The iterative reweighted Mixed-Norm Estimate for spatio-temporal MEG/EEG source reconstruction

Daniel Strohmeier, Yousra Bekhti, Jens Haueisen, Alexandre Gramfort

The iterative reweighted Mixed-Norm Estimate for spatio-temporal MEG/EEG source reconstruction

Daniel Strohmeier, Jens Haueisen, and Alexandre Gramfort

***Abstract*—Source imaging based on magnetoencephalography (MEG) and electroencephalography (EEG) allows for the non-invasive analysis of brain activity with high temporal and good spatial resolution. As the bioelectromagnetic inverse problem is ill-posed, constraints are required. For the analysis of evoked brain activity, spatial sparsity of the neuronal activation is a common assumption. It is often taken into account using convex constraints based on the l_1-norm. The resulting source estimates are however biased in amplitude and often suboptimal in terms of source selection due to high correlations in the forward model. In this work, we demonstrate that an inverse solver based on a block-separable prior with a Frobenius norm per block and a $l_{0.5}$-quasinorm over blocks addresses both of these issues. For solving the resulting non-convex optimization problem, we propose the iterative reweighted Mixed Norm Estimate (irMxNE), i.e., an optimization scheme based on iterative reweighted convex surrogate optimization problems, which are solved efficiently using a block coordinate descent scheme and an active set strategy. We provide empirical evidence based on simulations and analysis of MEG data that the proposed method improves on MxNE in terms of amplitude bias and support recovery.**

***Index Terms*—Electrophysical imaging, brain, inverse methods, magnetoencephalography (MEG), electroencephalography (EEG), sparse structured priors.**

I. Introduction

Source imaging with magnetoencephalography (MEG) and electroencephalography (EEG) delivers insights into the active brain with high temporal and good spatial resolution in a non-invasive way. It is based on solving the bioelectromagnetic inverse problem, which is a high dimensional ill-posed regression problem. In order to render its solution unique, constraints have to be imposed reflecting a priori assumptions on the neuronal sources. In the past, several non-linear source reconstruction methods were proposed, which have in common to favor sparse focal source configurations to explain the MEG/EEG signals [1]–[5]. Among these is regression with l_1-norm regularization, known as LASSO in statistics [6] and Minimum Current Estimate (MCE) in the MEG/EEG literature [1], [7]. It is a convex surrogate for the optimal, but NP hard regularized regression problem with a l_0-norm penalty. Due to its convexity, this approach allows for fast algorithms with guaranteed global convergence. However, the resulting source estimates are biased in amplitude and often suboptimal in terms of support recovery [8]. In the case of MEG/EEG, it is impeded particularly by the spatial correlation of the forward model. In contrast, regularized regression based on non-convex penalties, such as logarithmic or l_p-quasinorm penalties with $0 < p < 1$, yields sparser and less biased estimates than a standard LASSO [8] at the expense of convexity. The resulting optimization problem can be solved e.g. by iterative reweighted LASSO [8], [9].

Applying the aforementioned approaches to MEG/EEG inverse problems results in computing the source estimate for each time point independently discarding the temporal characteristics of neuronal activation. To estimate sources at multiple time points jointly, the Mixed-Norm Estimate (MxNE) was proposed in [10]. It is a variant of Group LASSO [11], i.e., regularized regression with a block-separable $l_{2,1}$-mixed-norm penalty, adapted for MEG/EEG inverse problems. In MxNE, each block represents the source activation at a specific source location. A Frobenius norm per block imposes stationarity on the source estimates, while a l_1-norm penalty over blocks promotes spatial sparsity [10], [12]. Due to this l_1-norm, MxNE estimates suffer however from the same problems as the standard LASSO in terms of amplitude bias and support recovery.

To address these limitations, we propose a MEG/EEG inverse solver based on a non-convex block-separable regularization functional, in which the l_1-norm used in MxNE is replaced by a $l_{0.5}$-quasinorm [13]. A preliminary version of this work was presented in [14]. In this paper, we start by introducing sparse MEG/EEG source imaging based on the proposed block-separable penalty and then propose an iterative optimization strategy based on reweighted MxNE to solve the non-convex optimization problem. We show that each weighted MxNE iteration can be solved efficiently by combining a block coordinate descent scheme [13], [15], [16] and an active set strategy with convergence controlled by means of the primal-dual gap [10], which is applicable to MEG/EEG inverse problems with and without orientation constraints. Finally, we provide empirical evidence using simulations and analysis of two experimental MEG data sets that the proposed method outperforms MxNE in terms of active source identification and amplitude bias, while running in a few seconds.

This work was supported by the German Research Foundation (DFG) (Ha 2899/14-1), and by the "Gaspard Monge Program for Optimization and operations research" offered by EDF and the Jacques Hadamard Mathematical Foundation (FMJH).

D. Strohmeier is with the Institute of Biomedical Engineering and Informatics, Technische Universiät Ilmenau, Ilmenau, Germany; e-mail: daniel.strohmeier@tu-ilmenau.de

J. Haueisen is with the Institute of Biomedical Engineering and Informatics, Technische Universiät Ilmenau, Ilmenau, Germany and the Biomagnetic Center, Department of Neurology, Jena University Hospital, Jena, Germany; e-mail: jens.haueisen@tu-ilmenau.de

A. Gramfort is with the Institut Mines-Telecom, Telecom ParisTech, CNRS LTCI, Paris, France and the NeuroSpin, CEA Saclay, Bat. 145, Gif-sur-Yvette Cedex, France; e-mail: alexandre.gramfort@telecom-paristech.fr

Notation: We mark vectors with bold letters, $\mathbf{a} \in \mathbb{R}^N$, and matrices with capital bold letters, $\mathbf{A} \in \mathbb{R}^{N \times M}$. The transpose of a vector or matrix is denoted by $\mathbf{a}^T$ and $\mathbf{A}^T$. The scalar $\mathbf{a}[i]$ is the i$^{\text{th}}$ element of $\mathbf{a}$. $\mathbf{A}[i,:]$ corresponds to the i$^{\text{th}}$ row and $\mathbf{A}[:,j]$ to the j$^{\text{th}}$ column of $\mathbf{A}$. $\|\mathbf{A}\|_{\text{Fro}}$ indicates the Frobenius norm, and $\|\mathbf{A}\|$ the spectral norm of a matrix.

II. Materials and Methods

A. The inverse problem

The MEG/EEG forward problem describes the linear relationship between the MEG/EEG measurements $\mathbf{M} \in \mathbb{R}^{N \times T}$ (N number of sensors, T number of time instants) and the source activation $\mathbf{X} \in \mathbb{R}^{(SO) \times T}$ (S number of source locations, O degrees of freedom of the current dipole per source location. O is typically 3 when source orientation is unknown or 1 when orientation is postulated, e.g. using the cortical constraint [17]). The model then reads:

$$\mathbf{M} = \mathbf{G}\mathbf{X} + \mathbf{E} \ , \tag{1}$$

where $\mathbf{G} \in \mathbb{R}^{N \times (SO)}$ is a known instantaneous mixing matrix referred to as the gain matrix, which links source and sensor signals, and $\mathbf{E}$ is the measurement noise, which is assumed to be additive, white, and Gaussian, $\mathbf{E}[:,j] \sim \mathcal{N}(0, \mathbf{I})$ for all j. This assumption is acceptable on the basis of a proper spatial whitening of the data using an estimate of the noise covariance [18]. As $SO \gg N$, the MEG/EEG inverse problem is ill-posed and constraints have to be imposed on the source activation matrix $\mathbf{X}$ to render the solution unique. By partitioning $\mathbf{X}$ into S blocks $\mathbf{X}_s \in \mathbb{R}^{O \times T}$, where each $\mathbf{X}_s$ represents the source activation at a specific source location s, we can apply a block-separable penalty term $\mathcal{P}(\mathbf{X})$ combining a Frobenius norm per block and a $l_{0.5}$-quasinorm penalty over blocks. The optimization problem reads:

$$\begin{aligned} \widehat{\mathbf{X}} &= \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{M} - \mathbf{G}\mathbf{X}\|_{\text{Fro}}^2 + \mathcal{P}(\mathbf{X}) \\ \widehat{\mathbf{X}} &= \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{M} - \mathbf{G}\mathbf{X}\|_{\text{Fro}}^2 + \lambda \sum_{s=1}^{S} \sqrt{\|\mathbf{X}_s\|_{\text{Fro}}} \ , \end{aligned} \tag{2}$$

where $\lambda > 0$ is the regularization parameter balancing the data fit and penalty term. Similar to the constraint applied in MxNE [10], $\mathcal{P}(\mathbf{X})$ promotes source estimates with only a few focal sources that have non-zero activations during the entire time interval of interest. Here, the Frobenius norm per $\mathbf{X}_s$, combining l_2-norm penalties over time and orientation as proposed in [2], [7], [12], imposes stationarity, while the $l_{0.5}$-quasinorm penalty promotes spatial sparsity.

B. Iterative reweighted Mixed Norm Estimate

The proposed block-separable regularization functional is an extension of the $l_{2,0.5}$-quasinorm penalty used for regularized regression in [8], [9], [13]. These works showed, based on the framework of Difference of Convex functions programming or Majorization-Minimization algorithms, that the resulting optimization problems can be solved by iteratively solving convex surrogate optimization problems. Each iteration is equivalent to a weighted Group LASSO with weights being defined by the previous estimate. For minimizing Eq. (2), we hence propose the iterative reweighted MxNE (irMxNE) optimization scheme. Given $\widehat{\mathbf{X}}^{(k-1)}$, the source estimate obtained in the (k-1)$^{\text{th}}$ iteration, the $\widehat{\mathbf{X}}^{(k)}$ is computed using Eq. (3).

$$\begin{aligned} \widehat{\mathbf{X}}^{(k)} &= \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{M} - \mathbf{G}\mathbf{X}\|_{\text{Fro}}^2 + \lambda \sum_s \frac{\|\mathbf{X}_s\|_{\text{Fro}}}{2\sqrt{\left\|\widehat{\mathbf{X}}_s^{(k-1)}\right\|_{\text{Fro}}}} \\ &= \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{M} - \mathbf{G}\mathbf{X}\|_{\text{Fro}}^2 + \lambda \sum_s \frac{1}{\mathbf{w}^{(k)}[s]}\|\mathbf{X}_s\|_{\text{Fro}} \end{aligned} \tag{3}$$

Due to the non-convexity of the optimization problem in Eq. (2), the results depend on the initialization of $\mathbf{w}^{(k)}[s]$. In this paper, we use $\mathbf{w}^{(1)}[s] = 1$ for all s as proposed in [9]. Consequently, the first iteration of irMxNE is equivalent to solving a standard MxNE problem. As each iteration of the iterative scheme in Eq. (3) solves a convex problem with guaranteed global convergence, the initialization of X has no influence on the final solution. Intuitively, sources with high amplitudes in the (k-1)$^{\text{th}}$ iteration will be less penalized in the k$^{\text{th}}$ iteration and therefore further promoted.

For sources with $\|\widehat{\mathbf{X}}_s^{(k)}\|_{\text{Fro}} = 0$, this optimization problem has an infinite regularization term. Typically, a smoothing parameter ϵ is added to avoid weights to become zero [8], [9]. Here, we reformulate the weighted MxNE subproblems to apply the weights by scaling the gain matrix as given in Eq. (4).

$$\begin{aligned} \widetilde{\mathbf{X}}^{(k)} &= \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{M} - \mathbf{G}\mathbf{W}^{(k)}\mathbf{X}\|_{\text{Fro}}^2 + \lambda \sum_s \|\mathbf{X}_s\|_{\text{Fro}} \\ &= \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{M} - \mathbf{G}^{(k)}\mathbf{X}\|_{\text{Fro}}^2 + \lambda \sum_s \|\mathbf{X}_s\|_{\text{Fro}} \end{aligned} \tag{4}$$

with $\mathbf{W}^{(k)} \in \mathbb{R}^{SO \times SO}$ being a diagonal matrix, which is computed according to Eq. (5):

$$\begin{aligned} \mathbf{W}^{(k)} &= \text{diag}(\mathbf{w}^{(k)} \otimes \mathbf{1}_{(O)}) \\ \text{with } \mathbf{w}^{(k)}[s] &= 2\sqrt{\left\|\widehat{\mathbf{X}}_s^{(k-1)}\right\|_{\text{Fro}}}, \end{aligned} \tag{5}$$

where $\mathbf{1}_{(O)} \in \mathbb{R}^O$ is a vector of ones and $\otimes$ is the Kronecker product. In each MxNE iteration, we restrict the source space to sources locations with $\mathbf{w}^{(k)}[s] > 0$, which reduces the computation time.

The irMxNE optimization scheme solves the non-convex optimization problem in Eq. (2) by solving convex surrogate optimization problems. The global convergence of the k$^{\text{th}}$ weighted MxNE subproblem can hence be controlled by computing the primal-dual gap $\eta^{(k)}$ [10]

$$\eta^{(k)} = \mathcal{F}_p\left(\widetilde{\mathbf{X}}^{(k)}, \mathbf{G}^{(k)}\right) - \mathcal{F}_d\left(\widetilde{\mathbf{Y}}^{(k)}, \mathbf{G}^{(k)}\right) \geq 0, \tag{6}$$

where $\widetilde{\mathbf{Y}}^{(k)} = \mathbf{M} - \mathbf{G}^{(k)}\widetilde{\mathbf{X}}^{(k)}$ is the dual variable associated to $\widetilde{\mathbf{X}}^{(k)}$, $\mathcal{F}_p(\mathbf{X}, \mathbf{G}) = \frac{1}{2}\|\mathbf{M} - \mathbf{G}\mathbf{X}\|_{\text{Fro}}^2 + \lambda \sum_s \|\mathbf{X}_s\|_{\text{Fro}}$ the primal problem (cf. Eq. (4)), and $\mathcal{F}_d(\mathbf{Y}, \mathbf{G})$ the associated dual problem. Using the Fenchel-Rockafellar duality theorem [19] and noting that the Fenchel conjugate of a norm is the

indicator function of its dual norm, the dual problem can be computed as follows [10]:

$$\mathcal{F}_d\left(\mathbf{Y},\mathbf{G}\right) = -\frac{1}{2}\|\overline{\mathbf{Y}}\| + \mathrm{Tr}\left(\left(\overline{\mathbf{Y}}\right)^T\mathbf{M}\right)$$
$$\text{with } \overline{\mathbf{Y}} = \mathbf{Y}/\max\left(\max_s\left(\left\|\mathbf{G}_s^T\mathbf{Y}\right\|_{\text{Fro}}\right)/\lambda, 1\right) \quad ,$$

where $\mathbf{G}_s \in \mathbb{R}^{N\times O}$ is the block of $\mathbf{G}$ corresponding to source location s. In practice, we terminate the optimization of the k[th] MxNE subproblem, when the current solution $\widetilde{\mathbf{X}}^{(k)}$ is ϵ-optimal with $\epsilon = 10^{-6}$, i.e., $\eta^{(k)} < 10^{-6}$. According to Gramfort et al. [10], this is a conservative choice provided that the data is scaled or normalized by pre-whitening. In order to compute the final estimate $\widehat{\mathbf{X}}^{(k)}$, we finally need to rescale $\widetilde{\mathbf{X}}^{(k)}$ with $\mathbf{W}^{(k)}$ after convergence, $\widehat{\mathbf{X}}^{(k)} = \mathbf{W}^{(k)}\widetilde{\mathbf{X}}^{(k)}$.

For solving the weighted MxNE subproblems, we apply a block coordinate descent (BCD) scheme [15], which, for the problem at hand, converges faster than the Fast Iterative Shrinkage-Thresholding algorithm (FISTA) proposed earlier in [10] (cf. section III-B). A BCD scheme for solving the Group LASSO was proposed in [13], [16]. The subproblem per block has a closed form solution, which involves applying the group soft-thresholding operator, the proximity operator associated to the $l_{2,1}$-mixed-norm [10]. Accordingly, the closed form solution for the BCD subproblems solving the MxNE problem can be derived, which is given in Eq. (7).

$$\begin{aligned}\overline{\mathbf{X}}_s^{(k)} &= \mathbf{X}_s^{(k-1)} + \boldsymbol{\mu}[s]\mathbf{G}_s^T\left(\mathbf{M} - \mathbf{G}\mathbf{X}^{(k-1)}\right)\\ \widetilde{\mathbf{X}}_s^{(k)} &= \tilde{\mathbf{X}}_s^{(k)}\max\left(1 - \frac{\boldsymbol{\mu}[s]\lambda}{\max\left(\left\|\overline{\mathbf{X}}_s^{(k)}\right\|_{\text{Fro}}, \boldsymbol{\mu}[s]\lambda\right)}, 0\right)\end{aligned} \quad (7)$$

The step length $\boldsymbol{\mu}[s]$ for each BCD subproblem is determined by $\boldsymbol{\mu}[s] = L_s^{-1}$ with $L_s = \|\mathbf{G}_s^T\mathbf{G}_s\|$ being the Lipschitz constant of the data-fit gradient restricted to the s[th] source location. This step length is typically larger than the step length applicable in iterative proximal gradient methods, which is upper-bounded by the inverse of $L = \|\mathbf{G}^T\mathbf{G}\|$. Pseudo code for the BCD scheme is shown in Algorithm 1.

Algorithm 1 MxNE with BCD

Require: $\mathbf{M}$, $\mathbf{G}$, $\mathbf{X}$, $\boldsymbol{\mu}$, $\lambda > 0$, $\epsilon > 0$, and S.
Initialization: $\mathbf{R} = \mathbf{M} - \mathbf{G}\mathbf{X}$, $\eta = \mathcal{F}_p\left(\mathbf{X}\right) - \mathcal{F}_d\left(\mathbf{R}\right)$
while $\eta \geq \epsilon$ **do**
  for $s = 1$ to S **do**
    $\widehat{\mathbf{X}}_s \longleftarrow$ Solve Eq. (7) with $\mathbf{X}$, $\boldsymbol{\mu}$, and $\mathbf{M}$
    $\mathbf{R} = \mathbf{R} - \mathbf{G}_s\left(\widehat{\mathbf{X}}_s - \mathbf{X}_s\right)$
    $\mathbf{X}_s = \widehat{\mathbf{X}}_s$
  end for
  $\eta = \mathcal{F}_p(\widehat{\mathbf{X}}, \mathbf{G}) - \mathcal{F}_d(\mathbf{R}, \mathbf{G})$
end while

The BCD scheme is typically applied with a cyclic sweep pattern, i.e., all blocks are updated in each BCD iteration. However, due to the constraint promoting spatial sparsity, most of the blocks should stay at zero. The computation time can thus be improved by using data-dependent sweep patterns (such as in Greedy Block Coordinate Descent [20]) or active set strategies [21], [22], which are designed to primarily update sources, which are likely to be active, while keeping the remaining sources at zero. The active set strategy, which we apply here, starts by estimating an initial active set $\mathcal{A}$ by finding the V (we use $V = 10$) sources violating the Karush-Kuhn-Tucker (KKT) optimality conditions [10], [22] the most. Subsequently, we estimate $\widehat{\mathbf{X}}^{\mathcal{A}}$ by solving Eq. (4) using $\mathbf{G}^{\mathcal{A}}$, i.e., $\mathbf{G}$ restricted to $\mathcal{A}$, with convergence controlled by means of the primal-dual gap. After convergence, we check if $\widehat{\mathbf{X}}$ with $\widehat{\mathbf{X}}_s = \widehat{\mathbf{X}}_s^{\mathcal{A}}$ if $s \in \mathcal{A}$, else $\mathbf{0}$ is ϵ-optimal for the full optimization problem (without restricting the source space to $\mathcal{A}$) by computing the corresponding primal-dual gap. If this is not the case, we update $\mathcal{A}$ by computing the support of $\widehat{\mathbf{X}}$ and adding at most V sources by re-evaluating the KKT optimality conditions and repeat the estimation of $\widehat{\mathbf{X}}^{\mathcal{A}}$ with warm start. Pseudo code for the proposed MxNE solver is provided in Algorithm 2.

Algorithm 2 MxNE-BCD with active set strategy

Require: $\mathbf{M}$, $\mathbf{G}$, $\lambda > 0$, $\epsilon > 0$, and S.
Initialization: $\mathbf{X} = \mathbf{0}$, $\mathbf{R} = \mathbf{M}$, $\eta = \mathcal{F}_p\left(\mathbf{X}\right) - \mathcal{F}_d\left(\mathbf{R}\right)$
for $s = 1$ to S **do**
  $\boldsymbol{\mu}[s] = \|\mathbf{G}_s^T\mathbf{G}_s\|^{-1}$
end for
$\mathcal{A} \subseteq \{s \mid \|\mathbf{G}_s^T\mathbf{R}\|_{\text{Fro}} > \lambda\}$
while $\eta \geq \epsilon$ **do**
  Define $\mathbf{G}^{\mathcal{A}}$ and $\mathbf{X}^{\mathcal{A}}$ by restricting $\mathbf{G}$ and $\mathbf{X}$ to $\mathcal{A}$
  $\widehat{\mathbf{X}}^{\mathcal{A}} \longleftarrow$ Solve Algorithm 1 with $\boldsymbol{\mu}$, $\mathbf{G}^{\mathcal{A}}$ and $\mathbf{X}_0 = \mathbf{X}^{\mathcal{A}}$
  $\widehat{\mathbf{X}} = \widehat{\mathbf{X}}^{\mathcal{A}}$ for $s \in \mathcal{A}$, else $\mathbf{0}$
  $\mathbf{R} = \mathbf{M} - \mathbf{G}\widehat{\mathbf{X}}$
  $\eta = \mathcal{F}_p(\widehat{\mathbf{X}}, \mathbf{G}) - \mathcal{F}_d(\mathbf{R}, \mathbf{G})$
  $\mathcal{A}^* \subseteq \{s \mid \|\mathbf{G}_s^T\mathbf{R}\|_{\text{Fro}} > \lambda\}$
  $\mathcal{A} = \mathrm{supp}(\widehat{\mathbf{X}}) \cup \mathcal{A}^*$
end while

We terminate irMxNE when $\max(|\widehat{\mathbf{X}}^{(k)} - \widehat{\mathbf{X}}^{(k-1)}|) < \tau$ with a user specified threshold τ, which we set to 10^{-6} in practice. The proposed optimization algorithm for irMxNE is fast enough to allow its usage on real MEG/EEG problems in the context of interactive data analysis. Full pseudo code for irMxNE is provided in Algorithm 3.

Algorithm 3 Iterative reweighted MxNE

Require: $\mathbf{M}$, $\mathbf{G}$, $\lambda > 0$, $\epsilon > 0$, $\tau > 0$, and K.
Initialization: $\mathbf{W}^{(1)} = \mathbf{I}$, $\widehat{\mathbf{X}}^{(1)}$
for $k = 1$ to K **do**
  $\mathbf{G}^{(k)} = \mathbf{G}\mathbf{W}^{(k)}$
  $\tilde{\mathbf{X}}^{(k)} \longleftarrow$ Solve Algorithm 2 with $\mathbf{G}^{(k)}$ and $\mathbf{X}^{(k)}$
  $\widehat{\mathbf{X}}^{(k)} = \mathbf{W}^{(k)}\tilde{\mathbf{X}}^{(k)}$
  if $\max(|\widehat{\mathbf{X}}^{(k)} - \widehat{\mathbf{X}}^{(k-1)}|) < \tau$ **then**
    break
  end if
  $\mathbf{W}^{(k+1)} \longleftarrow$ Solve Eq. 5 with $\widehat{\mathbf{X}}^{(k)}$
end for

C. Source constraints and bias

1) Source orientation: The proposed BCD scheme is applicable for MEG/EEG inverse problems with and without orientation constraint. For imposing a loose orientation constraint [23], we apply a weighting matrix $\mathbf{S} = \mathrm{diag}([1, \rho, \rho])$ to each block of the gain matrix $\mathbf{G}_s$ with $\mathbf{G}_s[:, 1]$ corresponding to the dipole orientated normally to the cortical surface, and $\mathbf{G}_s[:, 2]$ and $\mathbf{G}_s[:, 3]$ to the two tangential dipoles. The weighting parameter $0 < \rho \leq 1$ controls up to which angle the rotating dipole may deviate from the normal direction. The orientation-weighted gain matrix $\widetilde{\mathbf{G}}$ is hence defined as $\widetilde{\mathbf{G}} = \mathbf{G}\left(\mathbf{I}_{(S)} \otimes \mathbf{S}\right)$, where $\mathbf{I}_{(S)} \in \mathbb{R}^{S \times S}$ is the identity matrix. When the source orientation is postulated a priori (e.g. normal to the cortical surface), each block $\mathbf{X}_s$ corresponds to the activation of a fixed dipole. Consequently, the Frobenius norm per block can be replaced by the l_2-norm of the source activation of the corresponding fixed dipole.

2) Depth bias compensation: Due to the attenuation of the bioelectromagnetic field with increasing distance between source and sensor, deep sources require higher source amplitudes to generate sensor signals of equal strength compared to superficial sources. Consequently, inverse methods, which are based on constraints penalizing the source amplitudes, have a bias towards superficial sources. In order to compensate this bias, each block of the gain matrix is commonly normalized a priori, e.g. by its Frobenius or spectral norm. Here, we apply the depth bias compensation proposed in [24], which is based on computing a source covariance estimate $\mathbf{C} = \mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\mathbf{G}$ and weighting the gain matrix per source location with the inverse matrix square root of the corresponding $O \times O$ diagonal block of $\mathbf{C}$.

3) Amplitude bias compensation: Source activation estimated with source reconstruction approaches based on l_p-quasinorms with $0 < p \leq 1$, such as MxNE and irMxNE, show a varying degree of amplitude bias due to the inherent shrinkage. The amplitude bias is typically compensated by computing the least squares fit after restricting the source space to the support of $\widehat{\mathbf{X}}$, which is typically an over-determined optimization problem. However, this procedure modifies the time courses of the estimated neuronal activation. In contrast, we apply the debiasing approach proposed in [12], which preserves the source characteristics and orientations estimated with irMxNE by estimating a scaling factor for each active source.

D. Simulation setup

To provide a reproducible and reasonably fast comparison of MxNE and irMxNE, we generated a simulation data set with 20 sensors and 200 sources with fixed orientation. Active sources were randomly selected to be active, each with a random activation (20 time samples) drawn from a standard normal distribution. The simulations were performed with 2 and 4 active sources. For the first simulation (uncorrelated design), the columns of the linear forward operator were drawn from a multivariate standard normal distribution and normalized to 1, to make the comparison independent of the forward model and the sources' spatial configuration. For the second simulation (correlated design), we built a correlated gain matrix from a multivariate normal distribution $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$ with $\mathbf{\Sigma}$ being a Toeplitz matrix generated from a vector $\mathbf{v}$ with $\mathbf{v}_k = 0.95^{k-1}$, $k = 1 : 200$. White Gaussian noise was added to the sensor signals to set the signal-to-noise ratio (SNR), which we define here as $\|M_{\text{signal}}\|^2_{\text{Fro}} / \|M_{\text{noise}}\|^2_{\text{Fro}}$. For comparing the support recovery, we compute the F_1-score according to Eq. (8), where $\mathcal{A}(\mathbf{X})$ is the active set of X with $\mathcal{A}(\mathbf{X}) = \{j : \|\mathbf{X}_{j,\cdot}\|_2 > 0\}$, $|\mathcal{A}(\mathbf{X})|$ the number of elements in $\mathcal{A}(\mathbf{X})$, $\mathbf{X}^*$ the simulated source activation, and $\widehat{\mathbf{X}}$ the estimated source activation.

$$F_1 = 2\frac{|\mathcal{A}(\widehat{\mathbf{X}}) \cap \mathcal{A}(\mathbf{X}^*)|}{|\mathcal{A}(\widehat{\mathbf{X}})| + |\mathcal{A}(\mathbf{X}^*)|} \tag{8}$$

E. Experimental MEG data

We evaluate the performance of MxNE and irMxNE on two MEG data sets, auditory evoked fields (AEF) and somatosensory evoked fields (SEF), recorded using a 306-channel Elekta Neuromag Vectorview system (Elekta Neuromag Oy, Helsinki, Finland). A detailed description of the data and paradigms can be found in [25]–[27]. For the auditory data, we report results for AEFs evoked by left auditory stimulation with pure tones of 500 Hz. The analysis window for source estimation was chosen from 50 ms to 200 ms containing the N100m component. For the somatosensory data, we analyzed SEFs evoked by bipolar electrical stimulation (0.2 ms in duration) of the left median nerve. We used an analysis window from 18 ms to 200 ms. Following the standard pipeline from the MNE software [28], signal preprocessing for both data sets consisted of signal-space projection for suppressing environmental noise, and baseline correction using pre-stimulus data (from -200 ms to -20 ms). Epochs with peak-to-peak amplitudes exceeding predefined rejection parameters (3 pT for magnetometers, 400 pT/cm for gradiometers, and 150 V for EOG) were assumed to be affected by artifacts and discarded. Finally, the evoked fields were computed by averaging (AEF: 96 epochs, SEF: 294 epochs). The noise covariance matrix for spatial whitening was estimated from pre-stimulus data (from -200 ms to -20 ms). The gain matrix was computed using a set of 7498 cortical locations, and a three-layer boundary element model. The source reconstruction was performed without orientation constraint.

III. Results

A. Simulation study

MxNE and irMxNE were applied on both simulated data sets with 20 different regularization parameters normalized to λ_{max} (logarithmically spaced from 1 to 100 %). λ_{max}, which can be computed analytically [10], is the infimum of λ, such that if $\lambda \geq \lambda_{\text{max}}$, the active set is empty. 100 repetitions of the simulation were performed. The maximum F_1-scores for all repetitions are presented in Fig. 1.

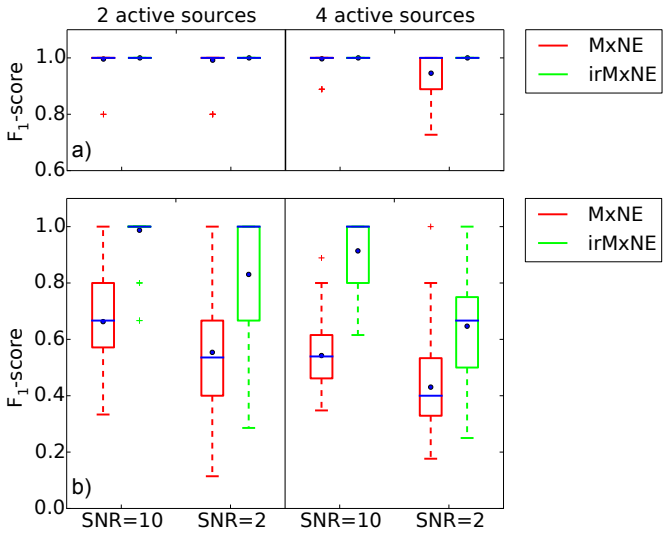


Fig. 1: F_1-scores for MxNE and irMxNE using simulated data with uncorrelated (a) and correlated design (b).

For the uncorrelated design, both methods reach high F_1-scores for both SNRs and active set sizes. While irMxNE allows for F_1-score = 1 in all repetitions, i.e., an exact reconstruction of the active set, MxNE shows individual outliers. For the correlated design, the F_1-scores decline with both decreasing SNR and increasing simulated active set size for both methods. However, the support recovery of irMxNE is less affected. Moreover, MxNE generally requires higher regularization parameters to reach its maximum F_1-score, which increases the amplitude bias.

B. Experimental MEG data

1) Auditory evoked fields: We first compare the performance of the proposed BCD scheme for solving the weighted MxNE with the Fast Iterative Shrinkage Thresholding Algorithm (FISTA) [29], an iterative proximal gradient method used in [10]. Both methods were applied with and without active set strategy. All computations were performed on a computer with a 2.4 GHz Intel Core 2 Duo processor and 8 GB RAM. The computation time as a function of the regularization parameter λ is illustrated in Fig. 2.

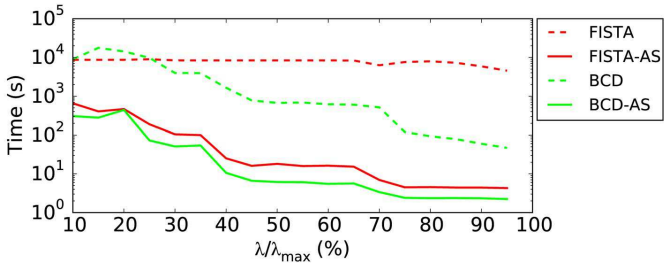


Fig. 2: Computation time as a function of λ for MxNE on real MEG data (free orientation) using BCD and FISTA with (solid) and without (dashed) active set strategy.

The BCD scheme outperforms FISTA both with and without active set strategy. Combining the BCD scheme and the active set strategy, which reduced the computation time by a factor of 100, allows to compute the MxNE on real MEG/EEG data in a few seconds. Since subsequent MxNE iterations are significantly faster due to the restriction of the source space, irMxNE also runs in a few seconds on real MEG/EEG source localization problems.

We then applied MxNE and irMxNE (with and without debiasing) with different regularization parameters λ to the AEF data and computed the mean Goodness of Fit (GOF) around the N100m component (from 90 ms to 150 ms) as well as the corresponding size of the estimated active set. Results are presented in Fig. 3.

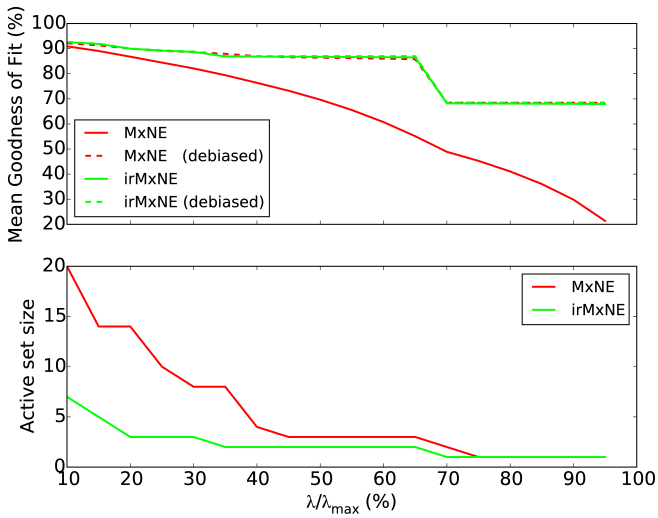
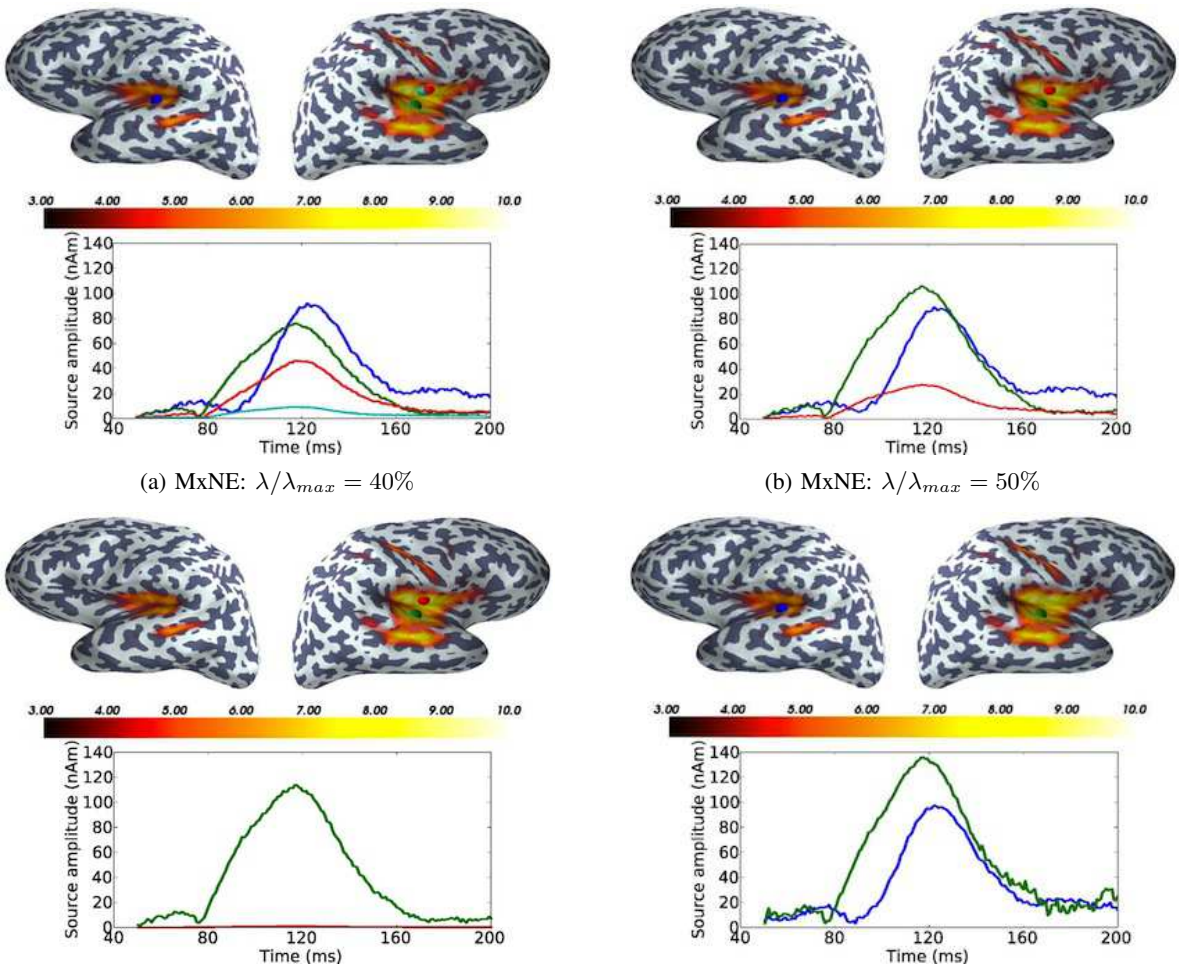


Fig. 3: Mean GOF before (dashed) and after debiasing (solid), and active set size obtained with MxNE and irMxNE as a function of the regularization parameter for the AEF data set.

The active set size estimated with irMxNE is smaller for all regularization parameters and less dependent on λ, particularly for small values of λ. For MxNE, the mean GOF is highly dependent on λ, when no debiasing is applied, and is improved significantly by debiasing. In contrast, the mean GOF for irMxNE is only slightly improved by debiasing, which indicates, that the irMxNE results are significantly less biased in amplitude. Compared to MxNE with debiasing, irMxNE yields a comparable mean GOF with similar stability plateaus with the advantage of providing a sparser model.

Source reconstruction results for MxNE and irMxNE are presented in Fig. 4. To compare results with a well established technique, we computed the dSPM [30] solution without orientation constraint using the MNE software [28] and added the mean dSPM estimates to Fig. 4. MxNE with $\lambda/\lambda_{max} = 40\%$ shows activation in both primary auditory cortices with its main peak around 110 ms corresponding to the N100m component. The activation on the right hemisphere



(a) MxNE: $\lambda/\lambda_{max} = 40\%$

(b) MxNE: $\lambda/\lambda_{max} = 50\%$

(c) MxNE: $\lambda/\lambda_{max} = 70\%$

(d) irMxNE: $\lambda/\lambda_{max} = 40\%$, convergence after 6 iterations

Fig. 4: Source reconstruction results using AEF data evoked by left auditory stimulation for MxNE (a, b, c) and irMxNE (d) with debiasing. The source locations, marked with dots, and the corresponding time courses are color-coded.

is however split into several highly correlated dipoles, which are partly located outside of the primary auditory cortex. Increasing λ, which is known to reduce the number of active dipoles, does not fix the latter issue, since dipoles in the left primary auditory cortex are eliminated before actually erasing the spurious activity from the right hemisphere. The loss of the active source in the left auditory cortex explains also the drop of the GOF in Fig. 3. In contrast, irMxNE with $\lambda/\lambda_{max} = 40\%$ reconstructs single dipoles in both primary auditory cortices. Intuitively, the green and blue sources, which are the strongest sources according to MxNE, are favored at the next iteration of the reweighted scheme. The estimated source locations are well inline with the maxima of the dSPM estimates, which are spatially smeared. Note also that source amplitudes obtained with irMxNE are moments of electrical dipoles expressed in nAm, which is similar to dipole fit procedures [31].

2) Somatosensory evoked fields: Fig. 5 presents the mean GOF and the corresponding size of the active set for MxNE and irMxNE (with and without debiasing) applied to the SEF data as a function of the regularization parameter λ.

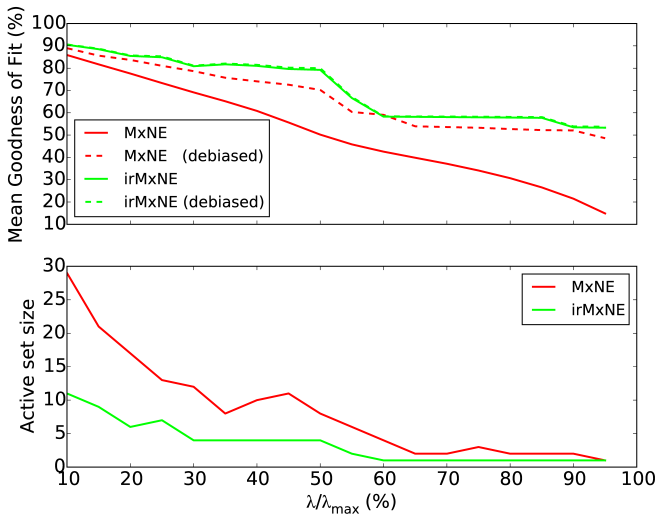


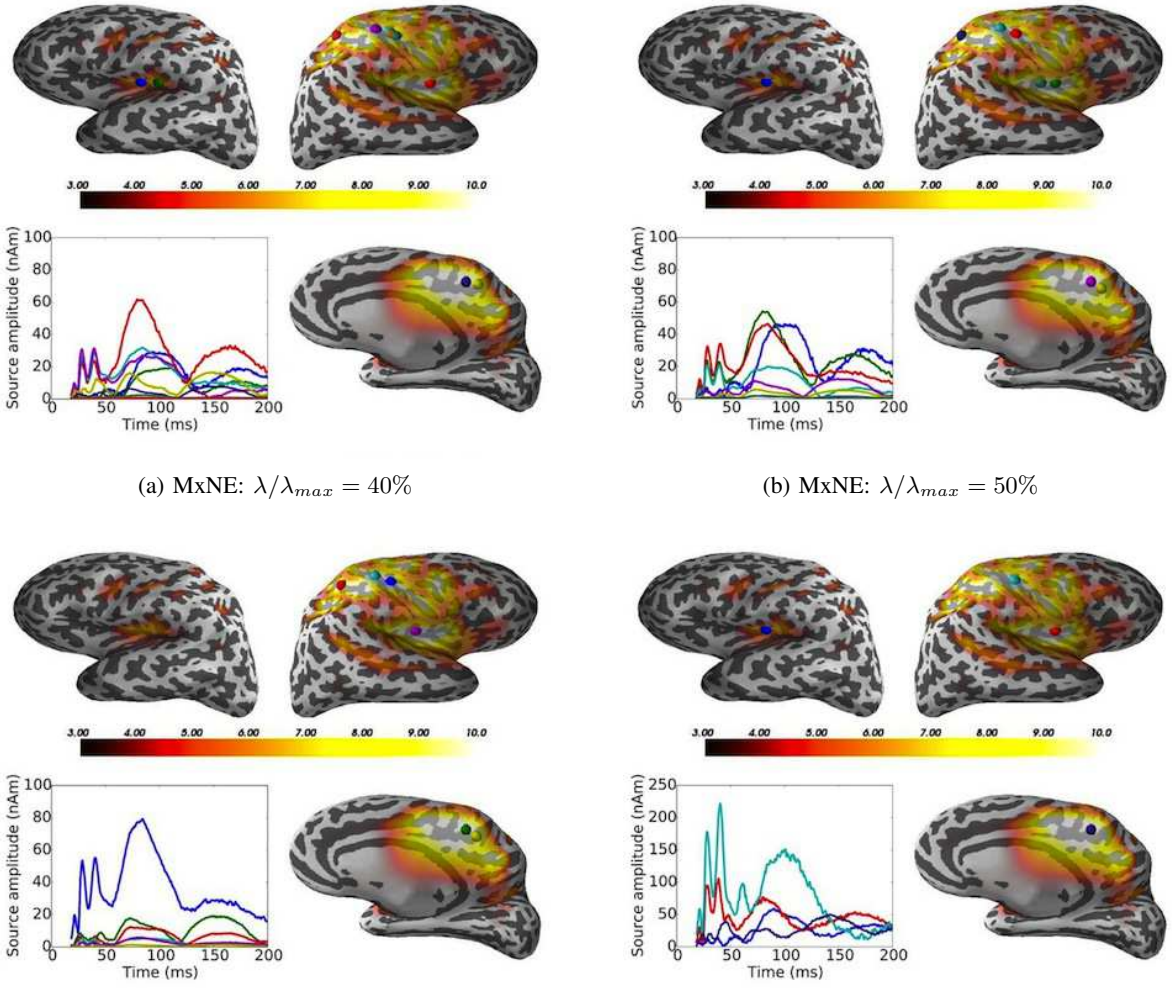
Fig. 5: Mean goodness of fit before (dashed) and after debiasing (solid), and active set size obtained with MxNE and irMxNE for the SEF data set.

The results confirm the findings obtained for the AEF data set in section III-B1. The source estimate obtained with irMxNE is sparser and less biased compared to the MxNE result. The mean GOF for irMxNE (with and without debiasing) is higher than for MxNE with debiasing, which indicates that the estimation of the source signal characteristics is improved due to the reduced bias of the iterative reweighted approach.

Fig. 6 presents source reconstruction results obtained with MxNE and irMxNE for selected regularization parameters as well as the mean of the corresponding dSPM estimate. For $\lambda/\lambda_{max} = 40\%$, both MxNE and irMxNE reconstruct dipoles in the contralateral primary somatosensory cortex (cS1), the contralateral and ipsilateral secondary somatosensory cortices (cS2 and iS2), and the contralateral medial wall. The source locations coincide with the maxima of the mean dSPM estimate. As for the AEF data set, the source activation computed with irMxNE is represented by a single dipole in each region, whereas the activation is split again into several correlated dipoles using MxNE. Increasing the regularization parameter in turn reduces the number of active sources as expected, but eliminates physiologically meaningful sources (iS2) before other activation (e.g. in cS1) is represented by a single dipole, which can also be seen by means of the drop of the GOF in Fig. 5. These results demonstrate that the proposed sparse solver can present a simple and easy-to-interpret full spatio-temporal picture of the active sources. In contrast, dSPM or similar linear inverse method (sLORETA, MNE, etc.) generally requires careful post-processing of the source estimates, e.g. by defining regions of interest, to present similar figures.

IV. Discussion and Conclusion

In this work, we presented irMxNE, a MEG/EEG inverse solver based on regularized regression with a non-convex block-separable penalty. The non-convex optimization



(a) MxNE: $\lambda/\lambda_{max} = 40\%$

(b) MxNE: $\lambda/\lambda_{max} = 50\%$

(c) MxNE: $\lambda/\lambda_{max} = 55\%$

(d) irMxNE: $\lambda/\lambda_{max} = 40\%$, convergence after 10 iterations

Fig. 6: Source reconstruction results using SEF data evoked by electric stimulation of the left median nerve for MxNE (a, b, c) and irMxNE (d) with debiasing. The source locations, marked with dots, and the corresponding source activations are color-coded.

problem is solved by iteratively solving (convex) weighted MxNE problems, which allows for fast algorithms and global convergence control at each iteration. The combination of a block coordinate descent scheme and an active set strategy significantly decreases the computation time making the proposed method applicable for the analysis of MEG/EEG data.

By choosing the initialization such that the first iteration is equivalent to computing a standard MxNE, the irMxNE source estimate is at least as sparse as the MxNE solution while reducing the amplitude bias. This was confirmed by empirical results based on simulations and analysis of two MEG data sets. Both experiments demonstrated that irMxNE outperforms MxNE in terms of active source identification and amplitude bias. As MxNE, irMxNE assumes that the locations of active sources is constant over time. Hence, it should be applied to data, for which this model assumption is approximately true, e.g., by selecting intervals of interest or applying a moving window approach. To go beyond stationary sources, the reconstruction of non-stationary focal source activation can be improved by applying sparsity constraints in the time-frequency domain such as in the TF-MxNE [12]. The use of non-convex regularization for such models will be part of future work. This solver will be made available in the MNE-Python package [32].

ACKNOWLEDGMENT

The authors would like to thank M. S. Hämäläinen for providing the experimental MEG data sets.

REFERENCES

[1] K. Matsuura and Y. Okabe, "Selective minimum-norm solution of the biomagnetic inverse problem." *IEEE Trans. Biomed. Eng.*, vol. 42, no. 6, pp. 608–615, June 1995.

[2] W. Ou, M. S. Hämaläinen, and P. Golland, "A distributed spatio-temporal EEG/MEG inverse solver," *NeuroImage*, vol. 44, no. 3, pp. 932–946, Feb. 2009.

[3] F. Lucka, S. Pursiainen, M. Burger, and C. Wolters, "Hierarchical bayesian inference for the eeg inverse problem using realistic fe head models: Depth localization and source separation for focal primary currents," *NeuroImage*, vol. 61, no. 4, pp. 1364–1382, Apr. 2012.

[4] A. Sorrentino, L. Parkkonen, A. Pascarella, C. Campi, and M. Piana, "Dynamical meg source modeling with multi-target bayesian filtering," *Human brain mapping*, vol. 30, no. 6, pp. 1911–1921, 2009.

[5] D. Wipf and S. Nagarajan, "A unified bayesian framework for MEG/EEG source imaging," *NeuroImage*, vol. 44, no. 3, pp. 947–966, Feb. 2009.

[6] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Royal Stat. Soc. (Series B)*, vol. 58, pp. 267–288, 1994.

[7] K. Uutela, M. S. Hämäläinen, and E. Somersalo, "Visualization of magnetoencephalographic data using minimum current estimates," *NeuroImage*, vol. 10, no. 2, pp. 173–180, Aug. 1999.

[8] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted l1 minimization," *J. Fourier Anal. Applicat.*, vol. 14, no. 5-6, pp. 877–905, Dec. 2008.

[9] G. Gasso, A. Rakotomamonjy, and S. Canu, "Recovering sparse signals with a certain family of nonconvex penalties and dc programming," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4686–4698, Dec. 2009.

[10] A. Gramfort, M. Kowalski, and M. S. Hämäläinen, "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.

[11] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *J. Royal Stat. Soc. Series B*, vol. 68, no. 1, pp. 49–67, Feb. 2006.

[12] A. Gramfort, D. Strohmeier, J. Haueisen, M. S. Hämäläinen, and M. Kowalski, "Time-frequency mixed-norm estimates: Sparse M/EEG imaging with non-stationary source activations," *NeuroImage*, vol. 70, pp. 410–422, Apr. 2013.

[13] A. Rakotomamonjy, "Surveying and comparing simultaneous sparse approximation (or group-lasso) algorithms," *Signal Process.*, vol. 91, no. 7, pp. 1505–1526, July 2011.

[14] D. Strohmeier, J. Haueisen, and A. Gramfort, "Improved meg/eeg source localization with reweighted mixed-norms," in *Proc. of the 4th Int. Workshop on Pattern Recognition in Neuroimaging 2014 (PRNI 2014)*, June 2014, pp. 1–4.

[15] P. Tseng, "Approximation accuracy, gradient methods, and error bound for structured convex optimization," *Math. Program.*, vol. 125, pp. 263–295, Oct. 2010.

[16] Z. Qin, K. Scheinberg, and D. Goldfarb, "Efficient block-coordinate descent algorithms for the group lasso," *Math. Program. Comput.*, vol. 5, no. 2, pp. 143–169, June 2013.

[17] A. Dale and M. Sereno, "Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach," *J. Cognitive Neurosci.*, vol. 5, no. 2, pp. 162–176, 1993.

[18] D. A. Engemann and A. Gramfort, "Automated model selection in covariance estimation and spatial whitening of MEG and EEG signals," *NeuroImage*, to be published.

[19] R. T. Rockafellar, *Convex analysis*. Princeton university press, 1997, no. 28.

[20] X. Wei, Y. Yuan, and Q. Ling, "DOA estimation using a greedy block coordinate descent algorithm," *IEEE Trans Signal Process.*, vol. 60, no. 12, pp. 6382–6394, Dec. 2012.

[21] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *J. Stat. Softw.*, vol. 33, no. 1, p. 1, 2010.

[22] V. Roth and B. Fischer, "The group-lasso for generalized linear models: uniqueness of solutions and efficient algorithms," in *Proc. of the 25th Int. Conference on Machine learning (ICML '08)*, Helsinki, Finland, 2008, pp. 848–855.

[23] F.-H. Lin, J. W. Belliveau, A. M. Dale, and M. S. Hämäläinen, "Distributed current estimates using cortical orientation constraints," *Hum. Brain Mapp.*, vol. 27, no. 1, pp. 1–13, Jan. 2006.

[24] S. Haufe, V. V. Nikulin, A. Ziehe, K.-R. Müller, and G. Nolte, "Combining sparsity and rotational invariance in EEG/MEG source reconstruction," *NeuroImage*, vol. 42, no. 2, pp. 726–738, Aug. 2008.

[25] C. J. Aine, L. Sanfratello, D. Ranken, E. Best, J. A. MacArthur, T. Wallace, K. Gilliam, C. H. Donahue, R. Montaño, J. E. Bryant, A. Scott, and J. M. Stephen, "MEG-SIM: a web portal for testing MEG analysis methods using realistic simulated and empirical data," *Neuroinformatics*, vol. 10, no. 2, pp. 141–158, Apr. 2012.

[26] M. Weisend, F. Hanlon, R. Montaño, S. Ahlfors, A. Leuthold, D. Pantazis, J. Mosher, A. Georgopoulos, M. S. Hämäläinen, and C. Aine, "Paving the way for cross-site pooling of magnetoencephalography (MEG) data," *International Congress Series*, vol. 1300, pp. 615–618, Aug. 2007.

[27] W. Ou, P. Golland, and M. Hämäläinen, "Sources of variability in MEG," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*. Springer, 2007, pp. 751–759.

[28] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, L. Parkkonen, and M. S. Hämäläinen, "MNE software for processing MEG and EEG data," *NeuroImage*, vol. 86, pp. 446–460, Feb. 2014.

[29] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. on Imaging Sci.*, vol. 2, no. 1, pp. 183–202, Jan. 2009.

[30] A. M. Dale, A. K. Liu, B. R. Fischl, R. L. Buckner, J. W. Belliveau, J. D. Lewine, and E. Halgren, "Dynamic statistical parametric mapping: Combining fMRI and meg for high-resolution imaging of cortical activity," *Neuron*, vol. 26, no. 1, pp. 55–67, Apr. 2000.

[31] M. Scherg and D. Von Cramon, "Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model." *Electroencephalogr. Clin. Neurophysiol.*, vol. 62, no. 1, pp. 32–44, Jan. 1985.

[32] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. S. Hämäläinen, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neurosci.*, vol. 7, no. 267, Dec. 2013.