Westfälische Wilhelms-Universität Münster

The Bayesian Approach to Inverse Problems: Hierarchical Bayesian Approaches to EEG/MEG Source Reconstruction

Invited Talk at the University of Cambridge, UK

wissen.leben
WWU Münster

Felix Lucka

15.11.2012

Outline

"The human brain undoubtedly constitutes the most complex system in the known universe" (Wolf Singer, Director of the MPI for Brain Research)

Major branches of neuroscience (by Wikipedia):

Affective neuroscience

Behavioral neuroscience

Cellular neuroscience

Clinical neuroscience

Cognitive neuroscience

Computational neuroscience

Cultural neuroscience

Developmental neuroscience

Molecular neuroscience

Neuroengineering

Neuroimaging

Neuroinformatics

Neurolinguistics

Neurophysiology

Paleoneurology

Social neuroscience

Systems neuroscience

Needs people from: Biology, chemistry, computer science, engineering, linguistics, mathematics, medicine, philosophy, physics and psychology.

"The human brain undoubtedly constitutes the most complex system in the known universe" (Wolf Singer, Director of the MPI for Brain Research)

Major branches of neuroscience (by Wikipedia):

Affective neuroscience

Behavioral neuroscience

Cellular neuroscience

Clinical neuroscience

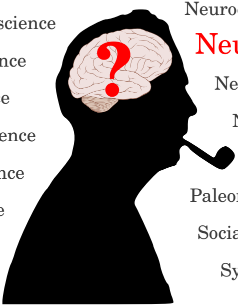
Cognitive neuroscience

Computational neuroscience

Cultural neuroscience

Developmental neuroscience

Molecular neuroscience



Neuroengineering

Neuroimaging

Neuroinformatics

Neurolinguistics

Neurophysiology

Paleoneurology

Social neuroscience

Systems neuroscience

Needs people from: Biology, chemistry, computer science, engineering, linguistics, mathematics, medicine, philosophy, physics and psychology.

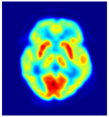
Major Modalities for Neuroimaging

X-ray imaging

- Projectional Radiography
- Computed Tomography (CT)

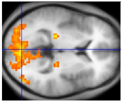
Nuclear imaging

- Planar Scintigraphy
- Positron emission tomography (PET)
- Single photon emission computed tomography (SPECT)



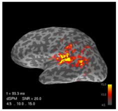
Magnetic resonance imaging (MRI)

- Basic structural scans
- Functional (fMRI)
- Diffusion weighted (DW-MRI)

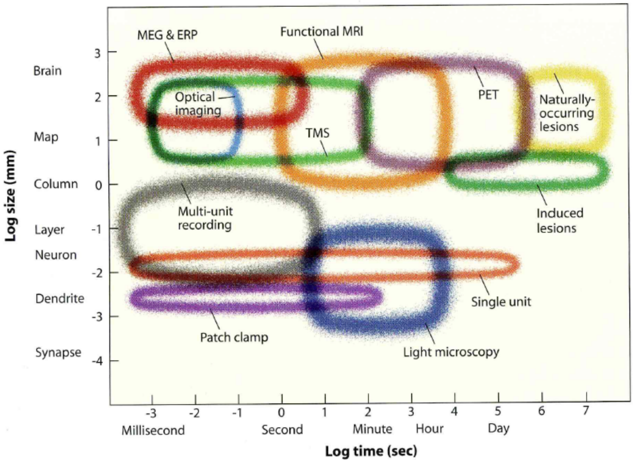


Bioelectromagnetic imaging:

- Electroencephalography (EEG)
- Magnetoencephalography (MEG)



Spatio-Temporal Resolution in Neuroimaging



source: Gazzaniga, Ivry & Mangun, Cognitive Neuroscience, 2nd ed., W.W.Norton & Company, 2002

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.

source: Wikimedia Commons

source: Wikimedia Commons

Institute for Biomagnetism and Biosignalanalysis

IBB
Institute for Biomagnetism and Biosignalanalysis

Focus on:

- Affective nsc
- Behavioral nsc
- Cognitive nsc
- Neuroimaging
- Clinical nsc
- Developmental nsc
- Neurolinguistics

Institute for Biomagnetism and Biosignalanalysis

IBB Institute for Biomagnetism and Biosignalanalysis

Focus on:

- Affective nsc
- Behavioral nsc
- Cognitive nsc
- Neuroimaging
- Clinical nsc
- Developmental nsc
- Neurolinguistics

Experimental devices used:

- MEG & EEG
- Behavioral laboratory
- MRI (Basic, fMRI, DW-MRI)
- tDCS & TMS

Current fields of research:

- Auditory system: Tinnitus, neuroplasticity;
- Emotion, attention and affection;
- Language & speech: Plasticity, cochlea implantation, aphasia;
- Visual system: Conscious vision;
- Neuromuscular disorders in stroke patients.
- Methodical development

Institute for Biomagnetism and Biosignalanalysis

IBB
Institute for
Biomagnetism and
Biosignalanalysis

Focus on:

- Affective nsc
- Behavioral nsc
- Cognitive nsc
- Neuroimaging
- Clinical nsc
- Developmental nsc
- Neurolinguistics

Experimental devices used:

- MEG & EEG
- Behavioral laboratory
- MRI (Basic, fMRI, DW-MRI)
- tDCS & TMS

Current fields of research:

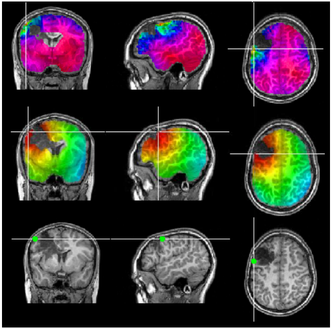
- Auditory system: Tinnitus, neuroplasticity;
- Emotion, attention and affection;
- Language & speech: Plasticity, cochlea implantation, aphasia;
- Visual system: Conscious vision;
- Neuromuscular disorders in stroke patients.
- Methodical development

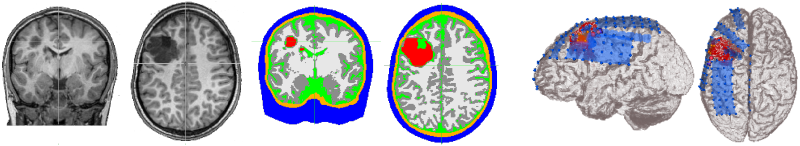
Workgroup “Methods in Bioelectromagnetism”

Aim: Improve quality, applicability and reliability of EEG/MEG based source reconstruction in the presurgical diagnosis of epilepsy patients.

Applications of EEG/MEG

- Diagnostic tool in neurology, e.g., Epilepsy.
- Scientific applications:
  - Examination tool in several fields neuroscience.
  - Validation of therapeutic approaches in clinical neuroscience.
  - Examination tool for neurophysiology.





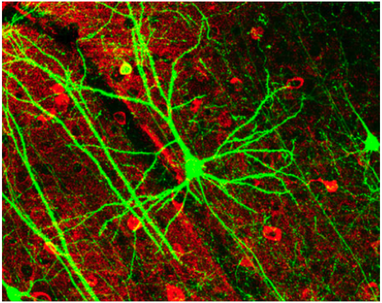
Challenges of Source Reconstruction: Mathematical Modeling

Mathematical modeling of bioelectromagnetism:

- Understand and model the transformation of the bio-chemical activity of the brain into ionic currents.
- Find reasonable simplifications to Maxwell's equations to formulate forward equations that relate ionic currents to measured signals:

$$\nabla \cdot (\sigma \nabla \phi) = \nabla \cdot \vec{j}^{pri} \quad + BC$$

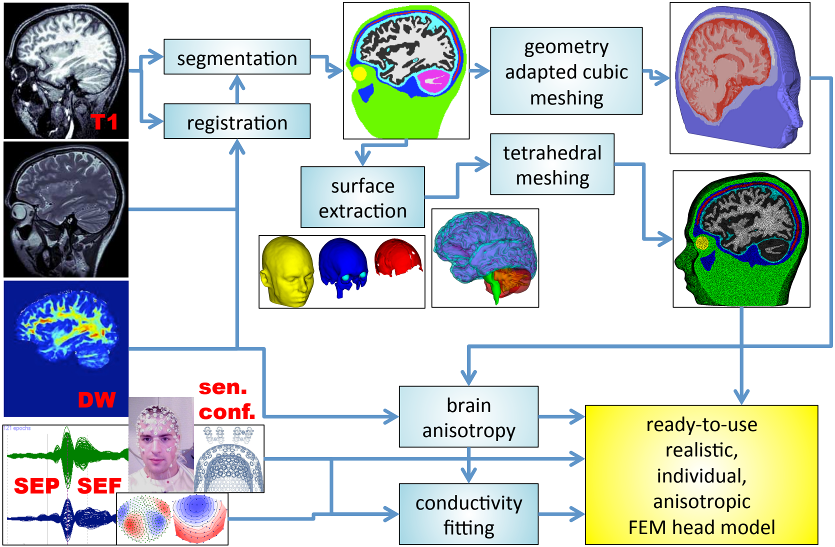
- σ: volume conductor model



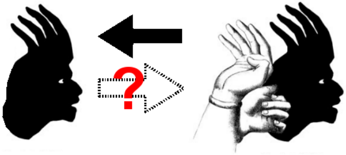
source: Wikimedia Commons

Challenges of Source Reconstruction: Head Modeling

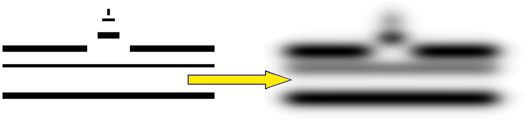
Development of realistic and individual head models for simulating the forward equations.



Challenges of Source Reconstruction: Inverse Problem

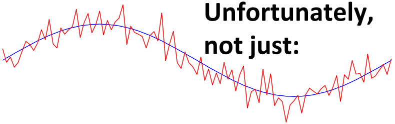


- (Presumably) under-determined



- Severely ill-conditioned

Unfortunately, not just:



- Signal is contaminated by a complex spatio-temporal mixture of external and internal noise and nuisance sources.

Demands on Source Reconstruction: Spatio-Temporal Aspects

- No principle limit of temporal resolution.
- Sampling rates up to 20000 Hz, i.e., timesteps of 0.05 ms.
- High sampling rates, $>$ 300 channels in combined EEG/MEG, long measurement times $\Longrightarrow$ Tons of data
- In principle, oversampling a temporal process gives useful additional information.
- However, the lowest temporal scale that contains valuable information is unknown.
- Spatio-temporal inversion can get tricky and computationally demanding.

Demands on Source Reconstruction: Group Studies

Neuroscientific studies with $n = 1$ subjects: Not really fancy.

Normally,

- Two matched groups A and B, each ~ 30 subjects.
- Different experimental conditions $C1$, $C2$, ...
- Collect data for all subjects and conditions
- Aim: Statistically significant inter-group differences.
- Problem: Large inter- and intra subject differences:
  - Individual head and cortex geometries;
  - Different SNRs;
  - Different cognitive constitution;

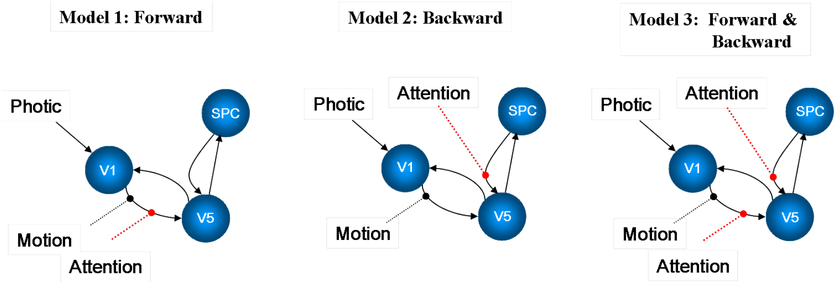
$\Longrightarrow$ A lot of variables and uncertainties.

Demands on Source Reconstruction: Multimodal Integration

- The brain is a complex bio-chemical information processing system.
- EEG/MEG measures only one correlate of "brain activity".
- Other imaging modalities measure other correlates.
- Multimodal integration sums up a lot of different approaches to fuse the different information.
- Might lead to big improvements of neuroimaging results
- Many open questions, the relation between the different correlates is subject to active research.

Demands on Source Reconstruction: Subsequent Analysis

- Result of source reconstruction: Temporal evolution of the spatial current distribution.
- Use these results to infer the causal architecture of the brain:
  - Structure of networks that pass and process information
  - Modulation of these networks
- Dynamical causal modeling (DCM): Bayesian model comparison procedure



Source: Andre C. Marreiros et al. (2010), Scholarpedia, 5(7):9568.

The Bayesian Approach in EEG/MEG Source Reconstruction

In summary, not the easiest problem, but a very interesting one!

Depending on the concrete application, we might have

- Plenty of variables and various sources of uncertainty
- Various sources and types of a-priori information
- Demand for statistical results and uncertainty quantification for subsequent processing.

The Bayesian approach seems appealing to deal with these issues!

Bayesian modeling: Determine priors and dependencies for all variables.

- Systematic approach due to the number of variables.
- Hierarchical Bayesian modeling (HBM): A specific modeling approach that emerged as a promising candidate for this.

A Complex Hierarchical Bayesian Model

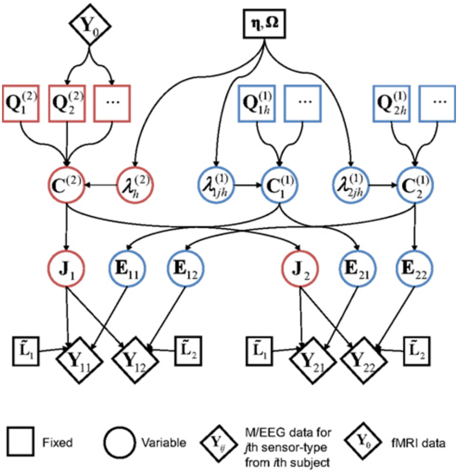
HBM for

- Multisubject
- Multimodal (EEG/MEG/fMRI)

Source reconstruction.

See Henson RN, Wakeman DG, Litvak V and Friston KJ (2011). A parametric empirical Bayesian framework for the EEG/MEG inverse problem: generative models for multi-subject and multi-modal integration. in *Frontiers in Human Neuroscience*, 5:76.

Multisubject fusion model



Source and sensor space

Outline

Cooperation with...

TAMPERE UNIVERSITY OF TECHNOLOGY

Dr. Sampsa Pursiainen
Department of Mathematics
Tampere University of Technology (TUT), Finland

Fachbereich 10 Mathematik und Informatik

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

IBB Institute for Biomagnetism and Biosignalanalysis

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

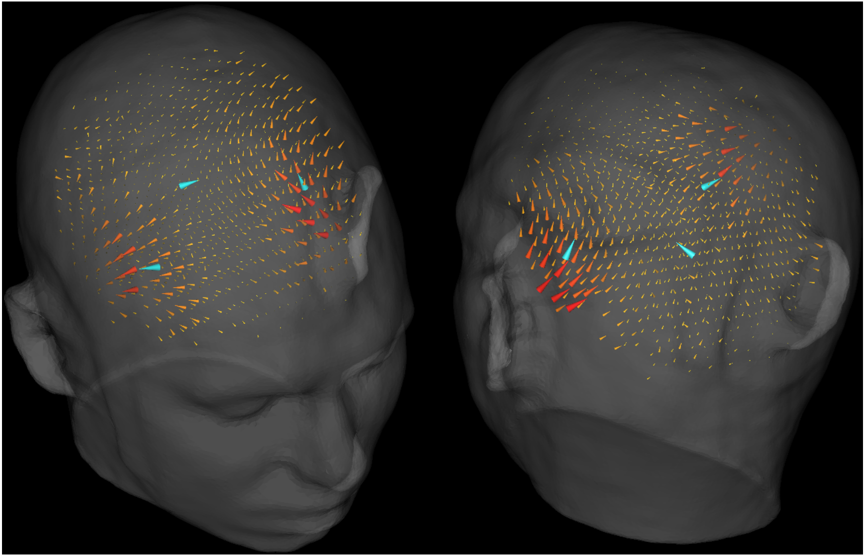
Background of the Talk

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.

Felix Lucka.
Hierarchical Bayesian Approaches to the Inverse Problem of EEG/MEG Current Density Reconstruction.
Diploma thesis in mathematics, University of Münster, March 2011

Current Density Reconstruction

Discretization of the underlying continuous current distribution by large number of elementary sources with fixed location and orientation.



Notation and Likelihood

Basic forward equation:

$$\boxed{b = L\, s}$$

- Up to now: Single time slice inversion;
- $b \in \mathbb{R}^m$: EEG/MEG measurements;
- $s \in \mathbb{R}^n$: Coefficients of the $d \in \{1, 3\}$ basic current sources at k different source locations; $n = d \cdot k$;
- $L \in \mathbb{R}^{m \times n}$: Lead-field matrix;

Likelihood:

$$p_{like}(b|s) \propto \exp\left(-\frac{1}{2}\|\Sigma_\varepsilon^{-1/2}\,(b - L\,s)\|_2^2\right)$$

Tasks and Problems for EEG/MEG in Presurgical Epilepsy Diagnosis

EEG/MEG in epileptic focus localization:

- *Focal epilepsy* is believed to originate from networks of focal sources.
- Active in inter-ictal spikes.
- Task 1: Determine number of focal sources (*multi focal epilepsy*?).
- Task 2: Determine location and extend of sources.

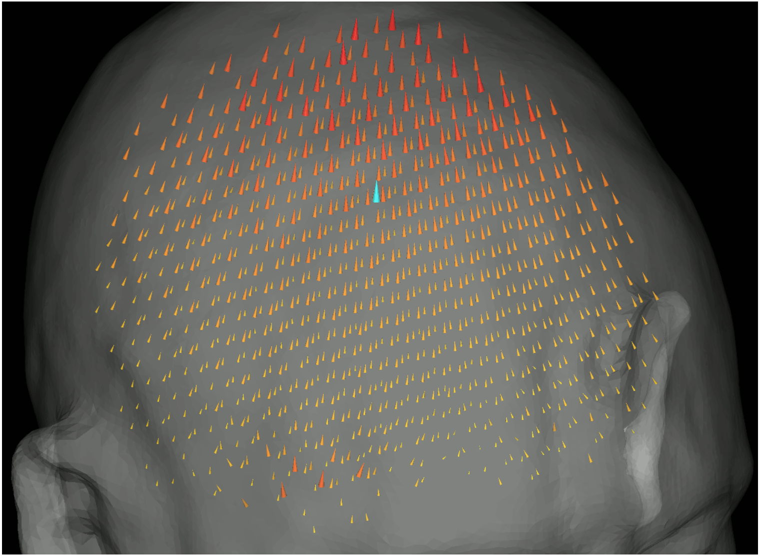
Problems of established CDR methods:

- Depth-Bias: Reconstruction of deeper sources too close to the surface.
- Masking: Near-surface sources “mask“ deep-lying ones.

Depth Bias: Illustration

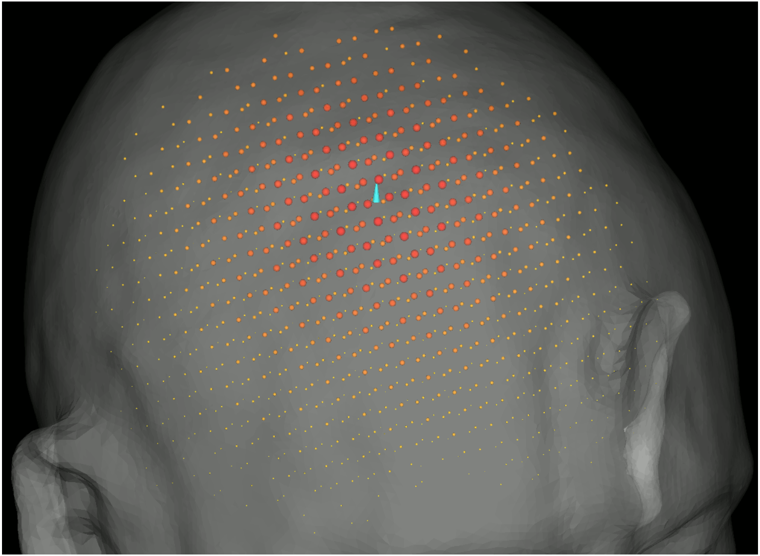
One deep-lying reference source (blue cone) and minimum norm estimate:

$$s_{\mathrm{MNE}} = \operatorname{argmin}\{\|\Sigma_{\varepsilon}^{-1/2}(b - L\,s)\|_2^2 + \lambda\|s\|_2^2\}$$



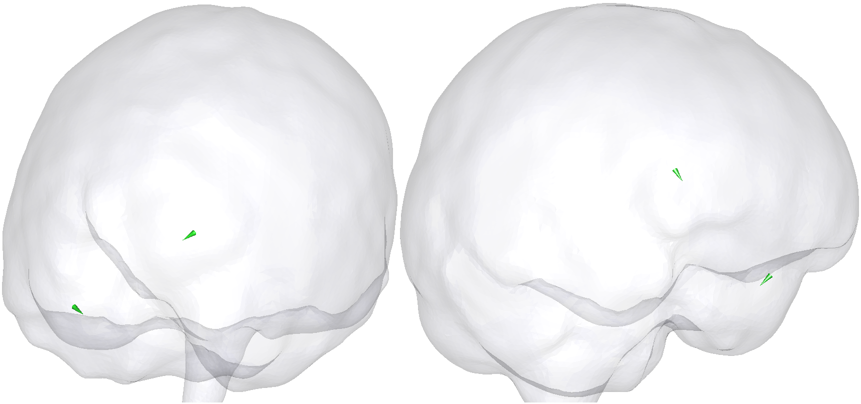
Depth Bias: Illustration

One deep-lying reference source (blue cone) and sLORETA result (Pascual-Marqui, 2002).



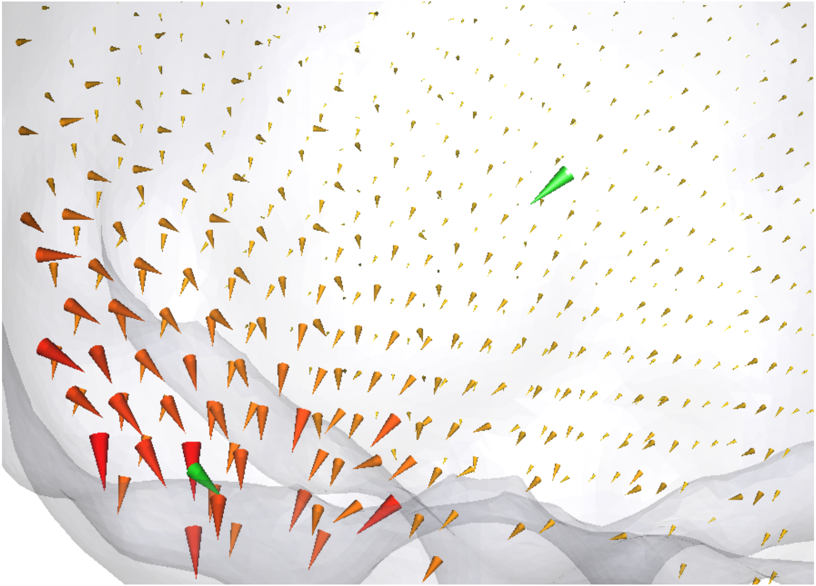
Masking: Illustration

Reference sources.



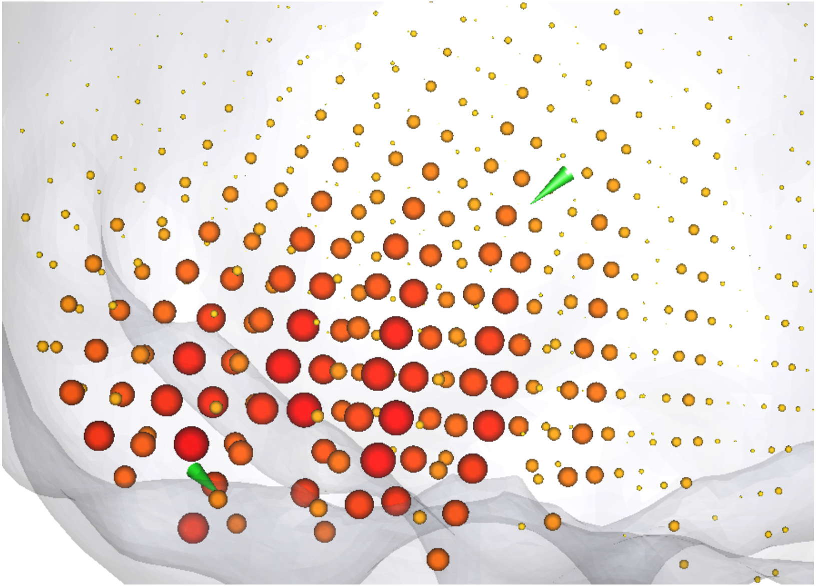
Masking: Illustration

MNE result and reference sources (green cones).



Masking: Illustration

sLORETA result and reference sources (green cones).



Problems of Classical Inverse Methods: Depth-Bias

- Using normal ℓ_2 and ℓ_1 type priors: MAP estimate has depth-bias.
- Heuristic reason: Deep sources have weaker signal; Signal of single deep source can be generated by extended patch of near-surface sources.
- Theoretical reason in simplified EEG example: $q \in \partial \mathcal{J}(s_{\text{MAP}})$ is harmonic function and, thus, fulfills maximum principle:
  - ℓ_2: s_{MAP} is harmonic $\Rightarrow$ maximum at boundary.
  - ℓ_1: sign of s_{MAP} is harmonic $\Rightarrow$ supported only at boundary.

Problems of Classical Inverse Methods: Depth-Bias

Introducing weighted norms ($\|s\|_2^2 \longrightarrow \|Ws\|_2^2$) to give deep sources an advantage.

- Partly solves depth-bias.
- Other draw-backs, e.g., larger spatial blurring $\Longrightarrow$ worse source separation.
- Critical from the Bayesian point of view: Would mean that deep sources usually have a stronger signal $\Longrightarrow$ unphysiological a-priori information.

Reweighting of the solution (e.g., sLORETA) also leads to problems w.r.t. source separation.

Starting Point for our Studies

D. Calvetti, H. Harri, S. Pursiainen, E. Somersalo, 2009.
Conditionally Gaussian hypermodels for cerebral source localization

- A specific hierarchical Bayesian model (HBM) aims to recover sparse source configurations.
- Calvetti et al., 2009 found promising first results for CM estimates for deep-lying sources and the separation of multiple (focal) sources.

Limitations of Calvetti et al., 2009 :

- (Full-) MAP estimates were not convincing; reason unclear.
- No systematic examination; only two source scenarios.
- Head models insufficient.

Contributions of our Studies

- Implementation of Full-MAP and Full-CM inference for HBM with realistic, high resolution Finite Element (FE) head models.
- Propose own algorithms for Full-MAP estimation.
- Examination of general properties, parameter choices, etc.
- Introduction of suitable performance measures for validation of simulation studies (Wasserstein distances).
- Systematic examination of performance concerning depth-bias and masking in simulation studies.

Outline

Gentle Introduction to Sparsity Promoting HBMs

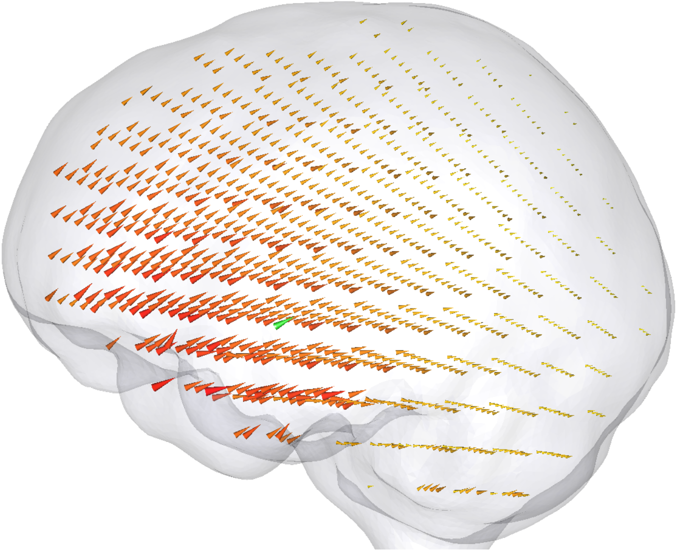
Wanted: A prior promoting sparse (focal) source activity.

First try:

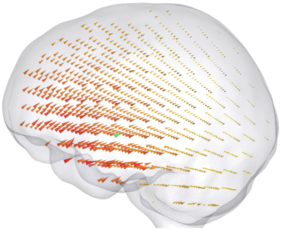
- Take Gaussian prior with zero mean and covariance $\Sigma_s = \gamma \cdot \mathrm{Id}, \quad \gamma > 0$ (*Minimum norm estimation*).
- Compute MAP or CM estimate (equal)!

$$\begin{aligned}\hat{s}_{\text{MAP}} :&= \underset{s \in \mathbb{R}^n}{\operatorname{argmax}} \left\{ \exp\left(-\frac{1}{2\,\sigma^2} \|b - \mathrm{L}\, s\|_2^2 - \frac{1}{2\gamma} \|s\|_2^2 \right) \right\} \\ &= \underset{s \in \mathbb{R}^n}{\operatorname{argmin}} \left\{ \|b - \mathrm{L}\, s\|_2^2 + \frac{\sigma^2}{\gamma} \|s\|_2^2 \right\}\end{aligned}$$

First try: NOT a focal reconstruction.



Gentle Introduction to Sparsity Promoting HBMs



What went wrong?

- Gaussian variables = characteristic scale given by variance. (not *scale invariant*)
- All sources have variance $\gamma \Longrightarrow$ Similar amplitudes are likely.

$\Longrightarrow$ Focal activity is very unlikely.

Gentle Introduction to Sparsity Promoting HBMs

Idea:

- Let sources at single locations i, $i = 1, \ldots, k$ have different variances γ_i.
- Let the data determine $\gamma_i \Longrightarrow$ New level of inference!
  - $\boldsymbol{\gamma} = (\gamma_i)_{i=1,\ldots,k}$ are called hyperparameters.
  - Bayesian inference: $\boldsymbol{\gamma}$ are random variables as well.
  - Their prior distribution $p_{hyper}(\boldsymbol{\gamma})$ is called hyperprior.

Gentle Introduction to Sparsity Promoting HBMs

Idea:

- Let sources at single locations i, $i = 1, \dots, k$ have different variances γ_i.
- Let the data determine $\gamma_i \Longrightarrow$ New level of inference!
  - $\boldsymbol{\gamma} = (\gamma_i)_{i=1,\dots,k}$ are called hyperparameters.
  - Bayesian inference: $\boldsymbol{\gamma}$ are random variables as well.
  - Their prior distribution $p_{hyper}(\boldsymbol{\gamma})$ is called hyperprior.
- Encode focality assumption into hyperprior:
  - Focality: Nearby sources should a-priori not be mutually dependent.
  - Focality: Most sources silent, few with large amplitude;
  - No location preference for activity should be given a priori.

Gentle Introduction to Sparsity Promoting HBMs

Idea:

- Let sources at single locations i, $i = 1, \ldots, k$ have different variances γ_i.
- Let the data determine $\gamma_i \Longrightarrow$ New level of inference!
  - $\boldsymbol{\gamma} = (\gamma_i)_{i=1,\ldots,k}$ are called hyperparameters.
  - Bayesian inference: $\boldsymbol{\gamma}$ are random variables as well.
  - Their prior distribution $p_{hyper}(\boldsymbol{\gamma})$ is called hyperprior.
- Encode focality assumption into hyperprior:
  - γ_i should be stochastically independent.
  - Focality: Most sources silent, few with large amplitude;
  - No location preference for activity should be given a priori.

Gentle Introduction to Sparsity Promoting HBMs

Idea:

- Let sources at single locations i, $i = 1, \ldots, k$ have different variances γ_i.
- Let the data determine $\gamma_i \Longrightarrow$ New level of inference!
  - $\boldsymbol{\gamma} = (\gamma_i)_{i=1,\ldots,k}$ are called hyperparameters.
  - Bayesian inference: $\boldsymbol{\gamma}$ are random variables as well.
  - Their prior distribution $p_{hyper}(\boldsymbol{\gamma})$ is called hyperprior.
- Encode focality assumption into hyperprior:
  - γ_i should be stochastically independent.
  - Sparsity inducing hyperprior, e.g., inverse gamma distribution.
  - No location preference for activity should be given a priori.

Gentle Introduction to Sparsity Promoting HBMs

Idea:

- Let sources at single locations i, $i = 1, \dots, k$ have different variances γ_i.
- Let the data determine $\gamma_i \Longrightarrow$ New level of inference!
  - $\boldsymbol{\gamma} = (\gamma_i)_{i=1,\dots,k}$ are called hyperparameters.
  - Bayesian inference: $\boldsymbol{\gamma}$ are random variables as well.
  - Their prior distribution $p_{hyper}(\boldsymbol{\gamma})$ is called hyperprior.
- Encode focality assumption into hyperprior:
  - γ_i should be stochastically independent.
  - Sparsity inducing hyperprior, e.g., inverse gamma distribution.
  - γ_i should be equally distributed.

Gentle Introduction to Sparsity Promoting HBMs

In formulas:

$$p_{prior}(s|\gamma) \sim \mathcal{N}\left(0, \Sigma_s(\gamma)\right), \quad \text{where} \quad \Sigma_s(\gamma) = \operatorname{diag}\left(\gamma_i \cdot Id_3, i = 1, \ldots, k\right)$$

$$p_{hyper}(\gamma) = \prod_{i=1}^{k} p^{i}_{hyper}(\gamma_i) = \prod_{i=1}^{k} p_{hyper}(\gamma_i) = \prod_{i=1}^{k} \frac{\beta^{\alpha}}{\Gamma(\alpha)} \gamma_i^{-\alpha-1} \exp\left(-\frac{\beta}{\gamma_i}\right)$$

$\alpha > 0$ and $\beta > 0$ determine *shape* and *scale*, $\Gamma(x)$ denotes the Gamma function.

Joint prior: $\quad p_{pr}(s,\gamma) = p_{prior}(s|\gamma)\, p_{hyper}(\gamma)$

Implicit prior: $\quad p_{pr}(s) = \int p_{prior}(s|\gamma)\, p_{hyper}(\gamma) d\gamma$

$$= \int \mathcal{N}(0, \Sigma_s(\gamma))\, p_{hyper}(\gamma) d\gamma \quad \rightsquigarrow \text{“Gaussian scale mixture”}$$

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

Gentle Introduction to Sparsity Promoting HBMs

Posterior, general:

$$p_{post}(s, \gamma | b) \propto p_{like}(b|s)\; p_{prior}(s|\gamma)\; p_{hyper}(\gamma)$$

$$\text{Comparison:} \quad p_{post}(s|b) \propto p_{like}(b|s)\; p_{prior}(s)$$

Posterior, concrete:

$$p_{post}(s, \gamma | b) \propto$$
$$\exp\left(-\frac{1}{2\sigma^2} \| b - L\,s \|_2^2 - \sum_{i=1}^{k} \left(\frac{\frac{1}{2} \| s_{i*} \|^2 + \beta}{\gamma_i} + \left(\alpha + \tfrac{5}{2}\right) \ln \gamma_i \right) \right)$$

Analytical advantages...

- Energy is quadratic with respect to s
- Factorizes over γ_i's.

and disadvantages...

- Energy is non-convex w.r.t. (s, γ) (posterior is multimodal)

Excursus: Full-, Semi-, and Approximate Inversion

Two types of parameters $\longrightarrow$ more possible ways of inference.

Full-MAP: Maximize $p_{post}(s, \gamma | b)$ w.r.t. s and γ.

Full-CM: Integrate $p_{post}(s, \gamma | b)$ w.r.t. s and γ.

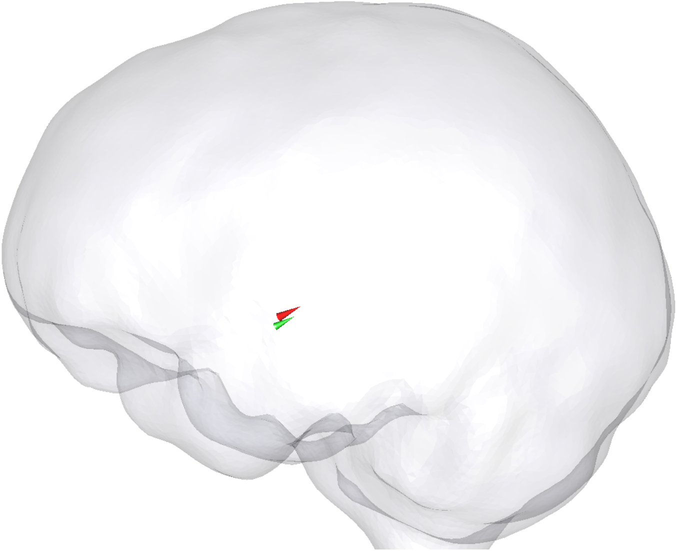
γ-MAP: Integrate $p_{post}(s, \gamma | b)$ w.r.t. s, and maximize over γ, first. Then use $p_{post}(s, \hat{\gamma}(b) | b)$ to infer s. (*Hyperparameter MAP/Empirical Bayes*)

S-MAP: Integrate $p_{post}(s, \gamma | b)$ w.r.t. γ, and maximize over s.

VB: Assume approximative factorization $p_{post}(s, \gamma | b) \approx \hat{p}_{post}(s | b)\, \hat{p}_{post}(\gamma | b)$; Approximate both with distributions that are analytically tractable.

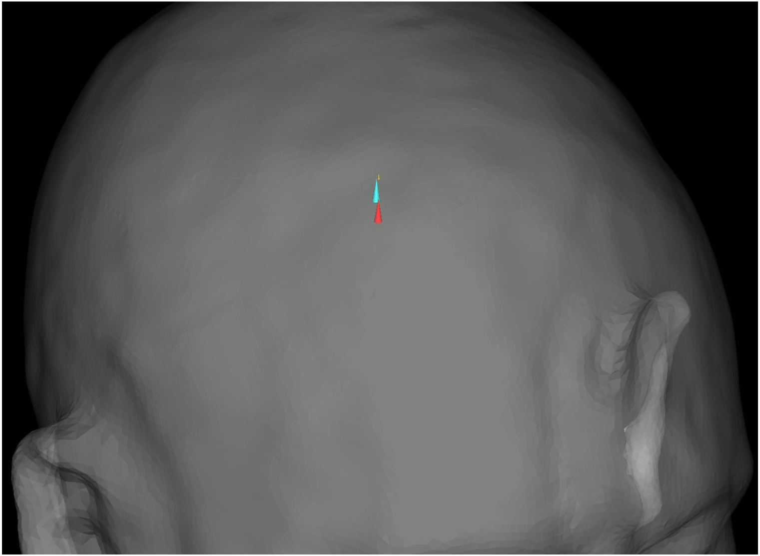
Focus of our work: Fully Bayesian inference.

Full-MAP estimate



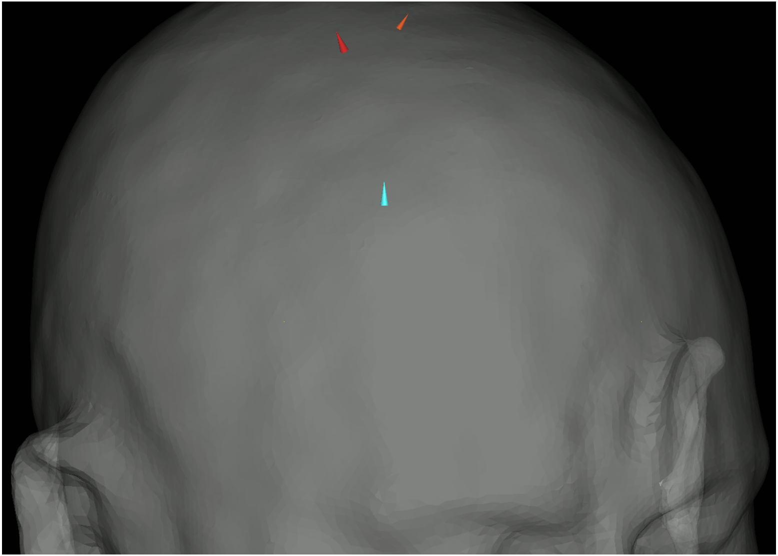
Results Depth Bias: Illustration

One deep-lying reference source (blue cone) and Full-CM result.



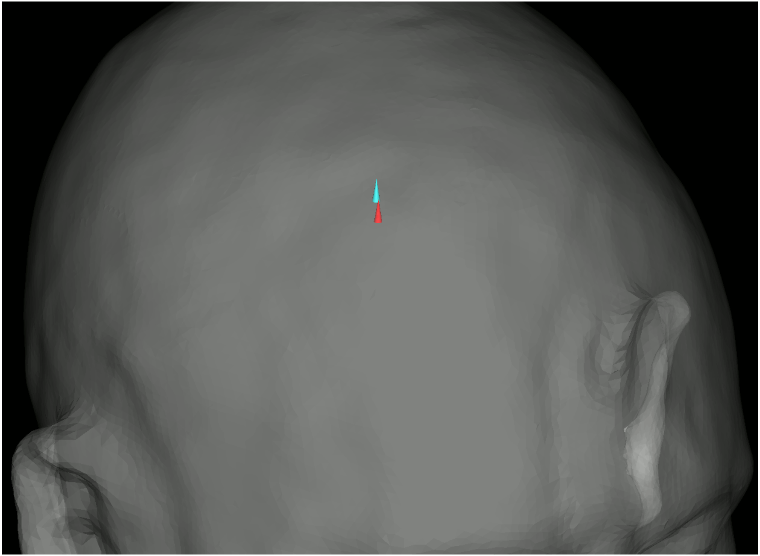
Results Depth Bias: Illustration

One deep-lying reference source (blue cone) and Full-MAP result proposed by Calvetti et al., 2009.



Results Depth Bias: Illustration

One deep-lying reference source (blue cone) and Full-MAP result proposed by us.

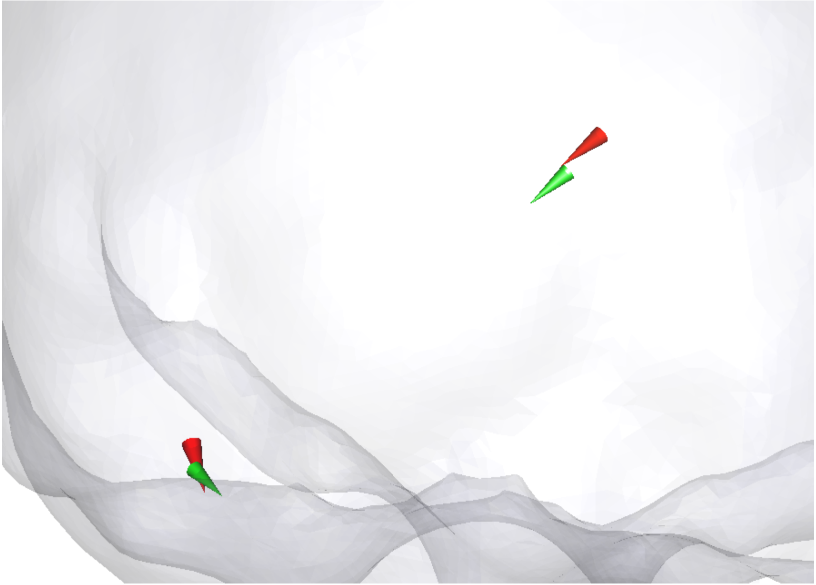


Results Masking: Illustration

Full-CM result and reference sources (green cones).

Results Masking: Illustration

Full-MAP result (by our algorithm) and reference sources (green cones).



Sparsity-Promoting HBM: Implicit Prior

Implicit prior on s is a Student's t-distribution on the (scaled) source amplitudes:

$$p(s) \propto \prod_{i=1}^{k} \left(1 + \frac{(s_i^{\text{amp}})^2}{2\beta}\right)^{-(\alpha+3/2)} = \prod_{i=1}^{k} \left(1 + \frac{t_i^2}{\nu}\right)^{-\frac{1}{2}(\nu+1)}$$

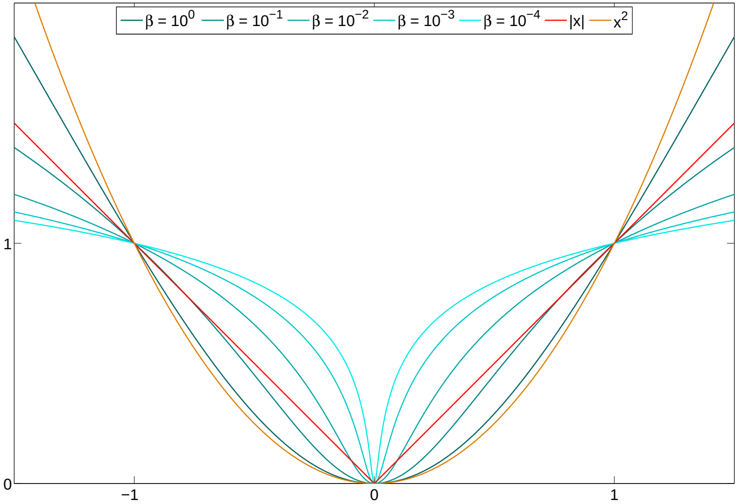
with $t_i = s_i^{\text{amp}}/\sqrt{\hat{\gamma}}$, $\hat{\gamma} = \beta/(\alpha+1)$, $\nu = 2(\alpha+1)$.

This corresponds to the regularization functional:

$$\mathcal{J}(s) = (2\alpha+3) \sum_{i=1}^{k} \log\left(1 + \frac{(s_i^{\text{amp}})^2}{2\beta}\right)$$

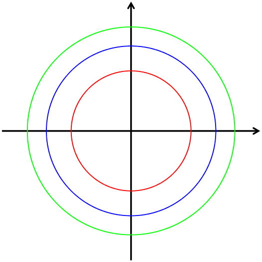
convex for $|x| < \sqrt{2\beta}$, concave for $|x| > \sqrt{2\beta}$.

Sparsity-Promoting HBM: Implicit Prior

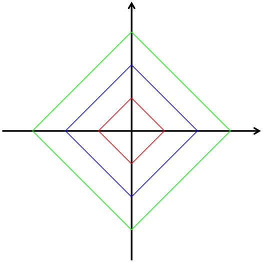


- J. Offtermatt & B. Kaltenbacher, 2011. *A convergence analysis of Tikhonov regularization with the Cauchy regularization term*.
- W. Alemie, 2010. *Regularization of the AVO inverse problem by means of a multivariate Cauchy probability distribution*.

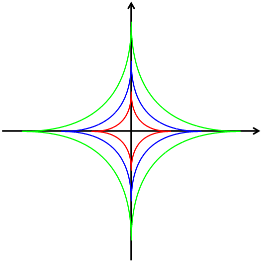
Isocontours of Different Priors



(a) ℓ_2



(b) ℓ_1



(c) $\ell_{(1/2)}$

Isocontours of Different Priors

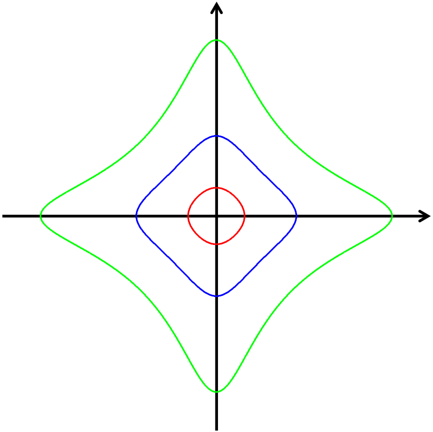


Figure: Students-t, $\nu = 3$

Strategies for Full-MAP Estimation and "Near-Mean" Estimates

- Optimization algorithm converges to the nearest mode $\Longrightarrow$ initialization is important.
- "Random initialization" (i.e. draw from hyperprior) does not help, it is not sparse enough.
- "Random sparse initialization": Combinatorial complexity.
- MCMC sampling suffers less from multi-modality: Initialization by γ_{CM}?
  - Full Near-Mean Estimates (NM): Initialize by γ_{CM}.
  - Heuristic for Full-MAP estimates: Initialize by various $\gamma_{\text{CM}}^{\text{approx}}$, pick the result with the highest probability.

Results:

- Initializations by $\gamma_{\text{CM}}^{\text{approx}}$ or γ_{CM} give higher posterior probability than uniform initialization.
- The reconstructions perform better.
- Full-NM vs. Full-MAP estimates? Not yet clear. Real data?
- Convert heuristic into proper reasoning? Alternative optimization schemes?

Outline

EEG/MEG Source Reconstruction: General Demands and Challenges

Depth Localization and Source Separation for Focal Sources

A Sparsity-Promoting Hierarchical Bayesian Model for EEG/MEG

Hierarchical Bayesian Modeling from an Empirical Bayesian Point of View

Take Home Messages & Conclusions

Empirical Bayesian Inference

Let the same data determine the prior used for the inference based on this data!

Empirical Bayesian Inference

Let the same data determine the prior used for the inference based on this data!

Sounds like...

Empirical Bayesian Inference

Let the same data determine the prior used for the inference based on this data!

Sounds like...

...but can be formulated into a consistent, statistical reasoning by adding a new dimension of inference: Hyperparameters and hyperpriors.

→ Parametric Empirical Bayesian inference

Top-down construction scheme → Hierarchical Bayesian modeling (HBM).

HBM as Empirical Bayesian Inference

$$p_{prior}(s|\gamma) \sim \mathcal{N}(0, \Sigma_s(\gamma)), \quad \text{where} \quad \Sigma_s(\gamma) = \mathrm{diag}(\gamma_i \cdot Id_3, i = 1, \ldots, k)$$

- Using the data to determine γ: Learn the prior from the data.
- Popular in machine learning.

Outline

Take Home Messages: EEG/MEG Source Reconstruction

- Measures one correlate of "brain activity" with a high temporal resolution.
- Various challenges and demands from the practical application.
- High uncertainty about various variables.
- Intrinsic challenges due to the spatial characteristics of the forward operator.

Take Home Messages: Hierarchical Bayesian Modeling

- Current trend in all areas of Bayesian inference.
- Extension of the prior model by hyperparameters γ and hyperpriors.
- Flexible framework for the construction of complex models with different levels for the embedding of different qualitative and quantitative a-priori information
- Gaussian w.r.t. s, factorization w.r.t. γ.
- Motivation 1: Capture the various variables and their dependencies in EEG/MEG in a systematic way.
- Motivation 2: Alternative formulation of sparse Bayesian inversion that has interesting features (no depth-bias, non-convex energy but the possibility to infer the support from CM estimate)
- Motivation 3: Use the element of adaptive, data-driven learning emphasized by the empirical Bayesian view.

Thank you for your attention!

D. Calvetti, H. Harri, S. Pursiainen, E. Somersalo, 2009.
Conditionally Gaussian hypermodels for cerebral source localization

D. Wipf and S. Nagarajan, 2009
A unified Bayesian framework for MEG/EEG source imaging.

F. L., S. Pursiainen, M. Burger and C.H. Wolters, 2012.
Hierarchical Bayesian Inference for the EEG Inverse Problem using Realistic FE Head Models: Depth Localization and Source Separation for Focal Primary Currents.