Nonlinear Bayesian Estimation of fMRI BOLD Signal under Non-Gaussian Noise

Ali Fahim Khan

Supervised by: Dr. Muhammad Shahzad Younis

Dr. Kashif M. Rajpoot  Dr. Khawar Khurshid  Dr. Amir Ali Khan
Introduction

- fMRI-Functional Magnetic Resonance Imaging
  - A window to the brain!
  - A non-invasive tool to study the neural activity
- Example: A blind individual reading Braille

Applications of fMRI
- Making brain atlas (HCP)
- Brain disease diagnosis such as Alzheimer's (Sterling R., 2011)
- Training patients having brain illness to cure them (Birbaumer et al., 2007)
- Lie detector (Langleben et al., 2005)

Image courtesy http://cortivis.umh.es/overview.htm
Introduction

- fMRI data is acquired using an fMRI scanner
- Time series data generated from voxels
- By studying this noisy data, one can make inferences on brain activity

- It is this brain activity we are interested in!

Photo Cour. devendradesmukh.blogspot.com

Literature Review

fMRI is most commonly performed using blood oxygenation level-dependent (BOLD) contrast (Ogawa et al., 1992)

Methods to study BOLD signal
- Statistical Parametric Mapping (SPM) (Friston, 1995)
- Methods based on the Hemodynamic Model
- Hybrid methods such as Genetic Algorithms and simulated annealing (Vakorin et al., 2007).
Literature Review

- The Hemodynamic Approach

- [Image of a brain and sensory organs with a flow inducing signal and BOLD signal]

- **Measurement Model**

\[ y = V_0 \left[ k_1 (1 - q) + k_2 \left( 1 - \frac{q}{v} \right) + k_3 (1 - v) \right] \]

\[ k_1 = 7E_0, \quad k_2 = 2, \quad k_3 = 2E_0 - 0.2 \]

- **Buxton et al. 1998**
- **Modified by** Mandeville et al. 1999
- **Completed by** Friston et al. 2000
Literature Review

- Friston et al. first solved the hemodynamic model using Volterra Kernels series (Friston et al. 2000)
- Later Gitelman et al. introduced Dynamic Causal Modeling that linked different regions of the brain together (Gitelman et al. 2003). Does not include physiological noise
- Riera et al. first included physiological noise to the hemodynamic model and performed blind deconvolution using Radial Basis Functions (Riera et al. 2004)
Literature Review

- Hu et al. inverted the model using the SR-UKF (Hu et al. 2009)
- Martin utilized the SR-CKF to perform blind deconvolution of the Hemodynamic model (Martin, 2010)
- All of the methods above assumed both the process noise as well as measurement noise to be Gaussian.
- Studies indicate Non-Gaussian noise in fMRI data
  - Gamma distribution (Stephen et al., 2001)
  - Rician Distribution (Arnold et al., 2005)
  - Impulsive noise (Josephs at al., 2007)
Literature Review

- Gaussian Sum Filtering can tackle non-Gaussian noise. (Alspach and Sorenson 1972). However computationally expensive.
- To alleviate the computational problem, (Plataniotis et al. 1997) proposed the Adaptive-GSF (AGSF). He and colleagues demonstrated its use in narrowband inference in presence of non-Gaussian noise (Plataniotis et al. 2000).
- (Miroslav et al., 2005) proposed the Sigma point GSF (SPGSF) having bank of SR-UKF and applied it to the non-Gaussian noise distribution problem.
Problem Statement

The problem boils down to finding the estimates of $x_k$ given a noisy times series data $y_k$. 

Parameters

The problem boils down to finding the estimates of $x_k$ given a noisy times series data $y_k$. 

$$f_{k-1}(x_{k-1}, u_{k-1})$$

$$h_k(x_k, w_k)$$

$\mathbf{v} \sim N(0, Q\mathbf{v})$

$\mathbf{w} \sim \text{Non-Gaussian}$
Methodology

- Modified Adaptive Gaussian Sum Filter
  EKF working in parallel each tuned to a component of the noise of interest. The posterior probability distribution is subsequently collapsed to yield single Gaussian term
Methodology

- Sigma Point Gaussian Sum Filter (SPGSF)

Quite similar to GSF. But here multiple SR-UKF work in parallel each tuned to a specific Gaussian term constituting the Gaussian mixture needed to obtain the closed loop Bayesian Recursive Relation.
Experiments

Synthetic Data
- MAGSF vs EKF
  - Impulsive noise
- SPGSF vs UKF
  - Gamma Noise

Real Data
- MAGSF vs EKF
  - Impulsive noise

Voxel Data Synthesis

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Statistical Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \varepsilon )</td>
<td>( N(0.54, 0.01) )</td>
</tr>
<tr>
<td>( \tau_2 )</td>
<td>( N(1.54, 0.0625) )</td>
</tr>
<tr>
<td>( \tau_f )</td>
<td>( N(2.46, 0.0625) )</td>
</tr>
<tr>
<td>( \tau_0 )</td>
<td>( N(0.98, 0.0625) )</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>( N(0.33, 0.002025) )</td>
</tr>
<tr>
<td>( \varepsilon_2 )</td>
<td>( N(0.34, 0.01) )</td>
</tr>
<tr>
<td>( V_1 )</td>
<td>( N(0.02, 0.000025) )</td>
</tr>
</tbody>
</table>

Table from (Friston et al. 2000)

Block design Exp. parameters
- TR=1.2s
- 13 seconds ON=1 and 13 OFF=0 periodic stimulus for 150 Seconds
- [0 1 1 1] Initial states
MAGSF vs EKF (Synthetic Data)

- Impulsive noise (e-mixture)
  \[ w \sim (1-e)N(\mu_1, \sigma_{n1}^2) + eN(\mu_2, \sigma_{n2}^2) \]

  where \( e \in (0, 1) \) is the mixing parameter and varies from 0.01 to 0.1 and the ratio of \( \sigma_2 / \sigma_1 \) is usually in the range of 10 to 10,000 (vatola, 1984)

- Due to flexible nature, e-mixture can be used to model several non-Gaussian distributions

  Process Noise
  \[ Q_v = 0.0223 \cdot \text{eye}(9) \]

  Measurement Noise
  \[ w \sim (1-e)N(\mu_1, \sigma_{n1}^2) + eN(\mu_2, \sigma_{n2}^2) \]

  \[ e = 0.01, \mu_1 = 0.01, \mu_2 = 0.02, \sigma_{n1}^2 = 0.0001, \sigma_{n2}^2 = 0.05 \]

Performance Criteria

\[ RMSE = \sqrt{\frac{1}{k} \sum_{i=1}^{k} (\hat{x} - x_{true})^2} \]
SPGSF vs SR-UKF (Synthetic Data)

- Impulsive noise (e-mixture), same as used for MAGSF
  
  \[ e = 0.01, \mu_1 = 0.02, \mu_2 = 0.03, \sigma^2_{n_1} = 0.0001, \sigma^2_{n_2} = 0.05 \]

- Gamma noise

- Process Noise
  
  \[ Q_v = 0.0223 \text{eye}(9) \]

- Measurement Noise
  
  \[ p(w_K) = 0.048 \times N(0.00299,3.07e^{-7}) + \]
  
  \[ 0.318 \times N(0.018,1.93e^{-5}) + \]
  
  \[ 0.109 \times N(0.334,5.31e^{-5}) + \]
  
  \[ 0.525 \times N(0.011,1.06e^{-5}) \]
State Estimation Results (Synthetic Data)

MAGSF vs EKF under Impulsive Noise

SPGSF vs SRUKF under Impulsive Noise

SPGSF vs SRUKF under Gamma Noise
Parameter Estimation Results (Synthetic Data)

MAGSF vs EKF under Impulsive Noise

SPGSF vs SR-UKF under Impulsive Noise

SPGSF vs SR-UKF under Gamma Noise
RMSE Evolution with Time (Synt. Data)

MAGSF vs EKF under Impulsive Noise

SPGSF vs SR-UKF under Impulsive Noise

SPGSF vs SR-UKF under Gamma Noise

Performance Criteria

\[ RMSE = \sqrt{\frac{\sum_{i=1}^{k} (\hat{x} - x_{true})^2}{k}} \]
Experiment using Real fMRI Data

Block design experiment.
Right hand clenched then left hand periodically
TR=1.92s
13 seconds ON=1 and 13 OFF=0
Data taken from mccauslandcenter.sc.edu

- Preprocessing was performed using SPM8 and brain activation map was generated.
- Voxels/ROI with highest activation were selected and their time series extracted
- Marsbar Used for Voxel/ROI management
MAGSF vs EKF, Impulsive Noise

State Estimates

Parameter Estimates
SPGSF vs SRUKF, Impulsive Noise

State Estimates

Parameter Estimates
Summary of Results

Results of our study using real data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Impulsive MAGSF</th>
<th>Impulsive EKF</th>
<th>Gamma SPGSF</th>
<th>Gamma SRUKF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon$</td>
<td>0.61</td>
<td>0.33</td>
<td>0.55</td>
<td>0.55</td>
</tr>
<tr>
<td>$\tau_z$</td>
<td>1.96</td>
<td>2.00</td>
<td>1.49</td>
<td>1.52</td>
</tr>
<tr>
<td>$\tau_f$</td>
<td>1.89</td>
<td>0.50</td>
<td>2.32</td>
<td>2.33</td>
</tr>
<tr>
<td>$\tau_0$</td>
<td>2.08</td>
<td>0.50</td>
<td>0.80</td>
<td>0.91</td>
</tr>
<tr>
<td>$E_0$</td>
<td>0.49</td>
<td>0.9</td>
<td>0.55</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Results of previously reported studies

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Friston et al.</th>
<th>Hu et al.</th>
<th>Johnston et al.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon$</td>
<td>$N(0.54,0.25^2)$</td>
<td>0.5985</td>
<td>0.069±0.014</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$N(1.54,0.25^2)$</td>
<td>1.4953</td>
<td>4.98±1.07</td>
</tr>
<tr>
<td>$\tau_f$</td>
<td>$N(2.46,0.25^2)$</td>
<td>2.4889</td>
<td>8.31±1.51</td>
</tr>
<tr>
<td>$\tau_0$</td>
<td>$N(0.98,0.25^2)$</td>
<td>1.0871</td>
<td>8.38±1.5</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td>$N(0.34,0.1^2)$</td>
<td>0.5272</td>
<td>0.0635±0.072</td>
</tr>
</tbody>
</table>
Conclusion and Discussion

- A filter designed to work under Gaussian noise may perform poor if made to work under non-Gaussian noise.
- We proposed a novel filter (MAGSF) and applied it together with SPGSF to the framework of the hemodynamic model. Our results match with those presented in earlier studies.
- Both are global filters and are less susceptible to getting stuck at a local minima.
- The parallel architecture enables noise of arbitrary distribution to be handled by the filters.
- The proposed filter can be applied to a number of applications such as sensor fusion, radar tracking applications, etc.
- Since the filters are recursive, they can be implemented on hardware with ease.
Contributions to the Scientific Community

Proposed a novel optimal filter and applied it to the problem of hemodynamic model under non-Gaussian noise environment

“Nonlinear Bayesian Estimation of BOLD Signal Under Non-Gaussian Noise” *(Accepted and in Press)* by Computational and Mathematical Methods in Medicine. **Impact Factor 1.018**

Applied SPGSF to the problem of hemodynamic model under non-Gaussian noise

“Nonlinear Bayesian Estimation of BOLD Signal Under Non-Gaussian Noise Using Sigma Point Gaussian Sum Filter” *(Under Review)* by EURASIP Journal of Advances in Signal Processing. **Impact Factor 0.81**
Human Connectome Project, http://www.humanconnectomeproject.org


Thank you!
Appendix
Noise Analysis

Noise Analysis done on both the active and inactive voxels
Noise Analysis

Most Active Voxels

![Graph showing normalized amplitude vs sample number for most active voxels.](image-url)
Noise Analysis

Least Active Voxels

![Graph showing least active voxels with normalized amplitude on the y-axis and sample number on the x-axis. The x-axis ranges from 0 to 300, and the y-axis ranges from 0.01 to 0.1. The graph displays multiple lines representing different segments of the data.]
Noise Analysis

Correlation Matrix between normalized adjacent most activated voxel time series.

<table>
<thead>
<tr>
<th>Voxel</th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
<th>V5</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>1.000</td>
<td>0.884</td>
<td>0.541</td>
<td>0.898</td>
<td>0.372</td>
</tr>
<tr>
<td>V2</td>
<td>0.884</td>
<td>1.000</td>
<td>0.486</td>
<td>0.842</td>
<td>0.419</td>
</tr>
<tr>
<td>V3</td>
<td>0.541</td>
<td>0.486</td>
<td>1.000</td>
<td>0.615</td>
<td>0.772</td>
</tr>
<tr>
<td>V4</td>
<td>0.898</td>
<td>0.842</td>
<td>0.615</td>
<td>1.000</td>
<td>0.504</td>
</tr>
<tr>
<td>V5</td>
<td>0.372</td>
<td>0.419</td>
<td>0.772</td>
<td>0.504</td>
<td>1.000</td>
</tr>
</tbody>
</table>

*V1 to V5 represents voxel 1 to voxel 5 respectively.*
# Noise Analysis

**Correlation Matrix between normalized adjacent most activated voxel and least activated voxel time series.**

<table>
<thead>
<tr>
<th>Voxel</th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
<th>V5</th>
<th>V6</th>
<th>V7</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>1.0000</td>
<td>0.7135</td>
<td>0.0873</td>
<td>0.0007</td>
<td>-0.0356</td>
<td>-0.2654</td>
<td>0.1059</td>
</tr>
<tr>
<td>V2</td>
<td>0.7135</td>
<td>1.0000</td>
<td>0.3408</td>
<td>0.5555</td>
<td>-0.0236</td>
<td>-0.2855</td>
<td>0.0331</td>
</tr>
<tr>
<td>V3</td>
<td>0.0873</td>
<td>0.3408</td>
<td>1.0000</td>
<td>0.2797</td>
<td>0.0736</td>
<td>-0.0825</td>
<td>0.0579</td>
</tr>
<tr>
<td>V4</td>
<td>0.0007</td>
<td>0.5555</td>
<td>0.2797</td>
<td>1.0000</td>
<td>0.4637</td>
<td>0.0946</td>
<td>0.1714</td>
</tr>
<tr>
<td>V5</td>
<td>-0.0356</td>
<td>-0.0236</td>
<td>0.0736</td>
<td>0.4637</td>
<td>1.0000</td>
<td>0.1036</td>
<td>0.5672</td>
</tr>
<tr>
<td>V6</td>
<td>-0.2654</td>
<td>-0.2855</td>
<td>-0.0825</td>
<td>0.0946</td>
<td>0.1036</td>
<td>1.0000</td>
<td>0.1367</td>
</tr>
<tr>
<td>V7</td>
<td>0.1059</td>
<td>0.0331</td>
<td>0.0579</td>
<td>0.1714</td>
<td>0.5672</td>
<td>0.1367</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

*V1 to V7 represents voxel 1 to voxel 7 respectively.*
**Noise Analysis**

**Correlation Matrix between normalized adjacent least activated Voxel time series.**

<table>
<thead>
<tr>
<th>Voxel</th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
<th>V5</th>
<th>V6</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>1.0000</td>
<td>-0.0286</td>
<td>0.0744</td>
<td>0.1054</td>
<td>-0.0058</td>
<td>-0.0499</td>
</tr>
<tr>
<td>V2</td>
<td>-0.0286</td>
<td>1.0000</td>
<td>0.7135</td>
<td>0.0007</td>
<td>-0.0356</td>
<td>0.1059</td>
</tr>
<tr>
<td>V3</td>
<td>0.0744</td>
<td>0.7135</td>
<td>1.0000</td>
<td>0.5555</td>
<td>-0.0236</td>
<td>0.0331</td>
</tr>
<tr>
<td>V4</td>
<td>0.1054</td>
<td>0.0007</td>
<td>0.5555</td>
<td>1.0000</td>
<td>0.4637</td>
<td>0.1714</td>
</tr>
<tr>
<td>V5</td>
<td>-0.0058</td>
<td>-0.0356</td>
<td>-0.0236</td>
<td>0.4637</td>
<td>1.0000</td>
<td>0.5672</td>
</tr>
<tr>
<td>V6</td>
<td>-0.0499</td>
<td>0.1059</td>
<td>0.0331</td>
<td>0.1714</td>
<td>0.5672</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

*V1 to V6 represents voxel 1 to voxel 6 respectively.*
Noise Analysis

Difference between two time series of voxels yields noise

![Graph showing signal amplitude over sample number and voxel vs adjacent voxel differences.]

![Graph showing amplitude difference over sample number with normal distribution and gamma distribution fitting.]

![Graph showing standard normal quartiles and amplitude difference density distribution.]

Data
Normal Distribution
Gamma Distribution
Noise Analysis

Q-Q plot to test Gaussianity

Amplitude Difference

Standard Normal Quartiles

Amplitude Difference

Standard Normal Quartiles

Amplitude Difference

Standard Normal Quartiles

Amplitude Difference

Standard Normal Quartiles