

# Open Database of Epileptic EEG with MRI and Postoperational Assessment of Foci—a Real World Verification for the EEG Inverse Solutions

Piotr Zwoliński · Marcin Roszkowski ·  
Jarosław Żygierewicz · Stefan Haufe ·  
Guido Nolte · Piotr J. Durka

Published online: 25 September 2010

© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** This paper introduces a freely accessible database <http://eeg.pl/epi>, containing 23 datasets from patients diagnosed with and operated on for drug-resistant epilepsy. This was collected as part of the clinical routine at the Warsaw Memorial Child Hospital. Each record contains (1) pre-surgical electroencephalography (EEG) recording (10–20 system) with inter-ictal discharges marked separately by an expert, (2) a full set of magnetic resonance imaging (MRI) scans for calculations of the realistic forward models, (3) structural placement of the epileptogenic zone, recognized by electrocorticography (ECoG) and post-surgical results, plotted on pre-surgical MRI scans in transverse, sagittal and coronal projections, (4) brief clinical description of each case. The main goal of

this project is evaluation of possible improvements of localization of epileptic foci from the surface EEG recordings. These datasets offer a unique possibility for evaluating different EEG inverse solutions. We present preliminary results from a subset of these cases, including comparison of different schemes for the EEG inverse solution and preprocessing. We report also a finding which relates to the selective parametrization of single waveforms by multivariate matching pursuit, which is used in the preprocessing for the inverse solutions. It seems to offer a possibility of tracing the spatial evolution of seizures in time.

**Keywords** Epilepsy · SVD · MMP · MUSIC · Beamformer · EEG · MRI

---

P. Zwoliński · M. Roszkowski  
Memorial Child Hospital, Warsaw, Poland

P. Zwoliński  
e-mail: pz@meddata.pl

M. Roszkowski  
e-mail: m.roszkowski@czd.pl

J. Żygierewicz (✉) · P. J. Durka  
Department of Physics, University of Warsaw,  
Warsaw, Poland  
e-mail: jarekz@fuw.edu.pl

P. J. Durka  
e-mail: durka@fuw.edu.pl

S. Haufe  
Berlin Institute of Technology, Berlin, Germany  
e-mail: stefan.haufe@tu-berlin.de

G. Nolte  
Fraunhofer FIRST, Berlin, Germany  
e-mail: guido.nolte@first.fraunhofer.de

## Introduction

How EEG Inverse Solutions can Help Epileptology

Electroencephalography is the most important examination in clinical decisions related to initiation of therapy, its continuation and all other clinical activity in epileptology (Loddenkemper and Kotagal 2005; Rosenow and Luders 2001). Unique and specific electric pathology underlying the paroxysmal discharges in different types of epilepsies gives a powerful tool to clinical evaluation and therapy planning (Karbowski 1990; Pillai and Sperling 2006). Especially, long-term monitoring, as the Videometric method introduced by Penin (1968), allows registering such activity on the long time-span periods, potentially including epileptic fits.

In certain drug-resistant types of epilepsy, removal of the epileptogenic zone is the only curative therapy. Methods of automatic scoring of spikes (e.g. Koffler and Gotman 1985), as well as mapping of amplitude or its derivatives (e.g. Wong 1998), are helpful, but not sufficient in localization of the epileptogenic zone and its activity. It is still very difficult to delineate the epileptogenic zone without invasive (intracranial) EEG registration. This is complicated and invasive so limits application of these methods in clinical practice (Rosenow and Luders 2001). Modern diagnoses of any kind tend to be less invasive. One of the major goals of this study is to reduce the need of invasive methods in delineation of the epileptogenic zone by using epileptogenic activity localization methods based on the EEG measurement—or, at least, to clearly identify the practical and theoretical limits of this approach.

#### How this Database can Help the Development of EEG Inverse Solutions

Localization of brain electrical activity estimated from the scalp (EEG) is a challenging problem and draws the ongoing attention of many researchers in the field. There are many parameters influencing the results, like tissue geometry and electrode locations used for the forward model, noise characteristics etc. An even more severe confounder is the fact that the inverse solution is not unique without making additional assumptions about the sources. Lacking the ground truth in almost all real cases the proper choice of assumptions has been under debate for decades.

In general, the various methods can be split into three groups: a) explicit models with fewer parameters than unknowns, b) regularized dipole field reconstructions, and c) beamformers. The most established model is the  $N$  dipole model, assuming that the electric potential is generated by a few pointlike activities (Scherg and Ebersole 1993). Generalizations include multipole expansion, useful for focal but not necessarily pointlike sources (Irimia et al. 2009) or well defined patches of activation (Cao et al. 2006). For dipole field reconstructions the brain activation is estimated over the entire brain or cortical surface using various additional assumptions. The assumptions typically constraint the norm of the dipole fields (Molins et al. 2008) or the norm of the derivative of the field, e.g. LORETA (Pascual-Marqui et al. 2002) to regularize the solution. Recently, nonquadratic regularizations have become popular, since they combine the focality assumption with estimates of the entire brain activity (Ou et al. 2009; Haufe et al. 2008). Beamformers can be con-

sidered as an intermediate between  $N$  dipole models and dipole field reconstructions. The source activity is estimated at each brain location as a linear combination of the sensor data, such that activity from a dipole at that location is recovered exactly and all other contributions are maximally suppressed (Veen et al. 1997; Sekihara et al. 2005).

The above examples of inverse methods cover only a tiny fraction of established variants. The saying goes, that there are as many methods as scientific groups working in this area. The main reason for this diversity is a) that for each solution it is possible to construct, i.e. simulate, a problem where this solution is optimal, and b) that the factual localization, that could serve as a reference for evaluation of accuracy of a solution, for real data is rarely known. Thus we believe that a set of relevant real data examples with ground truth known to reasonable accuracy may substantially shed light on the open problem which of the numerous methods should be recommended.

#### The Database

As the first step towards a solution to the problems discussed in the previous section, we created a database which can be used for both:

- Evaluation of the currently available EEG Inverse Solutions performance as an auxiliary tool for at least rough localization of the epileptic foci
- Comparing stability and accuracy of the plethora of different EEG Inverse Solutions.

The database contains records of 23 patients with severe epilepsy, mostly caused by different organic lesions. The patients are aged 1–18 (mean  $12 \pm 5$  years). Summary data for the patients is presented in Table 1. This data was collected during the routine of the Warsaw Memorial Child Hospital. All selected patients were diagnosed with and operated on for drug-resistant epilepsy. The subjects presented in the database were selected from a total of 140 cases screened. The criteria for the inclusion was that the data are sufficiently unambiguous to proceed with the surgical removal of the epileptogenic zone. In each case, surgery was performed with ECoG which enabled precise localization of the epileptogenic region. If there was resected tissue, its pathology was extensively examined. As a result, the database contains the most unambiguous cases available, which corresponds well both of the aims formulated above.

**Table 1** Summary information about cases available in the database

Subject	Description	Duration of the recordings	Number of inter-ictal events
CHIMIC	Male, 9 y.o. , with severe form of epilepsy of temporal lobe origin	40'12"	4
FRAANN	Female, 17 y.o., with severe form of temporal lobe epilepsy	28'24"	3
HRADAW	Male, 12 y.o., with drug-resistant generalized seizures	23'0"	4
JANPAT	Female, 10 y.o., with severe form of temporal lobe epilepsy	20'03"	4
JANPRZ	Male, 5 y.o., with severe form of epilepsy since 1 y.o.	32'36"	4
JATKAM	Female, 14 y.o., with severe temporal lobe epilepsy caused by very unique multiple pathology	20'00"	7
GILPAU	Female, 14 y.o., with severe frontal lobe seizures	20'03"	5
GREOSK	Male, 13 y.o., with intermittent epilepsy of temporal lobe	20'09"	3
JERKAT	Female, 15 y.o., with severe epilepsy and moderate cognitive impairment	20'03"	5
KOSPAW	Male, 16 y.o. , with occasional epileptic fits and mild cognitive impairment	20'03"	3
KOTLUK	Male, 17 y.o., with refractory epilepsy of temporal origin	20'58"	2
KROMIC	Male, 16 y.o., with severe epilepsy and mild cognitive impairment	20'06"	5
LADJAN	Male, 10 y.o., with drug resistant epilepsy of frontal origin	22'24"	5
MARPAW	Male, 16 y.o., with drug resistant epilepsy of temporal origin	26'12"	6
MATPAW	Male, 15 y.o., with generalized seizures and parietal lobe tumor	20'24"	3
MEZLUK	Male, 4 y.o., with severe epilepsy of left temporal lobe	0'59"	2
NASNAD	Female, 1 y.o., with severe infantile spasm and focal seizures due to large FCD	22'25"	4
NIZMAR	Male, 12 y.o., with motor seizures and frontal lobe tumor	21'27"	3
NOWJON	Male, 15 y.o., with typical picture of temporal lobe epilepsy	20'03"	3
PAKDAM	Male, 13 y.o., with gross, disseminated FCD of right hemisphere and contralateral hemiparesis and mild cognitive impairment	1h12'35"	4
SIEGRZ	Male, 18 y.o., with typical MTLE	20'03"	4
SNOKAC	Male, 3 y.o., with drug resistant epilepsy and right lobe FCD	20'03"	6
STAMIC	Male, 15 y.o., with epilepsy due to right frontal lobe tumor	20'28"	5

A screenshot of a sample record from the database, displayed in the Web interface, is given in Fig. 1. Each database record includes:

**EEG:** Clinically relevant EEG epochs collected during the pre-surgical period. The data consists in each case of 20–70 minutes of continuous inter-ictal EEG recording, containing inter-ictal discharges which are representative of symptomatology and are in spatial relation with the structural pathology. During the recording the subject was resting. Data was recorded using the 10–20 system (with 19 electrodes), stored in European Data Format (EDF) (Kemp et al. 1992). The Ag/AgCl electrodes 10 mm diameter were used with impedance <5k $\Omega$ . Two EEG systems were used for recordings: DigiTrack EEG System by Elmiko (sampling frequency 250 Hz), and Medelec-Profile system by Medelec, Oxford Instruments (sampling frequency 256 Hz). For both apparatuses, filters were set to pass frequency band 0.5–70 Hz; additionally, 50 Hz notch filter was applied. The hardware referencing was to electrode Fpz. The data contains no detailed fiducial documentation.

**Marked epileptogenic waveforms:** Scanned printouts of EEG epochs containing the epileptogenic structures explicitly marked by the epileptologist.

**MRI:** MRI recordings containing a T<sub>1</sub>, T<sub>2</sub> or fluid attenuated inversion recovery (FLAIR) weighted brain scans with morphologic substrate of the epilepsy (mostly cortical dysplasias, dysplastic tumors etc.). In some cases there is also a scan with the gadolinium (GAD) contrast. The images were collected by Siemens Sonata 1.5T scanner. The data is stored in the Digital Imaging and Communications in Medicine (DICOM) format. The name of the scan's folder indicates its weighting. The scans have the following resolutions:

T<sub>1</sub>: 512 × 512 pixels, pixel spacing 0.4687 mm, slice thickness 1.2 mm,

T<sub>2</sub>: 256 × 256 pixels, pixel spacing 0.9375 mm, slice thickness 2.5 mm.

**Reference region:** Recognized structural placement of the epileptogenic zone marked on pre-surgical MRI scans in transverse, sagittal and coronal projections. The placement was verified by ECoG and post-operational results.

**Case description:** Textual description of each case, containing demographic data of each patient, short anamnesis of epilepsy, essential symptomatology of epileptic fits, additional necessary diagnostic tests results, and a concordance report for the epilepsy surgery decision.

**CHIMIC**  
Male, 9 y.o., with severe form of epilepsy of temporal lobe origin

Duration of the recordings:	40'12"
Number of seizures included:	0
Age at recording:	9
Age at surgery:	9
Number of inter-ictal events:	4

**Epileptogenic structures marked by epileptologists**

**Scans of 3 MRI sections with location of foci**

9 y.o. male with temporal lobe epilepsy – seizures with simple partial onset (epigastric sensations) evolving to partial complex (motor automatisms) with/without secondary generalization. Fits occurring everyday, with maximal withdrawal for 1 week. None of the AED (anti epileptic drugs) effective. Interictal EEG revealed generalizing sharp waves and spikes, as well as some focal slow waves. Side and focus of discharges hard to determine. MRI revealed a mass in right temporal region (hippocampal region). Pathology after surgery: dual pathology – FCD and ganglioglioma in right hippocampal complex. Resection (tumor lesionectomy) in Aug 2006 – since then seizure free

[Download binary EEG file](#) [Download binary MRI file](#)

**EEG :**

EEG hardware used:	DigiTrack EEG System
Kind of electrodes:	Ag/AgCl 10mm
Hardware reference:	Fpz
Sampling rate:	250 Hz
Hardware filters:	0.5–70 Hz
Software filters:	50 Hz notch
Other hardware hooked up:	none
Patient state during recording:	resting

**MRI :**

MRI hardware:	SIEMENS Sonata 1.5T
Image processing prior to archival:	none

**Fig. 1** Illustrative screenshot of one record from the database. Links [Download binary EEG file](#) and [Download binary MRI file](#) lead to the corresponding files. The pictures provide links to PDF files with scans of: printouts of the EEG

epochs with epileptic structures marked by epileptologist (*left*), and scans of 3 MRI sections per each patient with the location of foci, verified by ECoG and post-operational results, marked by pen on the printout (*right*)

Personal information was removed from this data, cases are identified by six-letter codes.

### Example Analysis

This section presents preliminary results from analysis of only a few cases from the database—more extensive examination of these findings is an ongoing work.

### Methods

The accuracy of any source localization methodology relies heavily on the adequacy of every step in the data processing chain. These steps can be divided into the construction of the physical head model, the signal preprocessing, and the actual source localization technique. In the following paragraphs we will describe



what is needed to obtain a realistic EEG forward model and outline a number of preprocessing/feature extraction techniques and inverse solutions that are appropriate for localization of spike-like activity.

### EEG Forward Modeling

For the underlying physical model, also known as the forward model, it is common to assume a standard head geometry. However, for children no such template exists, and the deviation from a normal adult brain can be particularly large when dealing with long-term epilepsy patients. The repository provided here facilitates a more realistic modeling of the relation between brain sources and the EEG in these cases. This is because the patient's individual head geometry can be taken into account. The steps necessary for building such an individual model involve:

1. Segmentation of the MR image into homogeneous volumes representing skin, skull and brain tissues
2. Computation of the linear forward mapping in the obtained three-shell head model.

These steps were carried out here using (1) Curry 5.0 (Neuroscan) with segmentation resolution of 1mm and (2) a Matlab (The Mathworks) implementation of the semi-analytic leadfield expansion procedure described in (Nolte and Dassios 2005) using a model grid size of 5 mm and the following conductivities of the layers: skin 0.33 S/m, skull 0.004125 S/m and, brain 0.33 S/m. This three layer model is a standard one, and contains the biggest effect, namely the very different conductivities of the skull on the one hand and the skin and brain on the other. In principle, one could also include cerebrospinal fluid (CSF). The problem is that the brain is highly folded which would render the calculation inaccurate for the available resolution of the MRI segmentation. The EEG signals were measured in the standard 10-20 system (Niedermeyer and Lopes da Silva 2004). In the computation of the forward problem the standard positions of the electrodes are first transformed to match the size and orientation of the individual head model and later their positions are fine-tuned such that the electrodes are put exactly on the scalp.

### Signal Preprocessing

The quality and relevance of the estimate of activity localization, obtained from the inverse solutions, depends on the preprocessing of the data that is used as the input

to the inverse solution method (Durka et al. 2005). We present results obtained for three types of input:

1. Raw signal form selected epochs,
2. Topography of the first component of singular value decomposition (SVD) of the signal of the selected epochs,
3. Topography of the structures from multivariate matching pursuit (MMP) decomposition selected by epileptologist as relevant to the epileptic activity.

*Singular Value Decomposition* inter-ictal spikes strongly contribute to the EEG, that is, they account for a huge amount of variance in the epochs containing them. Thus epileptic activity can be most easily extracted by searching for EEG components with high-variance. This is achieved by Singular Value Decomposition (SVD), which is an orthogonal transformation of a multivariate signal into uncorrelated components. The EEG signal with  $N_c$  channels and  $T$  samples can be represented as  $T \times N_c$  matrix  $\mathbf{x}$ . The SVD is obtained as:

$$\mathbf{x} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (1)$$

where:

$\mathbf{U}$  is the  $T \times N_c$  matrix containing  $N_c$  normalized principal component waveforms ( $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ ),

$\mathbf{S}$  is  $N_c \times N_c$ , diagonal matrix of components amplitudes (singular values),

$\mathbf{V}$  is a  $N_c \times N_c$  matrix mapping components to original data such that  $V_{i,j}$  is the contribution of  $j$ th component to  $i$ th channel;  $\mathbf{V}^T\mathbf{V} = \mathbf{I}$ .

Hence, the columns of  $\mathbf{V}$  with highest corresponding singular values span a subspace of maximal signal energy. Sources contributing to this subspace can be localized using, e.g., the MUSIC algorithm (see MUSIC in Section “Inverse Solutions”). In the SVD implementation used in this study the components are sorted in the descending order of their singular values. The projection  $P_{i,j}$  of the  $j$ th component to  $i$ th channel can be computed as

$$P_{i,j} = U_{\cdot,j} S_{j,j} (V_{i,j})^T \quad (2)$$

*Matching Pursuit Algorithm* Matching pursuit (MP), proposed in Mallat and Zhang (1993), is a suboptimal solution to the NP-hard problem of finding an optimal representation of a signal in a redundant dictionary of functions  $D$ . To analyze multichannel EEG data, multivariate matching pursuit (MMP) was used—a version of the matching pursuit algorithm operating on multichannel signals. MMP can be realized in a variety of ways, depending on the constraints used to select related time-frequency waveforms (*atoms*) in different

channels (cf. Durka 2007). In this work we used the most straightforward implementation assuming that the waveforms  $x_i$  in different channels  $i$  must have the same phase. Maximizing the total energy explained by the atom  $g_\gamma$  in all the channels

$$\max_{g_\gamma \in D} \sum_{i=1}^{N_c} |\langle R^n x_i, g_\gamma \rangle|^2 \quad (3)$$

yields the following algorithm:

$$\begin{cases} R^0 \mathbf{x} = \mathbf{x} \\ R^n \mathbf{x}_i = \langle R^n x_i, g_{\gamma_n} \rangle g_{\gamma_n} + R^{n+1} x_i \\ g_{\gamma_n} = \arg \max_{g_\gamma \in D} \sum_{i=1}^{N_c} |\langle R^n x_i, g_\gamma \rangle|^2 \end{cases} \quad (4)$$

In each iteration, the multichannel residuum  $R^{n+1} \mathbf{x}$  is computed by subtracting from the previous residua in each channel  $i$  the contribution of  $g_{\gamma_n}$ , weighted by  $\langle R^n x_i, g_{\gamma_n} \rangle$ . Results of the applied MMP are given in terms of atoms  $g_{\gamma_n}$ , selected in consecutive iterations, and their weights in all the channels, determined for channel  $i$  by the real-valued products  $\langle R^n x_i, g_{\gamma_n} \rangle$ . In this work we used Gabor functions as atoms:

$$g_\gamma(t) = K(\gamma) e^{-\pi \left(\frac{t-u}{s}\right)^2} \cos(2\pi f(t-u) + \phi) \quad (5)$$

where  $\gamma = \{u, f, s, \phi\}$ , and  $K(\gamma)$  is the normalization coefficient such that  $\|g_\gamma\| = 1$ .

From the results of MMP one can easily retrieve the time course of function  $g_{\gamma_n}$ , projected on the given channel  $i$ :

$$P_{i,n} = \langle R^n x_i, g_{\gamma_n} \rangle g_{\gamma_n} \quad (6)$$

and the distribution of the amplitudes  $\langle R^n x_i, g_{\gamma_n} \rangle$  across channels. Time courses were used by an experienced epileptologist to visually identify the epileptogenic structures. The scalp distributions of selected  $g_\gamma$  were used as input to the localization algorithms.

### Inverse Solutions

From the spectrum of available inverse methods we select two particularly popular algorithms: Multiple Signal Classification (MUSIC) and Adaptive Spatial Filters (Beamformers). These algorithms may serve as a benchmark and starting point for more refined inverse solutions.

**MUSIC** The MUSIC approach is especially useful if topographies of individual sources can be identified only in unknown superposition, e.g. when the first two components of a SVD decomposition both contain topographies of two dipolar sources. Consequently, the input of the algorithm is a subspace spanned by  $N$  topo-

graphies. Finding more than one locations is not trivial and error prone using the standard MUSIC approach. This is because local minima of the error function can also arise from single sources. To avoid misinterpretations of these so called 'ghosts' modifications were formulated in the Recursively Applied MUSIC (RAP-MUSIC) approach (Ermer et al. 2000). Here, further locations are found by projecting out all previously found topographies both from the data and the dipole model and then again finding the brain location with the smallest error. It can be shown that this localization method is equivalent to a fit of an  $N$  dipole model if all sources are uncorrelated. For  $N = 1$ , MUSIC is a dipole fit using a global search. In general, the algorithm scans all points inside the brain and quantifies with an error function whether the topography of a dipole (with optimal orientation) at that point lies within the subspace found e.g. by the SVD decomposition (Mosher et al. 1992). The point with the lowest error is then considered as the location of the source. Special advantages, including applications to single topographies, are that the method scans the whole brain avoiding traps in local minima giving an indication of the stability of the method.

**Adaptive Spatial Filters** Adaptive Spatial Filters (Veen et al. 1997; Sekihara et al. 2005), a.k.a. Beamformers, are linear projections of the signal to be applied in the spatial dimension. Each filter is designed to preserve signals from a specific location in the brain, while maximally suppressing contributions from all other locations. As such, Beamformers can be thought of as frequency-filtering techniques, where pass- and stopband are defined spatially. The most popular Beamformer is the Linearly Constrained Minimum Variance (LCMV) spatial filter (Veen et al. 1997), which minimizes the variance of the filtered signal while keeping the passband-response constant. The LCMV  $N_c \times 3$  filter  $\mathbf{W}(q_0)$  for location  $q_0$  is obtained as the solution to

$$\arg \min_{\mathbf{W}(q_0)} \text{Tr}\{\mathbf{W}^\top(q_0) \mathbf{C}(\mathbf{x}) \mathbf{W}(q_0)\} \quad \text{s.t.} \quad \mathbf{W}^\top(q_0) \mathbf{H}(q_0) = \mathbf{I}, \quad (7)$$

where the  $N_c \times N_c$  matrix  $\mathbf{C}(\mathbf{x})$  is an estimate of the signal's covariance and the  $N_c \times 3$  forward matrix  $\mathbf{H}(q_0)$  describes the theoretical EEG-response of unit-length dipoles at location  $q_0$  pointing in x-, y- and z-direction, respectively.

It can be shown (Veen et al. 1997), that the optimal filter yields the variance estimate

$$\widehat{\text{Var}}(q_0) = \text{Tr}\{[\mathbf{H}^\top(q_0) \mathbf{C}^{-1}(\mathbf{x}) \mathbf{H}(q_0)]^{-1}\}, \quad (8)$$

which, when divided by the expected noise contribution  $\text{Tr}\{\mathbf{H}^\top(q_0)\mathbf{Q}^{-1}\mathbf{H}(q_0)\}^{-1}$ , serves as an index of estimated neural activity originating from  $q_0$ . Here,  $\mathbf{Q}$  is an estimate of the noise covariance matrix.

The estimation of  $\mathbf{C}(\mathbf{x})$  and  $\mathbf{Q}$  can crucially affect the performance of LCMV. If  $\mathbf{C}(\mathbf{x})$  contains substantial background activity, it is helpful to account for this by setting  $\mathbf{Q}$  to a covariance estimate obtained from a baseline measurement. If, on the other hand,  $\mathbf{C}(\mathbf{x})$  captures mostly signal activity (e.g., when computed from a number of spike-related MMP scalp distributions), the identity matrix should be used for  $\mathbf{Q}$ .

### *Suitable Combinations of Preprocessing and Inverse Methods*

The criterion by which MUSIC measures the fitness of a single dipole is based on the distance between the subspaces spanned by signal and dipole, respectively. This measure becomes more and more meaningless when the dimension of the signal space to be explained grows larger. In the extreme case of a full-rank covariance matrix, MUSIC is not applicable anymore. As an implication, MUSIC relies on a suitable dimensionality reduction preprocessing. If the signals of interest are expected to be among the strongest components, SVD is a natural choice. The covariance matrix given as input to SVD can be estimated from the raw EEG signal. An alternative is to derive the covariance matrix from an MMP decomposition, treating the individual topographies of epilepsy-related MMP components as vectors spanning the signal subspace. In our case, the selection of relevant MMP atoms was conducted by an epileptologist.

Beamformers, in contrast to MUSIC, are not based on a subspace-fitting criterion, but take the actual strength of the projected signal into account. It is desirable to supply a full-rank covariance matrix to the Beamformer, since otherwise the solution to cost function (Eq. 7) may not be unique. Since the full rank property cannot be guaranteed in all cases (e.g., if the number of MMP components is smaller than  $N_c$ ), we use a regularized covariance matrix  $\hat{\mathbf{C}} = \mathbf{C} + 1000 \cdot \mathbf{I}$  per default.

Based on these considerations, we present results obtained from four combinations of an inverse solution and a preprocessing method:

RAW-LCMV: LCMV applied to covariance of raw signals,

MMP-LCMV: LCMV applied to covariance of selected MMP topographies,

SVD-MUSIC: MUSIC algorithm applied to the topography of the first SVD component ( $V_{:,1}$ ) of raw signals,

MMP-SVD-MUSIC: MUSIC algorithm applied to the topography of the first SVD component of selected MMP topographies ( $V_{:,1}$ ). This method constitutes yet one level of preprocessing more. Here MMP is applied first in order to achieve a preselection of physiological relevant components. Next the one-dimensional subspace is sought by which the set of MMP topographies can be described best, and localized using MUSIC. In this way we achieve a signal aggregation priori to source localization that might offer more stability than the MMP-LCMV approach—which involves localization of single MMP topographies and latter aggregation in source space.

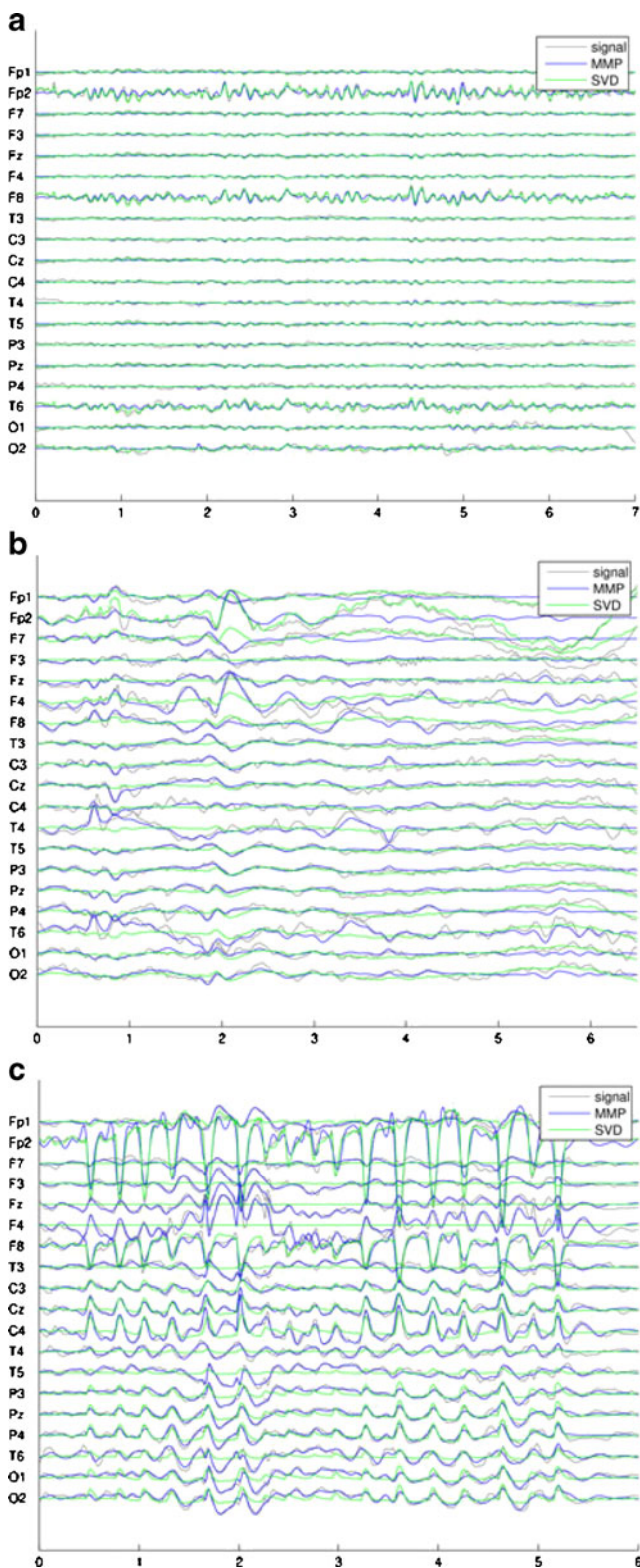
## Results

### *Features Extracted by SVD and MMP*

Both preprocessing methods used in this study can be understood as a filter which extracts components characterised by certain features from the signal. Some of those components are expected to have a better signal to noise ratio for the epileptic activity than the raw signal. The SVD and MMP decompositions, are both driven by the energy (variance) of the signal, however they follow different algorithms and focus on different features. As a result the SVD components have a more global scope, while the MMP tends to give a locally adaptive decomposition. This properties are illustrated in Fig. 2. In these plots the projections of selected MMP atoms (Eq. 6) and the first component of SVD decomposition (Eq. 2) are overlaid on the raw signal. This is done for three selected epochs containing the inter-ictal activity. In subfigure (a) one can see that the abnormal activity is expressed in the form of long oscillations, visible in channels Fp2, F8 and T6, with energy relatively high compared to the background EEG. In such a case the first component of SVD and the selected atoms of the MMP decompositions indicate almost identical features of the signal. Among the 26 analyzed epochs, an analogous situation, i.e. very similar features extracted by SVD and MMP, was observed in 15 cases.

Figure 2b shows another situation. The epileptic discharges are seen in the frontal and temporal channels. However, in channels Fp1, Fp2, F7, F4 and F8, between seconds 5 and 6, we see a high energy artifact. This artifact attracts the first component of SVD towards the





**Fig. 2** Example projections of features extracted in the first SVD component (green) and MMP (blue) overlaid on the raw signal traces: **a** subject FRAANN epoch 3, **b** subject JANPRZ epoch 3, **c** subject JANPAT epoch 3

frontal channels. As a result, the abnormal activity is only partially expressed by the first SVD component, and it is mixed with the artifact. MMP decomposition, due to the more locally adaptive properties, separates the epileptic and artifact activity into different time-frequency atoms. It makes possible the selection of the inter-ictal discharge itself. Analogous situations were observed in 3 of 26 epochs.

In Fig. 2c we see yet another possibility. There are a lot of features that are extracted by both algorithms. However, the first SVD component doesn't contain the activity seen in channels F3, Fz, F4 T3, T4 and T5. In this particular case that activity was separated as the second component in the decomposition. Analogous situations, when the first SVD component doesn't express all the abnormal waveforms in a given epoch, were observed in 8 of 26 epochs.

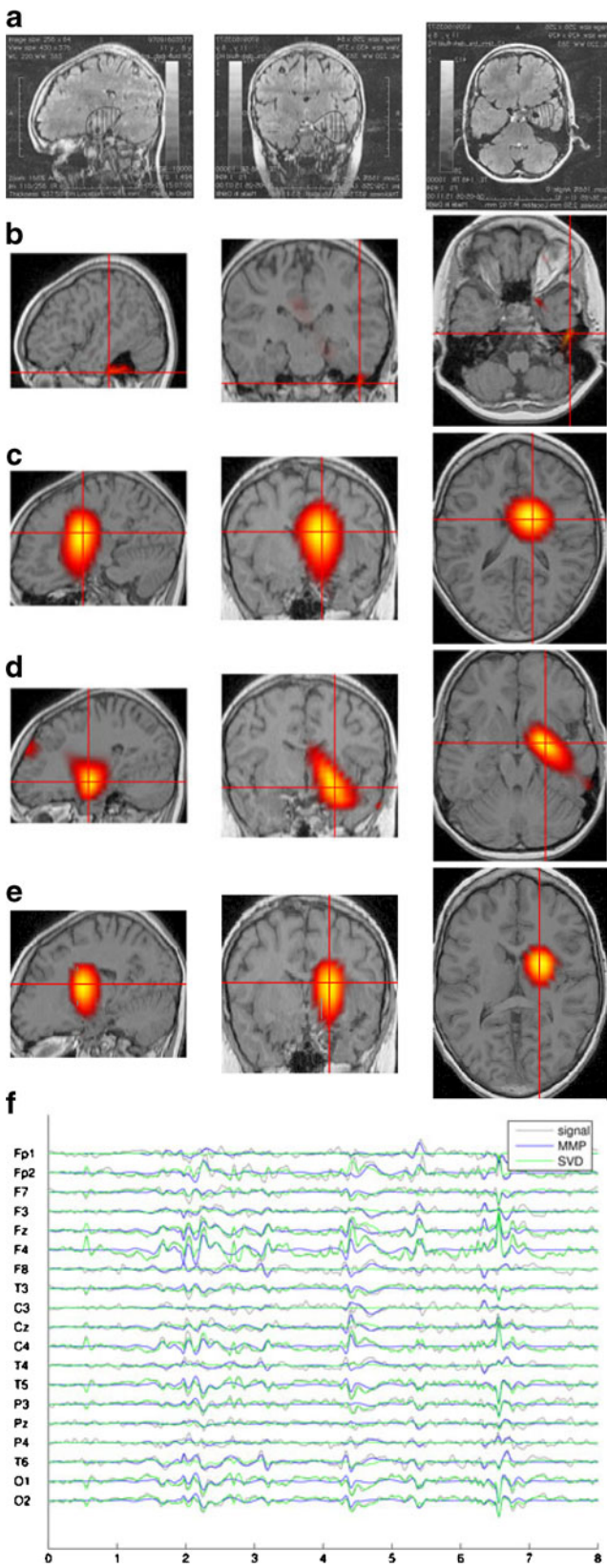
### Localizations

The six analyzed cases can be divided into two classes according to the relation of the computed and the reference localization. Below we describe localizations computed for combinations of the inverse solution and preprocessing methods described in Section “[Suitable Combinations of Preprocessing and Inverse Methods](#)” for the two classes of cases.

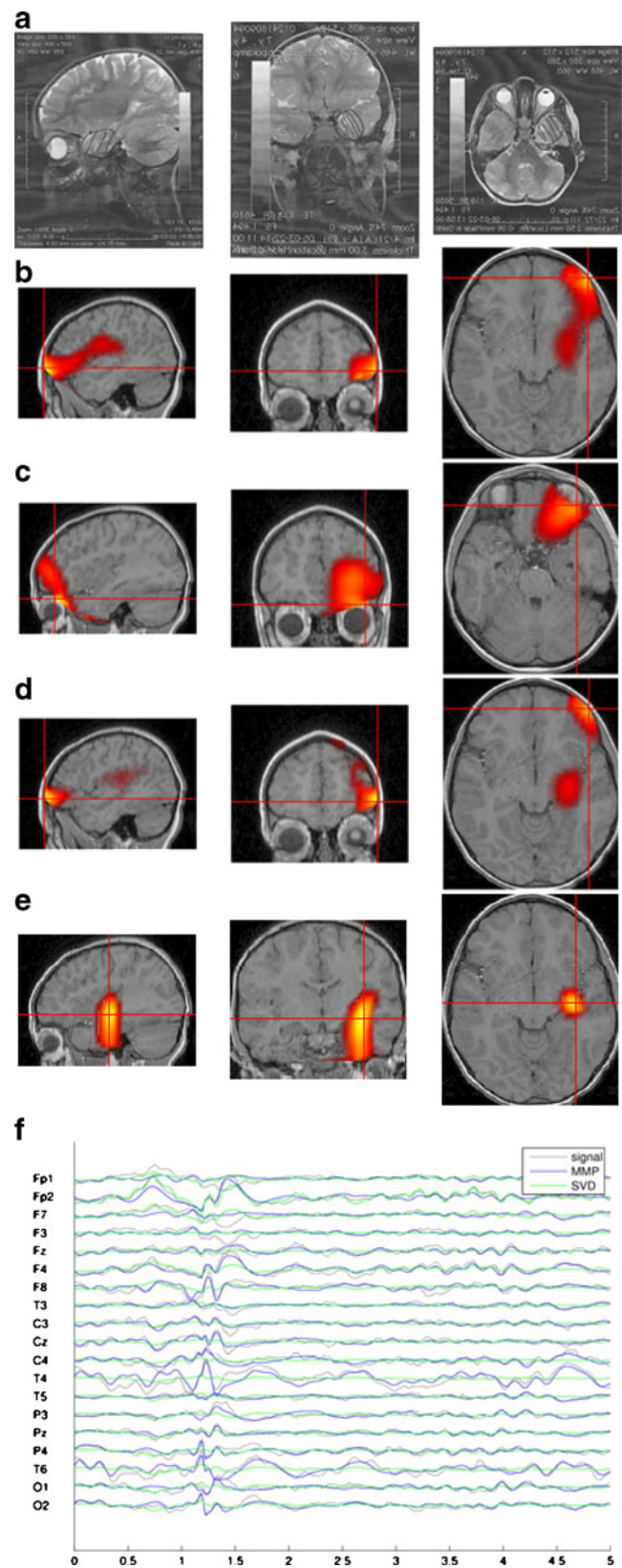
*Cases Where the Inverse Solutions are Compatible with the Epileptogenic Zone* Solutions compatible with the epileptogenic zone were obtained for subjects: CHIMIC, JANPRZ and JATKAM. Localizations obtained for subject CHIMIC are presented in Fig. 3. In this case the RAW-LCMV solution (panel b) indicates a small region on the very edge of the temporal lobe and on the edge of the reference localization (panel a). All other methods, using the pre-processed signal, indicate a much bigger region localized more centrally. The most concordant with the reference localization is the one computed with the MMP-LCMV method (panel d). The SVD-MUSIC and MMP-SVD-MUSIC methods indicate locations which are somewhat higher and more central relative to the reference one. The inter-ictal activity of this subject is in the form of individual discharges, an illustrative plot is presented in Fig. 3f.

Localizations obtained for subject JANPRZ are presented in Fig. 4. In this case the methods RAW-LCMV, SVD-MUSIC, and MMP-LCMV indicate a similar region located just above and behind the right eye. The reference localization is the right temporal lobe. A solution located very close to the reference localization was obtained by the MMP-SVD-MUSIC algorithm (panel e).



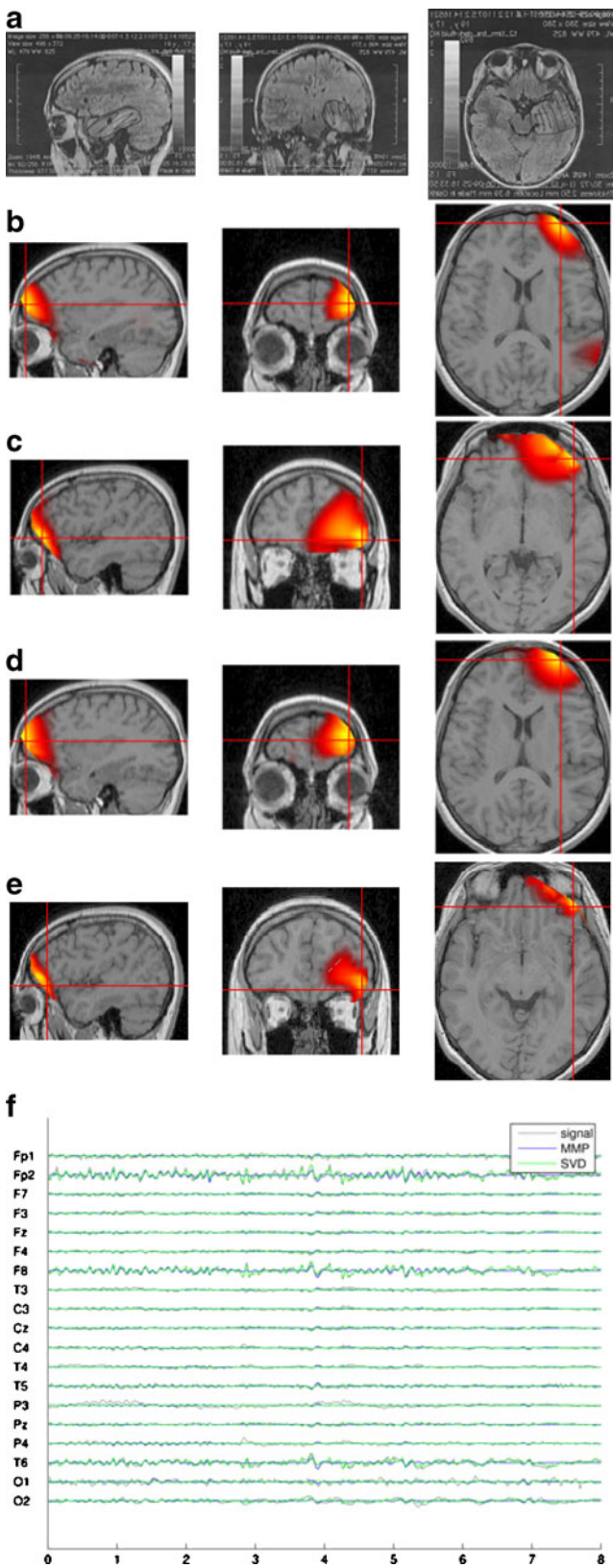


**Fig. 3** Subject CHIMIC: **a** Locations marked by epileptologist, **b** RAW-LCMV, **c** SVD-MUSIC, **d** MMP-LCMV, **e** MMP-SVD-MUSIC, **f** one of the epochs of signal and MMP and SVD components used to compute the inverse solutions

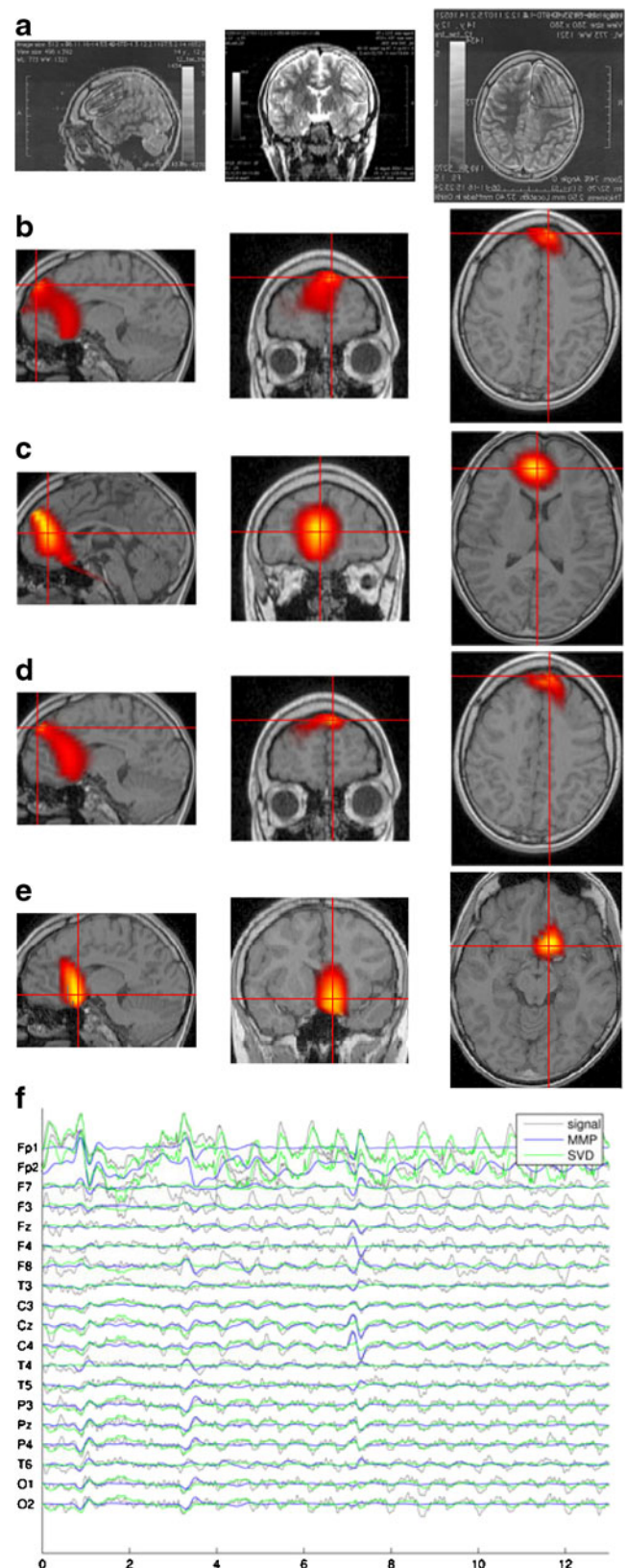


**Fig. 4** Subject JANPRZ: **a** Locations marked by epileptologist, **b** RAW-LCMV, **c** SVD-MUSIC, **d** MMP-LCMV, **e** MMP-SVD-MUSIC, **f** one of the epochs of signal and MMP and SVD components used to compute the inverse solutions





**Fig. 5** Subject FRAANN: **a** Locations marked by epileptologist, **b** RAW-LCMV, **c** SVD-MUSIC, **d** MMP-LCMV, **e** MMP-SVD-MUSIC, **f** one of the epochs of signal and MMP and SVD components used to compute the inverse solutions



**Fig. 6** Subject HRADAW: **a** Locations marked by epileptologist, **b** RAW-LCMV, **c** SVD-MUSIC, **d** MMP-LCMV, **e** MMP-SVD-MUSIC, **f** one of the epochs of signal and MMP and SVD components used to compute the inverse solutions

The inter-ictal activity is represented by discharges similar to that plotted in the Fig. 4f.

The situation for subject JATKAM is very similar to that of subject JANPRZ. RAW-LCMV, SVD-MUSIC, and MPP-LCMV methods all indicate a region which is shifted about 2 centimeters downwards in respect to the reference localization. The MMP-SVD-MUSIC solution is the one closest to the reference region. The inter-ictal activity is in the form of individual discharges.

*Cases for Which All the Inverse Solutions Consistently Indicate a Similar Region, not Compatible with the Epileptogenic Zone* This class of solutions was found for subjects: FRAANN, HRADAW and JATKAM.

Solutions obtained for subject FRAANN are presented in Fig. 5. In this case all four methods indicate a very similar region located in the right frontal cortex. The reference region is the right temporal lobe. It is worth noting the high concordance of all the inverse solutions. The typical abnormal activity for this subject—the oscillations around 8cps lasting a few seconds—is plotted in Fig. 5f and in Fig. 2a.

Solutions obtained for subject HRADAW are presented in Fig. 6. All applied methods indicate a consistent region in the central part of the frontal lobe. Algorithms using LCMV accentuate more the superficial part of that region, while the MUSIC based inverse solution gives a higher probability for sources located more centrally (Fig. 6b–e). In all solutions the computed localization is close to the edge of the reference localization. The typical inter-ictal activity of that subject is plotted in Fig. 6. It constitutes a mixture of low frequency oscillations and individual bursts.

The situation for subject JANPAT is very similar to that of subject FRAANN. Again, all four methods indicate with high concordance a source of the abnormal activity in the right frontal region. The reference localization is in the right temporal lobe. The illustrative activity of that subject is presented in Fig. 2c. It is dominated with low frequency oscillations.

#### *Tracing the Discharges Owing to MMP Preprocessing*

MMP decomposition combined with the MUSIC algorithm gives a straightforward localization of the *individual* time-frequency atoms. Choosing the waveforms related to epileptic discharges allows us to trace their time evolution.

Figure 7 illustrates results obtained for one of the epochs (epoch 2) of subject CHIMIC. We selected this data for the illustration as it reveals a relatively simple pattern. In this case we were able to trace the localization of the consecutive discharges. Separate

localization of subsequent waveforms revealed the following pattern:

- First activity starts in hippocampal region (Fig. 7a)
- Next we observe it in right frontal lobe adjacent to lateral venticle in anterior part (Fig. 7b, c)
- Finally, activity appears in the polar portion of frontal lobe, adjacent to rhino-optical region. (Fig. 7d)

This chain of consecutive locations of activity may correspond to the propagation of the epileptic discharge, migrating from the epileptogenic zone to adjacent structures.

#### **Discussion**

Localization of the epileptogenic zone is the most important clinical issue for the epileptologist. However, certain biological factors interfere with the possibility of such precise localization.

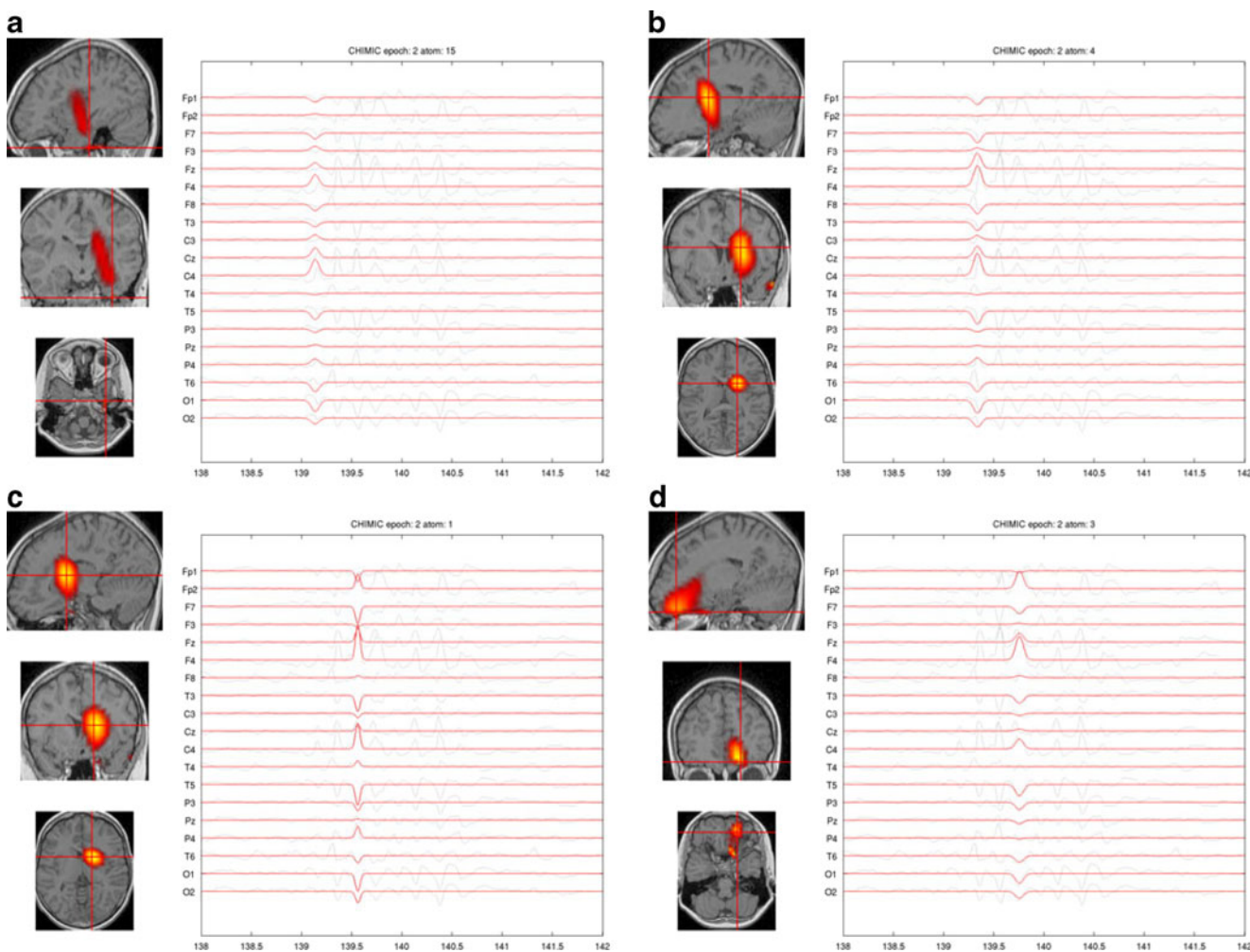
The symptomatogenic zone determines the semiologic sequel of the arising seizure. However, the semiological zone may be distant from the epileptogenic zone or may yield poor indicators in the sense that the eloquent area of the brain triggered by discharges may present few clinical symptoms in some cases, or present some uncertain and non-localizing ones like motionless stare. Moreover, we consider Video-EEG (VEEG, that is parallel registration of patient video and synchronously EEG) seizure registration of limited value if fits are rare. In pediatric epileptology this technique faces additional problems due to long examinations.

Another consideration comes from long-term observations of some drug-resistant patients with epilepsy, where the semiology and clinical picture, as well as the EEG pattern, changes in time. This may resemble the epileptogenic zone's maturation or degeneration.

Inter-ictal EEG can be relevant for localization of brain regions affected by the epileptic activity—however, one must consider the fact that the region of electrical discharge can vary, it is often a much larger area than those arising during ictal neuronal firing, and can be different from the region related to the behavioral observations. These different regions are schematically depicted in Fig. 8.

As mentioned in the Section “[Localizations](#)”, the six analyzed cases can be divided into two classes according to the relation of the computed and the reference localization. In the first class there is at least one method capable of indicating a region which is close to the reference one. In the other class all methods indicate a very similar region, which is significantly different





**Fig. 7** *Right part:* Localizations of successive (in time) MP atoms indicated as maxima of MUSIC inverse solution. *Left part:* traces of the localized atoms (*red*) superimposed on the EEG traces (*gray*)

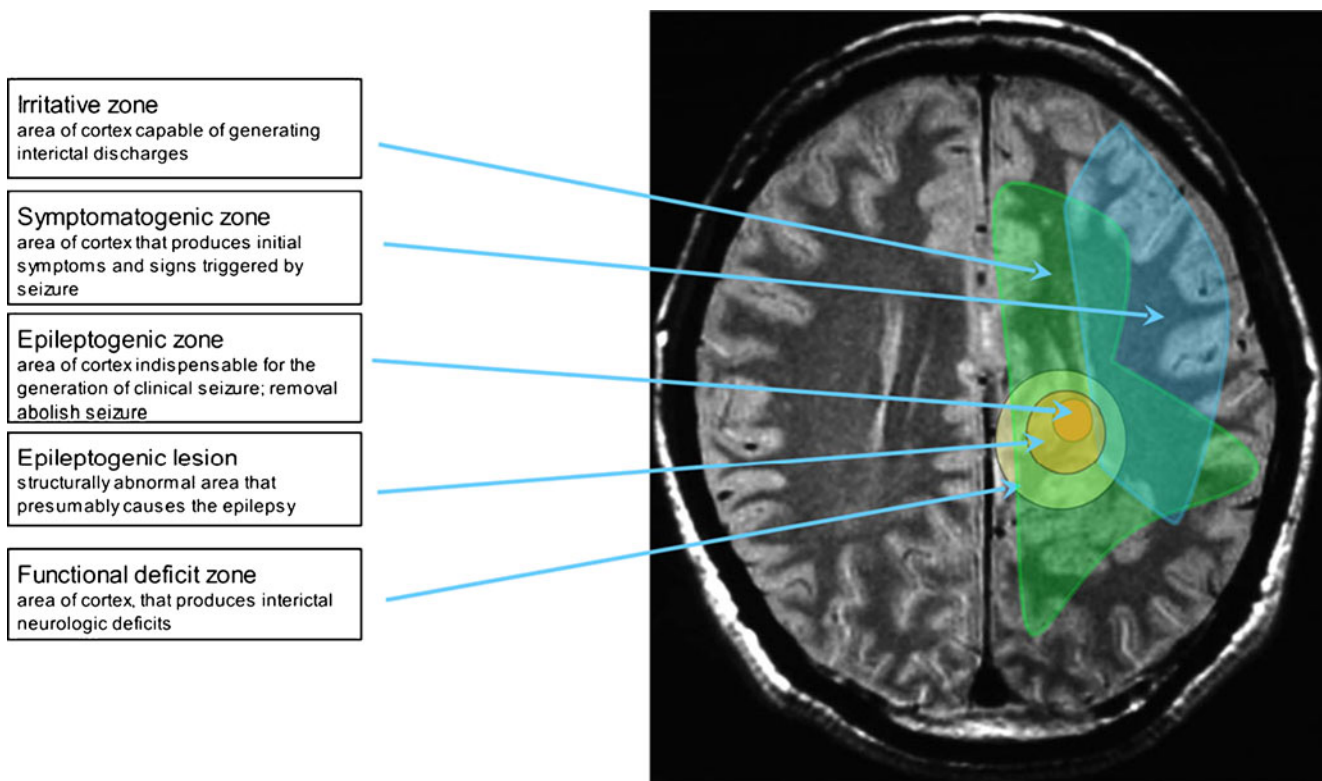
from the reference epileptogenic zone. The inter-ictal discharges are known (Pillai and Sperling 2006) to be located in the irritative zone (see. Fig. 8). Thus we can hypothesize that the second class may contain the cases where the epileptogenic zone is not concordant with the irritative zone. It is interesting that the two classes correlate with different characteristics of the abnormal inter-ictal activity. In the first class—the discharges are in the form of short bursts, in the second class—in the form of well defined oscillations.

The deep location of the epileptogenic zone poses additional problems. Many examples from the database presented deep epileptogenic foci within mesio-temporal structures of the brain, in such cases EEG is presumed to have low detection efficacy—often the only sign of zone activity are tracings of slow theta waves. Nevertheless, in some cases we can compute

the relevant localization as e.g. for subject JANPRZ (Fig. 4e).

In four of the six analyzed cases we observed that, for all the tested preprocessing schemes and inverse solution algorithms, the localizations are broadly consistent indicating similar regions (this is the case for subjects FRAANN, JANPRZ, HRADAW, JANPAT). In the two remaining cases (subjects CHIMIC and JATKAM), the localizations obtained from SVD-MUSIC and MMP-LCMV are similar, but differ from the RAW-LCMV solution. In these two cases the RAW-LCMV solution indicates a small region at the edge of the head, while in the other solutions that region is also present—but with a lower probability of the source being located there, some other, more extended regions with a high source probability are indicated in other locations (Fig. 3).





**Fig. 8** Schematic illustration of areas related to or affected by epileptic discharges

On the fine scale, however, there are more differences in the obtained solutions. The three tested methods differ mainly in the selectiveness of the input data, RAW-LCMV being least selective and MMP-LCMV and MMP-SVD-MUSIC being the most selective. This is indeed reflected in the obtained solutions. In cases where the epilepsy-related components are dominating, the solutions for all three preprocessings are similar (c.f. Fig. 5). The selectivity offered by MMP preprocessing is important in cases where the EEG is contaminated by artifacts. The artifacts can be of a similar power to the epileptic discharges. Therefore, e.g. the first component of SVD-MUSIC can mix the epileptic and artifact activity, like in Fig. 2b. In this case, the result of the inverse solution is biased towards the source of the artifact (e.g. Fig. 4c).

The presented solutions are just illustrations of the usage of the database. The obtained results point to perhaps trivial, but important observations. The significant contribution towards a solution to the inverse problem is through the careful selection of the interictal EEG epochs. In the analyzed cases, the activity in the form of individual bursts seemed to be the most adequate for the localization of the epileptogenic zone. Since the epileptic signals are very strong, methods of

preprocessing based on the selection of components according to their variance should give similar results.

Methods used in 3-D visualization of consecutive inter-ictal discharges (approximated by individual MMP atoms, see Section “Tracing the Discharges Owing to MMP Preprocessing”) give pictures that may resemble spatial electric arousal of different brain structures triggered by the epileptogenic zone. This may lead to recognition of individual patterns of discharge, as well as to the precise delineation of an epileptogenic zone if all the initial atoms of the discharge are concordant with the epileptogenic zone found during ECoG.

It is widely acknowledged that facilitating the widespread sharing of data and tools for neuroscientific research will accelerate the development of neuroinformatics (Eckersley et al. 2003). For example, in relation to epilepsy, one of the first research resources for complex biological signals—PhysioNet (Goldberger et al. 2000, <http://physionet.org>), offers “CHB-MIT Scalp EEG Database”, containing EEG recordings of 23 cases of pediatric subjects with intractable seizures collected at the Children’s Hospital Boston, freely available at <http://physionet.org/pn6/chbmit/>. University of Freiburg has developed a web-based database of

EEG recordings <http://epilepsy.uni-freiburg.de>, which has served as a common database for many seizure prediction studies—access to this data requires approval of the FDM and written consent to the terms and conditions. In the field of brain-computer interfaces, free access to data from BCI competitions (e.g. <http://www.bbci.de/competition/iv/>) stimulates development and makes possible comparison of various BCI algorithms. To our knowledge there is no database presenting EEG and MRI data with the location of the epileptic zone known suitable for testing inverse solution methods, therefore we feel it is valuable to make the reported data accessible to scientific community.

## Conclusions

Open access to the above described datasets offers a unique opportunity to verify and compare different methods of inverse solutions on the same real data. Spatial location of the epileptogenic zone was settled in each case with ECoG; nevertheless, it may not be understood as a spherical area, but rather a net of structures with unidentified shape. The algorithms of the inverse solution can be verified against the exact location of epileptogenic zone revealed by ECoG in operated subjects. However, we have to keep in mind that the main components of the signal measured by the scalp EEG are not expected to be generated in the epileptic foci. On the other hand, the knowledge of the location of foci can be used to verify methods developed to recover the activity at the foci itself, including new approaches like the tracing of the individual structures presented in this paper.

## Information Sharing Statement

All the datasets, described in Section “The Database”, are freely available from <http://eeg.pl/epi>.

**Acknowledgements** This work was supported from Polish funds for science 2007-2010 grant N407 041 32/1646, and the German Federal Ministry of Education and Research (BMBF) grant 01GQ0850 and the FP7-ICT Programme of the European Community, under the PASCAL2 Network of Excellence, grant ICT-216886.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Cao, N., Yetik, I., Nehorai, A., Muravchik, C., & Haeisen, J. (2006). Parametric surface-source modeling and estimation with electroencephalography. *IEEE Transactions On Bio-medical Engineering*, 53(12 Pt 1), 2414–24.
- Durka, P. (2007). *Matching pursuit and unification in EEG analysis*. Artech House, isbn 978-1-58053-304-1 edition.
- Durka, P., Matysiak, A., Martinez-Montes, E., Valdes-Sosa, P., & Blinowska, K. (2005). Multichannel matching pursuit and EEG inverse solutions. *Journal of Neuroscience Methods*, 148(1), 49–59.
- Eckersley, P., Egan, G., De Schutter, E., Yiyuan, T., Novak, M., Sebesta, V., et al. (2003). Neuroscience data and tool sharing. *Neuroinformatics*, 1(2), 149–165.
- Ermer, J., Mosher, J., Huang, M., & Leahy, R. (2000). Paired MEG data set source localization using recursively applied and projected (RAP) MUSIC. *IEEE Transactions on Bio-medical Engineering*, 47(9), 1248–60.
- Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. Ch., Mark, R. G., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23), e215–e220 (Circulation Electronic Pages; <http://circ.ahajournals.org/cgi/content/full/101/23/e215>).
- Haufe, S., Nikulin, V., Ziehe, A., Müller, K., & Nolte, G. (2008). Combining sparsity and rotational invariance in EEG/MEG source reconstruction. *Neuroimage*, 42(2), 726–738.
- Irimia, A., Swinney, K., & Wikswo, J. (2009). Partial independence of bioelectric and biomagnetic fields and its implications for encephalography and cardiography. *Physical Review. E, Statistical, Nonlinear And Soft Matter Physics*, 79(5 Pt 1), 051908.
- Karbowski, K. (1990). Sixty years of clinical electroencephalography. *European Neurology*, 30(3), 170–175.
- Kemp, B., Värri, A., Rosa, A. C., Nielsen, K. D., & Gade, J. (1992). A simple format for exchange of digitized polygraphic recordings. *Electroencephalography and Clinical Neurophysiology*, 82, 391–393.
- Koffler, D. J., & Gotman, J. (1985). Automatic detection of spike-and-wave bursts in ambulatory EEG recordings. *Electroencephalography and Clinical Neurophysiology*, 61(2), 165–180.
- Loddenkemper, T., & Kotagal, P. (2005). Lateralizing signs during seizures in focal epilepsy. *Epilepsy & Behavior*, 7(1), 1–17.
- Mallat, S., & Zhang, Z. (1993). Matching pursuit with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41, 3397–3415.
- Molins, A., Stufflebeam, S., Brown, E., & Hämäläinen, M. (2008). Quantification of the benefit from integrating MEG and EEG data in minimum l2-norm estimation. *NeuroImage*, 42(3), 1069–1077.
- Mosher, J., Lewis, P., & Leahy, R. (1992). Multiple dipole modeling and localization from spatio-temporal MEG data. *IEEE Transactions on Biomedical Engineering*, 39(6), 541–557.
- Niedermeyer, E., & Lopes da Silva, F. H. (2004). *Electroencephalography: Basic principles, clinical applications, and related fields* (p. 140). Lippincott Williams & Wilkins.
- Nolte, G., & Dassios, G. (2005). Analytic expansion of the EEG lead field for realistic volume conductors. *Physics in Medicine & Biology*, 50, 3807–3823.
- Ou, W., Hämäläinen, M., & Golland, P. (2009). A distributed spatio-temporal EEG/MEG inverse solver. *NeuroImage*, 44(3), 932–946.

- Pascual-Marqui, R., Esslen, M., Kochi, K., & Lehmann, D. (2002). Functional imaging with low-resolution brain electromagnetic tomography (LORETA): A review. *Methods and Findings In Experimental And Clinical Pharmacology*, 24(Suppl C), 91–95.
- Penin, H. (1968). Elektronische Patientenüberwachung in der Nervenklinik Bonn (Electronic patient monitoring in the neurologic hospital of Bonn). *Umsschau in Wissenschaft und Technik*, 7, 211–212.
- Pillai, J., & Sperling, M. (2006). Interictal EEG and the diagnosis of epilepsy. *Epilepsia*, 47(Suppl 1), 14–22.
- Rosenow, F., & Luders, H. (2001). Presurgical evaluation of epilepsy. *Brain*, 124, 1683–1700.
- Scherg, M., & Ebersole, J. (1993). Models of brain sources. *Brain Topography*, 5(4), 419–423.
- Sekihara, K., Sahani, M., & Nagarajan, S. (2005). Localization bias and spatial resolution of adaptive and non-adaptive spatial filters for MEG source reconstruction. *NeuroImage*, 25(4), 1056–1067.
- Veen, B. V., van Drongelen, W., Yuchtman, M., & Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Biomedical Engineering*, 44(9), 867–880.
- Wong, P. K. (1998). Potential fields, EEG maps, and cortical spike generators. *Electroencephalography and Clinical Neurophysiology*, 106(2), 138–141.