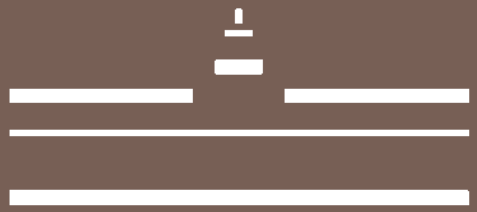
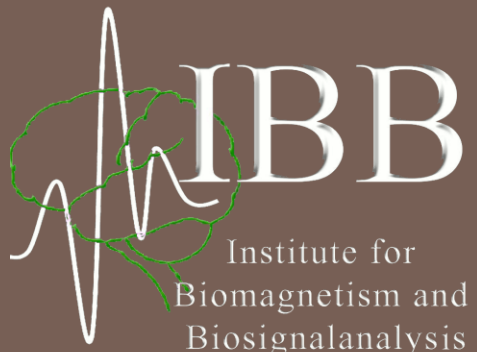


Westfälische
Wilhelms-Universität
Münster



EEG and MEG inverse problem, PART I: DIPOLE FIT AND DEVIATION SCAN

06.05.2025

Carsten Wolters

[Wolters, Vorlesungsskriptum, Chapter 8.1]

EEG/MEG inverse dipole fit approaches

One possibility is the restriction to a limited number of dipoles, the *focal source model* [336; 275; 214; 381]. The various spatio-temporal focal source models differ in the manner in which they describe the time dependence of the data. Generally, they are grouped into three classes, the unconstrained dipole model (so-called *moving dipole*), dipoles with temporally fixed location (*rotating dipole*) and dipoles with fixed location and fixed orientation (*fixed dipole*). If only one single time "snapshot" is taken into account, the three classes merge in a spatial dipole model, the so-called *instantaneous state dipole model* [391].

[Wolters, Vorlesungsskriptum, Chapter 8.1]

EEG/MEG inverse dipole fit approaches

8.3 Reconstruction of focal sources

Mosher et al. [214] showed how a common linear algebraic framework can be formulated for the three spatio-temporal focal source models, described in the introduction to this chapter. One can conclude from this formulation that measured fields depend nonlinearly on dipole location (and, for the fixed dipole, the fixed orientation) and linearly on dipole moment strength. Thus, after having chosen the number of sources, nonlinear algorithms should determine their locations (and possibly fixed orientations) and embedded linear methods their moment strength.

[Wolters, Vorlesungsskriptum, Chapter 8.1]

EEG/MEG inverse dipole fit and deviation scan approaches

It is thus the goal to find an influence node location tuple q for a chosen number of p dipoles (in practice p is always much smaller than 10) and the corresponding $r \times T$ strength matrix J_q such that

$$H(q) = ||\mathrm{U}^c - \mathrm{U}^{me}||_F^2 = ||\mathrm{L_q J}_q - \mathrm{U}^{me}||_F^2 \stackrel{!}{=} \min \tag{8.1}$$

where the $m \times r$ matrix $\mathrm{L_q}$ is the overdetermined lead-field matrix (thus $m > r$), corresponding to the location tuple q, the $m \times T$ matrix U^{me} are the noise-free measurements (EEG/MEG) and $||\cdot||_F$ is the Frobenius-norm. Using the normal-constraint, the number r of columns of $\mathrm{L_q}$ and rows of J_q equals the number p of dipoles, without this constraint it is $r = 3p$.

[Wolters, Vorlesungsskriptum, Chapter 8.1]

EEG/MEG inverse dipole fit and deviation scan approaches

The minimization task can thus be split into two problems. The first problem is to find the dipole location tuple q which gives a good approximation of the global minimum of H in a feasible calculation time. Second, a physiologically and mathematically suitable model should be developed for the shape of the functional graph H. Every evaluation of H for a given location tuple q contains the construction of the corresponding lead-field matrix $\mathrm{L_q}$ and the subsequent determination of the dipole direction and strength matrix J_q with respect to the noise U^{ε} in the measured data $\mathrm{U}_{\varepsilon}^{me} = \mathrm{U}^{me} + \mathrm{U}^{\varepsilon}$.

[Wolters, Vorlesungsskriptum, Chapter 8.3.1]

EEG/MEG inverse dipole fit approaches

8.3.1 A downhill simplex optimizer in continuous parameter space

Nelder and Mead's simplex method [217], formulated as an uphill approach in Section 4.5.1 (Algorithm 1), was described as a nonlinear optimization approach for the maximization of the mutual information between two MR images within the registration process. The concept of the Continuous Downhill Simplex (CDS) method used here is quite comparable. The CDS is used for the determination of the nonlinear parameters for simple focal source models in continuous parameter space, e.g., the continuous one dipole model. Main differences to Algorithm 1 are that the ordering is reversed and a downhill pivoting based on H instead of an uphill pivoting based on the mutual information is carried out. Furthermore, in order to speed up the optimization, the restriction to a Freudenthal triangulation was abolished. After each simplex reflection, which is still constructed to conserve the volume of the simplex and thus avoid degeneracy, the possibility of a moderate expansion in order to take larger steps was implemented.

[Wolters, Vorlesungsskriptum, see subsection "Optimization" in Chapter 4.5.1]

Reminder: Nelder-Mead optimization for image registration: The cost function

It is the goal of the registration of the PD- onto the T1-MRI to find transformation parameters $\underline{P}$, which maximize the mutual information between both images, i.e.,

$$\underline{P}^* := \arg\max_{\underline{P}\in\mathbb{R}^9} C_{X_{T1},X_{PD}}(\underline{P}) \tag{4.10}$$

with

$$C_{X_{T1},X_{PD}}(\underline{P}) := MI\left(X_{T1}, X_{T_{\underline{P}}(PD)}\right)$$

(see Equation (4.9) for the definition of $T_{\underline{P}}$).

[Wolters, Vorlesungsskriptum, see subsection "Optimization" in Chapter 4.5.1]

Reminder: Nelder-Mead optimization for image registration: The optimization algorithm

Optimization

the optimal transformation parameters are considered to maximize the function $C_{X_{T1},X_{PD}} : \mathbb{R}^N \to \mathbb{R}$ with, in our case, $N = 9$. Nelder-Mead's simplex method is an algorithm, which tries to determine a local extremum in multidimensional space [216]. The concept of the algorithm (see, e.g., [254]) was based on a Freudenthal triangulation, described by Allgower and Georg [6]

$$\begin{aligned} \underline{a}_1 &:= \underline{0}^N \in \mathbb{R}^N \\ \underline{a}_{i+1} &:= \underline{a}_i + \underline{e}_i, \ \forall i = 1, \ldots, N, \end{aligned}$$

with $\underline{e}_i$ the i^{th} unit basis vector of $\mathbb{R}^N$, in order to avoid degenerate simplices and to be successful for the considered application. The resulting method, presented

[Wolters, Vorlesungsskriptum, see subsection "Optimization" in Chapter 4.5.1]

Reminder: Nelder-Mead optimization for image registration: The optimization algorithm

in Algorithm 1, starts with a simplex with vertices $\underline{s}_1, \ldots, \underline{s}_{N+1} \in \mathbb{R}^N$, being images of Freudenthal's triangulation under an affine map, i.e.,

$$\underline{s}_i = d \cdot \mathrm{DIAG}(l_1, \ldots, l_N) \cdot \underline{a}_i + \underline{P}_{\mathrm{start}}, \; \forall i = 1, \ldots, N+1$$

with the user-given $N \times N$ diagonal matrix of scaling factors l_i, the resolution parameter d and the starting transformation vector $\underline{P}_{\mathrm{start}}$. In the next step, a permutation π from the set of all permutations of the integers 1 to $N+1$, S_{N+1}, is determined, arranging the vertices with respect to the corresponding value of mutual information. The simplex is then pivoted by reflection away from the minimal vertex value of $C_{X_{T1},X_{PD}}$ with the reflection function

$$\mathrm{refl}(\underline{s}_i) := \mathrm{pre}(\underline{s}_i) - \underline{s}_i + \mathrm{suc}(\underline{s}_i)$$

based on the cyclic left/right shift order-relation

$$\mathrm{pre}(\underline{s}_i) := \begin{cases} \underline{s}_{i-1} & \text{if} \quad i = 2, \ldots, N+1 \\ \underline{s}_{N+1} & \text{if} \quad i = 1 \end{cases}$$

$$\mathrm{suc}(\underline{s}_i) := \begin{cases} \underline{s}_{i+1} & \text{if} \quad i = 1, \ldots, N \\ \underline{s}_1 & \text{if} \quad i = N+1 \end{cases}$$

for the simplex vertices with $\mathrm{suc}(\underline{s}_i)$ the successor and $\mathrm{pre}(\underline{s}_i)$ the predecessor of $\underline{s}_i$ [6]. After pivoting, the algorithm restarts with a new shrunk version of a simplex at the actual maximum. It terminates, when the resolution parameter d falls under the predefined resolution bound d_{stop}.

[Wolters, Vorlesungsskriptum, see subsection "Optimization" in Chapter 4.5.1]

Reminder: Nelder-Mead optimization for image registration: The optimization algorithm

Algorithm 1 SIMP: $(T1, PD \in \mathcal{M}\,\mathcal{R}\,I, \underline{P}_{\text{start}}, \underline{l} \in \mathbb{R}^9, d_{\text{start}}, d_{\text{stop}} \in \mathbb{R}) \rightarrow (\underline{P}^* \in \mathbb{R}^9)$

Initialize: $\underline{P} = \underline{P}_{\text{start}},\ d = d_{\text{start}}$
while $d \geq d_{\text{stop}}$ **do**
  $\underline{s}_1 = \underline{P}$ /* starting simplex */
  $\underline{s}_{i+1} = d \cdot l_i \cdot \underline{e}_i + \underline{s}_i \quad \forall i = 1, \ldots, N$
  repeat
    Determine $\pi \in S_{N+1}$ with: /* ordering and uphill pivoting */
      $C_{X_{T1}, X_{PD}}(\underline{s}_{\pi(i)}) \leq C_{X_{T1}, X_{PD}}(\underline{s}_{\pi(i+1)})\ \forall i = 1, \ldots, N$
    for i=1,...,N+1 **do**
      if $C_{X_{T1}, X_{PD}}(\text{refl}(\underline{s}_{\pi(i)})) > C_{X_{T1}, X_{PD}}(\underline{s}_{\pi(i)})$ **then**
        $\underline{s}_{\pi(i)} = \text{refl}(\underline{s}_{\pi(i)})$
        Leave the for-loop
      end if
    end for
  until No more replacement of any $\underline{s}_i$
  $\underline{P} = \underline{s}_{\pi(N+1)},\ d = d/2$ /* translate and contract */
end while
$\underline{P}^* = \underline{P}$

[Wolters, Vorlesungsskriptum, Chapter 8.3.1]

EEG/MEG inverse dipole fit approaches

The CDS optimizer needs a-priori chosen seedpoints and, dependent on the seed dipoles, it can converge to local minima. This was shown for brain-stem auditory evoked potentials by Gerson et al. [101], where the CDS produced larger errors for even a simple one dipole model. Huang et al. [146] examined a multi-start CDS in order to imitate a global optimization technique for fitting multi-dipole, spatio-temporal MEG data.

[Wolters, Vorlesungsskriptum, Chapter 8.3.2]

Global optimization using Simulated Annealing (SA)

Algorithm 24 SA: $(H, p, \text{MAX} \in \mathbb{N}, t_{\text{step}}, \text{TOL} \in \mathbb{R}) \rightarrow q_{\text{opt}}$

DETERMINE t_{start} CHOOSE q_{utd}, $d_{\text{utd}} = d_{\text{opt}} = \text{BIGNUM}$, $t = t_{\text{start}}$, $n_{\text{sum}} = 0$
while $((d_{\text{utd}} > \text{TOL}) \wedge (n_{\text{sum}} < \text{MAX}))$ **do**
  $n_{\text{count}} = n_{\text{accept}} = 0$
  while $\left((n_{\text{accept}} < n_{\text{inf}}) \wedge (n_{\text{count}} < 10 * n_{\text{inf}})\right)$ **do**
    $q_{\text{tmp}} = RANDOM(q_{\text{utd}})$, $n_{\text{count}} = n_{\text{count}} + 1$
    $d_{\text{tmp}} = H(q_{\text{tmp}})$
    if $d_{\text{tmp}} < d_{\text{opt}}$ **then**
      $q_{\text{opt}} = q_{\text{tmp}}$, $d_{\text{opt}} = d_{\text{tmp}}$
    end if
    if $\text{dif} = (d_{\text{tmp}} - d_{\text{utd}}) > 0$ **then**
      $a = RANDOM([0, 1])$ /* Metropolis */
      if $e^{-\text{dif}/t} > a$ **then**
        $q_{\text{utd}} = q_{\text{tmp}}$, $d_{\text{utd}} = d_{\text{tmp}}$, $n_{\text{accept}} = n_{\text{accept}} + 1$
      end if
    else
      $q_{\text{utd}} = q_{\text{tmp}}$, $d_{\text{utd}} = d_{\text{tmp}}$, $n_{\text{accept}} = n_{\text{accept}} + 1$
    end if
  end while
  $t = t - t_{\text{step}}, n_{\text{sum}} = n_{\text{sum}} + n_{\text{count}}$
end while

[Wolters, Vorlesungsskriptum, Chapter 8.3.2]

Global optimization using Simulated Annealing (SA)

8.3.2 A simulated annealing algorithm on discretized influence space

The second presented optimization approach for the nonlinear source parameters tries to globally minimize the cost function H in a discretized parameter space and does not need any seedpoints. Theoretically, it would be possible to test all

$$\binom{n_{\text{inf}}}{\text{p}} = \frac{n_{\text{inf}}!}{(n_{\text{inf}} - \text{p})!\,\text{p}!}$$

combinations of a choice of p source locations without repetition out of n_{inf} influence nodes for their functional value H. In practice, this is generally not senseful because the number of elements in the configuration space is too large and cannot be explored exhaustively. The method of Simulated Annealing (SA) utilizes

[Wolters, Vorlesungsskriptum, Chapter 8.3.2]

Global optimization using Simulated Annealing (SA)

concepts of combinatorial optimization for searching the minimum of H in acceptable time [255; 24; 41; 101; 136; 381]. It simulates the process of a slow cooling (annealing) of a melted solid. If the annealing process is carried out slowly enough, then the crystal lattice of the solid finds the most regular and stress-free state, i.e., the state of lowest energy. If the heated solid is frozen too fast, then it can be that only a local energy minimum is achieved and stress remains in the crystal lattice. Analogous to the described annealing, the final state of the SA is determined by control parameters in the minimization algorithm. If the cooling is slow enough, the global minimum of the considered process will be found [149]. At high temperatures, the atomic mobility of the solid and therefore the probability of a displacement is increased. The mobility is lost proportionally to the cooling. In the minimization algorithm, this process is simulated with the Metropolis-criterion [206].

[Wolters, Vorlesungsskriptum, Chapter 8.3.2]

Global optimization using Simulated Annealing (SA)

Algorithm 24 SA: $(H, p, \text{MAX} \in \mathbb{N}, t_{\text{step}}, \text{TOL} \in \mathbb{R}) \rightarrow q_{\text{opt}}$

DETERMINE t_{start} CHOOSE q_{utd}, $d_{\text{utd}} = d_{\text{opt}} = \text{BIGNUM}$, $t = t_{\text{start}}$, $n_{\text{sum}} = 0$
while $((d_{\text{utd}} > \text{TOL}) \wedge (n_{\text{sum}} < \text{MAX}))$ **do**
  $n_{\text{count}} = n_{\text{accept}} = 0$
  while $\left((n_{\text{accept}} < n_{\text{inf}}) \wedge (n_{\text{count}} < 10 * n_{\text{inf}})\right)$ **do**
    $q_{\text{tmp}} = RANDOM(q_{\text{utd}})$, $n_{\text{count}} = n_{\text{count}} + 1$
    $d_{\text{tmp}} = H(q_{\text{tmp}})$
    if $d_{\text{tmp}} < d_{\text{opt}}$ **then**
      $q_{\text{opt}} = q_{\text{tmp}}$, $d_{\text{opt}} = d_{\text{tmp}}$
    end if
    if $\text{dif} = (d_{\text{tmp}} - d_{\text{utd}}) > 0$ **then**
      $a = RANDOM([0,1])$ /* Metropolis */
      if $e^{-\text{dif}/t} > a$ **then**
        $q_{\text{utd}} = q_{\text{tmp}}$, $d_{\text{utd}} = d_{\text{tmp}}$, $n_{\text{accept}} = n_{\text{accept}} + 1$
      end if
    else
      $q_{\text{utd}} = q_{\text{tmp}}$, $d_{\text{utd}} = d_{\text{tmp}}$, $n_{\text{accept}} = n_{\text{accept}} + 1$
    end if
  end while
  $t = t - t_{\text{step}}, n_{\text{sum}} = n_{\text{sum}} + n_{\text{count}}$
end while

The transfer of the SA algorithm into the application field of the focal inverse source reconstruction is shown in Algorithm 24, following [254; 24]. Within the inner loop, one source location of the up-to-date combination q_{utd} is randomly changed by the function *RANDOM* and saved in q_{tmp}, the temporary source combination. The functional H is then evaluated for the temporary sources. If the residual d_{tmp} is smaller than the current optimum, d_{opt}, then the sources are saved as the new optimal combination together with its discrepancy to the data. The temporary combination becomes the up-to-date one, if its defect to the data is smaller than d_{utd}. Especially in case of high system temperature, also a deterioration with regard to the defect is accepted. The coupling of the acceptance of a less favored source configuration to the temperature of the system is controlled by the Metropolis criterion. If the criterion is fulfilled, the less-favored source configuration becomes the up-to-date one. The starting high temperature results in a high probability of accepting a less favored configuration in order not to get stuck in a local extremum. The temperature is decreased by means of t_{step}, if either the number of accepted configurations, n_{accept}, reaches the number of influence nodes, or the number of all tested source tuples, n_{count}, gets larger than many times the number of n_{inf}. At the beginning of the algorithm, the starting temperature has to be determined. Therefore, an initial value for t_{start} is chosen and this value is increased during initial SA iterations as long as the condition $n_{\text{accept}}/n_{\text{count}} < 0.99$ is true. The algorithm thus adapts itself to the size of the influence space and, concerning the starting temperature, to the different scaling of EEG and MEG data. SA is terminated, if either the defect is falling under a given tolerance or a maximum of source configurations has been tested.

[Wolters, Vorlesungsskriptum, Chapter 8.3.3]

Dipole fit and deviation scan approaches: Determination of the linear parameters

$$H(q) = ||\mathrm{U}^c - \mathrm{U}^{me}||_F^2 = ||\mathrm{L_q J}_q - \mathrm{U}^{me}||_F^2 \stackrel{!}{=} \min \qquad (8.1)$$

8.3.3 Determination of the linear parameters

Assuming that a dipole location tuple q has been proposed, the problem (8.1) with noisy data $\mathrm{U}_\varepsilon^{me}$ should then be solved in order to compute the dipole strengths. Linear least square methods have to yield the "best" (will be defined) fit between measured and simulated fields. Refer to Zeidler [399, Chapter 37] or Lawson and Hansen [177] for a survey. If dipole components are proposed which "numerically" (nearly) project into the data null space, the corresponding lead-field matrix becomes ill-conditioned. In combination with noisy data, simple least square algorithms such as the generalized inverse can yield physiologically unexplainable results for dipole moment strengths, especially when overestimating the number of active sources. This problem can be alleviated with dipole fit regularization methods as shown in the following.

[Wolters, Vorlesungsskriptum, Chapter 8.3.3]

Dipole fit and deviation scan approaches: Determination of the linear parameters

$$H(q) = ||\mathrm{U}^c - \mathrm{U}^{me}||_F^2 = ||\mathrm{L_q J}_q - \mathrm{U}^{me}||_F^2 \stackrel{!}{=} \min \qquad (8.1)$$

Singular Value Decomposition and Generalized Inverse

The first three presented methods to solve the linear least square problem (8.1) do not take the noise in the measured data into account. They are based on different decompositions of the overdetermined $m \times r$ $(m > r)$ lead-field matrix $\mathrm{L_q}$. The first two methods, the *QR decomposition* and the *Complete Orthogonal Factorization* (COF), are discussed in [381; 373]. The third strategy is based on the

[Wolters, Vorlesungsskriptum, Chapter 8.3.3]

Dipole fit and deviation scan approaches: Determination of the linear parameters

Singular Value Decomposition (SVD) of the lead-field matrix

$$\mathrm{L}_q = \mathrm{VSW}^{tr} = (\underline{v}_1, \ldots, \underline{v}_m) \begin{pmatrix} \Sigma \\ 0_{m-r} \end{pmatrix} (\underline{w}_1, \ldots, \underline{w}_r)^{tr}$$

with the orthogonal $m \times m$ matrix V, the $m \times r$ matrix S with $\Sigma = \mathrm{DIAG}(\varsigma_1, \ldots \varsigma_r)$ and the orthogonal $r \times r$ matrix W [186; 365]. $\varsigma_i^2 \in \mathbb{R}$ are the eigenvalues and $\underline{v}_i \in \mathbb{R}^m$ the eigenvectors of $\mathrm{L}_q \mathrm{L}_q^{tr}$ and ς_i^2 the eigenvalues and $\underline{w}_i \in \mathbb{R}^r$ the eigenvectors of $\mathrm{L}_q^{tr} \mathrm{L}_q$. Furthermore, it is

$$\mathrm{L}_q \underline{w}_i = \varsigma_i \underline{v}_i \qquad i = 1, \ldots, r, \tag{8.2}$$

and

$$\begin{aligned} \underline{v}_i^{tr} \mathrm{L}_q &= \varsigma_i \underline{w}_i^{tr}, \qquad i = 1, \ldots, r, \\ \underline{v}_i^{tr} \mathrm{L}_q &= (\underline{0}^r)^{tr}, \qquad i = r+1, \ldots, m. \end{aligned}$$

Thus, the space spanned by $\{\underline{v}_1, \ldots, \underline{v}_r\}$ is called the *column space* and $\mathrm{span}\{\underline{v}_{r+1}, \ldots, \underline{v}_m\}$ is the so-called *left null space* of L_q. The *singular values* ς_i are automatically arranged by the SVD such that $\varsigma_1 \geq \varsigma_2 \geq \ldots \geq \varsigma_r > 0$, if full rank of L_q has been assumed. In practice, the *right singular vectors* $\underline{w}_i$ and the *left singular vectors* $\underline{v}_i$ are both arranged with increasing spatial frequency. In vector notation, the SVD can be written in the form

$$\mathrm{L}_q \underline{j}_q = \sum_{i=1}^{r} \varsigma_i (\underline{j}_q, \underline{w}_i) \underline{v}_i.$$

[Wolters, Vorlesungsskriptum, Chapter 8.3.3]

Dipole fit and deviation scan approaches: Determination of the linear parameters

Without respect to the noise in the data and using the above SVD, the least square problem (8.1) can be solved by means of the *generalized inverse* [186; 365] of the lead-field matrix, in matrix form

$$\begin{aligned} \mathrm{L}_q^{+}\mathrm{U}_\varepsilon^{me} &= \left(\mathrm{L}_q^{tr}\mathrm{L}_q\right)^{-1}\mathrm{L}_q^{tr}\mathrm{U}_\varepsilon^{me} = \mathrm{W}\mathrm{S}^{+}\mathrm{V}^{tr}\mathrm{U}_\varepsilon^{me} \\ &= (\underline{w}_1,\ldots,\underline{w}_r)\begin{pmatrix} \Sigma^{-1} & 0 \end{pmatrix}(\underline{v}_1,\ldots,\underline{v}_m)^{tr}\mathrm{U}_\varepsilon^{me}, \end{aligned}$$

or, in vector notation,

$$\mathrm{L}_q^{+}\underline{u}_\varepsilon^{me} = \sum_{i=1}^{r}\frac{1}{\varsigma_i}(\underline{u}^{me} + \underline{u}^{\varepsilon}, \underline{v}_i)\underline{w}_i. \tag{8.3}$$

As with the QR and the COF decomposition [381; 373], the only goal of the generalized inverse is to minimize the residual variance to the noisy data. Non-regularized dipole fit methods use this algorithm to solve the linear least square problems embedded in the nonlinear optimization process.

[Wolters, Vorlesungsskriptum, Chapter 8.3.3]

Dipole fit and deviation scan approaches: Determination of the linear parameters

Truncated singular value decomposition

In practice, L_q often suffers from large condition numbers $\mathrm{cond}_2(L_q) = \frac{\varsigma_1}{\varsigma_r}$ during the SA-optimization process. Thus, the singular values, ς_i, get very small and the high spatial frequency components of the noise in the data in Equation (8.3) can be extremely amplified. This has an effect on those spatial dipole components, which "numerically" (nearly) lie in the kernel of L_q. It can lead to source configurations, where dipoles with a large strength nearly cancel each other with regard to their surface field distribution and only explain noise (for both EEG and MEG inverse problem), the so-called *ghost sources*. When considering the MEG inverse problem, radial dipoles are such ghosts and they can get a big strength only to explain MEG noise. Both problems especially appear if the number of active sources is overestimated. The problem can be alleviated with a regularization T_λ of the generalized inverse

$$\underline{j}_q^\lambda = T_\lambda \underline{u}_\varepsilon^{me} := \sum_{i=1}^{r} \frac{1}{\varsigma_i} F_\lambda(\varsigma_i)(\underline{u}_\varepsilon^{me}, \underline{v}_i)\underline{w}_i,$$

where F_λ is called a *filter*, as described by Louis [186].

[Wolters, Vorlesungsskriptum, Chapter 8.3.3]

Dipole fit and deviation scan approaches: Determination of the linear parameters

The choice of

$$F_\lambda(\varsigma) = \frac{\varsigma^2}{\varsigma^2 + \lambda^2}$$

leads to the so-called *Tikhonov-Phillips regularization*, where the high spatial frequency components in the source space, strongly influenced by the noise in the data space, are attenuated. As discussed with more detail in Section 8.4.1, Tikhonov-Phillips regularization amounts to minimizing the functional

$$||\mathrm{L}_q \underline{j}_q - \underline{u}_\varepsilon^{me}||_2^2 + \lambda^2 ||\underline{j}_q||_2^2,$$

or, in equivalent denotation,

$$||\bar{\mathrm{L}}_q \underline{j}_q - \bar{\underline{u}}_\varepsilon^{me}||_2^2$$

with

$$\bar{\mathrm{L}}_q = \begin{pmatrix} \mathrm{L}_\mathrm{q} \\ \lambda \mathrm{Id}^r \end{pmatrix} \qquad \text{and} \qquad \bar{\underline{u}}_\varepsilon^{me} = \begin{pmatrix} \underline{u}_\varepsilon^{me} \\ \underline{0}^r \end{pmatrix}.$$

As shown by Hämmerlin and Hoffmann [133], the condition number of the Tikhonov regularized least square problem is then ameliorated to

$$\mathrm{cond}_2(\bar{\mathrm{L}}_q) = \sqrt{\frac{\varsigma_1^2 + \lambda^2}{\varsigma_r^2 + \lambda^2}},$$

This regularization concept for nonlinear dipole fit methods has recently been applied to source localization in [94].

[Wolters, Vorlesungsskriptum, Chapter 8.3.3]
[Wolters, Beckmann, Rienäcker, Buchner, *Brain Topography*, 1999]

Dipole fit and deviation scan approaches: Determination of the linear parameters

Another way is to choose the filter

$$F_\lambda(\varsigma) = \begin{cases} 1, \text{ if } \varsigma \geq \lambda \\ 0, \text{ if } \varsigma < \lambda \end{cases},$$

leading to a regularization called the Truncated Singular Value Decomposition (TSVD), which was proposed for source localization in [381]. This algorithm is simple to implement and has the effect of a lowpass filter. The high spatial frequency components of the data, $\varsigma_i(\cdot, \underline{w}_i)\underline{v}_i$, are lying below the noise level and, as a consequence, the high spatial frequency components in the source space cannot be reconstructed. Using regularization, information will be lost but otherwise the amplification of the high frequency data components would have a more negative effect on the solution, especially in combination with an overestimation of the number of active sources. Like Tikhonov-Phillips regularization, the TSVD ameliorates the condition of the problem.

[Wolters, Vorlesungsskriptum, Chapter 8.3.3]
[Wolters, Beckmann, Rienäcker, Buchner, *Brain Topography*, 1999]

Dipole fit and deviation scan approaches: Determination of the linear parameters

The SA-TSVD algorithm [381], described in the following, is an iterative procedure. In every step of the SA optimization, L_q is changing and thus the condition of the least square problem. Therefore, an automatic determination of the regularization parameter λ is essential. One possibility is provided by the so-called *discrepancy principle* (see, e.g., the overview article of Hanke and Hansen [137]) where the defect d to the measured data

$$
\begin{aligned}
d := \|(I - L_q T_\lambda)\underline{u}_\varepsilon^{me}\|_2^2 &= \|\underline{u}_\varepsilon^{me} - \sum_{i=1}^{r} F_\lambda(\varsigma_i)(\underline{u}_\varepsilon^{me}, v_i)v_i\|_2^2 \\
&= \|\underline{u}_\varepsilon^{me}\|_2^2 - \sum_{i=1}^{r} F_\lambda^2(\varsigma_i)|(\underline{u}_\varepsilon^{me}, v_i)|^2 \\
&\overset{\text{TSVD}}{=} \|\underline{u}_\varepsilon^{me}\|_2^2 - \sum_{\varsigma_i > \lambda} |(\underline{u}_\varepsilon^{me}, v_i)|^2
\end{aligned}
$$

is not only minimized, but chosen in dependence of the condition number of L_q and of the noise $\underline{u}^\varepsilon$ in the data.

[Wolters, Vorlesungsskriptum, Chapter 8.3.3]
[Wolters, Beckmann, Rienäcker, Buchner, *Brain Topography*, 1999]

Dipole fit and deviation scan approaches: Determination of the linear parameters

Let C be the $m \times m$ sample noise covariance matrix, determined e.g. from the signal-free pre-stimulus interval of the measurements, averaged over all epochs. This matrix reflects the spatial distribution and correlation of the measurement noise. C is a symmetric and positive definite matrix, which can be decomposed into $C = DD^{tr}$ by means of a singular value decomposition. If the noise statistics are known, i.e., if the number of epochs is sufficiently high in order to obtain a good estimate of the noise covariance matrix, the process of data *pre-whitening* can be used, as described by Knösche et al. [162]. Thus, we can restrict the theory to spatially uncorrelated noise where D is a diagonal weighting matrix.

[Wolters, Vorlesungsskriptum, Chapter 8.3.3]
[Wolters, Beckmann, Rienäcker, Buchner, *Brain Topography*, 1999]

Dipole fit and deviation scan approaches: Determination of the linear parameters

If we consider only one time point and Gaussian distributed and heteroscedastic (different in each channel) noise with zero mean, every channel i should be weighted according to its own noise standard deviation $\varepsilon_i = |(\underline{u}^{\varepsilon})^{[i]}|$ using the diagonal weighting matrix D^{-1} with entries $1/\varepsilon_i$. We thus get the weighted least square problem

$$H^w(q) \quad = \quad ||\mathrm{L_q}\underline{j}_q - \underline{u}^{me}_{\varepsilon}||^2_{\mathrm{C}^{-1}} = ||\mathrm{D}^{-1}(\mathrm{L_q}\underline{j}_q - \underline{u}^{me}_{\varepsilon})||^2_2 \stackrel{!}{=} \min,$$

the weighted regularized inverse

$$\underline{j}^w_{\lambda} = T^w_{\lambda}\mathrm{D}^{-1}\underline{u}^{me}_{\varepsilon} := \sum_{i=1}^{r} \frac{1}{\varsigma^w_i} F_{\lambda}(\varsigma^w_i)(\mathrm{D}^{-1}\underline{u}^{me}_{\varepsilon}, \underline{v}^w_i)\underline{w}^w_i,$$

where $\{\varsigma^w_i, \underline{v}^w_i, \underline{w}^w_i\}$ is the singular system of the weighted lead-field matrix $\mathrm{D}^{-1}\mathrm{L_q}$ and the weighted defect

$$d^w \quad = \quad ||\left(I - \mathrm{D}^{-1}\mathrm{L_q}T^w_{\lambda}\right)\mathrm{D}^{-1}\underline{u}^{me}_{\varepsilon}||^2_2 \stackrel{\mathrm{TSVD}}{=} ||\mathrm{D}^{-1}\underline{u}^{me}_{\varepsilon}||^2_2 - \sum_{\varsigma^w_i > \lambda} |(\mathrm{D}^{-1}\underline{u}^{me}_{\varepsilon}, \underline{v}^w_i)|^2.$$

[Wolters, Vorlesungsskriptum, Chapter 8.3.3]
[Wolters, Beckmann, Rienäcker, Buchner, *Brain Topography*, 1999]

Dipole fit and deviation scan approaches: Determination of the linear parameters

Algorithm 25 TSVD: $(R) \rightarrow d$

$\underline{j}_\lambda^w = 0,\ i = 0,\ \bar{d} = \|\mathrm{D}^{-1}\underline{u}_\varepsilon^{me}\|_2^2$
Compute $\{\varsigma_i^w, \underline{v}_i^w, \underline{w}_i^w\} = \mathrm{SVD}(\mathrm{D}^{-1}\mathrm{L_q})$
while $(i \leq r) \wedge (\bar{d} > R^2)$ **do**
  $i = i + 1$
  $\alpha_i = (\underline{v}_i^w)^{tr}\mathrm{D}^{-1}\underline{u}_\varepsilon^{me}$
  $\underline{j}_\lambda^w = \underline{j}_\lambda^w + \frac{1}{\varsigma_i^w}\underline{w}_i^w\alpha_i,\ \bar{d} = \bar{d} - \alpha_i^{tr}\alpha_i$
end while
$d = \bar{d}$

The instantaneous state ($T = 1$, α_i is a scalar) TSVD regularization procedure with so-called *a-posteriori* regularization parameter choice [186; 189], exploiting the discrepancy principle, is presented in Algorithm 25. Within this algorithm, the regularization parameter is coded by the free parameter R. The TSVD can be shown to be an order optimal regularization procedure [327; 186]. The larger R is chosen, the stronger the regularization will be. Because of the small number r of source components, good experience has been made with a choice of $R = 1.0$.

The extension of the SA-TSVD to spatio-temporal modeling ($T > 1$) is straightforward for the moving and for the rotating dipole model. In these cases, $\underline{\alpha}_i$ is the $1 \times T$ vector

$$\underline{\alpha}_i = (\underline{v}_i^w)^{tr}\mathrm{D}^{-1}\mathrm{U}_\varepsilon^{me},$$

$\frac{1}{\varsigma_i^w}\underline{w}_i^w\underline{\alpha}_i$ is an $r \times T$ matrix and the euclidian norm should be exchanged for the Frobenius-norm.

[Wolters, Beckmann, Rienäcker, Buchner, *Brain Topography*, 1999]

Studies in a simulated sulcus structure

Brain Topography, Volume 12, Number 1, 1999

Comparing Regularized and Non-Regularized Nonlinear Dipole Fit Methods: A Study in a Simulated Sulcus Structure

C.H. Wolters*, R.F. Beckmann^, A. Rienäcker#, and H. Buchner+

Summary: The inverse problem arising from EEG and MEG is largely underdetermined. One strategy to alleviate this problem is the restriction to a limited number of point-like sources, the focal source model. Although the singular value decomposition of the spatio-temporal data gives an estimate of the minimal number of dipoles contributing to the measurement, the exact number is unknown in advance and noise complicates the reconstruction. Classical non-regularized nonlinear dipole fit algorithms do not give an estimate for the correct number because they are not stable with regard to an overestimation of this parameter. Too many sources may only describe noise but can still attain a large magnitude during the inverse procedure and may be indiscernible from the true sources. This paper describes a nonlinear dipole fit reconstruction algorithm with a new regularization approach for the embedded linear problem, automatically controlled by the noise in the data and the condition of the occuring least square problems. The algorithm is stable with regard to source components which "nearly" lie in the kernel of the projection or lead field operator and it thus gives an estimate of the unknown number parameter. EEG simulation studies in a simulated sulcus structure are carried out for an instantaneous dipole model and spatial resolution in the sulcus and stability of the new method are compared with a classical reconstruction algorithm without regularization.

Key words: EEG/MEG; Nonlinear dipole fit; Simulated annealing; Regularization; Truncated singular value decomposition; Finite element method.

[Wolters, Beckmann, Rienäcker, Buchner, *Brain Topography*, 1999]

Studies in a simulated sulcus structure

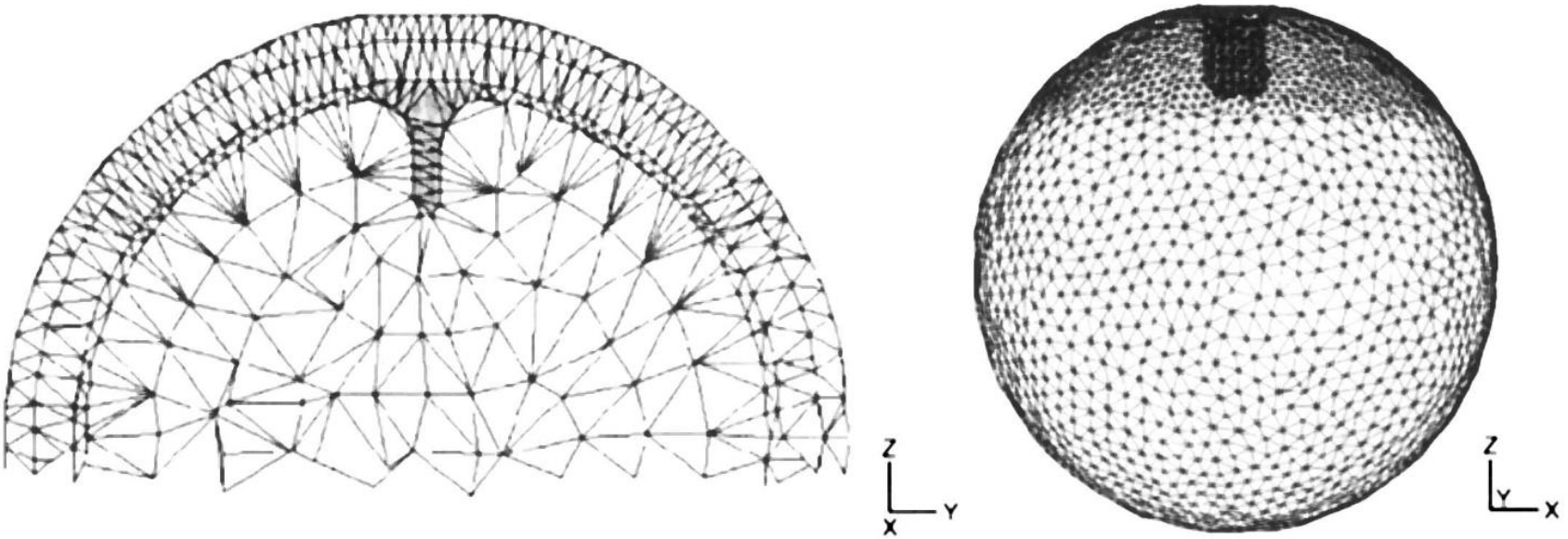


Figure 1. Tetrahedra mesh of the four layer sphere model with embedded sulcus structure. A cross-section through the finite element mesh (left) and the influence space mesh, used for simulations (right).

[Wolters, Beckmann, Rienäcker, Buchner, *Brain Topography*, 1999]

Studies in a simulated sulcus structure

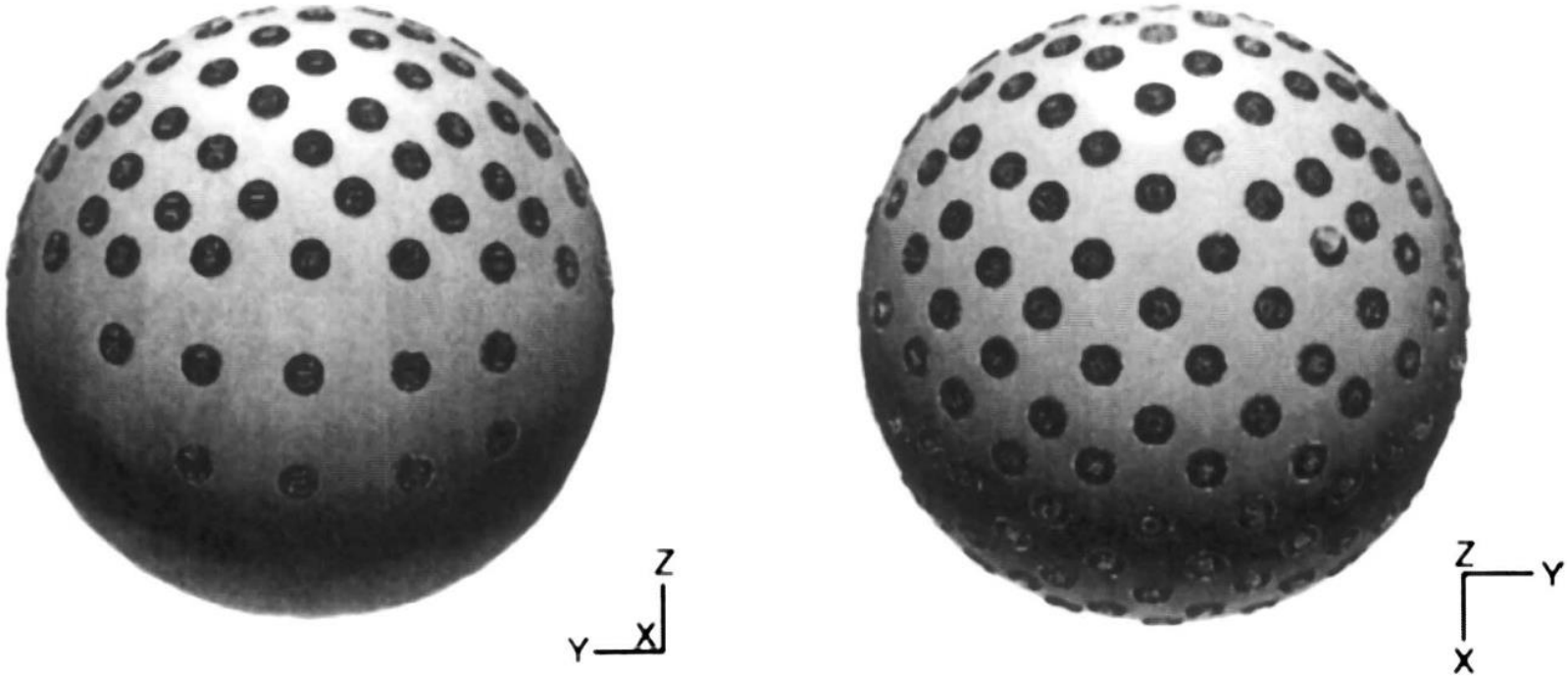


Figure 2. Right view (left) and top view (right) on the electrode positions on the outer sphere surface.

[Wolters, Beckmann, Rienäcker, Buchner, *Brain Topography*, 1999]

Studies in a simulated sulcus structure

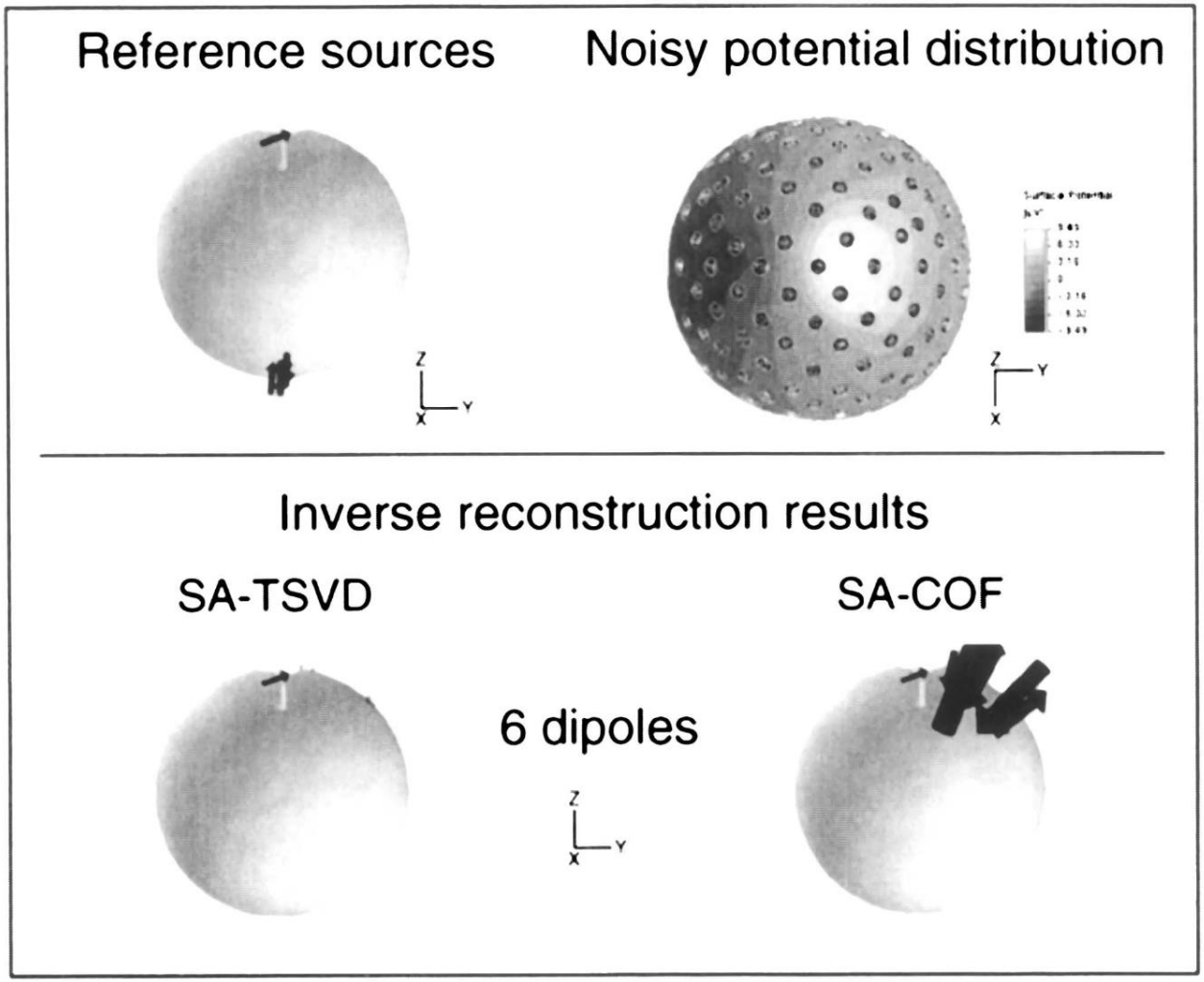


Figure 4. Activity on one sulcus wall: Simulation 1: Reference dipole configuration (top, left), the simulated noisy potential distribution on the outer sphere surface rereferenced to common average (top, right) and the reconstruction results of the SA-TSVD (bottom, left) and SA-COF (bottom, right) when searching for 6 dipoles.

Studies in a simulated sulcus structure

[Wolters, Beckmann, Rienäcker, Buchner, *Brain Topography*, 1999]

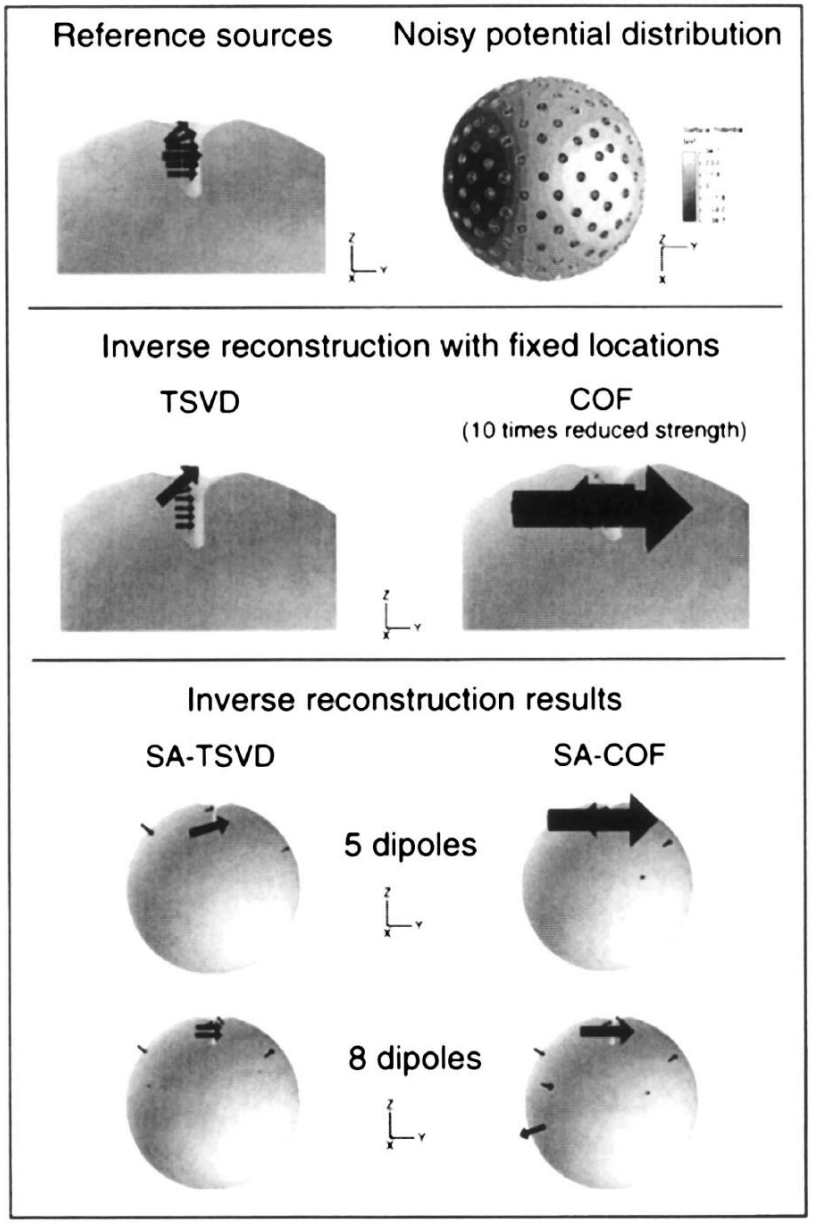


Figure 5. Activity on one sulcus wall: Simulation 2: Reference dipole configuration (top, left), the simulated noisy potential distribution on the outer sphere surface rereferenced to common average (top, right), the solution of the least square problem using the reference dipole locations with TSVD (middle, left) and with COF (ten times reduced dipole strength) (middle, right) and the reconstruction results of SA-TSVD (bottom, left) and SA-COF (bottom, right) when searching for 5 and 8 dipoles.

[Wolters, Beckmann, Rienäcker, Buchner, *Brain Topography*, 1999]

Studies in a simulated sulcus structure

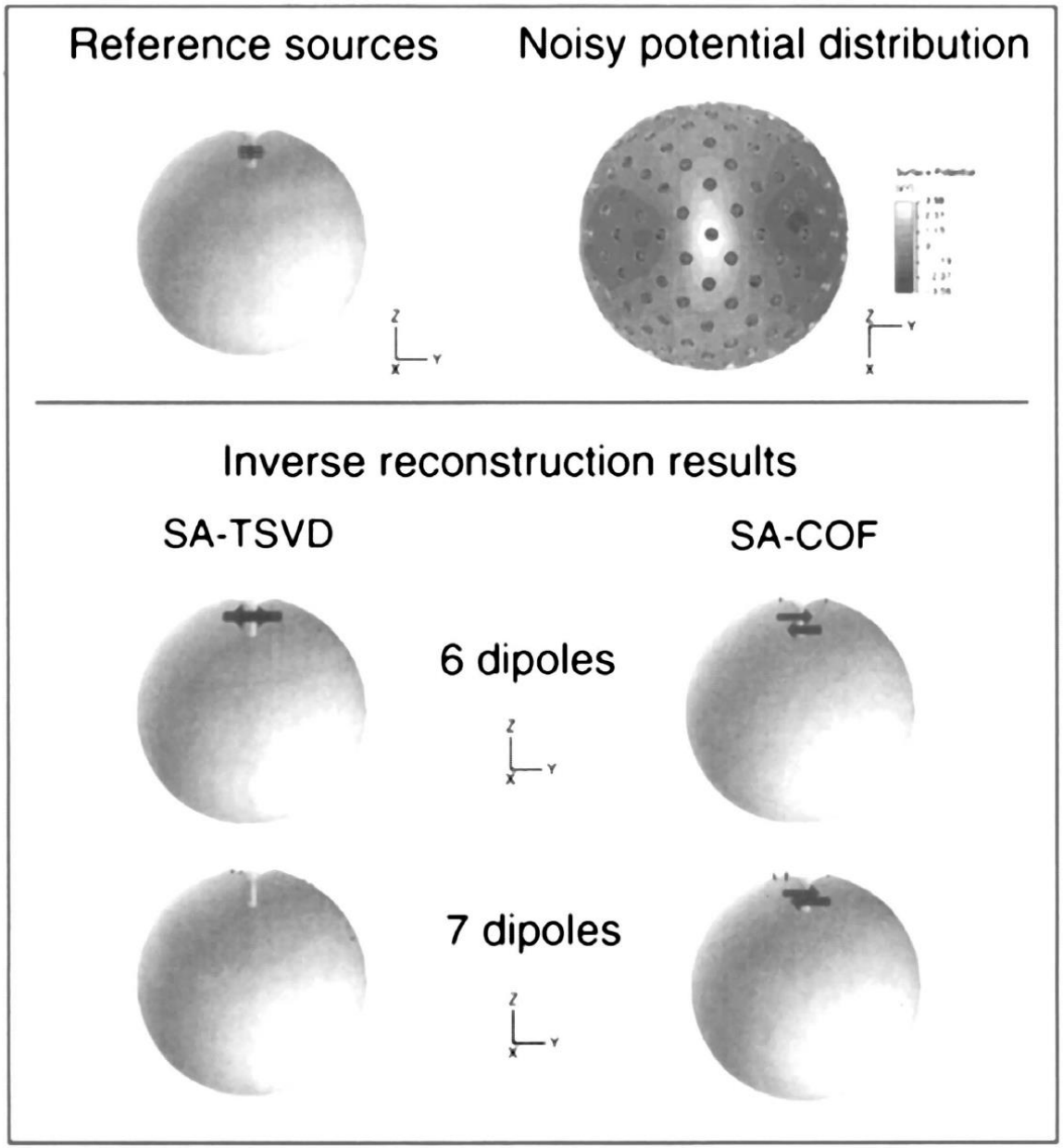
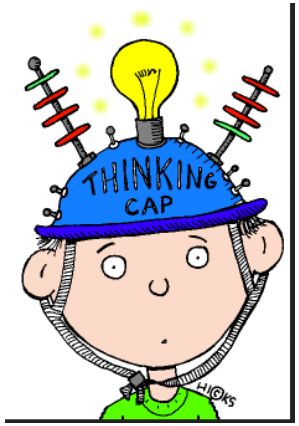


Figure 6. Activity on both sulcus walls: Reference dipole configuration (top, left), the simulated noisy potential distribution on the outer sphere surface rereferenced to common average (top, right) and the reconstruction results of SA-TSVD (bottom, left) and SA-COF (bottom, right) when searching for 6 and 7 dipoles.

Thank you for your attention!



2010-2016

2017-