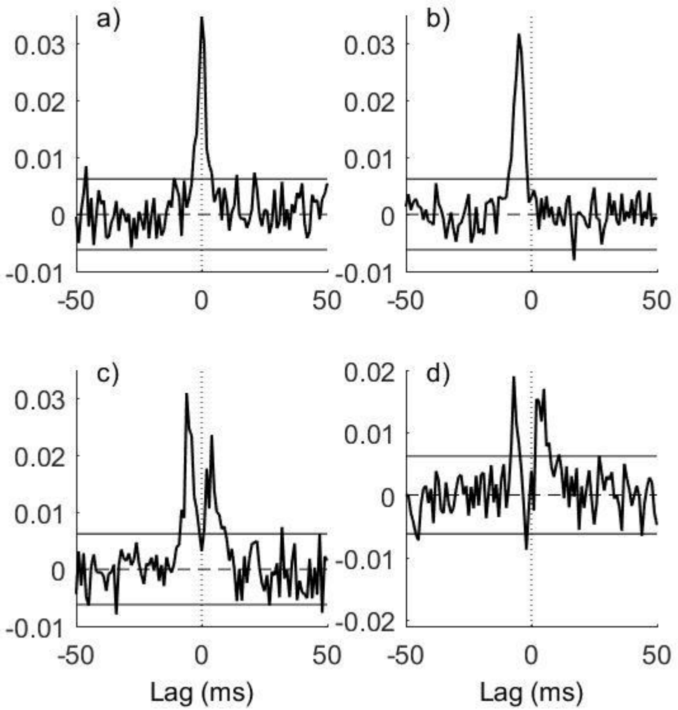
Reverse Zero-Lag Forward

Results: Conditional analysis

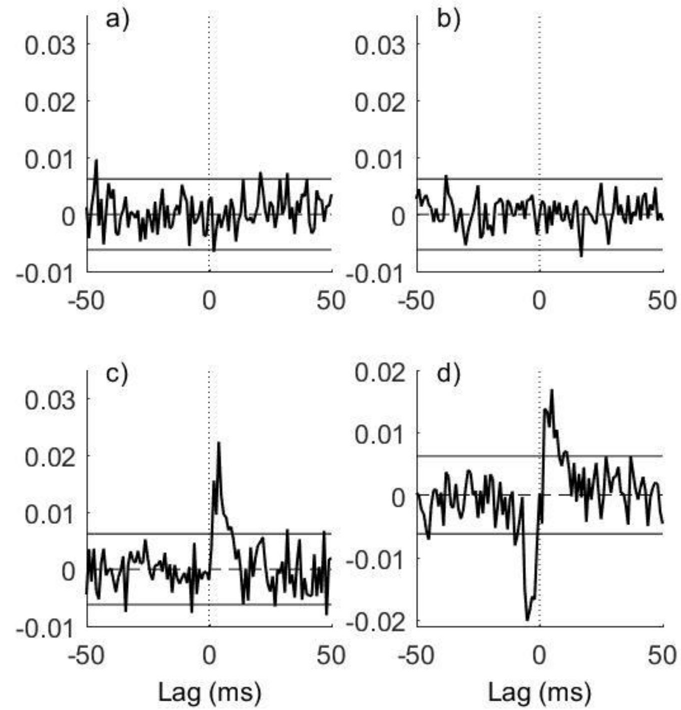
Simulated three cortical neuron network,
Unconditional directionality: $\hat{\rho}_{32}(\tau)$
Conditional directionality: $\hat{\rho}_{32|1}(\tau)$



$\hat{\rho}_{32}(\tau)$



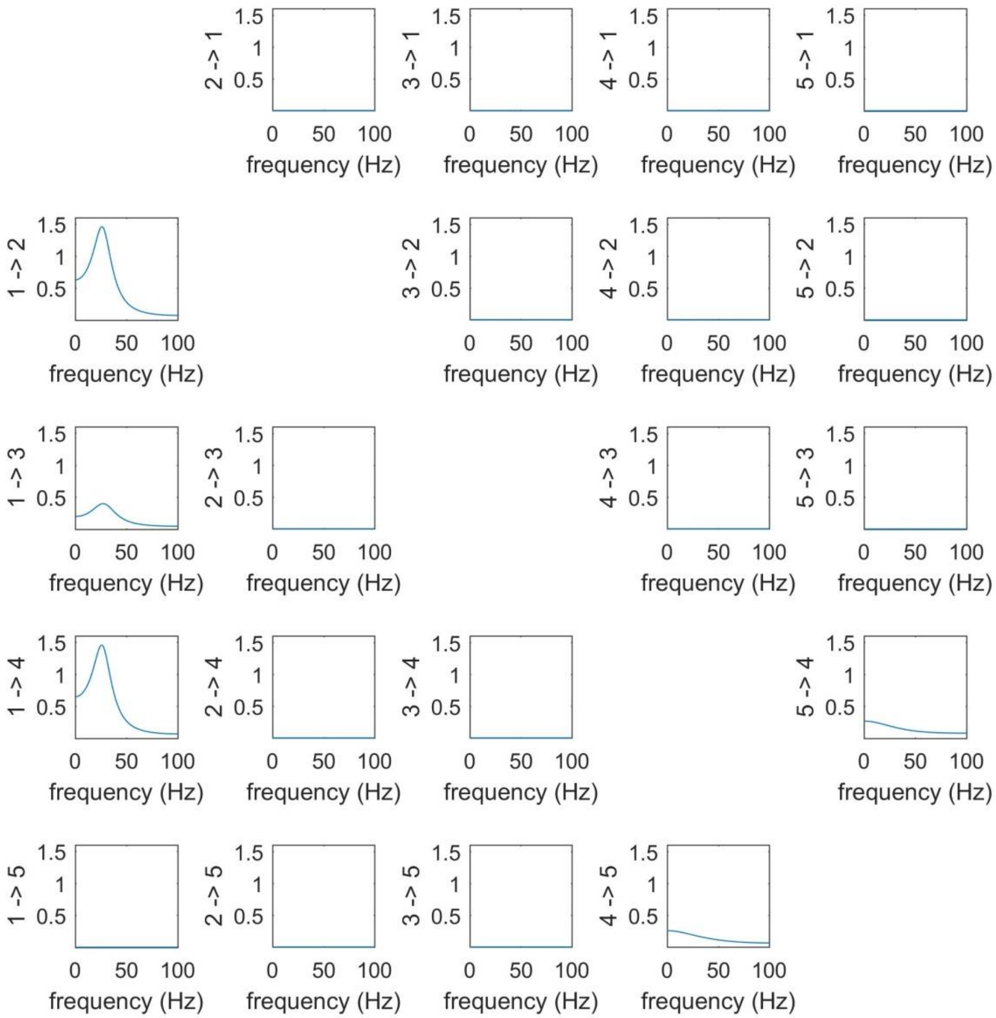
$\hat{\rho}_{32|1}(\tau)$



Figures: Halliday et al. (2016). J Neuroscience Methods 268:87-97.

Comparison with Granger causality: VAR(3) network

$\mathcal{F}_{Y \to X|Z}$

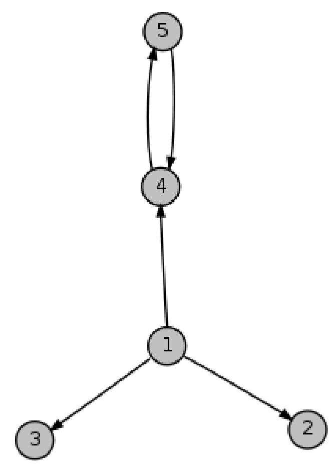


$\mathcal{F}_{X \to Y|Z}$

Granger causality: analysis using MVGC toolbox[1]

Simulated VAR(3) network.

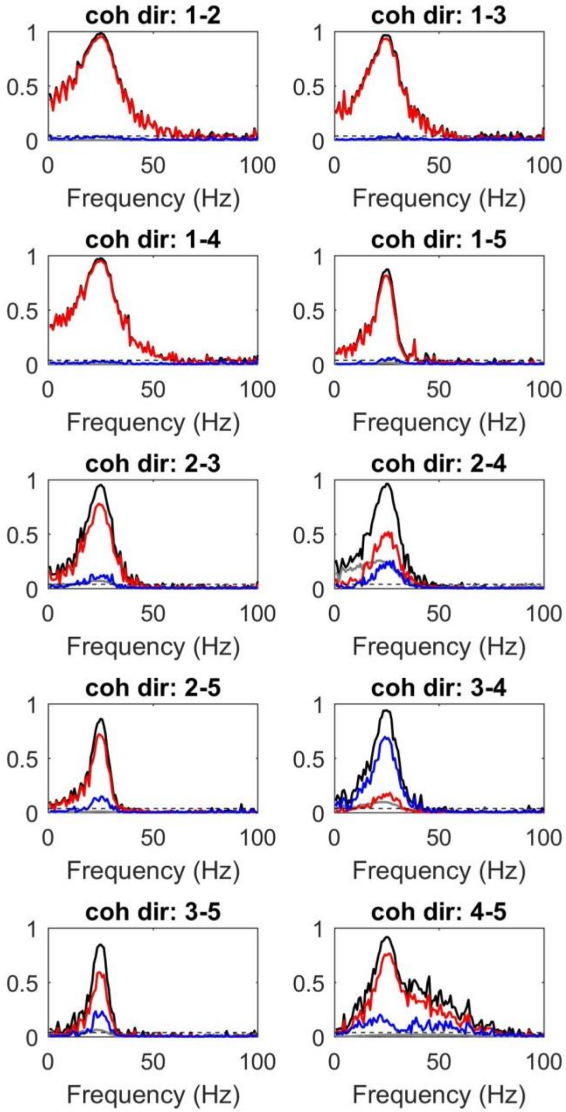
Matrix of frequency domain directional coefficients: $\mathcal{F}_{X \to Y|Z}$, $\mathcal{F}_{Y \to X|Z}$



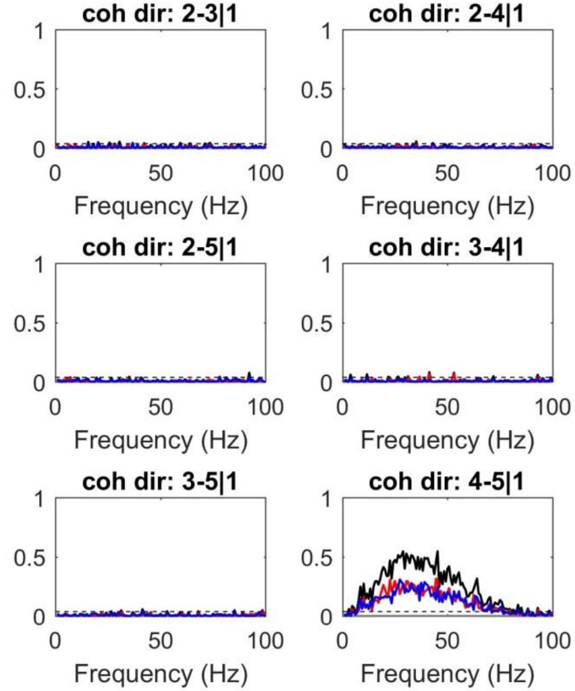
[1]http://www.sussex.ac.uk/sackler/mvgc/

Comparison with Granger causality: VAR(3) network

Unconditional



Conditional

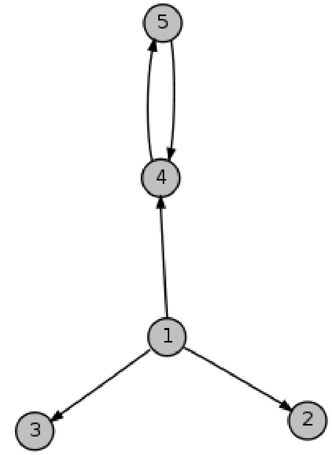


$$\left|\hat{R}_{xy}(\lambda)\right|^2 \quad \left|\hat{R}_{xy|z}(\lambda)\right|^2 : \text{black}$$

$$\left|\hat{R}'_{xy;+}(\lambda)\right|^2 \quad \left|\hat{R}'_{xy|z;+}(\lambda)\right|^2 : \text{red}$$

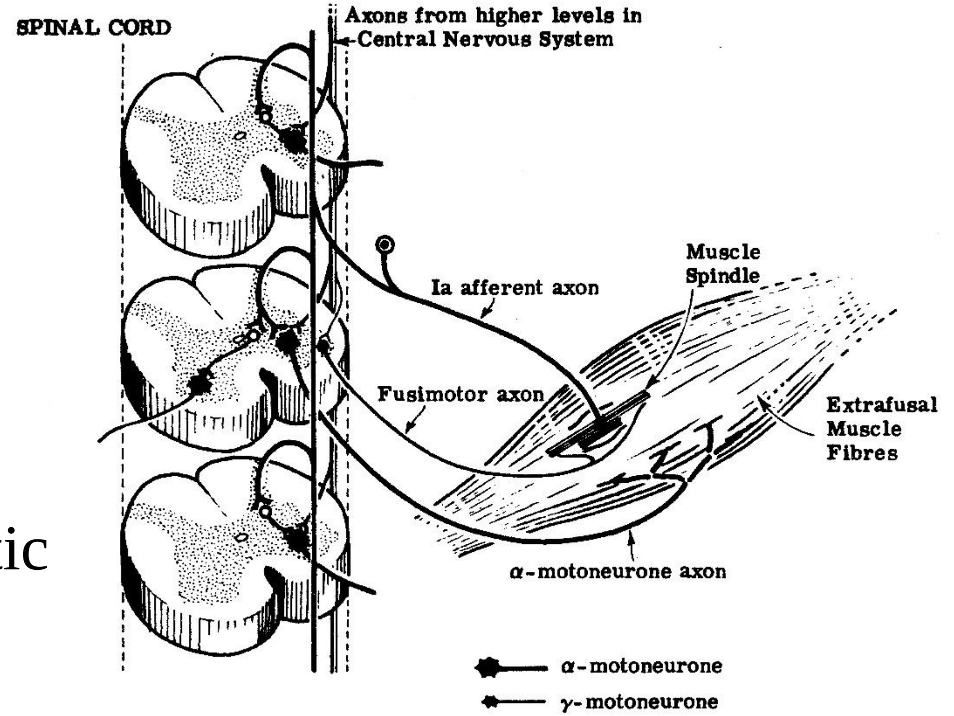
$$\left|\hat{R}'_{xy;-}(\lambda)\right|^2 \quad \left|\hat{R}'_{xy|z;-}(\lambda)\right|^2 : \text{blue}$$

NPD: Unconditional and conditional directionality analysis.



Spike train data – muscle spindle

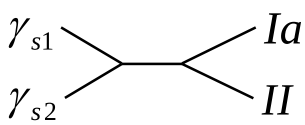
Simultaneous recording of spike timings from muscle spindle afferent receptors during efferent stimulation.



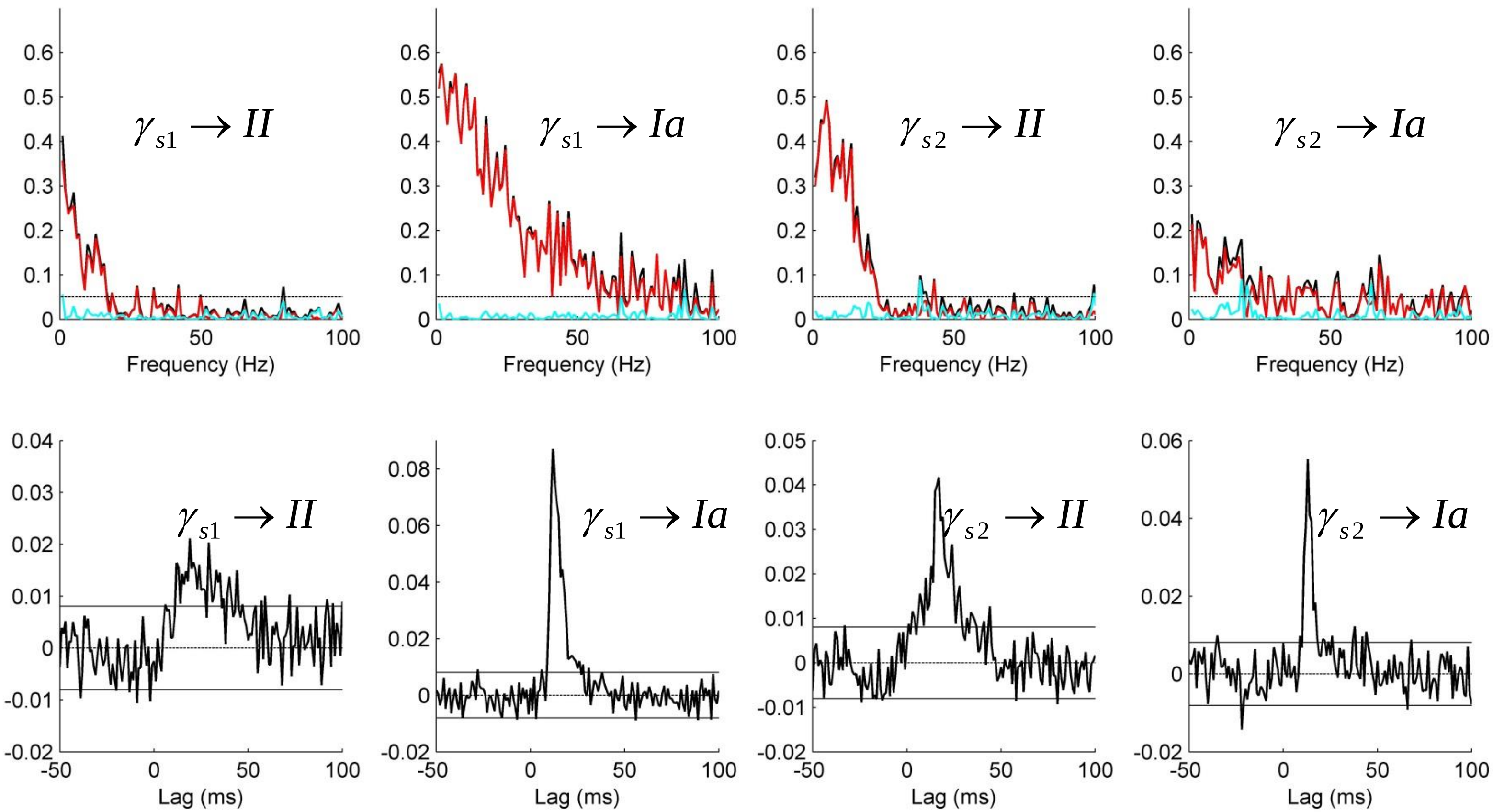
Input spike trains - random stimulation of two efferent static gamma fusimotor axons: γ_{s1}, γ_{s2}

Output spike trains - muscle spindle primary and secondary afferent spike trains: Ia, II.

Spike train data – muscle spindle

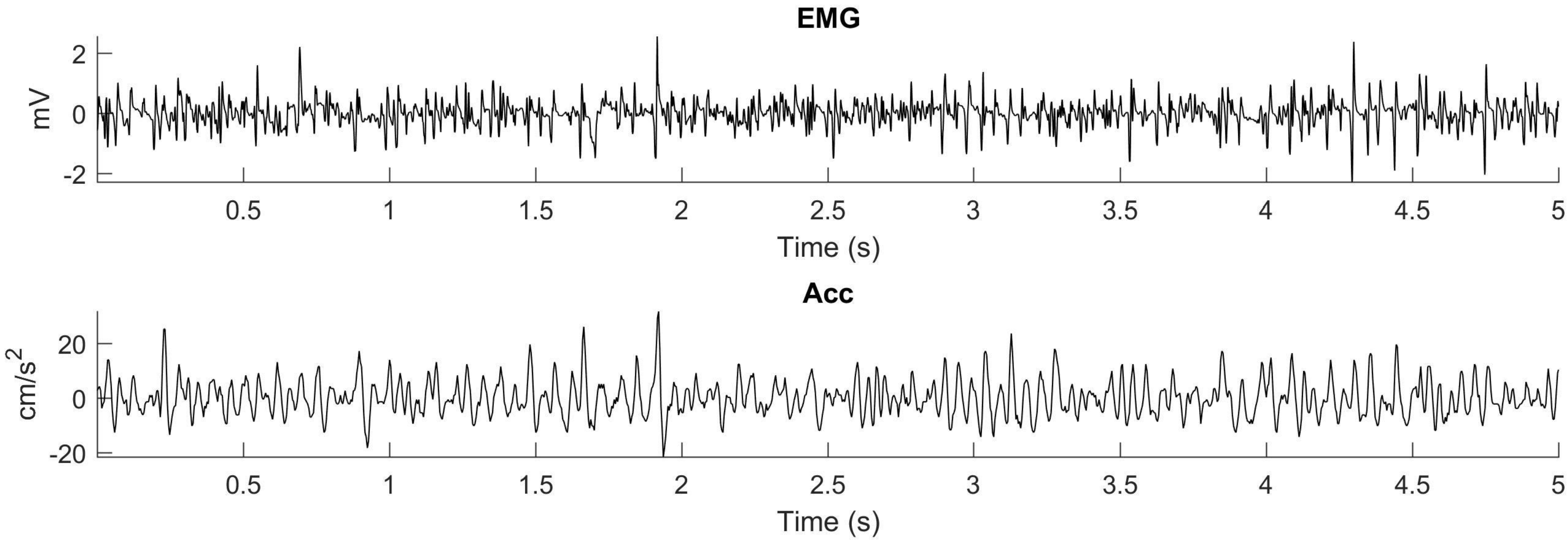


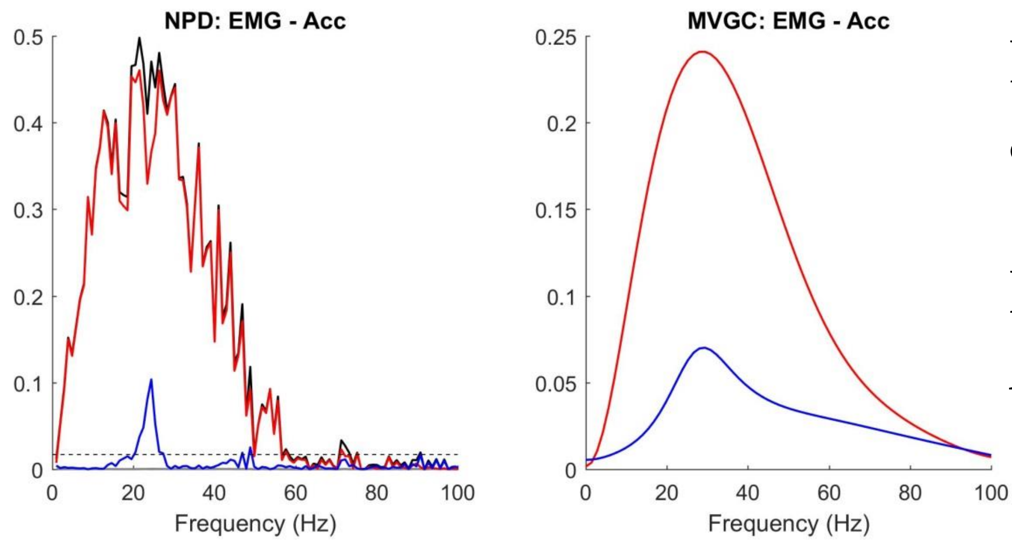
Random stimulation of γ_{s1} and γ_{s2} with fixed muscle length.



Data: Rosenberg et al. (1989). Prog Biophys molec Biol 53: 1-31.

EMG – Acceleration, upper limb postural contraction





Postural tremor: EDC and finger acceleration.

EMG → Acc: (red line)

Acc → EMG: (blue line)

Data: Halliday et al. (1995). Prog Biophys molec Biol 64: 237-278.

Hippocampal connectivity in a model of Kainic acid (KA) induced mesial temporal lobe epilepsy (mTLE) in rat.

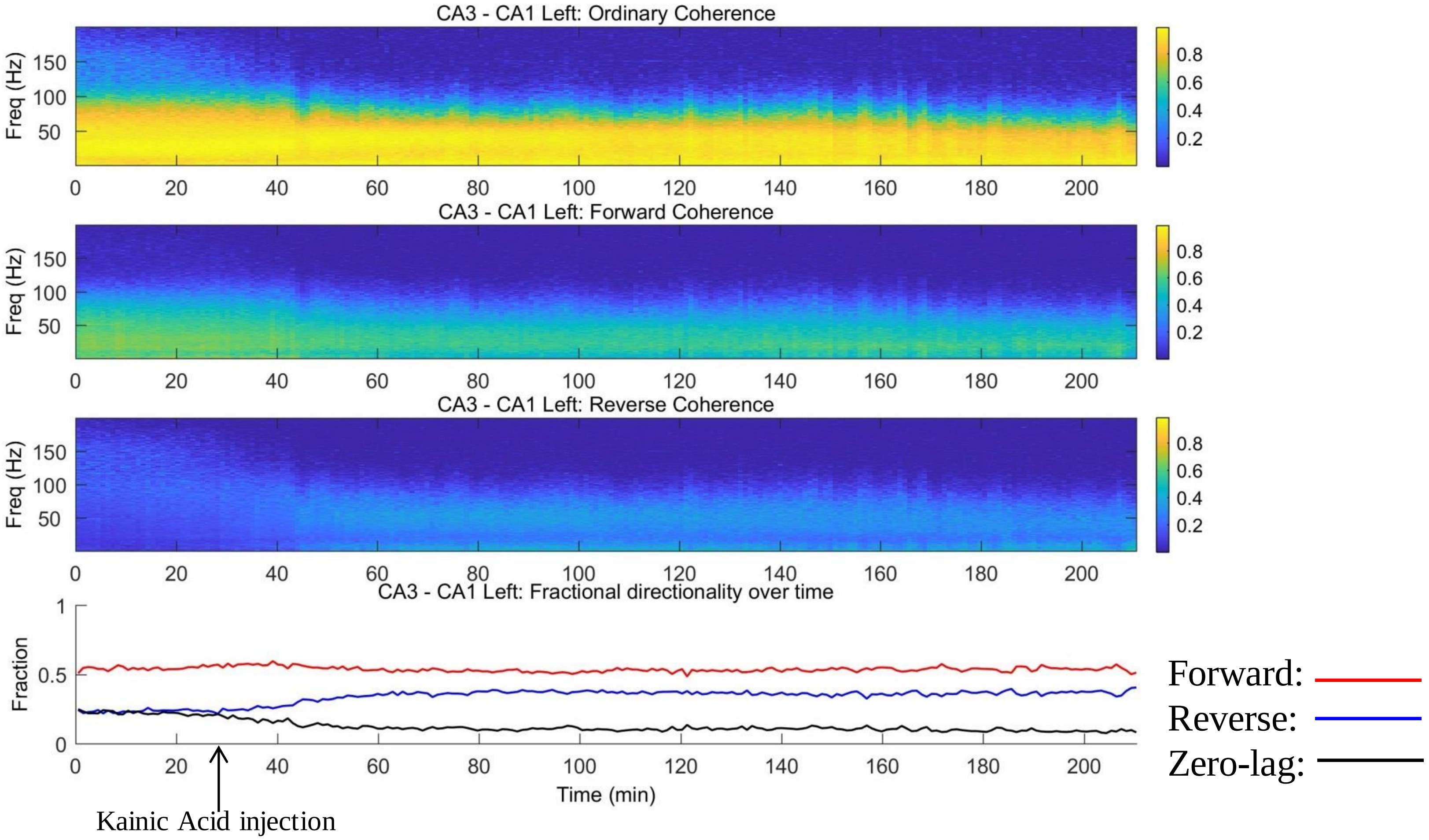


Figure: Halliday et al. (2016). J Neuroscience Methods 268:87-97.

Directional analysis of Basal Ganglia – Motor cortex interactions in Parkinsonian rat.

Recording of ECoG and LFP in Striatum (STR), Globus Pallidus (GPe) and Sub Thalamic Nucleus (STN).

Analysis: Unconditional and Conditional NPD between 4 regions.

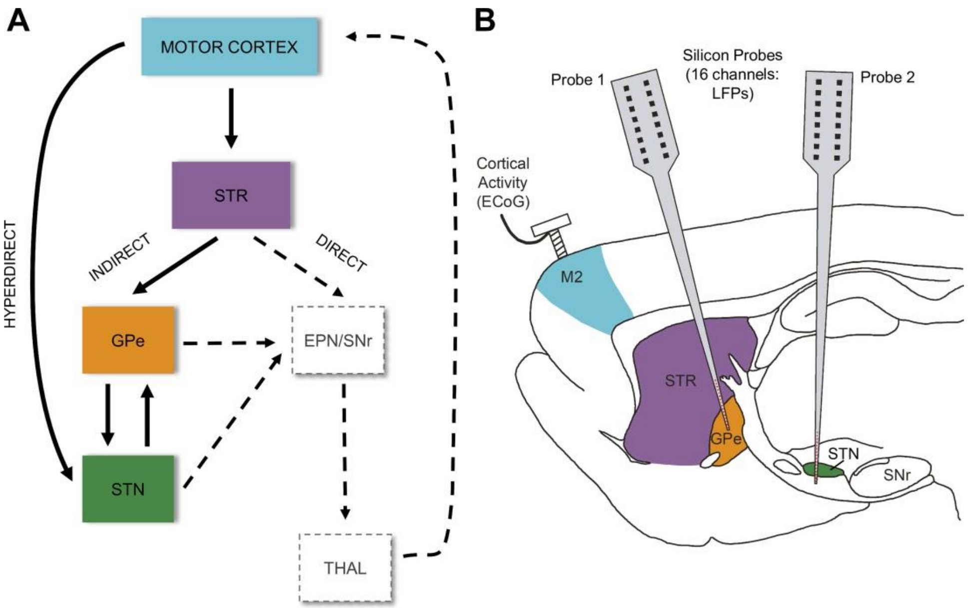
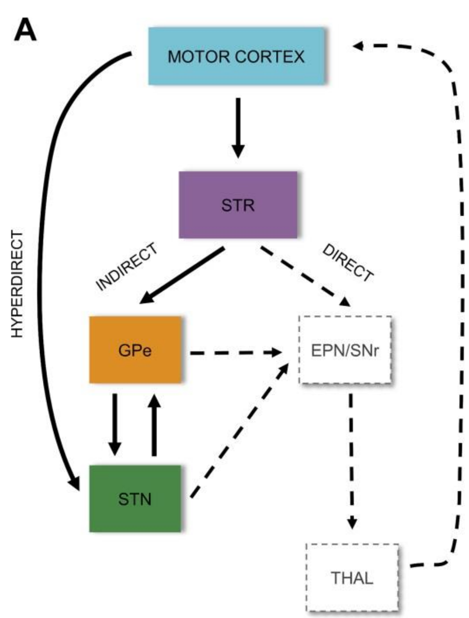


Figure: West et al. (2018). J Neurophysiol 119: 1608–1628.

Directional analysis of Basal Ganglia – Motor cortex interactions in Parkinsonian rat.

Unconditional analysis.

Propagation through indirect pathway.
Activation of long-loop in lesion state.



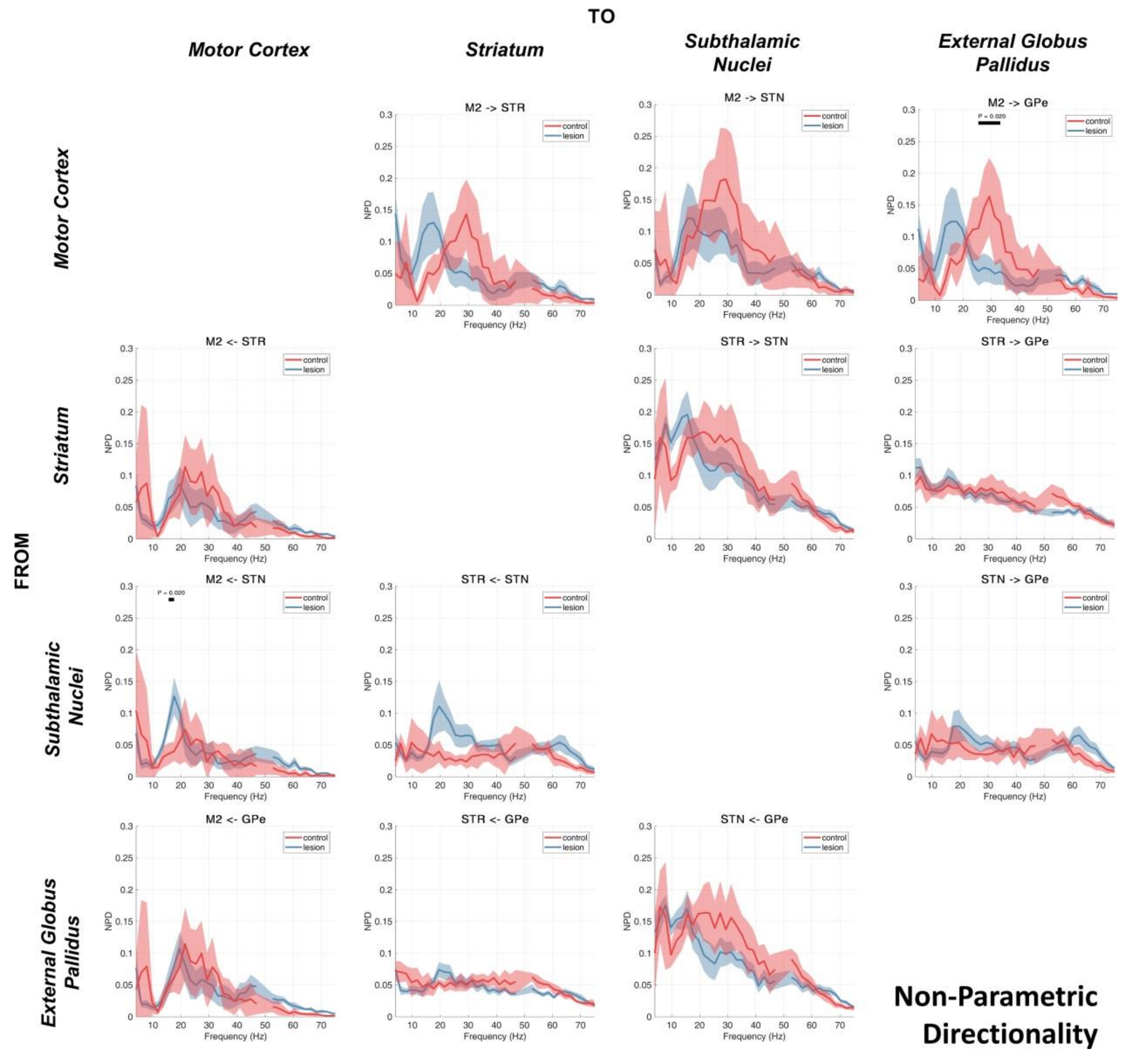
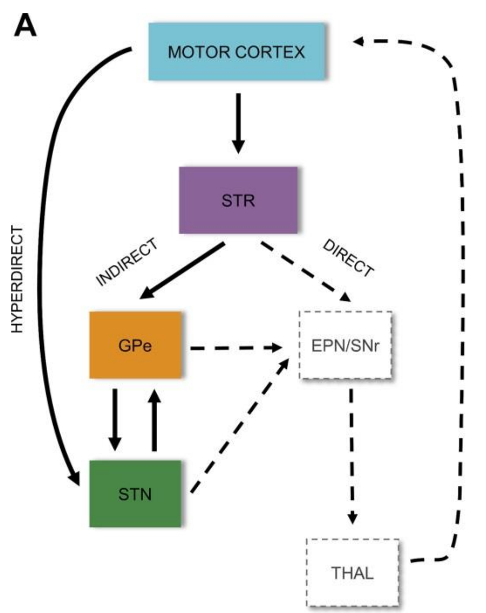


Figure: West et al. (2018). J Neurophysiol 119: 1608–1628.

Directional analysis of Basal Ganglia – Motor cortex interactions in Parkinsonian rat.

Conditional analysis – conditioned on GPe.

Control: Effects strongly reduced.
Lesioned: Hyperdirect pathway active.



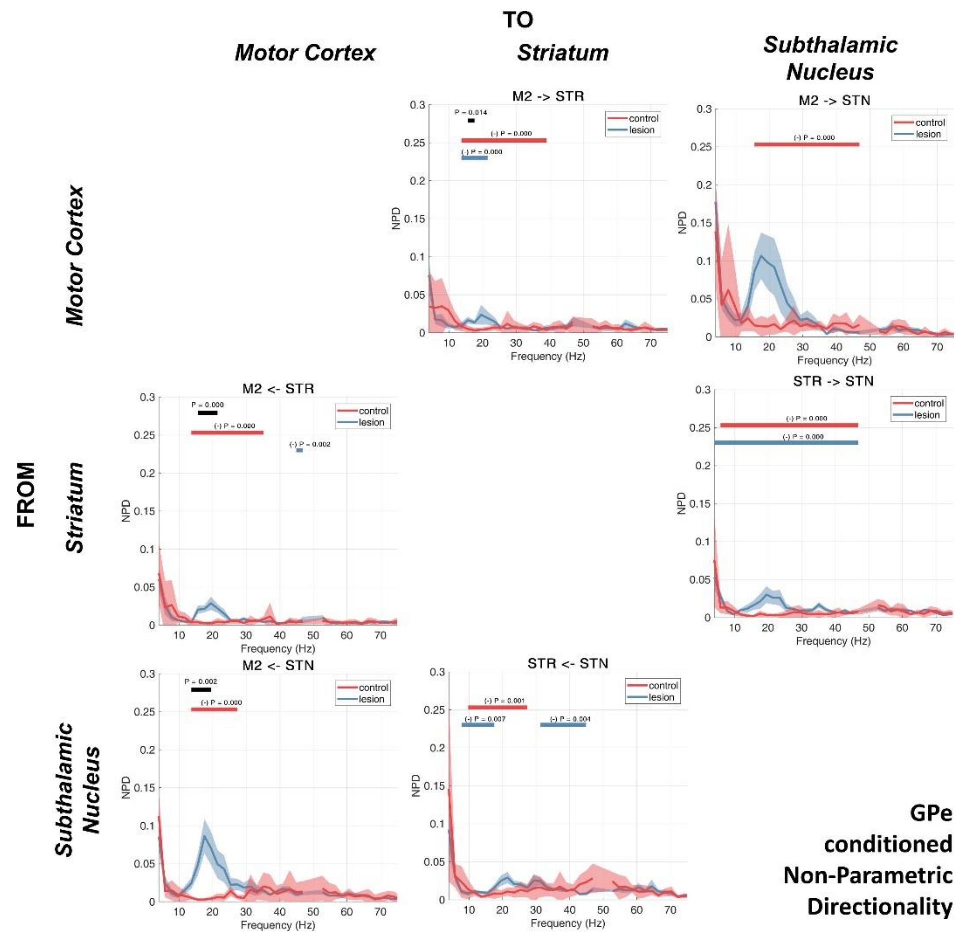


Figure: West et al. (2018). J Neurophysiol 119: 1608–1628.

Conclusions

•Brain connectivity: current approaches primarily parametric.

•Non-parametric approach:

Derived from standard linear regression model.

MMSE: reduces coherency to cross spectrum.

Straightforward addition to spectral analyse.

Summative directional metrics: $R_{yx}^2 = R_{yx;-}^2 + R_{yx;0}^2 + R_{yx;+}^2$

Summative decomposition of coherence:

$$\left|R_{yx}(\lambda)\right|^2 = \left|R'_{yx;-}(\lambda)\right|^2 + \left|R'_{yx;0}(\lambda)\right|^2 + \left|R'_{yx;+}(\lambda)\right|^2$$

•Extended to partial coherence (degree one): $\left|R_{yx|z}(\lambda)\right|^2$

•Future work:

Full multivariate model, Systematic validation against parametric approaches, Wider range of data.

Acknowledgements

Collaborators:

Jay Rosenberg, Margaret Gladden, Bernie Conway (Glasgow),
Rob Mason, Carl Stevenson (Nottingham),
Simon Farmer, Tim West (London).

Funding:

EPSRC, BBSRC,
Wellcome Trust/Centre for Future Health, University of York

Thank - you