

A Hierarchical Bayesian Approach in Distributed MEG Source Modelling

Aapo Nummenmaa*, Toni Auranen*, Matti S. Hämäläinen†, Iiro P. Jääskeläinen*†, Jouko Lampinen*, Mikko Sams*, and Aki Vehtari*

* Laboratory of Computational Engineering, Helsinki University of Technology, Finland (correspondence: Aapo.Nummenmaa@hut.fi)

† Massachusetts General Hospital, Harvard Medical School, Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA, USA

Introduction

We present a hierarchical generalization of the widely used Gaussian prior (\sim MNE) for an alternative Bayesian approach to the MEG inverse problem.

- Minimum-Norm Estimate (MNE) is the most traditional method used in MEG/EEG distributed source model estimation [1].
- The Bayesian interpretation of MNE: Zero-mean Gaussian prior with fixed variance at all source locations.
- Adjusting the prior variance corresponds to the degree of regularization.
- MNE is computationally convenient as an explicit inverse operator is available.
- Drawback: MNE is known to produce rather diffuse estimates also in case of focal sources (see Figure 1).

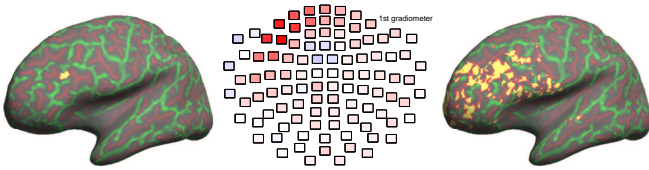


Figure 1: A simulated focal current source is depicted on the left. The corresponding simulated (gradiometer) MEG measurements are shown in the middle whereas the activity map produced by MNE is displayed on the right.

- We implemented a hierarchical generalization of the Gaussian prior: each of the source currents has a Gaussian prior with zero mean and individual variance.
- The source variances are then assumed to have a common inverse chi-square hyperprior distribution entailing a hierarchical nature to the model.
- The prior variances are considered as nuisance parameters and integrated out from the posterior distribution using Markov chain Monte Carlo (MCMC) methods.
- The limiting effect of such an integration is that it imposes a Student- t prior on the source currents.

Methods

The Forward Model

- The possible locations and orientations of the source currents were constrained by the individual cortical geometry [2]; the source space consisted of ~ 3000 points.
- A single-compartment boundary-element model comprising the inner skull surface was used in the MEG forward computations.
- Thus a linear observation model was obtained:

$$\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{n}, \quad (1)$$

where \mathbf{b} are the measured fields, \mathbf{A} is the forward matrix, \mathbf{s} is the distributed source and \mathbf{n} is noise.

- Simulated data were generated by adding Gaussian noise to the magnetic fields computed from simulated source using equation (1), giving rise to a Gaussian likelihood

$$\mathbf{b} - \mathbf{A}\mathbf{s} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}), \quad (2)$$

where \mathbf{C} is the noise covariance matrix.

- Data were created by using a finer ($\sim 30\,000$ points) discretization of the cortex and therefore the most obvious inverse crime was not committed.

The Hierarchical Prior

- The prior on the source current at cortical location i was

$$s_i \sim \mathcal{N}(0, V_i). \quad (3)$$

- An inverse chi-square hyperprior was imposed on the V_i 's:

$$V_i \sim \text{Inv-}\chi^2(\nu, \sigma^2). \quad (4)$$

- A uniform prior was assumed for $(\log \sigma, \nu^{1/3})$.
- A parameter expansion scheme [3, Ch. 11.8] was adopted in order to break the dependence between σ^2 and V_i 's and thereby improve mixing.

Posterior Simulation

- Fast Gibbs sampling was used to obtain numerical samples from the joint posterior distribution.
- An approximation was made: Non-diagonal entries of the conditional posterior covariance matrix corresponding to those source locations with almost zero variance V_i were assumed to be strictly zero.
- This accelerated the sampling drastically, by facilitating a block-diagonal matrix inversion, while leading to minimal numerical errors.

Results

- An interesting consequence of the hierarchical treatment is that the posterior distribution becomes multimodal (see Figure 2).

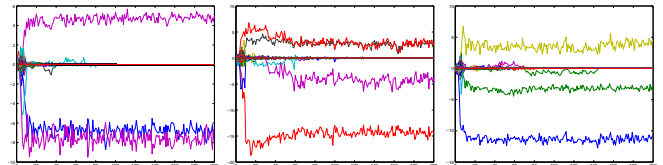


Figure 2: The trends of three different Markov chain simulation runs. Each chain converges to a different mode demonstrating the multimodality of the posterior distribution.

- Posterior modes represent different solutions of the source reconstruction problem (see Figure 3).

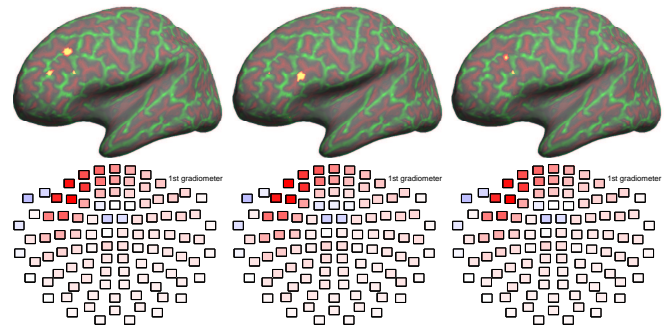


Figure 3: Activity maps representing some possible solutions of the inverse problem obtained from posterior simulations. All these solutions explain the data rather well; below each solution are the corresponding calculated measurements (compare with Figure 1).

- The number of different modes depends on the source and the dimension of the source space.
- Contrary to MNE, the hierarchical model leads to a heavy-tailed (Student- t distribution) prior and thus a "sparse" estimate with zero current at almost all locations.

Discussion

- Method seems to be especially suitable for focal source reconstructions.
- All parameters and hyperparameters are estimated from the data automatically, at the cost of increased computational load.
- Increasing the number of possible source locations leads to a more accurate representation of the individual cortical geometry, but also invalidates the assumption of *a priori* independent source strengths and leads to "overfitted" solutions.
- A spatial prior seems necessary if an accurate representation of the cortex plays significant role in the analysis (e.g., integration of fMRI and EEG/MEG).

References

- [1] Matti S. Hämäläinen and Risto J. Ilmoniemi. Interpreting measured magnetic fields of the brain: Estimates of current distributions. Technical Report TKK-F-A559, Helsinki University of Technology, Department of Technical Physics, 1984.
- [2] Anders M. Dale and Martin I. Sereno. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. *Journal of Cognitive Neuroscience*, 5(2):162–176, 1993.
- [3] Andrew Gelman, John B. Carlin, Hal S. Stern, and Donald B. Rubin. *Bayesian Data Analysis*. Chapman & Hall/CRC, second edition, 2003.