



HAL
open science

Mind the noise covariance when localizing brain sources with M/EEG

Denis Engemann, Daniel Strohmeier, Eric Larson, Alexandre Gramfort

► **To cite this version:**

Denis Engemann, Daniel Strohmeier, Eric Larson, Alexandre Gramfort. Mind the noise covariance when localizing brain sources with M/EEG. Pattern Recognition in Neuroimaging, 2015 International Workshop on, Jun 2015, Stanford, United States. hal-01183551

HAL Id: hal-01183551

<https://imt.hal.science/hal-01183551>

Submitted on 11 Aug 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Mind the noise covariance when localizing brain sources with M/EEG

Denis Engemann^{*†}, Daniel Strohmeier[‡], Eric Larson[§], Alexandre Gramfort^{†¶}

^{*}Neuropsychology & Neuroimaging team, INSERM UMRS 975, ICM, Paris

Email: denis.engemann@cea.fr

[†]NeuroSpin, CEA Saclay, Bat. 145, Gif-sur-Yvette Cedex, France

[‡]Institute of Biomedical Engineering and Informatics, Technische Universität Ilmenau, Ilmenau, Germany

[§]Institute for Learning and Brain Sciences, University of Washington, Seattle, WA, USA

[¶]Institut Mines-Telecom, Telecom ParisTech, CNRS LTCI, Paris, France

Abstract—Magnetoencephalography (MEG) and electroencephalography (EEG) are imaging methods that measure neuronal dynamics non invasively with high temporal precision. It is often desired in MEG and EEG analysis to estimate the neural sources of the signals. Strategies used for this purpose often take into account the covariance between sensors to yield more precise estimates of the sources. Here we investigate in greater detail how the quality of such covariance estimates conditions the estimation of MEG and EEG sources. We investigated three distinct source localization methods: dynamic Statistical Parametric Maps (dSPM), the linearly constrained minimum variance (LCMV) beamformer and Mixed-Norm Estimates (MxNE). We implemented and evaluated automated strategies for improving the quality of covariance estimates at different stages of data processing. Our results show that irrespective of the source localization method, accuracy can suffer from improper covariance estimation but can be improved by relying on automated regularization of covariance estimates.

Keywords—MEG; EEG; covariance estimation; source localization; statistical learning.

I. INTRODUCTION

Magnetoencephalography (MEG) and electroencephalography (EEG) allow imaging brain dynamics non invasively at a millisecond time scale. To image brain activity, sensor signals are projected onto a cortical model using inverse solvers. These solvers typically involve Euclidean L2 constraints which inherently assume Gaussian noise with zero mean and equal variances across sensors and therefore necessitate a spatial whitening step for input data [1]. The success of source localization thus depends on the quality of the covariance estimate from which the whitening is computed. Covariances, however, are subject to estimation errors when the number of samples is not sufficient. To address this problem, the covariance estimation can be regularized. The most common approach to regularize the covariance is to down-weight off-diagonal matrix coefficients, while more advanced regularization methods are based on shrinkage techniques [2] or generative models with low rank

assumptions: probabilistic PCA (PPCA) and factor analysis (FA) [3], [4]. Through the use of cross-validation, all of these models can be tuned and compared based on Gaussian likelihood computed on unseen data [1].

Here we investigate in greater detail the impact of proper regularization on source estimation obtained from three distinct inverse solvers: dynamic Statistical Parametric Maps (dSPM) [5], the linearly constrained minimum variance (LCMV) beamformer [6] and Mixed-Norm Estimates [7]. Each of these methods makes use of covariance estimates in different ways.

II. MATERIALS AND METHODS

A. Datasets

We used two datasets. The SPM-faces MEG dataset ¹ was recorded with a VSM MedTech Inc. whole-head axial gradiometer system with 274 channels (CTF/VSM) at the Functional Imaging Laboratory, London. The paradigm is detailed in [8]. The auditory sample dataset is shipped with the MNE software [9] and includes combined M/EEG recordings. This dataset was acquired using a Neuromag VectorView whole-head system with 306 channels (Elekta Neuromag, Finland), which are formed from 102 sensor triplets, each comprising two orthogonal planar gradiometers and one magnetometer. The protocol is detailed in [9].

B. Covariance estimation

The automated approach to covariance estimation is detailed in [1]. Here we used its implementation that is provided by the MNE software [9]. The best and the worst estimates are obtained out of a series of alternative strategies using cross-validation and log-likelihood computation. The methods compared were Factor Analysis (FA), Shrunken Covariance (SC) diagonal regularization, and unregularized, empirical covariance [1].

We evaluated one additional strategy for improving spatial whitening obtained from the covariance, which is based on empirical rank estimation and singular value decomposition

This work was supported by the ANR grant THALAMEEG, ANR-14-NEUC-0002-01, and the German Federal Ministry of Education and Research (BMBF) (3IPT605A)

¹<http://www.fil.ion.ucl.ac.uk/spm/data/mmfaces/>

(SVD) truncation. Spatial whitening amounts to multiplying the data by the square root of the inverse of the noise covariance [1], which is commonly computed using SVD. When the covariance is rank deficient, numerical instabilities appear in this inverse. Inverse solvers such as MNE must thus truncate small singular values obtained from the covariance matrix. The number of remaining singular values is the effective numerical rank of the data. This quantity can either be derived from the data processing history or estimated directly from the data. When the rank is estimated, a singular-value truncation threshold must be chosen. The result therefore depends on the scaling of the data. As MEG and EEG measurements data have different units and scales, how the data are rescaled before joint modeling ultimately affects the computed rank. As reducing the estimated rank achieves a form of regularization, choosing the rescaling factor can be used to stabilize source localization results.

Here we implemented an empirical rank estimator based on SVD. In a first step, this rank estimator was calibrated such that it provides correct rank estimates when using the empirical unregularized covariance. Rescaling factors were selected to transform all channel types to approximately unity scale (1×10^{15} for magnetometers, 1×10^{13} for gradiometers, and 1×10^6 for EEG). SVD values were then discarded if they were smaller than 1×10^{-5} . Subsequent tests² established that rank estimates matched the expected rank across differently processed data sets. The expected rank was computed as the number of channels minus the number of components that were removed from the decomposition of the data, e.g., when applying Signal Space Separation (SSS) [10]. To investigate how automated rank estimation can be used for regularization, smaller rescaling factors were also chosen (1×10^{12} for magnetometers, 1×10^{11} for gradiometers and 1×10^5 for EEG). For these values, the empirical covariance no longer produced the correct rank estimates. If the covariance is not regularized, the mismatch between the expected and the estimated rank increases as fewer trials are used. In other words, more singular values will be discarded, and more regularization will be applied when computing the whitening matrix. To investigate this relationship more closely, rank estimates were computed on the SPM faces dataset when varying the numbers of samples and the scaling factors.

C. Distributed inverse solvers

We used the SPM-faces MEG dataset to test distributed solvers. Data were bandpass filtered between 1-30 Hz using a zero-phase 4th order Butterworth filter. Baseline segments between -200 to 0 ms were used to estimate the noise covariance. The data covariance used by the LCMV beamformer

²The procedure is documented in the unit tests of the covariance code in MNE-Python. See functions `test_rank` and `test_cov_scaling`, https://github.com/mne-tools/mne-python/blob/master/mne/tests/test_cov.py

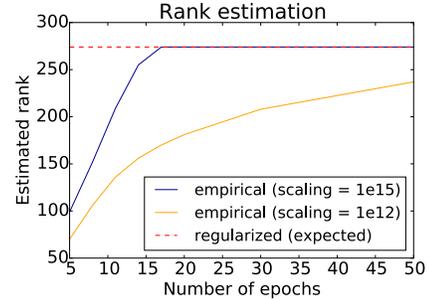


Figure 1. Rank estimation across epochs using different scaling factors. The dotted line indicates the expected rank. Here the expected rank equals the number of channels in the dataset (274). The regularized covariance estimators yielded the expected rank estimates. Rank estimation based on the empirical covariance led to lower than expected rank estimates where the discrepancy decreased as a function of the number of samples available. This trend was further modulated by the scaling factor: with a smaller rescaling more samples were required to produce a more accurate estimate.

was computed from 50 to 350 ms, reflecting a conservative choice regarding the face-selective evoked components. Source estimates were separately computed for the faces and the scrambled faces condition using MNE with dSPM noise normalization and LCMV [5], [6]. A paired contrast of the form `dSPM_faces - dSPM_scrambled` was then computed with positive values reflecting face-selective activation. This analysis was conducted over varying numbers of epochs using either the best or the worst covariance estimator, as defined by the cross-validated negative log-likelihood [1]. The worst estimate was consistently the empirical covariance, while the best was either SC and FA. Pearson correlations were then computed between activation maps based on the worst and the best estimator at a given number of epochs and then between each of these activation maps and the ones obtained from processing all available data. M/EEG data were processed using MNE [9]. For cross-validation and model selection of covariance parameters, the scikit-learn library for machine learning [11] was used. The computation of inverse solutions was based on fixed rank values and on the automated rank estimator, which is currently used as default in the development version of MNE-Python [12].

To evaluate the effect of noise covariance estimation on source estimates obtained with sparse solvers, we used the MNE auditory sample dataset. Data were processed in the same manner as the SPM face dataset. We used the iteratively reweighted mixed-norm inverse solver (IR-MxNE) [13], which promotes focal sources while enforcing temporal consistency. Instead of a contrast, source estimates for a single auditory condition were computed. To avoid redundancy, the best and the worst covariance estimators were directly compared for a reduced number of samples. The computation of inverse solutions was only based on the empirical rank estimator.

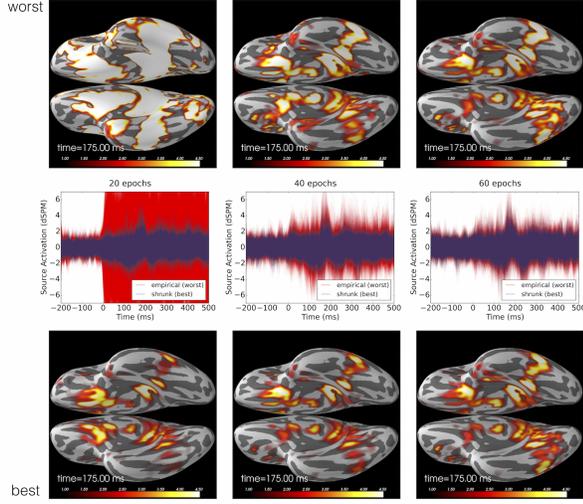


Figure 2. From left to right dSPM results with various numbers of epochs used for noise covariance estimation. Worse covariance (empirical) is compared to best regularized covariance. The mid-row represents average temporal dynamics for worst and best estimators superimposed. The lines refer to the average signal across vertices, the areas depict the standard deviation across vertices.

III. RESULTS

A. Empirical rank estimation

The rank estimates across varying numbers of samples and scaling factors are presented in Figure 1. The lower scaling factor consistently lead to stronger mismatches between the expected and the estimated rank as a function of the number of samples available. The lower scaling factor should therefore lead to stronger regularization when computing spatial whitening.

B. Distributed inverse solvers

For both MNE and LCMV, source amplitudes were negatively correlated to the number of samples if no regularization was applied to the covariance. (cf. Figure 2 and 3). If regularization was applied, the source amplitudes remained more similar across varying numbers of samples (cf. Figure 3). Third, the similarity between the worst and the best solution increased as a function of the number of epochs available (cf. Figure 4, panels (A) and (B)). Generally, the discrepancy between the worst and the best solution was higher for LCMV compared to MNE/dSPM with fewer numbers of samples. In other words, cross-validated covariance estimates yield more stable results and LCMV suffered more than MNE from the availability of limited data. Optimization benefits consistently seemed negligible beyond 50 epochs of data. The impact of improper regularization was mitigated if the whitening was based on the empirical rank estimation. Interestingly, the difference between results using the worst and the best covariance estimate at a given number of

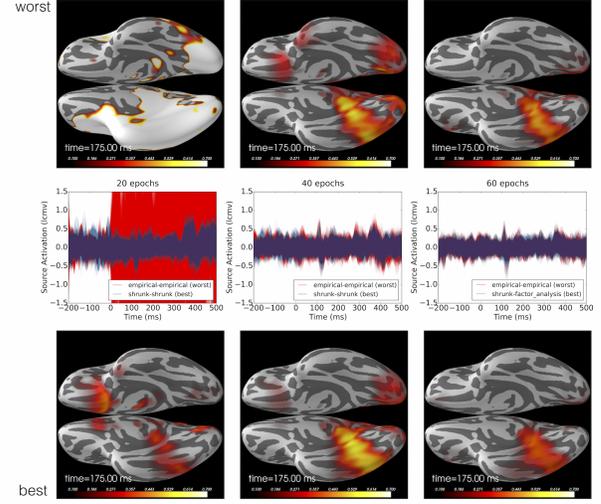


Figure 3. From left to right LCMV results with various numbers of epochs used for noise covariance estimation. Worse covariance (empirical) is compared to best regularized covariance. The mid-row represents average temporal dynamics for worst and best estimators superimposed. The lines refer to the average signal across vertices, the areas depict the standard deviation across vertices.

samples did not visibly change for LCMV beamformers when the empirical covariance was used (cf. Figure 4, panels (C) and (D)).

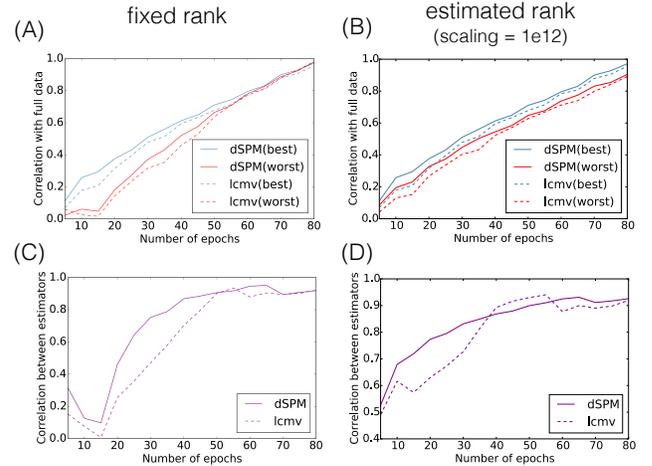


Figure 4. Correlations between MNE and LCMV sources estimates based on the best versus the worst covariance estimator. Panels (A) and (B) show results for a fixed (expected) rank value. Panels (C) and (D) show results based on the estimated rank using a reduced scaling factor. (A) and (C) depict the correlation between the source estimate based on a reduced number of epochs with the source estimate based on all epochs. Panels (B) and (D) show the correlation between the corresponding alternative source estimates at a given number of samples. In all cases, correlation increases as function of the number of samples available.

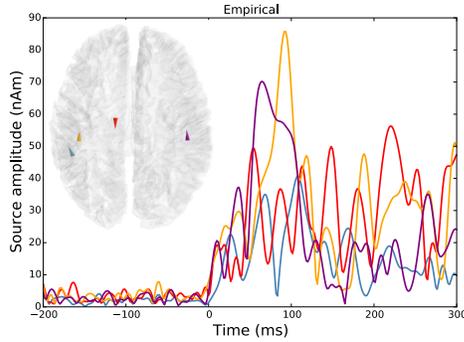


Figure 5. IR-MxNE with 15 epochs using empirical covariance with automated rank estimation (scaling = 1×10^{12}). One observes spurious active sources.

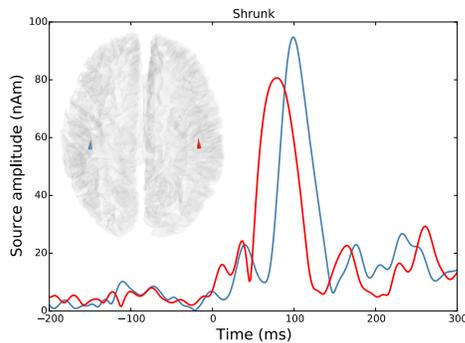


Figure 6. IR-MxNE with 15 epochs using best covariance with automated rank estimation (scaling = 1×10^{12}). One observes two expected active sources in bilateral auditory cortices.

C. Sparse inverse solvers

Time courses and source locations based on the IR-MxNE estimates are presented in 5 and 6. Although this analysis was based on the empirical regularized rank estimates, it suggested additional unexpected sources, beyond bilateral auditory cortices, when the worst covariance estimate was used. The corresponding time-courses were characterized by increased noise.

IV. CONCLUSION

Results demonstrate that proper regularization of covariance estimates stabilize source localization results. In particular our findings suggest that beamformers and sparse source estimates are more affected by incorrect covariance estimation and thus benefit more greatly from automated regularization. Importantly, while simpler models such as MNE/dSPM can be corrected post-hoc based on SVD truncation, complex models clearly benefit from covariance estimates that were optimized in the first place. The variance-reducing effect of automated regularization [1] can be said generalize across different inverse solvers and hence is expected to improve analysis of infrequent events, and comparisons across subjects in a general way.

REFERENCES

- [1] D. A. Engemann and A. Gramfort, "Automated model selection in covariance estimation and spatial whitening of MEG and EEG signals," *NeuroImage*, vol. 108, no. 0, pp. 328 – 342, 2015.
- [2] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of Multivariate Analysis*, vol. 88, no. 2, pp. 365 – 411, 2004.
- [3] M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 61, no. 3, pp. 611–622, 1999.
- [4] D. Barber, *Bayesian reasoning and machine learning*. Cambridge University Press, 2012.
- [5] A. Dale, A. Liu, B. Fischl, and R. Buckner, "Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity," *Neuron*, vol. 26, pp. 55–67, 2000.
- [6] B. V. Veen, W. V. Drongelen, M. Yuchtman, and A. Suzuki, "Localization of brain electrical activity via linearly constrained minimum variance spatial filtering," *Biomedical Engineering, IEEE Transactions on*, vol. 44, no. 9, pp. 867–880, Jan 1997.
- [7] A. Gramfort, M. Kowalski, and M. S. Hämläinen, "Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods." *Phys. Med. Biol.*, vol. 57, no. 7, pp. 1937–1961, Apr. 2012.
- [8] R. Henson and M. Rugg, "Neural response suppression, haemodynamic repetition effects, and behavioural priming," *Neuropsychologia*, vol. 41, no. 3, pp. 263–270, 2003.
- [9] A. Gramfort, M. Luessi, E. Larson, D. Engemann, D. Strohmeier, C. Brodbeck, L. Parkkonen, and M. Hämläinen, "MNE software for processing MEG and EEG data," *NeuroImage*, vol. 86, no. 0, pp. 446 – 460, 2014.
- [10] S. Taulu, J. Simola, and M. Kajola, "Applications of the signal space separation method," *IEEE Trans. Signal Proc.*, vol. 53, no. 9, pp. 3359–3372, 2005.
- [11] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [12] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. Hämläinen, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, no. 267, 2013.
- [13] D. Strohmeier, J. Haueisen, and A. Gramfort, "Improved meg/eeeg source localization with reweighted mixed-norms," in *Pattern Recognition in Neuroimaging, 2014 International Workshop on*. IEEE, 2014, pp. 1–4.