Hierarchical Bayesian Modeling for EEG/MEG:
From Simulated to Experimental Data

Mini-Symposium "Inverse Problems with Experimental Data"

Applied Inverse Problem Conference 2013 in Daejeon, Korea
Preamble...

Warning

This is not a talk about success!
Preamble...

Warning

This is not a talk about direct success!
Outline

Hierarchical Bayesian Modeling in EEG/MEG Source Reconstruction

First Results for Experimental Data

Sensitivity Studies

Current Results for Experimental Data

Conclusion & Next Steps
Source Reconstruction by Electroencephalography (EEG) and Magnetoencephalography (MEG)

Aim: Reconstruction of brain activity by non-invasive measurement of induced electromagnetic fields (bioelectromagnetism) outside of the skull.
One Challenge of Source Reconstruction: The Inverse Problem

Unfortunately, not just:

- Under-determined
- Severely ill-conditioned, special spatial characteristics.
- Signal is contaminated by a complex spatio-temporal mixture of external and internal noise and nuisance sources.

Felix Lucka (felix.lucka@wwu.de)
Discretization Approach: Current Density Reconstruction (CDR)

Continuous (ion current) vector field $\approx$ grid with 3 orthogonal elementary sources at each node.

\[ f = K \ u, \quad \implies \quad p_{li}(f \mid u) \propto \exp\left(-\frac{1}{2}\|\Sigma^{-1/2}(f - K \ u)\|_2^2\right)\]
Cooperation with...

Dr. Sampsa Pursiainen
Department of Mathematics and Systems Analysis,
Aalto University, Finland

Prof. Dr. Martin Burger
Institute for Applied Mathematics,
University of Münster, Germany

PD. Dr. Carsten Wolters
Institute for Biomagnetism and Biosignalanalysis,
University of Münster, Germany

Felix Lucka (felix.lucka@wwu.de)
Hierarchical Bayesian Modeling for Source Reconstruction

- Extend Gaussian prior model by flexible, individual source variances $\gamma_i$.
- Let the data determine $\gamma_i$ (hyperparameters).
- Use sparsity constraints on hyperprior $\sim$ by direct correspondence, we might get sparsity over the primary unknowns $u$ as well.
- Resulting regularization functional is non-convex / Posterior is multimodal.
- Compute Full-Conditional Mean (CM) or Full-Maximum A-Posteriori (MAP) estimates.

Our starting point:

Hierarchical Bayesian inference for the EEG inverse problem using realistic FE head models: Depth localization and source separation for focal primary currents

Felix Lucka, Sampsa Pursiainen, Martin Burger, Carsten H. Wolters

Institute for Biomagnetism and Biosignalanalysis, University of Muenster, Germany

Abstract

In its generic formulation, the inverse problem of EEG/MEG Current Density Reconstruction can be formulated as a linear relation between the magnetic field distribution on the head surface for a given source activity and the unknown source model. All configurations consisting of few, focal sources when used with realistic, high-resolution Finite Element (FE) head models. The main foci of interest are the correct depth localization, a well-known source of systematic error of many CDR methods, and the separation of single sources in multiple-source scenarios. Both aspects are very important in the analysis of neurophysiological data and in clinical applications.

Introduction

Hierarchical Bayesian modeling provides a promising framework and is able to improve upon established CDR methods such as the minimum norm estimator (MNE) and the minimum current estimator (MCE) by incorporating a priori information that can help to choose a particular solution from the set of likely solutions. This a posteriori information can reflect computational constraints as well as with extremely different properties, can explain the measured activity by means of non-invasive measurements of the associated fields. All established methods show crucial errors, promising results are attained. Additionally, we introduce Wasserstein distances as performance measures for the validation of inverse methods in complex source scenarios.

Implementation of HBM with realistic, high resolution Finite Element (FE) head models.

Improve Full-MAP estimation by utilizing MCMC.

Systematic examination of different aspects in extensive simulation studies.

EEG vs. MEG and EEG/MEG combination (EMEG)
From Simulated to Real Data...

Partial results in:

Felix Lucka., Sampsa Pursiainen, Martin Burger, Carsten H. Wolters.  
Hierarchical Bayesian Inference for the EEG Inverse Problem using Realistic FE Head Models: Depth Localization and Source Separation for Focal Primary Currents.  
Neuroimage, 61(4), 2012.

- Summary: Excellent results for sparse source configurations! Overcomes deficiencies of established inverse methods.
- Proceed to experimental data as soon as possible!
- Start with evoked responses, e.g. simple auditory activity.
- Proceed to interictal epileptic activity.
- Extend simple HBM in every possible way (spatio-temporal, multimodal, multi-resolution...).

Felix Lucka (felix.lucka@wwu.de)
Outline

Hierarchical Bayesian Modeling in EEG/MEG Source Reconstruction

First Results for Experimental Data

Sensitivity Studies

Current Results for Experimental Data

Conclusion & Next Steps

Felix Lucka (felix.lucka@wwu.de)
Auditory Data I: MEG Butterfly Plot

Aim: **Fast results** for Biomag 2012 conference.

1. Record responses to many stimuli.
3. Average responses of 109 stimuli.

Warning: Only illustration, different data set!
Auditory Data I: MEG Topographic Field Distribution

Warning: Only illustration, different data set!
Auditory Data I: Some Images

Full CM estimate for MEG data
Auditory Data I: Some Images

Full MAP estimate I for MEG data
Auditory Data I: Some Images

Full MAP estimate II for MEG data
Auditory Data I: Summary of Observations

- Full-MAP estimates are unstable and sometimes totally senseless.
- Not robust to parameter choice?!
- CM estimate seem more robust, but could also be better.
- Others report similar results:
  - "...these fancy non-linear, non-convex methods...good in simulations @#+! in practice"
  - "...too sensitive and not robust enough..."
  and give a lot of advice what to do better...
Outline

Hierarchical Bayesian Modeling in EEG/MEG Source Reconstruction

First Results for Experimental Data

Sensitivity Studies

Current Results for Experimental Data

Conclusion & Next Steps
What Shall We Do?

- Naive inverse problem guy’s view: $f = Ku$.
  - Give me $K$
  - Give me $f$ (and $\Sigma_\varepsilon$)
  - I’ll return $u$
  - Tell me if it is good.

- Works until you face a problem
What Shall We Do?

- Naive inverse problem guy’s view: $f = Ku$.
  - Give me $K \rightsquigarrow$ forward modeling, sensor registration
  - Give me $f$ (and $\Sigma_\epsilon$) $\rightsquigarrow$ preprocessing
  - I’ll return $u$
  - Tell me if it is good.

- Works until you face a problem

- HBM more sensitive to errors/uncertainties?
What Shall We Do?

- Naive inverse problem guy’s view: $f = Ku$.
  - Give me $K \leadsto$ forward modeling, sensor registration
  - Give me $f$ (and $\Sigma_\epsilon$) $\leadsto$ preprocessing
  - I’ll return $u$
  - Tell me if it is good.

- Works until you face a problem

- HBM more sensitive to errors/uncertainties?

I have control the whole pipeline to find out!
What Shall We Do?

▶ Naive inverse problem guy’s view: \( f = Ku. \)
  - Give me \( K \leadsto \) forward modeling, sensor registration
  - Give me \( f (\text{and } \Sigma_\varepsilon) \leadsto \) preprocessing
  - I’ll return \( u \)
  - Tell me if it is good.

▶ Works until you face a problem

▶ HBM more sensitive to errors/uncertainties?

I have control the whole pipeline to find out!

→ Replace commercial software by own pipeline: Extremely time consuming!
Example of What Might Go Wrong: Noise Modeling

Naive approach: \( f = K u + \epsilon \)

- Simulation studies: \( \epsilon \sim N(0, \sigma^2) \) (iid).
- Reality: \( \epsilon = \delta + \eta + \kappa \)
  - \( \delta \): Sensor noise and external nuisance fields → Empty room recordings
  - \( \eta \): Averaged internal nuisance fields → highly correlated.
  - \( \kappa \): Averaged background brain activity → highly correlated, in range of forward operator.

- Static temporal filtering?
- Blind unmixing? PCA, ICA?
- Model-based unmixing?
Example of What Might Go Wrong: Empty Room Recordings

Condition number of covariance matrix: 50
Largest ratio of variance: $\approx 5.1$
Example of What Might Go Wrong: Prestimulus Data

(a) Covariance
(b) Correlation

Condition number of covariance matrix: 6839
Largest ratio of variance: \(= 5.3\)
Example of What Might Go Wrong: Preprocessing and Noise Estimation

1. Static temporal bandpass (1-30 Hz) filtering: $\rightarrow$ temporal correlation.
2. Average data of 109 epochs to improve SNR.
3. Estimate the channel variances $\sigma_i^2$ from this data based on pre-stimulus interval ($\sim 300$ avg. sample).
4. Use $\Sigma_\varepsilon = \text{diag}(\sigma_i^2)$ (like CURRY does).

Abbildung: Timecourses of EEG channels 1 and 2 (left) and 20 and 50 (right).

Felix Lucka (felix.lucka@wwu.de)
Simulation Studies for Noise Sensitivity

Observations:

- Degenerate covariance structure, correlated noise components dominate.
- Preprocessing renders estimation of residual noise difficult.
Simulation Studies for Noise Sensitivity

Observations:
- Degenerate covariance structure, correlated noise components dominate.
- Preprocessing renders estimation of residual noise difficult.

Is HBM sensitive to these issues?

Examine by simulation studies:
- Reconstruct 1000 single sources from noisy measurements (SNR = 20).
- Use different noise covariance for noise simulation and reconstruction
- Use average localization error of reconstruction for validation.
Simulation Studies for Noise Sensitivity: Minimum Norm Solution

\[ u_{\text{MNE}} = \arg\min \left\{ \| \Sigma_{\epsilon}^{-1/2} (f - K u) \|^2_2 + \lambda \| u \|^2_2 \right\} \]

<table>
<thead>
<tr>
<th>Data Cov</th>
<th>( \bar{\sigma}^2 \cdot I_m )</th>
<th>( \text{diag}(\Sigma) )</th>
<th>( \Sigma )</th>
<th>( \Sigma_{\text{perm}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{\sigma}^2 \cdot I_m )</td>
<td>17.54</td>
<td>17.52</td>
<td>18.08</td>
<td>17.73</td>
</tr>
<tr>
<td>( \text{diag}(\Sigma) )</td>
<td>x</td>
<td>17.39</td>
<td>17.98</td>
<td>17.58</td>
</tr>
<tr>
<td>( \Sigma )</td>
<td>x</td>
<td>x</td>
<td>17.51</td>
<td>34.87</td>
</tr>
<tr>
<td>( \Sigma_{\text{perm}} )</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>17.38</td>
</tr>
</tbody>
</table>

Abbildung: Localization error for minimum norm solution

\( \Sigma \) is given by empty room data.
Simulation Studies for Noise Sensitivity: HBM CM

<table>
<thead>
<tr>
<th>Model Cov</th>
<th>Data Cov</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{\sigma}^2 \cdot I_m$</td>
<td>5.49</td>
</tr>
<tr>
<td>diag($\Sigma$)</td>
<td>x</td>
</tr>
<tr>
<td>$\Sigma$</td>
<td>x</td>
</tr>
<tr>
<td>$\Sigma_{perm}$</td>
<td>x</td>
</tr>
</tbody>
</table>

**Abbildung:** Localization error for CM estimate

$\Sigma$ is given by empty room data.

Felix Lucka (felix.lucka@wwu.de)
Simulation Studies for Noise Sensitivity: HBM MAP

<table>
<thead>
<tr>
<th>Model Cov</th>
<th>Data Cov</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{\sigma}^2 \cdot I_m$</td>
</tr>
<tr>
<td>$\bar{\sigma}^2 \cdot I_m$</td>
<td>5.47</td>
</tr>
<tr>
<td>$\text{diag}(\Sigma)$</td>
<td>x</td>
</tr>
<tr>
<td>$\Sigma$</td>
<td>x</td>
</tr>
<tr>
<td>$\Sigma_{\text{perm}}$</td>
<td>x</td>
</tr>
</tbody>
</table>

Abbildung: Localization error for MAP estimate

$\Sigma$ is given by empty room data.
Simulation Studies for Noise Sensitivity: Summary Noise

<table>
<thead>
<tr>
<th>Model Cov</th>
<th>Data Cov</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{\sigma}^2 \cdot I_m$</td>
<td>5.49</td>
<td>5.68</td>
<td>5.88</td>
</tr>
<tr>
<td></td>
<td>$\text{diag}(\Sigma)$</td>
<td>x</td>
<td>5.57</td>
<td>5.77</td>
</tr>
<tr>
<td></td>
<td>$\Sigma$</td>
<td>x</td>
<td>x</td>
<td>5.65</td>
</tr>
<tr>
<td></td>
<td>$\Sigma_{perm}$</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model Cov</th>
<th>Data Cov</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{\sigma}^2 \cdot I_m$</td>
<td>5.47</td>
<td>5.64</td>
<td>5.87</td>
</tr>
<tr>
<td></td>
<td>$\text{diag}(\Sigma)$</td>
<td>x</td>
<td>5.59</td>
<td>5.76</td>
</tr>
<tr>
<td></td>
<td>$\Sigma$</td>
<td>x</td>
<td>x</td>
<td>5.57</td>
</tr>
<tr>
<td></td>
<td>$\Sigma_{perm}$</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

(a) CM   
(b) MAP

Result: HBM estimates are surprisingly robust against noise miss-specification!

Felix Lucka (felix.lucka@wwu.de)
Simulation Studies for Noise Sensitivity: Summary Noise

<table>
<thead>
<tr>
<th>Model Cov</th>
<th>Data Cov</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{\sigma}^2 \cdot I_m$</td>
<td>$\text{diag}(\Sigma)$</td>
<td>$\Sigma$</td>
<td>$\Sigma_{perm}$</td>
</tr>
<tr>
<td></td>
<td>5.49</td>
<td>5.68</td>
<td>5.88</td>
<td>5.57</td>
</tr>
<tr>
<td>diag($\Sigma$)</td>
<td>x</td>
<td>5.57</td>
<td>5.77</td>
<td>5.74</td>
</tr>
<tr>
<td>$\Sigma$</td>
<td>x</td>
<td>x</td>
<td>5.65</td>
<td>5.89</td>
</tr>
<tr>
<td>$\Sigma_{perm}$</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>5.48</td>
</tr>
</tbody>
</table>

(a) CM

<table>
<thead>
<tr>
<th>Model Cov</th>
<th>Data Cov</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{\sigma}^2 \cdot I_m$</td>
<td>$\text{diag}(\Sigma)$</td>
<td>$\Sigma$</td>
<td>$\Sigma_{perm}$</td>
</tr>
<tr>
<td></td>
<td>5.47</td>
<td>5.64</td>
<td>5.87</td>
<td>5.56</td>
</tr>
<tr>
<td>diag($\Sigma$)</td>
<td>x</td>
<td>5.59</td>
<td>5.76</td>
<td>5.71</td>
</tr>
<tr>
<td>$\Sigma$</td>
<td>x</td>
<td>x</td>
<td>5.57</td>
<td>5.73</td>
</tr>
<tr>
<td>$\Sigma_{perm}$</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>5.52</td>
</tr>
</tbody>
</table>

(b) MAP

Result: HBM estimates are surprisingly robust against noise miss-specification!

Good to know, but we still don’t know what is going on.

Felix Lucka (felix.lucka@wwu.de)
What Else Might Go Wrong: Forward Modeling

- Approximation error modeling for EEG/MEG?
- Better model calibration?
Outline

Hierarchical Bayesian Modeling in EEG/MEG Source Reconstruction

First Results for Experimental Data

Sensitivity Studies

Current Results for Experimental Data

Conclusion & Next Steps

Felix Lucka (felix.lucka@wwu.de)
In the meantime...Auditory Data II

- Rebuild of own pipeline complete
- **Aim:** Use the same inversion procedure to reproduce the bad results
- Examine how change in pipeline affect results
- Different subject (no particular reason)
Auditory Data II: Full CM Estimate for MEG
Auditory Data II: Full MAP Estimate for MEG
Auditory Data II: Summary

- Not able to reproduce bad results!
- Only able to produce good results!
- Very robust to parameter changes!
Outline

Hierarchical Bayesian Modeling in EEG/MEG Source Reconstruction

First Results for Experimental Data

Sensitivity Studies

Current Results for Experimental Data

Conclusion & Next Steps
Summary and conclusions:

▶ Don’t decide to do too much work to soon...take a different data set first.
▶ However, leaving the ”only the inverse problem” comfort zone pays off:
  ▶ I learned a lot about EEG/MEG as a whole.
  ▶ I have my own complete processing pipeline.
▶ HBM estimates are surprisingly robust.
▶ HBM can also give good results in practice.

Next steps:

▶ Reexamine first data set with own pipeline.
▶ Examine more data sets from different activity.
Thank you for your attention!

and thanks to:

★ Institute for Biomagnetism and Biosignalanalysis (WWU, Münster):
  Carsten Wolters, Ümit Aydin, Johannes Vorwerk,
  Benjamin Lanfer & Andreas Wollbrink
★ Institute for Applied Mathematics (WWU Münster):
  Martin Burger
★ Donders Institute, Nijmegen:
  Arno M. Janssen, Sumientra M. Rampersad & Dick F. Stegeman
★ Martinos Center, Boston:
  Seok Lew

Felix Lucka (felix.lucka@wwu.de)
**Depth Bias: Illustration**

One deep-lying reference source (blue cone) and minimum norm estimate:

\[ u_{\text{MNE}} = \text{argmin}\{\|\Sigma_e^{-1/2} (f - K u)\|_2^2 + \lambda \|u\|_2^2\} \]
Depth Bias: Illustration
One deep-lying reference source (blue cone) and sLORETA result (Pascual-Marqui, 2002).
Masking: Illustration

Reference sources.
Masking: Illustration

MNE result and reference sources (green cones).
Masking: Illustration

sLORETA result and reference sources (green cones).
Depth Bias: Full-Conditional Mean Estimate (CM)

Computed by blocked Gibbs MCMC sampler.
Depth Bias: Full-Maximum A-Posteriori Estimate (MAP), Algorithm I

Computed by alternating optimization, uniform initialization.
Depth Bias: Full-MAP, Algorithm II

Computed by alternating optimization initialized at the CM estimate.
Masking: Result Full-CM
Computed by blocked Gibbs MCMC sampler.
Masking: Result Full-MAP
Computed by alternating optimization initialized at the CM estimate.
Basics of Hierarchical Bayesian Modeling

Full Posterior:

\[ p_{\text{post}}(u, \gamma | f) \propto \exp \left( - \left( \frac{1}{2} \| \Sigma^{-1/2} (f - K u) \|_2^2 + \sum_{i=1}^{k} \left( \frac{1}{2} \| u_{\text{amp}} \|_2 + \beta \| \frac{u_{\text{amp}}}{\gamma_i} \|_2^2 + (\alpha + \frac{5}{2}) \ln \gamma_i \right) \right) \right) \]

- Quadratic/Gaussian with respect to \( u \).
- Factorizes over \( \gamma_i \)'s.
- (Regularization) energy is non-convex w.r.t. \( (u, \gamma) \) / Posterior is multimodal.

Felix Lucka (felix.lucka@wwu.de)
Simulation Studies for Background Sensitivity

- EEG/MEG is **severely ill-conditioned** and **underdetermined**.
- Inverse Solutions for EEG/MEG are **prior dominated**.
- Sensitive to miss-specification of prior?

Method
- Pure
  - A, 10%
  - A, 20%
  - B, 10%
  - B, 20%

<table>
<thead>
<tr>
<th>Method</th>
<th>Localization Errors (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNE</td>
<td>17.54, 17.58, 17.60, 17.55, 17.68</td>
</tr>
<tr>
<td>sLORETA</td>
<td>5.20, 5.29, 5.57, 5.26, 5.59</td>
</tr>
<tr>
<td>HBM CM</td>
<td>5.55, 5.90, 6.12, 5.72, 6.62</td>
</tr>
<tr>
<td>HBM MAP</td>
<td>5.58, 5.81, 6.23, 5.69, 6.80</td>
</tr>
</tbody>
</table>

Felix Lucka (felix.lucka@wwu.de)
Simulation Studies for Background Sensitivity

- EEG/MEG is severely ill-conditioned and underdetermined.
- Inverse Solutions for EEG/MEG are prior dominated.
- Sensitive to miss-specification of prior?

A : Add signal of two focal sources with 10% / 20% main signal strength.
B : Add signal of Gaussian sources with 10% / 20% main signal strength.

<table>
<thead>
<tr>
<th>Method</th>
<th>Pure</th>
<th>A, 10%</th>
<th>A, 20%</th>
<th>B, 10%</th>
<th>B, 20%</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNE</td>
<td>17.54</td>
<td>17.58</td>
<td>17.60</td>
<td>17.55</td>
<td>17.68</td>
</tr>
<tr>
<td>sLORETA</td>
<td>5.20</td>
<td>5.29</td>
<td>5.57</td>
<td>5.26</td>
<td>5.59</td>
</tr>
<tr>
<td>HBM CM</td>
<td>5.55</td>
<td>5.90</td>
<td>6.12</td>
<td>5.72</td>
<td>6.62</td>
</tr>
<tr>
<td>HBM MAP</td>
<td>5.58</td>
<td>5.81</td>
<td>6.23</td>
<td>5.69</td>
<td>6.80</td>
</tr>
</tbody>
</table>

Tabelle: Localization errors for background activity, Σ : Empty room data.

Felix Lucka (felix.lucka@wwu.de)