Supplementary Material for the Tensor Granger Causality Algorithm with t-Product Algorithm used in the paper "Tensor Analysis and Fusion of Multimodal Brain Images"

This document is prepared for the explanation of the t-Product software presented in http:\\www.cneuro.cu\software\tensor.

I. NOTATION

$\mathbf{X}^T$ and $\mathbf{X}^H$ are the transpose and conjugate (Hermitian) transpose of the matrix $\mathbf{X}$, respectively. $\mathbf{I}_K$ is $K \times K$ dimensional identity matrix. $\breve{\mathcal{X}}$ denotes the Fourier transform of $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ along the third dimension. $\mathrm{diag}(\mathbf{X}, \mathbf{Y})$ is the block diagonal matrix in which the matrices $\mathbf{X}$ and $\mathbf{Y}$ are on its main diagonal blocks.

- **The** *MatVec* **operation** matricizes a three dimensional tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ as defined below:

$$\mathrm{MatVec}(\mathcal{X}) = \begin{pmatrix} \mathcal{X}(:,:,1) \\ \mathcal{X}(:,:,2) \\ \vdots \\ \mathcal{X}(:,:,K) \end{pmatrix}$$

- **The** *Tplz* **operation:**

A three dimensional tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ can be converted into a block toeplitz matrix of dimension $I \cdot K \times J \cdot K$ by rearrangement of its faces as follows:

$$\mathrm{tplz}(\mathcal{X}) = \begin{pmatrix} \mathcal{X}(:,:,1) & \mathcal{X}(:,:,2)^H & \dots & \mathcal{X}(:,:,K)^H \\ \mathcal{X}(:,:,2) & \mathcal{X}(:,:,1) & \dots & \mathcal{X}(:,:,K-1)^H \\ \vdots & \ddots & \ddots & \vdots \\ \mathcal{X}(:,:,K) & \mathcal{X}(:,:,K-1) & \ddots & \mathcal{X}(:,:,1) \end{pmatrix}$$

- **The** *embed* **operation:**

A block Toeplitz matrix $\mathbf{X} \in \mathbb{R}^{I \cdot K \times J \cdot K}$ of a tensor created by $\mathrm{tplz}(\mathcal{X})$ can be embedded in a larger circulant matrix of dimension $I \cdot (2K) \times J \cdot (2K)$ defined as follows (Dietrich & Newsam, 1997):

$$\mathrm{embed}(\mathcal{X}) = \begin{pmatrix} \mathcal{X}(:,:,1) & \mathcal{X}(:,:,2)^H & \dots & \mathcal{X}(:,:,2) \\ \mathcal{X}(:,:,2) & \mathcal{X}(:,:,1) & \dots & \mathcal{X}(:,:,3) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{X}(:,:,K) & \vdots & \ddots & \vdots \\ (\mathcal{X}(:,:,K)^H + \mathcal{X}(:,:,K))/2 & \mathcal{X}(:,:,K) & \ddots & \vdots \\ \mathcal{X}(:,:,K)^H & (\mathcal{X}(:,:,K) + \mathcal{X}(:,:,K)^H)/2 & \ddots & \vdots \\ \vdots & \mathcal{X}(:,:,K)^H & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{X}(:,:,2)^H & \mathcal{X}(:,:,3)^H & \ddots & \mathcal{X}(:,:,1) \end{pmatrix}$$

Recall that circulant matrices can be diagonalized with the normalized discrete Fourier transform (DFT). We will use the same property to diagonalize the block circulant matrices found by $\text{embed}(\boldsymbol{\mathcal{X}})$. Assume $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I \times J \times K}$ and the DFT matrix to be $\mathbf{F} \in \mathbb{R}^{2K \times 2K}$, the diagonalized tensor is formalized as:

$$(\mathbf{F} \otimes \mathbf{I}_I)\,\text{embed}(\boldsymbol{\mathcal{X}})(\mathbf{F}^H \otimes \mathbf{I}_J) = \text{diag}\left(\breve{\boldsymbol{\mathcal{X}}}(:,:,1), \breve{\boldsymbol{\mathcal{X}}}(:,:,2), \ldots, \breve{\boldsymbol{\mathcal{X}}}(:,:,K), (\breve{\boldsymbol{\mathcal{X}}}(:,:,K)^H + \breve{\boldsymbol{\mathcal{X}}}(:,:,K))/2, \breve{\boldsymbol{\mathcal{X}}}(:,:,K)^H, \ldots, \breve{\boldsymbol{\mathcal{X}}}(:,:,2)^H\right)$$

- **The** *unembed* **operation** undoes the *embed* operation:

$$\boldsymbol{\mathcal{X}} = \text{unembed}(\text{embed}(\boldsymbol{\mathcal{X}}))$$

- **Tensor-product (t-product)** is a special type of product modified from (Kilmer, Braman, Hao, & Hoover, 2013).

Let $\boldsymbol{\mathcal{X}}$ be a tridimensional tensor of size $I \times J \times K$ and $\boldsymbol{\mathcal{Y}}$ of size $J \times L \times K$, then the t-product $\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{X}} *_t \boldsymbol{\mathcal{Y}}$ is the $I \times L \times K$ tensor found by:

$$\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{X}} *_t \boldsymbol{\mathcal{Y}} = \text{unembed}(\text{embed}(\boldsymbol{\mathcal{X}}) \cdot \text{embed}(\boldsymbol{\mathcal{Y}}))$$

t-product is calculated very efficiently by using the diagonalization provided by DFT of circulant matrices as described above.

$$\boldsymbol{\mathcal{Z}}(:,:,k) = \text{unembed}\left(\text{ifft}\left(\breve{\boldsymbol{\mathcal{X}}}(:,:,k) \cdot \breve{\boldsymbol{\mathcal{Y}}}(:,:,k)\right)\right)$$

- The ***t-SVD*** of $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I \times J \times K}$ is: $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{U}} *_t \boldsymbol{\mathcal{D}} *_t \boldsymbol{\mathcal{V}}^T$, where $\boldsymbol{\mathcal{U}} \in \mathbb{R}^{I \times I \times K}$, $\boldsymbol{\mathcal{V}} \in \mathbb{R}^{J \times J \times K}$ are t-orthogonal tensors and $\boldsymbol{\mathcal{D}} \in \mathbb{R}^{I \times J \times K}$ is a tensor with diagonal faces. As with the usual SVD it provides an optimal approximation of a tensor in the Frobenius norm of the difference. (see Theorem 4.3 in (Kilmer & Martin, 2011)).
- The **tensor nuclear norm** (TNN) is defined from the t-SVD as: $\|\boldsymbol{\mathcal{X}}\|_{\circledast} = \sum_{i=1}^{\min(I,J)} \sum_{k=1}^{K} \breve{\boldsymbol{\mathcal{D}}}(i,j,k)$

II. Granger Causality With t-Products

Assume that the fMRI data is sampled from I_{Cx} voxels of the cortical grid at $I_{T\delta}$ time samples forming the data matrix as $\mathbf{B} \in \mathbb{R}^{I_{Cx} \times I_{T\delta}}$. Lagged data values are collected in the tensor $\boldsymbol{\mathcal{B}} \in \mathbb{R}^{I_{lag} \times I_{Cx} \times I_T}$ and connectivity tensor is denoted as $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{I_{Cx} \times I_{Cx} \times I_{lag}}$ as described in the Section IV of the paper.

Define the sample covariance $\boldsymbol{\mathcal{R}} \in \mathbb{R}^{I_{Cx} \times I_{Cx} \times (I_{lag}+1)}$ as follows:

$$\boldsymbol{\mathcal{R}}(i_{Cx}, i_{Cx}, i_{lag}) = \frac{1}{I_T} \sum_{i_T=1}^{I_T} \boldsymbol{\mathcal{B}}(i_{lag}, i_{Cx}, i_T)\mathbf{B}_t(i_{Cx}, i_T)$$

The Levinson-Durbin equations are represented in tensorial framework as follows:

$$\begin{pmatrix} \boldsymbol{\mathcal{R}}(:,:,0) & \boldsymbol{\mathcal{R}}(:,:,1)^H & \cdots & \boldsymbol{\mathcal{R}}(:,:,I_{lag}-1)^H \\ \boldsymbol{\mathcal{R}}(:,:,1) & \boldsymbol{\mathcal{R}}(:,:,0) & \cdots & \boldsymbol{\mathcal{R}}(:,:,I_{lag}-2)^H \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\mathcal{R}}(:,:,I_{lag}-1) & \boldsymbol{\mathcal{R}}(:,:,I_{lag}-2) & \cdots & \boldsymbol{\mathcal{R}}(:,:,0) \end{pmatrix} \begin{pmatrix} \boldsymbol{\mathcal{A}}(:,:,1) \\ \boldsymbol{\mathcal{A}}(:,:,2) \\ \vdots \\ \boldsymbol{\mathcal{A}}(:,:,I_{lag}) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mathcal{R}}(:,:,1) \\ \boldsymbol{\mathcal{R}}(:,:,2) \\ \vdots \\ \boldsymbol{\mathcal{R}}(:,:,I_{lag}) \end{pmatrix}$$

By using the notation introduced in Section I these equations can be expressed in tensor operations:

$$\mathrm{tplz}(\boldsymbol{\mathcal{R}}_1) \bullet_{\{I_{Cx} \cdot I_{lag}\}} \mathrm{MatVec}(\boldsymbol{\mathcal{A}}) = \mathrm{MatVec}(\boldsymbol{\mathcal{R}}_2)$$

where $\boldsymbol{\mathcal{R}}_1 = \boldsymbol{\mathcal{R}}(:,:,0:I_{lag}-1)$ and $\boldsymbol{\mathcal{R}}_2 = \boldsymbol{\mathcal{R}}(:,:,1:I_{lag})$.

The naïve solution of $\boldsymbol{\mathcal{A}}$ is:

$$\mathrm{MatVec}(\boldsymbol{\mathcal{A}}) = \mathrm{tplz}(\boldsymbol{\mathcal{R}}_1)^{-1} \mathrm{MatVec}(\boldsymbol{\mathcal{R}}_2)$$

This equation can be represented in t-product notation:

$$\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{R}}_1^{-1} *_t \boldsymbol{\mathcal{R}}_2$$

Since $\boldsymbol{\mathcal{R}}_1$ is calculated from sample covariance, we propose to use tensor nuclear norm to obtain a stable solution.

Tensor nuclear norm regularization of $\boldsymbol{\mathcal{R}}_1$ is estimated by:

$$\hat{\boldsymbol{\mathcal{R}}}_1 = \arg\min_{\Lambda} \left\{ \left\| \boldsymbol{\mathcal{R}}_1 - \Lambda \right\|_2^2 + \lambda \left\| \Lambda \right\|_{\circledast} \right\}$$

$\boldsymbol{\mathcal{R}}_1$ is found by using the t-product:

$$\hat{\boldsymbol{\mathcal{R}}}_1 = \boldsymbol{\mathcal{U}} *_t \rho(\boldsymbol{\mathcal{D}}) *_t \boldsymbol{\mathcal{V}}^T$$

For the $\rho(\cdot)$, we used the shrinking function suggested by (Chi & Lange, 2014). The details can be found in the Algorithm.

Then the connectivity tensor $\boldsymbol{\mathcal{A}}$ is found as:

$$\hat{\boldsymbol{\mathcal{A}}} = \boldsymbol{\mathcal{V}} *_t \rho(\boldsymbol{\mathcal{D}}) *_t \boldsymbol{\mathcal{U}}^T *_t \boldsymbol{\mathcal{R}}_2$$

III. ALGORITHM

Circulant Embedding of $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I \times J \times K}$

$\boldsymbol{\mathcal{S}} = \mathrm{embed}(\boldsymbol{\mathcal{X}})$

Inputs: $\boldsymbol{\mathcal{X}}$

Outputs: $\boldsymbol{\mathcal{S}}$

$\boldsymbol{\mathcal{S}}(:,:,1:K) = \boldsymbol{\mathcal{X}}$

$\boldsymbol{\mathcal{S}}(:,:,K+1) = 0.5 \cdot (\boldsymbol{\mathcal{X}}(:,:,K) + \boldsymbol{\mathcal{X}}(:,:,K)^H);$

$\boldsymbol{\mathcal{S}}(:,:,K+2:2K) = \boldsymbol{\mathcal{X}}(:,:,K:-1:2)^H;$

Granger Causality with t - Product

Inputs : $\mathbf{B}, I_{lag}, \lambda, \alpha$

Outputs : $\boldsymbol{\mathcal{A}}$

Normalize $\mathbf{B}$

$$\mathbf{B} = \frac{\mathbf{B} - \text{mean}(\mathbf{B})}{\text{var}(\mathbf{B})}$$

Calculate the Covariance

for $i = 0 : I_{lag}$

$$\mathbf{H}(i \cdot I_{Cx} : (i+1) \cdot I_{Cx} - 1, :) = \begin{pmatrix} \mathbf{0}_i & \mathbf{B} & \mathbf{0}_{I_{lag}-i} \end{pmatrix}$$

end for

$$\boldsymbol{\mathcal{R}} = \frac{1}{I_T} \text{fold}(\mathbf{H}(1 : \text{I}_{Cx}, :)\mathbf{H}^T)$$

$$\boldsymbol{\mathcal{R}}_1 = \boldsymbol{\mathcal{R}}(:, :, 0 : I_{lag} - 1)$$

$$\boldsymbol{\mathcal{R}}_2 = \boldsymbol{\mathcal{R}}(:, :, 1 : I_{lag})$$

Circulant Embedding of $\boldsymbol{\mathcal{R}}_1 : \boldsymbol{\mathcal{S}}_1 = \text{embed}(\boldsymbol{\mathcal{R}}_1)$

Circulant Embedding of $\boldsymbol{\mathcal{R}}_2 : \boldsymbol{\mathcal{S}}_2 = \text{embed}(\boldsymbol{\mathcal{R}}_2)$

TNN on $\boldsymbol{\mathcal{S}}_1$

$$\breve{\boldsymbol{\mathcal{S}}}_1 = \text{fft}(\boldsymbol{\mathcal{S}}_1, [], 3)$$

for $i = 0 : 2I_{lag} - 1$

$$[\mathbf{U}, \mathbf{S}, \mathbf{V}] = \text{svd}(\breve{\boldsymbol{\mathcal{S}}}_1(:, :, i))$$

$$d = [\text{diag}(S)]^2$$

$$d = \frac{-I_{Cx} + \sqrt{I_{Cx}{}^2 + 4\lambda\alpha(I_{Cx}d + \lambda(1-\alpha))}}{2\lambda\alpha}$$

$$\mathbf{S} = \text{diag}(\sqrt{\text{d}})$$

$$\text{t-inv}[\breve{\boldsymbol{\mathcal{S}}}_1(:, :, i)] = \mathbf{VSU}^T$$

end for

Calculate $\boldsymbol{\mathcal{A}}$

$$\breve{\boldsymbol{\mathcal{S}}}_2 = \text{fft}(\boldsymbol{\mathcal{S}}_2, [], 3)$$

for $i = 0 : 2I_{lag} - 1$

$$\breve{\boldsymbol{\mathcal{A}}}(:, :, i) = \text{t-inv}[\breve{\boldsymbol{\mathcal{S}}}_1(:, :, i)] \cdot \breve{\boldsymbol{\mathcal{S}}}_2(:, :, i)$$

end for

$$\boldsymbol{\mathcal{A}} = \text{ifft}(\breve{\boldsymbol{\mathcal{A}}}, [], 3)$$

$$\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{A}}(:, :, 0 : I_{lag})$$

References

Chi, E. C., & Lange, K. (2014). Stable estimation of a covariance matrix guided by nuclear norm penalties. *Computational Statistics and Data Analysis*, *80*, 117–128. Methodology; Numerical Analysis. doi:10.1016/j.csda.2014.06.018

Dietrich, C. R., & Newsam, G. N. (1997). Fast and Exact Simulation of Stationary Gaussian Processes through Circulant Embedding of the Covariance Matrix. *SIAM Journal on Scientific Computing*. doi:10.1137/S1064827592240555

Kilmer, M. E., Braman, K., Hao, N., & Hoover, R. C. (2013). Third-Order Tensors as Operators on Matrices: A Theoretical and Computational Framework with Applications in Imaging. *SIAM Journal on Matrix Analysis and Applications*, *34*(1), 148–172. doi:10.1137/110837711

Kilmer, M. E., & Martin, C. D. (2011). Factorization strategies for third-order tensors. *Linear Algebra and Its Applications*, *435*(3), 641–658. doi:10.1016/j.laa.2010.09.020